DTIC FILE COPY

AGARD-CP-420

AD-A217 612

AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT

7 RUE ANCELLE 92200 NEUILLY SUR SEINE FRANCE

AGARD CONFERENCE PROCEEDINGS No.420

Effects of Electromagnetic Noise and Interference on Performance of Military Radio Commmunication Systems

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

DTIC
ELECTE
FEB 01 1989
S D

NORTH ATLANTIC TREATY ORGANIZATION

DISTRIBUTION AND AVAILABILITY
ON BACK COVER

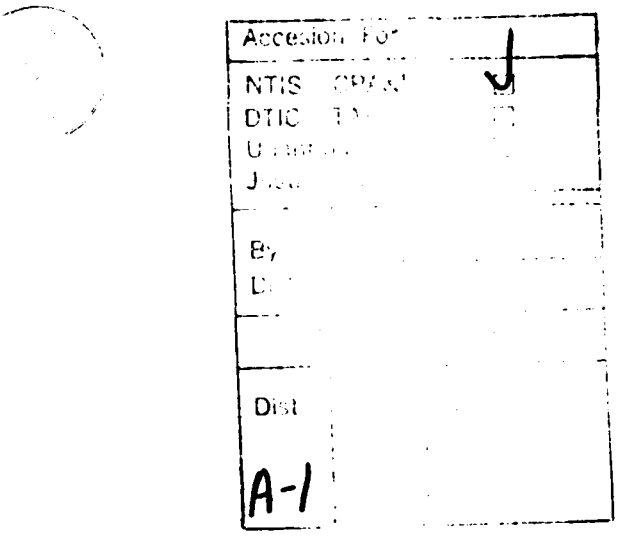
NORTH ATLANTIC TREATY ORGANIZATION

ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT

(ORGANISATION DU TRAITE DE L'ATLANTIQUE NORD)

AGARD Conference Proceedings No.420

EFFECTS OF ELECTROMAGNETIC NOISE AND INTERFERENCE ON PERFORMANCE OF MILITARY RADIO COMMUNICATION SYSTEMS



Papers presented at the Electromagnetic Wave Propagation Panel Specialists' Meeting held in Lisbon, Portugal, 26—30 October 1987.

THE MISSION OF AGARD

According to its Charter, the mission of AGARD is to bring together the leading personalities of the NATO nations in the fields of science and technology relating to aerospace for the following purposes:

- Recommending effective ways for the member nations to use their research and development capabilities for the common benefit of the NATO community;
- Providing scientific and technical advice and assistance to the Military Committee in the field of aerospace research and development (with particular regard to its military application);
- Continuously stimulating advances in the aerospace sciences relevant to strengthening the common defence posture;
- Improving the co-operation among member nations in aerospace research and development;
- Exchange of scientific and technical information;
- Providing assistance to member nations for the purpose of increasing their scientific and technical potential;
- Rendering scientific and technical assistance, as requested, to other NATO bodies and to member nations in connection with research and development problems in the aerospace field.

The highest authority within AGARD is the National Delegates Board consisting of officially appointed senior representatives from each member nation. The mission of AGARD is carried out through the Panels which are composed of experts appointed by the National Delegates, the Consultant and Exchange Programme and the Aerospace Applications Studies Programme. The results of AGARD work are reported to the member nations and the NATO Authorities through the AGARD series of publications of which this is one.

Participation in AGARD activities is by invitation only and is normally limited to citizens of the NATO nations.

The content of this publication has been reproduced
directly from material supplied by AGARD or the authors.

Published December 1988

Copyright © AGARD 1988
All Rights Reserved

ISBN 92-835-0487-9

Printed by Specialised Printing Services Limited
40 Chigwell Lane, Loughton, Essex IG10 3TZ

THEME

To achieve reliable radio communications, a thorough knowledge of the communications channel is required. It is important for system design to characterize the channel, which in general, consists of intentionally radiated signals, noise of either natural origin or radiated from man-made sources or composite variations of these signals. The various noise and interference sources existing in the communications channel impose a fundamental limit on achievable radio system performance. Communications equipment must be designed to cope with the effects of electromagnetic noise and interference.

The impact on communications systems can be: 1) degraded performance; 2) disruption in communications; or 3) equipment damage from severe electromagnetic disturbances.

To determine the effects of these disturbances it is essential that the characteristics of the unwanted signals be measured, modelled and particularly simulated, since the effects of electromagnetic noise and interference on system performance, and system optimization to cope with these interference effects are better studied in a laboratory environment employing simulated channels (signal, noise and interference).

The application of radio communications systems on ships and aircraft, and the resulting mobility of the ground forces results in a very dynamic ever-changing highly interactive environment. The effects of electromagnetic noise and interference on these systems performance and EMC is a most important subject for NATO.

Key words: Natural radio noises, man-made radio noise and interference, communications system performance, jamming, electronic countermeasures, lightning, EMP and thermal effects, EMC and spectrum management, NATO

Pour obtenir des communications radio fiables, il faut une connaissance complète de la voie de communication. Pour la conception du système il est important de caractériser la voie qui véhicule généralement des signaux volontairement émis, du bruit d'origine naturelle ou émis par des appareils fabriqués par l'homme, ainsi que diverses combinaisons de ces signaux. Les diverses sources de bruit et d'interférence existant sur la voie de communication imposent une limite fondamentale aux performances possibles d'un système radio. L'équipement de communication doit être conçu en tenant compte des effets du bruit et des interférences électromagnétiques.

Le bruit et les interférences peuvent avoir les effets suivants sur les systèmes de communication: 1) Dégradation des performances; 2) interruption des communications; ou 3) endommagement de l'équipement du fait des fortes perturbations électromagnétiques.

Pour déterminer les effets de ces perturbations, il faut absolument mesurer, modéliser et en particulier simuler les caractéristiques des signaux indésirables, car les effets du bruit et des interférences électromagnétiques sur les performances du système peuvent être plus facilement étudiés dans l'environnement d'un laboratoire utilisant des voies simulées (signal, bruit et interférences), et l'on peut ainsi optimiser plus efficacement le système.

L'utilisation des systèmes de communication radio sur les navires et les avions et la mobilité des forces terrestres créent un environment dynamique qui évolue sans cesse et qui est très interactif. Les effets du bruit et des interférences électromagnétique (EMC) de ces systèmes ont une grande importance pour l'OTAN.

ELECTROMAGNETIC WAVE PROPAGATION PANEL

Chairman: Prof. C.Goutelard
L.E.T.T.I.
Université Paris-Sud
9 Avenue de la Division Leclerc
94230 Cachan
France

Deputy Chairman: Ir. H.Vissinga
Physics and Electronics Lab. TNO
P.O. Box 96864
2509 JG The Hague
Netherlands

TECHNICAL PROGRAMME COMMITTEE

Dr J.S.Belrose (Co-Chairman)
P.O. Box 11490
Communications Research Centre
Station H, Ottawa, K2H 8S2
Canada

Dr J.Bishop (Co-Chairman)
Flight Systems Department
Royal Aircraft Establishment
Farnborough, Hants GU14 6TD
UK

Mr W.R.Lauber
Communications Research Centre
P.O. Box 11490, Station H
Ottawa, Ontario K2H 8S2
Canada

Dr A.D.Spaulding
Chief, Propagation Model Development
and Application
NTIS/ITS
325 Broadway
Boulder, CO 80303
USA

Dr Ing. D.Rother
Standard Elektrik Lorenz AG (SEL)
Abt PCM ERM
Ostendstrasse 3
D-7530 Pforzheim
Federal Republic of Germany

Ir H.Vissinga
Physics and Electronics Lab. TNO
P.O. Box 96864
2509 JG The Hague
Netherlands

Dr D.Yavuz
Communications Division
SHAPE Technical Centre
P.O. Box 174
2501 CD The Hague
Netherlands

AGARD EPP EXECUTIVE

AGARD
7 rue Ancelle
92200 Neuilly sur Seine
France

From the USA and
Canada:
AGARD-NATO
A.P.O. New York 09777

CONTENTS

	Page
THEME	iii
TECHNICAL PROGRAMME COMMITTEE/PANEL OFFICERS	iv
	Reference
SESSION I – NATURAL RADIO NOISE	
RADIO NOISE CHARACTERISTICS AND THEIR SYSTEM PERFORMANCE IMPLICATIONS by J.R.Herman	1
BALLOON MEASUREMENTS OF LOW FREQUENCY TE AND TM ATMOSPHERIC NOISE by J.P.Turtle and P.A.Kossey	2
LIGHTNING GENERATION OF LOW-FREQUENCY TE ATMOSPHERIC NOISE by E.C.Field Jr and C.Warber	3
VLF RADIO NOISE by M.D.Andrews, L.DeBlasio and F.J.Kelly	4
A NEW GLOBAL SURVEY OF ELF/VLF RADIO NOISE by A.C.Fraser-Smith, R.A.Helliwell, B.R.Fortnam, P.R.McGill and C.C.Teague	4A
SESSION II – MAN MADE RADIO NOISE AND INTERFERENCE	
MAN-MADE RADIO NOISE AND INTERFERENCE (invited) by G.H.Hagn	5
TRANSHORIZON COCHANNEL INTERFERENCE PATH MEASUREMENTS IN EUROPE by M.P.M.Hall	6
EXPERIMENTAL OBSERVATIONS OF SPECTRAL OCCUPANCY AT HF by G.F.Gott, P.J.Laycock, A.R.Ray and M.Morrell	7
NOISE MEASUREMENT SYSTEMS/TECHNIQUES by C.B.Christianson	8
SIGNAL NOISE/INTERFERER COMBINER UNIT-PROGRAMMABLE (SINCUP) by E.M.De Pison	9
SESSION III – COMMUNICATIONS SYSTEM PERFORMANCE AND DESIGN	
EFFECTS OF NOISE AND INTERFERENCE ON SYSTEM PERFORMANCE (invited) by A.D.Spaulding	10
THE DESIGN AND PERFORMANCE OF AN ADAPTIVE, PACKET-SWITCHED HF DATA TERMINAL by G.R.Nourry and A.J.Mackie	11
DESIGN CONSIDERATIONS FOR AN AUTOMATED HF DATA NETWORK WITH ADAPTIVE CHANNEL SELECTION by R.J.Goodwin and A.P.C.Reed	12
EXPERIMENTAL ROBUST HF MODEMS INCORPORATING ADAPTIVE EXCISION by G.F.Gott, P.Doany, S.W.Wong and E.P.Darbyshire	13
ARTIFICIAL INTELLIGENCE IN HF COMMUNICATION SYSTEMS by A.P.Jowett	14

Reference

PROCESSING TECHNIQUES FOR RADIO COMMUNICATIONS IN NON-GAUSSIAN NOISE ENVIRONMENTS
by M.Darnell 15

IMPROVED CODING AND CONTROL OF HF SYSTEMS IN A NON-GAUSSIAN NOISE ENVIRONMENT
by J.Hague 16

ELECTROMAGNETIC CODE FOR NAVAL APPLICATIONS
by F.Crescimbeni, F.Bessi and S.Chiti 17

NARRATIVE COMPRESSION CODING FOR A CHANNEL WITH ERRORS
by J.W.Bond 18

NEW METHOD FOR SUPPRESSION OF GONIOMETER DF ERRORS BY RADIO NOISE FROM OVERHEAD POWER LINES
by J.Benger, G.Deubach and G.Kosel 19

SESSION IV – JAMMING AND ELECTRONIC COUNTERMEASURES

JAMMING AND ELECTRONIC COUNTERMEASURES (invited)
by V.K.Bhargava and Q.Wang 20

ECM RESISTANCE IN THE HF BAND BY USE OF ADAPTIVE REACTION AND FREQUENCY HOPPING
by G.Greiner, P.Iselt and G.Müller 21

PREDICTION OF EFFECTS OF NOISE AND INTERFERENCE ON FAST FREQUENCY-HOPPED M-ARY NCFSK SYSTEMS
by E.B.Felstead, J.S.Bird and U.A.Tenne-Sens 22

INHERENT-JITTER EFFECTS IN THE NEW RAPID ACQUISITION METHOD FOR FREQUENCY-HOPPED PUSH–TO-TALK SPREAD SPECTRUM SYSTEMS: A JITTER-FIGHTING SCHEME
by J.M.N.Pereira and A.A.Albuquerque 23

SESSION V – LIGHTNING, EMP AND NUCLEAR EFFECTS

DEVELOPMENT OF A NEW HIGH ALTITUDE ELECTROMAGNETIC PULSE (HEMP) ENVIRONMENT AND RESULTING OVERHEAD LINE RESPONSES (invited)
by F.M.Tesche and P.R.Barnes 24

EMP-INDUCED TRANSIENTS AND THEIR IMPACT ON SYSTEM PERFORMANCE (invited)
by R.Sturm 25

Paper 26 withdrawn

SHIELDS FOR PROTECTING CABLES FROM THE EFFECTS OF ELECTROMAGNETIC NOISE AND INTERFERENCE
by L.O.Hoeft, J.S.Hofstra, R.J.Karaskiewicz and B.W.Torres 27

SESSION VI – EMC AND SPECTRUM MANAGEMENT

AUTOMATING SPECTRUM MANAGEMENT OF TACTICAL COMMUNICATIONS TO MINIMIZE MUTUAL INTERFERENCE AND MAXIMIZE FREQUENCY REUSE ON THE BATTLEFIELD
by S.M.Segner 28

COSITING OF RADIO TERMINALS
by G.Cucinotta and R.Azzarone 29

INTERMODULATION INTERFERENCES IN RADIO SYSTEMS
by P.S.W.Ho, L.H.Bevan, A.C.C.Tseung and W.S.Wilkinson 30

Reference

SESSION VI – CLASSIFIED SESSION

ASSESSMENT OF MILITARY USE OF THE BAND 20–108 MHz IN CENTRAL EUROPE
by G.Wyman — 31*

JAMMING THREAT MODEL AND ASSESSMENTS OF JAMMING EFFECTIVENESS IN RELATION WITH VHF COMBAT NET RADIO SYSTEMS
by K.S.Kho — 32*

SELF INTERFERENCE IN THE CO-SITE OPERATIONS OF COMBAT NET RADIO SYSTEMS
by K.S.Kho — 33*

CURRENT AND FUTURE TACTICAL AIR RADIO SYSTEMS: ECM ISSUES – ENHANCEMENT OPPORTUNITIES
by D.Yavuz and R.L.Nowill — 34*

AUTOMATING THE DECONFLICTION OF JAMMING AND SPECTRUM MANAGEMENT
by S.M.Segner — 35

THE UNMANNED AIRCRAFT COMMUNICATION LINK AND ITS OPERATING ENVIRONMENT
by J.M.Jones — 36*

INTERFERENCE LIMITATIONS ON HF SPREAD SPECTRUM SYSTEMS
by J.Pennington — 37*

A SOFTWARE-IMPLEMENTED ADAPTIVE ANTENNA SYSTEM FOR HF COMMUNICATIONS
by R.W.Jenkins — 38*

Paper 39 withdrawn

DETERMINING THE AZIMUTHAL DISTRIBUTION OF NOISE AND INTERFERENCE ACROSS THE HF SPECTRUM AT THE RAE RECEIVING STATION, COBBETT HILL, SURREY, ENGLAND
by H.L.Spong — 40

SYNCHRONIZATION BY CODE-REGENERATION IN SPREAD-SPECTRUM-SYSTEMS
by M.Ottka and U.Seier — 41*

go to p. iii for Keynote

* Printed in classified publication CP 420 (Supplement).

RADIO NOISE CHARACTERISTICS AND THEIR SYSTEM PERFORMANCE IMPLICATIONS

by
John R. Herman
Senior Scientist
GTE GOVERNMENT SYSTEMS CORPORATION
Strategic Systems Division
1 Research Drive
Westborough, MA
01581
U.S.A.

SUMMARY

This survey reviews the impacts of correlated-pulse impulsive noise on communication coding and system performance. Examples of atmospheric and man-made noise are provided to set the background. The statistical characteristics of radio noise are defined and several noise mitigation techniques are discussed.

1.0 INTRODUCTION

The ultimate limitation of radio systems performance is determined by the noise environment within which the received signal is imbedded. With an understanding of the noise characteristics, it becomes possible for the design engineer to optimize the radio system performance by considering an array of mitigation techniques ranging from a simple increase in transmitted signal power to sophisticated signal processing including interleaving and forward error correction.

In this survey, attention is restricted to noise external to the receiving system and includes natural and man-made sources. The basic statistical characteristics of noise are covered, including the amplitude probability distribution (APD), the voltage deviation parameter (V_d), the noise power factor (F_a), and higher order statistics, such as autocorrelation function (ACF), crossing rates (ACR) and pulse spacing and duration distributions (PSD, PDD). System performance for linear and nonlinear receivers with and without interleaving, and Hamming code performance in burst noise vis a vis Gaussian noise are then surveyed in the light of the affecting noise characteristics.

2.0 NOISE MEASUREMENT TECHNIQUE

Most of the measured noise characteristics were obtained using the dual-channel measuring system illustrated in Figure 1 (Herman et al., 1986, Ref. 1). An HF version of the receiver also exists.

The noise coming in through the antenna is filtered and amplified, quadrature demodulated, and further filtered, and then passed through a set of log amplifiers to increase the dynamic range.

The log compressed data is digitized at a rate of 250 kHz complex sampling and multiplexed to a 14-channel, high density digital tape recorder run at 120 in/s for a time resolution of 4 μs.

The high density measurement data is subsequently transcribed to computer tapes for analysis. The analysis software is used to obtain both first order and higher order noise statistics as well as time waveforms of the noise signatures.

3.0 EXAMPLES OF ATMOSPHERIC AND MAN-MADE NOISE

An illustrative 20-s digitized time sample of MF atmospheric noise is shown in Figure 2. HF noise is quite similar in structure. The amplitude has been normalized to the rms voltage level. Each point in the figure is an 8-ms average of the sampled noise.

Here it is evident that the noise in this sample varies over a dynamic range of about 60-70 dB (well within the 115-dB dynamic range of the equipment, so there is no saturation).

It is this continually fluctuating noise level over a wide dynamic range that poses

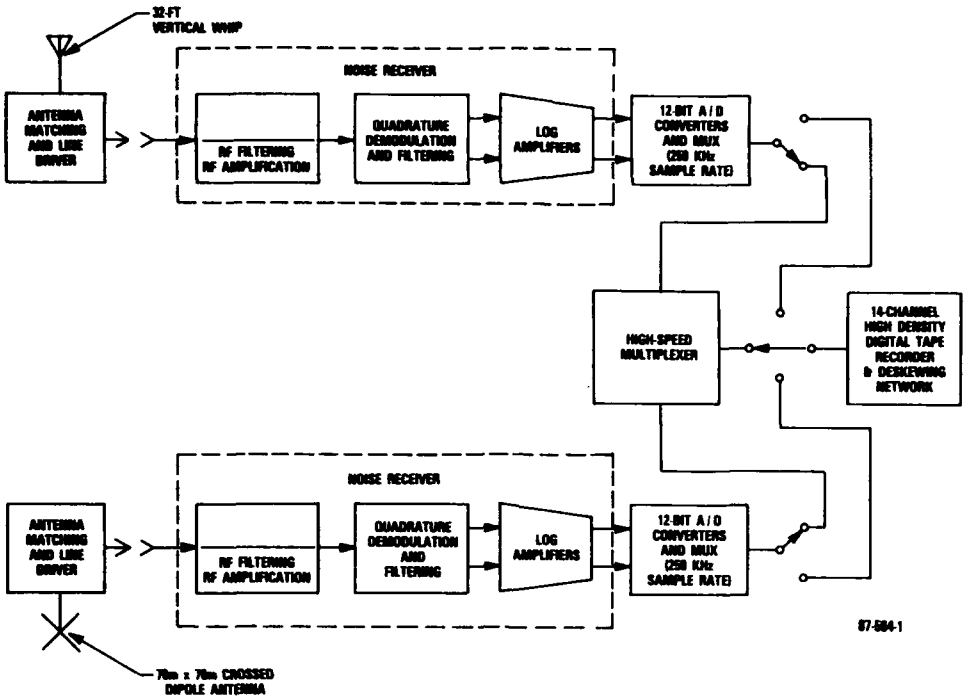


Figure 1. Noise-recording System.

many problems in digital communications. The quasi-periodicity in the data shown here is rather slow and sedate compared to a bit length, even though it does cover a large range of amplitude.

Let us, then, examine the finer structure of this noise by expanding the 3 indicated data points in Figure 2, representing a median level (Block 2 near V_{rms}), a high level (Block 1867) near the peak, and a low level, more or less minimum background (Block 282).

The 8-ms data point near V_{rms} in Figure 2 is shown in Figure 3 expanded by a factor of 1024, so that now each point on the plot is about 8 μs of data - that is, a single in-phase and quadrature sample of 4 μs each, taken at the 250 kHz sampling rate mentioned earlier.

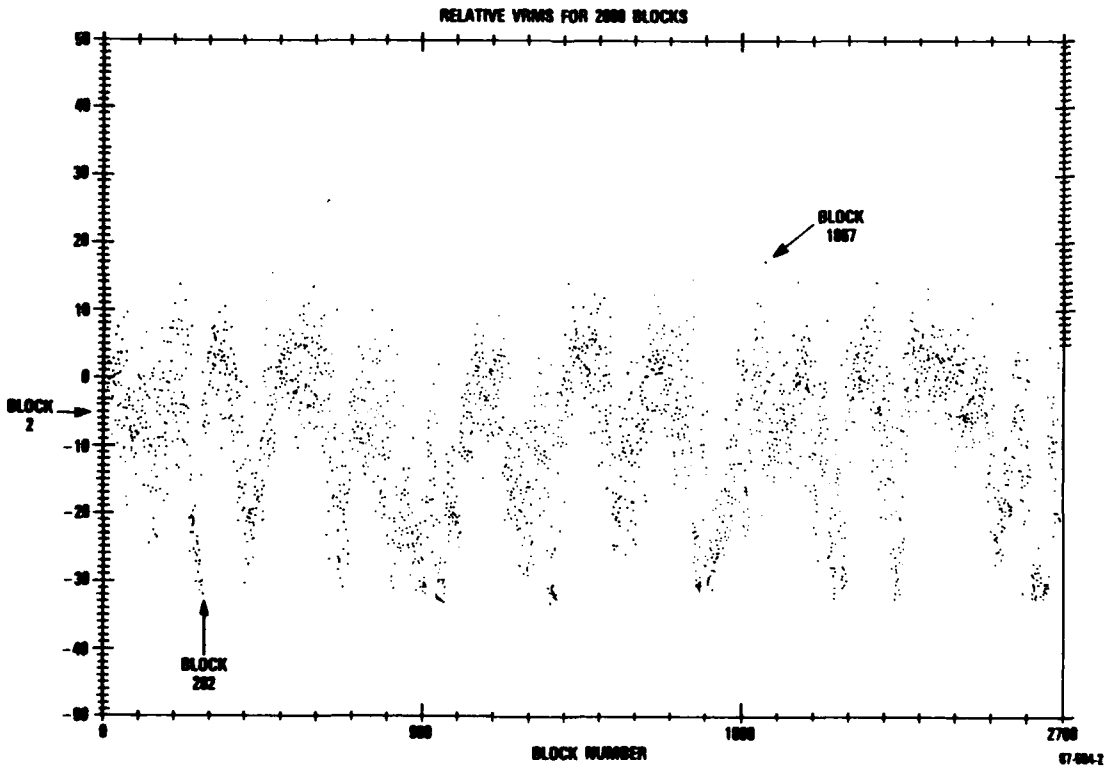


Figure 2. Twenty second digitized time sample of HF atmospheric noise.

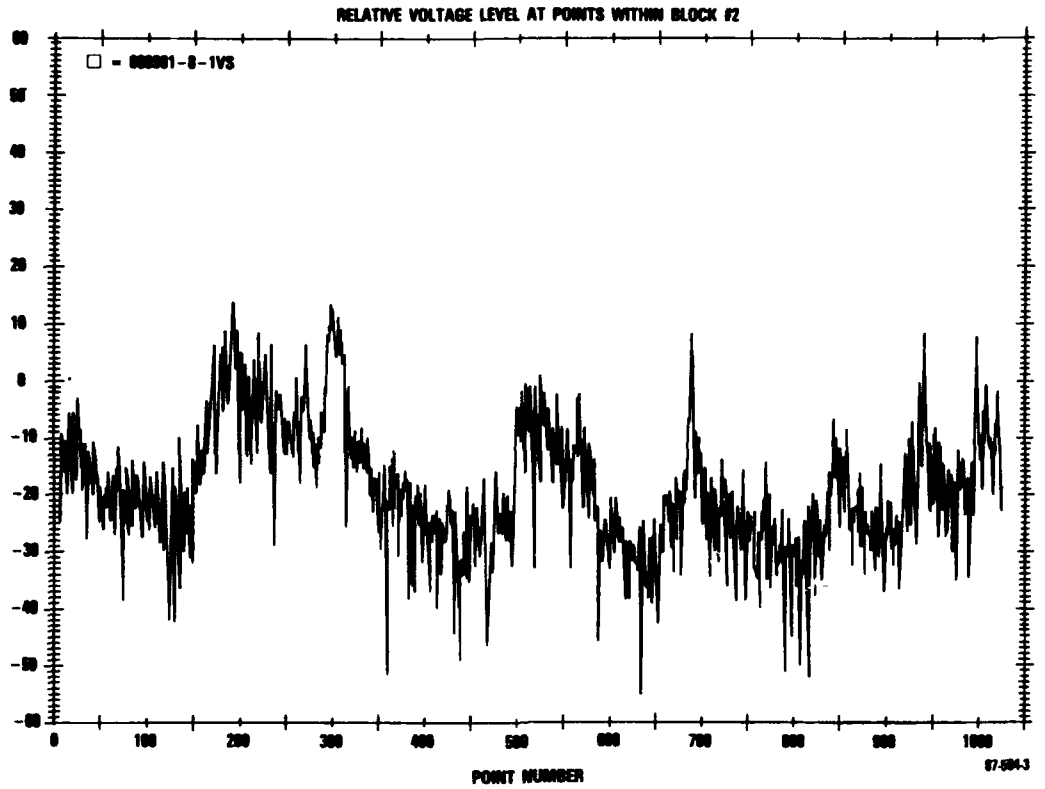


Figure 3. Eight millisecond time sample of block 2 of Figure 2.

The dynamic range here is approaching 80 dB, and it can be seen that the large peaks contribute most to the V_{rms} level of -5 dB relative to V_{rms} of the whole 20-s sample.

For several of the peaks, the rise time is less than the 8 μs resolution of the plot. But more importantly, from an interference viewpoint, the relatively broad peaks occurring in this sample would be sufficient to create a burst of errors in a digital signalling system.

This is not the case in the next 8-ms sample shown in Figure 4.

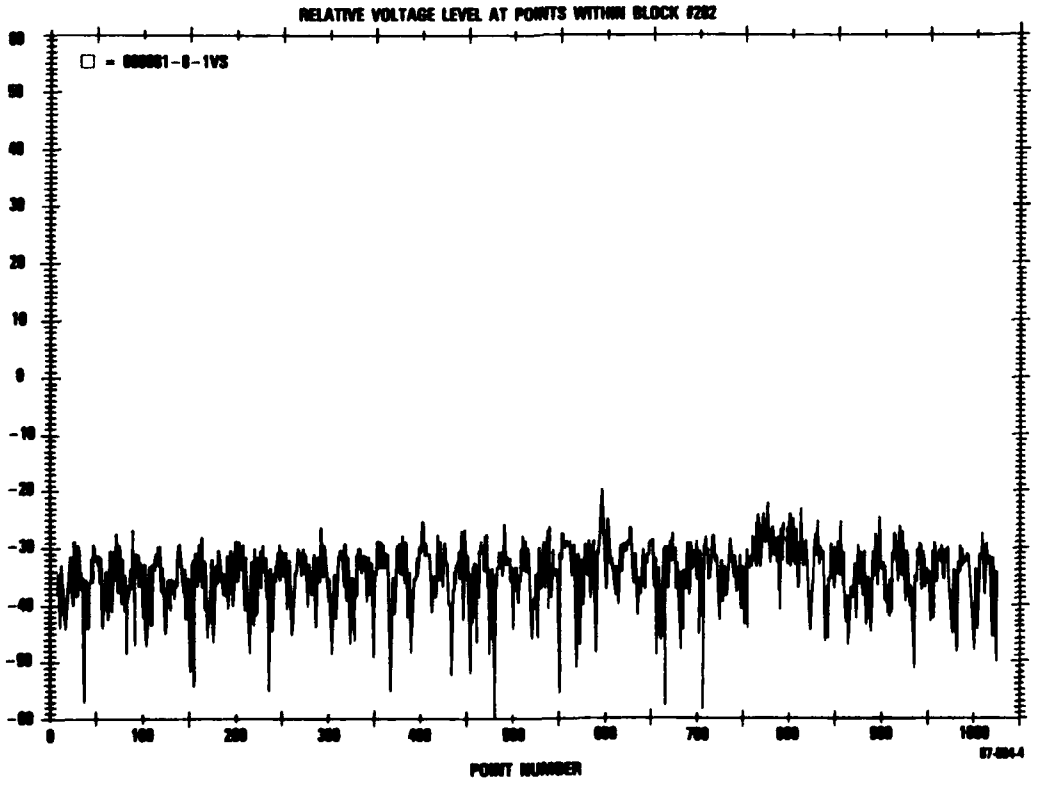


Figure 4. Eight millisecond time sample of block 282 of Figure 2.

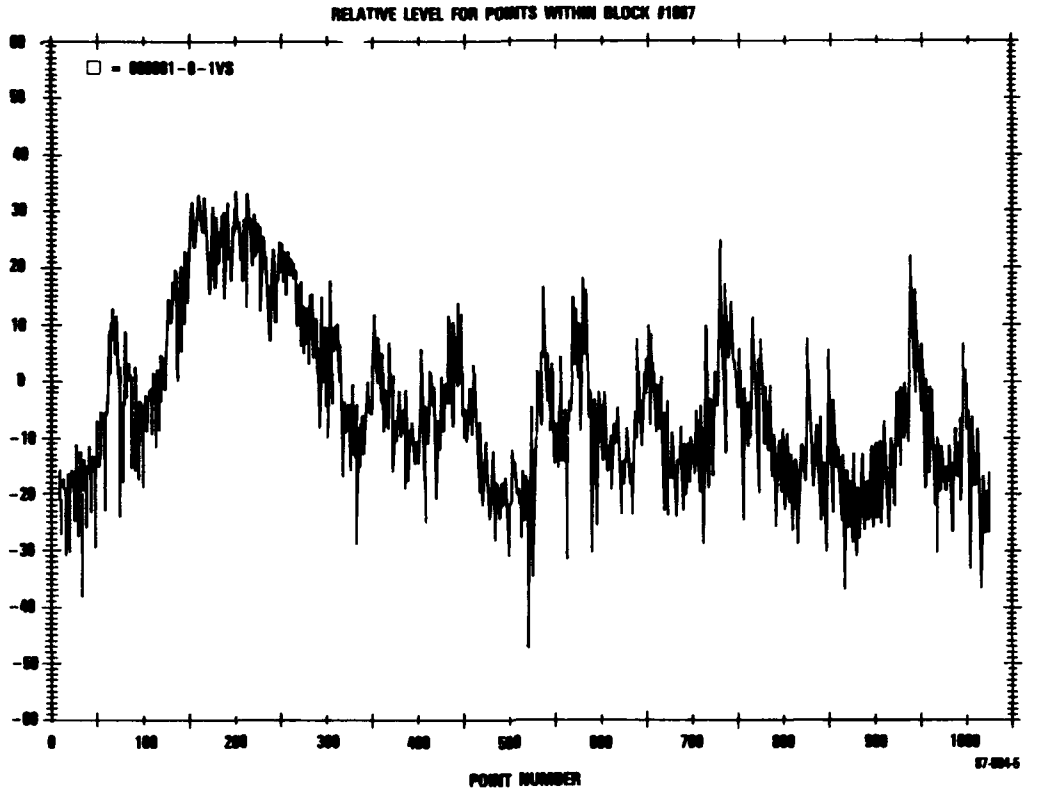


Figure 5. Eight millisecond time sample of block 1867 of Figure 2.

It will be recalled that the expanded data point (Figure 4) was one of the lowest in the 21.6 s sample shown in Figure 2. The structure you see in this plot is quite typical of Gaussian noise. The dynamic range is small, the overall level is low, and impulses occur rarely.

Digital signalling systems designed to operate in this kind of noise environment fare well when provided with an adequate signal-to-noise ratio (S/N). In this example, if the S/N is 12 dB, i.e., where the signal is at -18 dB and the rms noise level is at -30 dB on this relative scale, the transmission will be error free.

Figure 5 shows Block 1867 which is the third, high amplitude 8-ms data point shown in Figure 2. Again, the resolution is 8 μs, and now the dynamic range is approaching 80 dB. The large peak on the left (between point numbers 100 to 300) possibly represents the signature of an individual stroke from a lightning flash.

In this case, a 12-dB signal-to-noise ratio would be insufficient for error free transmission in a digital system. Error correction codes designed to correct single errors produced by independently occurring noise impulses (of the type seen in Figure 4) fail to account for a burst of errors that would arise from the wide noise peak shown here between point numbers 100-300.

Man-made noise from power lines, vehicular ignition systems, industrial electrical machinery and so on, is also non-Gaussian. A typical example of a 250 MHz analog recording taken in Colorado Springs is shown in Figure 6 (Spaulding and Disney, 1974, Ref. 2). The principal contributor of this noise is probably vehicular ignition systems. The bit errors in a digital system operating in this kind of noise would be a combination of single errors and burst (contiguous) errors due to the presence of both sharp spikes and wide pulses.

The noise samples shown in these figures are rather typical of variations in man-made and atmospheric noise. The noise process is stochastic rather than deterministic, so the impact on system performance must be approached from a statistical viewpoint.

4.0 STATISTICAL CHARACTERISTICS OF RADIO NOISE

4.1 DEFINITIONS

A complete description of the statistical characteristics of radio noise requires a number of parameters to assess the impact of noise on communication system performance (Spaulding, 1982, Ref. 3). Several of them are identified in Figure 7.

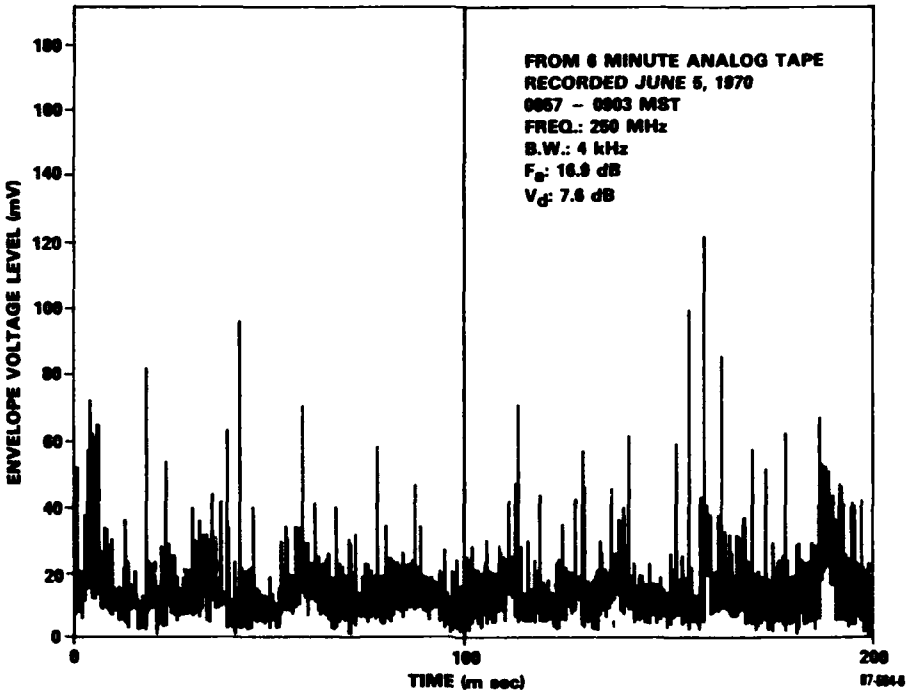


Figure 6. Randomly selected 200 ms sample of noise envelope from a 6-minute, 250 MHz central Colorado Springs, CO, recording.

The pulse duration varies both with the threshold level and with time at the same level. The characterization is generally made by analyzing a sufficiently long time sample of noise (at least several seconds) to obtain the "pulse duration distribution" PDD for a range of threshold levels. Normally -40 dB to +20 dB about the rms level is sufficient dynamic range according to the general consensus.

In like manner and for the same reason, the "pulse spacing distribution" PSD is derived for the same set of threshold values. Of interest also is the positive crossing rate of the noise impulses at the various threshold levels. This is generally termed the "average crossing rate", or ACR.

For some types of communication interference, e.g., television synchronization problems, the peak value of the noise is the most important parameter. A commonly used measurement instrument is a peak detector.

Implicitly contained in this sketch are several more statistical characteristics of value in system performance assessments which are commonly used as summarized by Herman (1979, Ref. 4).

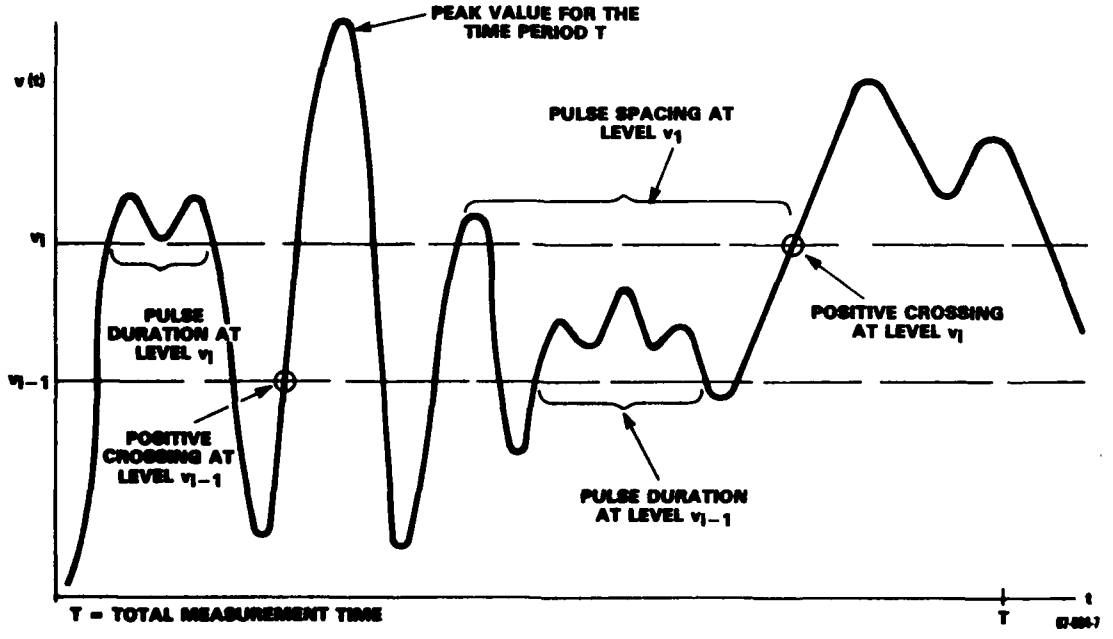


Figure 7. Noise envelope of a sample of man-made noise.

These are

1. Foremost, the amplitude probability distribution APD, which describes the fraction of time that the noise level is above specified levels;

2. The noise power factor, F_a, expressed in dB relative to $kT_o b$, which is a way of saying the power in watts per unit bandwidth relative to thermal noise kT_o where T_o = 288 K: F_a is proportional to the square of the rms voltage. The average power is always necessary to define the signal-to-noise ratio;

3. The rms (V_{rms}) and average (V_{ave}) voltages of the noise envelope;

4. The voltage deviation parameter V_d, defined as the ratio of the average to the rms envelope voltage; and

5. The voltage log deviation, L_d, defined as the ratio of the average of the logarithm of the envelope voltage to the rms voltage. L_d is especially useful for evaluating the performance of logarithmic detection devices.

The parameters V_d and L_d can be used to obtain the APD. The two deviation parameters are closely correlated, so that V_d can usually be regarded as a single-number representation of the APD curve.

4.2 AMPLITUDE PROBABILITY DISTRIBUTION (APD)

Figure 8 is taken from the new APD curves parametric in V_d for atmospheric radio noise

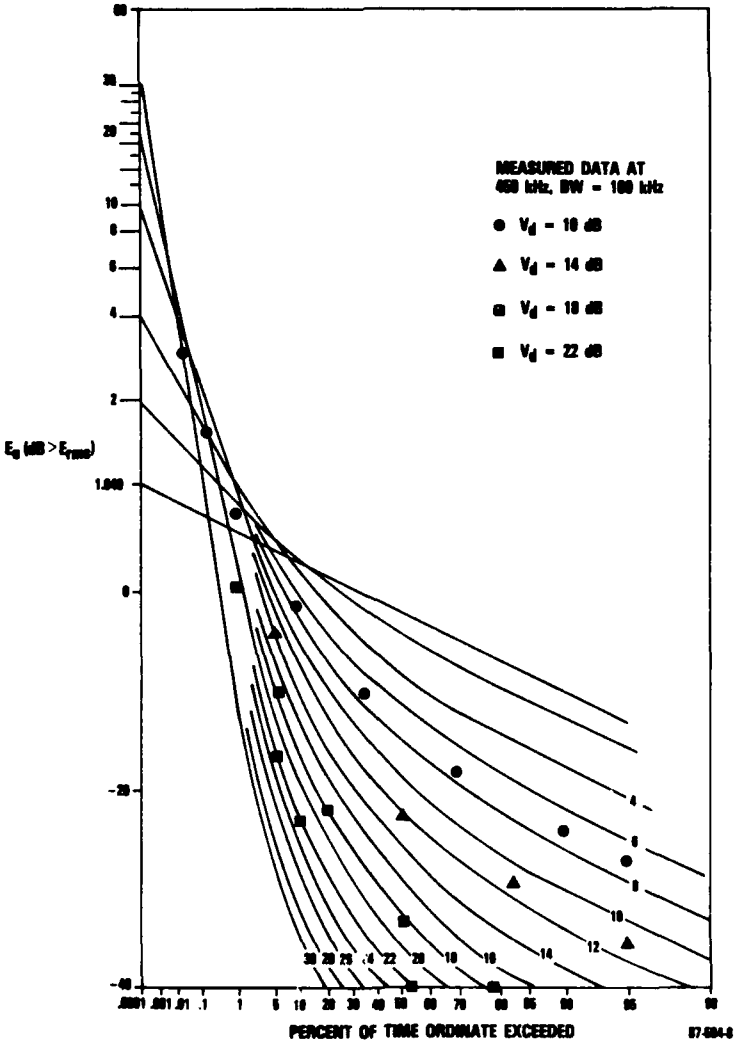


Figure 8. "New" set of amplitude probability distributions (APD) for atmospheric noise for various values of V_d.

reported by Spaulding and Washburn in 1986, as part of the proposed revision for CCIR Report 322. (CCIR, 1986, Ref. 5)

First, note that in the low amplitude, high probability portion, all the curves achieve or approach a slope of -2, consistent with the APD of Gaussian noise (as indicated by the straight curve with V_d=1.049).

Second, as V_d increases, indicating that the noise is becoming more impulsive, the Gaussian portion is suppressed to lower and lower levels.

Third, superimposed are some data points obtained from wideband MF noise recordings taken with the system mentioned in Section 1. As noted, the measured APD's are for separate noise samples with V_d = 10, 14, 18, and 22 dB. The measured data follow precisely the calculated APD's for V_d's of 22 and 18 dB, and although the form factor is preserved for the 10 dB measured points, they actually fit V_d = 7 dB better. This is not surprising, because the correlation between V_d and L_d is not quite as good in the midrange of V_d values (~ 10 dB), as it is for high and low values. The generally close correspondence between specific measurements and the calculated family of APD curves verifies the efficacy of using V_d to represent the APD of noise.

4.3 PULSE DURATION DISTRIBUTION (PDD)

Figure 9 shows a set of pulse duration distributions for 450-kHz atmospheric noise in a 100-kHz bandwidth, as derived from a noise sample with a V_d = 18.2 dB reported by Giordano et al. (1986, Ref. 6). The curves are parametric in threshold level relative to the rms voltage of the sample. In a general way, as anticipated from the waveform samples shown earlier, the durations get longer as the threshold level is decreased.

If the pulses were occurring independently with random widths, the plots would follow a Poisson distribution as indicated by the dashed straight line in Figure 9. The observations approach such a distribution on the short duration high probability side of the graph, but below the approximately 30% exceedance level, the durations are considerably longer. PDD plots for 48-MHz man made noise reported by Spaulding and Disney (1974, Ref. 2) show a similar behavior. For digital signalling systems, the consequence of long duration pulses is a string of consecutive bit errors, which are very difficult to correct using available error correction schemes.

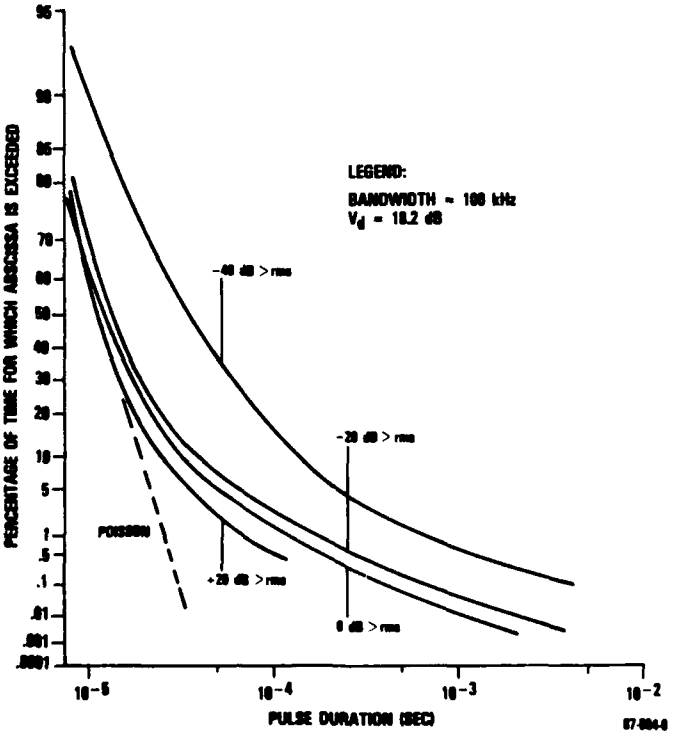


Figure 9. Median pulse distribution of a sample of MF radio noise.

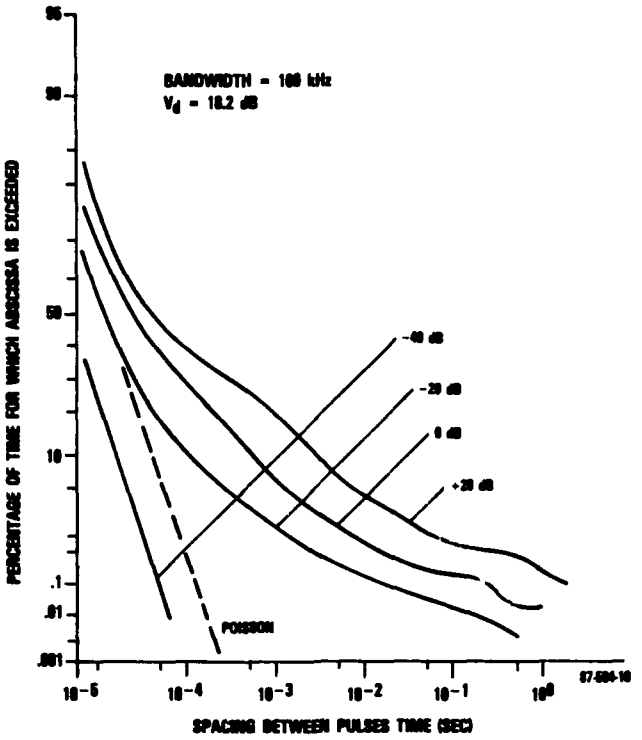


Figure 10. Median pulse spacing distribution of a sample of MF radio noise.

4.4 PULSE SPACING DISTRIBUTION (PSD)

A similar situation exists for the pulse spacing distributions of atmospheric and man-made noise. The set of PSD shown in Figure 10 are from the same MF noise sample (Figure 8), with the V_d = 18.2 dB representing a highly impulsive waveform. These data were reported by Giordano et al. (1986, Ref. 6). Again, Spaulding and Disney (Ref. 2) give similar plots for man-made noise. The general pattern here is for the distributions to shift toward greater spacings at the higher threshold levels, and only at the low threshold levels do the distributions approach a Poisson-like character.

In this example, the PSD's for the higher thresholds begin to depart from a Poisson distribution around the 30% exceedance level and show larger spacings than would be expected from random pulse occurrences.

The combined behavior of the PSD and PDD distributions describe atmospheric noise as a non-stationary noise process marked by relatively long duration high amplitude bursts of correlated noise pulses separated by relatively quiet levels where the noise character is Gaussian like.

For this reason, digital systems designed to operate in Gaussian noise, or in impulsive noise wherein the pulses are assumed to occur independentiy, will perform poorly in a correlated-pulse noise environment.

4.5 AVERAGE POSITIVE CROSSING RATE (ACR)

Now let us look at the situation in a slightly different way, by considering the average positive crossing rate (ACR), an example of which is shown in Figure 11 (adapted from Giordano et al., 1986, Ref. 6).

The data points are from a measured sample of 450-kHz atmospheric noise, 100-kHz bandwidth, and V_d of 7.6 dB. For comparison, the Hall model, which assumes independently occurring, random, infinitesimally-narrow noise pulses superimposed on a Gaussian noise background, has been used by Giordano et al. (1986, Ref. 6) to compute an ACR for the same conditions (BW = 100 kHz, V_d = 7.6 dB).

At the lower thresholds where the spaces between pulse occurrences approach a Poisson distribution, as noted in Figure 10, one would expect the number of positive crossings to behave likewise. In Figure 11 this is exactly the case.

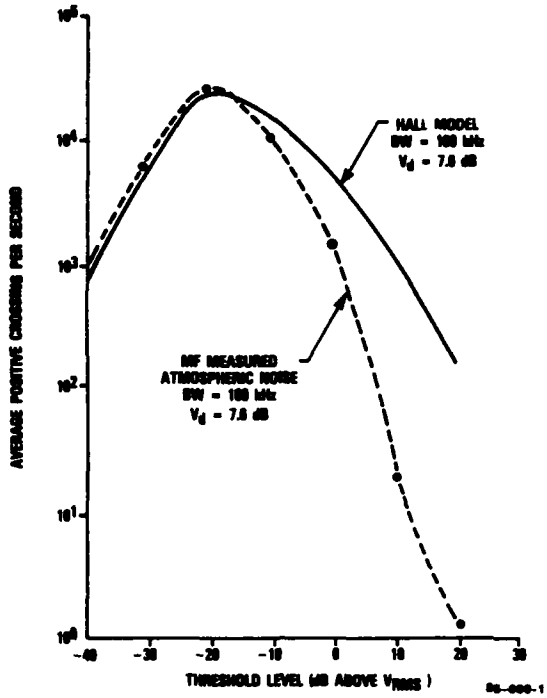


Figure 11. ACR for a sample of 450-kHz atmospheric noise.

At the higher thresholds, the actual number of crossings per second are fewer than predicted, because once the pulse level crosses a threshold it remains at the higher amplitude longer than would be anticipated by a simple Poisson approximation.

The impact on digital system performance remains the same. That is, the long duration high amplitude noise level would corrupt a succession of contiguous bits, and cause severe problems for an error correction code designed for single or double bit errors.

5.0 NOISE MITIGATION TECHNIQUES

A number of schemes have been developed over the years to mitigate the noise effects and improve the performance of radio communication systems. These have included the development of nonlinear receiver structures, and error detection and correction coding.

To illustrate, the simulated performance of nonlinear receivers in real and modeled noise will be described and then a bit error stream generator for testing error correction codes will be presented.

5.1 NONLINEAR RECEIVER STRUCTURES

The simulations described here are due principally to A. Giordano and F. Hsu, using the software approach shown in block form in Figure 12 (Giordano et al., 1986, Ref. 6). A transmitted information sequence is input to a BPSK generator whose N sample bits go to a mixer.

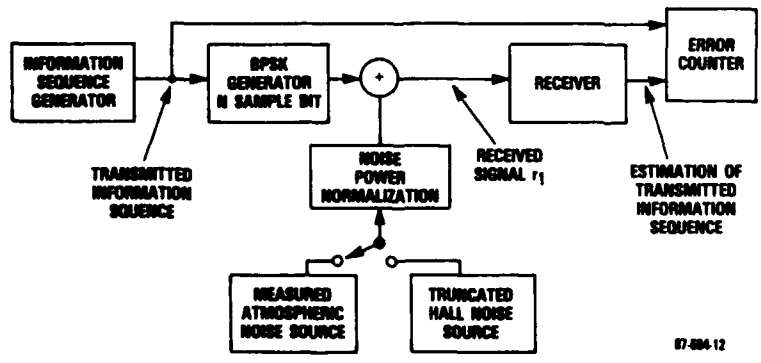


Figure 12. Simulation of error rate performance using measured and theoretical noise sources.

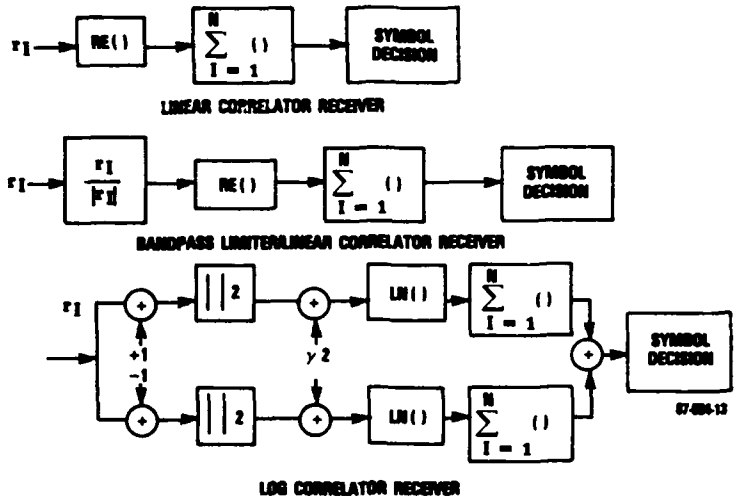


Figure 13. Receiver structures.

The noise source is selectable for measured atmospheric noise, truncated Hall model noise, or (not shown) a Gaussian noise model. The measured noise source is provided by our tape library of noise measurements obtained with the system described earlier. For input to the mixer, the noise waveform is normalized to provide a controllable signal-to-noise ratio in the mixer. The corrupted signal goes into a simulated receiver, and its output is analyzed for errors.

Simulations were made with a linear correlator receiver (Figure 13) to provide a basis for comparison, and two nonlinear receiver structures were analyzed. These were a bandpass/limiter (BPL) front end on a linear correlator receiver, and a log correlator receiver. In each case, the "symbol decision" examines each receive bit to determine whether or not it is in error and accumulates the number of errors in the received transmission.

Bit error rates (BER) were established for a variety of conditions, parametric in signal-to-noise ratio - more particularly for the ratio E_b/N_o, the energy per bit to noise spectral density. The conditions included a range of V_d values for impulsive noise, bandwidth-symbol time (BT) products, and measured or modeled noise.

Figure 14 illustrates the considerable improvement in BER predicted by modeled noise when nonlinear receivers are used. The linear receiver curve (X's) is typical of those seen in the literature. At a BER of 10^{-3}, use of a BPL or log correlator receiver would result in an apparent improvement of about 23 dB. That is, E_b/N_o = -8 dB rather than the +15 dB needed for BER = 10^{-3} in a linear receiver using a BT product equal to 10.

In measured noise the improvement is not so dramatic, as shown in Figure 15 taken from Giordano et al. (1986, Ref. 6). The linear receiver performs about the same (E_b/N_o = 16 dB for BER = 10^{-3}) as in the modeled noise case with the same V_d (about 10 dB); only minor differences arise due to the different choice of BT product (256 instead of 10).

The BPL receiver, though, does not perform nearly as well in the measured noise case compared to the modeled noise simulation shown earlier in Figure 14. It still gives better performance than the linear receiver in measured noise, but the improvement is 12 dB rather than 23 dB.

Introducing an interleaving technique produces several interesting results, as illustrated in Figure 16. The linear correlator results will be considered first.

With 1-s interleaving, the calculated BER performance in measured noise (x's) is better than the predicted performance in modeled (truncated Hall) noise at the higher error rates, but becomes worse at low error rates. Comparing the linear result here for truncated Hall noise with that shown in Figure 14, it is evident that interleaving introduces a 2-dB improvement in performance for BT = 10 and V_d = 10 dB.

In a similar manner, using interleaving with the BPL receiver in modeled noise provides about a 2-dB performance improvement (comparing the curves in Figures 14 and 16).

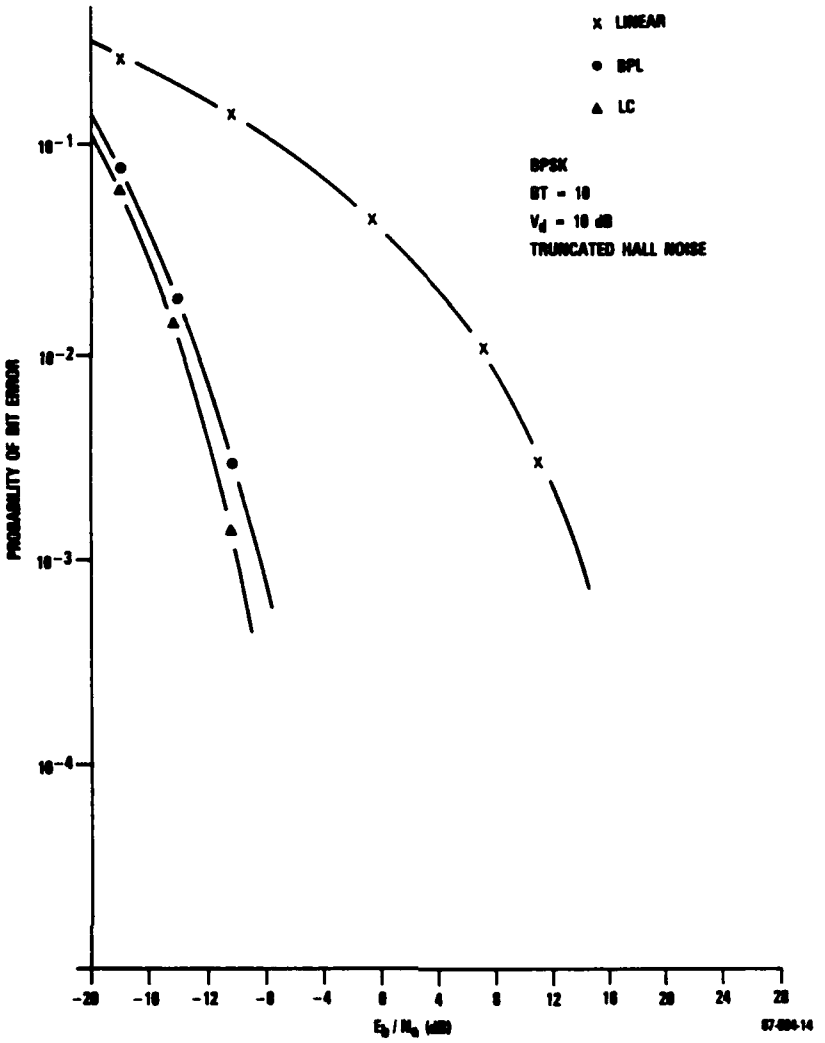


Figure 14. Graph of BER vs E_b/N_o for a linear, BPL, and log correlator receiver.

However, since real life atmospheric and man-made noise environments differ from even the best extant noise models, the measured noise result is of more practical importance. That is, the use of interleaving with a bandpass limiter front end provides a BER performance improvement of 20 dB or more over a linear receiver in real impulsive noise environments wherein the noise pulses are correlated (i.e., occur in bursts), rather than separate and independent.

5.2 ERROR DETECTION AND CORRECTION CODING

A large array of error detection and correction (EDAC) codes exist for mitigating gaussian noise effects (Michelson and Levesque, 1985, Ref. 7). Codes and algorithms that perform well in gaussian noise often fail in burst noise. To illustrate this point, let us examine the performance of Hamming codes in the presence of correlated-pulse impulsive noise vis a vis the Gaussian noise environment for which the codes were designed.

First, the method used to generate the bit error streams needed to evaluate the codes will be described followed by a few illustrative results. A functional flow of the simulation approach is shown in Figure 17.

Recorded (and modeled) noise samples are played into a transmitted bit stream to generate a corrupted message stream which is analyzed to obtain the error occurrence statistics, and these are stored in a data file as a function of specified noise parameters associated with the input noise.

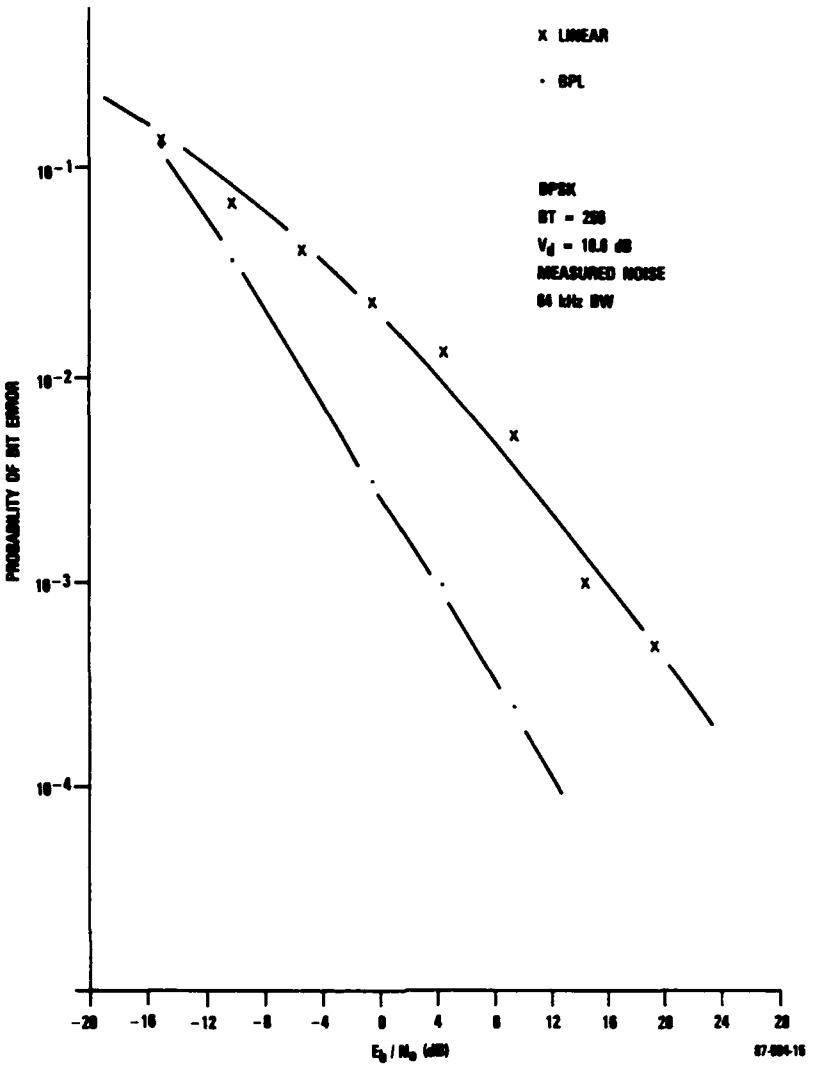


Figure 15. Graph of BER vs E_b/N_o for a linear and BPL receiver.

For the simulations, the run time inputs (communication parameters) are specified for the situation desired (V_d, E_b/N_o, data rate, BT). The appropriate error occurrence distributions (error spacings and durations) are then accessed in the data file and returned to the error sequence generator program. This "error engine" then generates a continuous binary data stream containing errors according to the occurrence pattern dictated by the error pulse distribution.

The bit error stream can then be input to any appropriate user device for evaluation. As noted earlier, the device to be treated is a Hamming code.

The data file was created as follows:

The simulation transmitter (Figure 18) can generate directly a binary bit stream with any specified sequence of zero's and ones, or it can accept a text message input which it converts to a binary stream at 7 bits per character.

Either BPSK or BMSK modulation can be selected; however, only the BPSK results will be discussed. The output signal S(t) goes to a channel simulator which mixes the signal with a selected noise input.

The noise-type choices are Gaussian, Hall 2 model or external. The external noise is provided by our tape library of recorded noise measurements to obtain a range of atmospheric noise environment conditions (e.g., F_a, V_d or APD, PSD, and PDD). It is preprocessed to achieve a range of signal-to-noise ratios (that is, E_b/N_o) with a variety of communication input parameters (e.g., data rate, BT).

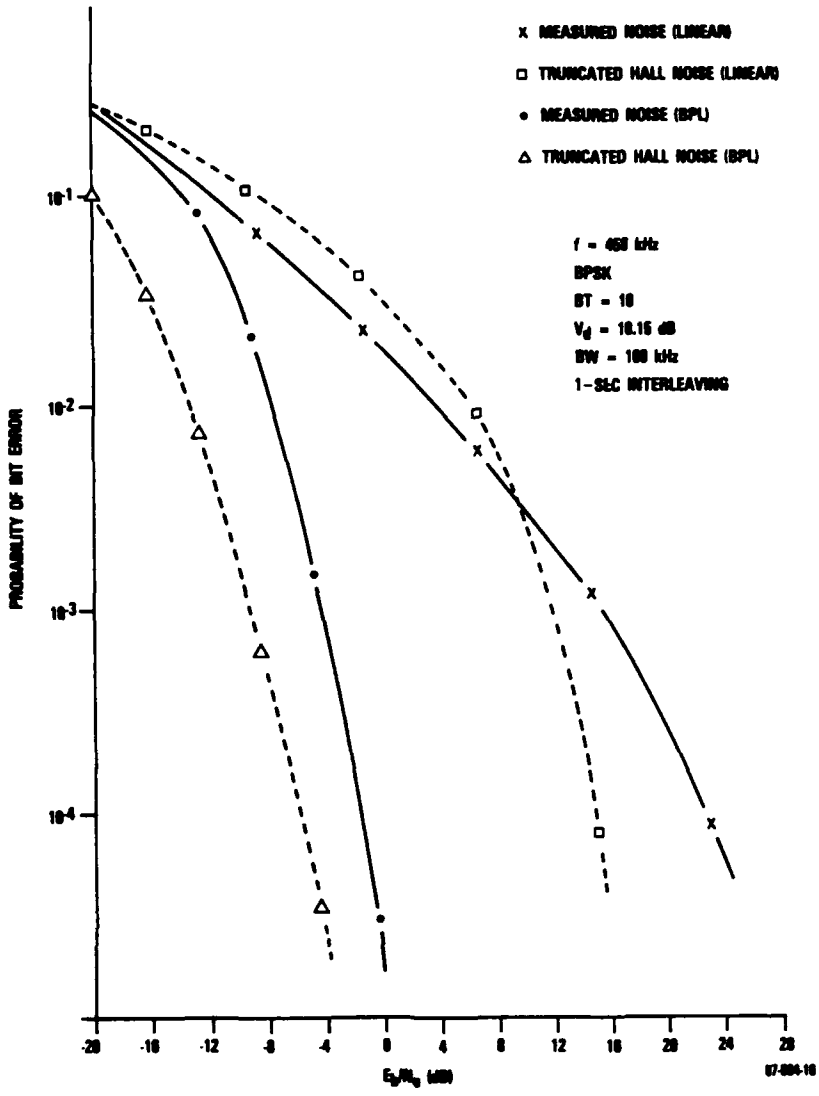


Figure 16. Graph of BER vs E_b/N_o for measured noise (linear), truncated Hall noise (linear), measured noise (BPL), and truncated Hall noise (BPL).

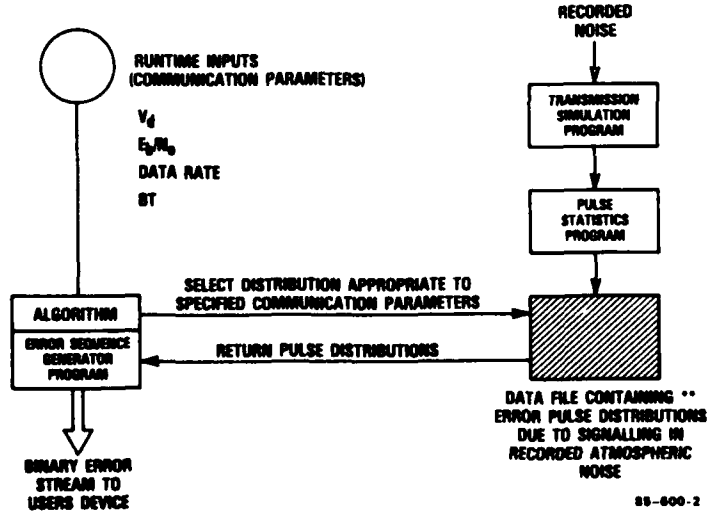


Figure 17. Functional diagram of simulation approach.

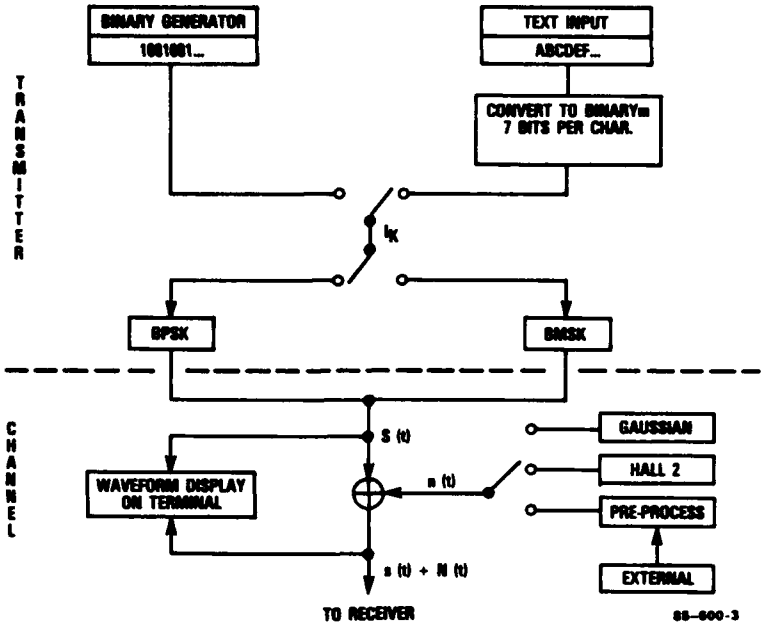


Figure 18. Functional block diagram of the simulation transmitter.

The corrupted signal S(t) + N(t) then goes to a simulated receiver.

The simulation receiver (Figure 19) demodulates the signal, writes to an output array, and compares the output to the input on a bit by bit basis. The error rate is calculated and the error sequence is written to a disk. From the comparator, one may also recreate the received stream and display a block of the corrupted signal.

Statistical analysis of the corrupted stream provides error spacing distributions (ESD) and error duration distributions (EDD). An example is shown in Figure 20. The measured atmospheric noise sample used to generate the bit stream on the left was moderately impulsive (V_d = 7 dB) so the longest error burst is only 3 bits for this particular example. Other runs with greater V_d's produced as many as 10 contiguous errors. By comparison (not shown), simulations using Gaussian noise generated bit error streams, wherein nearly all of the errors occurred singly and randomly, and the ESD, EDD statistics reflected that pattern.

A pair of ESD and EDD due to white Gaussian noise was selected to provide a bit error stream from the receiver output with an error rate of 0.054. The error occurrence pattern

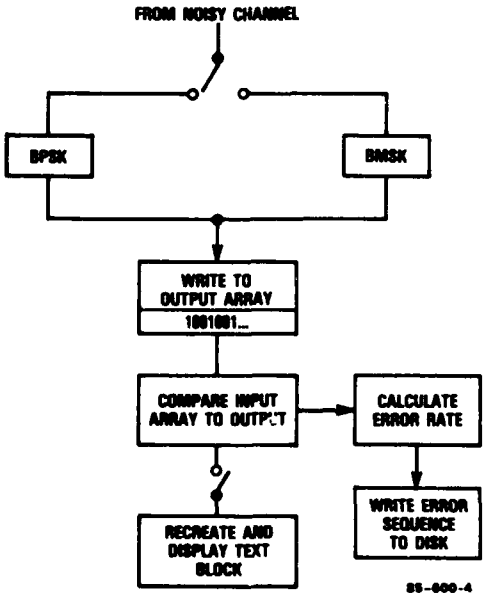


Figure 19. Functional block diagram of the simulation receiver.

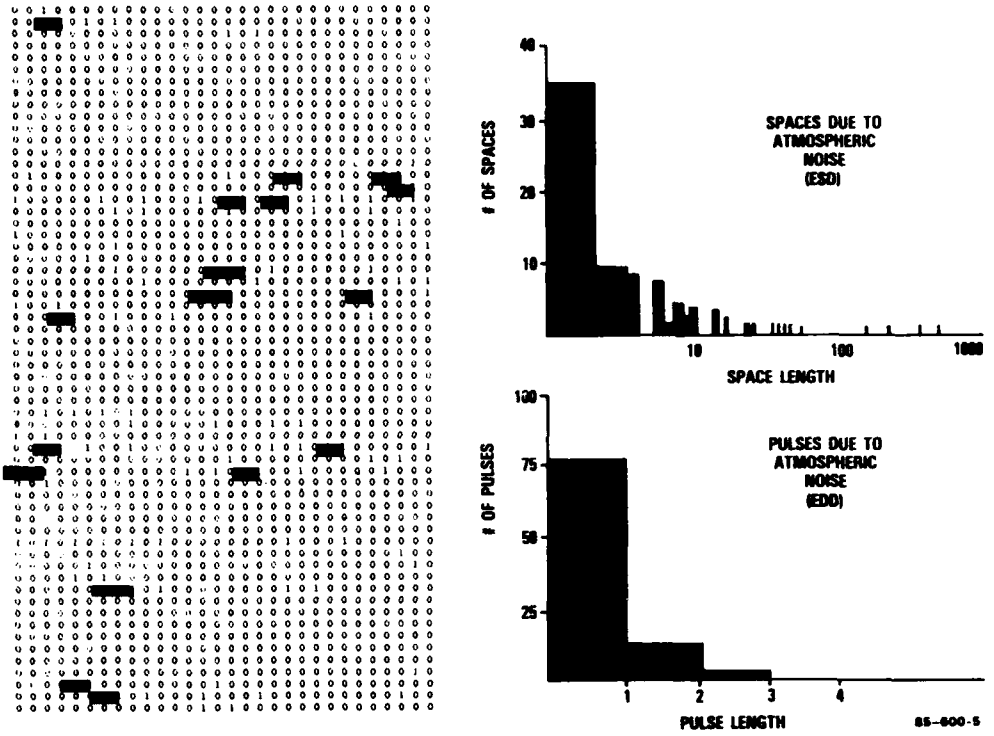


Figure 20. Example of a bit stream generated in atmospheric noise.

is shown in Figure 21. Here the zeroes representing correct bits have been suppressed, so that it is easy to see the error pattern represented by the ones. A few cases of double errors can be seen here. This stream was then played through a Hamming error detection/correction code.

The result of the Hamming decoded error pattern of the stream with error rate of 0.054 is shown in Figure 22. Note that most of the errors have been cleaned up, and the error rate is now only 0.00778, a factor of nearly 10 better than without the code. Thus, the Hamming code performs very well in Gaussian noise.

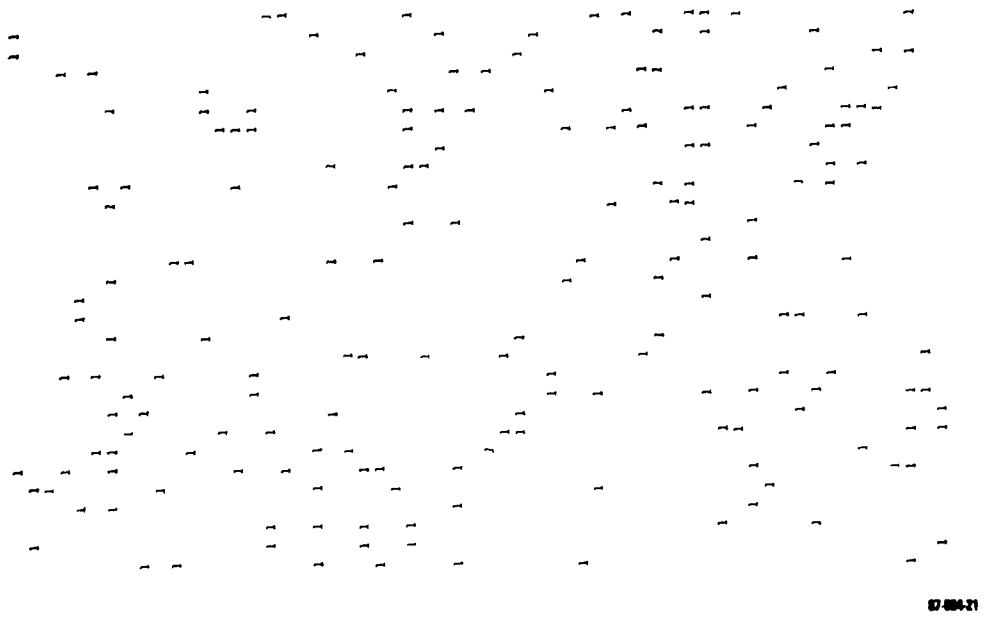


Figure 21. Error occurrence pattern for a white Gaussian noise with error rate of 0.054.

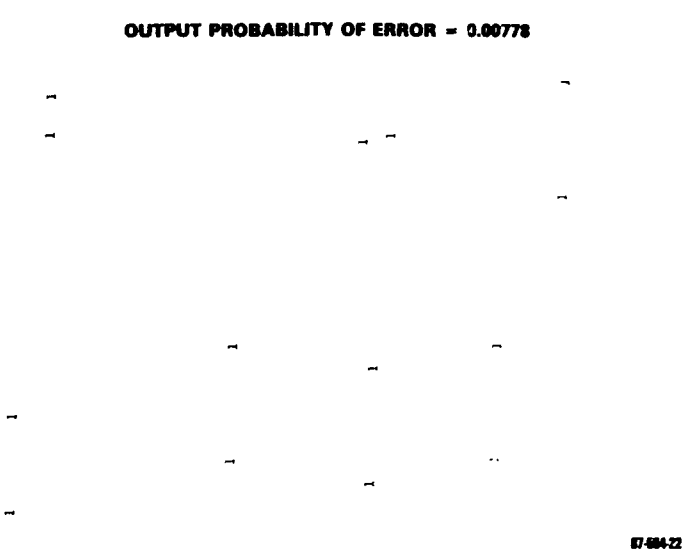


Figure 22. Error occurrence pattern after Hamming decoding the data stream of Figure 21. Input error occurrence pattern from a white Gaussian noise error rate of 0.054.

In a similar manner, a bit error stream with an error rate of 0.096 was generated in atmospheric noise for the example in Figure 23. Here it can be seen that many double errors occur, there are several error bursts of 4-bit lengths, and one near the bottom is 5 bits long. The result of playing this stream through the Hamming code is shown in Figure 24.

Compared to Figure 23, the error pattern shown here after Hamming decoding is scarcely improved. Though many of the single errors have been corrected, those strings of multiple errors still remain. The error rate, 0.08021, as shown at the top of the figure, is only slightly less than the 0.096 input error rate.

Another comparison, shown in Figures 25 and 26, further demonstrates the inability of this code to correct errors generated in a correlated-pulse, impulsive noise field. The input error rate, 0.133 is only slightly reduced to 0.1226 by the correction process.

The conclusion, obviously, is that an error correction code designed to function well in a Gaussian noise environment performs very poorly in the presence of impulsive, atmospheric noise.

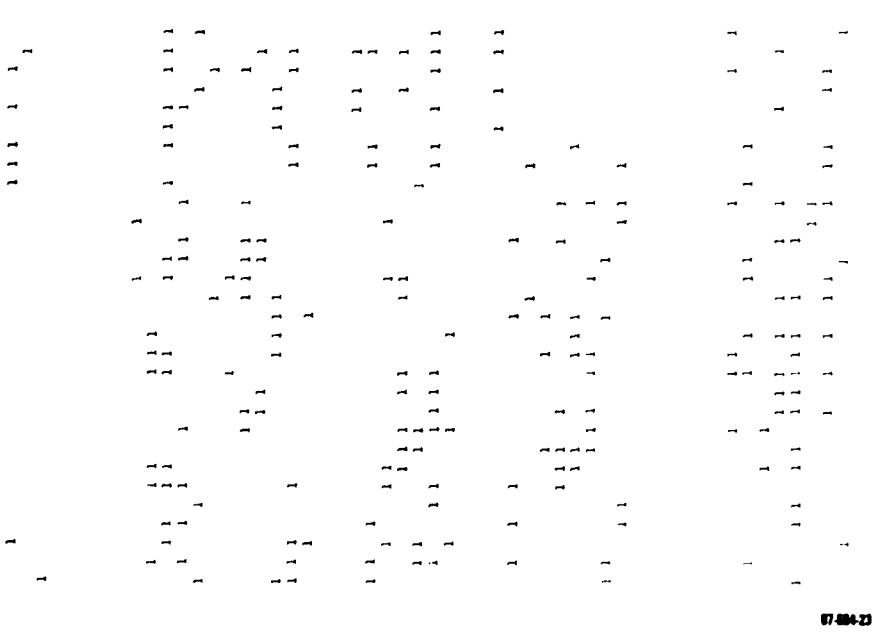


Figure 23. Error occurrence pattern with an error rate of 0.096 generated in atmospheric noise.

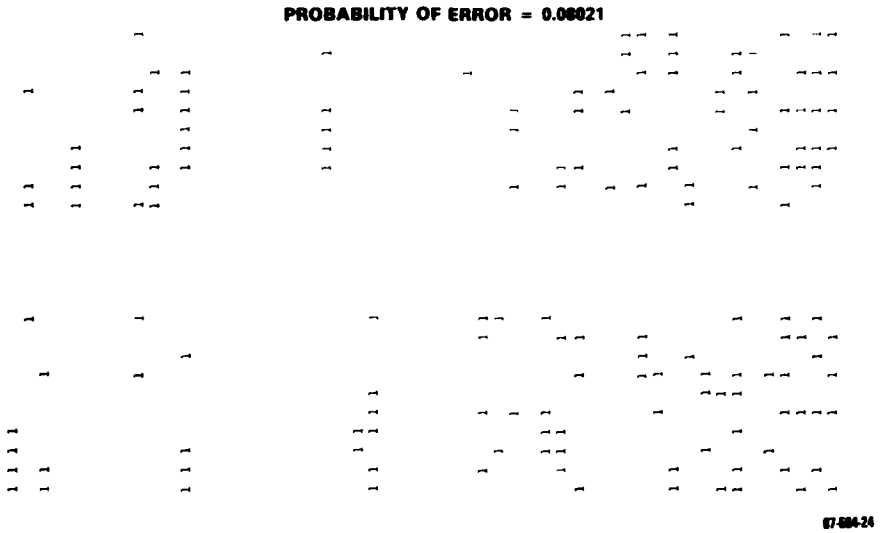


Figure 24. Error occurrence pattern after Hamming decoding the data stream of Figure 23.

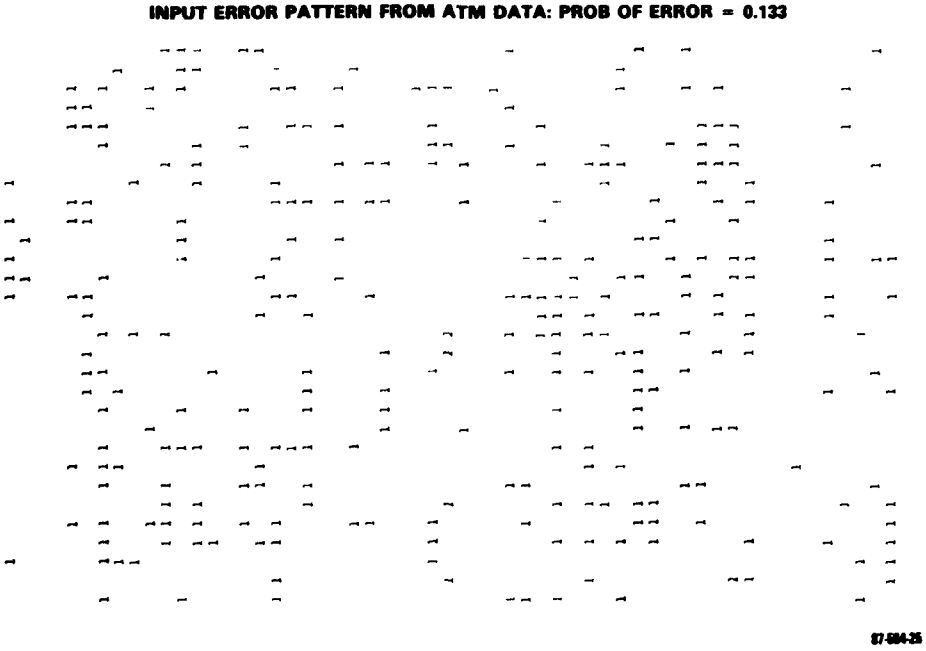


Figure 25. Error occurrence pattern with an error rate of 0.133 generated in atmospheric noise.

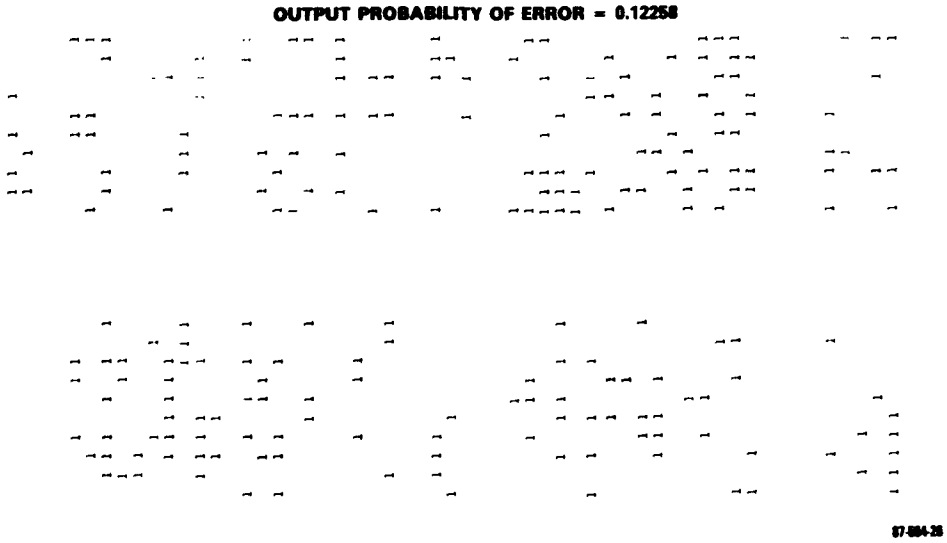


Figure 26. Error occurrence pattern after Hamming decoding the data stream of Figure 25.

6.0 CONCLUSION

This survey has reviewed the impacts of correlated-pulse impulsive noise on communication coding and system performance. To achieve communication system designs capable of meeting required performance specifications, the statistical characteristics of the noise environment must be taken into account.

7.0 REFERENCES

(1) Herman, J. R., A. A. Giordano, X. A. DeAngelis, K. F. Marzotto, and F. M. Hsu, Measurement and statistical analysis of wideband MF atmospheric radio noise, 1. Structure and distribution and time variation of noise power, Radio Science, Vol. 21, No. 1, 1986, pp. 25-46.

(2) Spaulding, A. D. and R. T. Disney, Man-made radio noise, Part I: Estimates for business, residential and rural areas, Report No. OT 74-38, U.S. Government Printing Office, Washington, DC, 1974.

(3) Spaulding, A. D., "Atmospheric noise and its effects on telecommunication systems performance," Handbook of Atmospherics, Vol. I, (H. Volland, ed.), CRC Press, Boca Raton, FL, 1982, pp. 289-328.

(4) Herman, J. R., Electromagnetic Ambients and Man-Made Noise, Don White Consultants, Inc., Gainesville, VA, 1979.

(5) CCIR, Characteristics and applications of atmospheric radio noise, Report No. 322-3, International Telecommunication Union, Geneva, 1986.

(6) Giordano, A. A., J. R. Herman, X. A. DeAngelis, K. F. Marzotto, and F. M. Hsu, Measurement and statistical analysis of wideband MF atmospheric radio noise, 2. Impact of data on bandwidth and system performance, Radio Science, Vol. 21, No. 2, 1986, pp. 203-222.

(7) Michelson, A. M. and A. H. Levesque, Error-Control Techniques for Digital Communication, John Wiley and Sons, New York, 1985.

BALLOON MEASUREMENTS OF LOW FREQUENCY TE AND TM ATMOSPHERIC NOISE

John P. Turtle
Propagation Branch
Rome Air Development Center
Hanscom AFB, MA 01731, USA

and

Paul A. Kossey
Ionospheric Interactions Branch
Air Force Geophysics Laboratory
Hanscom AFB, MA 01731, USA

SUMMARY

This paper describes measurements of Transverse Electric (TE) and Transverse Magnetic (TM) atmospheric noise, obtained during a five-hour flight of a free-floating balloon. The balloon carried three low frequency receivers and three (orthogonal) loop antennas to measure TE and TM atmospheric noise at 42.5 kHz, and the normal (TM) and converted (TE) signals from ground-based transmitters at 37.2 kHz, 48.5 kHz and 60 kHz. A description of the low frequency instrumentation is given, along with discussions of the TE and TM data acquired during the flight, most of which (four hours) occurred at a free-floating altitude of about 21 km. Emphasis is on the TE-to-TM signal and noise ratios observed during the flight. The use of such data to validate TE and TM atmospheric noise prediction codes is also discussed.

1. INTRODUCTION

Very Low Frequency (VLF) and Low Frequency (LF) waves propagate to great distances in a waveguide bounded by the earth and the lowest regions of the ionosphere. To the extent that the waveguide can be considered to be flat for VLF/LF waves, the distribution of field intensity with altitude may be thought of as resulting from the interference of elementary up-going and down-going plane waves reflecting from the boundaries at certain oblique (modal) incidence angles. At these angles the phase changes due to reflection are such that the up-going and down-going plane waves are consistent with each other. Owing to the boundary conditions associated with propagating waves having a Transverse Magnetic (TM) polarization or a Transverse Electric (TE) polarization (Figure 1a), the modal structures for TM and TE waves would tend to be the forms shown in Figure 1b. These indicate, for example, that for efficient coupling to TM modes, the source element should be vertical and preferably close to the ground; whereas, for efficient coupling to TE modes, the source element should be horizontal and at a high altitude. Similarly, the observation of TE waves would require receiver probes at altitudes amounting to an appreciable fraction of the distance between the ground and the ionosphere [1, 2 and 3].

Extensive studies of TM atmospheric noise have been conducted over the past thirty years, including attempts to characterize it on a global basis [e.g., 4 and 5]. At present there is an extensive ground-based measurements program being conducted to obtain a large variety and quantity of TM atmospheric noise data over the band of a few Hz to about 40 kHz [6]. To date, however, there has been little activity in the area of characterizing TE/TM atmospheric-noise, due (in part) to the relative difficulty in obtaining TE data, which requires a highly elevated, electrically quiet, observation platform. This paper describes TE/TM noise data recently obtained by the Rome Air Development Center (RADC) of the U. S. Air Force, using a balloon-borne technique designed to overcome some of the difficulties associated with measuring and characterizing TE atmospheric noise.

2. EARLY TE ATMOSPHERIC NOISE STUDIES

Gallenberger and Bickel [7] attempted to measure TE atmospheric noise up to altitudes of about 6 km, using an instrumented aircraft. Although they were able to obtain TM atmospheric noise data that were consistent with predicted noise levels, they were unable to obtain direct measurements of TE atmospheric noise. They did infer from the data, however, that the TE atmospheric noise at 15 kHz, in the altitude range from about 0.8 to 6 km, was 20 to 25 dB below the TM noise, for both day and night. They also noted two problems that point out some of the difficulties associated with such aircraft measurements: (1) if the TE noise is much weaker than the TM noise it is critical that the TE antennas be horizontal to avoid TM contamination; and, (2) aircraft generated noise may be much greater than the atmospheric noise to be measured. Hirst [8] encountered similar problems in a program to measure TE and TM signals at altitudes of 18 km in a U-2 aircraft, which is considered to be relatively quiet, electrically.

Harrison et al. [9] conducted measurements of TE/TM signals and noise, using a dart-like package instrumented with TE and TM antennas and receivers. The package was carried to an altitude of about 22 km by a meteorological balloon, and then released. During descent the axis of the probe was aerodynamically stabilized to within about one-degree so that the plane of the TE receiving antenna remained essentially horizontal. Although telemetry noise placed a limit on the level of the atmospheric noise that could be observed, it was concluded that the TE daytime atmospheric noise, at about 44 kHz, was at least 6 dB below the TM noise, even at 22 km. In later experiments, using more sensitive

equipment, the daytime TE atmospheric noise was estimated to be at least 12 dB below the corresponding TM noise at similar altitudes.

In a daytime flight of a TE- and TM-instrumented rocket Harrison et al. [10] obtained noise polarization ratios for 293 individual sferics. At 30 kHz the TE/TM noise ratio was -15 dB near the surface of the earth, but decreased with increasing altitude to about 0 dB at 60 km.

To date, TE and TM atmospheric-noise prediction techniques have been developed only to a very limited extent, primarily due to the lack of TE data required to validate them. In recent assessments of the potential performance of candidate airborne VLF/LF communication systems, estimates of TE atmospheric noise at 9 km altitude were made using the simple empirical formulas given below [11].

$$\text{TE Noise (dB)} = \text{TM Noise (dB)} \begin{cases} + 21.59 \log\{F(\text{kHz})/17\} - 31 & \text{(Day)} \\ + 32.56 \log\{F(\text{kHz})/17\} - 17.5 & \text{(Night)} \end{cases}$$

At 42.5 kHz, for example, these reduce simply to

$$\text{TE Noise (dB)} = \text{TM Noise (dB)} \begin{cases} - 22.4 & \text{(Day)} \\ - 4.54 & \text{(Night)} \end{cases}$$

Thus, at 42.5 kHz, the suggested nighttime TE/TM noise ratio at an altitude of 9 km is almost 18 dB larger than the daytime value. The degree to which these empirical results can be explained by polarization rotation effects in the ionosphere, which are much more significant at night than in the day, requires further investigation. The empirical expressions given above were developed from a small number of aircraft measurements of TE and TM atmospheric noise, at only two frequencies, 17 kHz and 29 kHz. As such, they employ extrapolations which require further validation, in accordance with appropriate experimental data.

Kelly [12] has developed a computer program for predicting horizontally and vertically polarized noise fields at any altitude in the earth-ionosphere waveguide, when the vertical electric fields (TM) in the propagating modes at the earth's surface are given. The technique covers the 10 to 30 kHz band, and uses the output of a previously developed program [5] to determine the required surface TM noise fields. Only contributions from vertically oriented noise sources are considered; i.e., the approach does not include possible contributions from horizontal components of lightning discharges. A further possible limitation of the model is that it does not compute TE noise fields that may result from the polarization conversion associated with the ionospheric propagation of the noise from the assumed, purely vertical, discharges. The lack of TE atmospheric noise data makes it impossible to adequately test such noise prediction models at this time.

3. RADC BALLOON-BORNE TE/TM ATMOSPHERIC NOISE PROGRAM

The Propagation Branch of RADC, in a joint program with the Defense Nuclear Agency and the Air Force Electronic System Division, has recently begun a program to obtain measurements of TE and TM atmospheric noise. The program, named ALFAN (Airborne Low Frequency Atmospheric Noise), is designed to obtain data at high altitudes, using a free-floating balloon to provide an electrically-quiet platform for the experiments.

ALFAN Measurement System

The ALFAN measurement system consists of three identical low frequency receivers and an octahedron-shaped antenna arrangement, made up of three loop antennas, orthogonal to each other. Each of the loop antennas are 1.73 meters on a side. In level flight (see Figure 2) two vertical loop antennas respond to horizontal magnetic fields, while a horizontal loop antenna responds to vertical magnetic fields. For low frequencies and at altitudes above about two kilometers, theoretical considerations show that these correspond essentially to the TM and TE fields to be measured, respectively. A block diagram of the receiver is given in Figure 3 and its characteristics are summarized in Table 1. The receiver is integrated with a digital sampling system and a downlink data telemetry system. In addition, an uplink command system is employed to allow various parameters of the receiver (including the frequency to be monitored) to be changed, via commands from the ground, during the flight of the balloon. The receiver, control and telemetry systems and the battery power supply are mounted on a horizontal "load-bar", with the antenna system suspended beneath the bar (see Figure 2). The payload (including antennas) weighs about 386 kg, and is connected to a recovery parachute having a diameter of about 18.3 meters. This is, in turn, connected to the free-floating balloon, which has a capacity of about 10,000 cubic meters. A block diagram of the balloon-borne and ground-based systems is given in Figure 4.

TABLE 1
ALFAN Receiver Characteristics

Frequency Range	30 kHz - 60 kHz
Bandwidth	1.5 kHz
Remote Tuning Capability	100 Hz steps
Antenna/Preamp Effective Height	10 meters
Amplifier Gains	0 dB, 20 dB, 40 dB, 60 dB
Sampling Rate	8 kHz
A/D Conversion	12 bit
In-Flight Calibration	1 mv/m RMS

Data Reduction Procedures

Data are recorded on high-speed, pulse-code modulated (PCM), telemetry tapes for later analyses using both analog and digital processing techniques, as outlined below. A diagram of the steps involved in the reduction of the data is given in Figure 5.

Analog Processing. For analog processing the telemetry tapes are first read through a PCM decommutator to retrieve the digital data. After undergoing digital-to-analog conversion the data passes through a RMS-to-DC converter and are recorded, using a strip chart recorder with individual channels allocated for the data from each of the balloon-borne receivers (i.e, one TE channel and two TM channels). The focus of the analog processing is on displaying signal data; specifically, the normal (TM) and ionospherically rotated, or converted (TE), signals from the ground-based (TM) low frequency transmitters that are monitored during the flights of the balloon.

Digital Processing. For digital processing of the data the telemetry tapes are first decommutated and converted from a serial to a parallel format. The data rate for these "new" digital tapes is the same as the sampling rate of the original data. Computer plots of various flight parameters, including altitude and antenna orientation (in azimuth), are then obtained from these tapes, along with plots of the RMS values of the TE and TM fields observed with the three balloon-borne receivers.
The focus of the digital processing is on displaying the noise data that are acquired during the balloon flights.

4. AUGUST 1986 ALFAN MEASUREMENTS OF TE/TM SIGNALS AND NOISE

The first ALFAN flight took place on 4 August 1986, under the direction of the Aerospace Instrumentation Division of the Air Force Geophysics Laboratory. The main purpose of the flight was to test the experimental procedures and the newly developed balloon-borne instruments. The balloon was launched at 1330 UT (0730 local time) from Roswell, New Mexico, and it carried the payload to an altitude of about 21 km, where it floated for about 5 hours. The flight was terminated at 1845 UT, and the payload parachuted safely to the ground in the vicinity of Alamogordo, New Mexico, approximately 121 km to the west of the launch site. The altitude of the balloon during the first 3.5 hours of the flight is given in Figure 6. Throughout the entire flight the weather was clear in the vicinity of the balloon. In addition, satellite photographs of cloud coverage and other meteorological records for the western half of the United States indicate that there was no thunderstorm activity within about 645 km of the balloon during the flight. The nearest thunderstorm activity was in central Texas and eastern Nebraska, and occurred during the latter portions of the flight. During portions of the flight TE and TM atmospheric noise data were acquired at 42.5 kHz; at other times the receiver was tuned to observe the normal (TM) and converted (TE) signals from ground-based TM-transmitters at 37.2 kHz, 48.5 kHz and 60 kHz.

TE/TM Signals From Ground-Based Sources

An important portion of the data analysis was the determination of the attitudes of the TE and TM antennas during the flight of the balloon. This knowledge was especially true with regards to the TE antenna, since even very small deviations of it from the "horizontal" due to swinging of the suspended antenna structure, could result in contamination of the data from the TM fields. Although this first ALFAN flight did not have any direct means of determining whether the antenna structure was swinging, it was possible to infer it from the processed data. Specifically, if the antenna structure was swinging, i.e., "oscillating", the TE data would, correspondingly, exhibit oscillations in accordance with the varying amount of TM contamination of the data. Figures 7-10 illustrate this effect.

Shown in Figures 7-10 are RMS signals from a ground-based, 48.5 kHz, transmitter located at Silver Creek, Nebraska, approximately 1125 km from the launch site of the balloon. The observation times correspond to just before and after launch (Fig. 7), at 2.4 km (Fig. 8), at 15 km (Fig. 9) and at the free-floating altitude of 21 km (Fig. 10). Each Figure gives observed output voltages (y-axes) from the receivers as a function of time (x-axes, UT). Also, in each Figure the top trace, designated "TE", gives the output from the horizontally oriented loop antenna, while the lower two traces, denoted as "TM1" and "TM2", give the outputs from the two vertically oriented loop antennas. Finally, because the TE signals were much weaker than the corresponding TM signals, the "TE" data were multiplied by a factor of 5x before being displayed.

Just before launch the antenna structure was swinging slightly in the wind, and accordingly, the TE signal (Fig. 7) shows some oscillations. Just after launch, however, the antenna structure was swinging rather wildly, as evidenced by the large scale variations in the data from the TE channel (Fig. 8). Conversely, the data from the two TM channels are not affected by the swinging antennas; rather, the more gradual variations observed in the TM data were caused by a slow turning of the payload and antenna structure about the vertical axis. The sensitivity pattern of each of the TM loop antennas (a figure-of-eight) causes the TM fields to go from a maximum to a minimum, as the vertically oriented loop antenna turns with respect to the azimuthal direction of the ground-based TM transmitter. Since the signals from the two orthogonal TM antennas are not in phase, a maximum signal from one of them corresponds to a minimum from the other. Conversely, the horizontally oriented TE antenna is not sensitive to such "turning" effects.

During ascent the antenna was swinging, as indicated by the oscillations seen in the TE signal traces in Figures 7, 8 and 9. At the free-floating altitude of about 21 km, which was above the tropopause (about 18 km for this flight), the TE antenna was essentially horizontal, as evidenced by the absence of oscillations in the TE signals (Fig. 10). This indicates that, at float, the balloon was drifting horizontally at the same speed as the wind around it, with very little turbulence so that the payload and the antenna structure remained essentially motionless with respect to their surroundings.

As indicated above, the TE channel was "contaminated" by leakage of the TM signals into the horizonatal loop antenna during the ascent of the balloon. Nevertheless, it is possible to estimate the true (uncontaminated) TE signal levels from the traces of Figures 7, 8 and 9 for they must lie somewhere between the maximums and the minimums of the oscillations that were observed. After accounting for the 5x gain-factor associated with the data from the TE channel, it is possible to describe qualitatively some of the "pure" TE and TM signal characteristics that were observed during the ascent phase of the flight. Specifically, throughout ascent the amplitudes of the TE signals increased with increasing altitude, whereas the amplitudes of the TM signals decreased. The behavior of the TE and TM signals from Silver Creek, during ascent, are summarized in Table 2. Of particular note is the significant increase in the ratio of the TE and TM signals that were observed as the altitude of the balloon increased. As indicated in Table 2, the TE-to-TM ratio increased by 28 dB over the 0-21 km altitude range of the flight.

A very large part of the increase in the TE signals occurred rapidly, after the balloon was released and it was still very near the earth. The explanation for this may

TABLE 2
Silver Creek (48.5 kHz) Signal Vs Altitude Data

Altitude (km)	Relative RMS Amplitude TE	TM	TE/TM (dB)
0.0	4	350	-39
0.9	38	350	-19
2.4	44	330	-19
15.2	70	310	-13
20.7	77	270	-11

be due, in part, to interference effects associated with the horizontal antenna and its image when it was within one or two loop diameters of the ground. However, this requires further investigation.

After the initial abrupt increase that occurred near the surface of the earth, the amplitude of the TE signal increased in a much more gradual manner, as the balloon rose to its free-floating altitude of 21 km. In general, the TE and TM amplitude-altitude trends that were observed are similar to those associated with the modal structures illustrated in Figure 1, which result from theoretical considerations of the propagation of low frequency waves in the earth-ionosphere waveguide.

During portions of the time that the balloon was floating at 21 km altitude, the signals from three ground-based (TM) transmitters were monitored. Table 3 gives TE/TM signal ratios which are typical of those obtained during that period of the flight. The data provide a measure of polarization conversion effects in the ionosphere.

TABLE 3
TE/TM Signal Ratios From Ground-Based Transmitters

Transmitter	Distance	Frequency	TE/TM (21 Km)
Hawes, Cal.	1125 Km	37.2 kHz	- 3 dB
Silver Creek, Neb.	1125 Km	48.5 kHz	- 11 dB
Fort Collins, Col.	825 Km	60.0 kHz	- 17 dB

TE and TM Atmospheric Noise at 42.5 kHz

Figure 11 shows RMS noise data at 42.5 kHz that were derived from data obtained over a 3.5 hour period of the flight. The RMS values for the TM data were obtained after first squaring the fields sensed by the two orthogonal TM loop antennas and then summing them, to provide a measure of the omni-directional TM noise. These can then be compared to the noise fields sensed by the single TE loop antenna which, by its horizontal orientation, already had an omni-directional receive pattern. Each of the RMS values in Figure 11 correspond to a minute's worth of data (about 500,000 samples of data). In the Figure, the y-axis is linear, and is in relative units. The x-axis is time (UT). The gaps in the noise data represent periods during which the receivers were tuned to monitor signals from ground-based transmsitters, as described above.

From about 1330 UT to 1430 UT, which corresponds to a period when the balloon was still ascending, the 42.5 kHz TM noise decreased whereas the TE noise increased. After about 1430 UT, when the balloon was floating at an altitude of 21 km, the TE noise was relatively constant, with only small fluctuations present, as seen in the data of Figure 11. The TM data showed larger fluctuations, although they occurred about a relatively constant level, as well.

Figure 12 and Figure 13 give TE-to-TM noise ratios (in dB), plotted as a function of time and altitude, respectively, as derived from the data in Figure 11. In addition, Table 4 provides a summary of the altitude data of Figure 13. The data indicate that the TE-to-TM noise ratio became significantly larger as the altitude of the observations increased. For example, over the altitude range of 1.5 km to 20.7 km, the TE-to-TM noise ratio increased by 12 dB. At the free-floating altitude of 20.7 km the TE noise was only

TABLE 4
TE-to-TM Noise Ratios at 42.5 kHz

Altitude (km)	Relative RMS Amplitude TE	TM	TE/TM (dB)
1.5	20	210	-20.4
4.6	28	203	-17.2
9.1	32	184	-15.1
16.8	46	170	-11.3
20.7	55	145	- 8.4

about 8 dB less than the TM noise, while at aircraft altitudes of about 9 km the TE noise was 15 dB less than the TM noise.

5. DISCUSSION

The experiment that was conducted on 4 August 1986 demonstrated that the ALFAN technique is a viable method for obtaining measurements of TE atmospheric noise. The results indicated that a free-floating balloon can provide a suitable, electrically-quiet, platform for obtaining measurements of TE fields that may be significantly weaker than corresponding TM fields. During the flight, for example, it was possible to obtain measurements of TE fields that were as much as 40 dB less than the corresponding TM fields. The experiment also showed that at the free-floating altitude of the balloon, the orientation of the TE loop antenna was essentially horizontal so that it was possible to obtain TE noise data that were remarkably free of contamination by TM fields.

During the flight TM and TE signals from three ground-based transmitters were monitored. Such data (see Tables 2 and 3) provide a measure of the polarization rotation effects in the ionosphere, caused by the geomagnetic field. Since those effects depend on such parameters as the signal frequency, the length of the propagation path, the direction of propagation (relative to the geomagnetic field) and solar illumination conditions, the data should be especially useful for validating low frequency propagation prediction techniques.

The TE-to-TM noise ratios at 42.5 kHz that were obtained during the first ALFAN flight, were appreciably larger than those reported from earlier experimental or theoretical efforts. For example, at an aircraft altitude of about 9 km, the ALFAN TE noise (see Table 4) was only about 15 dB weaker than the TM noise, more than 10 dB higher than earlier estimates [e.g., 7 and 12]. Of course, one cannot draw general conclusions from a few data points. It is hoped that data obtained from a number of planned ALFAN experiments will help to resolve such issues.

Factors which determine the TE/TM noise ratio at a given altitude include (a) the difference in altitude-amplitude profiles (the height-gains) which are, in turn, determined primarily by the boundary conditions at the earth and the ionosphere, and (b) the relative strength of the TM and TE patterns, which depends on the relative effectiveness of the average lightning stroke in exciting the TM and TE modes and the attenuation rates associated with the propagating modes. In addition, there are the "mixing" effects of polarization conversion in the ionosphere that affect the TE/TM ratios.

Unfortunately, because of the difficulty in obtaining TE noise data, an accumulation of a long-term statistical data base similar to those available for TM noise [4 and 6] is unlikely. Rather, the prediction of TE atmospheric noise will probably depend on theoretical models, validated in accordance with experimental data such as that being obtained in the ALFAN program.

REFERENCES

1. Kossey, P.A., E.A. Lewis and E.C. Field, Jr., Relative characteristics of te/tm waves excited by airborne vlf/lf transmitters, AGARD Conf. Proc., AGARD-CP-305 (paper 19), ADA113969, 1982.

2. Pappert, R.A. and J.E. Bickel, Vertical and horizontal vlf fields excited by dipoles of arbitrary orientation, Radio Sci., 35: 1445-1452, 1970.

3. Field, E.C., C.R. Warber, P.A. Kossey, E.A. Lewis and R.P. Harrison, Comparison of calculated and measured height profiles of transverse electric vlf signals across the daytime earth-ionosphere waveguide, Radio Sci., 21: No. 1, 141-149, 1986.

4. World distribution and characteristics of atmospheric radio noise, Proc. of the CCIR Xth Plenary Assembly, Geneva, 1963, Report 322, International Telecommunications Union, Geneva, 1964.

5. Maxwell, E.L., D.L. Stone, R.D. Croghan, L. Ball and A.D. Watt, Development of a vlf atmospheric noise prediction model, Westinghouse Geores. Lab. Rep.70-1H1-VLF NO-R1, 1970.

6. Fraser-Smith, A.C., and R.A. Helliwell, The Stanford University elf/vlf radiometer project: measurement of the global distribution of elf/vlf electromagnetic noise, Proc. IEEE Internat. Symp. on Electromag. Compatibility, Wakefield, Mass., 1985.

7. Gallenberger, R.J., and J.E. Bickel, Horizontal and vertical atmospheric noise measurements at vlf up to 20,000 ft altitude, Naval Electronics Laboratory Center Rep., NELC/TR 1793, AD889207, 1971.

8. Hirst, G.C., U-2 investigations of a new mode for lf air-to-air communications, Proc. AFSC Sci. and Engr. Symp., AFSC-TR-75-06, Vol. 1, ADA021660, 1975.

9. Harrison, R.P., G.C. Hirst and E.A. Lewis, Balloon measurements of te-and tm-polarization in lf waves from an airborne transmitter, Air Force Cambridge Research Laboratory Rep., AFCRL-TR-75-0613, ADA020545, 1975.

10. Harrison, R.P., E.A. Lewis, J.B. Donohoe and J.E. Rasmussen, TM/TE polarization ratios in a sample of 30 kHz sferics received at altitudes from 0 to 70 km, Rome Air Development Center Rep., RADC-TR-81-235, ADA108182, 1981.

11. Buckner, R.P., TCS Inc., Dallas, Texas, private communication, 1986.

12. Kelly, F.J., J.P. Hauser and F.J. Rhoads, Atmospheric vlf radio noise at elevated receivers: horizontal and vertical polarization, AGARD Conf. Proc., AGARD-CP-305 (paper 21), ADA113969, 1982.

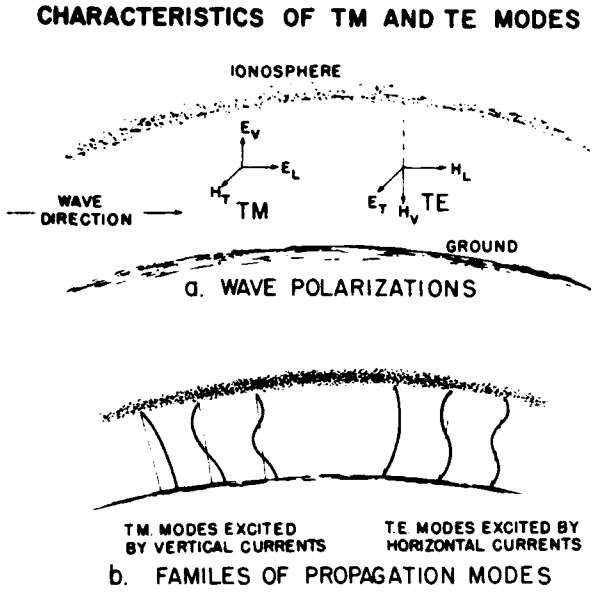


Figure 1. Characteristics of TM and TE modes.

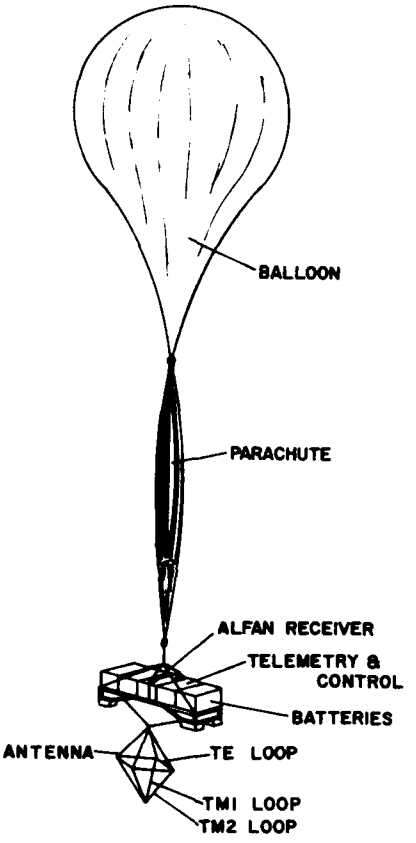


Figure 2. ALFAN balloon, parachute and payload configuration.

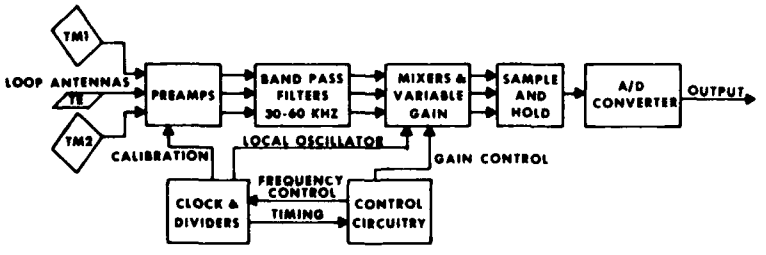


Figure 3. Block diagram of ALFAN LF receiver.

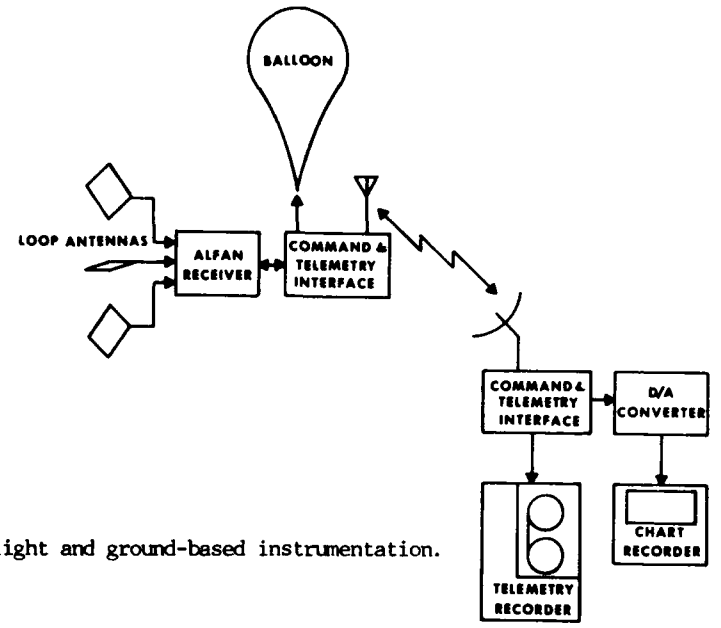


Figure 4. ALFAN flight and ground-based instrumentation.

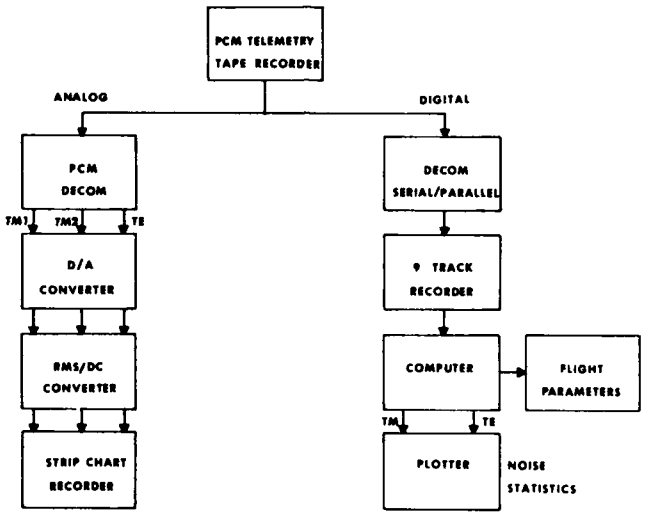


Figure 5. Flow chart for digital and analog analysis of flight data.

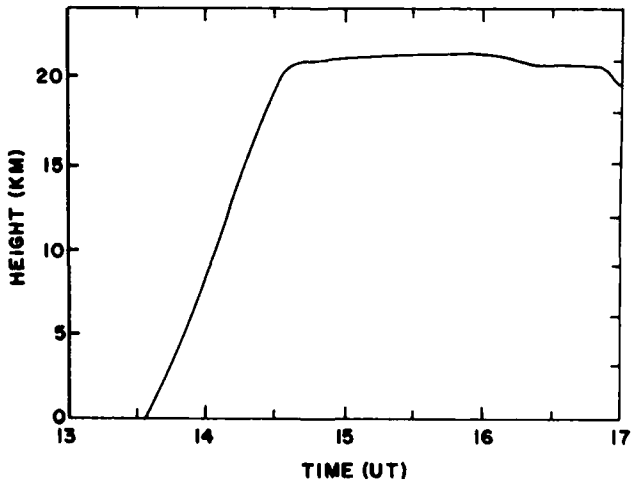


Figure 6, Payload altitude vs. time.

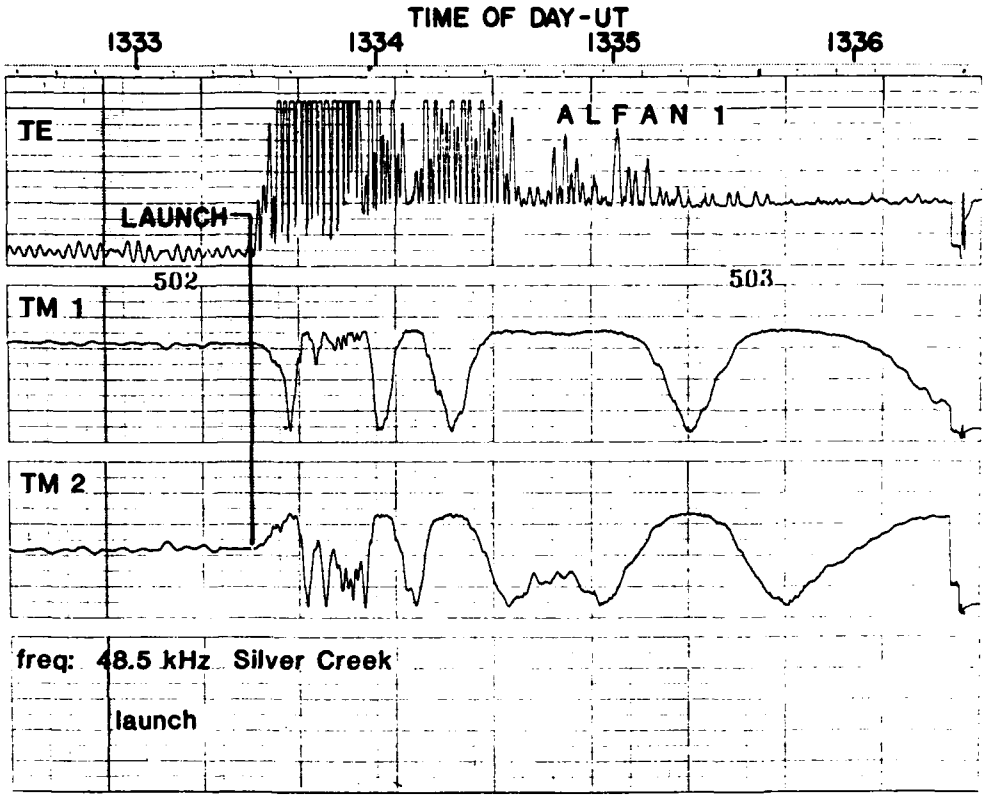


Figure 7. RMS signal strength for ground-based 48.5 kHz Silver Creek transmitter at balloon launch.

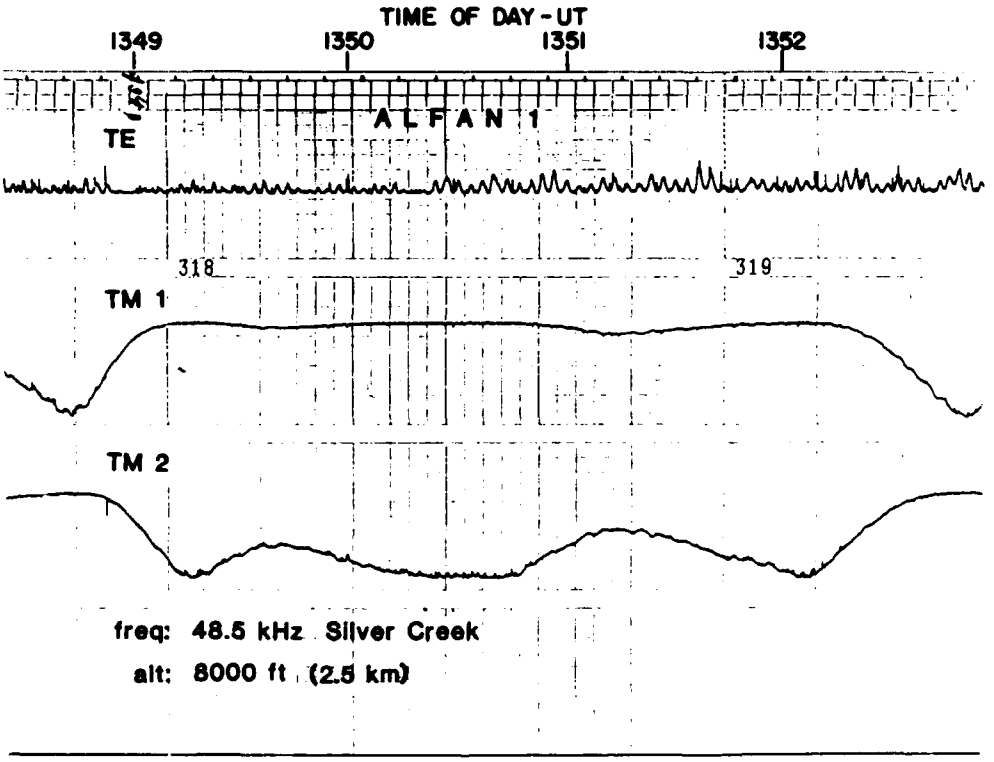


Figure 8. RMS signal strength fro ground-based 48.5 kHz Silver Creek transmitter at 2.5 km during balloon ascent.

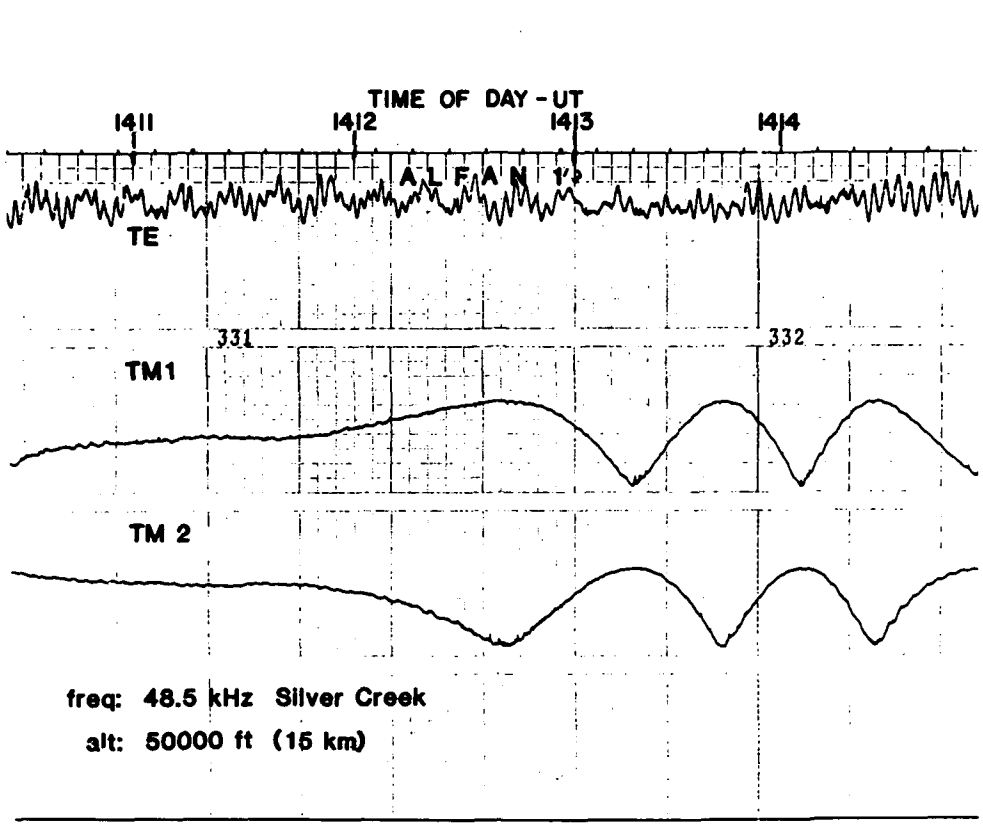


Figure 9. RMS signal strength for ground-based 48.5 kHz Silver Creek transmitter at 15 km during balloon ascent.

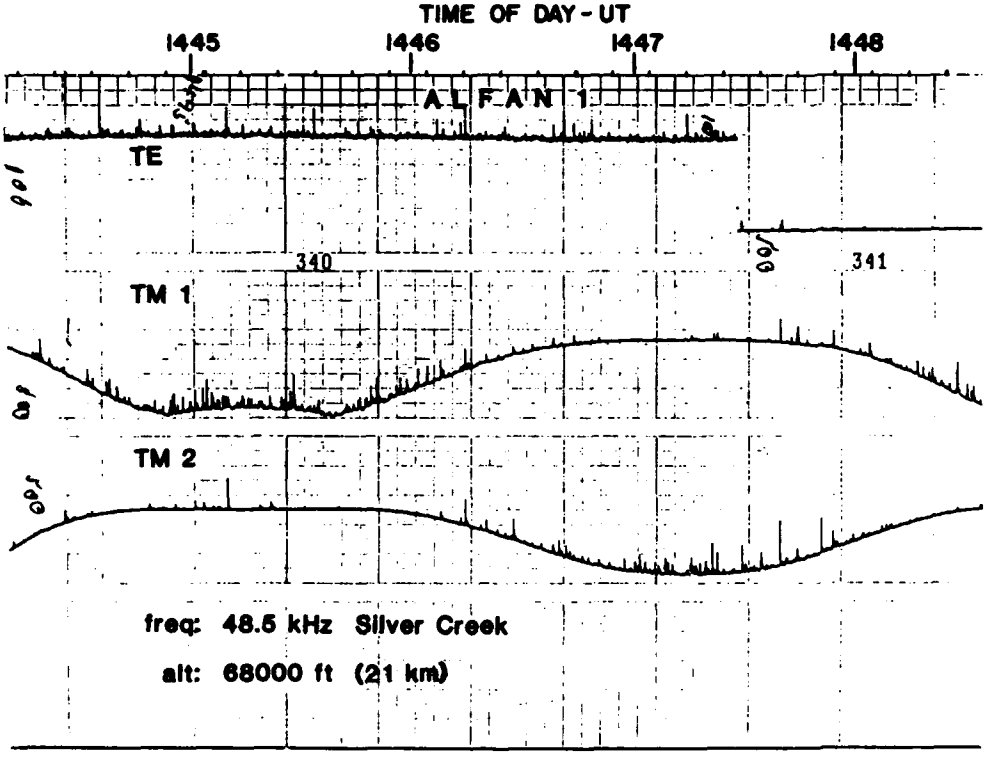


Figure 10. RMS signal strength for ground-based 48.5 kHz Silver Creek transmitter at 21 km, balloon floating altitude.

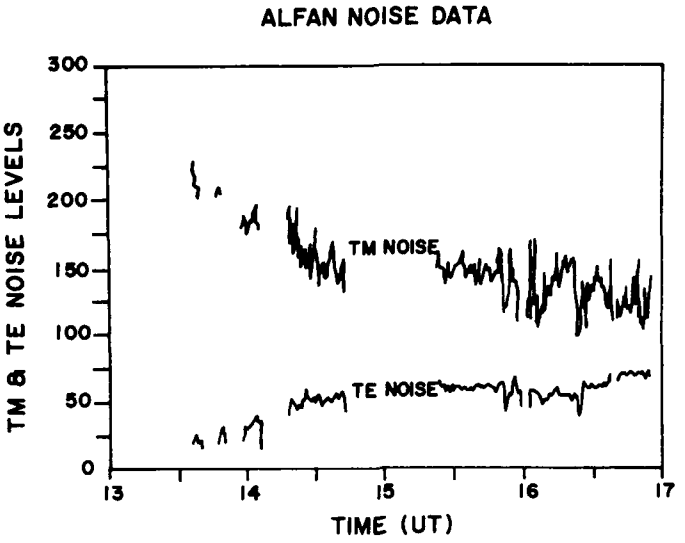


Figure 11. Relative RMS TE and TM 42.5 kHz atmospheric noise during flight.

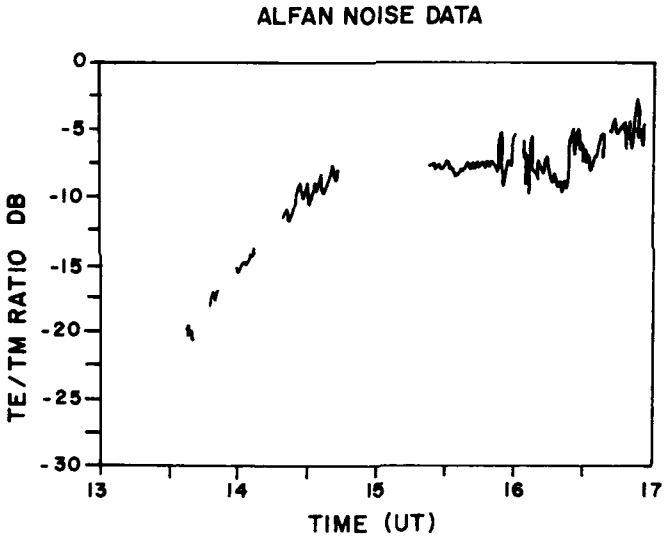


Figure 12. Ratio of TE to TM atmospheric noise during flight.

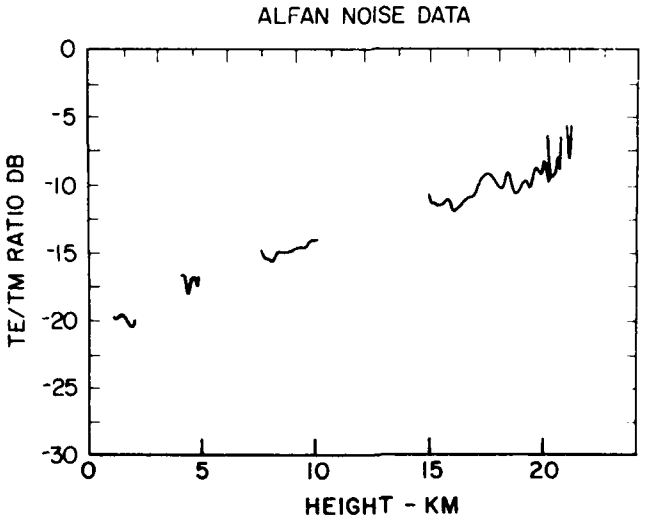


Figure 13. Ratio of TE to TM atmospheric noise as a function of balloon altitude.

LIGHTNING GENERATION OF LOW-FREQUENCY TE ATMOSPHERIC NOISE

Edward C. Field, Jr.
Chris R. Warber
Pacific-Sierra Research Corporation
12340 Santa Monica Boulevard
Los Angeles, California 90025

SUMMARY

This paper addresses the role of horizontal lightning strokes in producing transverse-electric low-frequency atmospheric noise. It calculates the fields generated by such strokes as a function of frequency, stroke altitude, stroke orientation, and the state of the ionosphere. Our results show that horizontal strokes are weak TE noise radiators unless their altitude exceeds a few kilometers. Above several kilometers, however, horizontal strokes can radiate TE noise almost as efficiently as vertical strokes radiate transverse-magnetic noise. Because TE noise does not propagate as well as TM noise, the TE/TM noise ratio tends to diminish with distance, particularly under disturbed ionospheric conditions. Most horizontal lightning channels do, in fact, occur above 3 km and are more frequent than vertical channels. Moreover, the horizontal structure of lightning is typically 2 or 3 times greater in extent than vertical structure, even for cloud-to-ground strokes. It therefore appears that horizontal lightning is high enough, occurs often enough, and has enough channel length to radiate substantial TE noise. TE atmospheric noise might therefore be stronger than previously believed. Recent balloon-borne measurements on the TE/TM noise ratio seem to confirm that conclusion.

1. INTRODUCTION

Energy radiated by very low frequency (VLF) or low frequency (LF) electromagnetic sources in the earth ionosphere waveguide propagates via transverse electric (TE) and transverse magnetic (TM) modes. The TM modes are best excited by vertically-polarized sources and can be radiated or received by terminals at virtually any altitude. The TE modes, on the other hand, are best excited by horizontally-polarized sources and are extremely weak unless both source and receiver are at least a few kilometers above the ground.

The advent of VLF/LF links between aircraft has made possible transmitting antennas that are elevated and horizontally oriented, and TE modes now play a role in strategic communications. Calculation of TE signals from airborne transmitters is straightforward because the source is well-defined and the theory is nearly the same as for TM signals. However, TE atmospheric noise is not as well understood as TM noise. Because TE noise must be measured aloft (on board airplanes or balloons), only a few brief measurements have been made. Accumulation of a long-term statistical data base like the one available for TM noise [CCIR, 1963; Fraser-Smith and Helliwell, 1985] is probably out of the question. Prediction of TE noise must therefore rely on theoretical modeling, as well as on empiricism.

The TE and TM modes are not strictly independent of one another, because geomagnetic cross-coupling in the ionosphere causes partial conversion between the two polarizations. Because of that conversion, a TM mode contains a horizontal component and a TE mode contains a vertical one. Conversion is strongest at night. In the daytime, or under nuclear-disturbed conditions, VLF/LF waves are reflected at low ionospheric altitudes, where electron-neutral collisions reduce the influence of the geomagnetic field, and the cross-coupling is usually weak. When coupling is important and the modes are not pure, they are often called quasi-TM and quasi-TE modes. We occasionally suppress the prefix "quasi" in this paper.

One would expect TE noise to be radiated mainly by the horizontal components of lightning strokes and TM noise to be radiated mainly by the vertical components. Nonetheless, computer models that predict VLF noise incorporate only vertical strokes and quasi-TM modes, and omit horizontal stroke components and quasi-TE modes [Maxwell et al., 1970; Kelly, Hauser, and Rhoads, 1982]. Such models therefore assume TE noise to arise solely from geomagnetically converted TM noise. That assumption could lead to substantial underestimates of TE noise levels. This paper calculates relative contribution of horizontal stroke components as a function of frequency, stroke altitude, stroke orientation, and the state of the ionosphere.

Horizontal lightning has received much less attention than vertical lightning in the literature because it typically occurs in-cloud and is of less socioeconomic importance than the more familiar cloud-to-ground lightning. Nonetheless, our results show that such lightning must be incorporated into models that predict VLF/LF atmospheric noise. We review the sparse data available on horizontally oriented lightning strokes and discuss their implications for TE noise modeling.

2. MATHEMATICAL FORMULATION

This section describes the structure of TE and TM modes and details how we calculate the fields generated by lightning strokes.

2.1. Fields Generated by Multiple Lightning Strokes

We begin by summarizing the equations governing TM and TE propagation, e.g., Galejs [1972]; Wait [1970]. The numerical solutions given below follow Budden's method [1961a,b], so as to account for the vertical inhomogeneity of the ionosphere and curvature of the earth. To define the notation and illustrate key dependences, we present the equations that govern the electric field when geomagnetic anisotropy is neglected. However, geomagnetic coupling among the modes is retained in the numerical results given in Sec. 4.

We use a spherical coordinate system (r, θ, ϕ) with its origin at the earth's center. We assume the wave to propagate in the θ-direction and define $E_V = E_r$ (vertical field); $E_H = E_\phi$ (transverse field); $E_L = E_\theta$ (longitudinal field).

2.2. Fields From Vertical Sources

The electric fields from the vertical component of a VLF/LF electric dipole are a superposition of TM modes. The usual propagation equations apply to radiation from coherent sources (transmitters) and allow for interference among the various propagation modes. That interference can be either constructive or destructive, and causes the well-known pattern of maxima and nulls that occurs on VLF/LF communication signals.

Lightning strokes, on the other hand, are neither sinusoidal in time nor coherent, but are short pulses. Moreover, because noise is treated statistically, various types of average values are of more interest than instantaneous values. We therefore ignore the phase terms and treat the mode sums in a RMS sense, which is tantamount to ignoring intermode interference. That approximation (essentially a random phase approximation) has been used in the formulas given below.

The vertical electric field is given by

$$E_{VV} = E_0 \sin\psi \left[\sum_\ell |S_\ell^3| \; |A_\ell^2| \; |G_\ell^2(h_T)| \; |G_\ell^2(h_R)| \times \exp\left(-\frac{2\alpha_\ell d}{8.7}\right) \right]^{1/2} \quad \text{V/m} , \tag{1}$$

and the longitudinal field (caused by wave tilt) by

$$E_{LV} = E_0 \sin\psi \left[\sum_\ell |S_\ell| \; |A_\ell^2| \; |G_\ell^2(h_T)| \; |A_\ell^{-2}| \; |G_\ell^2(h_R)| \times \exp\left(-\frac{2\alpha_\ell d}{8.7}\right) \right]^{1/2} \quad \text{V/m} , \tag{2}$$

where

$$E_0 = 120\pi \frac{IL}{(\lambda d)^{1/2}} \left(\frac{d/a}{\sin d/a}\right)^{1/2} \quad \text{V} . \tag{3}$$

The first subscript on E denotes the field component (V = vertical, L = longitudinal) and the second denotes the source component. We have assumed a sinusoidal frequency dependence and defined the various parameters as follows: ℓ = quantities associated with the ℓth TM mode, I = root-mean-square (RMS) current at antenna base, L = effective length of transmitting antenna, λ = free-space wavelength, d = distance from transmitter, a = radius of earth, $\sin\psi$ = factor accounting for inclined transmitting antennas (ψ being the angle between dipole orientation and the horizontal), always a value of 1 for vertical electric dipoles. Although most quantities are in meter-kilogram-second (MKS) units, distances (L, λ, d, a) are in megameters.

The quantity S_ℓ is the eigenvalue of the ℓth TM mode. At very low frequencies, S has a magnitude close to unity, so the term $S_\ell^{3/2}$ in Eq. (1) does not influence the field appreciably. The magnitude of the vertical electric field depends on the state of the ionosphere, mainly through the following parameters: A_ℓ, the excitation factor for the TM mode; α_ℓ, the attenuation rate in decibels per megameter of propagation (dB/Mm); and G_ℓ, the height-gain functions for transmitter and receiver heights h_T and h_R, respectively. The quantity A_ℓ in Eq. (2) is the wave admittance at the receiver altitude.

2.3. Fields From Horizontal Sources

The horizontal component of an elevated VLF/LF antenna radiates a complicated superposition of TE and TM modes. We avoid much of that complexity by considering broadside propagation only, where the great-circle path connecting transmitter and receiver is perpendicular to the plane containing the inclined electric-dipole transmitting antenna.

The broadside horizontal electric field is given by

$$E_{HH} = E_0 \cos\psi \left[\sum_m |S_m^{-1}| \; |\Lambda_m^2| \; |G_m^2(h_T)| \; |G_m^2(h_R)| \times \exp\left(-\frac{2\alpha_m d}{8.7}\right) \right]^{1/2} \quad V/m \quad . \tag{4}$$

The symbols are the same as in Eq. (1), except that the subscript m denotes the mth TE mode.

2.4. Mode Structure

The contribution of a given mode to the total field is proportional to its excitation factor Λ, its attenuation factor exp $(-\alpha d/8.7)$, and the source and receiver height-gain functions $G(h_T)$ and $G(h_R)$, respectively. We discuss those factors individually.

2.5. Excitation Factors

The factor Λ accounts for the efficiency with which a mode is excited by a ground-based transmitter. At lower VLF frequencies, the TM excitation factors Λ_ℓ are roughly equal to the reciprocal of the nominal reflection height. The TE modes are excited much less efficiently than the TM modes, with Λ_m being 4 to 5 orders of magnitude smaller than Λ_ℓ [Field et al., 1986]. That smallness prevents low-altitude sources from radiating TE modes efficiently in the VLF/LF bands, and it prevents low-altitude receivers from receiving TE modes.

2.6. Attenuation Rates

Higher-order modes are usually more severely attenuated than low-order modes and, as a result, can often be neglected at VLF over long path lengths. At LF, attenuation of the higher modes can be mitigated by efficient excitation.

The TE attenuation rate is virtually independent of ground conductivity. The TM rate, on the other hand, does depend on conductivity and exhibits a strong, broad maximum for conductivities between 10^{-5} and 10^{-4} mho/m, where the TM eigenangle is near the Brewster's angle of the ground [Pappert, 1970; Field, 1982]. Such a mode (usually the lowest TM mode) is very heavily attenuated and is called the Brewster mode.

The TE modes are usually attenuated somewhat more than TM modes, except for the Brewster mode. Both TM and TE modes become more heavily attenuated under disturbed conditions that depress the ionospheric reflection height. Such disturbances--of nuclear or natural origin--generally affect TE attenuation more adversely than TM attenuation.

2.7. Height-Gain Functions

It is the height-gain functions that are central to this paper, because they account for source and receiver elevations. Figure 1 diagrams idealized height-gain functions for the first three TM modes and the first three TE modes. The TE field is small at low altitudes, so the measurement of TE waves requires that receiver probes be at altitudes of at least several kilometers. Similarly, for efficient excitation of TE fields, the source must be at high altitude and have at least some horizontal orientation.

When the source is elevated above a few kilometers, the large height-gain function of TE modes mitigates the effects of the small excitation factor, so those modes can be excited nearly as effectively as TM modes. More precisely, the quantity $\Lambda_m G_m^2$ can be of the same order of magnitude as the quantity $\Lambda_\ell G_\ell^2$ at aircraft altitudes.

2.8. Geomagnetic Cross-Coupling Between TE and TM Modes

Strictly speaking, TE and TM polarizations are not independent because of geomagnetic cross-coupling in the ionosphere. Thus, TM waves partially convert to TE waves, and a vertically oriented source will excite a transverse field E_{HV} in addition to the primary fields E_{VV} and E_{LV} given by Eqs. (1) and (2). Similarly, a horizontal source will excite longitudinal and vertical fields E_{VH} and E_{LH}, even in the broadside direction.

Geomagnetic cross-coupling is usually strongest at night. In the daytime, VLF/LF waves are reflected at low ionospheric altitudes where electron-neutral collisions greatly reduce the influence of the geomagnetic field. Under disturbed conditions that

substantially depress the lower ionosphere, cross-coupling is virtually absent. We use a full-wave computer code that makes full allowance for geomagnetic anisotropy to calculate coupling between TM and TE modes, and therefore includes the cross-coupled fields E_{HV}, E_{VH}, and E_{LH} as well as the primary fields E_{VV}, E_{LV}, and E_{HH}.

3. NOISE-PREDICTION MODELS

At any point in time there is lightning activity somewhere on the earth. Each lightning stroke acts as a burst transmitter, and the radiated energy propagates to the receiver via earth-ionosphere waveguide modes. The total noise is the sum of the noise from all lightning strokes worldwide, weighted by the attenuation suffered propagating from source to receiver.

Depending on its height and inclination, a lightning stroke radiates both TM and TE modes. Figure 2 gives two examples of that process: a vertical cloud-to-ground stroke that radiates only TM modes; and an inclined in-cloud stroke that radiates both TM and TE modes. Figure 2 is oversimplified in two respects--first, cloud-to-ground strokes are seldom purely vertical; second, it neglects geomagnetic conversion and includes only the primary components E_{VV} and E_{HH} (E_{LV} is usually small).

Figure 3 shows schematically how geomagnetic coupling causes a vertical noise source at an undisturbed location to excite quasi-TM modes that have both vertical (E_{VV}) and horizontal (E_{HV}, E_{LV}) electric field components. If the ionosphere were disturbed, however, only pure TM and TE modes could exist, and the component E_{HV} in Fig. 3 would be very small. The fields would then approximate the classical form shown in Fig. 2.

Although more than 15 years old, the Westinghouse Georesearch Laboratory noise-prediction model is still the best available [Maxwell et al., 1970]. Developed in an era when no airborne VLF receivers were in operation, the WGL model treats only ground-level TM noise at frequencies from 10 to 30 kHz. It uses thunderstorm occurrence statistics as input data, divides the earth into many equivalent noise transmitters, and computes various noise properties. That model includes cloud-to-ground and in-cloud lightning strokes, but neglects the horizontal components of those strokes. The Naval Research Laboratory, Washington, D.C., has extended the WGL model by noting that a fraction of the horizontally polarized noise is due to geomagnetic conversion of the TM mode into a quasi-TM mode, as shown in Fig. 3 [Kelly, Hauser, and Rhoads, 1981]. Because the WGL model calculates E_{VV} for the lowest three quasi-TM modes, NRL was able to combine parts of Naval Ocean Systems Center WAVEGUID computer program with the WGL model to calculate quasi-TM-mode fields E_{HV} aloft, but did not include E_{HH}.

4. EFFICIENCY OF ELEVATED HORIZONTAL SOURCES

This section shows how the TE noise-generating efficiency of lightning strokes depends on the altitude and orientation of the stroke, and on the center frequency of the band in which the noise is observed. Although Eqs. (1) through (7) omit the geomagnetic field in order to simplify the discussion, all results given in this section were obtained with our WAVEPROP computer code, which accounts fully for the geomagnetic field and height gradients in the charged-particle densities and collision frequencies.

For all cases, we have assumed the fields to be measured at the nominal aircraft altitude of 30,000 ft [h_R = 9.1 km in Eqs. (1) through (7)]. We also assume propagation broadside to the horizontal component of the source, which we take to be an electric dipole of unit moment IL inclined at an angle ψ to the horizontal. Our use of a dipole rather than an extended source does not accurately represent a lightning stroke, which can be many kilometers in length. However, an extended stroke can be synthesized from a chain of elemental dipoles, properly weighted for intensity. Similarly, our use of specific frequencies should be interpreted as representing individual Fourier components of an actual lightning stroke.

4.1. Dependence on Source Inclination

Figures 4 through 6 illustrate how the vertical (E_V), transverse (E_H), and longitudinal (E_L) electric-field components depend on source inclination ψ for three different ionospheric conditions. Those three figures assume a source altitude of 5 km, a frequency of 45 kHz, a range from the source of 2 Mm, a geomagnetic dip angle of 60 deg, and propagation from geomagnetic south-to-north. Recall that ψ is measured from the horizontal, so ψ = 0 for a horizontal electric dipole (HED) and ψ = 90 deg for a vertical electric dipole (VED). Also note that because of geomagnetic conversion, the vertical field E_V contains the cross-coupled term E_{VH} as well as the primary term E_{VV}, and E_H contains E_{HV} and E_{HH}.

We begin by showing results for a disturbed ionosphere where the reflection heights are depressed far below their ambient levels. As discussed above, that case is simplest; geomagnetic cross-coupling is nearly absent, and the modes closely approximate classic TM and TE forms. In particular, the transverse field E_H is caused almost entirely by the horizontal component IL cos ψ of the dipole source and should be nearly identical to E_{HH}. The vertical and longitudinal fields E_L and E_V should be nearly equal to E_{LV} and E_{VV}, because they are caused almost entirely by the vertical source component IL sin ψ.

Figure 4 shows the three electric-field components versus dipole inclination for a severely depressed ionosphere, such as might occur in a moderate nuclear environment or

during a strong solar proton event. For the case shown, the ionospheric reflection height was depressed to nearly 50 km, well below its nominal daytime value of around 65 km, and far below the 70 to 75 km height range where collisional and geomagnetic effects are approximately equal. As expected, the fields have the classic form, where E_V and E_L are nearly proportional to sin ψ and E_H is nearly proportional to cos ψ. There is, however, a small amount of geomagnetic coupling, evidenced by the fact that the transverse field E_H does not quite vanish when the dipole is oriented vertically.

One measure of the strength of TM-to-TE coupling is the ratio of E_H to the vertical field E_V, both computed for ψ = 90 deg. That ratio, which is E_{HV}/E_{VV}, gives the relative transverse field generated by a VED. If coupling were absent, E_{HV} would be zero. Figure 4 shows E_{HV}/E_{VV} to be -80 dB, a very small number. It is therefore permissible to neglect the geomagnetic field when the ionosphere is depressed far below its daytime level.

Figure 5 applies to ambient daytime ionospheric conditions rather than disturbed conditions, so the geomagnetic coupling should be stronger than shown in Fig. 4, and the fields should differ somewhat from the classic TM and TE forms. The vertical and longitudinal components E_V and E_L still approximate a sin ψ dependence (although not so closely as in Fig. 4), but the transverse field E_H shows a dependence on ψ that bears little resemblance to cos ψ. The TM-to-TE conversion ratio E_{HV}/E_{VV} is -25 dB--still weak, but 55 dB stronger than indicated in Fig. 4 for the depressed ionosphere.

Figure 6 shows the three electric-field components under undisturbed nighttime ionospheric conditions, where geomagnetic cross-coupling is strongest. Here the ψ-dependence of E_V, E_L, and E_H, respectively, differs from the classic sin ψ and cos ψ forms. The strength of the TM-to-TE coupling at ψ = 90 deg is -7 dB, which means that under the conditions assumed, a VED will produce a transverse field E_{HV} that is nearly half as large as the primary field E_{VV}. Therefore, under ambient nighttime conditions, a vertical source can produce a substantial horizontal electric field.

4.2. Efficiency of Horizontal and Vertical Noise Sources

We now show how the noise-generating efficiencies of HED and VED sources depend on source elevation. Figures 7 and 8 illustrate an elevated HED (ψ = 0) at altitude h and a ground-based VED (ψ = 90). The ratio E_{HH}/E_{VV} is plotted in the figures, which gives the relative amplitudes of the primary field components generated by each source type. Stated simply, that ratio compares transverse noise excited by a horizontal source with vertical noise excited by a ground-based vertical source having the same moment. It reveals the TE-to-TM noise ratio if horizontal stroke components were as numerous and strong as vertical stroke components.

Figure 7 shows E_{HH}/E_{VV} versus the HED altitude for three types of ionosphere: ambient night, ambient day, and strongly disturbed. It assumes a ground conductivity of 10^{-2} s/m, south-to-north propagation, a frequency of 45 kHz, and ranges of 2 Mm and 4 Mm. The ratios are very small for HED altitudes below about 1 km because, as discussed in Sec. 2, horizontal sources radiate very poorly if they are too close to the ground. If the HED source is higher than a few kilometers, the ratios become much larger, particularly at night. In no case is the transverse field from the HED as strong as the vertical field from the VED, indicating that, in regions of high ground conductivity, TE noise is weaker than TM noise unless horizontal sources are more numerous or more intense than vertical sources. Note, however, that the ratio approaches unity (0 dB) at night if the source is above 5 km and within 2 Mm of the receiver. The ratio decreases as the range increases, because TE modes are usually more heavily attenuated than TM modes. That effect is especially pronounced when the ionosphere is depressed, which is why the ratio is so small as to be off-scale in Fig. 7b.

Figure 8 shows how the results depend on assumed ground conductivity. The HED is about 20 dB more effective over poor ground ($\sigma = 10^{-4}$ s/m) than over good ground ($\sigma = 10^{-2}$ s/m). Moreover, if the HED is more than 4 km above poor ground, the ratio E_{HH}/E_{VV} exceeds unity (0 dB), which means it is more efficient at exciting TE noise than the VED is at exciting TM noise. That result is due to two effects: (1) horizontal dipoles radiate better over poorly conducting ground than over highly conducting ground; and (2) TE modes, being earth-detached, propagate equally well over all types of ground; whereas, TM modes are more heavily attenuated over poor ground.

4.3. Importance of Horizontal Stroke Components in Generating TE Noise

Figure 9 illustrates the ratio E_{HH}/E_{HV}, where E_{HH} is the horizontal electric field radiated by an elevated horizontal source (lightning stroke) and E_{HV} is the geomagnetically coupled horizontal field radiated by a ground-based vertical source of the same strength. Both E_{HH} and E_{HV} are calculated for an altitude of 30,000 ft and a frequency of 45 kHz.

Recall that E_{HH} propagates via TE modes, omitted from certain noise models; whereas, E_{HV} propagates via quasi-TM modes, retained in those models. The ratio E_{HH}/E_{HV} shown in Fig. 9 is, therefore, the number of decibels that should be added to the horizontally polarized noise predicted by those models if horizontal stroke components were as numerous and intense as vertical ones.

As expected, the ratio E_{HH}/E_{HV} is small if the lightning stroke is below 1 or 2 km. It is larger for strokes at higher altitudes, but still only moderate under normal nighttime conditions, where geomagnetic conversion is strong. It is larger (up to 10 dB)

under normal daytime conditions because the daytime ionosphere suppresses geomagnetic conversion. For the disturbed environment, the ratio can exceed 30 dB. That huge ratio occurs because the nuclear environment depresses the waveguide height and virtually eliminates the ability of a vertical stroke to excite horizontal fields E_{HV}. Horizontally polarized noise is then dominated by horizontal strokes and TE modes.

5. SURVEY OF DATA ON HORIZONTAL LIGHTNING

Three conditions must be met for horizontal lightning to produce substantial global TE noise. First, horizontal lightning must be abundant; second, it must occur above 3 to 4 km; third, it must have adequate HED moments, which implies that the "IL" product (the product of current and channel length) should be sizable.

In order to assess how well horizontal lightning meets the above conditions, we surveyed numerous published sources on lightning phenomena. The materials reviewed, although not all-inclusive, adequately represent data now available. Only a small fraction of those sources contain quantitative data on horizontal lightning.

5.1. Types of Lightning

Lightning is divided into two major categories: (1) discharges that strike the ground, called ground flashes or cloud-to-ground (CG) lightning, and (2) discharges that occur within a cloud or between two clouds, usually classified together and called in-cloud (IC) lightning.

5.2. Cloud-to-Ground Lightning

Because of its economic and social importance, CG lightning has been studied extensively. The initial part of the flash is called the leader and is initiated by a preliminary breakdown within the cloud. As the leader nears the ground, its tip becomes connected to the ground potential. The leader channel is then discharged and the "return stroke" propagates up the ionized leader channel. Average leader currents are on the order of 100 A. However, peak currents of 30 kA occur in the return stroke, at least near the ground [Uman and Krider, 1982].

CG lightning traditionally has been considered primarily a vertical discharge. There is, however, mounting evidence that even a vertical return stroke becomes horizontal after entering a cloud, where it extends laterally for many kilometers [Pathak, Rai, and Varshneya, 1982]. CG strokes therefore can excite TE waveguide modes directly.

5.3. In-Cloud Lightning

Although IC lightning is more common than CG lightning, it has not been studied as extensively because detailed observations are difficult, and IC lightning is of little economic or social importance. Therefore, much less is known about the properties of IC lightning than earth flashes.

It is believed that IC discharges consist of a continuously propagating leader that generates a number of return strokes called recoil streamers. The duration of an IC discharge is approximately the same as that of a CG flash (0.5 ms). There is no stroke in IC lightning that compares in strength with the intense return strokes in CG lightning.

5.4. Abundance of Horizontal Lightning

Numerous data have been collected on the global distribution and frequency spectrums of CG lightning. The majority of those data are from single station measurements, which cannot provide precise information on the structure and location of the lightning channel. Pioneering work in the early 1970s [Few and Teer, 1974] demonstrated that the large-scale features of thunderstorms, including lightning structure within clouds, could be reconstructed from thunder recorded with an array of microphones.

Those techniques were then applied to IC lightning in a Tucson, Arizona, storm [Teer and Few, 1974], including the IC channel of the CG stroke as well as IC lightning paths. All channels exhibited considerable horizontal structure, with an average horizontal range of approximately 6.2 km, a minimum horizontal range of 2.6 km, and a maximum horizontal range of 11.6 km. Moreover, the predominant orientation of the discharge was horizontal in CG strokes as well as IC strokes. For CG strokes, the horizontal extent of the channel on the average exceeded the vertical by 2.1 ±0.8 times. The average ratio of horizontial extent to vertical thickness for the IC discharges was 2.9 ±0.9.
A more recent study [MacGorman, Few, and Teer, 1981] encompassed acoustic data from storms in Arizona, Colorado, and Florida. In all three storms, the structure of individual lightning flashes was primarily horizontal. Figures 10 and 11 give examples of lightning structure reconstructed from the acoustic data. The horizontal orientation of both the CG stroke (Fig. 10) and the IC stroke (Fig. 11) is evident.

Because IC flashes tend to be even more horizontal than CG strokes, the relative abundance of CG and IC flashes is important to our analysis. Prentice and Mackerras [1977] assembled much of the early data on the ratio of IC to CG flashes. With one exception, the number of IC flashes counted exceeded the number of CG flashes.

Southern hemisphere measurements also indicate an abundance of IC flashes. In a lightning study in Pretoria, South Africa, which lasted more than 10 years, the apparent

cloud flash density was 2.5 times the earth flash density [Anderson et al., 1984]. Throughout two thunderstorm seasons from September 1982 to May 1984 in Brisbane, Australia, the overall IC to CG ratio was 3.0 [Mackerras, 1985].

The studies cited above show that IC flashes are more abundant than CG flashes, and that, as expected, IC lightning is primarily horizontal. Moreover, the structure of CG lightning (long treated as vertical) seems to have a predominantly horizontal orientation within the cloud. Horizontal lightning therefore appears to be more abundant than vertical lightning, and to have greater average channel lengths.

5.5. Altitude of Horizontal Lightning

The data given by MacGorman, Few, and Teer [1981] indicate lightning tends to occur in layers 2 to 3 km thick. Figure 12 shows the altitude distribution of the average number of acoustic sources per lightning flash. The majority of activity occurs above 3 km, and much activity occurs above 6 km. Moreover, that high-altitude horizontal structure is by no means confined to IC flashes. Recall from Fig. 10 that (at least for the stroke depicted) the CG return stroke is mainly vertical only below about 3 km, above which it turns and runs horizontally for more than 10 km.

We conclude that horizontal channels for both CG and IC lightning occur at altitudes high enough to excite TE noise efficiently.

5.6. Strength of Horizontal Lightning

Equations (1) through (7) show that the radiation from an electric source is proportional to the current I and the channel length L. Data cited above indicate that horizontal lightning channels are typically 2 or 3 times longer than vertical ones, and, all other things taken equal, horizontal strokes should be better radiators than vertical ones. All other things are not equal, however, and we need data on horizontal current amplitudes in order to draw conclusions on horizontal source strength.

Lightning currents can be measured directly if the stroke contacts the ground, or can be inferred by inverting fields measured at some distance from the stroke. Horizontal strokes never contact the ground. Moreover, we showed in Sec. 4 that the TE fields needed to define horizontal lightning currents can be measured only at altitude. The TE noise-field data are therefore virtually nonexistent, and little information is available on the strength of horizontal lightning currents.

In the absence of data, the best we can do is assume the horizontal currents in IC strokes to be equal to the vertical currents, on which data are available. If that assumption is correct, the ratio of HED to VED moment in an IC discharge is simply the ratio of the horizontal to vertical channel lengths--about 2 or 3, according to data given above.

Maxwell et al., [1970] give peak currents on the order of 5 kA and average vertical channel lengths of 1 or 2 km for IC strokes. They use IC VED moments ranging from 2×10^3 to 5×10^3 A-km, and we infer the corresponding IC HED moments to be 2 or 3 times larger (on the order of 10^4 A-km). Maxwell et al., [1970] use 5×10^4 A-km ($1/2 \times 30$ kA $\times$ 3.5 km) for VED moments of return strokes in CG flashes. Horizontal IC strokes would therefore appear, on the average, to be about 14 dB (20 log 5) weaker than the vertical channels of CG return strokes. However, although IC strokes are weaker than CG return strokes, they are more abundant.

We found no data on the horizontal current strength in CG return strokes. That current--not included in the estimate presented above--could be the major contributor to TE noise. A plausable model for the complete discharge is that the horizontal current distributes the vertical current to charge concentrations within the cloud. We would then expect the total horizontal current to be approximately the same as the vertical current. The fact that horizontal currents cause thunder indicates the current is fairly large.

6. DISCUSSION

Vertically oriented lightning excites low-frequency TM noise efficiently, regardless of stroke altitude. Horizontally oriented lightning, on the other hand, is a weak radiator of TE noise if the stroke is within 1 to 2 km of the earth's surface. Only if a horizontal stroke altitude exceeds a few kilometers will it approach its full effectiveness as a TE noise radiator. In that situation, a horizontal stroke radiates TE noise almost as strongly as a vertical stroke (having the same current moment) radiates TM noise. It also produces much more TE noise than arises from geomagnetic conversion of TM noise. Because TE noise does not propagate as well as TM noise, the TE/TM noise ratio tends to diminish as the distance from the source increases, particularly under strongly disturbed ionospheric conditions.

Although data on horizontal lightning are scarce, our literature search showed that most horizontally oriented channels occur high enough (above 3 km) to be efficient TE noise radiators. Moreover, the horizontal structure of lightning--both IC and CG--is typically 2 or 3 times greater in extent than the vertical structure. It therefore appears that horizontal lightning is high enough, occurs often enough, and has enough channel length to radiate TE noise almost as strongly as vertical lightning radiates TM noise. The question remains how horizontal lightning currents compare with vertical currents. Data are virtually nonexistent, although we can conclude from TM field measurements that currents in IC strokes are smaller than the intense return-stroke currents in CG strokes.

The above discussion indicates that the TE/TM noise ratio at aircraft altitudes depends on several competing factors: TE noise sources have adequate altitude and greater length than TM sources; but TM sources can have extremely strong currents, and TM noise usually suffers less propagation loss than TE noise. A quantitative evaluation of all those factors is beyond the scope of this paper. However, two conclusions can be made at this time: First, horizontal lightning strokes must be included in TE noise-prediction models; second, TE atmospheric noise might well be stronger than previously believed.

There have been a few brief measurements of the TE/TM noise ratio aloft. Kelly, Hauser, and Rhoads [1982] cite data measured at midlatitudes on board an aircraft. The nominal TE/TM noise ratios at both 17 kHz and 29 kHz were between -25 dB and -35 dB in the daytime, and between -10 dB and -15 dB at night. In view of our findings, those ratios are surprisingly small, especially in the daytime.

More recently, the Rome Air Development Center (RADC), Griffiss Air Force Base, New York, used a balloon-borne package to measure the daytime TE/TM ratio at a frequency of 42.5 kHz [Turtle, 1987]. The location was New Mexico; the date was 4 August 1986; and the local time was from about 0900 to 1345. Most of the measurements were taken at the final balloon altitude of 68,000 ft, but a short measurement was made at 30,000 ft during ascent.

At the higher altitude (68,000 ft) the measured RMS TE/TM ratio was -8.4 dB. Assuming a linear dependence of noise amplitude on altitude,* the TE/TM ratio at the nominal aircraft altitude of 30,000 ft is -15.5 dB. That scaled number agrees closely with the actual -15.1 dB ratio measured during ascent. Results not presented here show that the daytime TE/TM ratio at 42.5 kHz should be only slightly higher than at 29 kHz. The ratios reported by RADC can thus be compared with those cited by Kelley, Hauser, and Rhoads [1982], and appear to be about 10 dB higher.

It is, of course, dangerous to draw general conclusions from a few data points because noise is variable, and any given data point could be sporadic. Accordingly, RADC has made another set of measurements in New Mexico. The data were not fully reduced at the time of this writing, but seem generally consistent with the August 1986 data.

The use of balloon-borne instrument packages to measure TE noise fields offers the potential of obtaining much needed data on the currents in horizontal lightning channels. If such a package were flown within 50 km of a storm, for example, it should be possible to resolve individual strokes and obtain TE and TM field time-histories that are relatively free of propagation distortion. Such data, now only available for TM noise sources, could then be used to infer the HED-to-VED ratio of various stroke types.

---

*Equivalent to assuming $\sin \pi h/H \simeq \pi h/H$, where h is balloon height and H is the height of the ionosphere (see Fig. 1).

REFERENCES

Anderson, R. B., et al., "The Development and Field Evaluation of a Lightning Earth-Flash Counter," Paper 2932 A, Proc. IEE, London, 1984.

Budden, K. G., Radio Waves in the Ionosphere, Cambridge Press, Cambridge, England, 1961a.

Budden, K. G., The Wave-Guide Mode Theory of Wave Propagation, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1961b.

CCIR (International Radio Consultative Committee), World Distribution and Characteristics of Atmospheric Radio Noise, International Telecommunication Union, Report 322, 1963.

Few, A. A., and T. L. Teer, "The Accuracy of Acoustic Reconstructions of Lightning Channels," J. Geophys. Res., Vol. 79, No. 3, 1974, pp. 5007-5011.

Field, E. C., "VLF Propagation in Disturbed Environments," AGARD Conference Proceedings 305: Medium, Long, and Very Long Wave Propagation, February 1982, pp. 10-1 to 10-13 [ADA 113969].

Field, E. C., et al., "Comparison of Calculated and Measured Height Profiles of Transverse Electric VLF Signals Across the Daytime Earth-Ionosphere Waveguide," Radio Sci., Vol. 21, No. 1, pp. 141-149, January-February 1986.

Fraser-Smith, A. C., and R. A. Helliwell, "The Stanford University ELF/VLF Radiometer Project," Proceedings of the 1985 International Symposium on Electromagnetic Compatibility, Wakefield, Massachusetts, August 1985, pp. 305-311.

Galejs, J., Terrestrial Propagation of Long Electromagnetic Waves, 362 pp., Pergammon Press, New York City, 1972.

Kelly, F. J., J. P. Hauser, and F. J. Rhoads, "Atmospheric VLF Radio Noise at Elevated Receivers: Horizontal and Vertical Polarization," AGARD Conference Proceedings No. 305, Paper 21, February 1982 [ADA 113969].

MacGorman, D. R., A. A. Few, and T. L. Teer, "Layered Lightning Activity," J. Geophys. Res., Vol. 86, 1981, pp. 9900-9910.

Mackerras, D., "Automatic Short-Range Measurement of the Cloud Flash to Ground Flash Ratio in Thunderstorms," J. Geophys. Res., Vol. 90, No. D4, 1985, pp. 6195-6201.

Maxwell, E. L., et al., Development of a VLF Atmospheric Noise Prediction Model, Westinghouse Georesearch Laboratory, Report 70-1H2-VLF-N0-R1, 1970.

Pappert, R. A., "Effects of Elevation and Ground Conductivity on Horizontal Dipole Excitation of the Earth-Ionosphere Waveguide," Radio Sci., Vol. 5, No. 3, 1970, pp. 579-590.

Pathak, P. P., J. Rai, and N. C. Varshneya, "VLF Radiation from Lightning," Geophys. J. R. Astro. Soc., Vol. 69, 1982, pp. 197-207.

Prentice, S. A., and D. Mackerras, "The Ratio of Cloud to Cloud-Ground Lightning Flashes in Thunderstorms," J. Appl. Meteor., Vol. 16, 1977, pp. 545-550.

Teer, T. L., and A. A. Few, "Horizontal Lightning," J. Geophys. Res., Vol. 79, No. 24, 1974, pp. 3436-3441.

Turtle, J., Rome Air Development Center, Griffiss Air Force Base, New York, private communication, June 1987.

Uman, M. A., and E. P. Krider, "A Review of Natural Lightning: Experimental Data and Modeling," IEEE Transactions on Electromagnetic Compatibility, Vol. EMC-24, No. 2, 1982, pp. 79-112.

Wait, J. R., Electromagnetic Waves in Stratified Media, Pergamon Press, New York City, 1970.

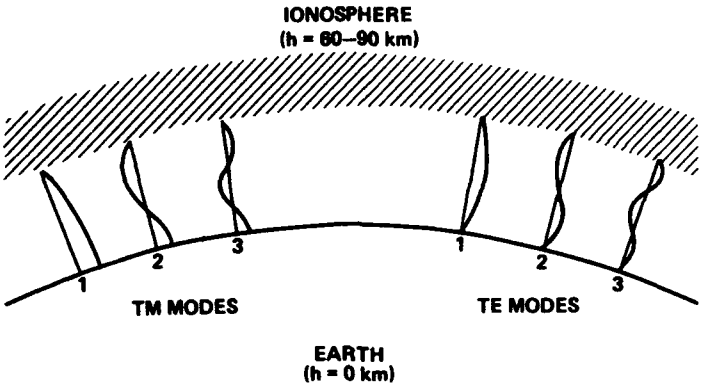


Fig. 1. Idealized height-gain functions in earth-ionosphere waveguide.

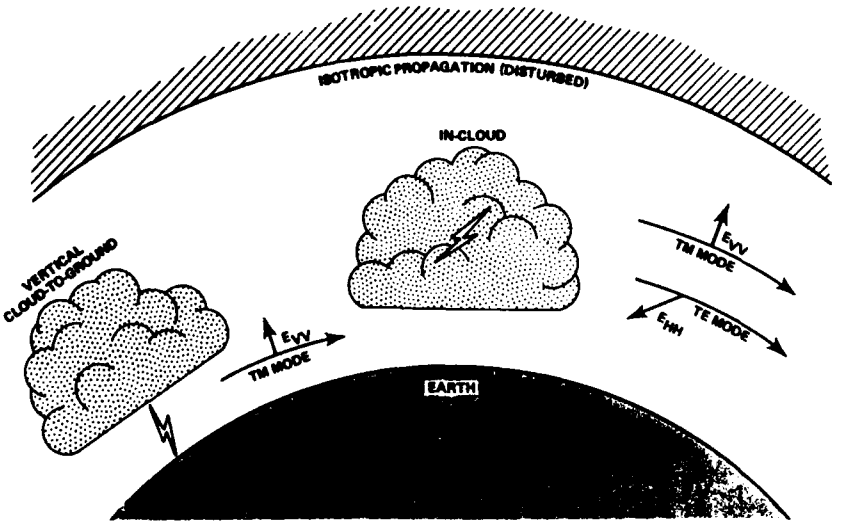


Fig. 2. TE- and TM-mode generation by lightning strokes (isotropic propagation).

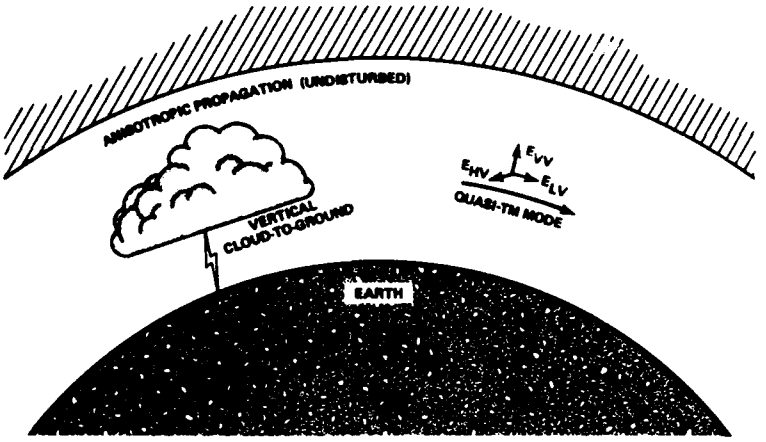


Fig. 3. Geomagnetic coupling between TM and TE noise in anisotropic waveguide.

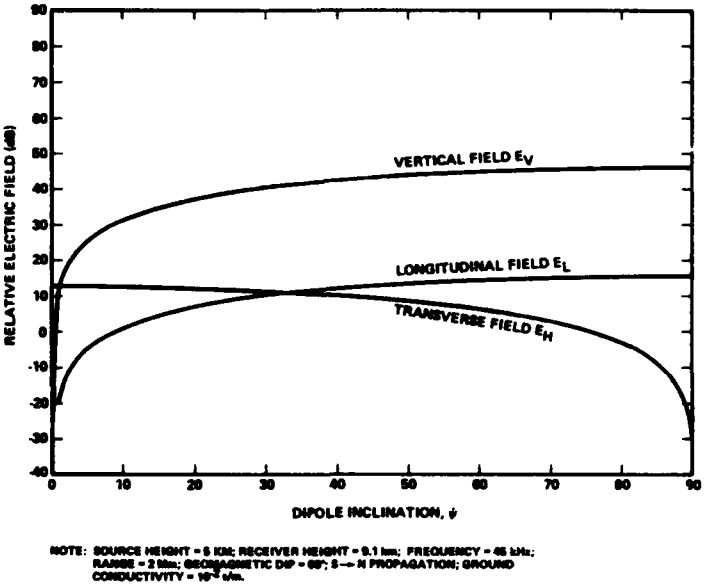


Fig. 4. Vertical and horizontal electric field components versus source orientation for disturbed conditions.

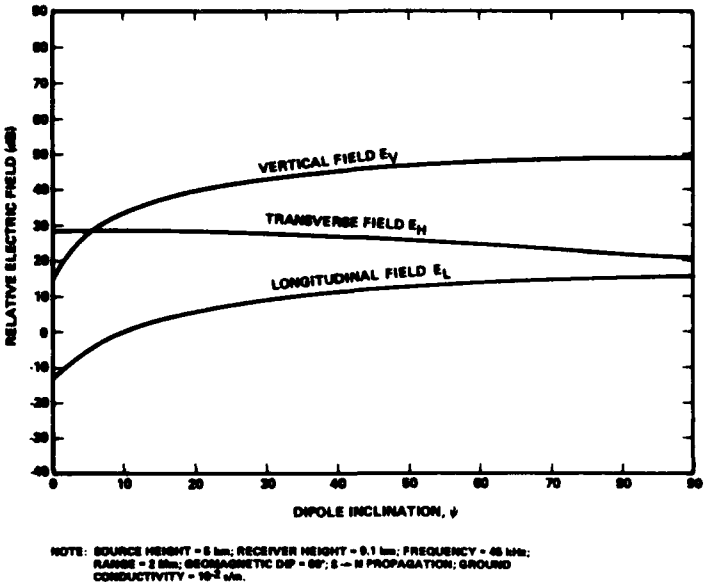
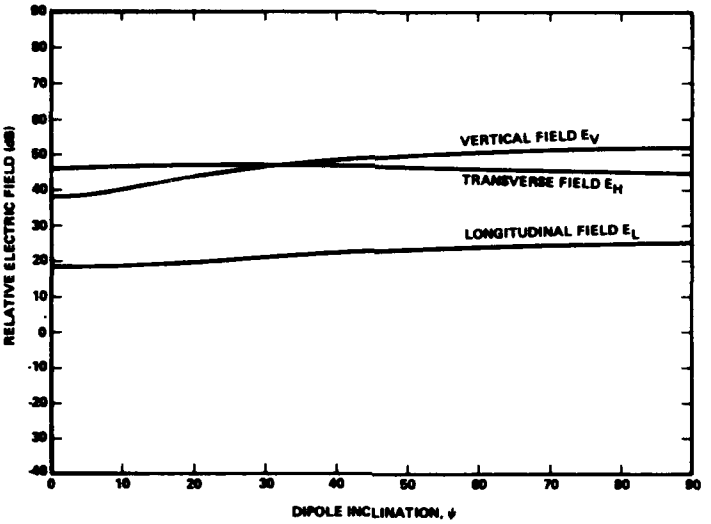
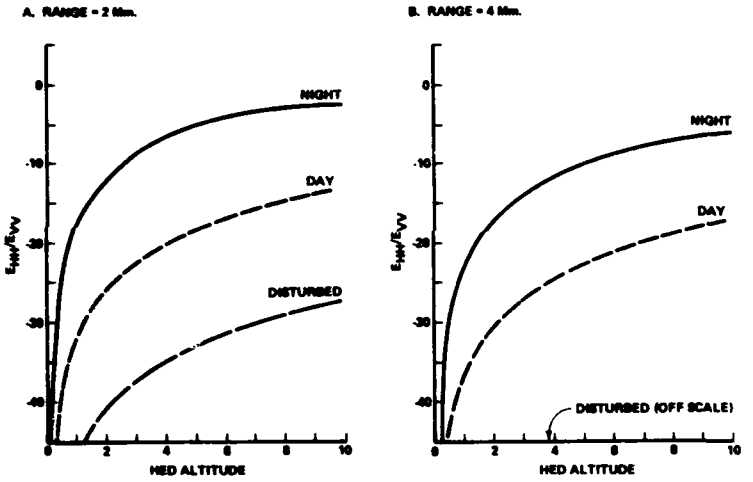


Fig. 5. Vertical and horizontal electric field components versus source orientation for ambient daytime conditions.



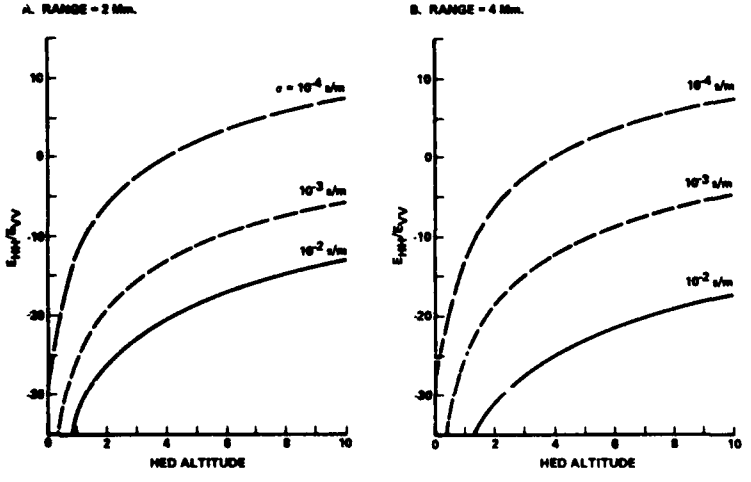
NOTE: SOURCE HEIGHT = 5 km; RECEIVER HEIGHT = 0.1 km; FREQUENCY = 45 kHz; RANGE = 2 Mm; GEOMAGNETIC DIP = 68°; S → R PROPAGATION; GROUND CONDUCTIVITY = 10^{-3} s/m.

Fig. 6. Vertical and horizontal electric field components versus source orientation for ambient nighttime conditions.



NOTE: RECEIVER HEIGHT = 0.1 km; FREQUENCY = 45 kHz; GEOMAGNETIC DIP = 68°; S → R PROPAGATION; GROUND CONDUCTIVITY = 10^{-3} s/m.

Fig. 7. Ratio of primary field E_{HH} produced by elevated horizontal dipole to primary field E_{VV} produced by ground-based vertical dipole for three ionospheric conditions.



NOTE: RECEIVER HEIGHT = 0.1 km; FREQUENCY = 45 kHz; GEOMAGNETIC DIP = 68°; S → R PROPAGATION.

Fig. 8. Ratio of primary field E_{HH} produced by elevated horizontal dipole to primary field E_{VV} produced by ground-based vertical dipole for ambient daytime conditions and various ground conductivities.

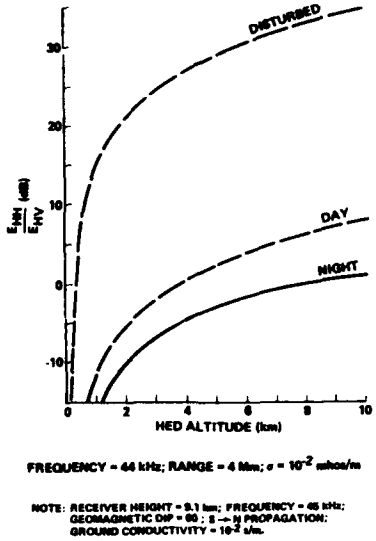


Fig. 9. Ratio of TE noise from horizontal source to TE noise from vertical source for three ionospheric conditions.

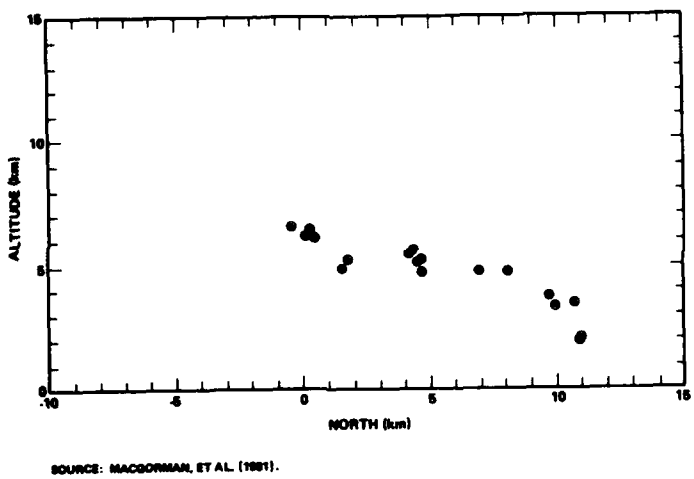


Fig. 10. Elevation view of acoustic sources in CG stroke in Arizona storm.

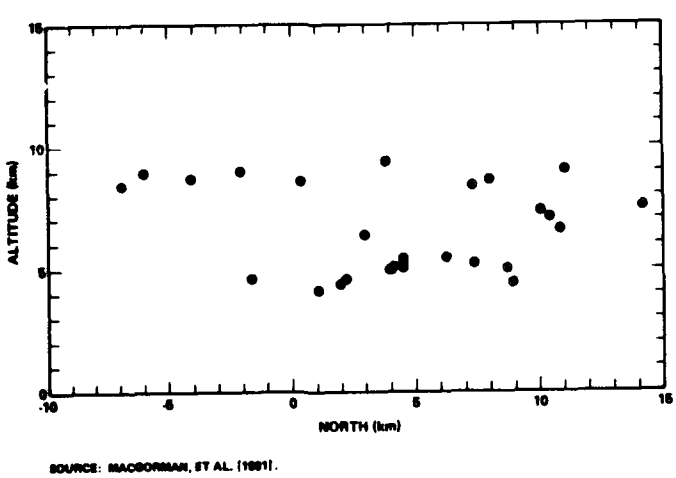


Fig. 11. Elevation view of acoustic sources in IC stroke in Arizona storm.

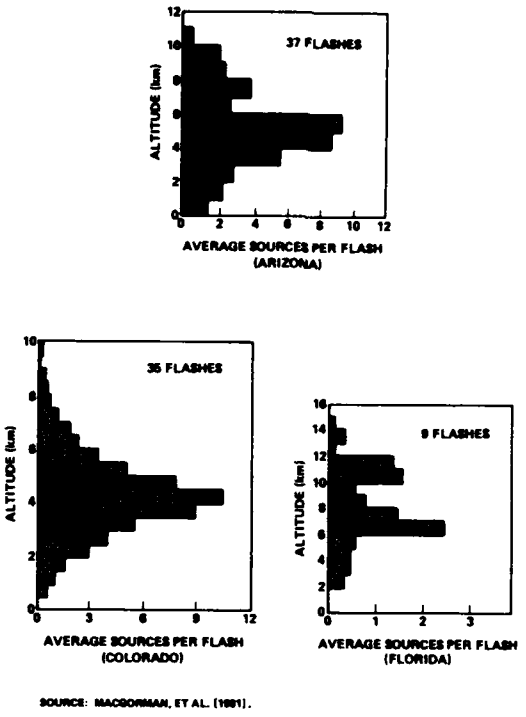


Fig. 12. Altitude distribution of sources for three storms at differing locations.

VLF RADIO NOISE

Michael D. Andrews
Interferometrics Inc.
8150 Leesburg Pike
Vienna, VA 22180

Lorraine DeBlasio and Francis J. Kelly
E. O. Hulbert Center for Space Research
U. S. Naval Research Laboratory
Washington, D.C. 20375

SUMMARY

The Ionospheric Effects Branch of the Space Science Division at the Naval Research Laboratory is engaged in an ongoing program to measure the natural noise environment at VLF. A Navy P-3 aircraft has been used to measure the ambient noise at 35 kHz. There are three orthogonal loop antennas so that both TE and TM components of the noise can be measured.

Mean noise levels and values for V_d will be presented for 1620 observations on seven days in July and August, 1986. Examples of the observed APDs are shown. This experiment observed values for V_d significantly lower than the C.C.I.R. values which probably indicates that these observation were heavily influenced by local noise rather than atmospherics.

INTRODUCTION

The Ionospheric Effects Branch of the Space Science Division of the Naval Research Laboratory (NRL) is engaged in an ongoing effort to study VLF/LF propagation. Figure 1, prepared by Mr. C. Smith of MITRE Corp, presents a summary of reported VLF/LF experiments. There is a clear lack of validation data for Arctic latitudes. This is particularly true for the TE component of signals. There are at least two areas in which additional data would aid in the design of communication systems. First, validation data would allow greater confidence in propagation modelling. Second, the noise characteristics must be known for proper implementation of noise compensation techniques.

At NRL, we are working to gather data to begin to fill the gaps in the data base. Ground stations were installed in September, 1985 at Andoya and Spitzbergen in Norway. Each station is equipped with a whip antenna and a loop antenna which are used to moniter several fixed site VLF/LF transmiters. In addition, several "empty" frequencies are monitered for measurement of noise level.

The experiment to be discussed in this paper was conducted on a Navy P-3 aircraft. A series of flights were conducted to measure the TE and TM components of both propagated signals and noise. To the best of our knowledge, this is the first report of the characteristics of TE noise at Arctic latitudes.

EXPERIMENT DESCRIPTION

A P-3 Orion aircraft was used to measure the TE and TM components of both noise and signals. The aircraft is operated by

NRL for use in scientific projects. NRL and Navy personnel supplied all of the support needed to complete this project.

Hardware

A block diagram of the experiment is shown in Figure 2. There were four antennas installed on the aircraft. The Long Wire (LW) antenna was a single wire strung from a point forward of the aircraft center to a point on the tail. The response of this antenna is similar to that of a short vertical antenna. However, the LW is somewhat sensitive to all components of the electric field. The unamplified signal from the LW antenna was of sufficient strength that a preamplifier was not necessary.

Three magnetic loop antennas were installed in an orthogonal configuration inside the tail near the rear of the aircraft. The loop antennas each contained twenty turns of wire inside a torus of conducting, nonmagnetic material. These antennas measured only the magnetic components of the field. The antennas were placed in a mounting in which the axes of the loops were oriented fore-to-aft, across the fuselage, and vertically. These antennas will be refered to as FA, CR, and VE respectively. The FA and CR antennas were sensitive to the Transverse-Magnetic (TM) field while the VE antenna measured the Transverse-Electric (TE) field. The loop antennas were connected to a specially designed low noise preamplifer with a nominal gain of 30dB mounted on the antenna rack in the rear of the aircraft.

The signals from all four antennas were passed through an NRL designed highpass filter. All frequencies below about 5 kHz were strongly attenuated by this filter which was necessary to suppress harmonics of 60 and 400 Hz which originate in the aircraft power system. The signals from all four antennas were passed in parallel to three HP Model 59307A VLF switches. Each switch, under computer control, passed one of the four antennas' signals.

There were three separate data paths, one from each of the switchs. To insure that the paths were indeed separate, the output of each switch was directed to the input of an Accudata 122 DC amplifier with a nominal gain of 34 dB. These amplifiers were required to properly buffer the signals as they were passed to the various receivers.

The signal from switch 1 went to an HP 3596B Selective Level Meter (SLM) which operated under computer control. The SLM was used to moniter several preselected frequencies with a bandwidth of 20 Hz.

The signal from switch 2 was sent to a Stodart NM12T noise receiver. The observing frequency of the NM12 was controlled by a Larch Adret frequency synthesizer which served as a stable external L.O. The analog output of the NM12 was passed to channel 1 of an NRL 12 bit A/D converter and, the digitized output was recorded by the computer. Switch 2 was never actually switched; the NM12 monitored the LW antenna at all times.

The signal on branch three was sent in parallel to two separate receivers. The first instrument was an HP 3562A Spectrum Analyzer (SA). The SA was controlled by the computer. The second instrument was a Stodart NM17/NM27A noise receiver. The NM17 was manually tuned to 35 kHz by turning a wheel on the front panel. The output of the NM17 took the form of analog data from a 100Hz bandpass in which the output voltage was proportional to the logarithm of the signal voltage. The analog output of the NM17 was passed through the second channel of the A/D converter and recorded by the computer.

There were several additional devices interfaced with the computer, an HP 2113B model of the HP1000. There was an NRL designed hardware interface to the Litton Internal Navigation System (INS). This allowed the data as recorded by the computer to be labeled with the aircraft's latitude, longitude, and heading. A HP 59309 system clock was in place to provide accurate time tagging of the

data. Finally, there was a line printer and a plotter available for printout.

The P-3 aircraft is a hostile environment for electronic equipment. Spares were carried on the aircraft of all the hardware except the SLM and the NM17. It was not believed that a spare SLM would be needed and there was no second NM17 available. The SA proved to be unreliable. At various times there were three of these instruments on the aircraft and all failed on more than one occasion. Fortunately, there was always one working.

Observing Strategy

The data collection was controlled by software on the HP1000 computer. There was a master program which called and controlled six other programs. The computer had three principal data taking programs and software which allowed the time and aircraft position to be entered into the data header.

The first program controlled the SLM, via an IEEE 488 bus, and the setting of switch 1. A set of sixteen frequencies, listed in Table 1, were measured with a 20 Hz bandwidth in order of increasing frequency. The switch was set to the next position after the highest frequency was measured. The meter reading for each frequency and the header information was recorded on disk files.

A second program controlled the data acquisition from the SA. The frequency range of 10-90 kHz was observed with 10 Hz resolution by recording a series of ten 8 kHz wide spectra. Switch 2 was advanced after each series was completed.

A third program controlled data aquisition from the NM17 so that it actually operated as a noise receiver. The 100 Hz bandwidth signal from the A/D converter was sampled at a 200 Hz rate. Every four seconds the accumulated data were written to tape. This program continued to read and record data for approximately 90 seconds while switch 2 remained on a single antenna. The data was also binned as read into 100 boxes covering the range of 60 dB. This program also processed the data so that at antenna switch a disk record was written that contained the header information, the mean, the ratio of the mean to the RMS, and the fraction of the data that fell into each of the 100 bins.

The NM17 was controlled by adjustments made on the front panel of the instrument. 35 kHz was selected as the observing frequency because there were no known transmitters operating near this frequency. The NM17 was tuned by injecting a 35 kHz signal from a frequency synthesizer and adjusting a knob on the front panel to maximum meter reading. The frequency of the NM17 did vary somewhat and the actual tuning may have varied by 100-200 Hz during any single flight. The gain of the NM17 was also adjusted from the front panel. The gain control was set so that the output voltage would not exceed the linear range of the instrument which was verified to be 60 dB. As a result the low level noise was not properly measured and in some cases the lowest bin contained a large portion of the data. It is our intention to obtain a noise meter with a substantially higher dynamic range for future experiments.

Flights

There were seven flights in which all of the equipment worked and high quality data were obtained. All of the data were collected during daylight hours. Flight 1 was the first successful flight and took place on 31 July 1986, day number 216. The aircraft flew northeast from the Pautuxant River Naval Air Station, Lexington Park, Maryland, to near the southern tip of Greenland and returned along the same path. The flight paths for six of the seven paths are shown in Figure 3. Flight 2 on 6 August, day 218, was from Pax to Puerto Rico and return, and cannot be easily displayed on one plot with the other six flights.

The actual series of Arctic flights began with flight 3 on 8 August, day 220. The flight path was up along the East coasts of the USA and Canada to the latitude of Iceland and then due East over Southern Greenland to Keflevic, Iceland. Flight 4 on 10 August, day 222, consisted of a ragged path over water in the North Atlantic. Flight 5 on 11 August, day 222, proceeded due North from Iceland to about 86 N. latitude and then roughly west to Fairbanks, Alaska. On 14 August, day 226, flight 6 was flown over Northern Alaska and the Arctic Ocean. Flight 7 on 15 August, day 227, proceeded on a direct path from Fairbanks to Pax River.

RESULTS

There was a substantial quantity of high quality data obtained from the seven flights. While the data from the SLM and the SA are not the subject of this presentation, a couple of examples will be presented before turning to the noise data which are the subject of the remainder of this paper.

There are roughly 50 hours of data from the SLM. Figure 4 is an illustrative plot produced from the SLM data. Shown is the variation of absolute field strength versus distance of the 21.4 kHz transmitter located in Annapolis, Maryland, for the flight on day 220. There were approximately 20,000 8 kHz spectra recorded from the SA. An example is shown in Figure 5. The signals from the Annapolis, Maryland, and Cutler, Maine, transmitters are clearly seen.

There were 1620 Amplitude Probability Distributions (APDs) calculated from the NM17. Each APD was calculated from roughly 18,000 individual measurements of the voltage amplitude. For each APD there is also a value of the mean noise level, and a value for V_d, the ratio in dB of the mean to the RMS.

Figures 6-12 are plots of mean noise level and V_d versus Universal Time, UT, for each of the individual flights. Each of the four antennas is identified by a different symbol as listed in the title. The noise level is usually highest on the LW antenna. The CR loop has the second highest noise level. The FA and VE loops have either comparable noise level or the FA loop is slightly higher.

The noise level is typically determined by local weather conditions. The weather was not good for any of these flights. The aircraft was often flying in the cloud layer. This resulted in increased noise level particularly on the LW antenna. The aircraft flew at lower altitude in the early and late portions of each flight. The aircraft was also over land at the beginning and end of each flight. The higher noise level seen at the beginning and end of several flights is probably due to the altitude and overland location of the aircraft. At this point in the analysis, local weather is the only factor that is clearly affecting the noise observations.

The V_d portions of these figures are more complicated. Often the LW antenna shows the lowest values of V_d. This may indicate that this noise is originating on the aircraft. This would not be surprising since the LW would be expected to "hear" the aircraft engines. The VE loop also often shows very low values of V_d. It is possible that the VE loop is not sufficiently sensitive to observe the ambient noise level and that most of the data represents instrument noise. The values of V_d on the FA and CR loops are much more variable. At certain times, such as day 212 around 1800 UT, the FA and VE loops show very similar noise levels but significantly different values of V_d.

The values of V_d observed on these flights are consistently lower the then the value contained in C.C.I.R. Report 322, "World Distribution and Characteristics of Atmospheric Radio Noise". This report lists average values of V_d of approximately 13 for

summer mornings to 8 for summer evenings. The values of V_d in Figures 6-12 are for a bandwidth of 100 Hz and should be multiplied by 4/3 before comparison to the C.C.I.R. values which are for a 200 Hz bandwidth. Most of our observed values of V_d are significantly less than 8 after this conversion. One possible explanation of this is that, if noise from thae aircraft dominates then, low values of V_d could be expected.

Figures 13-15 are three examples of APD plots on the long wire antenna. Figure 13 is a typical APD for temperate latitudes, and Figure 14 is a typical APD for the northern most positions. At northern latitudes, the noise level is lower and the dynamic range of the noise is smaller. The value of V_d at latitude 85 degrees is 1.7 dB compared with 2.2 dB at 40 degrees latitude. This may imply that airplane noise is a larger contributor at far northern latitudes.

Figure 15 is a plot of the APD recorded while flying near a severe thunderstorm near Boston, Massachusetts. The noise level is elevated and the APD curve is obviously distorted. The value of V_d for this plot was 2.1 dB which is typical of this antenna. However, as can be seen from Figure 8, at around 1600 UT many of the values of V_d fell in the range of 4-6 dB which after bandwidth correction would be in fair agreement with the C.C.I.R. values. This again may indicate that at most time the aircraft is making a large contribution of the noise, and that we are not actually monitoring atmospheric noise.

Table 1

Frequency (kilohertz)	Transmitter Location
16.40	Noviken, Norway
19.00	Anthorne, Scotland
21.40	Annapolis, Maryland
22.30	Australia
23.40	Lualualei, Hawaii
24.00	Cutler, Maine
24.80	Jim Creek
28.50	Aguada, Puerto Rico
37.20	Hawes AFB, California
48.50	Silver Creek
51.60	Annapolis, Maryland
55.50	Thurso, Scotland
57.40	Iceland
59.00	Greece
77.15	Driver, Virginia
88.00	Annapolis, Maryland

Figure 1

VLF/LF Model Validation Data

- ⌖ DCA TACAMO Air-Ground 1975-77
- □ DCA/ESDTE/TM Test Program, 1975-76
- △ RADC, 1974
- ○ RADC, 1970
- ◇ NRL, 1965
- ◆ RAF, 1970
- ● Multi-Frequency Sounder, (Hawaii) 1968
- ■ Morgan (WWVB/WWVL), 1964-66
- ▲ Other data collected by NELC
- ✳ NWC, 1970
- ✕ ITS, 1970

"Short" range = less than 3000 km

			TE						TM					
			Summer		Fall/Spring		Winter		Summer		Fall/Spring		Winter	
			Day	Night	Day	Night	Day	Night	Day	Night	Day	Night	Day	Night
14-30 kHz	Arctic	Long							▲⌖	⌖	⌖	⌖	⌖	⌖
		Short							▲				△	△
	Temperate	Long	□	□	□		□		◇□⌖✳	□⌖✳	□⌖●✳	●⌖✳	□⌖●	△⌖●
		Short	□		□		□	△□	◇▲⌖	⌖	△■□⌖	△⌖	△■□⌖	●⌖△□
30-45 kHz	Arctic	Long							✕	✕			○✕	○✕
		Short							◆				△	△
	Temperate	Long									●	●		△
		Short			□		□				□		△□	△
45-60 kHz	Arctic	Long												
		Short							◆				△	△
	Temperate	Long									●	●		△
		Short									■		△■	△

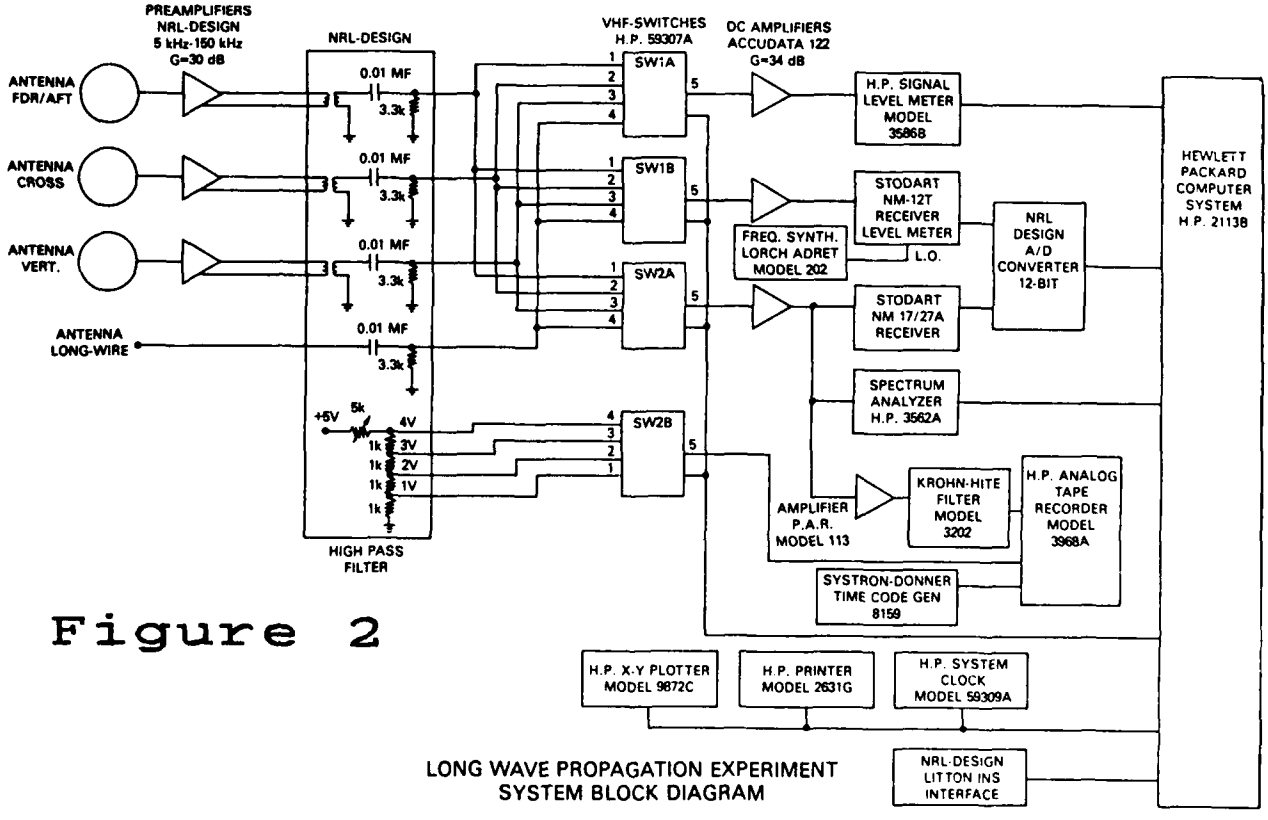


Figure 2

LONG WAVE PROPAGATION EXPERIMENT
SYSTEM BLOCK DIAGRAM

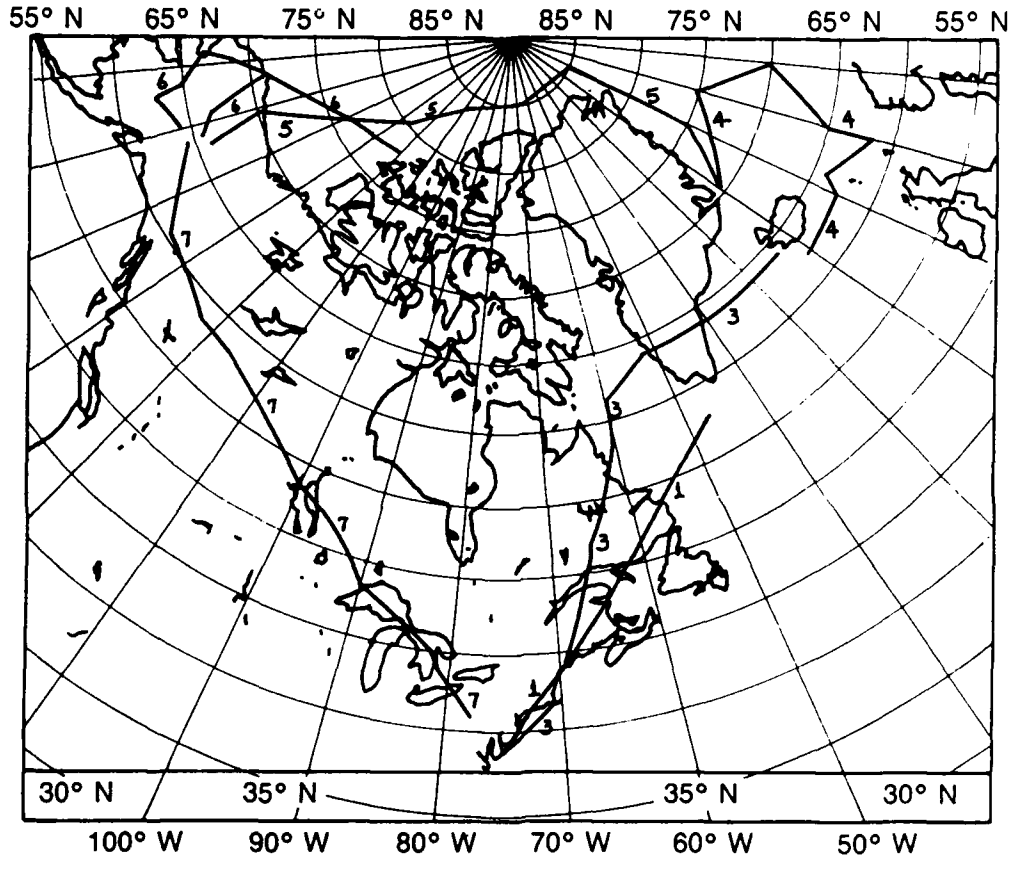

55° N
65° N
75° N
85° N
85° N
75° N
65° N
55° N
30° N
35° N
35° N
30° N
100° W
90° W
80° W
70° W
60° W
50° W

Figure 3

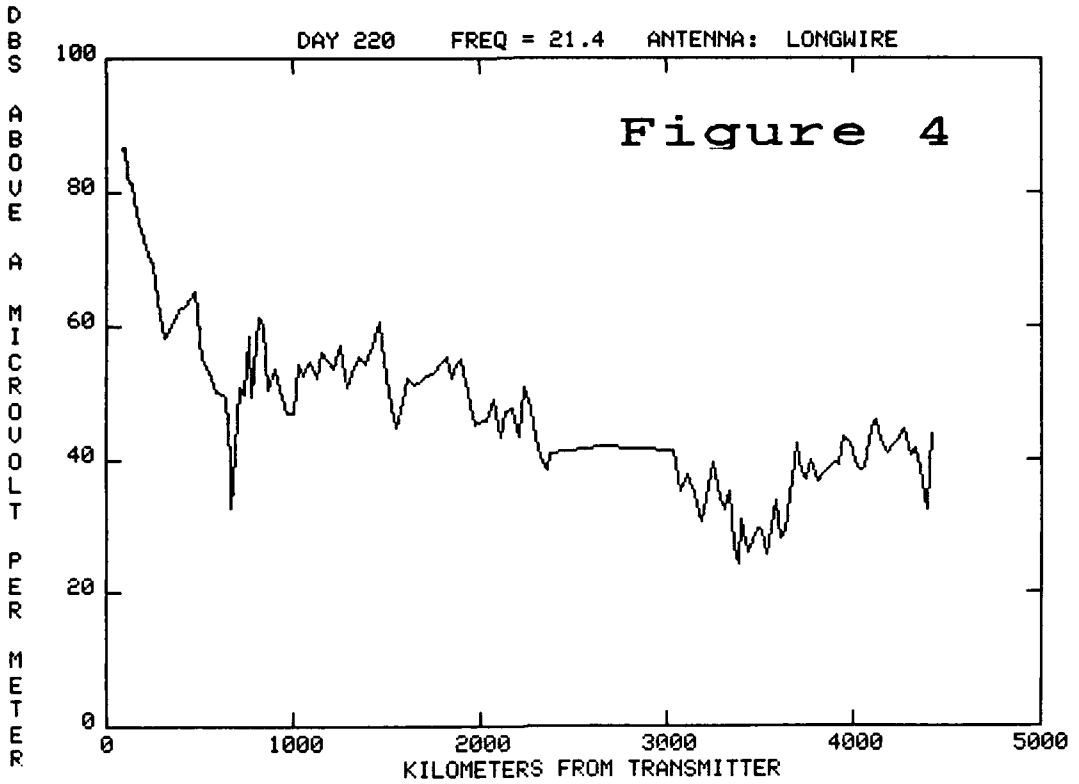


Figure 4

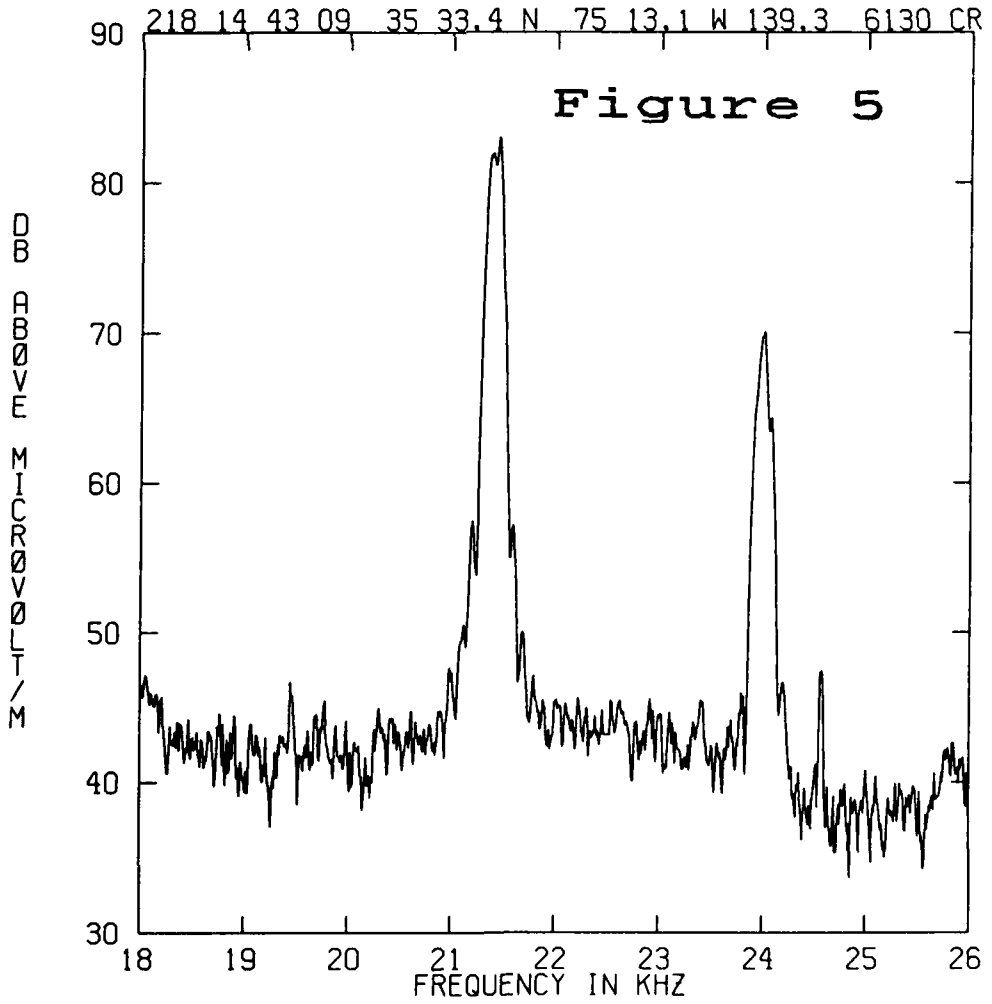


Figure 5

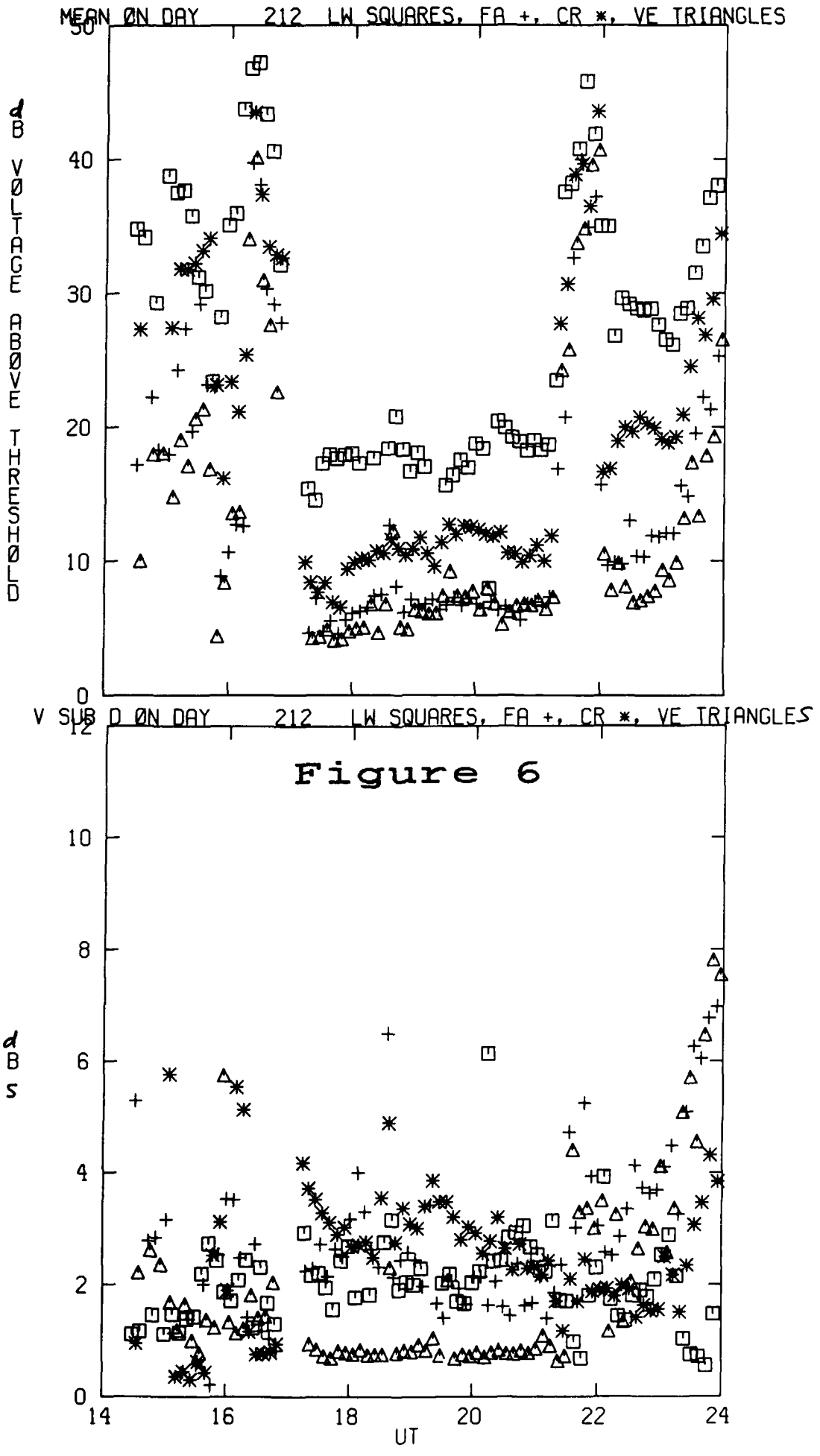

MEAN ØN DAY 212 LW SQUARES, FA +, CR *, VE TRIANGLES
dB VØLTAGE ABØVE THRESHØLD
0
10
20
30
40
50
V SUB D ØN DAY 212 LW SQUARES, FA +, CR *, VE TRIANGLES
dBS
0
2
4
6
8
10
12
14
16
18
20
22
24
UT

Figure 6

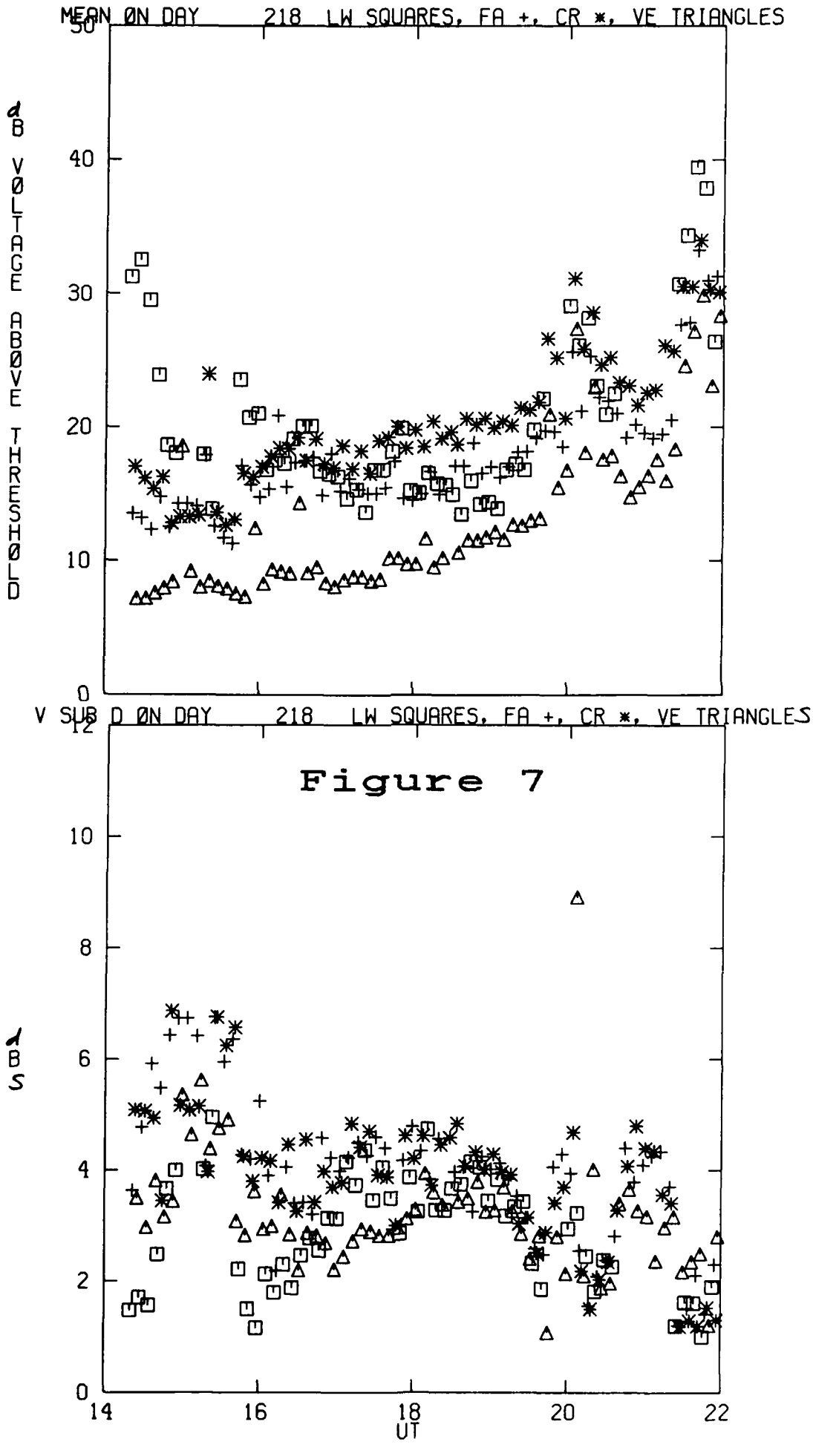

MEAN ØN DAY 218 LW SQUARES, FA +, CR *, VE TRIANGLES
dB VØLTAGE ABØVE THRESHØLD
50
40
30
20
10
0
V SUB D ØN DAY 218 LW SQUARES, FA +, CR *, VE TRIANGLES
dBS
12
10
8
6
4
2
0
14
16
18
20
22
UT

Figure 7

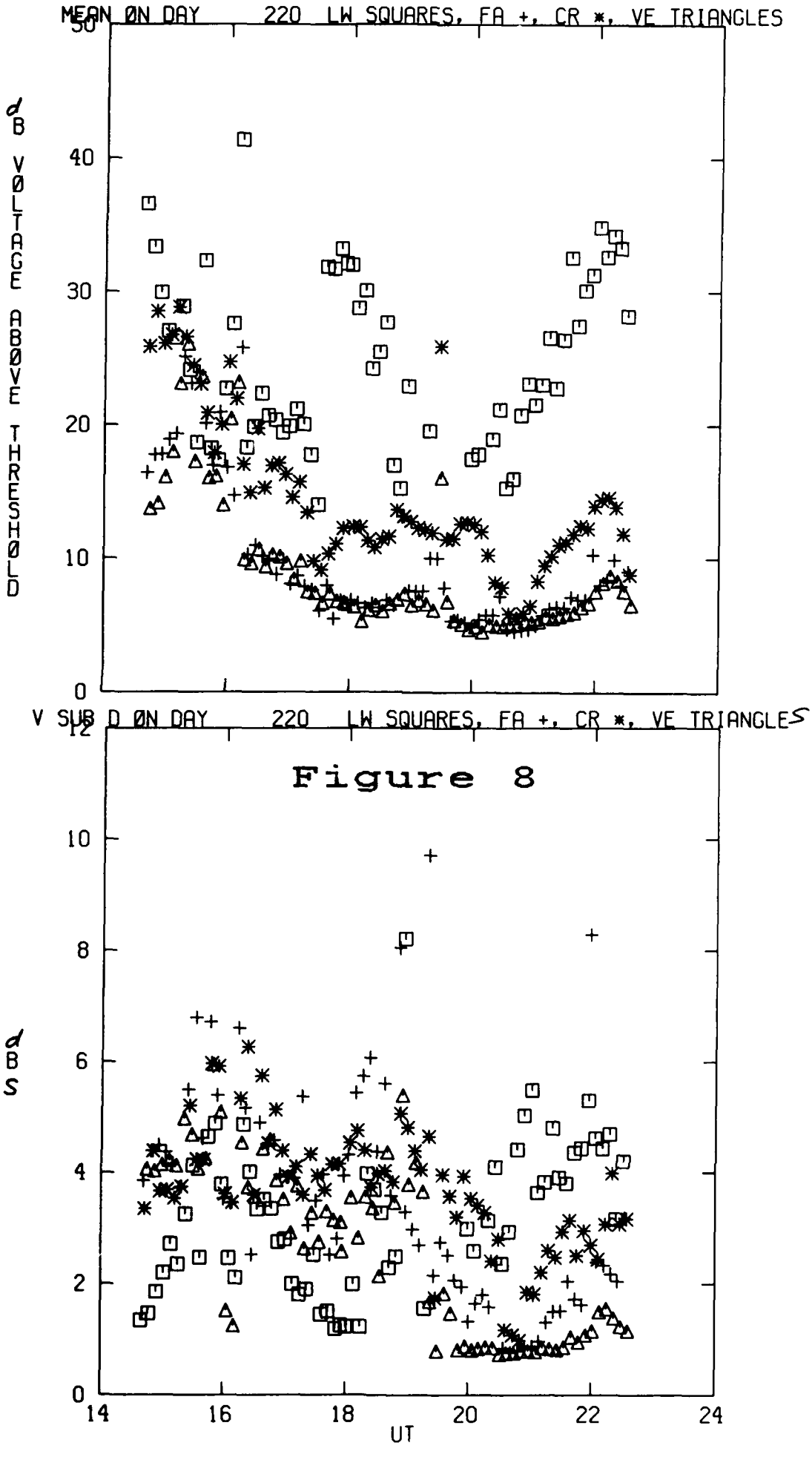

MEAN ØN DAY 220 LW SQUARES, FA +, CR *, VE TRIANGLES
dB VØLTAGE ABØVE THRESHØLD
50
40
30
20
10
0
V SUB D ØN DAY 220 LW SQUARES, FA +, CR *, VE TRIANGLES
dBS
12
10
8
6
4
2
0
14
16
18
20
22
24
UT

Figure 8

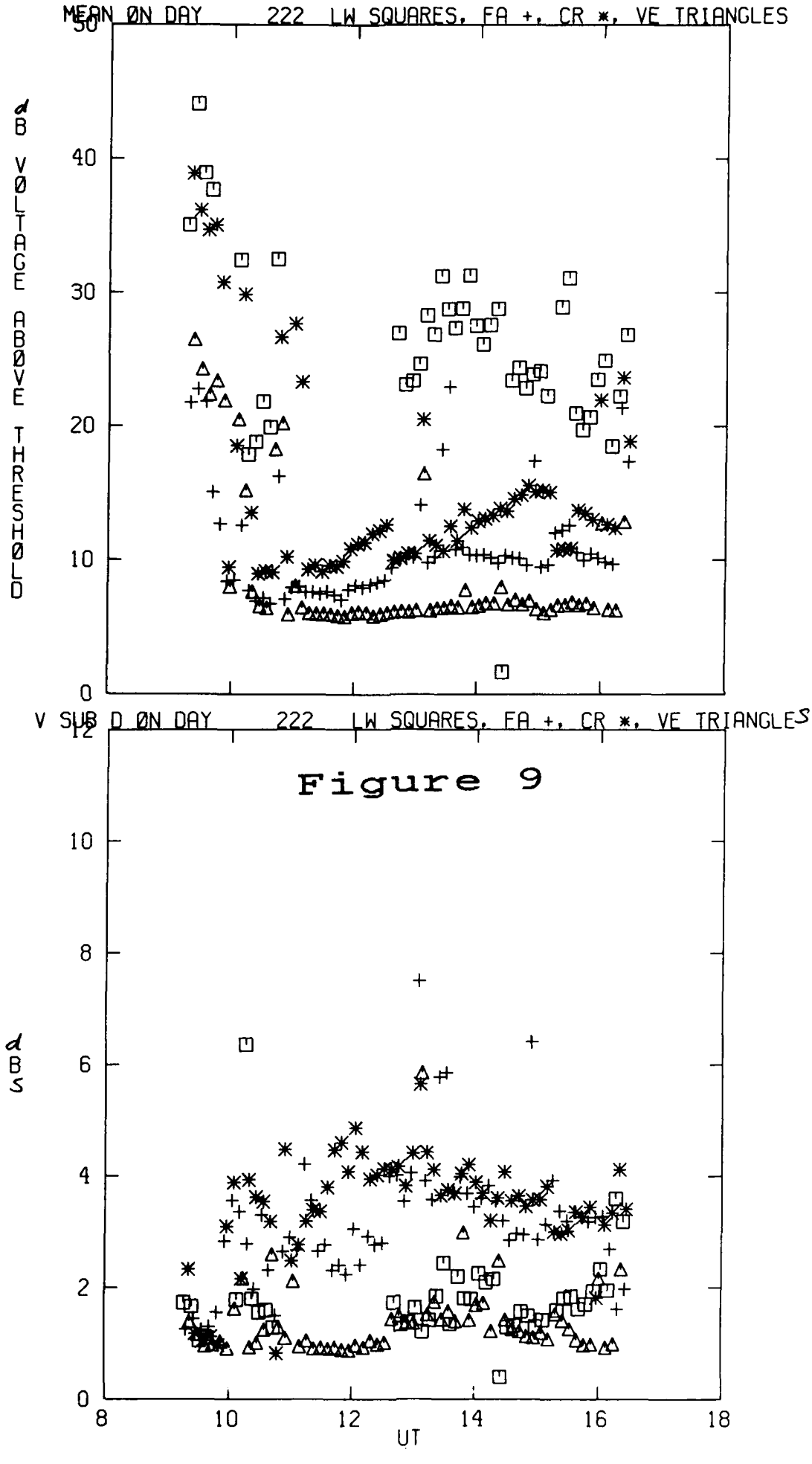
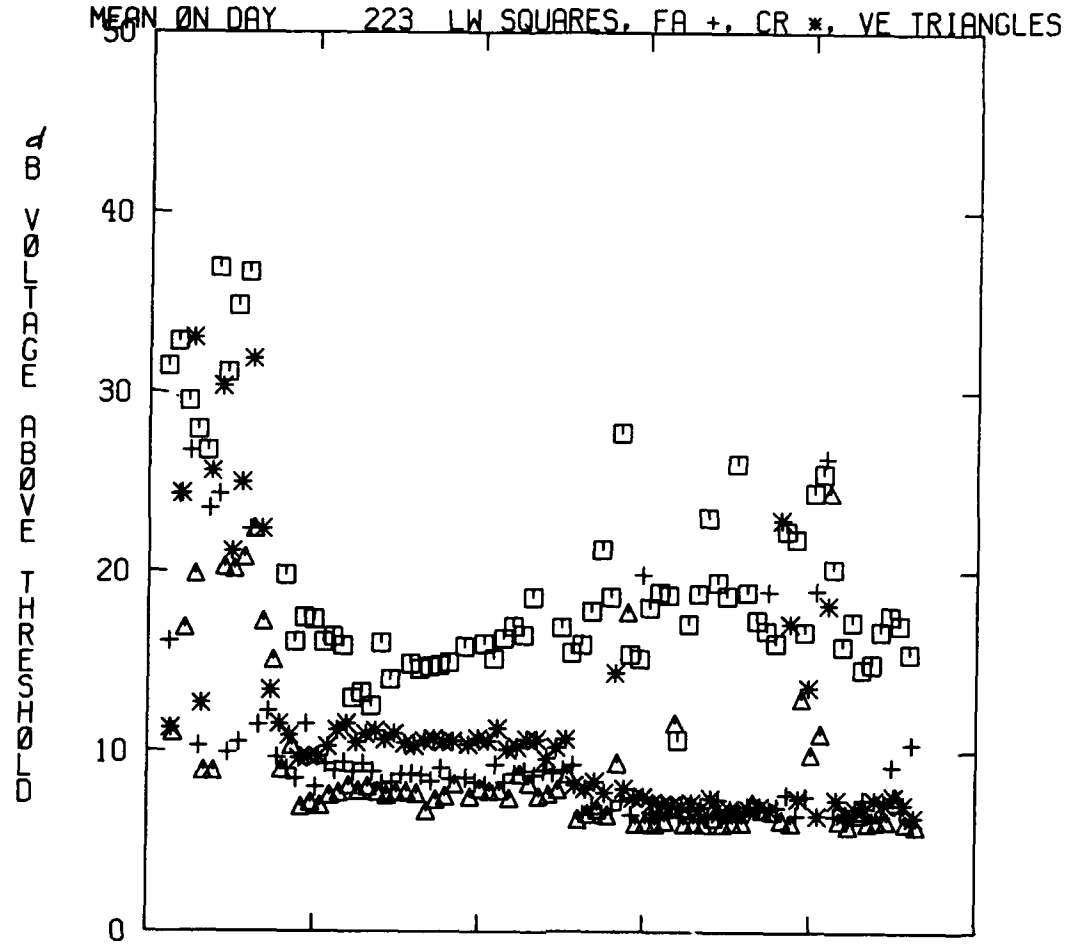

MEAN ØN DAY 222 LW SQUARES, FA +, CR *, VE TRIANGLES
dB VØLTAGE ABØVE THRESHØLD
0
10
20
30
40
50
V SUB D ØN DAY 222 LW SQUARES, FA +, CR *, VE TRIANGLES
dBS
0
2
4
6
8
10
12
8
10
12
14
16
18
UT

Figure 9


MEAN ØN DAY 223 LW SQUARES, FA +, CR *, VE TRIANGLES
dB VØLTAGE ABØVE THRESHØLD
50
40
30
20
10
0

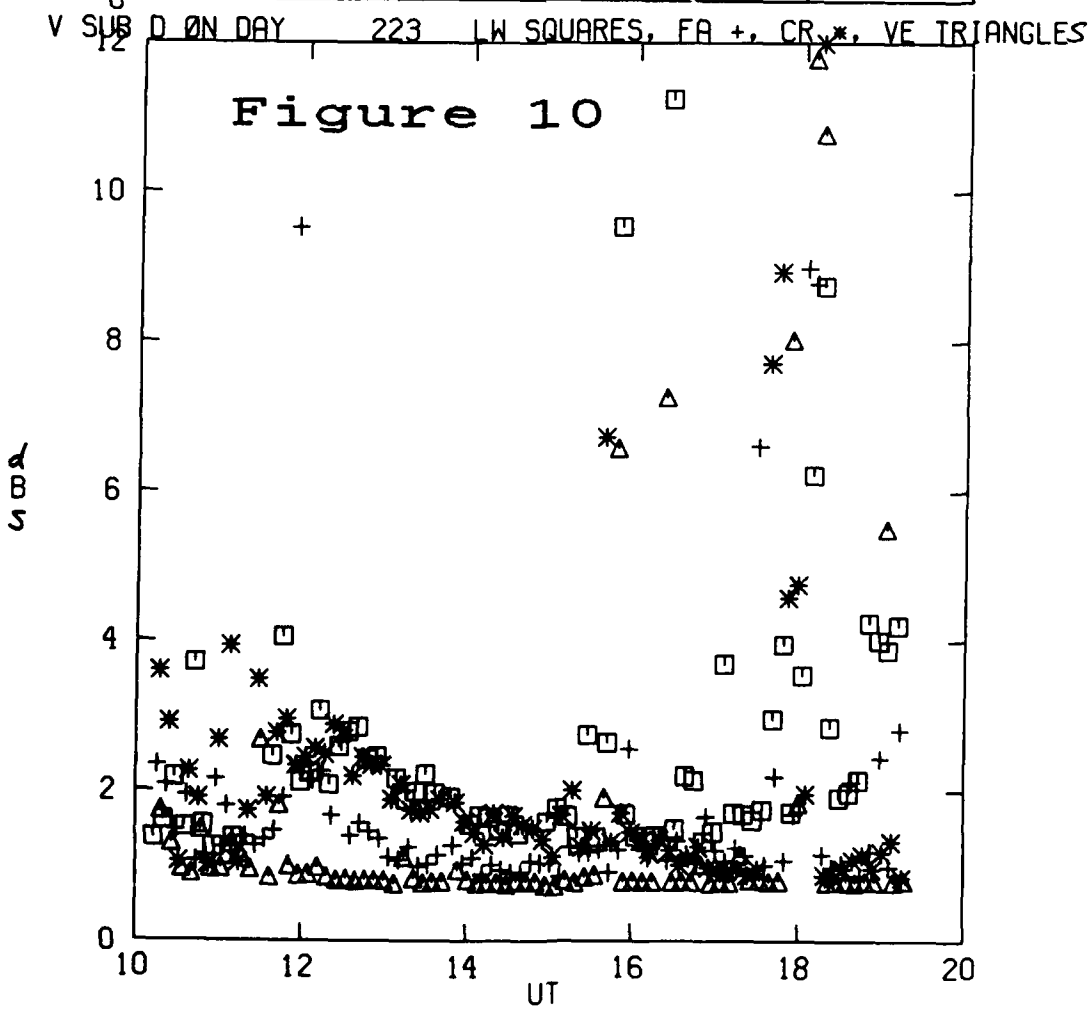

V SUB D ØN DAY 223 LW SQUARES, FA +, CR *, VE TRIANGLES
dB S
12
10
8
6
4
2
0
10
12
14
16
18
20
UT

Figure 10

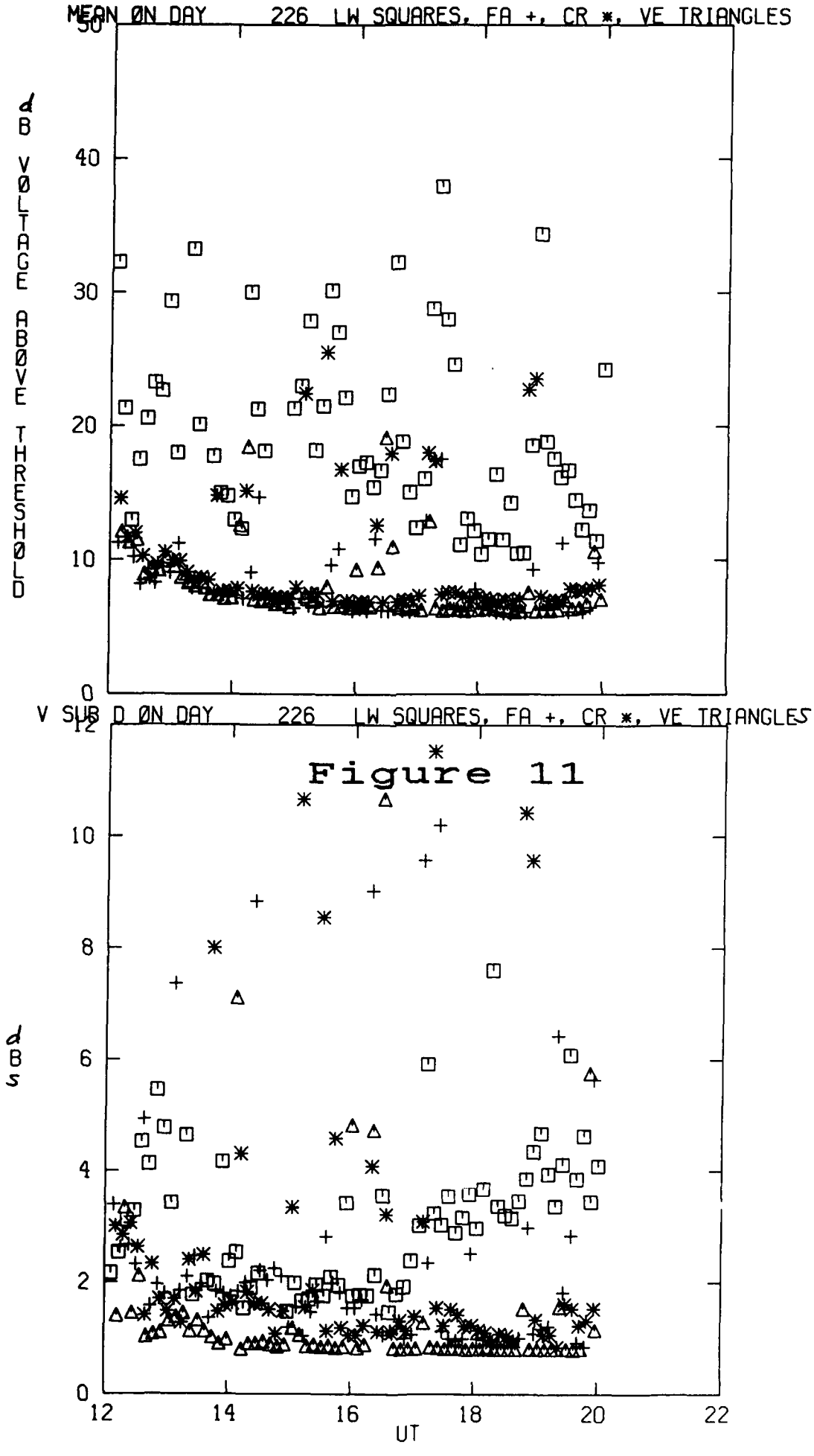

MEAN ØN DAY 226 LW SQUARES, FA +, CR *, VE TRIANGLES
dB VØLTAGE ABØVE THRESHØLD
0
10
20
30
40
50
V SUB D ØN DAY 226 LW SQUARES, FA +, CR *, VE TRIANGLES
dBs
0
2
4
6
8
10
12
12
14
16
18
20
22
UT

Figure 11

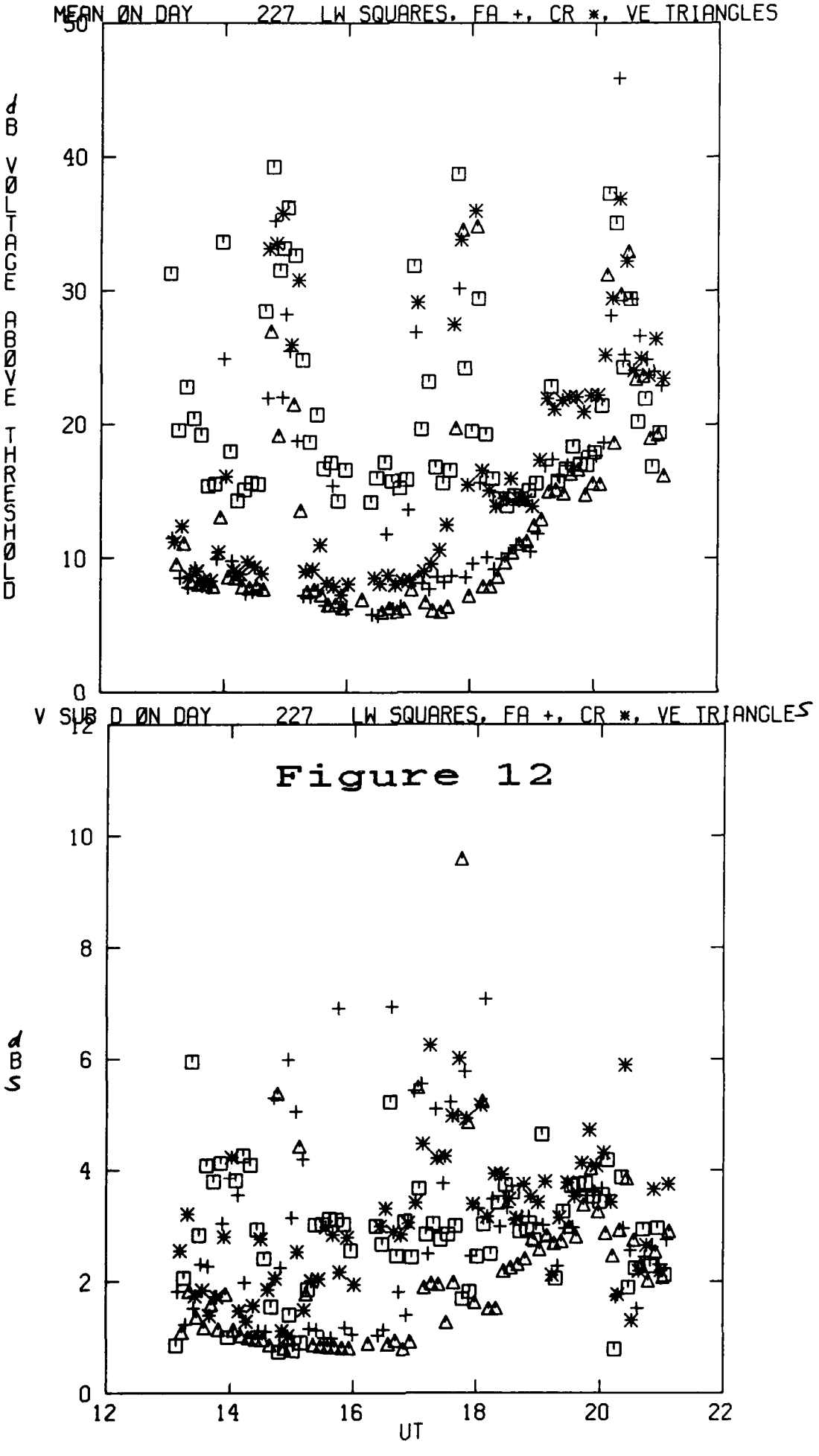

MEAN ØN DAY 227 LW SQUARES, FA +, CR *, VE TRIANGLES
dB VØLTAGE ABØVE THRESHØLD
0
10
20
30
40
50
V SUB D ØN DAY 227 LW SQUARES, FA +, CR *, VE TRIANGLES
dBS
0
2
4
6
8
10
12
12
14
16
18
20
22
UT

Figure 12

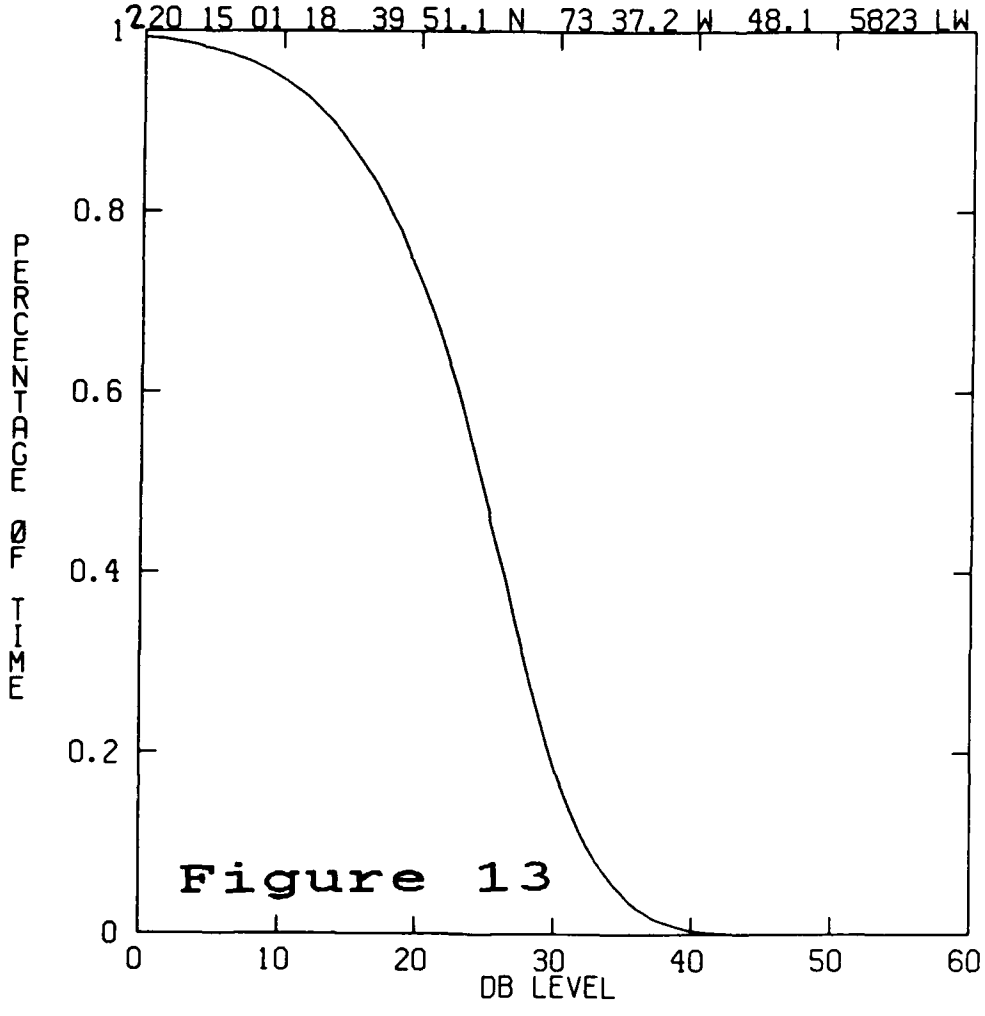

220 15 01 18 39 51.1 N 73 37.2 W 48.1 5823 LW
1
0.8
0.6
0.4
0.2
0
PERCENTAGE OF TIME
0
10
20
30
40
50
60
DB LEVEL

Figure 13

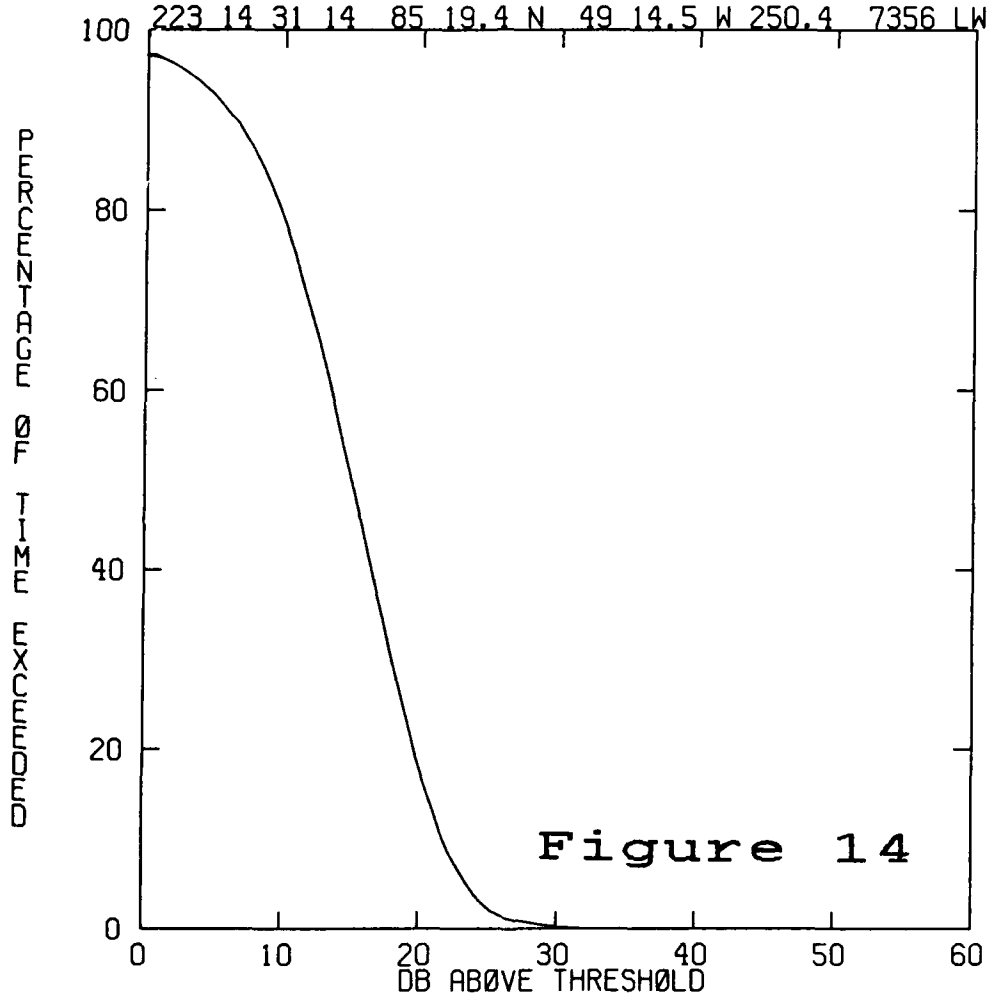

223 14 31 14 85 19.4 N 49 14.5 W 250.4 7356 LW
100
80
60
40
20
0
PERCENTAGE ØF TIME EXCEEDED
0
10
20
30
40
50
60
DB ABØVE THRESHØLD

Figure 14

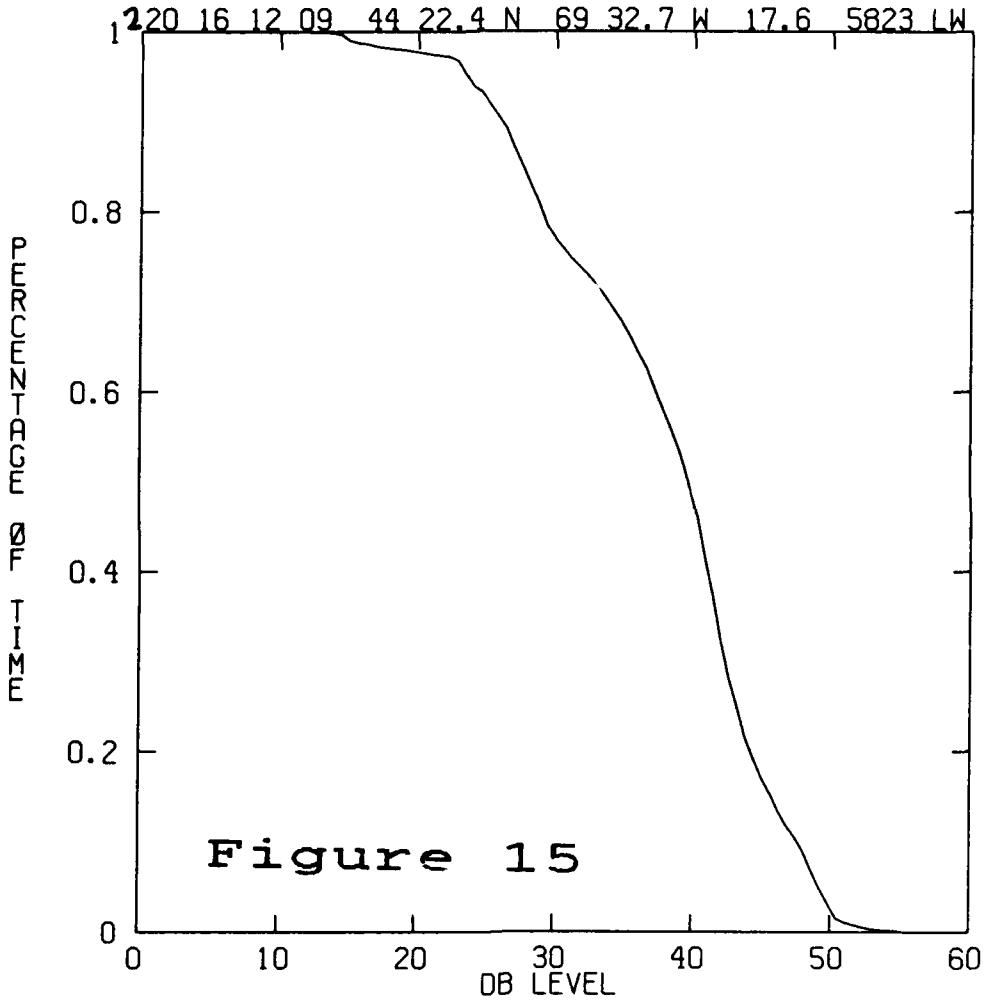

220 16 12 09 44 22.4 N 69 32.7 W 17.6 5823 LW
1
0.8
0.6
0.4
0.2
0
PERCENTAGE OF TIME
0
10
20
30
40
50
60
DB LEVEL

Figure 15

A NEW GLOBAL SURVEY OF ELF/VLF RADIO NOISE

A.C. FRASER-SMITH R.A. HELLIWELL, B.R. FORTNAM,
P.R. MCGILL, AND C.C. TEAGUE
Space, Telecommunications and Radioscience Laboratory
Stanford University
Stanford, California 94305

SUMMARY

Stanford University is presently operating a global network of eight computer-controlled receiving systems, or *radiometers*, for the measurement of electromagnetic noise in the 10–32,000 Hz (ELF/VLF) frequency band. Each radiometer consists of two dual-channel receivers, each with two crossed loop antennas (East-West, North-South). One of these receivers has a response covering the frequency range 10–500 Hz and the other covering the range 200–32,000 Hz. A bank of narrow-band filters (5% bandwidth) is used to monitor the noise present at 16 selected frequencies throughout the overall range of operation. The output of these filters is continuously sampled and statistical averages computed and recorded on magnetic tape, along with samples of the raw data. Broadband samples of the ELF/VLF noise are also taken periodically (typically one minute every hour) to provide a check on system performance and to aid in the interpretation of the statistical data. The digital data are being used primarily for statistical studies of the global distribution of ELF/VLF noise. Plots of minimum, maximum, average, and rms noise amplitudes versus frequency can be comparatively easily prepared for a wide variety of time intervals for each radiometer location. Plots of the noise statistic V_d and amplitude probability distributions are also easily prepared. The broadband data can be used for detailed studies of the spectral properties of the noise and for the identification of signals and sources.

1. INTRODUCTION

Three decades ago, W. Q. Crichlow began an article describing the inauguration of an ambitious study of radio noise by the U.S. National Bureau of Standards (NBS) with the following words: "Radio noise constitutes the basic limitation to radio reception. A determination of the weakest possible signal that will provide satisfactory service necessitates a knowledge of the noise with which it must compete" [*Crichlow*, 1957]. These words are still applicable today and make a fitting start to this description of a new survey of radio noise.

Recognizing that the available information on the surface distribution of terrestrial ELF and VLF radio noise was becoming outdated and that there were, in any event, large gaps in the information available, the Space, Telecommunications and Radioscience Laboratory at Stanford University commenced a new study of the global distribution of ELF/VLF noise in 1981. Support for this study was initiated by the Office of Naval Research (ONR); important additional support is now being provided by the Rome Air Development Center (RADC) and the National Science Foundation (NSF). Some equipment purchases were supported by a Department of Defense instrumentation grant.

The noise measurement systems to be used in the survey (called ELF/VLF *radiometers*) were designed and constructed at Stanford, the latter effort being assisted by a number of graduate and undergraduate student engineers. Originally seven radiometers were planned, but eight were ultimately constructed as a result of the support from RADC for the radiometer that is now installed at Thule, Greenland, near the north geomagnetic pole. The first radiometer to be completed was installed at Arrival Heights, Antarctica, during the austral summer of 1984-1985, with logistics support from the NSF. The final radiometer to reach operational status was installed at Stanford; it began its noise measurements early in November 1986, and the entire array has been in simultaneous operation since that time. Figure 1 shows the locations of the eight radiometers and Table 1 gives their coordinates.

TABLE 1. Geographical Coordinates of Radiometer Stations

Station	Coordinates
Thule, Greenland	77°N, 69°W
Sondre Stromfjord, Greenland	67°N, 51°W
New Hampshire	44°N, 72°W
L'Aquila, Italy	42°N, 13°E
Stanford, California	37°N, 122°W
Kochi, Japan	33°N, 133°E
Dunedin, New Zealand	46°S, 170°E
Arrival Heights, Antarctica	78°S, 167°E

Most of the radiometers are being maintained in operation by cooperating scientists under data sharing arrangements. In Kochi, Japan, the radiometer is being operated in cooperation with Professor Toshio Ogawa of Kochi University. In L'Aquila, Italy, the cooperative arrangement is with Drs Antonio Meloni and Paolo Palangio of the Istituto Nazionale di Geofisica. In Dunedin, New Zealand, the arrangement is with Dr Neil R. Thomson and Professor Richard L. Dowden of the University of Otago, and in Sondrestromfjord the radiometer is maintained by Mr Craig Heinselman as the result of an arrangement with Dr John D. Kelly of SRI International. Mr Michael Trimpi is responsible for the operation of the Grafton, New Hampshire,

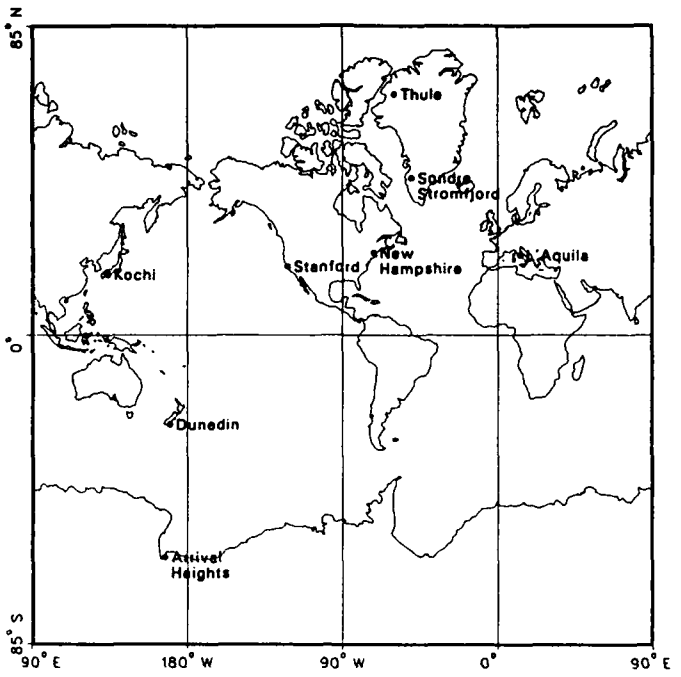


Figure 1. Locations of the eight ELF/VLF radiometer locations on a world map.

radiometer. Operation of the Thule, Greenland, radiometer is under the cognizance of Mr John P. Turtle of RADC, and the Arrival Heights, Antarctica, radiometer is maintained in operation under particularly rigorous conditions by scientific personnel provided by the Division of Polar Programs of the National Science Foundation.

The instrumentation included in each radiometer has been described fully elsewhere [*Fortnam*, 1985; *Fraser-Smith and Helliwell*, 1985], and no detailed description will be given here. Each radiometer consists of two dual-channel receivers, each with two crossed loop antennas (East-West, North-South). One of the receivers has a response covering the frequency range 10–500 Hz and the other covering the range 200–32,000 Hz. A bank of narrow-band (5% bandwidth) filters is used to monitor the noise present at 16 selected frequencies distributed approximately uniformly in a logarithmic sense through the overall frequency range of operation (Figure 2). These frequencies were carefully chosen to avoid harmonics of the 60 Hz and 50 Hz power line frequencies. The output of these filters is continuously sampled and a variety of statistical quantities calculated and recorded digitally on magnetic tape, along with samples of the raw data from the filters (typically one sample per second for all 16 filters). Operation of each radiometer is under the control of a mini-computer, which not only computes the statistical data, but also monitors all essential functions of the radiometer and automatically calibrates the response of the receivers at regular intervals.

The statistical quantities computed continuously during the radiometer operation consist of the root-mean-square (rms), average, maximum, and minimum amplitudes for each of the 16 selected frequencies. They are computed at the end of every minute from 600 amplitude measurements made at a rate of 10 per second on the envelope of the noise signal emerging from each narrow-band filter. These statistical data can be read and listed directly from the digital tape with no further processing required. Later processing of the data can, with little additional computation, give the V_d statistic, which is the ratio in dB of rms to average amplitude and is a measure of the impulsiveness of the noise data. Similarly, and again with only a little additional computation, the sampled data can be used to derive amplitude probability distributions (APD's).

In addition to the data from the narrow band filters, broad-band ELF data, sampled at a rate of 1000 samples per second during scheduled synoptic recording intervals (currently one minute each hour), are also recorded on the digital tape. These data can be converted to spectrograms and they provide an essential check on the operation of the narrow-band filters and their associated measurements. A similar synoptic picture of activity in the VLF range is provided by analog recordings of the range 200–32,000 Hz.

2. PREVIOUS ELF/VLF NOISE MEASUREMENTS

There is an extensive literature on radio noise measurements. *Spaulding* [1982], in a particularly wide-ranging review of the noise and its implications for telecommunication systems, suggests a starting date of 1896 for this literature. However, it was many years before global measurements could be made, and, even then, the frequency ranges covered by the studies typically did not extend far down into the VLF range. Our present knowledge of the worldwide distribution of radio noise is largely based on the results of the NBS study mentioned in the Introduction that was started in 1957 [*Crichlow*, 1957]. The results of this study, which include measurements of (1) the global distribution of average noise power levels, (2) the statistical quantity V_d, (3) APD's, and (4) the seasonal variation of some of the noise characteristics, form much of the basis for Report 322, published by the International Radio Consultative Committee (Comité Consultatif International des Radiocommunications, or CCIR) of the International Telecommunications Union [*CCIR*, 1964]. This report, together with certain updates [CCIR, 1982*a*,*b*], provides what might be termed the 'official' view of radio noise.

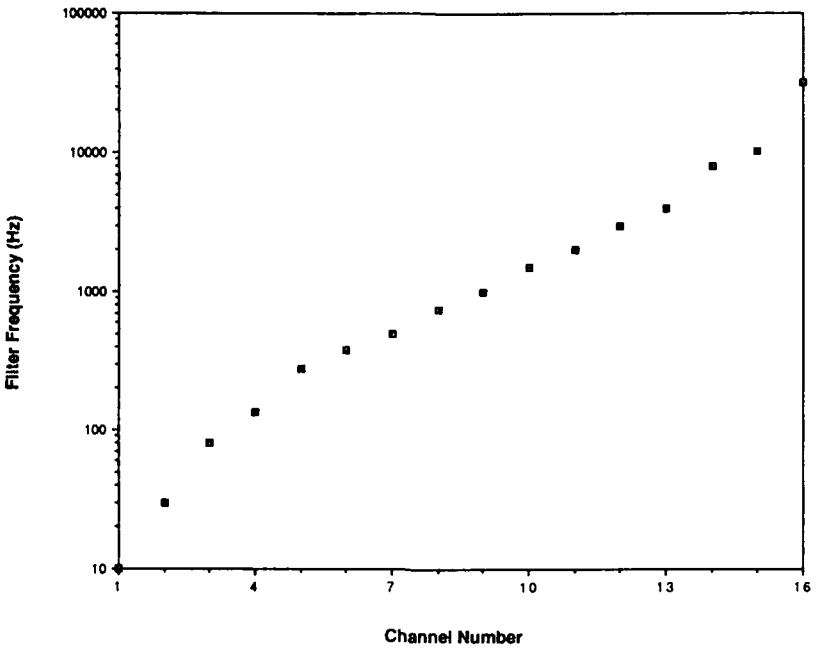


Figure 2. Distribution of the center frequencies of the 16 narrow band filters through the overall frequency range of operation of the ELF/VLF radiometers.

Unfortunately, insofar as the ELF/VLF band is concerned, the available CCIR radio noise information has three major weaknesses. First, it incorporates few measurements below 10 kHz. Second, although two high latitude stations were included in the original noise surveys, it is not clear that the contribution of magnetospheric noise to the high latitude measurements is adequately represented in the summary data. Finally, the timing of the noise data is quite coarse, since the intervals for which the data are presented are four hours long and the summary plots are organized according to the season.

More recent work has helped strengthen our knowledge of radio noise in the ELF/VLF band by partially filling some of the gaps in the CCIR data. For example, *Maxwell and Stone* [1963] and *Maxwell* [1966] provide important information concerning the electric field amplitudes of the noise below 10 kHz, as well as some indication of the variability of the fields. *Watt* [1967] discusses the *CCIR* [1964] data, but also presents additional results for some of the noise statistics [*Watt and Maxwell*, 1957*a,b*]. An independent series of measurements by *Dinger et al.* [1982] help provide a calibration of the VLF noise data available at that time as well as providing new information on the noise in the frequency range 1.0–4.0 kHz. APD's for the ELF/VLF range have been studied quite carefully and an analytical model developed [*Galejs*, 1966, 1967; *Field and Lewinstein*, 1978]. Perhaps most important from the point of view of the work reported here, syntheses of the available information on noise over a broad frequency range, including the ELF/VLF range as a small subsection, have been made which help to place the lower frequency noise into a broader perspective, i.e., the ELF/VLF noise can be viewed as small part of a larger picture of atmospheric radio noise in general [e.g., *Spaulding and Hagn*, 1978; *Smith*, 1982; *Spaulding*, 1982; *Flock and Smith*, 1984; *Fraser-Smith and Helliwell*, 1985].

Although this later work has contributed substantially to our knowledge of ELF/VLF noise, it nevertheless remains true that few measurements have been made of the APD's, V_d, or the other ELF/VLF noise statistics. Similarly, the contribution of magnetospheric and possibly even interplanetary ELF/VLF noise at high latitudes remains to be determined. Finally, with the advent of modern high-speed computers, the timing, statistical significance, and other computational features of the noise data can be greatly improved. The study discussed here was undertaken with the view of filling these gaps.

3. RECENT DATA

In this section we present examples of the various forms of data being acquired as part of the Stanford ELF/VLF noise survey. Figures 3 and 4 show some of the standard data produced by the radiometers, in this case the rms amplitudes for 3 February 1987 as measured at Arrival Heights, Antarctica. Each of the figures shows the rms amplitudes for 8 of the 16 narrow band channels; frequencies 10 Hz to 750 Hz are covered in Figure 3 and frequencies 1 kHz to 32 kHz in Figure 4. Each individual plot shows the variation of the 1440 one-minute rms amplitudes (computed from a nominal 600 envelope measurements made at a rate of 10 per second) over the 24-hour UT day 3 February 1987 for one particular narrow band frequency channel. Each individual plot incorporates a nominal 8.64×10^5 measurements, and the 16 plots for the one day incorporate a nominal total of 1.38×10^7 measurements. In practice some measurements are lost during the calibration intervals that occur automatically every five minutes; when the calibration losses are taken into account the actual measurements incorporated into each one day plot are roughly 5% less than nominal for the first six narrow band frequency channels and roughly 1% less than nominal for the remaining ten channels.

Since we use loop antennas, our calibrated amplitudes are given in magnetic field units. We normally use magnetic induction B to characterize the magnetic fields of radio noise, instead of the magnetic field intensity H (where $B = \mu_0 H$, and μ_0 is the permeability of free space), and thus our field units are given as femtoteslas (fT). The magnetic field amplitudes can of course be converted to electric field units by using the conventional conversion relation $E/H = 376.7$ ohms. However, as is well known, this last relation applies for plane electromagnetic waves propagating in free space, and it is not always valid for the noise fields observed on the earth's surface.

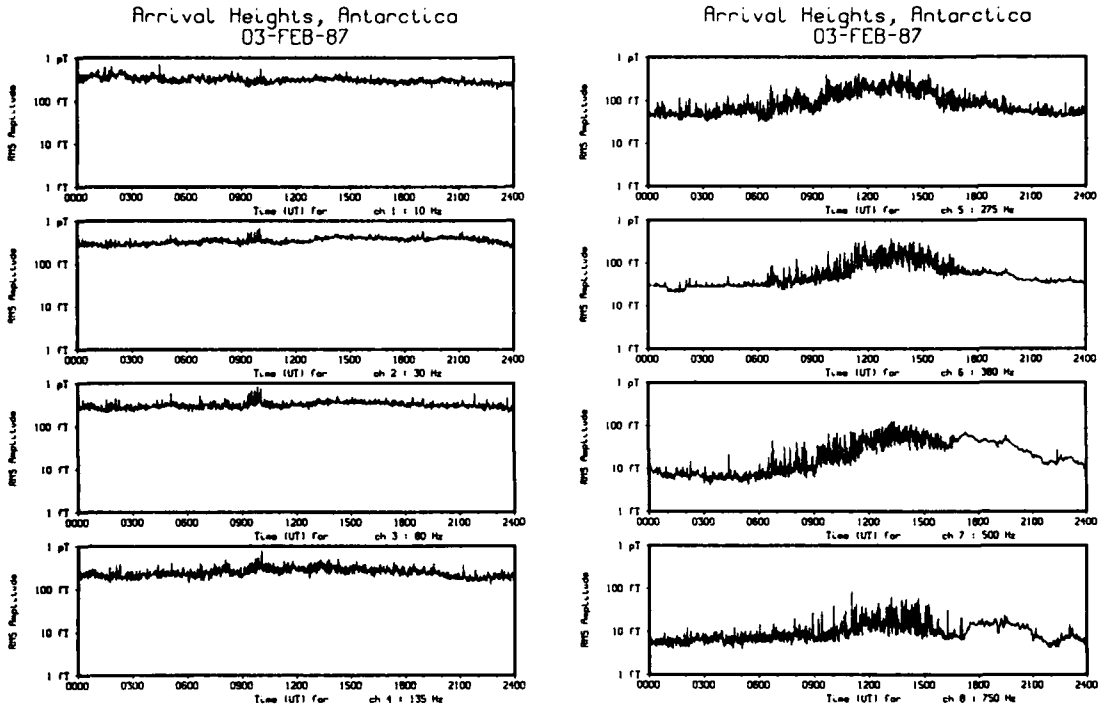


Figure 3. Variation of the one-minute rms average values over the UT day 3 February 1987 for each of the first eight narrow band channels of the Arrival Heights, Antarctica, radiometer. The channel frequencies are shown beneath each plot and cover the range 10–750 Hz. Arrival Heights was in continual daylight during the 24 hours covered by the display.

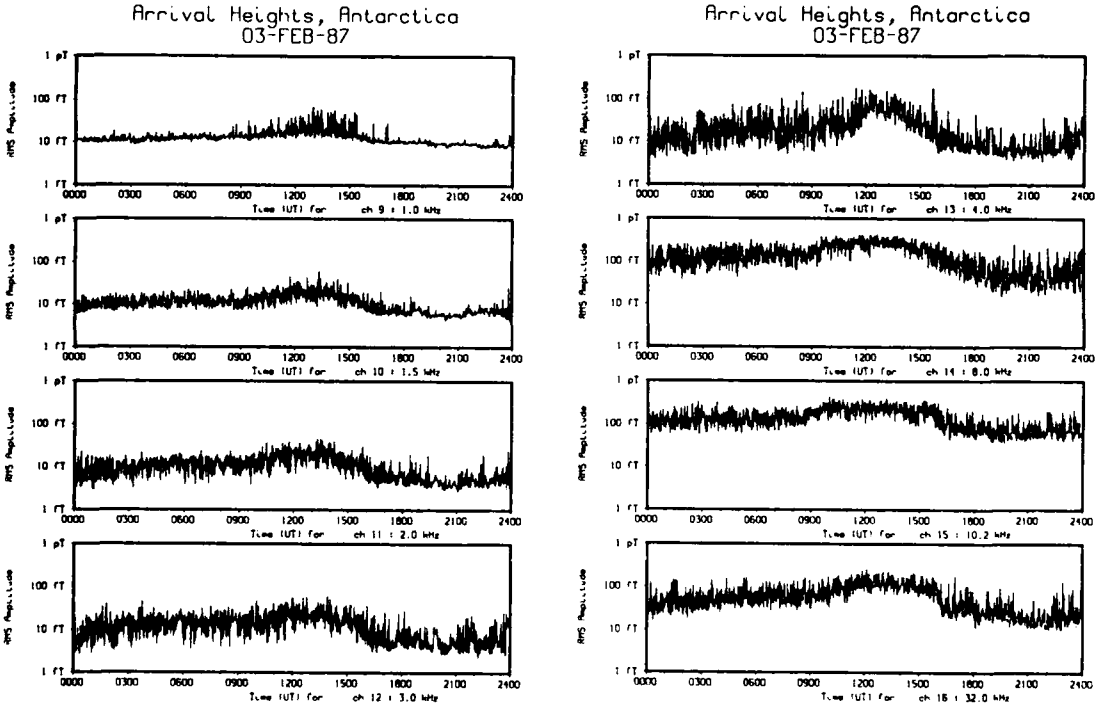


Figure 4. Variation of the one-minute rms average values over the UT day 3 February 1987 for each of the second eight narrow band channels of the Arrival Heights, Antarctica, radiometer. The channel frequencies are shown beneath each plot and cover the range 1–32 kHz. Arrival Heights was in continual daylight during the 24 hours covered by the display.

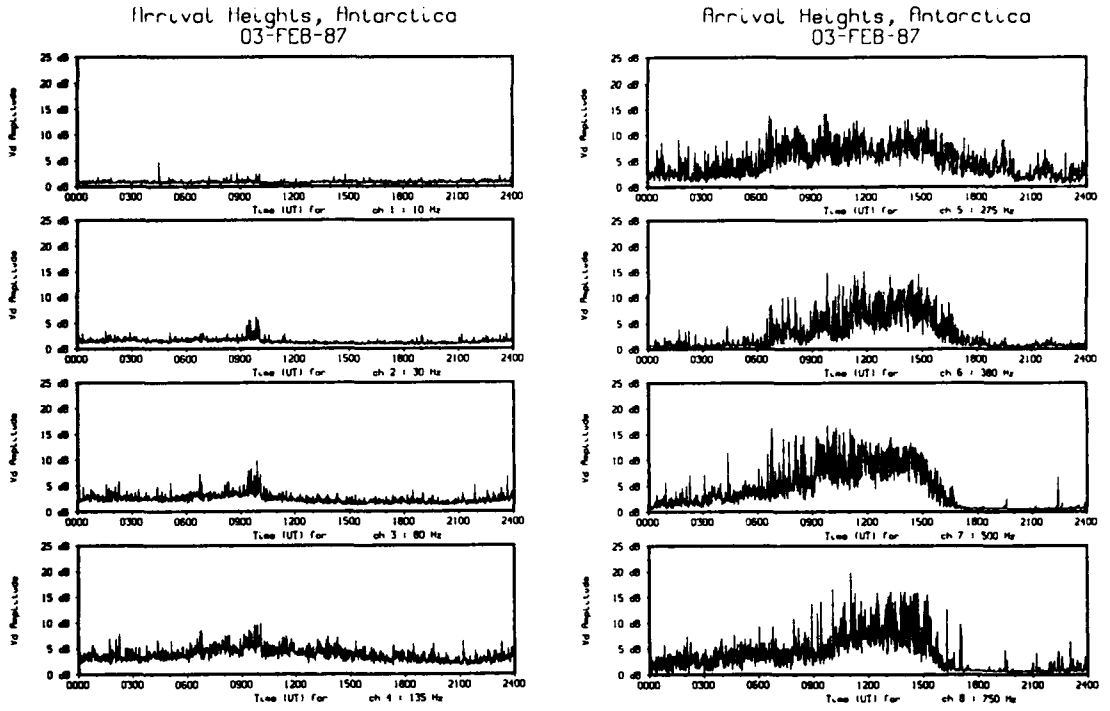


Figure 5. Variation of V_d at Arrival Heights, Antarctica, for the UT day 3 February 1987. The one-minute values of V_d are shown for the first eight narrow band radiometer frequencies and they are derived from, and may be compared with, the rms data shown in Figure 3. The channel frequencies are shown beneath each plot and cover the range 10–750 Hz.

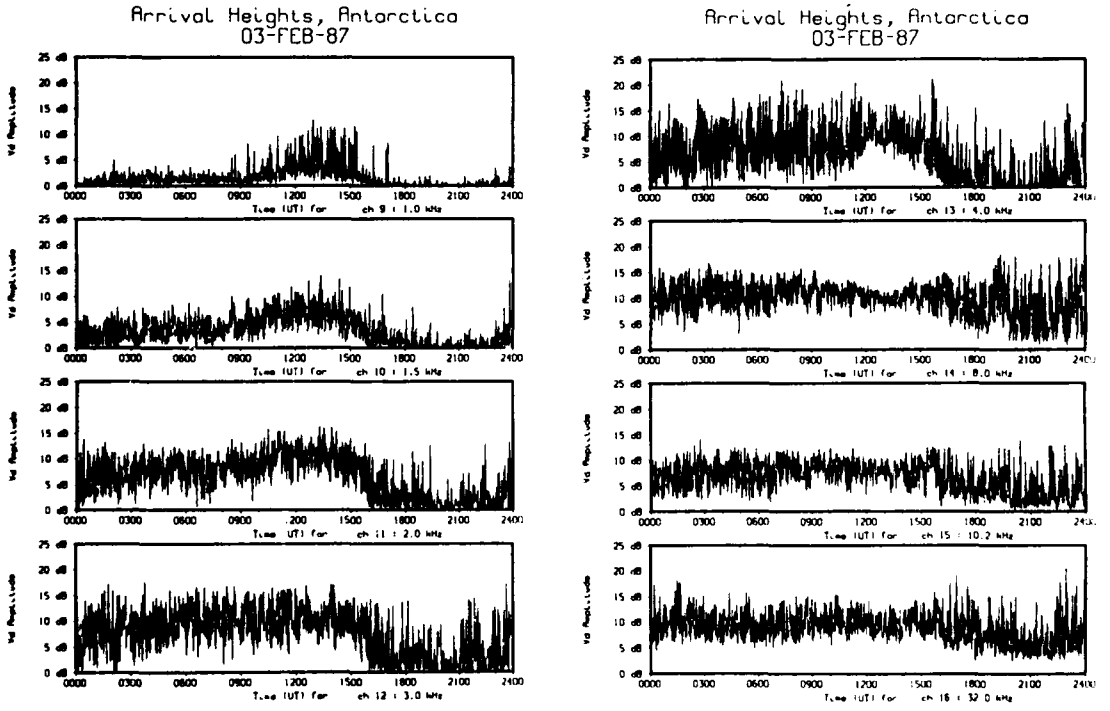


Figure 6. Variation of V_d at Arrival Heights, Antarctica, for the UT day 3 February 1987. The one-minute values of V_d are shown for the second eight narrow band radiometer frequencies and they are derived from, and may be compared with, the rms data shown in Figure 4. The channel frequencies are shown beneath each plot and cover the range 1–32 kHz.

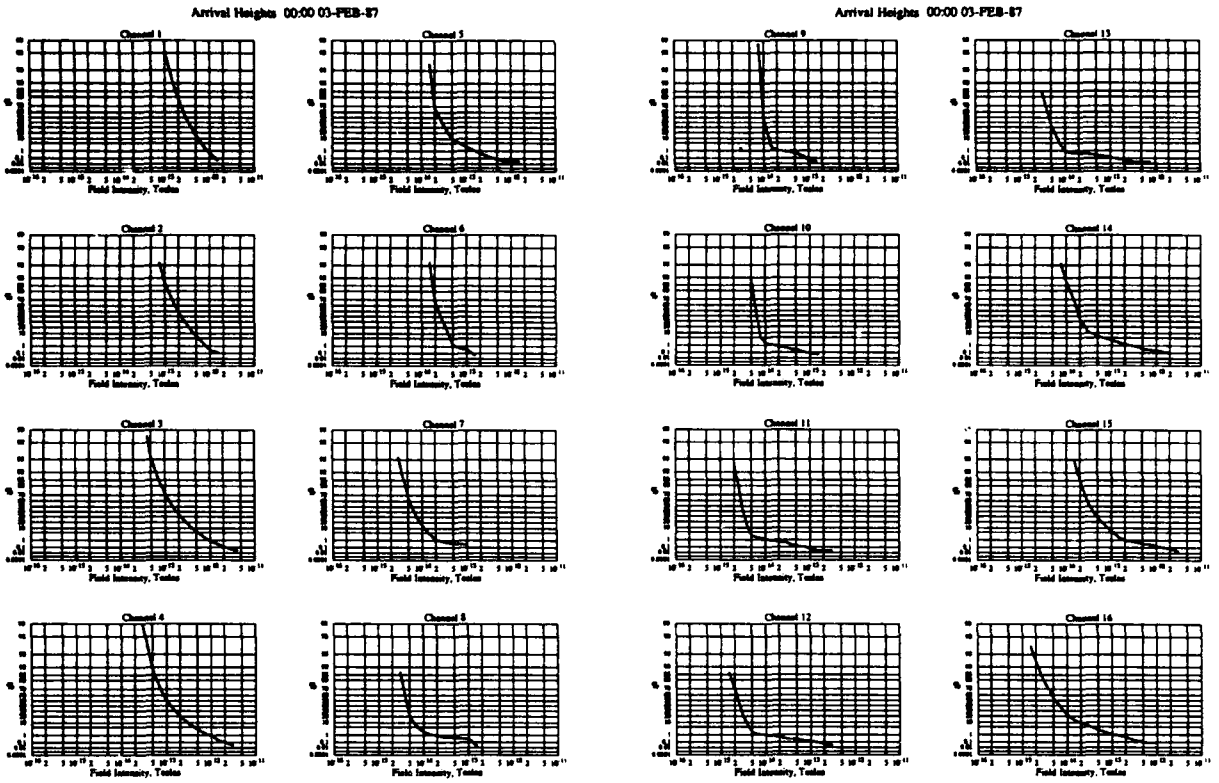


Figure 7. Illustrative amplitude probability distributions for Arrival Heights, Antarctica. The APD's are shown for the hour 0000–0100 UT on 3 February 1987 and cover each of the 16 narrow band channels; the rms amplitudes and V_d values for the same hour are shown in Figures 3–6. Some substantial differences between the various APD's can be observed.

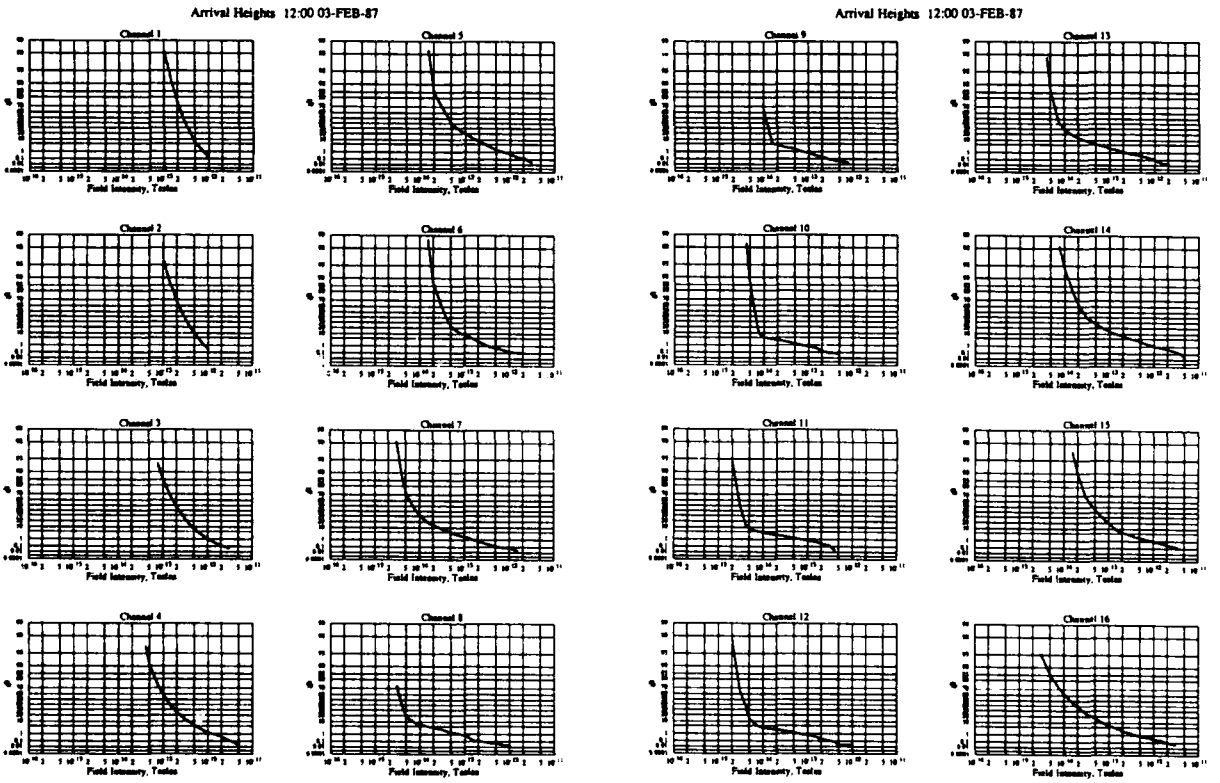


Figure 8. Further illustrative amplitude probability distributions for Arrival Heights, Antarctica. The APD's are shown for the hour 1200–1300 UT on 3 February 1987 and cover each of the 16 narrow band channels; the rms amplitudes and V_d values for the same hour are shown in Figures 3–6. Some substantial differences between the various APD's can be observed, but there do not appear to be large differences between the APD's shown here and the corresponding APD's in Figure 7.

Although there are notable differences between the various plots in Figures 3 and 4, there is nevertheless a measure of continuity as the frequency changes. Many of the plots show a peak close to 1200 UT, which is very close to local midnight at Arrival Heights. Also, the variability between the one-minute average values appears to increase around 1200 UT in some of the plots. This is particularly evident in the data for the 1.0 kHz channel. It is less noticeable in some of the other plots because of the logarithmic amplitude scales and the overall increases in amplitude, which in combination can lead to a visual impression of reduced variability. However, as we will see more clearly in Figures 5 and 6, there is also a true reduction in the variability around 1200 UT for some channels.

Figures 5 and 6 show the variations of V_d for the same day, times, and frequencies illustrated in Figures 3 and 4. Many of the features in these new figures could have been anticipated from the data shown in the preceding figures, but not all. The large variability of V_d in the data for the 4.0 kHz channel, for example, could not easily have been predicted, whereas the low levels of V_d in the 10–135 Hz channels confirm the impression of little variation given by the rms amplitude data.

Continuing the presentation of illustrative noise survey data, Figures 7 and 8 show some APD's for Arrival Heights on 3 February 1987. The APD's are shown for all of the 16 narrow band channels for the two hour-long intervals 0000–0100 UT and 1200–1300 UT; they are derived from the one-per-second samples that are stored on magnetic tape along with the one-minute average noise statistic data that were used for Figures 3–6. There are therefore a nominal 3600 measurements of the envelope amplitude included in the APD for each frequency. However, the actual number of measurements included is roughly 1–5% less due to the loss of data during the calibration intervals.

The percentages plotted for the APD's in Figures 7 and 8 is the percentage of time the envelope amplitude exceeds the amplitude given on the horizontal scale. Thus, except for the use of magnetic field instead of electric field amplitudes, the APD's are conventional. It can be seen that there are considerable differences between the APD's for the various narrow band frequencies, but there appear to be only small differences between corresponding APD plots for the two different hours.

To conclude this presentation of ELF/VLF noise statistic data, in Figure 9 we show an example of the rms amplitudes measured at Thule, Greenland, during the same UT day as the preceding Arrival Heights data. Only the rms amplitudes for the four highest frequency channels are shown, and the data for one channel are not fully illustrative: the 10.2 kHz channel includes signals from Omega navigation stations, which makes the noise measurements atypical for that channel (the reception of Omega signals at Thule is markedly different from that at Arrival Heights). Comparing the Thule and Arrival Heights data (Figures 4 and 9), we see that the Thule rms amplitude plots differ from those at Arrival Heights in almost every feature.

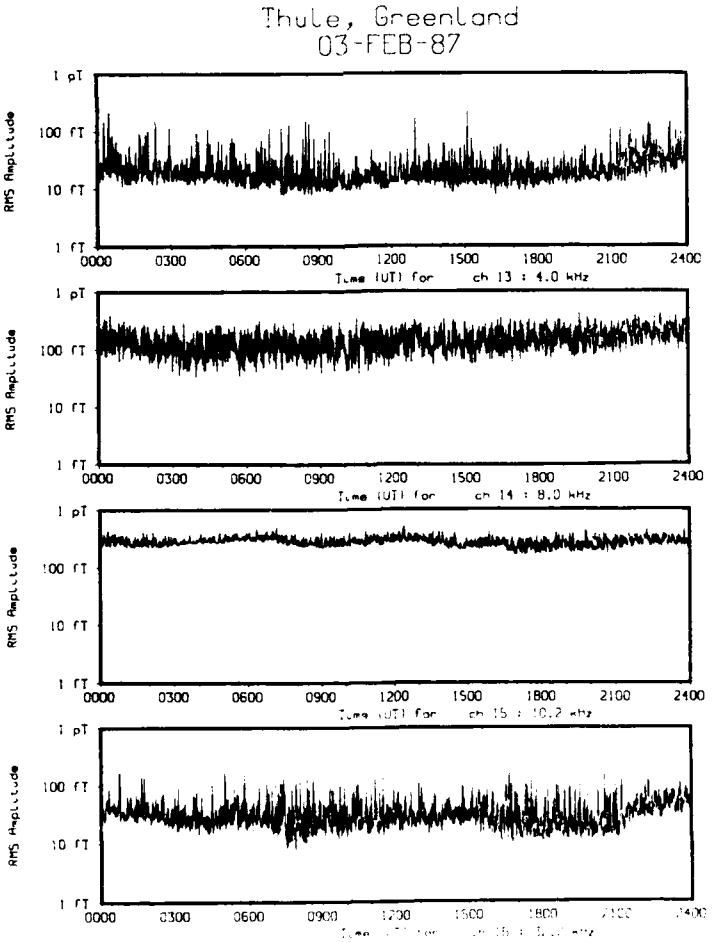


Figure 9. Some rms amplitudes for Thule, Greenland, on 3 February 1987. Only the amplitudes for the four highest frequency channels, 4–32 kHz, are shown.

In particular, the daily variations and variability for each channel are quite different. However, the average levels appear to be roughly the same. We note that Thule is in continual but variable darkness in early February, and Thule local time lags about 4 hours behind UT (Thule local midnight is around 0400 UT); differences in the rms amplitude plots for Thule and Arrival Heights would be expected for these two reasons alone.

Finally, in Figure 10 we show a digital spectrogram of the 0–400 Hz noise measured at Arrival Heights on 2 February 1987. The spectrogram was computed from the digital synoptic data that are recorded routinely at Arrival Heights as well as at all the other radiometer locations. We have not examined a large number of these spectrograms at the present time, but the data appear to be reasonably typical.

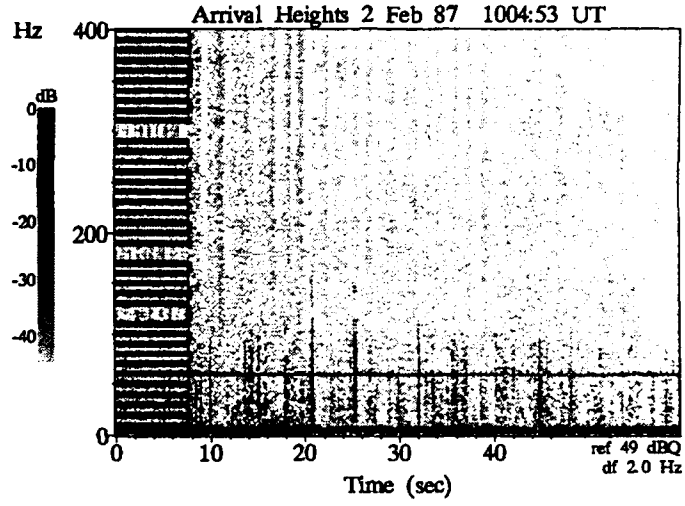


Figure 10. A digital spectrogram of the ELF activity (in the range 0–400 Hz) at Arrival Heights, Antarctica, during the one-minute synoptic interval 1004:53–1005:53 UT on 2 February 1987. The horizontal lines at the beginning of the spectrogram are calibration signals. Only a few sferics can be seen. There is a weak 60 Hz power line signal and a variable noise band centered on 5 Hz that may be due to vibration of the loop antennas caused by wind.

4. CONCLUSION

In this paper we have presented examples of ELF/VLF noise measurements and statistics for one day and mostly for just one location. Such data are of considerable scientific interest and can be used in many different kinds of cooperative studies. However, our goal, which is to characterize ELF/VLF noise on a global basis by simultaneous noise measurements over an entire year at our eight radiometer locations, is more ambitious. The total quantity of analog and digital data that must be acquired and processed to characterize the global distribution of ELF/VLF noise and its variations over an entire year will fill many hundreds of magnetic tapes and the data processing task will be formidable. However, modern computer systems have the capability of processing extraordinarily large quantities of data rapidly, and we expect soon to be able to move on from the processing of noise data on a daily basis to the preparation of summary data on a monthly basis for each location. These summary data will later be combined to give an appropriate picture of ELF/VLF noise over longer time scales. For much of the frequency range covered by our measurements, the picture will be new.

5. REFERENCES

C.C.I.R., "World distribution and characteristics of atmospheric radio noise," Report 322, International Radio Consultative Committee, International Telecommunication Union, Geneva, 1964.

C.C.I.R., "Worldwide minimum external noise levels, 0.1 Hz to 100 GHz," Report 670, pp. 224-229 in *Recommendations and Reports of the CCIR, 1982, 1*, International Radio Consultative Committee, International Telecommunication Union, Geneva, 1982*a*.

C.C.I.R., "Radio noise within and above the ionosphere," Report 342-4, pp. 184-196 in *Recommendations and Reports of the CCIR, 1982, 6*, International Radio Consultative Committee, International Telecommunication Union, Geneva, 1982*b*.

Crichlow, W.Q., "Noise investigation at VLF by the National Bureau of Standards," *Proc. IRE, 45*, 778-732, 1957.

Dinger, R.J., W.D. Meyers, and J.R. Davis, "Experimental investigation of ambient electromagnetic noise from 1.0 to 4.0 kHz in Italy and Norway," *Radio Sci., 17*, 285-302, 1982.

Field, E.C., Jr., and M. Lewinstein, "Amplitude-probability distribution model for VLF/ELF atmospheric noise," *IEEE Trans. Comm., COM-26*, 83-87, 1978.

Flock, W.L., and E.K. Smith, "Natural radio noise - a mini-review," *IEEE Trans. Antennas Prop., AP-32*, 762-767, 1984.

Fortnam, B.R., An automatic system for global monitoring of ELF and VLF radio noise phenomena, *Tech. Rept. E450-1*, 107 pp., STAR Laboratory, Stanford University, June 1985.

Fraser-Smith, A.C., and R.A. Helliwell, "The Stanford University ELF/VLF radiometer project: Measurement of the global distribution of ELF/VLF electromagnetic noise," *Proc. 1985 IEEE Internat. Symp. on Electromag. Compatibility, IEEE Cat. No. 85CH2116-2*, 305–311, 1985.

Galejs, J., "Amplitude distributions of radio noise at ELF and VLF," *J. Geophys. Res., 71*, 201–216, 1966.

Galejs, J., "Amplitude statistics of lightning discharge currents and ELF and VLF radio noise," *J. Geophys. Res., 72*, 2943–2953, 1967.

Maxwell, E.L., and D.L. Stone, "Natural noise fields from 1 cps to 100 kc," *IEEE Trans. Antennas Prop., AP-11*, 339–343, 1963.

Maxwell, E.L., "Atmospheric noise from 20 Hz to 30 kHz," pp. 557–593 in *Sub-Surface Communications*, AGARD Conf. Proceedings No. 20, 1966.

Smith, E.K., "The natural radio noise environment," *Proc. 1982 IEEE Internat. Symp. on Electromag. Compatibility*, 266–277, 1982.

Spaulding, A.D., and G.H. Hagn, "Worldwide minimum environmental radio noise levels (0.1 Hz to 100 GHz)," pp. 177-182 in *Proceedings: Effects of the Ionosphere on Space and Terrestrial Systems*, Ed. J.M. Goodwin, ONR/NRL, Arlington, Virginia, 1978.

Spaulding, A.D., "Atmospheric noise and its effects on telecommunication system performance," pp. 289–328 in *CRC Handbook of Atmospherics, I*, Ed. H. Volland, CRC Press, Boca Raton, Florida, 1982.

Watt, A.D., *VLF Radio Engineering*, 702 pp., Pergamon Press, New York, 1967.

Watt, A.D., and E.L. Maxwell, "Measured statistical characteristics of VLF atmospheric radio noise," *Proc. IRE, 45*, 55–62, 1957*a*.

Watt, A.D., and E.L. Maxwell, "Characteristics of atmospheric noise from 1 to 100 Kc," *Proc. IEEE, 45*, 787–794, 1957*b*.

6. ACKNOWLEDGEMENTS

The radiometers were designed, and their computers programmed, by Mr Evans W. Paschal. We thank Mr Paschal for this contribution, as well as for his continued advice throughout the course of the project. The contributions of our many collaborators, as listed above in the text, are an essential part of the noise survey; we greatly appreciate their collaboration. Support for this work is provided by the Office of Naval Research, through Contract No. N00014-81-K-0382 and Grant No. N00014-84-G-0202 (Department of Defense Equipment Grant), by Rome Air Development Center, partly through ONR Contract No. N00014-81-K-0382 and partly through Contract No. F19628-84-K-0043, and by the National Science Foundation, Division of Polar Programs, through Grant No. DPP-8316641.

MAN-MADE RADIO NOISE AND INTERFERENCE

G. H. Hagn
Information Sciences and Technology Center
SRI International
1611 North Kent Street
Arlington, VA 22209 (USA)

SUMMARY

This paper considers what the scientific community currently knows and does not know about man-made radio noise. Both empirical and analytical noise models are discussed. The measurement of man-made noise also is discussed. The measurement of man-made noise levels (and related statistics) during military exercises is suggested. A man-made radio noise program is described which would help fill the voids in current knowledge.

1. INTRODUCTION

1.1 Definitions and Scope

The problems of defining the terms noise and interference were considered by the author in 1974 when AGARD last met on this topic.[1] Although the matter of definitions is still somewhat confused, in this paper, noise is considered to be a source of potential interference and interference to be an effect--the degradation in system performance produced by the noise. This interpretation is consistent with that of the International Radio Consultative Committee (CCIR).[2] The composite electromagnetic noise environment observed at any given location can result from both natural and man-made sources (see Figure 1-1). The scope of this paper is limited to the man-made noise which tends to predominate over atmospheric noise during daytime in the MF and higher bands and in the VHF and higher bands at night (see Figure 1-2).

1.2 Background

A summary of the man-made noise from unintentional radiators was presented at the 1974 AGARD meeting in Paris by the author and R. A. Shepherd.[3] Several voids in our knowledge thirteen years ago were noted to be:

- Are man-made noise levels increasing, decreasing or staying about the same?
- What are the noise levels above about 250 MHz?
- How does one extrapolate noise data taken in different bandwidths?
- Is it possible to reach agreement on standardized measurement equipment and procedures?
- How best can we sample the more general classes of nonstationary random processes such as most classes of man-made noise?
- How do we relate noise environments to their effects on the performance of specific systems of interest?

In addition to the above, there are other important questions:

- Which noise models are best to use under which circumstances?
- What categories of noise environment (and related models) best describe a modern battlefield?
- How do we modify empirical man-made noise models (e.g., CCIR 258)[4] when we use different antennas?

We will try to address these questions, state what we know and do not know, and define a program to address the current voids.

1.3 Organization of This Paper

In Section 2, we will consider the available empirical noise models for man-made noise levels, and in Section 3 we will consider analytical models. Section 4 will discuss noise environment descriptions and the need for noise data (and models) for the environments of the modern battlefield as approximated by military exercises.

2. AVAILABLE EMPIRICAL MAN-MADE NOISE LEVEL MODELS

2.1 Introductory Comments

Man-made noise level predictions for an electrically-short, grounded, vertically-polarized monopole (whip) antenna have been given in CCIR Report 258 for several "typical" environments (see Figure 2-1).[4] These predictions generally apply to the 9-ft rod antenna on an aluminum groundplane on the ground (or mounted on a vehicle roof) for frequencies below about 20 MHz. The predictions for these antennas may apply at higher frequencies in the HF band. The noise parameter F_a for a given noise environment is a function of the type of antenna. This section summarizes the mathematical formulas for predictions of the median effective antenna noise factor, F_{am}, for electrically-short, grounded, vertically-polarized monopoles. THEY DO NOT APPLY DIRECTLY TO OTHER ANTENNAS.

2.2 Formulas for CCIR Report 258 Man-Made Noise Predictions

The environmental categories for which predictions are available in CCIR Report 258 include: business, residential, rural and quiet rural. These results were obtained for the first three man-made noise categories by Spaulding and Disney,[5] who also obtained some data on noise in parks and on university campuses. The data taken near interstate highways (see Figure 2-2) produced a level essentially identical to that for the residential category (constant = 73.0 dB). The definitions and formulas for these categories are:

$$\text{Business: } F_{am} = -27.7 \log_{10} f_{MHz} + 76.8, \text{ in } dB(kT_o); \quad (1)$$

where

F_{am} = median value of antenna noise figure, in $dB(kT_o)$:

$k = 1.38 \times 10^{-23}$ J/K = Boltzmann's constant;

T_o = 288 K = room temperature.

A business area was defined as any area where the predominant usage throughout the area is for any type of business (e.g., stores and offices, industrial parks, large shopping centers, main streets or highways lined with various business enterprises, etc.).

$$\text{Residential: } F_{am} = -27.7 \log_{10} f_{MHz} + 72.5, \text{ in } dB(kT_o). \quad (2)$$

Residential areas (urban or suburban) were defined as any areas used predominantly for single- or multiple-family dwellings with a density of at least two single-family units per acre and no large or busy highways.

$$\text{Parks and universities: } F_{am} = -27.7 \log_{10} f_{MHz} + 69.3, \text{ in dB } (kT_o). \quad (3)$$

Data were obtained in several parks and university campuses, and the values were between the values obtained for residential and rural areas.

$$\text{Rural: } F_{am} = -27.7 \log_{10} f_{MHz} + 67.2, \text{ in } dB(kT_o). \quad (4)$$

Rural areas were defined as locations where land usage is primarily for agricultural or similar pursuits, and dwellings are no more than one every five acres.

$$\text{Quiet rural: } F_{am} = -28.6 \log_{10} f_{MHz} + 53.6, \text{ in } dB(kT_o). \quad (5)$$

Quiet rural areas were defined as locations chosen to ensure a minimum amount of man-made noise. The data were obtained from selected ARN-2 receiving sites used to generate CCIR Report 322.

$$\text{Galactic: } F_{am} = -23.0 \log_{10} f_{MHz} + 52.0, \text{ in } dB(kT_o). \quad (6)$$

The formula for galactic noise from radio stars which is incident on the ionosphere is included here for comparison only. Galactic noise incident upon the ionosphere varies only a little at a given time of day, and the total diurnal variation due to the change in antenna orientation relative to the plane of the galaxy is less than $\pm$ 2 dB. On frequencies between about 10 and 30 MHz, the intervening ionosphere will reduce the level of galactic noise observed at the surface of the earth below the level predicted by the formula. During periods of very low sunspot number, galactic noise can be important at very quiet locations on frequencies down to 2 MHz.

The Final Acts of the HF Broadcasting World Administrative Radio Conference (WARC)[6] also specified a formula for noise to be used in analyses for HF broadcasting:

$$F_{am} = -28.1 \log_{10} f_{MHz} + 60.4, \text{ in } dB(kT_o). \quad (7)$$

HF broadcasting receiving areas vary widely, and this formula is intended to standardize computations rather than to describe a "typical" area.

2.3 Discussion of Data Upon Which the CCIR 258 Formulas Are Based

Data were obtained in 31 rural areas, 38 residential areas, and 23 business areas in the United States.[5] The area sampled varied from a few square blocks in the business areas to a few square miles in the rural areas. In the noise measurement method used, 10-s samples of the running average (time constant of about 50 s) of F_a were recorded. Thus, 360 samples of F_a were obtained each hour for a given measurement location and frequency. A least-squares fit was made to obtain the formulas given for the median value, F_{am}, versus frequency for each environmental category. The slope was found to be -27.7 dB/decade for business, residential and rural areas (at the 95% confidence level) over the band 250 kHz-250 MHz.

Few data of this type are known to the author which are available from countries other than the United States except for data taken by the U.S. Navy on various MF and HF noise and EMI surveys, and these data have not been analyzed extensively for comparison with the CCIR 258 noise model.

2.4 Noise Level Distribution Models

Three new statistical-distribution models of increasing complexity (simple Gaussian, composite Gaussian and chi square) have been developed which can be used to predict the exceedence probabilities for short-term (about 1 min) mean values of man-made radio noise "available power" levels for short, grounded, vertically-polarized antennas located near ground.[7] These models, which now have been included in CCIR

Report 258, are useful in predicting the probability that the short-term signal-to-noise ratio for a given communication system equals or exceeds a value required for successful communications. These models, one for each of the three environmental categories (business, residential, and rural), are based upon data used to develop the median noise level models in CCIR Report 258 for the frequency band 0.25-250 MHz. The models are most useful in the probability interval 0.1 to 0.9, but further checks against measured data are needed to determine the limits of their applicability.

Recall that a Gaussian distribution is completely specified by giving its mean value and its standard deviation. The simple and composite Gaussian models use the values of F_{am} given in para. 2.2 as estimates of the mean short-term antenna noise factor from which the mean noise power can be computed using:

$$P_n = F_a + 10\log_{10} b - 204, \text{ dB(1W)} \quad (8)$$

where

b = the noise power bandwidth, in Hz;

-204 = the noise power in a 1-Hz bandwidth at 288 K, in dB(1W).

Approximations for the standard deviations for the simple Gaussian model (σ_N) are summarized in Table 2-1. Figure 2-3 gives examples of the model predictions for 20 MHz in a residential area.

Table 2-1

MAN-MADE NOISE MODEL FOR σ_N
(after Hagn, et al., 1982)

For $1 \leq f_{MHz} < 100$:	
Business:	$\sigma_N = 10.5$ dB
Residential:	$\sigma_N = 8.5$ dB
Rural:	$\sigma_N = 6.5$ dB
For $1 \leq f_{MHz} \leq 20$:	
Quiet Rural:	$\sigma_N = 4.5$ dB
For $f_{MHz} > 20$:	
Quiet Rural:	$\sigma_N = 1.0$ dB (assumes galactic noise)
For $f_{MHz} > 100$, $F_{am} > 0$ dB, $\sigma_N > 2$ dB:	
Business:	$\sigma_N = 10.5 - 9 \log_{10} \frac{f_{MHz}}{100}$, in dB; or 2 dB
Residential:	$\sigma_N = 8.5 - 9 \log_{10} \frac{f_{MHz}}{100}$, in dB; or 2 dB
Rural:	$\sigma_N = 6.5 - 9 \log_{10} \frac{f_{MHz}}{100}$, in dB; or 2 dB
For $F_{am} = 0$ dB, all $\sigma_N = 0$ dB.	

Hagn, et al.[8] have suggested using the CCIR 258 values for frequencies below 200 MHz; but for frequencies above 200 MHz, better equations for F_{am} in dB(kT_0) for the following environmental categories are:

Business: $F_{am} = 49.4 - 15.8 \log_{10} f_{MHz}$; (9)

Residential: $F_{am} = 45.2 - 15.8 \log_{10} f_{MHz}$; (10)

Rural: $F_{am} = 39.2 - 15.8 \log_{10} f_{MHz}$. (11)

These equations have now been verified by Lauber and Bertrand[9] as being more accurate above 200 MHz than either CCIR 258 or the equations of Skomal.[10] For business areas, example comparisons are given in Figure 2-4.

3. ANALYTICAL MAN-MADE NOISE MODELS

3.1 Background

Recent (since 1972) development by Middleton of analytically tractable, canonical, and experimentally verified models[11-12] of electromagnetic noise and interference environments have significant implications, both for the measurement of EM environments and for the prediction and evaluation of the performance of communication systems.[13,14] Indeed, the availability of these new, non-Gaussian models for the first time makes possible a general, quantified approach to almost all classes of real-world noise and in particular their impact on telecommunication systems.

There are three canonical Middleton models, respectively for Class A, B, and C noise: Class A noise is essentially coherent in the receiver, in that it produces negligible transients therein, while Class B is "incoherent", generating only transient impulses. The former is non-impulsive, while the latter is impulsive. If Δf_R (see Figure 3-1) is the receiver bandwidth (assumed narrow-band here throughout), and T_N is the duration of the coherent portion of the noise, then Class A is distinguished by $T_N \Delta f_R >> 1$, while for Class B we have $T_N \Delta f_R << 1$. Class C is a combination of Class A and Class B. Middleton has shown recently[15] that, at least through the first-order statistics which are currently the subject of our analysis and experiment, the Class C cases are reducible to the basic Class B form.

A critical feature of these Class A and B models is that they are canonical: their analytic form remains invariant of the particular physical source mechanisms. Such diverse mechanisms as fluorescent lights, powerline noise, automobile ignition noise, ore-crushing machinery, etc., are readily accommodated without restriction. In fact, the only limitations on the statistical-physical foundations of our models are that they: (1) represent independent (or independent sets of) radiation "events"; and (2) that the output of the (linear) front-end stages of the receiver be narrowband. The result is that the basic statistics are fundamentally Poissonian,[15] and that essentially all noise may be canonically represented by these models, invariant of waveform and the details of propagation as far as formal analytical structure is concerned. An exception is the class of problem where the interference is completely deterministic and known at the receiver. These cases, of course, must be handled by the classical method of direct analysis and/or empirical study.

A second critical feature of these models is that they are analytically tractable, as well as computationally manageable. This is especially important when we go beyond the measurement program to apply the analysis to the prediction and evaluation of system performance, including optimum as well as suboptimum reception.[16,17] The model statistics of principal interest and use here as well as in the analysis of measurements and the parametric description of these models, are: (1) the various (first-order) moments; (2) the probability density functions (pdf's), and (3) the exceedence probabilities (or APD's). In any case, we emphasize that the model parameters are not ad hoc, but are physically derived from the model, involving source distributions, radiation properties and geometries, and front-end receiver structures. As has been noted before, the analytical model is in excellent agreement with the empirical data, as well as being canonical. These models are mathematically intricate and they are not reproduced here. Spaulding[18] provides a brief discussion of the probability density function (pdf) for the received instantaneous amplitude.

The canonical Class A noise model (in the first-order) is governed principally by these global parameters:

- A = the "overlap" or "unstructure" index, which is defined as the average number of radiation "events" per second times the mean duration of a typical source emission. The smaller the A, the more "structured" (in time) is the noise. Conversely, the larger the A, the more Gaussian and less structured is the noise. As A approaches infinity, the noise becomes Gaussian.
- Γ' = the Gaussian factor = ratio of average intensity of the Gaussian component of the noise to that of the non-Gaussian component.
- Ω = the mean intensity of the non-Gaussian (or "impulsive") noise component.

The Class B model involves several other parameters (e.g., α, related to the physical process causing the noise). These parameters are all measurable either at the input to the receiver or at the output of the initial linear front-end stages, before non-linear processing (see Figure 3-1). Currently, there is a limitation on some uses of the Middleton models because estimated values for the required parameters are not available. The parameters can be estimated through trial and error fits to measured APDs plotted relative to the RMS value of the envelope voltage, and this method has been used to produce the examples in the literature. Middleton has suggested using measured moments of the pdf, but practical parameter estimation techniques deserve considerable additional attention. Assuming that the model parameters can be estimated adequately from some data base, then the relative APD estimated from the Middleton model nicely supplements the empirical models discussed in Section 2.4 which can be used to estimate the absolute RMS envelope reference voltage required to convert the relative APD to an absolute APD. The resulting APD can be used to estimate system performance, as discussed by Spaulding.[18]

4. NOISE ENVIRONMENT DESCRIPTIONS AND THE MODERN BATTLEFIELD

4.1 Introductory Comments

Engineers and operations analysts who want to use the available man-made radio noise models in CCIR Report 258 for predicting the performance of tactical radios are faced with the problem of determining whether any given tactical environment is most like a U.S. business area, a U.S. residential area, or a U.S. rural area over a decade ago. The only CCIR model other alternative is to assume the environment is

similar to a quiet rural area near one of the quiet stations where the atmospheric noise data were gathered during the International Geophysical Year (IGY) to develop CCIR 322 over two decades ago. Those who want to use other noise models (e.g., Skomal)[10] are faced with similar problems. Added insight on noise levels in large tactical field exercises is needed to better use the existing models for military analyses and eventually to improve the models for military use -- perhaps by defining new environmental categories. It is necessary to identify candidate future tactical field exercises, in and outside CONUS over the next five years, at which man-made noise data could be obtained, and to develop a general noise measurement, analysis and data archiving plan for such exercises. The plan would describe the goals of such measurements, the types of equipment needed (including a list of the current owners of such equipment and its potential availability), and a generic equipment-deployment plan. The following section considers environmental descriptions pertinent to noise measurements and models. A program to accomplish this is outlined in Section 6.

4.2 Environmental Descriptions Pertinent to Noise Measurements and Models

It is important to be able to describe the environmental situation for which measured data were obtained or for which model predictions were made. Several types of environmental descriptions are pertinent to noise measurements and models (see Table 4-1). These include land use categories (e.g., business or rural), human activity categories (e.g., vehicular traffic densities), specific sources in proximity to a receiving antenna (e.g., powerlines), the weather (e.g., nearby thunderstorms), and scenarios (e.g., land battle scenarios or ship formations at sea).

Table 4-1

ENVIRONMENTAL DESCRIPTIONS

LAND USE CATEGORIES
- BUSINESS (OR URBAN) - RESIDENTIAL (OR SUBURBAN) - INDUSTRIAL PARKS AND CAMPUSES - RURAL - QUIET RURAL
ACTIVITY CATEGORIES
- VEHICULAR TRAFFIC DENSITY - POWER CONSUMPTION DENSITY - POPULATION DENSITY
SPECIFIC SOURCES
- MOTOR VEHICLES - POWERLINES - ELECTRICAL MACHINERY
WEATHER
- THUNDERSTORMS - PRECIPITATION
SCENARIOS
- MILITARY - CIVILIAN

Each of these types of environmental descriptions has its uses. For example, CCIR 258 provides median noise levels versus frequency for different land usage categories (see Section 2). Regression equations have been developed for certain activity categories such as vehicular traffic density.[5,18] Attempts to correlate noise level with power consumption have been successful for frequencies below 20 MHz, but reliable data on power consumption are scarce.[5] Allen[19] found a correlation between noise level and population density, but Spaulding and Disney found no correlation for groupings the size of the U.S. Census Bureau's standard location area (SLA).[18] In 1966, Lucas and Haydon attributed to Spaulding (apparently erroneously) a relationship between typical man-made noise levels and the relative population of the receiving area.[21] Finally, models have been developed for specific types of sources (e.g., powerlines) as a function of frequency, distance and polarization by the Electromagnetic Compatibility Analysis Center, Annapolis, MD. Nevertheless, some basic questions remain, such as which descriptions are most useful under any given set of circumstances (see Table 4-2). Meteorological conditions can cause noise directly (e.g., lightning) or indirectly (e.g., precipitation causing corona on powerlines with voltages > 100 kV).

4.3 Identifying and Selecting Exercises

To identify candidate tactical field exercises, within and outside of CONUS during the next five years, during which man-made noise data could be obtained, there must first be a method not only of

learning of the existance of plans for such tactical field exercises, but also of learning some details as to what types of information are included in those plans. Then, a set of selection criteria must be developed, based on the plans and on the requirements for noise information. When that task is completed, it will be known what information to search for in the planning documents of any particular tactical field exercise to determine whether that exercise might be suited for the collection of man-made radio noise data.

Table 4-2

SOME BASIC QUESTIONS

1) WHAT NOISE MODEL SHOULD ONE USE FOR SIMULATING COMMUNICATION SYSTEM PERFORMANCE ON A TACTICAL SCENARIO?

2) HOW DOES NOISE IN WARTIME RELATE TO NOISE IN A TACTICAL FIELD EXERCISE?

3) HOW DOES MAN-MADE NOISE IN A TACTICAL FIELD EXERCISE RELATE TO CCIR 258 LAND-USE CATEGORIES? EXERCISE VARIABLES?

4) ARE OCONUS NOISE LEVELS DIFFERENT FROM CONUS LEVELS USED IN CCIR 258?

5) ARE THE CCIR 258 LEVELS STILL VALID IN CONUS?

Actual access to the exercise plans will involve having the proper security clearance and the "need-to-know." It has been determined that there is one central source in the United States for the initial general information that will be needed; however, since these tactical field exercises might be either within CONUS or outside, such as in Europe or Korea, it will be necessary to maintain contact with several sources for the detailed information. A military sponsor or contract will be required to obtain approval for release of the required information. Any request for information on or for participation in any joint military exercises within CONUS or outside of CONUS should be made in writing and it should include the following information:

1) Sponsoring agency, point of contact, and contract number;

2) Objective of the exercise participation;

3) The general size and composition of the field measurement team;

4) The duration of the field measurements;

5) The type of equipment to be used; and,

6) A statement regarding the minimum level of security clearance to be possessed by the field measurement team.

The specific dates and locations of most of the exercises are classified prior to the exercise taking place, and in many cases this information remains classified after the exercise is completed. An example of a CONUS exercise series is the readiness exercise (REX). Examples of the types of exercises in Europe that should be considered include:

- Winter Exercise (WINTEX) -- generally in late winter or early spring;
- Return of Forces to Germany (REFORGER) -- generally in the late summer or autumn.

The test team should be aware that their presence and activities can be potentially disruptive to the primary objectives of the exercise. For example, if proper precautions are not taken, they can distract or bother exercise participants, disrupt physical security, compromise camouflage plans with odd antennas, etc. Careful coordination and planning before the exercise, and careful coordination and conduct during the exercise should permit the test team to overcome these types of potential problems.

Among the factors that should be considered as criteria in the selection of tactical field exercises suitable for the measurement of man-made radio noise are the following:

- The duration of the tactical field exercise. Will it last long enough to collect the amount of noise data that will be required? (Note that the amount is not yet determined.)

- The size of the land area or the physical extent of the tactical field exercises. What sorts of physical environments will be included? That is, will the environment include remote forests or deserts, suburban and/or urban areas, highways, and so on?

- To what extent will a tactical field exercise disrupt any normal civilian activities in the region--particularly those that might create man-made radio noise?

- The radio systems to be used, which will determine the frequency bands of greatest interest, the noise measurement bandwidths that are most germane, and the types of antennas that are the most important.

- The number (and activities) of the people involved in the tactical field exercise and the types of equipment (particularly vehicles) to be used.

- What movements of personnel and equipment are to be expected that might affect siting of measurement equipment and the type and amount of measurement equipment required to cover a given situation.

- The season of the year. There may or may not be differences in the noise environment caused by the seasonal operation of possible noise generating appliances or other systems. However, depending on the data collection process, that process may possibly be hindered by extremes in heat, cold, or humidity.

- The noise data sampling plan should include considerations of:

  - Space sampling grid
  - Time (per sample and duration of sampling process)
  - Frequencies and bandwidths.

- The data reduction and analysis planning should be performed prior to the exercise.

The noise parameters to be measured include the following, as discussed by The International Union of Radio Science,[22] Shaver, et al.,[23] Spaulding,[13,24] Hagn,[25] and Matheson[26]:

- V_{rms}, the root-mean-square noise envelope voltage;

- V_{ave}, the average envelope voltage;

- APD, the amplitude probability distribution;

- ACR, the average crossing rate for a specified set of thresholds;

- PSD, the pulse spacing distribution; and,

- PDD, the pulse duration distribution.

The last four parameters are especially useful at the higher frequencies, where F_{am} has dropped below the receiver noise figure, F_r. This typically occurs in the lower part or middle of the VHF band. The frequency range of most current interest is from HF up through SHF. Available instrumentation includes the DM-4[27] and several low-cost systems of the type developed by Christianson,[28] who has also reviewed the history of noise measurement equipment.[29] These instruments can measure all of the desired parameters; however, they require an appropriate receiver (with log-video output).

4.4 Changes in the Noise Level or Other Characteristics

It is important to ascertain whether there is a change in the man-made noise levels at appropriate locations during a tactical field exercise, whether the noise level increases or decreases, and if so, by what amount. Also, it would be useful to know what causes the change so that the results can be extrapolated to other situations.

The appropriate locations in question are, of course, the locations at which radio receivers will be operated during the tactical field exercise (e.g., command post receiving sites). It will be necessary to learn which types of radio systems are used in which types of locations. Army communication handbooks and manuals may provide instructions for siting radio systems during an exercise or an actual war; alternatively, this information may be provided by experienced military personnel who are able to describe these types of locations. On the other hand, it may be necessary to study tactical field exercises and interview participants to learn just which categories of locations are the most important. If it is not possible to obtain that information from past exercises, it might be necessary to observe a tactical field exercise to learn the locations of importance. Indeed, participation as an observer (perhaps making very limited measurements) would provide valuable insights to facilitate planning for more extensive noise measurements during a subsequent exercise. The number of locations that would be needed to characterize the change in the noise level over an area (vs at a point) is also a matter to be considered.

Once it has been learned which radio systems are used at what types of locations, the frequencies (and related bandwidths) that are most important at those locations can be determined. Then, if time or equipment constraints limit the amount of data that can be collected at a given place, the researcher will know where to concentrate his efforts.

To determine the change in the noise level at some location, it will be necessary to conduct measurements there during a tactical field exercise and also either before or afterwards to compare the results and determine the change. Probably the best sequence would be to make the first measurements during the tactical field exercise, since then the measurement locations would be certain to be determined by the essential military activity of the exercise. Also, exercise security would not be compromised by premature disclosure of planned locations.

4.5 Sampling Plan

Noise levels fluctuate with time, and it is most common to sample the noise at intervals and to record these samples. Then, the typical result of measurements at some location on a given frequency is a sample distribution of the noise. That distribution has some sample median level which is the number most commonly used to characterize the average noise power level. The sample distribution, developed from a number of measurements of the noise, can only be an approximation to the true distribution that would be obtained if the noise was examined continually; the approximation improves, however, as the number of samples of the

noise increases -- provided that the process is stationary (or at least stationary enough). It is possible to use nonparametric statistical techniques (e.g., Kolmogoroff-Smirnov)[30] to formulate and test hypotheses regarding the true distribution. In particular, confidence bounds surrounding a measured distribution can be determined such that a certain degree of confidence can be reached that the true distribution is within some specified range of the measured distribution. Then, when comparing the noise measured during the tactical field exercise with the noise measured later, it will be possible to state with known confidence that any differences seen in the median or distributional form are the result of actual differences in the underlying distributions and not just the result of chance.

Some guidance on field measurements of noise is available in Refs. 31 and 32. Before going to the field to make measurements, it is important to have a clear understanding of the amount of data that will be needed to compare the measurements against the models. At a given level of confidence, the bounds narrow as the number of samples increases. Thus, confidence that the measurements do (or do not) represent one of the models increases as more samples of the noise are taken. The Kolmogoroff-Smirnov statistical techniques can be used to determine just what degree of confidence can be obtained as a result of how many measurements. The discussion by Spaulding and Hagn[33] on spectrum occupancy definition and estimation contains some germane discussion regarding nonparametric statistics and sample size for occupancy data. Sample sizes of at least 100 or more are desirable for this type of analysis. Also, the distributions can change during a field exercise as activity in the vicinity of the noise measurement system changes. Therefore, tests of the stationarity of the distribution are required while processing the data. Measures like the Allen variance analysis can be performed during data acquisition to determine if the process being sampled is "stationary enough."[34] Finally, it is important to determine if the samples being acquired are independent before attempting to determine confidence bounds for a given sample size and level of significance.

4.6 The Tactical Environment and Existing Noise Models

For a given frequency, each of the CCIR 258 models yields a median noise level and a standard deviation so that the distribution of the expected average noise power level is known. The assumption is that the average noise power level in the measurement bandwidth is log-normally distributed. That is, the logarithm is normally distributed so that when the average noise power is expressed in decibels its distribution is normal. This is a reasonable assumption, but more accurate models have been developed. Hagn and Sailors[7] have suggested other models including a "composite Gaussian" model that is normal for a given half of the distribution. Thus, in comparing measurements made in the field against these models, a measured sample distribution can be compared against the models' distributions, and the following question can be asked: With what confidence can we state that the measured noise distribution is the same as that of one of the models? The answer can be found by using nonparametric statistical methods as mentioned above. Confidence contours will need to be developed that bound the measured distribution at various levels of confidence. Then, if those contours contain one of the model distributions, that model can be said, with a certain confidence, to represent the measured data.

4.7 Next Steps

The next steps in further documenting the noise during tactical field exercises should include:

- Acceptance of the idea that such data are needed, and funding to acquire the data;
- More detailed discussions to identify specific candidate exercises;
- Participation as an observer at a candidate exercise (perhaps attempting very limited measurements at a command post receiving location or some similar location);
- The formulation of a more detailed measurement plan, and execution of the plan; and,
- Analysis of the data and the development (as required) of noise models for the military environment.

5. VOIDS IN CURRENT KNOWLEDGE

Let us return to the questions posed at the beginning of this paper and assess what we know and do not know.

- Are man-made noise levels increasing, decreasing or staying about the same? To the author's knowledge, no data have been acquired over time at the same locations measuring the same parameters to be able to answer this question.
- What are the noise levels above about 250 MHz? The data of Lauber and Bertrand[9] in Canada have contributed to answering this question (see Figure 2-4).
- How does one extrapolate noise data taken in different bandwidths? Data by Herman and DeAngelis[35] have provided insights on the bandwidth extrapolation of V_d, and indeed, these data showed that the previous CCIR assumption was quite incorrect.
- Is it possible to reach agreement on standardized measurement equipment and procedures? It is certainly slow going. A two-day U.S. meeting on this topic held in 1984 did not resolve this issue.[36] The paper by Christianson in this meeting is germane.[29]
- How best can we sample the more general classes of nonstationary random processes such as most classes of man-made noise? The work of Middleton gives some insights on this, and the work of Spaulding and Hagn[33] provides some guidance for the measurement of

spectrum occupancy. Some comments in this paper on measurements during military exercises are germane.

- How do we relate noise environments to their effects on system performance. The work of Spaulding[13,14] provides useful input to this question.
- Which noise models are best under which circumstances? The new measurements proposed in this paper should help provide guidance on this question.
- What categories of noise environment (and related models) best describe a modern battlefield? Measurements during field exercises are the best way to obtain insights to answer this question.
- How do we modify empirical noise models when we use different antennas? Hagn and Shepherd[37] have provided some new insights into this problem, but it is not solved.

6. RECOMMENDED MAN-MADE NOISE PROGRAM

6.1 Steps in a Noise Program

A comprehensive man-made radio noise program would consist of the following steps:

1) Standardize the instrumentation and procedures for acquiring and archiving noise data, and prepare a noise measurement handbook.

2) Acquire man-made noise data over a group of frequency bands with representative antennas in a variety of environments of interest (including military communication station receiving sites, military exercise locations such as command posts, etc., as well as the standard civilian business, residential, and rural locations) in several countries.

3) Archive the noise data in a location (or locations) where it can be used for model development and checking.

4) Evaluate the utility of the CCIR 258 noise level models for use in countries beyond the United States for business, residential, rural and quiet rural locations, and determine if new categories are required.

5) Evaluate the extrapolation of the noise models to different types of antennas from data taken using different antennas in the same noise environment.

6) Analyze APDs from the same CCIR environmental category (e.g., rural) to determine if typical Middleton model parameters for Class B noise can be found as a function of category.

7) Develop a man-made noise simulator for use in laboratory studies of the effects of noise on system performance and on design.

8) Integrate the resulting man-made noise model into a more general composite electromagnetic noise environment model which would properly combine the contributions from both man-made and natural sources.

9) Prepare a radio noise model user's handbook.

Let us now consider each of these steps in more detail.

6.2 Standardize Instrumentation and Procedures

Proper measurements of radio noise are difficult to perform. The CCIR established Study Program 29B/6 and has produced Report 254-4, but this report does not provide guidance on measurement techniques. The U.S. National Telecommunications and Information Administration (NTIA) Manual of Regulations and Procedures for Federal Radio Frequency Management[38] recommends that the mean (average) noise power be used as the basic quantity for the measurement of radio noise, and that other measured quantities be relatable to this basic quantity. A radio noise measurement handbook could be developed to assist organizations planning noise measurements to obtain data of broader utility. According to U.S. DOD Directive 3222.3, this falls into the Army area of responsibility in the United States.

Hagn[31] has developed a draft handbook for MF and HF surveys of radio noise at the earth's surface using battery-operated equipment and a 9-ft rod antenna. The IEEE (1985) has published recommended practices for an electromagnetic site survey for the band 10 kHz to 10 GHz.[32] The 1979 handbook was intended primarily for organizations using the battery-operated Singer NM-26T analog receiver. This receiver is no longer manufactured. There is a need to update this survey handbook for use with currently available instrumentation. Such instrumentation incorporates digital techniques which make data acquisition, processing and analysis less time consuming and expensive. The IEEE document also is in need of revision. Vincent[39] has shown the variability of the environment to be measured, and this complicated the measurement requirements. A primary problem is achieving agreement on what parameters are to be measured and what type of instrumentation is to be used. Ideally, an agreement could be reached within NATO, and a NATO handbook could be written to guide future measurements.

6.3 Acquire Noise Data

An initiative is needed to measure noise in simulated battlefield environments (e.g., military exercises) in order to: 1) give guidance to current model users on existing options for man-made noise (i.e., business, residential, rural or quiet rural); 2) provide information for defining new scenario-related environmental categories that are more germane to military operations; and, 3) lead to more appropriate models for use in future studies. Also, data are needed to verify the CCIR 258 models in the United States in the current timeframe and in other countries. This data acquisition should use the standardized equipment and procedures.

6.4 Archive Noise Data

Several books have been published (e.g., Skomal[10] and Herman[40]) which summarized noise data available in the late 1970s. In the United States, the Defense Technical Information Center (DTIC) and the National Technical Information Service (NTIS) have numerous reports indexed by key words including man-made radio noise, but there currently is no one place where data, reports (including classified reports), papers and books on radio noise are available to U.S., or NATO, government researchers or their contractors. It has been suggested that an organization like the Institute for Telecommunication Sciences (ITS), Boulder, CO, be funded to establish and administer such a library using their current holdings as a point of departure. The intent would not be to duplicate the efforts of the DTIC, the NTIS, or other repositories. The ITS would publish a list of holdings and, at intervals, they would also publish an update. Researchers would be encouraged to add the ITS library to future noise report distribution lists.

The ITS already has a sizeable radio noise data base which supplements the holdings of the World Data Center in Boulder. This information has resulted in major savings in the current efforts to update the atmospheric noise model for worldwide predictions. Sailors and Brown[41] were able to use the ITS data base to develop a microcomputer version of CCIR 322 atmospheric noise maps which could only have been done at greatly increased cost had ITS not retained both the data and the germane computer codes. The ITS could serve as a repository of future selected noise data and computer codes of general interest. Also, perhaps organization(s) in Europe could be identified to serve such a function with some common holdings of data.

6.5 Evaluate Existing Noise Level Models

The CCIR models for business, residential, rural and quiet rural areas for the surface noise environment should be evaluated and improved. Skomal's model for noise vs distance from an urban center also should be evaluated, as well as his model for noise at aircraft altitudes above cities. The data base described in paragraph 6.4 could be used for this purpose.

6.6 Check Effect of Antenna Type

The noise data taken in the same environment at the same time with the same receiving system and different antennas should be used to evaluate the importance of using antennas different than those used to develop the noise model (short vertical monopole). Therefore, the short vertical monopole should be one of the antennas used. The other antennas used should be those typical of military systems.

6.7 Improve APD Estimation

The new bandwidth extrapolation rule for V_d developed for atmospheric noise at MF should be checked for applicability to man-made noise. The use of V_d to estimate the APD for man-made noise should be checked against measured APDs. Also, there should be an attempt to determine if stable Middleton model parameters exist for business, residential and rural areas. The orginal ITS data base of APDs could be used for this evaluation. If it is possible to obtain estimates of the Middleton model parameters for these man-made noise environmental categories, then it would be possible to make more refined estimates of APDs and other germane parameters.

6.8 Develop Man-Made Noise Simulators

Simulators are needed (both computer simulations and physical simulators) which can approximate various man-made noise environments.

6.9 Provide Unified Radio Noise Model

Currently, various organizations use different versions of existing models for noise from the same type of sources (e.g., atmospheric, man-made, and galactic). Models for noise from different sources frequently are not always combined properly. Spaulding and Stewart[21] have noted the need for a more unified approach and have provided an improved unified model for use in the IONCAP HF prediction model.[42] A new initiative in this area would involve the generation of a computer code in a standard, portable language (e.g., ANSI FORTRAN 77) that would combine (in the proper manner) the noise from various sources to give predictions of average noise power and the amplitude probability distribution (APD) in a specified bandwidth for a specified frequency, antenna type, geographical location, time, and environment. Versions of this code should be developed for both mainframe and mini/microcomputer applications. A users' guide should be provided for this model which discusses the rationale for the selection of any options which are offered. Such a code could be incorporated as a subroutine in telecommunication system performance models. There would also be some benefits in any comparative analyses of systems if different analysts used the same code for the noise prediction component of such system performance analyses.

6.10 Prepare Noise Model Users' Handbook

Currently, model users have a variety of noise models to select from, and model selection for specific problems is not always straightforward. Major and Segner[43] have noted that a model users' guide is needed. Such a guide should include a statement of the model's objective, a summary description of

each model, a list of inputs required and outputs available, a statement of any model limitations, and a model selection rationale for generic categories of problems with examples to illustrate each problem category.

6.11 Concluding Comment

This is the type of effort that is needed for progress in the area of man-made noise and to fill in the voids in our current knowledge.

REFERENCES

1. G.H. Hagn, "Definitions and Fundamentals of Electromagnetic Noise, Interference and Compatibility," presented at NATO Advisory Group for Aerospace Research and Development (AGARD) Conference on Electromagnetic Noise, Interference, and Compatibility, 21-25 October 1974, Paris, France, published in AGARD Conference Proceedings No. 159 (AGARD-CP-159), pp. 1-1 to 1-24, November 1975. Available NTIS, 5285 Port Royal Road, Springfield, VA 22151. AD-A-018 980.

2. CCIR, "Vocabulary (CMV)," Recommendations and Reports of the CCIR 1982, XVth Plenary Assembly, International Radio Consultative Committee, Geneva, Switzerland, 1982.

3. G.H. Hagn and R.A. Shepherd, "Man-Made Electromagnetic Noise from Unintentional Radiators: A Summary," NATO Advisory Group for Aerospace Research and Development (AGARD) Conference on Electromagnetic Noise, Interference, and Compatibility, 21-25 October 1974, Paris, France, published in AGARD Conference Proceedings No. 159 (AGARD-CP-159), pp. 3-1 to 3-24, November 1975. Available NTIS, 5285 Port Royal Road, Springfield, VA 22151. AD-A-018 980.

4. CCIR, "Man-Made Noise," Report 258-4, International Radio Consultative Committee, International Telecommunication Union, Geneva, Switzerland, 1982.

5. A.D. Spaulding and R.T. Disney, "Man-Made Radio Noise I: Estimates for Business, Residential and Rural Areas," OT Report 74-38, Office of Telecommunications, U.S. Department of Commerce, Boulder, CO, 1974.

6. ITU, "Report to the Second Session of the Conference," World Administrative Radio Conference for the Planning of HF Bands Allocated to the Broadcasting Service, First Session, International Telecommunication Union, Geneva, Switzerland, 1984.

7. G.H. Hagn and D.B. Sailors, "Empirical Models for Probability Distributions of Short-Term Mean Environmental Man-Made Radio Noise Levels," Electromagnetic Compatibility 1979, Proceedings of 3rd Symposium and Technical Exhibition on Electromagnetic Compatibility, Rotterdam, pp. 355-360, May 1-3, 1979.

8. G.H. Hagn, B.M. Sifford and R.A. Shepherd, "The SRICOM Probabilistic Model of Communication System Performance -- A User's Manual for Engineers, Applications Programmers and Systems Programmers," Final Report, SRI Project 3603, Contract NT-81-RC-16011, SRI International, Arlington, VA, May 1982, (with errata).

9. W.R. Lauber and J.M. Bertrand, "Man-Made Noise Level Measurements of the UHF Radio Environment," Symposium Record, 1984 IEEE National Symposium on EMC, San Antonio, TX, 14-26 April 1984.

10. E.N. Skomal, Man-Made Noise, Van Nostrand Reinhold, New York, 1978.

11. D. Middleton, "Statistical-Physical Models of Electromagnetic Interference," IEEE Trans. Electromagnetic Compatibility, EMC-19, August 1977, pp. 106-127.

12. D. Middleton, "Canonical Non-Gaussian Noise Models: Their Implications for Measurements and Prediction of Receiver Performance," IEEE Trans. Electromagnetics Compatibility, EMC-21, August 1979, pp. 209-220.

13. A.D. Spaulding, "Atmospheric Noise and Its Effects on Telecommunication System Performance," Chapter 7, Handbook of Atmospherics, Vol. 1, H. Volland, Ed., 1982, CRC Press, 2000 N.W. 24th Street, Boca Raton, FL, 33431.

14. A.D. Spaulding, "Effects of Noise and Interference on System Performance," AGARD, Lisbon,1987.

15. L.A. Berry, "Understanding Middleton's Canonical Formula for Class A Noise," IEEE Trans. EMC, Vol. EMC-23, No. 4, pp. 337-344, November 1981. See also, D. Middleton, "Canonical and Quasi-Canonical Probability Models for Class A Interference," IEEE Trans. EMC, Vol. EMC-25, No. 2, pp. 76-106, May 1983.

16. A.D. Spaulding and D. Middleton, "Optimum Reception in an Impulsive Interference Environment, Part I: Coherent Detection; Part II: Incoherent Reception," IEEE Trans. Com., Vol. COM-25, No. 9, pp. 910-934, September 1977.

17. D. Middleton, "Canonically Optimum Threshold Detection," IEEE Trans. Information Theory, Vol. IT-12, No. 2, pp. 230-243, 1966.

18. A.D. Spaulding, W.H. Ahlbeck and L.R. Espeland, "Urban Residential Man-Made Radio Noise Analysis and Predictions," Telecommunications Research and Engineering Report OT/TRER14, Institute for Telecommunication Sciences, Boulder, CO, June 1971.

19. E.W. Allen, "Man-Made Radio Noise," in The Radio Noise Spectrum, D.H. Menzel, Ed., Harvard University Press, Cambridge, MA, 1960.

20. D. Lucas and G. Haydon, 1966.

21. A.D. Spaulding and F.G. Stewart, "An Updated Noise Model for Use in IONCAP," NTIA Report 87-212, U.S. Department of Commerce, National Telecommunications and Information Administration, Boulder, CO, January 1987.

22. URSI, "The Measurement of Characteristics of Terrestrial Radio Noise," Special Report 7, International Union of Radio Science, Brussels, Elsevier Publishing Co., Amsterdam, 1962.

23. H.N. Shaver, V.E. Hatfield and G.H. Hagn, "Man-Made Radio Noise Parameter Identification Task," Final Report, Contract Number N00039-71-A-0223, SRI Project 1022-2, Stanford Research Institute, Menlo Park, CA, May 1972. AD 904 405.

24. A.D. Spaulding, "Man-Made Noise: The Problem and Recommended Steps Toward Solution," OT Report 76-85, Institute for Telecommunication Sciences, Boulder, CO, 1976.

25. G.H. Hagn, "Man-Made Radio Noise," Vol. I, Chapter 7, pp. 329-355, Handbook of Atmospherics, H. Volland, Ed., CRC Press, Boca Raton, FL, 1982.

26. R.J. Matheson, "Instrumentation Problems Encountered Making Man-Made Electromagnetic Noise Measurements for Predicting Communication System Performance," IEEE Trans. EMC, Vol. EMC-12, No. 4, pp. 151-158, November 1970.

27. R.J. Matheson, "DM-4 Operation and Maintenance Manual," NTIA-TM-80-50, Institute for Telecommunication Sciences, Boulder, CO, November 1980.

28. C.B. Christianson, "Low Cost APD Measurements of Noise," in DOD Radio Noise Symposium, Reston, VA, 7-8 March 1984, Ed., G.H. Hagn, et al.

29. C.B. Christianson, "Noise Measurement Systems Techniques," AGARD, Lisbon, 1987.

30. F.J. Massey, Jr., "The Kolmogorov-Smirnov Test for Goodness of Fit," J. Amer. Stat. Assn., Vol. 46, pp. 68-78, March 1951.

31. G.H. Hagn, "Interim DOD Radio Noise and Electromagnetic Interference (EMI) Survey Handbook," Final Report, Contract MDA 904-77-C-0555, SRI Project 6512, SRI International, Arlington, VA, July 1979.

32. IEEE, "IEEE Recommended Practice for an Electromagnetic Site Survey (10 kHz to 10 GHz)," IEEE Std 473-1985, Institute of Electrical and Electronics Engineers, Inc., 345 East 47th Street, New York, 18 June 1985.

33. A.D. Spaulding and G.H. Hagn, "On the Definition and Estimation of Spectrum Occupancy," IEEE Trans. EMC, Vol. EMC-19, No. 3, pp. 269-280, August 1977.

34. M. Kanda, "Time and Amplitude Statistics for Electromagnetic Noise in Mines," IEEE Trans. EMC, Vol. EMC-17, No. 3, p. 122, 1975.

35. J.R. Herman and X.A. DeAngelis, "Bandwidth Expansion Effects on the Voltage Deviation Parameter (V_d) of MF and HF Atmospheric Radio Noise," Radio Science, Vol. 11, No. 1, pp. 26-36, Jan.-Feb. 1987.

36. G.H. Hagn, R.A. Shepherd and J.H. Faulconer, "DOD Radio Noise Symposium, Reston, VA, 7-8 March 1984," Symposium Record, Contract NT-83-RAC-36001, SRI Project 5002, SRI International, Arlington, VA, June 1984. AD B085 493L.

37. G.H. Hagn and R.A. Shepherd, "Selected Radio Noise Topics," Final Report, Contract NT-83-RAC-36001, SRI Project 5002, SRI International, Arlington, VA, June 1984.

38. NTIA, "Manual of Regulations and Procedures for Federal Radio Frequency Management," National Telecommunications and Information Administration, Washington, D.C., 1987.

39. W.R. Vincent, "Examples of Signals and Noise in the Radio-Frequency Spectrum," IEEE Trans. EMC, Vol. 19, No. 3, Part II, p. 241, 1977.

40. J.R. Herman, "Electromagnetic Ambients and Man-Made Noise," Vol. 3, Multi-Volume EMC Encyclopedia Series, Don White Consultants, Gainsville, VA, 1979.

41. D.B. Sailors and R.P. Brown, "Development of a Minicomputer Atmospheric Noise Model," Radio Science, Vol. 18, pp. 625-637, 1983.

42. L.R. Teters, J.L. Lloyd, G.W. Haydon and D.L. Lucas, "Estimating the Performance of Telecommunications Systems Using the Ionospheric Transmission Channel: IONCAP User's Manual," Report 83-127, U.S. Department of Commerce, National Telecommunications and Information Administration, Institute for Telecommunication Sciences, Boulder, CO, July 1983.

43. P.A. Major and S.M. Segner, "Army Requirements for HF Noise Measurements and Models," in DOD Radio Noise Symposium, Reston, VA, 7-8 March 1984, Ed., G.H. Hagn, et al., SRI International, Arlington, VA, June 1984.

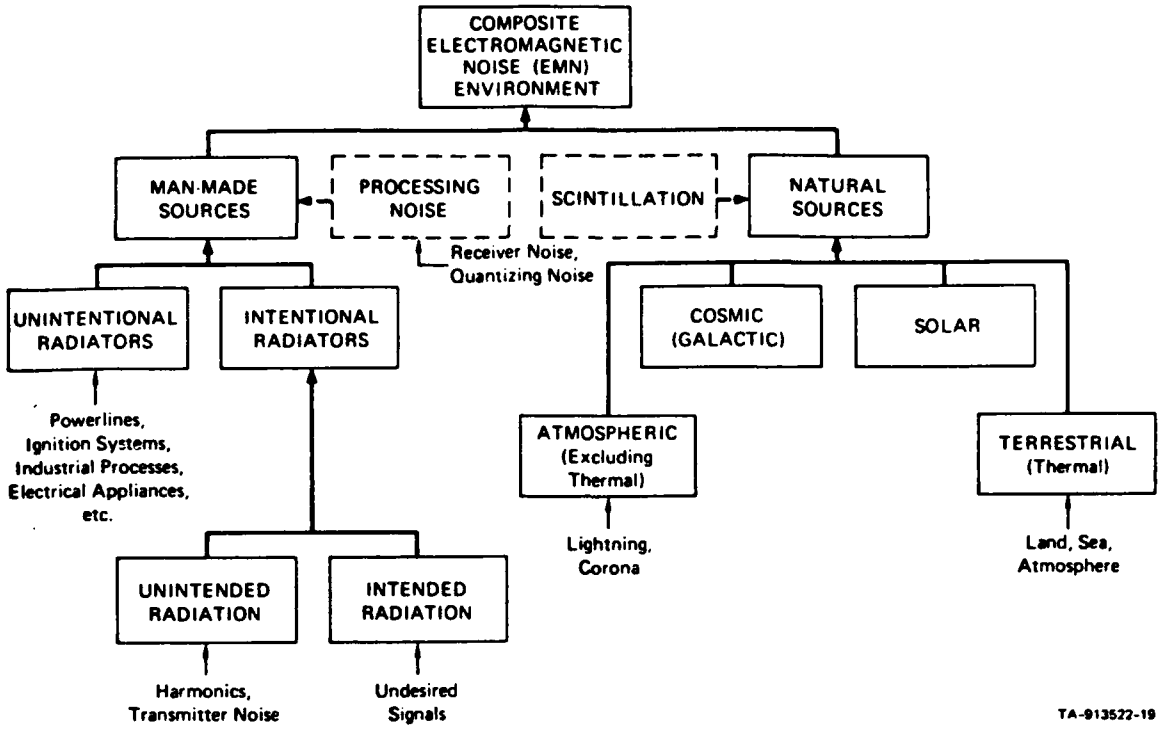


FIGURE 1.1 SOURCES OF THE COMPOSITE ELECTROMAGNETIC NOISE (EMN) ENVIRONMENT

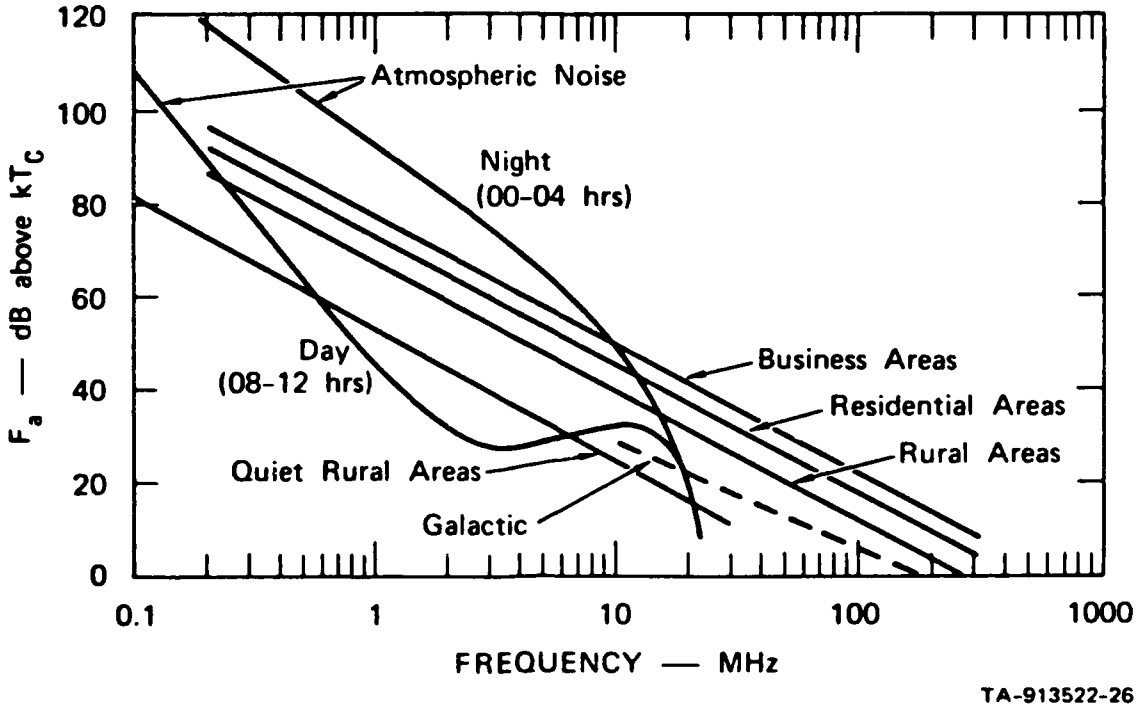


FIGURE 1.2 ESTIMATES OF MEDIAN VALUES OF MAN-MADE, ATMOSPHERIC, AND GALACTIC NOISE EXPECTED NEAR WASHINGTON, D.C. DURING SUMMER (AFTER SPAULDING AND DISNEY, 1974, AND CCIR, 1964)

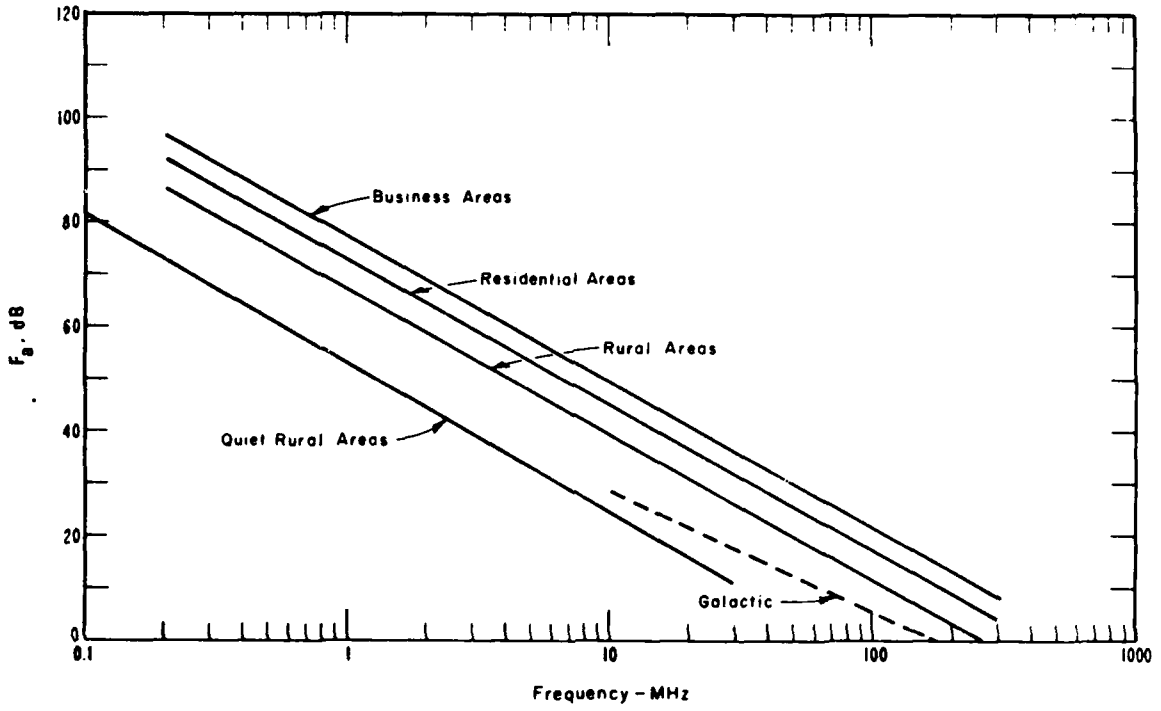
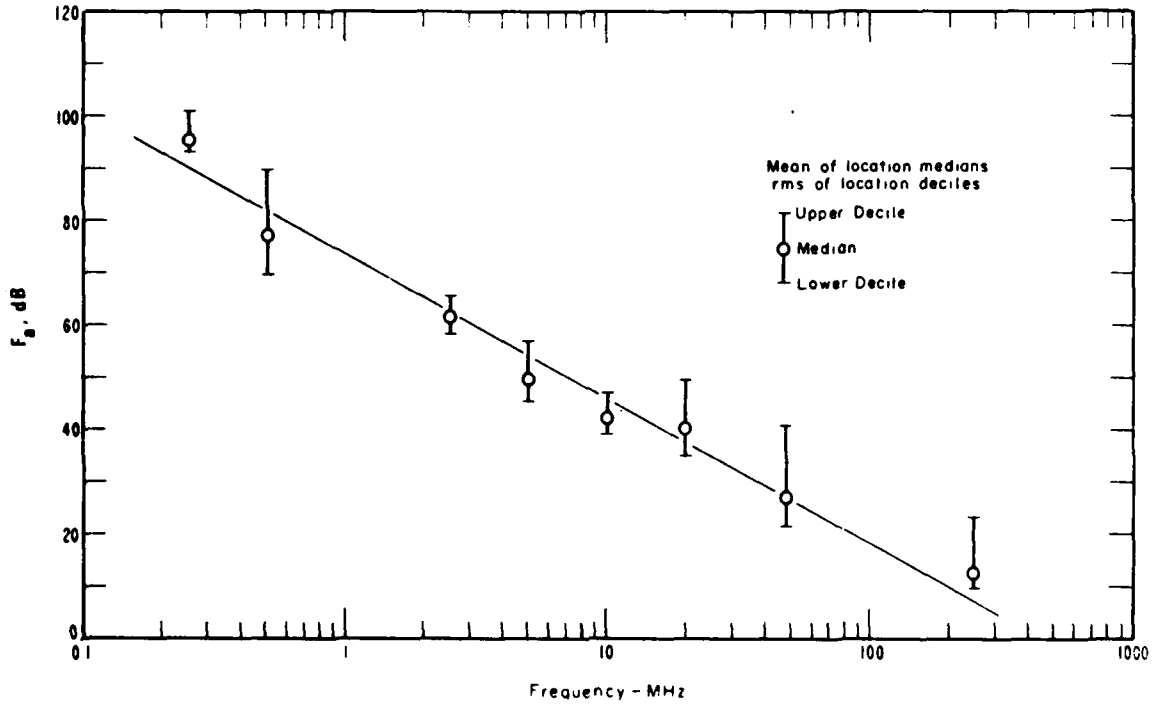


FIGURE 2.1 CCIR REPORT 258 ESTIMATES OF MEDIAN VALUES OF MAN-MADE NOISE EXPECTED AT TYPICAL LOCATIONS



(Spaulding and Disney, 1974)

FIGURE 2.2 ESTIMATES OF MAN-MADE NOISE LEVELS AND THEIR VARIATION WITHIN AN HOUR FOR INTERSTATE HIGHWAYS

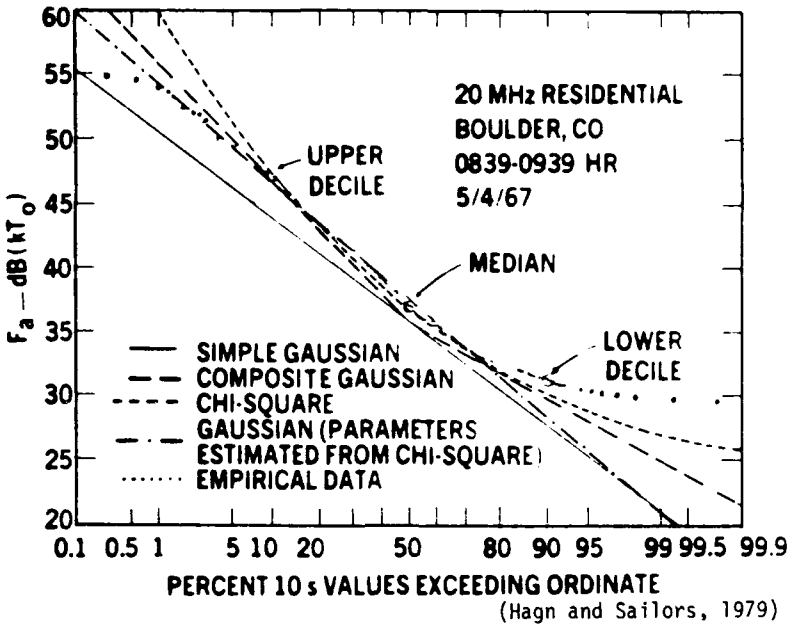


FIGURE 2.3 EXAMPLES OF MODEL PREDICTIONS FOR 20 MHz IN A RESIDENTIAL AREA

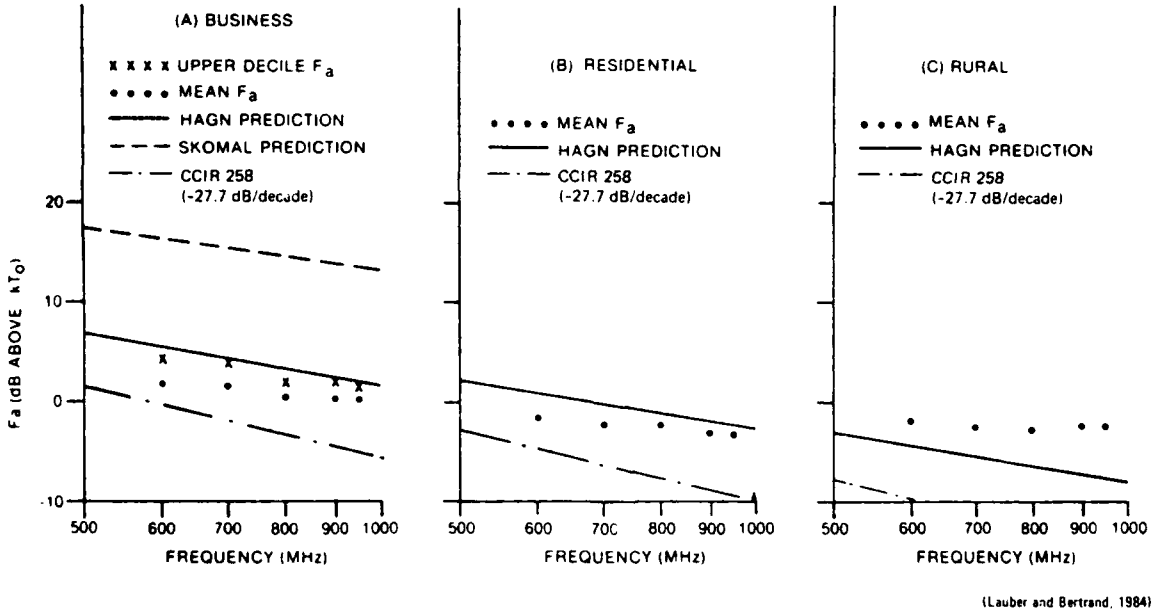


FIGURE 2.4 COMPARISON OF CANADIAN UHF MAN-MADE NOISE MEASUREMENTS WITH PREDICTIONS

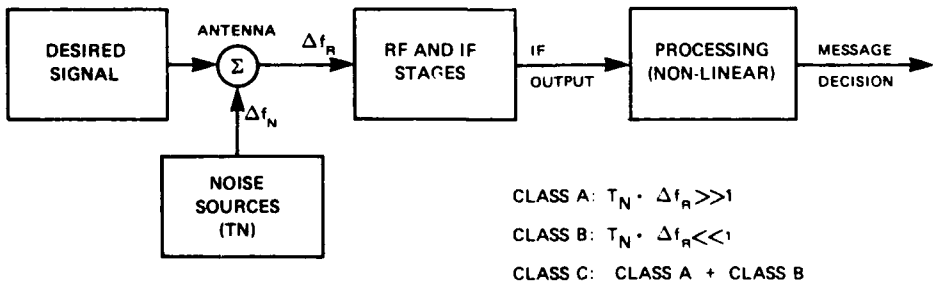


FIGURE 3.1 A TYPICAL NARROWBAND RECEIVER AND THE ELECTROMAGNETIC NOISE ENVIRONMENT FOR DIFFERENT CLASSES OF NOISE SOURCES

TRANSHORIZON COCHANNEL INTERFERENCE PATH MEASUREMENTS IN EUROPE

Martin P M Hall
Rutherford Appleton Laboratory,
Chilton,
Didcot,
Oxon OX11 0QX
UK

SUMMARY

Existing methods for predicting co-channel interference levels are very inadequate, and development of these methods has been severely limited by the lack of suitable data. There are particular problems within Europe because there is a wide range of climatic and geographical conditions, and COST Project 210 was set up to provide suitable data and prediction methods. The prime objective is to develop and evaluate models to serve as a basis for frequency planning co-ordination procedures and interference calculations, in order to improve the organisation of further radio communication systems within Europe.

After giving an up-to-date statement on the COST 210 activities set out above, the paper goes on to describe early results obtained from three representative areas of study:

i) Statistics of high signal level, and case studies, essentially in clear-air conditions, on paths of 150 to 300km from France to the UK (in collaboration with CNET (Paris), CNET (Lannion), Portsmouth Polytechnic, IBA and RSGB).
ii) Hydrometeor scatter studies using a 10cm wavelength dual-polarisation radar at RAL, where a radar-derived raincell database is being used to simulate interference paths and to examine relative effects of ice and rain, etc.
iii) Hydrometeor scatter statistics and case studies on a 131km overland path (in collaboration with DTI), and a 48km overland path, both in conjunction with the radar mentioned above, and two oversea paths of 201 and 302km (in collaboration with CNET)

1. INTRODUCTION

Under some conditions, microwave energy can propagate in the lower atmosphere over long distances and so produce co-channel interference between services. Engineers responsible for planning and co-ordinating radio systems must therefore have access to internationally-agreed interference prediction procedures in order to work effectively. These procedures are required for prediction of interference levels and establishing co-ordination distances, as laid down in the Radio Regulations of the International Telecommunication Union. The problems become particularly acute as high-density radio usage and interservice band sharing increases.

Current understanding of the various atmospheric mechanisms giving rise to transhorizon interference problems is far from complete. Propagation on transhorizon paths is influenced by more factors than are earth-space or line-of-sight terrestrial paths. At present, interference prediction models continue to have an essentially empirical approach which is seriously limited by the scarcity of measured radio data. The problems in Europe are especially difficult, because there is a wide range of climatic and terrain conditions. Also a co-ordination area may incorporate several small countries so that the practical and administrative implications of these coordination procedures are of serious concern. Although some work has been carried out throughout Europe, the study of interference mechanisms, and prediction of interference levels, has not been covered adequately. COST Project 210 has been established to provide the missing information.

The title of COST Project 210 is "The influence of the atmosphere on interference between radio communication systems at frequencies above 1 GHz". There is a wide range of technical problems needing to be investigated, where a radio transmission intended for one specific receiver location, or a location area, may produce interference to other receiving stations. Some of these situations are portrayed in Figure 1. The interference paths may be between independent systems or between independent services sharing frequencies (eg. Earth-space and terrestrial). The inter-service problems most of concern within Europe are considered to be potential interference between:

i) two stations on the Earth's surface, eg a terrestrial line-of-sight communication link terminal and a satellite-based communication Earth station (typically over a path of up to several hundred kilometres for large Earth stations, but now increasingly also over short paths for small Earth stations) or one terrestrial line-of-sight terminal and another (again usually over a long path, but not necessarily so);
ii) a terrestrial station and a space station (eg. an interference path travelling the whole atmosphere);
iii) two Earth-space communication paths (possibly both using relatively small antennas, when the spacing may be quite small).

These problem areas are being investigated on a world-wide basis within Study Group 5 (SG5) of the CCIR. This is the appropriate forum in which to reach agreement on interference prediction techniques that may be applied worldwide, and their current methods are contained within certain CCIR Reports. However, these reports are by no means a complete solution of the problems since many important gaps remain in the understanding of interference propagation. Moreover, the models in these reports cannot meet the accuracy required by the frequency managers in Europe in the coming years. However, whatever their strengths or weaknesses, the CCIR SG5 interference prediction models are the best that are currently available, and it is recognised that, if COST 210 is successful in producing a new and improved European model, then at some stage this will need to gain international agreement if it is to be used for formal co-ordination work, and it is likely that CCIR is the best place for this to be achieved.

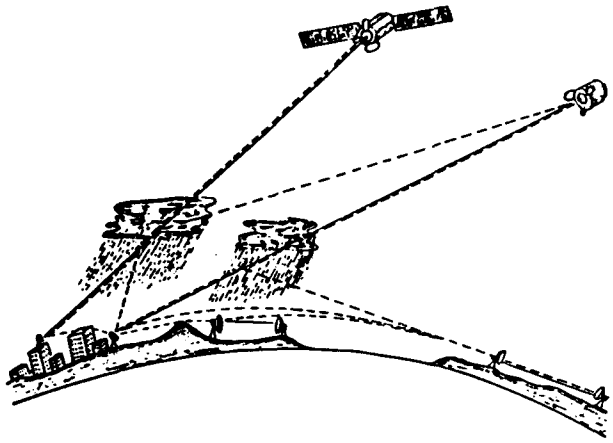


Figure 1: Paths of possible interference

Full curves: wanted paths B
Broken curves: unwanted paths

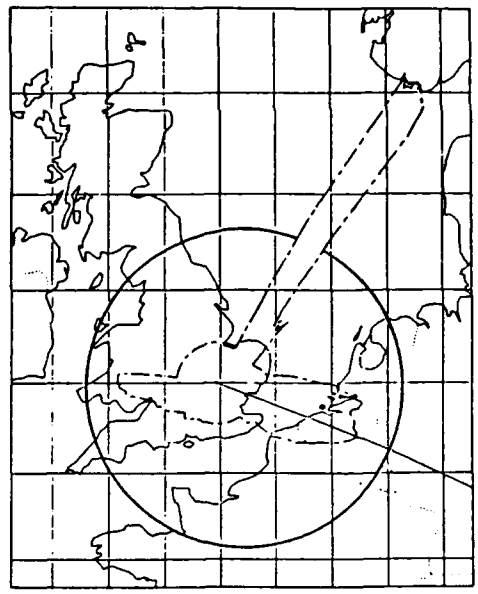


Figure 2: Co-ordination contour for an 11 GHz earth station

— - — - — clear air modes
——————— hydrometeor scatter

The modelling of propagation conditions likely to cause cochannel interference is usually separated into clear-air effects and hydrometeor scatter effects, ie scatter by rain and/or ice particles (eg hail or snow). Figure 2 illustrates contours of coordination distances drawn from the DTI station at Baldock using the CCIR SG5 procedures (1). Formal coordination, involving detailed computation of predicted interference levels, would be required if a new transmitter on a shared frequency were to be proposed within these contours. The dependence of the clear-air contour on length of land path and on horizon angle is clear, as is the large area for which the hydrometeor scatter contour is dominant. This figure illustrates the importance of having accurate propagation models.

Most attention has been given to clear-air effects, because the interfering signals and associated propagation conditions are fairly clearly recognised, and each event usually persists at a near-constant level for a considerable period of time. However, there are a very large number of factors influencing the received signal level on a transhorizon path, and accurate prediction has not always been achievable; indeed, quite large errors in prediction may occur. One of the networks of clear-air measurements being made is over paths of 150 to 300km from France to the UK. These enable special consideration to be made of the effects of antenna height (not yet considered in the CCIR model), frequency and coastal section of path. Some detail of this particular work programme is given in Section 3.

By contrast, hydrometeor scatter signals are considered much less easy to recognise; indeed, there is argument as to whether they have been observed at all on operational systems! Again it is difficult fully to take account of the many different factors which determine the received signal level, or even the nature of the dominant rain cells. However, modern radars are playing a major part in examining this problem, and RAL's dual-polarisation radar at Chilbolton is active in this field. In addition a short bistatic path is being used to examine effects of geometry, and to confirm the accuracy of the radar modelling, and long paths are being used to give statistics for operational path lengths. Some detail of this work programme is given in Section 4.

2. THE WORK OF COST 210

The broad objectives of COST Project 210 are (in a logical sequence) as follows:

i) to collate and assess the results of existing European investigations of propagation aspects relating to radio interference at frequencies above 1 GHz;
ii) to promote and co-ordinate new experiments where these are needed;
iii) to produce a comprehensive unified data bank from data already being obtained and data to be derived from new experimental paths;
iv) as a final result, to develop and evaluate models to serve as a basis for frequency planning co-ordination procedures and interference calculations with respect to future radio communication systems within Europe.

The prediction method to be developed should aim for inter-system distances to be a realistic minimum, particularly over land. However, it will need to take into account a wider range of propagation factors than are included in existing methods. The emphasis will be on accuracy under realistic practical interference conditions rather than extreme cases. Prediction of "worst month" as well as annual propagation conditions is considered to be of great importance.

Whilst the "full" method may need to be fairly complex in order to achieve the required accuracy, it is recognised that eventually a relatively simple method will be needed to fulfil the role currently occupied by the CCIR prediction method. Such simplification implies a larger standard deviation for the expected variability around the predicted transmission loss than would be needed for the more sophisticated method. The evaluation of these standard deviations is seen as an important part of the work.

In some circumstances, the dominant factor in interference will be a clear-air propagation mode on a great-circle path (diffraction, reflection, scattering or ducting). In other conditions (when beams from two stations intersect outside the great-circle path) hydrometeor scatter may dominate. In both cases, interference-reduction techniques will be important. For this reason, the technical work of the Project has been organised within three Working Groups:

WG1: Interference in Clear Air
WG2: Interference due to Hydrometeor Scatter
WG3: Interference Reduction Techniques

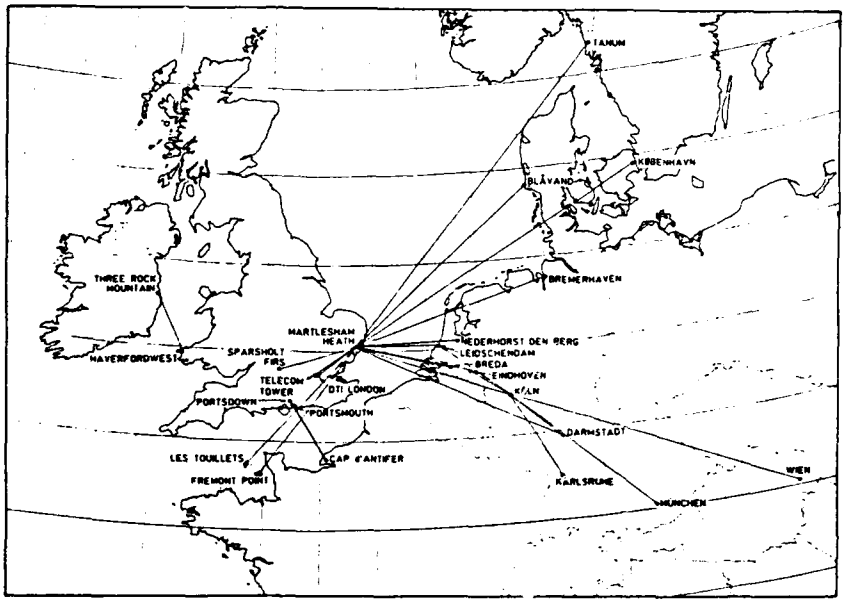


Figure 3 : 1.3 GHz paths for COST 210 clear-air studies.

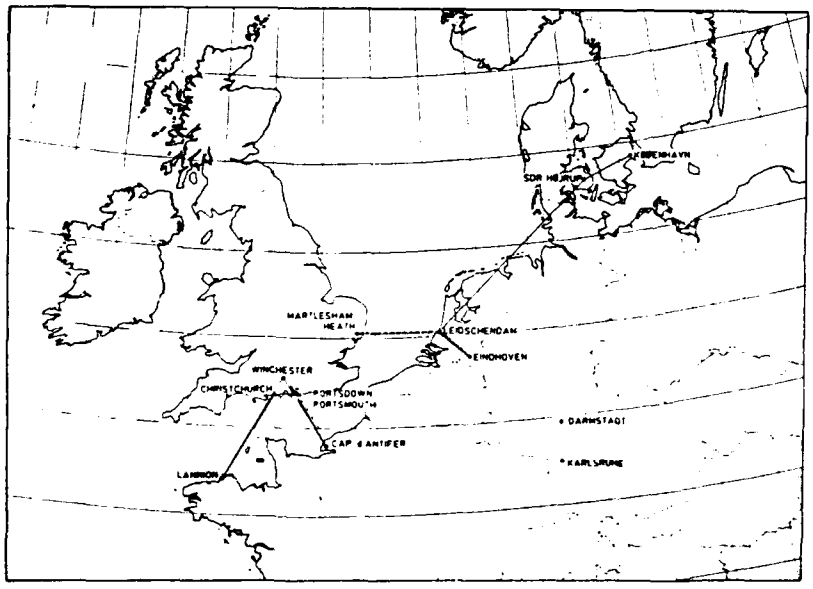


Figure 4 : 11 GHz and higher frequency paths for COST 210 clear-air studies.

———— 11 GHz
— — — — 19/22.3 GHz
—·—·—·— 24 GHz
..... 29 GHz

For the prediction of Interference in Clear Air (WG1) data available from earlier experiments are severely limited. In consequence, priority has been to provide experimental measurements on some 45 new paths early in the project since data must be collected over 2 or 3 years. Some of these are shown in Figures 3 and 4. These measurements are grouped into the 5 networks and two individual paths shown in Table 1.

Table 1: COST 210 clear air measurement networks, and additional paths

1	North Sea	1.3 GHz	5	Denmark/UK	11 GHz
2	European Mainland	1.3 GHz	a	Irish Sea	1.2 GHz
3	Cross-channel	1.3/11/24 GHz	b	Rhine Valley	30 GHz
4	European Coastal	11/19/22.3/29 GHz			

Work has also started on examining propagation models, and, in addition to the primary objectives and collecting transmission loss statistics to develop prediction methods, specific studies to progress this modelling include the following:

i) the frequency dependence of clear-air interference propagation mechanisms;
ii) the distance dependence, including the physical extent of atmospheric ducts;
iii) the influence of the surface of the earth, particularly the characteristics of the terrain;
iv) the particular problems of predicting propagation just beyond the horizon, where diffraction may be the dominant mode, and in coastal areas;
v) gaseous absorption during anomalous clear-air propagation conditions;
vi) antenna-to-propagation mechanism coupling;
vii) the possible roles for meteorological data in clear-air interference modelling and prediction.

Several studies of Interference due to Hydrometeor Scatter (WG2) are already in progress or are being planned. These comprise:

- eight short-path bistatic experiments (up to some 50km path length) covering the frequency range 11-34 GHz;
- five long-path bistatic experiments over 100-300km (four at 11 GHz, one at 30 GHz);
- five (monostatic) radar experiments (2.8-6 GHz).

Some of the experiments involve joint bistatic and radar studies. The locations of facilities are indicated in Figure 5.

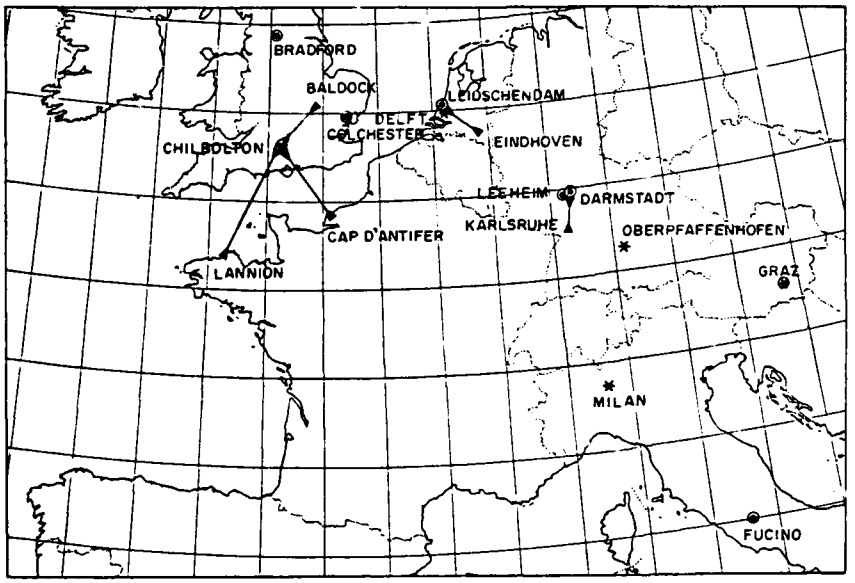


Figure 5 : Facilities for COST 210 hydrometeor scatter studies.

θ Short-path bistatic experiments
Long-path bistatic experiments
* Radar experiments

In some respects, the bistatic experiments may be regarded as providing propagation data most directly applicable to actual coordination work since the geometry and mode of operation can be close to those of real communication systems. Some of the experiments involve steerable antennas at one or both terminals to (a) collect more data, and hence arrive at more representative statistics than that obtained with a fixed geometry, (b) show spatial structures of the medium and (c) reveal direction-dependent properties of the scattering process. All short-path bistatic experiments are supplemented by line-of-sight links between the bistatic terminals and/or rain gauges below the path, the intention being to separate the attenuation and scatter contributions to the bistatic transmission loss. The long-path experiments are especially suited to collect long-term statistical transmission loss data which will be used to check and validate both the radar-derived data and prediction procedures.

Specialised meteorological radars operating alone may be used to (a) produce three-dimensional maps of precipitation from which to examine various geometries of off-beam rain scatter (computing the statistics of transmission loss on practical path lengths and taking account of both scatter and attenuation), and (b) to make studies of hydrometeor characteristics (eg phase, sizes, shapes, orientation) and mesoscale studies of precipitation (eg cell shape, height, etc).

In addition, theoretical studies and modelling are required to examine rain, melting-layer and ice scatter (also wet snow and wet ice), the temporal and spatial distributions of precipitation, and the effects of antenna diagrams on the sizes and shapes of scattering volumes (including cases of imperfectly intersecting antenna beams). Finally a clear procedure for prediction will be required.

Apart from preparing statistics of transmission loss on practical paths to develop prediction methods, specific studies to progress the necessary modelling include the following:

i) combined effects of scattering and attenuation as a function of radio frequency;
ii) effect of the spatial structure of precipitation;
iii) role of different types of hydrometeors (rain, hail, snow, water and ice cloud), and relative effects of ice (including snow) above and rain below the melting layer, again as a function of frequency;

iv) effect of path geometry (e.g. forward and back-scatter, etc);
v) dependence on polarisation (with scatter angle and antenna orientation), and influence of the extent of beam coupling and sidelobe coupling;
vi) fading characteristics of the scattered signal;
vii) validity of radar-derived (i.e. simulated) data.

Early consideration of Interference Reduction Techniques (WG3) concentrated on the problem of site shielding, but some activity has since been extended to other techniques. These studies are limited to aspects relating to the use of radio propagation data; it is not intended that emphasis be given to system studies. The topics being investigated are:

i) characteristics of the interfering signals;
ii) antenna sidelobe suppression;
iii) filters and cancellation networks;
iv) site shielding.

Investigation of these techniques will be essentially analytical, and no separate experiments are planned. The experimental input for this WG follows as a spin-off from the experiments of the other WGs, as well as from older investigations described in the literature.

COST 210 has a membership from the following ten countries: Austria, Belgium, Denmark, France, the F R of Germany, Ireland, Italy the Netherlands, Sweden and the United Kingdom. Some 25 organisations participate in the Project, which is scheduled to operate for six years, ending in mid 1990. So far the programme of work is well advanced. Further details may be found in the first and second Annual Reports (2, 3).

3. ANGLO-FRENCH STUDIES OF CLEAR-AIR PROPAGATION

A specific network within COST 210 is the third in Table 1, namely the Anglo-French cross-Channel network. CNET, in France, and IBA, Portsmouth Polytechnic, RAL and RSGB, in the UK, have set up over-sea paths and mixed-land-sea paths in a network operating primarily in the 10.7 - 11.7 GHz band allocated by WARC 1979 to the fixed, fixed-satellite and mobile services, but also at 1.3 GHz, 24 GHz (proposed) and 27 GHz. Not all are yet operational at the time of writing. The locations are indicated in Figure 6, and some details of the paths are given in Table 2.

Table 2: Anglo-French path characteristics

Path	Freq. (GHz)	Antenna terminal heights, m (asl)		Path length km
		Trans	Rec	
1 Cap d'Antifer (high)-Portsmouth (Eastney) (low)	1.296	120	13	153
2 Cap d'Antifer (low)-Portsmouth (Eastney) (low)	11.646	35	13	153
3 Cap d'Antifer (high)-Portsmouth (Eastney) (low)	11.647	115	13	153
4 Cap d'Antifer (high)-Portsmouth (Portsdown) (high)	1.296	120	110	160
5 Cap d'Antifer (low)-Portsmouth (Portsdown) (high)	11.646	35	110	160
6 Cap d'Antifer (high)-Portsmouth (Portsdown) (high)	11.647	115	110	160
7 Cap d'Antifer (high)-Portsmouth (Eastney) (low)	27.5*	120	13	153
8 Cap d'Antifer (low)-Winchester	11.646	35	175	192
9 Cap d'Antifer (high)-Winchester	11.647	115	175	192
10 Lannion-Christchurch	10.710	145	20	254
11 Lannion-Winchester	10.710	145	175	297

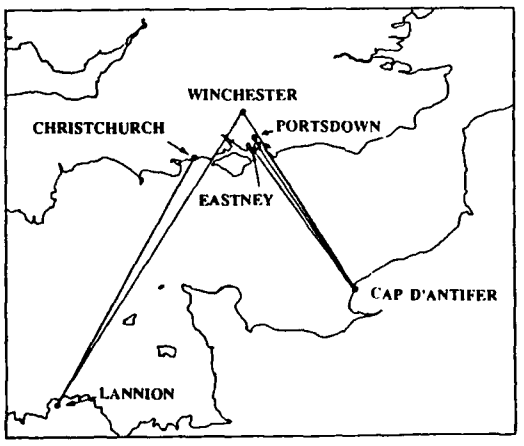


Figure 6: Map of radio paths

* Path 7 is also to be operated at 24.048 GHz.

The principal objectives of these clear-air studies are as follows:

i) To produce annual and monthly cumulative statistics (see Item v below).
ii) To produce statistics of the duration and diurnal times of occurrence of signal enhancements during ducting events, and joint statistics of transmission loss with duration etc.
iii) To study simultaneous path losses from Cap d'Antifer over low-low, low-high and high-high path geometries, the paths being almost identical in other respects. Studies will be made of both individual events and long-term statistics. At present, the CCIR model takes no account of terminal height. The paths will also allow some study of the effect of an additional 40 or 50km traversed over costal land (for Cap d'Antifer and Lannion paths, respectively).
iv) To produce 1.3/11/24/27 GHz frequency-scaling factors.
v) To compare the derived statistical data with results obtained from prediction models, eg CCIR Report 569-3 and 724-2, Appendix 28 of the ITU Radio Regulations, the CNET "angular distance model" and any new COST 210 method.
vi) To study the angle of arrival of signals using data obtained at Portsmouth (high terminal).
vii) To obtain frequency-dispersion statistics from measurements on the Lannion - Christchurch path.
viii) To study the physical structure of ducts over the sea, the land-sea interface and the land areas at differing heights using an aircraft-mounted refractometer.

Figure 7 illustrates an early event on the Cap d'Antifer to Eastney (Portsmouth, low) path. Preliminary inspection of the chart record suggests perhaps some enhanced troposcatter, lifting the signal level to within about 30dB of free space (at the maximum), a level predicted by CCIR to occur for about 6% of time (4).

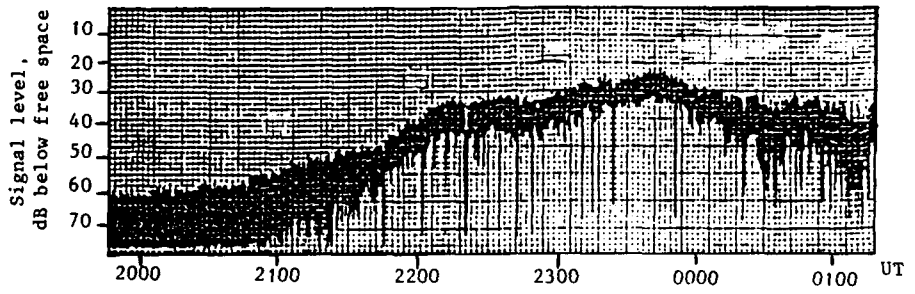


Figure 7: Signal enhancement on 20-21 October 1986, 20-01 hrs UT, on high-to-low Cap d'Antifer-to-Portsmouth path

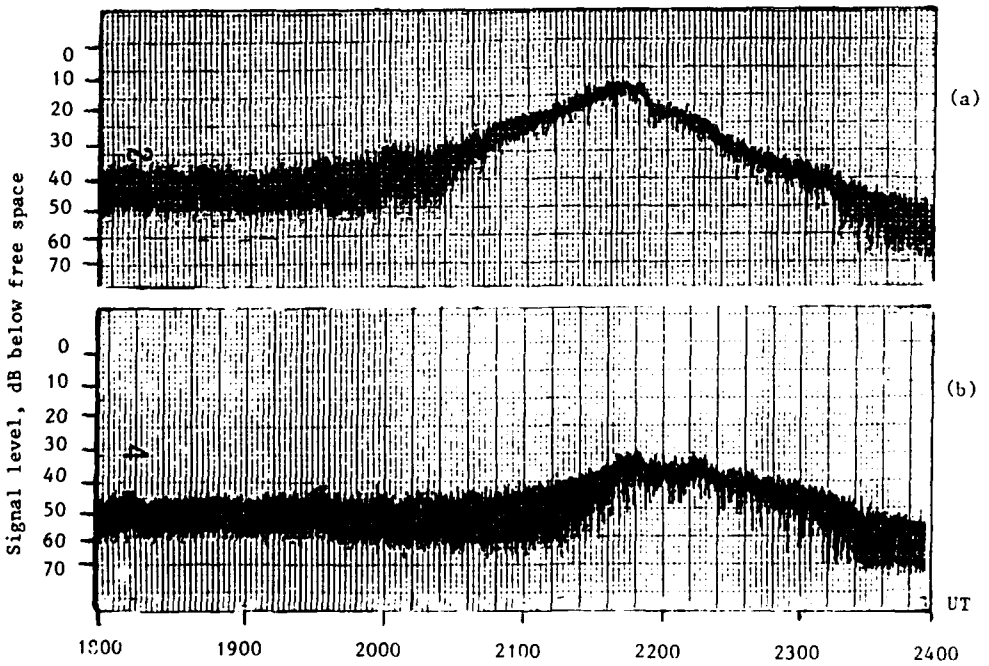


Figure 8: Signal enhancement on 20 April 1987, 18-24 hrs UT, on (a) high-to-low and (b) low-to-low Cap d'Antifer-to-Portsmouth paths

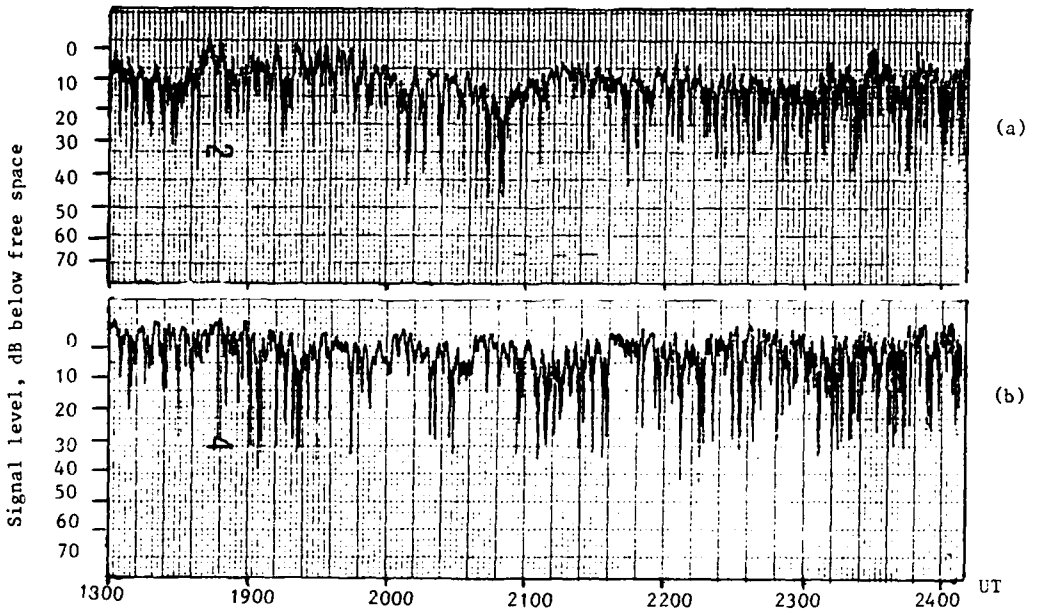


Figure 9: Signal enhancement on 21 April 1987, 18-24 hrs UT, on (a) high-to-low and (b) low-to-low Cap d'Antifer-to-Portsmouth paths

Figure 8 (a) shows a signal enhancement on the high-to-low Cap d'Antifer-to - Portsmouth path to within less than 10dB of free space, a level predicted to be exceeded for only 0.5% of time, though the whole event lasts only some two or three hours. Part b shows less enhancement on the low-to-low path, only to within about 35dB of free space, and only during the latter half of the high-to-low path event. It is too early to draw any firm conclusions on the cause of this example, but it would appear that a weak front crossed the path and had more effect on the upper path than on the lower path. Unlike Figure 7, the data in Figures 8 and 9 were subject to a 1s low-pass filter.

Figure 9 shows part of a much more interesting event, for which there was enhancement during 36 hours. In part a, the high-to-low path signal level is seen to be generally between 0 and 5dB below the free space level, free space being predicted to be exceeded for 0.05% of time, whilst the low-to-low path is generally between 0 and 10dB above free space. The fading rate is now quite low, but not as low as has been reported at lower radio frequencies. The fade rate may be governed by changes in the refractive index structure of the duct, or by more than one ducting mechanism being present. It is perhaps significant that the high-to-low signal level is only some 10 dB below the low-to-low signal level. The diversity of transmitter and receiver heights is expected to assist these studies greatly.

At the time of writing only preliminary statistical studies have begun. The duration of events study (item ii of objectives of this work) and frequency-scaling study (item iv) form part of the statistical package (item i). Item v) will flow from this. Preparations for aircraft-mounted refractometer studies and well advanced (5).

4. RAL STUDIES OF HYDROMETEOR SCATTER

4.1 Dual-polarisation radar studies

The main RAL facility for study of hydrometeor scatter is the dual-polarisation radar at Chilbolton. Apart from operating in dual-polarisation the main features of the radar are the large antenna (25m diameter giving 0.28° one-way 3dB beamwidth) and non-attenuating radio frequency (3.1 GHz). Differential polarisation is measured to within about ±0.1dB, and reflectivity to within about ±0.75dB. The radar is well tested, having been operating in dual-polarisation mode much of the time since 1978. The dual-polarisation allows a more clear distinction to be drawn between rain and ice phases, and a better estimate to be made of rain attenuation (6). The main data currently available are for 1981 and 1982, but a start has now been made on collecting a further two years of data.

Dual-polarisation radar data are being used to model coupling by scatter, using the radar data to predict the scatter from the common volume and attenuation on paths to this common volume from hypothetical Earth stations and terrestrial terminals. In this way the dependence of the coupling on the height of the common volume, and on the geometry is being investigated. Further, radar data are being used to provide statistics so that the likelihood of strong coupling may be estimated.

Figure 10 shows contour plots of (a) the radar reflectivity, Z, and (b) the differential reflectivity with polarisation, Zdr, from a vertical scan through a strong convective raincell (7 and 8). Bistatic scattering paths through the raincell were considered, with the centre of the common volume at two different heights, 1.5 and 3.0km, and at a series of ranges from the radar between 32 to 45km in 1km steps. Assuming a typical (radial) cell velocity of 5 m/s, this can be considered as a time series with approximately 3 minutes between samples in a sequence. Part d of Figure 10 shows the variation with the range of the common volume of the net coupling (scatter less attenuation) expressed as the ratio of scatter power received by the Earth station, Pr, to the power that would be caused by free-space (ducted)

propagation, Pf, into the Earth station sidelobes (see below). It can be seen that this coupling broadly follows the reflectivity contours, except for sharp decreases at around 36km and 39km for the 1.5 and 3km heights respectively. (The vertical dot-dash line links the patterns for 38 km range). Part c of the figure shows that, at these ranges, attenuation on the Earth station path, calculated for the frequency of 20 GHz, has become very large, reducing the coupled signal accordingly. The general pattern is for strong coupling to precede and follow the attenuation event.

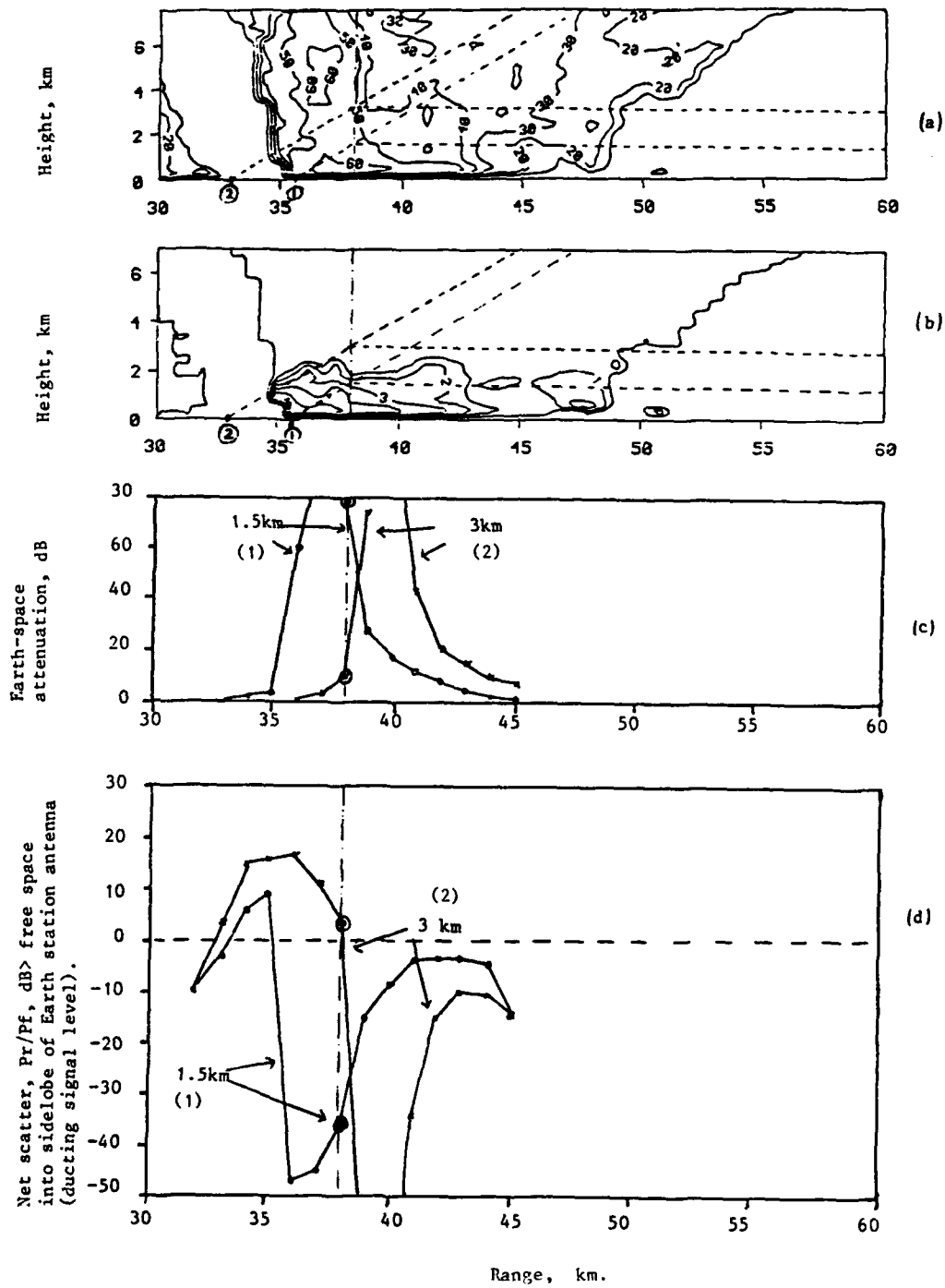


Figure 10: Contours of a) reflectivity, Z, and b) differential reflectivity, Z_{DR}, collected during a vertical radar scan (RHI) on 6 July 1983 at 1332 UT. Also shown is c) attenuation on wanted path up to the melting layer and d) signal on unwanted path, where the range is that of the intersection of the Earth station and terrestrial terminal beam axes.

Positions 1 and 2 are Earth station-locations with common volume at 38 km range from the radar, and the vertical line at 38 km range links parts a, b, c and d in this case.

The modelling assumed the gain of both the terrestrial terminal and Earth station antennas was 42 dB, and that the stations were 100km apart. The coupling was calculated for 20 GHz assuming the Earth station beam axis was at 30° elevation, as may occur with propagation by ducting. (The sidelobe level was assumed to be that for a CCIR model Earth station, which for 30° off the axis of a 42 dB gain antenna is -47 dB relative to the maximum gain (9).

The example shows that strong coupling is possible with the ratio Pr/Pf up to 17 dB while the attenuation of the wanted satellite signal may only be attenuated by a few dB. Also the level of coupling is very dependent on the height of the common volume. The accuracy of the radar prediction of the coupling is limited by experimental errors and assumptions in the technique, and is estimated to be up to 10 dB (7).

A study is currently being made of the effect of common volume height, the antenna beam widths and forward and backwards scatter, and statistics have been prepared from data collected in 1981 and 1982. Unfortunately it is still too early to report on these studies, but they are expected to add significantly to the modelling of hydrometeor scatter. The data bank is now to be extended by a further two years of data.

4.2 Short-path studies

Measurements of scatter on short (bistatic) paths are being made at 17.8 GHz on a 48km baseline terminating at Chilbolton. The 1.7° transmitter beam is steered on to a suitable raincell using the Chilbolton radar display. The 17.8 GHz receiver is mounted on the same 25m diameter antenna as the radar (3' beamwidth), and the net scatter is measured (ie scatter less attenuation). The radar is then scanned to evaluate the attenuation on transmit and receive arms of the bistatic path.

The objectives are:

i) To measure bi-static scatter from various types of rain, and from ice above the melting layer, over a wide range of scattering angles.
ii) To comment on the relative importance of the various forms of precipitation, as they affect calculations of interference and coordination distance.
iii) To estimate (from the radar data) the attenuation along the bi-static path, and hence to measure the scattering cross-section in rain and ice over a wide range of scattering angles.
iv) To compare the direct measurement of bi-static coupling with that estimated from the 3 GHz radar data, in order to comment on the validity of radar data alone as a means of investigating problems of interference from hydrometeorscatter.

Limited results from the comparative experiment between the radar and 18 GHz bistatic experiment have been obtained at the time of writing. Only weak rain at 1mm/hr was measured, but a coupling of -103dB relative to free space was detected and this agreed with the radar prediction to within experimental error. Great importance is placed on making many more measurements of this type, particularly in convective rain.

4.3 Long-path studies

As well as studies of hydrometeor scatter mechanisms by radars and short-path systems, and estimation of transmission loss statistics for practical path lengths from data banks collected by these means, there is a need to give a demonstration on the statistics of transmission loss on longer path lengths representative of possible interference paths. RAL is involved in two collaborative studies, one with DTI, and one with CNET.

The collaborative study with DTI involves an 11.2 GHz transmitter at Chilbolton, with a 1.6° beam pointed at the horizon, and a receiver 131km distant at Baldock, with an 11' beam pointed at a elevation of 20°, first towards the transmitter and then away from it. The fixed scattering height extends from 1.3 to 5.6 km. Statistical data are collected on a time sampling basis. The dual- polarisation radar is scheduled to examine the scattering common volume of these paths at selected times.

The collaborative study with CNET involves the transmitters at Cap d'Antifer and Lannion mentioned in Section 3, each with transmitter beam of 2.8° and path lengths of 201km and 302km, respectively. The receiver is mounted on the 25m diameter antenna at Chilbolton (5' beamwidth) pointed up to an elevation of 20°. The scattering height then extends from 1.5 km to 6.4 km for the Cap d'Antifer path, and 3.7 km to 11 km for the Lannion path. Again statistical data are to be collected on a time sampling basis.

5. CONCLUSIONS

The work of COST Project 210 is now well advanced. Many new radio paths and other facilities have been, or soon will be, put into operation to collect much-needed data. The Management Committee has decided on the format in which the data of the individual measurements should be collected and stored, to facilitate their later evaluation. Some new data have already become available and a large database can be expected to be built up within the coming 2-3 years. First results of the measurements have been discussed in recent meetings of the Management Committee.

As the work progresses, the emphasis is now changing gradually from the planning and implementation of new facilities to a critical review of existing models and interference prediction procedures, and to the development of new ones. Interesting ideas have been proposed, and these will have to be tested against the increasing amount of measured data.

It is to be expected that these studies will lead to considerable improvements in the prediction of interference levels and coordination distances. The ultimate goal of the study will be the development of a fully self-contained, comprehensive prediction procedure, including computer source listings, for application in European areas. Previous collaborative COST projects have been found to be highly stimulating and productive, and inputs from others working in the area would be most welcome.

6. REFERENCES

1 CCIR Report 724-2: "Propagation data required for the evaluation of coordination distance in the frequency range 1-40 GHz"; Recommendations and Reports of the CCIR, Volume V, CCIR, Geneva, 1986.

2 COST Project 210: First Annual Report 1984-85; DG XII/512/86; EUCO-TELE/210/4/85 EEC, Brussels, December 1985.

3 COST Project 210: Second Annual Report 1985-86; DG XIII/008/86; EUCOTELE/210/2AR/86; EEC, Brussels, December 1986.

4 CCIR Report 569-3: "The evaluation of propagation factors in interference problems between stations on the surface of the Earth at frequencies above about 0.5 GHz"; Recommendations and Reports of the CCIR, Volume V, CCIR, Geneva, 1986.

5 U M Yilmaz, G R Kennedy and M P M Hall (1985); "A GaAs FET microwave refractometer for tropospheric studies"; AGARD Conf Publn 346: Characteristics of the lower atmosphere influencing radio wave propagation, AGARD, Paris.

6 J W F Goddard and S M Cherry (1984); "The ability of direct polarisation radar (copolar linear) to predict rainfall rate and microwave attenuation"; Radio Science, 19, 201-208.

7 S M Cherry and J W F Goddard (1985); "The use of dual-polarisation radar in rainscatter interference studies"; Proc ICAP 85, IEE Conf Publn 248: Antennas and Propagation, 18-21.

8 J W F Goddard and S M Cherry (1985); Private communication.

9 CCIR Recommendation 465-2: "Reference earth-station radiation pattern for use in coordination and interference assessment in the frequency range from 2 to about 30 GHz"; Recommendations and Reports of the CCIR, Volume IV - 1, CCIR, Geneva, 1986.

7. ACKNOWLEDGEMENTS

Much of the material presented in this paper has previously been presented in COST Reports, and is presented here with the permission of COST 210 Management Committee which has 38 Members, Deputies and Experts. In addition, a large number of research workers in the 10 member countries contribute to the project from the following organisations: BTRL (UK), CNET, Lannion (F), CNET, Paris (F), Delft University of Technology (NL), DFVLR (FRG), Direction de la Meteorologie (F), DNL (NL), FI/DBP (FRG), FUB (I), IBA (UK), Polytechnic of Milan (I), Polytechnic of Wales (UK), Portsmouth Polytechnic (UK), Posts and Telegraphs of Denmark (DK), RAL (UK), DTI (UK), RSRE (UK), Swedish Telecomms Admin (S), Technical University of Graz (A), Telecom Eireann (IRL), University of Bradford (UK), University of Essex (UK) and University of Technology, Eindhoven (NL). The Chairman is Mr M P M Hall (UK), the Vice Chairman is Mr A Ochs (FRG) and the Chairmen of Working Groups 1, 2 and 3 are Mr M T Hewitt (UK), Mr A Ochs (FRG) and Mr P Scheeren (NL), respectively. Material relating to the Anglo-French studies is presented with the permission of the collaborating organisations (CNET, Paris, Mme M Juy and Mr M Rooryck; CNET, Lannion, Mr L Martin; IBA and RSGB, Mr D T Hayter; Portsmouth Polytechnic, Prof E Vilar and Miss C Spillard). RAL studies based on radar data from Chilbolton involve Mr S M Cherry and Mr J W F Goddard, as well as Dr A Awaka, currently visiting from RRL, Japan. RAL studies of clear-air events and bistatic hydrometeor studies involve Dr P C Barber, Dr G R Kennedy, Dr M Thurai and Mr J Eastment. At DTI, Mr J Warden and Mr K R E Dunk are primarily responsible for the Chilbolton-Baldock path.

EXPERIMENTAL OBSERVATIONS OF SPECTRAL OCCUPANCY AT HF

by
Geoffrey F Gott
Patrick J Laycock
Arup R Ray
Mark Morrell
University of Manchester Institute of Science and Technology
Manchester
England

SUMMARY

An experiment to measure spectral occupancy at HF has been undertaken since the sunspot maximum of 1982, and examples of the results are presented. Also, a mathematical model is being fitted to the experimental data, and initial results of this aspect of the work are included.

INTRODUCTION

A programme of work exists at UMIST, supported by the UK Science and Engineering Research Council and UK Ministry of Defence (RSRE), to investigate spectral occupancy within the HF band 1.5-30MHz.

The aim of the work is to provide data which may be used in conjunction with frequency predictions, to advise HF operators on the typical occupancy they may encounter, and how this may vary with threshold level, frequency, time, bandwidth, type of user allocation, and geographical location. Such information will also be useful to communication system designers, to HF ground wave users (who may then choose operating frequencies to avoid severe interference from sky wave users), and also to study groups who are concerned with the determination of international frequency assignments.

OCCUPANCY MEASUREMENTS

Spectral measurements are made twice yearly across the whole HF spectrum, approximately at the times of the winter and summer solstice, when the diurnal variation in the optimum working frequency is maximum and minimum, respectively. Such measurements have been made since 1982 (following the WARC 1979 changes), and correspond to a range of Belgian sunspot numbers varying from 130 to ?.

The experimental site is RSRE Pershore, a rural site in central England, with a low man-made noise level, and the receiving aerial is an active, wideband, vertical monopole.

By ITU regulations, the HF band is divided into frequency allocations which are dedicated to specific groups of users, e.g., aeromobile, fixed, broadcast, etc. In the measurement of occupancy for the different HF users, a Racal 1792 communication receiver, operated without AGC, and having a bandwidth of 1kHz, is stepped in 1kHz increments through each user defined allocation, spending one second at each increment. Each 1kHz channel is defined as occupied at a particular threshold level if the average signal value exceeds the threshold in the one second observation period. The percentage of such channels occupied across each user allocation then defines the 'congestion' for that allocation.

The communication receiver is used in preference to a spectrum analyser, because the receiver filters are more selective, and a filter bandwidth of 1kHz was chosen because it has been shown that congestion measurements are approximately independent for frequency separations greater than 1kHz (1). The observation period is one second, to accommodate signal variations due to modulation and fading.

This measurement of congestion across the whole HF spectrum takes several hours, and therefore only results corresponding to the stable ionospheric conditions that occur at about midday and midnight are taken. Even so, the measurement time exceeds the diurnal period of stability, and results for the whole spectrum are taken over three days and three nights, determining congestion across one third of the HF spectrum at each session over a period of about about three hours. It had been verified previously that very good correlation of such congestion measurements exists for results separated by several days, as long as the comparison is made for equivalent times of day or night (2).

Table 1 gives examples of congestion results for five threshold levels (as measured at the aerial terminals of the receiver), where the relationship between threshold and field strength is indicated. For the night results, it is apparent that the lowest threshold (-117dBm) intercepted the noise level at the lower part of the HF spectrum. Other results have been published previously (2).

Congestion will be strongly dependent on the bandwidth of the measurement filter, and

examples of this dependence have also been published (2).

The geographical area over which congestion measurements will be valid is not clear. Simultaneous measurements across the entire HF spectrum have been made at two sites in the UK, separated by 120km, with virtually identical values of congestion being achieved (2). Previous voice channel experiments have indicated that correlation of occupancy may extend to 500km, and theoretical models have supported this (3).

Also, congestion will be dependent on the type of antenna. Limited experiments have been made using four travelling wave Vee antennas, pointing north, south, east and west, from central England, and an effort is being made to compare these results with those determined using the active vertical monopole.

DIURNAL OCCUPANCY VARIATIONS

Since the measurement of congestion throughout the entire HF spectrum takes several hours, only results corresponding to the stable ionospheric conditions could be taken. It is of interest to know how congestion varies throughout the day, and therefore smaller parts of the spectrum were analysed more frequently.

For example, there are 31 allocations where fixed, or fixed and mobile operators are the primary users. The congestion was measured for each of these allocation, approximately every 50 minutes, over a 24 hour period, using a 1kHz bandwidth. Thus, for each allocation, a graph of congestion against time may be plotted for a particular threshold level, as shown in figure 1. Here the threshold level is -107dBm referred to the receiver aerial terminals, corresponding to a received field strength of 2uV/m. The 31 graphs form sections of a 3 dimensional surface, whose axes are congestion, frequency and time, and the surface shown has been smoothed by averaging over 3 adjacent points in time, and over 3 allocations in frequency.

Figure 2 shows the contour map for the surface of figure 1. The contours are shown as continuous curves, although they only have significance across the 31 fixed frequency allocations. Again, good correlation has been observed between sets of results taken for the same frequency allocations, but on different days.

Contour maps of congestion can be used in a prediction procedure to enable frequency planning, to ensure not only that a satisfactory grade of service on an interference free channel is expected, but also that the band in which there is the greatest probability of finding a clear channel is selected. An example of this procedure has been published, in which Barclay (4), using essentially the prediction techniques described in CCIR Report 894, determined the range of permissible frequencies that would provide an adequate signal-to-noise ratio for a specific 1500km east-west sky wave link, terminating at the occupancy measurement site. The range of frequencies which would provide a predicted adequate signal-to-noise ratio is shown in figure 3, in which the S/No=66dB curves correspond to a low-quality SSB circuit with a 100w PEP transmitter, and a transmitting antenna with a gain of 6dB above isotropic (4).

Overlaying this envelope on the occupancy statistics of figure 2 shows that, in this case, operation during the day should be at about 8MHz. This is well below the OWF but at a frequency where there is sufficient signal-to-noise ratio, and a small probability of occupancy. At night, the probability of occupancy in the acceptable bands is more broadly spread in frequency, and the choice is not so clearly defined. However, choice of frequency in this way implies that an operator must have a number of assignments across a wide frequency range.

A MATHEMATICAL MODEL FOR OCCUPANCY

An effort has been made to fit a mathematical model to the occupancy data. The model used was suggested by Dr P J Laycock, of the Department of Mathematics at UMIST, who has guided its application.

The available data comprises day and night measurements made at the times of the winter and summer solstices. For each set of measurements, the spectrum is divided into 95 frequency allocations, and congestion results are obtained for 5 threshold levels. Hence the total number of recorded values of congestion equals 9500 (corresponding to about 3 million channel observations) for a measurement bandwidth of 1kHz, with further limited measurements for different bandwidths. The Belgian sunspot number varied from 130 to 2 during this period.

Ideally we seek to represent the experimental data by a mathematical model, in which values of congestion Q, are given by

$$Q = f(x_1, x_2 \ldots)$$

where $x_1, x_2 \ldots$ represent parameters on which occupancy may be expected to depend, such as time, frequency, bandwidth, threshold level, sunspot number, and geographical location.

Consider the choice of a linear function of the variables.

i.e. $Q = \sum_n A_n X_n$

where X_n are functions of the parameters, and A_n are coefficients. Because of the linear structure, a 'best fit' to the data could be found by least squares regression.

It is apparent that the measured values of congestion are values of probability, and that their values must lie in the range 0 to 1. Since a significant proportion of the observed values of congestion lies on or close to these boundary values, there is a risk that a linear model as defined above may give estimated values of congestion which lie outside this range, and a logistic transformation (logit) has been used to overcome this, where

$$\text{logit } (Q) = \ln \left[\frac{Q}{1-Q} \right] = \sum_n A_n X_n = y$$

Hence $Q = \dfrac{e^y}{1+e^y}$ where $0<Q<1$

Thus the linear function of the variables has been included as the index of an exponent. There are other transformations which would restrict the range of Q between 0 and 1, such as the probit (5), and similar results have been obtained using both the logit and probit transformations.

Extensive statistical analysis has resulted in the use of the following linear function.

$$y = A_k + B \times \text{threshold(dBm)} + (C_0 + C_1 f + C_2 f^2) \times \text{sunspot no.} \qquad (1)$$

where A_k has 95 values, corresponding to the 95 frequency allocations.

B is a single coefficient to be multiplied by the threshold.

C_0, C_1, C_2 are coefficients in a quadratic expression for frequency, to be multiplied by the sunspot number.

There is no useful data for geographical location at present, except that the results apply to an area of at least 100km radius. Also, bandwidth has not yet been included in the model, but is presently being considered. Hence the present model applies for a bandwidth of 1kHz only.

Also the congestion results to be fitted were divided into 4 sets, corresponding to summer solstice day, summer solstice night, winter solstice day, and winter solstice night, such that 4 models were to be determined.

The model is fitted to the congestion results by iterative weighted least squares regression (5). The precise fitting procedures are complex, and are incorporated in the 'generalised linear model' section of the general statistical programme GENSTAT, developed by Nelder and associates, of the Statistics Department, Rothamsted Experimental Station, Harpenden, UK. Surprisingly good fits have been achieved, but the fit for the lowest threshold -117dBm is less accurate than for the higher thresholds, -107dBm to -77dBm. This may be due to the atmospheric noise occasionally rising above this threshold, and also because the sensitivity of the fit is greatest for the typical values of congestion encountered at this level.

Table 2 gives model coefficient values, for summer day congestion results. This model applies for the entire HF spectrum, and for threshold levels in the range -107dBm to -77dBm, corresponding to 1900 measured values of congestion. A separate model exists for the -117dBm threshold.

As a measure of the accuracy of the fit of the model, of the 1900 measured values of congestion, 56% are given by the model to an accuracy of ±.01, 91% to an accuracy of ±.05, and 98% to an accuracy of ±.1.

An operator at HF is likely to have little interest in the occupancy of the spectrum outside his particular type of user allocation. Hence, models have been determined for specific types of user, and have a slightly improved fit to the measured values corresponding to those particular allocations.

For example, from table 1, the fixed user has 48 allocations in which operation is permitted. This differs from the 31 allocations referred to previously because the listings of table 1 split wide allocations into smaller allocations, and also the 31 allocations does not include those shared by fixed and broadcasters. A model of the type defined in equation 1 would thus have 48 frequency allocation terms, and threshold and sunspot coefficients.

Simplification of the fixed user model, with some reduction in accuracy of fit, results when the 48 frequency allocation terms are replaced by a cubic function of frequency. This approximation is in keeping with the concept of a smooth surface for the single user, indicated in figure 1. A similar approximation to the 95 frequency allocation terms for the entire spectrum would not be easily possible, because of the complicated variation in the 95 band user coefficients, due to the inclusion of all users.

For example, for the 48 allocations where fixed operation is permitted, for summer day, and for threshold levels from -107dBm to -77dBm,

$$y = A_0+A_1f+A_2f^2+A_3f^3 + B \times \text{threshold(dBm)} + (C_0+C_1f+C_2f^2) \times \text{sunspot no.}$$

Only 8 coefficients are required, and are given by

$A_0 = -16.2075489$
$A_1 = 0.9156247$
$A_2 = -0.0618355$
$A_3 = 0.0010994$
$B = -0.1063141$
$C_0 = -25.7969626 \times 10^{-3}$
$C_1 = 3.5992532 \times 10^{-3}$
$C_2 = -0.1139105 \times 10^{-3}$

Of the corresponding 960 measured values of congestion, 54% are given by this model to an accuracy of $\pm.01$, 88% to an accuracy of $\pm.05$, and 96% to an accuracy of $\pm.1$.

CONCLUSIONS

Measurement of spectral occupancy in terms of the congestion of each of 95 ITU defined user frequency allocations has been achieved at the times of the winter and summer solstices, at a site in central England, having a low man-made noise level. The measurements have been made for a range of sunspot numbers from 130 to 2.

Analysis has shown that the results are highly structured statistically, and mathematical models have been derived, and continue to be developed.

During 1988, it is intended to install a dedicated equipment for the measurement of HF spectral occupancy, at the Department of Trade and Industry Monitoring Station at Baldock, England. The equipment will be linked to UMIST for data transfer and analysis, and this will result in a much improved data base.

ACKNOWLEDGEMENTS

The authors thank the UK Science and Engineering Research Council and the Procurement Executive, Ministry of Defence for supporting this work. In particular the assistance and suggestions given by J R Guest and C R Harding (RSRE), and L W Barclay (DTI) are acknowledged.

REFERENCES

1 Gott G F, Dutta S and Doany P; 'Analysis of HF interference with application to digital communications', Proc IEE, part F, vol 130, no 5, 1983.

2 Gott G F, Wong N F and Dutta S; 'Occupancy measurements across the entire HF spectrum' AGARD Conference 'Propagation Aspects of Frequency Sharing, Interference, and System Diversity', Paris, 1982.

3 Dutta S and Gott G F; 'Correlation of HF interference spectra with range', Proc IEE, part F, vol 128, no 4, 1981.

4 Gott G F, Wong N F and Barclay L W; 'Spectral occupancy and frequency planning', Proc IEE, part F, vol 132, no 7, 1985.

5 McCullagh P and Nelder J A; 'Generalised linear models', Chapman and Hall, 1983.

6 Harding C R, Guest J R and Gott G F; 'HF spectral occupancy measurements and their value to frequency planners and system designers', 5th Symposium of the Allied Radio Frequency Conference, NATO Headquarters, Brussels, 1986.

	FREQUENCY RANGE (MHz)	USER	DAY CONGESTIONS (%)					NIGHT CONGESTIONS (%)				
			-117dBm	-107dBm	-97dBm	-87dBm	-77dBm	-117dBm	-107dBm	-97dBm	-87dBm	-77dBm
1	1.606 - 1.810	FIXED/MOBILE	17	3	0	0	0	100	80	17	0	0
2	1.810 - 1.850	AMATEUR	12	2	0	0	0	100	85	15	0	0
3	1.850 - 2.045	FIXED/MOBILE	5	3	1	0	0	100	75	16	4	0
4	2.045 - 2.300	FIXED/MOBILE	7	1	0	0	0	99	78	21	3	1
5	2.300 - 2.500	FXD./MOB./BCST.	5	1	0	0	0	100	90	26	2	1
6	2.500 - 2.850	FIXED/MOBILE	10	3	1	0	0	100	94	52	11	5
7	2.850 - 3.155	AEROMOBILE	4	1	0	0	0	100	91	36	6	2
8	3.155 - 3.200	FIXED/MOBILE	9	2	0	0	0	100	87	57	17	2
9	3.200 - 3.400	FXD./MOB./BCST.	5	0	0	0	0	100	97	60	30	2
10	3.400 - 3.500	AEROMOBILE	3	2	2	2	0	100	92	46	13	6
11	3.500 - 3.800	FXD./MOB./AMTR.	16	6	1	0	0	100	99	77	34	5
12	3.800 - 3.900	FIXED/MOBILE	17	7	0	0	0	100	97	71	36	15
13	3.900 - 3.950	AEROMOBILE	10	2	0	0	0	100	96	61	18	4
14	3.950 - 4.000	FIXED/BROADCAST	2	0	0	0	0	100	100	94	76	53
15	4.000 - 4.438	MARITIME MOBILE	15	7	3	1	0	100	97	63	32	16
16	4.438 - 4.650	FIXED/MOBILE	18	10	4	1	0	100	96	67	42	17
17	4.650 - 4.750	AEROMOBILE	15	5	2	0	0	100	90	29	11	2
18	4.750 - 5.000	FXD./MOB./BCST.	14	7	1	0	0	100	95	64	35	13
19	5.000 - 5.480	FIXED/MOBILE	22	10	4	2	0	100	96	61	38	19
20	5.480 - 5.730	AEROMOBILE	8	3	1	0	0	100	89	35	10	3
21	5.730 - 5.950	FIXED/MOBILE	19	10	8	1	0	100	95	67	43	21
22	5.950 - 6.200	BROADCAST	63	41	23	9	2	100	100	93	75	51
23	6.200 - 6.525	MARITIME MOBILE	30	21	13	6	2	100	97	62	35	19
24	6.525 - 6.765	AEROMOBILE	19	5	1	0	0	100	93	46	10	4
25	6.765 - 7.000	FIXED/MOBILE	28	15	6	1	0	100	93	47	26	12
26	7.000 - 7.100	AMATEUR	47	21	8	1	0	100	93	53	33	6
27	7.100 - 7.300	BROADCAST	43	27	10	3	1	100	97	74	51	31
28	7.300 - 7.800	FIXED/MOBILE	20	10	3	1	0	100	95	58	28	15
29	7.800 - 8.195	FIXED/MOBILE	31	18	8	4	0	100	90	57	28	13
30	8.195 - 8.500	MARITIME MOBILE	24	11	5	1	1	100	92	52	29	13
31	8.500 - 8.815	MARITIME MOBILE	47	26	14	7	3	100	97	65	41	20
32	8.815 - 9.040	AEROMOBILE	20	5	0	0	0	100	84	33	14	8
33	9.040 - 9.500	FIXED	39	18	9	3	1	100	84	45	26	10
34	9.500 - 9.900	BROADCAST	67	43	24	11	4	100	99	92	79	49
35	9.900 - 10.000	FIXED	16	8	5	0	0	100	73	42	20	10
36	10.000 - 10.100	AEROMOBILE	53	19	8	4	1	100	80	33	16	10
37	10.100 - 10.150	FIXED/AMATEUR	80	45	25	10	6	100	94	65	47	24
38	10.150 - 10.600	FIXED/MOBILE	70	33	17	10	4	100	91	49	24	11
39	10.600 - 11.175	FIXED/MOBILE	67	27	12	6	3	100	83	33	19	7
40	11.175 - 11.400	AEROMOBILE	50	12	5	3	1	100	42	12	4	1
41	11.400 - 11.650	FIXED	70	31	16	8	2	100	80	53	25	15
42	11.650 - 12.050	BROADCAST	90	64	47	28	14	100	99	95	80	56
43	12.050 - 12.230	FIXED	82	50	24	11	5	100	88	60	35	16
44	12.230 - 12.730	MARITIME MOBILE	57	27	14	6	1	96	49	25	11	4
45	12.730 - 13.200	MARITIME MOBILE	68	34	20	9	3	95	52	31	15	6
46	13.200 - 13.360	AEROMOBILE	47	14	2	1	0	81	25	11	4	1
47	13.360 - 13.600	FIXED/MOBILE	66	37	22	11	4	89	37	20	11	6

Table 1a Congestion values for summer 1982

(-107dBm corresponds to a received field strength of 2uV/m)

	FREQUENCY RANGE (MHz)	USER	DAY CONGESTIONS (%) -117dBm	-107dBm	-97dBm	-87dBm	-77dBm	NIGHT CONGESTIONS (%) -117dBm	-107dBm	-97dBm	-87dBm	-77dBm
48	13.600 - 13.800	BROADCAST	66	32	21	9	1	77	38	17	9	1
49	13.800 - 14.000	FIXED/MOBILE	60	30	17	6	2	72	15	6	2	1
50	14.000 - 14.350	AMATEUR	55	28	7	2	0	89	46	22	7	1
51	14.350 - 15.000	FIXED/MOBILE	49	23	11	4	1	73	29	14	5	1
52	15.000 - 15.100	AEROMOBILE	57	31	10	4	1	97	63	32	7	5
53	15.100 - 15.600	BROADCAST	96	80	55	33	16	96	78	52	33	16
54	15.600 - 16.000	FIXED	54	32	19	9	2	31	17	9	4	1
55	16.000 - 16.360	FIXED	55	34	18	8	1	26	14	8	2	1
56	16.360 - 16.860	MARITIME MOBILE	28	11	3	1	0	13	7	2	1	0
57	16.860 - 17.410	MARITIME MOBILE	47	21	10	3	0	24	12	4	1	0
58	17.410 - 17.550	FIXED	44	21	8	5	1	3	1	0	0	0
59	17.550 - 17.900	BROADCAST	88	70	46	28	13	53	32	17	7	2
60	17.900 - 18.030	AEROMOBILE	25	9	4	2	1	12	4	2	0	0
61	18.030 - 18.068	FIXED	28	10	5	5	0	3	0	0	0	0
62	18.068 - 18.168	AMATEUR	28	16	7	0	0	9	4	0	0	0
63	18.168 - 18.780	FIXED	34	18	8	4	1	4	2	0	0	0
64	18.780 - 18.900	MARITIME MOBILE	37	23	11	2	1	3	0	0	0	0
65	18.900 - 19.300	FIXED	18	7	1	0	0	9	3	1	0	0
66	19.300 - 19.680	FIXED	23	8	2	1	0	5	1	1	1	1
67	19.680 - 19.800	MARITIME MOBILE	16	7	2	0	0	12	6	1	0	0
68	19.800 - 20.000	FIXED	21	7	3	2	1	7	4	1	0	0
69	20.000 - 20.500	FIXED/MOBILE	13	6	3	1	0	24	10	5	1	0
70	20.500 - 21.000	FIXED/MOBILE	10	4	1	0	0	2	1	0	0	0
71	21.000 - 21.450	AMATEUR	14	2	1	0	0	5	1	0	0	0
72	21.450 - 21.870	BROADCAST	71	48	29	16	7	19	9	4	1	0
73	21.870 - 22.000	AEROMOBILE	11	2	2	0	0	2	0	0	0	0
74	22.000 - 22.400	MARITIME MOBILE	17	7	1	0	0	1	0	0	0	0
75	22.400 - 22.855	MARITIME MOBILE	15	6	2	0	0	2	0	0	0	0
76	22.855 - 23.000	FIXED	9	4	1	1	0	6	1	0	0	0
77	23.000 - 23.200	FIXED/MOBILE	5	2	0	0	0	0	0	0	0	0
78	23.200 - 23.350	AEROMOBILE	1	1	0	0	0	0	0	0	0	0
79	23.350 - 24.000	FIXED/MOBILE	2	1	0	0	0	4	1	0	0	0
80	24.000 - 24.500	FIXED/MOBILE	1	0	0	0	0	0	0	0	0	0
81	24.500 - 24.890	FIXED/MOBILE	2	1	0	0	0	1	0	0	0	0
82	24.890 - 25.000	AMATEUR	0	0	0	0	0	0	0	0	0	0
83	25.000 - 25.210	MARITIME MOBILE	2	0	0	0	0	0	0	0	0	0
84	25.210 - 25.550	FIXED/MOBILE	0	0	0	0	0	0	0	0	0	0
85	25.550 - 25.670	RADIO ASTRONOMY	2	1	0	0	0	0	0	0	0	0
86	25.670 - 26.100	BROADCAST	3	2	0	0	0	0	0	0	0	0
87	26.100 - 26.175	MARITIME MOBILE	3	0	0	0	0	0	0	0	0	0
88	26.175 - 26.500	FIXED/MOBILE	3	1	0	0	0	0	0	0	0	0
89	26.500 - 27.000	FIXED/MOBILE	9	2	0	0	0	1	0	0	0	0
90	27.000 - 27.500	FIXED/MOBILE	22	6	0	0	0	12	4	1	0	0
91	27.500 - 28.000	FXD./MOB./METH.	19	3	0	0	0	36	13	5	2	1
92	28.000 - 28.500	AMATEUR	0	0	0	0	0	0	0	0	0	0
93	28.500 - 29.000	AMATEUR	0	0	0	0	0	0	0	0	0	0
94	29.000 - 29.700	AMATEUR	1	0	0	0	0	0	0	0	0	0
95	29.700 - 30.000	FIXED/MOBILE	0	0	0	0	0	0	0	0	0	0

Table 1b Congestion values for summer 1982

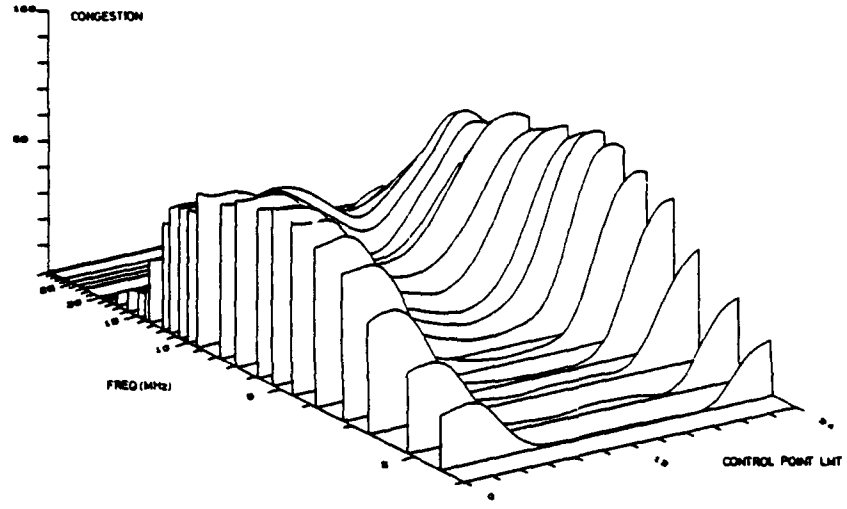


Figure 1 Congestion/frequency/time surfaces for fixed and mobile users (Summer 1982, threshold level -107dBm)

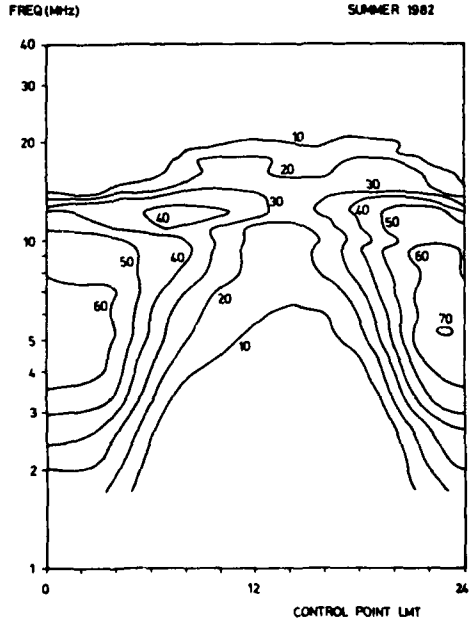


Figure 2 Contour map

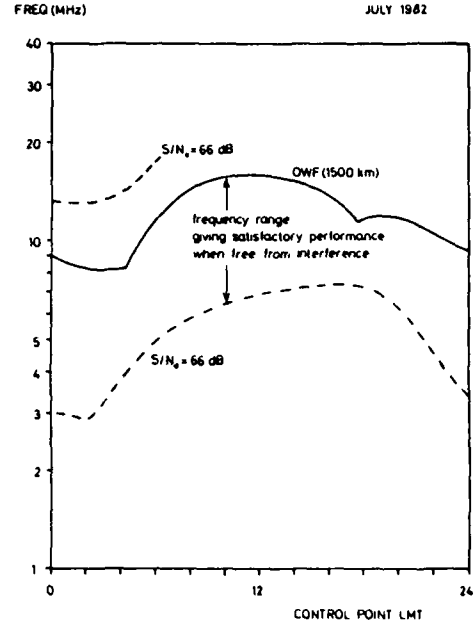


Figure 3 Frequency range

Congestion $Q = \frac{e^y}{1+e^y}$

$y = A_k + B \times \text{threshold level (dBm)} + (C_0+C_1f+C_2f^2) \times \text{sunspot number}$

k	A_k	k	A_k	k	A
1	-13.585664	33	-11.919495	65	-13.546634
2	-14.733032	34	-10.289421	66	-13.993226
3	-14.062625	35	-12.711416	67	-14.208674
4	-14.064345	36	-12.219176	68	-13.647821
5	-13.915349	37	-11.225493	69	-13.299901
6	-13.418279	38	-11.246493	70	-13.958850
7	-15.021687	39	-11.635492	71	-14.283216
8	-12.792868	40	-12.730354	72	-11.048516
9	-13.166584	41	-11.378533	73	-14.411120
10	-13.210117	42	- 9.410399	74	-13.973111
11	-12.712065	43	-10.779525	75	-13.829214
12	-13.263342	44	-11.663792	76	-14.733973
13	-13.962915	45	-10.963576	77	-15.514025
14	-13.965851	46	-12.744649	78	-15.221752
15	-12.034389	47	-11.050034	79	-15.906229
16	-11.952289	48	-10.546442	80	-16.478808
17	-12.839164	49	-11.426044	81	-16.213741
18	-12.507083	50	-11.958787	82	-20.534570
19	-12.261500	51	-11.729799	83	-15.904058
20	-13.413974	52	-11.292570	84	-21.632969
21	-11.921589	53	- 9.223130	85	-13.496772
22	- 9.867710	54	-11.633442	86	-15.292201
23	-11.288726	55	-11.693433	87	-15.529101
24	-13.433738	56	-12.808930	88	-15.102002
25	-11.865379	57	-11.698942	89	-14.065095
26	-11.570784	58	-12.420106	90	-13.117673
27	-11.026025	59	- 9.926302	91	-13.752702
28	-12.219094	60	-12.650106	92	-15.866187
29	-11.804916	61	-12.467779	93	-15.144994
30	-11.959665	62	-12.393797	94	-17.197009
31	-10.828862	63	-12.677131	95	-16.499500
32	-13.262643	64	-13.309509		

B = 0.100192

f = centre frequency of each allocation (MHz)

C0 = $-17.348957 \times 10^{-3}$
C1 = 1.910600×10^{-3}
C2 = -0.052359×10^{-3}

Table 2 All-users congestion model for summer day and for threshold range -107dBm to -77dBm

NOISE MEASUREMENT SYSTEMS/TECHNIQUES

Calvin B. Christianson
U.S. Army Information Systems
Engineering and Integrations Center
Fort Huachuca, Arizona 85163-7300

ABSTRACT

Almost from the start of radio science, it was understood that before radio systems could be adequately designed and performance statistically defined, it would be necessary to have statistical data and other characteristics of the signals and noise. Measurement systems that could measure signal levels and related characteristics fairly accurately were developed in the early 1920's. But the problem of measuring noise was considerably more of a puzzling matter to early investigators. As a result most work in noise investigation for many years was in comparing and determining the performance of various noise reducing techniques or the determination/correction of interference to radio listeners. It wasn't until the 1950's that measurement systems were developed that provided needed statistical noise data. Primarily this was the large fixed systems developed for measuring atmospheric noise during the International Geophysical Year (IGY). Since the early '70's a number of portable or mobile systems have become available that are limited in capabilities almost only by the amount of time or money available.

This paper will briefly trace the evolution of noise measurement systems, types of and capabilities of systems that have been used in the last decade, and a small portable system that was developed for noise measurements. The use of this small system for noise measurements in the field and resulting data will be presented.

EARLY NOISE MEASUREMENT SYSTEMS

In 1916 Austin(1), in describing some measurements made at a U.S. Naval Radio Station in Panama, provided a detailed description of the receiver used (Figure 1). The sensitivity of this early simple set was rather astonishing. Measurements down to 1.23x10^-15 Watts (.25uV at 50 Ohms) were possible in the Naval Radio Research Laboratory. In use an operator would adjust a shunt resistor across the headphone until the signal could no longer be copied. An audibility level would then be computed based upon the impedance of the headphones and the value of the shunt resistor. This measurement technique was initially developed to measure signal levels but starting in early 1918 it was also used for disturbance(static) measurements. For noise measurements, the shunt resistor was varied until the disturbance could be heard 3 times in 10 seconds. For actual measurements, Austin was of the opinion that measurements were usually accurate within 30% and occasionally 50%. He stated that this wasn't too good, but it was better than no measurements at all! Several years later, Austin(2) reported on the results of field strength measurements from 1915 to 1921 and noise measurements from 1918 to 1921 using the audibility meter. This report showed that when signal intensities were high, noise levels were also high. Although Austin provided the most information on the audibility method, it appears that this method was used even earlier. For example, Marriot(3) stated in 1914 that while making field strength measurements, audibility measurements of 20 dropped to 7 when a train passed by.

In discussions on Austin's work there was considerable debate on the merits of this system as it was subjective and dependent on an operator's ability or judgement. Suggestions were also made on other methods that could be used or were being used. Eccles(4) described a system with a local signal source (a buzzer) with adjustable coupling to the input of the receiver. Austin, in discussion of this method indicated that he had tested this method and similiar methods in his laboratory. But, due to static, these methods were not as good as the audibility method for actual measurements. Starting in 1922, Austin(5) changed the method of making measurements and used a tuning fork oscillator as a local reference. For signal measurements the receiver beat note was adjusted to the same frequency as the tuning fork. The local tone level was then adjusted until it equalled the received signal level. For disturbance measurements the local level was adjusted until the disturbances were just audible. For Austin's measurements, the overall measurement system, antenna to output, was frequently calibrated using received signals of known intensity.

In 1923, Bown, Englund and Friis(6) described apparatus and methods that they had developed since 1921 to measure the absolute value of field strengths and the effective value of radio noise. Three measurement sets were described. The first was a tuned radio frequency receiver for measurements at approximately 23.5 kHz. The other sets were both superheterodynes. One was designed for 57 kHz. The other, for shortwave measurements, was for 650 kHz to 950 kHz. All these sets had a local oscillator with a calibrated adjustable output. When static permitted a meter in the output could be used for comparison. When the meter could not be used, comparison was made again aurally. Practically, the block diagrams for the receivers are similiar to typical field intensity meters in present use.

Compared to the measurement of field strength, the use of these sets for measurements of noise was considered to be much more difficult. As measurements were being made in connection with the engineering of a low frequency transoceanic single sideband telephone system, the effect of noise on voice communication was of primary interest. The method adopted was to generate artificial speech at the receiver. The level of this was then adjusted until it was just audible through the noise. Using this method, a variation of less than 50% occurred for different observers. Results of some tests made in London on a

transatlantic path using the 57 kHz set were provided. Signal level, noise level and word list tests had been made. The results were preliminary as only a few tests had been made. However it appeared that a signal/noise voltage ratio of about 30 (=30 dB) was a transition level for voice transmission. Below 30, voice transmission rapidly degraded. Additional comprehensive measurements using this system were reported(7) in 1926 for signal and noise strengths at LF. A significant improvement in the signal/noise ratio had resulted using a Beverage antenna.

Friis and Bruce(8) in 1926 described a measurement set for frequencies up to 40 MHz. Some measurements had also been made of radio noise at higher frequencies by Heising, Schelleng, and Southworth(9). In 1931, Espenschied(10) outlined the various methods developed to measure noise. For radio, the artificial speech method was still being used for low frequencies. At higher frequencies, an observer would record the amplitude of peaks over a period of one minute. The average of the ten highest peaks was then taken as a noise value. Potter(11), in 1931, presented a comprehensive report on HF atmospheric noise using the measurement techniques just described. Noise levels versus time, direction, frequency, local storms, sunrise and sunset effects, solar eclipse, and a disturbance in the earth's magnetic field had all been measured.

In 1932, an automatic field strength and static measuring system was described by Mutch(12). A square law detector was used in this system. The detector only had a limited dynamic range so, very cleverly, rms voltage received over 9 seconds was averaged and system gain then automatically changed to stay in the square law region. Using this system, Jansky(13) made noise measurements at 14.6 MHz with a highly directional antenna in 1932. During these measurements a low level noise was detected which appeared to be originating from either the sun or the subsolar point on the earth. Continued investigation by Jansky(14) of this low level noise found that this noise was not sun related but was originating from a fixed point in space. Similiar measurements by Jansky(15) were repeated at 16.7 MHz several years later. Serious man-made noise, caused by diathermy machines, interfered with the measurements. It was so severe that at times hours were spent trying to find a frequency near 16.7 MHz on which interstellar or atmospheric noise could be measured.

In this same period, Peterson(16) reported a method by which the percentage of time that noise exceeded a reference level could be measured. The method used was essentially a constant current source triggered when the receiver input voltage exceeded a given level. An output meter with a very long time constant would then indicate the percentage of time the reference level was exceeded. A single amplitude probability distribution (APD) graph showed a typical measurement of noise using this method.

In the early '30's work was also started in the standardization of radio noise meters for measurement of noise interfering with broadcast services. The evolution of these types of instruments has been continuous and has led to the many CISPR type instruments presently available and frequently specified for interference measurements. Excellent reports on the early development, use, and accuracy of these instruments were provided in references (17), (18), (19) and (20). The problems of using this type of instrument to adequately statistically define noise were considered by Hoff and Johnson(21) in 1952. They also described a measurement system and results achieved in measuring LF atmospheric noise. Essentially, the system was the same as Peterson used in 1935 except reference levels were automatically changed and the percentage of time automatically graphically recorded for each level.

MAJOR ADVANCEMENTS THROUGH THE 1960's

A long needed major advancement occurred in the late '50's with the development of the ARN-2 atmospheric noise measurement system as described by Crichlow(22). This system measured the parameters specified by the International Union of Radio Science (URSI)(23) who had coordinated with the CCIR. Simultaneously on 8 frequencies, the rms voltage, average voltage, and the average logarithm of envelope voltage were measured. Data from these systems taken from 1957 to 1961 resulted in CCIR Report 322(24). The method developed for measuring rms envelope voltage in this system is very interesting. To square the input voltage, two tubes with a parabolic transfer characteristic were used. Drive to the tubes was push-pull with the plates in parallel. This essentially cancelled out the fundamental frequency and added the second harmonic at the output. The voltage level of the second harmonic was proportional to the square of input voltage. After filtering, amplification and rectification, an output voltage ranging from .1 V to over 1000 volts (>80 dB dynamic range) resulted. A few years later, another large system was described by Spaulding(25) to investigate man-made radio noise. This system measured the same three moments as the ARN-2 and also the ratio between quasi-peak (1 ms charge/160 ms discharge) and rms voltage. This system could also measure on 8 frequencies simultaneously and was installed in a large van so mobile measurements could be made. The 300 hours of measurements made over the period of 1966 to 1971 resulted in estimates of man-made noise for various areas(25). The results were also summarized in CCIR Report 258(26).

SMALL PORTABLE SYSTEMS OF THE 1970's

For many years there had been a requirement, especially by the military, for a small portable system that could measure noise power during surveys. Typical questions frequently asked were: 'Is site A better than site B?' and 'Is the noise getting worse or better?'. Using instruments that could only measure average, quasipeak and peak voltages were hopelessly inadequate. For example: In site surveys maybe site A was 3 dB lower than site B for an average voltage measurement. But site A quasipeak measurements were 5 dB higher than those of site B! A collective sigh of relief was heard when Mattheson and Beasley(24) modified the small, very portable NM-25T field intensity meter to measure rms

voltage and Vd (ratio in dB between rms noise voltage to average noise voltage). Subsequently, the manufacturer of NM-25T developed the NM-26T which provided the same features.

For these sets, Hagn(28) wrote a detailed handbook to assure standarized methods to acquire noise data and for making noise surveys. Typically, to measure noise with these instruments, 11 measurements 15 seconds apart were taken of Vrms and Vd on a signal free frequency. From each measurement series, a median value, upper decile, and lower decile were used in data reduction. Engles and Hagn (29) and Lauber(30) showed in detail how this instrument could be used for field measurements that previously were nearly impossible. As a noise measuring system is no better than its antenna, Hagn(31) also provided essential information on the use and calibration of short whips for use with the NM-26T and similiar systems. The NM-26T and the modified NM-25T has been used for many surveys by U.S. Government agencies for site selection and to investigate noise at existing sites. Most of these surveys were for internal use. However, many reports showed the man-made noise at a receiving site to be several dB higher than a quieter location a few kilometers away. Frequently, noise sources were also located during these surveys. When corrected, the overall site noise level usually decreased several dB.

SMALL COMPUTERIZED SYSTEM FOR THE 1980's

In the early '80's manufacture of the N-26T was discontinued due to difficulties in procuring components. Maintenence was also difficult due to the lack of spares. It thus became essential to develop a new system for field noise measurements. High on the list of desirables was a system that provided data comparable to that taken with the NM-26T. At that time small home computers became available that met FCC Class B requirements. Testing of the Commodore VIC20 computer found that it was relatively interference free, could be modified to operate off a `2 volt battery and was easily interfaced to external devices. The cassette system for the VIC20 was also found to be very reliable. The low power ADC0820 analog-to-digital half flash convertor could be easily interfaced with the VIC20 and to the log video output of available receivers. The measurement system that resulted is shown in figure 2. Calibration of the system was done using a noise generator. This eliminated any non-linearities in the receiver log video detector which had about a 75 dB dynamic range. In initial testing low-level noise from the computer could be aurally heard at the receiver output. The RF filters shown in figure 2 eliminated most of this. System noise still infrequently occurred above 15 MHz but could be easily aurally identified and eliminated by a slight change in frequency.

Two types of software were developed. For both, a short machine code was used for acquiring and storing counts versus level. To minimize memory space required for data storage only 15 levels were used. The 4 bit output from the analog digital convertor was used to specify a memory area to increment each time a reference level was exceeded. One million samples required approximately 22 seconds. The first software system was designed to maximize the number of measurements that could be made in a period of time. Thus, it only acquired data and stored data to tape for later analysis. The second software system after acquiring data, immediately computed Vrms, Vavg and Vd using appropriate correction factors. All data could then be stored and a another measurement made. Both systems, prior to making a measurement, would take a 10 second sample and graphically display an approximate APD on the screen. This would not show higher noise levels occurring infrequently, but it provided valuable information for the operator to set attenuators so that the most useful dynamic range was achieved. If attenuators were changed this sample could be repeated. A measurement would then be made. Normally 6 million samples (132 sec) were taken as a measurement. This required approximately the same time as used for making measurements with the NM-26T (150 sec for 11 measurements). Some comparison tests were made with an NM-26T. For 60 simultaneous measurements at two frequencies in an area with moderate power line noise, 92% were within 3 dB and 73% were within 1 dB.

The NM-17/27 field intensity meter used for measurements has one desirable feature (besides the essential log video detector and internal calibration). There is a second IF with a beat frequency oscillator. For a noise measurement on a hopefully signal free channel, this provides a very sensitive aural detector of CW and RTTY signals that otherwise could not be detected. For all measurements, this has been used to aurally monitor the frequency during a measurement. If, during a measurement, a signal is heard, the measurement can be immediately stopped. The percentage of the measurement done is then displayed. Options then available are to start a new measurement, continue the measurement or store data taken.

Figures 3 to 6 show typical results achieved. The measurements (1400) were taken at 2 sites on the East Coast of the United States in the spring of 1983. The separation between the two sites was approximately 80 miles (128 kilometers). This was the first time this system was used and for testing purposes, measurements were made in 5 time blocks. Both sites were essentially rural with 4kV power lines within several hundred feet. During most of the measurements, low level power line noise was identified. In one area, the noise generating sources on the power lines were located using an ultra-sonic detector. For all measurements a nine-foot antenna was used at ground level with eight 25' radials. During the surveys, serious lightning storms occurred 100 to 200 miles west of the measurement areas in the late afternoon almost very day. The increase shown on figures 3, 4, and 5 in the afternoon is attributed to these storms. Interestingly there was no increase at 16 MHz or at two higher frequencies (20 and 24 MHz). The atmospheric noise levels shown on figures 3 to 6 are based upon Spaulding and Washburn(32). For time blocks and frequencies (3,6 MHz) when atmospheric noise was substantially greater than man-made noise the mean noise level was close to predicted atmospheric noise levels. When atmospheric noise levels were low, the man-made noise levels were usually close to that estimated in reference(2) for rural areas.

Figure 7 is based on five days of measurements at another military base to select a site for a small HF station. Area A was immediately adjacent to a building with several 4kV power lines nearby. Area B was in the same general area but about 300 feet from the nearest power line. Both A and B were in residential areas but with a building density less then that defined in reference(25). Areas C and D were in a remote area but again with 4kV lines in the general area. Area C was near a vacant building with a power line to it. Area D was 200 feet from the building. The separation was not too effective in reducing noise levels except at 18 and 24 MHz. Area E was again in a remote area but with a buried power line to the site. Except at 3 MHz, this was the quietest area. At all areas again, low level power line noise was identified. All measurements were made between 0730 and 1300 local time.

The measurement antenna was a nine-foot whip mounted on the roof of a Dodge van. The antenna coupler was on an 18 inch by 18 inch aluminum plate with suction cups for attachment to the van roof. Grounding of the plate was accomplished with 4 wires, no longer than necessary, to bumpers at each corner of the van. The increase in noise at 24 MHz compared to 18 MHz for all areas was noted and later investigated. A ground plane of wire mesh 4 feet by 6 feet was laid on top of the van. When the aluminum plate was grounded to this, readings at 24 MHz (near quarter-wavelength resonance) dropped 5 dB. The change at other frequencies was less than 0.5 dB. This shows that a proper antenna installation is critical.

Two systems identical to this were constructed by another government agency and used for measurements at HF receiving facilities. At one important facility near a major city, measured man-made noise levels were found to be 6 dB higher at 3 MHz and 4 dB higher at 30 MHz than estimated for business areas. The linear regression line for all measurements between 3 and 30 MHz was 8 dB higher at 3 MHz and 3 dB higher at 30 MHz than for the estimated business regression line. As measurements had been in all time blocks during the winter season, comparisons were also made with CCIR 322 predictions. For all time blocks and frequencies (3-30 MHz), the man-made noise level was greater than predicted atmospheric noise.

Since the system above was developed, a limited investigation and some software development has been done to use a small laptop computer. This would have the advantage of being smaller, would require less power, could internally store data for over 100 measurements and require fewer interconnections.

CONCLUSION

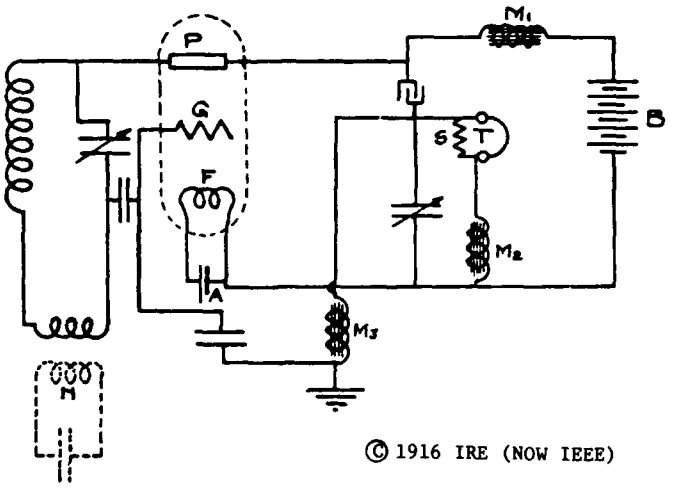
When describing measurements being made in 1923, Bown, Englund and Friis(6) wrote, 'To bring the engineering aspects of the whole subject more clearly into view let us consider the information which would need to be available to enable a a radio engineer to predict with certainty the cost of obtaining a given grade of service from a large and important radio project. This needed information may be itemized as follows:

1. A statement of the grade of service desired.
2. The relation between the cost of various sizes of transmitting apparatus and their power outputs into the antenna.
3. The relation between cost and radiation efficiency for various sizes and types of transmitting antenna.
4. Statistical data on the ether transmission efficiency from the transmitting station site to the receiving station site.
5. Statistical data on the absolute volume of radio noise at the receiving station site.
6. The relation between the cost and noise reducing capabilities of various receiving antenna and apparatus.

Radio engineering will have become pretty well standardized when a project may be estimated with all this information in hand. However, the respects in which the present knowledge fails to meet the list are an indication of the lines along which progress is needed. It is interesting to note that in all but the first two of these items experimental measurements of radio field strengths and of radio noise would be basically useful in getting the desired information in quantitive form.' © 1923 IRE (now IEEE)

The information needed for radio engineering as specified by Bown, Englund and Friis is as valid today as in 1923. Progress since then has been continous and substantial in the areas of receivers, transmitters, antennas, propagation techniques and signal processing. The development of radio noise measuring systems for over seventy years shows the continuing need for noise data. The large international effort to measure atmospheric noise resulted in predictions (recently improved by Spaulding and Washburn(32)) that are widely accepted and used. For man-made noise the measurements used for estimates are now over sixteen years old. Since those measurements were made, noise from silicon controlled rectifiers, computer systems, and other new devices that were not in general use sixteen years ago have all been reported. Their impact on noise in various areas is unknown but may be substantial. For military HF systems, the noise environment is usually unknown. As a result man-made noise estimates for business areas are frequently used for engineering. But, as shown, noise may be several dB higher at some facilities. For military systems, equipment is usually specified to standards controlling the noise generated. But the results of subsequent maintenance is unknown. In one measurement situation, all of the several power generators available caused interference. The noise would have seriously impacted radio communications in the area.

Spaulding's excellent report(33) in 1976 on man-made noise considered in detail the problems just outlined and provided recommendations for needed progress. Accurate noise measurements are necessary and methods to achieve this was discussed. The small portable system described in this paper is essentially the same as the software APD method in Spaulding's report. Athough simple, it has been invaluable for field noise surveys.



© 1916 IRE (NOW IEEE)

"The adjustment for greatest sensitiveness requires special skill on the part of the operator. Quantitative readings taken by untrained men will give considerably lower sensitiveness."

FIGURE 1. EARLY (1916) MEASURING SET

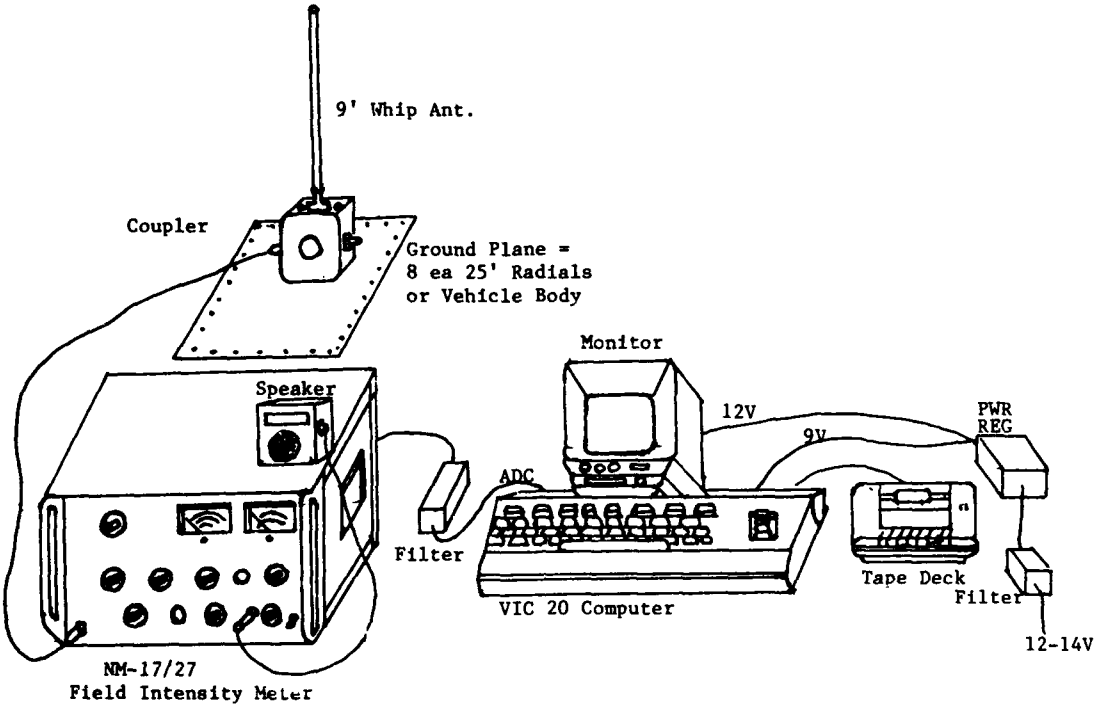


FIGURE 2. COMPUTERIZED SYSTEM TO MEASURE APDS.

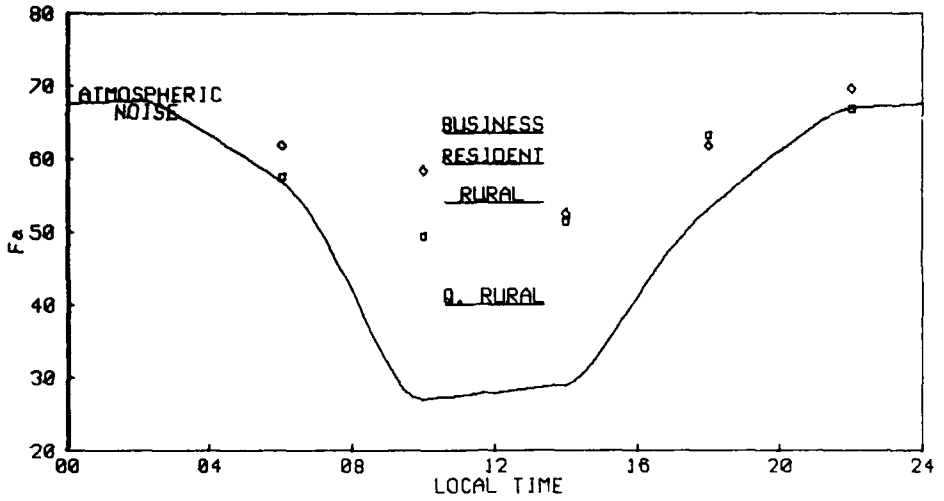


FIGURE 3. MEASURED NOISE LEVELS VERUS ESTIMATED LEVELS AT 3MHz

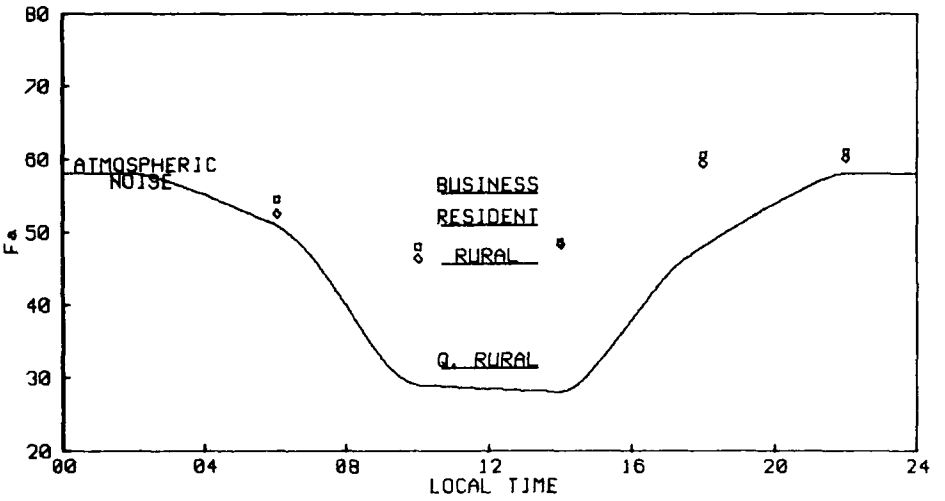


FIGURE 4. MEASURED NOISE LEVELS VERUS ESTIMATED LEVELS AT 6MHz

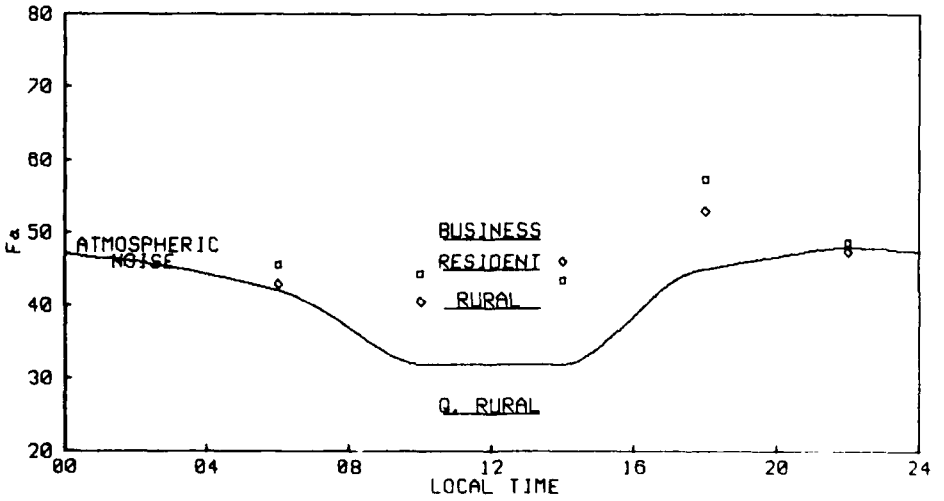


FIGURE 5. MEASURED NOISE LEVELS VERUS ESTIMATED LEVELS AT 10MHz

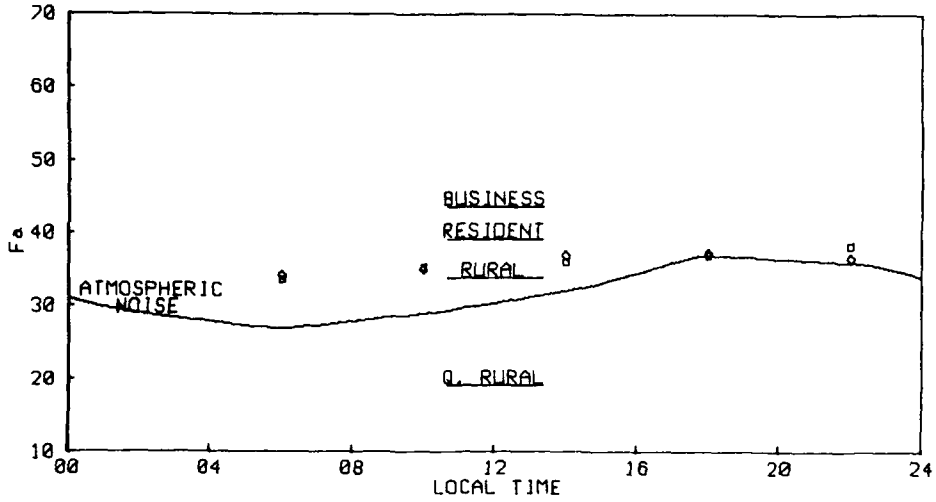


FIGURE 6. MEASURED NOISE LEVELS VERUS ESTIMATED LEVELS AT 16MHz

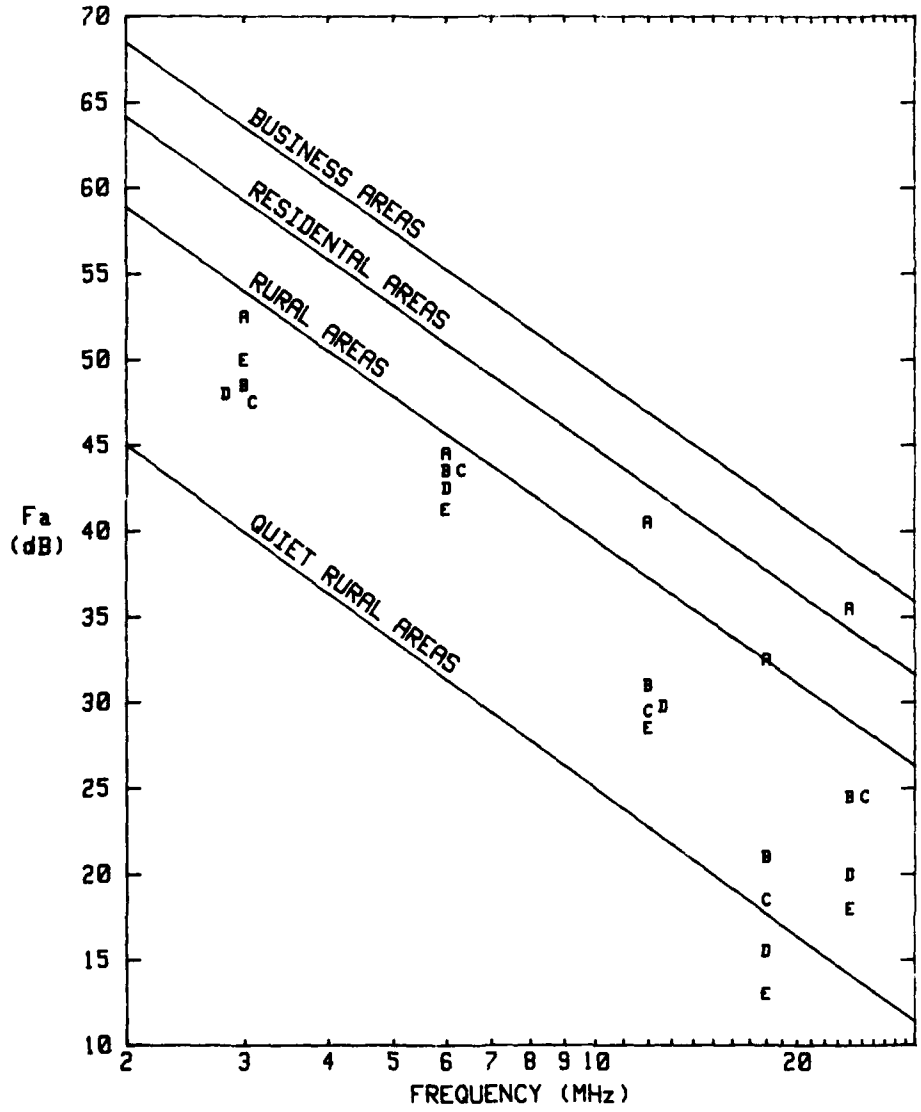


FIGURE 7. MEASURED VERSUS ESTIMATED NOISE LEVELS

REFERENCES

1. Louis W. Austin, "Experiments at the U.S. Naval Radio Station, Darien, Canal Zone," Proc. IRE, Vol. 4, No. 3, June 1916. pp. 251-270.
2. Louis W. Austin, "The Monthly Averages of Signal Strength of Nauen in Washington, 1915-1921 and the Monthly Averages of Atmospheric Disturbances in Washington 1918-1921," Proc. IRE, Vol. 10, No. 3, June 1922. pp. 153-158.
3. Robert H. Marriott, "Radio Range Variation," Proc. IRE, Vol. 2, No.1, Mar. 1914. pp. 37-58.
4. W. H. Eccles, "On Measurement of Signal Strength," Proc. IRE, Vol. 7, No.3, June 1919. pp. 267-278.
5 Louis W. Austin, "Receiving Measurements and Atmospheric Disturbances at the Naval Radio Research Laboratory, March and April 1922," Proc. IRE, Vol. 10, No. 4, Aug. 1922 pp. 239-243.
6. Ralph Bown, Carl R. Englund, and H. T. Friis, "Radio Transmission Measurements," Proc. IRE, Vol. 11, No. 2, April 1923. pp. 115-162.
7. Lloyd Espenschied, C. N. Anderson, and Austin Bailey, "Transatlantic Radio Telephone Transmission," Proc. IRE Vol. 14, No. 1, Feb. 1926. pp. 7-56.
8. H. T. Friis and E. Bruce, "A Radio Field-Strength Measuring System for Frequencies up to Forty Megacycles," Proc. IRE, Vol. 14, No. 4, Aug. 1926. pp. 507-520.
9. R. A. Heising, J. C. Schelleng, and G. C. Southworth, "Some Measurements of Short Wave Transmission," Proc. IRE, Vol. 14, No. 5, Oct. 1926. pp. 613-648.
10. Lloyd Espenschied, "Methods for Measuring Interfering Noises," Proc. IRE, Vol. 19, No. 11, Nov. 1931. pp. 1951-1954.
11. R. K. Potter, "High-Frequency Atmospheric Noise," Proc. IRE, Vol. 19, No. 10, Oct. 1931. pp. 1731-1765.
12. W. W. Mutch, "A Note on an Automatic Field Strength and Static Recorder," Proc. IRE, Vol. 20, No. 12, Dec. 1932. pp. 1914-1919.
13. Karl G. Jansky, "Directional Studies of Atmospherics at High Frequencies," Proc. IRE, Vol. 20, No. 12, Dec. 1932. pp. 920-1932.
14. Karl G. Jansky, "Electrical Disturbances Apparently of Extraterrestrial Origin," Proc. IRE, Vol. 21, No. 10, Oct. 1933. pp. 1387-1398.
15. Karl G. Jansky, "Minimum Noise Levels Obtained on Short-Wave Radio Receiving Systems," Proc. IRE, Vol. 25, No. 12, Dec. 1937. pp. 1517-1530.
16. H. O. Peterson, "A Method of Measuring Noise Levels on Short-Wave Radiotelegraph Circuits," Proc. IRE, Vol. 23, No. 2, Feb. 1935. pp. 128-131.
17. H. O. Merriman and F. G. Nixon, "Radio Interference- Investigation, Suppression, and Control," Proc. IRE, Vol. 27, No. 1, Jan. 1939. pp. 16-21.
18. Charles M. Burrill, "Progress in the Development of Instruments for Measuring Radio Noise," Proc. IRE, Vol. 29, No. 8, Aug. 1941. pp. 433-441.
19. Charles M. Burrill, "An Evaluation of Radio-Noise-Meter Performance in Terms of Listening Experience," Proc. IRE, Vol. 30, No. 10, Oct. 1942. pp. 473-478.
20. Harold E. Dinger and Harld G. Paine, "Factors Affecting the Accuracy of Radio Noise Meters," Proc. IRE, Vol. 35, No. 1, Jan. 1947. pp. 75-81.
21. Robert S. Hoff and Raymond C. Johnson, "A Statistical Approach to the Measurement of Atmospheric Noise," Proc. IRE, Vol. 40, No. 2, Feb. 1952. pp. 185-187.
22. William Q Crichlow, "Noise Investigation at VLF by the National Bureau of Standards," Proc. IRE, Vol. 45, No. 6, June 1957. pp. 778-782.
23. URSI. "The Measurement of Terrestrail Radio Noise," Special Report 7, International Union of Radio Science, Brussels, Elsevier Publishing Co., Amsterdam, 1962.
24. CCIR, "World Distribution and Characteristics of Atmospheric Radio Noise," Report 322, International Radio Consultative Committee (CCIR), International Telecommunication Union, Geneva, Switzerland (1964).
25. A. D. Spaulding and R.T. Disney, "Man-made Radio Noise, Part 1: Estimates for Business, Residential and Rural Areas," Institute for Telecommunications Sciences, OT Report 74-38, June 1974.
26. CCIR, "Man-Made Noise," Report 258-3, International Radio Consultative Committee (CCIR), International Telecommunication Union, Geneva, Switzerland (1978).
27. R. J. Matheson and K. R. Beasley, Institute for Telecommunications Sciences, U.S. Department of Commerce, Boulder Colo., "Field Intensity Meter Modified to Measure RMS and Average Noise Envelope Voltage," ITS Technical Memorandum No. OT TM-119, Nov. 1972.
28. G. H. Hagn, SRI International, 1611 No. Kent. St., Arlington, Virginia 22209, "Interim DOD Radio Noise And Electromagnetic Interference (EMI) Survey Handbook," Project 6512, July 1979.
29. J. W. Engles, Jr. and G. H. Hagn, "MF and HF Man-made Noise Measurements," Paper presented at the USNC/URSI 1973 Annual Meeting, Univ of Colorado, Boulder, Colo., Abstract, Aug. 21-24, 1973. pp. 107.
30. Wilfred R. Lauber, "Radio Noise Surveys at Canadian HF Communication Sites," IEEE Trans. on EMC, Vol EMC-19, No 2, May 1975. pp. 44-49.
31. G. H. Hagn, "Calibration of 9-Ft Rod Antenna for MF and HF EMC Measurements," IEEE International Symposium on EMC Record, June 20-22, 1978. IEEE Cat. No. 78CH1304-5EMC.
32. Arthur D. Spaulding and James S. Washburn, Institute for Telecommunication Sciences, U.S. Department of Commerce, "Atmospheric Radio Noise: Worldwide Levels and Other Characteristics," NTIA Report 85-73 April 1985.
33. A. D. Spaulding, Institute of Telecommunications Science, U.S. Department of Commerce, "Man-made Noise: The Problem and Recommended Steps Toward Solution," OT Report 76-85, April 1976.

ACKNOWLEDGEMENTS

The author gratefully appreciates the continuing technical discussions with and information received from Mr. G. H. Hagn, Mr. R. J. Matheson and Dr. A. D Spaulding.

DISCLAIMER

No endorsement of any manufacturer's product mentioned in this paper is made or implied by the U.S. Army. The equipment named or identified illustrates the type of equipment used. Equipment with equivalent technical characteristics would be equally satisfactory.

SIGNAL NOISE/INTERFERER COMBINER UNIT PROGRAMMABLE (SINCUP)

By Mr. Emilio Martinez De Pison
NAVOCEANSYSCEN, Code 832
271 Catalina Boulevard
San Diego, California 92152-5000

ABSTRACT

The Signal Noise Interferer Combiner Unit Programmable (SINCUP) has been developed to facilitate laboratory performance testing of Very Low Frequency (VLF)/Low Frequency (LF) receivers. To accomplish this, the unit allows the combining in controlled amounts of various real-world environmental and manmade interference with an information carrying signal. The externally modulated signal is combined with internally/externally generated Gaussian noise and/or with an internally/externally generated interferer. In order to test modern digital processing techniques, such as Adaptive Null Steering, Eigenvector Sorting, and Widrow-Hoff adaptive filters, SINCUP is capable of generating and meeting much higher signal-to-noise plus interference ratios than earlier channel simulators. The present software has been written to accommodate a dynamic signal-to-noise ratio (SNR) range from -60 to +60 db. Higher dynamic range units could be implemented.

BACKGROUND

SINCUP has been designed to test both analog and digital VLF/LF receivers. Long-range military communication channels are prone to noise perturbations, which can cause errors in received messages. The error rates are a function of signal-to-noise plus interference ratio and therefore extensive laboratory simulations are usually necessary to establish receiver performance as a function of this ratio. SINCUP provides a flexible test tool designed to support testing of VLF/LF receivers operating in the frequency range from 10 to 60 kHz and could be extended to 160 kHz with minor hardware modifications.

SINCUP was developed to replace earlier channel simulators which cannot function with today's advanced signal processing techniques that require testing formats of VLF/LF receivers at extremely low SNR's. SINCUP will also circumvent certain functionality problems that had become manifest, such as ground loop problems which arose while testing signal levels in the microvolt range. These problems are caused by the physical separation of various parts of the system under test using different ground points in a laboratory area where several other systems are under test.

A typical test configuration involves a signal source (usually transmitting a test message), noise and interferer sources, a true-root-mean-square (TRMS) meter to measure signal-to-noise interferer levels, the receiver under noise stress, and other monitoring equipment (such as bit or character error counters). A typical test setup using SINCUP is shown in figure 1. It consists of a communication signal source, a channel simulator, and the receiver under test. SINCUP is the central hardware component used to simulate the channel. The internally provided SINCUP capabilities that generate interference and Gaussian noise can be supplemented or supplanted through use of an external interferer source and/or an external Gaussian noise generator. The programmable features of SINCUP automate testing and provide repeatability of test results.

The communication signal source consists of a message source, such as a tape loop on a transmit distributor, connected to the input of a transmit modem. A message is then processed in one of several ways by the transmit modem;for example, it could be encoded to provide error detection and correction, encrypted, or multiplexed with other inputs. The resulting binary sequence is then converted to an analog signal through one of several modulation techniques, such as Frequency Shift Keying (FSK) or Minimum Shift Keying (MSK). The modulated signal is then amplified and broadcast.

The broadcast signal is normally propagated through the atmosphere and received along with other VLF/LF signals (manmade interferers) by the receiver. In the case of a laboratory measurement, the channel can be simulated by attenuating the communication signal to account for propagation losses of the signal and by a linear combining network to add to the communication signal both interference and environmental noise. In the operational environment, interference and noise levels range widely relative to the communication signal levels. Thus, in order to characterize receiver performance, simulations must reproduce a wide variety of conditions. Multiple channel effects, such as sea state effects, must be imposed by other techniques not represented in figure 1.

In order to characterize the channel as simulated through the addition of interference and noise, each of the noise and signal sources are monitored prior to the receive modem in a known and calibrated bandwidth to determine signal-to-noise and signal-to-interference levels. SINCUP allows the user to select signal, interference, and noise levels and exercise the receiver at these levels for selected periods of time. If a preformatted test message is used, it is easy to count the character errors that have occurred for the given signal-to-interference plus noise level, and thus obtain a point on the receiver operating curve (character error rate (CER) vs. SNR curve). After the completion of a series of tests at different signal-to-interference plus noise levels, it is easy to construct a waterfall, or CER vs. SNR performance curve for the receiver.

TYPICAL CHANNEL SIMULATOR

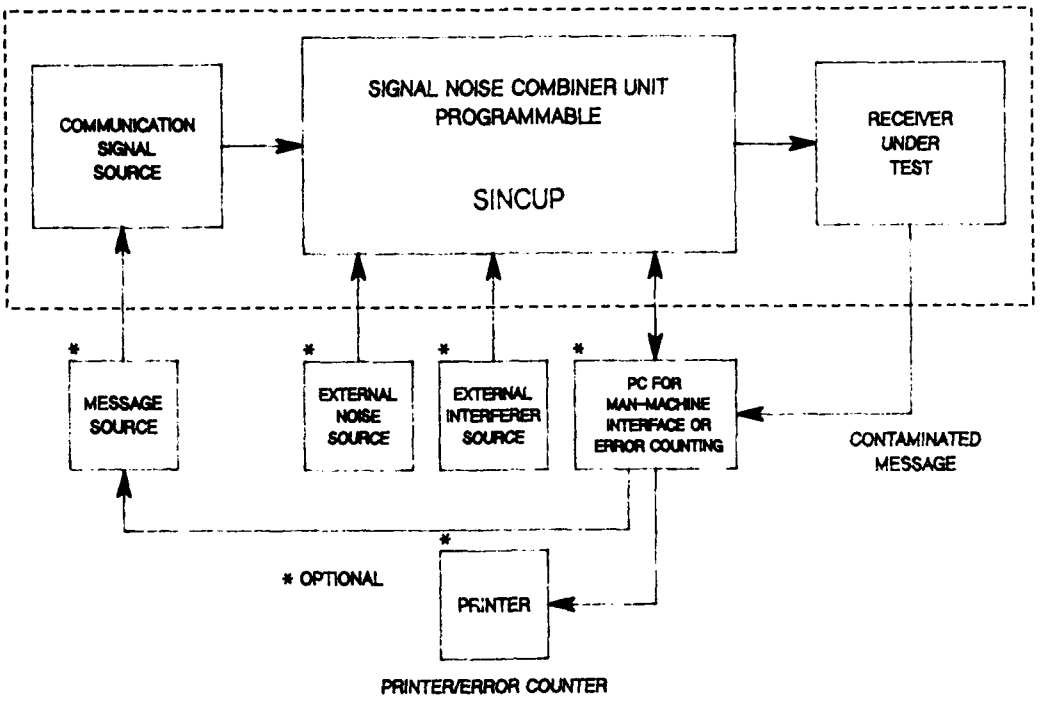


Figure 1. Receiver Test Setup with SINCUP

DESCRIPTION

SINCUP is an automated standalone test support tool conceived, designed, built and tested at NAVOCEANSYSCEN. SINCUP converts time-varying voltage values of analog signal and noise components to TRMS values so that appropriate measurable components can be attained for SNR computations. The TRMS values are digitized and processed in fixed-point binary format so that they are represented accurately.

A functional block diagram of SINCUP is shown in figure 2. Subsequent paragraphs describe the various SINCUP components.

The SINCUP's circuitry and related hardware mount in a 17" x 12" x 20" rack-mountable chassis. In order to minimize alternating current induced noise, a direct current (DC) powered fan is used for cooling. Also, contrary to the wide use of switching power supplies in today's state-of-the-art designs, it was decided to use linear power supplies due to their low noise characteristics (despite their considerably larger sizes). The multi-output linear power supply rack mountable is separately mounted beneath SINCUP's chassis to minimize self-noise.

SINCUP's controlling and processing power is provided over a STD Bus system. A front panel mounted Microterminal with an ASCII format keyboard and a 16-character light-emitting diode (LED) display provides the man-machine interface for SINCUP, while dual RS-232 I/O ports provide data and computer interfaces. The computer interface allows SINCUP to be controlled by an externally connected IBM-compatible PC XT/AT. The PC's processing and displaying power relieves SINCUP of generating system functions (which are not presently implemented) such as providing error counts, printouts, and most importantly, the ability to automate test scenarios with display presentations.

DIGITAL CONTROL

See figures 2 and 3.

The SINCUP's control and processing power is generated by the built-in STD Bus system consisting of six (6) STD BUS boards. The central processing unit (CPU) board has 22 Kbytes of programmable read-only memory (EPROM) for program code and 2 Kbytes of random access memory (RAM) to store program variables and provide input/output (I/O) buffers.

The man-machine interface between the CPU and either the microterminal or the external PC is achieved by one Dual Channel RS-232 SYNC/ASYNC Communications board. Each of three Parallel Output Boards are capable of driving six buffered ports for command-word outputs to the Signal, Noise and Interferer Attenuators, to the Frequency

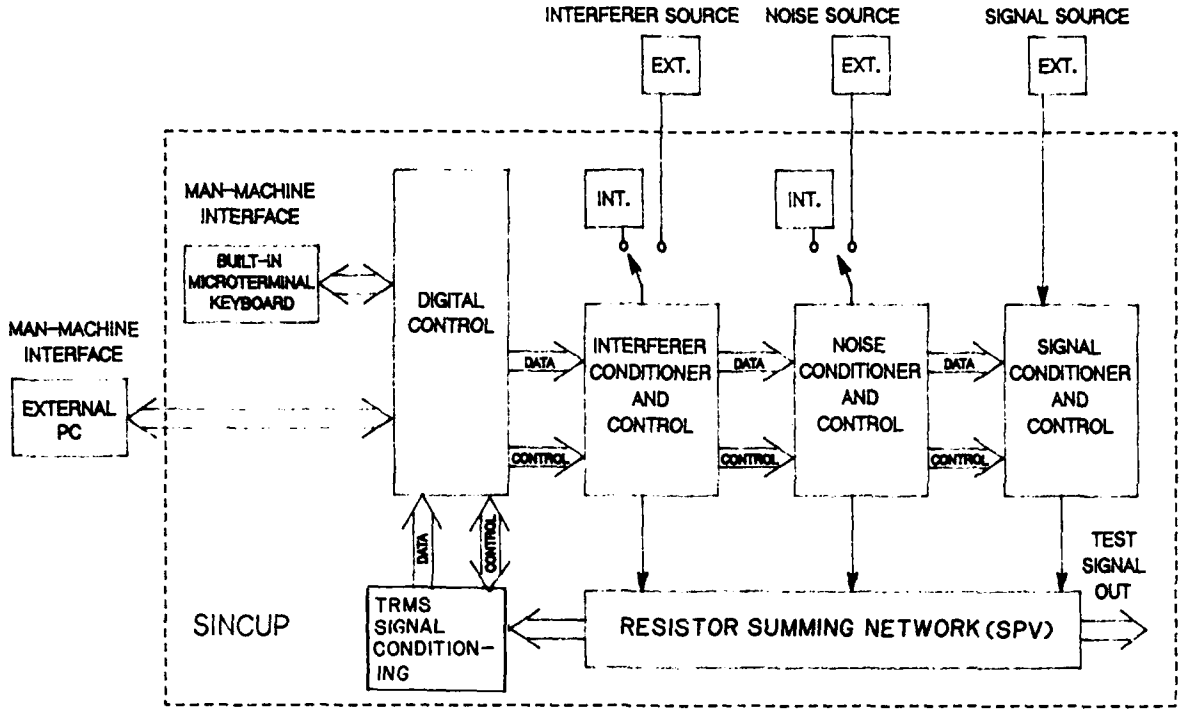


Figure 2. SINCUP Functional Block Diagram

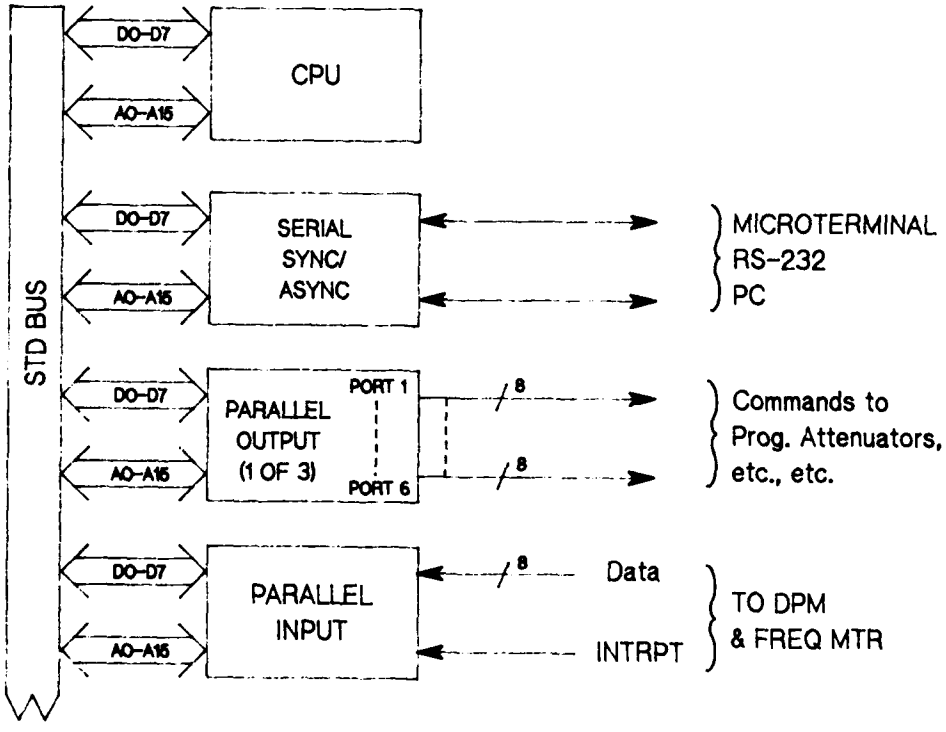


Figure 3. Digital Control

Synthesizer (interferer source), to the Dual Channel Programmable Buffer Amplifier (PBA) and to the various remotely located radio frequency (RF) relays.

The last STD BUS board provides six buffered parallel input ports used to recover the digitized data from the digital panel and the frequency meter.

SIGNAL CONDITIONING

See figures 4, 5, and 6.

SINCUP's remaining PC boards are non STD BUS type: there are three commercially acquired ones--"Gaussian Noise Source", "Frequency Synthesizer" (interferer source), and the "Sine Wave Converter" board, which converts the Frequency Synthesizer transistor-to-transistor logic (TTL) square wave output to a sine wave.

Additionally, there are three in-house designed PC boards:

(1) The "Programmable TRMS Signal Conditioner", which for the purpose of accuracy while converting to TRMS DC equivalent, preamplifies the selected signal or noise component prior to establishing a given SNR, converts this analog signal to a TRMS DC equivalent, and then attenuates this resultant DC signal to accommodate the DPM (digital panel meter/analog-to-digital converter) full scale range of 200 mv DC;

(2) the "Programmable Buffer Amplifier", which conditions the noise or interferer's sources in amplitude and line driving capabilities; and

(3) the "Isolation Systems Control" board, on which complete logic control and CPU interrupting capabilities are provided for the DPM and for the FREQUENCY METER (internal interferer's frequency display/interferer's analog-to-digital converter). Additionally, RF relay drivers, a 75 kHz low pass filter to produce band-limited Gaussian noise, and three isolation transformers are mounted on this board. The transformers ground isolate the external signal, noise, and interferer sources from SINCUP.

NOISE CHANNEL SIGNAL CONDITIONING FOR NOISE POWER BANDWIDTH CALCULATION

It is important to mention that the 75 kHz Butterworth low pass filter was a logical design choice. SINCUP needed a precisely defined frequency of 75 kHz, along with a filter providing constant amplitude in the pass band and sharp roll-off at the 3 db points. Such filter characteristics would allow operator-aided determinations of Noise Power Bandwidth by empirical means. Then adequate calibration Gaussian noise "K" factors could be obtained for use during SNR computations (see Signal Handling Computations to follow). SINCUP should also be capable of upgrades which allow the

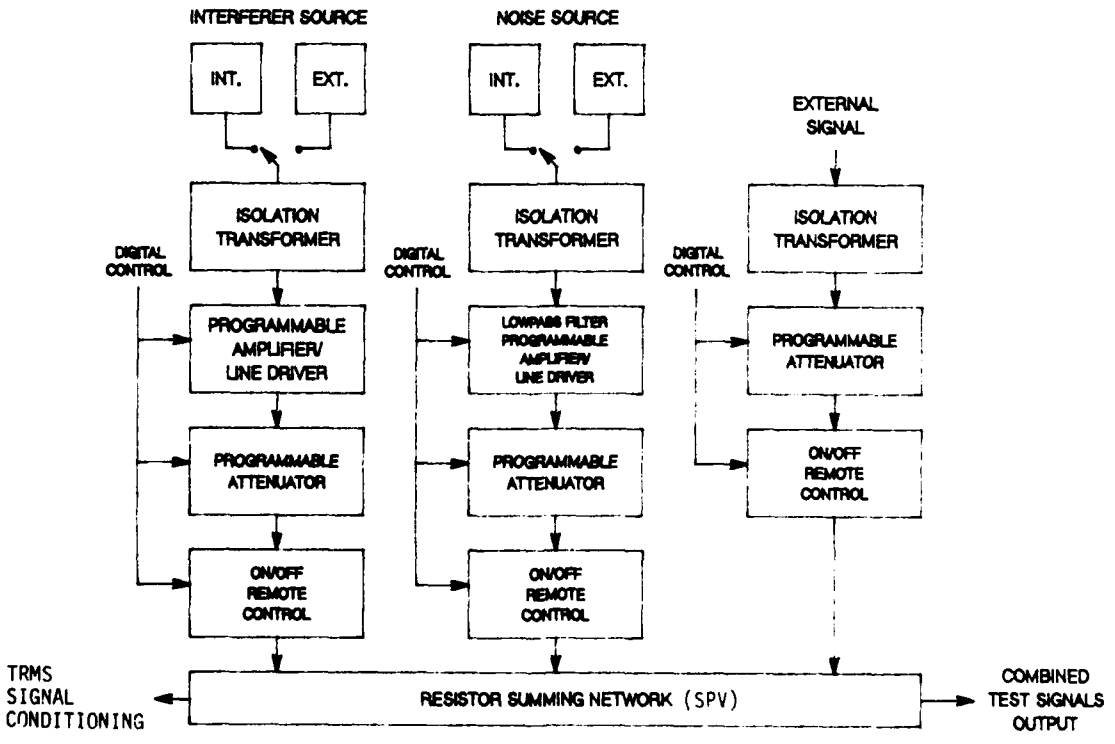


Figure 4. Signal Conditioners

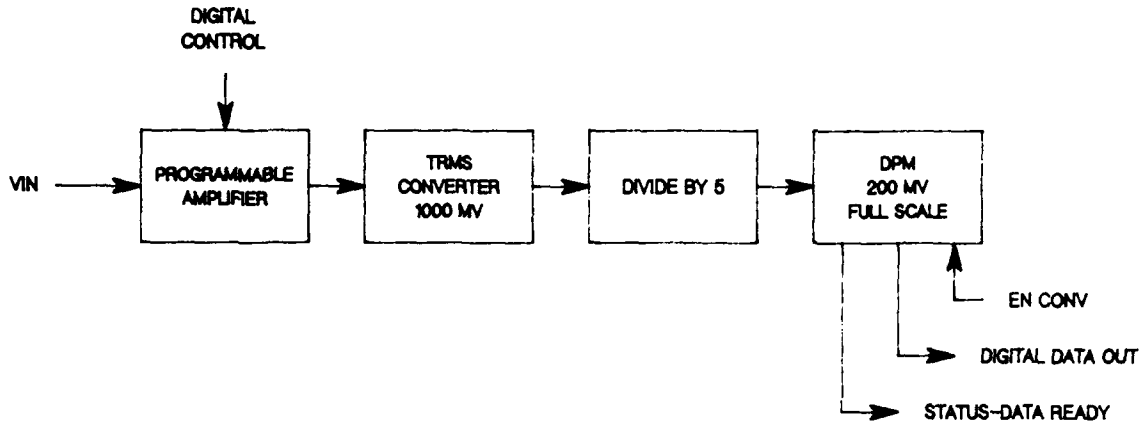


Figure 5. TRMS Converter Signal Conditioner

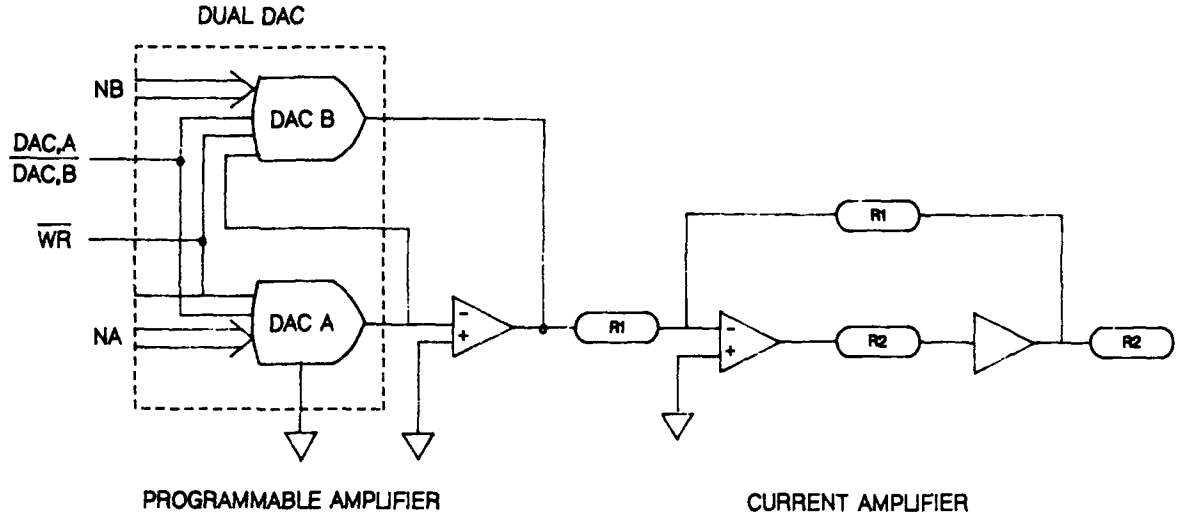


Figure 6. Programmable Buffer Amplifier (one-half shown)

frequency capability of around 160 kHz, even though present receivers under test have a 10 to 60 kHz frequency range.

By definition the Noise Power Bandwidth (NPB) of a transmission function is defined as the width of an ideal bandpass filter expressed in Hertz which has an absolute value of the transmittance Y(f) in its passband equal to the maximum absolute value Yo of the transmittance function of a non ideal filter:

$$NPB = \frac{1}{Yo^2} \int_0^{\infty} |Y(f)|^2 \, dt$$

Alternatively, the NPB is the filter bandwidth which delivers the same output noise power when driven by the same amount of white noise power under similar input and output terminations.

The corresponding SINCUP's NPB was obtained empirically by using standard laboratory methods: passing a signal with constant input amplitude through the transfer function in question (noise channel) and reading the corresponding TRMS output voltage magnitude at various selected frequency points of interest (10 Hz up to that frequency for which the maximum output voltage magnitude squared has decreased by 50 db). Then "k" factors (as shown below) can be obtained by the relationship K = 10log[NPB]/MR, where MR stands for modulation rate.

PROGRAMMABLE BUFFER AMPLIFIER PC BOARD

See figure 6

The Programmable Buffer Amplifier (PBA) consists of two identically designed PBAs; one amplifies the Gaussian noise and the other amplifies the interferer signal. The programmable features are possible due to a dual eight-bit digital-analog-converter (DAC), such as the model AD 7528, which is manufactured by Analog Devices. The equivalent resistance of each DAC from input to output is used to provide the input and feedback resistors in a standard inverting operation amplifier.

By loading the DAC by program control with suitable codes, programmable gain or attenuation over the range of -48 to 48 db is possible governed by the following relationship:

$$Vout = (Na/Nb)*Vin \quad (1)$$

where $1 < Na < 255$
$1 < Nb < 255$

where Na > Nb provides gain and Na < Nb provides attenuation. Na or Nb is an eight-bit word transferred into either of the two DACs' data latches via a common eight-bit port. Control input DAC/A or DAC/B with /WR determine which DAC is to be loaded.

Superior resistor matching and tracking performance is achieved with the dual DAC since both DACs are built in the same chip and therefore have similar properties because of common laser-trimmed fabrication characteristics. A two-stage preamplifier is used to assure optimum gain characteristics across the frequency band of interest. During measurement of a given summing point signal, the PBA is slaved to the digital processor whenever gain or attenuation values have been programmed. Additionally, these amplifiers provide the PBA with current driving capabilities to drive the 50 ohms input impedance offered by the noise and interferer programmable attenuators.

TRMS SIGNAL CONDITIONER PC BOARD

See figure 5.

The signal conditioner consists of four main building blocks: a TRMS preamplifier, a TRMS converter, a voltage divider, and a digital panel meter (DPM).

The TRMS converter is a model 442L manufactured by Analog Devices. This converter is a high performance true-RMS-to-DC converter, which provides conversion accuracies of +/-2 mv for input signals of up to 175 kHz with levels between 0 to 2 volt RMS. The model 442L provides one volt DC output for one volt TRMS input. It is this DC output that the voltage divider divides by five to accommodate the DPM maximum full scale range of 200 mv. See figure 5.

The DPM is basically an analog-to-digital converter with a visual readout. The DPM samples the input voltage periodically, converts that voltage to digital outputs, and displays the corresponding reading visually. The DPMs output consists of three binary coded decimal (BCD) bits plus an out-of-range overflow bit (a logical "1" indicates over 100 mv being converted and read). In addition, the DPM's output consists of: a polarity bit, a data ready bit to inform the digital processor that conversion of data has occurred, and the overload bit, which informs the digital processor that >199.9 mv is being processed. A LED display flashes "0"s to warn the user of an overload condition.

Note, built-in diagnostic software tests the DPM measuring and digitizing accuracy by commanding a multiplexer to switch in a calibrating reference voltage provided by a INTERSIL model ICL8069. This diagnostic applies a stable 110 mv DC voltage to the input of the DPM to establish its accuracy prior to use of SINCUP for performance testing for a given signal-to-noise ratio (SNR).

SYSTEMS CONTROL AND ISOLATION PC BOARD

The Systems Control and Isolation board provides a mounting platform for various components, such as:

a) the signal, noise and interferer RF relay drivers

b) the 75 kHz low pass filter used to band limit Gaussian noise

c) the signal, noise, and interferer isolation transformers. This board provides the hand-shaking necessary to interface the DPM and Frequency Meter with the digital processor. It controls data and interrupt exchanges during the analog-to-digital conversions of these meters as follows.

The main program initiates a conversion by sending the appropriate command word to a systems control and isolation board. To enable the DPM for conversion, a logical "one" must set a flip-flop on "high one" simultaneously with the arrival of a strobe pulse. This provides a signal which enables conversion and which triggers the DPM to start conversion. See figure 5. Once the DPM completes conversion, it raises its STATUS

BIT to a "high one", which in turn triggers a flip-flop so that it outputs a logical "one" causing a one shot to pulse shape a conversion complete STATUS signal from the DPM. Similarly, when the FREQ METER requires interferer data to be digitized so that the processor can digitally perform frequency measurement, the program must send the appropriate word-command-bit to set a logical "one" so that the digital processor can be interrupted. In so doing the processor initiates the reading of the FREQ METER digitized data. This is accomplished by comparing two consecutive frequency samples until they match, which indicates that the desired analog-to-digital conversion has taken place.

MANUAL/EXTERNAL FEATURES

The user has the option to recall a DIAGNOSTIC program named MANUAL by pressing the PFK "MANUAL" whenever the microterminal shows "READY" on its display and then, for the purpose of calibrating and servicing SINCUP when required, the summing point has been made accessible to external devices, such as a TRMS voltmeter, selective voltmeter/wave analyzer, oscilloscope or spectrum analyzer--via a BNC connector labeled COMBINE SIGNAL MONITOR.

A common application is that of checking out what particular functions the stored program has implemented upon request for a given SNR or signal-to-interference ratio (SIR) vs. that which is measurable via "COMBINE SIGNAL MONITOR". After a given test scenario has been programmably implemented, the user can access the manual mode by entering "ENTER" on the microterminal. Then, he enters MANUAL, to which the program asks "ENTER PORT ADDRESS"; the operator enters that corresponding command word address that controls ON/OFF of the desired function. The program then requests the control data bit that will turn on or off that particular function. An example would be to monitor the corresponding signal and noise being supplied to the summing point to achieve a specific SNR. Assuming a TRMS meter is connected to monitor the summing point, the user, via MANUAL, can maintain a dialogue with SINCUP such that the programmable noise relay can be turned off and the signal relay turned on, so that the corresponding signal magnitude (requested by the programmable implementation of the desired SNR) can be manually measured. The same approach is followed to determine the noise level. Once these magnitudes are available, methods described below can be used to ascertain that SINCUP provides the desired SNR within +/-0.1 db.

Two other DIAGNOSTICS are available. One diagnostic is initiated through "ATTEST" and allows an operator to verify correct performance of the attenuators. The other diagnostic initiated by "DISPLAY STATUS" displays information on the SINCUP configuration.

SIGNAL HANDLING AND COMPUTATION

Figure 7 summarizes the hardware components and mathematical factors involved in the process relating a voltage measurement at the summing point to a computational process by the computer software.

SINCUP software calculates SNRs using the following formula:

SNR (in dB and referred to a bandwidth equal to 1/(signaling interval) = 20log(ext signal)/noise + K (db) with

K = 10log (NPB/MOD RATE) where

NPB is the SINCUP's Noise Power Bandwidth (or Equivalent Noise Bandwidth)

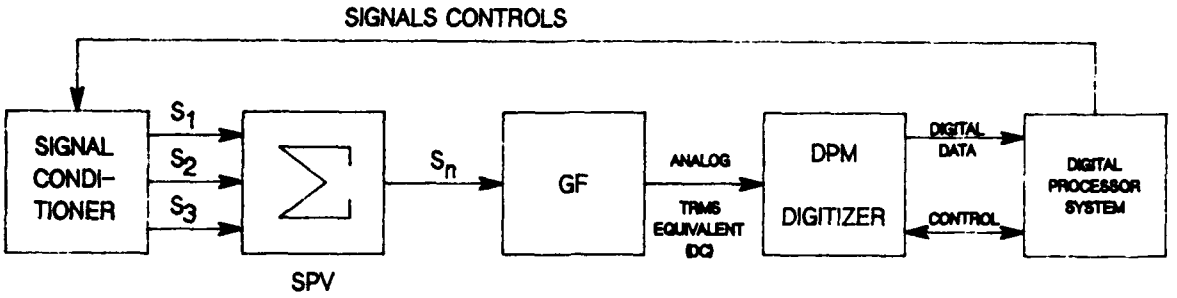
MOD RATE is the communications signal modulation rate, or 1/(signaling interval)

Note that "K" accounts for bandwidth differences between SINCUP and the receiver's intermediate frequency section. SINCUP has approximately 70 kHz of bandwidth and the receiver has a bandwidth of 200, 400, or 1000 Hz, approximately 1/(signaling interval). Had SINCUP been equipped with the same bandwidth as that of the receiver, measurement of its noise component would have become impractical due to variability of Gaussian noise and the inability of most TRMS meters to provide accurate readings under such conditions.

If an externally modulated signal is to be measured and conditioned to provide a desired SNR, the following software/hardware interactions occur:

1. To provide the signal at the summing point, where the mixing of the three signals occurs and at which each signal, individually, is measured by the Signal Conditioner.

A DAICO SWITCH model 100-C0878-12 is turned on by the remotely controlled digital word sent from the digital processor and via the parallel output STD BUS board to a Texas Instrument relay driver model 75468. Then, a Wavetek programmable signal attenuator model P127BB12-TTL is sent a digital control word that attenuates an input signal (provided that the signal level is not less than 50 or higher than 75 mv (TRMS)) sufficiently to allow this modulated signal to activate and initialize the intended Receiver System. Also, a Wavetek programmable combined signal output attenuator model P63BB12-TTL is commanded to let sufficient signal out of the summing point to the Receiver System under test.



Let S_n = one of three signals

Let GF = 5/G, G = Preamplifier Gain (known by the program)

Let 5 = Pre-DPM attenuation factor

Then the program computes "SPV" by SPV = GF · DDa
whereas DDa is the absolute (DPM_{out} − $DD_{initial\ DC\ offset}$) digitized
TRMS DC equivalent fed to the Digital Processor

Figure 7. Signal Measurement/Computation

2. Measurement and computation of a given signal is initialized by the program setting the corresponding attenuator to provide the desired SNR and the corresponding switching relay to its turned on state. The signal magnitude is measured by passing it into the combining network, the Signal Conditioner, the STD BUS parallel input board, and the digital processor, where all the measurement controlling algorithms reside. The formula, SNR db = 20log(Vin/Vmeas), is used to determine the input signal level Vin (at the front end of SINCUP) in terms of the measured voltage value, Vmeas, or at a 0 dB setting of the attenuator in question from which the desired signal attenuation (Vin/Vout) can be attained.

The variable terms "SNR" and the ratio S/N in the formula above are software manipulated to establish a given SNR. To accomplish this the signal conditioner measures the modulated signal value S and the Gaussian noise level N at the summing point. The computer program then sets and measures the largest of the components to be used in the SNR computation. Assume the Gaussian noise (N) has been programmed by the Noise Programmable Buffer Amplifier to supply 200 mv (TRMS) to the combining network. Then the program commands the appropriate circuitry for measurement as follows: in order to provide the TRMS converter with 1000 mv of input signal for optimum conversion, the pre-amplifier is sent a combined gain command of FIVE, so that the 200 mv appearing at its input and out of the summing point becomes amplified to 1000 mv. The program stores this gain command for future computations. The TRMS converter output (1000 mv DC) is attenuated by the resistor voltage divider and the resultant 200 mv DC is fed to the DPM input. Next, the DPM is commanded by the digital processor to convert this dc signal into digital data, by sending an enable conversion bit. This bit is stored momentarily in a flip-flop and starts analog-to-digital conversion in the DPM. The DPM generates a status bit once conversion is completed. This status bit must be pulse shaped by a one-shot before being sent to the digital processor. The end result is an interrupt (15 microseconds long) pulse sent to the digital processor to indicate that digital data awaits recovery. At this point, the program reads the digital data and computes the corresponding signal at the summing point. The magnitude of the noise signal to be mixed with other signals at the summing point is calculated as follows (see figure 7):

SPV = DDa*5/G where

SPV = summing point value

DDa = absolute DPM digitized reading
(instantaneous reading - DDi)

G = preamplifier gain.

5 = TRMS DC output attenuation factor to accommodate DPM full scale range

DDi = initial DPM DC offset voltage read under initial program control prior to all SNR measurements.

The magnitude of the signal at the summing point is calculated as follows: S = N*10**(SNR-K)/20, where N and SNR are known and K is a constant. This formula determines S in a computational sense; however, it is necessary to describe how the value of S is obtained in a hardware sense. Since Vin and S are known, the attenuation (A) in dB needed to reduce Vin to S is given by A (db) = 20 log(Vin/S). The quantity (A) is then in succession converted from decibel to linear and from linear to an eight-bit binary word. The eight-bit binary word controls the attenuator.

OPERATION

As shown in figure 2, the user interfaces with the SINCUP program through either the Microterminal or an IBM-compatible PC. The program is initialized at power-up time or by use of the "RESET" key of the Microterminal keyboard. Once initialized, a built-in diagnostics computer program automatically tests the hardware and performs self-calibration of its principal analog circuits. These circuits include the programmable amplifier and the digital panel voltage meter's initial DC offset voltages; the testing of the CPU instruction recognition; the testing of the CPU's built in RAM and Read-Only Memory (ROM); the testing of the DPM, FREQ. METER, and the internal FREQ. SYNTHESIZER's input and output hardware/software. Failure of any of the tests is displayed on the display with an error number. After passing the tests, a "READY" message on the Microterminal's display prompts the user to enter SNR values.

Up to three signal components (external modulated signal, internal or external Gaussian noise and an internal or external interferer of 10 to 60 kHz specified to within 1 Hz) can be combined linearly in ratios specified by the user.

Programmable function keys (PFK) are used to enter a selected SNR in dB, namely SIGNAL, EXT, and ENTER. The EXT SIG LED then lights and a READY message is displayed, which prompts the user to continue. The user then presses NOISE, IN., and ENTER. Again, the noise LED lights and a READY message reappears to acknowledge his selection. Next, the user presses COMBINE and ENTER and the user is prompted by "ENTER MODULATION RATE" to enter a two, three, or four digit number. The program then prompts the user with the message "ENTER SNR" to enter a number between -60 and +60 specified to within 0.1. SINCUP then verifies that the resulting signal components are within acceptable limits; if not, a message is displayed indicating that the amplitude of the signal is too high or too low. The user can then take corrective action, such as increasing or decreasing the SIGNAL EXTERNAL amplitude. Finally, after all requested parameters have been appropriately answered, the user is prompted to enter the desired combined signal output level that is to be output to the receiver under test (expressed in dB). The latter facilitates computing the attenuation value to be sent to the output attenuator relative to a maximum value of 10 millivolts. The output level can be changed by the user at any time during regular testing by pressing PFK CHANGE OUTPUT and entering a new value.

SUMMARY

Ground loops are virtually eliminated by SINCUP by built-in Gaussian noise and interference sources and by the specially designed "TRMS" meter. Electromagnetic interference (EMI) is further reduced through the use of linear DC power supplies. Because of the low resulting noise floor, SINCUP provides the capability to run tests over a large range of accurately generated SNR values.

SINCUP automates VLF/LF testing by expediting pre-testing scenario preparation by performing self-calibration, conducting hardware tests, and supporting operator selection of SNRs for testing. It supports unmanned computer-based closed-loop test scenario implementations providing high accuracy error counting and supports the generation of graphic summaries of performance data.

EFFECTS OF NOISE AND INTERFERENCE ON SYSTEM PERFORMANCE

Arthur D. Spaulding
U. S. Department of Commerce
National Telecommunications and Information Administration
Institute for Telecommunication Sciences
325 Broadway
Boulder, Colorado 80303
U. S. A.

SUMMARY

The real-world noise environment is almost never white Gaussian in character, yet receiving systems in general use are those that are optimum for white Gaussian noise (e.g., linear matched filter or correlation detectors), and, consequently highly suboptimum in the actual interference environment. It is well known that Gaussian noise is the "worst" kind of noise in terms of minimizing channel capacity or in its information destroying ability. This means that substantial improvement in system performance or signal detectability can be achieved if the actual statistical characteristics of the interference are properly taken into account. Early attempts to take account of the noise character involved ad hoc methods of trying to make the noise look "more Gaussian" and therefore better match the given (Gaussian) receiver. These attempts were usually not particularly successful and only recently (in the past 10 years or so) have detectors been designed based on the actual interference.

This paper starts with a summary of the two basic types of noise processes that confront us: "narrowband" processes composed of various collections of narrowband intentionally radiated signals, coherent pulse trains, various kinds of unintentionally radiated manmade noise, etc., all termed Class A; and "broadband" processes such as atmospheric noise from lightning and various forms of manmade noise such as automotive ignition noise, etc., termed Class B. After a short summary of these noise processes, stating only the "models" in current use (and which we use later in the paper), we review the performance of "normal" (linear) systems in Class A and Class B noise, not in great detail, but sufficiently to give a good feel for the diverse performance degradations obtained. Examples of both digital and analog systems are used. After reviewing the substantial degradation the real-world interference environment can cause the existing systems, we give results, some quite recent, of the performance of systems especially designed to take advantage of the actual interference, showing very substantial peformance improvements. Examples of both techniques currently being pursued, parametric and nonparametric, are used. Overall, the paper is hoped to represent a reasonably self-contained summary.

INTRODUCTION

In this basically review paper, we want to start by specifying the currently used noise and interference models, and then summarize the performance of example basic modulation systems, both digital and analog, in the actual interference, as expressed by these models, but also including measured results.

NOISE MODELS

In order to be able to determine the performance of a given communication system in interference or to determine the optimum receiving system and analyze its performance, a mathematical model for the random interference process is required. That is, for optimal system studies or for determining the performance of some of the existing suboptimum systems, more information about the noise process is required that can generally be obtained by measurement alone. The problem has been to develop a model that fits all the available measurements, is physically meaningful when the nature of the noise sources, their distribution in time and space, propagation, bandwidth, etc. are considered and is directly relatable to the physical mechanisms giving rise to the interference. Models that have been developed to date do not, with a few exceptions, satisfy the above criteria. Models can be categorized into two basic types: empirical models designed only to fit various measured statistics of the interference, and models which are designed to represent the entire random interference process itself. Almost all models that have been developed are for Class B noise (e.g., atmospheric noise) and a complete historical summary of these models has been given by Spaulding [1, 2]. The main ad hoc models for atmospheric noise (and some forms of man-made noise) is that given by the CCIR [3, 4]. Figure 1 shows this CCIR model, where the parameter V_d for "voltage deviation" is the dB difference between the envelope average and rms voltages. Figure 1 shows the exceedence probability of the received noise envelope. This has become the customary way of displaying noise data. We are assuming narrowband processes (i.e., expressible by envelope and phase) with the phase uniformly distributed. Unlike Gaussian noise, the noise characteristics are a function of receiver bandwidth, and Herman and DeAngelis [5] have given results on the effects of bandwidth on the received atmospheric noise process. Recent work by Middleton has led to the development of physical-statistical models for the entire interference process, both Class A and Class B [6, 7]. These models are physically derived, canonical in nature, and are analytically and computationally manageable. These models have been used to develop optimum detection algorithms for a wide range of communication situations [8, 9, and 10, for example]. Later in this paper, when we give examples of the peformance of these optimum systems obtained by simulation, we need to use simplified versions of the Middleton models. For our received interference process Z(t), the probability density function (pdf) for the received instantaneous amplitude, z, for Class B, is:

$$p_Z(z) = \frac{e^{-z^2/\Omega}}{\pi\sqrt{\Omega}} \sum_{m=0}^{\infty} \frac{(-1)^m}{m!} A_\alpha^m \Gamma\left(\frac{m\alpha+1}{2}\right) {}_1F_1\left(-\frac{m\alpha}{2}; 1/2; \frac{z^2}{\Omega}\right), \quad -\infty \le z \le \infty \qquad (1)$$

where ${}_1F_1$ is a confluent hypergeometric function. The model has three parameters: α, A_α, and Ω. The parameters α and A_α are intimately involved in the physical processes causing the interference and Ω is

a normalizing parameter. The corresponding expression for the envelope cumulative distribution (APD) is:

$$P(E > E_o) = e^{-E_o^2/\Omega} \left[1 - \frac{E_o^2}{\Omega} \sum_{m=1}^{\infty} \frac{(-1)^m}{m!} A_\alpha^m \Gamma\left(1 + \frac{m\alpha}{2}\right) {}_1F_1\left(1 - \frac{m\alpha}{2}; 2; \frac{E_o^2}{\Omega}\right) \right] \qquad (2)$$

$$0 \le E \le \infty.$$

Some time ago Hall [11] developed an ad hoc mathematically simple model for atmospheric noise (Class B), and Middleton has shown [6] that this Class B model reduces, approximately, for special parameter values, to expressions of the Hall Type. The Hall model has two parameters, θ and γ, and is given by

$$P_Z(z) = \frac{\Gamma\left(\frac{\theta}{2}\right) \gamma^{\theta-1}}{\Gamma\left(\frac{\theta-1}{2}\right)\sqrt{\pi} \left[z^2 + \gamma^2\right]^{\theta/2}} \qquad (3)$$

and

$$P[E>E_o] = \frac{\gamma^{\theta-1}}{\left(E_o^2 + \gamma^2\right)^{(\theta-1)/2}} \qquad (4)$$

The closest match between the Middleton model (1) and the Hall model (3) are for the Middleton parameters $\alpha = 1$, $A_\alpha = 1$, and $\Omega = 4 \times 10^{-4}$ with the corresponding Hall parameters $\theta=2$ and $\gamma=\sqrt{2} \times 10^{-2}$. For both the Middleton model ($\alpha=1$) and the Hall model ($\theta=2$) the second moment does not exist, so the normalizing prameters Ω and γ are set to match measured data. For $\theta=4$, however, the first 3 moments exist. We will review results of "optimum" system performance using the Hall model for Class B noise later.

The only model developed to date for Class A interference (defined above) is the physical-statistical model of Middleton. For Class A narrowband "impulsive" noise:

$$P_Z(z) = e^{-A} \sum_{m=0}^{\infty} \frac{A^m}{m!\sqrt{2\pi\sigma_m^2}} e^{-z^2/2\sigma_m^2}, \qquad (5)$$

where

$$\sigma_m^2 = \frac{m/A + \Gamma'}{1 + \Gamma'}, \qquad (6)$$

and, for the envelope

$$P(E>E_o) = e^{-A} \sum_{m=0}^{\infty} \frac{A^m}{m!} e^{-E_o^2/\sigma_m^2}. \qquad (7)$$

The Class A model has two parameters: A and Γ'. A is termed the overlap index, and as A becomes large (~10), the noise approaches Gaussian (still narrowband) and Γ' is the ratio of the energy in the Gaussian portion of the noise to the energy in the non-Gaussian components. Another model that is occasionally proposed is the Gaussian-Gaussian ϵ mixture model, which is given by

$$P_Z(x) = (1 - \epsilon)p_0(x) + \epsilon p_1(x), \qquad (8)$$

where p_0 and p_1 are both zero mean Gaussian densities, with $0 \le \epsilon < 1$, typically quite small, and $\sigma_1^2 > \sigma_0^2$. This (8) corresponds to Middleton's Class A model truncated after only two terms, with

$$\epsilon = \frac{A}{1+A}, \quad \text{and} \quad \frac{\sigma_1^2}{\sigma_0^2} = 1 + \frac{1}{A\Gamma'}. \qquad (9)$$

For actual implementation of detectors, we need to use an approximation to the Class A model (5), but a simple two term approximation, even when properly normalized, is not sufficiently accurate. The Gaussian-Gaussian ϵ mixture model, in any case, was originally proposed for Class B noise, for which it is quite inappropriate.

Figures 2 and 3 give an example of measured Class A and B interference, along with the corresponding Middleton Class A and B model. The above summarizes the currently used models, but, of course, leaves out all details. We next want to review the general characteristics of the performance of standard systems in the real-world environment, generally using the above models.

STANDARD SYSTEM PERFORMANCE

It has long been recognized that in most communication situations, the additive interference is not Gaussian in character, even though most existing systems are basically those known to be optimum in Gaussian noise (when such optimality can be determined, as in the case of simple digital systems). Correspondingly, there has been substantial effort to determine the performance both by theoretical calculations, and by measurement, of such systems. For surveys and summaries of early efforts see Fang and Shimbo [12] and the bibliography by Spaulding et al. [13] which contains some 315 references pertaining to system performance in impulsive noise. Summaries of more current studies can, perhaps, best be obtained from the International Union of Radio Science (URSI) Reviews [14], which occur each three years.

The desired signal can be subjected to additive interference, flat fading, and/or multiplicative interference which includes various signal distortions such as caused by frequency selective fading.

For digital signaling, the effect of multiplicative interference is to produce a probability of error, P_e, threshold; that is, a value of P_e which cannot be lowered by increasing signal power. For example, see Watterson and Minister [15]. For our examples here we will assume "properly designed" systems so that the irreducible P_e, which is always present, is low enough to be of no concern and we will only treat additive interference and flat fading signals.

The early work analyzed digital receivers in impulsive atmospheric noise by following the steps of the Gaussian analysis but using distributions appropriate for atmospheric noise, generally the CCIR model (Figure 1). In 1954, Montgomery [16] showed that for any arbitrary additive interference that is independent from an integration period (bit length) to the next and which has uniformly distributed phase, the P_e for binary symmetric NCFSK (noncoherent frequency shift keying) is given by one-half the probability that the noise envelope exceeds the signal envelope. This result is valid for current NCFSK systems (bandpass filters, discriminator receivers, and match filter, envelope detection receivers), but the "noise" in the signal-to-noise ratio, SNR, is the noise present in both filters, for example, rather than the current "textbook" answer where the SNR is defined using only that filter output which contains the signal (for dual filter systems). This 3 dB "difference" has caused much concern over the years. The Montgomery result means that the P_e for binary NCFSK can be read directly from the interference envelope exceedence distribution when these distributions are normalized to their rms level (Figures 1, 2, and 3, for example). For HF systems in general, Conda [17], using [16], has given results for NCFSK for the entire range of atmospheric noise conditions likely to occur and for a wide range of flat fading signal conditions. For digital systems and constant signal, white impulsive noise is much more harmful (causes more errors) than Gaussian noise of the same energy at the higher SNR's (i.e., lower P_e where we prefer to operate) while Gaussian noise is more harmful for the lower SNR's. Figure 4 illustrates this for binary NCFSK as an example, but all digital systems behave similarly. On Figure 4, Gaussian noise (Rayleigh envelope) is given by V_A=1.049 dB. When the signal is Rayleigh flat fading, Gaussian noise is more harmful at all SNR's. For diversity reception (used to overcome fading), however, impulsive noise is again more harmful at higher SNR's. Figure 5 illustrates this general result, again for binary NCFSK. Note that the nondiversity operation of binary systems with Rayleigh fading signal, performance for higher SNR's is independent of the noise statistics. That is, the Rayleigh fading statistic dominates. Other flat fading situations do arise for which impulsive noise will cause more errors than Gaussian noise at some SNR's as illustrated in Figure 6, where the signal is log-normal fading. Also, for diversity reception, impulsive noise, and Rayleigh fading signal, the degree of statistical dependence between the noise on the different diversity branches has a relatively minor effect on system performance for low orders of diversity. In addition to the basic P_e given in these examples, grouping of errors is important, since unlike with white Gaussian noise, the actual interference process is not independent from bit to bit, causing nonindependent errors. Also, differential systems (differentially coherent PSK, for example) have intersymbol dependency, even with independent noise. Halton and Spaulding [18] have given results for DCPSK, including the occurrence of various groupings of errors. All of this simply demonstrates the great suboptimality of standard systems (i.e., those optimum for Gaussian interference) when used in environments which are substantially different from Gaussian. Additional examples are included in the simulation results given next.

The situation for analog systems can be quite different from that above for digital systems, depending on the performance measure used. Figure 7 shows this for a standard AM system in atmospheric noise (approximately 250 kHz in this case). The performance measure used is word articulation score which means the understandability of standard word lists by a trained listener panel. Note that impulsive noise is less harmful than Gaussian noise in the sense that the understandability can be maintained at much lower SNR's, although the impulsive interference is quite bothersome (perhaps even painful). Limiting has little effect on Gaussian noise but can be used to further improve analog performance in impulsive noise as also shown in Figure 7. Additional examples for other analog systems are given in [29].

The above general performance characteristics in non-Gaussian noise have been known for many years and are very basic and have all be substantiated many times by measurement. They apply still, however, to modern systems. Even though the characteristics are (or should be) well known, they are still forgotten (or ignored) far too often during current system selection and performance determination activities.

The above results, all using atmospheric noise, are applicable to all Class B interference types. The situation with Class A interference is similar but with different performance chracteristics than Class B, although the general behavior noted above is the same. Figure 8 shows a range of Class A distributions and Figure 9 shows the corresponding performance (calculated) of a binary coherent phase shift keying (CPSK) system in this Class A noise. While the Class A model was designed to "fit" collections of narrowband signals from sources randomly located in space and emitting randomly in time, it is also appropriate for coherent pulse trains (e.g., radars) and appears to, sometimes at least, be appropriate for single interfering signals. Figure 10 shows measured (and in some cases simulated) results [19] of digital systems being interfered with by other systems, sometimes digital (in one case of the same type), and sometimes analog. Note that these P_e results have the same shape as the Class A results of Figure 9. Additional Class A results are also given in the simulation results discussed next.

In the examples above a single number (e.g., P_e) is used to specify performance. A "long-term" average, such as error rate gives a good measure of the performance only if we are dealing with completely stationary interference and signal processes. Since the actual environmental interference and actual signal processes are not stationary, additional information must be given. This usually takes the form of a "time availability" giving the percentage of time a given P_e, say, will be achieved or exceeded along with a statistical confidence factor. The time availability accounts for the change from hour to hour (or some other appropriate time scale) of the interference and signal distributions, while a statistical confidence factor is required to account for errors of prediction of the noise and signal distributions, propagation losses, antenna gains, and the like. Details can be obtained from CCIR Report 413 [20] and examples of specifing "complete" system performance are given in CCIR Report 322-3 [3].

As noted above, the real-world interference is almost never white Gaussian in character, yet typical receiving systems in use are those known to be optimum for Gauss. As we saw, these systems are very suboptimum in the actual non-Gaussian environments. We also noted that Gaussian noise is the worst kind of noise, so that very large improvements in the performance of systems can be achieved if the actual statistical characteristics of the noise and interference are taken into account, and there have been various significant efforts in the last few years in this area [8 - 11, 21 - 23]. (References [21-23] are three recent Ph.D. theses from Princeton University, and as such (along with [8, 9]) contain rather complete sets of additional references.) We want to review the current small signal approach to designing "good" receiving systems, and present some results (theoretical and obtained by simulation) for Class B and Class A interference examples.

When confronted with real-world noise, the earlier and usual approach was to precede the "Gaussian receiver" by various ad hoc nonlinearities (e.g., clipper, hole punchers, hardlimiters, etc.) in order to make the noise look "more Gaussian" to the given receiver. Later, optimum systems were derived using models of the actual noise (e.g., [8, 11]). These systems are adaptive in nature and usually very difficult to realize physically. Because of this, the following two assumptions are usually made:

1. The desired signal becomes "sufficiently" small. "Sufficiently small" is defined in [9, Section A.3-3] and also in Section 6.2.2 of [23].

2. The time-bandwidth product is large, so that a large number, N, of independent samples from the interfering noise process can be used in the detection decision process. This large time-bandwidth product requirement assumes only time waveform sampling. Under appropriate conditions, the required large N can be achieved using joint space-time sampling [24, 25], thereby reducing the required bandwidth.

When the above two conditions (signal sufficiently small and $N\to\infty$) are met, then a "locally optimum detector," usually termed a "locally optimum Bayes detector" or LOBD can be obtained. Under some rather strict conditions, these LOBD's are asymptotically optimum detectors (AOD) and approach true optimality for the non-Gaussian interference processes of concern.

The LOBD detectors are much simpler to construct than the true optimum detector and at a minimum show how to construct "good" suboptimum detectors. In actual use, however, the desired signal may not be always "sufficiently small" and N may not be particularly large. What we seek, in practice, is a detector (by necessity nonlinear), either the LOBD or a good suboptimum detector (perhaps nonparametric), which will perform acceptably well at all signal levels and values of N, approach optimality for small signal levels, and still be economical to construct and use.

The various techniques for driving the LOBD for various signaling situations are well known and covered in detail in [8, 9, 21, and 26, for example]. Here we simply review the results (only for binary CPSK and coherent signal - "pure" detection) in order to indicate where the two assumptions detailed above come into play, both for deriving the detector structure and for determining its limiting performance.

Our problem, for binary CPSK, is to decide optimally between the two hypotheses:

$$\begin{aligned} H_1 &: X(t) = S_1(t) + Z(t) \qquad 0 \le t \le T \\ H_2 &: X(t) = S_2(t) + Z(t) \qquad 0 \le t \le T. \end{aligned} \tag{10}$$

In (10) X(t) is our received waveform in detection time T and this waveform contains either the completely known signal $S_1(t)$ plus the noise Z(t) or the completely known, equi-probable, signal S_2 plus Z(t). To obtain our receiver structure we follow the standard procedure of replacing all waveforms by vectors of N samples from the waveforms ($X(t) \to \underline{X} = \{x_i\}$, etc.,) and forming the likehood ratio $\Lambda(\underline{X})$:

$$\Lambda(\underline{X}) = \frac{p(\underline{X}|H_2)}{p(\underline{X}|H_1)} = \frac{p_Z(\underline{X} - \underline{S}_2)}{p_Z(\underline{X} - \underline{S}_1)} \underset{H_2}{\overset{H_1}{\gtrless}} 1. \tag{11}$$

When Z(t) is non-Gaussian, we operate so as to generate independent noise samples, z_i, $i = 1, N$ in time T, so that only first order pdf's are required. We now use the LOBD or threshold operation which we know becomes asymptotically optimum as our signal S(t) becomes sufficiently small and $N\to\infty$. Increasing N corresponds to increasing the detection time T, since we cannot for any noise process sample more rapidly than the bandwidth and maintain independence. As noted earlier, spatial sampling can sometimes also be used to increase the effective N.

Using a vector Taylor expansion about the signals, $\underline{S}_j$, $j = 1, 2$ here, we get

$$p_Z(\underline{X} - \underline{S}_j) = p_Z(\underline{X}) - \sum_{i=1}^{N} \frac{\partial p_Z(\underline{X})}{\partial x_i} S_{ji} + \frac{1}{2} \sum_{i=1}^{N} \sum_{k=1}^{N} \frac{\partial^2 p_Z(\underline{X})}{\partial x_i \, \partial x_k} S_{ji} S_{jk} + \ldots \tag{12}$$

In this expansion, for coherent signaling, all signal terms of degree two and higher are discarded. This is the normal "small signal assumption." In general, simply discarding higher order terms can lead to receiver structures which are <u>not</u> locally optimum, or in the limit of infinitely large sample sizes ($N\to\infty$), are <u>not</u> asymptotically optimum detection algorithms (AODA's). The proper algorithms require a correct bias (obtainable from proper treatment of the higher order terms). The problem is, that without the proper bias, the higher-order terms in the expansion of $\Lambda(\underline{X})$ can be discarded only when

the sample size N is small. But N must be made large in order to obtain the required small probabilities of error for weak signals. This, of course, defeats the whole concept of a canonical and comparatively simple algorithm. One may as well use $\Lambda(\underline{X})$ itself, which is optimum for all signal levels. Sufficient conditions that the LOBD is an AODA as well as a LOBD ($N<\infty$) are given by Middleton and Spaulding [9] and Weiss and Schwartz [23].

For binary symmetric CPSK, and for independent noise samples (12) leads to

$$\Lambda(\underline{X}) \underset{H_2}{\overset{H_1}{\gtrless}} \frac{1 - \sum_{i=1}^{N} \frac{d}{dx_i} \ln p_Z(x_i) S_{2i}}{1 - \sum_{i=1}^{N} \frac{d}{dx_i} \ln p_Z(x_i) S_{1i}} \underset{H_2}{\overset{H_1}{\gtrless}} 1, \tag{13}$$

which gives the well-known receiver structure shown in Figure 11. In Figure 11, we see that the receiver is the standard memoryless Gaussian (i.e., degenerate matched filter) receiver preceeded by a particular nonlinearity given by

$$\ell(x) = -\frac{d}{dx} \ln p_Z(x). \tag{14}$$

Note that this is a completely canonical result in that we have not yet specified (in the above derivation) what $p_Z(z)$ is or what the signals, $S_1(t)$ $S_2(t)$, are except that they are completely known. Figure 11 is our receiver, which is adaptive in that it must change according to (14) for changing noise conditions. The receiver takes our received waveform samples x_i and uses them as shown to determine our decision variable δ. The pdf of δ is almost always impossible to obtain, however, unless we can invoke the Central Limit Theorem.

Although the nonlinearity $\ell(x)$ does not "Gaussianize" the noise, it does limit the amplitude excursions of the noise. Because of this, it is common to require N to be large (normally N must be relatively large to achieve any kind of processing gain over normal receivers as will be demonstrated later via simulation) so that we can apply the Central Limit Theorem. This means that we only need to compute the mean and variance of δ under each of the two hypotheses. We start with y_i, the output of the nonlinearity for input x_i. Suppose H_1 is true, then

$$E[y_i|H_1] = -\int_{-\infty}^{\infty} \frac{p_Z'(z)}{p_Z(z)} p_Z(z - S_{1i})\, dz.$$

and

$$E[y_i^2|H_1] = \int_{-\infty}^{\infty} \left[\frac{p_Z'(z)}{p_Z(z)}\right]^2 p(z_Z - S_{1i})\, dz. \tag{15}$$

In evaluating the above two integrals, the usual approach is to expand the $p_Z(z - S_{1i})$ and then discard all terms in S_{1i} of degree 2 and higher. This is the second use of the small signal assumption. The result is

$$E[y_i|H_1] = -S_{1i}L, \text{ where} \tag{16}$$

$$L = \int_{-\infty}^{\infty} \frac{[p_Z'(z)]^2}{p_Z(z)} dz, \text{ and} \tag{17}$$

$$E[y_i^2|H_1] = L. \tag{18}$$

The parameter L determines (for "small" signal) the processing gain achievable for any $p_Z(z)$, including Gaussian noise (for which $L = 1$).

Using the above we obtain, for binary symmetric signal with $S_1(t) = -S_2(t)$ (CPSK),

$$E[\delta|H_2] = -E[\delta|H_1] = 2L \sum_{i=i}^{N} S_{1i}^2, \text{ and}$$

$$\text{Var}[\delta|H_2] = \text{Var}[\delta|H_1] = 4 \sum_{i=i}^{N} (LS_{1i}^2 - L^2 S_{1i}^4). \tag{19}$$

An estimate of performance is then given by

$$P_e = \text{Prob}[\delta < o] = \frac{1}{2} \text{erfc} \left\{\frac{|E[\delta]|}{\sqrt{2\text{Var}[\delta]}}\right\} \tag{20}$$

If our two signals are, for example,

$$S_1(t) = \sqrt{2S} \cos(\omega_0 t), \qquad 0 \le t < T$$

and (21)

$$S_2(t) = -\sqrt{2S}\cos(\omega_0 t), \qquad 0 \le t < T$$

so that S is the signal power, then

$$E[\delta] = 2SLN, \text{ and} \tag{22}$$

$$\text{Var}[\delta] = 4SLN - 6S^2L^2N. \tag{23}$$

Since all our noise models are normalized so that the noise power = 1, S is also our signal-to-noise ratio. We note that SL must be such that the variance is positive. Since L is usually large (i.e., $\sim 10^3 - 10^4$), (21) defines, in a sense, the meaning of "small" signal in the above LOBD analysis. If SL << 1, then (20) becomes approximately

$$P_e \cong 1/2\ \text{erfc}\ (\sqrt{SNL/2}). \tag{24}$$

For LOBD's, the performance parameter L is ≥ 1, and is equal to 1 iff the noise is Gaussian.

The above reviews the LOBD approach for the simplest case. Suppose now that we have a LOBD detector based on the assumption that our interference is $\hat{p}_Z(z)$, and the actual interference is $p_Z(z)$. We can carry out the above analysis using $\hat{p}_Z(z)$ in place of $p_Z(z)$ where appropriate to determine the effects of "mismatching" the interference, or we can use this to determine the sensitivity of the LOBD performance to changing interference. This approach also gives results which can be easily used to evaluate the small signal performance of any ad hoc$_2$ nonlinearity. The result is that L is replaced by a parameter L_{eff}, for "L effective," where $L_{eff} = L_1^2/L_2$,

$$L_1 = \int_{-\infty}^{\infty} \left[\frac{\hat{p}'_Z(z)}{\hat{p}_Z(z)}\right] p'_Z(z)dz, \text{ and} \tag{25}$$

$$L_2 = \int_{-\infty}^{\infty} \left[\frac{\hat{p}'_Z(z)}{\hat{p}_Z(z)}\right]^2 p_Z(z)dz. \tag{26}$$

If $\hat{p}_Z(z) = p_Z(z)$, then $L_1 = L_2 = L = L_{eff}$.

We can quickly compute the performance of any arbitrary nonlinearity, g(x), used in the detector of Figure 11. For example, for the hard-limiter, g(x) = 1, if $x \ge 0$ and g(x) = -1, if $x < 0$. We can solve the resulting expression

$$g(x) = -\frac{d}{dx} \ln \hat{p}_Z(x), \tag{27}$$

to obtain the corresponding $\hat{p}_Z(z)$ to compute L_{eff} via (25) and (26) above. For the hard-limiter case, we obtain

$$L_{eff} = 4\,\hat{p}_Z^2(0) \tag{28}$$

where $p_Z(z)$ is the actual interference. Performance is given by (24), with L replaced by L_{eff}, so that the degradation (in the limit) caused by using the hard limiter is simply the difference between L for our actual interference (LOBD performance factor) and L_{eff} for the hard limiter (or similarly, for any-other nonlinearity).

The differential equation in (27) can be solved for any nonlinearity g(x) to determine the noise pdf for which the given g(x) is the locally optimum nonlinearity. For example, for the above hard limiter,

$$\hat{p}_Z(x) = \tfrac{1}{2} e^{-|x|}. \tag{29}$$

The noise process given by (29) is "LaPlace Noise" and is occasionally used although it is not particularly representative of actual noise processes encountered in practice. The result as given in (25) and (27) can be used to determine the robustness of a given detector, based on the LOBD for $p_Z(x)$, when the actual interference process is other than $p_Z(x)$; e.g., a not particularly good estimate of $p_Z(x)$, $\hat{p}_Z(x)$. An example of this for Middleton's Class A model is given in [27].

For any arbitrary nonlinearity, g(x), for g(x) antisymmetric, the L_{eff}, from (25) and (26) is

$$L_{eff} = \frac{\left[\int_{-\infty}^{\infty} g(x)p'_Z(x)dx\right]^2}{\int_{-\infty}^{\infty} g^2(x)p_Z(x)dx}. \tag{30}$$

Where L_{eff} is maximized by the LOBD nonlinearity, $\ell(x)$. For the linear receiver g(x) = x, and from (20), $L_{eff} = 1$ for any pdf $p_Z(x)$ which means the linear receiver (optimum for Gauss) has the same limiting performance for all noise processes. The L_{eff} (30) is the standard measure of the power of

statistical hypothesis tests (in the limit) termed "efficacy" E [28].

In the coherent signal detection (Neyman-Pearson) case, the hypothesis test is

$$\begin{aligned} H_1 &: X(t) = Z(t) \qquad 0 \le t < T \\ H_2 &: X(t) = Z(t) + S(t) \quad 0 \le t < T \end{aligned} \tag{31}$$

where S(t) is our completely known signal. The resulting receiver structure is as in Figure 11, but with $S_{2i} = 0$ and $S_{1i} = S_i$. Also, of course, the threshold is no longer 0, but is a function of the signal level and the specified false alarm probability. The small signal limiting performance result corresponding to (24) for $S(t) = \sqrt{2S}\cos(\omega_o t)$ is:

$$P_D = \frac{1}{2}\,\mathrm{erfc}\left[\mathrm{erfc}^{-1}(2\alpha) - \sqrt{\frac{NSL}{2}}\right], \tag{32}$$

where P_D is the probability of detection and α is the specified false alarm probability. The required threshold, λ is obtained from the signal level S and α by:

$$\lambda = \sqrt{2LNS}\,\mathrm{erfc}^{-1}(2\alpha), \tag{33}$$

and, as before, L is replaced in (32) and (33) by L_{eff} for suboptimum detectors. It is interesting to note that for the problem given by (31), the quantity L (17) is Fisher's measure of information [30].

In the noncoherent frequency shift (NCFSK) case the hypothesis test is:

$$\begin{aligned} H_1 &: X(t) = Z(t) + S_1(t, \phi) \qquad 0 \le t < T \\ H_2 &: X(t) = Z(t) + S_2(t, \phi) \qquad 0 \le t < T, \end{aligned} \tag{34}$$

where ϕ indicates the unknown phase angle. For the two equiprobable signals:

$$\begin{aligned} S_1(t, \phi) &= \sqrt{2S}\cos(\omega_1 t + \phi), \\ S_2(t, \phi) &= \sqrt{2S}\cos(\omega_2 t + \phi), \end{aligned} \tag{35}$$

the small signal limiting performance results corresponding to (15) is:

$$P_e = \frac{1}{2}\exp\left[-\frac{NLS}{2}\right]. \tag{36}$$

The locally optimum receiver for the noncoherent case (34) is the standard quadrature receiver preceded by the nonlinearity $\ell(x)$. As an example, Figure 12 shows theoretical results for coherent signal detection (31), using (32), for a sample Class A distribution.

The above approach of determining and implementing $\ell(x)$, or a good approximation g(x), is the parametric approach. Another approach is the nonparametric approach, that is, attempting to obtain a good detector that does not depend on the particular noise distribution but which will perform well (hopefully, near optimum) for a range of noise distributions of concern. There has been substantial effort in this area as recently summarized by Kassam and Poor [31]. In order to compare the limiting performance of nonparametric detection with each other and with parametric detectors, the efficacy corresponding to (30), but for nonparametric detection must be determined. Nonparametric detectors, except for the hard limiter, have to have "g(x)" to use in (30).

For any detector, D, the efficacy, E_D, for the test (31), is given by [28]

$$E_D = \lim_{N\to\infty} \frac{\left[\frac{\partial}{\partial\theta} E_{H_2}[\delta]\Big|_{S\to 0}\right]^2}{N\mathrm{Var}_{H_1}[\delta]} \tag{37}$$

where $E_{H_2}[\delta]$ and $\mathrm{Var}_{H_1}[\delta]$ are the expectation under H_2 and the variance under H_1 of the test statistic δ. The above concept of efficacy requires that a number of regularity conditions be met. The definition (37) leads to L_{eff} (30) for parametric detectors. For example, Gibson and Melsa [28] give a lengthy derivation, using (37) of the result (28) for the hard limiter (sign detector). Comparison of two detectors, 1 and 2, is almost always given by the asymptotic relative efficiency ARE, where

$$\mathrm{ARE}_{1,2} = \frac{E_{D_2}}{E_{D_1}}. \tag{38}$$

The ARE is a limiting result and can be very misleading for actual systems. What is required is the "relative efficiency," which is almost always impossible to obtain analytically or even estimate. As noted earlier, E_D for the hard limiter is given by $4p_Z^2(0)$ and for the linear receiver, $E_D = 1$ for all $p_Z(z)$.

The sign detector bases its decision on the likelihood that if a positive signal is received we will have more positive values than negative values in our sample vector $\{x_i, i = 1, N\}$. Information that is not used is the likelihood that these positive values will be larger than the negative values.

The amplitude data that are discarded by the sign detector are utilized by detectors that process both polarity and rank information. Rank detectors are based on rank statistics or order statistics. Nonparametric detectors other than the sign detector are rank detectors. The individual observations x_i of the input process are stored until all N samples have been taken, and the observations are then ranked in order of increasing absolute value

$$|x_{(1)}| < |x_{(2)}| < \cdots < |x_{(n)}|, \tag{39}$$

where we use $\{x_{(i)}\}$ to denote the order statistics for $\{x_i\}$, i.e., the x_i's rearranged according to (39). For the problem (31), detectors which use this rank information have test statistics of the general form

$$\delta = \sum_{i=1}^{N} f(R_i)\, u\,(x_i). \tag{40}$$

where R_i is the rank of the ith observation, $f(R_i)$ is some function of these ranks, and $u(x_i)$ is the unit step function. Of course, for the binary symmetric case we have

$$\delta = \sum_{i=1}^{N} f(R_i)\,\mathrm{sgn}(x_i). \tag{41}$$

As before, for (40), for coherent signals, δ is compared to a threshold λ which is a function of the signal level and the false alarm probability α and for (41), for coherent binary symmetric signals, δ is compared to zero. Two possible nonparametric detectors that should have better performance are the Wilcoxon detection for which

$$f(R_i) = R_i, \tag{42}$$

and the Sperman Rho detector for which

$$f(R_i) = iR_i. \tag{43}$$

These are perhaps the detectors which are the easiest to implement [after the hard limiter, $f(R_i) = 1$], especially the Wilcoxon detector [28]. Use of (37), gives the efficacy of the Wilcoxon detector as

$$E_W = 12 \left[\int_{-\infty}^{\infty} \left[p_Z(x)\right]^2 dx\right]^2. \tag{44}$$

As noted earlier, we can also attempt to develop a good suboptimum nonlinearity. The efficacy (37) or (30) is maximized by $\ell(x)$. Czarnecki and Thomas [22] have proven the theorem that if we have a set of candidate nonlinearities G the solution to the problem of finding the $g^* \in G$ which maximizes the efficacy is the same as the solution to the mean square estimation problem

$$\min_{g \in G} \int [g(x) - \ell(x)]^2 p_Z(x) dx. \tag{45}$$

Note that in (45), the square errors are weighted by the noise density so that points in the tail region are weighted much less heavily than those near the origin. The results (45) lend weight to intuitive ideas about what constitutes a "good" nonlinearity. It is important to match $\ell(x)$ closely in the regions where an observation is highly likely, while rougher approximations will suffice in the low probability regions. Also, once the approximation is fairly "close" to $\ell(x)$, further refinements gain little. However, as we will see in the examples given next, any nonlinear tail behavior will not suffice.

All of the above performance estimates (e.g., Figure 12) are based on N being large ($+\infty$) and S "sufficiently small." In actual situations, both of these conditions may not be true. The problem has been to use the above to develop detectors that do approach optimality for small signals, but still perform acceptably well for small N and any signal level. The example given next will illustrate this problem, especially for Class A interference.

SIMULATION RESULTS

Figure 12 showed samples of theoretical results for the Class A example, A = 3.5 and $\Gamma' = 0.5 \times 10^{-3}$. In general, for these Class A parameters, L (17) is 1340 (31.3 dB), which means that the optimum system (using $\ell(x)$) will be 31.3 dB superior to the linear (Gauss) system; i.e., for example, be able to detect a signal 31.3 dB smaller. These are, as stressed earlier, limiting results. Here, we want to show a few simulation examples to indicate how actual systems will perform.

The receiver of Figure 11 has been implemented on a large scale computer. Monte Carlo simulation results for the linear receiver, the bandpass limiter receiver, and the LOBD (bandpass) receiver are given here. Some results for the Wilcoxon detector and some "good" suboptimum nonlinearities are also included. By "bandpass" we mean that the nonlinearity acts on the received complex (magnitude and phase) sample of the received waveform. For example, the bandpass limiter nonlinearity is $X_i/|X_i|$, where X_i is the complex waveform sample (Figure 11). The simulation results for Class B noise (Hall, θ = 2,4) were given earlier [32] and a few of the results are repeated in order to contrast them with the Class A results.

The Class A example used for Figure 12 is used throughout, and Figure 13 shows the $\ell(x)$'s for the two Hall noise examples and the Class A example. The Class A example on Figure 13 includes the nonlinearity actually used (denoted by A = 0.4) based on the approximation used for Middleton's Class A

model. Note that the Class A $\ell(x)$ looks a lot like a "hole puncher."

Figure 14 shows simulation results for Hall noise, $\theta = 2$, normalized to represent Middleton's model. First note the interesting results for the linear receiver. Identical results were obtained for $N = 1, 10, 100$. This is, of course, not physically meaningful and is the result of using a model for which the moments do not exist. This "infinite power" problem does not exist whenever a nonlinearity is employed as with the other results of Figure 14. For a linear receiver, for $N = 10$, say, detection is based on a "noise sample" that is the sum of the ten noise samples from the underlying distribution. Except for Gaussian noise, the distribution of the "sum sample" is different from the distribution of each individual sample, and approaches a Gaussian distribution via the Central Limit Theorem. This makes it difficult to analytically determine the performance of linear systems in non-Gaussian interference for a time-bandwidth products other than 1. Also indicated on Figure 14 for $N = 1$ are results using the LOBD nonlinearity and the bandpass limiter. As indicated, these results were essentially identical to those obtained for the linear receiver, demonstrating the known result that for $N = 1$, no improvement can be obtained by using nonlinear receivers and in order for nonlinearities to be effective, N must be greater than 1. On Figure 14, the performance of the linear receiver for $N = 100$ is estimated. This is based on the parameter L (37 dB) and the simulation results for $N = 100$ for the LOBD receiver. [As discussed above, L is equal to the ARE (asymptotic relative efficiency) which gives the limiting ($N \to \infty$) performance difference between the LOB detector and the linear detector.] For $N = 100$, the simulation results match the analytical results (24) and (28). Computation of the ARE (38) to compare the hard limiter with the LOBD, for this noise example, indicates that the hardlimiter is only 1.5 dB inferior to the LOBD and the simulation results ($N = 100$) show this. In fact, computations indicate that the hard limiter is always, for the range of Class B parameters likely to be encountered in practice, within 2 dB of the LOBD. Note also, however, for $N = 10$ (Figure 14) the bandpass limiter begins to out perform the LOBD for larger signals.

Figure 15 shows simulation results for the Hall model, $\theta = 4$. First note that as before, use of nonlinearities for $N = 1$ gives no improvement over the linear receiver, but, of course, does give improvement for $N = 10$ and 100. For $N = 100$, the improvement is only 6 dB as predicted by L. Note that the LOBD nonlinearity here also is only slightly superior to the bandpass limiter. For $\theta = 4$, the moments (first three) exist (i.e., finite energy) and we obtain "normal" results for the linear receiver for $N = 1, 10, 100$. On Figure 15, a SNR of approximately -20 dB is required for P_e of 10^{-3} ($N = 100$), whereas from Figure 14, a SNR of approximately -53 dB is required for P_e of 10^{-3}. This difference was indicated by the two L values. Both distributions ($\theta = 2$ and 4) "look" highly non-Gaussian. This shows that we cannot arbitrarily say, by inspection, that a noise process that is "tremendously" non-Gaussian can result in "tremendous" improvement over the corresponding Gaussian or linear receiver situation.

We now turn to Figure 16 and Class A simulation results, using the same Class A example used for Figure 12. First, for the linear receiver, we obtain "normal" results for $N = 1, 10, 100$ with the $N = 1$ results matching analytical results as indicated by the solid curve on Figure 16. As expected, the nonlinearity results for $N = 1$ were essentially identical to the linear receiver results. Note, first, that for $N = 10$, the bandpass limiter becomes quite inferior to the LOBD. However, also note that as the SNR increases past approximately -27 dB, the LOBD performance degrades very rapidly and becomes inferior to even the linear receiver. From Figure 13, the Class A nonlinearity looks quite similar to a "hole puncher" nonlinearity with the punching level at approximately -27 dB or so. Therefore, the performance characteristic of very rapid deterioration above about -27 dB is not surprising. Also shown on Figure 16, for $N = 10$, are results for the Wilcoxon detector, which is far inferior to the bandpass limiter, and an "adaptive limiter" which follows the LOBD nonlinearity to its peak (--27 dB) and then limits at this level rather than "hole punching." This adaptive limiter performs no better than the bandpass limiter. Figure 17 shows additional results, mainly for $N = 100$. Note that the LOBD for $N = 100$ performs as theory (24, L = 31.3 dB) indicates for small signal, but, again, has quite poor performance as the SNR increases past approximately -25 dB. This dramatically illustrates that the "small enough" signal requirement that gives the LOBD must be met, at least for some cases, before any kind of "good" performance can be expected. The small signal part of the LOBD performance curve and the performance curve for the linear receiver for $N = 100$ differ by approximately 31 dB as given by the parameter L. Also shown on Figure 17 are the results for $N = 10$ for an adaptive hole puncher, which follows the LOBD, but "punches" at the - 27 dB level. This nonlinearity, of course, displays the same objectionable behavior as the LOBD. Results for $N = 100$, for the bandpass limiter, the adaptive limiter, and the Wilcoxon detector are also given in Figure 18. Calculations of the ARE show the hard limiter to be 3 dB inferior to the LOBD and the Wilcoxon detector (using (44)) to be 5 dB inferior to the LOBD. The simulation results for $N = 100$ confirm these results. Attempts to find detections that will give satisfactory results for Class A noise, small N, and any signal level are continuing.

CONCLUSIONS

We have attempted to summarize, in a general way, the kinds of performance observed for systems in the real-world interference environment, pointing out that all too often, the effects of the real environment are initially ignored, leading to very expensive fixes being required later on. Techniques for determining the performance of "normal" systems in the real environment are now well known and documented. We also have reviewed current attempts to take advantage of the real environment. When the interference is correctly accounted for, quite large processing gains can usually be realized. The techniques reviewed, or at least alluded to, have application in modern systems (e.g., spread spectrum). Finally, the interference models and the signal processing consequent upon them, are not confined to EM telecommunications alone. They are applicable to other domains and other physical media and propagation mechanisms (e.g., underwater acoustics [24] and remote sensing).

REFERENCES

[1] A. D. Spaulding, Stochastic modeling of the electromagnetic interference environment, Conf. Record, International Conf. on Communications, ICC '77, Chicago, June 12-15, 1977, pp. 42.4-114-123 (IEEE Catalog No. 77CH-1209-GCSCB).

[2] A. D. Spaulding, Atmospheric noise and its effects on telecommunication system performance, Chapter 7, Handbook of Atmospherics, Vol. 1, H. Volland, Ed., 1982 (CRC Press, 2000 N. W. 24th St., Boca Raton, Florida, 33431).

[3] Characteristics and applications of atmospheric radio noise data, CCIR Report 322-3, International Radio Consultative Committee, International Telecommunications Union, Geneva, 1986.

[4] A. D. Spaulding and J. S. Washburn, Atmospheric radio noise: Worldwide levels and other characteristics, NTIA Report 85-173, April 1975 (NTIS Order No. PB85-21942).

[5] J. R. Herman and X. A. DeAngelis, Bandwidth expansion effects on the voltage deviation parameter (V_d) of MF and HF atmospheric radio noise, Radio Science, Vol. 11, No. 1, pp. 26-36, Jan.-Feb. 1987.

[6] D. Middleton, Statistical-physical models of electromagnetic interference, IEEE Trans. Electromagn. Compat., EMC-19, August 1977, pp. 106-127.

[7] D. Middleton, Canonical non-Gaussian noise models: Their implications for measurement and prediction of receiver performance, IEEE Trans. Electromagn. Compat., EMC-21, Aug. 1979, pp. 209-220.

[8] A. D. Spaulding and D. Middleton, Optimum reception in an impulsive interference environment, Part I: Coherent detection, Part II: Incoherent reception, IEEE Trans. Comm. COM-25, Sept. 1977, pp. 910-934.

[9] D. Middleton and A. D. Spaulding, Optimum reception in non-Gaussian electromagnetic interference environments: II Optimum and suboptimum threshold signal detection in Class A and B noise, NTIA Report 83-120, May 1983, (NTIS Order No. PB 83-241141).

[10] D. Middleton and A. D. Spaulding, A tutorial review of elements of weak signal detection in non-Gaussian EMI environments, NTIA Report 86-194, May 1986 (NTIS Order No. PB 86-216496/AS).

[11] H. M. Hall, A new model of "impulsive" phenomena: Application to atmospheric noise communications channels, Electron. Lab., Stanford Univ., Stanford, CA, Tech Rep. 3412-8 and 7050-7, SU-SEL-66, 052 Aug. 1966.

[12] R. Farg and O. Shimbo, Unified analysis of a class of digital systems in additive noise and interference, IEEE Trans. on Commun., Vol. COM-21, No. 10, Oct. 1973, pp. 1075-1091.

[13] A. D. Spaulding, R. T. Disney and A. G. Hubbard, Man-made radio noise, Part II, Bibliography of measurement data, applications and measurement methods, Office of Telecommunications Report OT 75-63, May 1975, (NTIS Order No. COM 75-10799/AS).

[14] International Union of Radio Science, Review of Radio Science, 1972-1974, 1974-1977, 1977-1980, 1980-1983, 1983-1986, URSI, Avenue Albert Lancaster 32, B-1180, Brussels, Belgium.

[15] C. C. Watterson, and C. M. Minister, HF channel-simulator measurements and performance analysis on the the USCIO, ACQ6, and MX190 PSK modems, Office of Telecommunications Report 75-56, 1975, (NTIS Order No. COM75-11206/AS).

[16] G. F. Montgomery, Comparison of amplitude and angle modulation for narrow-band communications of binary coded messages in fluctuation noise, Proc. Inst. Radio Eng., 42, pp. 447, 1954.

[17] A. M. Corda, The effect of atmospheric noise on the probability of error for an NCFSK system, IEEE Trans. Commun. Technol., 13 (3), pp. 280, 1965.

[18] J. H. Halton and A. D. Spaulding, Error rates in differentially coherent phase systems in non-Gaussian noise, IEEE Trans. Commun. Technol. COM-124, No. 5, pp. 594-601, 1966.

[19] Communications/electronics receiver performance degradation handbook. The Office of Spectrum Management, National Telecommunications and Information Administration, U. S. Department of Commerce and the Electromagnetic Compatibility Analysis Center (ECAC), ESD-TR-73-014 (NTIS Order No. AD 764710).

[20] Operating noise-threshold of a radio receiving system, CCIR Report 413, International Radio Consultative Committee, International Telecommunications Union, Geneva, 1966.

[21] E. J. Modugno III and J. B. Thomas, The detection of signals in impulsive noise, Information Sciences and Systems Laboratory, Report Number 13, June 1983, Princeton University, Princeton, NJ 08544.

[22] S. V. Czarnecki and J. B. Thomas, Nearly optimal detection of signals in non-Gaussian noise, Information Sciences and Systems Laboratory, Report Number 14, February 1984, Princeton University, Princeton, NJ 08544.

[23] M. Weiss and S. C. Schwartz, Robust and nonparametric detection of fading narrowband signals, Information Sciences and Systems Laboratory Report Number 19, January, 1986, Princeton University, Princeton, NJ 08544.

[24] D. Middleton, Multiple-element threshold signal detection of underwater acoustic signals in non-Gaussian interference environments, Naval Ocean Systems Center Contractor Report 231, May 18, 1984, NOSC, San Diego, CA 92152.

[25] D. Middleton, Space-time processing for weak signal detection in non-Gaussian and nonuniform electromagnetic interference (EMI) fields, NTIA Contractor Report 86-36, Feb. 1986 (NTIS Order No. PB86-193406).

[26] Luciano Izzo and Luigi Paura, Asymptotically optimum space - diversity detection in non-Gaussian noise, IEEE Trans. on Commun., Vol. COM-34, No. 2, pp. 97-103, Feb. 1986.

[27] A. D. Spaulding, The robustness of the locally optimum detectors for non-Gaussian noise, Proceedings of the Fifth International Wroclaw Symposium on EMC, pp. 143-152, September 17-19, 1980.

[28] J. D. Gibson and J. L. Melsa, Introduction to Nonparametric Detection with Applications, Mathematics in Science and Engineering, Vol. 119, 1975, Academic Press, New York.

[29] A. D. Spaulding, Voice communication systems performance in the presence of automotive ignition noise, IEEE Trans. on EMC, Vol. EMC-24, No. 3, August 1982, pp. 344-348.

[30] P. J. Bickel and K. A. Doksum, Mathematical Statistics, Holden-Day, Inc., San Francisco, CA 1977.

[31] S. A. Kassam and H. B. Poor, Robust techniques for signal processing - a survey, Proc. IEEE, Vol. 73, No. 3, March 1985, pp. 433-481.

[32] A. D. Spaulding, Locally optimum and suboptimum detection performance in a non-Gaussian interference environment, IEEE Trans. on Commun. Vol. COM-33, No. 6, pp. 509-517, June 1985.

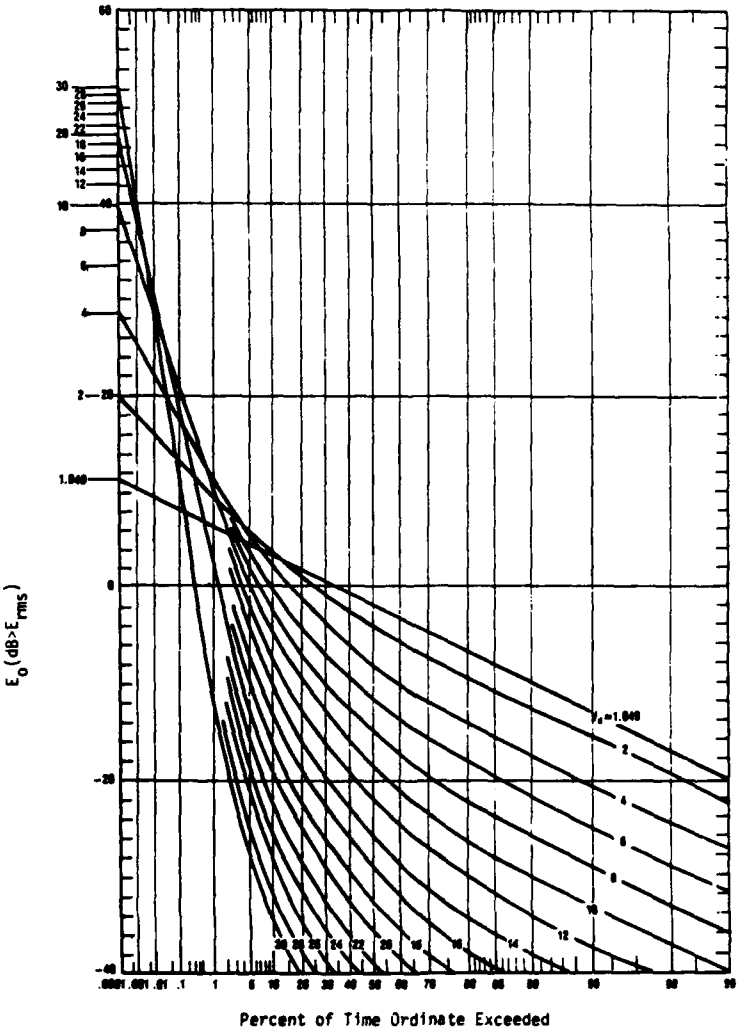


Figure 1. CCIR Report 322 set of amplitude probability distributions of atmospheric radio noise for various V_d values.

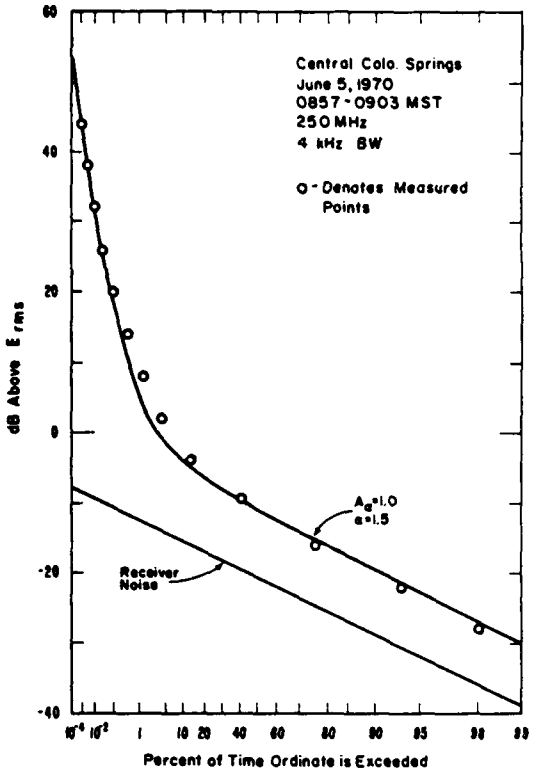


Figure 2. Comparison of measured envelope distribution of manmade noise (primarily automotive ignition noise) with the Class B model.

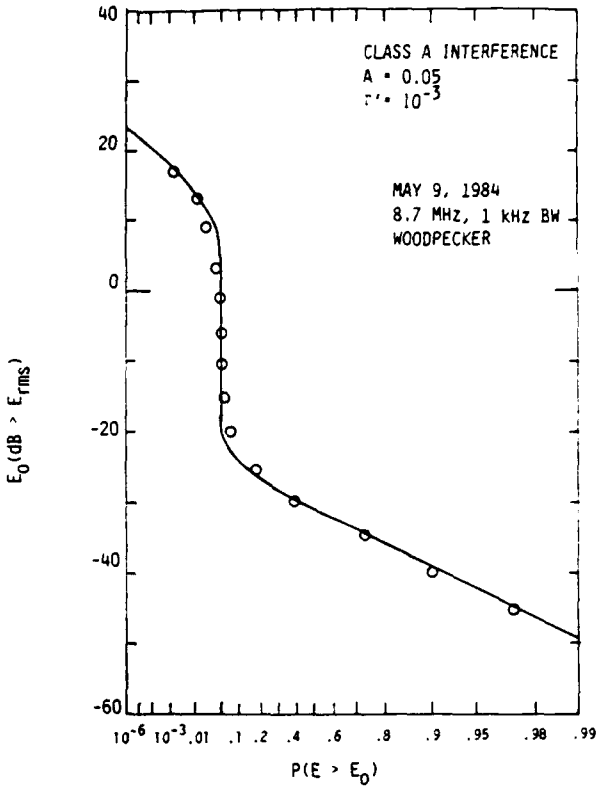


Figure 3. Comparison of measured envelope distribution with the Class A model. Interference is Soviet high power over-the-horizon radar (coherent pulses).

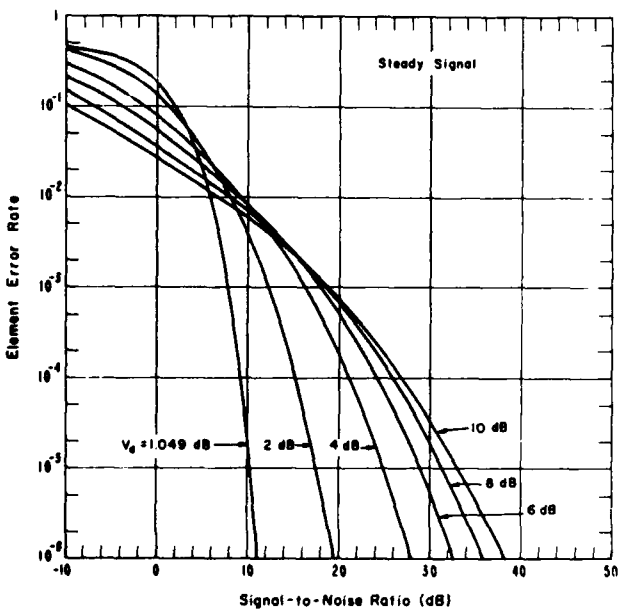


Figure 4. Performance of a binary NCFSK system, steady signal, in atmospheric radio noise.

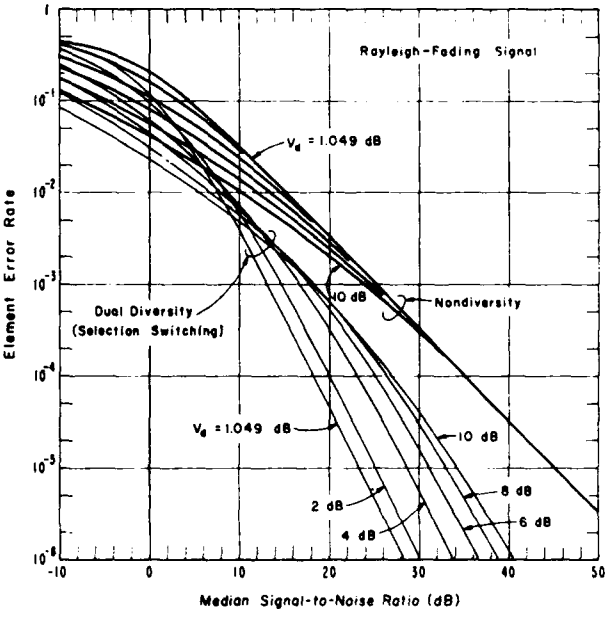


Figure 5. Performance of a binary NCFSK system, Rayleigh flat fading signal, in atmospheric radio noise.

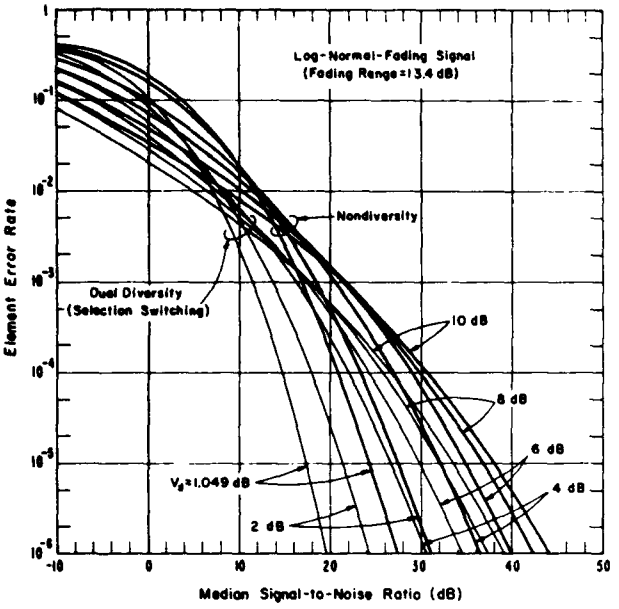


Figure 6. Performance of a binary NCFSK system log-normal fading signal, in atmospheric radio noise.

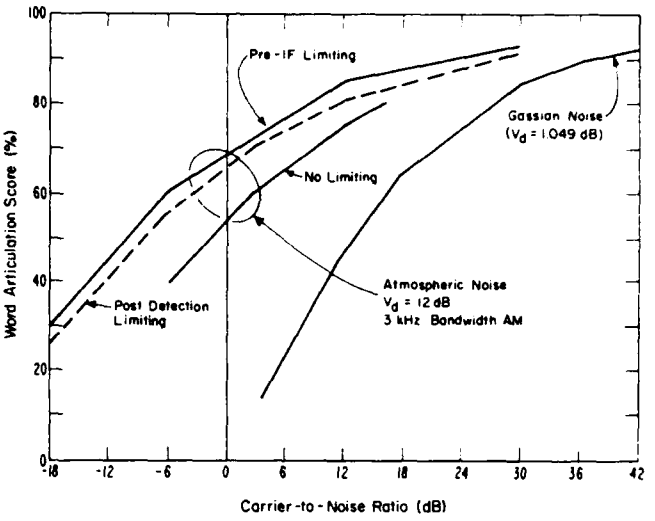


Figure 7. Comparison of the performance of AM noise system in Gaussian and atmospheric noise (V_d = 12).

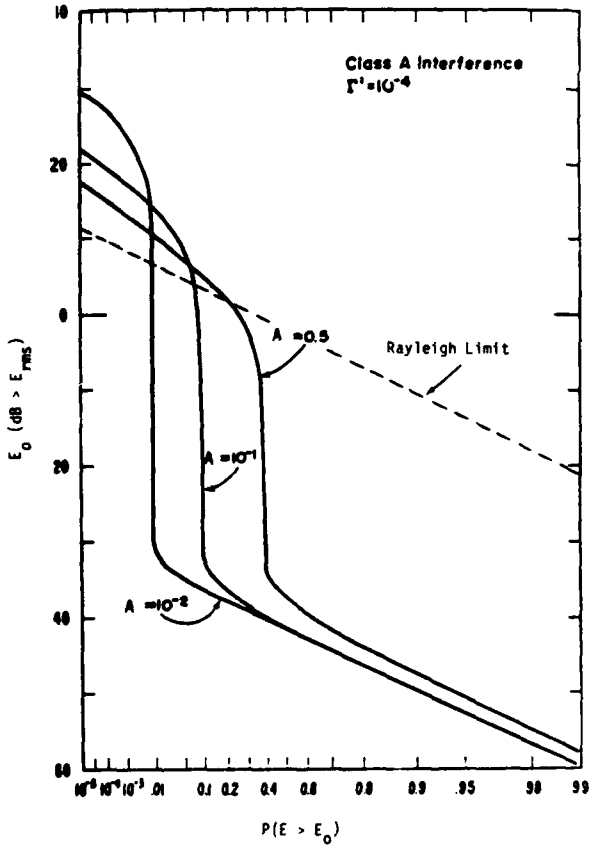


Figure 8. The envelope distribution for Class A interference for $\Gamma' = 10^{-4}$ and various A.

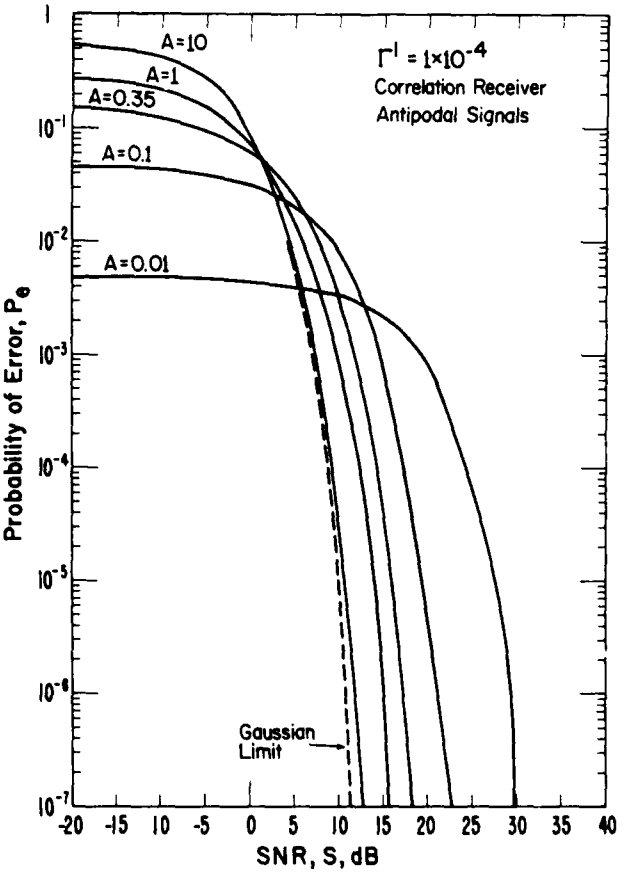


Figure 9. Performance of the correlation receiver for binary CPSK in Class A interference for $\Gamma' = 10^{-4}$ for various values of A.

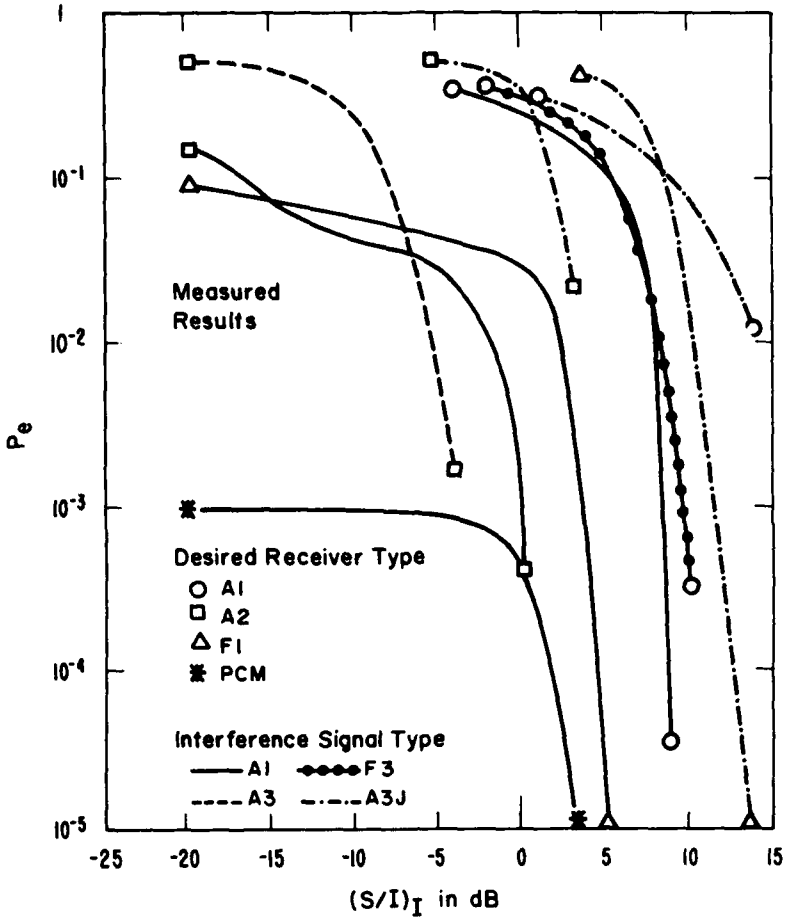


Figure 10. Measured performance of various receivers with various types of interfering signal as a function of input desired to interfering signal ratio.

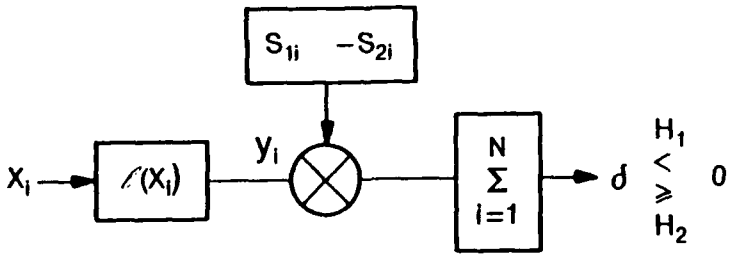


Figure 11. LOBD for binary symmetric coherent signals.

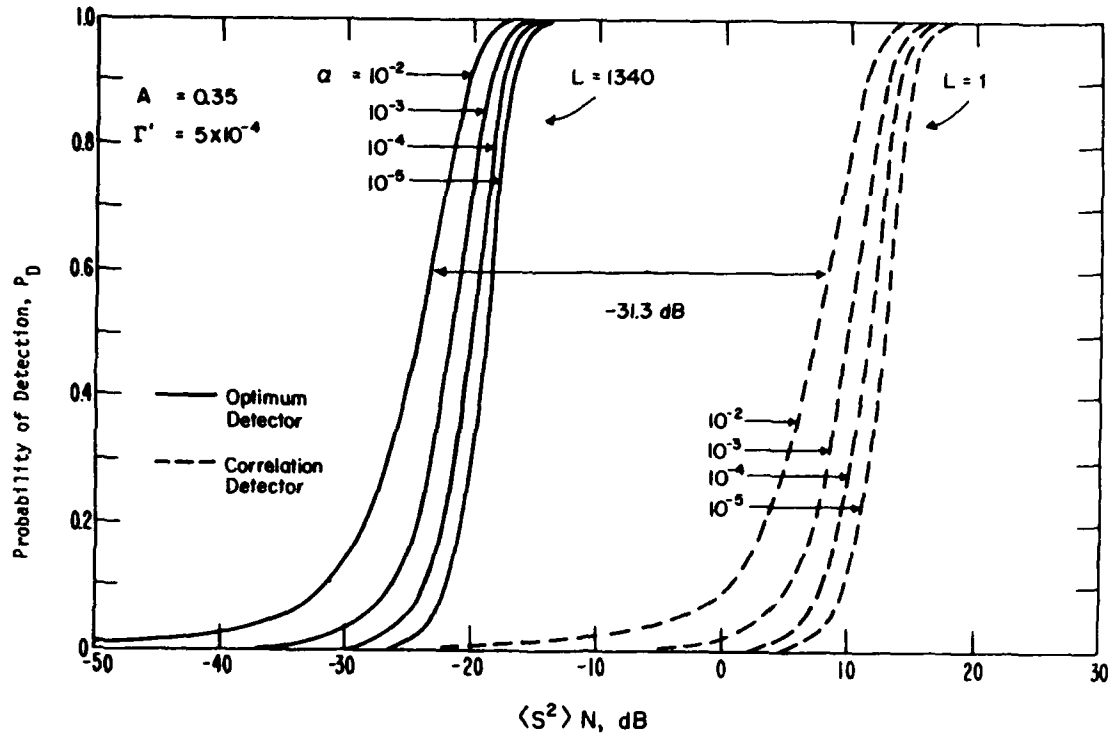


Figure 12. Probability of signal detection P_D for threshold coherent detection with fixed false alarm probability α for optimum and correlation receivers in a sample of Class A noise.

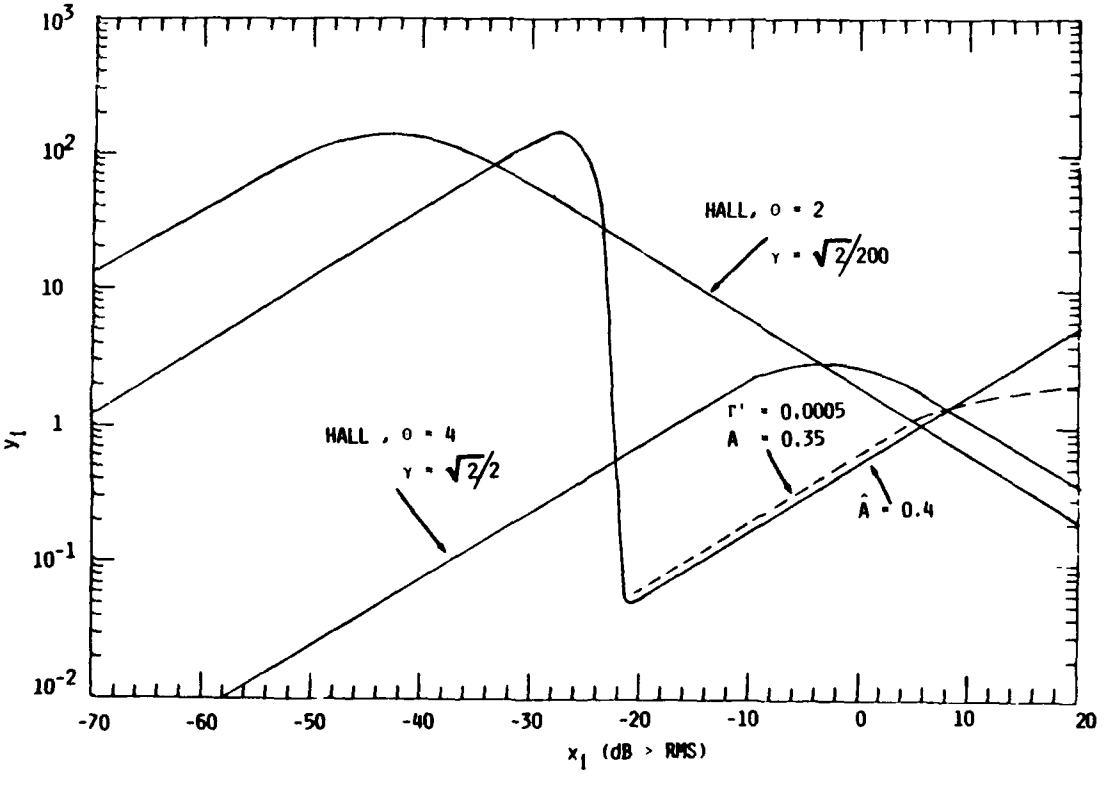


Figure 13. The LOBD nonlinearities.

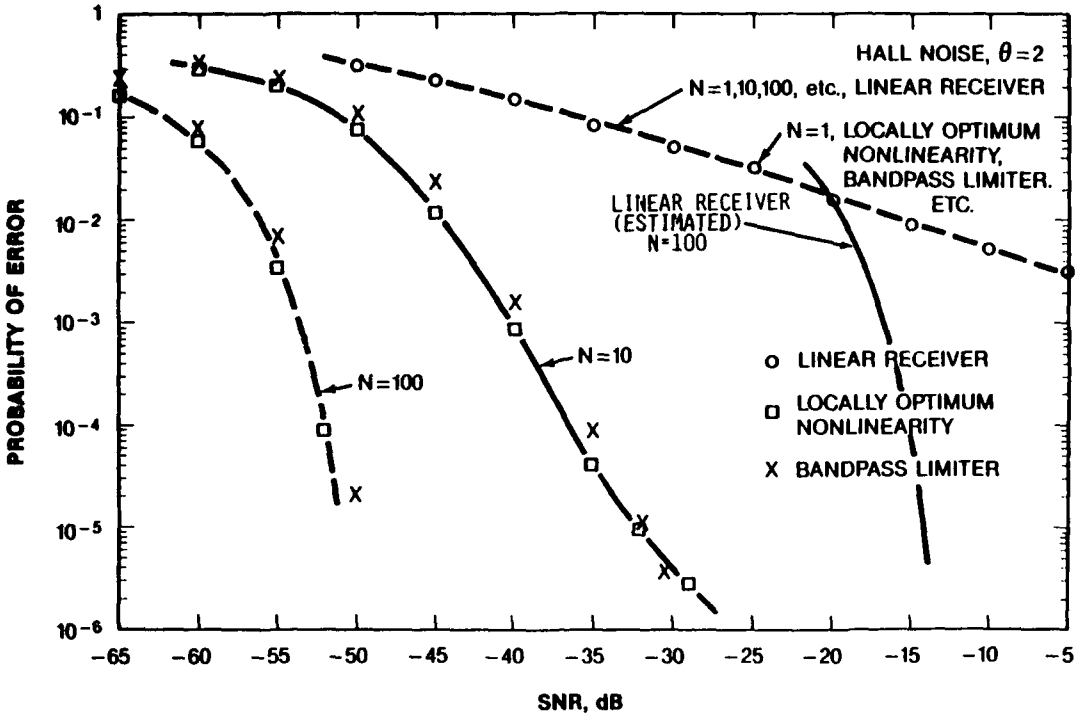


Figure 14. Simulation results with Hall noise, $\theta = 2$, and constant signal, binary CPSK, for a linear receiver and for the LOBD and bandpass limiter nonlinearities.

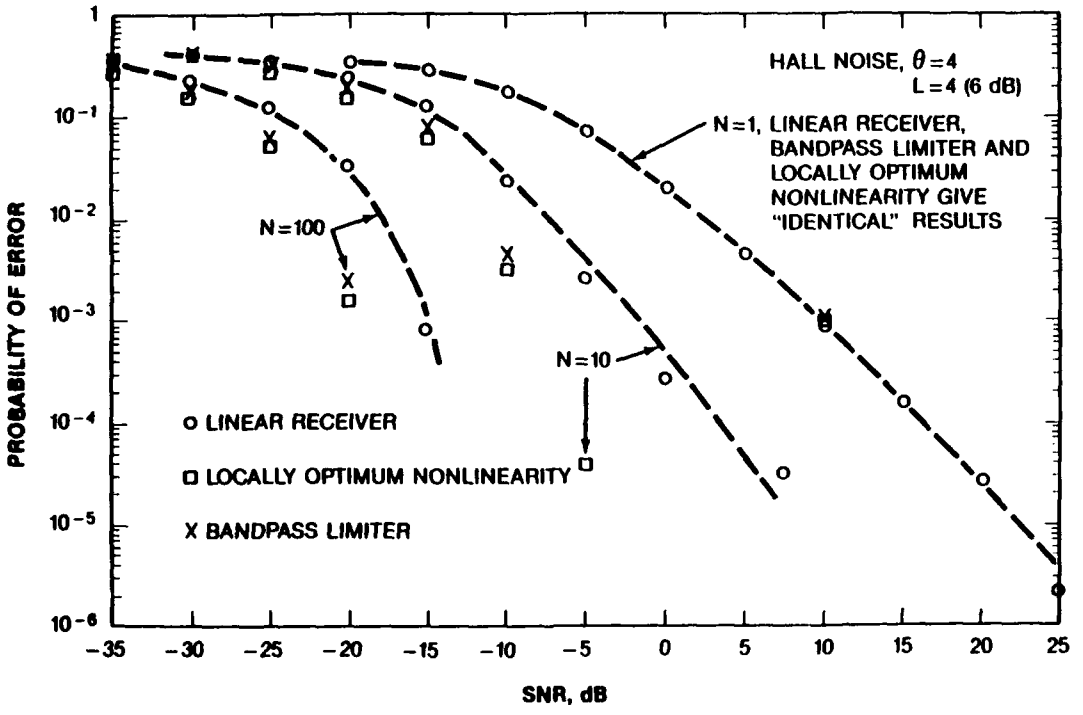


Figure 15. Simulation results with Hall noise, $\theta = 4$, and constant signal, binary CPSK for a linear receiver and for the LOBD and bandpass limiter nonlinearities.

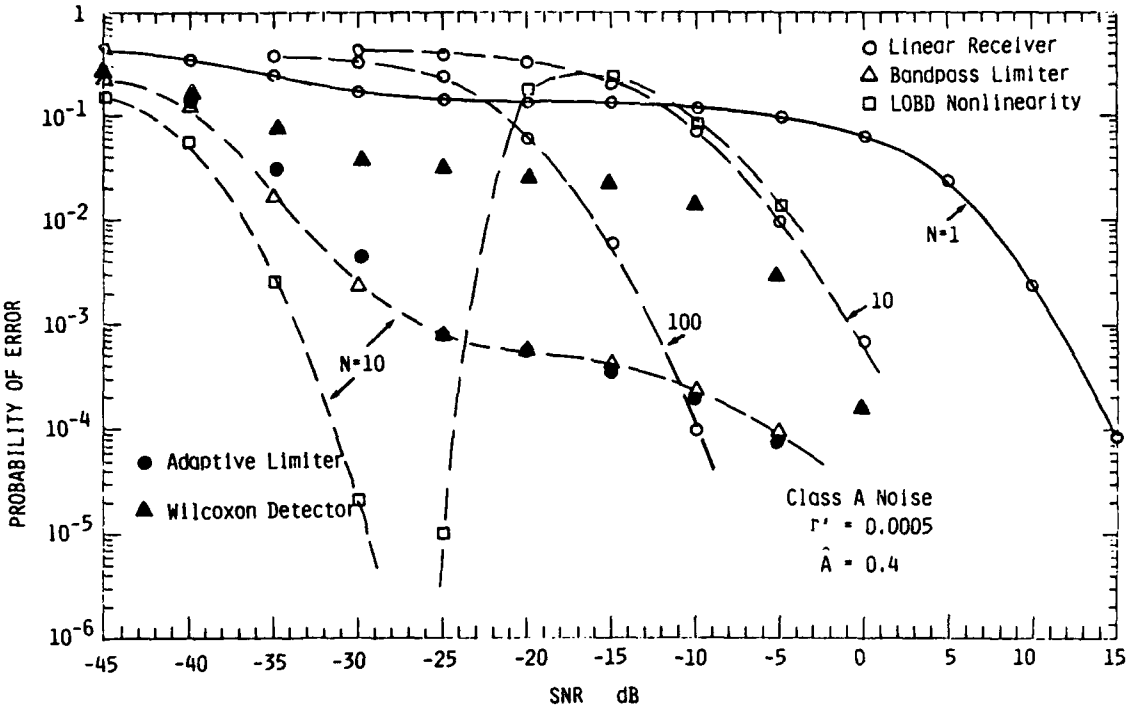


Figure 16. Simulation results with Class A interference, Binary CPSK and constant signal, for a linear receiver, various nonlinearities and nonparametric detectors.

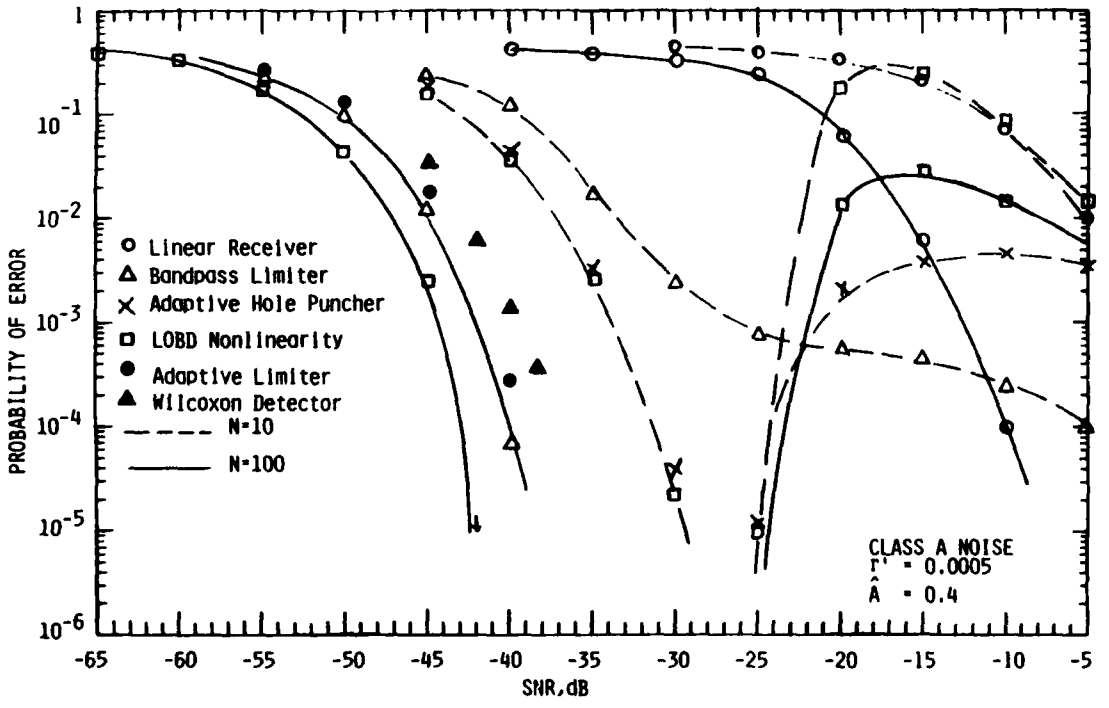


Figure 17. Simulation results with Class A interference, N = 10 and 100, binary CPSK and constant signal, for a linear receiver, various nonlinearities and nonparametric detectors.

THE DESIGN AND PERFORMANCE OF AN ADAPTIVE PACKET-SWITCHED HF DATA TERMINAL

by

G.R. Nourry
Communications Research Centre
Box 11490, Station H
Ottawa, Ontario, Canada K2H 8S2

and

A.J. Mackie
Intellitech Canada Ltd
81 Metcalfe Street, Suite 400
Ottawa, Ontario, Canada K1P 6K7

SUMMARY

Adaptive techniques form a useful set of tools for combatting the time and space variability of HF channels and the other factors adversely affecting HF communications. These techniques have been implemented together with packet-switching in a vehicle called the HF data terminal.

The terminal is characterized by adaptivity at all levels of its design, including a real-time channel evaluation and channel selection mechanism, an adaptive link protocol for channel optimization, and a fully distributed and adaptive routing algorithm for the selection of routes in an HF network. The adaptivity is implemented via a suite of robust, low-overhead, adaptive protocols that work even in low-bandwidth and error-prone environments. With the exception of channel evaluation/selection, all forms of adaptivity are dynamic and not negotiated over HF links.

The data terminal has been tested over both short and long HF links.The performance of this system is illustrated with results from these tests. It is shown that, in addition to fully automatic operation, the system provides a significant improvement in HF communications reliability and survivability. Throughput was demonstrated to be an order of magnitude greater than in existing HF communications systems.

1.0 INTRODUCTION

High-frequency (HF) radio represents a means of low-cost and long range communication. Some of the most significant attributes of the high-frequency radio channel are its inherent survivability, its ability to provide communications in areas where other communications means are non-existent, impossible and/or impractical to use, and its ability to provide reliable long distance communications via skywave propagation.

This unique combination of features is associated with communication problems, at times severe, due to the characteristics of the communication path and its variations in time and space. Frequencies which will propagate over a given path and the quality of communications depend strongly upon ionospheric conditions that are difficult to predict on an hourly or daily basis. The channel capacity exhibits large and largely unpredictable variations. Communications are affected by fading, co-channel interference and noise, and the signal exhibits dispersion and distortion.

HF radio communications offers significant operational advantages but these are offset by the difficulty in solving the propagation problems. The current HF communications methodology (e.g. the "two-frequency" operating schedule based upon propagation predictions) often aggravates the problems, and leads to a reduction in communications effectiveness. As a result, HF communications are often unreliable and the data traffic is limited to 75 bps or less in many systems.

Solutions must be found to improve HF communications in the areas of connectivity, speed, reliability and survivability in order to meet current and emerging requirements. Recently, much emphasis has been placed on the use of real-time channel evaluation (RTCE) techniques to improve such communications. Although very significant, RTCE is but one step toward better HF communications.

This paper discusses the main ideas in a different approach to HF communications - an adaptive-system approach. It describes the basic features of a system, the adaptive HF data terminal, based upon this approach. The performance of this terminal is illustrated with results from on-the-air experiments.

2.0 THE ADAPTIVE-SYSTEM APPROACH

The key to more reliable HF communications has long been limited to the use of more transmitted power, more powerful forward error-correcting code (FEC), repetition of messages and better HF modems. These methods of achieving better communications are not always permissible, practical, or affordable. For example, low probability of intercept (LPI) requires that transmitted power levels be kept as low as possible; powerful FECs and repetition of messages can be wasteful of bandwidth.

Since the resurgence of interest in HF and with the advent of powerful and inexpensive microprocessors and digital signal processors, significant equipment and systems advances have been made. Improved HF modems have contributed significantly to more reliable HF communications. However, good HF modems alone are not sufficient to guarantee a high degree of availability and reliability in HF communications. Real-time channel evaluation (RTCE) [e.g. 1-9] has emerged as a valuable tool in the selection of good channels to improve the performance of HF communication systems and, can now be found, in some form, in new commercially available systems and systems in development.

The adaptive-system approach goes one step further. It consists of trying to respond, in real time, to any changes in propagation conditions. A system based upon this approach makes extensive use of adaptive techniques

to improve on link or network connectivity and performance. It also optimizes, in real time, its operation to suit the prevailing conditions. This approach is not incompatible with the effort in developing better radio equipment and modems. The adaptive system can make use of them when they become available.

In order to improve HF communications as much as possible, the adaptive system should try to incorporate proven techniques such as time, frequency and space diversity, forward error correction, etc., in a manner that is not wasteful of resources (e.g. bandwidth, equipment, real estate, etc.) under any conditions.

Since the choice of adaptive techniques is wide [e.g. 10,11], a system must limit itself to a subset of techniques. A careful choice must be made to obtain the best possible improvement in HF communications performance as expressed by connectivity, speed, throughput, reliability and survivability.

3.0 THE HF DATA TERMINAL

3.1 Four characteristics of HF communications

In addition to the large and largely unpredictable time and space variations of the propagation medium characteristics and the associated variations in channel capacity, four general characteristics of HF communications are of utmost importance in the design of a data terminal. They are:

a) available low-bandwidth
b) error-prone environment
c) multicast nature of HF transmissions
d) non-reciprocity of HF communications

The effective bandwidth available on HF is variable and the current state-of-the-art is 2400 bps for digital communications. The probability of a bit error, often referred to as the bit error rate (BER), ranges from $5x10^{-1}$ to 10^{-6} on HF links. If there is such a thing as a typical BER range, it can best be set at $5x10^{-2}$ to $5x10^{-3}$ when the channel is available.

The nature of HF communication transmissions depends upon the directionality of the antenna being used. For short to medium (up to 1000 km) range communications, omni-directional antennas are often used and result in broadcast transmission. For medium (say > 500 km) and long-range communications, directional antennas are generally used. Typically, the beamwidths of directional antennas range approximately from 60 to 120 degrees. Such beamwidths produce a spatial coverage, at a distance "d" from the transmitter, of the order of d. This results in the possibility for a number of receiving stations (up to thousands of kilometers apart) to hear the transmitted signal. This type of transmission, being neither of a point-to-point (ppt) nor of a broadcast type, is hereafter referred to as being of a multicast nature.

Reciprocity cannot be assumed over an HF link. That is to say, even if user A receives user B's transmissions on frequency F_{AB}, it may well be the case that user B cannot receive user A on the same frequency. In fact, even if the best propagating frequencies (say F_{AB} and F_{BA}) in each direction are selected, one cannot assume bidirectional communications because of the local noise levels at each station. This non-reciprocity of HF communications, as will be seen below, has far reaching consequences on the design of a communication system.

3.2 The selection of adaptive techniques

In order to keep the level of difficulty within reason to ensure a successful outcome and to be able to produce a prototype system within a reasonable time frame, the data terminal has been limited to the use of a subset of adaptive techniques likely to produce the desired improvements in HF communications performance. Initially, the selection was limited to the use of a simple RTCE technique, channel selection, and adaptive link and network operation/management. This selection of techniques is not arbitrary.

RTCE techniques represent one of the most significant means to improve HF communications. A simple error-counting RTCE technique was included because of its ease of implementation and the good results it yields (see Section 4), but also because in-band link quality analysis is here considered to a large extent as a modem function. A high-speed adaptive serial modem has been developed at CRC and link quality analysis functions will be added to its design. This modem is currently being integrated in the HF data terminal design.

Frequency changes do not occur continuously and therefore a means to optimize the usage of a channel, once it has been selected, is required. This is done by adaptive link operation which, through optimization of the use of the available bandwidth, contributes to enhanced link connectivity and performance.

Packet-switching techniques are a main contributor to increased connectivity, robustness and survivability of HF communications. Figure 1 depicts a hypothetical HF network where individual links use different frequencies. Through routing of information in such a network and with the repetition of (unrecoverably) errored portions of transmissions, elements of time, frequency and path diversity are provided with a single antenna pair per link. This diversity is provided only as needed and without wasting of resources.

3.3 The HF data terminal design

The design of the HF data terminal has been subjected to a number of design guidelines. Some of the more significant were to produce a design that is as independent as possible of radio equipment and modems, to eliminate unnecessary (re)transmissions of information, to minimize time-synchronized operations and the negotiation of adaptive parameters over the HF link, and also to minimize the transmission overhead associated with link/network management and the adaptive nature of the system operations in general. The last three guidelines were essential in producing an efficient and robust communication system.

3.3.1 Hardware and software architectures

From a hardware point of view, the HF data terminal is simply a powerful microcomputer system. The current hardware design is depicted in Figure 2. The design is Multibus (IEEE 796) based. The most important components are a main processor (Intel 80286) board, a disk controller and a few serial communication boards. Figure 2 also shows the serial modem being incorporated in the terminal design.

The software architecture is based on the International Standards Organization (ISO) Open Systems Interconnection (OSI) model. The physical, link, network, and transport layers and a generalized application layer have been implemented. The main functions included in the data terminal architecture are shown in Figure 3. Only the spectrum management function at the network level has not been implemented yet. The data terminal architecture is characterized by adaptivity at each of its layers. Channel adaptation and transmission optimization occur mostly at the physical and link levels. Packet-switching functions (e.g. path selection or routing, priorities, etc.) take place at the network and transport levels..

3.3.2 The design characteristics

The HF data terminal is a complex system and a full description of its design is lengthy and inappropriate here. A discussion of the design decisions and a presentation of the design specifications are presented in [12]. Below, the main characteristics for each layer are briefly discussed.

3.3.2.1 The physical level

The main characteristic of the physical level design is the presence of a real-time channel evaluation (RTCE) and channel selection mechanism which permits the data terminal to adapt to changes in propagation conditions. End stations on a link periodically exchange test bit patterns on each assigned frequency and, upon reception, determine the best received frequency by counting the number of errors in the pattern. Based on this result and on the past history of the channel, the stations inform each other, on each assigned frequency, of the frequency to be used for transmission.

This simple error-counting RTCE technique is easy to implement and yields good results (see Section 4). The adaptive HF serial modem, currently being implemented in the HFDT design, will provide more adaptivity at the physical level as well as more sophisticated RTCE techniques. The serial modem includes adaptive channel equalization and will eventually provide extensive link quality analysis data dynamically.

Other features of the physical level include remote control of radio equipment and modems, and a hybrid FEC/ARQ scheme. Standard functions such as channel access, synchronization, etc. which are fundamental to any communication system, often become complex problems in the error-prone HF environment. Good schemes to accomplish these functions are essential to provide overall system efficiency. These issues are discussed in [12]..

3.3.2.2 The link level

The main characteristic of the link level is a specially designed robust, bandwidth-efficient and adaptive link protocol which allows dynamic optimization of the channel. This protocol implements a segment structure (Figure 4) [5,12,13] for its transmission unit (or frame) and can operate down to BER's as low as 10^{-2}.

The protocol varies the frame structure and length to adapt to the quality of the channel. When propagation conditions deteriorate, the protocol reduces the frame length and changes to a higher overhead state to ensure good error detection, to minimize potential data losses and to minimize the amount of information that might have to be retransmitted. When propagation conditions improve, the protocol decreases the overhead and lengthens the frames to maximize throughput on that link. This form of adaptivity is dynamic and not negotiated between stations.

Built into the protocol is a mechanism that allows a transmitting station, upon reception of data from the other end of the link, to find out how successful its own transmissions were and thus to take appropriate actions to improve upon them. Actions are taken dynamically and unilaterally by the transmitting stations. The receiving stations, upon decoding the frame header, know how to interpret the frame content and also know how successful its transmissions are.

This protocol is robust and bandwidth-efficient. Total loss of a frame of data can result only from unrecoverable errors in the frame-control field (a 6-byte subset of the frame header). Unrecoverable errors in any other segments of the frame result in loss of those segments only, and can be recovered through selective retransmission. The protocol uses a single frame type to accomplish all its functions including link establishment/closing and link management/maintenance functions. For comparison, the international standard HDLC (High Level Data Link Control) protocol [14,15] uses 19 types of frames for the same purposes and unrecoverable errors anywhere in the frame result in a total loss of that frame. HDLC is inefficient in a low-bandwidth environment and does not work when the error rate exceeds 10^{-3}.

3.3.2.3 The network level

The main network level adaptivity takes the form of a fully distributed, adaptive network routing scheme [16]. This routing algorithm allows each network node to operate autonomously without the use of a central control station and ensures that no node depends upon any other node to obtain routing information. In other words, no "master" station is needed with this concept and destruction, failure or addition of a network node, or loss of connectivity to a node does not affect the rest of the network.

The routing algorithm is based upon a minimum-relay (or minimum-hop) metric [17]. Each station constructs and updates its image of the network connectivity via the normal flow of packets passing through it. Only a few additional network management types of packets may be used to reflect lost links, node addition or node removal in the network connectivity matrix.

The network is self-configuring. At network startup time, the network level needs to know only the maximum possible number of nodes in the network and the maximum possible number of physical links at the local node. A "startup" packet is then sent on each possible physical link to establish initial connectivity status information (including whether the link is bidirectional or not) to its immediate neighbours. The network is then ready for operation. The normal data packet flow and a "looking-for-node" packet will suffice to determine all routes.

3.3.2.4 The transport level

The transport level introduces a priority scheme for data packets and a checkpoint/restart function. A packet priority scheme is a desirable feature under normal operating conditions and is a required feature under stressed conditions, especially in a military environment.

The checkpoint/restart feature allows the network node originating the traffic to keep track of the data succesfully delivered, across the network, to the destination. Should any (major) events occur making it impossible to reach the final destination through any network routes, the originator will stop sending data, periodically check if the problem has been resolved, and if so, will resume sending data at the point it was when the problem occurred. Such a mechanism is necessary because of the low bandwidth of the HF channel. Not doing so could result in a "snowball" effect where a system could spend hours or days to pass a message.

3.4 The functional aspect

The HF data terminal system alternates between two phases: a channel evaluation/selection phase and a data communication phase. Periodically, the system interrupts its data communication operations to proceed with an evaluation of its assigned channel, in each direction and on each link. End systems on a link then inform each other of the resulting selected channels and resume data transmission operations on the newly selected channels. This is the only time-synchronized link/network operation in the HF data terminal system. Other forms of adaptivity are dynamic and not negotiated over the air.

User data transmissions occur during the communication phase. During this phase and after having selected the best communication channel the system attempts, until it succeeds, to deliver user data to its destination in an error-free manner. While doing this, the system dynamically optimizes its operation for the quality of the communication channel, and dynamically makes the appropriate decisions, based on the current network connectivity, to reach the final destination in the most efficient manner. The sequence of operations taking place in this process can be best understood by following the data flow in the system (see Figure 4).

The sender's data terminal accepts messages typed on its keyboard or files from almost any external devices (including hosts) and stores them. Files or messages are fragmented into packets each prefixed with a transport header containing the destination, priority, etc. and all the information required to reconstitute the original message.

When ready to transmit a new packet, the transport level selects the packet with highest priority and passes it to the network level. This level examines the destination address and determines the best route to reach it. If a route cannot be found, the network level will circulate a "looking-for-node" packet and upon receipt of an answer from the destination node, will be informed of the route. Once a route has been established, the network level adds its own header (e.g. containing a network priority) and passes the packet to the link level.

The link level transmits the packet to this node's nearest neighbour along the route to the destination node. According to its current information (updated dynamically) on the quality of the channel, this level may decide to transmit this packet as one or more frames (transmission units), each of possibly different structures and lengths. Each frame is given to the physical level for application of an FEC code and transmission as a bit stream over the link.

Each node on the path will, upon receipt of an error-free packet by its link level, pass the packet to its network level which will update its connectivity matrix (based upon the information in the packet's network header) and will take one of the following actions:

- if this node is the destination for this data packet, the network level will pass it on to its transport level after having updated its connectivity matrix.
- if this node is the destination for a network management packet, the network level will absorb it and take the appropriate action.
- if this node is not the intended destination for a packet, the network level will route the packet to another node (using the same sequence of operations described above and after having appended this node's address to the packet header.

The transport header of the destination data terminal will send an (end-to-end) acknowledgement back to the sender's transport level for every packet or group of packets (group size is a system parameter) it has successfully received. It will then use this packet to reconstruct the original message or file.

It is important to note that each packet may be transmitted in a different manner on each leg of its journey to the final destination. This results from each link layer involved on this journey always optimizing the transmissions for that link. Likewise, groups of packets belonging to the same file/message, may take entirely different routes through the network since the same route selection process takes place, on a packet basis, at each intermediary node. There is thus no predetermined route across the network. Lastly, at any point during the transmission, the sender may request to checkpoint the packet stream.

3.5 The operational characteristics

The main operational characteristics of the HF data terminal are that it is automatic, adaptive, high-speed (2400 bps, can be downgraded), may operate half or full-duplex links and supports packet-switched network operations. All important system parameters can be changed by the operator and extensive link/network online

monitoring facilities are provided to the operator. The HF data terminal can interface to a wide variety of receivers, exciters, transmitters and parallel or serial HF modems. The consequences of these operational features are discussed below.

The HF data terminal system is automatic and thus eliminates the need for skilled HF operators. It currently requires one operator when, and only when, the time comes to load or extract data to/from the system. Messages, files or other kinds of data are fed into the system which will then ensure their virtually error-free delivery to the final destination without further operator intervention. The data terminal can be used as an automatic (half or full-duplex) link controller and/or as an HF network node.

The HFDT prototype design is flexible and provides extensive monitoring functions. All important system parameters are fed into the system and can generally be changed during system operation. For example, the number of frequencies and the particular frequencies for each link are parameters for each network node. The frequency of channel evaluations/selections, the link data rate, the frequency of or the number of packets between checkpoint/restart operations are also system parameters. Changes in those parameters are reflected in configuration data which, with status and performance data for each link, are available online.

The HF data terminal is, for all practical purposes, independent of the radio equipment and HF modems used. It only requires that such equipment be computer controllable and this is a general characteristic of modern RF equipment. The operational differences between HF vendors' equipment can generally be accomodated by changes in the system parameters or, for significant differences, by simple software modifications. The terminal can thus make use of better radio equipment as it becomes available. In particular, better and/or faster, parallel or serial HF modems can be interfaced to the system as they become available.

The current network design is, for demonstration purposes, for a broadcast TDMA network. The HFDT design will, for operational purposes, accomodate FDMA network operations.

3.6 Summary

The HF data terminal is a generic, adaptive communication system and can thus be tailored for operation in a wide range of HF applications. Its design is characterized by adaptivity at each OSI level implemented. The main form of adaptivity includes a real-time channel evaluation and selection scheme at the physical level, an adaptive link level protocol and a fully distributed, adaptive network routing mechanism. With the exception of channel evaluation/selection, all forms of adaptivity are dynamic and not negotiated over HF links. Channel evaluation/selection is the only link/network time-synchronized operation between communicating data terminal systems.

The HF data terminal is adaptive and thus allows optimum utilization of propagation conditions and equipment resources. Its built-in adaptivity solves the limitations associated with the current method of operation of HF systems. In particular, it solves the problems associated with the "two-frequency" operating schedule by responding in real time to changes in propagation conditions. Through real-time adaptation, advantages can be taken of unpredictable phenomena such as sporadic-E propagation to provide better HF communications. Network route selection contributes to being able to reach the maximum connectivity possible between stations by providing elements of path and frequency diversity. Dynamic link operation optimization results in the best possible utilization of the bandwidth available.

The HF data terminal can be used as an automatic, adaptive, high-speed (currently up to 2400 bps), half or full-duplex HF link controller and/or HF network node. In an HF network, the data terminal does not require a network "master" station and allows node addition and node removal to be done without network interruption. A node failure/recovery is also dealt with automatically and, depending on the network architecture, generally has little impact on network operation and performance.

4.0 THE HFDT EXPERIMENTAL EVALUATION

In addition to the laboratory back-to-back system tests following each development phase, the HF data terminal system was evaluated in a series of experiments. The experimental programme has consisted of:

- short HF link experiments
- a network emulation
- long HF link experiments.

The emphasis in this section is on the connectivity and throughput results from on-the-air tests of the system.

4.1 The short-link experiments

The short-link experiments were conducted from July to early October 1984 between CRC and Petawawa. At this stage in the terminal design, the physical layer, link layer and link monitoring functions as well as the user interface had been developed. The HF data terminal thus had the functionality required to act as an automatic and adaptive link controller.

The main operational and link parameters for these experiments were:

- 136-kilometer link
- no forward error correction
- half-duplex, 2400-bps link operation
- 8 frequencies from 5.1 to 11.6 MHz
- omni-directional, inverted-V dipole antenna
- 100 watts average power
- 16-tone Kineplex-type HF modem (~ 6 watts/tone)
- solar cycle nearing sunspot minimum.

The experimental configuration is depicted in Figure 5. In this half-duplex setup, all equipment is co-located. Transmit and receive equipment share the same antenna. The HF data terminal did not include forward error correction and thus relied upon the adaptiveness of the link protocol and on the automatic repeat request (ARQ) mechanism built into the protocol to provide error-free communications.

Two basic experiments were conducted during this test. The first one was a channel evaluation/selection experiment where the system, every 3-5 minutes, would sound all channels and select the best one. This experiment lasted for three months from July to October. The second experiment was a throughput experiment and lasted through the month of September.

4.1.1 The results

For the duration of the short-link experiments, the daytime average bit error rates (BER) were 10^{-2} at CRC and $4x10^{-3}$ at Petawawa. This difference is easily explained by the high background noise and interference levels at CRC. Nighttime propagation was generally not possible with the range of frequencies used. A channel around 2 MHz would have been needed for nighttime communications. The experimental period included a two-day HF blackout period (reported by other experimenters) during which the HF data terminal was still able to pass information at a reduced rate.

The main daytime results from the experiments were that:

- more than 90% of the time, a communication channel is available at 2400 bps,
- the typical 8-hour average error-free user throughput, in each direction, is 150 bps. It ranges from 80 to 400 bps.

A typical throughput curve and one for a good communication day, are shown in Figures 6 and 7, respectively. The curves are drawn from fifteen-minute averages of the throughput. The variations in throughput on each curve result from varying propagation conditions and communication environment (e.g. noise or interference level), with the system automatically adapting to them.

Although those results were very encouraging, they served to pinpoint deficiencies in the system's initial design, in particular in its frame synchronization mechanism and in the channel access strategy. For example, data analysis revealed a frame-loss rate as high as 50% at times. The experiment was followed by numerous enhancements in the data terminal design to correct the problems above and to enhance its performance in general.

4.2 The Carp-Penhold experiments

The short-link experiments were followed by various design enhancements, by the development of full-duplex link operation and by the development of the network and transport functions. The correctness of the implementation of the networking functions was demonstrated in a network emulation which took place in the summer of 1985.

The long-link experiments took place from November 1985 to February 1986 between Carp (near Ottawa) and Penhold (Alberta). The same two basic experiments conducted over the short link were repeated over the long link. As before, the systems relied strictly upon their adaptiveness and ARQ mechanism to provide error-free user data. Forward error-correction was not applied to the user data.

The experimental configuration for the long-link experiments is shown in Figure 8. The configuration is characterized by the distribution of equipment over transmitter, receiver and control centre sites. Control of the entire configuration is done via the system console at the control centre. Remote receive and transmit control units were required to control HF modems and radio equipment under the direction of the HF data terminal. More modern radio equipment would alleviate the need for such remote units.

Another important characteristic of this configuration is the placement of the HF modems at the transmitter and receiver sites. Communication to the transmit/receive sites is done over landlines. In Penhold, approximately five and ten miles of underground twisted-pair cables link these sites to the control centre. Because the quality of these cable pairs and the effects of variable soil conditions (e.g. frost, water, etc.) on them were unknown, the decision was made to remote the HF modems to the transmitter and receiver sites. This safe approach was made at the cost of reducing system throughput by 50% owing to the delays introduced by digital transmissions over 4.8 kbps landlines.

The length of the frame synchronization detection bit-pattern was increased from 16 bits for the short-link tests to 24 bits for the long-link experiments. This is the maximum length permissible by the enhanced version of the frame detector. Although better than a 16-bit pattern, the 24-bit pattern is still too weak. A reliable frame detection pattern for HF communications needs to be substantially longer [12].

The main operational and link parameters for this experiment were:

- 2260-kilometer link
- full-duplex, 2400-bps link operation
- forward error correction applied to the frame header only
- 5 frequencies from 7.9 - 19.2 MHz, in each direction
- horizontally polarized, log-periodic antenna
- less than 400 watts average power (typically 200 watts)
- 16-tone Kineplex-type HF modems (~ 20 watts/tone)
- solar cycle nearing sunspot minimum.

4.2.1 The results

For the three-month duration of the Carp-Penhold experiment, the channel-averaged daytime BER values, corrected for missed frame detection, were approximately 10^{-1} at Carp and $2x10^{-2}$ at Penhold. The best (selected) channel average daytime BER values, also corrected for missed frame detection, were $4x10^{-2}$ at Carp and $4x10^{-3}$ at Penhold, over the same period. The differences between those results is primarily due to receiver problems at Carp; a secondary factor is the higher background noise level at Carp.

Nighttime propagation was generally not possible (see Figure 9) with the range of frequencies used. Channels in the 4-6 MHz range would have been required for nighttime communications. Two major geomagnetic events, including a severe geomagnetic storm, occured during the experimental period and resulted in HF radio blackout. The Carp-Penhold link (as well as other strategic HF links in Canada) lost connectivity for roughly two days in each case. Had the systems been using more power, there would still have been a loss of connectivity between Carp and Penhold although probably for a shorter period of time. An interesting observation was made during those events and is described below.

The main daytime results from the Carp-Penhold experiment were that:

- more than 90% of the time, a communication channel is available at 2400 bps (using little power)
- the typical daily average, error-free user throughput, in each direction, ranged from 100-300 bps (it can be doubled if HF modems are located at the control centres)
- instances were found, associated with the geomagnetic events, where Carp-Penhold communications were impractical, but shorter links (e.g. Carp-Shilo, Penhold-Shilo) recovered quickly and would have been utilizable.

Figure 9 shows the percentage of the time, over the three-month interval, when the best frequency was available as a function of Ottawa local time. It can be seen that the best individual frequency was available, on average, approximately 50% of the time. That is to say, a communication system using only that frequency would have been able to communicate only 50% of the time over this three-month interval.

Figure 10 shows the same type of curve when the best channel, out of the five available frequencies, is selected after each channel evaluation phase. The percentage of availability then jumps from 50 to 90%. This is achieved through channel evaluation based on a simple error-counting RTCE technique. This result also illustrates, as is well known, that a principal mechanism to improve HF communications, is real-time channel evaluation.

The benefit of routing information in an HF network was recognized during the HF radio blackout events. At these times, all HF links were inoperational. However, short to medium (< 1000 km) range links recovered nearly one day sooner than longer links. Although it was a partial recovery, i.e. communications were not great, it would have been possible to communicate over shorter links (e.g. Carp-Shilo, Shilo-Penhold) and to achieve connectivity over the longer link by routing the information over shorter links.

The results presented above were achieved with low power, no forward error correction of the user data, a weak frame detector and an elementary RTCE technique for channel evaluation/selection. The connectivity obtained with this adaptive system was superior to that of strategic links using lower data rate (75 bps) and more power (1 kw). The error-free throughput achieved with this adaptive system, even with no forward error correction, exceeds that of existing HF communication systems by at least an order of magnitude. The operation of the system is automatic rather than being man-intensive.

4.3 Discussion

The results from on-the-air tests of the HF data terminal indicate the existence of a high-speed (2400-bps) HF communication channel at least 90% of the time in the sub-auroral region. Although the tests were done near sunspot minimum, the long-link experiments were performed during the winter when HF communications are more difficult. The experimental results demonstrate the validity of the adaptive-system approach as a means to improve HF communications.

The adaptivity built into the link protocol, combined with the physical-level channel evaluation/selection mechanism, is able to provide better connectivity and throughput than in existing HF systems using lower signalling rates, higher power levels and error correction schemes. The improvement in error-free throughput provided by the HF data terminal is at least an order of magnitude beyond that of existing systems. The error-free throughput value can be doubled simply by changing the HF modems' location in the experimental configuration.

Experimental evidence has been found showing increased connectivity (survivability) provided by adaptive routing of information in an HF network. A similar result was noted by BR Communications in tests with polar-cap and auroral circuits [18]. BR found the existence of good north-south (latitudes of approximately 62° and 37°) routes at times when polar-cap east-west routes were non-existent.

Although very encouraging, the experimental results of this prototype version of the HF data terminal can be significantly improved with the use of better equipment and some minor design modifications. In particular, the HF modems used, in addition to having an unreliable carrier detect signal, introduced fixed delays of typically 75% or more in each frame transmission time. The fixed delays reduce the effectiveness of the data terminal (i.e. decrease its throughput) by reducing the frequency of packet transmissions. They are the main contributors to the difference in the link data rate and the actual user throughput.

5.0 CONCLUSIONS

This project has used an adaptive-system approach to HF communications. It has made extensive use of adaptive techniques and of packet-switching techniques in particular. The vehicle implementing these techniques, the HF data terminal, not only includes RTCE techniques but is characterized by adaptivity throughout its design. It

includes the functions for packet switching over HF and the management of a fully distributed and adaptive network. The terminal's adaptivity is implemented via a suite of robust, low-overhead, adaptive protocols that work even in low-bandwidth, error-prone environments.

A system based upon adaptive techniques improves significantly the connectivity, speed, throughput and reliability of HF communications. The experimental data has shown the availability of high-speed (2400-bps) HF channels more than 90% of the time in the sub-auroral region. The HF data terminal, in its current version, already offers at least a one-order-of-magnitude improvement in error-free throughput over existing operational HF communications systems.

The HF data terminal design, because it adapts to channel quality, is less vulnerable to errors and gracefully degrades its throughput under deteriorating channel conditions. Further, the experimental data suggests that the HF data terminal could efficiently be used at higher data rates over strategic links. It would thus be worth conducting system tests at, for example, 4800 bps.

Packet switching over low-bandwidth, error-prone HF channels is feasible. A system using this technique improves the reliability and the survivability of HF communications. More experiments are needed to quantify this improvement.

Overall adaptive HF communication systems probably represent the next generation of HF communication systems. Further experimentation is required across trans-auroral and polar-cap circuits and in periods of high solar activity, to find out if such systems, in an adaptive network configuration, can meet all new HF military requirements. They are, it is believed, the most serious contender to meet these requirements.

REFERENCES

1. G.W. JULL, D.J. DOLE, G.W. IRVINE and J.P. MURRAY, Frequency sounding techniques for HF communications over an auroral zone path. Proc. IRE, Vol. 50, 1962.

2. E.E. STEVENS, The CHEC sounding system. In Ionospheric Radio Communications (Ed. K. Folksted), Plenum Press, New York, 1968.

3. B.D. McLARNON, Real-time channel evaluation in an automatic HF radiotelephone system. Second Conf. on HF Communication Systems and Techniques, London, England, February 1982.

4. S.M. CHOW, G.W. IRVINE, B.D. McLARNON and A.R. KAYE, Communications for small communities in developing countries. Pacific Telecommunications Conference, Honolulu, Hawaii, January 1981.

5. N.M. SERINKEN and S.M. CHOW, High-frequency data-message terminal. 15th Annual Off-Shore Technology Conference, Houston, Texas, pp. 415-420, May 1983.

6. N.M. SERINKEN and S.M. CHOW, Automatic HF radio message terminal system for sparsely populated communities. Pacific Telecommunications Council Seventh Annual Conference, Honolulu, Hawaii, 1985.

7. M. DARNELL, Real time channel evaluation. AGARD Lecture Series No. 127 "Modern H.F. Communications", AGARD-LS-127, 1983.

8. M.D. GROSSI, H.F. Propagation measurements for real time channel evaluation schemes. AGARD Lecture Series No. 127 "Modern HF Communications", AGARD-LS-127, 1983.

9. M. DARNELL, Embedded real-time channel evaluation techniques. AGARD Lecture Series No. 145 "Propagation Impact on Modern HF Communications System Design", AGARD-LS-145, 1986.

10. J. AARONS, High frequency communication: an introduction. AGARD Lecture Series No. 145 "Propagation Impact on Modern HF Communications System Design", AGARD-LS-145, 1986.

11. G.R. NOURRY, An adaptive, packet-switched HF data terminal - Part I. Department of Communications (Canada), CRC Report No. 1423, 1987.

12. G.R. NOURRY, An adaptive, packet-switched HF data terminal - Part II. Department of Communications (Canada), CRC Technical Report No. 1424 in preparation, 1987.

13. L.R. BODE, S.M. CHOW, N. SERINKEN, B. TEPPER and R. YAMANITO, Experimental evaluation of an HF data terminal in ship/shore mode. Department of Communications (Canada), CRC Report No. 1375, August 1984.

14. A.S. TANENBAUM, Computer networks. Prentice-Hall Inc., Englewood Cliffs, New Jersey, USA, 1981.

15. D.W. DAVIES, D.L.A. BARBER, W.L. PRICE and C.C. SOLOMONIDES, Computer networks and their protocols. John Wiley & Sons, Wiley Series in Computing, 1980.

16. K. BRAYER, An adaptive computer communications network designed with decentralized control. IEEE Communications Magazine, Vol. 21, 1983.

17. D.H. CHYUNG and S.M. REDDY, A routing algorithm for computer communication networks. IEEE Trans. Comm., Vol. COM-23, 1975.

18. BR COMMUNICATIONS, HF propagation in the arctic: High performance communication channels are available if you know where to look. Presentation at the AFCEA show (Canada), April 1986.

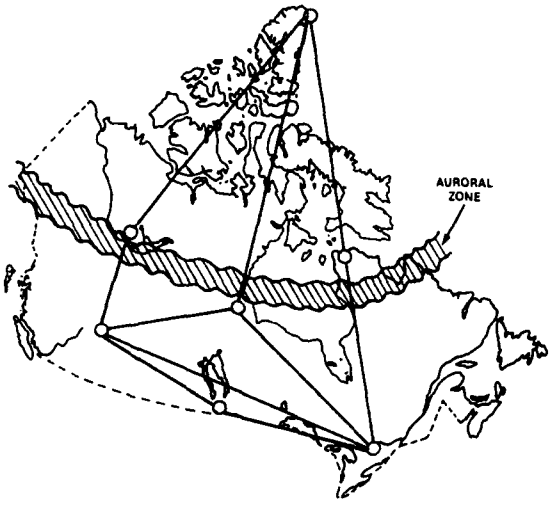


Figure 1: AN EXAMPLE HF NETWORK

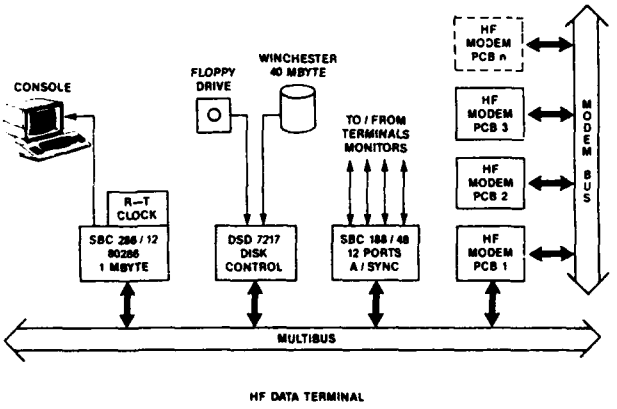


Figure 2: THE HFDT HARDWARE ARCHITECTURE

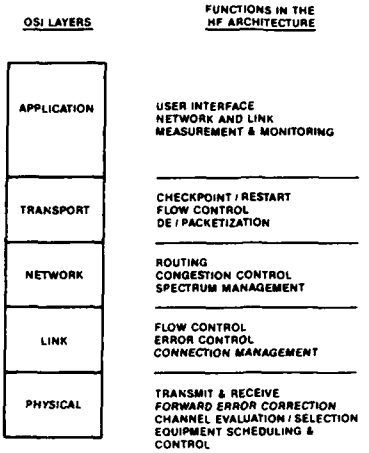


Figure 3: THE HFDT SOFTWARE ARCHITECTURE

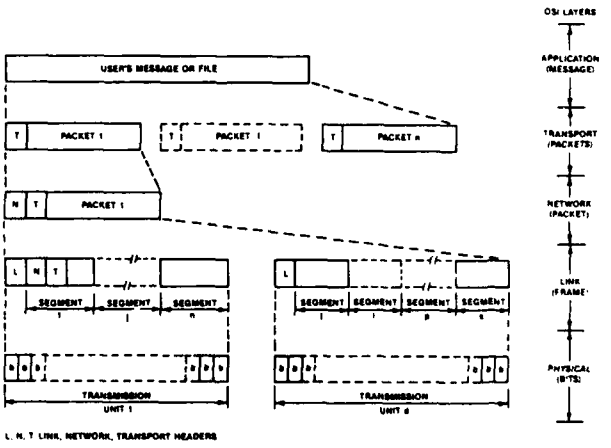


Figure 4: THE HFDT FUNCTIONAL DATA FLOW DIAGRAM

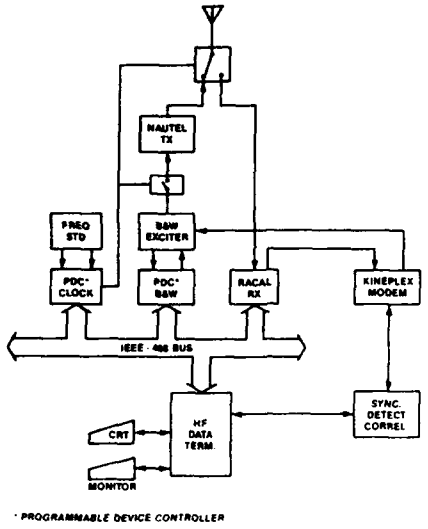
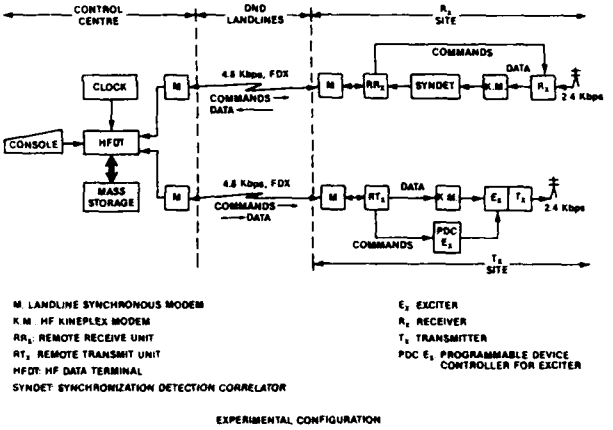


Figure 5: THE SHORT-LINK EXPERIMENT
- HDX EXPERIMENTAL CONFIGURATION



EXPERIMENTAL CONFIGURATION

Figure 8: THE CARP-PENHOLD FDX EXPERIMENTAL CONFIGURATION

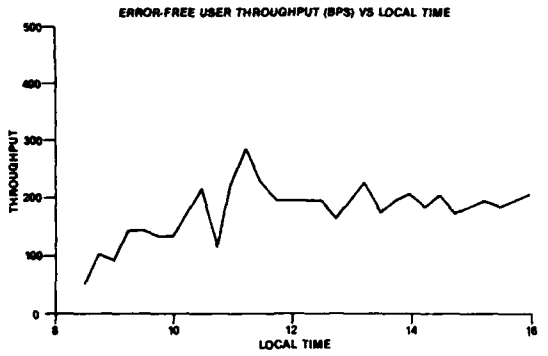


Figure 6: THE SHORT-LINK EXPERIMENT
- TYPICAL THROUGHPUT CURVE (DAY 273)

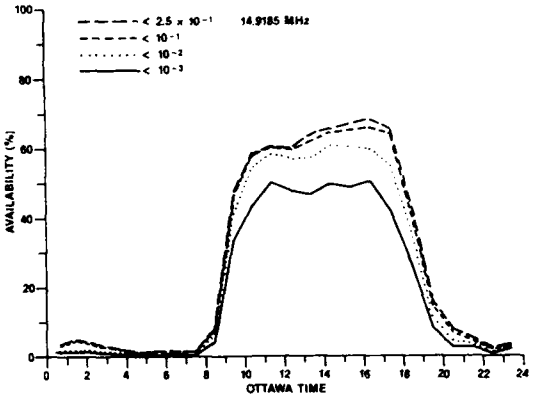


Figure 9: THE CARP-PENHOLD EXPERIMENTS
- CHANNEL 3 AVAILABILITY (PENHOLD)

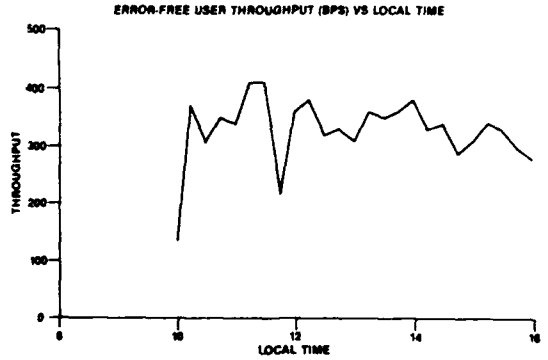


Figure 7: THE SHORT-LINK EXPERIMENT
- THROUGHPUT CURVE ON A GOOD DAY (DAY 275)

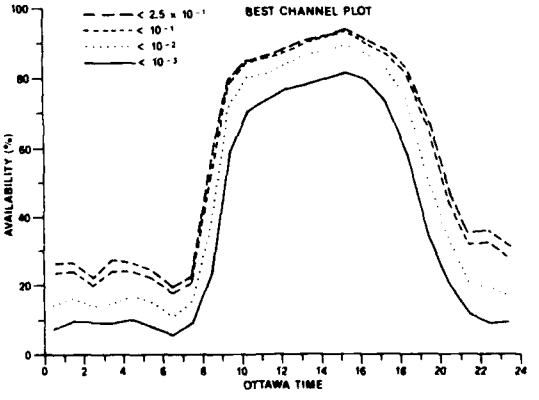


Figure 10: THE CARP-PENHOLD EXPERIMENTS
- AVAILABILITY OF BEST CHANNEL (PENHOLD)

DESIGN CONSIDERATIONS FOR AN AUTOMATED HF DATA NETWORK WITH ADAPTIVE CHANNEL SELECTION

R J Goodwin
A P C Reed
Plessey Research Roke Manor Ltd
Romsey
Hampshire SO51 0ZN

SUMMARY

A fully automatic HF data network has been designed which adapts to changing propagation and interference conditions. Frequency management at network level derives real time propagation data which is supplied to all network stations. At link level, propagation prediction, interference assessment and link sounding together provide adaptive channel selection for each call. Efficient and reliable message transmission is achieved using a novel FEC/ARQ protocol which also provides for adaptive changes in data rate or channel reselection during the call. Frequency Management and Network Management functions can be executed by any subset of stations in the network and no special sounding equipment is required.

1 INTRODUCTION

1.1 An Automatic HF Network

A new generation of HF systems is emerging in which traditional operator skills are being replaced by processor functions able to respond to the time-varying propagation conditions, noise and interference levels experienced in the HF band.

It is now possible to provide real time frequency management for such systems which can track changing propagation conditions. It is also possible to implement adaptive channel selection algorithms which set up calls on viable channels, not blocked by noise or interference. The system architecture can ensure that in a large, busy network, which may be expanded, member stations do not cause mutual interference on the network's allocated channels. Low cost processor techniques allow the use of powerful error control coding and repeat facilities which offer efficient data transfer over time-varying channels without compromising message integrity.

This paper describes aspects of the design of a large, fully automatic data network which will shortly undergo system trials. Much of the design has already been tested through extended use of a skywave trials link in the UK.

The network design conveniently separates into Frequency Management, Network Management and Link Management Subsystems and these are each covered in some detail later in the paper.

1.2 HF Trials Link

A fully automatic HF trials link has been operated over a 132km skywave path in the South of England during an extended period between October 1986 and July 1987. The two stations operate 24 hours per day, in an unattended mode, and performance data is continuously logged for subsequent off-line analysis.

Key features of the adaptive HF system design described have been exhaustively tested using this trials link and relevant results will be presented in subsequent sections of this paper. The trials link has been especially useful in optimising the Frequency Management and Link Management Subsystems described in Chapters 3 and 5 respectively.

2 SYSTEM DESIGN CONSTRAINTS

Prior to discussing aspects of the system design, it is helpful to summarise the major constraints within which that design progressed.

The HF network is designed to carry telegraph traffic only and is required to link some 80 communications centres distributed throughout a reasonably compact geographical area. Direct communication, without rebroadcast, is required between any two sites for Single Address Messages (SAMs) or between one site and defined groups of sites for Multiple Address Messages (MAMs). The majority of connections will be made via single hop skywave paths with just a few links able to operate using groundwave transmissions.

With a large network of this type, equipment costs present a major constraint. In particular the available budget rules out the possibility of using all but the simplest of modems and a maximum transmission rate of 300 baud has been accepted at the outset. Nevertheless a high traffic capacity is required in the network (it is expected that between 10 and 20 simultaneous calls will be in progress at all times of day). Furthermore short delivery times are specified, particularly for high precedence messages.

Each active link must achieve the maximum possible data throughput in order to meet the traffic requirements. This conflicts with a second important requirement for high data integrity (ie low character error rates). The two opposing requirements have been met through the use of highly adaptive frequency management and channel selection algorithms, plus carefully optimised call set-up and data transmission protocols.

Fortunately, this large HF network will receive a generous allocation of channels distributed across the HF band. This has permitted subsets of channels to be reserved for call set-up, engineering and frequency management purposes and thus careful control of mutual co-channel interference problems within the network is possible. Figure 2.1 summarises the use of available channels in the network. Co-channel interference from other HF users remains a significant challenge and is a major influence throughout the system design.

A final constraint, imposed by the network user, is the need for complete hardware and software commonality at all stations in the network. Perhaps the most important result of this constraint is to eliminate the possibility of using special purpose HF sounding equipment. The provision of such equipment at every site would be prohibitively expensive.

3 FREQUENCY MANAGEMENT SUBSYSTEM (FMS)

3.1 FMS Overview

Although frequency management in a radio network is only one part of the communications management function, its importance at HF is paramount. It is therefore appropriate to discuss frequency management in advance of the higher level network architecture considerations.

The required network must offer a high grade of service in terms of throughput and message delivery times at all times of day. Operation must continue, possibly at a reduced level, in disturbed propagation conditions. A system which adapts automatically to prevailing conditions is thus required, rather than one relying on propagation prediction alone. In fact, a compact prediction model, which assesses both skywave and groundwave paths, is provided within the station software, but this model is updated regularly in accordance with measurement data derived by the network.

The FMS subsystem divides into two functional areas at network and link levels respectively (Figure 3.1). The objectives at network level are to establish the current state of the ionosphere, by means of channel soundings, and to disseminate this information to all network stations. The objective at link level is to select viable channels for each traffic transmission.

3.2 Description of Network Level FMS

A small subset of stations within the network, and spaced throughout the deployment area, are selected to act as frequency management stations (Figure 3.2). Since all stations are identical this selection may vary with time, either to counter an ECM threat or to provide replacements in the event of station failures.

At predetermined times, one FMS station will coordinate a sounding routine, which scans a set of channels in the HF band. The coordinating station transmits probes on each of the channels in turn and then awaits responses from the remaining FMS stations, in ordered timeslots (Figure 3.3). From the replies received, the coordinating station can estimate the current Maximum Usable Frequency (MUF), assess current absorption levels and determine whether disturbed conditions are present. The same station can then compare measurement data with predicted values from its local model and compute correction factors for distribution to all stations in the network. Embedded prediction models, at stations throughout the network, can thus be updated to take account of both unpredictable day-to-day variations in the ionosphere and the existence of severe disturbances when appropriate.

The coordinating station will also use the results of channel sounding to construct an assignment of calling channels for use throughout the network (Section 5.2 will describe the use of calling channels). This assignment must also be distributed to all stations in the network.

The time occupied by the sounding routine will depend on equipment tuning times, the number of channels in the scan and the number of FMS stations used. Where the FMS stations are also required to pass traffic, it is obviously important to minimise this time and a significant improvement can be achieved by synchronising the stations concerned. For example if the following parameters are assumed:

a) Minimum dwell time per channel = 2 secs
b) Number of channels in the sounding routine = 40
c) Number of FMS stations = 4

in the non-synchronised case, each FMS station must cycle its receiver through the 40 channels awaiting a probe transmission and this probe transmission, from the coordinating FMS station, must occupy at least 82 seconds to ensure detection by the

scanning receivers. Allowing 6 secs for replies gives a time of 88 secs per channel, or a total scan time of 3520 secs.

If the stations are time synchronised then the probe transmission reduces to 2 secs, the total time per channel becomes 8 secs and the scan occupies 320 secs - an order of magnitude improvement!

Reducing the scan time in this way not only releases more time for the FMS stations to handle traffic, it also allows the FMS scan channels to be used for other purposes in the network, including call set-up (Section 5.2). These channels will not normally be selected for traffic thus avoiding co-channel interference due to sounding.

3.3 Performance of Network Level FMS

The HF trials link, described in Section 1.2, includes software designed to evaluate the performance of the network level FMS algorithms. A bidirectional sounding of 37 channels distributed across the HF band is performed every 30 minutes. The results of each scan are stored for subsequent analysis.

Following each scan, a software algorithm estimates the maximum usable frequency and applies a correction to the embedded prediction model at the terminal station. Examination of the results on both normal and disturbed days shows that this algorithm correctly estimates the MUF for 99% of scans.

It has also been possible to compare the scan results with vertical incidence soundings performed at the UK observatory at Slough which apply to a point some 48km east of the centre point of the trials link.

Figures 3.4 to 3.6 show that the sounding mechanism will correctly track the propagating window in a variety of conditions and detect significant departures from mean predictions from an uncorrected model.

3.4 Link Level FMS

Each station in the network is provided with a compact propagation prediction model which uses a reduced CCIR data base relevant to the geographical deployment area (1). The model is updated regularly using correction data supplied by the network level FMS previously described in Section 3.2.

A station wishing to originate a SAM call will compute link range using the geographical coordinates of the single destination station and then derive the viable propagation window or windows using its updated prediction model. For shorter links groundwave propagation may be shown as feasible and all links may occasionally benefit from additional spectrum availability due to anomalous modes such as sporadic E.

The task is slightly more complex for a station wishing to originate a MAM call since the model will restrict the viable propagation window to normal E or F layer skywave propagation and must take account of the range of link distances involved.

The originating station will next select allocated traffic channels randomly within the viable propagation window(s) and reject any with noise or interference levels likely to disrupt data transmission. This assessment is based on passive monitoring of the channels using the station receiver. Channels not rejected in this passive assessment form a small set of candidate traffic channels which the source station ranks in ascending order of noise level (ie quietest first) and offers to the destination station(s) during the call offer protocol.

The call offer sequences for both SAM and MAM calls incorporate active sounding of the candidate traffic channels before a single channel is chosen on which to begin the call. Since active sounding would cause interference to any existing call on a channel, it is clearly important that the earlier passive assessment rejects channels already in use in the network.

For SAM calls, test transmissions are sent on each of the candidate traffic channels, allowing the destination to measure error rate as an indication of channel quality. The destination station readjusts the ranking order of the candidate channels and replies on its first choice. The message transmission then begins on this preferred channel. Additional candidate channels are held in reserve in case the first choice degrades during the call and a channel change becomes necessary.

For MAM calls, the source station transmits test signals on the candidate traffic channels and allows time for the destinations to reply with quality scores in ordered timeslots. As soon as an adequate number of replies is received on a channel the transmission begins without further active sounding of remaining channels. This curtailment of active sounding for MAM calls is necessary to reduce the set up time for calls addressed to a large number of destinations.

3.5 Performance of Link Level FMS

Section 3.3 described the sounding routines performed every 30 minutes on the trials link in order to update the propagation prediction models at the terminal stations. In the intervals between these soundings representative calls are set-up

across the link. A full link level FMS routine is performed, prior to each call, in which the source station calculates the propagating window for the link using its updated model, selects candidate traffic channels via passive interference assessment and finally performs active sounding to allow the destination to choose a preferred channel.

The propagation prediction and passive assessment aim to supply a list of candidate traffic channels which will propagate and which are not blocked by interference. Figure 3.7 indicates that this is successfully achieved.

The active sounding stage aims to select a preferred channel which will allow the successful transmission of at least the first part of the call. During the trials typically 86.5% of segments in the first message packets were received correctly without a need for repeats. Overall, some 15% of segments in the first 5 message packets were repeated under the ARQ protocol.

4 NETWORK MANAGEMENT SUBSYSTEM (NMS)

4.1 NMS Overview

The most critical aspect of HF networ' management has been identified as the Frequency Management Subsystem and this has already been covered in Section 3. The Network Management subsystem discussed in this section provides the following functions:

a) Network initialisation.
b) Station affiliation (both during network initialisation and as late entrants to the network).
c) Distribution of engineering information including FMS data and time synchronisation.
d) Network monitoring and control.

4.2 NMS Architecture

The earlier description of the Frequency Management Subsystem identified the need for a network level FMS based on a small subset of stations performing regular channel sounding routines. Correction factors and channel assignments, based on the results of this sounding, must be distributed to all stations in the network. Distribution of this data at HF is most effectively handled by a single control station (CS) which logically should also be the FMS coordinating station described in Section 3.2. The single CS also becomes the master reference for network synchronisation and the coordinator for station affiliation and network connectivity.

A single CS architecture may, in some circumstances, be considered vulnerable to ECM, but it should be remembered that all stations are identical and, if affiliated to the network, will have a common engineering data base (Section 4.4). Thus the CS role may be transferred from site to site in the network either when an existing CS becomes ineffective, or at regular intervals according to a prearranged sequence.

A second disadvantage would arise if certain stations in the network are unable to receive transmissions from the single CS. This is unlikely to occur at HF but the system design does provide for optional relaying of control messages by the other FMS stations.

4.3 Network Management Channels

Network initialisation and late entrant affiliation can only be achieved efficiently in terms of spectrum usage and time if a common set of network management channels are pre-stored at all stations prior to deployment. These channels should be distributed across the HF band and should be sufficient to cover the majority of propagation conditions. All stations which are not affiliated to the network will scan these network management channels, using their local receivers, until a broadcast signal is received from the CS.

The network management channels are part of the set used for the FMS scan (Figure 2.1). Mutual interference between FMS and NMS transmissions is readily avoided by time scheduling at the CS. The network management channels are not normally used as traffic channels as this would lead to co-channel interference problems.

The CS is able to use data gathered from the network FMS sounding routines to save time by choosing preferred management channels for its control transmissions. Stations not affiliated to the network will not be aware of this choice and must therefore scan the total set.

4.4 Engineering Updates

Stations affiliated to the network will receive regular engineering updates in the form of messages broadcast by the CS at known times. These updates contain FMS information (correction factors and calling channel assignments), a time reference, active station list and control instructions. Error correction and diversity techniques are used to ensure the reliable distribution of engineering data.

5 LINK MANAGEMENT SUBSYSTEM

5.1 Link Management Overview

Link level FMS, which effectively provides an automatic and adaptive channel selection procedure, has already been described in Section 3.4. Link management also requires a robust protocol for establishing a call to the required destination (or destinations) which is effective even in adverse conditions and which does not itself cause interference to other users in the network.

The management of a call in progress is essentially the operation of a secure and efficient data link protocol which achieves the required balance between traffic throughput and data integrity.

5.2 Call Placement Procedures

A rapid and reliable call placement procedure is an essential factor in providing adequate network capacity and in achieving efficient equipment usage at busy stations. Considerable attention has been paid to this aspect of the system design and in comparing the various techniques available.

In any dynamic radio network, stations in the idle mode must listen on one or more channels and await a call offer. In most systems a successfully received call offer will lead to a subsequent call transmission on the same channel. Once a call is in progress the channel is unavailable for further call offers or calls and indeed if the channel is inadvertently seized by another station, co-channel interference will result.

It is more efficient, in a network in which a number of calls must be handled simultaneously, if idle stations scan a number of channels looking for call offers. In a large network, time is then wasted visiting channels which have already been seized for traffic. There is also a significant risk of interference due to stations offering calls on channels which they have failed to detect as occupied.

One possible remedy would be to subdivide the network into station groups and to assign a channel or small set of channels to each group. This reduces the scan time for idle stations and speeds up call placement. However spectral efficiency is reduced because unused channels in one part of the network cannot be made available for use in more congested areas. In addition, call delivery times would increase because calls would be delayed due to the lack of a channel, even though the destination is available.

In large networks, the solution is to separate the calling and message transfer transmissions by assigning dedicated calling channels which are not used for traffic. This immediately removes the possibility of call offers interfering with existing transmissions and, because the calling channels are relatively uncongested, call placement becomes rapid and reliable.

Stations in the network continuously scan a set of calling channels awaiting a call offer. A station wishing to place a SAM call transmits an offer message on each of the calling channels in turn until an acknowledgement is received from the destination. Immediately a successful handshake is completed on a calling channel the call set-up transfers to the active sounding of candidate traffic channels.

For MAM calls, the source station transmits unacknowledged but similar call offer signals on each of a set of 5 calling channels and then begins trial transmissions on each of the candidate traffic channels. A different procedure is required for MAMs because the call is offered to multiple destinations. By offering the call on each calling channel there is a high overall probability of contacting each available destination and it is then possible to dispense with acknowledgements. This saves considerable time and reduces the loading on the calling channels.

5.3 Data Link Protocol

This paper has already described the selection procedures which aim to set up calls on channels which will propagate over the required link or links and which will not be blocked by noise or interference. Almost inevitably at HF, conditions on the chosen channels will suffer fading and time varying interference and it is essential that the chosen data link protocol should offer a high degree of protection against these effects.

A fixed strategy, designed against worst-case conditions, would be highly inefficient, especially in Western Europe where co-channel interference is such a dominant factor. The adopted approach has been to design a protocol which achieves a constant (high) level of data integrity while adaptively varying data throughput in response to perceived changes in channel condition.

HF channels are well known to be characterised by good periods (low error rate) interspersed with periods of markedly poorer conditions (high error rates). Both fading and co-channel interference contribute to these time-varying properties which have been widely observed on the trials link. Interference profiles reported by Gott (2) show that fine features change on a timescale of <1s (coherence approximately 0.25s) while major features persist rather longer.

In the network described in this paper, equipment constraints and spectral efficiency demand single channel simplex operation. A new hybrid Forward Error Correction (FEC)/Automatic Repeat Request (ARQ) protocol has been implemented with variants optimised for SAM and MAM calls.

Message characters are first protected using a short block code. Groups of codewords are assembled into numbered repeat blocks which represent the shortest message segments which can be repeated in the ARQ protocol. The repeat block length represents a trade-off between the overheads required to number and identify blocks and the wasted capacity in retransmitting a proportion of correctly received codewords within otherwise corrupted blocks. A bit-interleaving algorithm is employed to spread burst errors due to rapid fading and impulsive noise. The interleaving depth is also chosen with regard to the statistics of error burst duration typical of HF channels (3). The majority of bursts are spread such that the coding corrects the corruptions. Longer bursts, which would require deeper interleaving for immediate correction, are constrained within blocks and ARQ operation provides correct versions. This combination of an interleaving strategy (no throughput penalty) with an efficient random error correct/detect code is effective on HF channels which are characterised by error bursts and intervening random errors.

The repeat blocks are assembled, together with header information, into forward transmission packets lasting approximately 20s. In SAM calls a short, highly protected acknowledgement is sent by the destination station, following each received packet, requesting repeats where necessary. A time frame of 20s allows a timely response to be made to gross changes in channel conditions. To improve efficiency in MAM calls, several 20s forward packets are concatenated before the destinations are given an opportunity to respond.

During the detection process only part of the power of the block code is used to correct errors, the remaining power being reserved for detecting a greater number of errors which cannot be located. This further error detection is the first level input to the ARQ protocol. At high channel error rates, errors may escape detection and therefore affect message integrity. To guard against this a logical decision mechanism can request repeats of suspect message sections.

An intelligent overlay strategy is used to compare repeated blocks with the originally received version. Even where both versions contain detected (but uncorrectable) errors in the constituent codewords, so long as these do not coincide, the valid portions of each are combined to produce a wholly correct block. Again this matches the error patterns found on HF channels and means that more than one repetition of a block is most unusual.

Finally, for SAM calls, the ARQ protocol will request adaptive changes of transmission rate (every packet if necessary) to combat gross and prolonged changes in channel conditions and ultimately will request transfer to a new channel if the existing one fails.

Figure 5.1 summarises how the data link protocol offers a graded response as the channel deteriorates.

5.4 Performance of the Data Link Protocol

The HF trials link includes an implementation of the full FEC/ARQ link protocol, including the facility for adaptive changes of data rate.

As an illustration of a) how the channel conditions can change dramatically over very short periods and b) why coding is justified, Figure 5.2 shows the bit error rate per block measured during a single message. The average channel bit error rate is measured to be 1%, a figure often quoted as typical of HF channels although it is immediately obvious that the error distribution is far from Gaussian and hence such figures should be treated with obvious caution. In this example, on three occasions the error rate resulted in detected (ie uncorrectable errors) necessitating block repetition. It is noticeable that even in the case where the error rate rose to 50%, the next block was error free indicating the volatility of channel conditions. Further, for all the other blocks sent, FEC overcame the corruptions on the first transmission.

Analysis of trials results shows that the average throughput over a 24 hour period (measured over many continuous 24 hour periods) is equivalent to a continuous 32,800 characters per hour in the final system (64 bits/s net) against a maximum possible 53,000 characters/hour. With this throughput the character error rate has been maintained at better than 1 in 10^5.

6 CONCLUSIONS

From a careful analysis of the system constraints, and a thorough understanding of the propagation and interference characteristics of HF channels, it has proved possible to design an automatic and adaptive network. Low cost processing power has been used throughout to replace traditional operator skills, and extensive link trials have demonstrated a high level of performance which could not be achieved in a manual system.

A unique opportunity has been exploited to design, from fundamental principles, Frequency Management, Network Management and Link Management subsystems in an integrated design, which will provide a robust, reliable and yet low cost solution to a wide variety of user requirements.

Extensive link trials have provided early validation of key aspects of the system design.

7 REFERENCES

1) **E T Clarke**
Real Time Frequency Management in an Embedded Microcomputer
IEE Conf Proc 245 (1985) 23 - 26

2) **N F Wong, G F Gott, L W Barclay**
HF Spectral Occupancy and Frequency Planning
IEE Proc F, 132 (1985) 548 - 557

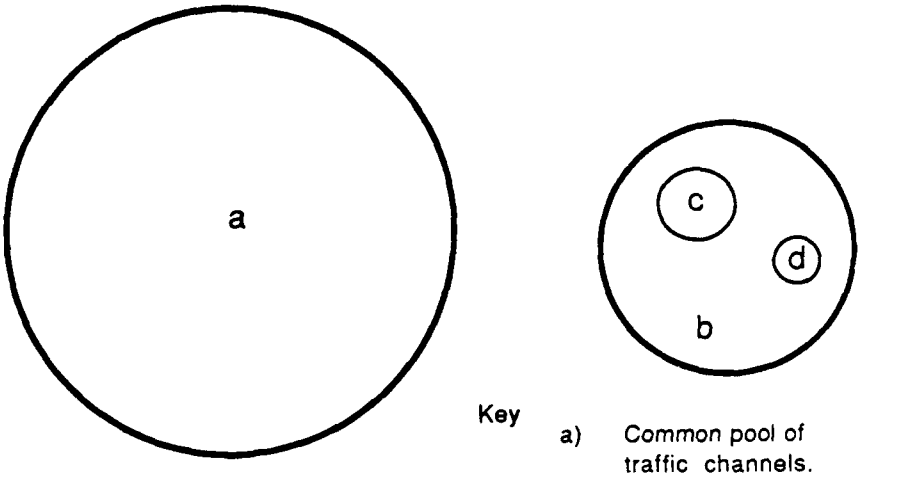
3) **D R Isaac, C T Spracklen**
Error Patterns and Real Time Correction Procedures for Data Transmission over HF Radio Links
IEE Conf 206 (1982) 106 - 110.

8 ACKNOWLEDGEMENTS

The authors would like to thank all members of the project teams, within both Plessey and the UK MOD, for their contributions to the design of the system described. Particular thanks must go to the staff at the Plessey research establishments at Roke Manor and Caswell for providing trials link facilities over an extended period.

The daily results from the soundings performed by the Slough ionospheric observatory and supplied by the Rutherford Appleton Laboratories (UK) have provided a most useful reference point when assessing the FMS subsystem.

©1987, The Plessey Company plc.



Key

a) Common pool of traffic channels.
b) Frequency management sounding channels.
c) Network management channels
d) Calling channels.

FIG. 2.1 Channel Usage To Minimise Interference Within The H.F. Network

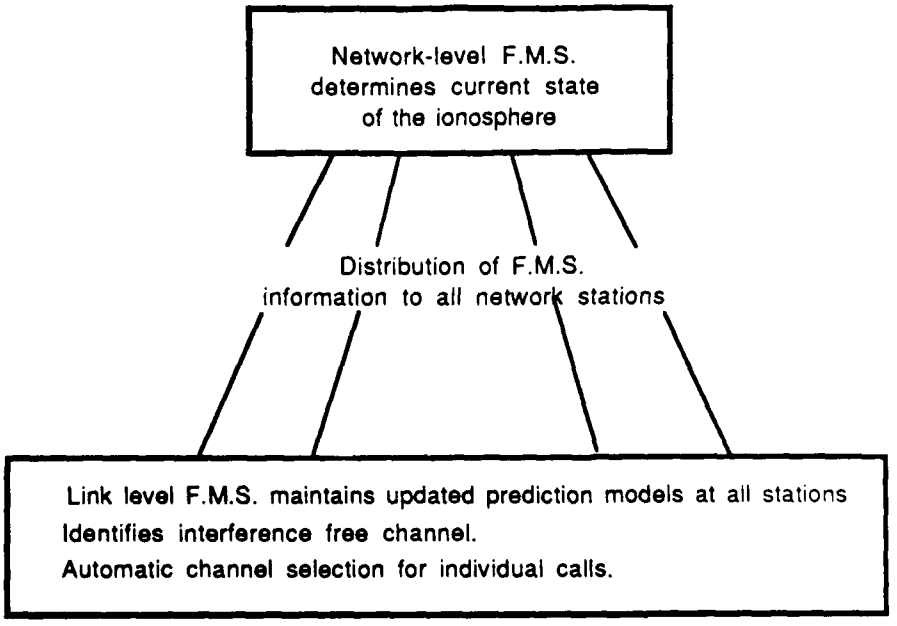


FIG. 3.1 Two Tier Structure Of H.F. Frequency Management Sub - System (F.M.S.)

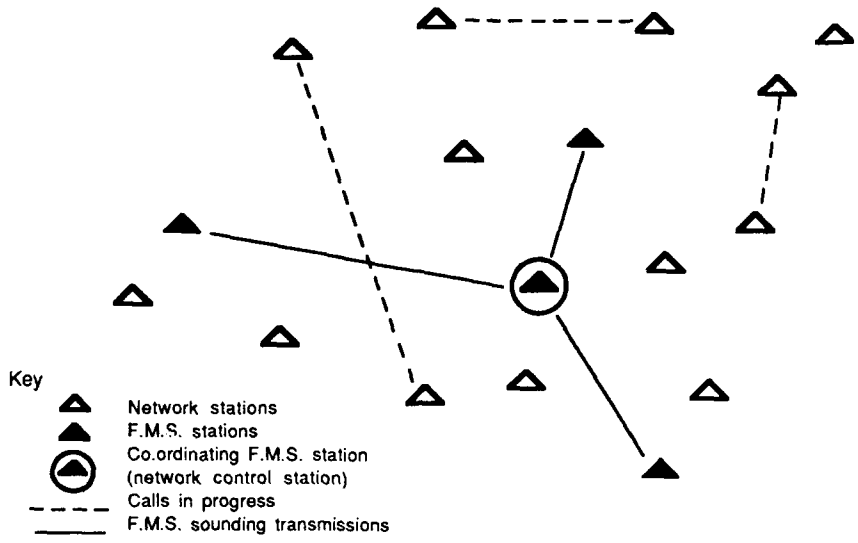


FIG. 3.2 Network Level Frequency Management Architecture

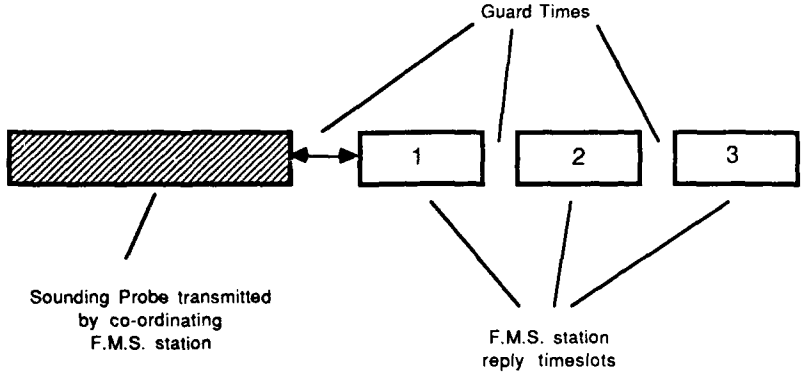


FIG. 3.3 Network Level Frequency Management System Sounding Protocol

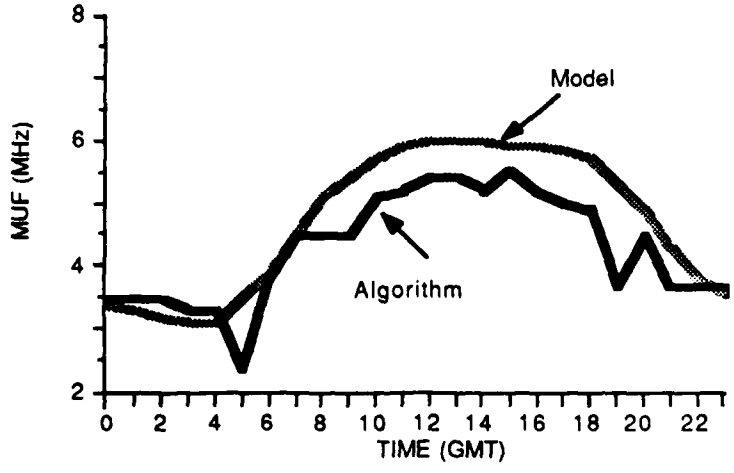


FIG. 3.4 Disturbed Day - Muf Below Normal (31/3/87)

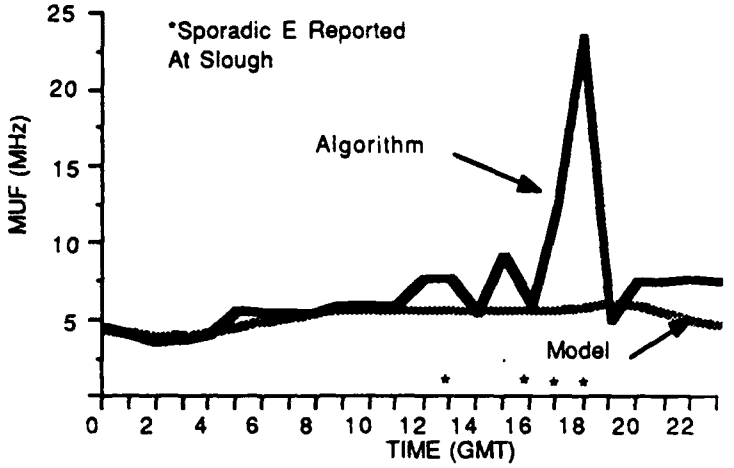


FIG. 3.5 Disturbed Day - Sporadic E (28 /5/87)

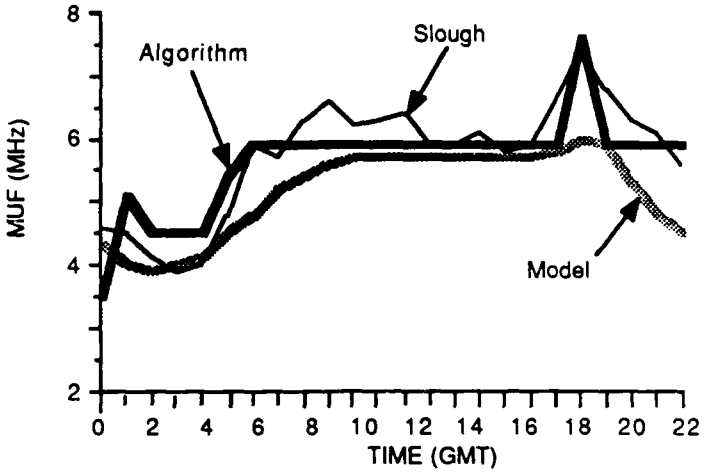


FIG. 3.6 Quiet Day (22 /5/87)

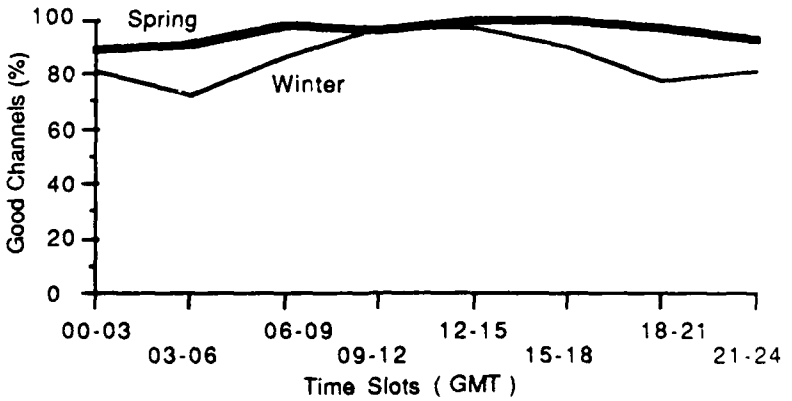


FIG 3.7 Proportion Of Candidate Traffic Channels Which Propagate And Are Not Blocked By Interference

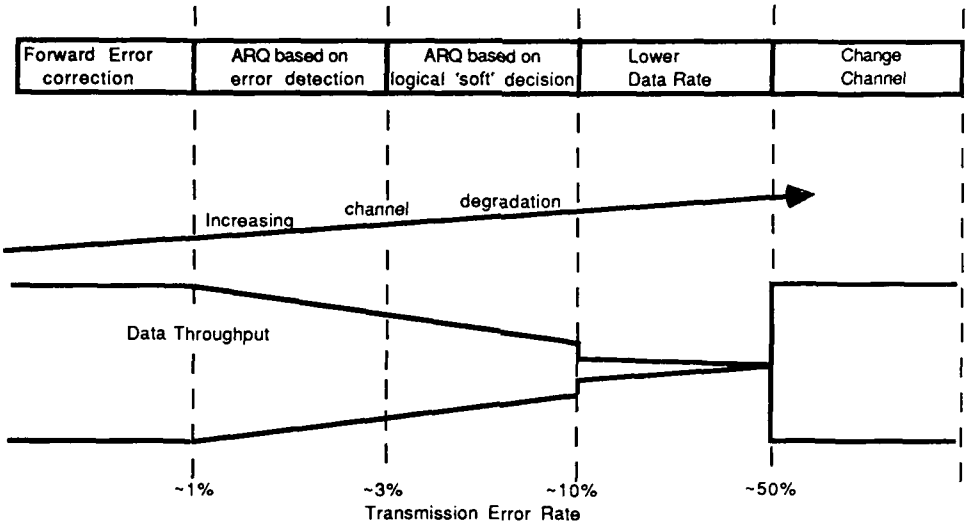


FIG. 5.1 Graded Response Of Data Link Protocol As A Channel Degrades.

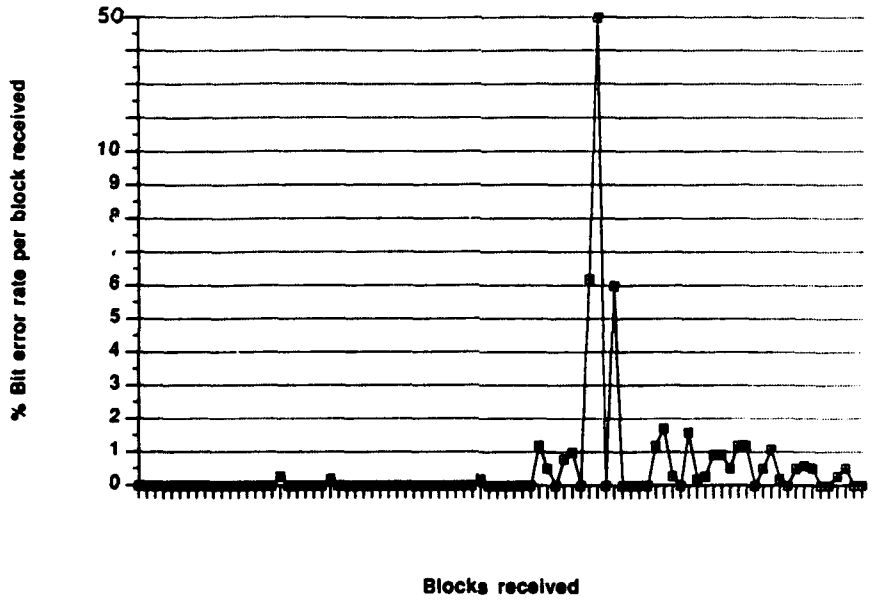


FIG. 5.2 Measured Error Rate Per Received Block

EXPERIMENTAL ROBUST HF MODEMS INCORPORATING ADAPTIVE EXCISION

by
Geoffrey F Gott
Paul Doany
Siu W Wong
E Philip Darbyshire
University of Manchester Institute of Science and Technology
Manchester
England

SUMMARY

The paper describes two 75 bits/sec modems which are able to operate in the presence of severe interference from other HF users, frequency selective fading, and noise. One modem is a frequency hopper (with a total bandwidth of about 200 kHz in our case), and the other occupies a single voice channel.

INTERFERENCE PARAMETERS

Statistical properties of spectral occupancy on short range (1000 km) HF links have been studied by Dutta (1), who examined many 50 kHz sections of spectra for fixed users. One aspect of this work was the introduction of parameters that provide a basis for determining the performance of communication systems in the presence of HF interference. An important but elementary parameter is a measurement of occupancy called 'congestion', which is defined below.

Congestion, Q, is the probability of finding, by random, a given bandwidth within each 50 kHz spectrum, where the average interference level exceeds a defined threshold. Dutta initially chose the bandwidth to be 100 Hz, and the threshold level to be -125 dBm.

Examples of congestion values are summarized in table 1. These results were taken over a period of one week, approximately at the winter solstice, during a period of low sunspot activity. The antenna was an active horizontal dipole.

It is evident that congestion was consistently high at night and dusk as indicated by the low value of standard deviation with respect to the mean. It was less at dawn, and at frequencies near the OWF during day, and least at daytime at frequencies well below the OWF, where the non-OWF frequency indicated is halfway between the OWF and the LUF, for 1000km range. This latter point suggests that it may be better to operate below the OWF during the day, and accept some loss in propagation quality in exchange for less interference from other users.

RELEVANCE TO DATA TRANSMISSION

The parameter congestion may be used to determine the probability of HF signals avoiding interference, assuming that such signals may be placed anywhere within each 50 kHz spectrum examined. Since the spectral measurements were made with a resolution bandwidth of 100 Hz, the analysis is restricted to slow rate data signals (50 - 100 bauds) which use such bandwidths.

For example, the probability of finding a 100 Hz bandwidth for a slow rate DPSK signal, where the interference is less than -125 dBm, is simply equal to 1-Q. This is defined as the availability of the signal.

Similarly, for frequency exchange keying (FEK), where the frequency shift exceeds about 1kHz, which is sufficient frequency separation for the tones to be affected independently by interference (1), the availability is $(1-Q)^2$.

For FEK with a frequency shift significantly less than 1 kHz, the tones will not be affected independently by the interference, and the availability will be greater than $(1-Q)^2$. For example, for a frequency shift of 85 Hz, values of availability have been determined experimentally (2) and are plotted along with those for DPSK and wide-shift FEK in figure 1, for the range of measured congestion values. Figure 1 shows that narrow-shift FEK is only marginally worse than DPSK.

THE USE OF FREQUENCY DIVERSITY

The application of the parameter congestion has been extended to the availability of robust formats with a high level of frequency diversity. For example, if M diversity tones are transmitted simultaneously, with at least 1 kHz separation between adjacent tones, then the probability that at least two will be received free of interference is given by (2)

$$P(2,M) = 1 - Q^M - M(1-Q)Q^{M-1}$$

The significance of at least two tones being interference free is that frequency diversity is preserved, and protection achieved against frequency selective fading, after the interfered tones are rejected. For M = 6 and Q = 0.5, P(2,6) = 0.9, inferring a 90% availability in severe congestion.

AN ADAPTIVE FREQUENCY HOPPING MODEM

SIGNAL FORMAT

Thus a signal format which may be expected to operate well in interference could comprise of 6 narrow-shift FEK pairs, with 1 kHz separation between adjacent pairs, as shown in figure 2. Narrow-shift FEK is preferred DPSK because it is easier to implement, and more robust in operation.

To obviate the problem of total signal loss in severe and unfortunately structured interference, the 6 tone pairs may be hopped in frequency. The signal format will then encounter the average congestion value across the hopping bandwidth. If the tone pairs are hopped as a single group, then a 6 kHz hop in carrier frequency would ensure uncorrelated interference conditions, and the use of forward error correcting codes would protect data on hops where 5 or 6 tone pairs may be lost.

It is important that severe interference is not permitted to enter the decision circuit of the detector. This may be achieved by examining the interference spectrum on each hop, prior to transmission occuring on that hop, and rejecting those FEK signals which are interfered with, when data is subsequently received on that hop. This 'look ahead' interference analysis technique may be realised by a second receiver, which hops one step ahead of the main receiver, and whose purpose is to examine interference only (3).

The FEK tones may be phase locked to give a lower peak-to-average power ratio than might be expected from a conventional parallel tone format. Indeed, by correct phase locking, the peak-to-average power ratio of the 6 transmitted tones may be reduced from a worst case value of 10.8 dB, to a value of less than 6 dB (2).

THEORETICAL DIVERSITY PERFORMANCE

The theoretical performance of a square-law combiner has been derived for binary non-coherent FEK in slow Rayleigh fading and white Gaussian noise, the probability of error being given by (4)

$$Pe(M) = \sum_{m=0}^{M-1} \frac{(2M-1)!(-1)^m}{(M-1)!(M-1-m)!m!(M+m)} (2+x)^{-M-m}$$

where M = number of diversity branches

x = long-term average normalised SNR per diversity branch

The analysis may be extended to the sixth order diversity excision detector (2). This includes the effect of interference, and gives the probability of error after excision as

$$Px(6) = 0.5Q^6 + 6(1-Q)\,Q^5 Pe(1) + 15(1-Q)^2 Q^4 Pe(2)$$
$$+ 20(1-Q)^3 Q^3 Pe(3) + 15(1-Q)^4 Q^2 Pe(4)$$
$$+ 6(1-Q)^5 Q\, Pe(5) + (1-Q)^6 Pe(6)$$

This is plotted for various values of congestion in figure 3.

While it may be necessary to operate with 6 channels for Q = 0.5, lower redundancy is acceptable at lower congestion. For Q = 0.1, for example, dual diversity is sufficient, and hence the data rate may be increased by a factor of 3. It is thus possible to operate with an adaptive data rate chosen relative to the severity of interference. This is simplified because the average values of congestion are predictable.

HF LINK TESTS

A frequency hopping modem with 6 FEK channels has been tested at low power on an 800 km HF link (5), when it was compared with a standard FEK system (75 bauds keying rate, single tone pair, 850 Hz frequency shift). Messages of about 2000 bits were transmitted alternately via the standard FEK modem, and via the frequency hopper. The hopper used 30 hops, each representing a 6 kHz increment in carrier frequency, and of duration 1 second. The sixth-order diversity format was thus incremented uniformly once only across a total bandwidth of about 200 kHz during the message transmission. This is a simple hopping pattern, but the hopping parameters are largely arbitrary.

Figure 4 shows very severe interference across a bandwidth of 200 kHz, with the frequency hopping signal indicated where it fell, on one specific hop, into a gap in the interference spectrum. The spectrum has been measured rapidly (500 mS) with a wide resolution bandwidth (1 kHz), and the fine resolution of the spectrum gaps is therefore not visible. The reason for such measurement parameters is that the hopping period was 1 second, and therefore the hopping signal will be caught on a single hop position by the 500 mS sweep. It can be seen that the hopping signal level is much lower than the majority of interference peaks - about 50 dB below the maximum value - but frame synchronisation was acquired and the message error rate for this particular example was about 0.1.

The hopper has a different type of performance to that of the standard FEK signal. In particular, the FEK signal tends to have good performance in the absence of significant interference, and very poor performance in the presence of severe interference. The hopper on the other hand cannot encounter severe interference for the whole message and tends to have a low error rate under severe interference conditions, although zero error rate per 2000 bit message is not frequently achieved.

This type of performance is illustrated in the example of results shown in figure 5. This plots the error count for each 2000 bit message against number of messages for both the standard FEK system and for the hopper, over about a two hour period. The FEK signals and hopping signals were interleaved in time, and were transmitted at an equal average power.

Forward error correcting codes are effective on the hopping format, since the error rate is not typically very high. Also the errors typically occur in bursts of duration equal to the hopping period, so that the degree of interleaving becomes self evident. Excellent improvement has been achieved using the Golay (23,12,3) code, interleaved relative to the hopping boundaries (5).

CONCLUSIONS ON THE FREQUENCY HOPPER

The signal design for the adaptive frequency hopping modem has been derived by using elementary parameters of HF interference, and the link test has confirmed that the adaptive hopper offers a method of achieving useful communication under very severe interference conditions, without the use of a control link. The frequency diversity also offers good performance in fading.

AN ADAPTIVE CHIRP MODEM

INTRODUCTION

A second adaptive modem uses differential phase shift keying (DPSK) of swept frequency chirp signals (6,7). The data rate is 75 bits/s and the swept bandwidth is 2.7kHz, occupying a voice channel.

The DPSK format and detector are shown in figure 6 (7). The detector includes an adaptive filter which can selectively reject parts of the signal spectrum. This is followed by a conventional matched filter for the chirp signal, which in turn is followed by the differential phase measurement.

The detector also includes an interference assessor, which measures interference levels in each of 16 contiguous sub-channels, spread uniformly across the voice channel. Sub-channels containing severe interference are then omitted from the overall adaptive filter response, thus suppressing the interference and the corresponding parts of the chirp signal spectrum. However, in severe interference, the signal to interference ratio will be enhanced, and thus the rejection will reduce the error rate when the power spectral density of severe interference within the voice channel is significantly non-uniform.

CHIRP SIGNAL PARAMETERS

```
 Start frequency  = 300Hz
  Stop frequency  = 3000Hz
         Sweep F  = 2.7kHz
   Bit duration T = 13.3ms
Processing gain FT = 35.94 (15.6dB)
```

CHIRP SIGNAL DETECTOR

In the initial implementation of the chirp detector of figure 6, the interference assessment is achieved prior to chirp signal transmission, and the rejection configuration remains unchanged during the 2000 bits of each chirp message. Thus the interference assessment filter and the adaptive rejection filter are not needed simultaneously, and may be synthesised by the same transversal filter.

The assessment of the interference takes place at the beginning of each message. Each of the 16 contiguous rejection filter responses is loaded in turn from the microprocessor memory into the transversal filter memory. The interference level at the output of each of the 16 filters is determined, and stored. The microprocessor then computes the rejection pattern required, and synthesises the required network.

EXCISION NETWORK

The system uses 512 point programmable transversal filters to achieve rejection and matched filter operations. Interference rejection in excess of 50db is achieved with a linear phase characteristic and low passband ripple.

Examples of rejection responses are shown in figure 7. The effect of the notches is to disperse the energy of the compressed pulse. However, this is generally accommodated in the integration window at the multiplier output of the differential phase detector of figure 6, together with multipath components. In white noise only, the use of a wide integration window degrades the system (2), but with multipath propagation, multipath components will be coherently combined in the resultant integration, with their amplitudes squared, which is advantageous.

In the experimental system, an integration window of 3.3ms was used. In white noise alone and with no rejection, this gives the chirp system a performance which is approximately equal to that of the non-coherently detected FEK signals. Ideally of course, the DPSK chirp has a 3dB advantage over non-coherent FEK in noise.

CHIRP MODEM LINK TESTS

The system is being tested over a 170km sky-wave HF link, in which 2000 bit chirp messages are interleaved with frequency exchange keyed (FEK) messages. The FEK detector incorporates no interference rejection, and the received chirp signals are detected both with and without rejection.

Without rejection, the overall chirp performance is worse than that of FEK. This is because the processing gain of the chirp signal is typically too small to give useful rejection of interfering signals falling within its bandwidth. With rejection, the overall performance of the chirp system is better than that of FEK.

The transmitter power is almost always less than 1w, and occasionally as low as 20mw, as measured at the input of a Racal RA1003 Difan wideband antenna.

The receiver is a Racal RA1792 fed by a wideband active horizontal dipole antenna. This antenna is close to the ground, and has a polar diagram appropriate for the reception of high angle skywaves.

PRESENTATION OF RESULTS

The field trial results are displayed as accumulative error curves, an example of which is given in figure 8. This displays the total number of errors against number of messages received for a 2 hour test period. In addition, the adaptive filter response is indicated for each message, by a chart below each set of error curves, and which follows the same time scale as the error curves. For these filter response plots, a '*' indicates that the frequency slot in question, (1 to 16), was removed from the overall response, whereas a '-' signifies inclusion. The corresponding FEK mark and space frequencies (m and s), are also indicated.

The system having the best performance is that which possesses the longest, lowest, and most flat curve. This signifies a good syncronisation performance (the length of the plot indicating number of messages for which synchronisation was achieved), together with a consistently low error rate.

The interference spectrum is often remarkably stable with time, with virtually identical rejection patterns being determined for periods up to 2 hours. However, figure 8 shows more variable interference conditions, when continuous assessment of the interference and more frequent updating of the adaptive filter would be advantageous.

CONCLUSIONS ON THE CHIRP MODEM

The excision of interfering signals from within the swept bandwidth of a chirp signal has been achieved on an experimental HF link, using modern signal processing techniques. In comparison to FEK, the adaptive chirp system is instrumentally complex, and the use of DPSK implies further complexity due to the requirement of accurate mixer oscillators for signal generation and reception. However, the adaptive chirp system seems to reliably overcome the two main problems of HF data transmission, namely, interference from other users, and multipath propagation.

The modem also has the property of being able to coherently combine multipath components which fall within the detector integration window, weighted by the square of their amplitudes, although this advantage must be traded against reduced performance in additive white Gaussian noise.

An important improvement which has now been realised is the continuous assessment of the in-band interference during chirp signal reception, and the frequent updating of the adaptive rejection filter, also during chirp signal reception. Also, further work which has reached an advanced stage, is an increase of the data rate to 600 bits/s.

ACKNOWLEDGEMENTS

The authors thank the Procurement Executive, Ministry of Defence for supporting this work. The hopper was supported by RAE, and the chirp modem by RSRE, and we acknowledge the help given by members of these establishments.

REFERENCES

1 Gott G F, Dutta S and Doany P; 'Characteristics of HF interference with application to digital communications', Proc IEE, part F, August 1983.

2 Doany P and Gott G F, 'Theoretical performance of an adaptive frequency hopping modem at HF', Proc IEE, part F, August 1986.

3 Gott G F, Doany P and Sepehri M, Patent GB2071464B relating to frequency hopping communication systems, published May 1984.

4 Pierce J N; 'Theoretical diversity improvement in frequency shift keying', Proc IRE, vol 46, May 1958.

5 Wong S W; 'A frequency hopping modem at HF', PhD Thesis, UMIST, 1985.

6 Gott G F and Karia A J; 'Differential phase shift keying applied to chirp data signals', Proc IEE, November 1974.

7 Darbyshire E P and Gott G F; 'Robust data transmission at HF', IERE International Conference on 'Radio Receivers and Associated Systems', Bangor, July 1986.

	congestion	standard deviation
Dawn (OWF)	0.06	0.025
Day (OWF)	0.26	0.10
Day (non OWF)	0.09	0.11
Dusk (OWF)	0.43	0.15
Night (OWF)	0.46	0.06

Table 1 Experimental values of congestion

(OWF infers 1000 km range)

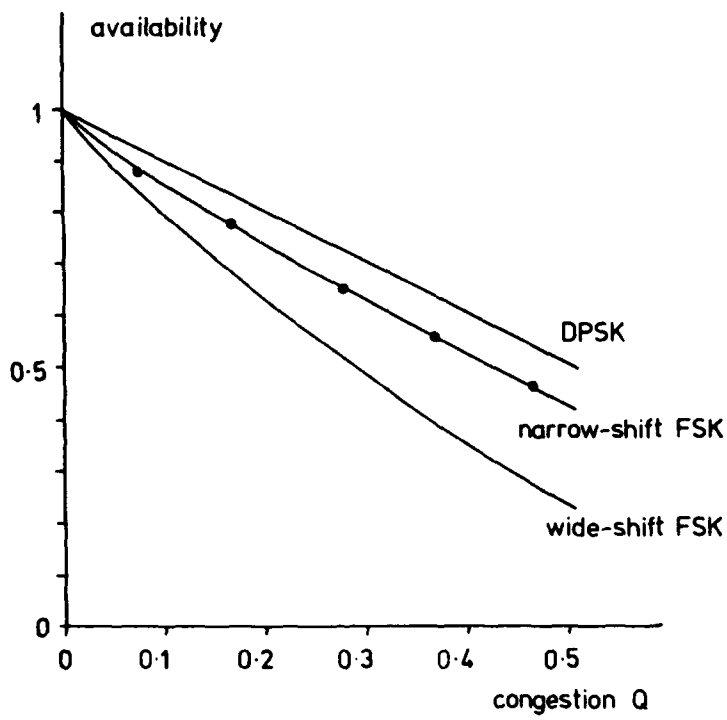


Figure 1 Availability of several signal formats

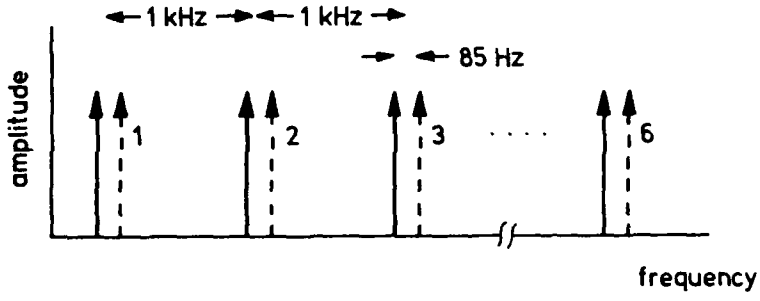


Figure 2 Sixth-order frequency diversity FEK format

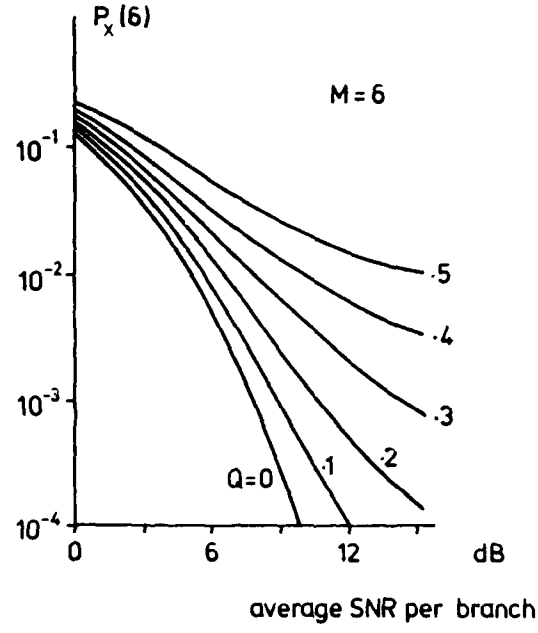


Figure 3 Error probability for the frequency hopping system in the presence of interference (excision assumed), slow frequency selective fading, and noise

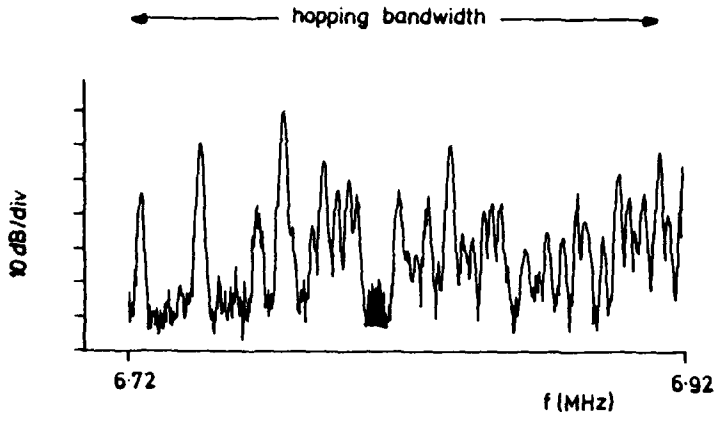


Figure 4 The frequency hopping signal (shaded) in the presence of severe interference, as received on an 800 km link

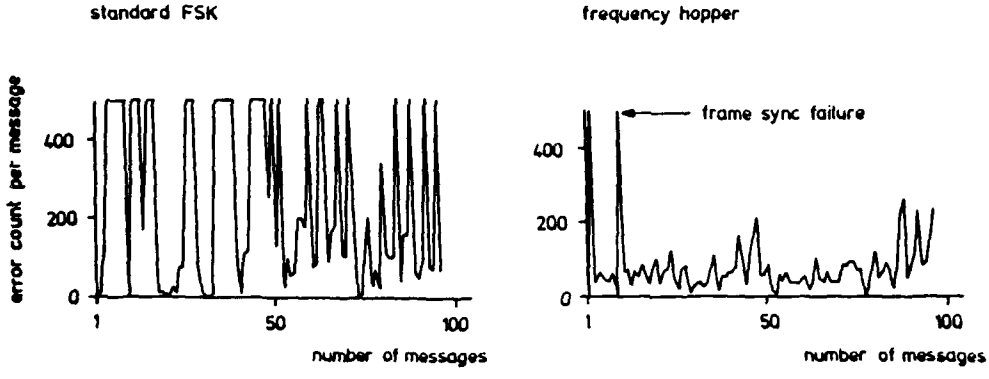


Figure 5 An example of the comparison of standard two tone FEK and the frequency hopping modem, on an 800 km link

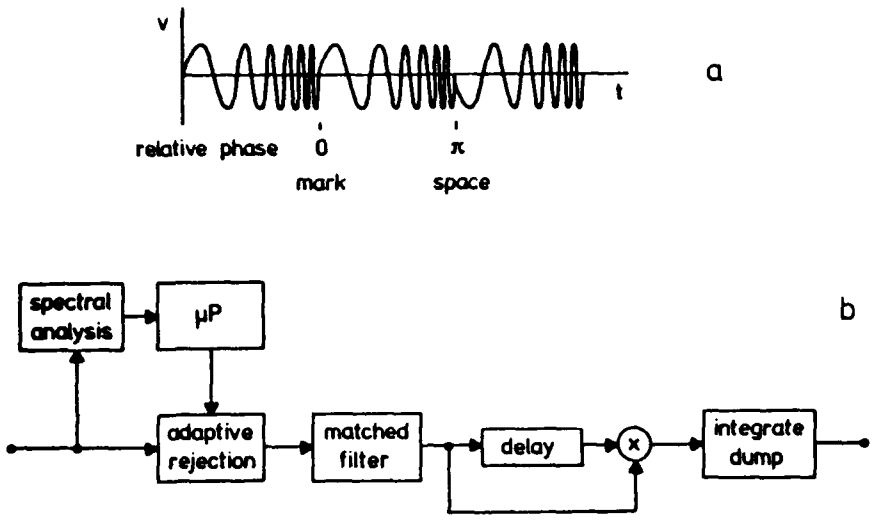


Figure 6 a) DPSK chirp signals b) Detector

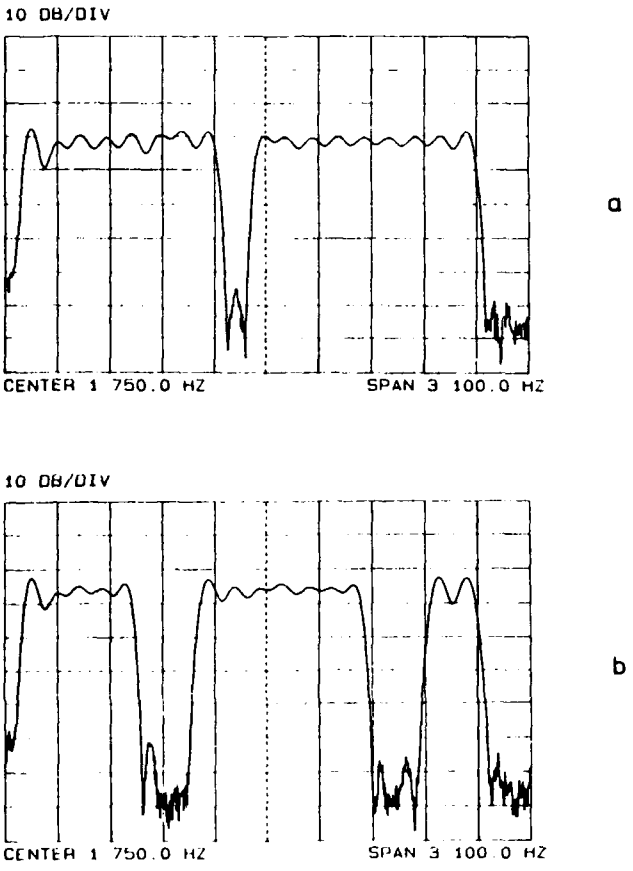


Figure 7 Experimental amplitude/frequency responses for the rejection filter

a) One filter rejected b) Four filters rejected

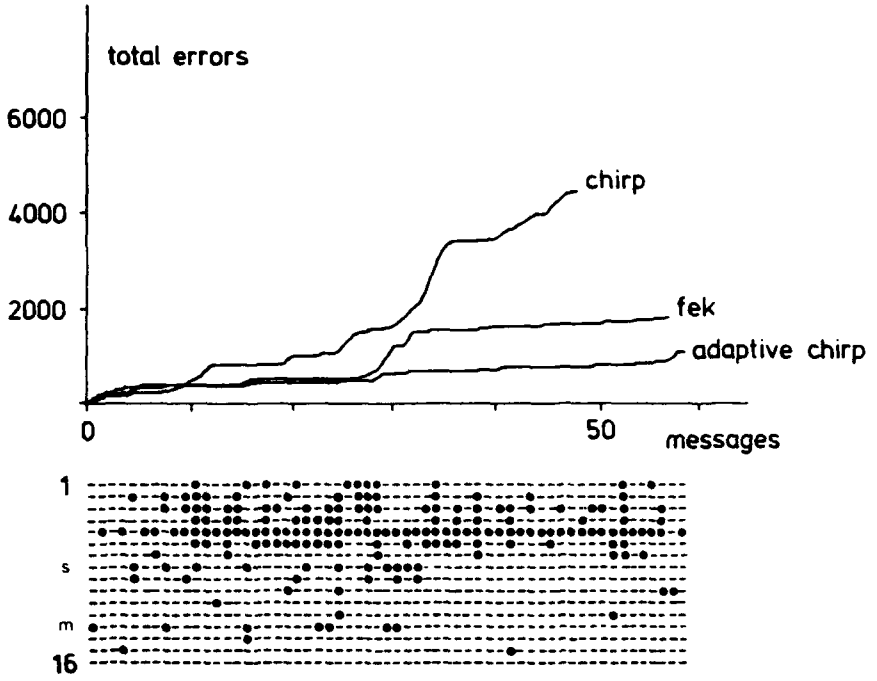


Figure 8 Chirp and FEK comparisons on a 170 km link

ARTIFICIAL INTELLIGENCE IN H.F. COMMUNICATION SYSTEMS

A.P.Jowett
Department of Electronic Engineering
University Of Hull
Hull HU6 7RX
U.K.

SUMMARY

The current status of H.F. communications system design is reviewed. Attention is given to the constituent components of H.F. systems. The need for automatic system control and frequency management is presented. A new method of in-band R.T.C.E. is described and a proposed co-channel interference model is included. A formalised H.F. systems design philosophy is presented in order to achieve optimum performance in noise and interference environments, and compatibility with other types of communication system.

1. INTRODUCTION

1.1 Overview

Currently, the majority of H.F. radio communication systems are controlled by human operators. Control data is supplied via long-term, off-line channel analysis procedures. Such communication systems operate with limited or no adaptivity, e.g. fixed transmitter power, a single modulation format, and a fixed transmission rate. System control and frequency management exercised using these techniques is non-optimum, as is outlined below (Darnell, 1986a).

The channel capacity of an H.F. link can vary from tens to a few thousand of bits/s within a short time interval. A non-adaptive control regime cannot react to , or exploit, these variations. Periods of high channel capacity would thus be wasted. Also, human operators are prone to errors and have a slow response to changes in path conditions. This again results in wasted capacity.

Long-term, off-line channel analysis procedures give reasonably accurate monthly median parameters predictions for a path. However, significant departures from these monthly values can be expected on a minute-to-minute or second-to-second basis. Thus a system will not be operating efficiently if it relies upon such long-term predictions alone.

There is no adequate model of man-made, co-channel interference, which can be the limiting factor on system performance in spectrally congested, mid-latitude regions.

In order to manage an H.F. communications link efficiently it is necessary to employ automatic frequency management and system control techniques via :-

1) Embedded propagation, noise, and interference models.

2) Embedded real-time channel evaluation (R.T.C.E.) techniques.

3) Adaptive system components to facilitate variation in transmission and reception parameters, e.g., transmitter power, modulation format.

Effective system control and frequency management also requires the integration of the above components in a structured and systematic way. This will facilitate easier maintenance, and will allow testing of new system components and control methods. Such an approach may also allow compatibility with other communication systems to be achieved. This paper describes a structured approach to H.F. communication system design in order to incorporate the above features.

The overall system heirarchy adopted is that of the International Standards Organisation - Open Systems Interconnection (I.S.O.-O.S.I) 7-layer model for data transfer between computers. In the proposed H.F. system, the individual layers of the O.S.I. model are each implemented as an Expert System. This design philosophy allows modularity to be achieved at macro and micro layer levels. As stated in the above paragraph such a structured approach will allow easier testing of new channel analysis and prediction methods, interference models, system components etc. in order to combat the harsh noise and interference environment associated with H.F. radio communication links.

1.2 Description Of The Paper

Section 2 outlines the current status of embedded analysis techniques. Section 3 looks at the updating of such models via passive and active channel measurements. Section 4 describes the current status of embedded R.T.C.E techniques with attention being given to the various types. Section 5 outlines a new and potentially effective method of in-band R.T.C.E., called Template Correlation. Section 6 deals with the current status of co-channel interference models. Section 7 presents one such model with system design applications. Section 8 attempts to integrate the above elements of H.F. communication system design in a structured and effective design methodology involving Artificial Intelligence techniques.

2. CURRENT STATUS OF EMBEDDED ANALYSIS TECHNIQUES

The nature of embedded analysis procedures and the current 'state of the art' are given by (Bradley, 1979), and (Thrane, 1986). A description of the current status of such techniques now follows.

In order to control H.F. sky-wave communication links successfully, it is necessary to be able to predict propagation, noise, and interference conditions. Such channel assessment is carried out by channel analysis procedures and models. Currently, these models perform predictions via a combination of information from large databases of ionospheric measurements and empirical equations governing the behaviour of the ionosphere, coupled with models of the ionosphere. The databases are compiled by large international organisations , such as The International Union Of Radio Science (U.R.S.I.) or The International Radio Consultative Committee (C.C.I.R.).

The output of analysis procedures is in the form of monthly median estimates of path parameters as a function of time or frequency, or both, e.g., maximum usable frequency (M.U.F.), signal strength, signal-to-noise ratio. Often a statistical measure of circuit reliability is provided e.g, predicted signal-to-noise ratio versus required signal-to-noise ratio. Noise and interference data is provided by means of a database of measurements for the time of year, sun-spot number, site classification (urban or rural) etc. An example output from an analysis procedure is shown in Figure 1. Examples of current analysis procedures are IONCAP, MUFLUF, FTZ, LIL 252, and MINIMUF. They vary in size, accuracy, and execution speed.

The limitations of current models are :-

1) They are derived from limited databases.

2) They treat effects such as sporadic E, and sudden ionospheric disturbances (S.I.D.) in a simplistic manner.

3) There is no adequate method of modelling co-channel interference.

4) Output is in the form of long-term, monthly median figures.

The H.F. medium is time varying as explained in Section 1. Hence a further requirement for such models is that they should possess the ability to be updated via link specific data from, for example, real-time channel evaluation routines. More is said about this in Section 3 of the paper.

The term "embedded" implies that the channel analysis procedure should be of a size that is compatible with incorporation into the overall control algorithm. Hence, large and complex procedures requiring correspondingly large computing resources for execution are not considered to be useful in this context.

3. POTENTIAL FOR UPDATING OF CHANNEL ANALYSIS MODELS

As stated in Section 2, one limitation of channel analysis procedures is that they use information from limited databases when making predictions. For any specific H.F. radio communication link (other than that from which the data was derived), it is necessary to increase the accuracy of such predictions.

One method of enhancing the accuracy of analysis procedures is to update them via specific channel measurements. An example of this is described by (Uffelman et al, 1984). The method of updating a propagation analysis model in this example is outlined below.

Chirpsounder data from several oblique sounders is used to measure the M.U.F. between the base station and the sounders. A comparison is made between the measurements and the predictions from the propagation prediction procedure. The predicted M.U.F. curve is then made to match the measured M.U.F. curve for a point in time by iteratively

altering the input sunspot number to the prediction routine. This so-called "pseudo sunspot number" is then used for all predictions until the next update, 24 hours later.

The results from this example show that the error between predicted and actual M.U.F. can be reduced to that of the numerical error associated with the prediction algorithm used. The following improvements to this technique are suggested by the author :-

1) Make use of known broadcast stations to provide M.U.F. data via passive monitoring of the station's frequency allocation. Figure 2 illustrates this point. Ionospheric conditions are relatively stable over periods of up to one hour. If broadcast stations were monitored approximately up to 1-2 hours earth's rotation to the east, it should be possible to estimate the M.U.F. between the passive receiver and the broadcasting station. Thus estimation of ionospheric conditions could be achieved up to one hour before they occur.

2) Investigate the use of parameters other than sunspot number for manipulation when updating the analysis model.

3) Update more frequently then once every 24 hours. This should further improve the accuracy of the prediction.

Other methods of updating channel analysis procedures and models include time and frequency shifts of M.U.F. profiles. This again requires measurement of the actual M.U.F. to be made.

4. EMBEDDED R.T.C.E. TECHNIQUES

4.1 Introduction

The current status of embedded R.T.C.E. techniques is described by (Darnell, 1986b). R.T.C.E. is the name given to the process of path parameter identification and modelling. It has been shown to be a pre-requisite for optimal H.F. communication system control. Embedded R.T.C.E. requires that the R.T.C.E. algorithms are incorporated into the system at the design stage, and that they use the same R.F. units and antennas as the actual communications traffic. This type of R.T.C.E. is thus more economically viable than R.T.C.E. systems involving dedicated equipment.

Long-term, off-line channel analysis procedures cannot provide communication circuit parameter forecasts for H.F. communications where high reliability and availability are a requirement, as they have several important limitations (see Section 2). Such requirements call for accurate characterisation of the communication path in real-time. This can be provided via embedded R.T.C.E.

4.2 General Points

1) The term "real-time" implies that the measured channel parameter values are updated at intervals less than the overall response time of the communications system to control inputs.

2) The output must be in a form that is meaningful to the overall control algorithm.

3) Both propagation and noise/interference data should be produced by the R.T.C.E. algorithm.

4) An R.T.C.E. algorithm should give information on the optimum working frequency (O.W.F.) and optimum start and finish times for transmission.

5) The R.T.C.E. algorithm should make use of the "other than normal" propagation modes, e.g., sporadic E, and meteor burst.

6) The channel modelling method employed should be appropriate to the class of user. Hence the model for 75 bits/s telegraphy will be less complex than that for 2.4 kbits/s digitised speech.

A generalised R.T.C.E. algorithm is shown in Figure 3.

Embedded R.T.C.E. algorithms tend to be of two types :-

- Those operating within a limited set of channels assigned to a system for communications purposes.

- Those operating within a given assigned channel.

There now follows an overview of the various methods of R.T.C.E.

4.3 R.T.C.E. Algorithms

When choosing an R.T.C.E. algorithm, the following points should be considered :-

1) Can the necessary R.T.C.E. input information be gleaned from the normal operating signals for the communications channel ?

2) If mixing of R.T.C.E. signals with traffic signals is required, what proportion of the transmission power, bandwidth and time will these occupy ?

3) To what extent can passive, rather than active, R.T.C.E. techniques be employed? (such techniques reduce spectral intrusion.)

The R.T.C.E. algorithms presented below have been divided into two groups :-

- Alternative channel R.T.C.E.

- Current channel R.T.C.E.

Alternative channel R.T.C.E involves the assessment of the communication system's other frequency assignments. These techniques involve active probing of the channels.

Current channel R.T.C.E. involves the assessment of the channel that is being used for communication purposes. It is advantageous if the channel evaluation information can be derived from the normal operating signals of the channel. The following points apply to current channel evaluation :-

1) Assessment of the current channel in use relative to the other communication channels must be performed.

2) The optimum positioning of signals within the channel needs to be determined.

3) The optimum signal processing to be applied to a traffic signal for transmission over the channel must also be determined.

4.3.1. Alternative Channel R.T.C.E

CHEC - CHANNEL EVALUATION AND CALLING

This was developed for mobile-to-base communications. On each of the assigned channels a five second probing sequence is radiated from the base station containing a selective calling code, data on the average noise level at the base in the channel, and a C.W. section (see Figure 4). The remote receiver is alerted by the calling sequence and it uses the noise data and the C.W. section to compute the average signal-to-noise ratio for it's own transmissions to the base (assuming propagation reciprocity, and allowing for the differences in antenna gains and transmitter powers). The optimum channel is thus defined to be the one with the highest predicted signal-to-noise ratio. In practical trials, this method gave significant improvements in channel availability and reliability. There are now many similar systems, e.g., RACE and ALQA, which give multi-parameter analysis of channels.

R.T.C.E. BY PILOT TONE MEASUREMENT

A simple C.W. pilot tone is inserted at a suitable position within the channel bandwidth. At the receiver the pilot tone is detected in a narrow band-pass filter. It's phase variations are analysed over a period of time, and use is then made of theoretical relationships between phase instability and data error rate to infer the suitability of the channel for various types of traffic. This method is intended to be used for mobile-to-base station transmissions. The same assumptions are made as for the CHEC system (path reciprocity, and allowance for different effective radiated power of the base station and the mobiles). It's main advantage is it's simple implementation. However, the time required for channel evaluation is relatively long.

R.T.C.E. BY ERROR COUNTING

A digital sequence which is known at both transmitter and receiver is sent over the channel . The number of errors occurring at the receiver is counted. The optimum channel is thus the one with the lowest error count. Again reciprocity of propagation is assumed. This technique requires synchronisation of both ends of the link in terms of time and frequency. It can be used together with most forms of communications traffic. The main disadvantage of this method is the time taken to accumulate R.T.C.E. data in low data rate systems. This can be overcome to some extent by a technique known as "pseudo-error counting". This uses an over-sensitive detection method so that the error rate used for error analysis is greater than that for the actual traffic. The accuracy of this method can be improved by use of soft-decision thesholds.

4.3.2 Current Channel R.T.C.E.

It is advantageous if R.T.C.E. data can be derived from normal operating signals in the channel to allow assessment of the channel currently in use. This section describes such R.T.C.E. algorithms.

IN-BAND R.T.C.E.

Each 3 khz-wide channel is monitored and periodically the channel spectrums are obtained. Regions of lowest power content are then looked for and labelled as being useful for reception if the received power level is below a threshold value. This is a useful technique when the signal format employed divides the channel into sub-channels, e.g. with multi-tone F.S.K. For F.S.K. systems, individual sub-channels can be evaluated when the tone corresponding to that sub-channel is switched off.

R.T.C.E. IN A.R.Q. SYSTEMS.

Here the number of packet repeat requests in a time interval is used as a measure of the channel's quality.

R.T.C.E BY TRAFIC SIGNAL MODIFICATION

This method adds extra error control redundancy to the data by encoding into codewords which are designed to yield R.T.C.E. information at the error control decoder.

R.T.C.E. BY ASSESSMENT OF SYNCHRONISATION QUALITY

In order to achieve synchronisation of transmitter and receiver, a synchronisation preamble is often included at the start of each transmission period. This usually takes the form of a digital sequence with a highly impulsive autocorrelation function, e.g. a Barker sequence. The level of the correlation peak from such a pre-amble can be used to provide channel evalution data.

4.4 Summary - The Potential Advantages Of R.T.C.E.

The potential advantages of R.T.C.E. can be identified as :-

1) The elimination of channel analysis procedures and models for operational purposes. They are however still required for planning and initial start-up.

2) It can characterise co-channel interference to some extent.

3) Effects such as sporadic-E, and meteor-burst can be included in the channel analysis.

4) Due to the 'real-time' nature of R.T.C.E., it facilitates use of channels higher in frequency than the predicted M.U.F. from off-line channel analysis procedures. Thus, spectral congestion is reduced.

5) It provides an automatic, computationally-viable method of ranking channels.

6) It allows more efficient usage of transmitter power, which in turn reduces spectral pollution from spurious component generation effects.

) It provides data for adaption of the communication system.

5) TEMPLATE CORRELATION - A NOVEL METHOD OF IN-BAND R.T.C.E.

A new method of channel selection and signal optimisation is described below.

5.1 Introduction.

Many channel evaluation routines assess occupancy on an allocation-by-allocation basis, using say 3 kHz as the frequency resolution of the assessment. However, it has been shown that H.F. band co-channel interference is predominately narrow-band in nature (Gott et al, 1983), i.e., less than 3 kHz wide. Thus it may be that potentially useful channels will be discarded as being occupied when they actually possess significant amounts of spectral space. This would be particualrly true of channels containing high power or broadcast transmissions. Thus there is good reason for assessing the suitability of a channel to pass traffic, at a sub-channel level. A technique known as 'Template Correlation' has this capability.

5.2 Theory Of Template Correlation.

It has been shown that for maximum information rate over a noisy channel, the power spectrum of the signal should be adjusted so that sum of the signal and noise power spectral densities is frequency independent (Goldman, 1953). The proof of this theorem is included in Appendix 1, and the principle is illustrated by Figure 5.

Hence to achieve maximum information rate one should shape the signal so that it has greatest spectral density where the noise spectral density is least. The proof of this theorem (shown in Appendix 1) states that the sum of the spectral densities should be a constant. If the noise power spectral density becomes greater than the constant value, then in order to satisfy this theorem, the signal spectral density would have to have a negative value. In practical terms this is clearly a nonsense. Thus, to apply this theorem, regions of high interference should be avoided completely, and the signal spectrum should be shaped into regions of lower interference. Figure 6 illustrates this point.

For a practical communications system there is the problem of establishing points in a noisy channel where the noise spectrum is matched to the modulation format in this way. An original method of locating such positions termed "Template Correlation" has been developed. The principle of this is outlined below.

Consider the situation where it is required to position a signal within a noisy channel: if one performs a crosscorrelation between the tone and the interference spectrum, (over the whole of the spectrum) then points of greatest similarity will show up as maxima in the correlation function, whilst points of least similarity will be displayed as minima in the correlation function. The principal minimum will be where the condition described by Figure 6 is met, i.e. a condition of least similarity between interference spectrum and modulation spectrum.

By performing the above correlation and looking for minima in the correlation result, the points within a channel that are most suitable for use by a particular modulation format are located.

5.3 A Computer Simulation Of Template Correlation.

A computer simulation of Template Correlation has been performed. This was written in Pascal on an I.B.M. p.c. Details of this and the results achieved are described below.

For the purposes of this simulation, interference spectra and modulation templates were sampled every 10 Hz. Hence a 3 kHz spectrum could be stored in 300 samples. Sampling spectra and templates in this way was thought to be realistic as, in a practical communications system, Template Correlation would most likely be performed digitally (via a computer). The mechanism of this discrete correlation is described by Figure 7.

Figure 8 represents the results of two sample correlations. Figure 8a shows the result of correlating 330 Hz wide rectangular modulation template (e.g. that required by a Piccolo M.F.S.K.-type modem) with the spectrum at the top of the diagram. The interference spectrum was produced using a random number generator within the I.B.M. p.c. The gap in the spectrum was produced by reducing the maximum amplitude of the random number generator over the gap frequency range. This gives a region of lower interference (in this case 330 Hz wide) which is the most suitable position for the modulation template.

Figure 8b is a more realistic example. This spectrum is a modified and digitised version of one produced by Dr. G. Gott at U.M.I.S.T. in the U.K. A sinc(x) shaped tone of bandwidth 200 Hz was used as the modulation template (the bandwidth of the template is defined as that being between the first two nulls). The lowest correlation in this case corresponds to the region of lowest interference around 750 Hz. Hence this is the optimum position for such a signal.

Figure 9 depicts other simulated correlations and shows a table of available modulation templates.

These and other simulation results show that the technique of Template Correlation is capable of finding the optimum position for a modulation format within a crowded or noisy channel. Practical trials over real H.F. links are required to verify this; however the results obtained so far are promising.

5.4 Reduction Of Required Computing Resources.

The correlations described so far have been of an analogue nature in that the interference spectra and modulation templates can consist of an infinite number of amplitude levels. Storage and processing time would be reduced if the spectra and templates were in binary form. One way of achieving this is to "threshold" them before correlating, as described by Figure 10.

The level at which the threshold is set can be related to system performance parameters such as minimum required signal to noise ratio, and (in digital modulation schemes) bit error rate. The following shows how this would be performed for a system employing simple F.S.K. modulation.

Figure 11a shows one F.S.K. tone of bandwidth B, and peak amplitude A. B is the detection bandwidth of the corresponding matched filter and is the first null-to-null bandwidth of the signal. Figure 11b is the power spectral density function of the same signal. The total power of the signal tone in the detection bandwidth B is represented by the shaded area of Figure 11b. Assuming the logic 1's and 0's are equi-probable, then the mean signal power in the detection bandwidth will be half that area, i.e., S/2. The probability of error for matched filter detection of F.S.K. in Gaussian White Noise is given by (Schwartz, 1981) :-

$$Pe = 1/2.\exp(-\Psi/2) \qquad (1)$$

where Ψ = (mean signal power) / (mean noise power)

The mean signal power for a practical F.S.K. transmitter could be determined at the receiver by performing a Fast Fourier Transform on the incoming signal, and then calculating the area as depicted by Figure 11b. Practical calculation of the area can be achieved by adding up the samples in bandwidth B, and using a look-up table to translate this into standard power units. This is then used together with the minimum required probability of error for the system and equation (1) above to give the maximum mean noise or interference power allowable. This figure then represents the level at which the interference spectrum and modulation format are thresholded.

In an adaptive system, (such as that described in the introduction), one would have a series of thresholds corresponding to different combinations of source and channel encoding, each having it's own maximum mean permissible noise/interference level.

Further evaluation of the effects of thresholding is necessary. However, it is desirable from a computational point of view as it allows Template Correlation to be performed more efficiently.

6) THE DEVELOPMENT OF CO-CHANNEL INTERFERENCE MODELS

As mentioned in previous sections of this paper, there is as yet no adequate model of co-channel interference for the H.F. radio communications environment. This type of interference can be the factor limiting communication system performance in spectrally congested, mid-latitude regions. Such a model would be incorporated into the communications system control algorithm in order to provide interference data for channel selection decisions. Work has been performed on theoretical models, such as that described by (Gott et al, 1983), and (Riley, 1986). Further development and evaluation is necessary to decide how useful these models are. Generally, they tend to be based upon databases of measurements from a single site only; thus, these models may have limited applicability.

From an H.F. radio communication system design point of view, there is a need for a computationally viable model of co-channel interference to be embedded into the system control and frequency management algorithms. Such a model is proposed in the next section of the paper.

7) A COMPUTATIONALLY VIABLE MODEL OF H.F. RADIO CO-CHANNEL INTERFERENCE

A simple co-channel interference model has been conceived, and this is outlined below.

The proposed model relies initially upon occupancy data for the whole H.F. spectrum, such as that contained in (Gott et al, 1982). A number of threshold levels are generated for the different combinations of source and channel encoding available with the particular system in use. The threshold levels represent the maximum permissible mean noise or interference power to enable a specified combination to provide the required class of service, e.g., bit error rate, transmission rate, etc. Thus, if there are n combinations of source and channel encoding, then there will be n threshold levels. Therefore, for each channel there are n interference models. Using the occupancy data for the whole H.F. band, the initial model is generated according to the following procedure :-

- If the mean occupancy level in a given channel is greater than the threshold, then that channel is said to be occupied at that threshold level.

The model for each channel-threshold combination can be stored as a single bit at this stage. These models will be embedded into the system prior to the initial

establishment of a communication link. They would provide the basis of interference estimation at this stage.

An obvious requirement for an interference model is that it should have the ability to be updated as the interference environment changes. Updating of the above model is achieved via passive monitoring of the system's frequency allocations. The proposed updating technique is described by Figure 12. For each 3 kHz allocation, a Fast Fourier Transform is performed to obtain the baseband power spectrum of the channel. This is then thresholded at the n levels specified for the initial model. A binary waveform results from this on the basis that if a sample is above the threshold level, then a '1' is stored in the threshold result at the corresponding position; otherwise a '0' is stored. The models for each channel are updated by repeating the above process and performing an averaging operation on the stored model and the threshold result thus generated. Run-length encoding would be used to reduce the storage associated with the models.

The question remains as to how often the model should be updated. It is thought that it will be necessary to update most often during changing ionospheric conditions. A variable rate of update would seem logical. The way in which new data is averaged with the old model also needs further work. A form of weighted averaging with bias towards most recent data could be performed.

8) APPLICATION TO H.F. COMMUNICATION SYSTEM DESIGN

8.1 H.F. System Design Philosophy.

A structured and coherent design methodology has been conceived for H.F. radio communication systems. This is described below.

For the proposed system there are two major aspects to it's operation :-

i) PRE-CALL PHASE - channel analysis prior to transmission.
This involves an assessment of which channels will provide the maximum probability of establishing a link of acceptable quality within the minimum period. Use is made of long-term, propagation, noise, and interference analysis routines and models, and the passive monitoring of known broadcast stations.

ii) IN-CALL PHASE - channel analysis during transmission over an established link.
This involves assessment of the channel in use to pass the required traffic. The other channels are ranked in their order of ability to provide the required grade of service. A prediction of when the current channel will become unusable is made. Here, embedded R.T.C.E. techniques are used.

Initially, control information would be supplied via stored data and long-term models. These models would be updated via embedded R.T.C.E. measurements and passive monitoring of allocated channels in the in-call phase above. Figure 13 demonstates how control is transferred from a basis of stored data and unmodified models to a basis of updated models and data from passive monitoring and R.T.C.E., as a function of time. By employing such techniques, it is predicted that efficient system control and frequency management will be achieved.

It is desirable that the architecture of the communications system allows :-

- Easy maintenance.
- Easy testing of new system components prediction routines and models.
- A degree of compatability with other types of communication systems, e.g., P.T.T. networks.

A control system architecture based on the I.S.O.-O.S.I. 7-layer model for data transfer bewteen computers has been developed. The O.S.I. model is shown in Figure 14 (Zimmerman, 1980).

8.2 Aims Of The O.S.I. 7-Layer Model.

The O.S.I. model was developed to provide a framework for the definition of standard protocols for heterogeneous informatic networks. The need for O.S.I arose from individual manufacturers developing their own architectures, which were largely incompatible with each other. Hence a common standard was required if such large networks were to be realised. The basic objective is to standardize the rules of interaction between interconnected systems. Only the external behaviour should conform to the O.S.I. architecture as the internal operation is not visible from other such open systems.

8.3 The Principle Behind The Layered Approach

Each layer adds to the services provided by the lower layers. A new layer is created when a new level of abstraction is required. Interfaces between layers are precisely defined. This allows alteration and maintenance of the internal functions of a layer without affecting other layers, as the same set of services must still be offered to the other layers.

8.4 Definition Of The Layers (Tanenbaum, 1984)

Referring to Figure 14:

PHYSICAL
This layer is concerned with the transmission of raw bits of information over a communications channel. It provides mechanical, electrical, and procedural interfacing to the communications network.

DATA-LINK
This transforms the physical communications path into an error-free channel, using error control techniques, e.g. A.R.Q. It is also responsible for set-up procedures, link management and link protocols.

NETWORK
This is concerned with the routing of the communications traffic through the network.

TRANSPORT
This layer translates user addresses to network specific addresses. It is also responsible for translation of user requirements into a routing and channel handling strategy.

SESSION
This is concerned with user authentication, call logging, and billing.

PRESENTATION
This layer perfroms data translations, e.g., ASCII to EBCDIC, and data encryption.

APPLICATION
This is the users process that requires network communications.

8.5 Description Of The Model Applied To H.F. Communication Systems

Figure 15 shows how the seven layer O.S.I model has been applied to an H.F. communications system. A description of each of the layers now follows. The following paragraphs assume that data is to be transmitted. For the case of the reception of data, the proceses outlined are reciprocal, except where shown not to be.

LAYER 7 (APPLICATION LAYER)
This encompasses the user's processes wishing to employ the communication medium. These may be computers, vocoders, speech recognition systems etc. Users may also be multiplexed at this level.

LAYER 6 (PRESENTATION LAYER)
Any data translations or data compression would be performed by this layer. Examples of layer 6 operations are run-length encoding, encryption and use of message dictionaries (for limited vocabulary communication systems).

LAYER 5 (SESSION LAYER)
This layer performs recovery strategies from equipment failures , and communication path failures. This layer is also responsible for extracting control information from the presentation layer. The destination of the data and it's operational application are output from the layer e.g., one way, low speed telegraphy with no security to base station.

LAYER 4 (TRANSPORT LAYER)
This layer takes the user requirements extracted by the session layer, and translates them into a channel handling strategy. If speech were to be transmitted, then the channel requirements in terms of required bandwidth would be much greater than for low speed telegraphy. Data is input to the Transport Layer in the form of complete messages. If necessary, these messages are broken down into smaller units, termed Message Packets. The Transport Layer is responsible for the re-formatting of Message Packets into complete Messages. Outputs from this layer are data, the address for the data in terms of network nodes, channel requirements and planned outage, e.g., maintenance schedules.

LAYER 3 (NETWORK LAYER)
Resident in this layer are the channel analysis procedures and models and passive monitoring control. When data is received for transmission from the Transport layer, the

Network Layer uses the channel analysis information and network status information (which nodes are operational, which currently provide the fastest routing to other nodes etc) to decide the routing for the data. The Network Layer breaks up message packets into frames (if need be) and is responsible for ensuring that these frames are re-compiled into message packets in the correct order. Outputs from this layer are as shown . Inputs to the layer from the Data Link Layer include R.T.C.E update information and network performance figures.

LAYER 2 (DATA-LINK LAYER)
This layer is responsible for transforming the bit-level communication facilities offered by Layer 1 into an error-free channel for the upper layers. Embedded R.T.C.E., error control techniques (e.g. Golay Encoding) and A.R.Q. schemes are resident in this layer. The Data-Link Layer passes bits to the Physical Layer for transmission, and also is responsible for control of the communications equipment.

LAYER 1 (PHYSICAL LAYER)
This layer encompasses the necessary transmission equipment. It has been divided into three sections, although it is expected that the embedded R.T.C.E. algorithms will use the same R.F. equipment as the communications traffic.

8.6 Internal Implementation Of The Model Layers.

The implementation of the individual layers should also be structured. Employing the seven layer model at a sub-layer level was found to be excessive and irrelevant. A formalised method of constructing the individual layers was found by implementing each layer as an expert system. This branch of Artificial Intelligence will now be discussed.

8.6.1. Expert Systems - An Overview.

A definition of an expert system is given by (Sell, 1985) as :-

> "An expert system is a knowledge-based system that emulates expert thought to solve significant problems in a particular domain."

In H.F. communication systems the "expert" is the human operator of the system. Figure 15 shows the three main components of an expert system to be :-

1) An Elicitation System - in order to acquire expertise in the domain of interest.

2) A Knowledge Base - containing a meaningful representation of the relevant knowledge.

3) An Inference Engine - which uses the knowledge base to make intelligent decisions.

The key feature of expert systems that makes them useful for providing an orthogonal design methodology for H.F. radio, is the seperation of domain specific knowledge and the inference engine. Thus an inference engine can be applied to many different fields of knowledge. Alterations to the operation of an expert system are made simply by altering the knowledge base, the inference mechanism remaining the same. In this way, other operations of the control sysetm will not be unwittingly affected.

8.6.2. Knowledge Representation Methods And Applications To H.F.Communication System Control.

The main methods of representing knowledge in expert systems are described below.

1) Complex record-like data structures which allow the expression of the interdependency of data items, e.g., Semantic Networks.

2) Rules, e.g.
IF channel n < M.U.F. THEN it will propagate.
IF only 2kHz available spectrum THEN analogue speech no good.

3) Fuzzy Logic.
This allows use of terms such as "probably true" rather than just "true" or "false".

4) Uncertainty, e.g., The M.U.F. will be < 10MHz with a probability of 0.85. Expert systems also allow such probabilities to propagate through decision making algorithms.

9) CONCLUSIONS

The current status of H.F. communication system design has been reviewed. A novel method of in-band R.T.C.E. has been presented with computer simulation results showing it to be potentially useful. A systems orientated co-channel interference model has been

outlined with attention being given to the practical requirements for such a model. A structured design methodology for H.F. systems has been presented involving the I.S.O.-O.S.I seven layer model for data transfer between computing machinery and Expert System design techniques. It is predicted that such techniques will allow the development of effective, and adaptive H.F. systems with the ability to operate in noise and interference ridden environments.

The author would like to thank all concerned at Plessey Research Ltd, U.K., for their assistance with this research programme.

10) FUTURE PROGRAMME

The future activities of the research programme at The University Of Hull are as follows :-

1) The implementation and testing of the ideas presented in this paper. This will require the construction of an H.F. radio link/network.

2) The formalisation of H.F. communications system design.

3) An investigation of the role of passive monitoring techniques and embedded R.T.C.E. measurements for the updating of embedded channel analysis models.

4) The development of co-channel interference models.

5) The evaluation and comparison of different embedded R.T.C.E. algorithms.

11) REFERENCES

1) Darnell, M., "The design of static and mobile H.F. communication systems." 1986a. AGARD Lecture Series No. 145 on "Propagation Impact On Modern H.F. System Design".

2) Thrane, E.V., "Propagation 1: State of the art of modelling and prediction in H.F. propagation." 1986. AGARD Lecture Series No. 145 on "Propagation Impact On Modern H.F. System Design".

3) Bradley, P.A., "Propagation at medium and high frequencies, 2: Long and short term models". 1979. AGARD Lecture Series No. 99 on "Aerospace Propagation Media Modelling And Prediction Schemes For Modern Communications, Navigation, And Surveillance Systems".

4) Uffelman, D.R., et al, "Real-time update of two H.F. channel evaluation models by oblique sounding". 1984. Naval Research Laboratory Memorandum Report 5246, Washington, U.S.A.

5) Darnell, M., "Embedded real-time channel evaluation". 1986b. AGARD Lecture Series No. 145 on "Propagation Impact On Modern H.F. Communication System Design".

6) Gott, G.F., et al. "Analysis of H.F. interference with application to digital communications". 1983. I.E.E. Proceedings Part F, Vol 130, No.5, pp 452-458.

7) Riley, N.G., "An H.F. interference model - early steps". 1986. I.E.E. Colloquium Digest on "H.F. Frequency Management".

8) Gott, G.F., et al , "Occupancy measurements across the entire H.F. spectrum". 1982. AGARD Conference on "Propagation Aspects Of Frequency Sharing, Interference, And System Diversity", Paris.

9) Goldman, S., "Information Theory". New York. Prentice_Hall International. 1953. pp 111-161.

10) Schwartz, M., "Information Transmission, Modulation, And Noise". Singapore. McGraw-Hill. 1981. p386.

11) Zimmerman, H., "O.S.I. reference model - The O.S.I. Model of architecture for Open Systems Interconnection". 1980. I.E.E.E. Transactions on Communications, Vol. COM-28, No.4. pp 425-431.

12) Tanenbaum, A.S., "Computer Networks" Englewood Cliffs, N.J. U.S.A. Prentice-Hall International. 1981. pp 15-21.

13) Sell, P.S., "Expert Systems - A Practical Introduction". London, Macmillan Publishers Ltd. 1985. p15.

12) APPENDICES

12.1 Appendix 1

The following is proof that for maximum information rate in a noisy channel, the power spectrum of the desired signal should be adjusted so that the sum of the signal and noise power spectral densities is frequency independent (Goldman, 1953).

Consider (refer to Figure 5) :-

$$\int_{f_1}^{f_2} S(f).df = P \qquad (1.1)$$

$$\int_{f_1}^{f_2} N(f).df = N \qquad (1.2)$$

where S(f) = power spectral density of the signal.
N(f) = power spectral density of the noise.
P = total signal power in range f_1 to f_2 .
N = total noise power in range f_1 to f_2.

Assume that the noise is Gaussian, i.e., there is no intersymbol influence and that the noise follows a Gaussian amplitude distibution. For maximum information rate the signal will also have Gaussian properties.

The frequency band f_1 to f_2 is divided into smaller bands so that N(f) is approximately constant in each.

Consider also :-

The entropy of any continuous distribution is given by

$$H = -\int_{-\infty}^{\infty} p(x).\log(p(x)).dx \qquad (1.3)$$

and the probability density function, p(x) for a Gaussian process is

$$p(x) = \frac{1}{(2\pi\sigma^2)^{0.5}}.\exp(-x^2/2\sigma^2) \qquad (1.4)$$

where σ is the standard deviation of the Gaussian Distribution.

By taking anti-logs and performing the entropy integral one gets the result shown below :-

$$H = \log(2\pi e\sigma^2)^{0.5} \qquad (1.5)$$

So for a Gaussian source sampled at n points we have :-

Entropy of each point,

$$H = \log(2\pi e\sigma^2_I)^{0.5} \qquad (1.6)$$

where σ_I is the root mean square value of x about the mean at the sample point . (Mean square of x about the mean = variance, σ^2.)

For our noise source, let $N = \sigma^2_I$ i.e., let the mean square noise power be the variance of the distribution (as per definition of variance).

Thus the entropy of each sample point is

$$\log(2\pi eN)^{0.5} \qquad (1.7)$$

This is the entropy per degree of freedom per sample. If we assume Nyquist sampling and the number of degrees of freedom is 1, as we either sample in frequency or time, then the total entropy is given by

$$H = T.W.\log(2\pi eN) \qquad (1.8)$$

where T = duration of signal.
W = bandwidth of signal.

Hence the entropy of our combined signal and noise in a narrow band Δf is :-

$$T.\Delta f.\log(2\pi e[S(f).\Delta f + N(f).\Delta f]) \qquad (1.9)$$

Entropy of the noise alone in the bandwidth Δf is :-

$$T.\Delta f.\log(2\pi e[N(f).\Delta f]) \qquad (1.10)$$

The information rate in the bandwidth Δf is given by :-

(1/T).[(Entropy of signal + noise) - (entropy of noise)]

as $T \rightarrow \infty$ (1.11)

Thus the information rate in our bandwidth Δf is

$$\Delta f . \log((S(f) + N(f))/N(f)) \quad (1.12)$$

Hence the information rate in the whole bandwidth f_1 to f_2 is

$$\int_{f_1}^{f_2} \log((S(f) + N(f))/N(f) . df \quad (1.13)$$

We wish to maximise this rate subject to the following constraint

$$P = \int_{f_1}^{f_2} S(f) . df \quad (1.14)$$

This is achieved via the calculus of variations, which yields :-

$$(\delta / \delta S(f)) \log.(1 + (S(f)/N(f))) + \lambda = 0 \quad (1.15)$$

where λ is a constant.

Hence

$$N(f)/(S(f) + N(f)) \times 1/N(f) = -\lambda \quad (1.16)$$

This gives :-

$$S(f) + N(f) = -1/\lambda \quad (1.17)$$

i.e. for maximum information rate the sum of the noise and signal power spectral densities should be frequency independent (see Figure 5).

FIGURE 1 EXAMPLE PROPAGATION ANALYSIS PROGRAM OUTPUT.

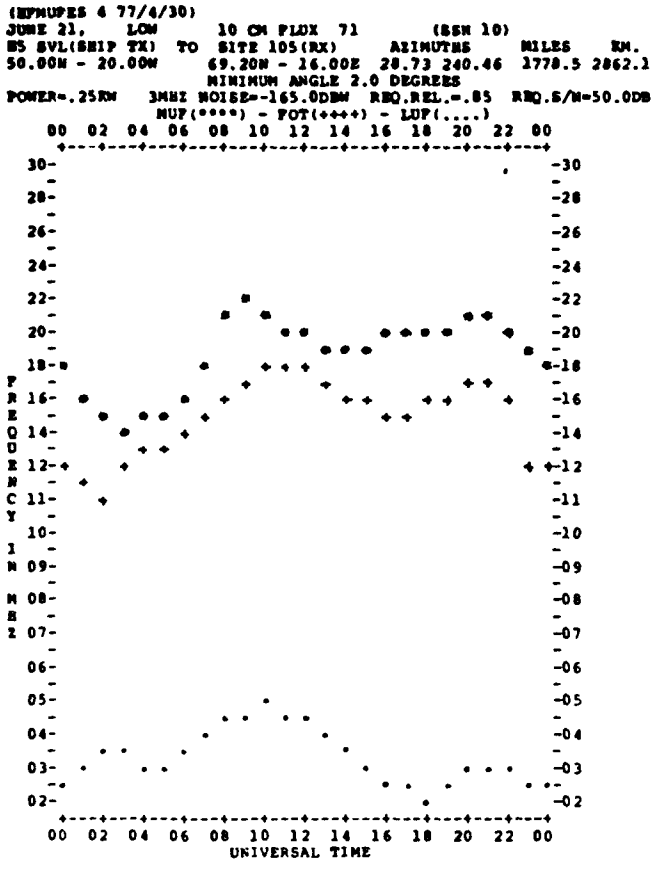


FIGURE 2 . PROPOSED PROPAGATION MODEL UPDATING METHOD.

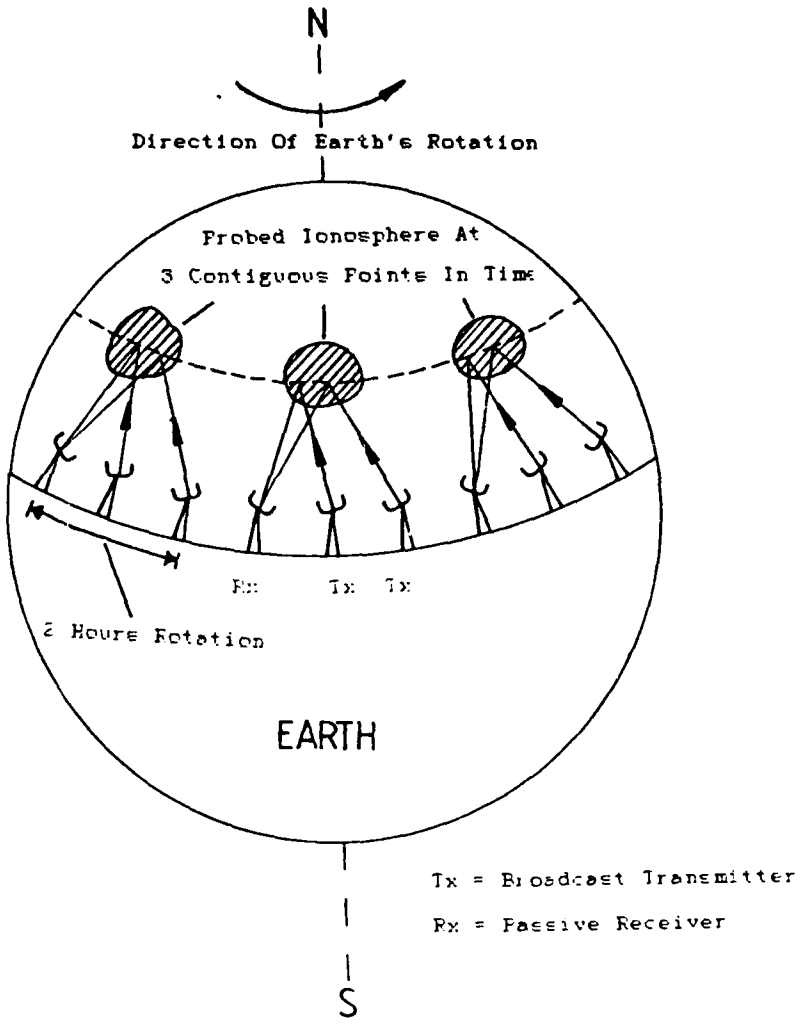
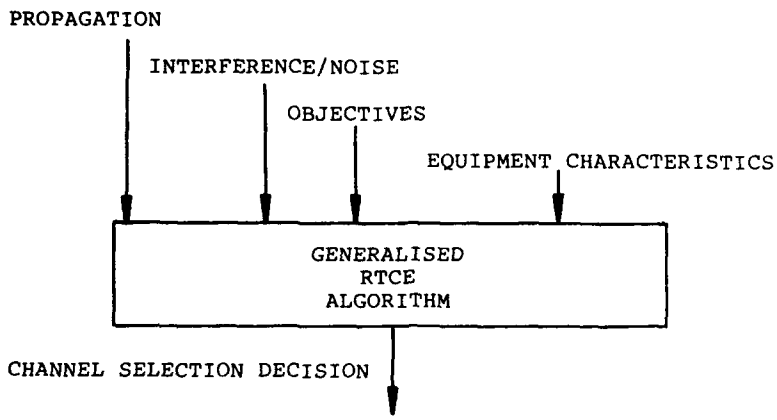


FIGURE 3. A GENERALISED R.T.C.E. ALGORITHM.



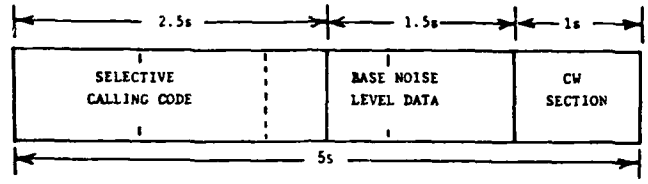


FIGURE 4 THE C.H.E.C. R.T.C.E. FORMAT.

FIGURE 5. SPECTRAL SHAPING REQUIREMENTS FOR MAXIMUM INFORMATION TRANSFER IN NOISY CHANNELS (GOLDMAN,1953).

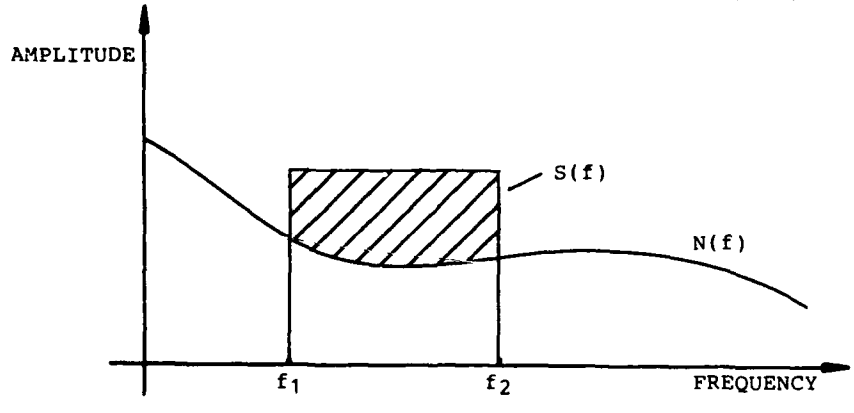


FIGURE 6. IDEAL SIGNAL POSITIONING WITHIN A NOISY CHANNEL

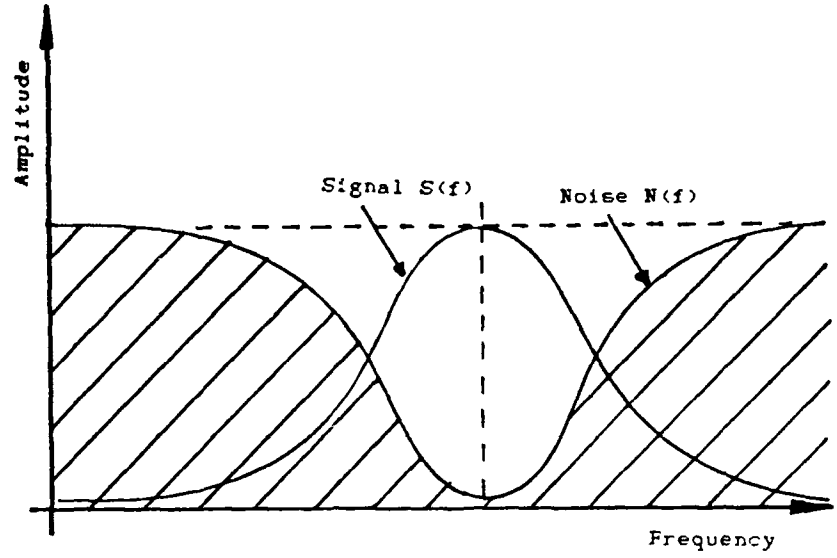


FIGURE 7. TEMPLATE CORRELATION SIMULATION

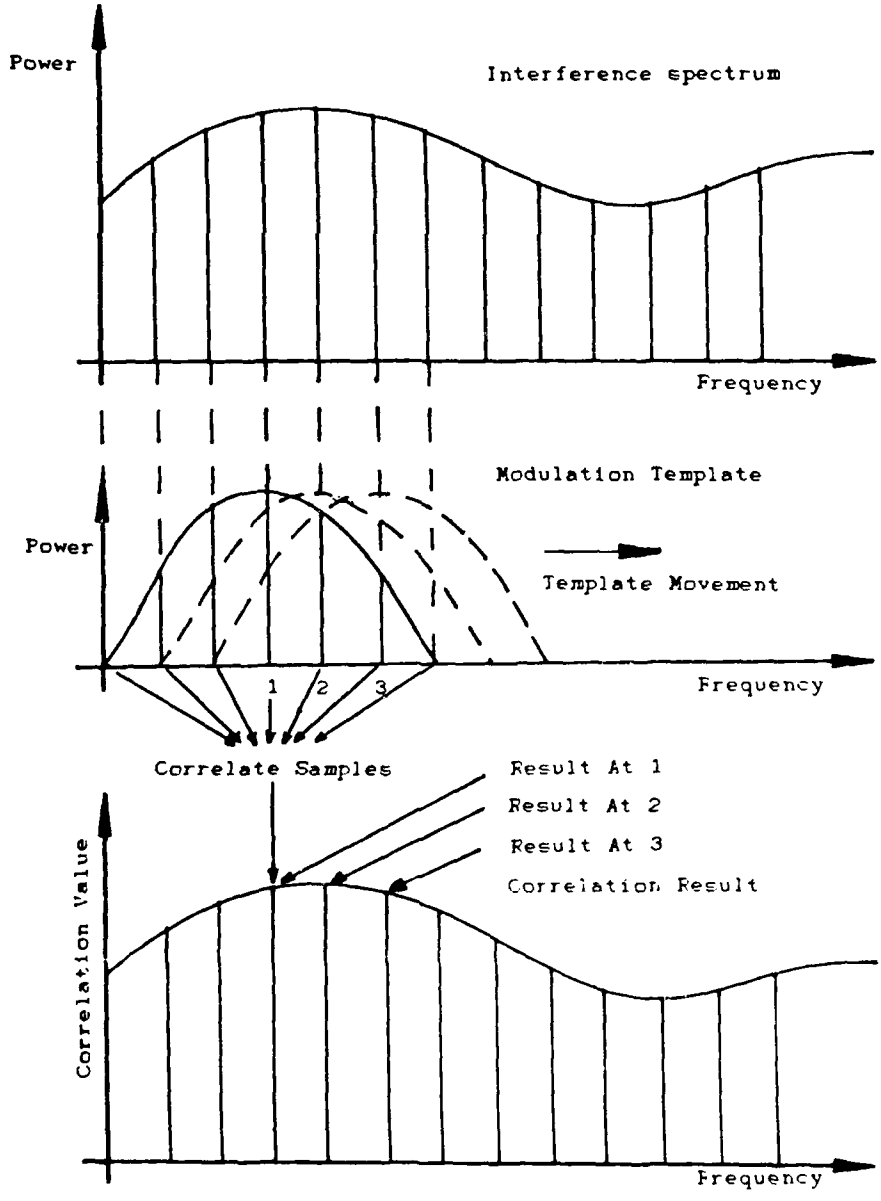
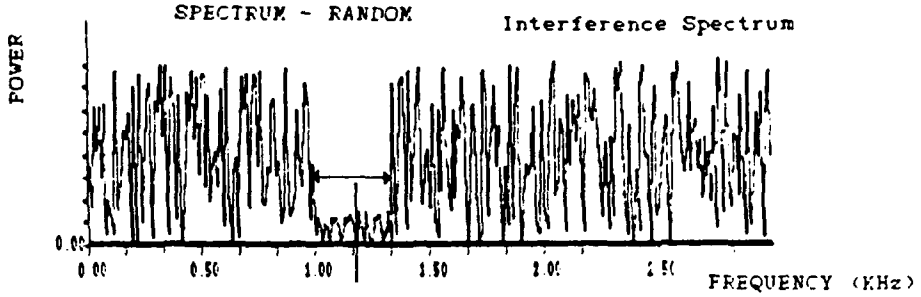


FIGURE 8. TEMPLATE CORRELATION SIMULATION EXAMPLES.

FIG 8a. TEMPLATE - 330Hz WIDE RECTANGLE

SPECTRUM - RANDOM



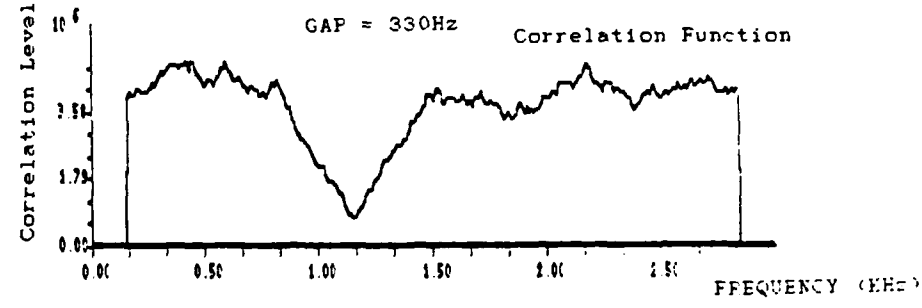
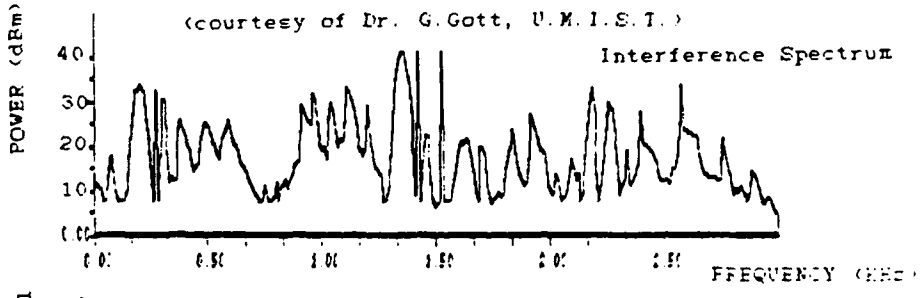


FIG 8b. TEMPLATE - SINGLE F.S.K. TONE

SPECTRUM - ENHANCED H.F. CHANNEL

(courtesy of Dr. G.Gott, U.M.I.S.T.)



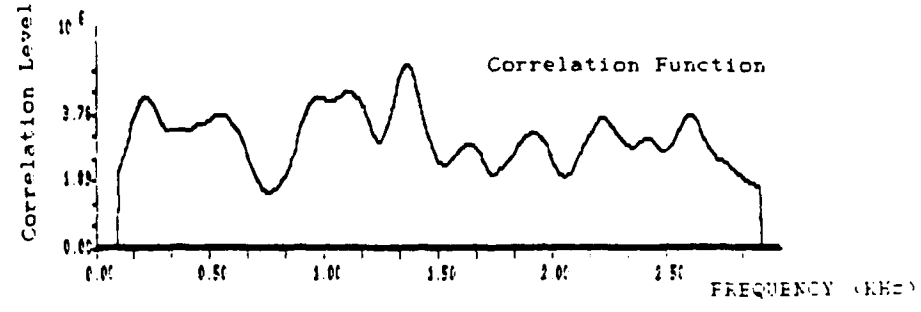


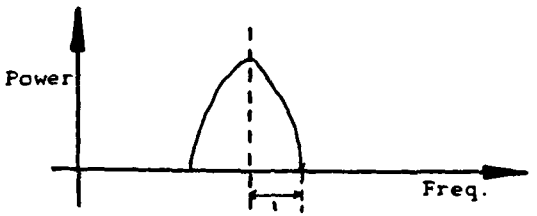
FIGURE 9

CONTENTS

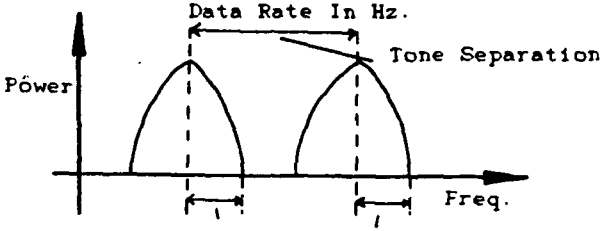
FIGURE NUMBER	DESCRIPTION
9a	Table of modulation formats.
9b	Correlation of F.S.K. modulation with a random spectrum. Gaps in spectrum match exactly the signal.
9c	As 7b but thresholding is used.
9d	Adaptive F.S.K. with H.F. spectrum in dB. (Courtesy Dr. G. Gott, U.M.I.S.T.)
9e	As 7d but the spectrum is linear.
9f	The spectrum is random with reduced amplitude over 2700 Hz. The modulation format is the C.C.I.R. standard raster.
9g	The spectrum is random with three progressively lower amplitude regions with width 330 Hz. The modulation format is Piccolo.

FIG. 9a TABLE OF MODULATION FORMATS

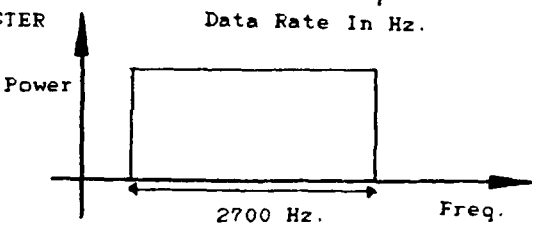
SINC(X) TONE



F.S.K. SIGNAL



C.C.I.R. MULTI-TONE RASTER

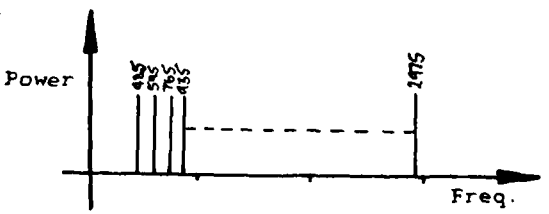


ADAPTIVE F.S.K. SIGNAL

Correlate Sinc(x) Tone

At 425 + n(170) Hz

Where n = 0 -> 15



PICCOLO

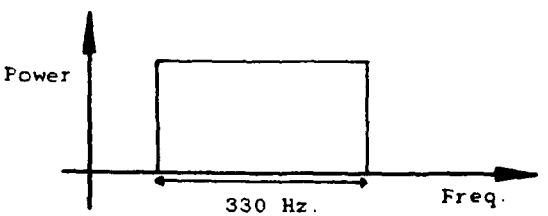


Fig 9b

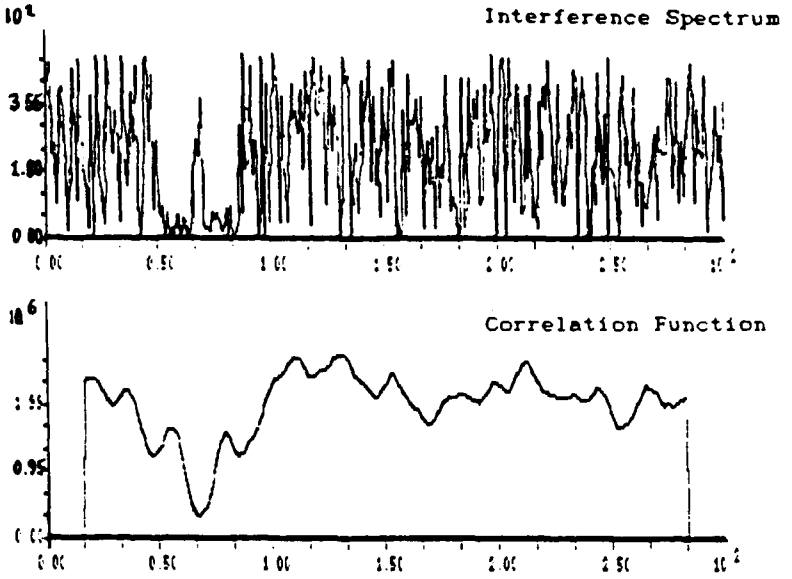

Interference Spectrum
Correlation Function

Fig 9c

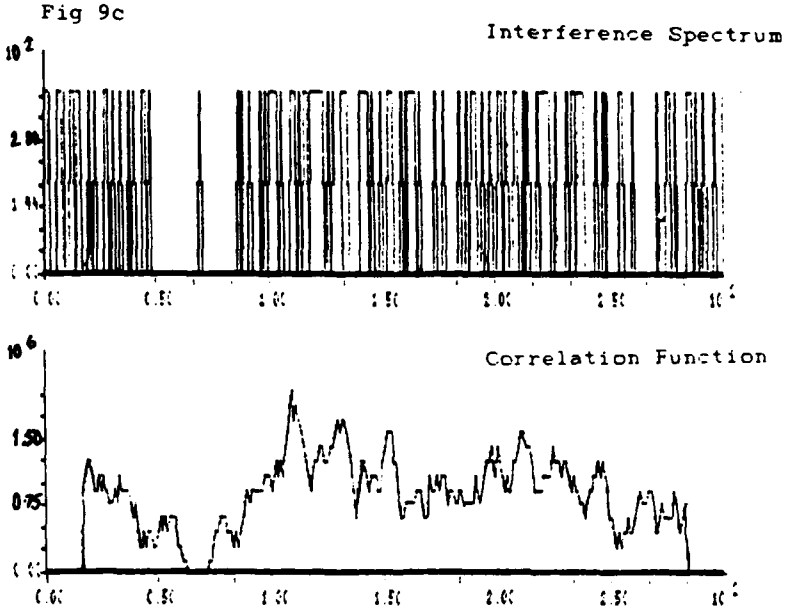

Interference Spectrum
Correlation Function

Fig 9d

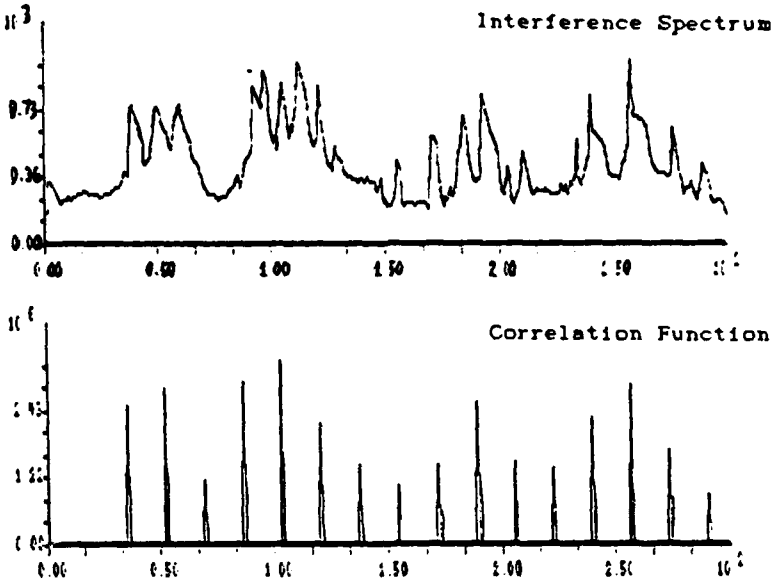

Interference Spectrum
Correlation Function

Fig 9e

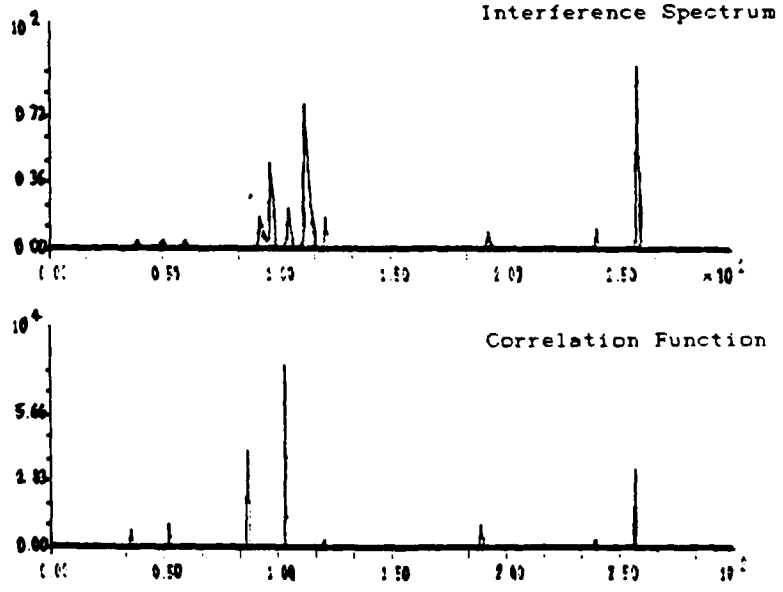

Interference Spectrum
Correlation Function

Fig 9f

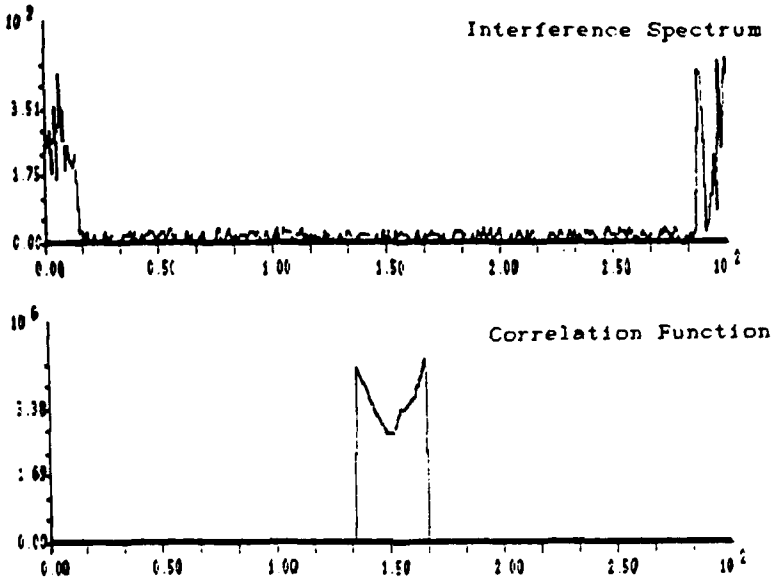

Interference Spectrum
Correlation Function

Fig 9g

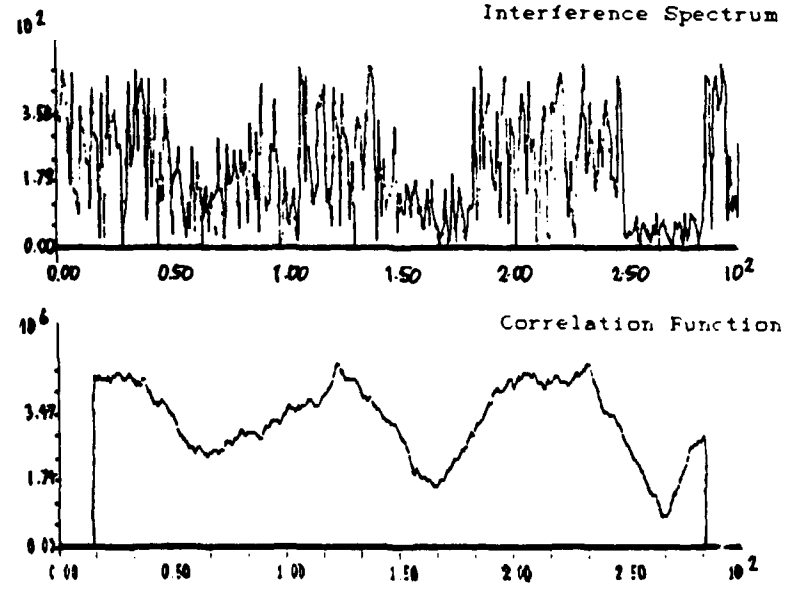

Interference Spectrum
Correlation Function

FIGURE 10 THE THRESHOLDING OF INTERFERENCE SPECTRA AND MODULATION TEMPLATES

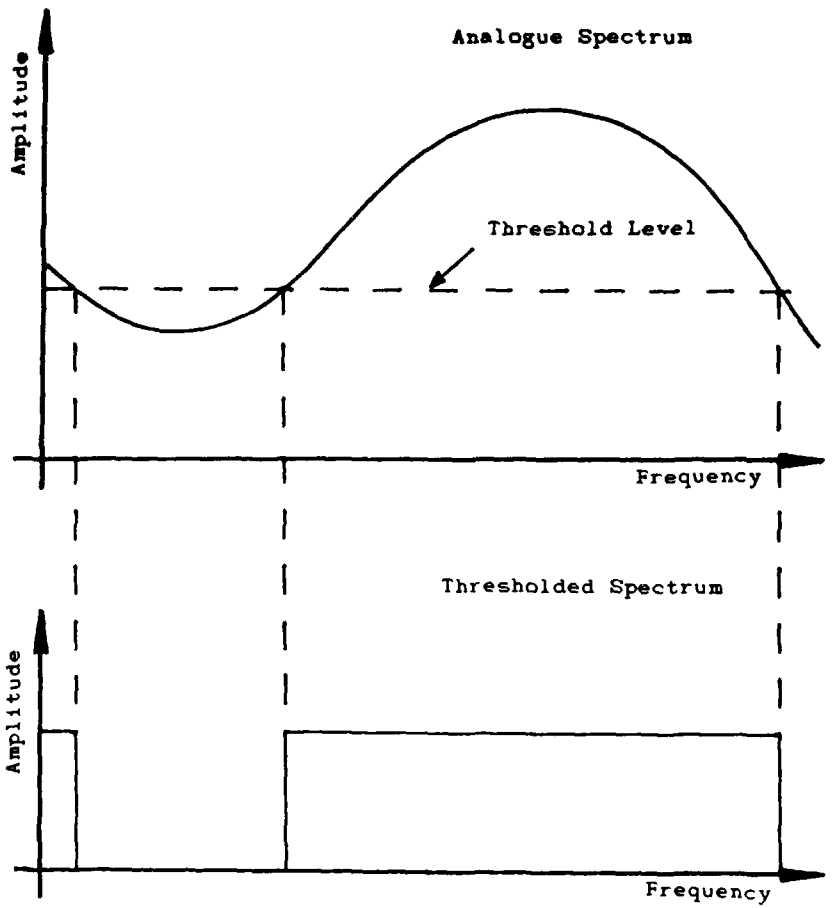


FIGURE 11a AMPLITUDE SPECTRUM OF SINC(X) TONE

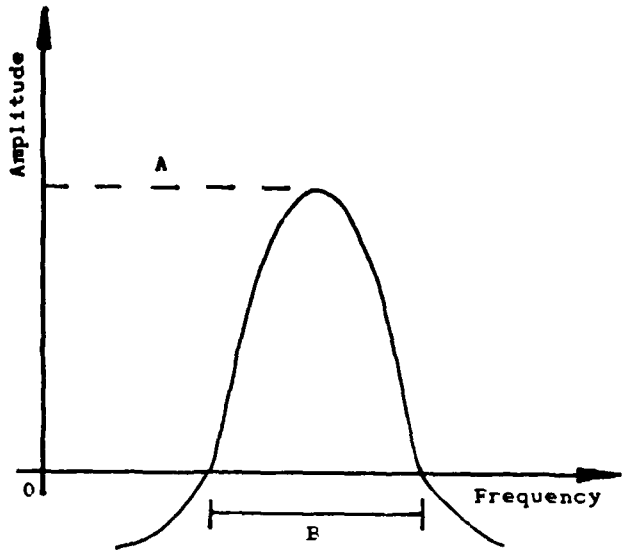


FIGURE 11b POWER SPECTRUM OF SINC(X) TONE

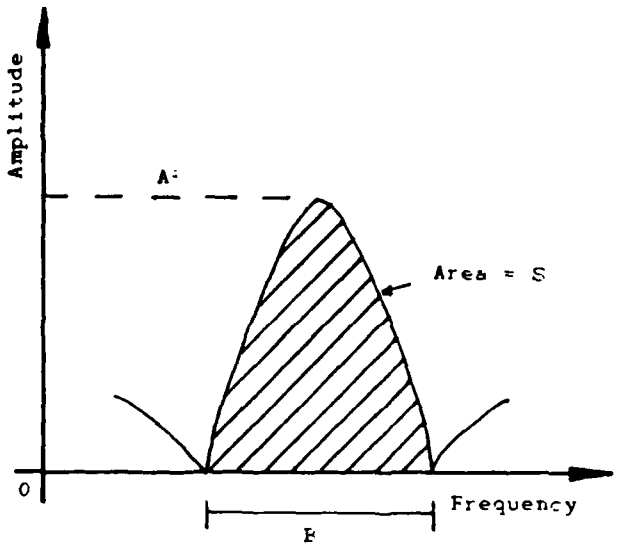


FIGURE 12

THE GENERATION AND UPDATING OF THE INTERFERENCE MODEL

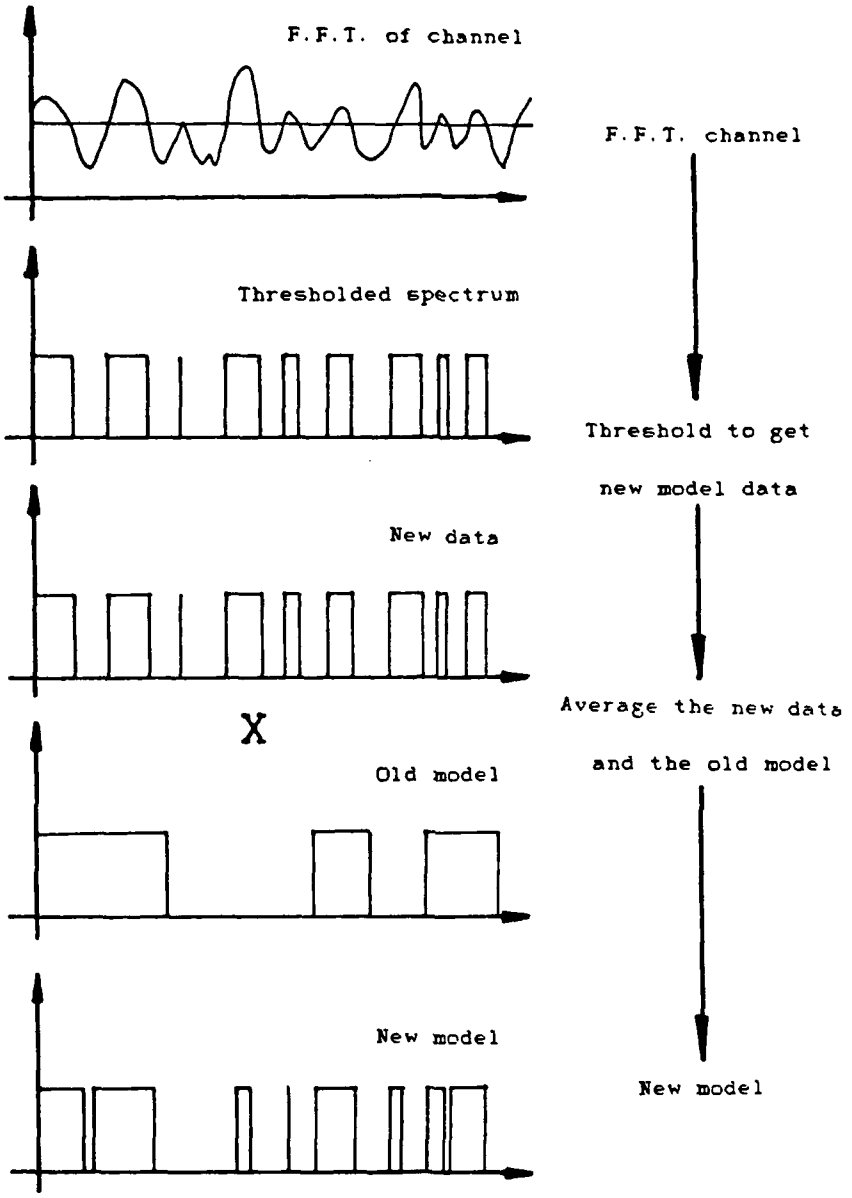


FIGURE 13

THE TEMPORAL TRANSFER OF CONTROL FOR THE PROPOSED AUTOMATIC H.F. LINK.

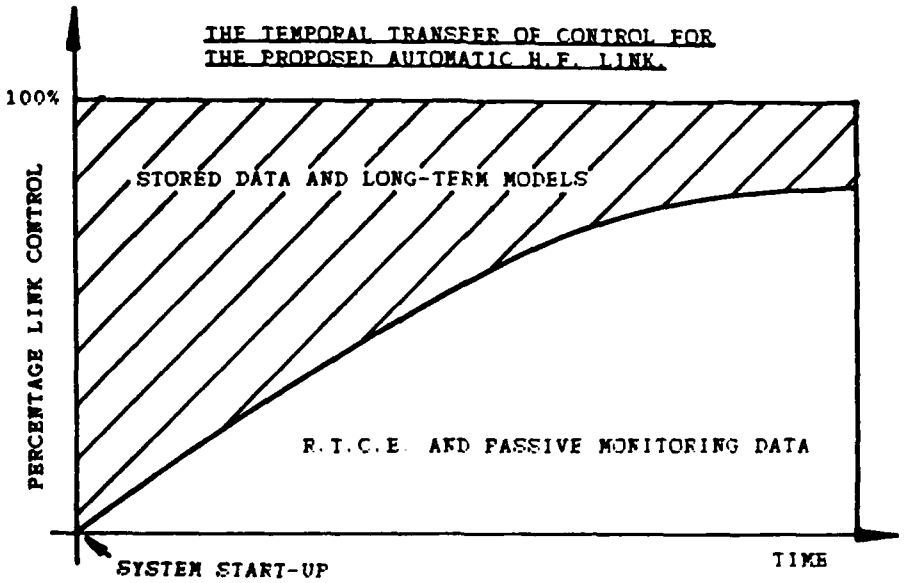
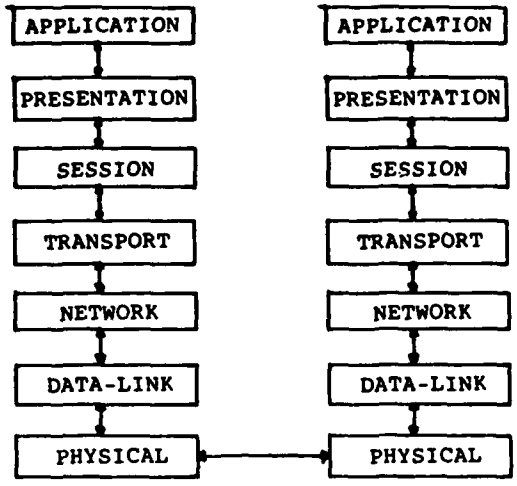


FIGURE 14 THE I.S.O.-O.S.I. 7-LAYER MODEL.



FOR FIGURE 15, SEE NEXT PAGE.

FIGURE 16. A BLOCK DIAGRAM OF AN EXPERT SYSTEM.

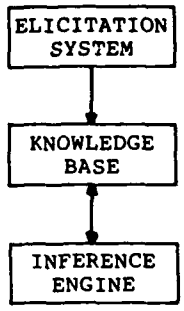
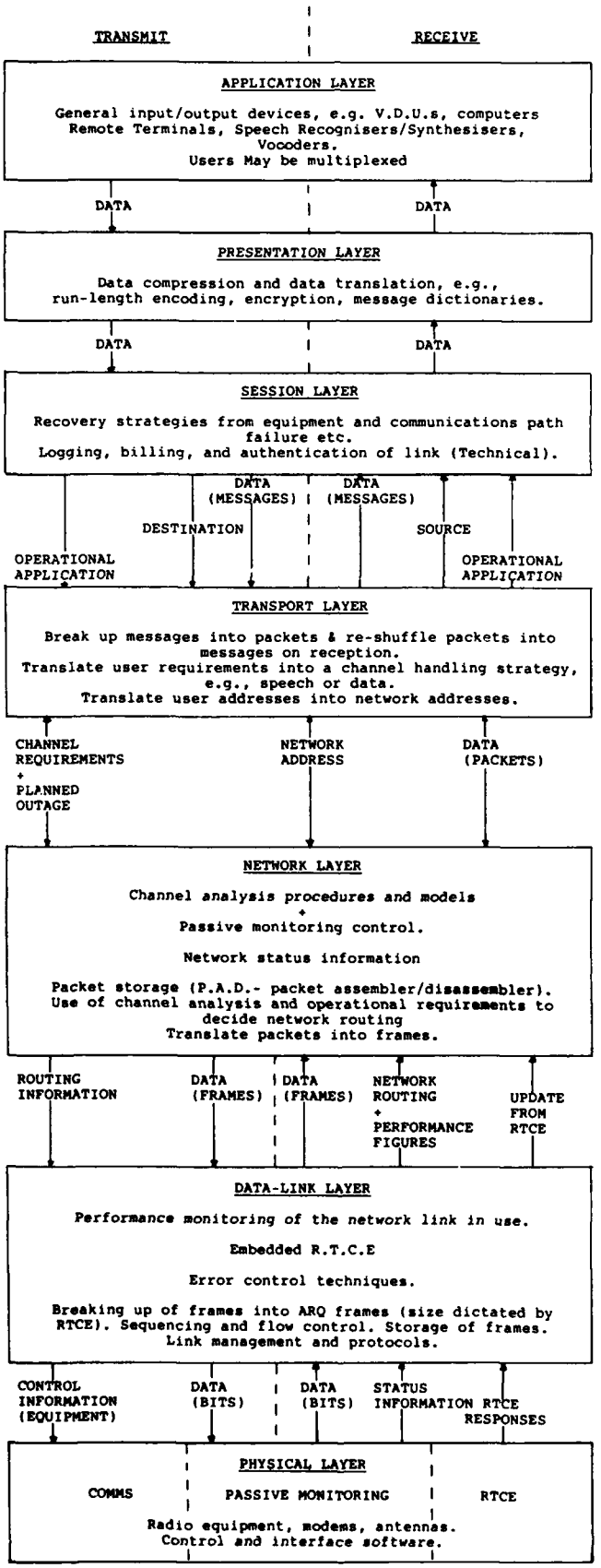


FIGURE 15 THE O.S.I. MODEL APPLIED TO H.F. COMMUNICATIONS



PROCESSING TECHNIQUES FOR RADIO COMMUNICATIONS IN NON-GAUSSIAN NOISE ENVIRONMENTS

M. Darnell
Department of Electronic Engineering
University of Hull
Hull HU6 7RX
UK

SUMMARY

The paper first discusses the different forms of background noise/interference environment encountered by various types of radio communication systems. Shortcomings of the analytical models available for these environments are then identified. The fundamental characteristics of radio paths which affect their ability to pass information are then examined, and certain general design principles established. Finally, several procedures and techniques with the potential to enhance radio system performance in non-Gaussian environments are described.

1. INTRODUCTION

Much of the published information on radio communication system performance relies on analysis techniques applicable to Gaussian white noise (GWN) environments. In practice, the designers and operators of radio systems frequently have to contend with noise backgrounds, both natural and man-made, that are predominantly non-Gaussian. Certain problems arise from this situation, ie:

(a) performance ranking orders for channel coding (modulation and error control) techniques obtained under GWN conditions may well be invalid for non-Gaussian conditions;

(b) conventional optimum detection schemes for GWN environments, eg matched filtering, may be sub-optimum in non-Gaussian noise;

(c) inadequate characterisation of the non-Gaussian noise environment for various classes of radio channels may give rise to uncertainty in the choice of system design parameters.

1.1 Content of the Paper

This paper attempts to take a co-ordinated approach to the problem of designing signal processing and system control techniques to allow reliable digital communication over non-Gaussian radio channels. Specifically, the following topics are considered:

- the nature of noise/interference for various types of radio channels, together with the analytical models which are available;
- the fundamental nature of the radio communication problem;
- forms of real-time channel evaluation (RTCE) data which can be obtained from the normal operating and control signals of a radio system in order to derive both on-line models of the noise/interference environment and to allow adaptation of communication system parameters;
- "look-ahead" processing, using buffer delays, and quenched filtering to provide more effective initialisation of the signal processing procedures at the receiver;
- embedded modulation formats, whereby composite modulation schemes have certain modulation components which are more resilient to noise than others: under high noise conditions, only the most robust elements will be demodulated successfully;
- embedded error control coding, whereby the same data is transmitted simultaneously at several distinct rates and, at any time, the receiver decodes at the highest rate commensurate with available channel capacity: the results of practical tests are presented;
- multi-functional coding schemes, by which a number of distinct signal processing functions, eg synchronisation, multiple-access coding and error control, can be considered in a connected manner, thus making the overall transmission procedures more efficient and responsive to channel conditions.

For reasons to be explained in Sections 2 and 3, the basic radio system architecture considered is automatic repeat request (ARQ).

2. RADIO NOISE & NOISE MODELLING

2.1 The Gaussian White Noise Model

The noise model used almost invariably in the analysis of communication system performance is that of Gaussian white noise, having a Gaussian amplitude probability density function (pdf), a uniform power density spectrum (pds) and an impulsive autocorrelation function (acf) (Lee, 1960). In practice, most communication systems are subject to noise/interference which does not conform to the GWN model - although, in some cases, such a model is a reasonable approximation to the actual conditions. Thus, there is a potential problem if system design parameters, eg modulation type and error control coding algorithm, are chosen on the basis of their performance in GWN, whilst the conditions experienced during operation over the real radio channel are non-Gaussian; in the latter case, a different ranking order for candidate modulation and coding schemes may well be applicable. It cannot be assumed that GWN performance can be extrapolated to non-Gaussian channels.

Clearly, there are very good reasons for the use of the GWN model, ie:

(a) it can be treated analytically;

(b) a reasonable approximation to GWN can be generated under laboratory conditions and in software simulations;

(c) for many types of radio channel, noise and interference are substantially unquantified and representative models are not available.

Making a simplified generalisation: as the frequency of operation of a radio system increases, so the noise background becomes a better approach to true GWN. This is primarily because, above a frequency of about 2 GHz, receiving system thermal noise becomes the dominant mechanism, as is shown in Fig. 1. At lower frequencies, noise may become non-Gaussian due to various natural and man-made effects. Examples of these include:

(a) energy from thunderstorm activity, propagating over extremely long distances in the earth-ionosphere spherical waveguide at very low frequency (VLF) (Omura & Shaft, 1971) - usually having a high peak-to-mean ratio and often described with acceptable accuracy by a log-normal amplitude pdf;

(b) similar effects, generally known as "static", propagating by both surface-wave and ionospheric skywave modes in the HF band;

(c) noise from electrical machinery, car ignition systems, etc which may affect any communications receiver in radio line-of-sight if sufficient energy is being radiated from the noise source in the passband of the receiver: VHF/UHF urban mobile communications are subject to this type of disturbance.

The other major factor influencing communication system performance is co-channel interference from other users of the spectrum. In this respect, those radio systems which rely upon long-range (beyond line-of-sight) propagation mechanisms are vulnerable to this form of interference, particularly when the propagation can only be characterised by statistical models which may have limited validity at a given instant in time. Fig.2 illustrates, in general terms, the variation of range achievable by radio systems operating at different frequencies. World-wide ranges can be achieved via sky-wave propagation in the HF band, where the natural screening properties of the ionosphere are employed to isolate services and to allow frequency re-use in different parts of the world. This isolation may be more or less effective, depending on the state of the ionosphere. From Fig.2, it is seen that the other very long range propagation mechanisms correspond to E/VLF and satellite communication systems: propagation for these classes of system can be modelled with reasonable precision for frequency planning and assignment purposes, and therefore co-channel interference normally presents no problems.

Although line-of-sight propagation at VHF and higher frequencies can be modelled with considerable accuracy, urban environments and anomalous propagation effects cause uncertainties and can cause frequency re-use problems due to unpredicted interference. Cellular radio systems depend critically upon the efficiency of frequency re-use and, as these systems proliferate, it can be expected that the potential for mutual interference will increase.

Possibly the most widely-used radio communication medium in which noise and interference effects are substantially uncharacterised is the HF band from 2 - 30 MHz. Some attempts at modelling these effects has been made (Gott et al, 1982) (Darnell, 1986) but, to date, there are no generally applicable models for co-channel interference. What can be stated with certainty is that to a communicator attempting to use the HF band, co-channel interference will appear to have non-Gaussian characteristics. In this paper, attention will be concentrated on the problems of efficient communication over the HF non-Gaussian channel; however, it is also anticipated that the techniques described will be applicable to non-Gaussian channels in other frequency bands.

3. FUNDAMENTAL NATURE OF THE RADIO COMMUNICATION PROBLEM

The fundamental characteristics of a typical radio channel will now be examined from two viewpoints, ie physical and information-theoretic.

3.1 The Physical Viewpoint

Looking at the radio channel firstly with a physical perspective: in a typical radio system, some or all of the following effects may influence the characteristics of the received signal:

(a) multipath;
(b) time dispersion;
(c) noise;
(d) co-channel interference;
(e) doppler spreading (frequency dispersion);
(f) fading - either flat or frequency-selective;
(g) polarisation variations;
(h) attenuation;
(i) reflection;
(j) refraction;
(k) diffraction;
(l) ducting.

Taken together or individually, these effects may give rise to a time variation in the received signal-to-noise ratio (SNR), as illustrated in Fig. 3. This SNR variation could equally well be translated to a time variation in channel capacity, C, where

$$C = B \log_2 [1 + S/N] \text{ bits/s} \qquad [1]$$

B being the channel bandwidth in Hz and S/N being the power SNR.

Normal communications system design methods involve the specification of a fixed rate of data transmission, eg 75 bits/s, 2.4 kbits/s, 16 kbits/s, etc. Clearly, this gives rise to a fundamental "mismatch" between the characteristics of the transmitted signal and those of the propagation path; the result of such a mismatch will normally be errors in the received information. Another way of looking at the channel is in terms of its error characteristics, as will be discussed in the following section.

3.2 The Information-Theoretic Viewpoint

In practice, a variation in channel capacity of the type mentioned in the previous section, when viewed at the receiver of a constant rate transmission system, tends to give rise to bursts of errors; this is in contrast to the situation occurring with GWN when the majority of errors would be of a random isolated nature. Certain techniques, such as adaptive equalisation, diversity processing, etc, can reduce the range of capacity variation to a limited extent.

One method of quantifying the "burstiness" of a radio channel is via a concept known as "μ-memory". This measure compares, at the receiver, the average conditional entropy, H, of a discrete channel having given bit transition probabilities with that of a discrete memoryless channel (DMC) having the same bit transition probabilities, H_o. The μ index is defined by (Wolfowitz, 1967)

$$\mu = \frac{H_o - H}{H_o} \qquad [2]$$

For a random (non-bursty) error channel

$$H \simeq H_o \qquad [3]$$

and

$$\mu \simeq 0 \qquad [4]$$

This implies that error bursts are extremely infrequent.

For a bursty channel

$$H < H_o \qquad [5]$$

because, at times, the data will be totally corrupted at the receiver for several bit intervals. Thus, the parameter μ is an indication of the burstiness of the channel errors as perceived at the receiver.

The critical factorr in this form of channel analysis is the "bit dependency", m, of the bit transition probabilities. If m is unity, the transition probabilities depend cnly on the previous bit state; if m = 2, the transition probabilities depend upon the previous two bit states, etc.

The analysis of this situation is relatively complex (Zolghadr, 1987). In general, it

can be demonstrated that as burstiness increases and H decreases, so the value of μ for a given m will also increase.

Tables 1 and 2 below show the result of calculations of the index μ for a GWN and an HF radio channel respectively using three different data sample lengths. The results for the HF channel were derived using recorded error sequences taken from an HF skywave trial. Data sample lengths of 500, 1000 and 10000 bits were used, together with values of m in the range

$$1 \leq m \leq 10 \qquad [6]$$

Bit Dependency m	μ values		
	500 bits	1000 bits	10000 bits
1	0.015	0.005	0.002
2	0.029	0.019	0.005
3	0.045	0.030	0.008
4	0.045	0.031	0.008
5	0.054	0.040	0.010
6	0.070	0.049	0.012
7	0.132	0.089	0.017
8	0.105	0.092	0.024
9	0.114	0.100	0.032
10	0.122	0.107	0.040

Table 1: μ-index values for GWN channel

Bit Dependency m	μ values		
	500 bits	1000 bits	10000 bits
1	0.047	0.050	0.015
2	0.106	0.108	0.034
3	0.123	0.120	0.048
4	0.162	0.128	0.055
5	0.236	0.214	0.086
6	0.297	0.310	0.125
7	0.363	0.384	0.174
8	0.378	0.414	0.201
9	0.422	0.456	0.232
10	0.419	0.476	0.266

Table 2: μ-index for HF channel

From Table 1, it is seen that for the GWN random error channel, the calculated μ values are low (less than 0.05 for 10000 bit samples). In contrast, Table 2 shows considerably higher μ values for all sample lengths (up to about 0.5), suggesting that the HF channel analysed had significant burstiness.

If μ can be monitored continuously for a given radio channel, then the channel encoding procedures can be varied responsively. It should be noted that the accuracy of the μ-index estimate is dependent upon the source state probabilities; thus, the characteristics of the source also need to be monitored continuously to assist the optimisation of the data encoding process prior to transmission over the channel.

The μ-index description of a discrete communication channel shows a clear distinction between GWN and non-Gaussian links, when the non-Gaussian characteristics arise from memory in the channel, eg as a result of multipath or noise bursts. It is apparent that the non-Gaussian channel should therefore be treated in a different manner from the GWN channel in terms of adaptation of source and channel encoding algorithms in response to path state.

4. RTCE PROCEDURES FOR RADIO CHANNELS

The μ-index technique outlined above is an example of an RTCE algorithm which makes use of data derived from the normal operating signals of the radio system. Such techniques are known collectively as "embedded" RTCE procedures, ie they form an integral component of the system design and employ essentially the same equipment for their operation as is used for communication purposes. This is in contrast to "stand-alone" RTCE procedures, such as pulse and chirp sounding, which require separate and expensive RF units.

The topic of RTCE is discussed extensively in the literature, eg (Darnell 1983) (Darnell 1986) (CCIR, 1982), and will not be reviewed in any detail here. It should be pointed out, however, that RTCE is particularly valuable - if not essential - in non-Gaussian environments where even rudimentary models of interference may not be available; this is the current state in the HF band. In this case, it is not possible

to predict in advance what the effects of co-channel interference on the received signal might be: therefore, real-time monitoring is the only option if adaptation of the system parameters in response to channel conditions is envisaged.

In general, embedded RTCE techniques can be classified under the following three headings:

(a) those utilising the normal traffic transmissions of a communication system;

(b) those utilising the normal control signals of a communication system;

(c) those involving a small degree of modification to the traffic or control signals of a communication system, but not involving significant additional cost.

Each of these three classes will now be discussed.

4.1 RTCE via Normal Traffic Transmission

As stated previously, the μ-index technique is an example of RTCE derived from the normal data transmissions of a communication system. Other possibilities include (Darnell 1986):

(i) monitoring of the received traffic error rate;

(ii) phase or zero-crossing measurements on the received signal;

(iii) soft-decision data monitoring at the receiver;

(iv) pseudo-error counting at the receiver;

(v) baseband spectral analysis for noise/interference characterisation, eg template correlation (Jowett 1987).

Clearly, all the above procedures require feedback from receiver to transmitter if adaptation of the signal format is to be carried out at the transmitter. Otherwise, all adaptation must be carried out on an "open-loop" basis at the receiver, which is normally less effective. Section 5 of this paper outlines two techniques which can be applied to enhance the efficiency of open-loop RTCE.

4.2 RTCE via Normal Control Signals

Potentially, some types of control signals present within the communication system also provide a valuable source of RTCE information. Examples include:

(i) the rate at which an error control decoder is detecting and/or correcting errors;

(ii) the number of block repeats requested per unit time in an automatic request repeat (ARQ) system;

(iii) the output of a receiver matched filter/correlator in response to a synchronisation preamble code, eg in terms of the peak-to-sidelobe ratio.

4.3 RTCE via Modification of Traffic or Control Signals

The key consideration for this class of RTCE techniques is that the additional costs of modification should be kept very small relative to the costs of the basic communications system equipment. In this category are procedures such as:

(i) insertion of low-level pilot tones into baseband traffic signal formats (Betts & Darnell 1975);

(ii) use of auxiliary error control in encrypted systems (Darnell 1983);

(iii) use of separate low-cost transceivers for probing alternative assigned channels (Hague 1987).

In practice, two RTCE functions must be implemented in an HF radio system:

(a) monitoring the state of the channel currently carrying traffic;

(b) monitoring the states of alternative assigned channels, not currently in use for traffic.

Different forms of RTCE may be required for these two functions.

The communications system design problem involves selecting and coalescing the information from a number of available RTCE sources in order to produce appropriate and reliable channel models which can then be used as a basis for system parameter adaptation. The updating interval for the RTCE-derived model must be commensurate with the rate at which system parameters can be adapted.

The following four sections of the paper describe specific techniques for enhancing the performance of radio systems operating in non-Gaussian environments. Section 5 introduces signal pre-processing techniques to assist reception and RTCE on an open-loop basis; Sections 6, 7 and 8 discuss signal generation and processing techniques designed to be responsive to RTCE-controlled adaptation in non-Gaussian environments.

5. PRE-PROCESSING TECHNIQUES

In a largely uncharacterised noise/interference environment, it is important that the signal processing procedures at the receiver are given the best possible chance of operating effectively on the corrupted received signal. The two pre-processing techniques described below appear to offer promise in this respect.

5.1 Quenched Filtering

A problem occurs with digital communication systems when the data detection system is biassed, or initialised, by high levels of noise or interference. A typical detector will employ some form of stringent band limitation resulting from the requirement for selectivity; the degree of band limitation will be determined by the frequency domain transfer function of the detector, H(jw). Using basic Fourier analysis theory, it can be shown that the corresponding time domain unit impulse response function of the detector, h(t), and the transfer function form a Fourier transform pair, ie

$$h(t) \rightleftharpoons H(jw) \qquad [7]$$

Therefore, the band limitation in the frequency domain will give rise to time constants in the time domain which will tend to limit the rate of response of the detector to changes in the incoming signal or noise/interference. The actual response obtained at the detector output will also depend upon the nature of the input waveform. Using a simplified example: if the detector is modelled as a 1st-order low-pass filter, the response to a positive-going unit impulse at its input will be as shown in Fig. 4(a) whereas, if the input is a negative-going step function, the response will be as illustrated in Fig. 4(b). Suppose now that a composite input comprising a positive impulse closely followed by a negative step is applied to the detector; the corresponding response will be as shown in Fig. 4(c). In this case, the time before the step detector threshold, t", is increased in comparison with the corresponding interval, t', before the same threshold is crossed in Fig. 4(b); this time increase is due to the detector being initialised by the preceding impulse.

The situation is analogous to the one occurring when impulse noise is present at the input of a receiver. If the detector in Fig. 4(c) had been quenched and then opened at the start of the step, then the detrimental initialisation would have been removed.

In general, the time domain response of the detector, y(t), to an arbitrary input, x(t), would be given by the convolution

$$y(t) = \int_{-\infty}^{\infty} h(u)\, x(t - u)\, du \qquad [8]$$

where u is a variable with the dimensions of time.

This quenching procedure is only feasible if the symbol timing is known at the receiver, ie the system is bit-synchronous. In asynchronous systems, the detector most be kept open continously and is therefore subject to initialisation effects. The Piccolo system (Ralphs, 1985) is an example of a system which essentially employs this form of quenching.

5.2 Buffer Delay

Considering the reception arrangement shown in Fig. 5: a variable buffer delay, τ, is inserted into the signal processing path as indicated, whilst an RTCE analysis unit accepts the received signal directly off-air. This allows a time τ for the RTCE analysis to be applied prior to the received signal being processed. In this way, the parameters of the signal processor can, to some extent, be optimised in response to RTCE analysis; as τ increases, so the precision of the RTCE analysis will tend to improve. As indicated in Section 4, this is an example of open-loop RTCE in which all adaptation is carried out at the receiver and no feedback from receiver to transmitter is involved. Examples of the forms of modification to the signal processing procedures, which could be carried out in response to buffered RTCE data, include:

(a) adaptive band-stop filtering to remove narrowband interfering signals;

(b) peak-clipping of impulsive noise;

(c) identification of signal segments with low reliability, say due to error bursts.

If the transmitted signal format is designed appropriately, ie to facilitate the

extraction of RTCE data, this form of buffered processing can be made more effective. Techniques such as data interleaving and time-diversity are also valuable in this respect.

The maximum allowable value of τ will depend upon the type of traffic being passed over the link; with speech, there will normally be an an upper limit of about 0.5 second before the delay becomes annoying and inhibiting.

6. EMBEDDED MODULATION SCHEMES

An aspect of channel encoding which can be made adaptive in response to path conditions is modulation/demodulation. One way in which this can be achieved is via a technique which will be termed "embedded" modulation. Here, the different elements of the overall modualtion format have different levels of resilience to noise/interference and distortion. Fig. 6(a) illustrates the concept by means of a simple example of an embedded phase-shift keyed (PSK) scheme. In the phase-plane shown, the transmitted signal can assume 8 possible phase states (corresponding to an information content of 3 bits/state):

$$\left.\begin{array}{rcl} 0 & \pm & \Delta\phi \\ \pi/2 & \pm & \Delta\phi \\ \pi & \pm & \Delta\phi \\ 3\pi/2 & \pm & \Delta\phi \end{array}\right\} \qquad [9]$$

where

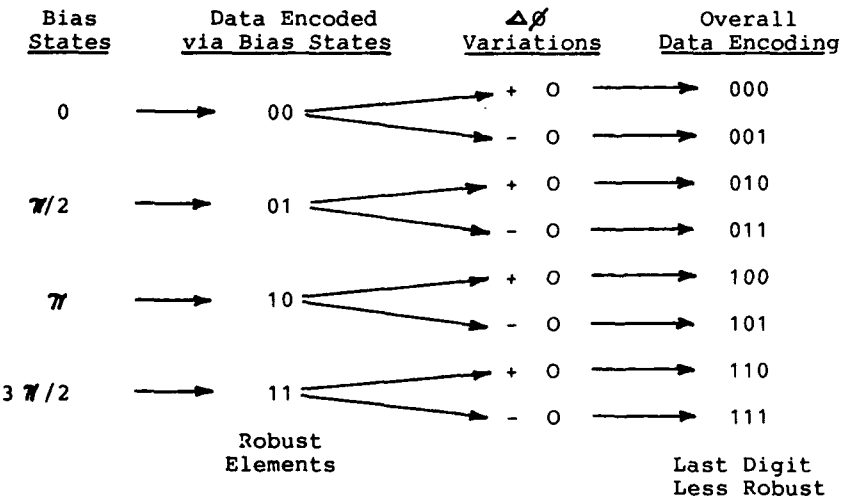
$$\Delta\phi \ll \pi/2 \qquad [10]$$

Under stable, low-noise channel conditions, all phase states will be completely resolvable. As the channel degrades, however, phase changes of the order of $\Delta\phi$ may no longer be detectable and only the "bias" states

$$\left.\begin{array}{r} 0 \\ \pi/2 \\ \pi \\ 3\pi/2 \end{array}\right\} \qquad [11]$$

will be resolvable. Thus, the more important data elements would be associated with these bias states, with the less important data being transmitted via the $\pm\Delta\phi$ variations when channel quality permits.

The data mapping might be as shown below:



The first two digits of each 3-bit encoded word would correspond to the more important data, whilst the third digit would be assigned to data of less importance.

Fig. 6(b) shows how an amplitude modulated (AM) component can be added to what would otherwise be a constant envelope transmission. This low-level AM again would only be detectable under relatively favourable channel conditions.

Components of the overall modulation format would be designed to provide specified levels of tolerance to noise, with this tolerance depending primarily upon:

(a) type of modulation, ie ASK, FSK, PSK or DPSK;

(b) detection scheme, ie synchronous or asynchronous;

(c) magnitude of the modulation state changes.

The objective is to allow the information rate achievable to vary in a controlled manner in response to channel conditions.

In the following section, the concept of embedded error control coding is introduced: practically, it is anticipated that embedded error control coding and modulation schemes will be designed in a co-ordinated manner.

7. EMBEDDED ERROR CONTROL CODING

Embedded error control (Darnell et al, 1987) is a technique designed to adapt the instantaneous information rate in an ARQ system to the available channel capacity. The principle is illustrated in Fig. 7(a): here, the same information stream is simultaneously encoded at different rates and then put into a composite format which is transmitted a block at a time over the communication channel. At the receiever, decoding of each block takes place at the highest rate which the channel can support at that time, with the ARQ protocol indicating at which rate decoding has been possible in the previous block.

The system has been implemented using a (52, 12) array code format, an example of which is shown in Fig. 7(b). In addition to two-dimensional array coding, the information is also obviously protected by repetition coding since blocks K1 and K2 occur more than once in the array. The disadvantage of this particular format is its asymmetry in that protection for K1 is greater than that for K2, which in turn is greater than for K3 and K4. This has been overcome to some extent by the modified (64, 23) array code format shown in Fig. 7(c), in which a more symmetric protection is achieved.

Fig. 8 shows the results of tests over a simulated HF path for

(a) a normal (52, 36) array code,

(b) the embedded (52, 12) array code,

(c) the embedded (64, 23) array code,

all transmitted at a symbol rate of 200 bits/s.

Fig. 9 illustrates the input-output error rate transfer characteristics for the same systems over the same simulated HF channel.

It is interesting to note that the embedded coding technique provides significant performance gains over conventional array coding under bad channel conditions, ie decoder input error rates of 1-in-10 or worse. Both embedded codes continue to provide usable throughput in an ARQ configuration when the normal array code has locked-up. Similarly, the embedded codes produce about an order of magnitude improvement in output error rate, for a given input error rate, as compared with the normal array code. Thus, the embedded error control technique is tending to provide performance improvements under conditions where constant-rate error control schemes would fail.

Further development of the embedded error control coding concept is currently proceeding; trials over real HF paths are being carried out and will be reported shortly. In principle, there is no reason why the number of coding levels should not be increased over the number employed in the preliminary tests; this would allow a more precise matching of transmission rate to channel capacity - possibly at the expense of throughput.

8. MULTI-FUNCTIONAL CODING

The concept of multi-functional coding (Darnell & Honary, 1986) seeks to amalgamate the following source and channel coding functions in a systematic manner

(a) multiple/random access techniques;
(b) security/privacy;
(c) speech digitisation/data compression;
(d) synchronisation;
(e) error control;
(f) modulation;
(g) diversity processing.

The aim is to generate a single integrated coding procedure in which coding power can be assigned adaptively between the various functions in response to channel conditions; elements to enable this to take place are:

(h) RTCE for channel characterisation;

(i) signal generation and processing architectures which can be adapted in response to RTCE data;

(j) a systematic approach to coalescing and weighting elements (a) to (g) above

in any given radio environment, based upon the application of artificial intelligence (AI) and expert system (ES) methodologies.

To date, the elements (a) to (g) above have largely been considered as separate aspects of communication system design. Substantial work on the combination of modulation and error control coding, via the vehicle of soft-decision decoding, has been published eg (Chase, 1973). However, as yet, such combination does not appear to have been extended to more than two functions in a systematic manner.

Taking an example of data security: a combination of functions might be particularly attractive since security is not necessarily based upon the function of encryption alone, but may also depend upon other aspects of source and channel encoding together with system control procedures. There is little point in providing high levels of cryptographic security if transmission reliability is so low that the operational requirement cannot be met. In attempting to counteract this unreliability, the system security may well be compromised in other ways, eg by distinctive signatures, lack of radiated power control, excessive synchronisation intervals, susceptibility to direction finding, etc. Thus, it can be seen that the components of the design problem are extremely interactive.

Since the channel characteristics affect all aspects (a) to (g) above, they cannot be considered as truly independent in that they all contribute to total system performance. It is logical, therefore, to view them as connected and interactive elements of a wider system design problem.

Source and channel characteristics will normally be specified probabilistically; it is thus convenient to apply formalised AI techniques to the problem of integrating, weighting and adapting the various functions in a radio environment as sensed by RTCE. This approach, which is computationally intensive, is rendered technically and economically feasible by the ability to embed powerful processors into radio terminals; thus each terminal becomes an expert system in its own right.

9. CONCLUDING REMARKS

In this paper, a number of concepts and techniques intended for application to radio systems operating under non-Gaussian conditions have been discussed. They are currently being investigated via several co-ordinated research programmes, with the aim of integrating them progressively into the architectures of automatic HF radio systems. However, it should be stressed that the techniques are not specific to HF radio: in principle, they could also be applied to other forms of radio system with non-Gaussian noise/interference backgrounds.

In conclusion, the following points are emphasised:

(a) The ability to embed compact and relatively powerful processors within the terminals of radio systems at reasonable cost allows the incorporation of AI and ES techniques to assist system control and adaptation.

(b) Non-Gaussian radio channels exhibit a time-varying information transmission capacity due to both propagation and noise/interference effects. Because interference in particular tends to be largely uncharacterised, RTCE is an essential element of the radio system design if efficient parameter adaptation is to be achieved.

(c) The ARQ mode of transmission recognises the variable-capacity nature of the radio path, since it provides a variable throughput rate. For this reason, the techniques described in this paper are assumed to be applied to a system operating in an ARQ mode, and are intended to make such an architecture more effective.

(d) The unifying concept, encompassing the range of individual channel and source coding techniques, is that of multi-functional coding. The objective of this procedure is to allow the embedded processing power of the terminals to be distributed systematically between the various signal generation and processing functions in a way which is responsive to path conditions at any time, and also optimises the overall system performance in that environment.

In many respects, the techniques discussed here are at an early stage of development. However, they do represent an attempt to consider the fundamental nature and requirements of radio communication over channels which, in a number of important aspects, are not characterised with any precision.

10. REFERENCES

1. Lee, Y.W., 1960: "Statistical theory of communication", Wiley.

2. Omura, J.K. & Shaft, P.D., 1971: "Modem performance in VLF atmospheric noise", IEEE Trans. Vol. COM-19, No. 5.

3. Gott, G.F., Wong, N.F. & Dutta, S., 1982: "Occupancy measurements across the entire HF spectrum", AGARD Conf. Proc. CP-332.

4. Darnell, M., 1986: "The influence of spectral congestion on HF (2 - 30 MHz) communication system design", IERE Internat. Conf. on "Electromagnetic compatibility", York, IERE Pob. No. 71.

5. Wolfowitz, J., 1967: "Memory increases capacity", Information & Control, Vol. 11.

6. Zolghadr, F., 1987: "Improved coding and control for communication over dispersive radio channels with non-gaussian noise", Coventry Polytechnic Research Report No. CEES 2/87.

7. Darnell, M., 1983: "Real-time channel evaluation", AGARD Lecture Series No. 127 on "Modern HF communications".

8. Darnell, M., 1986: "Embedded real-time channel evaluation", AGARD Lecture Series No. 145 on "Propagation impact on modern HF communications system design".

9. CCIR, 1982: "Real-time channel evaluation of ionospheric radio circuits", Report 889 in Doc. of XVth Plen. Assy., ITU, Geneva.

10. Jowett, A.P., 1987: "Artificial intelligence in HF communication systems", AGARD Specialists' Meeting on "Effects of electromagnetic noise and interference on performance of military radio communication systems", Lisbon.

11. Betts, J.A. & Darnell, M., 1975: "Real-time HF channel estimation by phase measurements on low-level pilot tones", AGARD Conf. Proc. CP-173.

12. Hague, J., 1987: "Improved coding and control for HF systems in a non-Gaussian noise environment", as 10.

13. Ralphs, J.D., 1985: "Principles and practice of multi-frequency telegraphy", Peter Perigrinus.

14. Darnell, M., Honary, B. & Zolghadr, F., 1987: "Embedded coding technique: principles and theoretical studies", to be published in Proc. IEE, Part-F.

15. Darnell, M. & Honary, B., 1986: "Multi-functional coding schemes applicable to secure communication", IEE Internat. Conf. on "Secure communication systems", CP-269, London.

16. Chase, D., 1973: "A combined coding and modulation approach for communication over dispersive channels", IEEE Trans., Vol. COM-21, No.3.

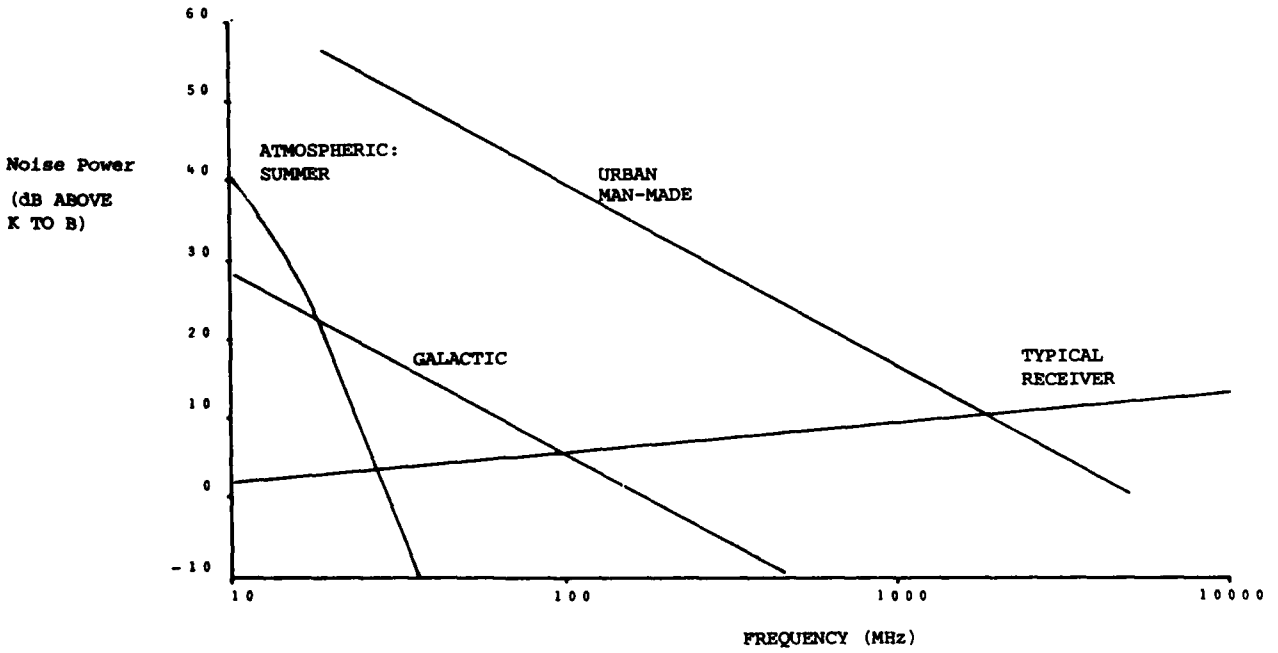


FIG.1 AVERAGE NOISE POWERS FROM VARIOUS SOURCES

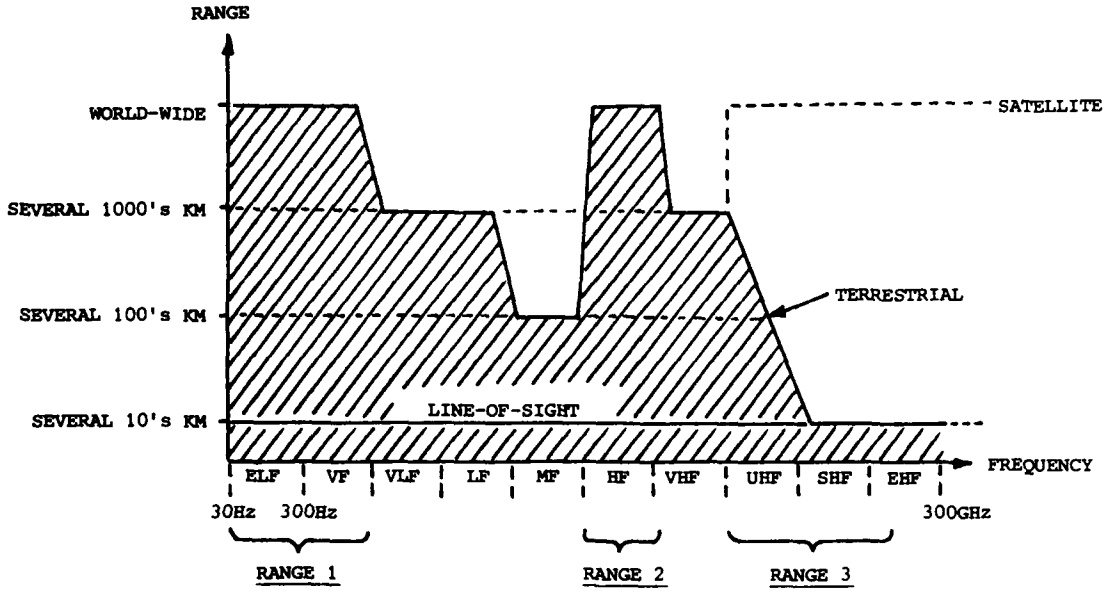


FIG.2 VARIATION OF RANGE WITH FREQUENCY

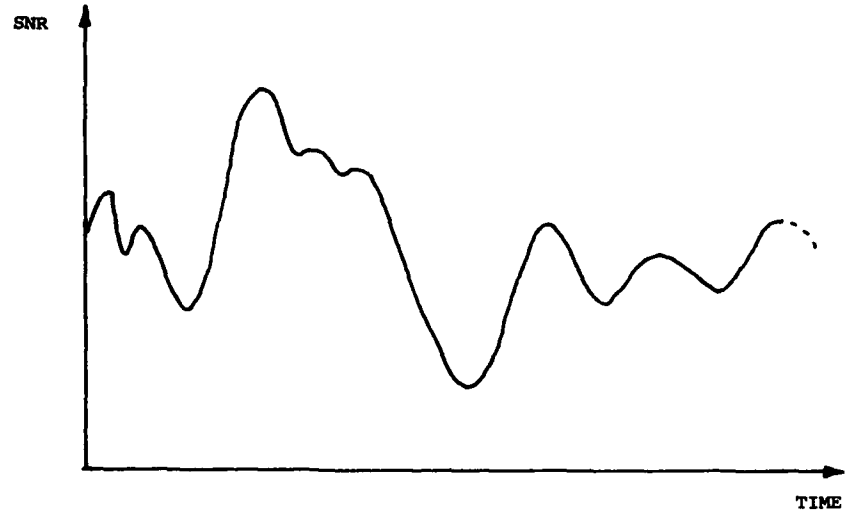


FIG.3 VARIATION OF SNR WITH TIME FOR RADIO CHANNEL

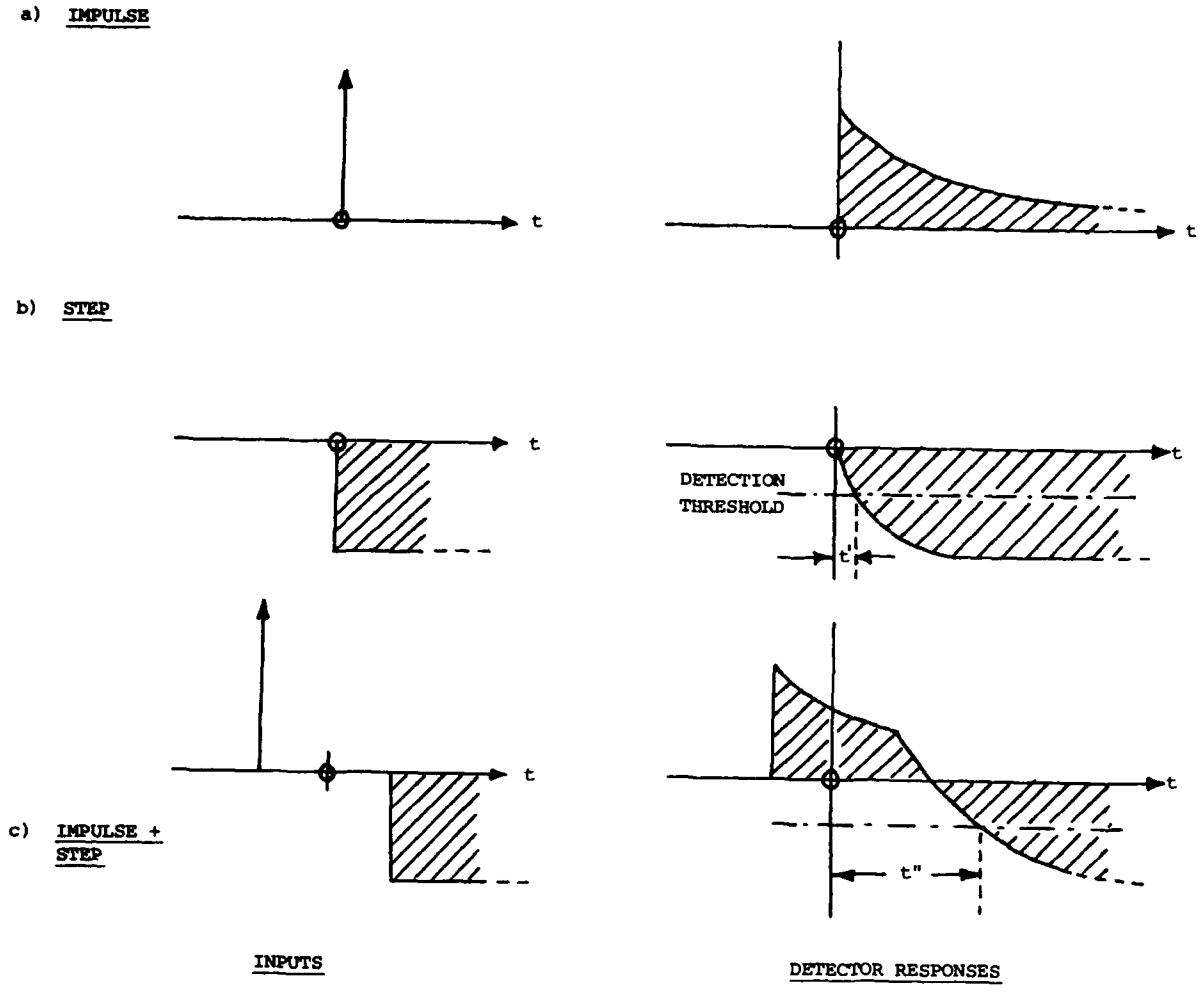


FIG.4 INITIALISATION OF DETECTOR BY IMPULSIVE SIGNAL

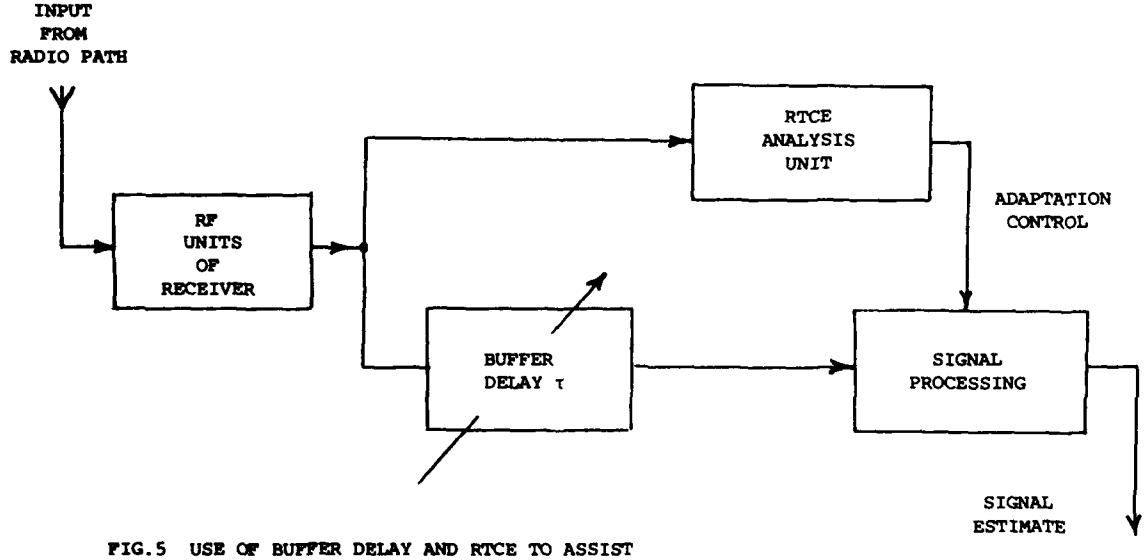


FIG.5 USE OF BUFFER DELAY AND RTCE TO ASSIST RECEIVED SIGNAL PROCESSING

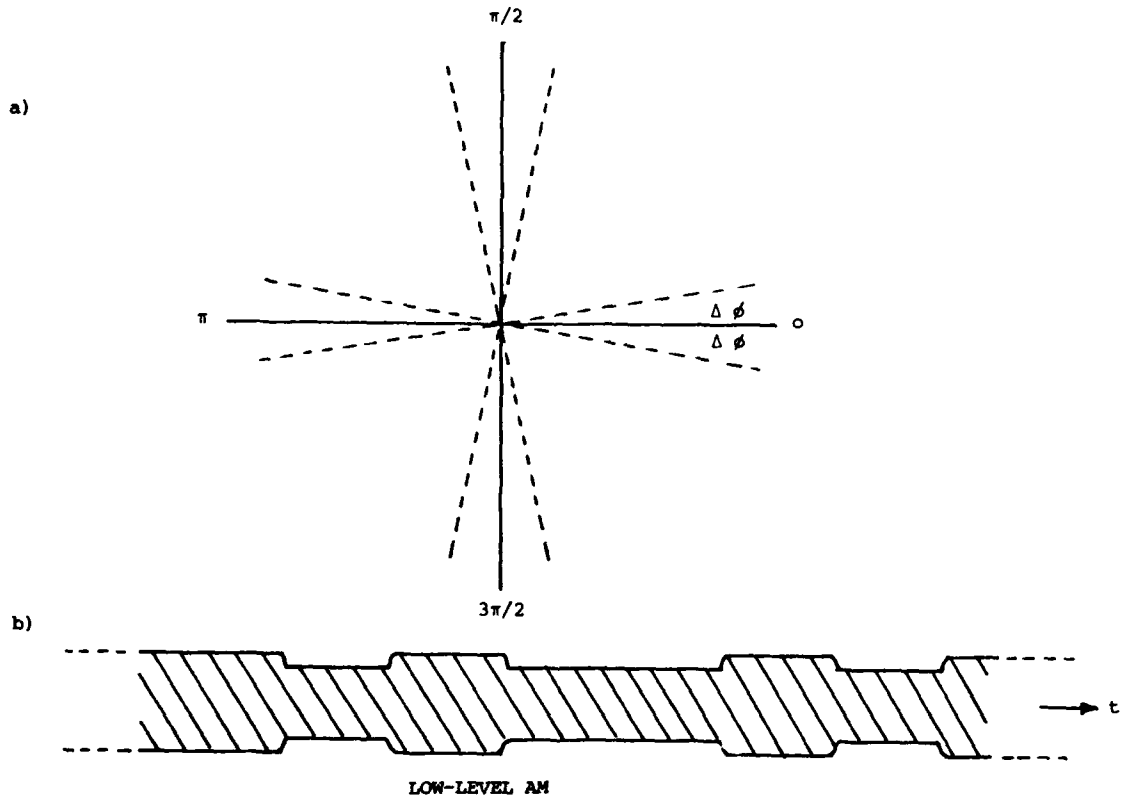
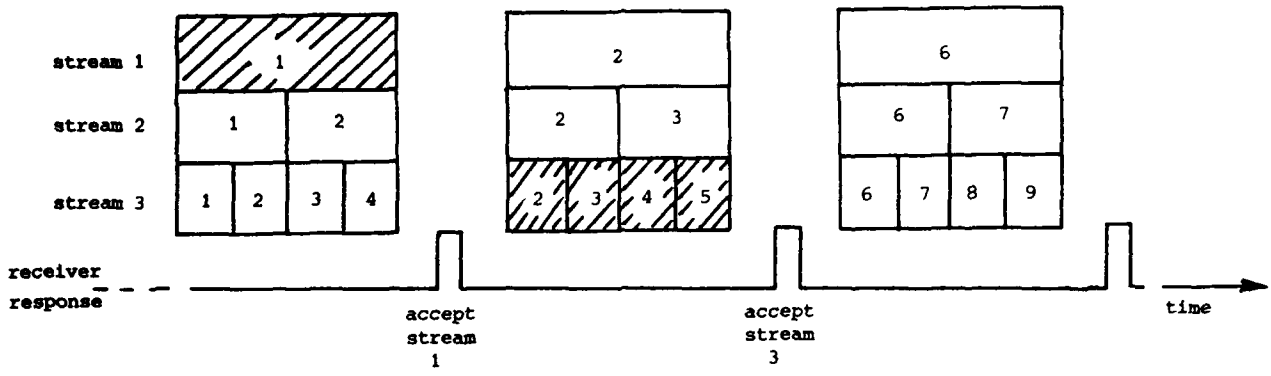
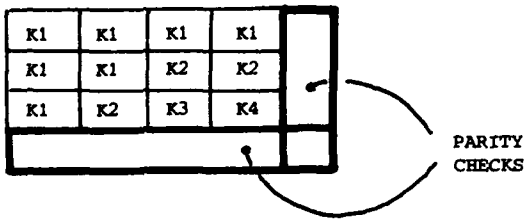


FIG.6 EXAMPLE OF AN EMBEDDED MODULATION SCHEME



(a) Principle of embedded encoding.



a_1	a_1	a_1	a_1	b_1	b_1	b_1	b_1	c_1	c_1	c_1	c_1	P_1
a_1	a_1	b_1	b_1	c_1	c_1	a_2	a_2	b_2	b_2	c_2	c_2	P_2
a_1	b_1	c_1	a_2	b_2	c_2	a_3	b_3	c_3	a_4	b_4	c_4	P_3
P_{16}	P_{15}	P_{14}	P_{13}	P_{12}	P_{11}	P_{10}	P_9	P_8	P_7	P_6	P_5	P_4

(b) (52,12) embedded array code block structure.

k_1	k_1	k_1	k_1	k_1	k_2	k_2	k_2	k_2	k_2	k_3	k_3	k_3	k_3	k_3	P_1
k_4	k_4	k_4	k_5	k_5	k_5	k_6	k_6	k_6	k_7	k_7	k_7	k_8	k_8	k_8	P_2
k_9	k_{10}	k_{11}	k_{12}	k_{13}	k_{14}	k_{15}	k_{16}	k_{17}	k_{18}	k_{19}	k_{20}	k_{21}	k_{22}	k_{23}	P_3
P_{19}	P_{18}	P_{17}	P_{16}	P_{15}	P_{14}	P_{13}	P_{12}	P_{11}	P_{10}	P_9	P_8	P_7	P_6	P_5	P_4

(c) (64,23) modified embedded array code block structure.

Fig. 7 Embedded error control coding.

Simulated throughputs under hf
conditions, rms fade rate 2.74
Baud rate 200 Bits/sec.

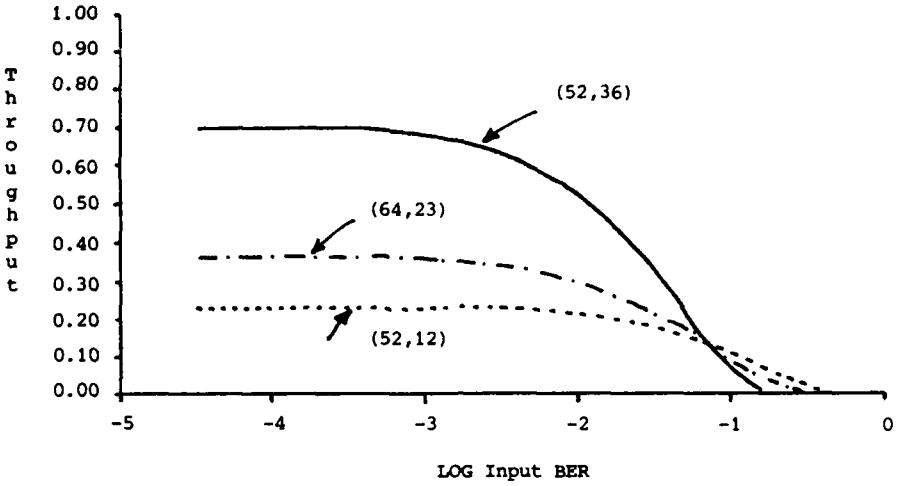


Fig. 8 Simulated throughput comparison of array, embedded array and modified embedded array codes.

Simulated reliability under hf
conditions, rms fade rate 2.74
Baud rate 200 bits/sec.

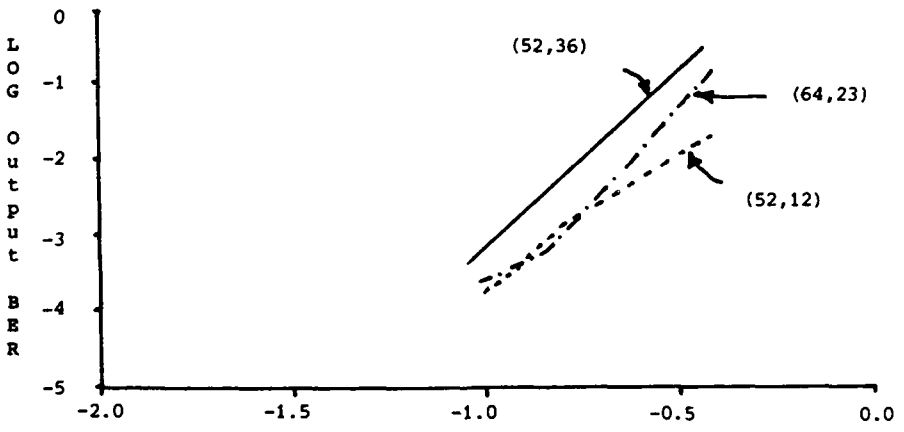


Fig. 9 Simulated reliability comparison of array, embedded array and modified embedded array codes.

IMPROVED CODING AND CONTROL OF HF SYSTEMS IN A NON-GAUSSIAN NOISE ENVIRONMENT.

J.Hague, AMIEE
Department of Electronics,
University of York,
Heslington,
York YO1 5DD
UNITED KINGDOM.

SUMMARY

This paper describes an HF communication system which is capable of adapting various parameters, such as frequency and coding scheme, to provide the user with the most reliable HF path.

The system outlined in Figure 1 is to be employed. Each transceiver/ATU is of amateur-grade, and is capable of delivering approximately 150W over the range 2 to 30 MHz. The transceivers are fully computer-controllable, and this enables automatic channel selection to be achieved. When used with 10m whip antennas, a cheap and portable HF system results; however, for some channels, the 10m antenna will not present an ideal match to the system, and thus the automatic tuning facility of the ATU is used to achieve a match, which may take up to 3 seconds. Due to this constraint, broadband antennas for the probing channel become desirable, at the expense of making the system less portable.

The coding/control software runs on two 8085 systems, and analyses current/alternative channel status to set the optimum frequency and coding scheme to use. Each 8085 system is equipped of 8 Kbytes of ROM and 8 Kbytes of RAM, and therfore reasonably complex algorithms may be accomodated. Additional memory cards may be added to the system at a later date as required.

Presently, Golay (23,12) and other half-rate codes (the Golay (23,12) code is in fact implemented as a (24,12) code for ease of system timing), are supported by the system, but it is intended to utilise other coding schemes, such as the embedded-array code (Darnell, Honary & Zolghadr, 1986) at a later date by modifying the software.

1 INTRODUCTION

The HF spectrum provides the user with a unique set of channels in that it is difficult to design a system suitable to cope with all its different, varying parameters and yet still present acceptable data to the user.

There are several ways of overcoming the unsuitable nature of such a medium:

(i) Overdesign the system by using very high powers, and large antenna arrays - typically tens of kilowatts, and arrays occupying several acres of land.

(ii) Incorporate some artificial intelligence into the system by utilising propagation prediction programs to set the optimum working frequency.

(iii) Employ operators to monitor the traffic, and other channels to determine an optimum channel to use.

The above three methods all suffer drawbacks:

In the overdesigned system, signals may still not propagate satisfactorily, and interference may still be caused to adjacent users. This type of system is also not an appropriate basis for an easily installed, portable station for field use, due to the fact that high power output amplifiers are generally physically large and, as mentioned above, the antenna will require a large amount of space.

In (ii), the calculated optimum channel is used, but the calculation is unable to take into account effects such as SIDs, auroral blackout and man-made interference; therefore it may still not provide a satisfactory working channel for the user.

Whilst solution (iii) seems to provide a satisfactory answer to the problems associated with HF, it is very man-intensive, thus adding excessive life-cycle costs associated with manning, and is still prone to operator error/fatigue.

A better system would be one without operator intervention, and could itself adjust the traffic channel / error control coding scheme to suit the prevailing conditions. This type of system may use a propagation prediction program to define a starting point on which to set-up the system's data flow.

2 AN ADAPTIVE SYSTEM CONCEPT

One such system is outlined in Figure 2: it can be seen to comprise of two amateur-grade transceivers and associated software routines. One of the transceivers is devoted to probing and assessing the channels available to the user, according to some appropriate algorithm, whilst the other carries data on a channel determined by the control software. It can also be seen that the error-control coding scheme is also under the control of the same software, thus providing a variable-redundancy coding scheme for the user data.

The usefulness of being able to vary the amount of redundancy added to data can be seen from the plot of any block code's performance, shown in Figure 3; for very high error rates, the bit error rate for the coded case is worse than that for the uncoded case, and it can therefore be seen that the added redundancy serves no purpose except to reduce the throughput. Work has been done on simulating a variable-redundancy coding scheme (Hellen, 1985) and Figure 4 shows the variation in throughput achieved when using such a scheme. Real tests of such a scheme have also been performed (Goodman and Farrell, 1975)

In order to take full advantage of the channels available, the data channels at each end of the link may transmit on different frequencies, as the HF path may not be reciprocal between the two sites, and therefore some means of informing each end of the link of the correct channel to use must be devised.

Referring to Figure 5, it can be seen that the channel for data transmitter A to use, and the coding scheme for coder A to employ, is determined by the results of probing/error measurements taken by receiver B, which have been compiled into a table indicating the best channel and code to use at any one time; similarly for transmission in the opposite direction. It can be seen that by responding to these commands to change channel and coding scheme, then optimum use can potentially be made of the channels available.

3 CHANNEL ENCODING/DECODING PROCEDURES

The channel encoding/decoding procedures can basically be split into 2 parts; error-control coding and modulation.

3.1 Error-control coding

The error-control coding schemes are implemented in software on an 8085-based system, with 8K ROM and 8K RAM for memory, although this may be upgraded at a later date by the inclusion of additional memory cards. The execution time associated with a software-based system is not a limiting factor since, for this application, it is intended to operate only at up to 150-300 bits per second maximum source rate.

The algorithm used to code/decode the data is based upon matrix generation of the code (see Appendix 1); thus, in order to specify a code for the system to employ, it is only necessary to define the generator matrix, the length of the code and the syndrome look-up table for the decoder. This may be performed via the RS232 serial interface between the coding and the control boards. At present only half-rate codes have been implemented in order to simplify system timing, although it is intended to cater for other codes in the future by employing the system's internal timers and interrupt structure to clock data in and out of the software buffers at differing rates.

Initial results for the Golay (23,12) code have been promising (Figure 6) in that a coding gain of approximately 2dB was achieved using a source rate of 150bps. The Golay (23,12) code is in fact implemented as a (24,12) code, to simplify system timing, by adding a extra, unused zero to each 23-bit codeword.

Basic error rates for the channel may be deduced from the operation of the decoder, since a look-up table technique is used. This means that for any given syndrome produced by the decoder, there will be associated with it an error pattern, which is used to correct the incoming word. Knowledge of this error pattern enables the user to gain an estimate of the current error-rate. However, due to the nature of the look-up table, ie. all the error patterns used are of minimum weight, some larger error patterns may be mistaken for minimum weight patterns and thus will give a false indication of error-rate. The look-up table for the larger error patterns is also known, and therefore may be switched with the minimum-weight table, to provide more reliable error-correction for channels with a higher error-rate.

3.2 Modulation

A 2-tone FSK modem has been designed for use by the system (Figure 7). A simple 2-tone design was felt to be sufficient to cope with the low-rate traffic, both control and data, expected to be passed by the system. However, due to the modular construction of the modem, with each bandpass filter contained on a single Eurocard, future upgrade to more than 2 tones is possible. The important features of this design are:

3.2.1 Frequency agility

The corner frequencies of both the lowpass and the highpass filters which make up each bandpass filter are computer-controllable, as are the two modulator tones. Thus it can be seen that, in conjunction with a suitable algorithm, shifting the tones to regions of low interference within a given channel is possible.

3.2.2 Variable threshold

The threshold, with which the summed outputs of the bandpass filters is compared, is made automatically variable to enable it to lie midway between the two signal levels, thereby helping to combat frequency-selective fading. This is achieved as follows:

(the higher of the two tones is taken to represent a 'space', i.e. a logic '0')

The rectified output from the 'space' filter $x_s(t)$, is integrated over 10 bit intervals,
$t_1 = 10T$, where T equals the bit period, to give

$$y_s(t) = \int_{t_1} x_s(t)\, dt \qquad (3.1)$$

Similarly, for the 'mark' filter:

$$y_m(t) = \int_{t_1} x_m(t)\, dt \qquad (3.2)$$

The difference between these is taken to be the time-varying threshold, $t_h(t)$:

$$t_h(t) = y_s(t) - y_m(t)$$

$$= \int_{t_1} x_s(t)\, dt - \int_{t_1} x_m(t)\, dt \qquad (3.3)$$

A decision is then made on the output as:

$$\text{Output} = 0 \text{ if } (x_s - x_m) > t_h$$
$$= 1 \text{ if } (x_s - x_m) < t_h \qquad (3.4)$$

(in practice, the output is also converted to RS232 levels)

Also included is the option to have a fixed threshold of zero volts, as the output can become noisy whilst using a variable threshold; for example, under 'mark' only, or 'space' only conditions, the threshold moves towards the corresponding level.

An on-board microcontroller is to be added to enable full control of the modem over a serial network, together with the sending of A-D conversion values for the signal and threshold levels, to allow use of soft-decision decoding if required. It is also intended to provide a memory facility in the controller to enable standard rates / tone spacings to be called from memory, thus allowing for the possibility of quickly changing the rate of data transmission.

Initial results for the modem are presented in Figure 8, though for the purposes of future calculations, the performance is taken to be:

$$p_e = 1/2 \exp [-E_b/N_o] \qquad (3.5)$$

Where, E_b = signal energy per bit, N_o = noise power spectral density for white noise

The tones are presettable (as are the filters) to within 12 Hz, over the range 500 Hz to 3.0 kHz, and should thus support the normal range of HF bit-rates.

4 SYNCHRONISATION AND CONTROL OF EACH END OF THE LINK

The control software assumes that the current data channel is satisfactory, since this will have been the best choice given by the control software and therefore appends control data for the other end of the link, in the form of new channel and/or new coding scheme, in the header of the user data. The data format is of the form shown in Figure 9.

The header serves various purposes, these being:

(i) The reversals inform the receiver software of the prescence of a signal, and also serve to enable the decoding software to derive clock synchronisation. This is done by running an internal clock in the decoder at a multiple of the incoming data clock frequency and using this to sample the incoming reversals. If the sampling rate is 16 times the data rate, then the decoder will fill a buffer of length 16 with the incoming samples. Clock synchronisation can then be derived by waiting for the weight of the buffer to reach a predefined threshold, say 15 (thus allowing for a margin for noise without severely degrading the clock's phase). This will indicate that a full bit reversal has been received, and that the transmitter clock is about to change state. Therefore the receiver clock may now be started at the expected data rate.

(ii) The synchronisation sequence used is the concatenated Barker (7,11) sequence (see Figure 10) which enables frame synchronisation to be achieved down to a high error rate (Barker, 1953). To determine the robustness of such a sequence, assume that cross-correlation in the polar-binary case is taken. Thus, at zero shift ie. maximum correlation, the peak height is reduced by 2 for each bit error. For an (m,n) sequence, the peak height h, with no errors is given by:

$$h = mn \qquad (4.1)$$

Thus, with r errors added, the height is now:

$$h = mn - 2r \qquad (4.2)$$

Therefore, for a probability of error, p,

$$h = mn - 2pmn$$

$$= mn\,(1-2p) \qquad (4.3)$$

In a Gaussian white noise environment, using equation (3.5), equation (4.3) becomes,

$$h = mn\,(1 - 2\exp[-E_b/2N_o]) \qquad (4.4)$$

Frame synchronisation may be found by waiting for the cross-correlation function value to exceed some predefined threshold, determined by the maximum error-rate that can be satisfactorily handled by the following coding scheme(s). It is possible to make this threshold variable to accomodate for different signal-to-noise conditions, although this is not implemented in the system described.

(iii) After the frame synchronisation has been sent, the next 48 bits are encoded in Golay (24,12) format, and comprise of two 12-bit contorl words. The control words inform the contorl software on which channel to trasmit data, or with which code to encode the user data. These control codes have to be robust to minimise system hold-ups caused by changing to the wrong channel/code. The method chosen to encode these control words comprises the following steps:

(a) Choose a control word of either CHANNEL or CODE which is minimally spaced by a distance of 4 - the words 1010 and 0101 have been chosen. For future use with more commands, the minimum spacing of 4 will not apply, and therefore the control will be degraded if left unmodified.

(b) Repeat the word 3 times to form a 12-bit word. This enables a majority vote to be carried out at the decoder to minimise the probability of false decoding.

(c) Encode in Golay (24,12) format, thereby providing an additional error-correction facility of up to 3 errors out of each block of 24 bits. This also provides an error-detection facility of up to 6 errors in 24 bits.

Obviously, we cannot provide minimum distance decoding for the control parameter, x eg. CHANNEL X, due to its nature, thus step (a) will be missed out for the second of the two control words.

4.1 Probability of a command being incorrectly decoded

The probability of incorrectly decoding a control word may be calculated: (assuming an ideal modem is used) by:

(a) The probability of decoding a bit in error over the channel, using a half-rate code, and a 2-tone FSK modem is:

$$p_e = 1/2 \exp [-E_b/4N_o] \qquad (4.5)$$

(b) A Golay decdoing error occurs if 4, or more errors in a block of 23 bits, although error-detection may still be possible, depending on the error pattern (Shanmugam, 1979); thus:

$$p_g = \sum_{i=4}^{23} \begin{bmatrix} 23 \\ i \end{bmatrix} p_e^i (1 - p_e)^{23-i} \qquad (4.6)$$

where p_e is defined in (4.5)

(c) For the repetitions of the 4-bit 'word', a bit error occurs if 2, or more bits out of the triplet are in error:

$$p_r = \begin{bmatrix} 3 \\ 2 \end{bmatrix} p_g^2 (1 - p_g) + \begin{bmatrix} 3 \\ 3 \end{bmatrix} p_g^3$$

Thus,

$$p_r = 3p_g^2 - 2p_g^3 \qquad (4.7)$$

where p_g is defined in (4.6)

(d) For the minimum-distance decoding of the first control word only, a decoding error occurs if 2, or more bits from the 4 are in error:

$$p_m = \begin{bmatrix} 4 \\ 2 \end{bmatrix} p_r^2 (1 - p_r)^2 + \begin{bmatrix} 4 \\ 3 \end{bmatrix} p_r^3 (1 - p_r) + \begin{bmatrix} 4 \\ 4 \end{bmatrix} \qquad (4.8)$$

where p_r is defined in (4.7)

A similar expression exists for the probability of incorrectly decoding a control parameter, since an error will occur if 1 or more bits from the 4 are in error. In fact, the expression can be written in terms of equation (4.8):

$$p_p = p_m + \begin{bmatrix} 4 \\ 1 \end{bmatrix} p_r (1 - p_r)^3 \qquad (4.9)$$

A plot of the theoretical performance of the above stages is presented in Figure 11, and it can be seen that the system collapses at very-low signal-to-noise ratios.

By using this graph, a value for the frame synchronisation cross-correlation threshold can be determined, by ensuring that frame synchronisation is not achieved for an unacceptable command word error-rate.

5 CHANNEL EVALUATION

Two forms of channel evaluation have to be used in order to achieve a high level of performance from the system. These are evaluation of the channel in use, and evaluation of alternative channels available.

5.1 Current channel evaluation

On start-up of the system, passive probing of all the channels may take place (up to the predicted MUF) to determine a starting channel, depending on the channel which appears to be the clearest. Passive probing will typically take the form of scanning all available channels, and awarding them a figure of merit depending on which has the lowest unwanted signal level.

After start-up, an estimate of the channel's usability may be found in several ways:

(i) The (peak-minus-sidelobe) height of the frame synchronisation correlation is directly proportional to the probability of error for the channel (Figure 12) and therefore, by calculating this value, a rough estimate of the channels' bit error rate may be ascertained.

(ii) As stated in section 3.1, an estimate of the channel bit error-rate may be obtained from looking at the error pattern which the decoder adds to the incoming data in order to correct errors. This will also enable the control software to acsertain whether the errors come in bursts, or are random, and may thus alter the coding scheme to suit.

(iii) The maximum usable data rate may be found in an ARQ system by using embedded-array coding (Darnell, Honary & Zolghadr, 1986). Figure 13 shows a typical flow of data for such a scheme, and it can be seen that by accepting the appropriate data stream, 1, 2, or 4, data block throughput is available, since data transmission continues after the last correctly received block. Thus if stream 3 were accepted, then this is equivalent to 4 blocks in which, the lower rate stream 1 was accepted; hence the system is essentially self-adjusting to the channel's available capacity (Figure 14).

(iv) Use of the signal and threshold levels from the modem enables the control software to determine the relative strengths of each tone, and thus provides an indication of frequency-selective fading, since the signal level in each bit-interval will give the level for each tone; ideally, these two levels should be the same.

5.2 Alternative channel evaluation

The process involved in evaluating alternative channels is basically the same as that for the current channel evaluation. However, this process will be faster, since only a probing sequence, for example the Barker (7,11) sequence, will be sent. The reception of such a sequence will enable steps (i) and (iv) above to be performed on the alternative channel. Adding data in the form of embedded-array coding may also enable the probe to determine the maximum data rate on that channel, at the expense of more time taken to probe all channels.

Once all the channels have been probed, a table is compiled indicating the best channel / code which is available to the receiver. This information is then transmitted to the other end of the link in the user's data header, as detailed in Section 4, and a channel / code update will then occur, hence maximising throughput.

6 OVERVIEW AND PRELIMINARY RESULTS

A flow chart of the operations involved in probing and selecting channels is given in Figure 15, and it is hoped to implement the full system in the near future. This flow chart does not detail specific operations involved in the probing and channel assessment algorithms, since these are not yet fully developed. However, the basis for channel assessment via the probing channel has been presented in the above text.

Basic tests conducted so far have involved 'off-air' testing of the data header format detailed in Section 4, and preliminary results are shown in Figure 16. As can be seen, due to the length of the synchronisation sequence and control codes, the length of each trial was limited to only 255, and therefore provides only a rough indication of the results. Indeed, for some of the higher signal-to-noise ratios, the resolution over 255 trials was not enough to enable a consistent result to be achieved.

It is hoped to conduct some live trials of the control system in the near future, over a path length of approximately 100 miles. Further work can then begin on providing a reliable means of probing alternative channels which, when coupled with a suitable control algorithm, will provide the basis for meaningful test.

Initial tests will be carried out using just one transceiver to carry both the data and probing information, but this will not limit the effectiveness of a fully operational two-transceiver system as detailed. It is anticipated that the results of practical tests will be presented at the Specialists Meeting.

Once a system comprising the above blocks is completed, it should provide the user with a reliable, portable and cheap means of utilising the HF spectrum.

APPENDIX A

GOLAY (23,12) ENCODING/DECODING USING MATRICES.

1.1 Before a matrix description of the coding scheme is given, a description using the codes' generator polynomial will be given, as the two are related.

Assume that there exists a generator polynomial, g(x) for an (n,k) cyclic code, where k incoming bits are encoded into an n bit word. For a message given by m(x), the parity p(x) may be calcuated by:

$$p(x) = \text{rem}\ [\ x^{(n-k)}.m(x) \ /\ g(x)\ ] \qquad (A.1)$$

where rem is the remainder resulting from the division, and m(x) and g(x) are the polynomial representations of the message and the generator respectively.

For the Golay (23,12) code, the generator polynomial is given by,

$$g(x) = x^{11} + x^{10} + x^6 + x^5 + x^4 + x^2 + 1 \qquad (A.2)$$

The codeword polynomial, c(x) is now formed by appending the parity to the message:

$$c(x) = x^{(n-k)}.m(x) \ , \ p(x) \qquad (A.3)$$

The message is output to the channel first, since this is the most significant.

At the decoder, a polynomial, s(x), known as the syndrome is produced, which is used in the error correction process.

$$s(x) = \text{rem} [\ c(x) \ / \ g(x)\ ] \qquad (A.4)$$

When s(x) = 0, then either the codeword is uncorrupted, or an undetectable error pattern has occurred.

When s(x) <> 0, then it may be used to find the most likely error pattern that has occurred, to enable error correction to take place.

1.2 Using a matrix description

Since a polynomial calculation will be relatively complex to be performed in real-time on a computer, another means of calculating the parity and the syndrome must be used.

For a given code there also exists an (n,k) generator matrix G, which is given by:

$$G = [I \ P] \qquad (A.5)$$

where I is the k identity matrix, and P is the parity matrix. G for the Golay (23,12) code is given by:

$$G = \begin{array}{ll} 100000000000 & 10101110001 \\ 010000000000 & 11111001001 \\ 001000000000 & 11010010101 \\ 000100000000 & 11000111011 \\ 000010000000 & 11001101100 \\ 000001000000 & 01100110110 \\ 000000100000 & 00110011011 \\ 000000010000 & 10110111100 \\ 000000001000 & 01011011110 \\ 000000000100 & 00101101111 \\ 000000000010 & 10111000110 \\ 000000000001 & 01011100011 \end{array} \qquad (A.6)$$

G may be formed from the generator polynomial, by using the identity matrix elements as input to equation 1.1. This will produce the corresponding parity matrix. It may also be formed by only generating a subset of the total elements, and by shifting and adding (in modulo-2) the elements of this subset, the remaining elements in G may be found. (It is possible to shift the codewords, for a cyclic code, and produce another valid codeword; this is the definition of a cyclic code).

The codeword C is produced from a message word M by:

$$C = M \ G \qquad (A.7)$$

where the above is a matrix multiplication.

At the decoder, the syndrome is again calculated, by multiplying out the transpose of the parity-check matrix, H, by the incoming codeword R, to form the syndrome word S:

$$H = [\ P^T \ I\ ] \qquad (A.8)$$

where P^T = parity section of (A.6) transposed
I = n-k identity matrix

thus,

$$S = R \ H^T \qquad (A.9)$$

In practice, the received codeword, R may not be equal to the transmitted codeword, C, due to errors added by the channel, and thus S will be non-zero (for the system described, the syndrome is used as an offset into a lookup-table containing correctable error-patterns, thus error-correction is reasonably straightforward).

1.3 Half-rate codes.

To make the system timing simple, half-rate codes were chosen. The Golay (23,12) code is 'padded-out' by the inclusion of an extra bit to change it to a (24,12) code. The extra bit may be used as an overall parity check, but has not been implemented as such in the system described.

To 'pad-out' the code, a column of zeros is added at the right-hand side of the P matrix in (A.6), and similarly for the H matrix. The position of the extra bit is at the least-most significant bit and does not affect the performance of the Golay (23,12) code, except for the slight increase in the channel data rate.

Codes other than half-rate ones may be implemented later by using two system clocks; one clock to input data from the user into a buffer, and the other clock to output data from the codeword buffer. This is done at present, but by inputting data on one edge of the clock, and outputting data on both edges; therefore providing the necessary half-rate.

REFERENCES.

[1] Darnell M, Honary B K, and Zolghadr F, 1986, 'Embedded coding techniques : principles and theoretical studies'. (submitted to Proc. IEE - F for publication: available in draft form).

[2] Hellen P, 1985, 'The provision of computer quality data transmission on HF', IEE Conf. Proc., No 245.

[3] Goodman R M F, and Farrell P G, 1975, 'Data transmission with variable-redundancy error control over a high-frequency channel', Proc. IEE, Vol. 122, No. 2, February.

[4] Barker R H, 1953, 'Group synchronising of binary digital systems', Communication Theory, Butterworth, London, pp 273-287.

[5] Shanmugam K S, 1979, 'Digital and analogue communication systems', Wiley, New York, pp 486-490.

ACKNOWLEDGEMENTS.

I would like to thank Dr. E. D. Chesmore at the University of Hull for his assistance in producing the results shown in Figures 10 and 12.

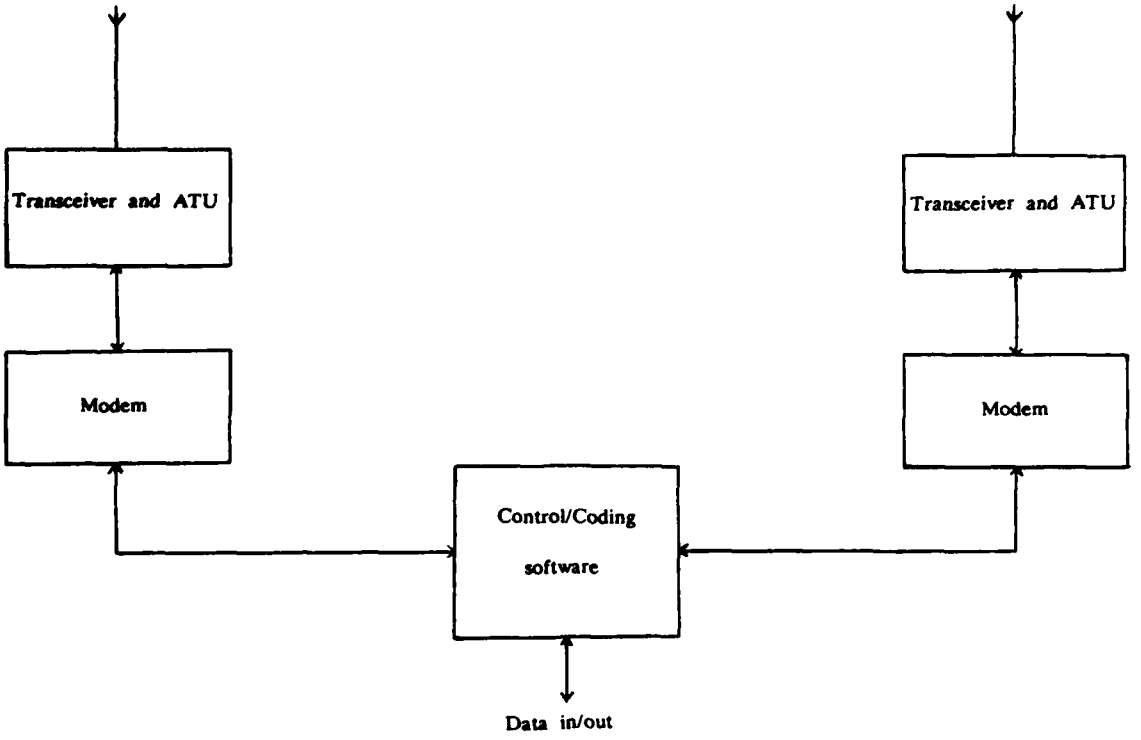


Figure 1 Basic system block diagram

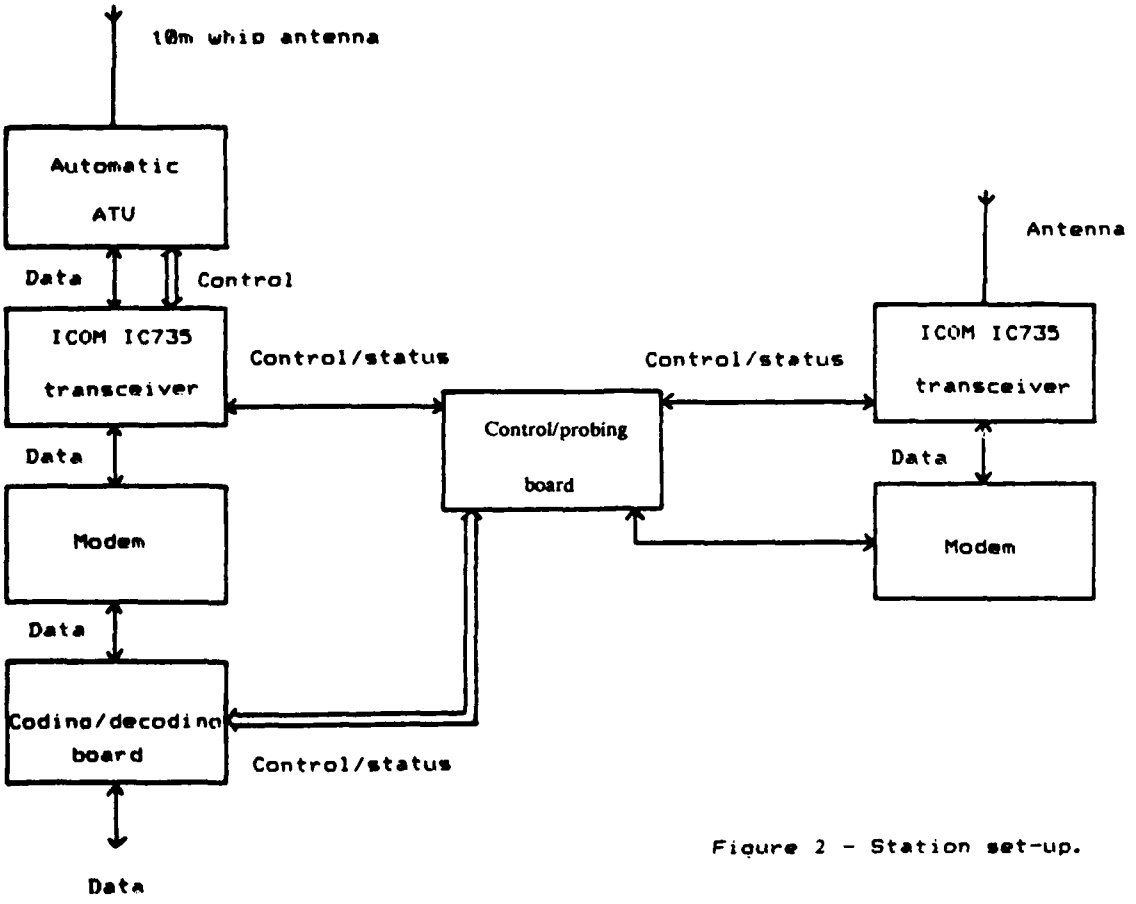


Figure 2 - Station set-up.

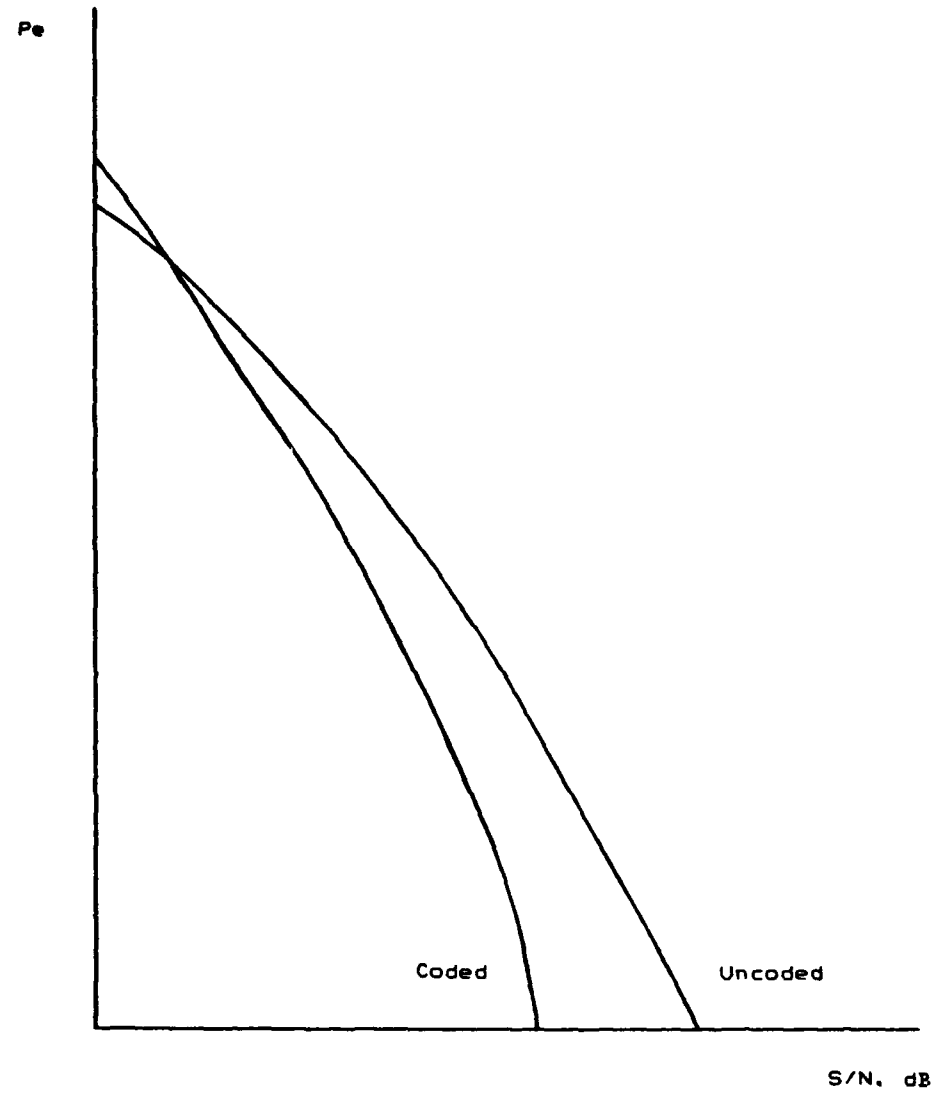


Figure 3 - Block code performance.

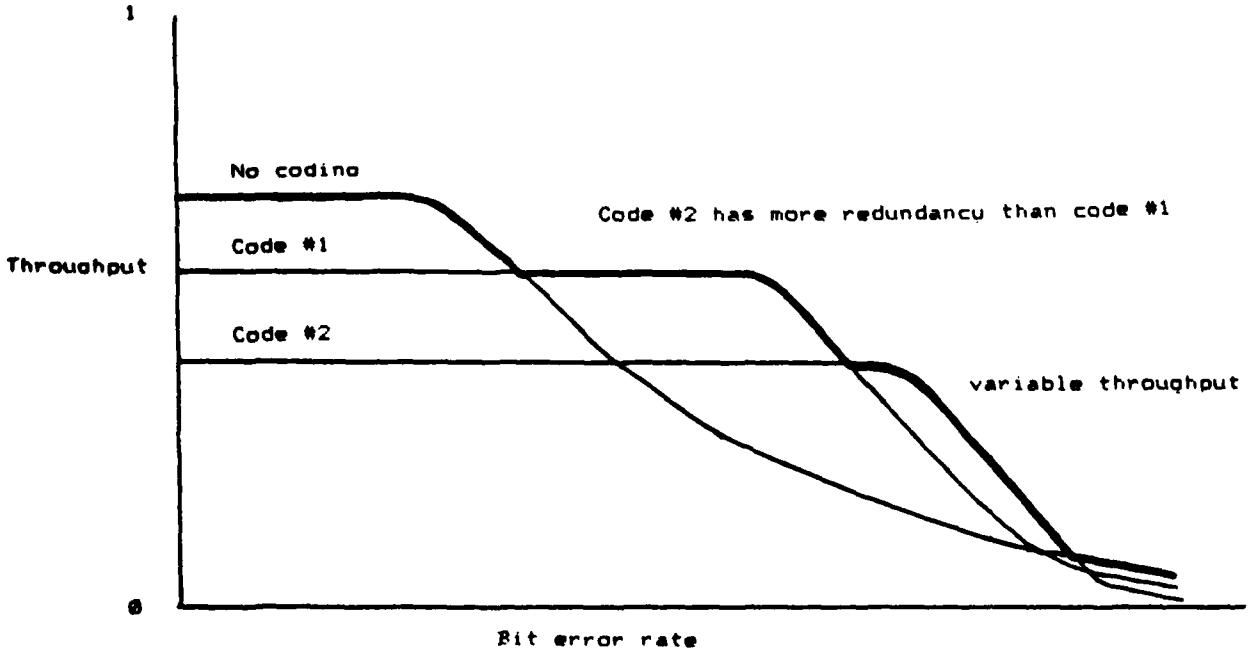


Figure 4 - Variable-redundancy coding.

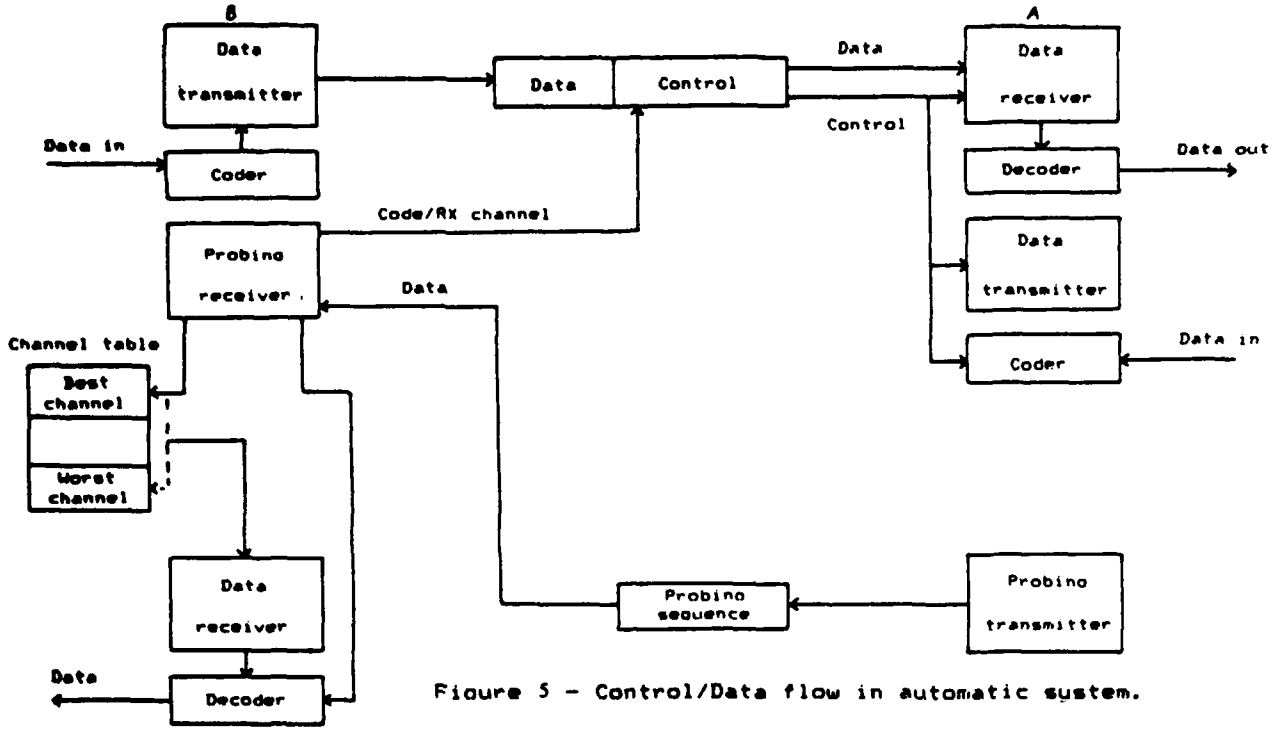


Figure 5 - Control/Data flow in automatic system.

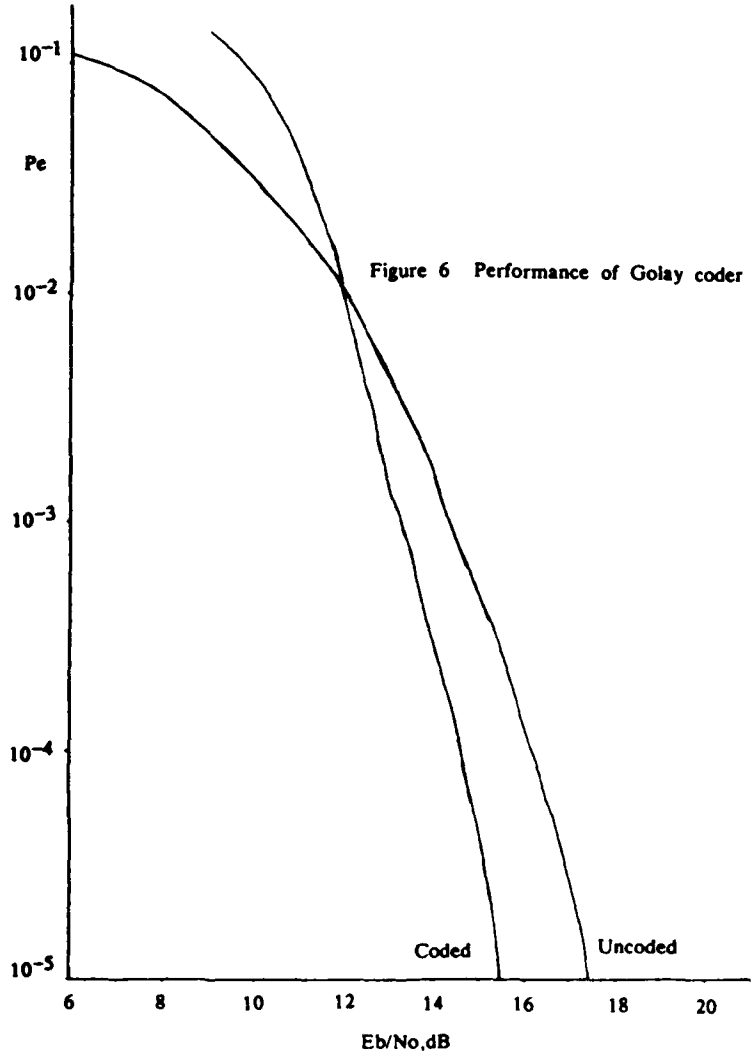


Figure 6 Performance of Golay coder

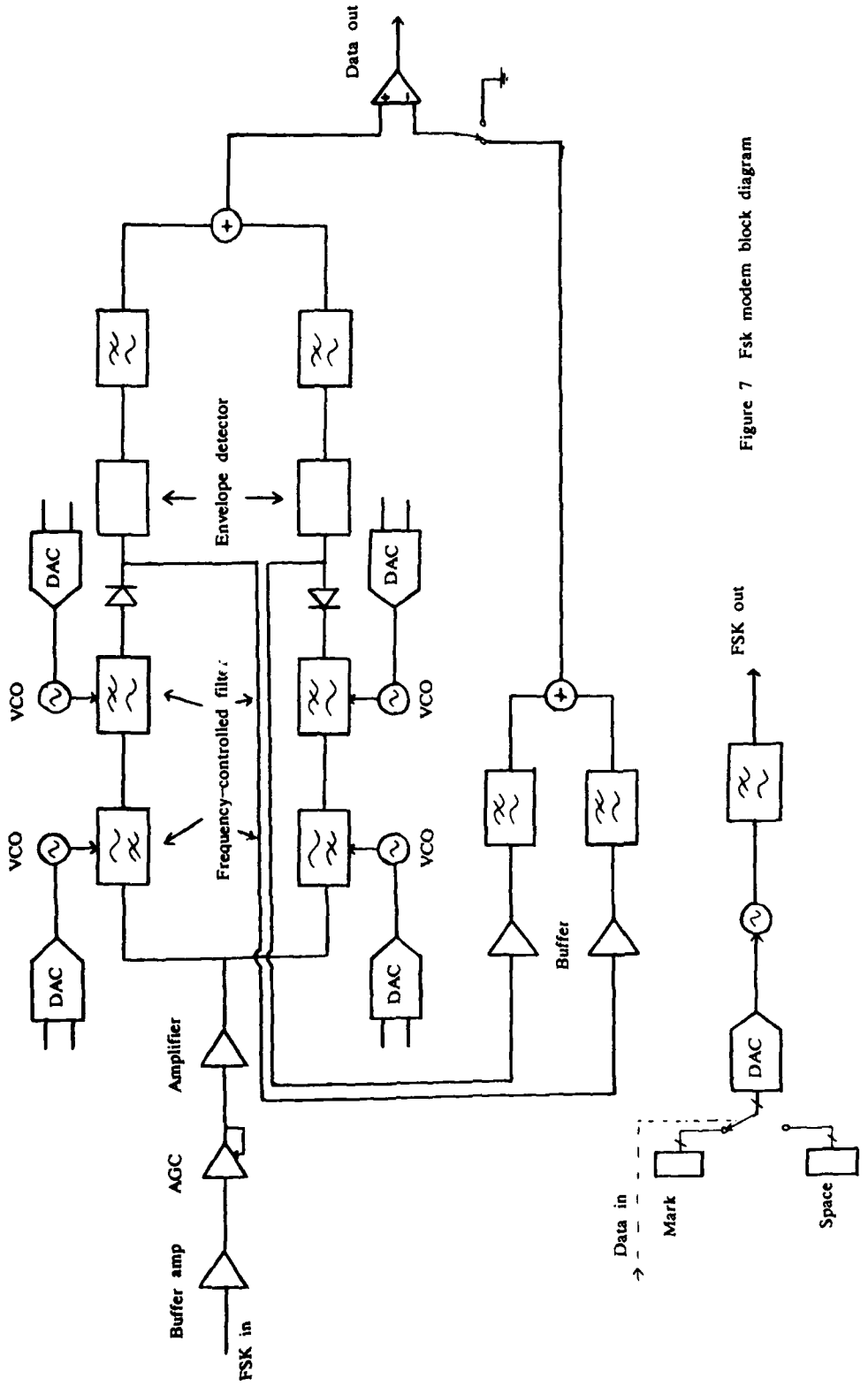


Figure 7 Fsk modem block diagram

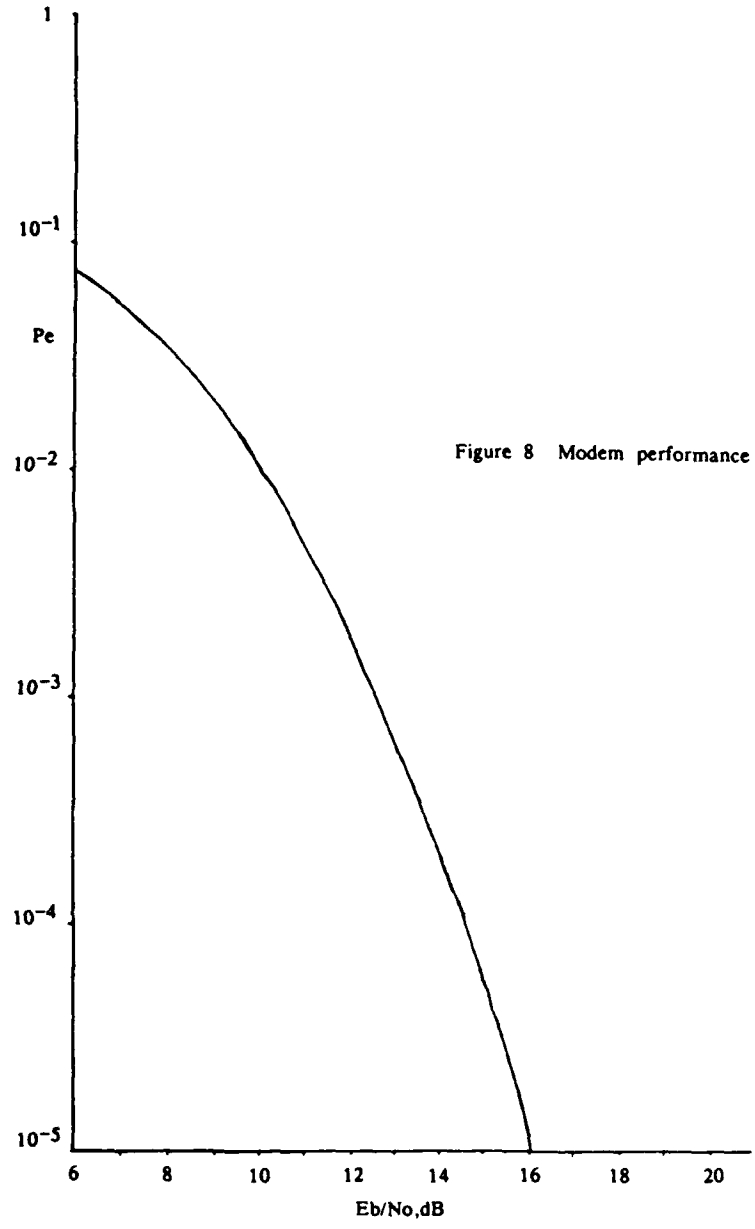


Figure 8 Modem performance

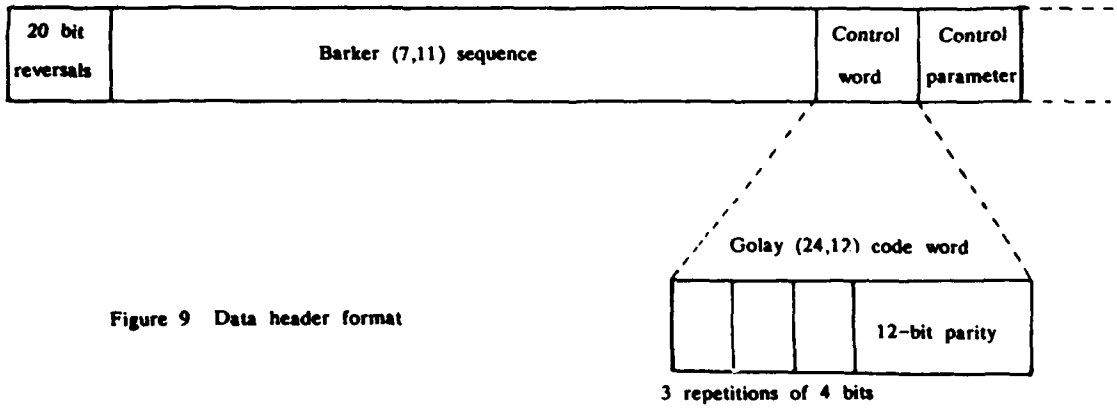


Figure 9 Data header format

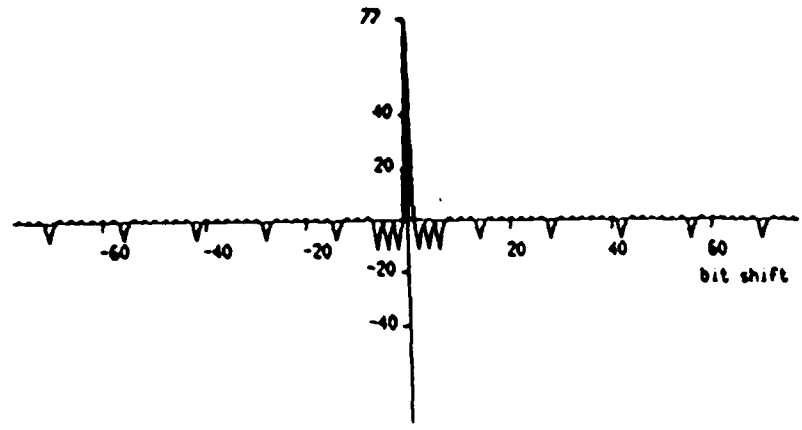


Figure 10 Correlation of concatenated Barker Sequence B(7,11)

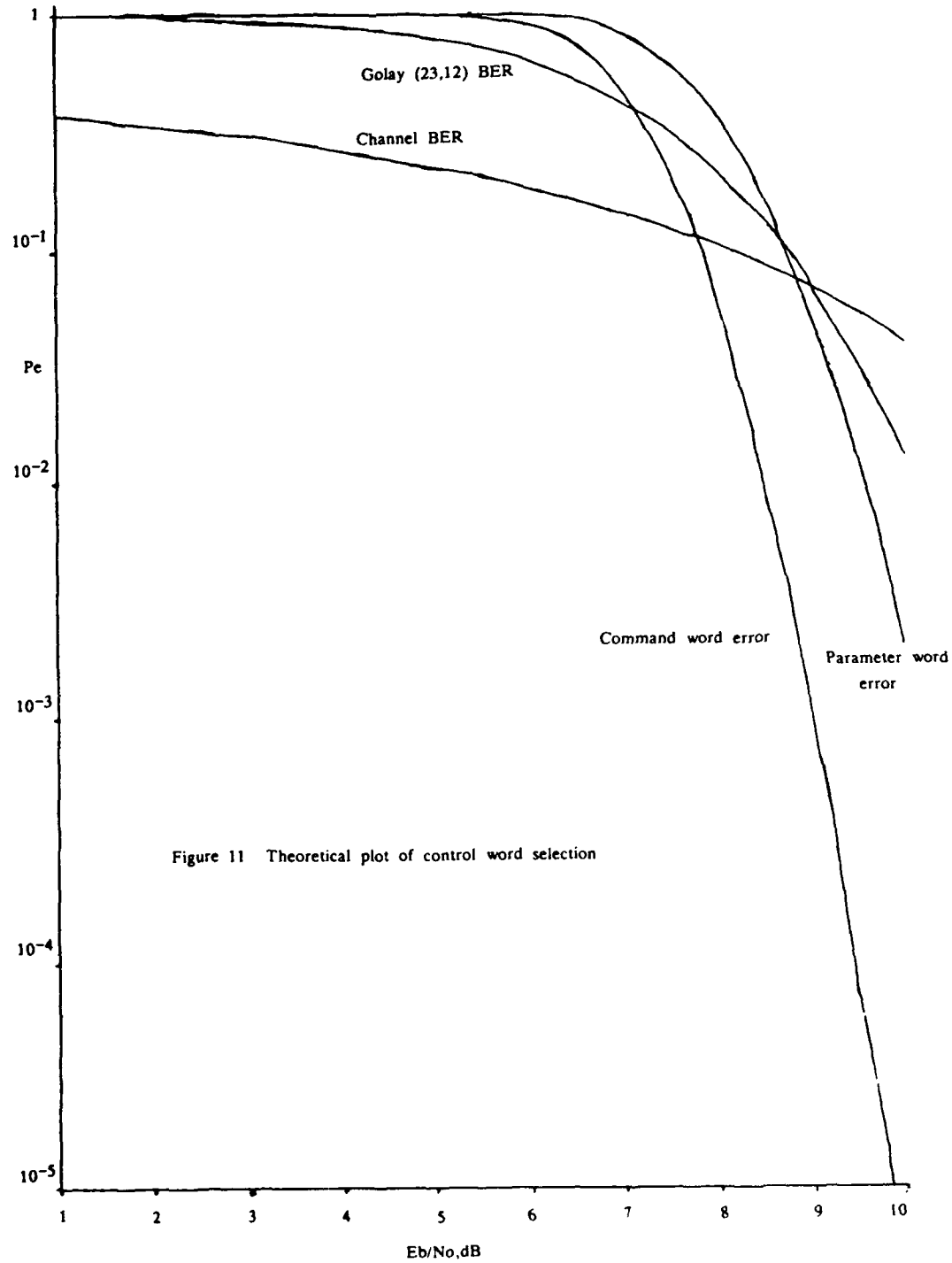


Figure 11 Theoretical plot of control word selection

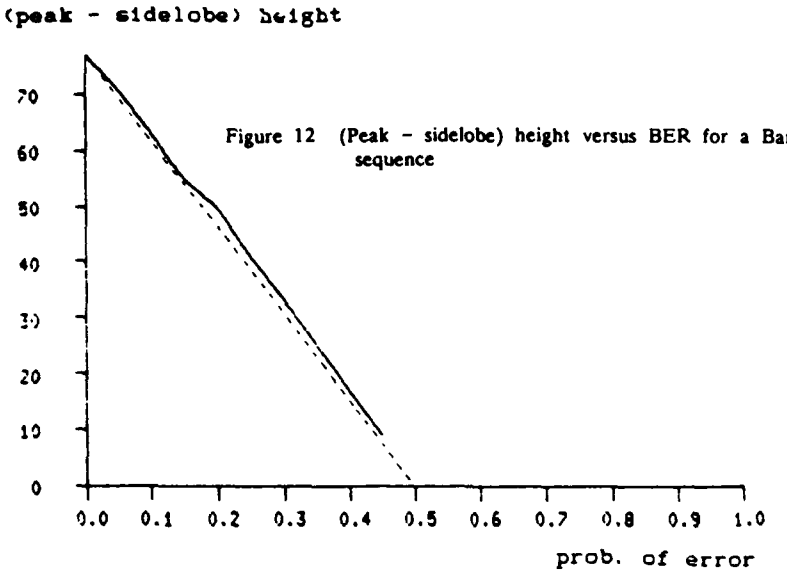


Figure 12 (Peak - sidelobe) height versus BER for a Barker (7,11) sequence

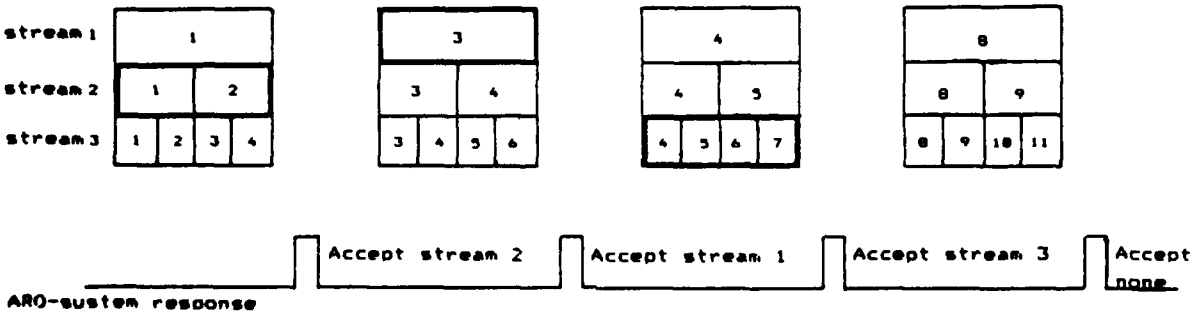


Figure 13 - Embedded array code operation.

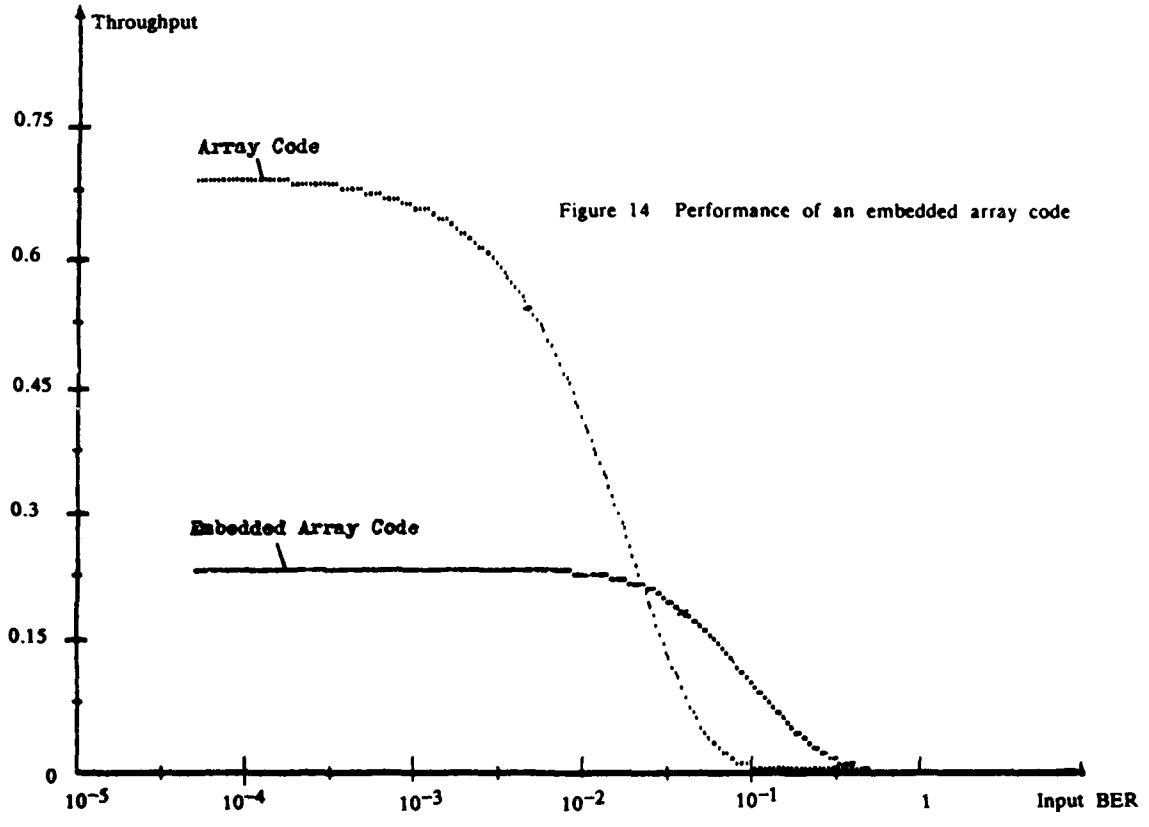


Figure 14 Performance of an embedded array code

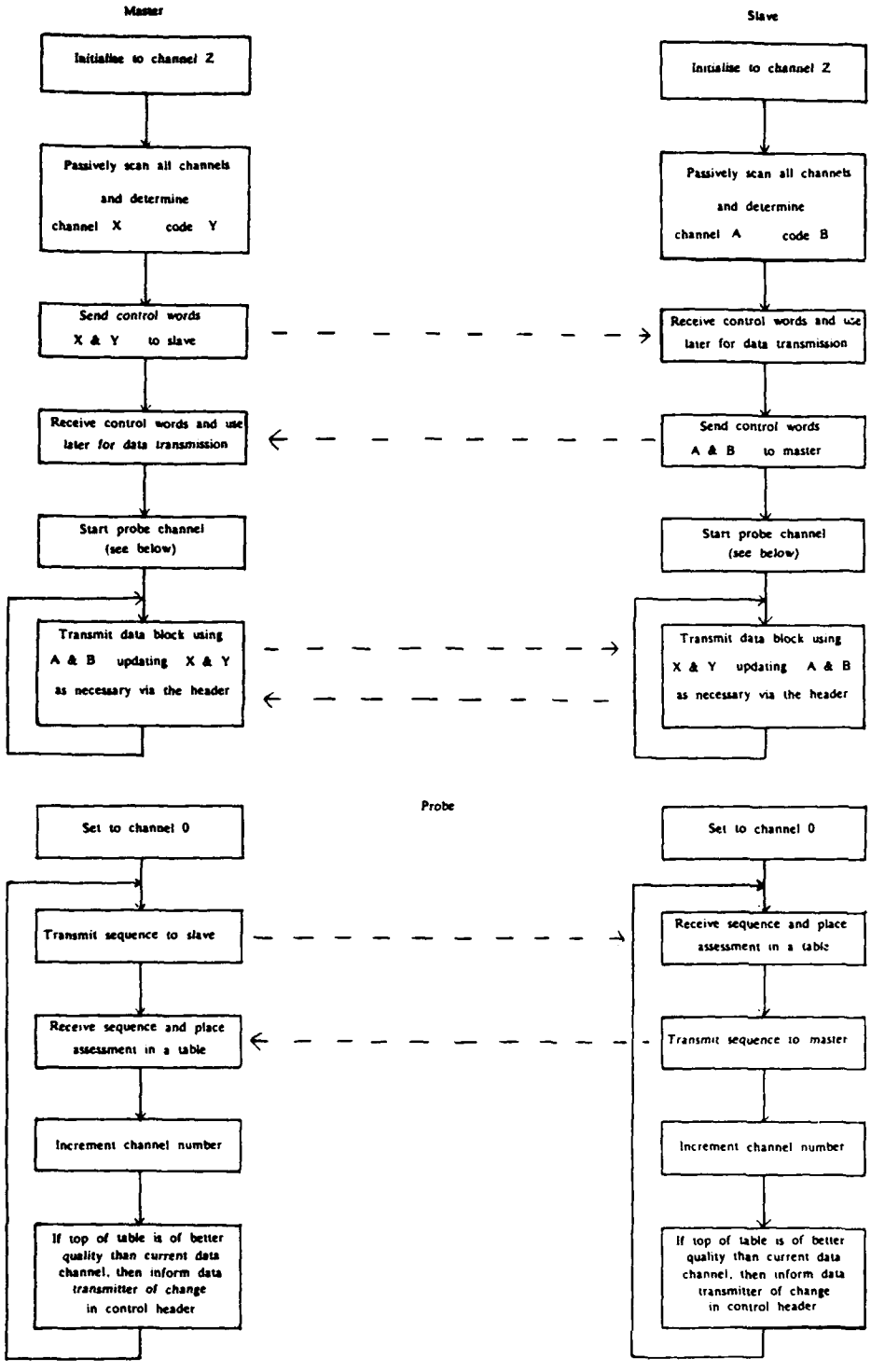


Figure 15 Flow-chart of system operation

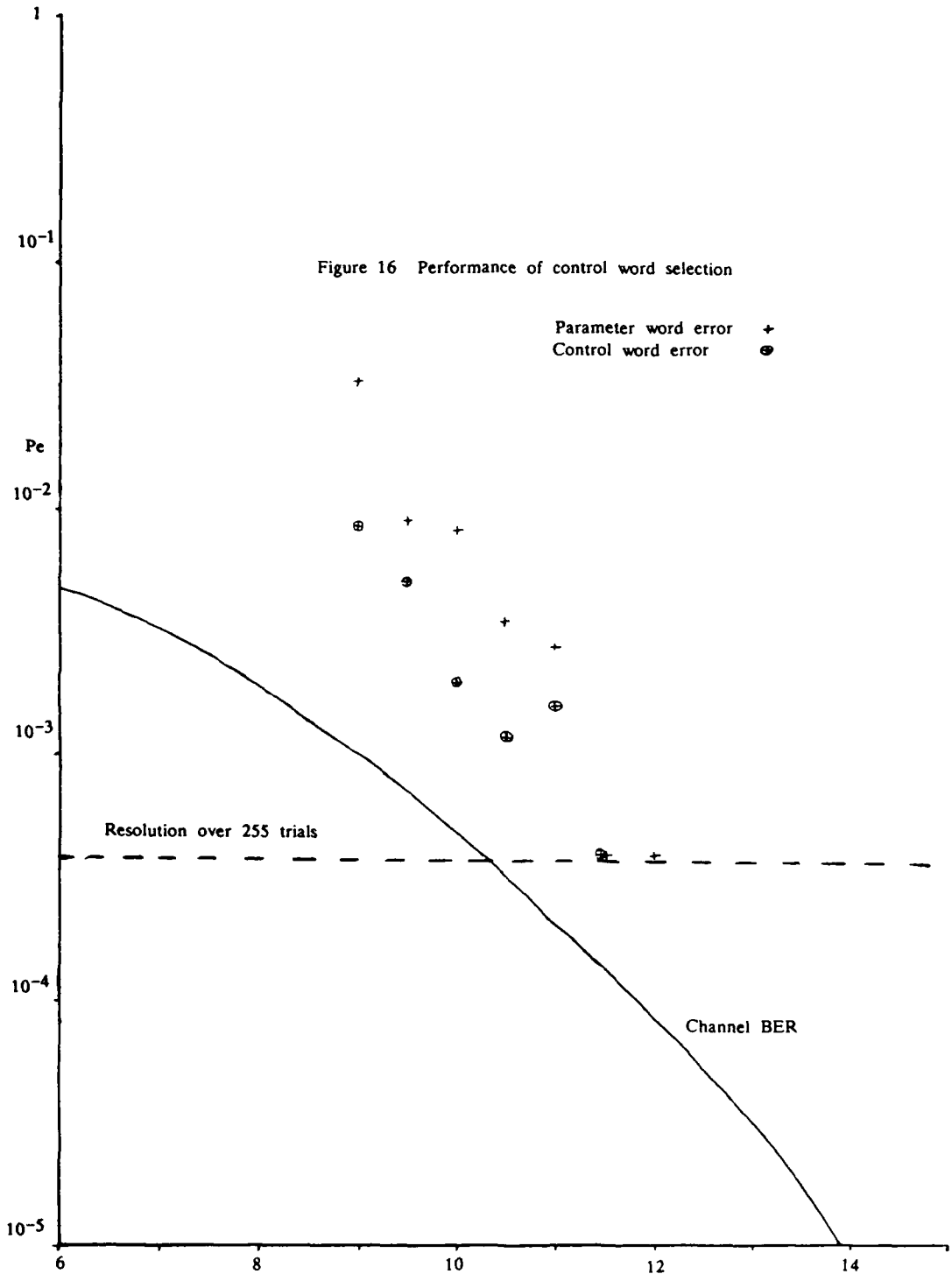


Figure 16 Performance of control word selection

ELECTROMAGNETIC CODE FOR NAVAL APPLICATIONS

by

F. CRESCIMBENI*, F. BESSI**, S. CHITI**

* Istituto per le Telecomunicazioni e l'Elettronica G. Vallauri
Viale Italia 57100 LIVORNO ITALIA

** I.D.S. Ingegneria dei Sistemi S.p.A. Via Roma 50 56100 PISA

SUMMARY

The use of an increasing number of electronic apparatus became vital to meet the high performances required for military navy applications. Thus the number of antennas to be mounted on shipboard greatly increased.

As a consequence of the high antenna density, of the complexity of the shipboard environment and of the powers used for comunication and radar systems, the EMC (Electro-Magnetic Compatibility) problem is playing a leading role in the design of the topside of a ship.

As other Navies, also the Italian Navy has acquired a numerical code for the antenna siting and design. This code, together with experimental data measured at the Italian Navy test range facility, allows for the evaluation of optimal sitings for antenna systems on shipboard, and the prediction of their performances in the actual environment. The structure of this code, named "Programma Elettromagnetico per Applicazioni Navali", PEAN (Electromagnetic Code for Naval Applications) is discussed in this paper, together with its capabilities and applications. Also the results obtained in some examples are presented and compared with the measurements.

1. PEAN DEVELOPMENT

During the last years previsional numerical methods became very important, and the development of more and more advanced computers broadened their applicability. Also in the field of electromagnetic applications and particularly in that of the antenna performance evaluation a large number of research activities has been devoted to the development of numerical techniques for the solution of practical problems.

The Italian Navy has realized the need and the utility of a numerical facility to evaluate the performances of antennas in the actual environment before the siting on ships. To this end the Italian Navy has sponsored co-ordinated research activities at the Istituto per le Telecomunicazioni e l'Elettronica "G. Vallauri", the Ingegneria dei Sistemi I.D.S. S.p.A., and the Department of Electronic Engineering of the University of Florence. Their joint effort lead to the development of the PEAN code. A development plan [1] was deviced so that the final goal has been reached through the achievement of partial aims. During the development of PEAN routines, the numerical results have been continuously compared with experimental data. The measurements, carried out at the antenna pattern test range facility of the Italian Navy, have played a key role in checking every new developed routine.

The development of the whole PEAN code has been completed in about three years, this achievement has been possible also thanks to the US Navy [2,3] that delivered to the Italian Navy some of their numerical codes, useful for the development of PEAN.

2. PROGRAM PEAN DESCRIPTION

Program PEAN is composed of many numerical routines which apply modern electromagnetic techniques to the solutions of EMC problems on shipboard. So the code has the following feature:

1. capability of analysis at all the frequency used on shipboards (0.5 MHz-40 GHz);
2. capability to solve different problems in the naval EMC field (radiation patterns, couplings, hazard);
3. greater accuracy when it is required to meet accurate verifications in critical cases of installation;
4. conformity to the requirements of integrated applications to the naval paltform system design.

The first three goals have been reached developing a modular structure of the code; such structure allows the users to apply the program to a wide field of applications.

Moreover it makes it possible to modify, to change and to add routines to the code so that PEAN can remain at the state of art about electromagnetic methods.

A set of routines of PEAN allow for the user to process and display the input and output data; so it is also possible to obtain an efficient representation of the results.

PEAN can evaluate (fig. 1):

1. the efficiency and radiation patterns of antennas mounted on ships;
2. the coupling between antennas on ships;
3. near fields, hazard regions;

these results are evaluated on the base of geometrical and electrical data of:

1. the antennas,
2. the superstructures of the ship,
3. the layout of antennas.

Moreover the program PEAN can evaluate the optical coverage of weapons, to obtain a complete evaluation of the topside arrangement (fig. 1).

The program PEAN contains also a set of routines for specific simulation of tipical communication antennas, as such as [4-8]:

1. whip and twin whip,
2. wire and fan antenna,
3. dipole, yagi-uda, discone, ground plane, log-periodic dipole array, dipole array,
4. reflector and aperture antenna.

It is obviously possible to model antennas of other kinds, but not in an automatic way.

About the modeling of superstructures, PEAN code contains various modeling techniques, according to the frequency of interest. At low frequency wire-grid and patches models are used [9,10]; while at high frequency many-sided flat plates and elliptical cylinders 10 are used.

The electromagnetic methods employed in the PEAN code are the following.

1. Method of Moments (MM).
   This method is used to simulate antennas and structures, when their dimensions are smaller than or comparable to a wavelenght. Usually, for naval applications, MM is used for HF systems. The formulations of MM used in the PEAN are the following:
   a) for wire grid model
      . point matching with sine, cosine and constant expansions functions [11],
      . Galerkin method with piece-wise sinusoid expansions functions [12],
   b) for patches model
      . point matching with pulse expansions functions [11],
      . Galerkin method with piece-wise sinusoid expansions functions [13].
2. Physical Optics (PO), Geometrical Optics (GO), Geometrical Theory of Diffraction (GTD).
   These methods are employed to analyze antennas and structures when their dimensions are large compared with a wave lenght. The Uniform formulation of GTD has been used (UTD) [14,15]. The leading terms of the interactions among antennas and structures (direct fields, reflected fields, single edge and convex surface diffracted fields ...) are included in the PEAN code as well as higher order terms such as doubly diffracted fields, plate-cylinder interactions, etc. The capability of evaluating higher order interactions allows the user to analyze very critical problems, such as coupling between antennas shadowed by a mast, and provides an improvement with respect to other available GTD codes.
3. Hybrid Techniques (HT).
   HT make use of a suitable combination of MM and GTD [16,17] and are applied at VHF and UHF frequencies. Usually HT applications are based on MM formulation for antenna analysis, and GTD evaluations for interactions with the obstacles.

3. PEAN APPLICATION

The use of PEAN code for the evaluation of antenna performance in real environment requires a deep experience both from a theoretical (electromagnetic methods) and a technical point of view. For this reason the development of (numerical) geometrical models of the environment (structures) involved in the radiative phenomena becomes very important. It should be noted that when the geometrical model of the whole ship has been completed and stored on the computer, PEAN is able to extract numerical models of portions of the ship, suitable for each antenna performance evaluation in its frequency range. Moreover the development of numerical models often needs checks and comparisons with experimental results to assure the reliability of the model.

For a general evaluation of antenna performances, the program use can be split into

three main phases (fig. 1):

1. development of the geometrical model for antennas and structures; definition of both electrical and geometrical input data;
2. execution of electromagnetic evaluations; results written on data file;
3. processing of results and display on graphic or alphanumeric devices.

A set of numerical routines has been also developed to generate the input data and to process the output data.

The development of geometrical models is performed by a modeling code which can fastly accomplish complex models, because of two important reasons:

1. automation in the definition of complex structure by means of a digitizer which "reads" drawings or by means of merge of simple canonical objects;
2. capability of debugging of the model by performing real time checks on graphic display.

The definition of input electrical data is performed by means of routines that enable the user to store a record of the input (data-base) and to check the validity of the input data before starting with the evaluations.

The routines for electromagnetic evaluations are subdivided into libraries for each e.m. method. The output data are written on data-files and can be printed or displaied on alphanumeric or graphic devices. Moreover output data can be processed for other applications, as in the case of near field data to evaluate radiation hazard.

The PEAN code will be resident in a host computer at the Istituto per le Telecomunicazioni e l'Elettronica G. Vallauri (Mariteleradar) of the Italian Navy (fig. 2). Also the following items will be installed on the computer:

1. a data base for antennas and devices,
2. a data base for geometrical models of actual ships;
3. a data base for electronic systems on shipboard.

Moreover the computer will be linked to a graphic station, to terminals at Mariteleradar and to other external users, such as Navy Minister, ship yards and so on. So they will be able to read data and results on data base in real time. A link to the antenna pattern test range facility is also scheduler, to include measurement data in the data base.

4. NUMERICAL EXAMPLES

PEAN has been successfully used to solve many practical problems. Up today PEAN has been applied to studies on:

- six Italian Navy ships,
- Saclant's research vessel (Alliance),
- helicopters and aircrafts.

The PEAN application fields are as follows:

a. design of HF communication antenna layout,
b. design of radar antenna layout,
c. evaluation of radar and communication antenna performances in the actual environment,
d. evaluation of hazard regions,
e. design of wide band antennas,
f. evaluation of radiation inside rooms (towards windows, etc.),
g. evaluation of current and charge distribution on complex structures, due to trasmitting antennas or incident plane waves (frequency domain),
h. evaluation of current and charge distribution excited by impulsive excitations (time domain).

To explain the above, some examples are now provided.

The first example (fig. 3) deals with the evaluation of azimuthal radiation pattern of a whip HF antenna mounted near a mast. In this case the Method of Moments (MM) has been applied to a numerical wire grid model of the antenna and the structure.

The numerical results (fig. 3) have been compared with measurements accomplished at the Italian Navy test range facility on a scale brass model. Numerical evaluations and physical measurements have also been performed out of HF band, to check for the validity of the numerical model out of the frequency range. The results show good performances up to 40 MHz (fig. 3).

The second example deals with the design of a wide band HF fan antenna [18]. In this case it is important to achieve VSWR values below 3 on the entire frequency band of use. For this reason we have to control the input impedance of the antenna during the design process. An example of this study, developed with MM, is shown in fig. 4.

The third example shows a typical application of the PEAN in the case of near e.m.

field evaluations; these analyses are very useful to solve problems concerning hazard radiation or compatibility with weapons or devices on shipboard. In this case the MM has been used too. Some results are shown in fig. 5; the field has been evaluated (on a numerical model) in the region (a) of fig. 5a, under the fan antenna, with a transmitting power equal to 1 Kw, such as many communication systems on actual shipboard.
In fig. 5a an example of numerical preliminary model is shown; fig. 5b represents the diagrams of electric field components on a plane in region (a) and fig. 5c provides comparisons between numerical results and data measured on actual ship at 4 MHz.

The evaluation of microwave antenna performances in a complex environment is another important PEAN application field. An example is shown in fig. 6. The problem deals with the evaluation of the radiation pattern of a search radar slot array antenna at 10 GHz, mounted in the presence of a reflector of another radar antenna (the obstacle). A numerical model of the slot array has been developed using a set of current elements.

The radiation pattern in the presence of the obstacle has been carried out by means of Geometrical Theory of Diffraction. The numerical results have been compared with measurements on actual antennas (fig. 6); this comparison show the accuracy of the method.

The last example presented deals with the application to ElectroMagnetic Pulse (EMP) problems. It is known that the problem of system resistance (aircraft, ships, etc.) to EMP excitation, can be related to two aspects:

1. "external" response; i.e. the current and charge distribution on the external surface of the system;
2. "internal" response; currents and charges induced on internal cables, which also depends on the direct (external) response via the geometrical configuration of the object and the displacement of openings on it.

Concerning the first aspect, the PEAN can be used [19] to evaluate current and charge distributions on the system in the EMP frequency band. The results are processed (fig. 7b) first by weighting the frequency samples with the EMP spectrum and next by applying an algorithm of inverse Fast Fourier Trasform (FFT) to obtain time domain response.

For example the current diagram excited on the tail wing of an aircraft is shown in fig. 7c. This analysis, carried out on the whole surface of the aircraft, enables to:

a. evaluate the "critical surface regions" on which there are the peaks or RMS maximum value of current and of charge;
b. evaluate the stress amplitude;
c. evaluate the stress time extent.

5. FUTURE DEVELOPMENTS

At present the PEAN code is able to carry out most of the evaluations required for the practical treatment of EMC problems on ship-board within the limits of state of art of the electromagnetic methods. In addition, it had been developed with a particular emphasis to naval engineering problems. Also, some topics have been identified in order to improve the code accuracy and efficiency.

The topics, that have currently being considered, are:

1. the numerical optimization, to reduce the execution time and computer storage when large volumes of data are involved as in the case of Method of Moments matrices (1000 x 1000 complex elements);
2. the development of electromagnetic algorithms for solving problems involving structures (plane or curved) covered or composed of non-conducting materials;
3. the development of a direct interface between the PEAN code and CAD facilities of the ship yards.

6. REFERENCES

[1] F. BESSI: "Analisi critica della teoria e dei metodi numerici di simulazione per lo studio delle prestazioni delle antenne per TLC installate in ambienti densi di ostacoli", Tech. Report IDS SpA n. 83/029, Pisa 1983, prepared under contract n. 2225 for It. Navy.

[2] G.J. BURKE, A.J. POGGIO: "Numerical Electromagnetics Code (NEC) - Method of Moments", Lawrence Livermore Lab., NOSC TD116, Jan 1981.

[3] R.J. MARHEFKA, W.D. BURNSIDE: "NEC-Basic Scattering Code V.2", Tech. Report ESL 712242-14/15, OSU 1982.

[4] F. BESSI, S. CHITI: "Il Metodo dei Momenti per lo studio di antenne TLC, limiti applicativi e tecniche di superamento", Tech. Report IDS SpA n. 84/058, Pisa 1984,

prepared under contract n. 13397 for It. Navy.

[5] F. BESSI, A. NALDINI: "Studio di Compatibilità e.m. per antenne LPDA (log-periodic dipole array) a bordo di navi", Tech. Report IDS SpA n. 82/026, Pisa 1982.

[6] M. BANDINELLI, S. CHITI: "Analisi numerica di antenne LPDA per GE-TLC", Tech. Report IDS SpA n. 86/058, Pisa 1986.

[7] P.E. LAW: "Shipboard Antennas", Washington, Artech House, 1983.

8] W.L. STUTZMAN, G.A. THIELE: "Antenna Theory and Design", New York, John Wiley and sons, 1981.

[9] R.F. HARRINGTON: "Field Computation by Moment Methods", New York, Mac Millan Co, 1968.

[10] R. MITTRA (Ed.): "Numerical and Asymptotic Techniques in Electromagnetics", New York, Springer Verlag, 1975.

[11] A.J. POGGIO, E. MILLER: "Integral Equation Solutions of Three Dimensional Scattering Problems", in "Computer Techniques for Electromagnetics", Pergamon Press, 1973.

[12] J.H. RICHMOND: "Radiation and Scattering by Thin Wire Structures in the Complex Frequency Domain", OSU, ELS Report 1973.

[13] E.H. NEWMAN: "Electromagnetic Surface Patch (ESP) Code V.2", OSU-ESL, Tech. Report 717067-4, 1985.

[14] R.G. KOUYOUMJIAN, P.H. PATHAK: "A Uniform Geometrical Theory of Diffraction for an Edge in a Perfectly Conducting Surface", Proc. of the IEEE, Nov. 1974.

[15] R.G. KOUYOUMJIAN, P.H. PATHAK, W.D. BURNSIDE: "The Uniform Geometrical Theory of Diffraction and its Applications to E.M. Radiation and Scattering" in "Acoustic, E.M. and Elastic Wave Scattering - Focus on the T-matrix Approach", V.K. Varadan and V.V. Varadan ed.s, Pergamon Press, New York, 1980.

[16] G.A. THIELE, T.H. NEWHOUSE: "A Hybrid Technique for the Combining Moment Methods with the Geometrical Theory of Diffraction", IEEE Trans. on AP, Jan. 1975.

[17] W.D. BURNSIDE, C.L. YU, R.J. MARHEFKA: "A Technique to combine the Geometrical Theory of Diffraction and the Moment Method", IEEE Trans. on AP, July 1975.

[18] S. CHITI, L. GIOVANNONI: "Antenne filari e fan: tecniche di simulazione numerica e confronti con misure sperimentali", Pisa 1985, Tech. Rep. IDS SpA n. 85/030 prepared under contract n. 13970 for It. Navy.

[19] S. CHITI: "Applicazione del Metodo dei Momenti a studi EMP su oggetti complessi - Analisi in frequenza", Pisa 1986, Tech. Rep. IDS SpA n. 86/096 (interno).

7. ACKNOWLEDGEMENT

The authors wish to express their appreciations to prof. R. Tiberio for useful discussions and advices.

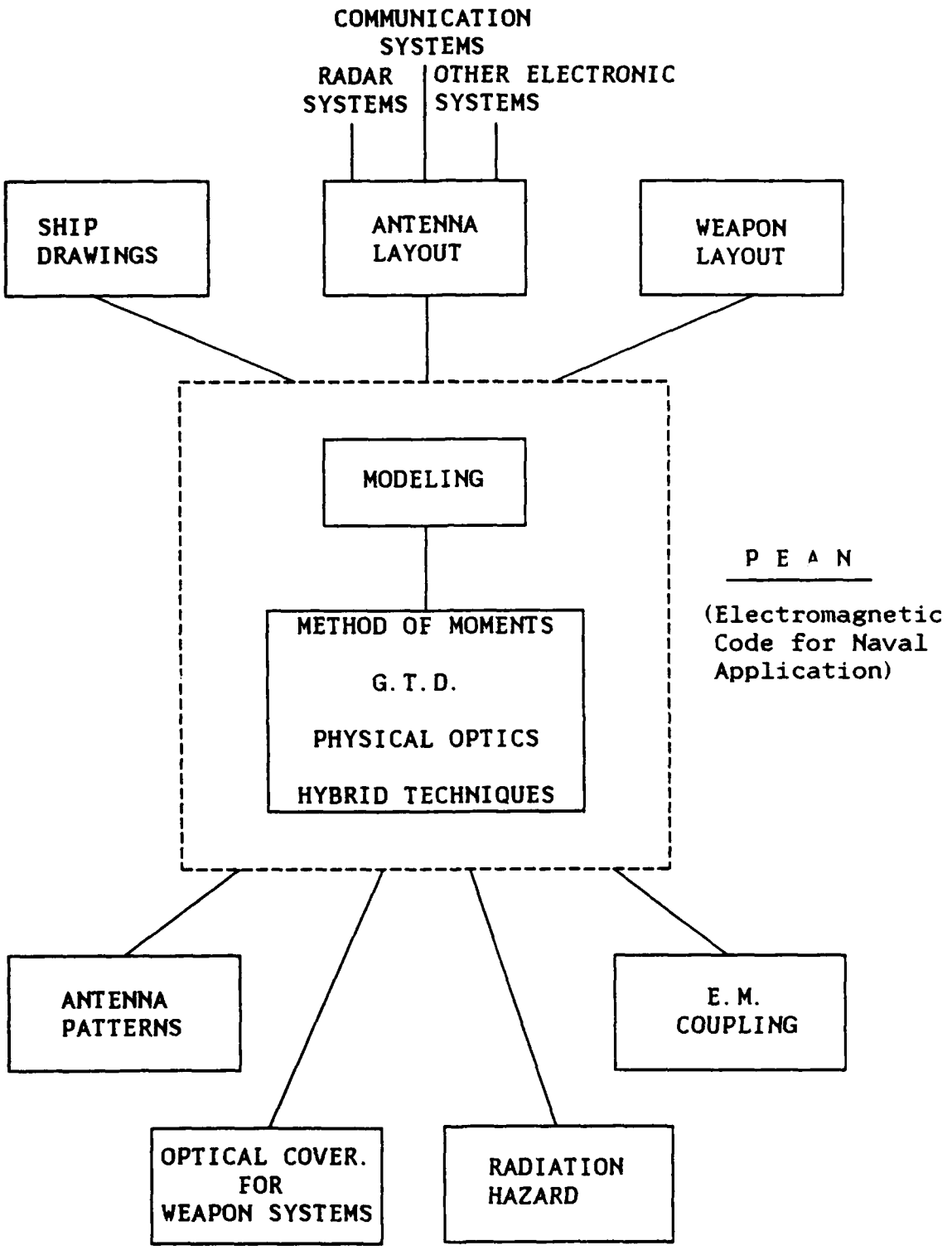


FIG. 1. I/O OF ELECTROMAGNETIC CODE FOR NAVAL APPLICATION

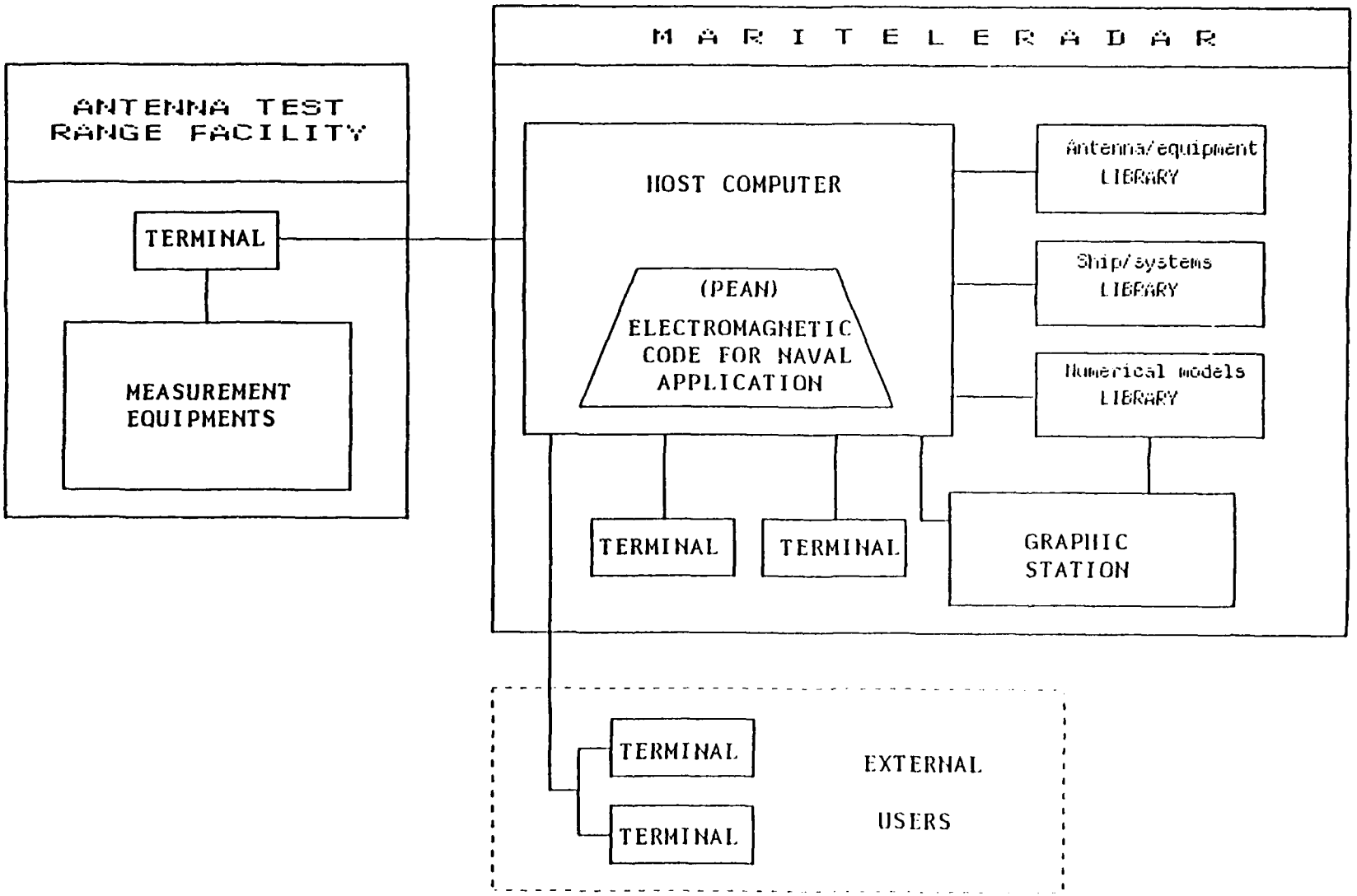
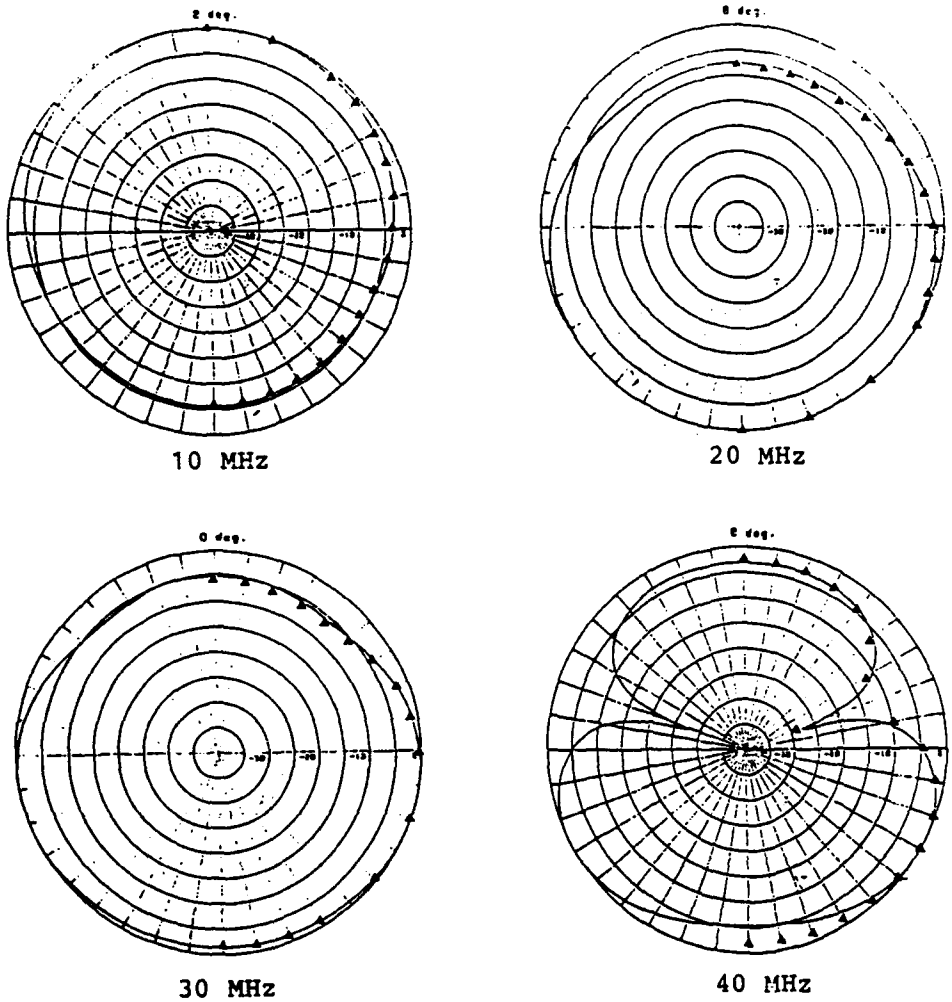


FIG. 2. E.M. CODE UTILITY



▲▲▲▲ measured at the antenna radiation pattern facility of Italian Navy (Mariteleradar)

——— Method of Moments

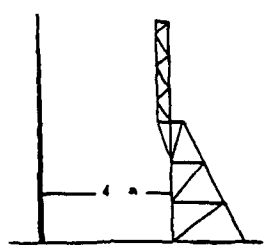


FIG. 3. RADIATION PATTERNS OF HF ANTENNA MOUNTED NEAR A MAST

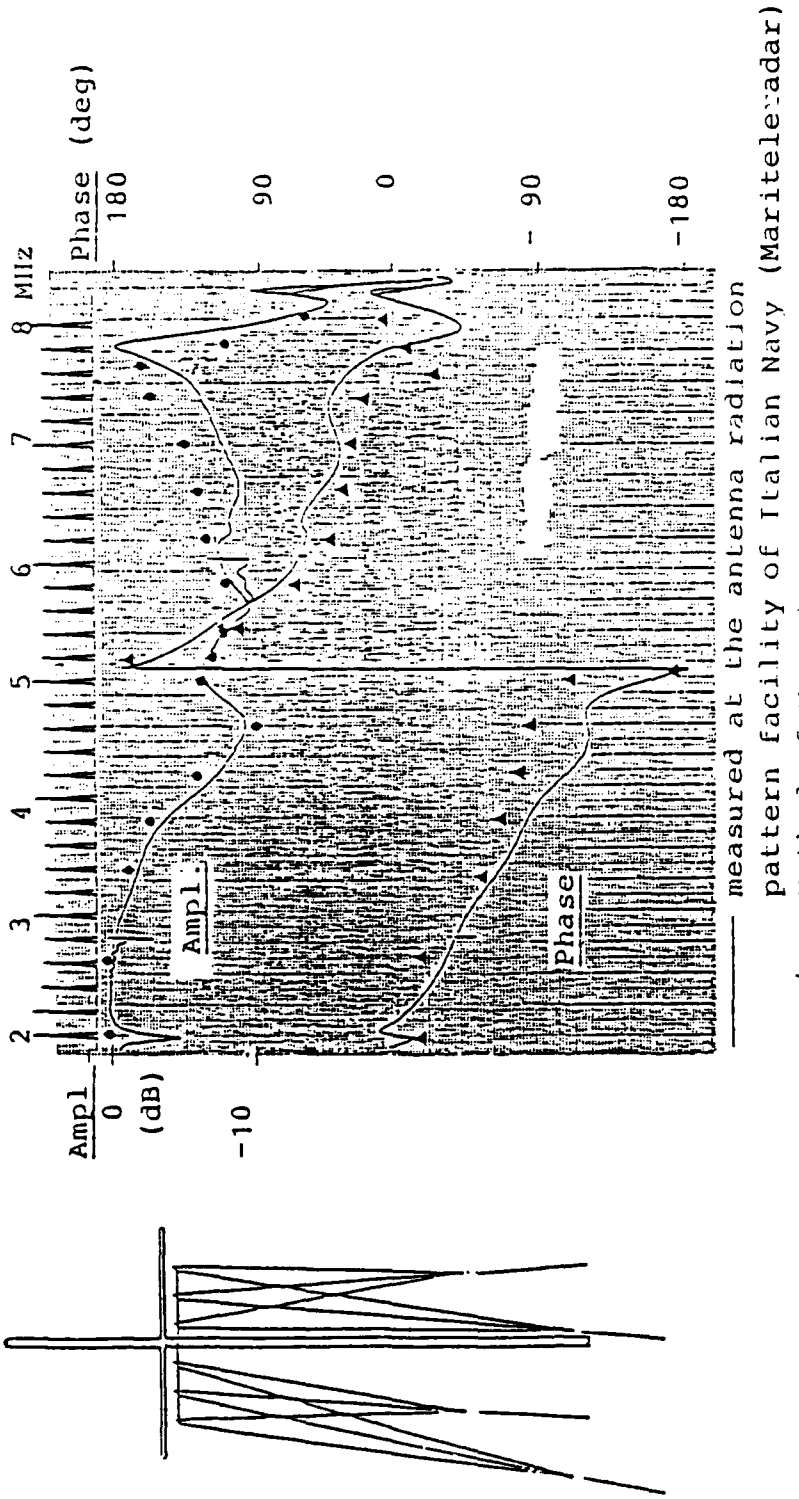
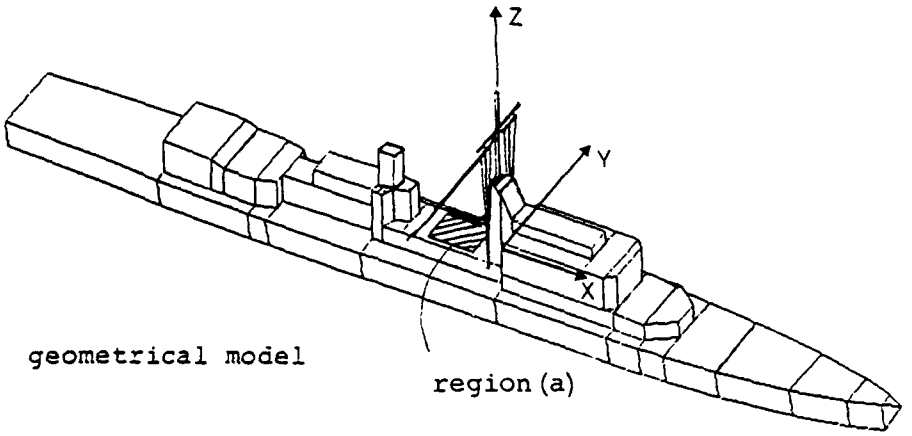
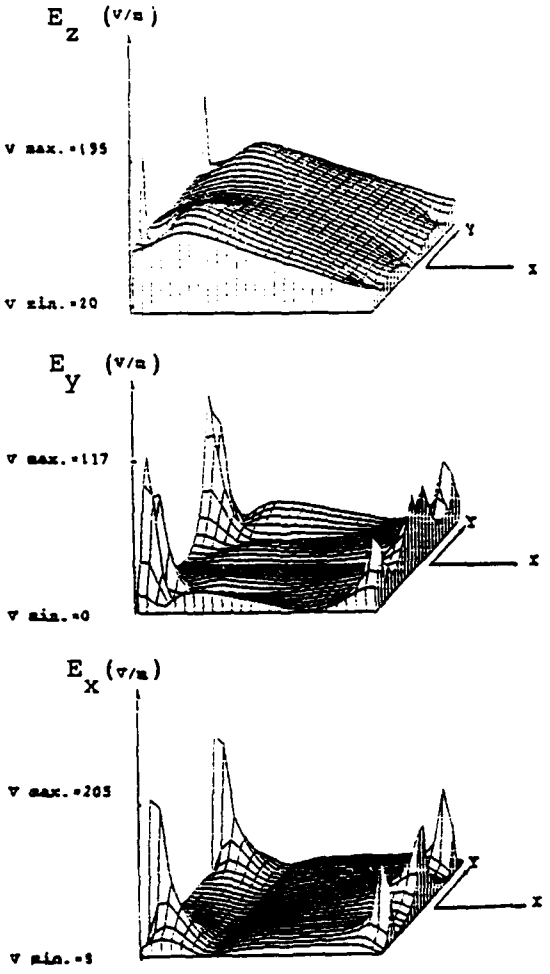


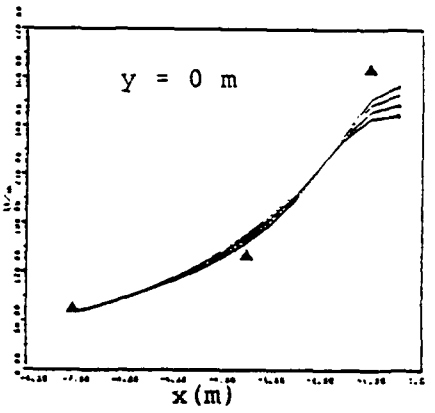
FIG. 4. RETURN LOSS OF FAN ANTENNA ON GROUND PLANE

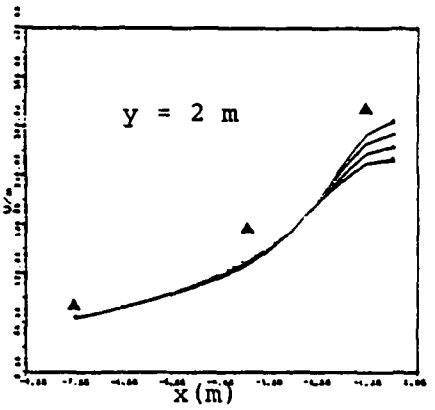


a) geometrical model



b) Electric field components computed on region (a)



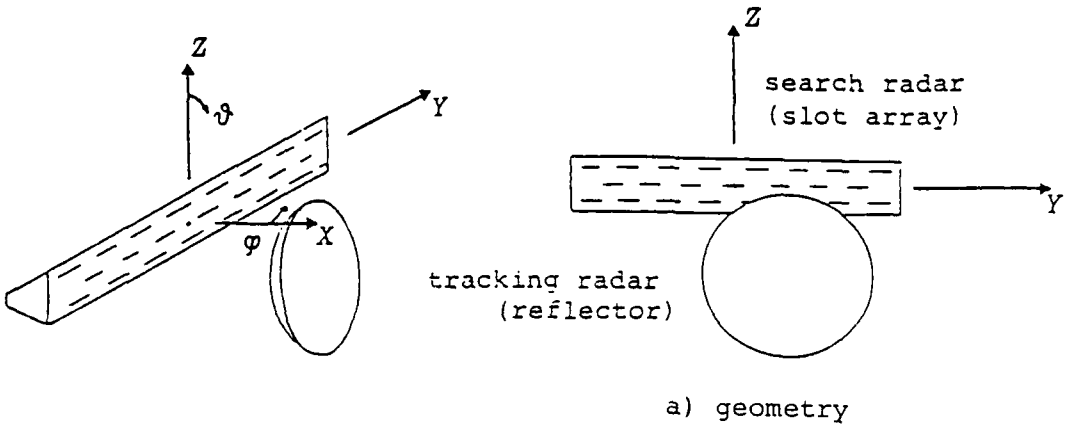


c) Comparison between computed and measured data on shipboard in region (a)

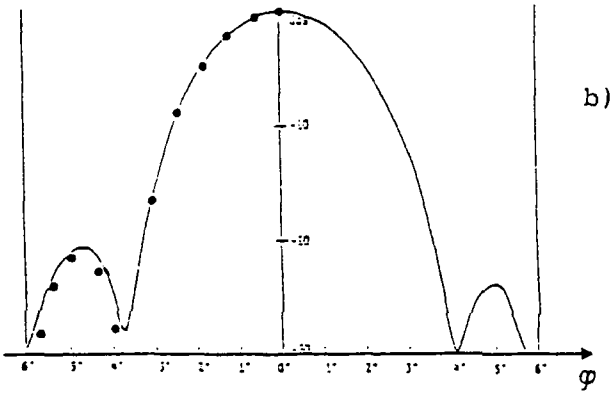
▲▲▲ measured

_____ computed along lines near measure points

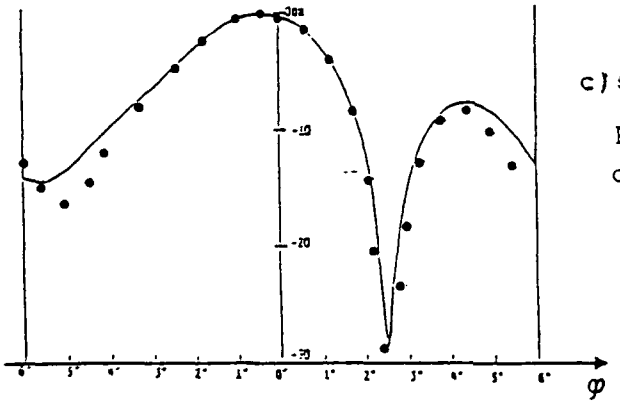
FIG. 5. NEAR FIELD RADIATION ON SHIPBOARD



a) geometry



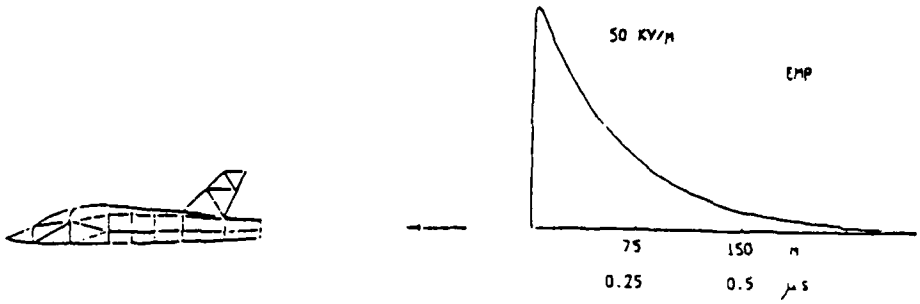
b) slot array:
free space pattern
ϑ = 0 deg.



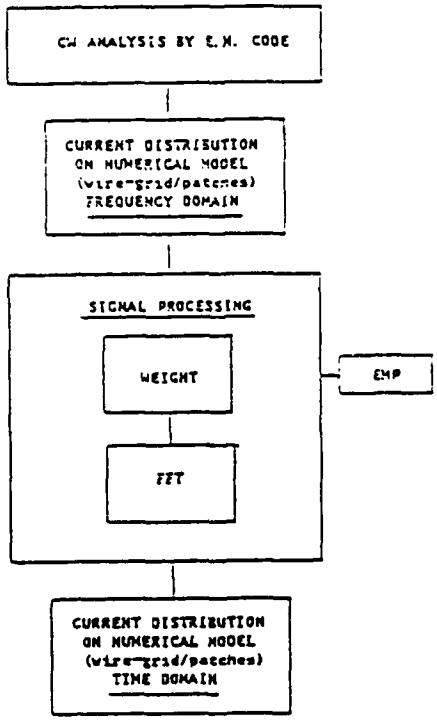
c) slot array:
pattern in presence
of radar reflector
ϑ = + 106 deg.

——— measured • • • • GTD calculations

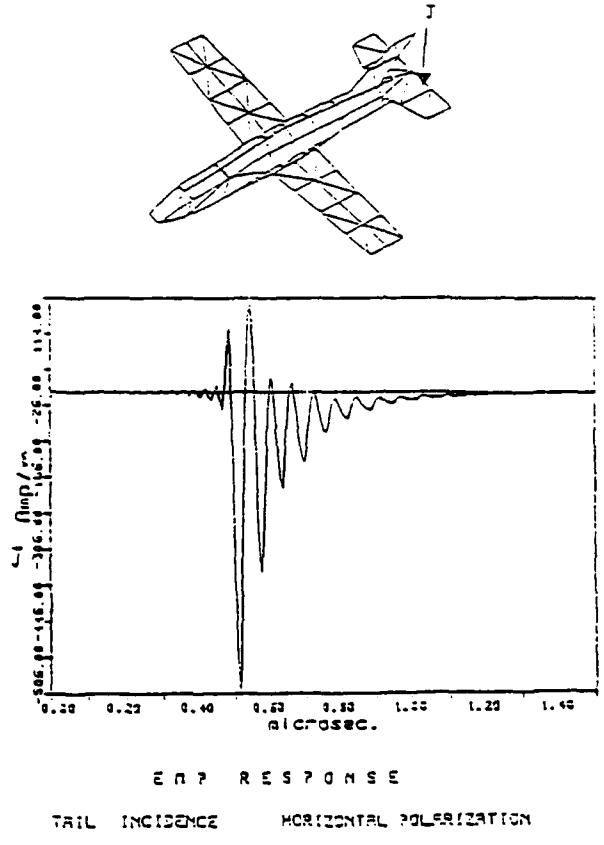
FIG. 6. E-PLANE RADIATION PATTERN OF SEARCH RADAR ANTENNA



a) EMP problem (diff. scale)



b) Flow chart for EMP analysis



c) current density on tail wing

FIG. 7. EXAMPLE OF EMP ANALYSIS: CURRENT INDUCED ON AIRCRAFT MODEL

NARRATIVE COMPRESSION CODING FOR A CHANNEL WITH ERRORS

By Dr. James W. Bond, Staff Scientist
NAVOCEANSYSCEN, Code 83
271 Catalina Boulevard
San Diego, California 92152-5000

ABSTRACT

Data compression codes offer the possibility of improving the thruput of existing communication systems in the near term. This study was undertaken to determine if data compression codes could be utilized to provide message compression in a channel with up to a .10 bit error rate.

The data compression capabilities of codes were investigated by estimating the average number of bits-per-character required to transmit narrative files. The performance of the codes in a channel with errors (a noisy channel) was investigated in terms of the average numbers of characters-decoded-in-error-per-bit-error and of characters-printed-in-error-per-bit-error.

Results were obtained by encoding four narrative files, which were resident on an IBM PC and use a 58 character set. The study focused on Huffman codes and suffix/prefix comma-free codes. Other data compression codes, in particular, block codes and some simple variants of block codes, are briefly discussed to place the study results in context.

Comma-free codes were found to have the most promising data compression because error propagation due to bit errors are limited to a few characters for these codes. A technique was found to identify a suffix/prefix comma-free code giving nearly the same data compression as a Huffman code with much less error propagation than the Huffman codes. Greater data compression can be achieved through the use of this comma-free code with code word assignments based on conditioned probabilities of character occurrence.

INTRODUCTION

Data compression encoding offers an option for increasing the channel capacity of existing communications systems by efficiently encoding the narrative portions of messages. A data compression code assigns short binary code words to characters with a high frequency of occurrence and long code words to characters with a low frequency of occurrence. Difficulties arise when data compression codes are used in noisy channels because one bit error can lead to multiple character errors due to temporary loss of character synchronization.

This study focused on the investigation of Huffman and comma-free data compression codes which could be used to encode characters based on their probabilities of occurrence. The comma-free code results were then extended to encoding of characters based on their conditional probabilities of occurrence. In addition, several coding approaches using block codes and simple variants of block codes are discussed.

The data compression provided by a code is measured by the average number of bits-per-character of the encoded narratives files; the performance of the code in noisy channels is measured by the average number of characters-decoded-in-error-per-bit-error and the average number of characters-printed-in-error-per-bit-error. Generally speaking, as the number of bits-per-character decreases (that is, as data compression increases), the numbers of characters-decoded and printed-in-error-per-bit-error increase. Observe that under the assumptions of a fixed bit-error-rate and random bit errors, the ratio of the average number of character (decoded or printed) errors for two codes is equal to the ratio of the product of the average number of bits-per-character and the average number of characters (decoded or printed)-in-error-per-bit-error for the two codes.

Results are presented for a 58 character alphabet derived from the 95 character set of the personal computer and for processing narrative files stored on its hard disk. These files were edited to use only capital letters and certain seldom used symbols were deleted to obtain an alphabet emulating the Military Standard of the American Variation of the International Telegraph Alphabet No. 2 (hereafter called the Military Baudot Code) in use for Navy communications.

The error properties of both Huffman and comma-free codes depend on the specific choices of bits and code words, respectively, used to construct the codes. A main thrust of this paper is to identify the Huffman codes and the comma-free codes giving the lowest average number of character (decoded or printed)-errors-per-bit-error for a given compression gain.

APPROACH

Huffman codes are known to provide the best data compression possible for variable length coding of individual characters [reference 1]. This property is ensured by the code construction process because it is based on the probabilities of occurrences of the characters to be encoded. We began our investigation by establishing the properties of Huffman codes in noisy channels using character encoding.

The comma-free codes analyzed in this report are constructed using a sequential procedure found by R. A. Scholtz [references 2 and 3]. His procedure does not utilize probabilities of occurrence to guide the construction process. We developed a way to most nearly match the word lengths of a comma-free code to those of an optimum Huffman code in order to maximize the data compression performance of the selected comma-free code.

Even after specifying the distribution of word lengths of Huffman and comma-free codes, there are degrees of freedom in the construction processes. It was discovered that the error properties of a code depended on the use made of these degrees of freedom.

The insights provided by the investigation of Huffman codes and comma-free codes led to the identification of certain natural extensions of the presently used Military Baudot code. A comparison of the performance of these codes with those of Huffman and comma-free codes provides an additional performance gauge against which the latter codes can be assessed.

The Huffman code construction process has a great number of degrees of freedom. The impact of bit errors on character synchronization and character errors is very context-dependent; therefore, an analytical study of the dependency of error statistics on the Huffman construction process could not be performed. Therefore a simulation program was written and exercised to search for the best Huffman codes.

The number of degrees of freedom in the comma-free code construction process depends on the number of sequential steps in the process and not on the character set size. The performance of codes constructed in a few steps, which are the codes of most interest, are established analytically by an exhaustive treatment of the available codes.

Huffman and comma-free coding could also be applied to character encoding based on conditional probabilities of character occurrence. We obtained results for comma-free codes, which are placed in perspective by considering block coding of the words in large dictionaries.

DATABASE

Four narrative files were used to analyze the performance of data compression codes. These files (labeled I, II, III, and IV) contain 31,744, 28,672, 12,288, and 13,312 bytes, respectively, and are resident on the hard disk of the personal computer. All of the narrative files were technical documents involving some equations.

Table 1 presents the probabilities of occurrence for the different characters for each of the four narrative files. Note that the probabilities of occurrence of the characters are similar for the four narrative files.

BLOCK CODES AND GENERALIZED BAUDOT CODES

This section describes block codes and a family of compression codes which have a structure very similar to that of the presently used Military Baudot code.

Block Codes

A code consisting of code words of equal length is called a block code. The length of the block code used depends on the number of different symbols being encoded. A 58 character set is used by the United States Navy. Six-bit code words are necessary to encode this character set.

If words are encoded, then the dictionary size determines the length of the code words necessary to encode the words: 14 bits are necessary to encode a 16,384 word dictionary; 15 bits to encode a 32,768 word dictionary; and 16 bits to encode a 65,536 word dictionary. (Note for an average word length of five characters, the dictionary schemes use 2.33, 2.50, and 2.67 bits per character encoding, respectively, by not requiring the encoding of spaces.) Larger dictionaries appear to be prohibited by the difficulty of implementing encoding and decoding by table look-up operations. One final note: if block codes are used, a bit error leads to a single character error if characters are encoded or a single word error if words are encoded.

Since the English language is known to contain over 300,000 words some provision must be made to handle words not in the dictionary. It is undesirable to limit the vocabulary of the writer of Navy messages so that words occur which are not in the dictionary. In order to take the most advantage of a fixed-size dictionary some way should be found to handle variants of the same word, such as alternate spellings, misspellings, and abbreviations. We do not analyze block encoding of large dictionaries in this paper. Rather we chose to analyze the use of data compression coding of characters based on conditional probabilities of character occurrence.

Generalized Baudot Codes

The Military Baudot code consists of information carrying characters and two shift characters, which change the decoding of the next code word. The Military Baudot alphabet consists of 56 information characters and shift characters. One shift character shifts letters to figures and the other from figures back to numbers. If a simple block code was

TABLE 1. CHARACTER PROBABILITIES OF OCCURRENCE FOR FOUR NARRATIVE FILES

CHARACTER	NARRATIVE FILE I	NARRATIVE FILE II	NARRATIVE FILE III	NARRATIVE FILE IV
" "	0.3085	0.3171	0.2851	0.4095
E	0.0855	0.0865	0.0882	0.0664
T	0.0636	0.0677	0.0585	0.0492
N	0.0543	0.0537	0.0483	0.0407
O	0.0416	0.0518	0.0505	0.0420
I	0.0513	0.0511	0.0537	0.0426
A	0.0455	0.0450	0.0499	0.0434
R	0.0392	0.0421	0.0471	0.0438
S	0.0372	0.0391	0.0440	0.0367
H	0.0277	0.0263	0.0291	0.0162
C	0.0193	0.0232	0.0245	0.0216
L	0.0233	0.0218	0.0262	0.0226
D	0.0191	0.0206	0.0327	0.0165
U	0.0142	0.0185	0.0147	0.0165
P	0.0163	0.0170	0.0171	0.0164
M	0.0120	0.0162	0.0230	0.0150
F	0.0151	0.0141	0.0190	0.0169
G	0.0117	0.0121	0.0163	0.0109
B	0.0049	0.0101	0.0017	0.0067
V	0.0136	0.0099	0.0078	0.0077
W	0.0147	0.0082	0.0080	0.0030
.	0.0073	0.0076	0.0075	0.0087
Y	0.0050	0.0058	0.0074	0.0026
,	0.0062	0.0001	0.0036	0.0060
)	0.0016	0.0039	0.0024	0.0024
(	0.0016	0.0039	0.0024	0.0024
-	0.0047	0.0034	0.0044	0.0053
1	0.0078	0.0024	0.0015	0.0054
K	0.0040	0.0023	0.0014	0.0008
/	0.0003	0.0016	0.0003	0.0019
J	0.0016	0.0014	0.0011	0.0000
X	0.0018	0.0013	0.0017	0.0008
2	0.0059	0.0013	0.0016	0.0031
+	0.0031	0.0011	0.0000	0.0000
=	0.0019	0.0010	0.0002	0.0000
Z	0.0002	0.0009	0.0011	0.0004
Q	0.0014	0.0008	0.0012	0.0004
3	0.0015	0.0007	0.0010	0.0018
0	0.0009	0.0005	0.0014	0.0050
~	0.0000	0.0005	0.0000	0.0000
"	0.0003	0.0003	0.0003	0.0000
4	0.0000	0.0003	0.0002	0.0001
^	0.0014	0.0002	0.0000	0.0000
:	0.0006	0.0002	0.0008	0.0004
8	0.0001	0.0002	0.0005	0.0015
@	0.0000	0.0002	0.0000	0.0000
5	0.0001	0.0002	0.0005	0.0014
9	0.0000	0.0002	0.0007	0.0034
*	0.0021	0.0002	0.0000	0.0000
;	0.0000	0.0002	0.0006	0.0000
6	0.0000	0.0001	0.0002	0.0015
>	0.0000	0.0001	0.0001	0.0000
7	0.0000	0.0001	0.0001	0.0001
'	0.0000	0.0001	0.0002	0.0004
<	0.0000	0.0000	0.0002	0.0000
!	0.0000	0.0000	0.0001	0.0000
#	0.0018	0.0000	0.0000	0.0000
%	0.0000	0.0000	0.0001	0.0000

NOTE: " " denotes space

used, six bits would be required; if a five-bit code is used instead, and one of the 32 code words is used as a shift character, 31 information characters can be transmitted using five bits and the remaining 25 information characters can be transmitted by using the five-bit code word reserved for a shift character followed by a five bit code word.

We model the Military Baudot code in terms of a code using a single one-character shift character. The Military Baudot code should perform somewhat better than predicted by the model because of the tendency for numbers and characters in the alphabet to occur sequentially and the actual implementation of the shift as a toggle operation.

More than one shift character can be used and these can be used sequentially to provide a whole family of different Baudot-like codes, which we call generalized Baudot codes. A generalized Baudot code is specified by its basic code length and the number of characters used as shift characters for each multiple of the block length. The generalized Baudot codes of most interest for Navy messages use code lengths which are multiples of 3, 4, or 5 (2 allows too few code words to build upon and 6 will block encode 58 characters).

The bits-per-character for the best single shift code is obtained as 5 + 5P, where P is the probability of occurrence of any of the 27 least commonly occurring characters. The average number of bits-per-character turns out to be 5.08, 5.06, 5.06, and 5.06, for narrative files I, II, III, and IV, respectively. For this code, one character is decoded in error per bit error and on the average 1.01 to 1.02 printed characters are in error per bit error, depending on the training file.

Consider a generalized Baudot code using more than one shift symbol. Suppose, in particular, that the shifts were used to produce a code with 15 words of length 4, 15 of length 8, 15 of length 12, and 13 of length 16. One of the first 16 code words is a shift, i.e., leads to a different interpretation of the next code word; one of these code words is reserved to lead to still another interpretation of the next code word; and one of these is reserved to lead to still another interpretation of the next code word. The average number of bits-per-character required to transmit information using this code is 4.68, 4.52, 4.54, and 4.52, for narrative files I, II, III, and IV, respectively. For this code, one character is decoded in error per bit error and on the average 1.13 to 1.17 printed characters are in error per bit error, depending on the training file.

A Baudot code based on length three code words provided about the same compression as one of length four, at the cost of greatly increased complexity.

HUFFMAN CODES

Using only the probabilities of a set of characters being transmitted, Huffman provided an organized technique for constructing efficient codes, i.e., using a minimum number of bits (on the average) to transmit characters. The procedure for constructing a Huffman code is illustrated in the following example drawn from reference 1.

Suppose that five characters, a, b, c, d, and e, with probabilities of occurrence 0.125, 0.0625, 0.25, 0.0625, and 0.5, respectively, are to be encoded (see figure 1).

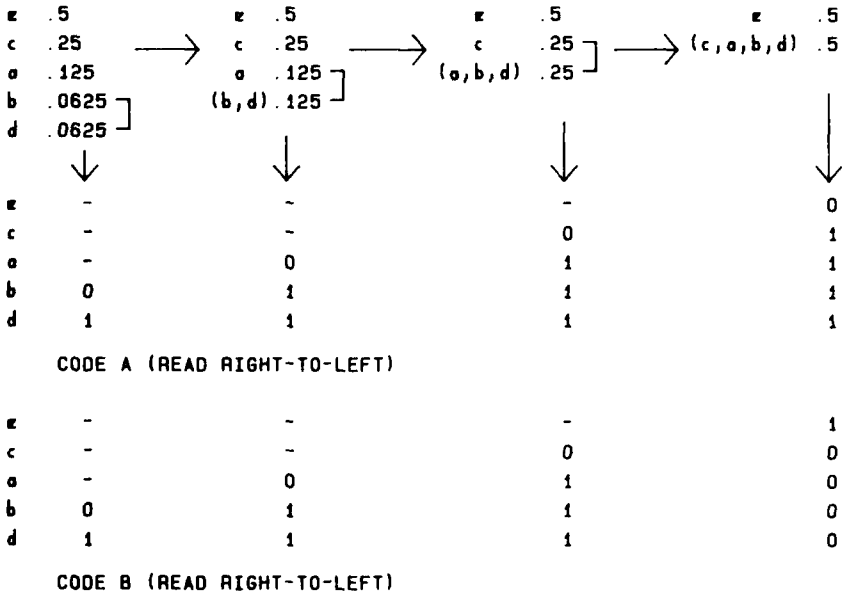


Figure 1. Two Examples of Huffman Coding

For this example, the Huffman procedure involves three regroupings of five characters. Grouped characters are indicated by (b,d), (a,b,d), and (c,a,b,d) along the top of figure 1. At each stage in this step, the two characters or group of characters with the lowest probabilities are grouped and the group is assigned the probability obtained by summing the probabilities of its members. The Huffman code is constructed based on the character groups by proceeding from right to left. Two of the many possible codes which can be assigned to the original character set by the Huffman construction process are illustrated in figure 1.

We discuss the construction of Code A first. Step 1: assign "0" to the most likely character "e" and "1" to the character set (c,a,b,d). These bits are the first bit in the code words assigned to the characters. The character "e" is distinguished from the characters "c", "a", "b", and "d" by the fact that the code for "e" begins with "0" and the others begin with "1". Step 2: no bit is assigned to "e", and a second bit is assigned to the remaining characters. This bit is chosen to distinguish "c" from "a", "b", and "d"; "0" is assigned to "c" and "1" is assigned to the other characters. Step 3: no additional bits are assigned to "e" and "c"; additional bits are assigned to distinguish "a" from "b" and "d". Step 4: no bits are assigned to "e", "c", and "a"; bits are assigned to distinguish "b" and "d".

Code B, also shown in figure 1, differs from code A in that at step 1, the character "e" is assigned "1" and the characters "c", "a", "b", and "d" begin with "0". The remaining steps are the same. Note that "0" and "1" can be assigned in either way at each step, leading to the construction of 16 different codes for the example shown in figure 1.

The example in figure 1 is very regular in that no reordering is necessary during the grouping of characters at the different stages of the construction process. This is not always the case. It is also worthwhile to note that the Huffman coding procedure can lead to block coding when all of the character probabilities are the same. For example, consider the case of eight characters: a,b,c,d,e,f,g, and h, each having a probability of 0.125. The first step leads to grouping g and h, the next step to grouping e and f, the next to grouping c and d, and the fourth step to grouping a and b. Each group is assigned a probability of 0.25. The next two steps lead to grouping e,f,g, and h, and to grouping a,b,c, and d. Each of these groups is assigned a probability of 0.5. It is easy to see that in this case each character is assigned a three-bit code word. In general, the Huffman code construction process for characters with differing probabilities of occurrence leads to a code with some characters having code words of the same length and other characters having code words of differing lengths.

Figure 2 shows the impact of introducing a single bit error into the code word assigned "a" for Code A and Code B. For Huffman codes, and other variable length codes, the impact of an error depends on the characters following "a". In the example, "abcde" is being transmitted. The impact of the single bit error is enclosed by brackets and an error count shown to the right for each of the two Huffman codes. For code A, an error in the first bit of the code word for "a" leads to it being incorrectly decoded into the two characters "e" and "c"; i.e., one character decoded in error and two characters printed in error. For code B, an error in the first bit of the code word for "a" leads to the next three characters being decoded in error for a total of 10 characters printed in error. For code A, the bit error does not lead to loss of character synchronization; while for code B, it does.

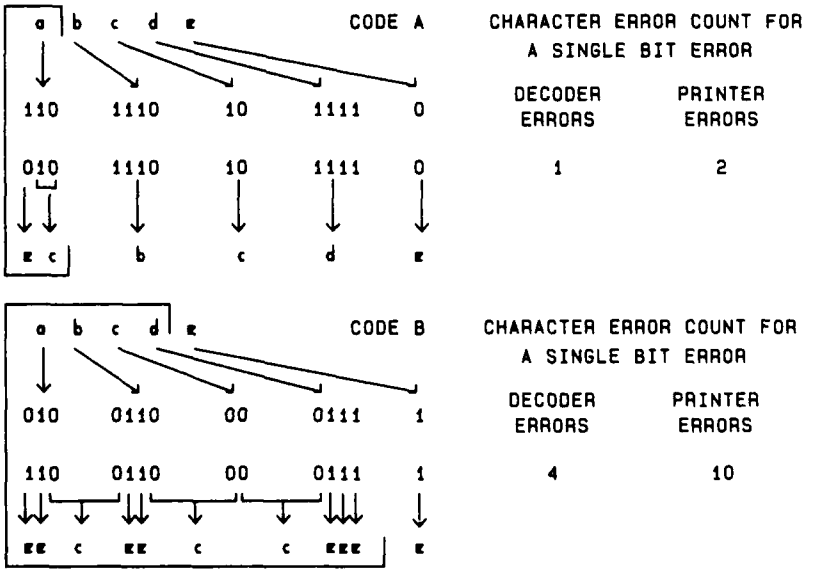


Figure 2. Examples of Error Propagation for Two Huffman Codes Providing the Same Compression

We turn now to the analysis of Huffman encoding of the four narrative files previously described. The structure of a Huffman code in the sense of its distribution of lengths of code words is determined by the probabilities of occurrence of the 58 characters in the encoded narrative file (provided some convention to treat equi-probable sets in the construction process is adopted).

Table 2 summarizes the code words assigned by the particular computer implementation of the Huffman constructed process that we used in our study. (It also contains a column of word lengths for a comma-free code. This column will be discussed later.) The code word lengths for Huffman encoding were obtained using the probabilities of occurrence of the characters presented in table 1 for the four narrative files. The order of the characters is the same in table 2 as that in table 1 and the characters are partitioned into sets of 15 characters to facilitate discussion.

The probability of occurrence of any of the first 15 characters listed in table 2 exceeds .84 for all the narrative files. The word lengths assigned to the first 15 characters based on the probabilities of occurrence of the characters in the different narrative files never differ by more than one bit. The word lengths are nearly the same for the next 15 characters and tend to differ greatly only for the least probable characters. Some of the differences between word lengths presented in table 2 for low probability of occurrence characters could have been lessened by adopting a different convention for equi-probability character sets than that used in our simulations. Nevertheless, use of any of the four narrative files as a training file should lead to similar compression results for Huffman (and later, comma-free) encoding of the narrative files.

Huffman code data compression performance is summarized by the average number of bits-per-character required to transmit the narrative files depending on the particular file (the training file) whose probabilities of character occurrence were used to construct the code. Table 3 summarizes the results of the Huffman code average bits-per-character calculations. Note that using narrative files II and III as training files gave nearly the same results. The maximum difference between two entries of the tables occurred when narrative file IV was used as a training file for narrative file I; however, the difference was only .23 bits-per-character.

TABLE 2. HUFFMAN AND COMMA-FREE CODE WORD LENGTHS FOR FOUR NARRATIVE FILES

CHAR	WORD LENGTHS FOR NARRATIVE FILE I	II	III	IV	COMMA-FREE	CHAR	WORD LENGTHS FOR NARRATIVE FILE I	II	III	IV	COMMA-FREE
" "	2	2	2	1	2	J	9	9	10	22	9
E	3	3	4	4	3	X	9	9	9	11	9
T	4	4	4	4	3	2	10	8	10	9	9
N	4	4	4	5	4	+	10	8	18	15	9
O	4	5	4	5	4	=	10	9	13	21	9
I	4	4	4	5	4	Z	10	12	10	12	9
A	4	4	4	5	5	Q	10	10	10	12	10
R	5	5	4	5	5	3	10	9	10	10	10
S	5	5	5	5	5	0	11	10	9	8	10
H	5	5	5	6	5	~	11	25	19	27	10
C	5	6	5	6	6	"	11	12	11	18	10
L	6	5	5	6	6	4	12	15	12	13	10
D	6	6	5	6	6	^	12	9	19	27	10
U	6	6	6	6	6	:	12	11	10	11	10
P	6	6	6	6	6	8	12	14	11	10	10
M	6	6	6	6	7	@	12	18	16	24	11
F	6	6	6	6	7	5	12	12	11	10	11
G	6	6	6	7	7	9	13	22	10	9	11
B	7	8	6	8	7	*	12	9	17	26	11
V	7	6	7	7	7	;	12	24	11	19	11
W	7	6	7	9	7	6	13	19	13	10	11
.	7	7	7	7	8	>	13	17	13	23	11
Y	7	8	7	9	8	7	14	20	13	14	11
,	7	7	8	8	8	'	15	21	12	11	11
)	8	7	9	9	8	<	16	25	12	20	11
(	8	7	9	9	8	!	17	23	13	17	12
-	8	8	8	8	8	#	18	13	15	25	12
1	9	7	9	8	8	%	18	16	14	16	12
K	9	8	9	11	9						
/	9	11	11	9	9						

TABLE 3. AVERAGE NUMBER OF BITS-PER-CHARACTER FOR HUFFMAN CODES AND DIFFERENT TRAINING FILES

TRAINING FILE	NARRATIVE FILE I	NARRATIVE FILE II	NARRATIVE FILE III	NARRATIVE FILE IV
I	4.04	4.05	4.24	3.77
II	4.12	3.95	4.19	3.73
III	4.12	3.96	4.15	3.72
IV	4.27	4.05	4.32	3.63

The entries in table 3 required to encode the training files are calculated directly as the sum of the probabilities of occurrence of a character with the length of the code word assigned to it by the Huffman construction process. The remaining table entries are obtained by multiplying each character probability of occurrence in the narrative file under consideration by the length of the Huffman code word assigned to that character and summing the results.

A computer program was written to search among the possible Huffman codes, which give the bit-per-character values presented in table 3, for the one performing best in a noisy channel. The original program reads a text file and counts the number of occurrences of each character; from this, a Huffman code is constructed using the construction process first described earlier. In order to search for a good code the specific choices made in the Huffman construction process were randomized. The probabilities of occurrence for each character and the assigned Huffman code words for each character were written to files so that the particular codes could be recovered if desired. Trials were run using the probabilities of occurrence of the characters in each of the four narrative files. The simulations were run by randomly introducing bit errors at a rate of 3 per 1000 bits. No attempt was made to model the impact of burst errors on the channel. It was felt that should burst errors pose a problem in the implementation of a particular code, it would always be possible to superimpose interleaving after data compression encoding and deinterleaving prior to data compression decoding.

We found that, regardless of the narrative file used as a training file, the poorest error performance results were obtained when processing narrative file IV (the most compressible) and the best error performance results were obtained for narrative file III (the least compressible). Figure 3 illustrates the dependency of the results upon the code selection process by presenting the average numbers of characters-decoded-in-error-per-bit-error for experiments run using narrative file II as a training file.

A best and worst code for each narrative file as a training file was selected for further analysis. Figure 4 presents the distributions of lengths of successive printed characters in error obtained for the best and worst of the eight cases analyzed. Note that any of the printed character error sequences may involve more than one bit error. However, the likelihood of two bit-error induced error sequences merging is very small for a bit error rate of 3 in 1000 and can be neglected; therefore, printed output character performance is summarized in terms of the average number of printed character errors per bit error, which we estimated by dividing the total number of character errors by the total number of bit errors introduced during a simulation run.

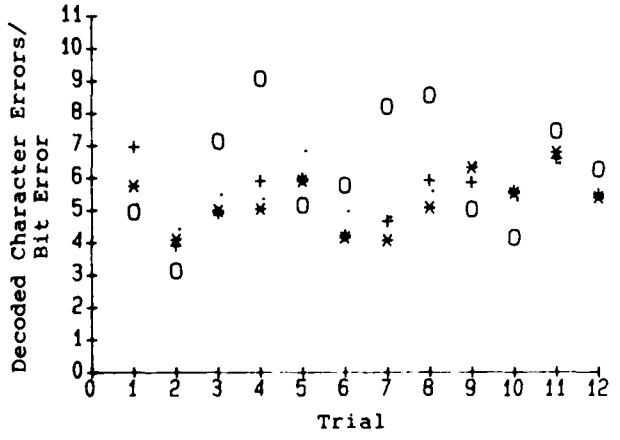


Figure 3. Huffman Code Error Propagation Results with Narrative II as Training File (. = Narrative I, + = Narrative II, * = Narrative III, o = Narrative IV)

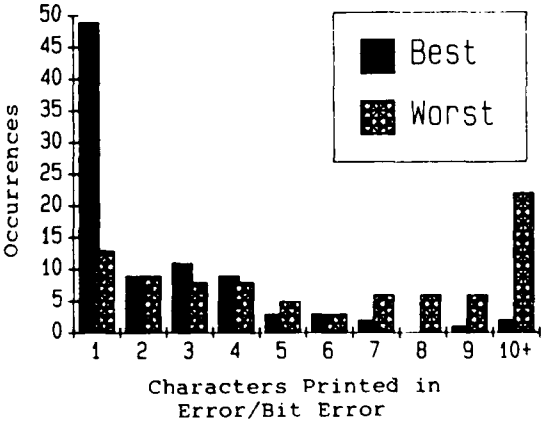


Figure 4. Distributions of Printed Characters in Error per Bit Error for Best and Worst Huffman Codes

There is a dramatic difference in the structure of the distributions for each of the narrative files used as a training file for the Huffman codes found to give the best and worst performance in a channel with errors. The best distributions have a preponderance of short length sequences (lengths one, two, and three) while the worst distributions tend to be relatively flat with the occurrence of extremely long character error sequences (35, 104, 65, and 90 for narratives I, II, III, and IV, respectively). The simulation for narrative IV shown in figure 4 only resulted in three printed character sequences longer than 7 characters, one each of 9, 10, and 14 characters. The average length of a printed sequence of character errors for this Huffman code, presented in table 4, was 2.4 printed character errors.

Some experiments were run using an operator-interactive program to determine the percentage of errors introduced into a text file through Huffman decoding of bit errors that could be corrected through narrative context. It appears possible to change a bit likely to be in error and then to use a standard spell check program to check whether reinitialization of Huffman decoding by the change leads to more reasonable results. The potential of such an algorithm could be assessed by using an operator-interactive program--with the operator choosing the decoding which provided text which made the most sense.

A 4271 character narrative file consisting of 88 lines and 4456 bytes was chosen to assess operator-interactive correcting of narrative character errors. Bit errors were introduced randomly at a rate of .005. This error rate would lead to an estimated 90 characters containing a bit error ((.005)(4271 characters)(4 bits/character)). These 90 bit errors led to 362 character decoding errors. After the interactive session the operator was able to reduce the number of character decoding errors to 85 errors (that is the number of character errors were reduced by 76 percent).

TABLE 4. HUFFMAN CODE WORDS FOR NARRATIVE FILE IV WHICH PROVIDED THE BEST PERFORMANCE IN A CHANNEL WITH ERRORS

CHAR	CODE WORD	CHAR	CODE WORD
" "	01	@	100010100010
!	10001010000000011	A	1110
"	11111110101	B	0001101
#	100010100000000100	C	11110
$	100010100000000101	D	000011
'	100010100000001	E	110
(	00100001	F	100011
)	00100000	G	111110
*	111111101001	H	10000
+	0010101000	I	1010
,	1111110	J	111111110
-	10001011	K	001010101
.	0010001	L	000010
/	111111111	M	001001
0	00101011101	N	1001
1	001010110	O	1011
2	0010101111	P	001011
3	1111111011	Q	1000101001
4	001010111000	R	00000
5	100010100011	S	00010
6	0010101110010	T	0011
7	10001010000001	U	000111
8	100010100001	V	0001100
9	0010101110011	W	0010100
:	100010101100	X	111111100
;	111111101000	Y	1000100
<	1000101000000000	Z	1000101010
=	0010101001	^	100010101101
>	1000101000001	~	10001010111

COMMA-FREE CODES

Comma-free codes are binary codes so constructed that it is possible to identify individual code words prior to decoding the received bit stream. We analyze a family of comma-free codes, known as "suffix/prefix" codes, found by R. A. Scholtz [reference 2].

In order to illustrate the ideas involved in Scholtz's construction process, we discuss a particularly simple example of the Scholtz construction process. Scholtz constructs sets of code words sequentially. We begin with the set of two code words {0,1}. The next code set is obtained from this set by choosing "1" as a suffix. This set consists of {0,01,011,0111,01111,...}. A next set is obtained by choosing an element of this set as either a suffix or a prefix. If "0" is chosen as a prefix, the code set becomes { 0...01...1 with at least one "0" and one "1") (we call this code the suffix-prefix comma-free code); if "0" is chosen as a suffix, the code set becomes { 01...10...0 with at least one "1") (we call this code the suffix-suffix comma-free code); if "01" is chosen as a suffix, the code set becomes {0...01...101...01 with zero or two or more "1"s). Generally speaking, new code words can be constructed by either using suffixes or prefixes. The process can be carried out any number of times.

Figure 5 shows the operation of the three-step process used to insert "commas", which corresponds to the two-step construction process used to construct the suffix-suffix code described in the last paragraph. (Individual code words are enclosed in brackets and the bits maintained in alignment in figure 5 to aid the reader. The transmitted and received bit stream would consist simply of the bits enclosed in these brackets with no indication of where one code word ended and another began.) The comma insertion process parallels the code construction process and the reader could readily verify that it reconstructs the correct code words in the absence of errors, except possibly at the beginning or end of the decoded sequence. It proceeds by first inserting commas between all the bits and then successively deleting those according to rules based on the suffix choices. For example, corresponding to choosing "1" as a suffix, commas are removed preceding "1"s in the second step of the comma insertion process. Other algorithms are available to insert commas for some of the comma-free codes. For example, for a suffix "1" prefix "0" code, one need only insert a comma between every string of "1"s and string of "0"s to identify the code words.

We turn now to the selection of the comma-free code which would provide the best data compression for a narrative file with particular character probabilities of occurrence. R. A. Scholtz does not discuss how to match his comma-free code construction process to the probabilities of occurrence of the characters to be encoded to provide the best compression. We found how to survey the possible codes in terms of the distributions of their code word lengths. The survey can be conducted without specifying the particular code word chosen at each step of the construction process, or whether the chosen word at each step is used as a suffix or a prefix. All that need be specified is the lengths of the words chosen for suffixes and prefixes.

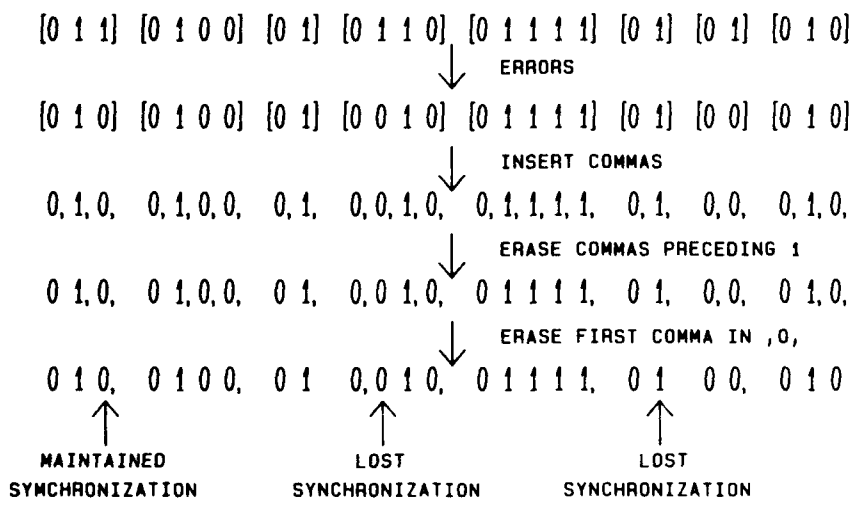


Figure 5. An Example of the Comma-Free Algorithm to Insert "Commas" in a Channel with Errors

A natural way to survey the codes is to survey them inductively based on the construction steps. Toward this end, let C[k] denote the set of code words produced after the first k steps of the construction process; let C[1] = {0,1} be the starting point in the construction process; and let n[k](j) denote the number of code words of length j in set C[k]. In the suffix/prefix construction process, a word used as a suffix or prefix can no longer be used as a code word. To take this into account let n^[k](j) = n[k](j) - 1, if j = s, and = n[k](j), if j ≠ s, where s is the length of the suffix or prefix chosen to construct the (k+1)-th code set from the k-th code set.

The inductive formula for the number of comma-free code words of length j resulting from the choice of a suffix or prefix of length s in the (k+1)-th construction step is given by n[k+1](j) = n^[k](j) + n^[k](j-s) + ... + n^[k](j-ns), with the convention that n^[k](j-ns) = 0 if j - ns < 1.

The above formula allows easy compilation of tabular summaries of distributions of code word lengths for available comma-free codes constructed using the suffix/prefix process. Table 5 illustrates its use. With the exception of the first column, the numbers of code words in a code are only summarized up through the length of code word needed to allow the coding of 58 characters. The codes summarized in table 5 begin with C[1] = {0,1} and each new code set C[2] through C[6] is obtained from the previous one by using one of shortest available code words as a suffix or prefix in the construction process.

We turn now to the selection of a comma-free code to encode the narrative files used earlier to assess Huffman codes. The probabilities of occurrence for the character set used for our 58 character simulations of the Huffman code, has the property that the character probabilities fall off rapidly from the most used characters to the least used characters (as shown in table 1 which was presented earlier). In such a situation, if we could closely match the code word lengths provided by the Huffman code constructed for the given character probabilities of occurrence for the first 10 to 15 characters, we would expect very similar compression performance from that comma-free and a Huffman code.

Table 2 (previously discussed) shows how closely the simplest suffix/prefix candidate code word lengths match those provided by the Huffman codes. The first four columns present Huffman word lengths and a fifth column presents the word lengths for any of the suffix-prefix comma-free codes obtained by use of "0" and "1" as suffixes or prefixes in a two-step construction. The assignment of code words to characters is optimum for narrative file II. However, as can be seen from table 2, this assignment leads to excellent word length agreement through the first 40 characters for all of the narrative files so no other word assignments were studied.

Table 6 presents a comparison between Huffman code bits-per-character values and comma-free code bits-per-character values for the code word to character assignments shown in table 2. The next best comma-free code appears to be the code (1,1,3) for which similar calculations revealed a penalty of .114 (rounded down to three significant places) bits-per-character for using this comma-free code instead of the Huffman code for narrative file II.

There are choices in the construction of comma-free codes leading to the same distribution of code word lengths. The behavior of the code in a noisy channel depends on these choices. We single out the suffix-prefix and suffix-suffix codes for detailed study. These codes represent the two fundamentally different codes providing the compression summarized in table 6.

TABLE 5. PARTIAL SURVEY OF THE DISTRIBUTIONS OF CODE WORD LENGTHS FOR COMMA-FREE CODES

WORD LENGTH	COMMA-FREE CODE					
	C[1]	C[2]	C[3]	C[4]	C[5]	C[6]
1	2	1				
2		1	1			
3		1	2	2	1	
4		1	3	3	3	3
5		1	4	6	6	6
6		1	5	8	9	9
7		1	6	12	15	18
8		1	7	15	21	27
9		1	8	18	27	
10		.	9	24		
11		.	10			
12		.	11			

TABLE 6. COMPARISON OF BITS-PER-CHARACTER VALUES OF HUFFMAN AND COMMA-FREE CODES

NARRATIVE FILE	AVERAGE NUMBER OF BITS-PER-CHARACTER	
	HUFFMAN CODE	COMMA-FREE CODE
I	4.04	4.10
II	3.95	4.00
III	4.15	4.26
IV	3.63	3.81

The impact of bit errors on character errors can be determined analytically for the suffix-prefix and the suffix-suffix codes. The basic observation is that for these two codes a bit error in the middle word of three code words w[1]w[2]w[3] always leads to a bit sequence which can be expressed as three other code words w^[1]w^[2]w^[3] with at most two of the code words in error. The probability that a particular bit in w[2] is in error is given by the probability that the character w[2] represents occurs times the probability that bit in w[2] is in error, which is just 1 divided by the length of w[2].

Generally speaking, four probabilities determine the error propagation properties of a comma-free code: P[D], the probability that a bit error leads to the deletion of the comma separating two code words; P[M], the probability that a bit error leads to the misplacement of a comma; P[A], the probability that a bit error leads to the addition of a comma; and P[N], the probability that a bit error leads to no change in the placement of the commas. The impact of a bit error on the comma determines the number of characters decoded in error and the number of characters output by the decoder in error. In particular: (1) if a bit error leads to comma deletion then two characters are incorrectly decoded as a single character or not decodable; (2) if a bit error leads to comma movement then two characters are incorrectly decoded into two characters; (3) if a bit error leads to the insertion of a comma (always within the code word with the bit error) then one character is incorrectly decoded into two characters; and (4) if there is no change in the commas then one character is incorrectly decoded into a single character.

The required calculations for a particular assignment of the suffix-prefix code words to a 58-character set are easy but tedious. We omit the majority of the details (see reference 4 for them) and summarize the results of the calculations. Calculations were carried out for the probabilities of occurrence of the characters in narrative file II. Similar results are expected for the remaining three narrative files.

For the assignment of suffix-prefix codes to characters described above and summarized by table 2, the following statistics were obtained: P[D] = .42, P[M] = .21, P[A] = .16, and P(N) = .21. Note that about three-quarters of the contribution to P[D] is that provided by the code word "01" assigned to the space character with probability of occurrence 0.317. The average number of characters decoded in error per bit error is (.42)(2) + (.21)(2) + (.16)(1) + (.21)(1) = 1.63. The average number of incorrect printed characters per bit error is given by (.42)(1) + (.21)(2) + (.16)(2) + (.21)(1) = 1.37. Note that these values are obtained by treating words too long to be decoded because they exceed the longest word assigned one of the 58 characters as being incorrectly decoded. (Such characters could be decoded into a 59th character indicating an error has occurred.) The calculations presented clearly indicate that the performance of the suffix-prefix code in an error channel is considerably better than the performance of any Huffman code that we found.

A two-step comma-free code construction using a one-bit prefix and a one-bit suffix, no matter what choices are made, leads to code words either of the form 0...01...1 or 1...10...0 with each code word containing at least one "1" and one "0". One of these codes can be obtained from the other by interchanging "1"s and "0"s. If this were done to the assignment of code words, the same probabilities would be obtained as for the code word assignment before the interchange. Thus all the prefix-suffix codes using a one-bit prefix and a one-bit suffix would for these assignments have the same error statistics.

We turn now to estimating the impact of errors on the suffix-suffix code discussed earlier. Recall that the code words for this code have the structure 01...10...0 with at least one "1". This code differs from the suffix-prefix code in that (1) the impact of an error in the first bit position of a code word depends on the ending of the previous code word and (2) the impact of an error in the second bit position depends on whether or not the second bit is the only "1" in the code word. Again calculations were only carried out for the probabilities of occurrence of the characters in narrative II.

An error in the first bit leads to the movement of a comma or the deletion of a comma depending on whether the first code word ends in "0" (probability of 0.51) or ends in "1" (probability of 0.49). It follows that the probability that an error in the first bit leads to the movement of a comma is .158 and the probability that an error in the first bit leads to the deletion of a comma is .152.

An error in the second bit leads to the deletion of a comma or the addition of a comma, depending on whether or not the second bit in error was the only "1" in the code word. It turns out that the probability that an error in the second bit will lead to the deletion of a comma is 0.217 (rounded to three places) and that the probability that an error in the second bit leads to the addition of a comma is 0.95.

The remaining calculations are similar to those for the suffix-prefix case. We found that the probability that a bit error leads to the addition of a comma through changing other than the first or second bit is .18 and the probability that a bit error leads to no change in the commas through changing other than the first or second bit is .20.

From these calculations, it follows that P[D] = .37, P[M] = .16, P[A] = .27, and P[N] = .20 so that the average number of characters decoded in error per bit error is 1.51 and the average number of printed characters which are incorrect per bit error is 1.45. These statistics can be seen to apply to all suffix-suffix and prefix-prefix codes with length one suffixes or prefixes.

The result that a single bit error leads to at most two character errors (decoded or printed), established for the two simplest kinds of comma-free codes, can be extended to other comma-free codes. Some additional terminology is needed to facilitate the discussion of general comma-free codes. Let k denote the kernel of the code under construction, p(i), i = 1, 2, ... denote the prefixes used in the code under construction, and s(j), j = 1, 2, ... denote the suffixes used in the code under construction. Suppose that the codes under discussion satisfy: (1) k = "0" or "1", (2) both "0" and "1" are used as either prefixes or suffixes, and (3) the length of the prefix or suffix used in k-th construction step is less than or equal to the length of the prefix or suffix used in the (k+1)-th construction step. A code is called exhaustive if for each of the steps in the code construction process the code word chosen as either a prefix or suffix is one of the shortest code words available.

For an exhaustive comma-free code, a single bit error can lead to at most two characters decoded in error. To establish this result, consider (1) an incoming sequence of bits as a sequence of kernels, prefixes, and suffixes, and (2) the comma-insertion algorithm (after the first step) consists of deleting commas between the kernels, prefixes, and suffixes. Now, let us discuss the potential impact of a single bit error occurring in a kernel or in a prefix or suffix of the code words.

Let us denote the word with a bit error by use of "^". Consider the incoming sequence of binary bits parsed into codewords w(1)w(2)w^(3)w(4)w(5). Under what conditions will the comma separating w(1) and w(2) or the comma between w(4) and w(5) be altered as a result of a bit error somewhere in the codeword w(3)? Each of these words is constructed from the kernel and prefixes and suffixes, as described above so that for the comma between w(1) and w(2) to be erased by the comma-insertion algorithm, the prefix or kernel beginning w(2) must be transformed into a suffix through a bit error in w(3). Since none of the bits in w(2) are in error, this can only happen if the addition of bits to the bits of w(2) has created a suffix used in the construction process; i.e., there exists a code word of shorter length in the code than some suffix in the code, a contradiction. For the comma between w(4) and w(5) to be erased by the comma-insertion algorithm, the suffix or kernel ending w(4) must be transformed into a prefix through a bit error in w(3). Since none of the bits in w(4) are in error, this can only happen if the addition of bits to the bits of w(2) has created a prefix used in the construction process, a contradiction.

It is clear that one could improve upon the results by examining the non-exhaustive codes to see if either of the above phenomena can occur for a particular selection of prefixes or suffixes. A cursory examination allowed us to establish that for the codes with sequences of suffixes or prefixes with the lengths indicated by (1,1,3), (1,1,3,3), (1,1,2,4), and (1,1,4), a single bit error can never lead to four or more character decoding errors.

It is also possible to use Huffman or comma-free codes to encode characters based on one or more of the previous characters encoded. We refer to this as encoding based on conditional probability of occurrence of characters. Since the error propagation properties of Huffman codes were so much worse than those for comma-free codes, we restricted our attention to comma-free encoding of characters conditioned on the occurrence of previous characters.

The structure of the comma-free codes limits the impact of bit errors. However, given an error has occurred in an encoded character, then it will be decoded in error, the next character will be decoded in error if conditioned on it, the next code word decoded in error if conditioned on either of the previous characters, and so on. To prevent decoding errors from propagating in this manner, it is necessary to reinitialize the coding process fairly often.

Military and commercial messages are transmitted using characters of various types, categorized as the set of letters { A,B,C, ..., X,Y,Z }, the set of numbers { 0, 1, 2, ..., 9 }, and the set of symbols { punctuation symbols, special characters, control characters }. The basis of the conditional probability encoding approach is to reinitialize the encoding process whenever c is a symbol. Furthermore, in order to limit the propagation of decoding errors, symbols are encoded independent of previously encoded characters.

Off-line processing would be used to determine the assignment of comma-free code words to characters. The overall probabilities of occurrence can be used to assign code words to symbols (although we found that this was not optimum). The crucial assignment is the code word assigned to the space. (We found that the assignment of a three-bit code word to the space gave the best compression.) The variable length code words assigned to fixed length words representing characters or numbers are the remaining code words.

We assign a code word to a letter depending on whether or not it is the first letter of a word and to a number depending on whether or not it is the first number in a sequence of numbers. The comma-free code words assigned unconditionally to letters or numbers are based on the probabilities that a character occurs as the first letter in a word (including one letter words) or the first number in a sequence of numbers (including single numbers); the conditional assignment of comma-free code words to letters and numbers is based on the probabilities that a character follows a specific letter or follows a specific number.

The data compression provided by the comma-free encoding approach described above was estimated by the following formula: $3/(L+1) + N[Uncond]/(L+1) + N[Cond](L-1)/(L+1)$, where N[Uncond] = the average number of bits assigned the starting letter in a word, N[Cond] = the average number of bits encoding a character conditioned on the receipt of a previous character, and L = the average length of a word (with symbols treated as length 1 words). This formula neglects the contributions of numbers conditioned on other letters or numbers, which are extremely rare in the narratives. The first term, $3/(L+1)$, would be accurate if blanks separated all the words, another fairly valid assumption because of the absence of short sentences in the file manuscript and the usual practice of separating them from letters and numbers using spaces. Table 7 summarizes the results obtained using the above formula. By using a length 3 code word for the blank, and making effective use of the length 2 code word for the most commonly occurring character starting a word, the table shows that the bits to encode spaces and beginning of words are about the same as the bits to encode the remainder of words.

We estimate the average number of characters decoded in error for the suffix-prefix comma-free encoding of characters conditioned on the previous character for narrative file II by considering the impact of bit errors on an average word followed by a space. The expected structure of a five-character word followed by a space is (4.4 bits)(3.5 bits)(3.5 bits)(3.5 bits)(3.5 bits)(3 bits), where we have rounded 3.55 down to 3.5 bits to compensate for rounding 4.69 up to 5 letters. Thus, a word is expected to consist of 18.4 bits and a word followed by a space of 21.4 bits. We proceed by estimating the number of characters in error as a result of bit errors in the different bit positions. A bit error in the first bit of a code word of a letter leads to the previous character being decoded in error; such errors contribute $(1/21.4)(6+5+4+3+2) = .93$ characters in error. A bit error in any of the other bits of the letter code words, with the exception of the very last bit of the code word of the last letter in the word, leads to the rest of the word being decoded in error; these bit errors contribute $(1/21.4)[(3.4)5+(2.5)4+(2.5)3+(2.5)2+(1.5)] = 1.92$ characters in error. If the last bit of the code word of the last letter of the word, or the first bit of the space is in error, both the last letter and the space will be decoded in error, and therefore the following word will be decoded in error; these bit errors contribute $(2/21.4)(7) = .65$ characters in error. Bit errors in the remaining two bit positions of the space lead to it and the following word being decoded wrong, so they contribute $(2/21.4)(6) = .56$ characters in error. Totalling these contributions leads to an estimate of 4.1 characters in error per bit error.

Table 7 allows us to estimate the potential payoff of encoding the third through last letters of words based on the occurrence of two previous letters. Since there is a single remaining comma-free 2-bit code word, a single 3-bit code word, and the remaining code words are of length 4 or more, the least length that one might expect for the doubly conditioned encoded characters is $(1/2)2 + (1/4)3 + (1/4)4 = 2.75$. Then a lower bound on the expected overall average number of bits would be $(1/2)(3.6) + (1/2)(2.8) = 3.2$ for this encoding approach. We would expect about 3.3 to 3.4 bits per character performance if we carried out the calculations more exactly.

TABLE 7. AVERAGE NUMBER OF BITS AND WORD LENGTHS FOR COMMA-FREE CODES USING CONDITIONAL PROBABILITIES

	AVERAGE NUMBER OF BITS			
NARRATIVE FILE	START OF WORD	REMAINDER OF WORD	OVERALL	AVERAGE WORD LENGTH
I	4.51	3.55	3.62	4.64
II	4.44	3.55	3.61	4.69
III	4.71	3.64	3.71	5.15
IV	4.53	3.54	3.62	4.91

The estimated number of character errors per bit error for a doubly conditioned comma-free encoding scheme using a suffix-prefix code is the same as for a singly conditioned comma-free encoding scheme; all the letters or numbers after an error are in error until the encoding process is reinitialized whether singly or doubly conditioned.

SUMMARY AND CONCLUSIONS

The performance of Huffman codes, suffix/prefix comma-free codes, block codes, and some variants of Baudot codes was obtained for the encoding of narrative files of a personal computer for a 58-character set. Figure 6 summarizes the results of this investigation. The performance of each encoding approach is summarized by its compression performance measured by the average number-of-bits per character (the y-axis) and by its error propagation statistics measured by the average number of characters decoded in error per bit error (the x-axis). A word error was treated as equivalent to five character errors to allow comparison of character and word encoding approaches.

The Military Baudot code uses two shift keys ("LTRS" and "FIGS"), which we model mathematically as a single shift key generalized Baudot code, to reduce the number of bits required to transmit information from 6 to about 5.06. Generalizations of this construction can further reduce the average number of bits required to around 4.5 bits-per-character while maintaining a basic block structure.

A suffix/prefix comma-free code can be constructed which provides nearly the same data compression as a Huffman code, about 4 bits-per-character. For the character set and probabilities of occurrence of the characters of the set used in the Huffman simulation, the penalty varied from a low of .05 bit per character to a high of .18 bit per character for four narrative files. A single bit error can lead to at most two character errors for the above prefix/suffix codes while a single bit error was found to lead to as many as 90 character errors for a Huffman code. Relative to a 1-shift Baudot code, the best Huffman code requires an average 24% fewer bits for encoding with 43% more decoded character errors, while the best comma-free code requires 22% fewer bits for encoding with 18% more decoded character errors.

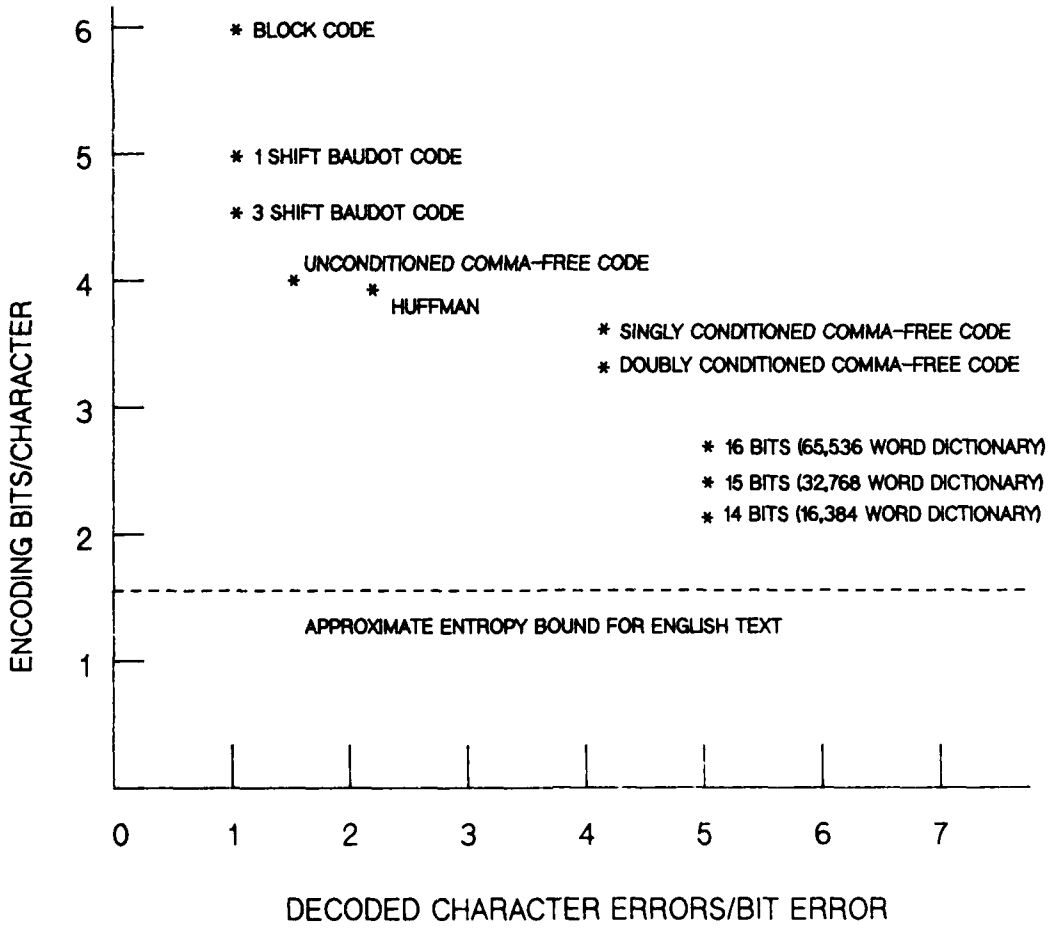


Figure 6. Data Compression and Error Propagation in Noisy Channels

The use of comma-free encoding for words with non-letters and non-numbers encoded independently of the previous character led to an estimated 3.6 bits per character along with 4.1 characters in error per bit error for a singly conditioned encoding scheme, and about 3.3 bits per character along with 4.3 characters in error per bit error for a character for a doubly conditioned encoding scheme. Relative to a 1-shift Baudot code, the singly conditioned scheme requires 29% fewer bits for encoding and the doubly conditioned scheme about 35% fewer bits both with 289% more decoded character errors.

Block encoding of words for dictionaries ranging in size from 16,384 to 65,536 words requires 2.33 to 2.67 bits per character with one word error per bit error.

The following conclusions were drawn as a result of the investigation:

(1) For unconditioned character encoding, comma-free codes significantly outperform Generalized Baudot codes and Huffman codes in a noisy channel. They provide nearly the same compression and have significantly fewer decoded or printed character errors than Huffman codes.

(2) Additional compression is achievable by the use of comma-free encoding of characters based on their conditional probability of occurrences.

LIST OF REFERENCES

1. Huffman, D., "A Method for the Construction of Minimum Redundancy Codes", Proceedings of the Institute of Radio Engineers, Vol. 40, pp. 1098-1101, September 1952.

2. Scholtz, R., "Codes with Synchronization Capability", IEEE Transactions on Information Theory, Vol. IT-12, No. 2, April 1966.

3. Scholtz, R., "Maximal and Variable Word-Length Comma-Free Codes", IEEE Transactions on Information Theory, Vol. IT-15, No. 2, March 1969.

4. Bond, J., "The Performance of Data Compression Codes in Channels with Errors", personal correspondence of author, 19 February 1987.

5. Bond, J., "Variable Length Data Compression Encoder", Patent Disclosure, 1987. (available from author)

NEW METHOD FOR SUPPRESSION OF GONIOMETER DF ERRORS BY RADIO NOISE FROM OVERHEAD POWER LINES

J. Benger, G. Deubach, Dr. G. Kosel
FGAN-IFE
Neuenahrer Str. 20
5307 Wachtberg-Werthhoven
Federal Republic of Germany

SUMMARY

Overhead power lines below 70 kV often generate impulsive radio noise by sparking which may heavily interfere on the performance of military VHF communication systems. The first part of this presentation deals with that kind of radio noise, its characteristics and possible sources and shows some measurements.

Measurements also show that radio noise from sparking overhead power lines may influence automatic goniometer direction finding; one gets a wrong bearing, that of the noise source. Therefore, a new method for eliminating this noise influence was investigated. It determines the lower envelope of the DF signal and is only applicable in the digital area of signal processing. The method is described and its effectiveness is presented.

1. INTRODUCTION

In areas with high population as in many parts of Europe we find dense nets of overhead power lines. These power lines and associated equipment often generate radio noise, caused by partial electrical discharges, such as corona, or by complete electrical discharges across small gaps (1). Radio noise from lines above about 100 kV is principally due to conductor corona while lines below 70 kV are generally free of corona-type radio noise, and only gap-type discharges may occur. According to the purpose of this meeting radio noise from gap-type discharges on overhead power lines and its interference on the performance of DF systems with circularly disposed antenna arrays is described in detail.

Furthermore there will be shown a new cost effective method to eliminate goniometer DF errors caused by radio noise from sparking on overhead power lines.

2. RADIO NOISE FROM SPARKING ON OVERHEAD POWER LINES

If radio noise from overhead power lines is experienced in VHF range, it is probably produced by sparking. The gap-type radio noise source is a complete electrical discharge. Unbonded conducting parts of a power line can become electrically charged and the potential difference between adjacent conducting parts will increase even if both parts are floating, not connected to a line conductor or to earth.

If the distance separating the conducting parts is small, the increasing field strength in the space between may reach a critical level and lead to a complete breakdown of the gap. Avalanche ionization initiates the development of an arc, gap discharge occurs, the potential difference across the gap then falls to a low level and the arc extinguishes. The whole sequence is repeated when the conducting parts become re-charged. The repetition rate depends on the charging and discharging time constants of the circuit and the value of the surrounding electric field as well as on the length of the gap.

Basically one has to distinguish between two different cases of sparking on overhead power lines: In most cases sparking is due to bad, that means loose or imperfect, contacts. Fig.1 shows an example of an imperfect contact between metallic parts of an insulator (point 1). When the surfaces of these metallic parts have weathered, oxides and sulphides may coat the surfaces and, mainly in dry weather, may lead to gap-type discharges. In wet weather, the comparatively small gaps involved usually become bridged with moisture, thereby establishing a conducting path. The radio noise generated is called dry noise. Sometimes sparking occurs just in wet weather on the surface of a heavily polluted insulator (Fig.1, point 2). Then, because of heating due to leakage currents, dry bands are created. The potential difference between the ends of such a small dry band increases and initiates the development of an arc.

The form of the radio noise generated by sparking is impulsive as it is produced by current pulses. The amplitude and number of these pulses depend on the voltage stress across the insulating part, its shape, dimensions and conductivity. Individual sparks can occur at many hundreds to a few thousands of times per second. The repetition rate is, however, one order of magnitude lower than the range of repetition rates of corona discharges. As the voltage of an a.c. power line changes steadily, a train or burst of sparks is generated when the line voltage and consequently the field strength in the gap exceeds a critical value; sometimes this happens only during one of both half-cycles of the line voltage, but usually the repetition rate of these bursts is twice the line frequency.

A significant factor in the shape of a single discharge pulse is its steep rise time and, therefore a broad range of high frequencies is produced and emitted. Fig.2 shows that the spectrum of radio noise by sparking may reach from medium wave band up to several hundred megahertz. If the discharge process excites a fitting or component which is capable of oscillating at a particular frequency, due to its geometrical dimensions, then even higher energetic narrow-band radiation at this frequency may interfere on the performance of military radio communication systems.

Significant levels of the radio noise by sparking are normally confined to the vicinity of the power line reaching out to several hundred metres on either side. If a resonant condition occurs in a fitting or component, a narrow band of radiation may be measured at distancies up to a few kilometres.

Numerous measurements of radio noise by sparking were made in VHF range. Fig.3 shows part of such a spectrum taken by a spectrum analyzer within 30 msec. When sparking occurred the noise floor increased by about 15 dB. The repetition rate of groups of sparks was 100 Hz, i.e. double the line frequency, while individual sparks repeated several thousand times per second. All measurements showed following: The noise floor remarkably increased in frequency range up to 180 MHz. Mostly the weather was dry and the repetition rate of groups of sparks was twice the line frequency. Sometimes, different noise sources sparking at the same time could be determined.

3. IMPACT OF SPARKING NOISE ON AUTOMATIC GONIOMETER DIRECTION FINDING

Radio noise by sparking may heavily impede automatic goniometer direction finding. This method of direction finding by using a mechanical rotating goniometer, whose inputs are switched to the elements of a circularly disposed antenna array (CDAA), is explained (Fig.4). The goniometer simulates a continuous rotation of a high gain beam. This RF beam is fed into the antenna input of the DF receiver; there the signal is band selected whose direction of arrival has to be found. The IF signal is rectified, low pass filtered and digitized (DF signal). The north pulse signal of the goniometer indicates the north transit of the rotating RF beam and the bearing tone should be used as timer for the A/D converter. The DF signal is processed for automatic bearing or may be displayed for manual bearing. The goniometer contains a movable rotor assembly and a stationary stator assembly (2). The stator supports input networks and stator coupling combs, the rotor supports rotor coupling combs and beamforming networks. When the rotor combs are in mesh with the stator combs, RF signals are capacitively coupled to the beamforming networks. Fig.5 and Fig.6 demonstrate how the antennas of a sector of the associated CDAA are properly combined to a sum and a difference beam; each half of the circular antenna sector is transformed into a linear antenna array. The sum and difference beam are capacitively coupled from the rotor through rotary couplers to the appropriate output connectors. Fig.7 shows an overall view of a mechanical goniometer with its inputs and outputs.

The scan rate (rotor speed) of the goniometer is about ten revolutions per second. For each revolution one gets an amplitude distribution over azimuth (DF signal) which shows the antenna diagram of the rotating high gain beam when there is one signal within IF bandwidth. An example is given in Fig.8. The peaks of the discrete samples (900) were combined. As the sidelobes are more than 10 dB down to the mainlobe, and the signal-to-noise-ratio is only 10 dB, they can't be seen here. The period for getting one complete amplitude distribution is 100 msec.

The task is automatically finding the position of the centre of the antenna diagram as true bearing. At first the position of the maximum has to be found before fine bearing is done in a way which shall not be described here. Bearing accuracy will be better than one degree when the DF signal is undisturbed. But sometimes the DF signal is interfered by sparking noise. If this non-Gaussian radio noise is remarkably stronger (Fig.9) than the interesting signal (most signals are weak), then known methods for reducing the influence of unwanted signals, i.e. filtering the DF signal with a bandwidth equal to that of the wanted signal, really reduce but do not eliminate their influence as is shown in Fig.10. As result one gets a completely wrong bearing, that of the noise source.

4. NEW METHOD PREVENTING AUTOMATIC DF ERRORS BY SPARKING NOISE

There has been found a reliable method to prevent goniometer DF errors caused by radio noise from sparking overhead power lines. Fig.9 shows an example of a measured DF signal with heavy interference by sparking noise. Looked at more closely it is obvious that this interference may be suppressed by forming the lower envelope. This is not possible with analog, but with digital signal processing in the following way: Superimpose a window, the width of which is n samples (n integer, $n > 1$), seek the lowest amplitude (minimum) within the window and substitute the amplitude in the middle of the window by that minimum. Then shift the window by k samples (k integer, $1 \leq k \leq n$) and again substitute the amplitude in the middle of the window by its minimum. Repeat this operation as long as one reaches the first set window. So one gets a new (minimized) DF signal with suppressed sparking noise (Fig.11), which then may be further processed up to the correct fine bearing.

The optimum width of the window depends on the characteristic of the radio noise and mainly on the beamwidth of the rotating antenna beam. Measurements showed, that the window should be as wide as 0.5 to 0.8 times the beamwidth.

By this new method forming the lower envelope the number of samples with different amplitude values is reduced; it corresponds to a reduction of bandwidth and therefore causes an increase of signal-to-noise-ratio. By use of this method the number of automatic goniometer DF errors could be lowered remarkably.

5. CONCLUSION

Military VHF communication systems operated in areas with overhead power lines (in particular below 70 kV) may be disturbed by sparking noise. This kind of noise consists of groups of pulses repeating with single or double the line frequency and appears mainly in dry weather.

In DF systems with circularly disposed antenna arrays sparking noise sometimes causes completely wrong bearings. A new method has been found which prevents these goniometer DF errors. It forms the lower envelope of the measured signal.

This effective method may also be helpful when other kinds of measurements are superimposed by impulsive noise, as digital signal processing means today are available at low cost.

6. REFERENCES

(1) Commission Électrotechnique Internationale, Comité International Spécial des Perturbations Radioélectriques (C.I.S.P.R.)
Publication 18-1, Première édition, 1982:
Caractéristiques des lignes et des équipements à haute tension relatives aux perturbations radioélectriques,
Première partie: Description des phénomènes
CEI 1982
Bureau Central de la Commission Électrotechnique Internationale
3, rue de Varembé
Genève, Suisse

(2) Olektron Corporation Framingham Division,
Framingham, MA 01701, USA: Goniometer Model N-AS-36SV-1-S1,
Technical Manual, Operating and Maintenance Instructions

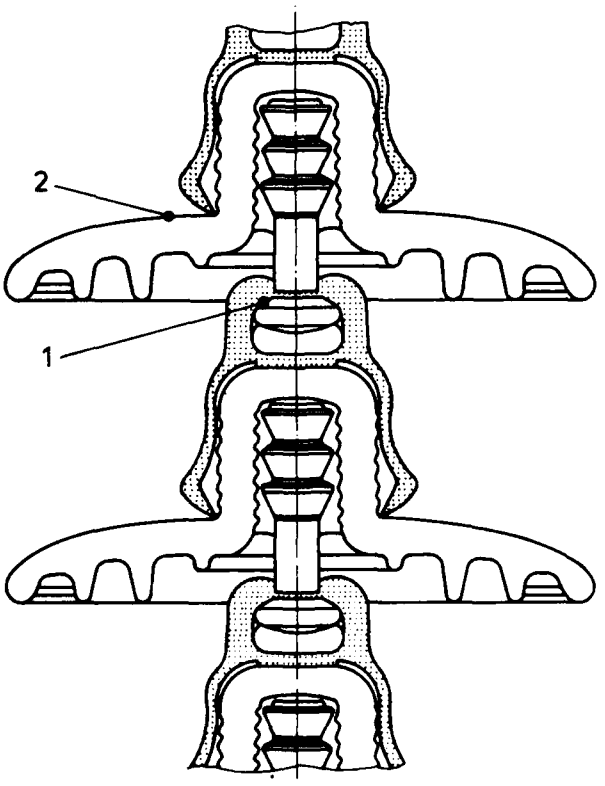


Fig. 1: Part of 20 kV-insulator: Examples of possible radio noise source

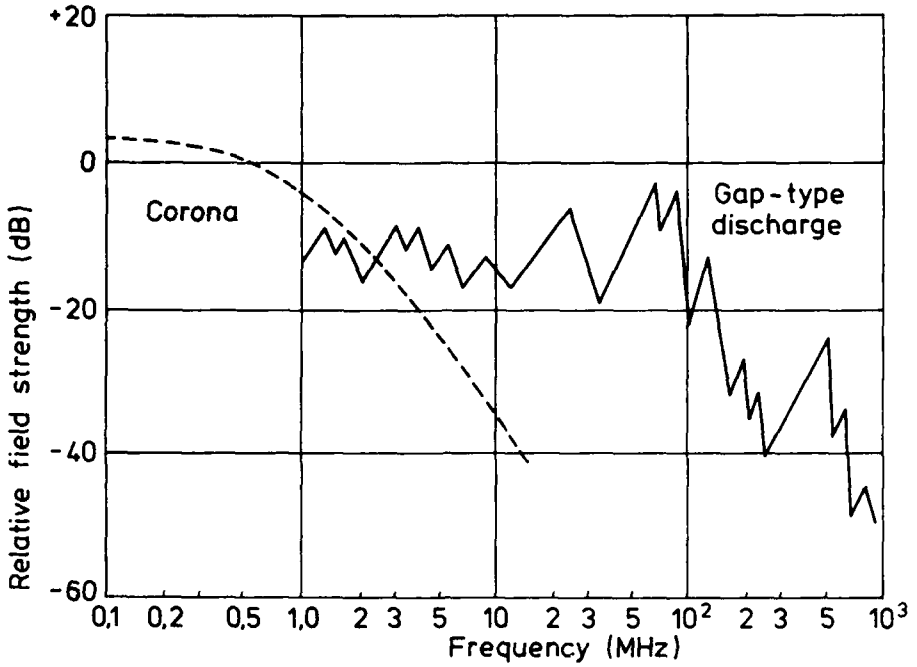


Fig. 2: Example of relative strength of radio noise field as a function of frequency

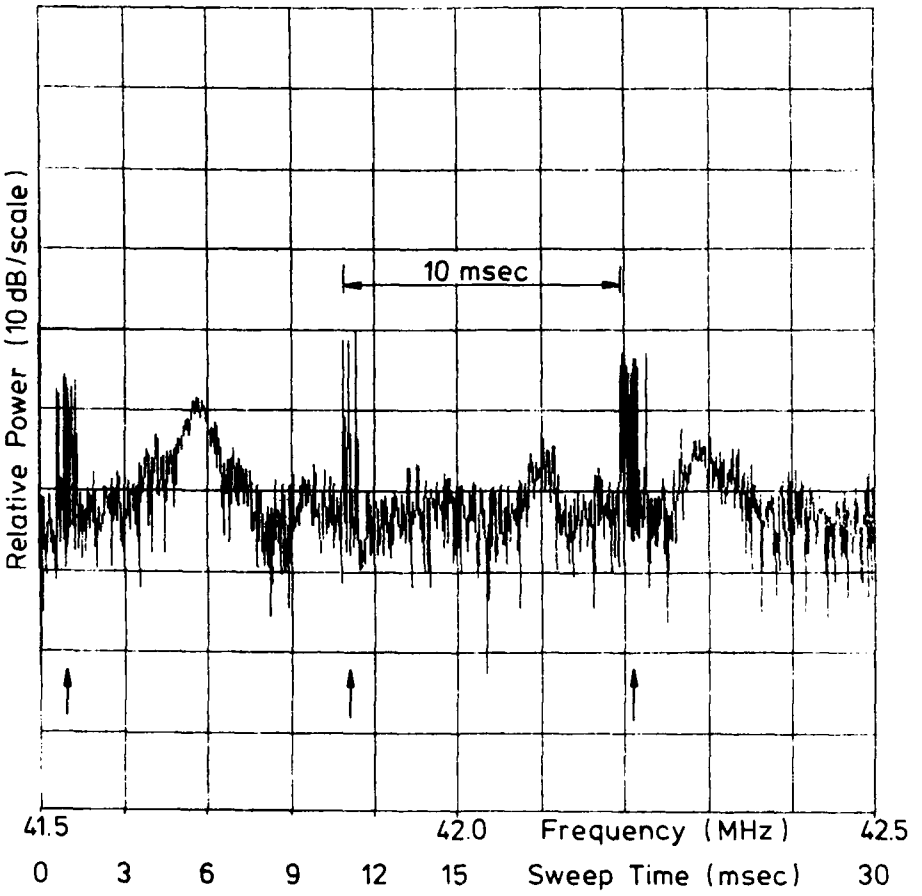


Fig. 3: Oscillogram of Spectrum 41.5 - 42.5 MHz
Sweep Time 30 msec / Bandwidth 30 kHz
Radio noise by sparking at arrows

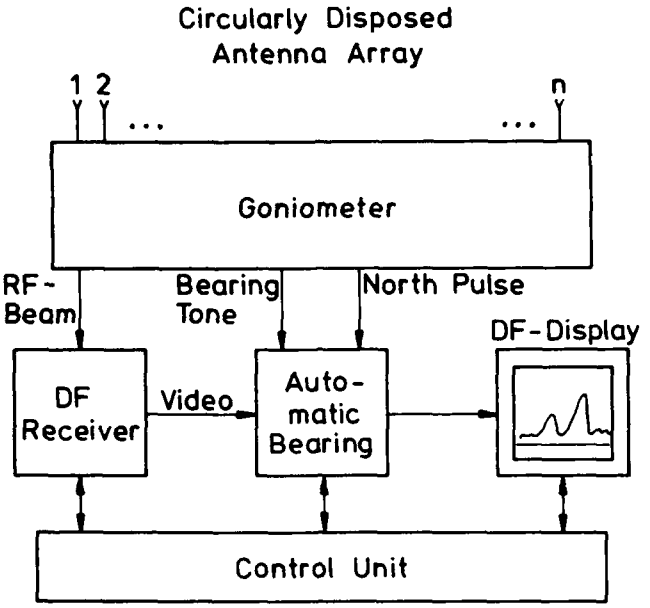


Fig. 4: Block diagram of goniometer direction finder

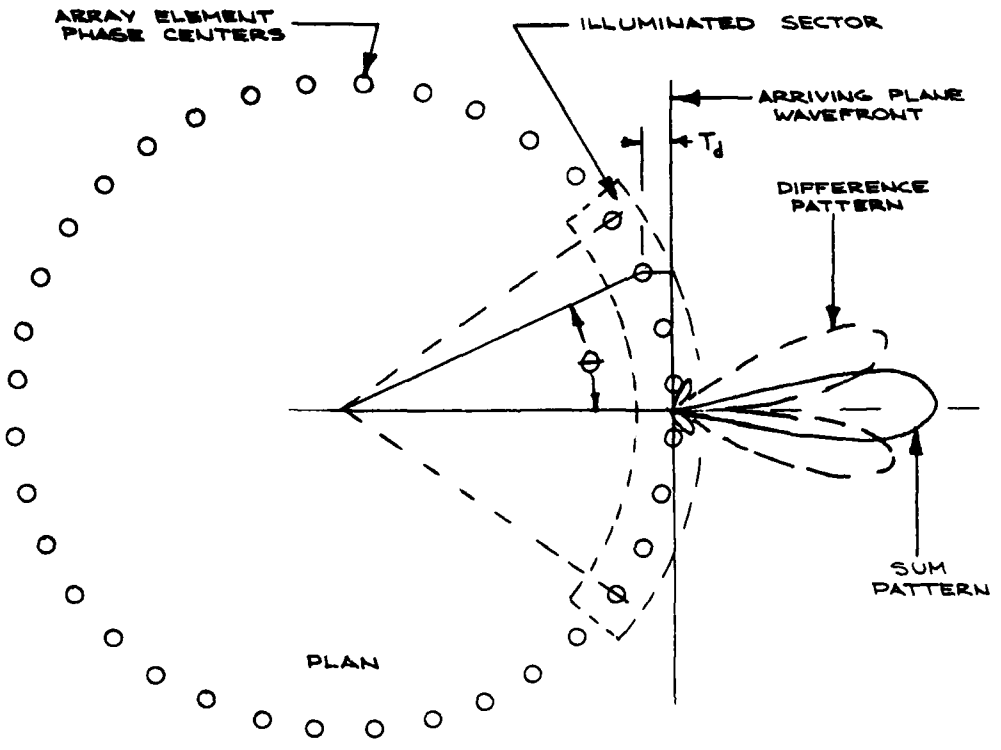


Fig. 5: CDAA geometry and radiation pattern

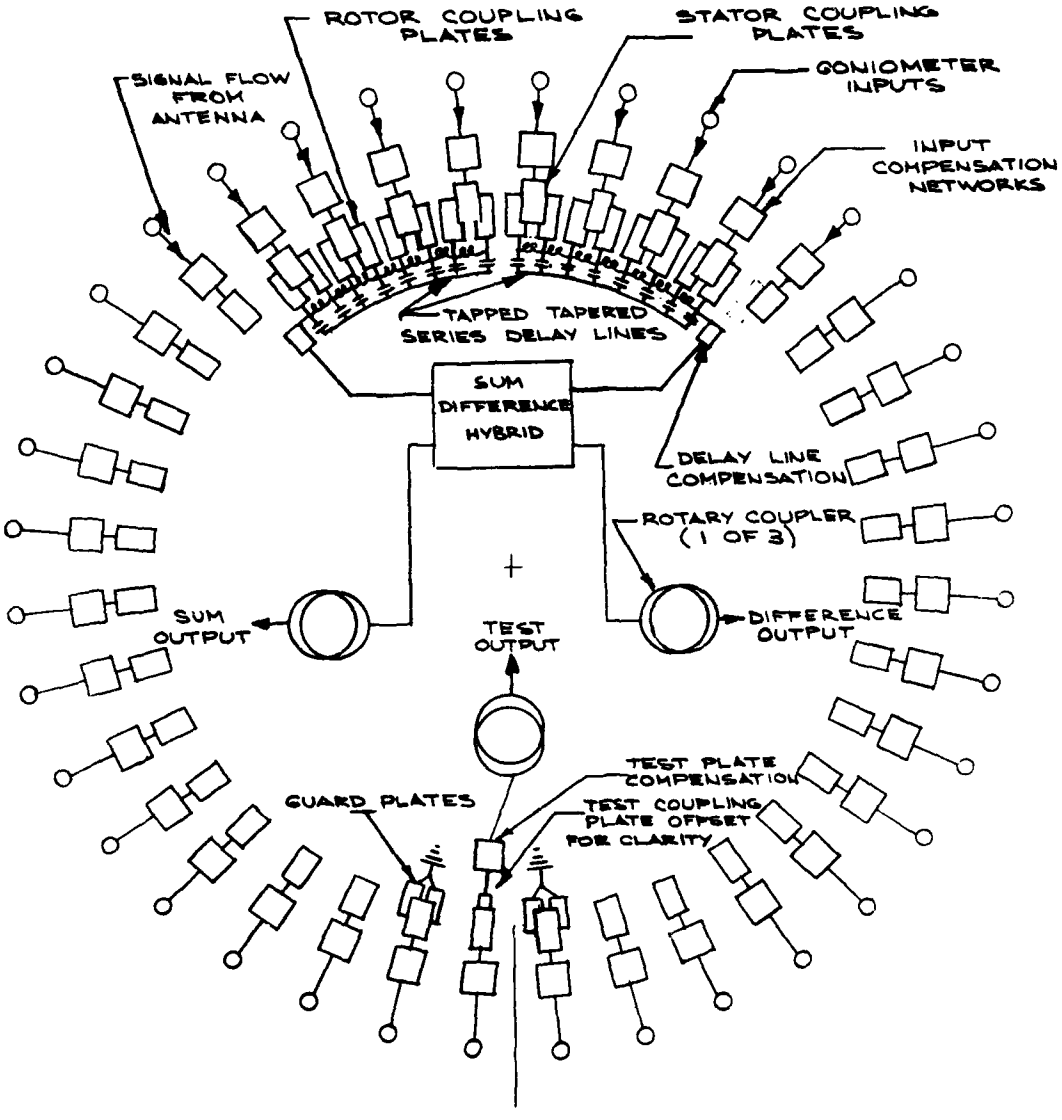


Fig. 6: Simplified diagram of a goniometer RF head

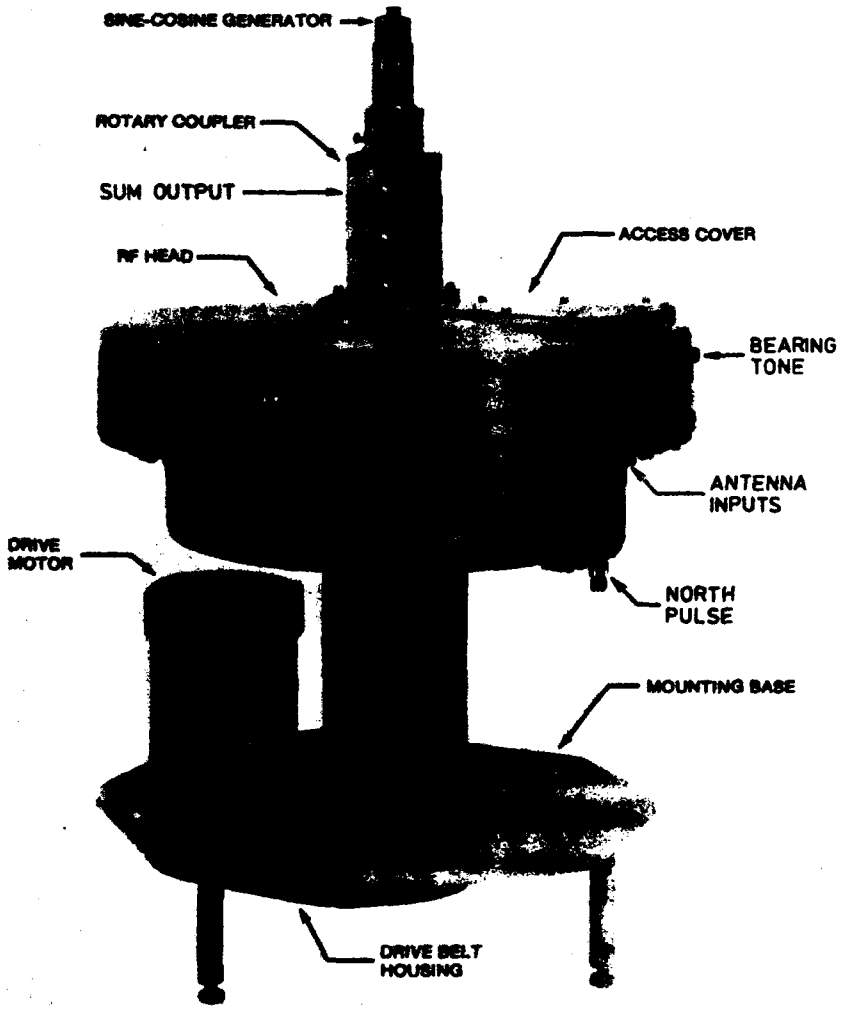


Fig. 7: Goniometer

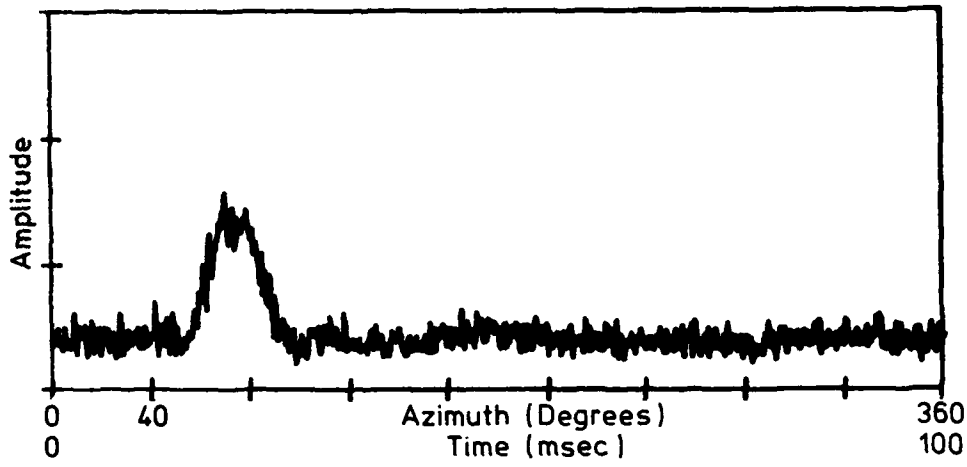


Fig. 8: DF-Signal without radio noise by sparking
RF = 55 MHz, S/N = 10 dB, BW = 30 kHz

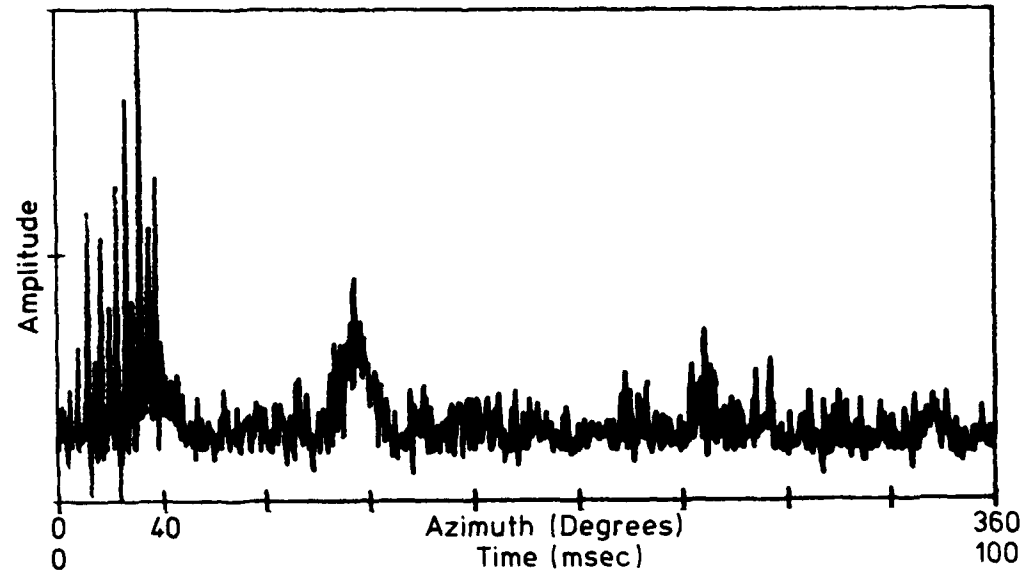


Fig. 9: DF-Signal with radio noise by sparking
RF = 75 MHz, S/N = 5 dB, BW = 30 kHz

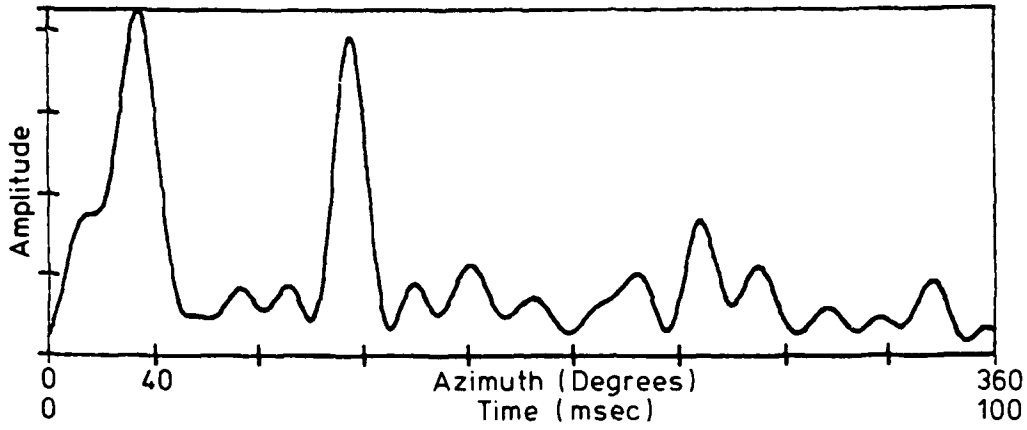


Fig. 10: DF-Signal according to Fig. 9 after optimum low pass filtering

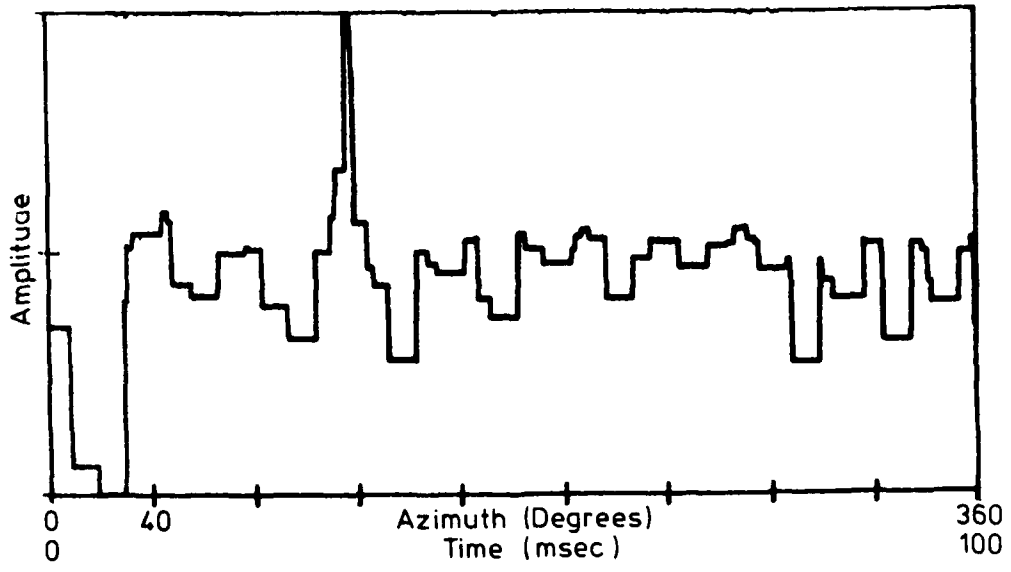


Fig. 11: DF-Signal according to Fig. 9 after determination of the lower envelope

Jamming and Electronic Countermeasures

by

Vijay K. Bhargava and Qiang Wang
Department of Electrical and Computer Engineering
University of Victoria
P.O. Box 1700
Victoria, B.C.
Canada V8W 2Y2

Summary

There is no single jamming strategy that is worst for all spread spectrum systems and there is no single spread spectrum system that is best against all jamming waveforms. In this paper we examine the following potential jammers: (i) Broadband and partial-band noise jammers, (ii) continuous wave and multitone jammers, (iii) Pulse jammers, and (iv) Repeat-back jammers.

In order to operate in the presence of jamming, military communication systems must incorporate one or more of the primary countermeasure techniques in their designs, such as spread spectrum, error control coding, and steerable-null antennas. This paper will introduce the basic design approach for such anti-jam communication systems. Ground rules for system performance are specified by stating assumptions regarding jammers and anti-jam systems along with definitions of the fundamental system parameters. Two useful examples of such systems, namely, coherent direct sequence spread, phase-shift-keying (DS/PSK), and non-coherent frequency-hopped frequency-shift-keying (FH/FSK) systems will be presented, opposed by various jammer waveforms. Because coding and interleaving are extremely important in anti-jam system design, the impact of these techniques is illustrated with examples.

The concepts presented in this paper will be illustrated using JTIDS as an example. JTIDS is an advanced system which provides jam-resistant communication and location. Jamming resistance is achieved using a hybrid DS/FH strategy. The basic channel is 32-ary. The particular coding scheme used is a (31,15) Reed-Solomon code providing error/erasure correction and error detection.

List of Symbols

W_{ss}	=	spread spectrum bandwidth
R_b	=	information bit rate
R_h	=	hop rate
S	=	tatal average signal power
J	=	total average jammer power
J_O	=	equivalent jamming power spectral density
L_{opt}	=	optimum diversity
E_b	=	energy per information bit
E_s	=	energy per symbol
P_b	=	bit error rate
k	=	number of information symbols
n	=	number of coded symbols
k/n	=	code rate
$GF(Q)$	=	Galois Field of Q symbols
N_O	=	one sided power spectral density of additive white Gaussian noise

1 Introduction

Communication warfare has been and will be an element of warfare that pits potential communicators against hostile personnel who seek to intercept and/or disrupt their communications. Hence, in contrast to ordinary communication systems, military communication systems must be designed so as to be able to operate in severe environments. These hostile environments make special techniques crucial for the design of communication systems that are viable. Since the development of covert communication techniques has greatly curtailed the possibility of intercepting and interpreting communications, it seems inevitable that military communications will be forced to operate in a jamming environment. This paper is mainly concerned with jamming and anti-jam communications. It should be noted, however, that to best utilize the available power, the potential jammer must

usually first intercept the communications to obtain the relevent parameters of the signal and perhaps locate the receivers. Of course there may be other ways for a jammer to obtain these parameters, such as stealing or capturing a copy of the transmitter and receiver used. It is always possible that the jammer will have complete *general* knowledge of the communication system so that he can optimize his jamming strategy. This kind of jammer is usually called an intelligent or smart jammer. It is this kind of jamming that we consider in this paper. Due to the rapid development of signal processing technology, the potential intelligence of the jammer is increasing as is the potential existence of a smart jammer.

Another feature in the game of jamming and anti-jam is the competition in generating large signal and jamming powers. However, a more important and intelligent part of the game is how to best utilize these powers (by both the jammer and the communicator).

In this paper we examine the following types of jammers: (i) Broadband and partial-band noise jammers, (ii) continuous wave and multitone jammers, (iii) Pulse jammers, and (iv) Repeat-back jammers. In order to operate in the presence of jamming, communication systems must incorporate one or more of the primary countermeasure techniques in their designs, such as spread spectrum (SS), error control coding, and steerable-null antennas. These techniques will be considered with the emphasis on spread spectrum and error control coding. Since there is no single jamming strategy that is worst for all spread spectrum systems, and there is no single spread spectrum system that is best against all jamming waveforms, we have to consider different spread spectrum systems opposed by various jamming waveforms.

Spread spectrum communications has a relatively long and interesting history [8]. Although the problem has an obvious military flavour, much of the spread spectrum material has appeared in the public domain. A three-volumn book by Simon et al [3] and half of the text by Ziemer and Peterson [15] are devoted to spread spectrum communications. The book by Torrieri [4] covers more areas in military communications. Among the most recent tutorials on spread spectrum communications are those by Blake [5] and Bird and Felstead [2]. Unfortunately, there appears to be little standardization in the literature in terminology, definitions, and performance criteria. This lack of standardization, together with the variety of system assumptions, has led to difficulties in making comparisons. However the working principles are the same. In this paper, we emphasize these fundamental principles.

In the next section, we will introduce basic anti-jam techniques including adaptive antenna systems, spread spectrum and error control coding. In Section 3, different jamming strategies are examined. In Section 4, we evaluate the anti-jam performance of spread spectrum communication systems with the emphasis on frequency hopping systems. In Section 5, an example of a real spread spectrum system is considered.

2 Basic Anti-jam Techniques

In this section, basic anti-jam techniques such as adaptive antenna systems, spread spectrum and error control coding, will be briefly introduced.

2.1 Adaptive Antenna System

An adaptive antenna system functions like a gateman to keep as much interference as possible from entering into the communication receiver. It automatically monitors its output (to the receiver) and adjusts its parameters accordingly. It does so in order to reduce the impact of the interference that enters through the sidelobes, or possibly the mainlobe, of its antenna radiation pattern, while still allowing reception of an intended transmission. The design of an adaptive antenna system requires little *a priori* knowledge of the signal or interference characteristics. The detailed design of an adaptive antenna system depends upon the performance criteria selected. Two of the most useful criteria are the signal-to-noise ratio and the mean-square error.

A classical example of an adaptive antenna system is the sidelobe canceller. It has been shown that the adaptation in a sidelobe canceller can be interpreted not only as noise cancellation, but also as adaptive beam forming and null steering. The sidelobe canceller is a special type of adaptive noise canceller. The adaptive notch filter, which removes periodic interference signals from wideband desired signals, is another variation of the adaptive noise canceller.

An adaptive antenna system sometimes exhibits unintentional cancellation of the desired signal. To preclude such a possibility, the constrained minimum power criterion can be used in the design of an adaptive system. The resulting system automatically limits the cancellation of the desired signal, while still adaptively filtering the interference.

There is a vast amount of literature available on the subject of adaptive antenna systems. Interested readers are refered to [4,6] for details. The net effect of an adaptive antenna system on the performance of communication systems is that the effective jamming power injected into the receiver is reduced.

2.2 Basic Spread Spectrum Techniques

The anti-jam capability of a spread spectrum system, as its name suggests, is due to the spectrum spreading. A basic and important measure of the performance improvement of a spread spectrum system is the *processing gain*. While there is no universally accepted definition of the processing gain, the essential meaning of it remains the same.

Suppose the total bandwidth of a spread spectrum system is W_{ss} and the information bit rate is R_b. Let the total average power of the signal and jamming be S and J, respectively. Define the equivalent jamming power spectral density as $J_O = J/W_{ss}$ and denote the energy per information bit as E_b. Then we have

$$\frac{S}{J} = \frac{E_b R_b}{J_O W_{ss}}$$

and,

$$\frac{E_b}{J_O} = \frac{S}{J} \times \frac{W_{ss}}{R_b} \tag{1}$$

The bit error rate (BER) performance of a spread spectrum system depends on E_b/J_O. Equation 1 says that E_b/J_O is larger than the signal to jamming power ratio S/J by a factor W_{ss}/R_b. For a non-spread spectrum system, W_{ss} is close to R_b. For a spread spectrum system, W_{ss} is much larger than R_b, so E_b/J_O is much larger than S/J. By processing gain we mean this increase in signal-to-noise ratio. In fact, as argued in [3], the factor W_{ss}/R_b itself may be a good definition of the processing gain. At this point, it seems that the jammer is already at a disadvantage in the jamming and anti-jam game due to the large processing gain, but he can still try to maximize the BER, P_b vs. E_b/J_O. It is this function $P_b(E_b/J_O)$ that we evaluate for various jamming and anti-jam strategies.

A general spread spectrum communication system model is shown in Figure 1. In this model, spectrum spreading is realized by the spread spectrum modulator and demodulator. The Channel encoder and decoder implement some forward error control (FEC) scheme. Most FEC schemes are designed to cope with channel errors that are independent from one symbol to another over the code alphabet. In jamming environments, channel transmission errors may occur in bursts. In this case interleaving can be used to distribute channel errors randomly throughout the decoder input sequence. This is accomplished by changing the order in which the coder output symbols are transmitted over the channel. With the transmission order changed, it is unlikely that a burst of contiguous channel errors will affect contiguous coder output symbols. Prior to decoding, the reordering performed by the interleaver is inverted by the de-interleaver. In the de-interleaving process, the burst of channel errors is distributed throughout the decoder input sequence.

In a jamming environment, the channel is nonstationary, and side information is valuable for efficient data reception. It should be noted, however, that a system heavily dependent on side information may not be robust, because it may not always be obtainable or its quality may vary greatly. Side information may be generated by monitoring the channel state. It may also be generated during the data demodulation process. In general, the channel encoder output alphabet and channel decoder input alphabet may be different to allow for soft-decision decoding, which may include the use of side information.

The modulating signal of SS modulation is a pseudorandom (PN) sequence. At the receiving end, SS demodulation is performed to remove the spreading modulation from the received signal. This despreading operation is the key function of any SS system. Despreading can be accomplished only if accurate synchronization information is available. This information is derived in the timing and synchronization hardware. The operations of the spreading waveform (PN sequence) synchronization, symbol synchronization and frame synchronization, are all included in the synchronization block of Figure 1.

There are two basic methods used to spread the spectrum. One is called direct sequence (DS) SS modulation. The other is frequency hopping (FH). Time hopping will not be considered in this paper.

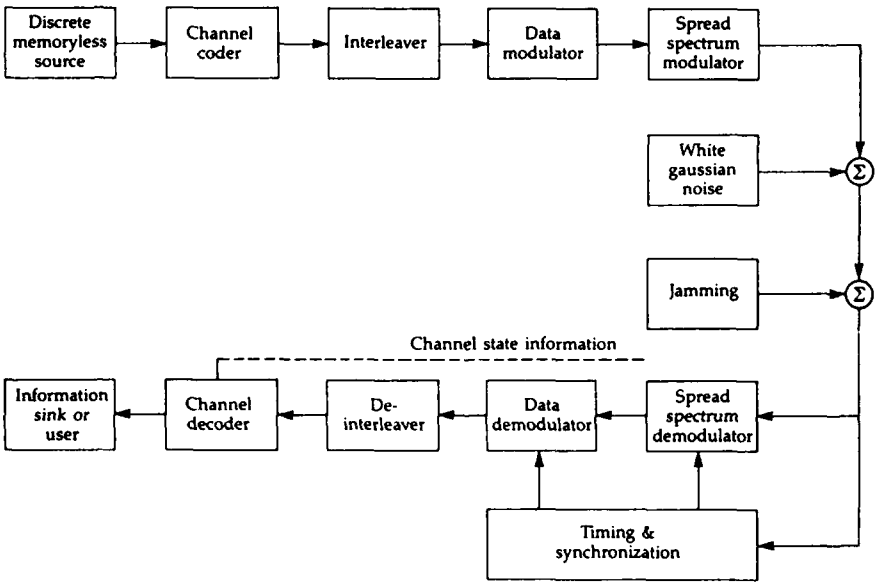


Figure 1: Spread spectrum communication system model.

2.2.1 Direct Sequence Spread Spectrum Modulation

Direct sequence spread spectrum modulation directly modulates the data-modulated signal a second time with a PN sequence at a much higher rate than the information bit rate R_b. It is well known that coherent modulation is superior to noncoherent modulation in noise-resistance, and so, to allow the data modulation to be coherent, the DS SS modulation must be coherent also. Thus DS SS modulation may be any coherent modulation scheme such as BPSK, QPSK, MSK, etc. Since the rate of the PN sequence is much higher than R_b, the bandwidth of the transmitted SS signal, W_{ss}, depends primarily on the PN sequence rate. The exact PN sequence used for SS must be kept strictly confidential (unknown to the jammer), to ensure the effectiveness of the system. Thus the only information available to a jammer is that a very wideband signal is being transmitted. In this way, a large processing gain is realized. The higher the rate of the PN sequence, i.e., the rate of the SS modulation,

the larger is W_{ss}, and thus the larger the achieved anti-jam capability. The high speed of DS modulation and demodulation is a limiting factor on the available SS bandwidth. As an alternative, a frequency hopping system can be considered.

2.2.2 Frequency Hopping

In a frequency hopping SS system, the spectrum spreading is achieved by changing the carrier frequency of the transmitted signal from time to time in accordance with a PN sequence. Frequency hopping can be viewed as wideband FSK over an RF band (transmission band). The hopping frequencies are generated by a frequency synthesizer. Because of the difficulty in maintaining the phase coherence of the different hopping frequencies, there is usually a phase incoherence from one hop to another. For this reason, the data modulation schemes used are different from the coherent data modulation in DS SS systems. In general, a noncoherent modulation scheme such as FSK is used for data modulation in a FH system. If over one hop there is more than one data symbol transmitted, the FH system is called a Slow Frequency Hopping system (SFH); otherwise, the system is called a Fast Frequency Hopping system (FFH). For SFH systems, differentially coherent modulation schemes can also be used for data modulation. Such schemes include DPSK and differential quadrature amplitude-shift-keying (DQASK). Only noncoherent modulation schemes can be used for data modulation in FFH systems. A commonly used system is FFH/MFSK, which we will consider in detail.

2.2.3 Hybrid DS/FH Spread Spectrum

Another method for spectrum spreading is to employ both DS and FH spreading techniques in a hybrid DS/FH system. One reason for using hybrid techniques is that some of the advantages of both systems are combined into a single system. Hybrid techniques are widely used in military SS systems and are currently the only practical way of achieving extremely wide spectrum spreading. Many methods of combining DS and FH spreading are possible. One method is to DS modulate the data modulated signal first and then hop the center frequency of the DS signal. As discussed earlier, because noncoherent frequency hopping is used in a hybrid DS/FH system, the data modulation must be either noncoherent or differentially coherent, the same as in a pure FH system. This partially explains why more attention is given to FH systems over DS systems in this paper.

2.3 Forward Error Control Coding

In error control coding, redundancy is added to the information to be transmitted so that the transmission errors can be detected and/or corrected at the receiving end. Thus the reliability of the communication system is improved. Error control coding has its own history and has been used in many areas. In fact, wherever errors occur error control coding can usually be applied to reduce errors. This point has been verified by the fact that FEC coding has been effectively used to reduce the error rate in many communication systems. In the game of jamming and anti-jam, the jammer tries to cause as many transmission errors as possible, but the communicator can use error control coding as an effective tool to reduce the number of transmission errors.

There are two types of codes used for forward error control. One is block codes, and the other is convolutional codes. A block encoder groups k input information symbols into a block and outputs an n-symbol codeword. The code rate of this (n, k) code is k/n. The Hamming distance between a pair of codewords of a block code is the number of positions where the codewords differ. A measure of error control capability is the minimum distance, d, of a code, which is defined as the minimum Hamming distance between all pairs of codewords in a code. A commonly used class of block codes is the binary BCH codes. These codes have large minimum distances and efficient decoding algorithms are known for them. If a binary BCH code is used over a nonbinary channel in an anti-jam communication system, interleaving must be used to eliminate bursts of errors. Another class of codes that are often used, especially over nonbinary channels, are Reed-Solomon codes. Length $n = Q - 1$ Reed-Solomon (RS) codes are Q-ary codes over $GF(Q)$. For (n, k) RS codes, $d = n - k + 1$. Since RS codes meet the Singleton bound, they are called maximum-distance codes[16].

Over a jammed channel, an important technique to improve the performance of a block code is to use error-erasure-decoding. In this case, an erasure is declared when a unreliable received symbol is sensed (in fact it is likely that it has been jammed). Because the locations of these erasures are known, the decoder can handle more transmission errors.

A simple, yet quite effective block coding scheme often used in anti-jam communications is a repetition code. With this code, an input symbol is transmitted over a channel more than one time. Note that the channel over each transmission should have an independent jamming state, which may be realized with the help of interleaving. Repetition coding is also known as time diversity or simply diversity. An important advantage of this type of coding is that it can be used in concatenation with some other code.

A convolutional code is similar to a block code in that each time k input symbols are collected, the encoder outputs n symbols. In contrast to the block encoder, the mapping of input k-tuples to output n-tuples are not independant from one mapping to the next, that is the convolutional code possesses memory. The amount of memory varies for different codes and is specified by the constraint length of the code (which is as important as the minimum distance of a block code). The constraint length of a convolutional code is usually defined as the number of output n-tuples which are influenced by a particular input k-tuple.

Maximum likelihood decoding of a convolutional code can be accomplished by the Viterbi algorithm for small constraint lengths. When the contraint length increases, the complexity of the usual Viterbi decoder becomes prohibitive. To relieve this limitation, parallel processing techniques have been considered for the implementation of the Viterbi algorithm [9,10].

A clear advantage of a convolutional code over a block code is that the full use of soft-decision information can easily be incorporated into the decoder.

To provide more error correction power, error control codes can be used in concatenation. To concatenate two codes, the information symbols are encoded through two encoders. First the information symbols are encoded in the outer code encoder, the output of which is then encoded a second time in the inner code encoder. At the receiving end, the received coded symbols are first sent into the inner code decoder, the output of which is then sent into the outer code decoder. One unavoidable cost of using concatenated codes is the processing delay.

2.4 Evaluation Methods

For a coded SS system we can, in general, consider three methods to evaluate the BER performance.

2.4.1 Computer Simulation

Monte Carlo simulation is the most universal but most time consuming method to evaluate the BER. For the computing power commonly available at present, simulation for a BER $= 10^{-5}$ is near the limit. Therefore, this brute force method is not suitable for a comprehensive investigation of error control codes. It may be considered to evaluate the BER performance when a system configuration and most parameters have been adopted. Since this method has the highest credibility, it is worth examining possible ways to reduce computation time.

2.4.2 Exact Computation

It is generally difficult and cumbersome to compute the exact performance of coded systems. This is because exact performance computation involves a complicated statistical process of combinatorial feature. Since these results are usually very complex, in most cases they do not yield insights as readily as the closed form expressions such as those derived from the Chernoff union bounds.

2.4.3 Chernoff Union Bound

The Chernoff union bound gives an upper bound for the BER. The computation involved in this method is usually much simpler than the other two methods. Though the general credibility of this method remains controversial, for most systems of concern here it has been shown to provide useful and reliable information[3]. Due to its relative simplicity it is especially suitable for a comprehensive study of various codes. Also, this method may provide a kind of unified approach to evaluating the BER performance. This provides a clear relationship between the system parameters and BER.

3 Jamming Strategies

A number of jamming signals have been considered against SS systems. Some of these are clearly real threats, whereas others are considered only because of their optimal nature. The most benign jammer is the barrage noise jammer. This jammer transmits white Gaussian noise with one-side power spectral density (psd) J_O, as shown in Figure 2a. It is usually assumed that the jammer power spectrum covers exactly the same frequency range as the SS signal. The effect of the barrage noise jammer on the system is simply to increase the Gaussian noise level at the output of the receiver down-converter. Actually, barrage noise jamming is difficult to generate due to the large bandwidth of the SS signal. It is also possible that a barrage noise jammer will also jam its own communications. Thus it is not a good choice, but it does provide a reference system from which others can be compared.

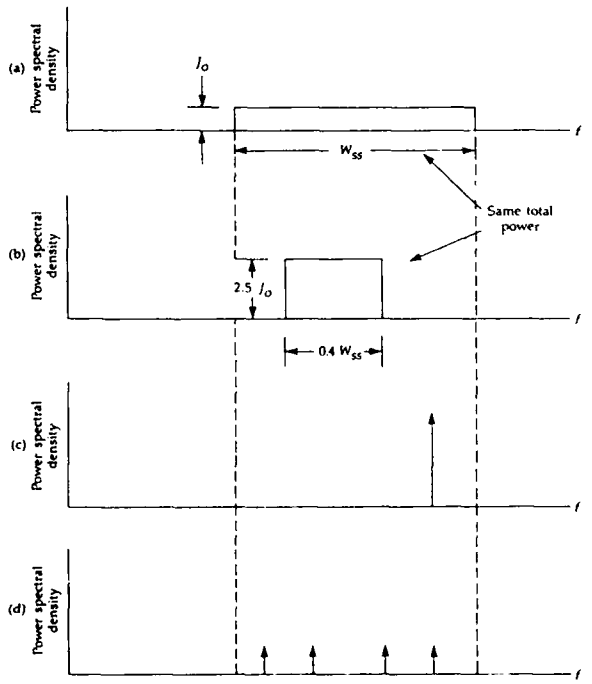


Figure 2: Typical jammer one-sided power spectral densities: (a) barrage noise jammer, (b) partial band noise jammer, (c) single-tone jammer, and (d) multiple-tone jammer.

When the SS modulation has a frequency-hop component, jamming power can be more efficiently used by transmitting all the available power in a limited bandwidth which is smaller than the SS signal bandwidth. A jammer using this strategy is called a partial band jammer and the fraction of the SS signal bandwidth which is jammed is denoted by ρ. Over a jammed band of width ρW, the partial band jammer psd is J_O/ρ, where J_O is the equivalent wideband jammer psd. Illustrated in Figure 2b is a partial band jammer for the case $\rho = 0.4$. The partial band jammer is particularly effective against a FH SS system because the signal will hop in and out of the jamming band and can be seriously degraded when in the jamming band. It can be shown that there is an optimum (for the jammer) ρ which is a function of E_b/J_O.

A third type of jammer is the single-tone jammer. The single-tone jammer transmits an unmodulated carrier with power J somewhere in the SS signal bandwidth. The one-sided power spectrum of this jamming signal is shown in Figure 2c. The single-tone jammer is important because the jamming signal is easy to generate and is rather effective against DS SS systems. In this case, analysis shows that the jammer should place the tone at the center of the SS signal bandwidth to achieve maximum effectiveness. The single-tone jammer is somewhat less effective against a FH system since the FH instantaneous bandwidth is small, and for large processing gains the probability of being jammed on any one hop is minimal. For FH systems, a better tone jamming strategy is to use several tones which share the power of the single-tone jammer. A jammer using this technique is called a multitone jammer. The one-sided power spectrum of a typical jammer is shown in Figure 2d for a four-tone jammer. Multitone jammers will be further discussed in conjunction with FH/MFSK. It should be noted that a multitone jammer is also effective against hybrid DS/FH systems.

Another technique for concentrating the jamming power is to pulse the jammer on for only a fraction of the time. This jamming philosophy is the same as for partial band and multitone jamming. Specifically, the jammer turns on with just sufficient power to degrade SS performance significantly so that the average performance degradation is large. The pulsed noise jammer transmits a pulsed Gaussian noise whose psd just covers the SS system bandwidth W_{ss}. The duty factor for the jammer is the fraction of time during which the jammer is on and is denoted by ρ (the same as for the partial band jammer). Figure 3a and b compare the transmitted waveforms of the full-time barrage noise jammer and the pulsed noise jammer. The pulse duty factor used in Figure 3b is $\rho = 0.5$ and has an RMS amplitude $\sqrt{2}$ larger than the RMS amplitude of the full-time jammer. It is possible to use partial band techniques simultaneously with pulse techniques, although they are usually considered independantly. Note that pulsed jamming assumes the jammer final power amplifier is average power limited rather than peak power limited. In practice, the maximum peak power limits the smallest possible duty factor.

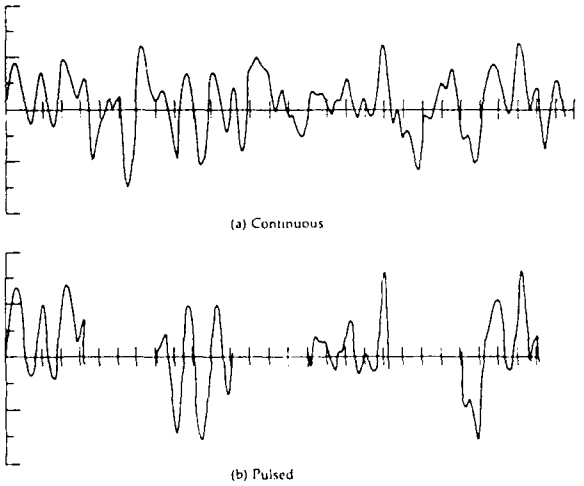


Figure 3: Typical waveforms for continuous and pulsed noise jammer.

Another potential threat is the repeat-back jammer. It first estimates parameters from the intercepted SS signal and then transmits a distorted version of the signal at high power. Such jammers are primarily effective against FH systems when the hop rate is slow enough for the repeat-back jammer to respond within the hop duration. This type of jammer can be neutralized by increasing the hop rate or independantly hopping each tone in the case of MFSK signals. In this case such jammers are sometimes refered to as "frequency-following jammers". Repeat-back jammers will not be discussed further, but it should be emphasized that they are a threat against only FH SS systems with a low hop rate. Generally the effectiveness of a repeat-back jammer depends on the hop rate and the distances between the transmitter, receiver and jammer.

Apart from jamming, there always exists receiver thermal noise which is typically modelled as additive white Gaussian noise with single-sided spectral density N_O. With this receiver noise taken into account, the BER should have the form $P_b(E_b/J_O, E_b/N_O)$. Generally speaking, in this case, the optimum communication system design is one somewhere between a design against jamming and one against thermal noise. However, thermal noise is usually neglected in the analysis due to the following consideration. Since the jamming power and waveform are under the jammer's control, E_b is forced to be large in order to obtain the required P_b. This large E_b will make the contribution to P_b of E_b/N_O very small. Thus P_b will be mainly determined by E_b/J_O and the thermal noise can be neglected.

4 Anti-jam Performance of Spread Spectrum Systems

In this section, we illustrate the anti-jam performance of SS systems by considering two examples. First we examine an uncoded DS/BPSK system under pulsed noise jamming. With this example, we show that a worst case jammer can degrade the exponential dependance of BER on E_b/J_O to an inverse linear relationship. This severe degradation necessitates the use of error control coding. Then we will consider a coded FFH/MFSK system to show the effectiveness of various error control codes against two worst case jammers, a partial band noise jammer and a multitone jammer. Through this example, we will elaborate various design parameters in more detail.

4.1 An Uncoded DS/BPSK System Under Worst Case Pulsed Jamming

Consider a DS/BPSK system where the SS modulation is direct sequence modulation and data modulation is BPSK. Suppose first that the system is opposed by a full-time barrage noise jammer with psd J_O. It can be shown that after despreading, the psd of the noise at the input to the data demodulator is still appoximately J_O [15]. Thus, under a barrage noise jammer, with thermal noise neglected, the BER is given by

$$P_b = Q(\sqrt{2E_b/J_O}) \tag{2}$$

where the Q-function is defined as

$$Q(x) = \int_0^x \frac{1}{\sqrt{2\pi}} exp(-t^2/2)dt. \tag{3}$$

Using an approximation of this Q-function, for reasonably large E_b/J_O, Equation 3 is approximated by

$$P_b \approx \frac{1}{\sqrt{4\pi E_b/J_O}} exp(-E_b/J_O) \tag{4}$$

Now suppose the jammer pulses its jamming waveform with a duty factor ρ against the communicator. Corresponding to Equation 3, we now have

$$P_b = \rho Q(\sqrt{2\rho E_b/J_O}) \tag{5}$$

which, similar to Equation 4, can be approximated as

$$P_b \approx \frac{\rho}{\sqrt{4\pi\rho E_b/J_O}} exp(-\rho E_b/J_O) \tag{6}$$

The maximum of this function is achieved when $\rho = \frac{0.5}{E_b/J_O}$, for which the maximum P_b is given by

$$P_b \approx \frac{1}{\sqrt{2\pi e}} \frac{1}{2E_b/J_O} \tag{7}$$

where the assumption is made that $E_b/J_O \geq 0.5$ (which is slightly restrictive). Observe that the exponential dependence of BER on E_b/J_O in Equation 4 has been replaced by an inverse linear relationship in Equation 7. This is a characteristic of all worst case jammers against uncoded systems. Equations 3 and 7 are plotted in Figure 4, where it can be seen that the pulsed noise jammer causes a degradation of approximately 31.5 dB at a BER of 10^{-5} relative to full-time barrage noise jamming. This large degradation is achieved by jamming only a small fraction of the signal. The above results also apply to partial band noise jamming and multitone jamming against FH systems. In general, under worst case jamming, most transmitted symbols will not be jammed and are likely to be received correctly. A very small number of transmitted symbols will be jammed with large interference and they are very likely to be in error. That is, there are a large number of symbols of very good quality and a small number of symbols of very poor quality. This implies two things. First, error correcting codes can be used to reduce the jamming effectiveness, because they can correct a small number of errors quite effectively. Second, information about whether a transmitted symbol is jammed or not can often be easily derived, and this is very useful because if a symbol is jammed it can simply be assumed to be in error. These countermeasures are discussed only in conjuction with a FH system below, but they are equally effective in other systems as well.

4.2 A Coded FFH/MFSK System Under Worst Case Jamming

The system considered employs the fast frequency-hopping (FFH) noncoherent M-ary frequency-shift-keying (NCFSK) technique. We will show the performance of various known error correcting codes in such a system under different kinds of jamming. Although one set of assumptions are made, the results shown are also valid under other assumptions.

4.2.1 System Model

The system model shown in Figure 5 is similar to that under consideration and very typical. System assumptions are as follows.

Signalling and Detection

The transmitted signals are MFSK orthogonal signals which hop over a total spread spectrum bandwith W_{ss}. Noncoherent soft energy detection (square-law reception) of each hop is assumed without quantization. In practice, this can be approximated with finite level quantization. Note that this soft decision is for energy detection. While the soft energy decision may be passed to a soft decision decoder such as that for a convolutional code, the decision made after diversity combination can be a hard decision such as that for hard decision decoding

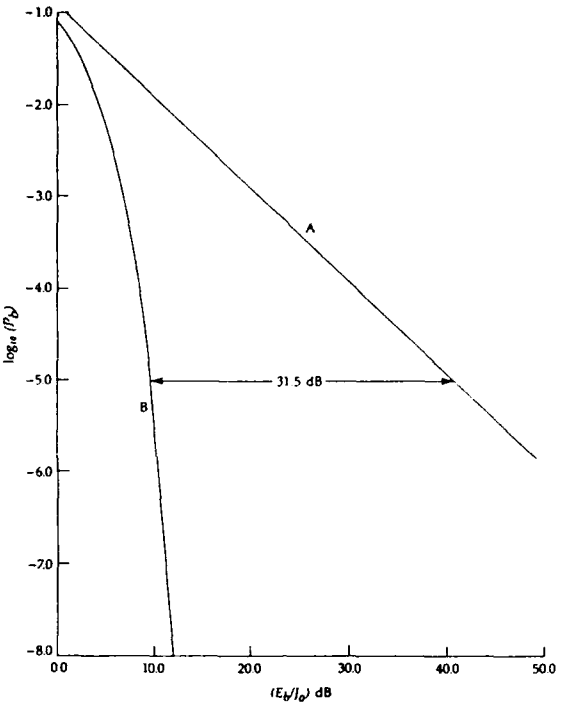


Figure 4: Bit error probability: (A) worst-case pulse noise jammer, and (B) continuous noise jammer.

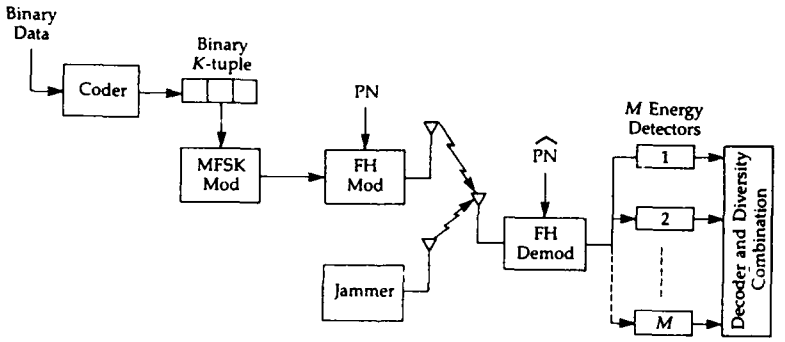


Figure 5: Block diagram of a FH/MFSK system under jamming. Transmit one of $M = 2^K$ tones; carrier is hopped in the pattern determined by the PN code; dehopping requires derived PN reference ($\widehat{\text{PN}}$); non-coherent detection; diversity combination is included in the decoder.

of a block code.

Diversity, Combination and Side Information

By diversity we mean transmit one coded M-ary symbol in L hops. For a given M and error correcting (EC) code, when the energy per symbol E_s is fixed, the energy per bit E_b is fixed. In this case there is usually an optimum L, denoted as L_{opt}, at which the final bit error rate (BER) can be minimized for a given signal to noise ratio. This may not be true when the hop rate R_h is fixed and the information bit rate R_b is varied. However, since the objective here is to show the effectiveness of various EC codes, we assume E_b is fixed. This will help identify effective EC codes. It appears that an effective code can work well in both situations. In analysis, L_{opt} provides an indication of how efficient a specific code is. A smaller L_{opt} implies that the code is more efficient against jamming.

As mentioned earlier, in a jamming environment, side information is valuable for efficient data reception. One method used in practice to derive the side information is to implement automatic gain control (AGC) in the receiver, which may be monitored to determine whether jamming power is corrupting a given hop[1]. Based on this implementation, we may assume that the receiver knows with certainty whether each hop of an M-ary symbol is jammed or not. If any of the L hops is not jammed, an error free M-ary decision can be made (the thermal noise is neglected, as discussed previously); otherwise select the largest of the linear combinations (direct sums) of the energy of L hops. Note that the assumed perfect side information is for the diversity combination.

4.2.2 Types of Jamming

For reasons discussed in Section 3, we neglect the received thermal or non-hostile background noise. In this section we consider two types of worst case (WC) intelligent but non-repeat-back jammers, namely, partial band noise and multitone interference.

Partial Band Noise Jamming

Recall that partial band noise (PBN) jamming occurs when the total jamming power is restricted to a fraction ρ ($0 < \rho \leq 1$) of the full spread spectrum bandwidth. It is equivalent to pulse jamming for frequency hopping systems, and in this case ρ is the fraction of time that the jamming is on. ρ is a parameter that the jammer can optimize against the communicator.

Multitone Jamming

Multitone jamming (MT) includes band multitone jamming and independent multitone jamming. It has been shown that the worst case multitone jamming is band multitone jamming with a single jamming tone per jammed band[2]. We consider only this worst case multitone jamming. In this case the jammer has one parameter to optimize, namely the ratio of signal power of one hop to the power of the jamming tone, denoted as α.

4.2.3 Fading and Uniformity

We do not consider fading of the signal due to propagation, but in worst case jamming the signal already suffers from a kind of fading. Thus results without considering propagation fading may be indicative of the case of propagation fading. This point has been verified in previous work. For instance, the broadband jammer is the worst noise jammer in a Rayleigh fading channel[3].

A satellite communication link over frequency bands presently in use can usually be viewed as a uniform channel over W_{ss}.

4.2.4 Performance of EC Codes in a FH system

Although performance of error-correcting codes in FH systems has been widely studied, there is no single reference in the literature providing complete information on the BER performance (rather than other criteria, such as the cutoff rate) of various codes for jammed M-ary NCFSK, and hence a convincing comparison of them. Ma and Poole's paper[7] and the book by Simon et al.[3] may be the most comprehensive. Only partial band noise jamming was considered in [7] and only Reed-Solomon codes and several convolutional codes were considered in [3].

Under the assumptions discussed above, the BER performance of a FH/MFSK system with various EC codes can be obtained under worst case (WC) jamming. The evaluation method used here is based on the Chernoff union bound, as in [7,3]. We present only the final results here, as details can be found in [3] and, in a more compact form, in [11]. All BER performances are given under worst case jamming and with optimum diversity.

In anti-jam communications, a good code should perform well under any kind of jamming. Thus good codes are those with the best BER performance for the most effective type of jamming, WC MT jamming or WC PBN jamming, at low BER. We only consider M-ary ($M = 2^K$) signalling for K up to 5.

Clearly, the final BER, P_b, depends on the signal to noise ratio, the EC code, M (or K) and the jamming strategies. The worst case ρ_{wc} for partial band jamming, and the worst case α_{wc} for multitone jamming depend on the signal to noise ratio, the EC code, and M (or K). The optimum diversity factor L_{opt} depends on the signal to noise ratio, the EC code, M (or K) and the jamming strategies. It is implied that a good design is a good combination of MFSK, an EC code and diversity with all possible jamming waveforms taken into account.

We can consider three kinds of codes: convolutional codes, block codes and cancatenated codes. Figure 6 shows the performance of two rate 1/2 convolutional codes discovered by Odenwalder and Trumpis, respectively, over Binary FSK and 4-ary FSK respectively. The constraint length of both codes is 7. Soft decision Viterbi decoding is assumed for these convolutional codes. Figure 7 shows the performance of two Reed-Solomon codes with bounded-distance decoding over 4-ary FH/FSK and 8-ary FH/FSK. In all cases, worst case jamming and optimum diversity are assumed.

From these figures, it seems that the convolutional codes are more powerful than Reed-Solomon codes, but the performance of Reed-Solomon codes can be improved with the use of error-erasure-correction decoding.

5 Joint Tactical Information Distribution System

The Joint Tactical Information Distribution System (JTIDS) is a tactical military spread spectrum radio network. It has been designed to provide jam resistant communications and location, and is being developed with support from all US military services and NATO. Most of the information transferred is expected to be digital data. A secure, jam protected digitized voice capability is provided as a tradeoff for a number of data channels (on the basis of equivalent total bits/second required). As shown in Figure 8, jamming resistance and low probability of intercept are achieved using a hybrid DS/FH transmission strategy.

All JTIDS carrier frequencies are within the frequency bands 969 to 1008 MHZ and 1113 to 1206 MHZ [15]. The choice of carrier frequencies in these bands limits transmitted energy to roughly the band 960 to 1215 MHZ and avoids interference with the IFF frequencies of 1030 and 1090 MHZ. Within this band, about 20 JTIDS nets can be operated simultaneously in the same geographical area (without degrading AJ characteristics), by using code-division techniques. Thus JTIDS is a spread spectrum multiple access system.

With reference to Figure 1, the basic channel is a 32-ary channel since each of the 32 phases of the minimum-shift-keyed spreading code burst is transmitted. The direct sequence spreading chip rate is 5×10^6 chips/second. A 5 bit message is transmitted with each burst by associating each message with a different phase

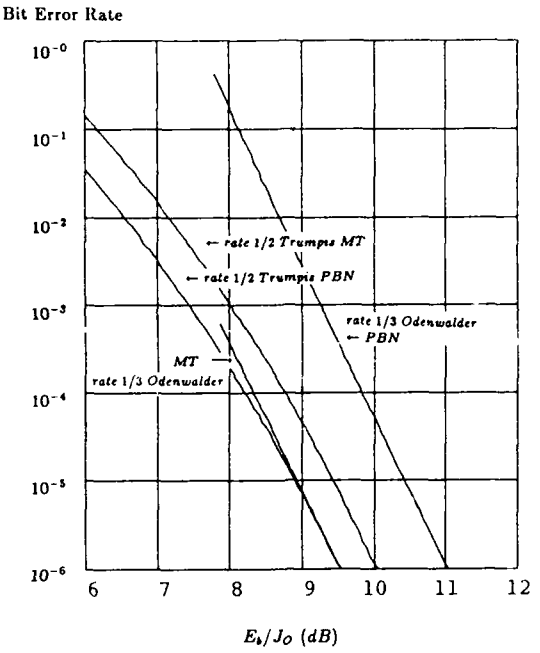


Figure 6: BER performance of rate 1/3 binary Odenwalder convolutional code over Binary FH/FSK and rate 1/2 Trumpis code over 4-ary FH/FSK. The constraint length is 7.

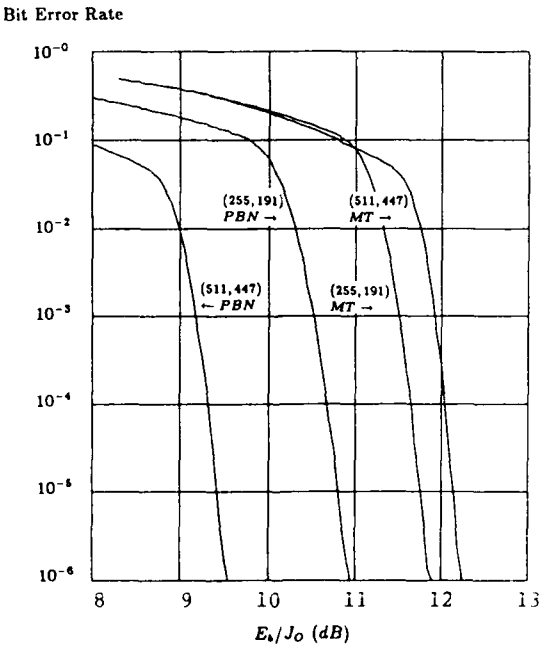


Figure 7: BER performance of the (255,191) RS code over 4-ary FH/FSK and (511,447) RS code over 8-ary FH/FSK. Worst case jamming and optimum diversity are assumed.

of the 32 bit direct sequence spreading code burst. Frequency hopping is implemented by hopping the carrier to a new frequency for each transmitted pulse. Error control coding is used by all JTIDS members. The particular code used is a n=31, k=15 Reed-Solomon block code where each code symbol consists of five bits. This code provides both error correction and detection.

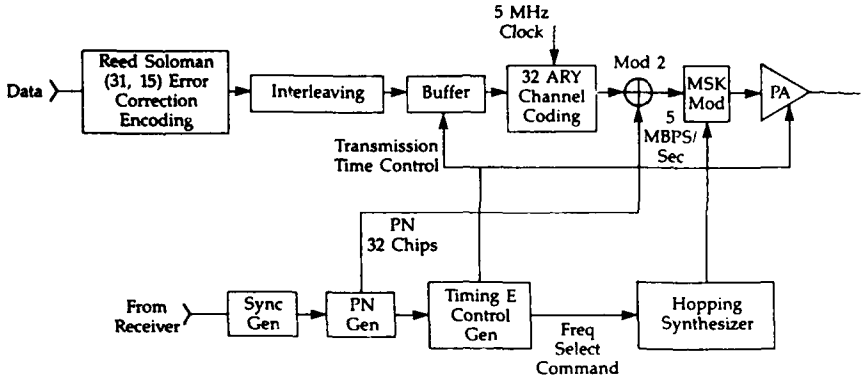


Figure 8: JTIDS transmitter functional block diagram.

6 Conclusion

In this paper we have surveyed various jamming and electronic countermeasure techniques. Spread spectrum techniques were examined in detail with an emphasis on frequency hopping systems.

JTIDS was introduced as an example of a real system which uses many of the concepts discussed within this paper.

Acknowledgement

The authors would like to thank Dr. E. Barry Felstead and Mr. T. Aaron Gulliver for helpful discussions.

References

[1] Viterbi, A. J. and Jacobs, I. M.: 'Advances in Coding and Modulation for Noncoherent Channels Affected by Fading, Partial Band, and Multiple Access Interference', in *Advances in Communication Systems*, Vol. 4, New York: Academic Press, 1975, pp 279-308.

[2] Bird, J. S. and Felstead, E. B.: 'Antijam Performance of Fast Frequency-Hopped M-ary NCFSK – An Overview", *IEEE J. on Selected Areas on Commun.*, Vol. SAC-4, No. 2, March 1986, pp 216-233.

[3] Simon, M. K., Omura, J. K., Scholtz, R. A. and Levitt, B. K.: *Spread Spectrum Communications, Vol. I - III*, Rockville, Maryland: Computer Science Press, 1985.

[4] Torrieri, D. J.: *Principles of Military Communication Systems*, Dedham, Maryland: Artech House, 1981.

[5] Blake, I. F. and Poor, H. V., Ed.: *Communications and Networks*, New York: Springer-Verlag, 1986, pp 246-265.

[6] Steinberg, B. D.: *Principles of Aperture and Array System design*, New York: Wiley, 1976.

[7] Ma, H. H. and Poole, M. A.: 'Error Correcting Codes Against the Worst-Case Partial-Band Jammer', *IEEE Trans. on Commun.*, Vol. COM-32, No. 2, February 1984, pp 124-133.

[8] Scholtz, R. A.: 'The Origins of Spread Spectrum Communications', *IEEE Trans. on Commun.*, Vol. COM-30, pp 822-854, 1982.

[9] Kailath, T.: 'VLSI Array Processors for Communications, Control and Signal Processing', *Proceedings of IEEE Tencon 87*, Seoul, Korea, Vol. I, pp 96-98, August 1987.

[10] Yuen, J. H.: 'New Codes for Deep Space Communications', *Proceedings of ICC'87*, Seattle, U.S.A., pp 20.7.1-20.7.5, June 1987.

[11] Wang, Q., Gulliver, T. A., Bhargava, V. K. and Little, W. D.: 'Coding for Frequency Hopped Spread Spectrum Satellite Communications', *Annual Report prepared for the Department of Communications of Canada under DSS contract No. 27ST. 36001-6-3539*, April, 1987.

[12] Odenwalder, J. P.: 'Optimal Decoding of Convolutional Codes', Ph.D. Dissertation, University of California, Los Angeles, 1970.

[13] Trumpis, B. D.: 'Convolutional Coding for M-ary Channels', Ph.D. Dissertation, University of California, Los Angeles, 1975.

[14] Viterbi, A. J. and Omura, J. K.: *Principles of Digital Communications and Coding*, New York: McGraw-Hill, 1979.

[15] Ziemer, R. E. and Peterson, R. L.: *Digital Communications and Spread Spectrum Systems*, New York: Macmillan, 1985.

[16] Blahut, R. E.: *Theory and Practice of Error Control Codes*, Addison-Wesley, 1983.

ECM RESISTANCE IN THE HF BAND BY USE OF ADAPTIVE REACTION AND FREQUENCY HOPPING

by

G.Greiner, P.Iselt and G.Müller
Rhode & Schwartz GmbH
Postfach 80.14.69
D-8000 Munich 80
Federal Republic of Germany

Abstract

In shortwave communications, change of frequency either by adaptive reaction or by frequency hopping as a means of protection against jamming must be adapted to the particular transmission characteristics of shortwaves. Adaptive reaction to a new frequency is a technique used for maximizing the throughput and optimizing reliability, and is an Electronic Counter Counter Measure (ECCM) against conventional jamming systems. As a result of the analysis and reaction times of follower jammers, frequency hoppers with 10 to 20 hops/s can provide adequate protection against ECM also in the 90's. In this case, effective error correction techniques are required because of the high proportion of disturbed hop channels. Hopping systems with hop rates of more than 100 or 1000 hops/s are discussed. An HF communication system with adaptive reaction and frequency hopping of up to 20 hops/s is presented. Results are presented of field trials with adaptive reaction and frequency hopping with a dwell time of 160 ms corresponding to 6 hops/s. A test system with reproducible channel characteristics of the HF channel is described.

1. Introduction

The use of highly integrated low-cost components opened up new ways in automation and computer control; the importance of communication via the ionosphere is continually on the increase. (1,2,3)

Shortwave communications are characterized by the following features:

+ Channel quality and useful bandwidth variable in time
+ Shortage of available frequency due to large number of users
+ High interference potential as a result of multiple assignment of frequencies and worldwide transmission
+ Complex operation of system

These features call for the following requirements which up-to-date shortwave communication systems fulfill:

1. Reliable transmission even on poor links (4)
2. Maximum throughput matched to the transmission channel
3. Automated and thus simple and reliable operation (5,6,7)
4. Operability even in the face of Electronic Counter Measure (ECM) (8)

The increasing importance of shortwave communications makes it necessary that this communication medium, too, is protected against jamming (ECM). From a general analysis of threat scenario, follower jammers gain a special importance over broadband jammers and will therefore be examined in the following.

2. ECM scenario

EC measures can only be effective if three prerequisites are fulfilled:

1. An interfering link must be intercepted
2. The link must be identified
3. The link must be disturbed

2.1 Interception

Dense occupancy of the shortwave spectrum requires the use of narrowband detectors with excellent control characteristics so that the time of detection is lengthened.

Conventional detection equipment in the case of which the operator searches a given frequency band and jams particular frequencies are replaced by state-of-the-art automatic equipment which, in conjunction with fast and automatic receivers, are able to scan wide frequency bands in a short time. The development is towards powerful multichannel receivers and, in the future, towards very fast FFT analyzers; these units will be able to analyze bandwidths of, for example, 1 MHz at a resolution of 1000 channels in a short time, i.e. in 1 ms minimum as in the present example.

2.2 Identification

When a signal is detected in the shortwave spectrum, the signal has to be identified whether it is meant to be interfered. The identification can be made by means of electronic fingerprinting or by additional results of direction finding. Electronic fingerprinting would, for example, be the determination of system-specific transmit spectra, particular synchronisation processes or characteristic modulation response. For this purpose, time-consuming analysis methods are, however, required.

For the identification of a transmitter, it is sufficient to determine its geographical location. Adequate accuracy of the latter is possible with conventional DF techniques in the available time and using sky waves only at a high level of equipment investment. In the development of fast jamming systems, the identification of a signal places high demands on the processing speed of the evaluation unit and is thus the limiting factor in the performance of a follower jammer.

2.3 Jamming

The effectiveness of a follower jammer depends on its output power, antenna gain, bandwidth and/or the modulation type of the jamming signal and the time during which a signal can be jammed. The hardware requirement for a fast, powerful jammer is not the limiting factor of the total system but more important are questions relating to the geographical location of the jamming system, the latter's influence on the transmit power and the antenna configurations, and side effects of jamming such as interference on interception or on own HF communications.

3. ECCM

3.1 Protection against broadband jamming

An FSK signal can usefully be received if its S/N ratio is in the order of 10 dB. The intentional interference of a broadband jammer reducing the S/N ratio can be encountered only by an increase of the process gain since the transmitter power is generally limited.

Process gain is enhanced by increasing the transmission bandwidth (9).
The transmission bandwidth can be increased by direct sequence modulation (Spread Spectrum) or by frequency hopping (FH). In the shortwave range, the frequency hopping technique is preferred to direct sequence modulation because of the dense occupancy by a variety of signal types and the problem of group delay (10,11), the upward interoperability and co-location problems.
The process gain G of the frequency hopping technique is proportional to the number of used frequencies N

$$G \approx N$$

The attainable gain is 27 dB for 500 free frequencies.

The following example is presented for assessing the effectiveness of frequency hopping versus broadband jammers:

Transmitter: 100 W, FSK modulation, antenna gain 3 dB, radiated power 200 W
Jammer: 1 kW, CW modulation, antenna gain 10 dB, radiated power 10 kW
Reception: Useful/signal/interfering signal = 200 W/10 kW = 0.02 = -17 dB
Required S/N ratio: 10 dB
Required process gain: 17 dB + 10 dB = 27 dB

The required process gain of 27 dB can be obtained by a frequency hopping system which uses 500 frequencies. The example illustrates that broadband jammers can be effectively encountered in a simple way.

3.2 Hop rate

Due to the finite velocity of propagation of electromagnetic waves and the reaction time t(react) required for interception, identification and jamming, a signal cannot be disturbed if its duration t(hop) is subjected to the following condition:

$$t(hop) \leq d2/c + d3/c + t(jam) - d1/c$$

c = velocity of light, t(jam) = time of interception and identification of signal and setting time of jammer, for d1, d2, d3, see Fig. 1.

With t(jam) = 0, this leads to the known representation of groundwave links (the problem is three-dimensional for sky-wave links) according to which the jammer must be located within an ellipse with the major axis d2 + d3 (Fig. 1).

For the path transmitter - reciever of d1 = 200 km and transmitter - jammer - receiver

of d2 + d3 = 280 km, the difference in time of travel is 0.26 ms; for d2 + d3 = 1000 km, the difference in time of travel is 2.6 ms.

If the follower jammer is driven by a classical, fast search receiver that searches signals step-by-step on different channels and provided that the receiver is intelligent enough to search only the hop channels (number of channels m), the percentage (j) of channels found for a hopping time of t(hop) and reaction time t(react) per channel is

$$j \approx t(hop)/mt(react).$$

Figure 2 shows the relationship between hop rate r and reaction time t(react) at a channel number of m = 100, m = 500, m = 1000 and a relative number of channels found j = 10 %, which, in case of jamming, can still be corrected effectively by error correction.

It is shown that for a given reaction time t(react) the immunity to jamming of a frequency hopping system can be increased either by a higher hop rate or by increasing the number of channels. In practice, hop bandwidths in the shortwave range lie between 1 and 3 MHz for short to medium distances. A hop bandwidth of 1.5 kHz can be split up into, for example, 500 channels each with a bandwidth of 3 kHz, latter being the typical value for the shortwave range.

A hopping system with a hop rate of 20 hops/s can be disturbed by jammers using single-channel receivers according to Fig. 2 only if the reaction time per channel is less than 1 ms. Detection systems of this kind cannot, however, be realized by single-channel receivers and are the equipment of the next generation.

The reaction time includes the interception time as well as the identification time. Even if, in the future, the interception time falls significantly below 1 ms, the computer remains the limiting factor of jammers as far as identification in the shortwave range is concerned, e.g. if a receive system were capable of analyzing 1000 channels each with 1 kHz bandwidth in 1 ms, a data flow of 1 Mbyte/s has to be processed in real time assuming that the information required for identification is 1 byte. For this purpose, a mainframe computer would be required.

Due to the computing time required, the reaction time of a follower jammer is not expected to drop below 30 ms for some time to come. Signal travel times lie substantially below the reaction time of jammers and can therefore be ignored.

3.3 Hop rates

3.3.1 Adaptive reaction (0.1 to 1 hop/s)

The adaptive reaction (change of frequency on jamming, adaptive power matching to increase interception immunity) is the 1st ECCM step which offers an effective protection against conventional jammers when using a large number of channels. Systems of this type afford optimization regarding reliability and speed of transmission, and offer a certain ECM protection and are indispensable for normal radio operation.

3.3.2 Slow hopping (3 to 10 hops/s)

As shown above, slow hopping at a sufficient number of channels proves to be an effective ECCM measure. If the dwell time per channel is 100 ms at a hop rate of 3 hops/s, then this has the effect of 10 hops/s on the intercepting adversary. The enhanced ECM protection is obtained at the expense of a reduced throughput. Slow hopping can be realized with remote controlled, conventional radio equipment modified for hopping.

3.3.3 Medium hop rates (20 to 100 hops/s)

The shortwave transmission rate for sky-wave and multipath propagation is limited to about 200 Bd without the use of modems. Signal duration has therefore a minimal value of approximately 5 ms. Taking into account the transients required for switching as 5 ms, then the upper limit of frequency hopping lies at 100 hops/s in conjunction with mulitpath propagation with conventional methods.

A hop rate of 20 hops/s is attained with radio equipment of the new generation which is designed for fast frequency switching. The latter concerns synthesizer, receive/transmit switch and ATU. Hop rates of 20 hops/s can be realized with most components and using conventional techniques, whilst maintaining outstanding characteristics such as spectral purity of signal or high S/N ratio of radio equipment.

A hop rate of 100 hops/s requires radio equipment especially tailored for frequency hopping. To achieve this, known problems of fast frequency switching must be solved: fast frequency processing with the disadvantages of an unfavourable spectrum, rapidly switching harmonics filter and ATU, IF filter optimized for fast transients.

3.3.4 Fast hopping (1000 hops/s)

With hop rates of over 100 hops/s, multipath effects occurring with sky-wave propagation must be removed by the reception time-slots or a method has to be selected that

compensates for the effect of multipath propagation. As the time of travel of multipath components varies, the removal of these components and thus the hop rate depend on the range and are time-variable.

A compensation of multipath propagation can be achieved by signal equalization requiring very fast digital filters. Due to the short signal duration at high hop rates, training pulses cannot be used in the determination of filter coefficients and signal shapes are to be utilized whereby the signal consists of the part first arriving without multipath superposition and subsequent part with multipath superposition.

The above mentioned problems do not occur with ground-wave links (e.g. communications at sea). If a transmitter, however, uses a special communication technique such as fast hopping, it cannot be reached by another method, i.e. slow hopping.

Technological solutions for fast hoppers are feasible (12), their realization is, however, not expected in the long term because of the enormous technical expenditure required and is not needed since the reaction times of jammers of a few milliseconds cannot be reduced for a long time.

4. System realization

A dwell time of 40 ms (corresponding to 25 hops/s) offers sufficient protection against jamming also in the long term; 10 ms switching time allows a design to yield radio equipment with excellent characteristics (e.g. spurious suppression). Requirements for reliable transmission even on poor links and simple operation have to be fulfilled and lead to a radio system with the following characteristics (') :

- A switching synthesizer with 2 slow and one fast loops permitting changes of frequency at 50 ms interval for a transmit signal of high spectral purity
- A PIN-diode-controlled harmonics filter allowing fast frequency change even beyond the limits of the harmonics filter

For 1 kW-type:
- A broadband antenna in combination with an intelligent amplifier affording a change of frequency in about 2 MHz wide bands without tuning the ATU

For 150 W-type:
- A concept same as for 1 kW allowing a fast frequency change over broadband ranges without tuning the ATU. For frequency changes beyond the range limits, the ATU tunes silently in 7 ms with fast mechanical relays having a lifetime of 10exp9 switching operations, leading to an average MTBF of several years under normal operating conditions.
- An intelligent controller (") controlls the automatic operation with automatic channel selection following passive and/or active channel analysis, automatic addressed link set-up in different network configurations, adaptive reaction, frequency hopping and error correcting facility matched to the type of transmission.

5. Field trials

Field trials were intended to test the concept Automatic Link Set-up with automatic channel selection (ACS), adaptive reaction (AR) and frequency hopping (FH). Accessibility, transmission reliability and attainable throughput were to be quantified.

The field trials were carried out in November 1985 between Hook of Holland and Munich (700 km) (''') . The transmitting power was 400 W, both stations were equipped with broadband antennas. Results taken over the duration of a day are presented and are typical for the whole series of measurements.

5.1 Methods using acknowledgement

Methods using acknowledgement are superior to techniques with forward error correction in the case of strongly fluctuating channel quality since a reaction to the changing channel quality is possible and the redundance required for error correction is adapted to the channel quality.

The trials were carried out with simplex ARQ and the maximum transmission rate was 100 Bd. Thus, the throughput being between 0 and 100 Bd is directly proportional the efficiency and/or the relative number of useful channels.

In the case of technique with adaptive reaction, 12 ACS frequencies between 3.2 and 8 MHz were employed and distributed practically uniformly over the frequency range. In trials with frequency hopping, 29 FH frequencies were used between 3.1 and 5.5 MHz repeatedly over a period of 24 hours.

Footnotes:
' Further development of equipment described in (5,6)
'' ECCM and control unit jointly developed by Rohde & Schwarz and Siemens
''' The field trials were conducted by B. Rittenauer, SAG; SHAPE Technological Center kindly placed fequencies and radio equipment infrastructure at our disposal

The trials were performed as follows:
Every minute, a link setup was initiated followed by the transmission of a test string. A link was set up ten times with ACS on the ACS frequencies and a message was transmitted with AR. A link set-up was subsequently made ten times in FH mode and a message transmitted on the FH frequencies. The following parameters were measured: time for link set-up, input bit error rate, residual error rate and throughput of message transmission.

When the quality of the link deteriorated during transmission in AR mode, a new frequency was selected with active channel analysis (ACA) and the transmission continued on this frequency. In FH mode, 29 FH frequencies were repeatedly used, their selection being pseudo-random.

5.1.1 Adaptive reaction

Fig. 3 illustrates the results of automatic link set-up with ACS and transmission with AR in a 24-hour period from 14.00 h on 30.11.1985 to 14.00 h on 01.12.1985. The top curve represents the throughput for message transmission: it lies at 75 Bd on average and fluctuates between 50 and 90 Bd. The maximum throughput of 100 Bd could not be attained because of protocol overheads. It is therefore expected that for normal operation the average throughput would increase by 10 % from 75 to 82.5 Bd. The lower curve shows the corresponding times for the link set-up: the values lie between 2 and 6 seconds with an average of 3 seconds. There were no undiscovered residual errors.

5.1.2 Frequency hopping

Fig. 4 illustrates the results of transmission with FH in a 24-hour period from 14.00 h on 30.11.1985 to 14.00 h on 01.12.1985. The top curve represents the troughput for message transmission: it fluctuates in a range between 20 and 60 Bd and has an average value of 43 Bd. If hop sets matched to the time of day are used, the throughput may well be 10 to 20 Bd higher.
The lower curve shows the corresponding times for the link setup: the values lie between 2 and 9 seconds with an average of 4.5 seconds. The residual error rate for both techniques lie at around 10exp(-5). There were no undiscovered residual errors.

5.2 Methods without acknowledgement

As results under 5.1.2 show, high redundancy is required for frequency hopping in combination with techniques without acknowledgement in order to enable transmission in spite of a high percentage of disturbed channels. This redundancy is attained in the described system by coding the information with an error correcting code and by the multi-transmission of the information on different channels. It will therefore be possible to transmit information reliably and free of errors, although with reduced throughput, even in the case of broadcasting techniques using frequency hopping.

6. System test with communication simulator

In order to test and compare radio systems, system measurements on channels with reproducible parameters are required in addition to the link trials. This task is best performed by a channel simulator (') that simulates the characteristics of the HF transmission over a broadband. This means that a broadband channel of say 5 to 25 MHz is assigned to the tester which possesses different preset characteristics at different frequencies.

Fig. 5 presents a communication simulator for channel and ECCM evaluation. The transmitter and receiver to be tested are connected to the transmission channel. The channel is simulated to include all known effects of HF transmission, i.e. delayed transmit signals for the simulation of multipath propagation, phase and group delay fluctuations of the different propagation paths, selective fading and absorption fading and the simulation of the Doppler effect of the transmission path. In addition, jammers such as single channel, multichannel, pulse, sweep jammers and others can be simulated by applying different types of interfering signals.

7. Summary

Tests on frequency hopping have shown that at the present level of occupancy of the shortwave band a high percentage of disturbed hop channels i.e. 50% is to be expected. Frequency hopping systems occupy in addition a large number of frequencies and raise the interference level of the available frequency range that is already scarce in Europe. Although, if required, frequency occupancy could be lowered by greater radio discipline and less traffic, the situation remains basically the same. This means that the effectiveness of FH systems regarding throughput will always be lower and the main application of these systems can only lie in the field of strong jammers. Since, in the long term, search times of single channel receivers will not drop below 1 ms per

Footnote:
' The cannel and jam simulator was promoted by the BMVg and developed in joint cooperation of E81 and Rohde & Schwarz

channel and the identification time of multichannel or FH analyzers is not expected to be below 30 ms, a hop rate of 20 hops/s is regarded as adequate.

Normal radio operation will, also in the future, use the non-hopping mode, it will however be significantly improved with state-of-the-art techniques such as adaptive reaction. Frequency hopping remains unquestionably as a fall-back facility.

Acknowledgement

The authors wish to express their gratitude to Herr B. Rittenauer of Siemens AG for the fruitful cooperation in the development of the control concept and for conducting the field trials

SHAPE Technical Center for their support of the field trials

Dr. K.H. Petry of Rohde & Schwarz Radiomonitoring Division for his suggestions in questions relating to ECM and ECCM

Dr. L. Schirm of Rohde & Schwarz Radiocommunications Division for his contribution to Figs. 1 and 2.

* ROHDE & SCHWARZ Gmbh & Co KG
Postfach 80 14 69
D-8000 MüNCHEN 80
F.R.G.

References:

1. Conference on "Recent Advances in HF Communication Systems and Techniques", Colloquium Digest No. 1979/48, 27-28 February 1979, IEE, London

2. Second International Conference on "HF Communication Systems and Techniques", 15-17 February, 1982, IEE, London

3. Third International Conference on "HF Communication Systems and Techniques", Conference Publication No. 245, 26-28 February, 1985, IEE, London

4. "Reliable Shortwave Communications using ALIS", G. Greiner, to be published in News from Rohde & Schwarz, No. 116, 1987, Munich

5. "Transceiver Family HF 850 for Shortwave", G. Greiner et al News from Rohde & Schwarz, No. 105, 1984, Munich

6. "ALIS Processor GP 853 for Automated and Adaptive Shortwave Radiocommunication", P. Iselt News from Rohde & Schwarz, No. 109, 1985, Munich

7. "CHP 200 - Ein Kommunikationsprozessor für HF-Funknetze", R. Mahla, B. Rittenauer, Telcom Report 9 (1986), Sonderheft "Nachrichtenübertragung auf Funkwegen", S. 354-360

8. "Improvements in Reliability and Security for Future HF and VHF/UHF Links", W.-R. Lange, Middle East Electronics, May 1983

10. "Experiments with Spread Spectrum Modulation on Radiowaves Reflected from the Ionosphere", R. Skaug, AEÜ, Band 35, Heft 4, S. 151, 1981

11. "Factors affecting Use and Design of Spread Spectrum Modems for the HF Band", A.N. Ince, R.E. Schemel, IEE Proceedings, Vo. 133, Pt. F, No. 2, April 1986

12. "Technological Solutions for HF Communications", U. Rohde, Conference Publication No. 245, 26-28 February, 1985, IEE, London

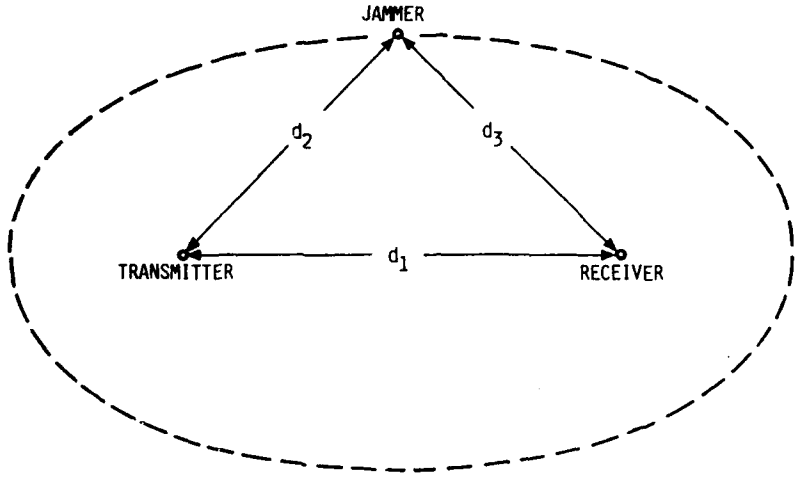


Fig. 1: GEOMETRICAL CONFIGURATION OF COMMUNICATORS AND JAMMER

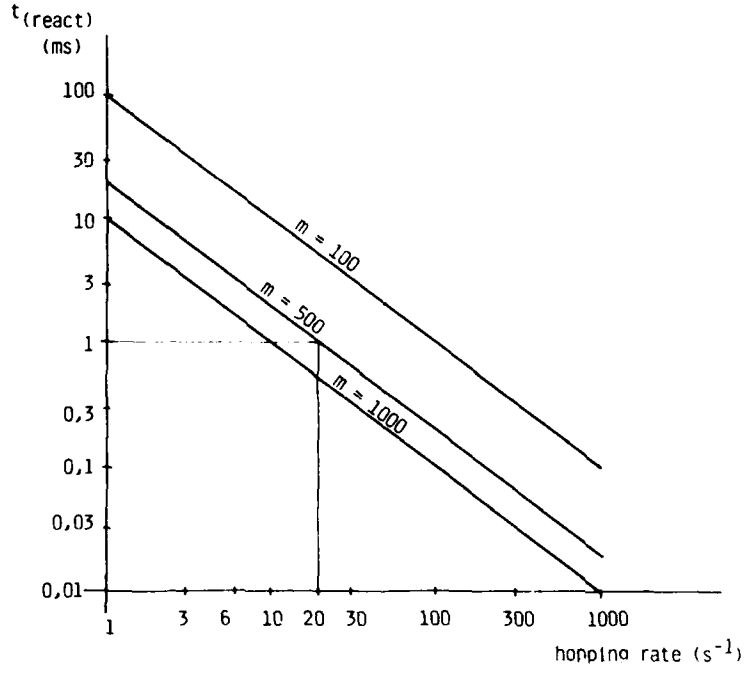


Fig. 2: Maximum permissible reaction time $t_{(react)}$ per channel for 10% of detected channels as a function of hopping rate; m=number of hop channels

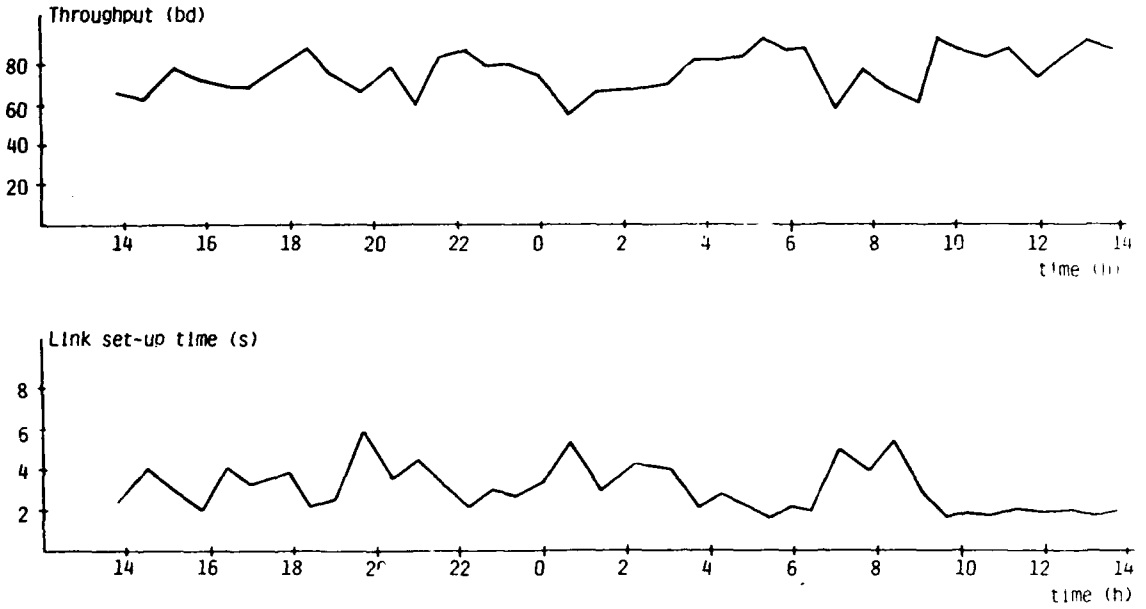


Fig. 3: Typical curve for the throughout of a transmission with adaptive reaction over 24 hours (measurement: B. Rittenauer, SAG)

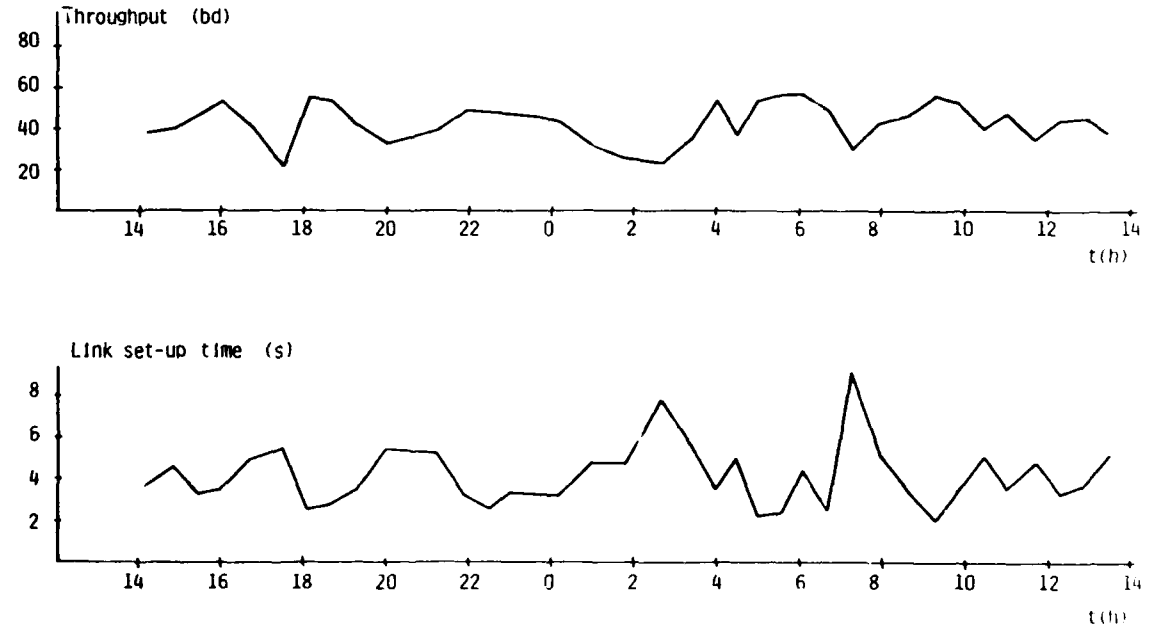


Fig.4: Typical relationship between time-of-day and throughput (=% of interference-free channels) for FH transmission with constant frequency pool (measurement: B. Rittenauer, SAG)

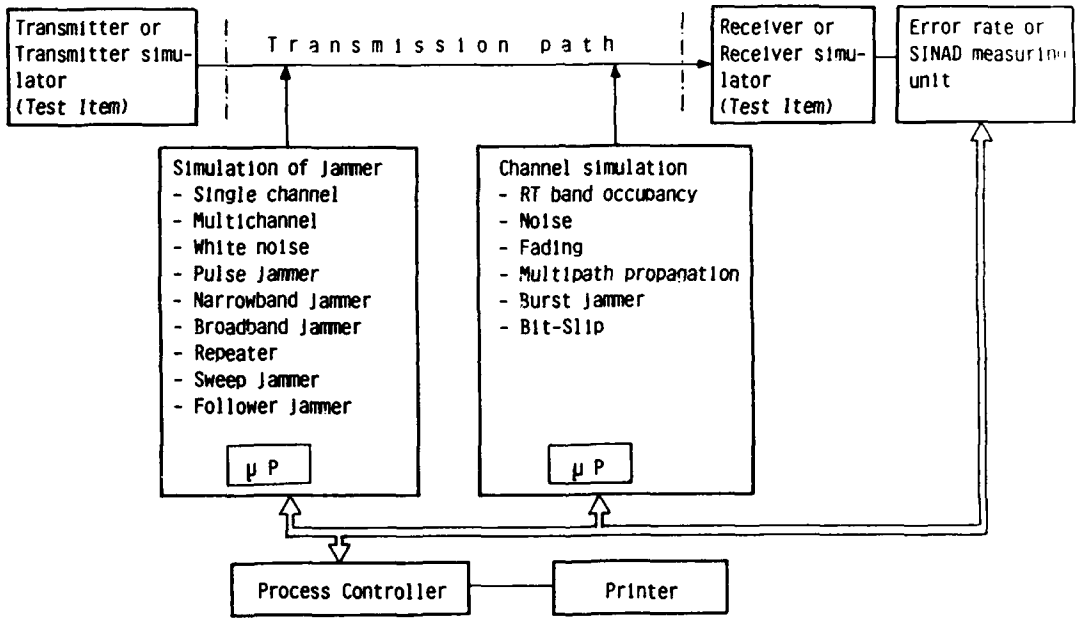


Fig. 5: Block diagram for a computer-controlled communications simulator

PREDICTING THE EFFECTS OF NOISE AND INTERFERENCE ON FAST FREQUENCY-HOPPED M-ARY NCFSK SYSTEMS

by

E. BARRY FELSTEAD
JOHN S. BIRD
U. ANDREJ TENNE-SENS

Communications Research Centre
P.O. Box 11490, Station H
Ottawa, Ontario, Canada K2H 8S2

SUMMARY

This paper discusses calculation of the performance, in terms of received bit error probabilities, of fast frequency-hopped M-ary non-coherent frequency-shift-keyed (NCFSK) systems. It introduces a general method of calculating accurately the performance of such systems for arbitrary M, degraded by system noise and both partial-band-noise and multi-tone jamming. Some typical results are presented and it is shown for a fixed hop rate how the use of time diversity, through repetition of each symbol on several hops, can dramatically improve performance; M may be simultaneously increased to keep the data rate high. This technique in conjunction with only simple error-correction coding can make the system quite robust. System degradation due to multipath can fortuitously be overcome by the same anti-jam techniques.

1. INTRODUCTION

Frequency hopping (FH) is a useful spread-spectrum technique for providing protection from jamming of digital radio links, both in satellite and terrestrial applications. In comparison to the other main spread-spectrum technique, direct-sequence (DS) spreading, FH has several practical advantages. Most importantly, it tends to make it easier to achieve wider spread bandwidths. Also, synchronization tends to be simpler and it can easily be adapted to multiple-user operation through frequency-division multiplexing.

The terms "slow hopping" and "fast hopping" must be carefully distinguished because they have two different senses. In the absolute sense, the terms relate vaguely to the hop rate required to avoid follower jammers which in turn relates to the spatial geometry of the transmitter, receiver, and jammer. In the relative sense, the terms relate the hop rate to the channel symbol rate so that fast hopping is defined as hopping with one or more hops per transmitted-channel M-ary symbol and slow hopping has more than one transmitted symbol per hop. In this paper, only fast hopping in the relative sense is considered.

Much confusion has been generated in the past by analysis in which the data rate is kept constant and the hop rate is allowed to float. This approach is a carryover from analyses of non-spread systems where, indeed, a constant data rate is a good basis of comparison. However, in a FH system there are important practical reasons why the hop rate, R_h, once selected, should remain constant or at least have very few selectable rates that are integer multiples of the lowest rate. Even on a point-to-point single-channel communications link, a continuously variable hop rate would be difficult to implement especially from a synchronization point of view. With a satellite communications system or a terrestrial combat-net radio system, there will usually be multiple users and it would be very impractical to allow each user pair to choose and vary its own hop rate independently. The rate itself is chosen by a trade-off between needing a high hop rate to avoid follower or repeat-back jammers and a low rate to accommodate frequency-synthesizer technology and ease of synchronization, and to achieve as large a value of energy per hop as possible. In this paper, R_h is held constant and the data bit rate, R_b, is allowed to vary. The relevant signal energy quantity becomes the energy per hop, E_h, in place of the more usual energy per data bit, E_b.

It is common to define a FH processing gain as W_T/R_b where W_T is the total hopped bandwidth. This definition is valid only when the interference is wideband additive white gaussian noise (AWGN). Since the fast-hopping systems considered here have a fixed R_h, the achievable processing gain from the FH is limited to

$$PG_{fh} = W_T / R_h, \qquad (1)$$

independent of R_b. Thus, decreasing R_b well below R_h gives no increase in PG_{fh}. However, a decrease in R_b gives an increase in redundancy which has an associated processing gain defined as

This processing gain can be implemented through the use of diversity (repeating the same transmitted symbol on L hops) and/or error-correction (EC) coding. This value of PG_r will not necessarily be achieved against wideband AWGN interference. Note that PG_{fh} is defined for wideband AWGN interference but can degrade considerably (tens of dB) for other types of interference such as partial-band noise (PBN) and multiple-tone (MT) jamming. It is strongly emphasized that in the presence of such interference, the diversity/coding can not only provide gain on its own approaching PG_r, but can compensate for most of the losses suffered to PG_{fh}.

It is usually impractical to maintain phase coherence between hops which means that in fast-hopped systems, there is no coherence between transmitted symbols. As a result, the modulation of choice tends to be M-ary non-coherent frequency-shift keying (NCFSK).

In this paper we will briefly describe a performance analysis method which is capable of calculating exactly the probability of bit error, P_b, of fast-hopping M-ary NCFSK systems under partial-band noise and multiple-tone jamming, with the system noise taken into account. The jammer's strategy to cause worst-case degradation in performance will be described. The significant improvement in performance possible through the use of diversity and increased M (to maintain the data rate) will be presented. Similar improvement will be shown to be obtainable through the use of error-correction coding. By combining diversity and EC coding, it will be shown how a relatively simple and robust system can be implemented to mitigate all known forms of worst-case jamming. There follows a brief discussion on adaptively varying L and M in the field to maximize throughput as a function of jamming level.

Finally, it will be shown that multipath affects a fast-hopping system in a manner that is similar to the way it is affected by partial-band noise jamming. Therefore, the diversity and EC coding used against such jamming is fortuitously useful in mitigating the effects of multipath.

2. GENERAL PERFORMANCE ANALYSIS FOR BROADBAND-NOISE AND TONE JAMMING

The basic elements of the transmitter are shown in Fig. 1. The input binary data has a period T_b s corresponding to a rate R_b bits/s. The data may be EC encoded at code rate r = (no. of data bits in)/(no. encoded bits out) so that the encoded period is $T_c = rT_b$ s and encoded bit rate is $R_c = R_b/r$ bits/s. This binary information is converted $k = \log_2 M$ bits at a time into one symbol tone of frequency f_1, where f_1 has one of $M = 2k$ possible values. The symbol duration is $T_s = kT_c$ s, and the symbol rate is $R_s = R_c/k$ symbols/s. Finally, the symbols are mixed with a frequency-hopping tone of frequency f_h and duration T_h and hop rate $R_h = 1/T_h$. L is the number of hops/symbol and is a positive integer. Since L>1 corresponds to repeating a symbol, it is a form of time diversity. The number L is sometimes called the "order" or "level" of diversity.

The signal is hopped over a total bandwidth W_T which is divided into N_b bins (typically > 1000) as shown in Fig. 2. The bins are further partitioned into "channels" consisting of M bins; the channels normally do not overlap. We assume in this paper that the M signal bins are assigned contiguously, as contiguous assignment eases implementation of frequency-division multiplexing of multiple users. (It is found in [1] that noncontiguous spacing is generally not recommended, so it will not be discussed in this paper). On a given hop, the signal will be in one of the M bins corresponding to the transmitted symbol. The channels are hopped in frequency in such a way that the jammer cannot predict where the next channel will be.

The receiver is shown in Fig. 3. Its input is the vector sum of the amplitudes of the signal, jamming, and system thermal noise. The signal amplitude is assumed to be $\sqrt{2}s$ so that rms power is s^2. The signal energy received on a single hop is $E_h = s^2T_h$ joules. The dehopped signal goes to a bank of M bandpass filters and detectors. The output of each filter/detector is sampled every T_h s and a decision is made as to which of the M samples is largest. Usually diversity combining is done before the decision device. However, in this paper the diversity combining is assumed to take place after the decision device by simple majority vote over L decisions to determine the most likely transmitted. This M-ary symbol is then converted to a binary representation. Finally, if EC coding is used, decoding is performed.

As discussed in Section 1, throughout this paper R_h is held constant. If increased levels of diversity or coding are required to maintain links, it is the data rate R_b that is sacrificed. These assumptions are the opposite to the ones most often seen in the literature.

We make the common assumption that the jammer has a maximum average power constraint of J_{tot} as seen at the receiving antenna and this power can be distributed over W_T in any manner. The jammer attempts to distribute J_{tot} so as to cause the maximum possible degradation in communications performance. In this paper, the communicator's point of view rather than the jammer's is taken, so that, such jamming will be known as "worst case". It is also assumed that the communicator's transmit power is limited so that s^2 is the maximum signal available at the receiver. It is further assumed that s^2 can be sufficiently small that the receiver noise, N, cannot be neglected. Therefore, it is assumed that R_h, s^2, J_{tot}, and W_T are fixed. The communicator's strategy is to select practical values of M, L and EC coding to get the best performance possible. If the performance is not acceptable, the data rate R_b is sacrificed so that L can be

increased. The jammer's strategy is to try to choose the jamming type (noise or tone) and the distribution over W_T that will cause the communicator's performance to be unacceptable.

Fig. 4 illustrates the important combinations of signal and jammer distributions. The jammer can use noise, assumed here to be white and Gaussian, or tones. Since there is practical difficulty in generating noise over separated individual bins, such situations are not considered here. Therefore all noise jamming is assumed to be distributed over n_j contiguous bins (Fig. 4a) so that a fraction

$$\gamma = n_j/N_b = n_j W_{bin}/W_T \tag{3}$$

of the bins contain jamming power. Such jamming is called partial-band noise (PBN) jamming.

For multi-tone (MT) jamming it is feasible to distribute the tones on any contiguous or noncontiguous distribution over a fraction γ of the bins (Figs. 4b,c). It is assumed that a bin may contain at most one tone. If the distribution is noncontiguous, then it is assumed that the n_j tones are distributed at random over the total N_b bins, i.e., the probability that a bin contains a tone is equal for all bins.

The last two combinations shown in Figs. 4d and e have only one jammer tone per channel. This distribution is of special interest because it is related to Houston's worst-case MT jamming [2]. Since it degrades performance more than the distribution of Fig. 4b, the latter will not be further considered. It can also be shown [1] that the configuration of Figs. 4d and e degrade system performance identically. The fraction of channels jammed is

$$\beta = M\gamma$$

The number of bins being very large implies that for PBN- and contiguous-MT jamming, all M bins of the signal channel are either jammed or none of them are jammed because the probability that the signal channel may span across the boundary between jammed and unjammed bins is negligible. This assumption was found to give results very close to those when exact calculations were performed under the assumption of even a small number of bins ($N_b < 50$).

Some useful powers and power ratios are now defined for the input to the receiver in Fig. 3. For the receiver, the signal-to-thermal-noise ratio (SNR) is defined in the usual manner

$$SNR = s^2/2\sigma_n^2 \tag{5}$$

where σ_n^2 is the variance of each component (in-phase and quadrature) of the complex noise in one bin. Thus if the one-sided noise-power spectral density is N_o W/Hz then $\sigma_n^2 = N_o W_{bin} = N_o/T_h$ so that the familiar form

$$SNR = E_h/N_o \tag{6}$$

is obtained where E_h is the received energy per hop and

$$E_h = krE_b/L \tag{7}$$

It is usual to define an "effective" or "normalized" jamming power as

$$J = J_{tot}/N_b = J_{tot} W_{bin}/W_T \tag{8}$$

which is the power that would be placed in each bin if the jammer's power were spread uniformly across the hopping band. If jamming power is represented through the effective *signal-to-jammer power ratio*

$$SJR = s^2/J, \tag{9}$$

the general analysis can proceed without specific knowledge of s^2, J_{tot} or N_b. The designers of a specific system then merely use the jammer threat characteristics and the FH parameters to find the particular SJR at which they must operate.

The SJR is now defined for both PBN and MT jamming. For PBN jamming, n_j of the N_b bins are jammed with noise that has power σ_j^2 in each bin and the remaining bins have no jamming. Thus

$$2\sigma_j^2 = J_{tot}/n_j = J/\gamma \tag{10}$$

Notice that the jamming power in each of the jammed bins varies inversely with γ; the jammer attempts to optimize γ by trading off jamming power in the jammed bins versus the number of bins jammed. Combining (9) and (10) we can also write

$$SJR_{PBN} = s^2/(\gamma 2\sigma_j^2) \tag{11}$$

For MT jamming, n_j of the bins are jammed with a continuous-wave (CW) tone of amplitude a, or power

$$a^2 = J_{tot}/n_j = J/\gamma \tag{12}$$

so that

$$SJR_{MT} = s^2/(\gamma a^2) \tag{13}$$

From the above assumptions and definitions, one can calculate for each jamming and signal combination of Fig. 4, the bit error probability, P_b, given only the SNR, SJR and γ or β as applicable. It has been recently found [4], [1] that exact analytic solutions for the tone-jamming cases, with system noise present, may be found using the so-called Fourier-Bessel series [3]. It is also possible to evaluate these series to within well-defined accuracy limits. This method was used to calculate the P_b curves herein for tone jamming.

Interleaving is a method to overcome burst errors. However, for the signal and jamming combinations considered in this paper, the hopping causes the jammed and unjammed hops to be randomly distributed so that no interleaving is needed. Interleaving could be useful for ameliorating fast fading that is frequency independent if the interleaving depth is longer that the average fade time. For slow fading, such as rain fading at EHF, interleaving becomes impractical and message repetition is more practical.

3. TYPICAL PERFORMANCE OF FAST FH SPREAD-SPECTRUM SYSTEMS

In this section we present some typical examples of the behaviour of P_b under jamming and discuss their implication to the jammer's strategy to disrupt communications as much as possible. Numerous representations of P_b can be used. One useful representation, which will be used here, is a plot of bit error probability P_b versus the fraction, γ, of the bins (or fraction, β, of the channels) jammed, as this gives insight into the jammer's best strategy. We will also assume a baseline SNR of 13.35 dB, which for binary NCFSK in the absence of jamming gives a P_b of exactly 10^{-5}. Curves for M=2 and 8 will be shown for comparison . For purposes of discussion, a P_b of 10^{-5} will be deemed as acceptable in the presence of system noise only, and a P_b of 10^{-3} will be considered as just tolerable in the presence of jamming and system noise.

The full FH processing gain PG_{fh} as defined by [1] is achieved only against AWGN jamming spread uniformly across the hopping band, i.e. γ=1. In practice, the jammer selects a value of γ<1, and possibly uses multitones, that results in P_b degrading to a value well above that for γ=1. Thus the jamming strategy forces the PG_{fh} actually achieved to a value much less than that given by (1).

Fig. 5 shows P_b as a function of γ , when partial-band-noise jamming occurs. There is no diversity, i.e. L=1. For a high SJR of 20 dB and binary communications, it is seen that the jammer will succeed in his objective to make the $P_b > 10^{-3}$ if he spreads his power over less than 0.15 of the hopping bandwidth. He must also, of course, jam more than 2 times 10^{-3} of the bandwidth to ensure a $P_b > 10^{-3}$. His optimum γ is sharply peaked at 0.3 where $P_b = 0.004$. Thus, a jammer which has power enough to reduce SJR to only 20 dB must control the jamming bandwidth fairly stringently to damage communications. If 8-ary communications are used, it can be seen that performance is not as good as with binary communications. It can also be shown that as SNR improves, the upper limit of γ for $P_b > 10^{-3}$ decreases, although there is only a modest effect on the value of worst-case γ.

When the jammer power is increased such that SJR = 10 dB, the curves for both M=2 and 8 flatten out considerably. More importantly, $P_b > .001$ for all $\gamma > .002$. If,however, the jammer does not know what the SJR is at the receiver, it is still in his best interest to keep γ at around 0.05.

When SJR = 0 dB, the jamming power is ample enough that maximum damage is done when the power is spread over the full band, although as for SJR = 10 dB, any $\gamma > .002$ achieves $P_b > 10^{-3}$.

Figure 6 shows the effects of tone jamming for M=2 and 8 when the tones are randomly distributed and more than one tone may fall in a channel M bins wide (Fig.4c). Here the same type of behaviour is noted as for PBN jamming. For high SJR, the optimum γ for the jammer is close to 0. As SJR decreases (jamming power increases), the peak moves up and to the right, and broadens at the same time. For SJR below about 13.5 dB for the binary case, and 16.0 dB for the 8-ary case, $P_b > .001$ for all values of γ. Again, if the jammer does not know the SJR, he should keep below 0.1 to increases his chances of at least causing P_b to be > .001. Just as for PBN jamming, the 8-ary system does not perform as well as the binary, although it should be kept in mind that each 8-ary symbol successfully received (in one hop) conveys three times as many bits of information. Later we will see that it can be useful to increase M to counteract the loss in data rate due to diversity (L > 1).

Figure 7 shows the effect of MT jamming when the jammer is careful to place at most one tone per M-bin channel. Again, results for both M=2 and 8 are plotted. Comparing the binary case with Fig. 6 wherein the jammer paid no regard to the number of tones per channel, we see that the damage done is always greater for $\beta < .5$, but may not be for larger β. When M=8, the jammer generally requires less power to achieve its goal. Most significantly, the peaks for optimum β broaden, so that the jammer need not control the fraction of the channels jammed as closely. He must, of course, be aware

that 8-ary communications is in fact taking place, which may not be easy to detect.

It is of considerable interest as to whether PBN or MT jamming causes more degradation in performance. We consider only the worst-case MT jamming (Fig. 7). In the absence of system noise, it can be shown [1] that

$$SJR_{MT} = (M/2ck)SJR_{PBN} \tag{14}$$

which gives the SJR required to achieve equal worst-case P_b, denoted P_{bwc}. Here, c is a constant dependent upon M and is derived in [2]. Thus, to achieve equal performance, SJR_{MT} must be larger that SJR_{PBN} by the factor M/2ck which is given, in dB, in Table 1. The extra power requirement goes from a bothersome 4.3 dB at binary to a decidedly unpleasant 12.6 dB at 32-ary. There are two obvious reasons for the disparity in performance. First, worst-case MT jamming uses a jamming tone whose amplitude equals the signal in any jammed hop, whereas PBN jamming has an amplitude that can be less that the signal, thereby usually not causing any error; or can be greater, thereby wasting jamming power. Second, only one bin out of M is jammed by MT jamming whereas the PBN power is distributed across all M bins. This unbalancing effect makes the effect of MT jamming worse as M increases. However, if system noise is assumed to be present, as is usually the case, it is demonstrated in [1] that the power advantage the jammer has using MT instead of PBN jamming is significantly reduced.

TABLE 1

Values of $SJR_{MT}|_{dB} - SJR_{PBN}|_{dB}$ to Obtain Equal P_b for Worst-Case Operation in the Absence of System Noise

M	10 log (M/2ck)
2	4.3 dB
4	6.3 dB
8	8.3 dB
16	10.5 dB
32	12.6 dB

4. PERFORMANCE IMPROVEMENT BY DIVERSITY COMBINING AND EC CODING

As discussed in Section 3, the FH processing gain PG_{fh} given by (1) can be degraded considerably by a jammer's use of choosing a worst-case γ. The values of P_b in the examples calculated were in general unacceptable. The anti-jam performance of a FH system can be significantly improved by the use of time redundancy at the sacrifice of data rate. The redundancy "gain" (see (2)) can be used not only to retrieve the part of PG_{fh} lost to jammer strategy but to give performance improvement beyond that of FH alone. This performance enhancement is particularly important in the presence of very strong jamming.

Two forms of redundancy are considered here, namely EC coding and diversity combining. It can be argued that diversity combining is merely a simple form of EC coding. Nonetheless, it will be seen to be useful to distinguish the two and to use both in concatenation. Generally speaking, a diversity combiner can yield performance gains with lower values of input probability of symbol error P_{sp} than can EC coding. Also, a diversity combiner can easily be changed to higher levels of diversity, L, as a system becomes more stressed whereas it is difficult to change the code or code rate, r. However, EC coding tends to be much more efficient with respect to preserving data rate. Therefore, a good combination for the receiver end is shown in Fig. 3. wherein following the demodulator there is a diversity combiner, which can handle a variable diversity, and can bring the error performance up to the level that the EC decoder can handle adequately. The subsequent EC coder would be at a fixed rate. In unjammed or lightly jammed operation, the diversity would likely be set to L=1 for maximum throughput but the EC coding would likely be left in place. The application usually considered in the literature for EC coding is for worst-case γ (peak of the curves in Figs. 5 to 7) but light jamming, eg. SJR >10 dB.

Reference is sometimes made in the literature to use of "side information" to determine which hops are jammed during PBN or MT jamming. This information is then used to improve the error-correction process by discounting those jammed hops. Impressive performance enhancement appears to be achieved. However, we believe that such a technique is not recommended for practical systems for several reasons. Firstly, the detection of which hops are jammed is unreliable since a sophisticated jammer does not use excessive power in any particular hop channel. Secondly, the detection of such side information can add to the system's complexity. Most importantly, these techniques are prone to attack by a revised jammer strategy with the risk of worse communications performance than if side information had not been used.

Diversity combining at the receiver can be implemented in a number of ways. The usual form seen in this literature is to follow each of the M matched filters in Fig. 3 with a nonlinearity and a summer that adds L values. After this addition, the one bin out of M with the largest value is declared to be the received symbol. If the nonlinearity is merely an envelope detector such as the square law used in [5], then the diversity combining has a serious weakness against an intelligent jammer. The jammer could use few relatively large jamming tones so that even one "hit" in L hops could make the decision go in favour of the jammed bin. In order to avoid this problem of giving undue weight to a jammed hop, a variety of other nonlinearities have been looked at. Some of the many nonlinearities are considered in [6]. They usually involve some form of limiting or AGC based upon total power in the M bins in each hop. The ones involving a form of AGC introduce some complexity. An approach that appears to perform well at the expense of complexity [7], generates a "quality" bit based upon whether the ratio of the two largest bins in a hop exceeds a threshold. A very simple form of nonlinearity is to make a hard decision (choose 1 of M) on each hop and then declare the received signal to be the one with the most "hits" in L hops (majority vote). In the event of a tie, either one is arbitrarily chosen. The hard-decision-majority-vote approach protects against the large-tone jamming problems, is very simple to implement, and for L>2 probably performs nearly as well as the more complex methods.

The hard-decision-majority-vote diversity combining method is discussed in [1] and is the one used for illustration in Fig. 3. The output of the initial decision has a symbol error probability of P_{sD} which is calculated from the methods discussed earlier. The output symbol error P_{sv} is found by considering the combinations of events that can occur and calculating the probability of all those events that result in an error. Ties are handled by making an arbitrary decision between the tied symbols. Methods of calculating P_{sv} as a function of P_{sD}, M and L are discussed in [1] and found to have a general form of

$$P_{sv} = \sum_{i=[L/2]}^{L} a_{MLi}\, P_{sD}^{\,i}\, (1 - P_{sD})^{L-i} \qquad (15)$$

where [L/2] denotes "next integer above L/2". Selected values of a_{MLi} are given in Table 2 and are given in [1] for L from 1 to 9 and M from 2 to 16.

TABLE 2

Coefficients a_{MLi} for various values of L, i and M.

L	i	M=2	M=4	M=8	M=16
3	2	3	2.33	2.14	2.07
	3	1	1	1	1
6	3	10	1.11	0.204	0.044
	4	15	10.6	5.55	2.80
	5	6	6	5.85	5.53
	6	1	1	1	1
9	5	126	9.33	0.840	.090
	6	84	50.6	13.6	3.36
	7	36	36	28.64	18.36
	8	9	9	9	8.90
	9	1	1	1	1

We now give specific examples of performance enhancement possible through diversity combining. As discussed in Section 2, it is assumed that the hop rate R_h is constant and that the communicator, in the face of an inadequate SJR = E_h/J_o, varies M and L at the sacrifice of R_b.

In the first example, the SJR = 15 dB curve from Fig. 7 for binary NCSFK in worst-case MT jamming had diversities L=3, 6 and 9 applied to it and the resulting curves are shown in Fig. 8. They were calculated from P_b of Fig. 7 by applying the appropriate polynomials derived from [15] and Table 2. A modest value of 3 for L brings the peak P_b to within the 10^{-3} objective, and at L=9 the peak is impressively low 2.2 x 10^{-7}. The data rates are $R_b = R_h/L$. Furthermore, the peaks are made sharper as L increases. This increased sensitivity to the value of β makes the jammer's work more difficult.

In the second example, the above calculations were repeated for 8-ary NCFSK for the SJR = 15 dB curve of Fig. 7. The results are shown in Fig. 9. The data rates R_b are $3R_h/L$. For this example, the improvement due to diversity is not as dramatic as for the binary example because the performance for no diversity (L=1) is so poor. L must be at least 6 to decrease P_b below the 10^{-3} objective for all β, but R_b now equals $R_h/2$, an improvement over the binary case considered above. Thus, increasing M can at

least partially compensate for the decrease in symbol rate.

The application of EC coding to fast FH has been considered in the literature. A useful summary of results is presented in [5] where the "coding gain" of a variety of EC codes is determined. The reader is cautioned that they use a fixed R_b and floating R_h. Rather impressive coding gains (20 to 30 dB) are reported but are perhaps misleading because of the method of determining coding gain. To explain this statement consider the P_b vs β curves of Fig. 7 for binary signals and suppose SJR = 15 dB and is the worst-case value so that $P_b = 1.8 \times 10^{-2}$. Such an error rate can be corrected to 10^{-5} or better by even simple EC codes. To calculate coding gain, one then must determine the increase in SJR required to achieve 10^{-5}. In [5] it is assumed that as SJR is increased, the jammer varies β to its new worst-case value so that an SJR of 45.7 dB is required for $P_b=10^{-5}$. The corresponding coding gain is said to be 20.7 dB. This approach is misleading for two reasons. Firstly, at SJR = 45.7 dB the peak is very sharp so that if the jammer does not choose β very precisely the value of P_b falls dramatically. Secondly, in [5] a region of input P_b is chosen that is easily correctable by EC coding. Had the authors looked at the region nearer SJR = 0 dB where the P_b vs β curves are flatter and P_b is larger, the EC coding would not give anywhere near the impressive gains and in fact could break down and perform worse than no coding. In short, it is best not to attempt to define a coding gain at all; it probably arises because of the affection that communications engineers have for "gain" in dB. Instead, what is important are the initial and final values of P_b.

In summary, one should use diversity to whatever level is necessary to bring P_b to 10^{-1} to 10^{-2}. Then the EC coding can be used to bring P_b down to the desired level (10^{-5} in our example).

5. ADAPTIVE DIVERSITY

From the foregoing, it is clear that one may adaptively increase the values of L and M as the threatened or existing jamming is increased. An increase in L will significantly increase the anti-jamming margin, as illustrated in Figs. 8 and 9, and an increase in M can offset at least some of the decrease in data rate.

It is emphasized once more that, for practical reasons discussed earlier, the hop rate should be held constant. In addition, there is another reason why the hop rate should not be changed adaptively. It can be assumed that the enemy's EW capabilities allow it to monitor the hop rate. A change in hop rate would tell the enemy that the jamming strategy is having some effect. It is better for the communicator not to give such feedback. By contrast, changes in L and M are not detectable by an interceptor.

In an adaptive system, the actual received SJR must be determined so that the required value of L and M can be selected. Usually, L and M will have relatively few selectable values so that SJR need only be known approximately. Thus, the technique used at the receiver to determine average SJR can likely be quite simple such as tapping off the samples in the decision device of Fig. 3 and performing averaging in each frequency bin over many hops followed by a calculation estimating SJR. This estimate then determines the values of L and M required. The more difficult part of the adaptive process is the need to inform the transmitter that a change in L and M is needed and to perform the ensuing system changeover. Care is needed in avoiding the introduction of points of attack by jammers on the adaptive system itself. For example, if the adaptive system reacts too fast, the jammer that turns on and off could conceivably cause the system to be reconfiguring continuously and never communicating. One approach for critical missions would merely be to set L and M to the largest values necessary to overcome the estimated jamming threat with recognition that information throughput is suboptimal.

6. MITIGATING THE EFFECTS OF MULTIPATH

The performance of frequency-hopping systems can be degraded by multipath propagation. In a VHF combat-net radio system, for example, multipath may be caused by reflection off large objects such as buildings, hills, or aircraft. We show here how it is possible to analyze the effects of multipath by methods similar to those used for MT jamming.

We will use the commonly employed two-component model of multipath. The larger of the two components is sometimes denoted the "direct" one and the smaller the "reflected". A reflection coefficient is defined as

$$\rho = |A_r|/|A_d|, \quad 0 \leq \rho \leq 1 \tag{16}$$

where A_r and A_d are the complex-valued amplitudes of the reflected and direct components, respectively. The phase difference between the two components is a uniformly distributed random variable. Although in the following discussion only two components are considered, the results are easily extended to more than two.

During one hop period, the reflected component is at the same frequency as the direct component but at random phase and offset in time. It acts as a tone interferer. The overlap region, shown in crosshatching in Fig. 10, has a duration gT_h where g is the overlap coefficient. It is defined such that for no overlap (and therefore no

interference) g=0, for complete overlap g=1, and for partial overlap it is proportional to the magnitude of the overlap; therefore $0<g<1$.

The effect of the interfering reflected tone can be seen better in the Fourier transform domain. Recall that a Fourier transform followed by envelope detection and sampling at the tone frequencies is mathematically equivalent to the matched-filter operation. A typical result for a binary signal with minimum orthogonal spacing is shown in Fig. 11 where the solid curve represents the direct component and the dashed curve represents the reflected component. The spectrum of the reflected components is spread because only gT_h of the input tone is processed. This spreading causes "leakage" into the adjacent bin. System noise is not shown. The reflected component interferes in two ways. In the signal bin itself there is an interfering tone of amplitude ρg and random phase, relative to the direct component. In the other frequency bin, the leakage results in another interfering tone whose amplitude is $(\rho/\pi) \sin \pi g$ relative to the normalized direct signals. This amplitude is 0 at g=1 because there is no leakage, is 0 at g=0 because there is no reflected signal in the hop period, and is a maximum of ρ/π at g = 0.5.

In the absence of system noise, these interfering tones would cause a bit error only when the interference in the signal bin is of such an amplitude and phase to reduce the combined direct plus reflected signal to below the leakage level in the adjacent bin. In the presence of system noise, the analysis techniques provided in Section 2 are useful. One merely specifies the SNR, and the amplitudes of the two interfering tones. For values of M beyond 2, the analysis is slightly more complicated because the leakage tones appear in M-1 bins and their amplitudes depend upon the location of the signal bin within the channel.

A specific example was calculated for binary signals and is shown in Fig. 12 where P_b is plotted against the overlap, g, for the worst-case reflection coefficient, ρ=1. It is assumed that SNR = 13.35 dB, and there is no jamming. When there is no overlap (g = 0), P_b is determined only by the system noise and is 10^{-5}. As the amount of overlap increases, P_b rises sharply to a peak of 0.056 at an overlap of 0.7, and then settles down to 0.042 at full overlap. Thus, for SNR = 13.35 dB, the worst-case performance is 0.056 and occurs at ρ = 1, and g = 0.7. EC coding and diversity can be selected to correct this value of P_b down to the desired one. For other values of ρ, the value of P_b falls between the ρ=0 and ρ=1 curves of Fig. 12.

A general conclusion can be drawn from the calculated binary example. Since the worst value of P_b ever encountered for SNR = 13.35 dB and binary signalling is 0.056, and this worst value will be rarely obtained, then the EC coding or diversity used for AJ purposes would be more than adequate to compensate for the multipath degradation. Thus, fast FH systems with EC coding or diversity are robust against multipath effects. A similar conclusion can likely be drawn for values of M beyond 2.

7. CONCLUSION

A fast FH system with constant hop rate can provide processing gain of the order of W_T/R_h against full-band jammers. However, for light jamming, say SNR > 15 dB, much of this gain can be lost if the jammer takes an appropriate PBN or MT approach. Fortunately, such degradation in performance can be corrected by EC coding techniques.

For heavy jamming, say SJR $\leq$ 0 dB, the jammer will spread its power across the band (γ= 1) since this value of γ causes the highest P_b and this value of P_b is so large that most EC codes will not be able to improve the error rate significantly or might even degrade it. Fortunately, not all is lost. At the sacrifice of data rate, the diversity, L, can be increased indefinitely and almost any value of input P_b can be corrected to the desired levels. Although the processing has no theoretical limits, other practical factors begin to limit the minimum value of SJR that can be handled. Some examples of these factors are the ability to synchronize, mixer break-through for ground-based point-to-point radio, and receiver saturation for satellite communications.

In short, fast FH in combination with EC coding and diversity can provide considerable protection against a powerful and intelligent jammer.

REFERENCES

1. J.S. Bird and E.B. Felstead, "Antijam performance of fast frequency-hopped M-ary NCFSK -- an overview," IEEE J. Selected Areas in Communications, Vol. SAC-4, pp. 216-233, Mar. 1986.

2. S.W. Houston, "Modulation techniques for communication, Part 1: Tone and noise jamming performance of spread spectrum M-ary FSK and 2, 4-ary DPSK waveforms," in IEEE NAECON '75 Rec., 1975, pp. 51-58.

3. J.S. Bird and D.A. George, "The use of Fourier-Bessel series in calculating error probabilities for digital communication systems," IEEE Trans. Communications, Vol. COM-29, pp. 1357-1365, Sept. 1981.

4. J.S. Bird, "Error performance of binary NCFSK in the presence of multiple tone interference and system noise," IEEE Trans. Communications, Vol. COM-33, pp. 203-209, Mar. 1985.

5. H.H. Ma and M.A. Poole, "Error correcting codes against the worst-case partial-band jammer," IEEE Trans. Communications, Vol. COM-32, pp. 124-133, Feb. 1984.

6. K.S. Gong, "Performance of diversity combining techniques for FH/MFSK in worst case partial-band noise and multiple-tone jamming," in IEEE Milcom 83, 1983, IEEE Cat. no. 83CH 1909-1, pp. 17-21.

7. A.J. Viterbi, "A robust ratio-threshold technique to mitigate tone and partial-band jamming in coded MFSK systems," IEEE Milcom 82, 1982, IEEE Cat. No. 82CH 1734-3, pp. 22.4.1-5.

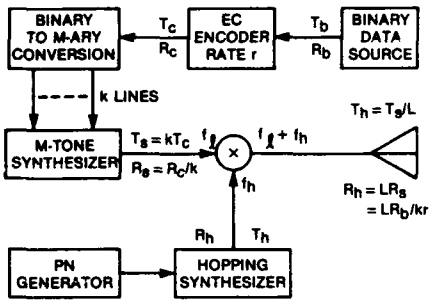


Fig. 1. Block diagram for a fast-hopped M-ary NCFSK transmitter.

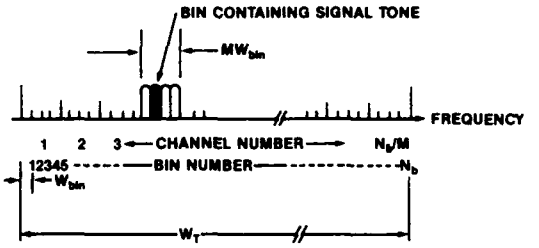


Fig. 2. A representation of the signal spectrum during a given hop, for a system with M=4. The particular channel selected for this hop is symbolized by the group of four outlined signal bins, each representing a different dibit. The receiver tries to determine which of these bins contains the signal tone: here it is the second one.

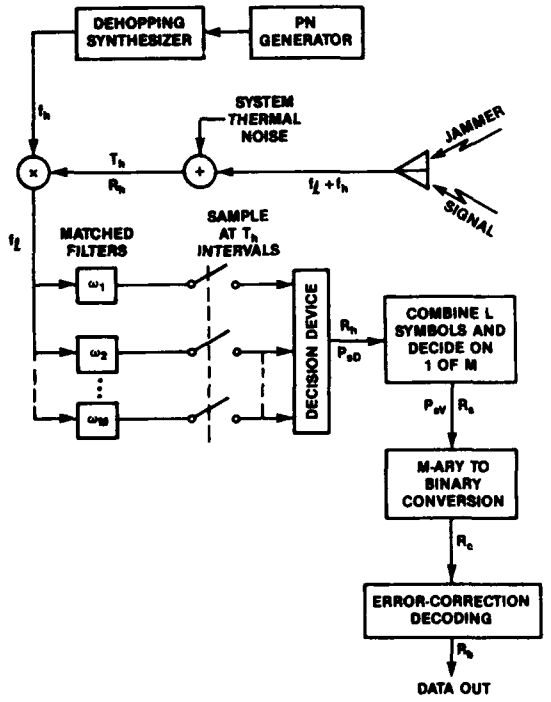


Fig. 3. Receiver block diagram.

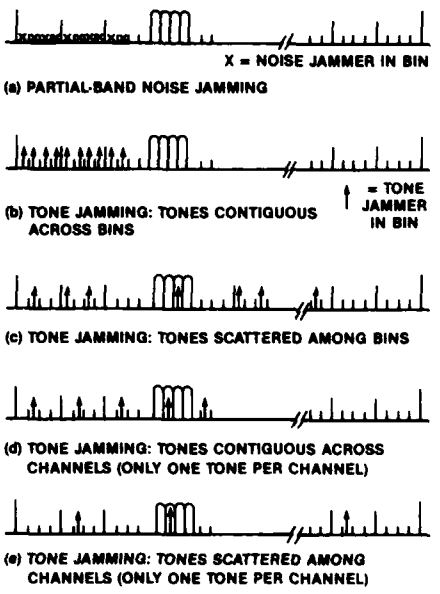


Fig. 4. Five different types of jamming configurations. Configurations (d) and (e) damage communications equivalently, and always more than (b).

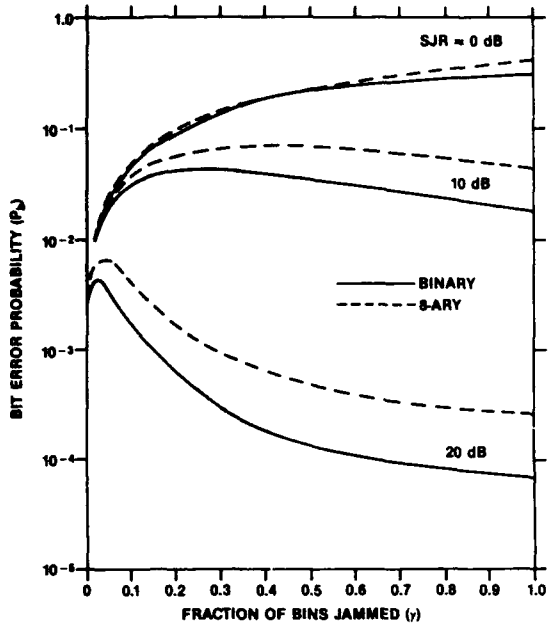


Fig. 5. P_b versus γ for SNR = 13.35 dB under partial-band noise jamming as in Fig. 4a.

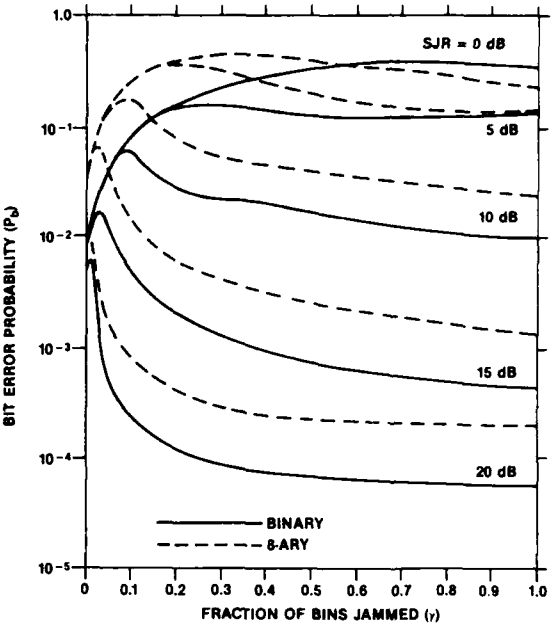


Fig. 6. P_b versus γ for SNR = 13.35 dB under multitone jamming as per Fig. 4c.

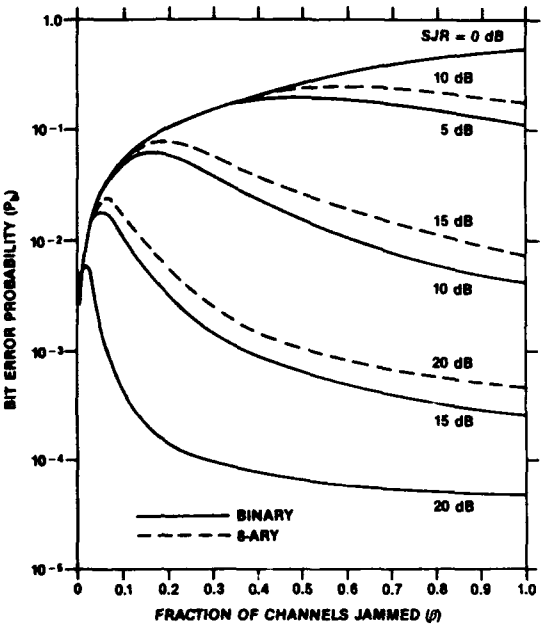


Fig. 7. P_b versus β for SNR = 13.35 dB under multitone jamming as per Figs. 4d or e.

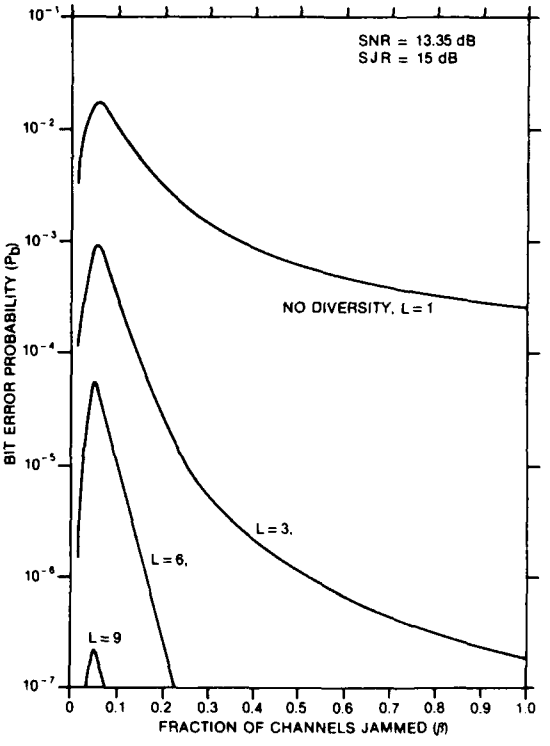


Fig. 8. P_b as a function of $\beta = 2\gamma$ for binary NCFSK, with diversity combining, in the presence of MT jamming (Figs. 4d or e) plus system noise of SNR = 13.35 dB.

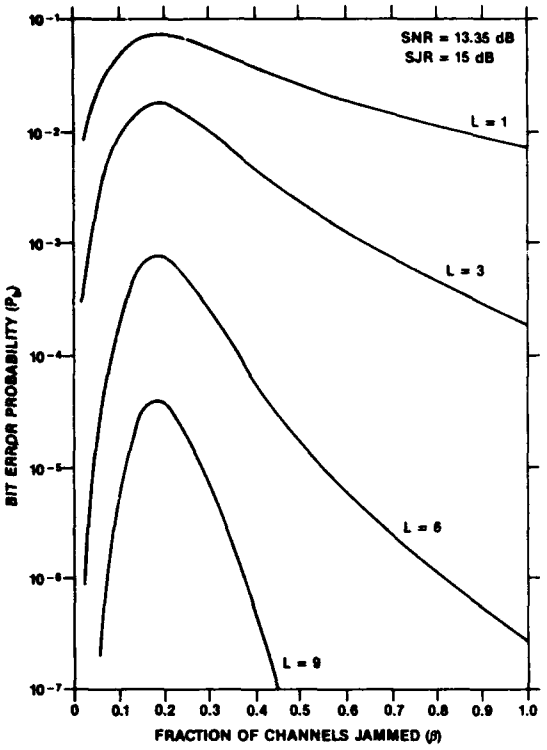


Fig. 9. P_b as a function of $\beta = 8\gamma$ for 8-ary NCFSK with diversity combining at the symbol level, MT jamming (Figs. 4d and e) plus system noise of SNR = 13.35 dB.

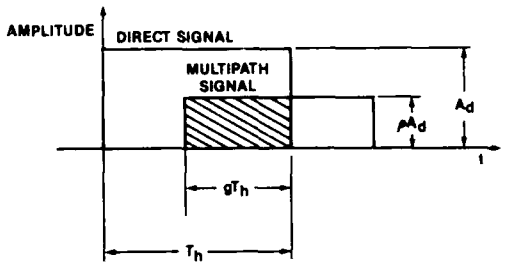


Fig. 10. Relative amplitudes and delays of direct signal and reflected signal in a given hop.

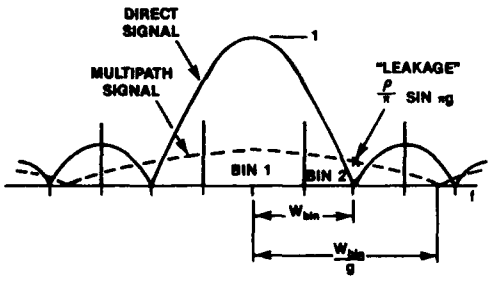


Fig. 11. Amplitude spectrum of direct and reflected signals in a binary channel. The transmitted signal is in bin 1 but spreading of the reflected signal causes power leakage into the adjacent bin.

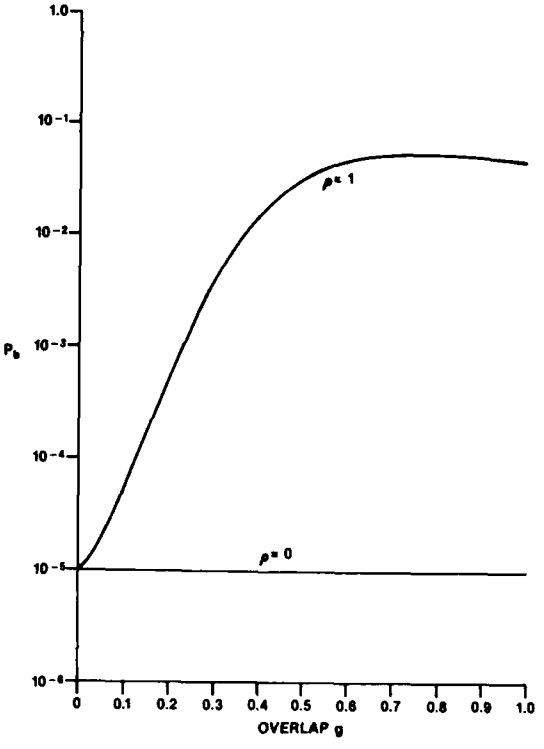


Fig. 12. P_b vs overlap g for SNR = 13.35 dB, no jamming, and binary signals.

INHERENT-JITTER EFFECTS IN THE NEW RAPID ACQUISITION METHOD FOR FREQUENCY-HOPPED PUSH-TO-TALK SPREAD SPECTRUM SYSTEMS : A JITTER-FIGHTING SCHEME

JORGE M. N. PEREIRA*
AUGUSTO A. ALBUQUERQUE
Instituto Superior Técnico
1096 LISBOA CODEX, PORTUGAL

Abstract

A critical problem in Push-to-Talk Frequency-Hopped Spread-Spectrum systems is its vulnerability to eavesdropping during the necessary preamble, or to possible reproduction by a would-be interferer [3].

A new Acquisition method has been envisioned, and is presented, to reduce this weakness, permitting the receiver to work in lower Signal-to-Noise Ratio than the usually refered [4],[10], with minimum complexity, and, even, to acquire faster than in any known system.

The essence of the new method is the use of Sliding-Window Integration non-Coherent Detectors and a Decision Rate much higher than the Hopping one, and coincident with the Window-Relocation Rate [8].

The Inherent-Jitter effects in the new Method are described, and results presented for an Optimal (**omniscient**) system, after "Jitter-Bypassing". A Jitter-Fighting scheme is, then, proposed and its performance and adequacy analised.

INTRODUCTION

Spread Spectrum (**SS**) techniques are, more and more, a valuable option in the Communication Systems field (see the recent **FCC** regulations on Spread Spectrum Communications).

This popularity is not entirely due, either to the immunity from interference it provides, or to the protection against eavesdropping, but, mainly, in the civilian environment, to the flexibility of usage of the already over-poluted Radio Frequencies Spectrum : transmission below noise level and multiple usage of the communication channel without prohibitiv interferences.

Let us just remember that, already in 1959, J. P. **COSTAS** [1] asserted : "... for congested-band operation, broad-band systems appear to offer a more orderly approach to the problem, and a potentially higher average traffic volume than narrow-band systems.".

Frequency-Hopped (**FH**) **SS** Systems are particularly vulnerable to jamming during the Code Acquisition stage, specially when communication is bursty and includes random periods of silence.

The resulting time uncertainty between transmitter and intended receiver (added to distance uncertainty in *mobile radio* units) implies the necessity of code synchronization for every single arriving message, independently of previous transmissions [4].

*This contribution was made possible through a grant from the "**Luso-American Foundation for Development**" complemented by an **INVOTAN** (NATO *'Science for Stability'* program) grant

Synchronization Preambles are one of the most effective Acquisition techniques [3] : it is, by far, the **least critical, easiest to implement, least complex, and best for all around use.**

Nevertheless, a critical problem in Push-to-Talk systems is their vulnerability during the necessary preamble. Therefore, Acquisition must be a rapid one, in order to minimize successful interference probability, i.e. to maximize the Probability of Overall Detection, P_D^{ov}.

The typical approach is to employ a Synchronization (*no data*) Preamble to be repeated a sufficient number of times (only sufficient) to ensure Acquisition, with an acceptably high probability, within some specified stopping time limit T_s [5], which can never exceed 1*s* (usually between 0.1 and 0.5*s*).

A new Acquisition Method has been envisioned, and is presented, to reduce this weakness of the Push-to-Talk system, permitting the receiver to work in lower SNRs than the usually refered, with minimum complexity, and, even more, to acquire at least one order of magnitude faster than in any known system [6].

The essence of the new Method is the use of one Hop time, T_H, Sliding-Window Integration non-Coherent Detectors, and a Decision Rate much higher than the Hopping one, and coincident with the Window-Relocation Rate [8].

During the study of the Acquisition system a problem arose : the existence of an **Inherent-Jitter**, due to the low SNRs, which could degrade performance, appropriate measures not being taken.

The Inherent-*Jitter* characterization enabled a first **Jitter-bypassing** approach, which led to the **Optimal Results** presented in **Section 1.3.2**, for an **Omniscient System**. The real (average) performance of a feasible system will be worst, but is, nevertheless, expected to outperform in celerity any known system with the same P_D^{ov}, with minimum complexity.

A **Jitter-Fighting** scheme was, then, considered, and is presented. Interesting is its adequacy to the Rapid Acquisition system under consideration (generically to any, even quick, *Jitter*-troubled system), once the *Jitter* characterized.

This paper is organized as follows. In the first Section, the new Rapid Acquisition Method is presented and the Inherent-*Jitter* analised. Optimal results are presented for an Omniscient system after "*Jitter*-bypassing". The Inherent-*Jitter* degradation in feasible systems is shown.

In the following Section, a *Jitter*-Fighting scheme is proposed, and the performance enhancement obtained is analised.

Finally, conclusions are drawn on the interest and praticability of the new Acquisition Method and of the *Jitter*-Fighting scheme.

1 INHERENT-JITTER IN THE ACQUISITION PROCESS

1.1 BRIEF REVIEW OF CLASSICAL METHODS

The acquisition receiver is typically a **Two-Dwell** system, as defined in [2], discarding quickly most of the incorrect cells. It employs a **Matched-Filter (MF)**, implementing the **Passive Correlation**, consisting of M non-Coherent Detectors[1], each one producing a binary decision by comparison of the Detector output to its Threshold (all equal). The individual decisions are accumulated and compared to a second, Summation Threshold. Once the latter is exceeded the system enters the **Verification Mode - Active Correlation** : the local code starts hopping with the incoming one. Otherwise the next candidate code offset is examined [9].

To verify Sync indications, the Verification Mode performs **A** independent tests, each associated with a MT_H delay. If at least **B out of A** are positive, Synchronization is declared and the Tracking Loop is activated. Otherwise Passive Search is resumed. The cost of a First-Dwell False Alarm is, therefore, AMT_H seconds.

This system is specified (optimized) by the choice of the parameters M, A and B, and by the determination of the two Thresholds 4 .

POLYDOROS , WEBER presented a better and unified approach to Serial Search Code Acquisition, applied to Direct Sequence (**DS**) **SS** systems, permitting a quicker Acquisition.

[1] M is a design parameter, a trade off between decision reliability and cost and complexity [4]

As expressed in [9], the essence of Rapid Acquisition through Passive Correlation is an MT_c Observation Time per decision, while the Decision Time, T_D, is just a small fraction of the Chip Time, $T_D = \Delta T_c$ ($\Delta = 2^{-n}$).

The results presented are conclusive : their method is, indeed, an important contribution, and the obtained performance is a goal (yet) to be achieved, specially in FH systems, where things are not that "easy" to implement.

Another quite different approach, although more suitable for permanent communication links, is the Sequential Acquisition of the PN code [12],[7]. Here one tries to minimize the Acquisition Time using a special technique : Wald Sequential Detection. The objective is to set the Acquisition Time, with a very high probability, within the Preamble duration T_s.

The essence of Sequential Detection, which enables so quick an Acquisition, is that the Mean Time to Dismiss an Incorrect Cell is smaller than in any other detector, once the Detector is designed so that the output of the post-detection integration, under False Sync conditions, decreases at an average rate toward a Dismissal Threshold [12].

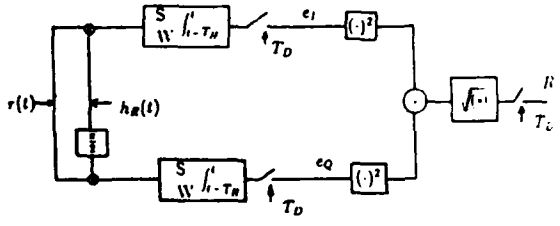
One problem of this method is that, for a given Design Point SNR, if the Operating Point SNR decreases to zero when *Signal + Noise* is being evaluated, the Average Sample Number necessary to make a decision **increases significantly**, *at first*, and, *then*, **most of the decisions become incorrect** : mostly missed detections occur [13],[11].

The major problem, however, is its complexity, specially in a Rapid Acquisition environment.

1.2 THE NEW RAPID ACQUISITION METHOD

As it was already suggested, the problem with **FH-SS** systems is that, in order to accelerate the decision process, one will have to work with lower post-Detection SNRs.

In order to overcome this difficulty, we will use Sliding-Window T_H integration non-Coherent Detectors, whose output will be analised at a much higher rate, $T_D^{-1} = NT_H^{-1} (N = \Delta^{-1} \gg 1)$, coincident with the Window-Relocation one :



with

$$r(t) = \sqrt{2S}\cos(\omega_k t + \Theta_k)\, rect\left(\frac{t - kT_H}{T_H}\right) + n(t) \quad , \quad (1)$$

$$h_R(t) = \sqrt{2}\cos\omega_k t \quad , \quad (2)$$

Figure 1: Sliding-Window non-Coherent Detector

where

- S - Signal Power at the receiver;
- ω_k - angular frequency of the K^{th} tone;
- Θ_k - random phase, uniformly distributed in $[0, 2\pi]$;
- $n(t)$ - Aditiv White Gaussian Noise.

This will enable us to have, with such a Rapid Decision Rate, a post-Detection SNR, γ_D, even equal to the received SNR, $\gamma_H = E_H/N_o = ST_H/N_o$, where

- E_H - one hop Signal energy;
- N_o - single-sided power spectral density of the AWGN.

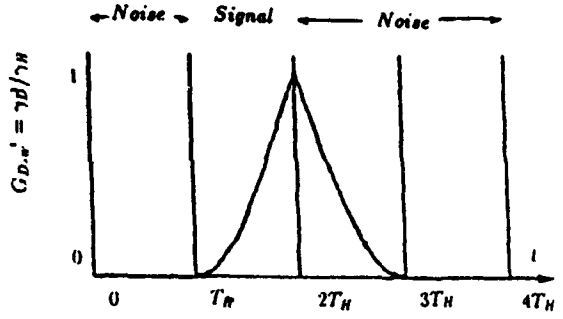


Figure 2: Demodulation Gain

As a matter of fact, the expression of the Demodulation Gain is

$$G_{D_{sw}} = \left(\frac{T_{ef}}{T_H}\right)^2 \qquad (3)$$

where T_{ef} is the Effective Integration Time, and whose evolution is reflected in the following *Figure* :

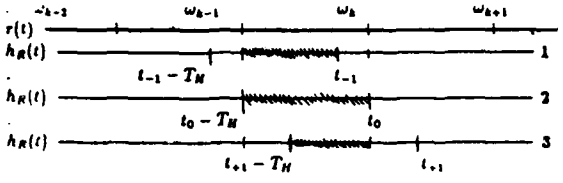


Figure 3: T_{ef} definition

From *Fig. 2*, it is obvious that decisions must be taken with a variant γ_D with **Signal Present** and even with **Noise Alone** after a "correct" tone. Therefore, a careful data processing is needed to achieve correct decisions.

Let us, now, enunciate the characteristics of the system we are going to describe. Our system will be Two-Dwell, hybrid Passive(Matched-Filter)/Active(Verification Block) *Integration*, *Unanimous* (Matched-Filter)/Majority(Verification Block) Decision Logic [9].

In order to permit the receiver to work in low SNR environments and to maximize the Probability of Overall Detection, P_D^{Ov}, during the Preamble, the new Acquisition System will consist of a special **Pre-Synchronizer Matched-Filter** (M Passive non-Coherent Sliding-Window Detectors) and a **Verification Block** (*at least* one Active non-Coherent Detector and associated Decision Logic).

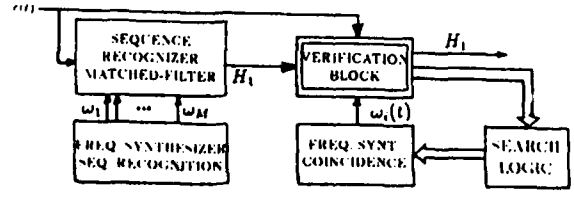


Figure 4: Block Diagram of the Rapid Acquisition Receiver

All the following considerations are made under the assumption that a Block of C Active Verification Detectors is available, C being such that no Tentative Sync indication is lost, once there will always be one idle Detector [10].

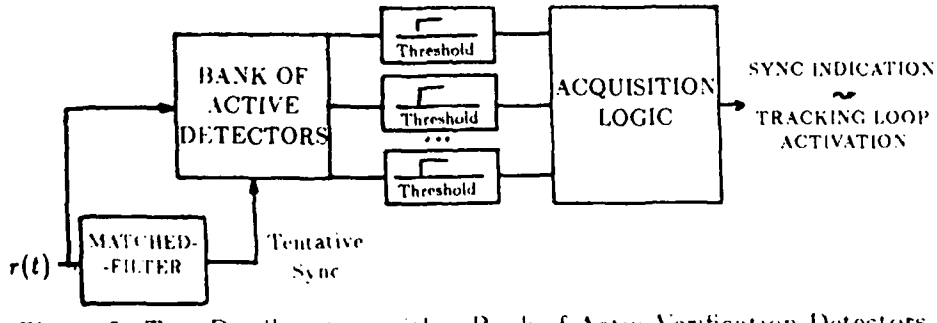


Figure 5: Two-Dwell system with a Bank of Activ Verification Detectors

Pertinent considerations on this optimal behaviour are drawn in [6]. For the moment, it is enough to say that C is an affordably small number of Detectors, once "wrong" Sync indications will have a very low probability.

1.2.1 PRE-SYNCHRONIZER MATCHED-FILTER

The Pre-Synchronizer Matched-Filter detects a short subsequence (length M) of frequencies of the Preamble (Recognition sub-Sequence, RS) and can be designed to achieve a Probability of False Recognition $(P_{FR})_{MF}$ and a Probability of Missed Recognition $(P_{MR})_{MF}$ almost as low as wanted.

This is achieved by setting the appropriate Thresholds for the Sliding-Window non-Coherent Detectors, and the number M of Detectors of the Matched-Filter, therefore the length of the RS.

The Pre-Synchronization function of the Matched-Filter stems from the Sliding-Window Integration over a period T_H, window changing at a rate much higher than the Hopping one : Decision/Changing Rate $R_D = NR_H$, with, tipically, $N = 2^n >> 1$, and $R_H = T_H^{-1}$.

In order to implement the Sliding-Window Integration, a very simple scheme can be used :

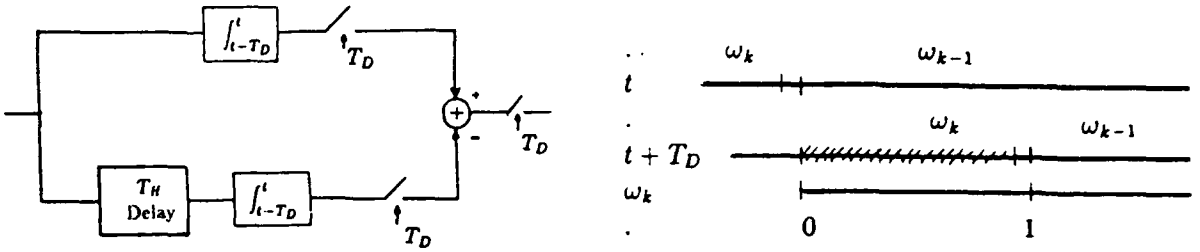
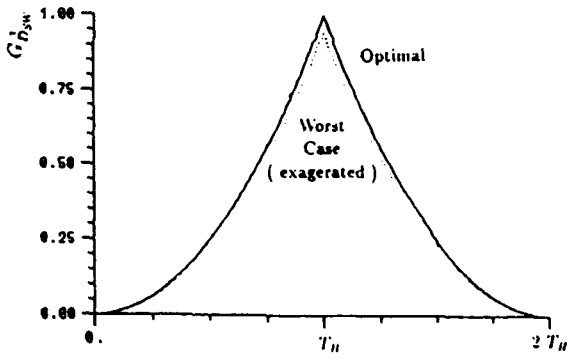


Figure 6: Sliding-Window Integration; T_{ef} imprecision

An imprecision up to $T_D/2$ results from this implementation. Therefore, the Demodulation Gain is slightly degraded (once $T_D \ll T_H$).



The output of the Sliding-Window non-Coherent Detector is processed in order to detect the end of (each frequency - Tone - of) the Recognition sub-Sequence.

Figure 7: $G_{D_{SW}}$: Optimum and Worst cases (exagerated)

Once we want to work in low SNRs, the necessary signal processing is almost impressive, although the always growing computational power makes it feasible, if not straight forward.

First of all, we must realize that the R samples at the output of the Sliding-Window non-Coherent Detector have a Rayleigh (**Noise Alone**) or Rice (**Signal Present**) distribution.

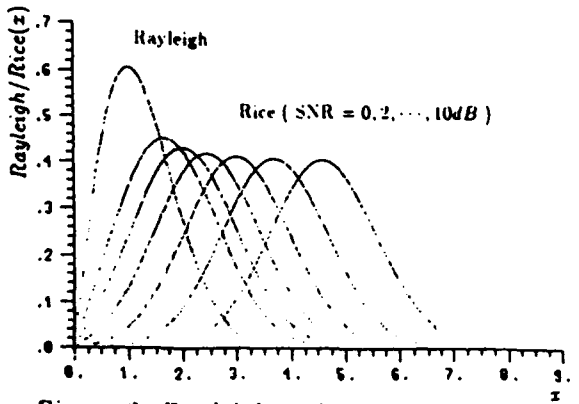


Figure 8: Rayleigh and Rice Distributions

The curves above refer, obviously, to the same Noise Single-Sided Power Spectral density.

One sample is not enough to decide on Signal Presence/Absence. So, we will begin by adding, over a period T_H, all R samples, as it is shown immediatly.

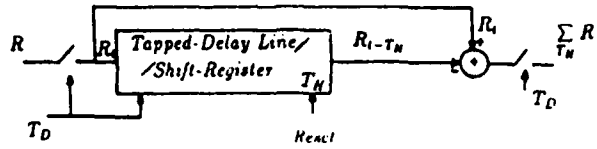


Figure 9: T_H Accumulator

Let us study the behaviour of the system for a given tone sequence, including one "correct" tone :

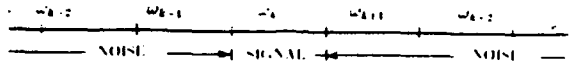


Figure 10: Tone Sequence to be analised

The respective Demodulation Gain is:

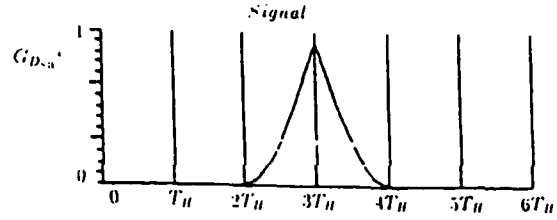


Figure 11: Demodulation Gain for the chosen Tone Sequence

The output of the Accumulator follows directly, with an "average behaviour" given by

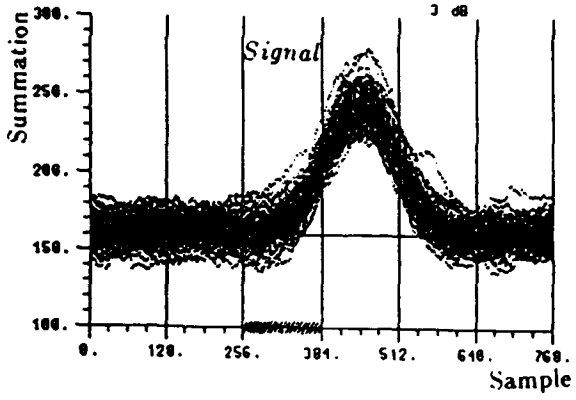


Figure 12: Accumulator Output - $\gamma_H = 3dB$ (128 samples/ T_H)

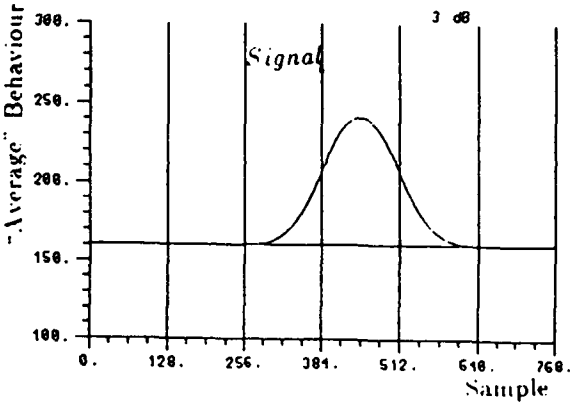


Figure 13: Accumulator "average" Behaviour

The curves, obtained by **1024** runs of the chosen Tone Sequence, have to big an imprecision to enable a good decision.

Further processing is necessary. We used **Least Square Linear Regression** followed by **Smoothing**.

The result was :

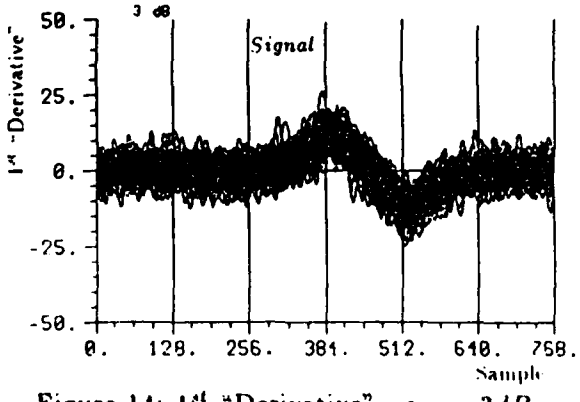


Figure 14: 1st "Derivative" - $\gamma_H = 3dB$

as expected.

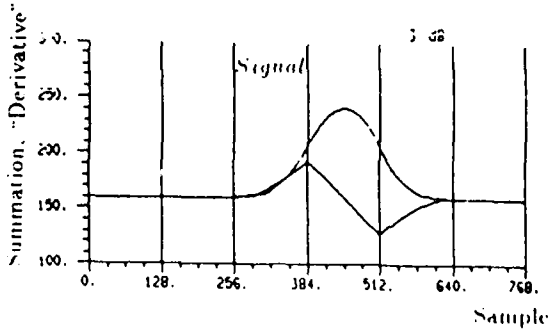


Figure 15: Expected "average" Behaviour of the 1st "Derivative"

Once more, no "correct" decisions are to be expected.

Repeating the above process, we obtain

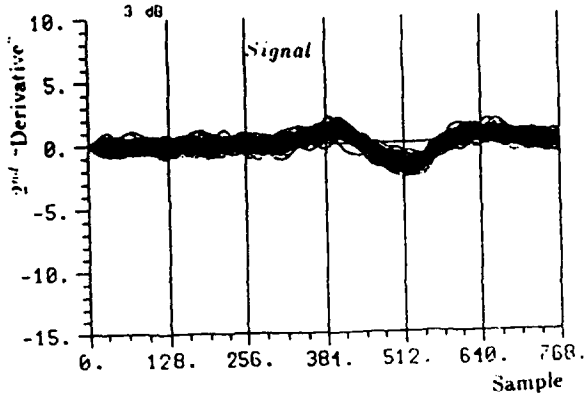


Figure 16: 2nd "Derivative" - $\gamma_H = 3dB$

with a little distortion and some delay refering to what was expected.

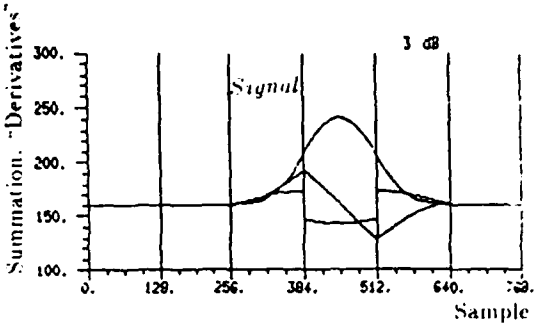
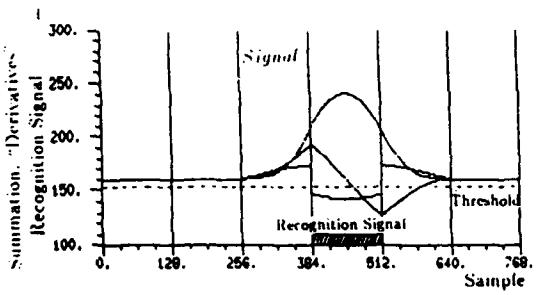


Figure 17: Expected "average" Behaviour of the 2nd "Derivative"

Now, we have a signal on which decisions can be taken, once a **Threshold** is defined. Threshold setting was done through successive trials.



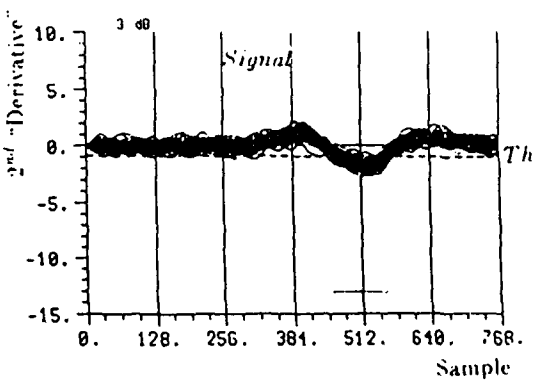
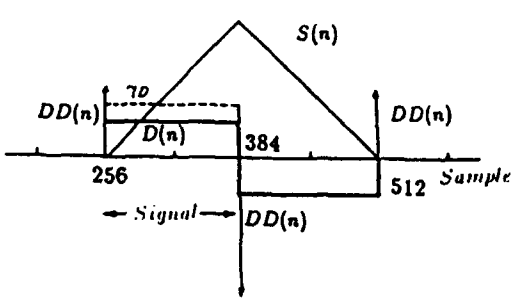


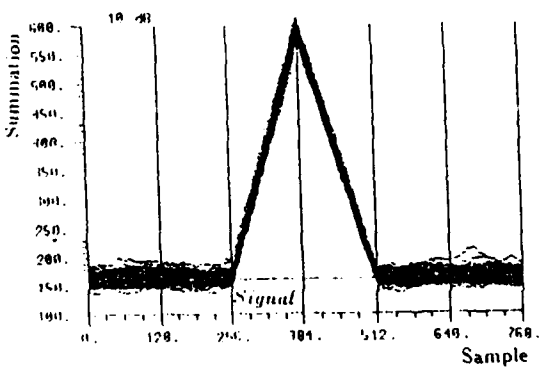
Figure 18: Summation, "Derivatives", Threshold and the Recognition Signal

The idea behind this scheme came arose from similar systems working with very high SNRs. There, the rapid Decision Rate was affordable : the post-Detection SNR was high enough.

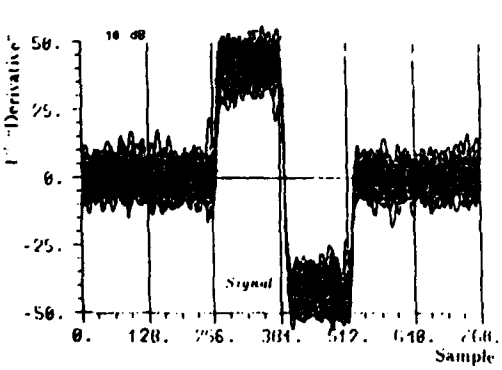
The next *Figures* presents the results at all stages for $\gamma_D = 10dB$.



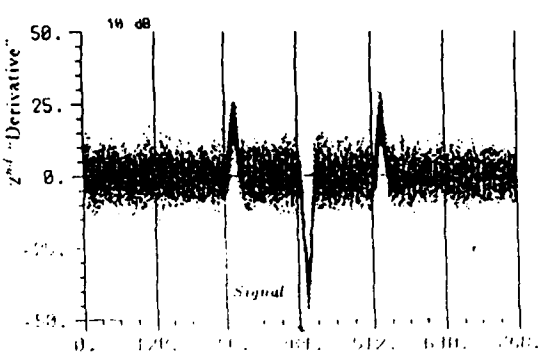
a) Theoretical Approach;



b) Accumulator Output;



c) 1st "Derivative".



d) 2nd "Derivative"

Figure 19: High SNR performance - $\gamma_D = 10dB$

The Block Diagram of the Tone Recognizer is, then :

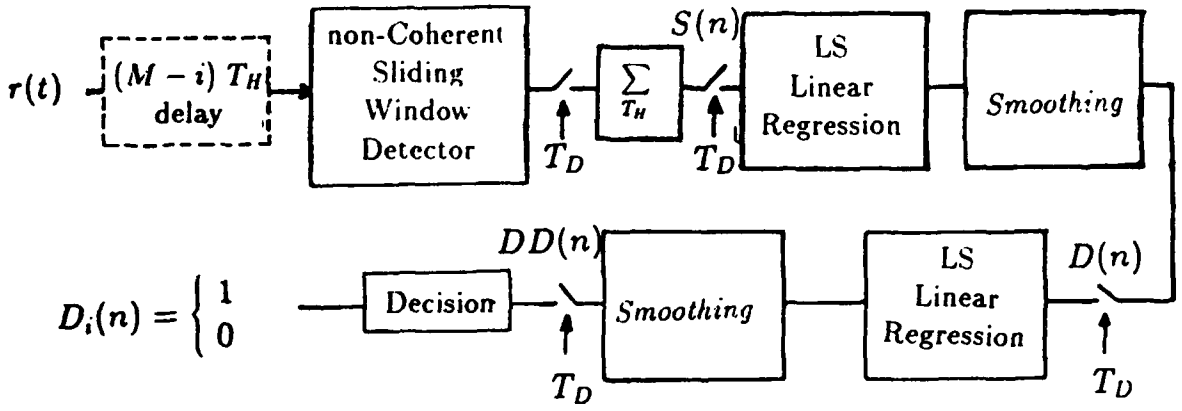


Figure 20: Tone Recognizer Block Diagram

In fact, the "slow" Hopping Rate enables the computation of the Linear Regressions. However, if time constraints exist, the Linear Regression may be substituted by a less time consuming algorithm, just sufficiently accurate.

This new Sliding-Window method achieves the Detection of each Tone of the *RS* with a certain delay, within a very little fraction of T_H (nT_D, with $n << N$ - *Fig. 16*), through the comparison of *DD(n)* with the Recognition Threshold : Inherent-Jitter. Let us just say that the distribution of the Tone-Recognition Sample, S_{TR}, is, approximately gaussian.

The simultaneous consideration of *M* channel (Tone Recognizer) indications,

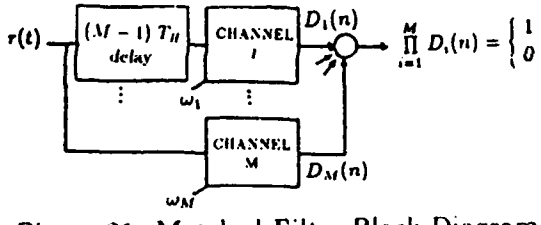


Figure 21: Matched-Filter Block Diagram

will lead to the recognition of the RS with a very high precision, once the decision is unanimous : in order to recognize the RS, all channels must recognize their respective Tone.

Sequence-Recognition is achieved, obviously, only when the latest Tone-Recognition occurs .

The distribution of the Sequence-Recognition Sample, S_{SR}, assuming a normalized gaussian distribution of the Tone-Recognition Sample, follows, as a function of the number of Channels, *M* .

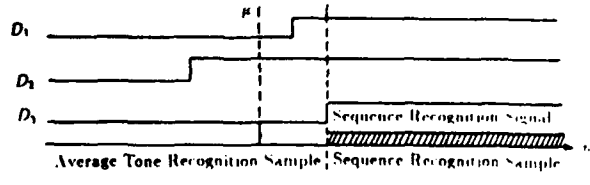


Figure 22: Sequence-Recognition Sample

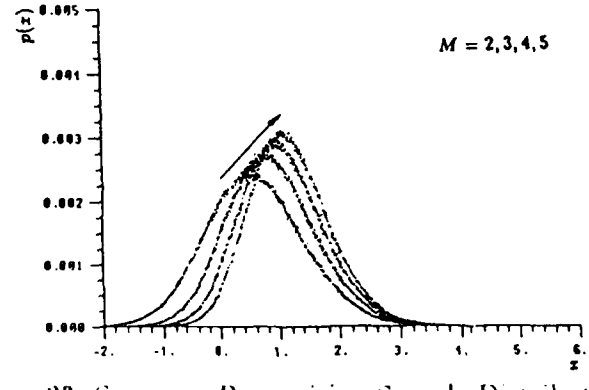


Figure 23: Sequence-Recognition Sample Distribution

From the above *Figure* one can see that the Sequence-Recognition Sample will be (coherently) delayed refering to the individual Tone-Recognition Samples, and will have a smaller variance - an interesting feature to be exploited in the following.

Now that Sequence recognition was achieved, the system will enter the Verification Mode.

1.2.2 VERIFICATION BLOCK

The Verification Block associated with the Active non-Coherent Detector performs the validation of the Alarms triggered by the Matched-Filter through a pair of tests, represented by finite state Markov chains with absorbing boundaries [13], designed to achieve maximum Probability of Acquisition, $P_{Acq} \equiv P_{Conf|H_1}$, and minimum Probability of False Alarm, $P_{FA} \equiv P_{Conf|H_0}$. The tests are performed at the same high rate, R_D.

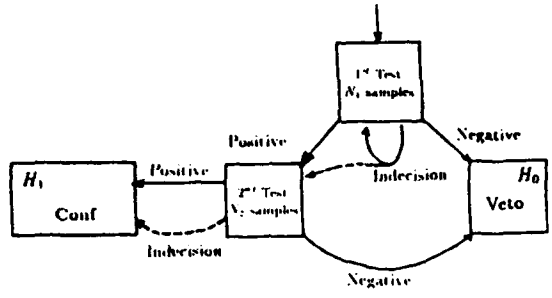
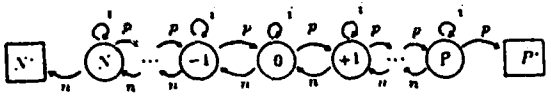


Figure 24: Verification Test

Verification will consist in the (eventual) detection of *coincidence* between the incomming Tone Sequence and the expected one ($\omega_{M+1}, \omega_{M+2}, \cdots$: Verification Test ≡ Coincidence Test.

Some remarks are necessary. Verification must be done quickly. Therefore, N_1 and N_2, number of Samples to be processed in each test, must be a fraction of N ($NT_D = T_H$). We chose $N_1 + N_2 = N$ and $N_1 < N_2$, whereby we expect to "get rid", quickly, of most False Recognitions, confirming in a longer Test all positive indications.

Another aspect, closely related to the Test implementation,



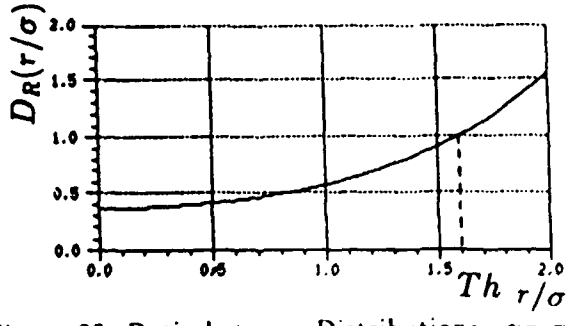
is the treatment of the Indecision hypothesis : in the First Test it determines its repetition; in the Second it will be considered a Positive (non-Negative) indication.

Figure 25: Validation Test State Transition Diagram

But the State Transition Diagram, itself, considers an Indecision Probability : the reason stems from the low SNRs one is expected to work with.

As a matter of fact, the individual decisions on the R Samples at the output of the Validation Active Detector are based on a Likelyhood Test, a Ratio between Rice and Rayleigh probabilities :

$$D_R(r) = \exp^{-\gamma_D} I_0\left(\frac{2r}{S}\gamma_D\right) \tag{4}$$



Th is the minimum of the Sample values corresponding to a probability of **Signal Present** bigger than that of **Noise Alone**. But, for low SNRs

$$Pr\{R > Th|H_1\} < Pr\{R < Th|H_1\} \ , \tag{5}$$

indeed an indesirable situation.

Figure 26: Ratio between Distributions - $\gamma_D = 1$

Therefore, one must find another threshold, Th_{inf}, so that

$$Pr\{R > Th|H_1\} > Pr\{R < Th_{inf}|H_1\} \ , \tag{6}$$

defining an Indecision Range as shown :

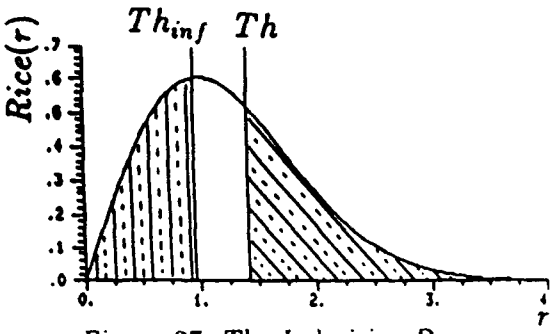


Figure 27: The Indecision Range

So, refering back to the State Transition Diagram, *Fig. 25*,

- p - probability of *positive* individual decision ($R > Th$);
- n - probability of *negative* individual decision ($R < Th_{inf}$)
- i - probability of *indecision* ($Th_{inf} < R < Th$).

Let us precise the *modus operandi* of the Verification Mode. It was already said one will process the R Samples at the output of the Active Detector.

What is the situation at the end of the **RS** ?

The Active Detector is already positioned (waiting) at the ω_{M+1} tone, and, once Sequence-Detection occurs with a delay refering to the end of the **RS**, the post-Detection SNR will be

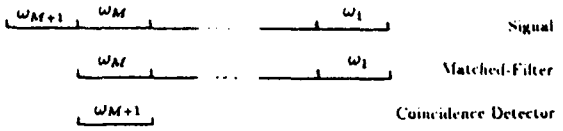


Figure 28: Frequency positioning of the Detectors

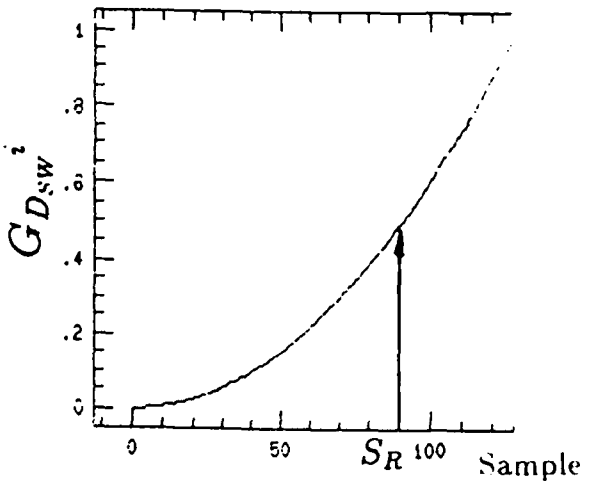


Figure 29: γ_D evolution after the end of Recognition sub-Sequence

All individual decisions in the State Transition Diagram will have to be taken with a variant γ_D', therefore with variable/adjusting thresholds $(Th, Th_{inf})_{\gamma_D'}$.

Our objective is to make all the decisions with the best γ_D' possible.

1.2.3 SERIAL SEARCH

The repetitive structure of the Synchronization Preamble is perfectly suited for **Straight Serial Search** of the *Uncertainty Region* [9].

The generic Circular State Transition Diagram of Straight Serial Search follows :

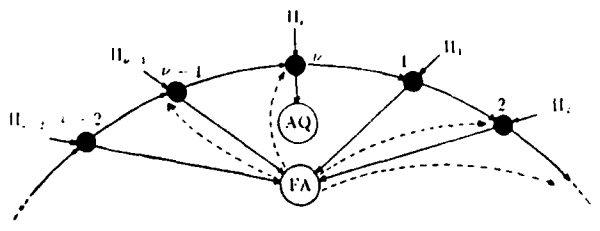


Figure 30: Straight Serial Search Circular State Transition Diagram

where

- **AQ** - Acquisition State;
- **FA** - False Alarm State;
- ν - length of the Preamble Code;
- Π_i - *a priori* probability of begining the Preamble at the ω_i^{th} Tone.

The False Alarm state, corresponding to a prohibitive Tracking Mode Activation (from an Acquisition time viewpoint), was considered an *absorbing state* in the State Transition Diagram of the discrete time Markov process representing the Serial Search triggered by the Matched-Filter [9].

Expliciting the Verification Test in the State Transition Diagram, we obtain :

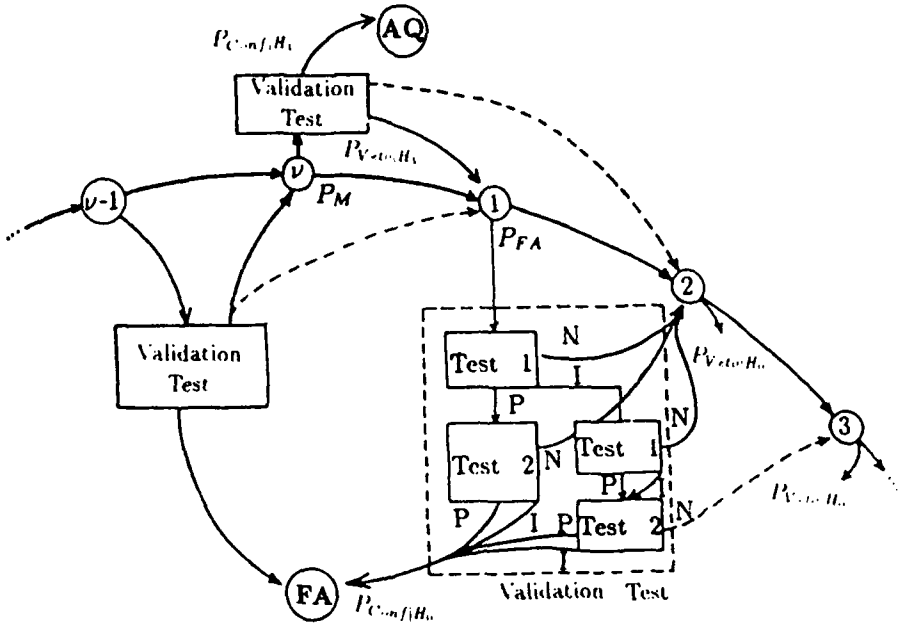


Figure 31: State Transition Diagram : Verification Test

ν is the State corresponding to a "correct" Tentative Sync indication (H_1), i. e. to the **RS**.

In every Verification Test, if an **Indecision** occurs at the end of the First Validation Test, which is consequently repeated, it may happen that a *Veto* occurs at the end of the Second Validation Test. In that case, and only then, the commulative duration of all those tests exceeds T_H.

Would it happen in the $(\nu - 1)^{th}$ State, a chance of Acquisition would be lost, had no diversity been antecipated (**Section 1.2** - *Fig. 4*).

But, as it will be shown lately, the probability associated to this event is negligible comparing to all other involved probabilities. Therefore, nothing will be lost if one considers only the "significant" State Transitions.

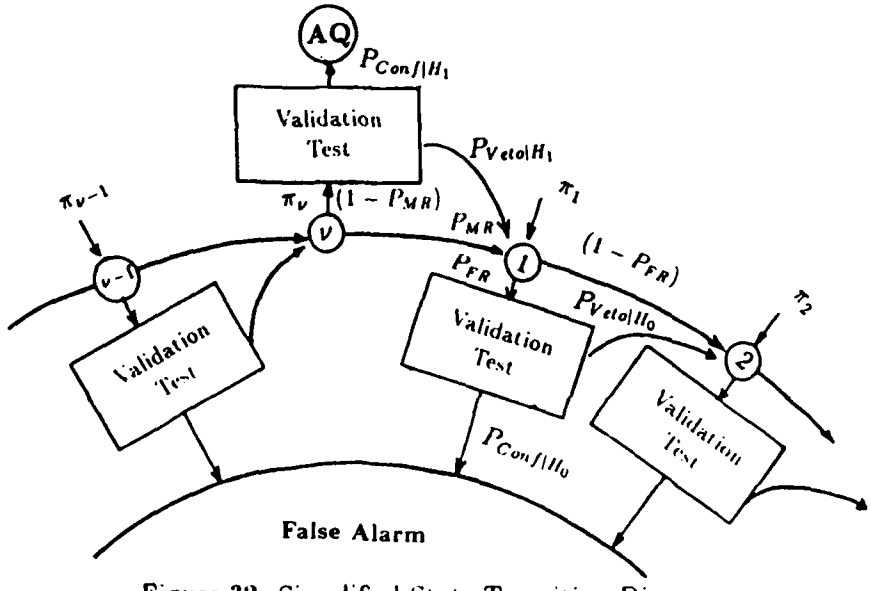


Figure 32: Simplified State Transition Diagram

Allowing for a certain number of repetitions of the Preamble sequence, whose length L_P is usually much smaller than that of the coded sequence of frequencies used in normal transmition, L, one can obtain a Probability of Overall Missing, $P_M^{(ov)} = 1 - P_D^{(ov)}$, almost as low as whished, by adjusting the Preamble sequence length P, the Preamble duration T_S, the Decision/Analysis Time T_D, and the precision of the Linear Regression calculation, all of this in very low SNRs.

For a certain time limit, T_S, the maximum number of repetitions is $R_{max} = T_S/(L_P T_H)$. We will consider R_{max} as the number of complete repetitions of the Preamble, after a first run of the RS.

The **False Alarm Probability per H_0 cell** is

$$P_{FA,r} = P_{FR} \cdot P_{conf|H0} \quad , \tag{7}$$

and the **Detection Probability for one run of the H_1 region** is

$$P_{d,r} = (1 - P_{MR}) \cdot P_{conf|H_1} \quad . \tag{8}$$

The Probability of Overall Detection, $P_D^{(ov)}$, following [9], will be

$$P_D^{(ov)}|^i = \underbrace{P_{d,r}(1 - P_{FA,r})^{L-i}}_{\mathbf{0}} + \sum_{n=1}^{R_{max}} \underbrace{P_{d,r}(1 - P_{FA,r})^{L-1}}_{\mathbf{I}} \underbrace{(1 - P_{d,r})(1 - P_{FA,r})^{L-i} \prod_{0}^{n-1}(1 - P_{d,r})(1 - P_{FA,r})^{L-1}}_{\mathbf{II}} \quad , \tag{9}$$

where

- **0** is the Detection Probability after the first incomplete run of the Preamble, which begins at the i^{th} Tone;
- **I** is the Probability of Missed Detection in the first run;
- **II** the Detection Probability in the n^{th} run of the Preamble.

The hypothesis $i = 1$ is the **worst case** : there are more chances of False Alarm ($\Pi_1 = 1; \Pi_i = 0, i \neq 1$). In the **uniform distribution** hypothesis all the values of i give their contribution ($\Pi_i = 1/L_P, i = 1, \ldots, L_P$) [9].

Final expressions for $P_D^{(ov)}|_{Worst}$ and for $P_D^{(ov)}|_{Uniform}$ are available in [6].

1.3 INHERENT-JITTER EFFECTS

1.3.1 JITTER CHARACTERIZATION

As it was already stated, Tone-Recognition is obtained with a certain delay refering to the end of the Tone, and within a certain Sample range.

S_{TR} $\gamma_H(dB)$	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5	6	7	8	9	10
Mean	109	108	102	97	93	91	87	84	82	80	77	74	72	70	68	67
St. Dev.	24	22	20	18	17	15	13	12	10	10	8	7	6	5	4	3

Table 1: Tone-Recognition Sample Distribution

The following *Figure* shows explicitly the characteristics of the Tone-Recognition Sample Distribution.

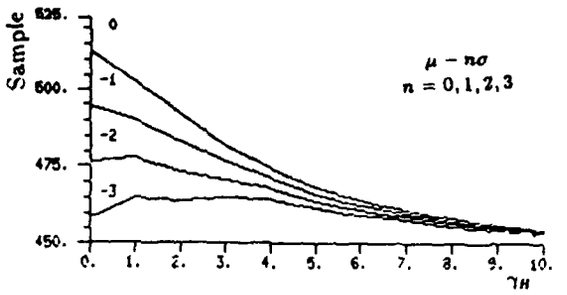


Figure 33: Tone-Recognition Sample Distribution Characteristics

The Mean of the Distribution, $\overline{S_{TR}}$, decreases with γ_H, as expected:

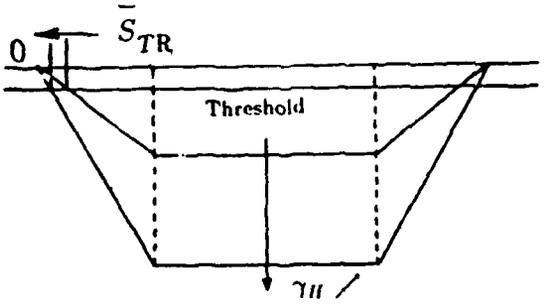


Figure 34: Evolution of $\overline{S_{TR}}$

and the Standard Deviation follows closely.

As it is evident from the above *Figure*, the Mean is a function of the Threshold.

One can not make the delay smaller without reducing the Threshold, consequently increasing the number of **False Recognitons**.

It was empirically found that the Threshold should be

$$Threshold \approx \mu(Noise\ Peak) + 1.8 \times \sigma(Noise\ Peak) \ , \tag{10}$$

where *Noise Peak* refers to the maximum value of $DD(n)$ during an Hop time, for **Noise Alone** and for **1024** runs of the chosen Tone Sequence.

For this Threshold, the following **False Tone-Recognitions** and **Missed Tone-Recognitions** occurred in **1024** runs.

$\gamma_H(dB)$	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5	6	7	8	9	10
Missed Recog.	112	67	40	7	3	0										
False Recog.	43	53	46	45	51	49	53	48	45	53	49	52	47	47	53	43

Table 2: **False** and **Missed** Tone-Recognitions

Missed Tone-Recognitions are determinant : once unanimity is the rule, one Missed Recognition is enough to miss the **RS**.

The Tone-Missing probability, $P_{TM_{(\cdots)}}$, is

$\gamma_H(dB)$	0	0.5	1	1.5	2
$P_{TM_{(\cdots)}}(10^{-3})$	109	65	39	7	3

Table 3: Tone-Missing Probability

and may be extrapolated for greater values of γ_H as follows :

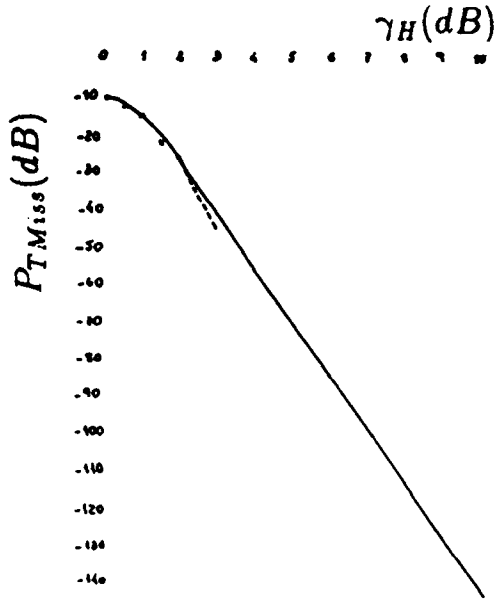


Figure 35: Tone-Missing Probability (extrapolation)

The consequent Sequence-Missing Probability Missed Recognition) is

$$P_{MR} = 1 - (1 - P_{TMiss})^M \quad , \tag{8}$$

and, for $M = 3$, we obtain :

$\gamma_H(dB)$	0	0.5	1	1.5	2	* 3	4	5	6	7	8	9	10 *
$P_{TMiss}(dB)$	-10	-12	-14	-22	-25	*-41	-55	-70	-85	-99	-113	-128	-142*
$P_{MR}(dB)$	-5	-7	-9	-17	-21	*-36	-50	-65	-80	-94	-108	-123	-137*

* → extrapolation ← *

Table 4: Missed Recognition Probability

Refering, now, to False Recognitions, due to the unanimity rule, they will only happen when all M channels wrongly decide for **Signal Present**[2].

After a somehow heuristical approach [6], one can situate the False Recognition Probability :

$$0.38 \times 10^{-7} < P_{FR} < 1.95 \times 10^{-7} \tag{12}$$

for $M = 3$.

As one can see, $M = 3$ (minimum complexity) led to very interesting results, and no more channels are necessary.

1.3.2 OPTIMAL RESULTS BY *JITTER-BYPASSING*

Fig. 29 shows the evolution of $\gamma_D{}'$ after the end of the RS, and also our problem : once S_{SR} follows the distribution of *Fig. 23*, how can one decide when to move to the next Tone (ω_{M+2}), in order to maintain the maximum $\gamma_D{}'$?

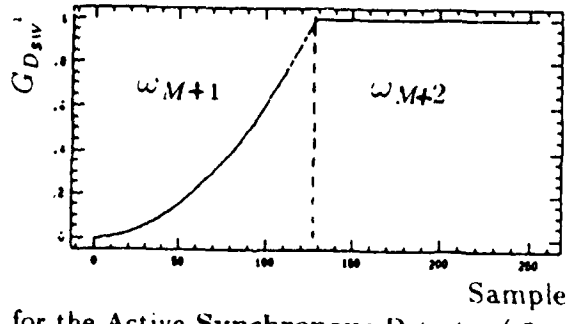


Figure 36: $G_{D_{SW}}$ for the Active **Synchronous** Detector (Omniscient System)

[2]In order to minimize the Probability of False Recognition, P_{FR}, one should **not use any Tone** of the **Preamble** in normal communications, not to make it easier a simultaneous (that way favoured) "mistake".

To find a solution, let us review the implementation of the Verification Mode.

The Validation Tests will be done with a variant γ_D', therefore with variant thresholds, $(Th, Th_{inf})_{\gamma_D'}$.

So, using the worst case thresholds, $(Th, Th_{inf})_{W.C.}$, corresponding to the event of Sequence-Recognition 2 $\sigma_{\overline{S_{SR}}}$ in advance refering to the $\overline{S_{SR}}$, one can be sure to take "all" individual decisions with appropriate thresholds[3].

Using this *low* thresholds, one must expect a somewhat higher False Confirmation probability, $P_{Conf|H_0}$, but, due to the small value of $\sigma_{\overline{S_{SR}}}$, nothing excessive is foreseenable.

Once one is working with low γ_H, the Validation Tests, *Fig. 25*, must not be symmetrical, in order to enable a higher Correct Confirmation probability, $P_{Conf|H_1}$. This gain goes with a similar, but acceptable, one of $P_{Conf|H_0}$.

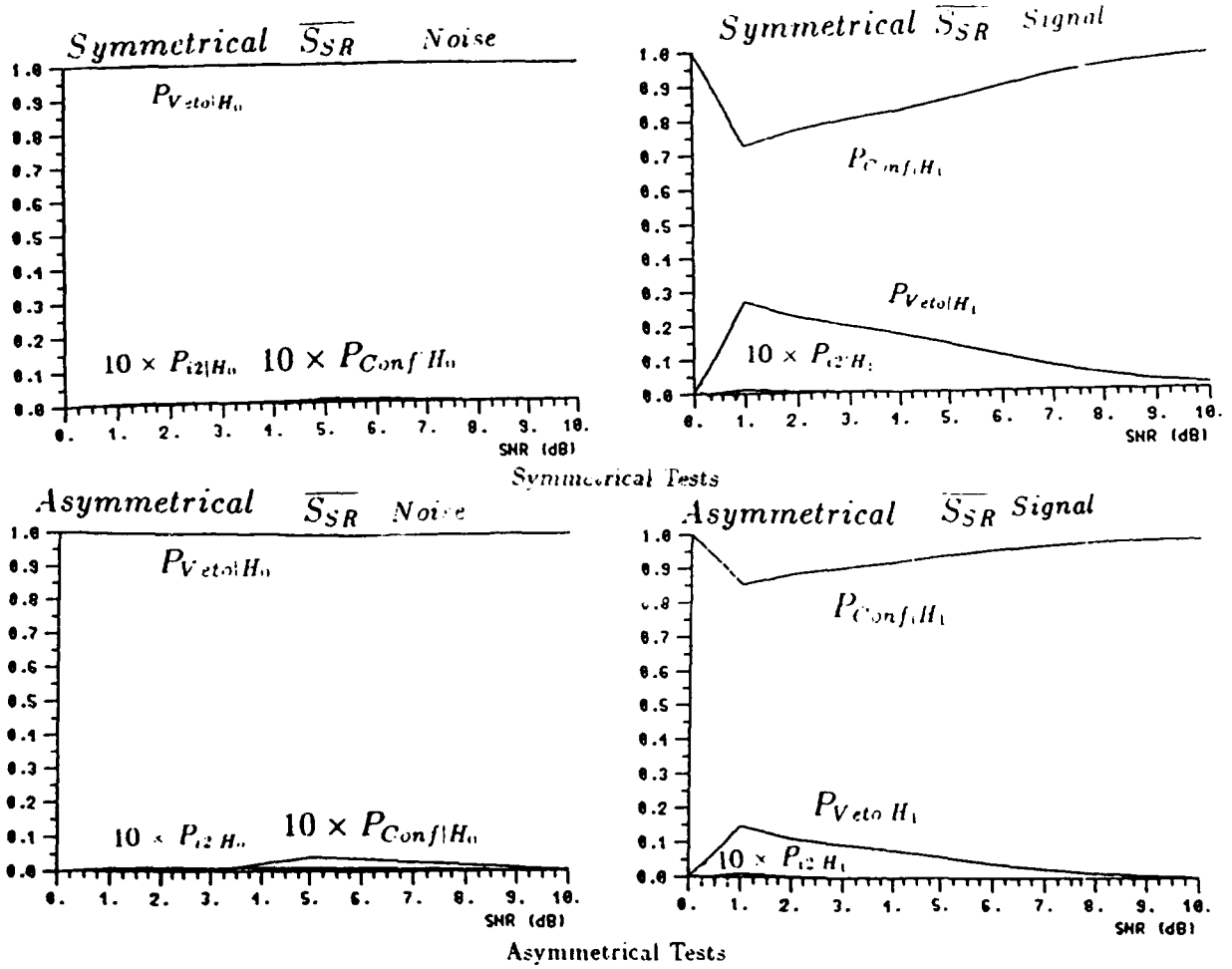


Figure 37: Probabilities for $\overline{S_{SR}}$

The above *Figure* shows the obtained Validation probabilities assuming a recognition coincident with the Mean of Sequence-Recognition Sample distribution, using the appropriate thresholds, and not the worst case ones.

Using, now, the thresholds suited for the Reference Sample, $S_{R,f} = \overline{S_{SR}} - 2\,\sigma_{\overline{S_{SR}}}$, one obtains, for the Mean Sample, $\overline{S_{SR}}$, better probabilities :

[3] Refering back to *Fig. 23*, the Distribution asymmetry assures a minimum probability for higher advances than 2 $\sigma_{\overline{S_{SR}}}$.

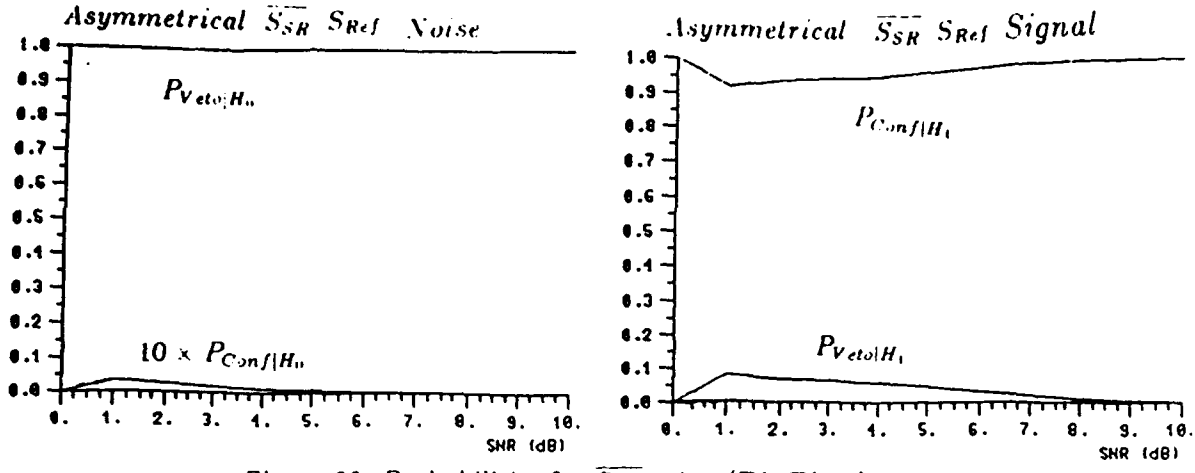


Figure 38: Probabilities for $\overline{S_{SR}}$ using $(Th, Th_{inf})_{S_{Ref}}$

Although the thresholds refer to the Reference Sample, advanced Recognitions will lead to poorer performances, but this will happen to only a few ones.

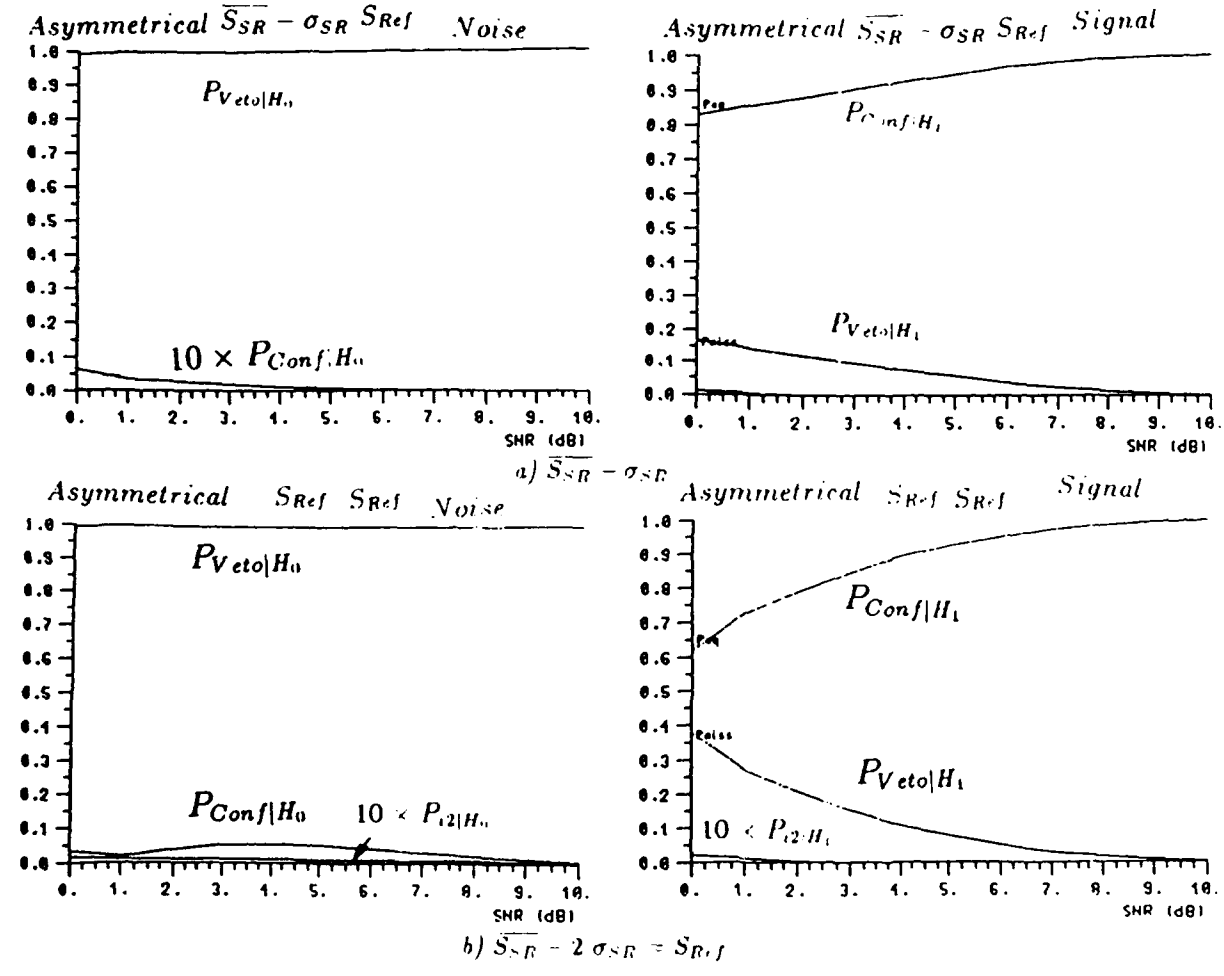


Figure 39: Probabilities for advanced Recognitions using $(Th, Th_{inf})_{S_{Ref}}$

Let us note that *Fig. 38* can be considered a good estimate of the "*average*" Validation *probabilities* of the Coincidence Test, while *Fig. 39-b)* is a good estimate of the *worst case probabilities*.

Finally, we must remember that all this probabilities were obtained assuming an Omniscient System, able to change conveniently its Tone, in order to maximize γ_D.

The results for the Verification Block, obatined for a Push-to-Talk FH-SS system working at $1Khop/s$, with $N = 128$ samples per hop, at $\gamma_H = 3\ dB$, and detecting a 3 frequencies Recognition sub-Sequence ($M = 3$), were :

- $P_{Verif|H_1} > 0.935$;
- $P_{Verif|H_{nc}} < 2 \times 10^{-3}$;

with a First Validation Test with 5 and a Second one with 11 states, being $N = P + 2$ (Asymmetrical Tests).

With the approximation already refered, the Verification does not take more than on Hop time (Verification Time $T_V \leq T_H$).

The Pre-synchronization is established with a standard deviation $\sigma = 5.6$ samples, i.e. $n \simeq 6$, for 16 and 48 points considered in the Least Square Linear Regressions computation, respectivelly ($n \ll N$, as wanted).

For $T_S = 1s$ and for a Preamble code length $L_P = 128$, and assuming *uniform distribution* of the Preamble begining frequency [9], $P_{MF}^{(n)} < 10^{-7}$.

1.3.3 PERFORMANCE DEGRADATION

What happens due to the Inherent-*Jitter* ?

The evolution of $\gamma_D{}'$ will suffer some degradation, leading to Validation tests using thresholds suited for higher $\gamma_D{}'$ than the available ones :

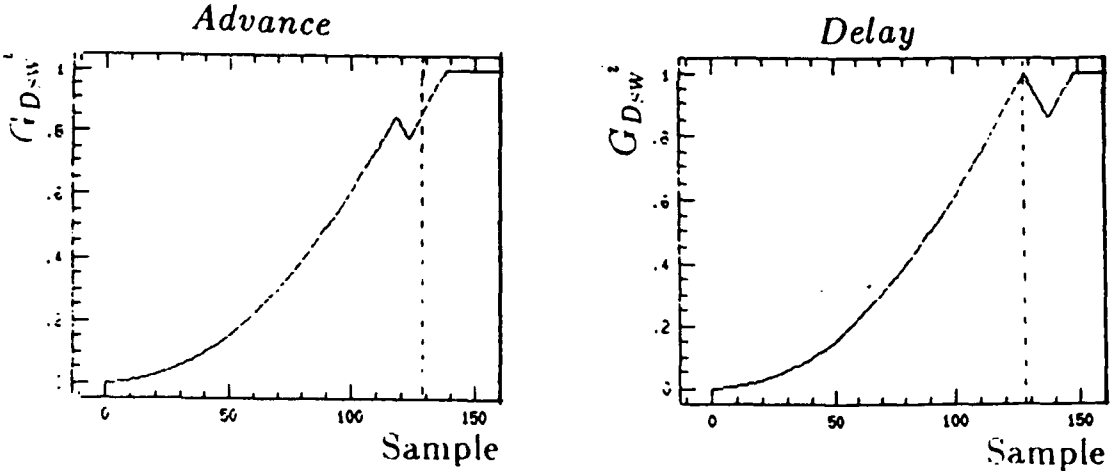


Figure 40: *Jitter*-troubled $\gamma_D{}'$

The performance will be degraded, although, once one uses worst case thresholds, $(Th, Th_{inf})_{\sigma_n, \ldots}$, things will not be severely affected.

Something else to be noticed is that, once $\sigma_{\leq R}$ decreases with γ_H, the inherent degradation is negligible for $\gamma_H > 3dB$.

The *Advance* hypothesis is the worst, once all individual decisions will be taken with smaller $\gamma_D{}'$ than the expected ones, while the *Delay* hypothesis will take decisions with bigger $\gamma_D{}'$, except for the latest ones, the least important [13] in the final decision.

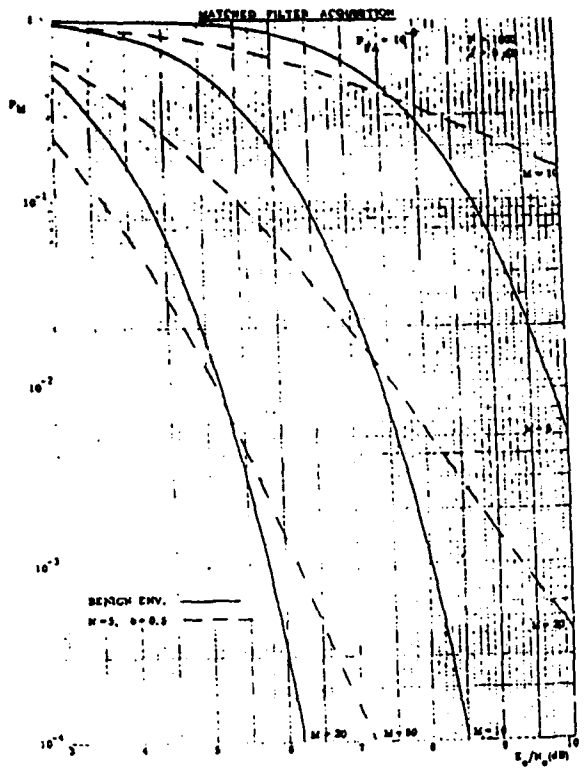


Figure 41: Performance of a Classical Method

Numerical results are expected in the near future to quantify the Inherent-*Jitter* effects. Nevertheless, a feasible system is expected to outperform, at least by one order of magnitude, any known Acquisition system.

Refering to the results of **PUTNAM, RAPPAPORT, SCHILLING**, specially to the curves concerning a **benign environment**[4], one can notice that, for $\gamma_H = 3dB$, with $M = 20$, $P_M^{(1)} \approx 0.5$, while our **Optimal Result** is $P_M^{(1)} < 10^{-7}$ for $M = 3^5$.

2 A JITTER-FIGHTING SCHEME

2.1 DIVERSITY : *THE SOLUTION*

There is a way to overcome the Inherent-*Jitter* effects : **diversity** and **additional signal processing**.

In the next *Figure*, T_δ is the delay associated to $\overline{S_S R}$, and $T_\delta' = T_H - T_\delta$.

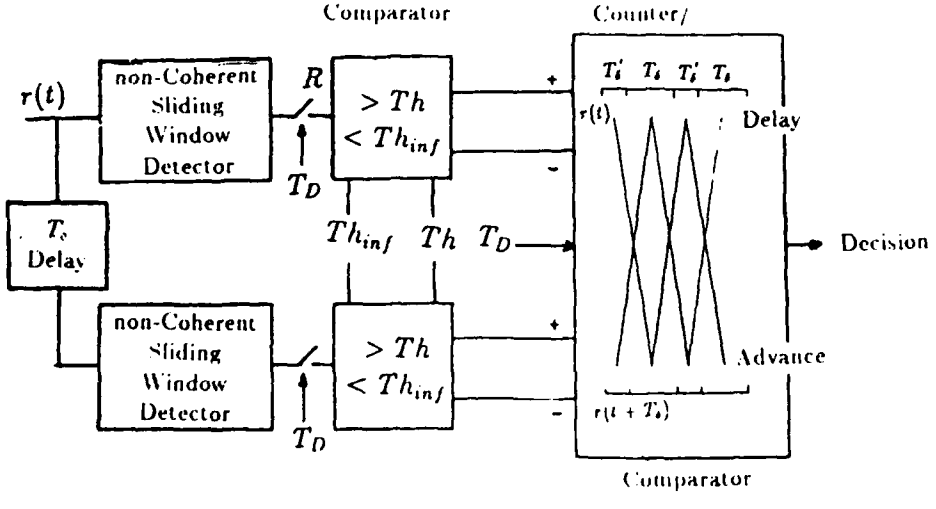


Figure 42: A *Jitter*-Fighting Scheme

The *Jitter* elimination is achieved choosing the appropriate alternance of signals to be analised.

[4]Exactly our case : no Jammers, no Interference, AWG Noise only.
[5]We assumed a Bank of C Active Detectors in the Verification Block.

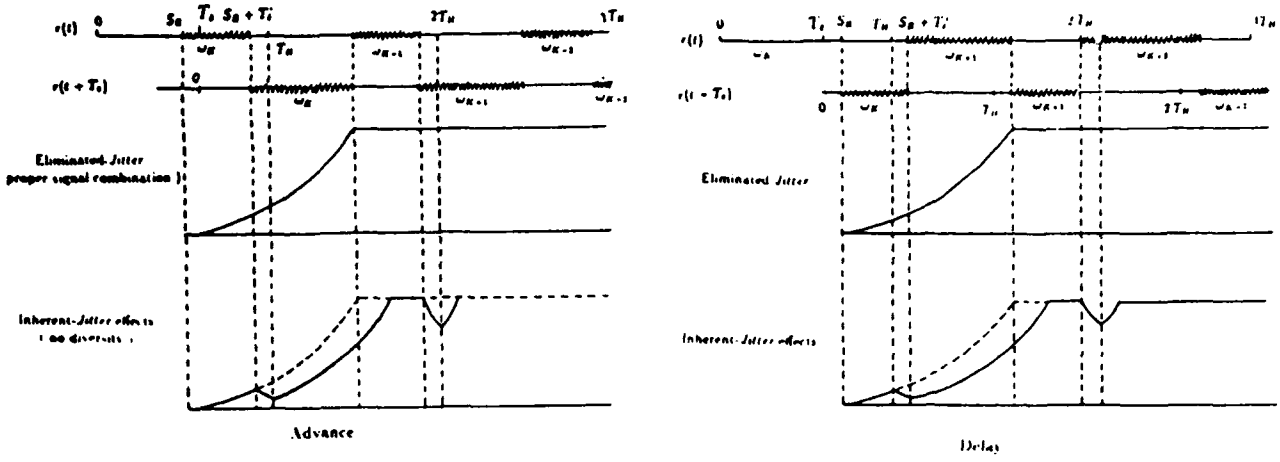


Figure 43: *Jitter*-Fighting Effects

The **wrong choice** will lead to poorer results (with **Signal Present**), therefore to higher $P_{V_{Test}|H_1}$.

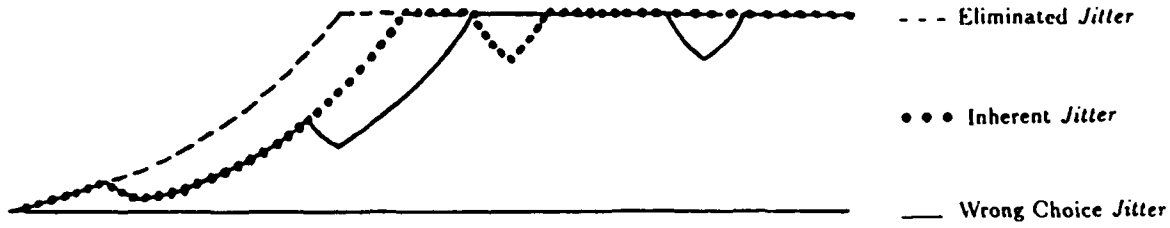


Figure 44: Wrong Choice Degradation

The Decision Device (Counter/Comparator) will have to decide whenever conflictual "opinions" occur. Nevertheless, if one of the decisions is *Positive*, some thought must be given to the low γ_H, emphasizing **Signal Presence** probability.

As soon as a conclusion (one of the absorving boundaries), Positive or Negative, is reached, even before the time limit, if the alternative Test is far from "decided", it is accepted (the absolute difference between the two Counters must exceed a certain value). Otherwise, the time limit will settle the dispute, even if the decisions become contraditory. In that case, evaluation will be done to both Tests, counting all Positive (+1) and Negative (−1) individual decisions. Any difference will decide.

2.2 PERFORMANCE ENHANCEMENT

The performance enhancement made possible through this *Jitter*-Fighting scheme is obvious from *Fig. 44*.

The expected results will approximate reasonably the Optimal Results here presented, leading to a quite interesting, although *software* intensive, system.

Once more, the first numerical results for this scheme are expected soon.

3 Conclusions

This new Acquisition Method for Push-to-Talk FH-SS system achieves a very high $P_D^{'s}$ with minimum hardware complexity (low M) and an affordable number of computations per Decision Time T_D, in low SNRs (3 dB and even under - compare with [4] and [10]).

Unlike in the system discribed in [2], the First-Dwell stage is a Matched-Filter like the one used in [9] - Part II], therefore permitting Rapid Acquisition through the Passive Correlation implemmented. The Second-Dwell stage is a double test on the output of an Active non-Coherent Detector, once more making a quick decision : each Sync alarm is verified in less than one Hop Time ($N_1 + N_2 = N$), outperforming, again, the above refered systems.

Another advantage of this new method is the easier job for the Tracking Loop, once Acquisition is achieved within a small fraction of T_H.

The presented Optimal Results refer to an ideal system - Omniscient System - which *knows* the exact timing of the Recognition Sample, in spite of the variance of the Pre-Synchronization indication, therefore adjusting itself in order to perform the Validation tests with optimum thresholds, conveniently changing the Tone.

A feasible version of this system will present a worst performance, although we consider it will be, nevertheless, better, by a few orders of magnitude, than that of classical methods.

A *Jitter*-Fighting scheme, involving additional computational power necessity, is presented to enable a feasible system to achieve a performance as near the optimum as possible.

References

[1] J. P. COSTAS, 'Poisson, Shannon and the Radio Amateur', *IRE Proc.*, Dec 59

[2] D. M. DiCARLO, C. L. WEBER, 'Multiple-Dwell Serial Search : Performance and Application to Direct-Sequence Code Acquisition', *IEEE Trans. on Comm.*, May 83

[3] R. C. DIXON, *Spread-Spectrum Systems*, John Wiley and Sons, 76

[4] P. R. PAWLOWSKY, A. POLYDOROS, 'Worst-case Jamming for Frequency-Hopping Code Acquisition', MILCOM 83, Washington D.C.

[5] P. R. PAWLOWSKY, A. POLYDOROS, 'Optimization of a Matched-Filter Receiver for Frequency-Hopping', MILCOM 85

[6] J. N. PEREIRA, *A New Rapid Acquisition Method for Frequency-Hopped Spread Spectrum systems using Synchronization Preambles*, M. Sc. Thesis, IST, Lisbon Technical University, Nov 86

[7] J. N. PEREIRA, F. B. CERCAS, A. A. ALBUQUERQUE, 'Synchronization of a Frequency-Hopping Spread Spectrum System', 2nd Symposium on Telecommunication Electronics, Lisbon, May 86

[8] J. N. PEREIRA, A. A. ALBUQUERQUE, 'A new Acquisition Method for Frequency-Hopping Push-to-Talk Spread Spectrum Systems', 1st *International Conference on Communication Technology*, Nanjing, CHINA, Nov 87,

[9] A. POLYDOROS, C. L. WEBER, 'A unified approach to Serial Search Spread Spectrum Code Acquisition', Part I and II, *IEEE Trans. on Comm.*, May 84, pp 542-560

[10] C. A. PUTNAM, S. S. RAPPAPORT, D. L. SCHILLING, ' A comparison of schemes for Coarse Acquisition of Frequency-Hopped Spread Spectrum Signals', *NTC' 81 Conference Record*, paper 34.2, New Orleans, LA, Dec 81, (*IEEE Trans. on Comm.*, pp 183-189, Feb 83)

[11] M. K. SIMON, J. K. OMURA, R. A. SCHOLTZ, B. K. LEVITT, *Spread Spectrum Communications* - vol III, Computer Science Press, 85

[12] F. L. RICKER, P. L. McADAM, B. D. TRUMPIS, 'Sequential Acquisition of PN Codes for TDRSS', TRW Defense and Space System, *internal report*, 77

[13] R. E. ZIEMER, R. L. PETERSON, *Digital Communications and Spread-Spectrum Systems*, Macmillan Publishing Company, 85

DEVELOPMENT OF A NEW HIGH ALTITUDE ELECTROMAGNETIC PULSE (HEMP) ENVIRONMENT AND RESULTING OVERHEAD LINE RESPONSES

F.M. Tesche
LuTech, Inc.
P.O. Box 796012
Dallas, TX 75252

P.R. Barnes
Energy Division
Oak Ridge National Laboratory
P.O. Box X
Oak Ridge, TN 37831-6366

SUMMARY

Being able to accurately determine the response of an electrical system which is excited by a high altitude nuclear electromagnetic pulse (EMP) requires a knowledge of the time history of the incident EMP field strength, as well as its angle of incidence and polarization. A commonly used, unclassified, description of this environment is provided by the "Bell Laboratory waveform". Recent studies have shown, however, that this EMP waveform tends to over estimate the response of an above-ground transmission line by more than an order of magnitude. As a result, other unclassified high altitude EMP environments have been developed.

This paper discusses the development of two alternate unclassified EMP environment descriptions: one arising from a simple radiating dipole moment model, and the other resulting from curve-fitting the calculated fields from a computer code named CHAP. For both of these EMP models, the electric field at two earth observation points are compared. These fields are then coupled to an above-ground line and the resulting open-circuit voltage responses are compared. Using the CHAP EMP environment, a limited parametric study of the peak positive and negative open-circuit line voltage is then performed and surface plots of these peak voltages are presented.

I. INTRODUCTION

In order to evaluate the effects of a radiated electromagnetic pulse (EMP) from a high altitude nuclear burst on a ground-based system, it is first necessary to have an estimate of the incident EMP field. Frequently, a "worst case" EMP field is specified for use in a particular situation, and this environment may not be developed from a consideration of the physics of the EMP production. Such is the case with the unclassified Bell Laboratory HEMP waveform [1], which is viewed as a "bounding" waveform, and is commonly used for predicting system response to EMP.

The concept of performing a worst case analysis is often acceptable for designing a small, compact system which is to be hardened against EMP. In such systems, which are typically military in nature and in which a failure cannot be tolerated, a worst case analysis leads to an inherent hardness margin in the design. For performing an assessment of the effects of EMP on a system, however, the use of a worst case environment is usually not appropriate, since it is desired to develop an accurate, quantitative measure of the system response, not just to make a statement that the system is hard.

In other instances, the use of a single EMP environment may not be appropriate, as in the case of a large, distributed power or communication system. In these cases, the incident EMP field may vary in polarization, angle of incidence and wave shape over the extended system. For accurate assessments of these types of systems, it is often required to have a more realistic estimate of the EMP excitation of the system than that provided by the Bell Laboratory waveform. Such is the case in performing an assessment of the effects of EMP on a commercial power system as described in [2].

For assessing an electrical power system, one EMP coupling problem which is useful to solve is the EMP interaction with a semi-infinite, above-ground line. For this case, a response of interest is the open-circuit voltage at the end of the line and this may be used to infer the overall behavior of the power system under HEMP excitation [2]. For lines having dimensions typical of power transmission lines, preliminary calculations have been performed using the Bell Laboratory waveform as an excitation, and peak open-circuit voltages on the order of 15 to 17 MV were computed. This suggests that EMP might pose a problem to equipment attached to the line.

These large line responses are not reasonable estimates, however, due to the fact that the Bell Laboratory EMP waveform has an unrealistically long tail and this tends to provide a line response which is too large. This observation ultimately lead to the development of an alternate, unclassified EMP environment in [2] for use in the DOE Power System EMP Assessment Program. This resulted in an incident field whose polarization, rise and fall times, and peak amplitude varied as a function of position under a high altitude burst, with a spatial variation being described by a simple radiating magnetic dipole moment. Many of the typical line responses presented in [2] used this environment.

The difficulty with this alternate EMP environment is that there is some uncertainty as to the parametric values which enter into the model. In a recent report [3], Longmire took a different approach for obtaining a suitable unclassified description for the high altitude EMP fields. This involved the use of the CHAP code to compute the fields from a nominal, large yield burst at a height of 400 km, and resulted in plots of the electric field components at several different locations on the earth. These results were then fit to analytic expressions and these may be used to predict the EMP environment at an arbitrary point on the earth's surface.

With these different EMP environments, a question arises as to how these EMP environments compare with each other. In addition, it is important to understand how the corresponding line responses compare. Furthermore, it is useful to understand how the positive and negative peak values of the HEMP-induced voltage responses vary as a function of position and line orientation of the earth. These issues are explored in this paper.

After this introduction, Section II provides a brief review of the EMP environments provided by the dipole moment model and by the CHAP model. In Section III, a simple transmission line coupling model is introduced for estimating the open-circuit response of an above-ground line subjected to these HEMP environments, and several different line responses are illustrated. In Section IV, the CHAP environment is used to perform a parametric study of variations in the peak positive and negative line voltages with line location, line height and earth conductivity as parameters. As a useful result of this study, curves showing the probability of occurrence of different voltage levels on above-ground lines is presented. Finally, in Section V, a brief summary and conclusions are presented.

II. DEFINITION OF THE HEMP ENVIRONMENTS

The geometry of the problem under consideration here is illustrated in Figure 1. A high altitude nuclear burst is detonated at an altitude h_b over the earth's surface. This results in a downward-propagating EMP which interacts with a transmission line located on the earth's surface at an observation point as shown in Figure 2.

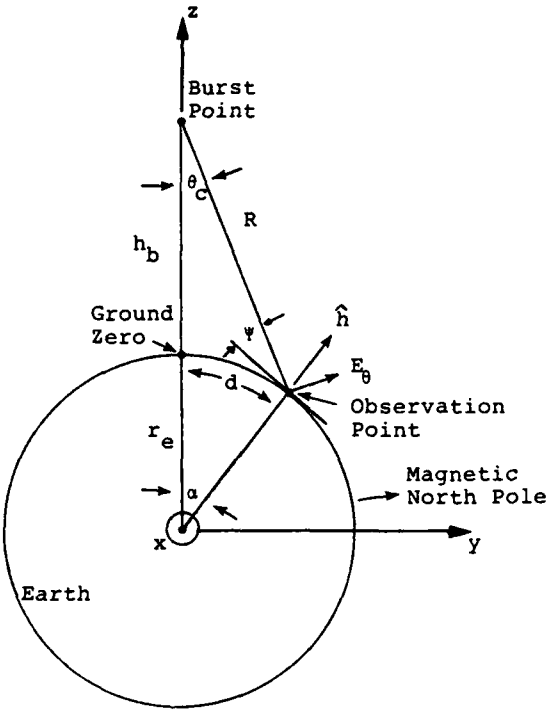


Figure 1. Geometry of the Problem.

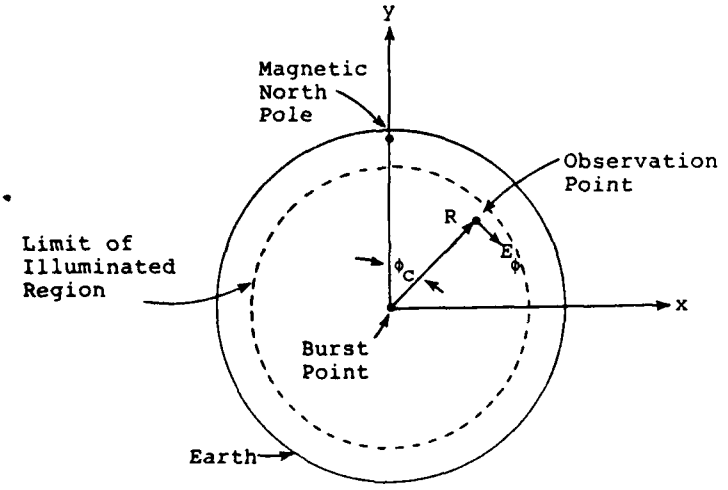


Figure 2. Earth Geometry as Seen From the Burst Point.

The location of the observation point is defined by the angle θ_c which is the angle between the earth's normal at the burst location and the position vector from the burst location to the observation point. This angle is shown in Figure 1, and was referred to in [3] as the "CHAP angle". The other angle which defines the observation location is the azimuthal angle ϕ_c which is measured as being positive in the clockwise direction from the magnetic north pole, as illustrated in Figure 2, which is a view of the earth as seen looking down from the burst point. Given a fixed burst point, only a finite region of the earth's surface is directly visible region. In this illuminated region, a fixed angle θ_c defines a locus of possible observation points which is a circle intersecting the earth's surface, as shown in Figure 2.

At the observation point, it is possible to define a local elevation angle of incidence ψ with respect to the earth tangent line which is given in terms of θ_c as

$$\psi = \frac{\pi}{2} - \arcsin\left[\frac{r_e + h_b}{r_e}\sin\theta_c\right] \tag{1}$$

where r_e is the radius of the earth. The surface arc distance d from ground zero under the burst to the observation point is given as

$$d = r_e \alpha \tag{2}$$

where the interior angle α is

$$\alpha = \frac{\pi}{2} - \theta_c - \psi \quad . \tag{3}$$

The distance from the burst point to the observer on the earth is denoted by R and is

$$R = R_t \cos\theta_c \left[1 \pm \sqrt{1 - \frac{R_t^2 - r_e^2}{R_t^2 \cos^2\theta_c}}\right] \tag{4}$$

where $R_t = r_e + h_b$.

As is described in [1], the incident EMP can be divided into vertically and horizontally polarized components, as shown in Figures 1 and 2. As illustrated in Figure 1, the vertically polarized component of the electric field is denoted as E_θ and lies in the plane of incidence (i.e., the plane formed by the $\bar{R}$ vector and its projection in the ground plane. Figure 2 shows the direction of the horizontal component of the incident electric field, E_ϕ.

The reason for making the distinction between the field components in this manner is that it is possible to use the Fresnel reflection coefficients for these field components to easily compute the effects of the lossy earth on the total field above the ground. The total response of the line to the EMP field can then be considered as the superposition of the responses of these two field components.

In reference [3], the incident EMP fields at 5 different angles of ϕ_c were plotted for several different angles of θ_c. This corresponds to observation locations #1 through #5 as depicted in Figure 3. At each of these locations, it is possible to specify the local orientation of the transmission line being excited by the incident EMP. This is equivalent to specifying the angle ϕ' in Figure 4. For $\phi' = 0^\circ$ it is noted that the line responds only to the vertically polarized component of the field. For other angles of ϕ', the response is a combination of the two polarizations. In this study, we will first consider the two angles ($\phi' = 0^\circ$ and 90°) separately to illustrate the relative levels of responses from the two fields.

A. Overview of the Magnetic Dipole Model for HEMP Fields

The model used in reference [2] for determining the spatial and temporal behavior of the EMP from a high altitude burst is essentially that of a radiating magnetic dipole moment located at the burst point. This is done in an attempt to account for the effects of the Compton electrons' motion in the earth's magnetic field and the resulting electromagnetic radiation.

Using the geometry shown in Figure 1, a time-dependent magnetic dipole moment of strength $\bar{m}(t)$ is assumed to be located at the burst point and is oriented in the direction of the local geomagnetic field. For this dipole, the transient electric field at a point $\bar{r}$ on the surface of the earth is given by the vector relation

$$\bar{E}(\bar{r},t) = \frac{\mu_o}{c}\frac{1}{4\pi r}\left[\hat{r} \times \frac{\partial^2}{\partial^2 t}\bar{m}(t)\right] \tag{5}$$

where $\hat{r}$ is the unit vector from the burst point to the observation point. Note that this expression neglects near-field (or equivalently, late time) corrections to the field, neglects the extended nature of the EMP source region, and assumes that the equivalent source dipole moment is located at the burst point and not at the source deposition region which is located at a height of approximately 20 km.

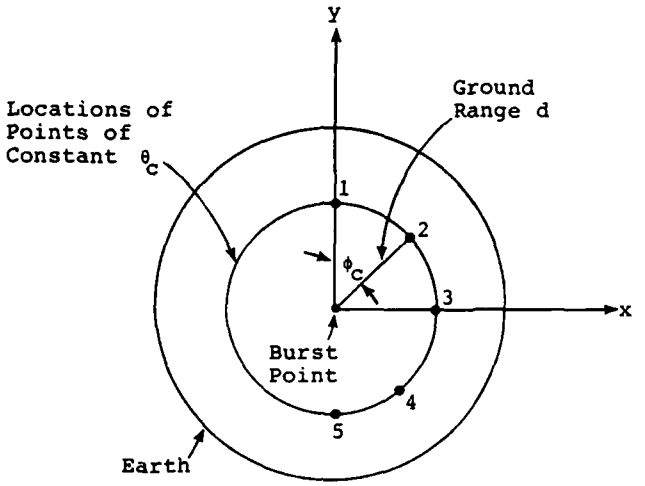


Figure 3. Location of Field Observation Points of Reference [2] as Seen From the Burst Point.

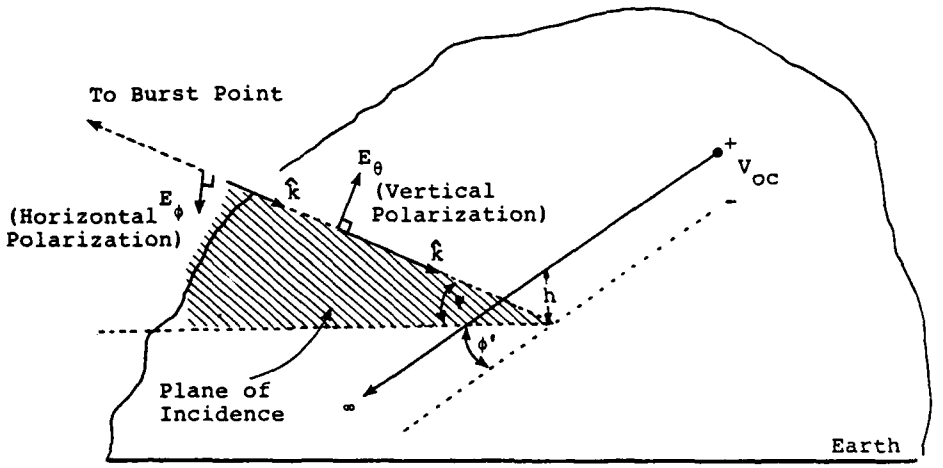


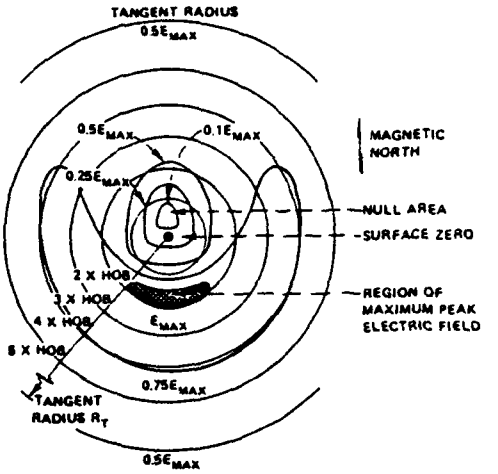
Figure 4. Local Line Geometry Showing Incident EMP Field and Polarizations.

The vector cross product in equation (5) gives the spatial variation and polarization aspects of the incident EMP. Figure 5a plots the variations of the HEMP peak amplitude over the earth's surface as given in reference [1], and is often referred to as the "smile" diagram. Figure 5b shows a sample surface plot of the magnitude of the incident EMP field computed using equation (5).

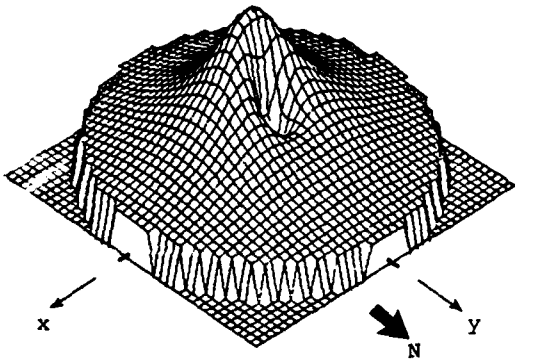
The electric field in equation (5) can be expressed as

$$\bar{E}^{inc}(t) = E_o\, S(\theta_c,\phi_c)\, [\, a_v\hat{\theta} + a_h\hat{\phi}\,]\, f(t) \qquad (6)$$

where θ_c and ϕ_c define the position of the observation point on the earth's surface, $\hat{\theta}$ and $\hat{\phi}$ are unit vectors defined in Figures 1 and 2, and S is the magnitude function of the field amplitude variation which has been normalized to have a peak value of unity. The parameters a_v and a_h are the fractions of vertically and horizontally polarized vector field components and which are related by $a_v^2 + a_h^2 = 1$.



a. Variations in high altitude EMP peak electric field on ground surface from [1].



b. Surface contour plot evaluated from $(\hat{R}x\bar{M})$

Figure 5. Example of HEMP Intensity Variations Over the Illuminated Region of the Earth's Surface.

The specification of the time history of the incident field in equation (6) is still unknown, since the time history of $m(t)$ in equation (5) is not specified. For the dipole moment model, it is assumed that the incident electric field has temporal behavior which is represented by a double exponential waveform. Thus, the term $f(t)$ is assumed to be given by

$$f(t) = \Gamma \left(e^{-\alpha t} - e^{-\beta t} \right) \qquad (7)$$

where α and β are constants and Γ is chosen so that the peak value of $f(t)$ is unity.

Specific unclassified values for the parameters α and β are given in [1] for the bounding Bell waveform as

$$\alpha = 4.0 \times 10^{6} \text{ (1/sec)}$$

and

$$\beta = 4.76 \times 10^{8} \text{ (1/sec)}$$

For the present study, we desire a specification of α and β as a function of position below the burst. Statements in [1] suggest that directly under the burst, the incident field has a rise time of about 5 ns, and a fall time of about 20 ns. At the horizon, however, the field has a rise of 10 ns and a fall of 200 ns. From these statements it is possible to infer two sets of waveform parameters, namely:

At ground zero:

$\alpha = 3.46 \times 10^7$ (1/sec)

$\beta = 4.40 \times 10^8$ (1/sec)

At the horizon:

$\alpha = 3.46 \times 10^6$ (1/sec)

$\beta = 2.00 \times 10^8$ (1/sec)

For intermediate points, a linear interpolation between these points can be used to provide an estimate of these waveform parameters.

The only remaining parameter to be chosen is the value of E_o which is the global maximum electric field. For this model, E_o is chosen so that the maximum field strength on the smile diagram is 50 KV/m. This completes the determination of the dipole moment model which is used in [2] for power system assessments.

B. Overview of the CHAP Code Calculations for the HEMP Environment

In an attempt to produce a HEMP environment which is more representative of what would be expected in an actual high altitude nuclear explosion, Longmire has used the CHAP code [4] to compute the transient fields produced by a large-yield burst at 400 km over the central US. For his study, nominal unclassified weapon output parameters were used, along with unclassified EMP theory and calculation methods. The resulting calculations have been documented in reference [3], and can serve in developing a more accurate understanding as to the effects of HEMP on electrical power systems.

The HEMP environments computed in [3] consist of both the horizontally polarized (E_ϕ) component and the vertically polarized component (E_θ) as a function of time at the five observation positions illustrated in Figure 3 which are defined by magnetic azimuthal angles of $\phi_c = 0^\circ$, 45°, 90°, 135°, and 180°. For these calculations, a local geomagnetic field dip angle of 70° was assumed, and 6 different polar observation angles θ_c were used, corresponding to 6 different ground range distances for the observation points. These ground distances varied from just under the burst point (i.e., ground zero) to the horizon point, which is the limit of the directly illuminated region.

In order to provide this HEMP environment for an arbitrary ground range and magnetic azimuthal angle, ϕ_c, reference [3] developed analytic fits which are continuous functions of these variables, and these curve-fit results were presented in the report along with the CHAP-computed results for comparison purposes.

As may be seen from a comparison of the CHAP and the curve-fit results in [3], there is a reasonably good agreement between the two for early times, but the agreement deteriorates somewhat in late times. This discrepancy, however, appears to be much worse than it really is, due to the fact that the results are plotted on a log scale. As will be shown in the next section, the effects of these two different HEMP environments when coupled to an above-ground power line, provide virtually identical responses.

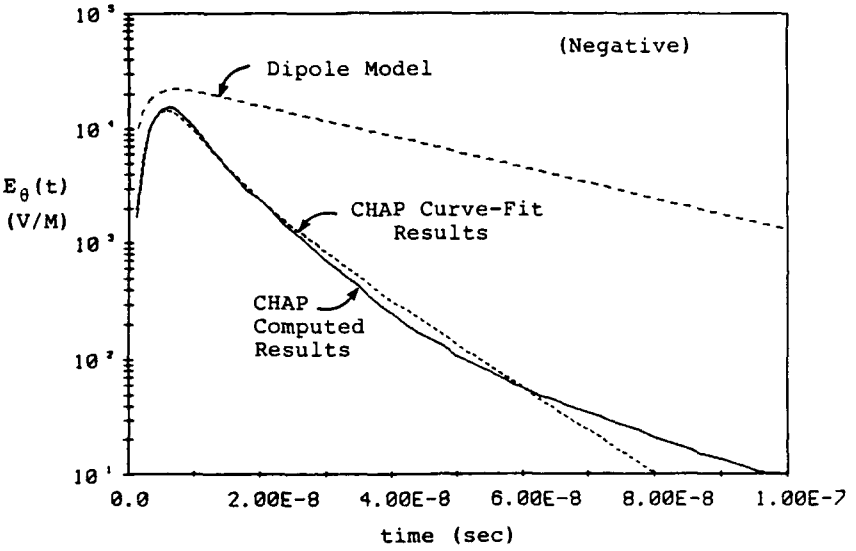
C. Comparison of the Dipole Moment and CHAP HEMP Environments

In comparing the HEMP environments computed from the dipole moment model and the CHAP calculations, observation location #3 with a ground range of 233.5 km was selected. This corresponds to angular positions defined by $\theta_c = 30^\circ$ and $\phi_c = 90^\circ$. The dipole moment model was given the same geomagnetic field dip angle as used in the CHAP calculations and plots of the various HEMP electric fields were made.

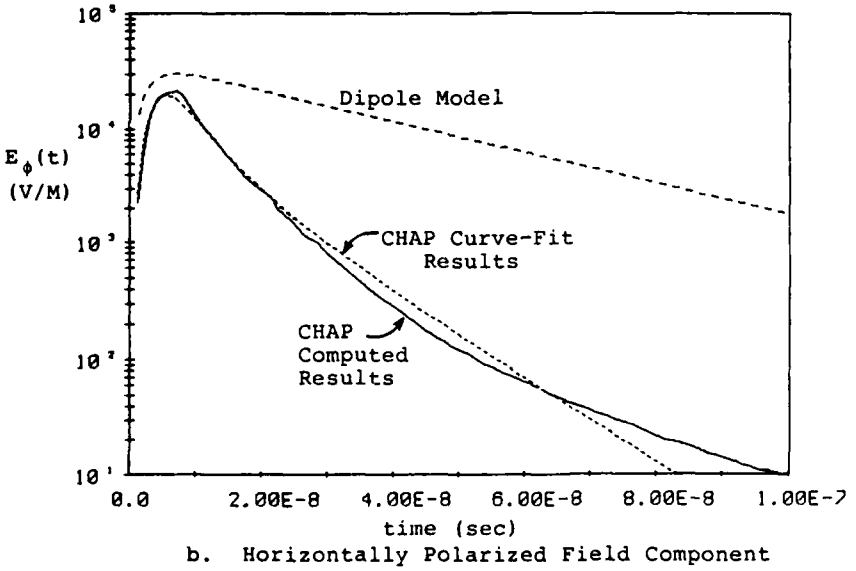
Figures 6a and 6b present overlays of the various fields for the vertical and horizontal components of the HEMP environment. In Figure 6a, the values of E_θ are negative with respect to the $\hat{\theta}$ direction defined in Figure 1. The solid curves represent the CHAP-computed data, and the dotted lines are the curve-fit results, both taken from reference [3]. As previously indicated, the early-time agreement between these two environments is good, but with some apparent deviations occurring at late time. It is important to point out that the solid curves were hand digitized from the plots presented in [3], and this ccounts for the "jittery" nature of the curves.

The dashed lines in the figures represent the calculated results from the dipole moment model using the Bell Laboratory information on the rise and fall times, and with the assumption of a global maximum of 50 kv/m for the radiated field. As may be noted, the peak value is about 1.5 times larger than that of the CHAP results, and the fall time is significantly longer. As will be seen in the next section, the area under the incident HEMP waveform is important in determining the coupled response to power lines, and that the dipole moment environment will provide a larger line response than would that of the CHAP environment.

As another comparison of these different environments, observation point #3 was moved to near the horizon where the ground range was 2201.0 km. This corresponded to a "CHAP Theta angle" of 70.2° and an azimuthal angle of $\phi_c = 90°$.



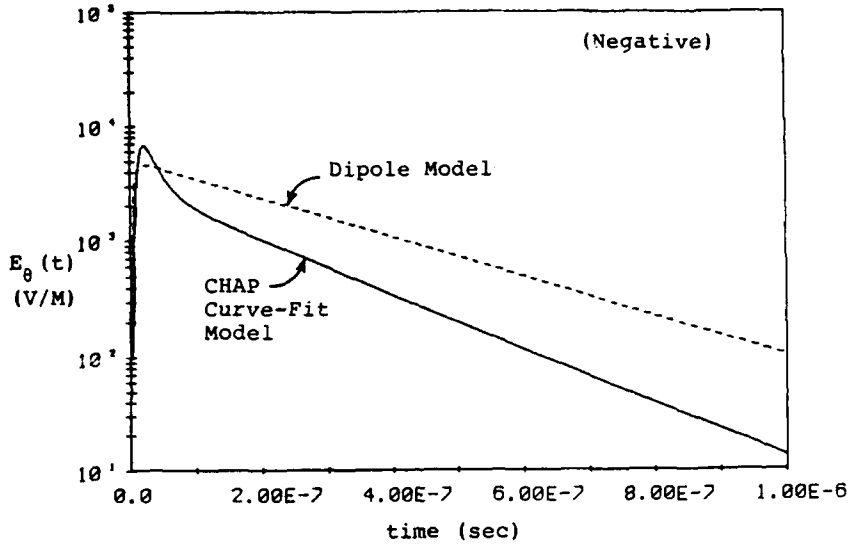
a. Vertically Polarized Field Component



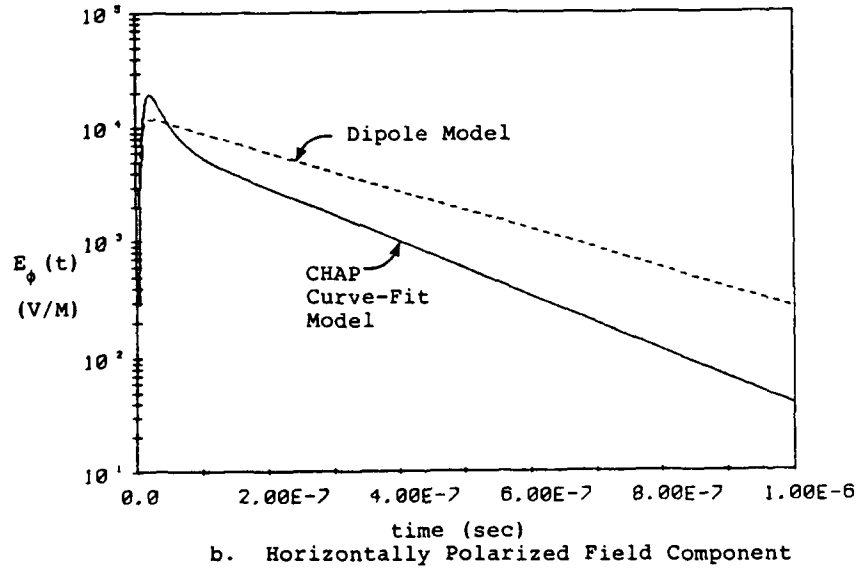
b. Horizontally Polarized Field Component

Figure 6. Plots of the Transient HEMP Fields at Point #3 (d = 233.5 km, $\phi_c = 90°$, $\theta_c = 30°$).

Figures 7a and 7b show the vertical and horizontal components of the computed fields. As in the previous case, the data in Figure 7a are negative. In this instance, the solid lines denote the CHAP curve-fit environments, and the dashed lines are from the dipole moment mode. In this case, it is seen that the initial CHAP results provide a larger peak value of the incident field than found for the dipole moment mode. However, as time progresses, the CHAP results fall off rapidly and eventually become less than the dipole results.



a. Vertically Polarized Field Component



b. Horizontally Polarized Field Component

Figure 7. Plots of the Transient HEMP Fields at Point #3 on the Horizon (d=2201 km, $\phi_c = 90°$, $\theta_c = 70.2°$).

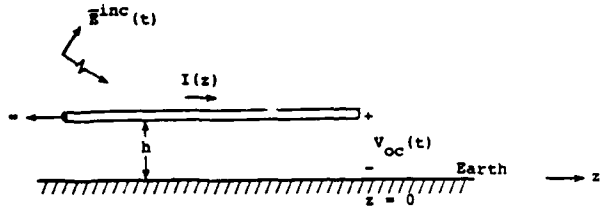
III. COUPLING OF THE HEMP ENVIRONMENTS TO AN ABOVE-GROUND LINE

The determination of the HEMP-induced currents and voltages on above-ground lines has been discussed by a number of investigators [5], [6], [7], and [8]. In this section a simple transmission line model for determining the line response is reviewed, and the responses of a particular line to the HEMP environments in the previous section are illustrated.

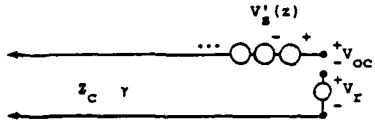
A. Determination of the Line Response

The problem to be discussed here is illustrated in Figure 8, which depicts a conducting line of radius a at a height h above a lossy earth, which has a conductivity of σ_g and a relative dielectric constant ϵ.

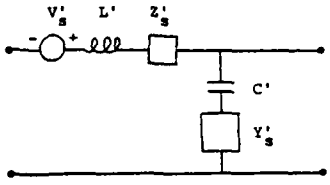
The incident HEMP field induces currents on the line, and as a consequence of the spatial variation of the line current, there are also charges on the line. At sufficiently low frequencies (or, equivalently, at late times) it is possible to define a potential difference between the line and the earth's surface, which is viewed as the return conductor for the line current. For EMP studies, the voltage at the open end of the line is of particular importance, since at this point one might locate a transformer or some other power system component which might be affected by the HEMP surges on the line.



a. Semi-infinite line over lossy earth



b. Transmission line model with distributed sources



c. Per-unit-length, lumped parameter model of line

Figure 8. Semi-Infinite Line and Equivalent Models.

A frequently used model for computing the line response is that of a dispersive transmission line, as indicated in Figure 8b. The propagation constant γ on the line is generally a complex quantity and accounts for the dispersive nature of wave propagation on the line.

The incident plus ground-reflected HEMP fields induce a distribution of excitation voltage sources along the line which are equal to the tangential component of these electric fields on the line. At the end on the line there is a lumped voltage source which arises from the vertical component of these fields acting on the "riser" of the line.

A lumped parameter model for a differential section of the line can be developed as shown in Figure 8c. The per-unit-length inductance and capacitance elements have the values

$$L' = \frac{\mu_o}{2\pi} \text{arccosh } (h/a) \approx \frac{\mu_o}{2\pi} \ln(h/a) \tag{8}$$

and

$$C' = \frac{2\pi\epsilon_o}{\text{arccosh } (h/a)} \approx \frac{2\pi\epsilon_o}{\ln (2h/a)} \tag{9}$$

which are the corresponding values for the line over a perfectly conducting earth.

The effects of the lossy earth on the transmission line model is accounted for through the Z'_s and Y'_s elements. As discussed by Vance [5], these elements can be expressed approximately as

$$Z'_s \approx -\frac{j\gamma_g}{4\pi h\sigma_g} \frac{H_o^{(1)}(j\gamma_g 2h)}{H_1^{(1)}(j\gamma_g 2h)} \tag{10}$$

and

$$Y'_s \approx \frac{\gamma_g^2}{Z_g} \tag{11}$$

where

$$\gamma_g = \sqrt{j\omega\mu_o(\sigma_g + j\omega\epsilon_g)} \tag{12}$$

and $H_0^{(1)}$ and $H_1^{(1)}$ are cylindrical Hankel functions.

If the line itself were not considered to be perfectly conducting, an additional internal impedance of the line could be added in the model. However, investigations have indicated that for realistic wire conductivities, this effect is unimportant, and consequently this is neglected here.

Using the above values for the line parameters, it is possible to define total per-unit-length impedance and admittance values as

$$Z' = j\omega L' + Z_s' \tag{13}$$

and

$$Y' = j\omega C' \, Y_s' \, / \, (Y_s' + j\omega C') \tag{14}$$

With these definitions, the frequency dependent line propagation constant and characteristic impedance becomes

$$\gamma(\omega) = \sqrt{Z'Y'} \tag{15}$$

and

$$Z_c(\omega) = \sqrt{Z'/Y'} \tag{16}$$

The distributed voltage sources along the line are equal to the tangential component of the incident electric field on the line. This field contains contributions from both the vertically and horizontally polarized field components plus reflections of these fields from the earth. As described in [5], the distributed voltage source has the form

$$V_s'(z) = E^{inc}(\omega)\, e^{-\gamma_o z \cos\psi \cos\phi'} \left[\cos\psi \left(1 + R_v e^{-\gamma_o 2h\sin\psi} \right) + \sin\psi \cos\phi' \left(1 - R_h e^{-\gamma_o 2h\sin\psi} \right) \right] \tag{17}$$

where $\gamma_o = \omega/c$ and is the free space propagation constant, and the angles ψ and ϕ' are defined in Figure 4. In this equation the zero phase location is taken to be at $z = 0$, implying that the incident field arrives at this location at $t = 0$.

The terms R_v and R_h are the Fresnel reflection coefficients for the vertically and horizontally polarized field components respectively. These terms have the following representations:[2]

$$R_v = \frac{\epsilon_r \left(1 + \frac{\sigma_g}{j\omega\epsilon_g} \right) \sin\psi - \left[\epsilon_r \left(1 + \frac{\sigma_g}{j\omega\epsilon_g} \right) - \cos^2\psi \right]^{1/2}}{\epsilon_r \left(1 + \frac{\sigma_g}{j\omega\epsilon_g} \right) \sin\psi + \left[\epsilon_r \left(1 + \frac{\sigma_g}{j\omega\epsilon_g} \right) - \cos^2\psi \right]^{1/2}} \tag{18}$$

and

$$R_h = \frac{\sin\psi - \left[\epsilon_r \left(1 + \frac{\sigma_g}{j\omega\epsilon_g} \right) - \cos^2\psi \right]^{1/2}}{\sin\psi + \left[\epsilon_r \left(1 + \frac{\sigma_g}{j\omega\epsilon_g} \right) - \cos^2\psi \right]^{1/2}} \tag{19}$$

The induced field source at the riser at the end of the line as shown in Figure 8b is related only to the vertically polarized component of the field, since the horizontally polarized field never is tangential to the riser. Assuming that the height of the line is electrically small, which is an approximation consistent with the use of transmission line theory for this problem, this voltage source takes the form

$$V_r \approx E^{inc}(\omega)\, h \sin\psi \left(1 + R_h \right) e^{-\gamma_o z_o \cos\psi \cos\phi'} \tag{20}$$

for a riser located at an arbitrary point z_o. For the riser at the end of the line, z_o is to be set to zero.

With these excitation sources and line parameters defined, it is possible to use transmission line theory to determine the open circuit voltage response at the end of the line. The solution takes the form

$$V_{oc}(\omega) = \int_{-\infty}^{0} V_s'(z)\, e^{\gamma z}\, dz + V_r \tag{21}$$

Using the expression for the distributed line voltage sources in equation (17) and the riser voltage in equation (20), the above expression can be evaluated as

$$V_{oc}(\omega) = E^{inc}(\omega) \left\{ \cos\psi \left(1 + R_v e^{-\gamma_o 2h\sin\psi} \right) + \sin\psi \cos\phi' \left(1 - R_h e^{-\gamma_o 2h\sin\psi} \right) \right\} \left(\gamma - \gamma_o \cos\psi \cos\phi'\right)^{-1}$$

$$+ E^{inc}(\omega)\, h \sin\psi \left(1 + R_h \right) \qquad (22)$$

In this last expression, the bracketed term accounts for the excitation of the line by the distributed sources along the line, and the last term accounts for the vertical riser excitation.

Equation (22) is a time harmonic expression and may be transformed into the time domain using a numerical Fourier transform. The resulting time domain voltage, along with the corresponding input impedance of the line, forms a Thevenin equivalent circuit which is useful in determining the behavior of the line when connected to various equipment loads.

It is important to note that the expression for line response due to the distributed field excitation terms in equation (22) involves a term in the denominator which varies approximately as $j\omega$. In the absence of the reflected field from the earth, this implies that the portion of the line voltage resulting from this incident field component is proportional to the integral of the incident field. Although the ground reflected field influences the total response, it is apparent that this component of the line response will be more sensitive to the integral of the incident waveform than to the rate of rise. Consequently, for incident HEMP waveforms which have a lingering late-time component, one can expect a larger response than for an incident field which dies out rapidly.

The riser term, however, has a direct dependence on the incident field. Hence, the component of the initial open circuit voltage on the line arising from this term will closely follow the initial time dependence of the HEMP field, and will, therefore, have a rise time on the order of that of the incident field.

B. Examples of Line Responses

As an example of line responses using the CHAP and dipole moment models, the semi-infinite line model discussed above has been used to predict the open circuit line voltage. It is difficult to make an absolute generalization regarding the differences between these two HEMP environments using such a limited sampling, but nevertheless, the present study does give a preliminary indication of the expected trends in the data.

Consider first a line located at point #3 at a range of 233.5 km, and at a height of 10 m above the earth which has a conductivity of 0.01 mhos/m and a relative dielectric constant of 10. Figure 9a presents the open circuit voltage at the end of this line for the case of end-on incidence. Referring to Figure 4, this implies that the angle of incidence is $\phi' = 0^\circ$. In this configuration, only the vertically polarized field component will excite the line.

The first contribution to the line's open circuit voltage response arises from the interaction with the incident field with the vertical riser. Since this field is a negative quantity (i.e., in the $-\hat{\theta}$ direction), this provides a positive pulse at the line end. Later, the effects of the distributed sources begin to arrive at the line end, and the direction of the incident field is such that their effects have a negative polarity and the sign of the induced line voltage is eventually reversed.

The solid curve represents the line response to the computed CHAP environment shown in Figure 6. The dotted curve, which is virtually co-incident with the solid curve, is the response to the curve-fit CHAP data. As previously discussed, the difference between the line responses for these two environments is negligible.

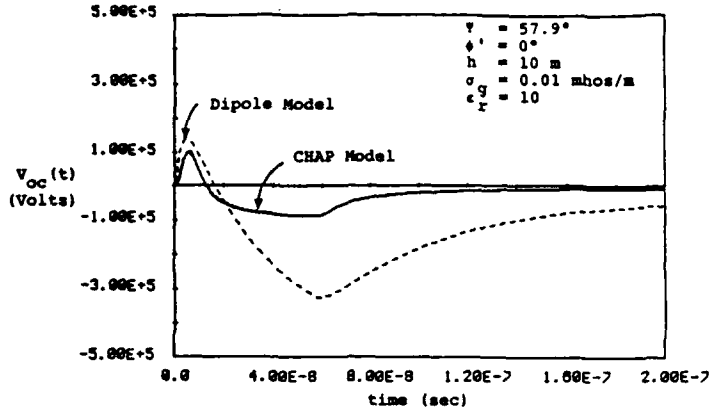
The dashed curve represents the response of the line to the dipole moment environment. Clearly it is larger than the CHAP results, primarily due to the differences in the late time tails of the HEMP environments.

Figure 9b presents similar data for the line being struck in the broadside direction with $\phi' = 90^\circ$. In this case, the horizontally polarized component of the incident field excites the horizontal line, and the vertically polarized field excites the vertical riser. The overall behavior of the line response at this line angle is very similar to that of Figure 9a.

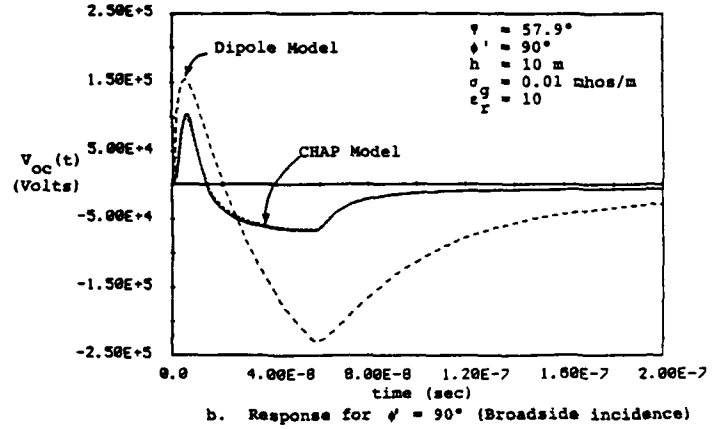
The second line location corresponds to point #3 at the near-horizon, with the HEMP environments as defined in Figure 7. For these coupling calculations, the elevation angle ψ was artificially set to 2°, implying that the lines were actually moved slightly toward the burst point, but with the HEMP environments being those computed at the horizon. This change in the angle of incidence was applied consistently to both of the HEMP field environments and, thus, does not affect the relative comparison between the two results.

The data in Figures 10a and 10b present the line voltage in these cases. Note that the time scales in these figures are significantly longer than those in Figures 9a and 9b. Figure 10a presents the line response for the end-on incidence case ($\phi' = 0^\circ$). Initially, there is a small positive-going spike in the response on the order of that in Figure 9, but it is rapidly overwhelmed by the large negative pulse arising from the horizontal field interaction with the line. In this figure, the time scale is such that the small initial positive spike due to the riser interaction is not evident in the curves. The solid curve represents the results using the CHAP curve-fit environment and the dashed curve is for the dipole moment model. Because the previous results indicated that the computed CHAP results provided a line

response virtually identical to those using the curve-fit data, the computed CHAP results were not used here.

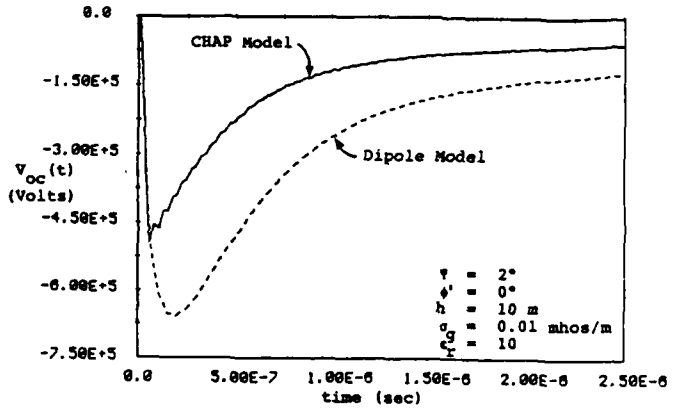


a. Response for $\phi' = 0°$ (End-on incidence)

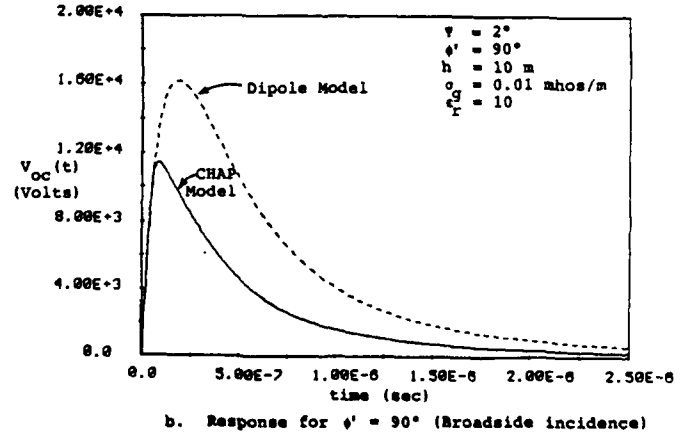


b. Response for $\phi' = 90°$ (Broadside incidence)

Figure 9. Open Circuit Voltage Responses for a Semi-Infinite, Above-Ground Line at Point #3 at Ground Range = 233.5 km.



a. Response for $\phi' = 0°$ (End-on incidence)



b. Response for $\phi' = 90°$ (Broadside incidence)

Figure 10. Open Circuit Voltage Response for a Semi-infinite, Above-ground Line at Point #3 at the Horizon.

Note that although the CHAP curve-fit electric field data has a larger peak than does the dipole model, the coupled line response is larger for the dipole model because of the lingering late-time dipole moment HEMP environment.

The corresponding results for the line at the near-horizon location (point #3) and illuminated in the broadside direction are shown in Figure 10b. Note that this response is significantly lower than that of Figure 10a, and has only a positive peak. It is interesting that the negative line response in Figure 10a is significantly larger than for the other cases. This is due to the so-called "bow-wave" effect in which a large amplitude traveling wave is induced on the line for a grazing incident field [5]. For the particular ground range chosen, the local angle of incidence on the line was $\psi \approx 2^{\circ}$. In this case, the line response is clearly larger than that for the other case, but this angle of incidence is not the one which will maximize the open circuit voltage of the line. Local excursions of this angle would provide a significantly larger line response, as will be evident in the next section..

IV. A PARAMETRIC STUDY OF LINE RESPONSES

As may be observed from the sample results of line responses presented in the previous section, the functional behavior of V_{oc} with variations in the line location on the earth's surface is not simple. It is not possible to examine equation (22) to determine how the response will vary with the angles ψ or ϕ', and the time histories of the various responses are impossible to determine without performing a detailed calculation for each case.

As a result, a limited parametric study of the line responses was performed to illustrate their behavior, and to provide an indication of the maximum and minimum open-circuit voltages which might be expected on the line. For this study, the analytic curve-fit data from the CHAP runs provided in [3] were used to provide the HEMP environment at an arbitrary location on the earth's surface.

Considering first the line located at point #3 at a ground range of 233.5 km, (the same point as for the results of Figure 9), Figure 11 presents the open circuit line voltage for 12 different values of the local line orientation angle ϕ'. The responses for $\phi' = 0^{\circ}$ and 90° correspond to those in Figure 9.

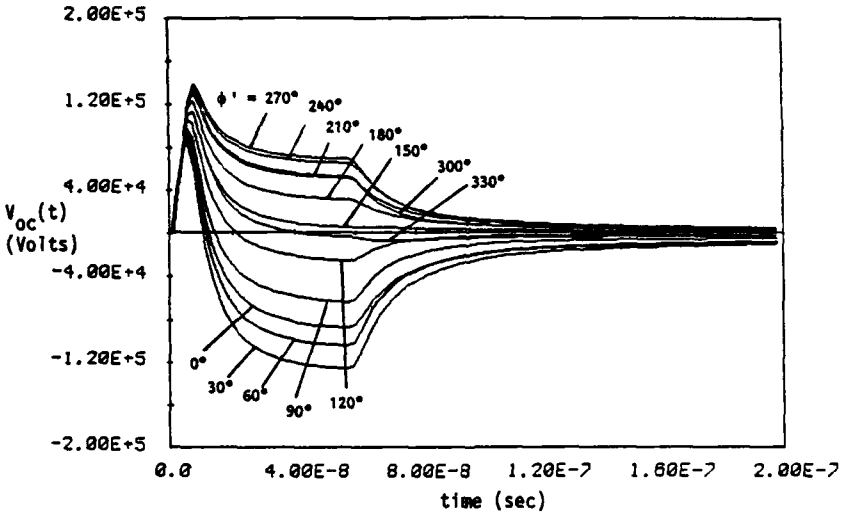
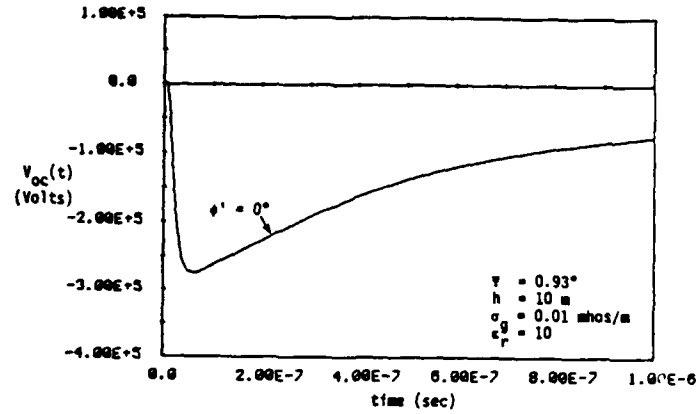


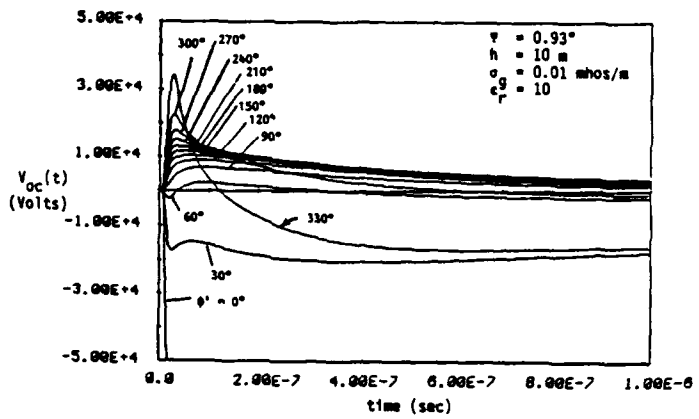
Figure 11. Variations of $V_{oc}(t)$ on a Semi-Infinite Line at Point #3 at Ground Range = 233.5 km with Line Orientation Angle, ϕ'.

As is evident in this figure, the overall open circu't voltage response is by no means simple, and a single double exponential representation of the response is not adequate. The early-time portion of the waveform is positive for all values of ϕ', and as previously indicated, is due to the interaction of the incident field with the vertical riser. Later, the interaction effects of the incident field with the horizontal portion of the conducting line contribute to the line response, and as seen from the figure, these can contribute either a positive or negative effect, depending on the line orientation. Finally, at a time about 57 ns after the incident field strikes the end of the line, the response is modified by a reflected field from the earth.

The voltage response for lines near the horizon, where the incident field arrives with a near-grazing angle, has a much larger range of values than found in the previous case. This is due to the "bow wave" effect previously described. As an example of this, Figure 12 shows the open-circuit line voltages for the line located at point #3 at a ground range of 2100 km, just near the edge of the illuminated region. For this range, the elevation angle is $\psi = 0.93^{\circ}$. Figure 12a presents the transient response for the line oriented with $\phi' = 0^{\circ}$, and Figure 12b shows the same response for all of the 12 values of ϕ'. The response for $\phi' = 0^{\circ}$ is about an order of magnitude larger than that for the other lines. Additional calculations have indicated, however, that this large response occurs only over a limited range of the angle ϕ'.



a. Line response for end-on incidence ($\phi' = 0°$).



b. Line response for different values of ϕ'.

Figure 12. Variations of $V_{oc}(t)$ on a Semi-Infinite Line at Point #3 at Ground Range = 2100 km with Line Orientation Angle ϕ'.

In an attempt to describe these waveforms, it is possible to tabulate or otherwise display the salient features of the responses. One possible set of parameters is the maximum and minimum voltages in the waveform. Another might be the times to these peak values.

In order to illustrate the behavior of the maximum and minimum line voltages, a parametric study was performed in which the illuminated region shown in Figure 3 was inscribed in a rectangular region 4400 km on a side. Within this rectangle, there were 100 individual observation points uniformly distributed in the area. At each point, a set of 36 separate line calculations were performed, corresponding to stepping the line orientation angle ϕ' from 0° to 360° in 10° increments.

For each angle ϕ' the maximum and minimum line voltages were calculated, and these were then used to determine the overall maximum and minimum values for each observation point. The resulting maximum and minimum voltage values have then been plotted as a surface on the calculational grid.

For this study, line heights of 10 and 20 meters have been used, and earth conductivities of ∞, 0.1, 0.01, and 0.001 mhos/meter were considered. In all of the calculations, the conductor radius was 2.5 cm, and the earth relative dielectric constant was 10. Results from this study are presented in Figures 13 through 20, and Table 1 summarizes the cases considered, the respective parameters used, and references the figure in which the voltage plots are presented. Typical computation time for each case was on the order of 22 hours on an IBM-PC.

TABLE 1

Parameters Used for Line Calculations and Corresponding Figure Numbers.

		Earth Conductivity (Mho/M)			
		∞	0.1	0.01	0.001
Line Height (M)	10m	13	14	15	16
	20m	17	18	19	20

As an example of the results of this parametric study, consider the plots in Figure 13. The top plot is that for the maximum positive values of the voltage responses, while the bottom plot is for the maximum negative responses. Here the negative responses have been plotted in the positive direction for convenience. In these plots, ground zero directly under the burst point is in the center of the grid, and the direction to the magnetic north pole in indicated by the arrow.

Because the rectangular calculational region extends outside the circular region of illumination, there are small areas near the corners of the grid where the line voltage responses are identically zero. Generally, in all of the plots, it is seen that the line response directly under the burst is smaller than at points out away from the burst.

It is apparent that the response tends to peak for points near the horizon due to the bow wave effect previously mentioned. This is especially evident in Figure 14 which shows the maximum and minimum voltage values for a 10 meter high line with an earth conductivity of $\sigma = 0.1$ mhos/m. For this case, a large positive peak is observed for lines on the western horizon, and a correspondingly large negative peak is seen for lines on the easterly horizon. Ideally, it would be expected that these plots would be somewhat smoother, but due to the limited samplings in the angle ϕ' and on the earth grid, the response surfaces have some slight slope discontinuities in them.

In looking at the results presented in Figures 13 through 20, it is evident that there are several important trends in the data. These are summarized below:

1. For a higher line, both the maximum and minimum line voltages are larger than those computed for a lower line height.

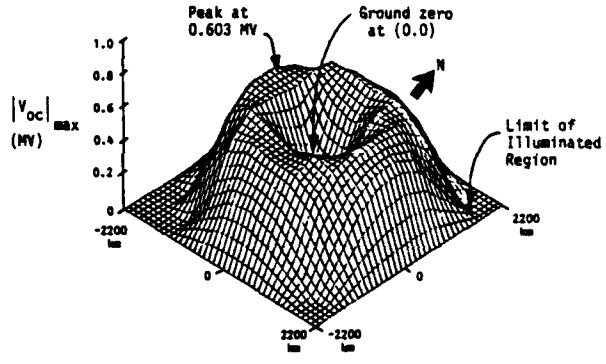
2. For a perfectly conducting earth, the values of the maximum and minimum voltages seem to be relatively insensitive to the magnetic azimuthal angle ϕ_c of the observation point.

3. For a lossy earth, the peak positive open circuit voltage occurs for lines near the western horizon, and the peak negative voltages are found for lines near the eastern horizon.

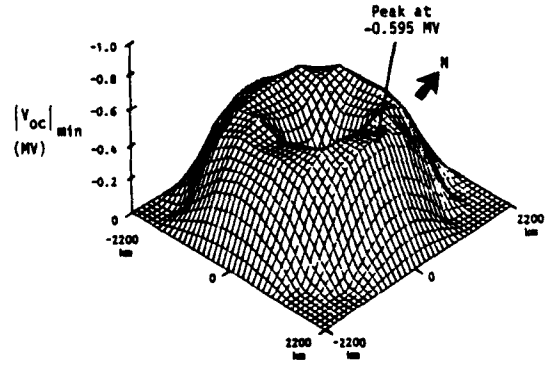
4. For the case of a lossy earth and a specific line geometry, the largest line response appears to occur for the largest earth conductivity.

5. The line response directly under the burst point is much smaller than at other points away from the burst.

6. The surface plots of the positive and negative responses appear to be roughly mirror images of each other when reflected through the North-South direction. That they are not exact images can be attributed to the fact that only a limited number of ϕ' angles have been considered in the calculation.

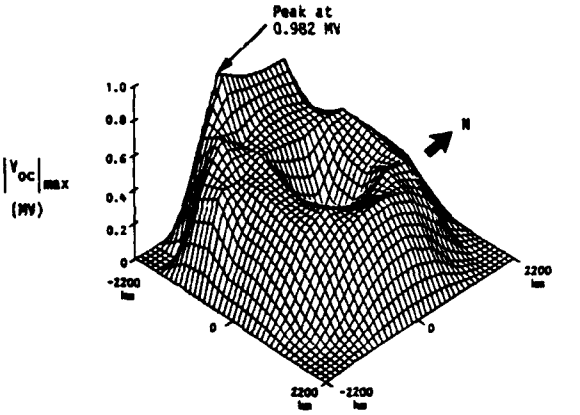


a. Maximum Positive Voltage

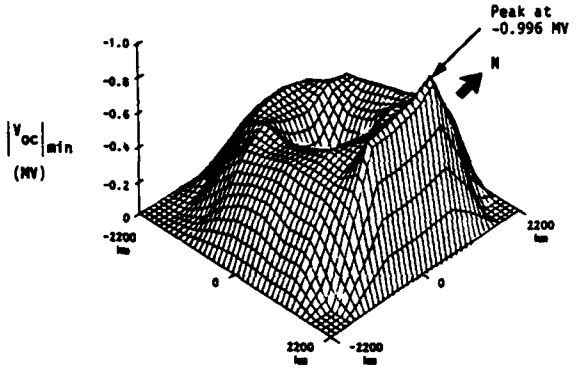


b. Maximum Negative Voltage

Figure 13. Surface Plot of Maximum Positive and Negative Voltages Induced on a Semi-Infinite Line Located at h = 10 Meters Over a Perfectly Conducting Earth.

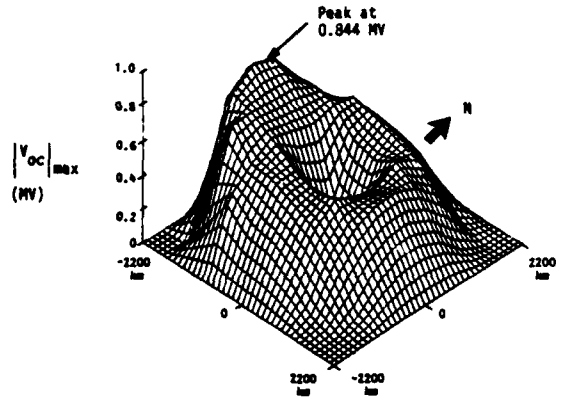


a. Maximum Positive Voltage

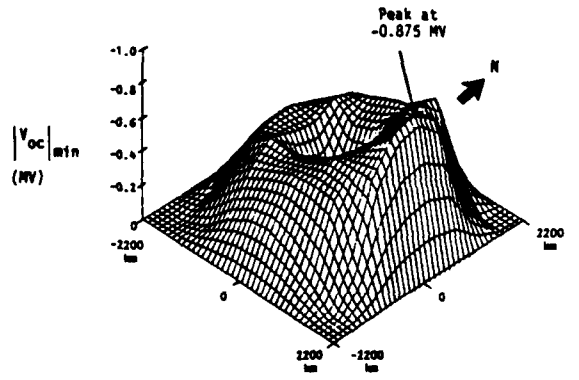


b. Maximum Negative Voltage

Figure 14. Surface Plot of Maximum Positive and Negative Voltages Induced on a Semi-Infinite Line Located at h = 10 Meters Over an Earth with σ = 0.1 mhos/m.

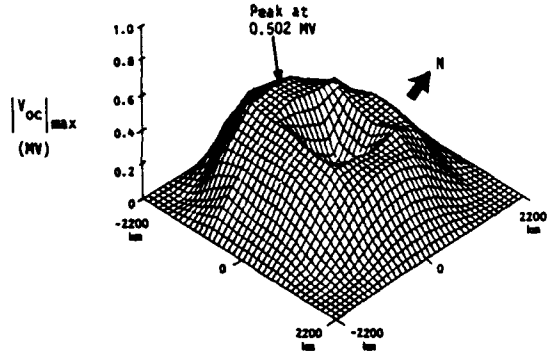


a. Maximum Positive Voltage

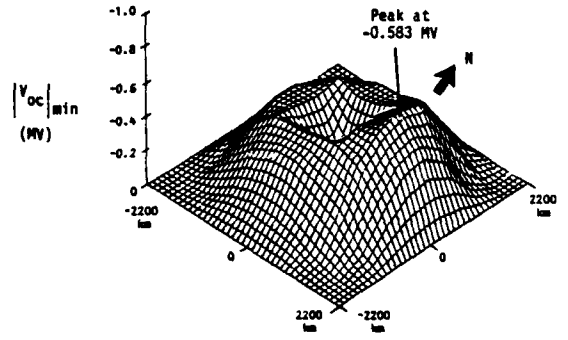


b. Maximum Negative Voltage

Figure 15. Surface Plot of Maximum Positive and Negative Voltages Induced on a Semi-Infinite Line Located at h = 10 Meters Over an Earth with σ = 0.01 mhos/m.

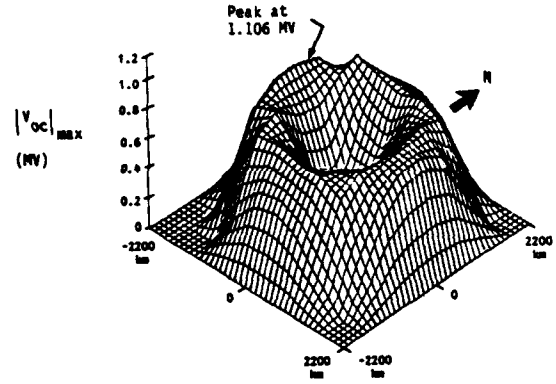


a. Maximum Positive Voltage

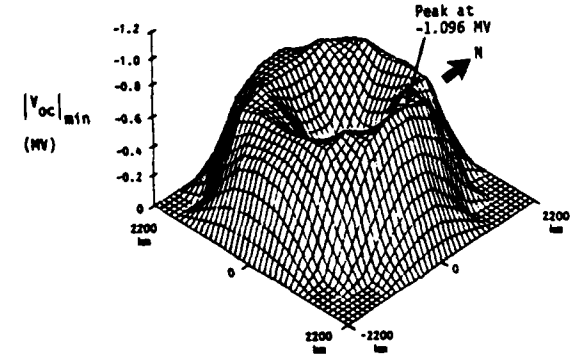


b. Maximum Negative Voltage

Figure 16. Surface Plot of Maximum Positive and Negative Voltages Induced on a Semi-Infinite Line Located at h = 10 Meters Over an Earth with σ = 0.001 mhos/m.

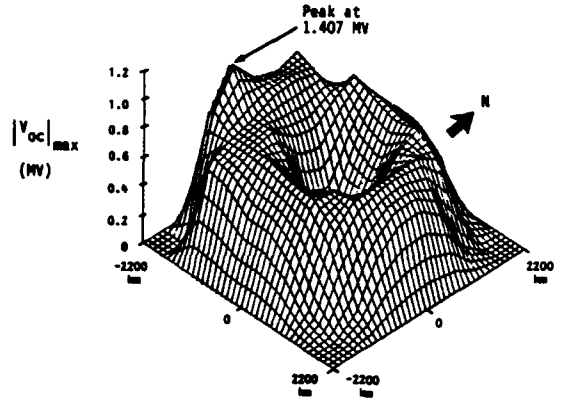


a. Maximum Positive Voltage

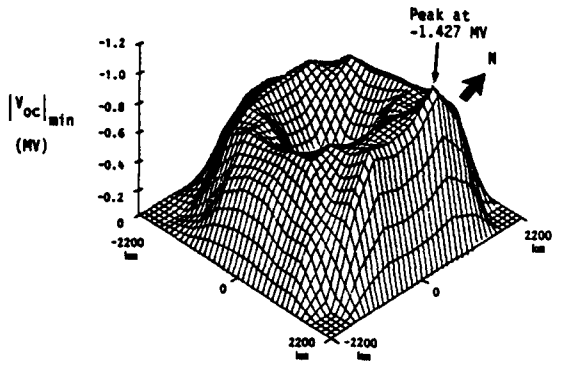


b. Minimum Negative Voltage

Figure 17. Surface Plot of Maximum Positive and Negative Voltages Induced on a Semi-Infinite Line Located at h = 20 Meters Over a Perfectly Conducting Earth.

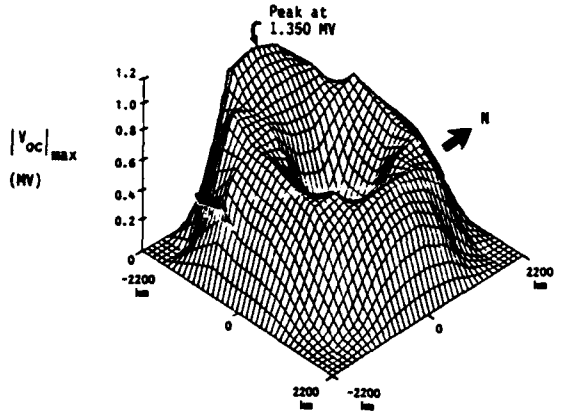


a. Maximum Positive Voltage

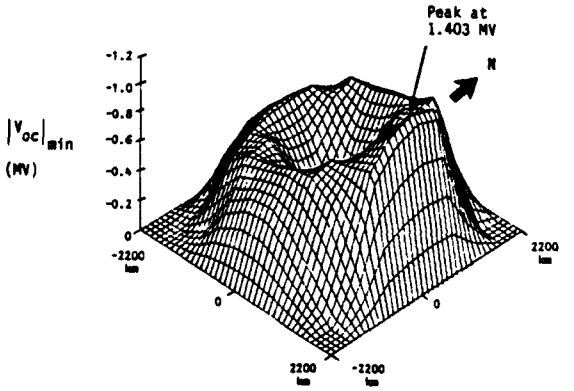


b. Maximum Negative Voltage

Figure 18. Surface Plot of Maximum Positive and Negative Voltages Induced on a Semi-Infinite Line Located at h = 20 Meters Over an Earth with σ = 0.1 mhos/m.

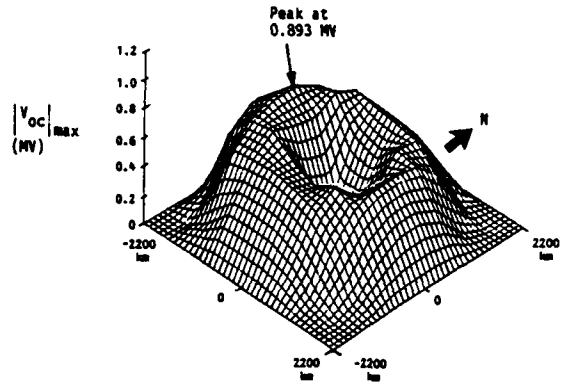


a. Maximum Positive Voltage

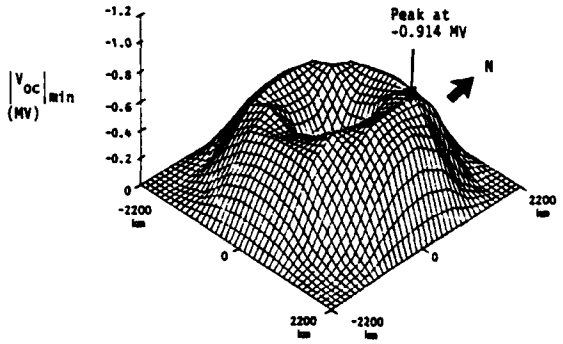


b. Maximum Negative Voltage

Figure 19. Surface Plot of Maximum Positive and Negative Voltages Induced on a Semi-Infinite Line Located at h = 20 Meters Over an Earth with σ = 0.01 mhos/m.

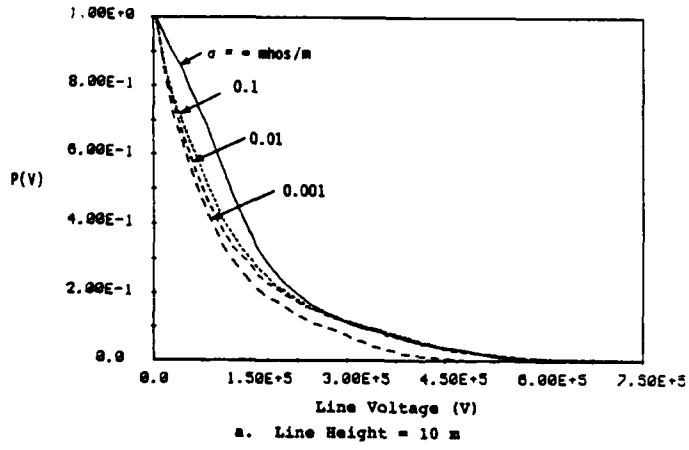


a. Maximum Positive Voltage

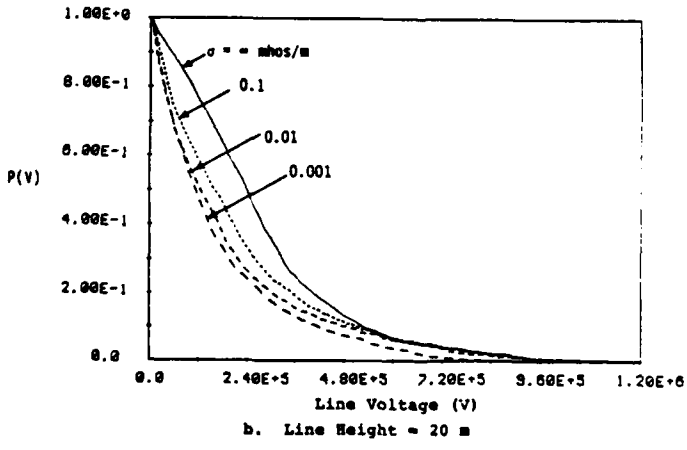


b. Maximum Negative Voltage

Figure 20. Surface Plot of Maximum Positive and Negative Voltages Induced on a Semi-Infinite Line Located at h = 20 Meters Over an Earth with σ = 0.001 mhos/m.



a. Line Height = 10 m



b. Line Height = 20 m

Figure 21. Probability of Peak Voltage $|V_{oc}|$ Exceeding the Abscissa.

In examining the calculated waveforms in Figures 11 and 12, it is apparent that although there might be an occasional large value for the open circuit voltages of the lines, on the average, the line responses will be significantly smaller than the maximum positive and negative values plotted in Figures 13 through 20. Consequently, during the computations of the surface plots, a distribution of observed maximum and minimum voltage levels was computed. These data involving the frequency of occurrence of individual voltage levels have been used to construct the probability curves presented in Figure 21. These data show the probability of the peak voltage on the line (either positive or negative) exceeding the voltage level specified on the abscissa. Figure 21a presents the data for the line height of 10 meters and Figure 21b is for the 20 meter line. Different curves are found for each of the four conductivities. In performing these calculations, the positive and negative voltages were found to have the same probability distributions, consequently, only a single set of curves is presented here.

V. CONCLUSIONS

This paper has briefly examined two different ways of specifying the HEMP environment produced by a high altitude nuclear explosion. The first, referred to as the dipole moment model, is derived from considering the EMP source region to be a simple magnetic dipole moment aligned in the direction of the earth's magnetic field, and with certain specified constants being specified from the Bell Laboratory Handbook.

The second model arises from a detailed computation of the electromagnetic fields produced within the source region using the CHAP code, and a subsequent propagation of these fields out to the observer on the earth's surface.

In comparing the HEMP environments produced by these two models, it is found that the dipole moment model generally has a longer fall time than does the CHAP mode. At times, the dipole moment model provides an incident field with a higher amplitude that of the CHAP model, but this is not true for all cases.

In a limited investigation of the field-induced response of above-ground lines, it was found that the dipole moment model tends to predict open-circuit voltages which are larger than those predicted by the CHAP results. This observation is attributed to the fact that the dipole moment field has a longer tail than does the CHAP environment, and this gives rise to a larger response.

Using the CHAP model, a parametric study of the variations of the maximum positive and negative peaks in the open-circuit voltage waveform on a line was carried out. The results of this study indicate that rather large positive peaks in the voltage response can be expected for lines located near the western horizon from ground zero, and correspondingly large negative responses are found on lines near the eastern horizon.

However, the probability of occurrence of these maximum values is small. An investigation into the probability of finding various open-circuit voltage levels was made, and these data have been tabulated. For a 10 meter high line, only about 10% of all possible lines will have an open circuit voltage level exceeding ± 300 KV. For a 20 meter high line, only about 10% of the lines will exceed ± 500 KV.

REFERENCES

1. EMP ENGINEERING AND DESIGN PRINCIPLES, Bell Laboratories Publication, Whippany, NJ, 1975.

2. Legro, J.R., et. al., "Study to Assess the Effects of High Altitude Electromagnetic Pulse on Electric Power Systems, Phase 1 Final Report", Oak Ridge National Laboratory Report ORNL/Sub/83-43374/1/V2, February 1986.

3. Longmire, C.L., et. al., "A Nominal Set of High Altitude EMP Environments", ORNL/Sub/86-18414B/1, Oak Ridge National Laboratory Report, Oak Ridge, TN, to be Published.

4. Longley, H.J., and C.L. Longmire, "Development of CHAP - A High-Altitude EMP Code", Mission Research Corp. Report MRC-R-375, January 1978.

5. Vance, E.F., Coupling to Shielded Cables, Wiley Interscience, New York, 1978.

6. Lee, K.S.H., et. al., "Interaction of High-Altitude Electromagnetic Pulse (HEMP) with Transmission Lines: An Early-Time Consideration", AFWL EMP Interaction Note 435, December, 1983.

7. Tesche, F.M., and T.K. Liu, "Recent Developments in Electromagnetic Field Coupling to Transmission Lines", Proceedings of the 1981 EMC Conference, Zurich, Switzerland, March 10-12, 1981.

8. Flammer, C. and H.E. Singhaus, "The Interaction of Electromagnetic Pulses with an Infinitely Long Conducting Cylinder Above a Perfectly Conducting Ground", AFWL EMP Interaction Note 144, July 1973.

ACKNOWLEDGEMENT

Research is sponsored, in part, by the Office of Energy Storage and Distribution, Electric Energy Systems Program, U.S. Department of Energy, under Contract DE-AC05-84OR-21400 with Martin Marietta Systems, Inc., for the Oak Ridge National Laboratory, and by LuTech, Inc., IR&D funds.

We would like to acknowledge the support and technical interest extended by the staff of the Oak Ridge National Laboratory during the course of this work. In addition, thanks go to Dr. C. Longmire of MRC for discussions pertaining to the CHAP calculations. Finally, thanks go to Sheila Kelly for her help with the preparation of the manuscript.

EMP-induced transients and their impact on system performance

Richard J. Sturm

Federal Armed Forces Defense Science
Agency for NBC Protection, D-3042 Munster
Fed. Rep. Germany

Summary

The interaction of electromagnetic waves with complex systems is still a puzzling phenomenon. This paper presents an attempt how information about the sequence of interaction steps can be extracted from the whole set of data which are recorded during an EMP test of a complex system. The basic idea is to separate the interaction process in two or more steps, the external interaction, coupling through e.g. an aperture, and the internal interaction. The internal interaction results in the all important pin currents-/voltages which endanger the proper function of the system. In case of EMP the induced currents can reach peak values of more than 100 A even on short cables in compact systems (helicopters, tanks etc.). The typ of upset and damage which have been observed after illumination with EMP-like fields are reported. The basic concept of the determination of the EMP-vulnerability is discussed, and explained by examples.

Introduction

EMP induced transients in complex systems are still impervious to quantitativ model calculations. All the more it seems to be useful to analyse measured transients to bring some order in the sea of data which are produced during EMP-tests.

The transients recorded during EMP-tests are usually scanned with respect to the information which is important for the judgement upon EMP-vulnerability. The other information which is contained in the data is ignored to a great extent. One of the reasons might be the lack of standardized evaluation methods. This mainly concern the parameters which are related to the external and internal interaction problem.

The following considerations are a first step towards an empiric evaluation of large samples of transients. The method is demonstrated for a tank. The only reason for this choice is the fact, that for the tank the available set of data is more complete than for other systems. The method as such should be applicable to several classes of objects including helicopters and aircrafts. Moreover the tank has the advantage that the electromagnetic topology is easier to survey. It has a well defined but not perfectly closed metallic case with no long unprotected exterior cables penetrating the (conducting) envelope.

For an incident field the first interaction with the system occurs mainly at the surface of the system. In principle this process is well understood by scattering theory. In practice it is impossible for most real cases to solve the scattering problem analytically. The available numerical methods are very powerful tools for solving special cases, but they mask the fundamental understanding of the problem. The induced transients are always dependent on the properties of the testobject (scatterer) and on the properties of the field. It would be highly desirable to develop easy applicable methods which allow the separation of both parts. The principle how both parts contribute to the transient is described by the singularity expansion method (SEM) which was introduced by Baum /1/ into EMP analyses. Combined with Prony's method /2/ or the method of complex demodulation /2/ SEM has been applied with some success on transients induced by EMP-like fields. The difficulties in using these methods however become severe when the characteristic quantities (e.g. complex singularities) of the object under test are completely unknown. It is an objective of this paper to show that a systematic examination of the whole set of data can yield such information. It is however worthy to note that the information one can extract in this way is incomplete on principle. The number of singularities which characterize a scatterer is usually infinite. What we observe are at best the main resonances (dominant peaks). In particular the oscillations which correspond to singularities far away from the imaginary axis are strongly damped and therefore almost unobservable. The corresponding peaks in the Fourier spectra are broad and unresolved. They more or less build up a homogenious background.

The Interaction Scheme

Based on the assumption that our system has 'good' shields and no unprotected shield penetrating wires it is plausible that the all important pin currents and voltages are not excited directly by the incident field but step by step in an interaction sequence /3/. After the first interaction with the exterior envelope (external interaction) the incident field is superimposed with the scattered field resulting in a total field which penetrates through the aperture into the interior zone. The coupling through the aperture may or may not show resonance character. (In our particular case we are probably below the first resonance of the hatch). The penetrating field excites the cavity's natural modes.

There upon the cavity field couples to the sheaths of the internal cables. Finally the currents and charges on the cable shield couple to interior lines of the cable which transmit the signal to the pins.

The sequence described here contains several steps. The quantities (field, current, charge) excited in a certain step are always the driving force for the next scatterer or oscillator and so on. Using the transfer function or operator notation /3/ a coupling path can be decomposed in the following way.

(1) $(V, I) = Z^{cable} \cdot Z^{internal} \cdot Z^{hatch} \cdot Z^{external} (\vec{E}^{inc}, \vec{H}^{inc})$

with

V, I pin voltage, current

Z^{index} transfer function of the individual interaction steps

$\vec{E}^{inc}, \vec{H}^{inc}$ incident E and H field

Equ.(1) can be written in the form

(2) $(V, I) = (\prod_i Z_i) (\vec{E}^{inc}, \vec{H}^{inc})$

The product goes over all steps of the interaction sequence. $\prod_i Z_i$ can be considered as the global transfer function over all steps.

In case that more than one linear coupling path contributes equation (2) gets the more general from

(3) $(V, I)_{total} = \sum_j (\prod_i Z_i)_j \cdot (\vec{E}^{inc}, \vec{H}^{inc})$

j denotes the coupling path, the sum goes over all of them.

Supposed again this model is an appropriate description we have gained a very powerful tool. If we measure the sequence step by step we can decide in what interaction step a new peak comes in. Peakes which are common to all steps must consequently result from the first interaction step (external interaction in our picture). Missing a peak in a certain step does not mean that the respective mode has not been excited. The measured quantities (field, current and charge density) show a characteristic pattern in space. By accident it may happen that the measurement was done in a node. (In general a complete mapping is impossible for practical reasons)

Test conditions

A tank was placed in a transmission line typ EMP-simulator. The polarization of the test field is vertical (Fig. 1 and 2). The pulse shape of the testfield is shown in Fig. 3, the corresponding Fourier transform in Fig. 4. The field propagates in z-direction (note the coordinate system in Fig. 1). The principal components are E_y for the E-field and H_x for the H-field.

The tank seems to meet the condition of 'good' shielding fairly well. The principle points of entry are the hatch at the top and (may be) the cover of the engine. The hull of the tank is fairly flat with some metallic structure attached to the top. Typical dimension of this structure is 1 to 3 m. During this part of the test the front end of the tank pointed towards the apex of the simulator.

With respect to interaction the cables of the tank can be grouped in two classes, inside and outside of the hull. Primarily the bulk currents of all cables, as far as accessible, were measured. Following the strategy explained earlier also field measurements were done, close to the hatch. All cable current measurements were done with the hatch closed. The field measurements were done with hatch open and hatch closed.

Experimental results

In the view of our model there are three classes of measurements available
1. bulk currents on exterior cables
2. fields (E and H) behind the hatch (at 0.2 and 0.7 m distance)
3. bulk currents on interior cables
(4. pin voltages which are neglected here)

According to the coupling scheme we first consider the external interaction. A randomly selected set of transfer functions of measured bulk currents on exterior cables is shown in Fig. 5 a through f. The dominant common peaks are approximately at 75, 120 and 170 MH. They obviously correspond to modes of the attachment at the top.

The fields measured 0.2 m and 0.7 m behind the closed hatch show a different behaviour. H_x, the main component of the magnetic field measured at a distance of 0.2 m is shown in Fig. 6, the magnitude of the corresponding Fourier transform is shown in Fig. 7. It is more or less smooth with no pronounced peak. The pulse shape in time domain is similar to the incident field with some short peak superimposed at the front end of the pulse.

In contrast to the magnetic field strong oscillations were observed for the y-component of the electric field, measured at the same position (see Fig. 8). The pertinent Fourier spectrum (Fig. 9) shows strong peaks at 52 and 120 MHz and, though less pronounced, peaks at 93 and 170 MHz.

For comparison a randomly selected sample of transferfunctions of bulk currents measured on interior cables is shown in Fig. 10 a through f. As one would expect the peaks at (52), 75, (93), 120, and 170 MHz are common to all transfer functions. The peak at 170 MHz dominates in all cases shown here. This is also true for most of the other transfer functions not shown.

Interpretation

The experimental results seem to support the basic idea of the coupling sequence. The peaks at 75, 120 and 170 MHz obviously correspond to modes of the attachement. The peaks at 52 and 93 MHz which are missing in the transferfunctions of the bulk currents on exterior cables can be attributed by way of trial to surface modes on the hull. Further analysis is necessary to get additional evidence for that. The driving force of the internal interaction is therefore composed of these two parts, in other words two coupling paths contribute.

The interior system consists of quite a number of individual resonators (mainly cables and the cavity). However the idea that the field, penetrating through the hatch, interacts with the individuals independent of the other cables seems to be too simple even for a first order approximation. There is some evidence that we are faced with a kind of collective excitation of the interior system.

Fundamental considerations on the determination of EMP susceptibility

The all important quantities which are responsible for the impact on complex systems are the pin currents and voltages which are induced by the incident field. Unfortunately the number of pins ranges between a thousand and ten thousands (and more) in complex systems. Moreover measurements at the pins require special preparation of the system. Usually break boxes are inserted between the plug of the cable and that of the component, which makes this kind of measurement costly and time consuming. Measurements at all or most of the pins of a system are therefore selfinhibitory. Measurements on pins are an exception. As a substitute usually bulk current measurements /4/ are performed on as many cables as possible. (Typical several hundreds). Besides the significant reduction of the number of measurements this way has the advantage that the system under test remains virtually unchanged. If done properly no additional coupling path is established by the probe and the test set-up. On the other hand we have lost direct information about the pin voltage/current and we have to reconstruct it in an indirect way.

In doing so one has to keep in mind that the relationship between the bulk current, which is mainly the current on the shield, and the current on interior lines is not unequivocal. From a negligible bulk current one cannot deduce a negligible current on interior lines. For illustration the coupling scheme of a helicopter is shown in Fig. 11. The shielded cables which run in the electrically open cockpit can catch much more current then the cables in protected areas. Worst case values are 100 A. In case of a common braided cable shield the transients on the interior lines can be fairly high (rough guess 100 V). At the penetration through the external shielding layer the sheath current is shunted. Behind the shielding layer the bulk current is usually negligible. In contrast to the bulk current the current on the interior lines still exists but it remains undetected by bulk current measurements. An assessment which does not account for that can therefore end up in a complete misjudgement.

Even full threat level tests are not conclusive per se. Besides other effects which can fool the analyst, upset sometimes appears to be of random nature. Under exactly the same external test conditions an upset is sometimes observed sometimes it is missing. This is particularly (but not exclusively) true for digital circuits. Systems obviously have time periods during their working cycle where they are much more susceptable to distortion by transients than during the rest of the time. When a short transient (some microseconds) coincides with the time window of high susceptibility upset may be observed, otherwise nothing will happen. (This is not meant as counter-evidence against a deterministic view of the damage mechanisme. It just reflects the tactical situation, where EMP field and operational state of the system must be assumed to be uncorrelated).

Impact on unhardened systems

During the last years quite a number of unhardened systems (tanks, helicopters, radar systems etc.) have been tested in our facilities. Although there are significant differences between the individual systems what the electromagnetic topology is concerned one can draw some general conclusions which might be of interest.

Tanks with exterior attachement and medium size helicopters show similar response in the sense that the statistical distribution of the peak values of the bulk currents are similar. Fig. 12 shows an example. The sample does not include the cables in the shielded case. Peak values of 10 A are fairly numerous. The 100 A threshold is not exceeded in this case. (On exterior cables of tanks more than 100 A were observed for several times).

Before drawing conclusions about the vulnerability of the system we have to make an assumption on the qualitiy of the cable shields /5/.

None of the systems was EMP hardened, none was brand new. All of them have been in use several years. Years after their installation in military equipment standard braided cable shields have a typical transfer impedance of 0.1 to 1 ohm (total cable, plugs included). Based on the upper value of 1 ohm we estimate that on about 20 per cent of the plugs the pin voltage exceeds the 10 V threshold.

In a few cases the 30 Volt threshold may be exceeded. This is certainly more than enough to cause upset in particular in digital units. This finding is confirmed by the full threat level tests. A frequently observed reversible type of upset manifests itself in a stop of the unit with or without an error indication. After resetting and restarting the system it's function usually recovers. In some cases power must be turned off for a moment for full recovery. The system is obviously caught in an undefined state with no regular return to normal operation.

For a tank this kind of upset might be tolerable if the down time is short enough. For a helicopter in flight it can be fatal if the failure occurs in a flight critical component. For mission critical components, e.g. the fire control unit, a down time of several minutes might be tolerable.

If we confine ourselves to interface circuits permanent damage on hardware is not very likely below peaks values of 20 V. This is one result of our study on standard computer interfaces. Some of them were hard up to 130 V. In some cases even higher voltages were tolerated. From a comparison we conclude that there is some probability of hardware damage. We are just in the range where the highest voltages exceed the lowest damage thresholds. In a somewhat more refined model one has to take into account that vital cables are double screened in many cases for EMC reasons. Putting all facts together we expect single hardware failures on unhardened systems. Electromechanical are not considered to be susceptable in the voltage range discussed above. So far no direct EMP induced damage has been observed on helicopters or tanks.

Conclusions

In unhardened helicopters and tanks EMP is expected to generate pin currents and voltages much above the normal signal level. However, judging from the test results, this condition seems not to be sufficient for an upset in many cases. One (or more) additional condition must be satisfied. The key for understanding seems to be that different operational states show different susceptibility to transients. As a consequence upsets preferably occur by coincidence of the transient with the susceptable operational state. Nevertheless functional upset by EMP fields is very likely to occur.

Permanent damage is difficult to predict without detailed analysis, but one has to face the fact that it can happen.

For further developments the tests yield valuable information for hardening measures.

References

1. C.E. Baum in Felsen (edit), 'Transient Electromagnetic Fields', Chapt. 3 'The Singularity Expansion Method'; Springer-Verlag 1976
ISBN 3-540-07553-4

2. H.T. Davis, 'Use of Complex Demodulation in Analysing Transient Responses'; Mathematics Note 82, Air Force Weapons Laboratory, Kirtland AFB, NM, March 1984

3. Lee, K.S.H. (edit) 'EMP Interaction', Air Force Weapons Laboratory, Kirtland AFB, NM. Dec. 1980 (unclassified)

4. R.J. Sturm, 'Shielding Against EMP Induced Currents and Voltages', AGARD Lecture Series No. 144, AGARD-LS-144

5. E.F. Vance, 'Coupling to Shielded Cables'; Wiley and sons, New York 1978

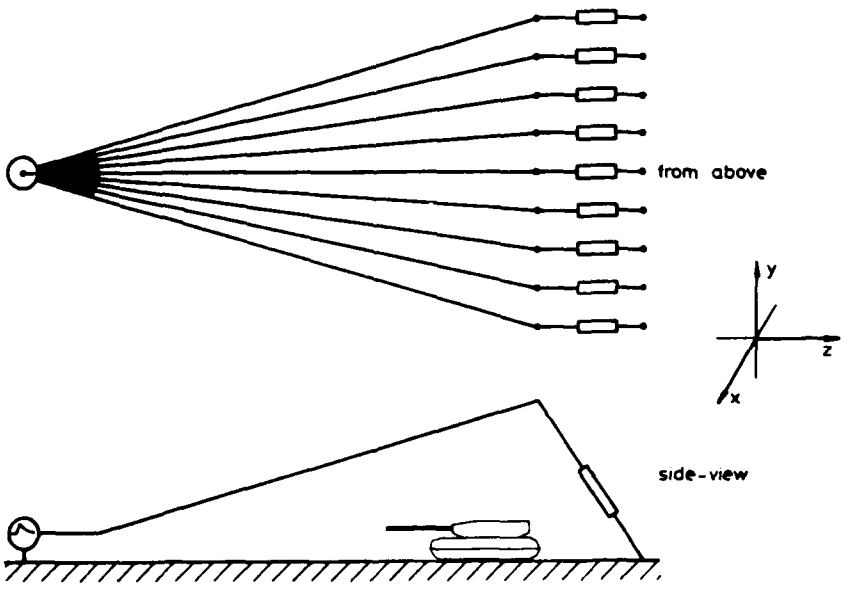


Fig. 1 and 2 Schematic drawing of the transmissionline EMP simulator at WWDBw (top and side view)

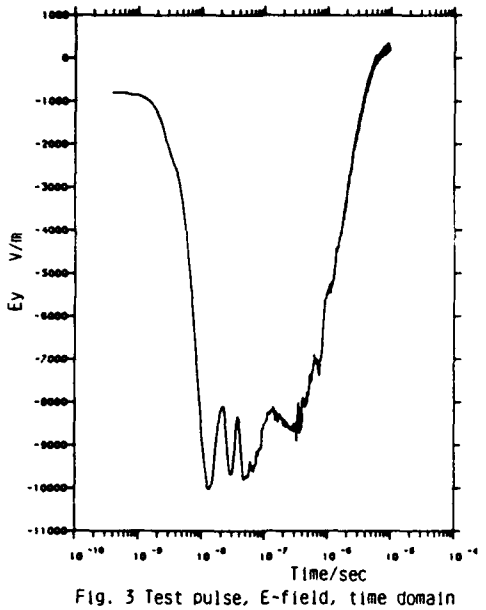


Fig. 3 Test pulse, E-field, time domain

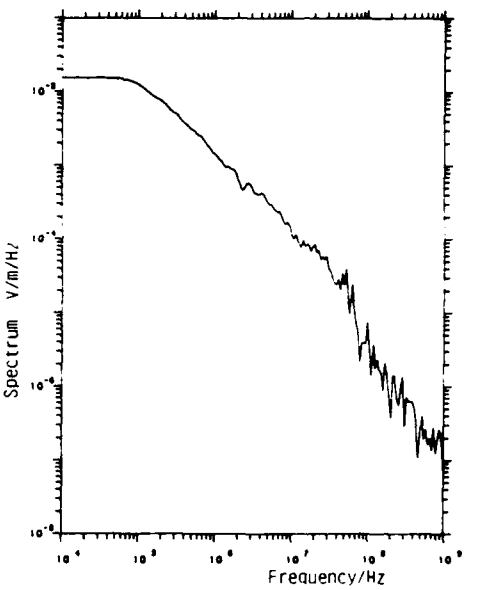


Fig. 4 Test pulse, E-field, frequency domain

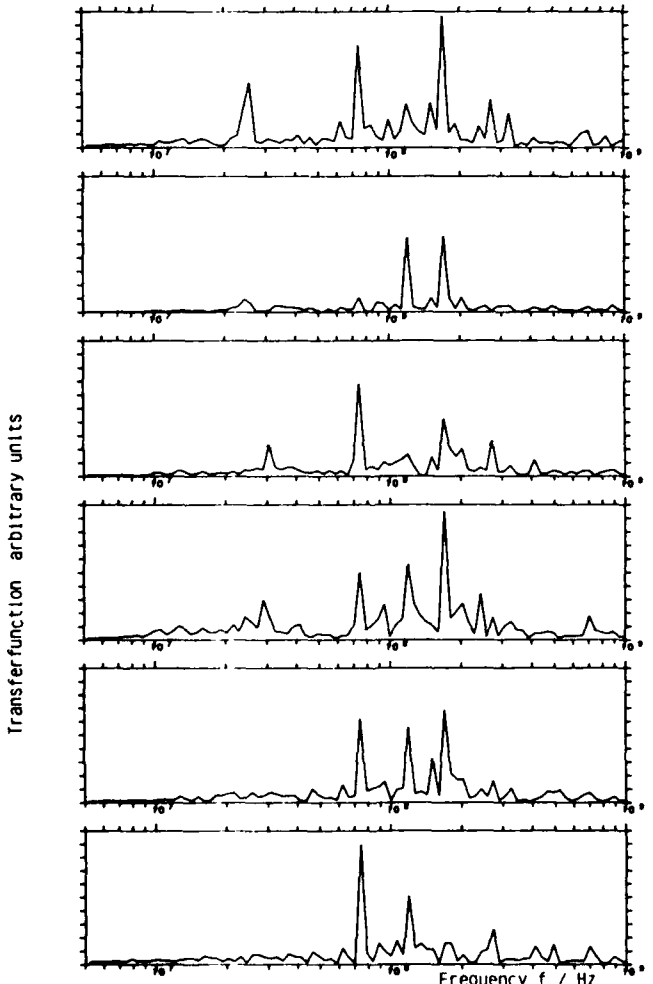


Fig. 5 Transferfunction (bulk current/incident field) of a sample of exterior cables

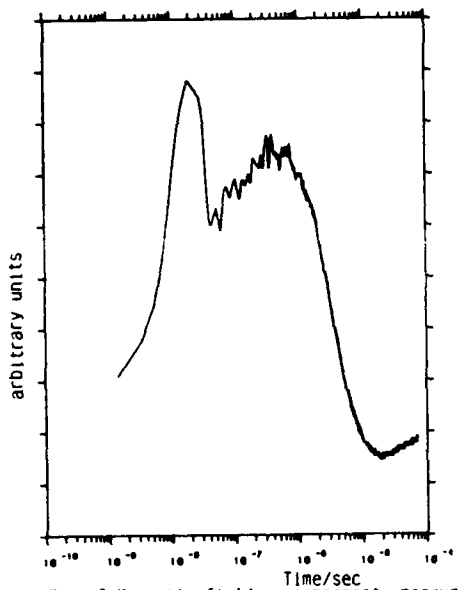


Fig. 6 Magnetic field, x-component, measured 0.2 m behind the hatch, time domain

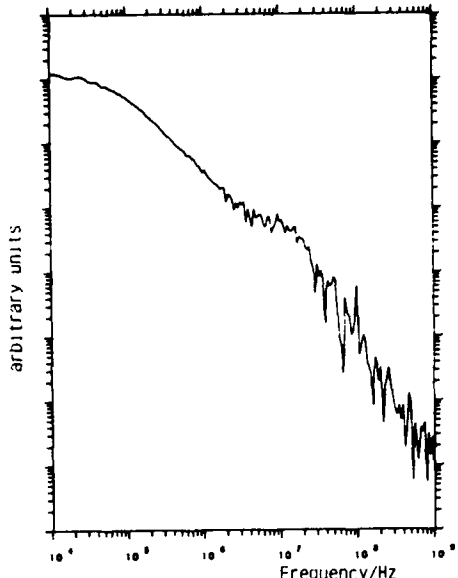


Fig. 7 Magnetic field, x-component, 0.2 m behind the hatch, frequency domain

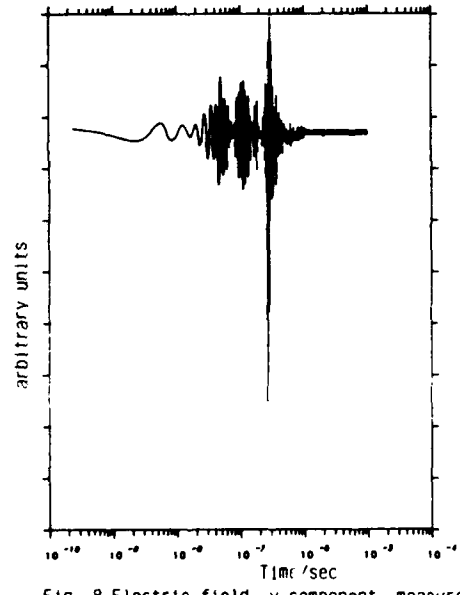


Fig. 8 Electric field, y-component, measured 0.2 m behind the hatch, time domain

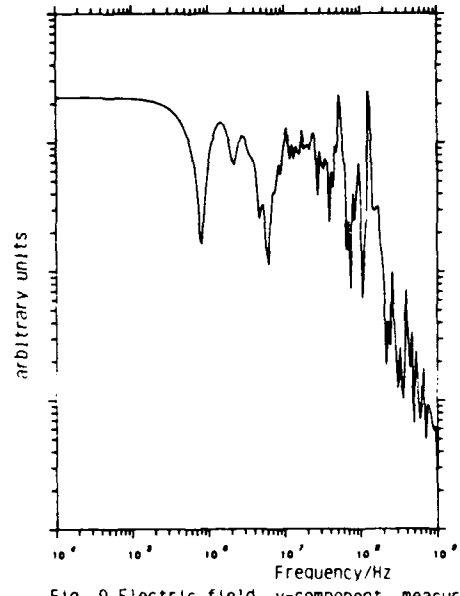


Fig. 9 Electric field, y-component, measured 0.2 m behind the hatch, frequency domain

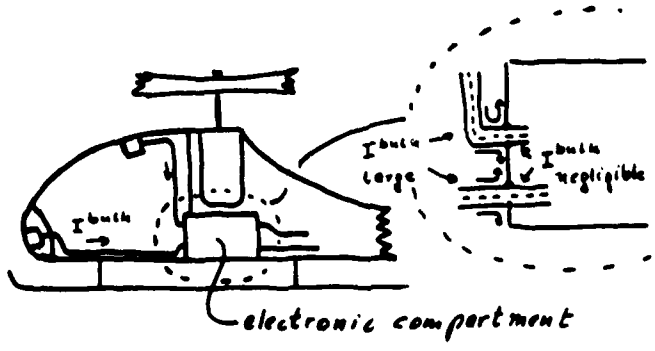


Fig. 11 Partial coupling scheme for a helicopter

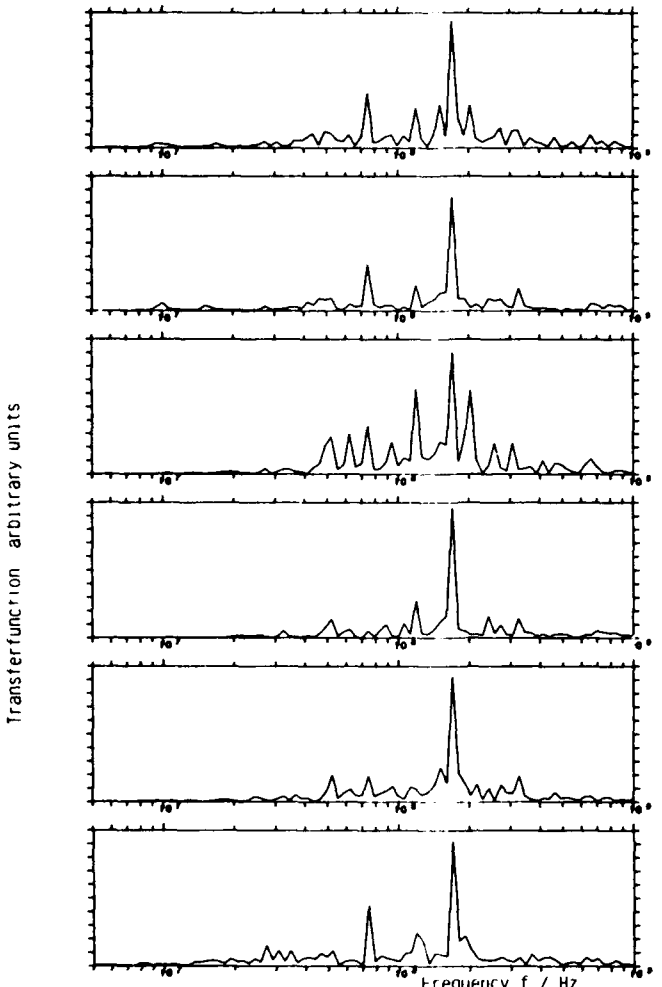


Fig. 10 Transferfunction (bulk current/incident field) sample of interior cables

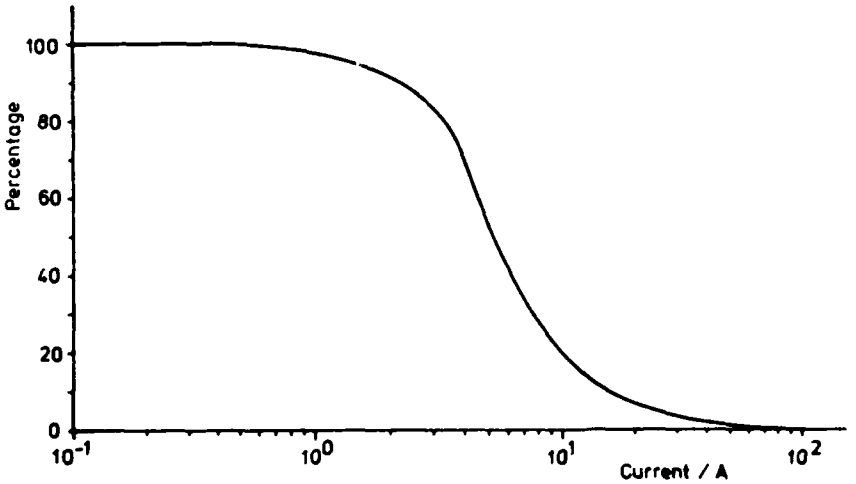


Fig. 12 Statistical distribution of the peak values of bulk currents on a medium size helicopter

SHIELDS FOR PROTECTING CABLES FROM THE EFFECTS OF ELECTROMAGNETIC NOISE AND INTERFERENCE

L. O. Hoeft, J. S. Hofstra,
R. J. Karaskiewicz and B. W. Torres
The BDM Corporation
1801 Randolph Road, S.E.
Albuquerque, New Mexico 87106
(505) 848-5399

SUMMARY

The intrinsic electromagnetic property of a cable or connector shield is its surface transfer impedance. This is the ratio of the longitudinal open circuit voltage measured on one side of the shield (normally the inside) to the axial current on the other side (normally the outside). In cases where a high electric field is present at the surface of the shield, the transfer admittance or charge transfer elastance is also important. Measurements of typical cables, connectors, backshells and cable terminations will be presented and explained in terms of simple models.

INTRODUCTION

Electromagnetic shielding is used to protect electronic equipment and systems from the effects of lightning, nuclear electromagnetic pulse, and electromagnetic interference (EMI or EMC). In general, the hostile electromagnetic environment is outside the system and the electromagnetic shielding is used as a barrier between the hostile environment and the system which must be hardened. Electromagnetic shielding is also used to reduce emanating radiations which could cause electromagnetic compatibility or TEMPEST problems. In this case, the hostile environment is on the inside of the cable and system, and the protected environment is external to the system. In both cases, the shielding acts as a barrier to the electromagnetic radiation. Shielding is added to cables, connectors, and cable assemblies in order to provide this electromagnetic barrier.

Traditionally, some form of shielding effectiveness has been used to specify cable shields and connectors or backshells. The IEEE defines shielding effectiveness as the ratio of the field at a point with and without the shield in place. This has the advantage of being conceptually simple and easy to measure. Unfortunately, there are a number of problems with the concept of shielding effectiveness. The first is that there is a lack of a unique definition. The IEEE definition of shielding effectiveness (the ratio of the field at a point with and without the shield in place) does not fit all situations. In fact, one definition, namely the ratio of the current on the shield to the current on the core wires, has become quite popular. Second, it is not an intrinsic property of the shields because it depends on the external and internal impedances. This has been pointed out by a number of authors, including that of reference 1. Thirdly, there is a lack of an independent calibration of shielding effectiveness. There is no standard shield whose shielding effectiveness can be derived from first principles knowing the geometry and material properties; thus, a shielding effectiveness measurement could be in error due to differences in measurement setup, and the experimenter would not be able to know that he has a measurement problem. In many situations, a particular product can be made to meet specifications simply by varying the reference antenna rather than improving the electromagnetic performance of the product. This is certainly an undesirable situation.

In the 1930s, Schelkunoff showed that the surface transfer impedance was the intrinsic shielding property of cables, connectors/backshells and cable assemblies.[2] Initially, he treated only solid, cylindrical shields. However, his work has been extended to include imperfections in the shield such as apertures, porpoising, etc.

This paper will discuss surface transfer impedance of cables and connectors by first giving the definitions of surface transfer impedance and surface transfer admittance. Then, a brief discussion of surface transfer impedance theory will be presented, followed by typical results.

DEFINITIONS

The surface transfer impedance of a cylindrical shield such as is found on cables, connectors, backshells, and cable assemblies is defined by the relationship shown in Equation 1; namely, the voltage drop on the inside of the shield divided by the current flowing on the external surface:

$$Z_T = \frac{1}{I_o}\frac{dV}{dz}\bigg|_{I=0} \qquad (1)$$

where I_o is the current flowing on the shield and dV/dz is the voltage per unit length on the inside of the shield. Since surface transfer impedance is an intrinsic property of the shield, the reverse configuration can also be used. The current can be on the inside, and the voltage per unit length can be measured on the exterior surface. In most cases, Equation 1 is simplified to the relationship shown in Equation 2 where V_{oc} is the open circuit voltage on the inside of the shield, I_o is the current flowing on the shield, and ℓ is the length of the cable sample. Equation 2 is obtained by integrating Equation 1 along the z axis.

$$Z_T = \frac{V_{oc}}{I_o \ell} \qquad (2)$$

The complementary coupling quantity is the short-circuit current induced on the center conductor for an electric field on the external surface of the shield. This can be calculated using the surface transfer admittance. Traditionally, this is defined by Equation 3:

$$Y_t = \frac{1}{V_o} \left. \frac{dI_{sc}}{dz} \right|_{V = 0} \tag{3}$$

where dI_{sc}/dz is the short-circuit current per unit length flowing on the internal conductor of the cable, and V_o is the voltage between the shield and the external electrode. In most cases, the transfer admittance is related to the transfer capacitance by the relationship shown in Equation 4:

$$Y_t = j\omega C_t \tag{4}$$

where the angular frequency is equal to $2\pi f$, and C_t is the transfer capacitance (i.e., the capacitance between the external electrode and the center conductor of the cable).

Transfer admittance is important when the electric field at the shield is significant. This is usually not the case since the shield is normally grounded. The electric field will be small as long as the cable is electrically small. Under these conditions, transfer admittance can be neglected. In addition, for cables with high optical coverage, the transfer admittance is so small that its contribution can be neglected. Note that the transfer admittance depends on the external circuit as well as the electromagnetic characteristics of the shield. Thus, it is not an intrinsic shielding property.

Several authors have suggested other parameters such as a charge transfer frequency or charge transfer elastance.[3,4] The charge transfer elastance or S_s parameter is the ratio of the transfer capacitance to the internal and external capacitance. This is generally an intrinsic property of the shield. Unfortunately, measurements of transfer admittance are seldom reported. Reference 5 presents one of the few laboratory measurements of surface transfer admittance.

THEORY

For a thin, solid cylindrical shield, only current diffusion is important. In this case, the surface transfer impedance is given by Equation 5:[6]

$$Z_T = R_o \frac{(1 + j)\, T/\delta}{\sinh (1 + j)\, T/\delta} \tag{5}$$

where T is the wall thickness, and δ is the skin depth. The d.c. resistance of the shield, R_o, is given by the following equation:

$$R_o = \frac{1}{2\pi a \sigma T} \tag{6}$$

The skin depth, δ, is the distance which the current can diffuse into the shield material during each cycle. The skin depth can be calculated using Equation 7:

$$\delta = \frac{1}{\sqrt{\pi f \mu \sigma}} \tag{7}$$

where f is the frequency, and μ is the permeability of the shield material which is equal to $4\pi \times 10^{-7}$ μ_r. Note that the skin depth is frequency dependent, varying inversely with the square root of the frequency. Another way to characterize current diffusion is to define its diffusion time constant, namely the time it takes the current to diffuse from the outer surface of a cylindrical shield to the inner surface. The diffusion time constant is given in Equation 8:

$$\tau_s = \mu \sigma T^2 \tag{8}$$

When analyzing surface transfer impedance in the frequency domain, the current diffusion break frequency is often a useful way of characterizing this type of coupling. The current diffusion break frequency, f_δ, is the frequency where the skin depth is equal to the thickness of the shield. Its relationship to the quantities discussed earlier is given in Equation 9:

$$f_\delta = \frac{1}{\pi \tau_s} = \frac{1}{\pi \sigma \mu T^2} \tag{9}$$

The surface transfer impedance of a cylindrical shield, such as described by Equation 5, is shown in figure 1. The surface transfer impedance is normalized by dividing by its d.c. resistance. For frequencies up to the break frequency, the surface transfer impedance is equal to the cable's d.c. resistance. Above the current diffusion break frequency, the surface transfer impedance drops rapidly, indicating that the current no longer diffuses through the shield, and the shield is acting more and more like an impervious electromagnetic barrier. The preceding discussion pertains primarily to cylindrical shields. For shields of other geometries, such as rectangular or shields of arbitrary cross section, the physical principles presented in the preceding discussion are still valid. However, the equations for calculating the sample's d.c. resistance would change. The d.c. resistance can usually be calculated by considering the shield as a piece of metal wrapped around the cable core, and the d.c. resistance of such a piece of metal is the length divided by the product of the cross sectional area and the conductivity.

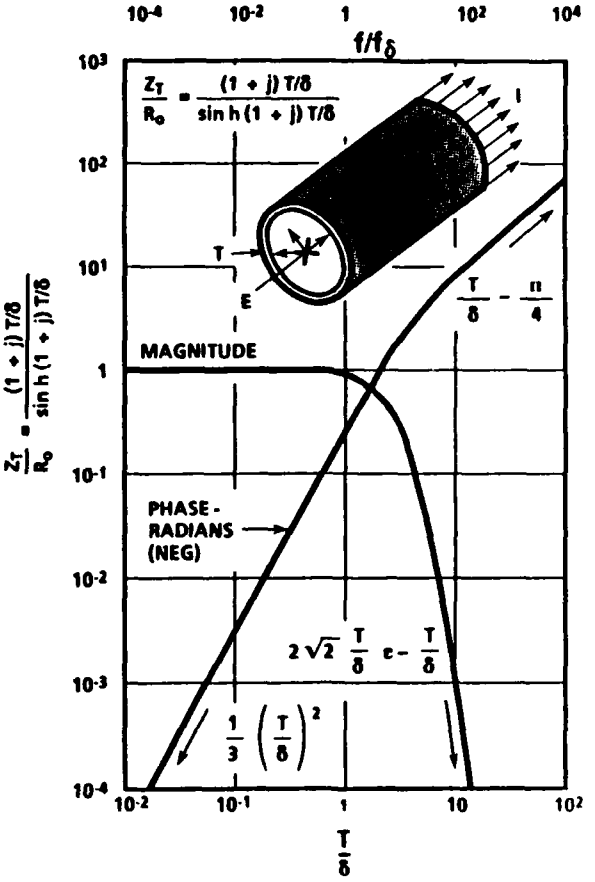


Figure 1. Surface Transfer Impedance of a Solid Cylindrical Shield

If the shield has imperfections such as apertures or penetrations (for example, the carriers of a braided shield), Equation 5 must be modified to account for the coupling due to the imperfections. These imperfections are modeled as a mutual inductance. Thus, Equation 5 can be rewritten by adding a mutual inductance term. The surface transfer impedance then becomes:

$$Z_o = R_o \frac{(1+j)\,T/\delta}{\text{Sinh}\,(1+j)\,T/\delta} + j\omega M_{12} \qquad (10)$$

where ω is the angular frequency and M_{12} is the shield mutual inductance. A braided shield may have a mutual inductance due to both apertures and porpoising. The apertures are formed by the intersections of the carriers. If the braid does not completely cover the exterior of the shield and the optical coverage is less than 100 percent, the braid will have small diamond-shaped apertures at the intersections of the carriers.

Porpoising coupling occurs because of the finite contact resistance or impedance between the carriers as they pass from the outside to the inside of the cable. When the carrier is on the outside of the cable, it carries the external shield current. Because of the finite impedance between the carriers, some of this current remains on the carrier as it reaches the inside of the cable shield. Porpoising coupling is characterized by a surface transfer impedance that increases at 10 dB per decade in the vicinity of 1 MHz and eventually behaves as a mutual inductance. Note that the imperfections, such as aperture and porpoising coupling, are both high frequency effects; they cause the surface transfer impedance to increase with frequency. These imperfections are usually important only above 1 MHz.

Aperture coupling depends on the magnetic polarizability of the apertures. The mutual inductance of an aperture can be predicted using Equation 11:

$$M_{12} = \frac{\mu_o \alpha_m}{(\pi D)^2} \qquad (11)$$

where α_m is the magnetic polarizability. The mutual inductance of a complete cable is the mutual inductance of a single aperture multiplied by the number of apertures. The magnetic polarizability for a circular hole or rectangular slot is shown in Equations 12 and 13 respectively:

$$\alpha_m = \frac{4}{3} r^3 \text{ (Circular Hole)} \qquad (12)$$

$$\alpha_m = \frac{\pi}{16} W^2 \ell \text{ (Rectangular Slot)} \qquad (13)$$

In certain cases where the apertures are significant, such as in a calibration pipe which may have rather large holes, the electric field contribution can be included through the use of an effective magnetic polarizability. This is shown in Equation 14:

$$\alpha_{m_{eff}} = \left(1 + \frac{\alpha_e}{\alpha_m}\right) \alpha_m \qquad (14)$$

In general, the electric polarizability is equal to one-half the magnetic polarizability. Figure 2 shows a typical transfer impedance measurement of a shield with an aperture. This particular sample is a copper pipe with a single hole of various diameters. A similar frequency dependence would be expected of a braided cable where the high frequency coupling was dominated by porpoising coupling.

There are no general relationships for calculating porpoising coupling. In general, however, porpoising coupling is opposite in phase compared to aperture coupling. Thus, aperture coupling can be balanced against porpoising coupling to give an optimized braid. The high frequency performance of most cables is determined by porpoising coupling because it is evidence of more than optimum braid, and most cable designers tend to err on the side of too much optical coverage rather than too little.

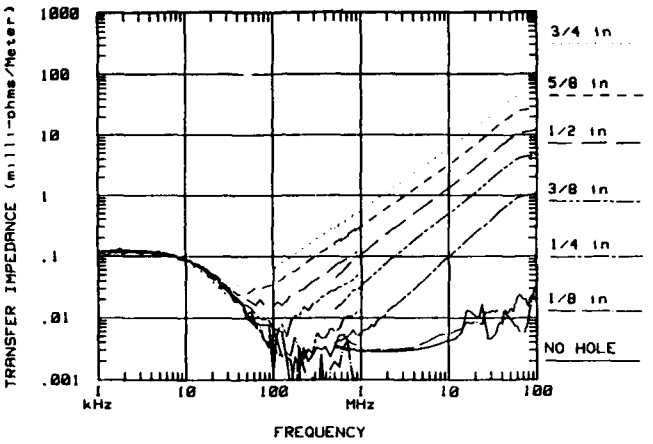


Figure 2. Surface Transfer Impedance of a Shield with Imperfections

MEASUREMENT TECHNIQUES

Surface transfer impedance may be measured using a variety of techniques. In general, the more sophisticated the test fixture, the broader its bandwidth, and the easier it is to use. A review of surface transfer impedance measurement methods is given in reference 7.

All surface transfer impedance measurement methods work at low frequencies. In general, the terminated methods have an octave more bandwidth than the unterminated ones. The use of d.c. resistance measurements is important for establishing the credibility of the surface transfer impedance measurement. The care with which the cable ends are attached to the test fixture usually limits the measurement of very good cables such as solid pipes and conduits. When the cable sample is electrically long, transfer impedance is no longer being measured. Rather, a voltage response that is related to the long line response of a cable is being measured. Computations can be used to derive the transfer impedance from this voltage response. However, such calculations are difficult and full of chances for error.

TYPICAL RESULTS

Solid shields, such as rigid electrical conduit or water pipe, can be extremely cost-effective cable shields for EMP-hardened ground facilities if they are installed correctly. The surface transfer impedance of solid walled nonferromagnetic cable shields, such as copper water pipe can be accurately predicted using equation 5. A 22 mm (7/8-inch) diameter copper pipe with a wall thickness of 1.6 mm (1/16-inch) has a d.c. resistance of less than a milliohm/meter and a diffusion break frequency of less than 10 kHz. A corresponding rigid thin-walled steel conduit has a slightly higher d.c. resistance (a few milliohms/meter) but much lower diffusion break frequency (a few hundred Hertz).[8] Black iron water pipe (33 mm diameter, 3.34 mm wall thickness) has a lower d.c. resistance and diffusion break frequency because of the increased wall thickness. Some measurements suggest that the relative permeability of steel and iron is frequency dependent while other measurements suggest that it is frequency independent at least for frequencies below 1 MHz. However, even when observed, the dependence is never great enough so that it causes an increase in the surface transfer impedance at high frequencies. When relatively new, the compression joint fittings used to join sections of thin-walled steel conduit displayed a very low transfer impedance, being equivalent to a very short section of conduit[8]. The surface transfer impedance of joint fittings using set screws increased as the square root of frequency, suggesting that its performance was limited by contact impedance. Above a kilohertz, the set screw fitting was significantly worse than the compression fitting. Both could be expected to degrade with time. Welded, soldered, brazed, or threaded joints should be more stable.

Figure 3 shows a typical measurement of the surface transfer impedance of a 1 meter-long, tinned-plated, copper braided shield. This figure shows the measured transfer impedance for single, double, and triple overbraids.[9] At low frequencies, the surface transfer impedance, which is really the transfer resistance, is inversely proportional to the number of shields. At high frequencies (above

.5 MHz), the surface transfer impedance decreases about an order of magnitude (20 dB) as each shield is added. Careful examination of the region between 500 kHz and 5 MHz shows that the transfer impedance of the single braid is increasing at the rate of 10 dB per decade rather than the expected 20 dB per decade. This indicates that the primary coupling mechanism was porpoising coupling, at least for the single tubular braid sample, in this set of measurements. Reference 10 presents a worst case model for predicting the transfer resistance and mutual inductance of braided cable shields using jacket diameter as the independent parameter. This simple model was based on the theory presented earlier and was adjusted to fit a large number of transfer impedance measurements performed on a wide variety of braided cables.

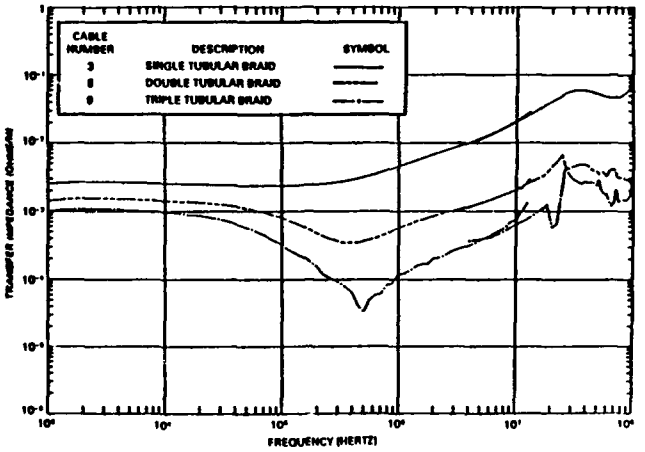


Figure 3. Measured Transfer Impedance of 1-Meter Long, Tin-Plated Copper, Tubular-Braided Shields

Figure 4 shows the measured transfer impedance of three flexible metal core conduit assemblies showing the effect of added tinned copper braid.[11] The measurement labeled "Bronze Overbraid No. 1" is a brass metal core conduit with bronze overbraid. It is typical of a thin, solid cylindrical shield made of relatively low conductivity brass and bronze. Its surface transfer impedance is not particularly good. The addition of a tinned copper overbraid (sample 4) decreases the high frequency transfer impedance so much that it is below the noise level of the system for frequencies above 1 MHz. Figure 5 shows the comparison between brass convolute samples covered with both a tinned copper braid and with either a SnCuFe or tinned copper braid.[11] Notice that the higher permeability of the SnCuFe braid is evident in the measured transfer impedance between about 10 and several hundred kHz. Above 1 MHz, the transfer impedance of the shield is below the measurement capability of the system. SnCuFe has a lower conductivity than tinned copper. This is evident in the higher surface transfer impedance of the SnCuFe sample below 5 kHz.

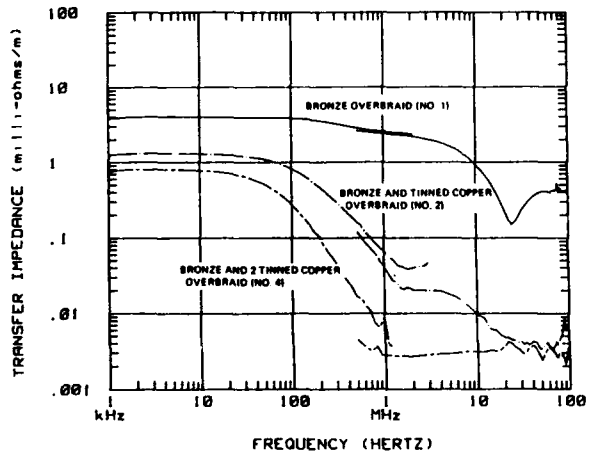


Figure 4. Measured Transfer Impedance of Three Flexible Metal-Core Conduit Assemblies Showing the Effect of Adding Tinned Copper Braids

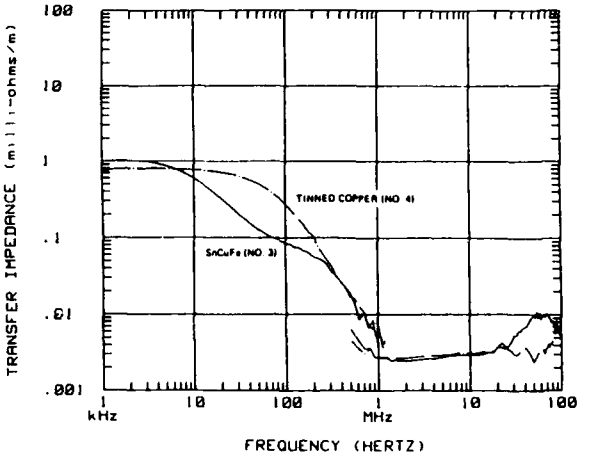


Figure 5. Comparison of Samples With SnCuFe Tinned Copper as the Middle Braid

Figure 6 shows the measured transfer impedance of samples that incorporated ferromagnetic conduits.[9] Sample 58, a molypermalloy annular hose did not have an overbraid, and therefore had a rather high (60 milliohms/meter) transfer resistance. However, it had essentially negligible surface transfer impedance above a couple of MHz. In Sample 15, the high permeability convolute had a current diffusion break frequency below 1 kHz. Therefore, a comparison between the measured surface transfer impedance and its d.c. resistance could not be made. Sample 48, which used mu-metal tape between 2 layers of nickle plated copper overbraid, was one of the lowest transfer impedances measured at our laboratory.

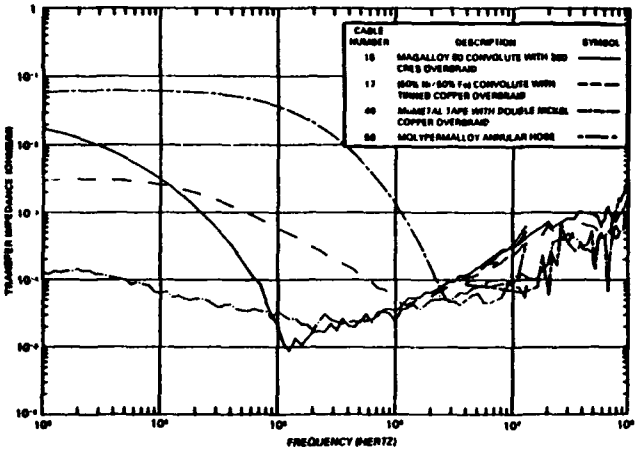


Figure 6. Measured Transfer Impedance of 1-Meter Long Magnetic Shields

Most of the flexible conduits described in the preceding paragraphs were manufactured by forming a spiral strip of metal and soldering the assembly together so that it forms a solid conduit without apertures. The solder ensures that the impedance between the turns of the spiral is very low. Non-soldered spiral conduit is sometimes used where mechanical protection is the primary design requirement. Figures 7 and 8 show the measured transfer impedance of such conduit made of aluminum and stainless steel, with and without an overbraid.[9] Without an overbraid, these conduits have very high transfer impedances since there is little turn-to-turn contact. They can be modelled as a long strip of metal wound into a solenoid. The result is a high transfer resistance and a high transfer mutual inductance. Adding an overbraid reduces the transfer impedance by orders of magnitude.

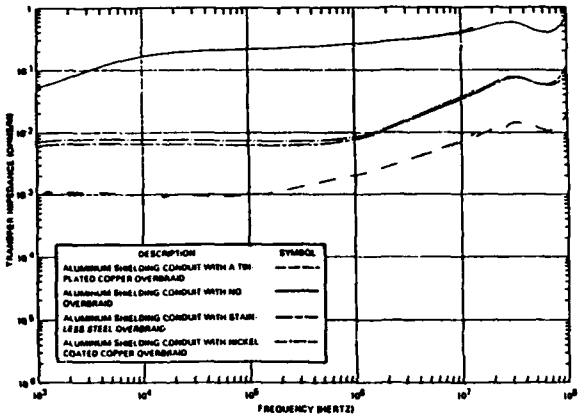


Figure 7. Comparison of the Measured Transfer Impedance of an Aluminum Shielding Conduit Illustrating the Effects of Different Overbraid Materials

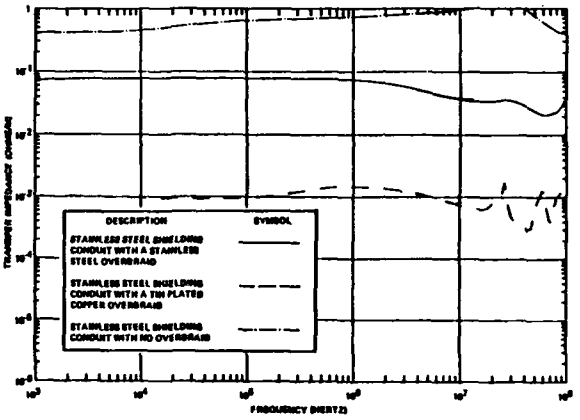


Figure 8. Comparison of the Measurement Data for a Stainless Steel Conduit Showing the Effects of Different Overbraid Materials

In some cases, a shield must be placed over an existing cable or the cable harness is so complicated that a normal machine braid would be too expensive. Metalized plastic tape or knitted wire mesh is sometimes suggested for these applications if the shielding requirements are not too stringent. Figures 9, 10 and 11 show the measured transfer impedance of these types of cable shields.[12] The transfer impedance of circumferentially wound metalized plastic tapes, such as those shown in figure 9, were very high and showed no evidence of turn-to-turn contact. Figure 10 compares the performance of circumferentially and longitudinally applied copper tape. The surface transfer impedance of longitudinally applied tape (cigarette wrap) was surprisingly good (10-15 milliohms/meter). Figure 11 shows that the surface transfer impedance of knitted wire mesh was a few tens of milliohms/m and was frequency independent.

Figure 12 shows the surface transfer impedance of a typical modern connector with an RFI/EMI backshell and a short length of braid measured using a quadraxial test fixture.[13] In this case, the connector was degraded in various ways. The resulting changes in the surface transfer impedance are clearly evident. Figure 13 shows the surface transfer impedance of a connector/backshell/shield termination measured using an inside-out triaxial test fixture.[14]

This connector was one of 22 that were measured in a test series. The mean transfer resistance of the series of measurements was 0.7 milliohms with a standard deviation of 0.3 milliohms. The mean surface transfer impedance at 20 MHz was 1.1 milliohms with a standard deviation of 0.5 milliohms. Note that the transfer impedance shown in Figure 16 does not increase at 20 dB/decade, as would be expected of a shield with an aperture. Instead, the transfer impedance increases at 13 dB/decade, which is closer to the 10 dB/decade expected from coupling due to contact impedance. The mean slope for this set of connectors/backshells/cable terminations was 14 dB/decade with a standard deviation of 5 dB/decade. This suggests that the electromagnetic performance was limited by contact impedance rather than apertures.

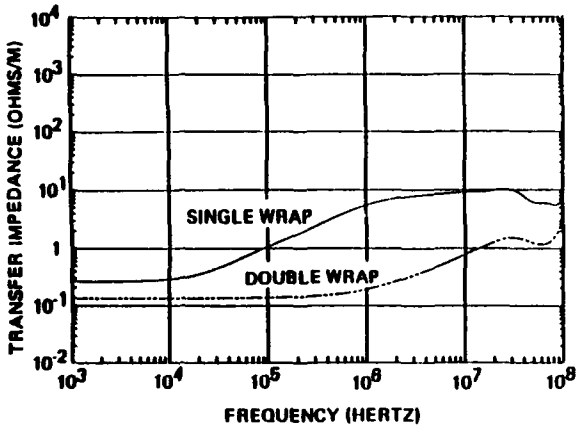


Figure 9. Measured Transfer Impedance of Single and Double Wrap Cable Shields Constructed Using Double Sided Aluminized Polyester Tape (Sample 1)

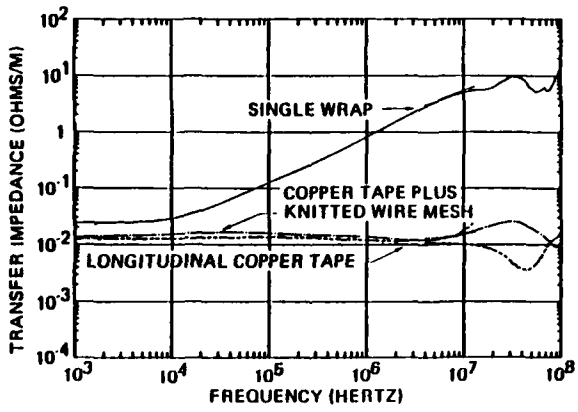


Figure 10. Measured Transfer Impedance of Special Cable Samples (Copper) Tape/Knitted Mesh and Longitudinal Copper Tape)

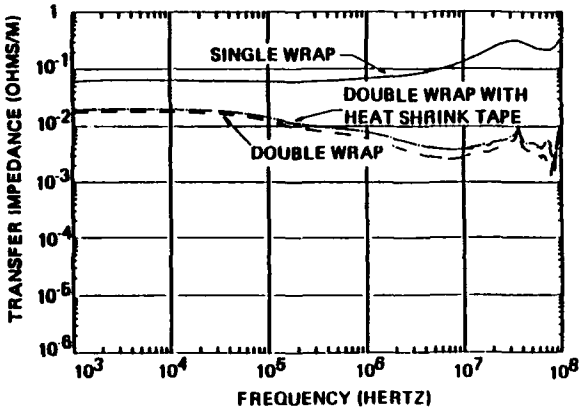


Figure 11. Measured Transfer Impedance of and Double Wrapped Cable Shields Constructed Using Knitted Wire Mesh (60 Percent Overlap)

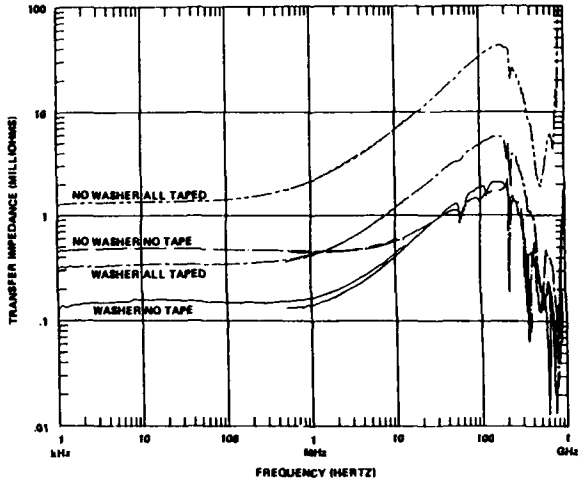


Figure 12. Effect of the Wavy Washer on the Measured Transfer Impedance on a MIL-C-388999-500 Series IV Connector

The first inch or so of braid next to the backshell usually dominates the high frequency measurements that incorporate a braid termination. Small changes in the tension or position of the sample can change the high frequency transfer impedance by at least an order of magnitude because changes in the contributions of the braid apertures and porpoising (which depends on contact impedance and therefore braid tension) will result in large changes in the transfer mutual inductance. Figure 14 shows an example of the changes that can result from small changes in the position and tension of the braided shield.[14] In operational systems, the most common degradation is the loosening of the backshells or improper reassembly of the backshell/braid interface. These degradations can increase the transfer impedance of a cable assembly by orders of magnitude.

Surface transfer impedance measurements are not limited to cables and connectors. They can also be used to characterize any type of shield which is essentially longer than it is wide. Figure 15 shows the measured transfer impedance of a 12-inch diameter graphite epoxy tube in which four copper current diverters were added to its sides.[15] The difference between resistive and inductive current division is clearly evident. Figure 16 shows the measured transfer impedance of an 8" x 24" aluminum cableway in which the transverse screw spacing is varied.[16] Measurements such as those shown in figure 16 can be used to guide the design of such cableways so that they provide sufficient shielding and are economical to produce.

CALCULATION OF SHIELDING EFFECTIVENESS

Irrespective of the problems alluded to earlier in this paper, a shielding effectiveness number is sometimes required to meet a specification. In some cases, this can be calculated from the surface transfer impedance and the geometry and impedances of the test set-up. A popular definition of shielding effectiveness is the ratio (in dB) of the current carried on the core to the current flowing on the shield. Reference 17 calculates shielding effectiveness according to this definition for the case of an electrically long cable which has a uniform distribution of imperfections such as apertures. The theoretical treatment given in this reference assumes matched terminations and calculates a ratio of current shielding effectiveness. This reference shows that the shielding effectiveness can be written as Equation 15. At low frequencies, the shielding effectiveness depends on length, whereas at high fre-

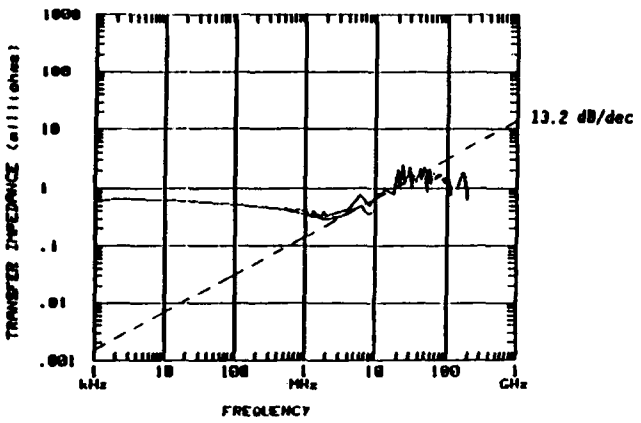


Figure 13. Typical Measurement of a Connector/ Backshell/Braid Termination Using an Inside-Out Triaxial Test Fixture

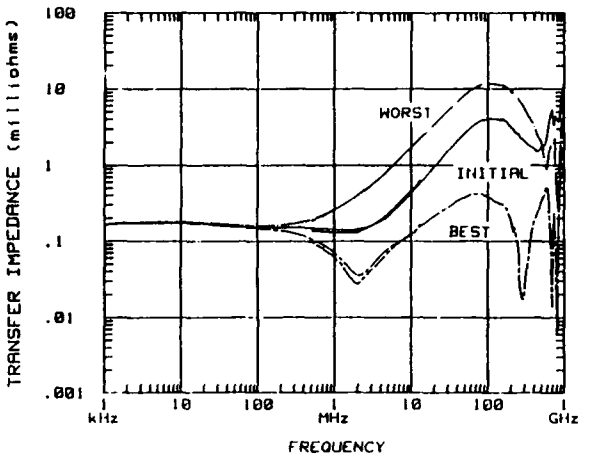


Figure 14. Several Measurements of the Surface Transfer Impedance of a Dual Cone RFI-EMI Backshell Termination, After Braid Tension Was Changed

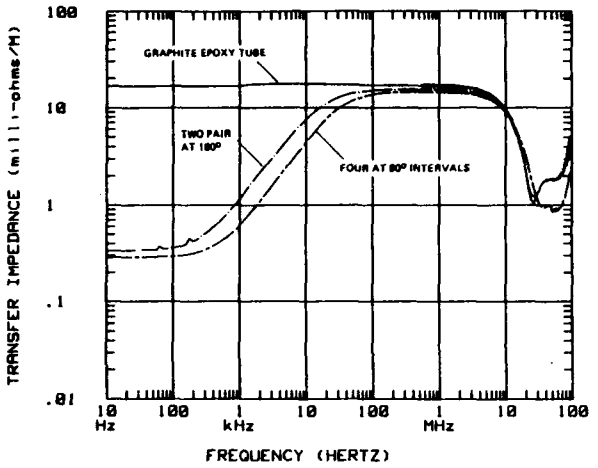


Figure 15. Effect of Current Diverter Placement on Measured Transfer Impedance of Graphite Epoxy Tube With Current Diverters

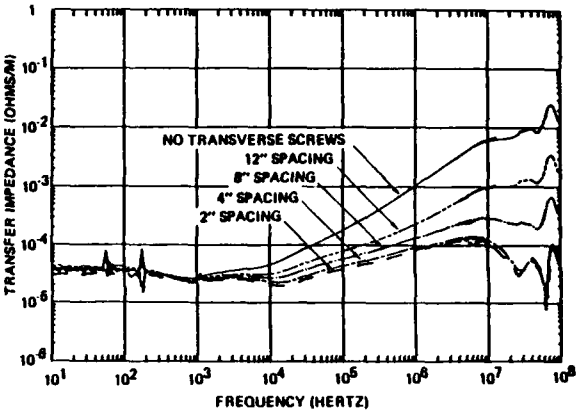


Figure 16. Effect of Transverse Screw Spacing on Measured Transfer Impedance

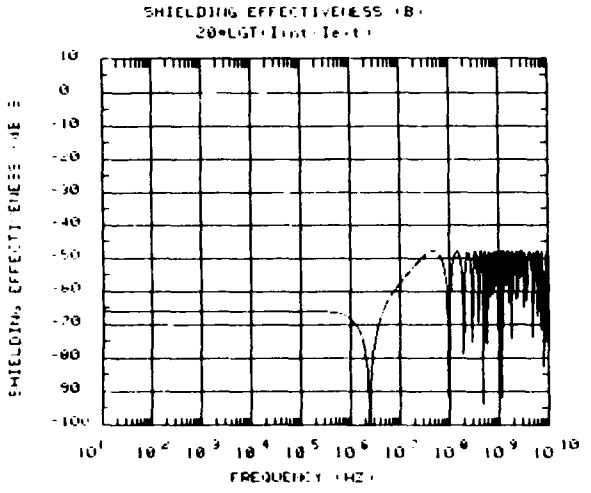


Figure 17. Calculated Shielding Effectiveness Using the Ratio of Currents Definition (SEI) for a 1.2-Meter-Long STP, Type STME 815

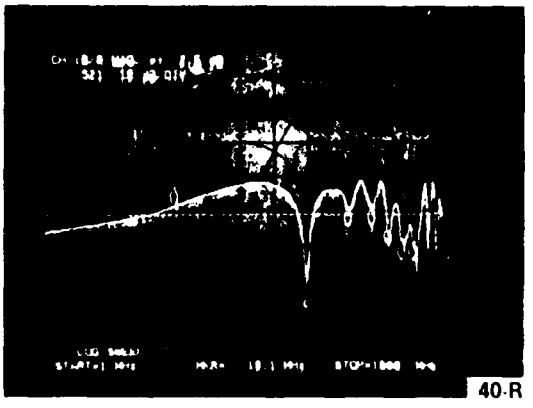


Figure 18. The Measured Voltage in Decibels, Referenced to the Test Fixture Input Voltage, for a Shielded Twisted Pair STME 815 is Shown as a Function of Frequency on a Log Scale for a 1.2 Meter Length Sample. The Center Reference Graticule is -70 dB, which is Equivalent to an SEI of -50 dB. The First Marker is at 10 MHz. All Other Markers Occur at 100 MHz Steps Beginning at 100 MH. The Vertical Sensitivity is Set at 10 dB/DIV.

quencies the mutual inductance coupling is limited by interference between multiple sources. The measured and calculated shielding effectiveness is shown in figures 17 and 18. The shielding effectiveness of a 1.2 meter long shielded twisted pair at low frequencies is about 70 dB, whereas at high frequencies, the shielding effectiveness can decrease to only 48 dB. The surface transfer resistance and mutual inductance of the actual cable can be used to calculate the worst case values for the shielding effectiveness of such a cable.

$$SE = 20 \log \frac{Z_T \ell}{2} + 20 \log \frac{\sin\theta}{\theta} - 10 \log(R_T)^2 \qquad (15)$$

$$\theta = \frac{W}{c}(\varepsilon_r + 1)\frac{\ell}{2}$$

CONCLUSION

Schelkunoff showed that surface transfer impedance is the intrinsic property for describing electromagnetic shields. The transfer impedance measurement techniques are well established. Resistance measurements provide much of the desired information. Calibration samples establish the credibility of the measurement systems. Transfer impedance measurements are available on a variety of shields.

REFERENCES

[1] P. J. Madle, "Cable and Connector Shielding Attenuation and Transfer Impedance Measurements Using Quadraxial and Quintaxial Test Methods," 1975 IEEE Electromagnetic Compatibility Symposium Record.

[2] S. A. Schelkunoff, "The Electromagnetic Theory of Coaxial Transmission Lines and Cylindrical Shields," Bell Sys. Tech. J., Vol. 12, pp. 532-579, Oct. 1934.

[3] K. S. H. Lee and C. E. Baum, "Application of Modal Analysis to Braided-Shield Cables," IEEE Trans. on Electromagnetic Compatibility, Vol. EMC-17, pp. 159-169, August 1975.

[4] R. W. Latham, "Small Holes in Cable Shields," Interaction Notes, Note 118, Air Force Weapons Laboratory, Kirtland Air Force Base, N.M., September 1972.

[5] L. O. Hoeft, and J. Hofstra, "A Simple Technique for Measuring Transfer Admittance/ Capacitance," Proceedings of the 1984 International Symposium on Electromagnetic Compatibility, p. 835, October 1984

[6] E. F. Vance, "Coupling to Shield Cables," Wiley Interscience, John Wiley and Sons, New York, 1978, p. 168.

[7] L. O. Hoeft, "Measurement of Surface Transfer Impedance of Cables and Connectors," Symposium Record of the EMC-EXPO 86 International Conference on Electromagnetic Compatibility, Washington, D. C., p. T17.1, June 1986.

[8] Hoeft, Christine M., "The Measured Frequency Dependance of the Relative Permeability of Ferromagnetic Materials," presented at the New Mexico Science and Engineering Fair, Soccorro, NM, April 11, 1987.

[9] J. R. Hofstra, M. A. Dinallo, L. O. Hoeft, "Measured Transfer Impedance of Braid and Convoluted Shields," published in the Proceedings of the IEEE Electromagnetic Compatibility Symposium, p. 482, September 1982.

[10] L. O. Hoeft, "A Model for Predicting the Surface Transfer Impedance of Braided Cable," Proceeding of the 1986 IEEE International Symposium on Electromagnetic Compatibility, p. 402, September 1986

[11] L. O. Hoeft, J. Hofstra, and J. E. Merrell, "Measured Transfer Impedance of Metallic and Non-Metallic Conduits Covered with Tinned Copper and SnCuFe Braids," Proceedings of the International Aerospace and Ground Conference on Lightning and Static Electricity 1984, pp. 44-1. June 1984.

[12] L. O. Hoeft and J. S. Hofstra, "Measured Transfer Impedance of Metallized Plastic Tape and Knitted Wire Mesh Cable Shields," Proceedings of the IEEE 1983 International Symposium on Electromagnetic Compatibility, p. 296, August 1983.

[13] L. O. Hoeft and J. S. Hofstra, "The Effect of Removing Spring Fingers and the Wavy Washer on the Measured Transfer Impedance of a MIL-C-38999 Series IV Connector," Proceedings of the 1986 IEEE International Symposium on Electromagnetic Compatibility, pp. 155-157, September 1986.

[14] L. O. Hoeft, J. S. Hofstra, J. R. Evert, and J. E. Merrill, "Selected Measurements of Connectors, Backshells, and Cable Terminations," IEEE Transactions on Nuclear Science, Vol. NS-33, No. 6, December 1986.

[15] L. O. Hoeft, V. A. Gieri, and Joseph S. Hofstra, "Measured Transfer Impedance of a 11.5-Inch Diameter Graphite Epoxy Composite Tube with Current Diverters," Proceedings of the 1984 IEEE National Symposium on Electromagnetic Compatibility, p. 35, April 1984.

[16] J. S. Hofstra and L. O. Hoeft, "Measured Transfer Impedance and Current Reduction Rations of a Generic Cableway," Proceedings of the IEEE 1983 International Symposium on Electromagnetic Compatibility, p. 521, August 1983.

[17] M. Dinallo, L. O. Hoeft. and J. Hofstra, "Shielding Effectiveness of Typical Cables from 1 Mhz to 1000 MHz," Proceedings of the 1982 IEEE International Symposium on Electromagnetic Compatibility, p. 489, September 1982.

AUTOMATING SPECTRUM MANAGEMENT OF TACTICAL COMMUNICATIONS TO MINIMIZE MUTUAL INTERFERENCE AND MAXIMIZE FREQUENCY REUSE ON THE BATTLEFIELD

Samuel M. Segner
Center for Command, Control and Communications Systems
Fort Monmouth, New Jersey 07703
USA

Summary.

Automating spectrum management of tactical multichannel LOS and TROPO communications systems began in the early 80's with the establishment of the Automated Battlefield Spectrum Management and Engineering (ABSME) studies and analysis program and its test bed for software evolution, the Army Automated Tactical Frequency Engineering System (ATFES) pilot program in US Army Forces, Europe (USAREUR). The ATFES is a minicomputer based, four echelon pilot system deployed in Europe as a vehicle to research potential battlefield spectrum management capabilities through operational evaluation of software and operating procedures during major USAREUR exercises. The pilot system provides a "go to war" asset for use until the software is transitioned to the communications system control centers/elements for evolving switched systems; it is mounted in existing tactical communications control shelters during exercises. ATFES EMC software and work in HF demand access and VHF CNR management is described in this paper. Interoperable distributed data base processing is the key to spectrum management responsive to highly maneuverable forces.

Background: (See Fig 1)

Up to the mid 70's, spectrum management was essentially propagation charts or calculations, mutual interference charts or calculations, and noise charts or calculations, and some rules of thumb, sometimes called engineering judgement. Although spectrum management algorithms had been fairly well researched, the state of the art in small computers kept the soldier in the field in the manual mode using grease pencils on transparent overlays.

The Automated Spectrum Management and Engineering (ABSME) program which started in the late 70's is aimed at automating spectrum management of all communications and eventually all electronic systems; electronic systems are also called non-communications systems. The international regulations call the combination "telecommunications" systems. In the ABSME program, only the equipment which is seen to be excluded is that of the Intelligence/Electronic Warfare (I/EW) community. It is considered as a major coordination interface and is being explored as "deconfliction" or "fracticide" resolution; this subject will be discussed in another paper at this conference, ref 1a.

The actual work in progress since 1980 is listed in Figure 2. Only the first three will be reviewed to meet the time limitations and theme of this conference. Those listed below DECONFLICTION are only in the study and analysis status.

In the late 70's, the state of the art in transportable data processing equipment reached the poirt where the US Army Chief Signal Officer and his staff recognized the practicalit of field testing spectrum management and communication network engineering software; they took the bold step of using off-the-shelf shock mounted commercial hardware.

Figure 3 outlines the hardware evolution of the Army Automated Tactical Frequency Engineering System (ATFES) pilot project which was fielded in US Army Europe Command (USAREUCOM) forces in the early 80's as a vertical slice of a field army; namely at a division, corps, and echelon-above-corps (EAC) communication engineering facility and at the US Army theater communications headquarters, 5th Signal Command.

Luckily, the system engineers of today recognize that hardware has to be changed to match the needs of the software evolution. During this year, the 16 bit DEC PDP 11/70s are being changed to the 32 bit DEC MicroVAX minicomputers. Theoretically, software development in a given higher order language such as FORTRAN 4 or 77 should be machine independent, but, there is interaction because of evolution of operating system software, peripherals, and data base management and inquiry systems.

Recent successes by other US programs in the use of ADA have convinced USACECOM that all new software for spectrum management should be written in ADA. Due to cost, it will be a while before the older software is rewritten into ADA. Fortunately, both FORTRAN and ADA can run on the MicroVAX simultaneously.

As shown in Figure 4, although the physical presence of the ATFES program is hardware, it is basically a research and development program in software algorithm evolution to optimize use of the spectrum in the presence of fierce competition for spectrum, electromagnetic noise, and interference. The US Army and some allied armies recognize t[illegible] an interim operational capability is available in the current ATFES hardware suite and the test bed is being expanded to a worldwide program

while offering the worldwide participants an interim capability.

It is also recognized that selected modules of the ATFES software may be integrated into other tactical command control facilities which have a responsibility for managing the spectrum-dependent tactical communications and/or electronic systems used in their weapon systems.

Figure 5 shows some of the evolving candidate hardware systems which may use the software algorithms developed via the ATFES program. Without attempting to decode all the "alphabet soup", suffice it to say that management of all tactical communications systems is usually in the system control centers of the multichannel systems while management of avionic, air traffic control, radar, sensor and satellite communications systems is usually in the C-E section of the G3 (operations) shop. Management of deconfliction concerns G2, the EW section of G3, and the I/EW management facilities. The point being made is that the integrated management of tactical communications and electronic systems is foreseen to be accomplished by interoperable distributed processing facilities each of which maintains relevant parts of a distributed data base.

This leads the spectrum engineer and/or frequency manager into a most difficult area of computer applications today, namely interoperability and its ripples as depicted in Figure 6. The sophistication of the electromagnetic compatibility (EMC) engineering algorithms are more than adequate for tactical situations; but, they are only as useful as the accuracy of the distributed data base in reflecting the tactical situation vs the planned maneuvers.

Now that the hardware evolution has been explored, the software evolution can be explained with regard to Figure 7. Based on priorities of US Army use of the spectrum, studies are undertaken to identify how a frequency band should be managed. The user community of the test bed hardware is continually consulted to establish details of software capability and priority of needs. Software configuration control is a major effort and cost item. Even after the software is considered to be mature, changes in equipment and/or concept and doctrine of use can force a new study and development effort.

This year, the software capabilities developed in the ATFES program, most of which are applicable to spectrum management of multichannel line of sight (LOS) and tropospheric (TROPO) radio systems, are being transferred to the Communication System Control Elements (CSCE) of the US Army TRI-TAC switched systems and into their software configuration control. Hence, the feedback loop back to the R&D group has yet to be tried.

Discussion:

The next two figures, Figures 8 and 9, list all the software which has been developed or specified. As previously stated, several which relate to the theme of this conference will be explored. The appendix to a referenced paper (ref 1b) provides details of each software capability for readers interested in what is available and/or in development.

Worldwide Topgraphic Loader (WOTL).

For several decades, personnel concerned with LOS propagation in the 200 to 2000 MHz have been plotting curves showing the difference between theory (propagation loss = K + 20 log (frequency) + 20 log (distance)) and actual results which could be 40 to 60 dB higher as shown in Figure 10.

The US Department of Defense Electromagnetic Compatibility Analyses Center (DOD/ECAC) has used this data to develop a computer program which computes propagation loss for 20 MHz to 20 GHz based on a profile of the terrain between the transmitter and the receiver, Terrain Integrated Rough Earth Model (TIREM).

The WOTL capability as summarized in Figure 11 is to use digitized terrain maps and produce subsets for lower echelon systems which are only interested in a division combat area (nominally a 30 km front by 40 km deep) or a corps combat area (nominally 150 km front by 250 km deep). Level 1 data corresponds to 3 seconds of arc which is about 93 meters at the equator or along a circle of longitude and about 66 meters at 45° latitude.

The current use of digitized terrain maps does not include cultural features; the user usually adds 15 to 30 meters to take trees into account.

Terrain Resources Analysis Program (TRAP).

This program is a major work saver in the process of engineering transportable line-of-sight radio relay and tropospheric scatter radio systems. It provides the path profile and the propagation loss between a transmitter and a receiver, each mounted on an antenna mast of selectable height. The user selects the antenna polarization. As an example of output, the plot in Figure 12 shows the line-of-sight path and under it a plot of the first Fresnel Zone. The software enables the user to rapidly analyze many alternatives about choice of antenna site and height of antenna mast. Of particular interest to military communications is minimizing

profile while maintaining sufficient Fresnel Zone clearance. The user can choose whether he wants a 60% or full Fresnal Zone clearance plotted since this is an engineering judgement factor. Environment parameters which the user can enter or use default values are atmospheric refractivity, ground conductivity, and ground permitivity. Tropospheric refractivity affects the bending of radio waves. Typical sea level values are between 200 and 450 N units. The default valve for the FRG is 314. The choice of ground type (marsh vs average vs desert vs fresh water vs sea water) sets the conductivity and permitivity. The user can enter new types of radio or choose one already in the program and can enter the frequency or use the default to center of tuning range. The stored radios are the AN/GRC-50 (LO and HI band), AN/GRC-103 (bands I, III, and IV), the SHF LOS AN/GRC-143 and the AN/GRC-144 TROPO system.

Other uses of TRAP are natural extensions. Examples are studies of an airborne relay or ground or airborne electronic systems.

The TIREM propagation module is kept up to date by DOD ECAC so it is always as accurate as the state of the art permits. Conversion modules enable input coordinates in latitude and longitude or grid coordinates. Although the computer can handle VHF propagation, this use is rare because of the mobility of the combat net radio (CNR) user.

Offsite Signal to Interference Model (OFFSITE S/I).

This capability, outlined in Figure 13, is an obvious outgrowth of the TRAP module. Although there are many applications for this software capability, it was originally driven by the need to compute compatibility of a ground based civil communications or electronics system with ground based military components of a spaceborne system and vice-versa. Figure 14 is an example of the possible compatibilty issues which may involve military spectrum managers with space based systems in the 1 GHz to 2 GHz band.

OFFSITE S/I is a recent development. It can analyze single transmitter-to-receiver link signal-to-noise and convert that to receiver performance by standardized charts of articulation index or bit error rate based on published standards. It can analyze and compute signal-to-interference and convert that to receiver performance by published curves. It includes a data base of equipment and antenna characteristics to minimize user input for standard equipment.

OFFSITE S/I switches to the smooth earth model, Integrated Propagation System (IPS11) if terrain data is not available. For HF systems, the Ionospheric Communications Analysis and Prediction program (IONCAP11) is used. Consideration is being given to include the International Telecommunications Union (ITU) Appendix 28 computation for calculating the need for coordination between a satellite earth station and a terrestrial communication stations. Median values of man-made noise are stored based on published standards.

While OFFSITE S/I is an accumulation of decades of propagation, interference, noise, articulation and bit-error-rate analyses, it is difficult to use in the peacetime environment despite all its built-in defaults and data bases. Its use in wartime may be only during the planning stages. It may have to be greatly simplified to use it in combat operations where timely response is an overriding consideration.

Line-of-Sight Area Coverage (LOSAC).

As shown in Figure 15, once a computer powerful enough to rapidly process digitized terrain maps becomes available, a large number of three dimensional views become practical. Maps can be produced of radio line of sight for a particular centerpoint on the ground or airborne. By overlaying two maps, the terrain of mutual visibility to two sites can identify sites where a relay can be located. Other uses in hilly terrain involve selecting sites with best natural protection against enemy electronic countermeasures (ECM).

The latest versions of this program include signal level contours. These can be used to evaluate many items of interest to operational commanders. Examples are vulnerability of a multichannel or single channel communications system, radar coverage, electronic warfare coverage, air traffic control envelopes, and affiliation capabilities of mobile subscriber systems. The caution to be noted with this particular software capability is that it does tax computer ability and hence may only be used during the planning stages of combat maneuvers.

Frequency Assignments Capability for Tactical Systems (FACTS).

One of the earliest S-W capabilities to be developed was the FACTS software as outlined in Figure 16. It has been rewritten several times due to its popularity and expansion of applications. This computer module takes into account the usual cosite restrictions for the line-of-sight, multichannel radio relay and tropospheric scatter equipment. Two early added features were to take into account parallel links (e.g. 2 ea 12 channel links to make a 24 channel link where each receiver is in the main beam of two transmitters) and overshoot to avoid reusing a frequency assignment for a receiver in the area of a given transmitter-receiver link which

may be in the main beam coverage of the transmitter.

The most recent version of this S-W capability, FACTS III, has reached a degree of sophistication where it is not practical to use it on a 16 bit machine such as the PDP 11/70. The program takes into account previous assignments at a site and minimizes impact of changing the number and types of radios at a site. In many cases, the S-W program cannot make an assignment at a very complicated node until the interaction with the user permits the software program to remove some of the restrictions. The use of a relational data base provides greatly increased flexibility without consuming extra run time. The latest versions are designed to rapidly run simple changes in the network as a result of an unforeseen redeployment during an exercise when timely response is more important than best EMC engineering.

A planned extension of the program, outlined in Figure 17, would be capable of providing map overlays of the total network for a scale of 1:50000 or 1:250000 military maps. Future planning is to add the antenna beamwidths shown in Figure 18 along the path and annotate those where a large part of the main lobe of the receiving antenna faces the areas where enemy ECM might be employed so that the signal officers can advise commanders of system vulnerability.

HF Systems Engineering.

At tactical echelons, communications propagation often forces a commander to compromise on his maneuvers to maintain command control and timely availability of combat support. Up to the mid 60's, highly maneuverable units had to rely on HF radio to assure coverage beyond line-of-sight. In the mid 60's, tactical satellite communications became available and the problems of the volatile and crowded HF band and its frequency dependent performance made it a "loser".

In the late 70's and early 80's, reality set in:

a. There was not enough channel capacity available in spaceborne elements.

b. Command control of US services were becoming so satellite communications dependent that jamming and destruction was cost effective regardless of the sophistication of the in-built defenses of the space elements.

c. Satellite repeaters were subject to one on many jamming, i.e., one jammer could simultaneously disrupt all users forcing them all to reduce traffic to the minimum obtainable from their ECCM systems.

The pendulum swung the other way in the early 80s and now there is great demand for HF systems, especially those with frequency hopping (FH) as a built in electronic counter-countermeasure (ECCM) capability. The introduction of vertical incidence HF has solved the skip zone distance although the multipath reception has yet to be cleanly resolved.

Studies indicated that there were some management opportunities available with some risk in implementing them. The current practice of providing skywave users with a daytime frequency assignment, a night time frequency assignment, and a transition frequency assignment, meant two frequency assignments went unused at any given time. The success of the satellite communications systems to increase overall throughout by a factor of two to four times by using demand access was noted and gave impetus to investigate HF pooling based on accurate propagation and interference prediction. The first hurdle though was to improve the short term propagation prediction.

HF Propagation Prediction.

There are two approaches at USACECOM. The approach favored by our EUROCAP engineering group believes that new algorithms based on theater-specific measurements should be written. The EUROCAP program uses the vertical incidence pulse sounders as a vertical and oblique sounder. The geography of the involved universities and scientific organizations is shown in Figure 19.

A different approach is used in the ABSME program which has recently passed leadership from the US Army to DOD ECAC. This approach is to take advantage of the existence of a worldwide network of chirpsounders. It is hoped that by measuring propagation on intelligently selected skywave paths and by using the US Navy PROPHET HF prediction algorithms (essentially derived from MINIMUF and IONCAP), an artificial psuedo sunspot number can be computed which is accurate for a region near the mid point of the skywave and that this psuedo sunspot number will be stable for the next three to six hours. The research to prove this is based on inserting one to three chirpsounder receiving systems in this theater. The US Naval Research Center (NRL) in the Washington, DC area and the US Naval Ocean Systems Center (NOSC) in the San Diego, California area lead this scientific effort; the data they collect is being integrated by NOSC with the EUROCAP data.

Frequency Assignment System-HF (FAS-HF).

Software has been developed in accordance with the outline in Figure 20 to

experiment with demand assignment techniques. Ref 4 provides details on this approach which is still evolving. Temporarily, two problems have been put off to future years:

1. Cosite type intermodulation and cosite harmonic interference into radios in the VHF and UHF band; and

2. A means of disseminating new assignments via systems.

VHF Field CEOI.

This software capability was developed to enable personnel in the field who, for some operational reason, did not want to use the paper/hard copy Communications Electronics Operational Instructions (CEOI) (generated and furnished in the US by the NSA) to generate their own CEOI. The EMC cosite ground rules for single frequency combat net radios (SF-CNR), such as the AN/VRC-12 and AN/PRC-77 are shown in Figure 21. The capability has not been extensively used; after the computer has generated the requisite CEOI, a substantial reproduction facility is required to make enough copies.

Figure 22 is an example of a chart which could be manually manipulated or programmed into a computer and can be used to establish frequency-reuse-denial-areas around members of a combat network. There is a set of about a dozen charts like these in current US Army field manuals.

There are several factors to be taken into account when automating combat net radio systems. Cositing/colocation in the same vehicle or command shelter is very important. However, cositing among vehicles is really a judgement factor and many times a severe cosite problem can be allieviated by a small relocation on the part of one vehicle.

In the case of the typical US Army division, there are 300 to 700 VHF nets involving 2000 to 3000 VHF combat net radios. The network structure is planned based on a judgement about deployment and very rapidly changes as the maneuvers begin. The value of charts, such as Figure 22, therefore seems more applicable to setting a denial area around a set of fixed sites rather than between mobile units.

International Efforts.

Nine quadripartite standardization agreements (QSTAGs) have been developed in the ABCA Quadripartite Working Group on Combat Communications (QWG/Comms). These are listed in Figure 23. These have been passed to NATO as an example of the types of agreement which should be made to standardize on a common approach to spectrum management and allow exchange of algorithms, and perhaps software, among Armies.

Conclusion.

The US Army has found the automation of spectrum management necessary, but it must be tempered with prudent reasoning to avoid a degree of sophistication which is not necessary for the uncertainties found in the operational electromagnetic environments as listed in Figure 24. Although many organizations have coded more or less the same EMC algorithms with academically "correct" results, the problem in the field hinges on how accurately the distributed data base can provide good, current information on the tactical situation.

References.

1. S.M. Segner, USACECOM, "Automating the Deconfliction of Jamming and Spectrum Management", AGARD, Oct 1987.
2. S.M. Segner and J. Barry, USACECOM, "The Army Tactical Frequency Engineering System (ATFES)", AFCEA, May 1987.
3. International Radio Consultative Committee (CCIR) Recommendation 527, Annex 1, "Electrical Characteristics of the Surface of the Earth".
4. P.A. Major, USACECOM, "High Frequency Demand Assignment/Pooling", AFCEA, May 1987.

Pig 1

AGARD CONFERENCE, 26-30 OCTOBER 1987

THEME: EFFECTS OF EM NOISE AND INTERFERENCE ON PERFORMANCE OF MILITARY RADIO COMMUNICATIONS SYSTEMS

PAPER # 28: AUTOMATING SPECTRUM MANAGEMENT OF TACTICAL COMMUNICATIONS TO MINIMIZE MUTUAL INTERFERENCE AND MAXIMIZE FREQUENCY REUSE ON THE BATTLEFIELD.

SAMUEL M. SEGNER, USACECOM, FORT MONMOUTH, NEW JERSEY

US ARMY PROGRAM: AUTOMATED BATTLEFIELD SPECTRUM MANAGEMENT AND ENGINEERING (ABSME).

US ARMY TEST BED: ARMY AUTOMATED TACTICAL FREQUENCY ENGINEERING (ATPES) PILOT PROGRAM.

PIG 2

ABSME STUDIES AND SOFTWARE DEVELOPMENT

- MULTICHANNEL LOS AND TROPO RADIO
- HF CNR (SP AND FH)
- VHF CNR (SP AND FH)
- DECONFLICTION VS I/EW AT VHF
- MSRT/SCRA SYSTEMS
- TACTICAL SHF SATCOM SYSTEMS
- RADAR SYSTEMS
- AIR TRAFFIC CONTROL/MANAGEMENT

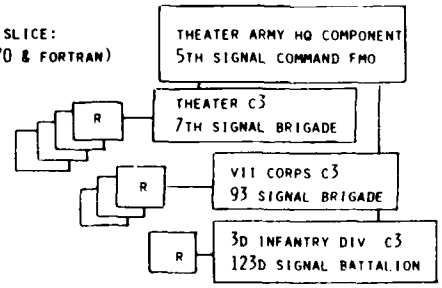
LEGEND: CNR = COMBAT NET RADIO
SP = SINGLE FREQUENCY MODE
FH = FREQUENCY HOPPING MODE
MSRT = MOBILE SUBSCRIBER RADIO TERMINAL
SCRA = SINGLE CHANNEL RADIO ACCESS

FIG 3

H-W IN PILOT SYSTEMS

LATE 70'S: 35 SIGNAL GROUP
WANG 2200 & BASIC)

EARLY 80'S: VERTICAL SLICE:
(PDP 11/70 & FORTRAN)



LATE 80'S: ADD SEVERAL SYSTEMS
(MICROVAX & FORTRAN/ADA)

FIG 4

GOALS AND TOOLS

ATFES IS A TEST BED FOR AN R&D PROGRAM IN SOFTWARE (S-W) EVOLUTION.

GOAL: ACHIEVE "EXPERT" STATUS IN SPECTRUM MANAGEMENT.

CURRENT TOOLS:

FIELD PILOT SYSTEM IN US ARMY EUROPE TO:

A. R&D S-W CAPABILITIES.

B. PROVIDE INTERIM OPERATIONAL CAPABILITY.

FUTURE TOOLS:

ADP HARDWARE (H-W) OF VARIOUS US ARMY C3I AND WEAPONS SYSTEMS WITH NEED TO MANAGE THEIR SPECTRUM DEPENDENT COMMUNICATIONS AND ELECTRONIC (C-E) EQUIPMENT.

FIG 5

ARMY C3I PROGRAMS WITH BFSM NEEDS

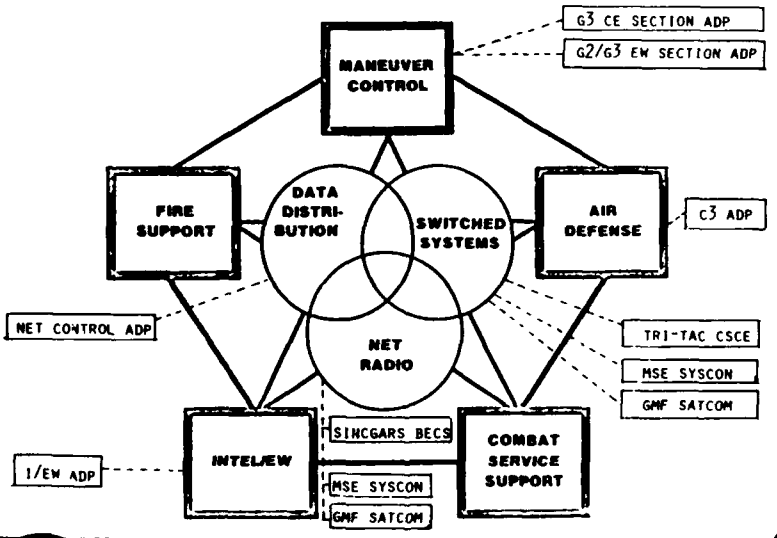


FIG 6

INTEROPERABILITY RIPPLES

- HETERGENOUS H-W SYSTEMS
- STANDARDIZED PROTOCOLS, PROCEDURES & MESSAGE FORMATS
- DATA BASE FORMAT, PROCESSING & UPDATING
- LOCAL AND WIDE AREA NETWORKING
- ACCESS AND SECURITY CONTROL
- ELECTRICAL STANDARDS

FIG 7

ABSME S-W EVOLUTION PROCESS

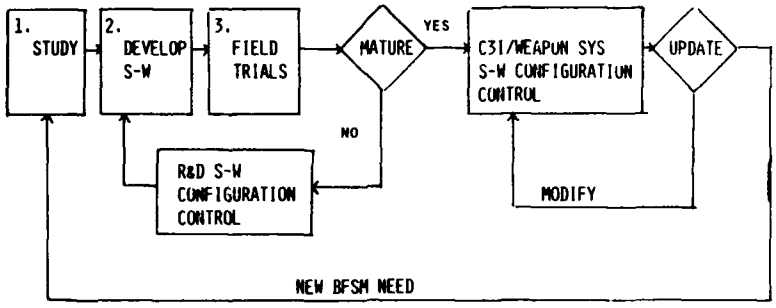


FIG 8

ATFES APPLICATIONS SUFTWARE IAW APPENDIX A

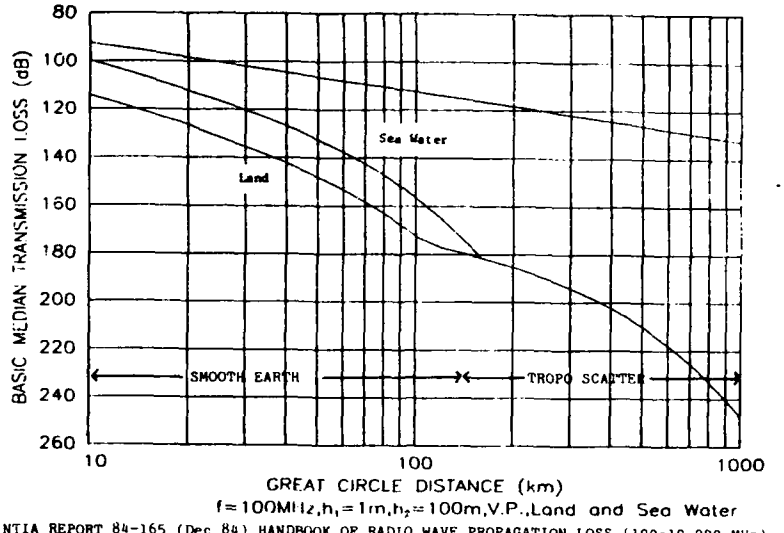
1. FREQUENCY ASSIGNMENT CAPABILITY FOR TACTICAL SYSTEMS (FACTS)
2. WORLDWIDE TUPOGRAPHIC LOADER (WOTL)
3. TERRAIN RESOURCES ANALYSIS PROGRAM (TRAP)
4. LINE OF SIGHT AREA COVERAGE (LOSAC)
5. BASE HILL-PATH HILL (BH-PH) ARCHIVING
6. HIGH FREQUENCY S-W CAPABILITIES
7. FIELD CEOI FOR SINGLE FREQUENCY (30 TO 88 MHZ)
8. COSITE EMC ANALYSES
9. OFFSITE SIGNAL/INTERFERENCE ANALYSES (OFFSITE S/I)
10. DECONFLICTION (VHF-DECON)
11. RADAR FREQUENCY ENGINEERING SYSTEM (RAFES)
12. BATTLEFIELD ELECTRONIC EQUIPMENT DATA BASE (BEED)
13. CIRCUIT ROUTING AND STATUS SYSTEM (CRASS)
14. C-E EQUIPMENT AND ORGANIZATIONAL STATUS (CEEOS)
15. TACTICAL AUTOMATIC SWITCH LOAD (TASL)

FIG 9

ATFES APPLICATIONS SUFTWARE IAW APPENDIX A (CONT)

16. GENERAL UTILITY (GUT)
17. FIELD OPERATIONS ORDER GENERATOR (FOOG)
18. LOG ON/MATRIX ACCESS (LOMA)
19. SYSTEM RECOVERY (SYSREC)
20. AUTOMATED FREQUENCY ASSIGNMENT INFORMATION PROCESSING SYSTEM (AFAIPS)
21. PROJECT DIARY FILE (PDF)

FIG 10 PROPAGATION VS FREE SPACE LOSS



f=100MHz,h_1=1m,h_2=100m,V.P.,Land and Sea Water

NTIA REPORT 84-165 (Dec 84) HANDBOOK OF RADIO WAVE PROPAGATION LOSS (100-10,000 MHz)
REP: CCIR SG 5 Report 714.

FIG 11

WORLDWIDE TOPOGRAPHIC LOADER (WOTL)

INPUT: US DEFENSE MAPPING AGENCY LEVEL 1 TOPOGRAPHIC DATA TAPES.

PROCESS: CONVERT TO TOPUGRAPHIC DATA FILES FOR SELECTED SUB-AREAS.

OUTPUT: DIGITIZED TERRAIN FOR USE BY TRAP AND LOSAC.

FIG 12

TERRAIN RESOURCES AHALYSIS PROGRAM (TRAP)

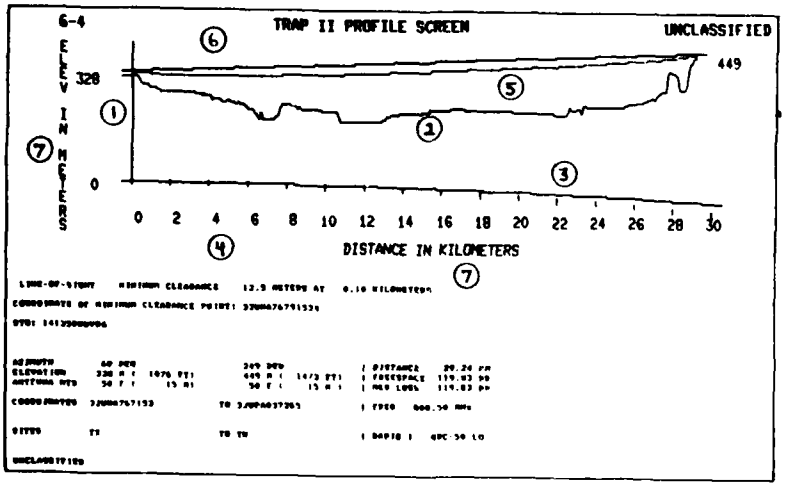
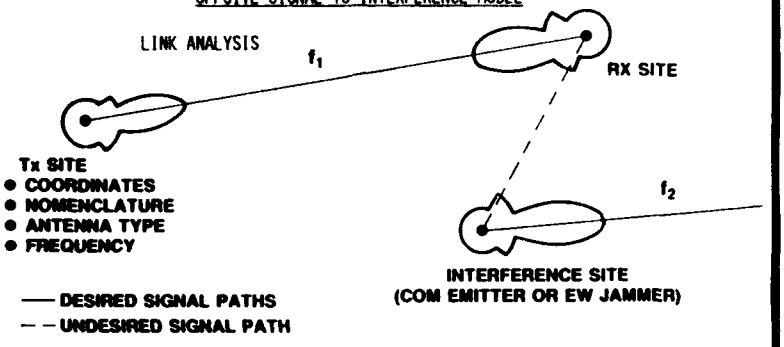


FIG 13

OFFSITE SIGNAL-TO-INTERFERENCE MODEL



- **USES DIGITIZED TERRAIN DATA OVER CONSIDERED PATH(S)**
- **DISPLAYS RESULTS IN THE FORM OF DESIRED LINK PERFORMANCE**
- **MODEL CAN ALSO BE USED TO EXAMINE EFFECTS OF FRIENDLY EMITTERS ON EW RECEIVERS**

FIG 14

EXAMPLE OF (1 GHZ TO 2 GHZ) OF POTENTIAL SPACE VS TERRESTRIAL COMPETITION IN REGION 1

MEGAHERTZ	SPACE ALLOCATION
1215 - 1260	RNAV (SE)
1400 - 1427	RES (P), EE (P)
1427 - 1429	OPS (ES)
1525 - 1530	OPS (SE), FIXED, EE
1530 - 1535	OPS (SE), MMOB (SE), EE
1535 - 1544	MMOB (SE)
1544 - 1545	MOBILE (SE)
1545 - 1559	AMOB (SE)
1559 - 1610	RNAV (SE)
1626.5 - 1645.5	MMOB (ES)
1645.5 - 1646.5	MOBILE (ES)
1646.5 - 1660.5	AMOB (ES)
1660.5 - 1668.4	RES (P)
1670 - 1710	MET (SE)

LEGEND:

RES	RESEARCH
ES	EARTH TO SPACE
RNAV	RADIO NAVIGATION
MET	METEOROLOGICAL
OPS	OPERATIONS
EE	EARTH EXPLOITATION
MOBILE	COMMUNICATIONS
AMOB	AERONAUTICAL MOBILE
MMOB	MARITIME MOBILE
P	PASSIVE
FIXED	COMMUNICATIONS
SE	SPACE TO EARTH

US ARMY COMMUNICATIONS ELECTRONICS COMMAND

FIG 15

LINE-OF-SIGHT AREA COVERAGE II LOSAC II

- **PRODUCE MAP OVERLAYS SHOWING CONTOURS OF:**
  - **RADIO LINE OF SIGHT**
  - **POWER DENSITY**
  - **PATH LOSS**
  - **TERRAIN ELEVATION**
- **UTILIZE WOTL-CREATED TOPOGRAPHIC DATA BASES**

FIG 16

FREQUENCY ASSIGNMENT FOR TACTICAL SYSTEMS (FACTS)

INPUT: FREQUENCY RESOURCES UHF (225 TO 1850 MHZ) BAND AND SHF (4.4 TO 5.0 GHZ) BAND.

MULTICHANNEL MODE STRUCTURE AND SITE COORDINATES.

EXTERNAL PREASSIGNMENTS AT SITE.

CONSTRAINTS: LOS RADIOS: AN/GRC-50; AN/GRC-103; AN/GRC-144; USER CHARACTERIZED.

TROPO RADIOS: AN/GRC-143; USER CHARACTERIZED.

EMC:

XMTR VS XMTR	RCVR VS RCVR
XMTR VS OWN RCVR	XMTR VS CoSITE RCVR
IMAGE FREQUENCY	3,5,7 I.M. PRODUCTS
SPURIOUS FREQUENCY	INTERSITE XMTR VS RCVR
OVERSHOOT	PARALLEL LINKS

OUTPUT: FREQUENCY ASSIGNMENTS AZ/EL
MAP OVERLAYS OF TOTAL NETWORK

FACTS I - MAY 81 (ABM USING FM 24-21/AUTO CRITERIA REDUCTION)
FACTS II - JAN 84 (NETWORK PREDJUDGE/USER CONTROL CRITERIA)
FACTS III A/B LATE 87/EARLY 88 (TIME ORIENTED RUNS/RELATIONS DATA BASE)

FIG 17

MULTICHANNEL DIAGRAM

- **AUTOMATICALLY GENERATE A BUBBLE DIAGRAM OF A TACTICAL MULTICHANNEL DEPLOYMENT FROM A STORED DEPLOYMENT DATA BASE**
- **VARIOUS MAP SCALES**
- **ONE-PAGE OUTPUT ILLUSTRATING NETWORK CONNECTIVITY**

FIG 18

NOMINAL ANTENNA DIRECTIVITY

RADIO	MEGAHERTZ	GAIN (DBI)	BEAMWIDTH (DEGREES)
AN/GRC-103			
1 (NATO I)	220.0 - 404.5	11	67
2 (NATO II)	394.5 - 705.0	13	43
3 (NATO II)	695 - 1000	14	32
4 (NATO III)	1350 - 1850	17	13
AN/GRC-144			
SHF (NATO IV)	4400 - 5000	37	3
AN/GRC-143			
TROPO	4400 - 5000	40	1

PIG 19

HF CHANNEL ESTIMATOR AND EUROCAP* DEVELOPMENT

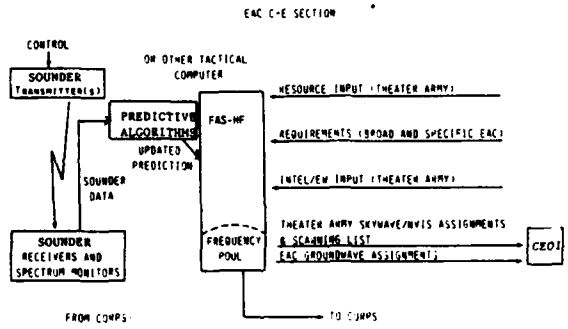
STATION	LATITUDE	LONGITUDE	VERTICAL SOUNDINGS BEGIN	OBLIQUE SOUNDINGS BEGIN
SLOUGH, ENGLAND	51.5N	359.4E	DEC 86	JUL 87
DOURBES, ENGLAND	50.1N	4.6E	DEC 86	JUL 87
HAMBURG, GERMANY	54N	9.2E	DEC 86	JUL 87
MUNICH, GERMANY	48N	11.5E	JUL 87	JUL 87 *GFE SOUNDER
TURIN, ITALY	45.2N	8.3	DEC 86	JUL 87
ROME, ITALY	41.8	12.7E	DEC 86	JUL 87
TORTOSA, SPAIN	40.8N	0.5E	DEC 87	DEC 87 *NEW SOUNDER

*EUROPEAN COMMUNICATIONS ANALYSIS PROGRAM
CONTRACTOR: ALCOA DEFENSE SYSTEMS, SAN DIEGO, CALIFORNIA

US ARMY COMMUNICATIONS ELECTRONICS COMMAND

PIG 20

OPERATIONAL SYSTEM CONCEPT - EAC



*HF RECEIVERS/SPECTRUM MONITORS ARE LOCATED IN CORPS/DIVISION AREA. EACH PROVIDES THE DATA SYSTEM & PROVIDES THE DATA TO EAC. IF THE DATA COMES DIRECTLY TO EAC, EAC PROVIDES THE DATA SYSTEM & PROVIDES THE DATA TO CORPS.

US ARMY COMMUNICATIONS ELECTRONICS COMMAND

PIG 21

MINIMUM ANTENNA DISTANCE SEPARATION (WHIP ANTENNA OR RC-292)

Minimum Frequency Separation Required	Between AN/VRC-12 Series Radio on HIGH Power	Between AN/VRC-12 Series Radios on LOW Power or AN/PRC-25 and AN/PRC-77
10 MHz	[illegible] feet	5 feet
7 MHz	60 feet	10 feet
4 MHz	150 feet	50 feet
2 MHz	400 feet	200 feet
1 MHz	800 feet	350 feet

Other rules: Avoid image frequency: $f_z = f_1 \pm 11.5$ MHz.

Avoid multiples of image (½ and 2):

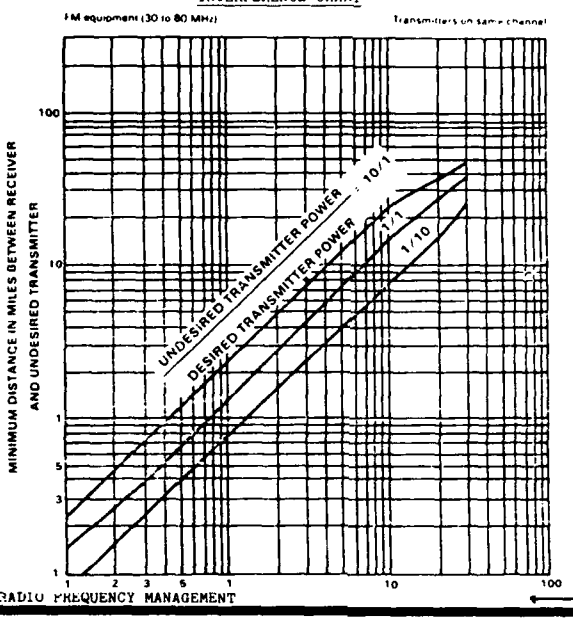
Avoid second harmonics, $f_z = 2f_1$

FM 24-2 Radio Frequency Management

US ARMY COMMUNICATIONS ELECTRONICS COMMAND

FIG 22

INTERPERENCE CHART



FM 24-2 RADIO FREQUENCY MANAGEMENT

US ARMY COMMUNICATIONS ELECTRONICS COMMAND

FIG 23

ABCA QSTAGS ON BATTLEFIELD SPECTRUM MANAGEMENT

QSTAG	TITLE	CUSTODIAN ARMY	STATUS
679	MILLIMETER WAVES AND LASERS (30GHZ & ABOVE)	CAN	APVD
715	HF (1.5 - 30 MHz BAND)	AUS	APVD
716	VHF (30 - 88 MHz BAND)	UK	DRAFT
717	VHF (88 - 225 MHz BAND)	US	DRAFT
718	UHF (225 - 400 MHz BAND)	CAN	APVD
719	TACTICAL RADIO RELAY AND TROPOSPHERIC SCATTER SYSTEMS	CAN	APVD
720	GMFSATCOM SYSTEMS	US	DRAFT
721	RADAR, POSITION NAVIGATION, TACTICAL SENSOR SYSTEMS AND MULTIFUNCTION INFORMATION DISTRIBUTION SYSTEMS	US	APVD
723	DECONFLICTION	US	APVD

US ARMY COMMUNICATIONS ELECTRONICS COMMAND

FIG 24

JUDGEMENT FACTORS VS SOPHISTICATION
(HENCE COST) OF SOFTWARE CAPABILITY

- PROPAGATION STANDARD DEVIATION IS USUALLY BETWEEN 6 ANT 10 DB.
- MOBILE USERS CAN BE ABOUT 10 KM FROM THE MODELED SITE DUE TO MANEUVER VARIATIONS.
- COMPUTERS CAN SLOW UP TIMELY RESPONSE AND RESULT CAN BE NO BETTER THAN SOLDIERS ENGINEERING JUDGEMENT.
- FOLIAGE CAN STOP INTERFERENCE ABOVE 100 MHz.
- ENEMY USES SAME SPECTRUM IN WARTIME.

US ARMY COMMUNICATIONS ELECTRONICS COMMAND

APPENDIX A

COMPONENTS OF RESTRICTED FREQUENCY LIST (Ref 3)

TABOO. TABOO frequencies are any friendly functions/frequencies of such importance that they must never be deliberately jammed or interfered with by friendly forces. Normally, these functions/frequencies include international distress, stop buzzer, safety, and controller frequencies. These functions/frequencies are generally long standing; however, they may be time-oriented in that, as the combat/exercise situation changes, the restriction may be removed by the originating headquarters.

PROTECTED. Protected functions/frequencies are those friendly functions/frequencies used for a particular operation, identified and protected to prevent them from being inadvertently jammed by friendly forces while active EW operations are directed against hostile forces. These functions/frequencies are of such critical importance that jamming should be restricted unless absolutely necessary or until coordination with the using unit is made. They are generally time-oriented, may change with the tactical situation, and must be updated periodically.

GUARDED. Guarded functions/frequencies are enemy functions/frequencies that are currently being exploited for combat information and intelligence. A guarded function/frequency is time-oriented in that the list changes as the enemy assumes different combat postures. These functions/frequencies may be jammed after the commander has weighed the potential operational gain against the loss of technical and tactical information.

COSITING OF RADIO TERMINALS

Giuseppe Cucinotta
ELMER
Industrie per lo Spazio e le Comunicazioni
S.p.A.
Viale dell'Industria, 4 - 00040 Pomezia (Roma)
Italy
and
Raffaele Azzarone
EPP Member
Mariteleradar - Livorno
Italy

SUMMARY

The causes of the radiocommunication channel degradation are identified when receivers and transmitters or transceivers are colocated in mobile platforms or in sites where space is at a premium.

System and equipment parameters are evaluated to minimize the degradation of the receivers performance in cosited radio installations, together with additional technical and operational resources necessary to support this target. Furthermore the parameters necessary for computer simulations of cositing effects and associated frequency management are identified.

1. INTRODUCTION

The military radio communication systems are to be designed and implemented under rigorous criteria of overall protection including:

- information security (COMSEC)
- compliance with Tempest requirement
- ECM protection and associated radio transmission security (TRANSEC)
- robust codec processes
- signal processing
- EMP and INR protection
- EMC/EMI protection from system self-interference

It can be said that of all the above aspects relevant to the protection of the radiocommunication channel, EMC/EMI protection does not seem to be sufficiently investigated in the current technical literature.

The importance of EMC/EMI protection is reflected in the diversified engineering problems associated with the cositing of radio terminals in diversified platforms and sites.

The problems relate to the correct operation of receiving systems operating simultaneously with transmitters (in some cases high-power transmitters) in installations which do not allow high physical isolation between Tx and Rx antennas.

This can result in heavy degradation of the receivers' performance and in some cases this "friendly threat" is more disrupting than the enemy's threat unless proper measures are implemented.

This requires that the cositing problems be identified and deeply investigated.

The cositing situations for military radicommunications systems can be reduced to the following most significant cases

- surface ships fitted with a large number of 1KW HF Transmitters (10 or more transmitters per ship) and HF Receivers plus 10 to 20 UHF Transceivers(225 to 400MHz).
- helicopters tasked with missions that require the use of two simultaneous HF channels.
- vehicular platforms fitted with HF (2-30 MHz) and/or VHF (30-88 MHz) radios. The number of VHF radios can be as high as 8 in the SCRA (Single Channel Radio Access) applications.
- air defense sites with a large number of VHF/UHF Transceivers for Ground to Air communications.

In all the above situations, the cositing problems are aggravated to the limit of practicability whenever ECCM spread spectrum techniques (e.g. frequency hopping, FH) are to be adopted.

The requirement for FH operation involves a deep study of the cositing and frequency management aspects in order to verify the conditions necessary to prevent channel degradation and/or to identify the frequencies at which a certain amount of degradation is acceptable.

2. SYSTEM PARAMETERS AND COSITING FACTOR

When cositing of radio terminals occurs in a communication system, the requirement is that the performance of the receiver component be not degraded with respect to non cositing conditions.

To verify this condition, the two following factors must be assessed:
- environmental noise power density typical of the installation site;
- isolation between the system's antennas.

A fundamental parameter is the minimum frequency distance Δf between the radio terminals necessary to prevent receivers degradation in the system.

The value of Δf, generally indicated as a percent figure, is a limiting factors for the allocation of colocated channels in a given frequency band f2-f1; more specifically the number of such channels is

$$N = \frac{f2-f1}{\Delta f . fm} \cdot 100$$

where : Δf is a percent figure

$$fm = \sqrt{f1 . f2}$$

In addition to the Δf parameter, a new useful parameter, the Cositing Factor, can be considered.

This factor is defined as the ratio between the number No of available channels in a given f2-f1 band and the number of usable channels (i.e. on which co-sited operation is allowed, indicated by (1) above).

If Bch is the bandwidth assigned to a radio channel in the operating frequency band (e.g. 3 kHz in the HF band, 25 kHz in the V/UHF bands), then No = (f2-f1)/Bch

Consequently the cositing factor, CF, is given by

$$CF = \frac{No}{N} = \frac{\Delta f . fm}{100 . Bch}$$

For example, in the 2 to 6 MHz band, No=1333 channels of 3 KHz are available, with Δf =5%, the cositing factor is CF=57,66 and the number N of usable channels is 1333/57,66=23, a value that can be easily derived from equation (1).

3. CAUSES OF DEGRADATION

The causes of degradation in the system performance are of different nature and are based on:
- the characteristics of the radio equipment, namely the power rating and linearity of the transmitters, sensitivity and dynamic range (intercept point) of the receivers;
- the system parameters such as radiator characteristics, isolation between Rx and Tx antennas, operating frequencies and frequency spacings;
- environmental conditions such as natural noise and man-made noise;
- simultaneous transmission and reception.

Besides the non-essential radiations (spurious and harmonics) of the individual transmitters and the image and IF rejection of the receivers, there are four fundamental causes of degradation of a system in cositing conditions.

3.1 Radiated out-of-band noise

Military radiocommunications in the HF, VHF and UHF bands often employ wideband configurations at both system and equipment level.

In such conditions, the level of out-of-band noise may be quite high and spread over the channels assigned to the colocated receivers.

This noise radiated by the Tx antennas, if higher than the environmental noise, produces desensitization of the receivers.

3.2 IMP generated by coupling among transmitters

If fA is the fundamental frequency of a transmitter, the radiated harmonics have frequencies

n.fA where n = 1, 2, 3 .....

In the case of two colocated transmitters operating over two different antennas or over the same antenna, in addition to the harmonics intermodulation products, IMP, having frequencies mfA $\pm$ nfB will be generated.

The frequencies of the intermodulation products up to the fifth order are listed below

2nd order	fA ± fB				
3rd order	2fA ± fB	2fB ± fA			
4th order	2fA ± 2fB	3fA ± fB	3fB ± fA		
5th order	3fA ± 2fB	3fB ± 2fA	4fA ± fB	4fB ± fA	

In applications using wideband transmitters operating into wideband antennas it is necessary to reduce the level of such components down to values that do not degrade the performance of the colocated receivers tuned to any of the above frequencies.

The reduction of harmonics and even-order IM products is an easy task that can be achieved by use of selectivity resources.

The situation is different for the odd-order IM products, in particular when low values of Δf are used between the operating frequencies.

For the case in subject, the products

2fA-fB 2fB-fA
3fA-2fB 3fB-2fA

have a frequency spacing Δf = fA-fB among each other and from the fundamental frequencies fA, fB.

Consequently, the reduction of such products is a problem of increasing difficulty as Δf decreases.

The assessment of the level of such products can also be a problem.

For two transmitters, operating into the same antenna through a multicoupler, the level of intermodulation products caused by mutual coupling can be assessed by use of;
- conventional two-tone test (IMo)
- selectivity characteristics of the multicoupler units

Such level, below the rated power level, is IMP = IMo + 2A where A is the attenuation at the n. $\Delta f(\%)$ point of the selectivity curve of the multicoupler unit located between each transmitter and the wideband antenna.

The factor "n" is 1,2,.... for the IM components respectively of 3rd, 5th,.... order.

For example, if

IMo = -40dB
A = -35dB for Δf = 5%
A = -50dB for Δf = 10%

the level of the IM products will be

-40dB -70dB = -110dB below rated power level for the 3rd order
-40dB -100dB = -140dB below rated power level for the 5th order

The above data are consistent with the test results.

3.3 Receiver desensitization

Receiver desensitization occurs when strong out-of-band interfering signals are applied to the receiver front-end together with the desired signal.

The level of the interfering signals is dependent on
- the isolation between the antenna of the co-located transmitter and the antenna of the receiver;
- the selectivity introduced between the receiving antenna and the receiver;

If Io is the level of interfering signal relativee to the 1dB compression point of the receiver, no sensitivity degradation occurs for interfering signals of level less than Io.

3.4 IMP and cross-modulation generated in the receivers front-end

The strong interfering signals from colocated transmitters generate intermodulation products, IMP, in the receivers front-end due to the front-end non linearity.

These IM products have the same frequencies indicated in paragraph 3.2 for the transmitters IMP and in this case too the problem is to eliminate front-end IMPs located in the receiver channel, thereby capable of producing receiver desensitization.
Again, odd-order IMP are to be considered and in particular 3rd order components.

In case of two interfering signals of equal level Pin (dBw), the level of

intermodulation products is

$$IM(dBw) = 3\ Pin(dBw) - 2\ IP(dBw) \quad (2)$$

where IP is the 3rd order intercept point of the receiver front-end.

If the interfering signals are attenuated by a preselector filter inserted between antenna and receiver

$$IM(dBw) = 3\ Pin\ eq\ (dBw) - 2\ IP(dBw)$$

where

$$Pin\ eq = M - \frac{1}{3}(M-m)$$

$$M = Pin - A(\Delta f)$$
$$m = Pin - A(2\Delta f)$$

being $A(\Delta f)$ the attenuation of the preselector at Δf and $A(2\Delta f)$ at $2\Delta f$.

In addition to the 3rd order intermodulation products, crossmodulation is a different type of interference which eventually results in receiver desensitization.

Crossmodulation, whose effects are most important in radio systems operating in the AM mode, consists in the translation of the modulation sidebands of a strong interfering signal aroung the carrier frequency of the desired signal.

4. RESOURCES FOR COSITING SOLUTION

Resources are available for a solution to the cositing problem at both technical and system engineering level, such as:

- High linearity of the transmitters power amplifiers
- High linearity and high dynamic range of receivers front-end
- Selectivity at both transmitter and receiver sides
- Optimization of isolation between transmitting and receiving antennas
- Exciters with low out-of-band noise power density

From the system point of view, other resources are available such as the use of Interference Cancellation Systems (ICS) particularly suited for platforms where the isolation between antennas is close to 0dB (Helicopters, submarines, ground vehicles).

Another resource to be used, particularly for SCRA type systems, is the enforcement of a discipline for the time dedicated to transmission and reception, i.e. the adoption of a time-frequency division concept on transmission and reception.

In this case, all transceivers are simultaneously on transmission in a given time Δt and simultaneously on reception in the subsequent Δt time interval.

This discipline prevents degradation of the receivers performance; the information must necessarily be of digital type. The problems associated with the time-frequency division architecture are balanced by the fact that Δf can be reduced to a minimum (consequently the cositing factor is equal to 1) and that the system has no degradation even in ECCM-FH operation for all or part of the transceivers.

5. SYSTEM ARCHITECTURE FOUNDED ON SELECTIVITY AND LINEARITY RESOURCES

5.1 System configuration

Many configurations of communication systems operating in cositing conditions can be sketched as shown in fig. 1, relative to a naval system using wideband antennas and multicouplers for both transmission and reception functions.

As an example, the 2-12 MHz band will be considered.

The resources available to such system are:

- on the transmit side
  - . exciter with low out-of-band noise power density
  - . use of high selectivity characteristics post-selectors at low rf level
  - . use of high linearity WB power amplifiers
  - . use of high selectivity characteristics antenna multicoupler at high rf level

- on the receiving side
  - . passive multicouplers
  - . high selectivity pre-selectors
  - . front-ends with high linearity and dynamic range.

5.2 System and equipment parameters

In the evaluation of the system in cositing conditions the following assumptions

are made
- simplex operation
- isolation between transmitting and receiving antennas: -40dB
- minimum Δf between operational frequencies : 5%
- colocation factor, CF
  . in the 2 to 6MHz range: 57,6
  . in the 2 to 12MHz range : 81,6
- channel availability
  . N = 23 in the 2 to 6MHz range
  . N = 41 in the 2 to 6MHz range
- environmental noise

	Noise power density dBw/Hz	Noise power in the 3KHz BW dBw
2 MHz	- 150	- 115
4 MHz	- 160	- 125
12 MHz	- 170	- 135

-Transmitting system
  . RF power level (50 ohm) 1000 W (+30dBw)
  . Two tone test IM (IMo) -40dB
  . Exciter rf power level 100 mW (-10dBw)
  . Exciter out-of-band noise -140 dBw/Hz
  . Post selector and antenna multicoupler selectivity characteristics
- 35dB at Δf = 5%
- 60dB at Δf = 10%
- Receiving system
  . Preselector selectivity characteristics: same as post-selector
  . Front-end 3rd order intercept point (IP) -10dBw

5.3 Radiated out-of-band noise (fig. 2)

The exciter's out-of-band noise is reduced by the post-selector selectivity, amplified by wideband amplifier and subsequently reduced by the multicoupler selectivity

The levels at the output of the above units are given in the table below

	RF power dBw	noise power density (Δf=5%) dBw/Hz
Exciter	-10	-140
Post-selector	-10	-175
P.A.	+30	-135
Antenna MCPR	+30	-170

If a 40dB isolation is assumed between Tx and Rx antennas the noise power density at the receiver antenna is -210dBw/Hz, which is substantially lower than the environmental noise.

Consequently, in the adopted configuration no degradation effects on the receiver performance are present even in the case of simultaneous operation of more than one transmitters.

5.4 Intermodulation products among transmitters (fig. 3)

The level of intermodulation products generated by mutual interference among transmitters is given by

IMP = IMo + 2A = -40dB -70dB = -110dB (3rd order)
= -40dB -120dB = -160dB (5th order)

related to the rated 30dBw power level.

The level of the radiated IMP is -80dBw and -130dBw respectively for 3rd order and 5th order products.

The level at the Rx antenna are -120dBw and -170dBw respectively.

Considering the noise power levels in a 3 KHz bandwidth, it is apparent that:
- the 5th order IMP do not produce any degradation;
- the 3rd order IMP are 5dB below noise in the lowest part of the band, 5dB above at 4 MHz and 15dB above at 12 MHz.

In any case, degradation is introduced only if the receiver is tuned to the

frequency of the IM product.

Frequency management can be used to avoid such occurrence; shift to the adjacent channel is sufficient to prevent any degradation.

5.5 Harmonics

If the transmitter operating into 50 ohm features a 60dB harmonic attenuation due to the selectivity of the antenna multicoupler, the radiated harmonics have a -120dB level with respect to 30dBw, i.e. -90dBw.

At the receiving antenna the level is -130dBw consequently, in the band of interest the situation is comparable with that already examined for the IM products.

5.6 Receivers Desensitization (fig. 4)

The RF power radiated by the transmitting antenna is attenuated by 40dB down to -10dBw at receiving antenna.

The preselector selectivity reduces the level of the interfering signal at the receiver front-end down to -45dBw.

The receiver sensitivity is not degraded if the receiver 1dB compression point is better than -40dBw.

5.7 Receiver front-end IMP (fig. 5)

If two transmitters operate at frequencies fA, fB such that the 3rd order IM product has the same frequency of the desired signal, the receiver is desensitized whenever the level of such product exceeds the noise power level in a 3 KHz bandwidth.

As seen before, the level of the interfering signal at the receiver antenna is -10dBw.

If fA is spaced by Δf =5% from the desired frequency and fB by Δf = 10%, bearing in mind that

M = Pin -35dB = -10dBw -35dB = -45dBw
m = Pin -60dB = -10dBw -60dB = -70dBw

then IM = -140dBw.

This intermodulation level is lower than the noise power level in 3 KHz bandwidth consequently no degradation occurs.

5.8 Comments

The study of the potential causes for performance degradation in the receivers colocated with transmitters, even though limited to the HF band, is in general applicable to the other frequency bands.

However, the following comments are in order.

As it concerns the out-of-band noise, as shown in 5.3, its level has been reduced to negligible values for Δf = 5%; it is safe to assume that normal operation is possible with Δf = 2,5%, halving the cositing factor.

It is to be borne in mind however, that such values have been obtained by introducing the post-selector selectivity at low RF level and the antenna multicoupler selectivity at high RF level.

The IM products are the most critical aspect of co-siting since they are generated by the mutual interference between transmitters and non-linearity of the receivers front-end.

Another source of IMP can be identified in the metal structures of the platform which contain zones and elements that behave as non linear devices for the impinging RF signals producing and radiating therefore harmonics and IM products.

However, the effect of the selectivity and linearity resources used for IMP reduction has been clearly assessed.

Luckily, the identification of the radio channels victim of IM products is an easy task and consequently frequency management measures are not difficult to implement.

6. IMPACT OF ECCM OPERATION WITH THE COSITING OF RADIO TERMINALS

The number of usable channels, i.e. non degraded channels, is drastically reduced in cositing conditions by a factor which has been indicated as the "cositing factor".

This situation is difficult to handle also in operation at fixed frequency and in

frequency bands where the choice of frequencies is constrained by other factors such as propagation conditions, frequency assignment by international regulatory agencies, use of the channel by other parties.

Obviously, the difficulties increase as the number of operating channels approaches that of the available channels.

When ECCM techniques (such as frequency hopping) must be used in systems under cositing conditions, all the above problems are magnified and in some frequency ranges, a limit situation may be reached for the system feasibility.

On the other side, cositing may be a conditioning factor in the selection of the ECCM techniques or parameters, when protected ECCM radio channels must coexist with fixed-frequency radio channels, a situation typical of naval platforms in the HF and UHF band.

The ECCM-FH parameter which most suffers from the constraints set by the cositing is the segregation of the hopping set in a limited band (e.g. 10-15% of central frequency) protected by the selectivity of an "ad-hoc" multicoupler unit.

Another ECCM area influenced by cositing is that of the hopping sequences orthogonal in the time/frequency domains, even though frequency management can be implemented to achieve coexistence among the different ECCM nets and the fixed frequency channels.

7. IDENTIFICATION OF COSITING PARAMETERS FOR COMPUTER SIMULATION

7.1 General

In many istances, the cositing situations can be very effectively handled by means of computer simulations, particularly in cases where the cositing factor is much greater than 1 and the number of operating radio channels is very close to that of the available channels : N = No/CF

The computer simulation objective is the assignment to all users, located on the same platform, of fixed frequency values (or sets of frequencies for FH nets) selected among a set of available frequencies so that the levels of the interfering signals do not exceed some specified values.

In the previous paragraphs these values and the target characteristics of the transmitting and receiving equipment in a cositing environment have been assessed by using simple mathematical calculations.

The impact of these parameters on the choice of the assignable frequencies will be briefly analysed for the various fields of this specific type of modelling.

7.2 Models description for different objectives

The overall production of the frequency management software can be divided into three main different types, as it will be shown in the following:

a) frequency assignment :
   it consists of the optimum management of the overall available frequency band among different users operating on the same platform.
   It is a typical operative research problem in which the resources are the frequencies the users are the nets and the constraints that drive the assignments are constituted by the minimization of the probability of interference between the nets.
   In order to optimize the assignments the order of the requests, that is the order in which the various nets enter the research tree, is very important. For this purpose a sort of "incompatibility matrix" is constructed, in which in the i-j element a coefficient that qualifies the compatibility between the ith and jth request is present.
   The minimum value is attributed if the two equipment i and j are colocated in the same platform whereas the maximum corresponds to a situation in which the distance between the two nets is larger than the frequency reuse distance.
   Through a sort of processing of these coefficients it is possible to determine the order of frequency assignment among the N nets.
   In the general case in which both fixed-frequency and FH operations coexist, the fixed frequency assignment is performed first and then the hopsets for the FH nets are allocated, taking into account the lockout frequencies utilised by the fixed-frequency equipment.

b) up-dating :
   this procedure is performed in order to assign individual frequencies to further requests when a planned situation is already operational for the previous users.
   It also provides a verification on the impact produced on an already optimized system by mandatory changes of frequencies.
   This is only a particular case of the general procedure of the above frequency assignment and it is performed with the same criteria, with the exception that no construction of the incompatibility matrix is necessary since no priority is necessary.

c) forecasting
This is an aspect quite different from the previous two and it represents a useful tool when an operational scenario has to be analysed in order to set suitability for a class of equipment.
In this case, the frequency management is only the first phase of the algorithm, whereas in the second phase the level of interference for a certain radio link is evaluated assuming the worst-case conditions if the activity factor for the other radios is a random variable.

7.3 Input/output characterization

The algorithms can be defined through their lists of input and output parameters and through the characterization of the constraints that determine the assignment or non-assignment of a frequency (or an hop-set) to each net.

a) Input data :
The input data can be subdivided into technical data and operational data : the first related to the type of equipment and the second,to the type of assignment. As it concerns the technical data, they have been discussed in para. 5 and they could be also considered embedded program parameters, depending on how the computer model is addressed to the specific problem.

A list of such parameters includes:
- transmitter power
- antenna gain
- linearity characteristics of the transmitting equipment
- sensitivity
- selectivity
- coupling factors
- out-of-band noise
- simplex or duplex operation
- receiving system parameters
- fixed-frequency or FH operation
- location of the equipment

As it concerns the operational data, they are generally related to the utilization of the link and typically, they specify, if it is permanent, temporary or preassigned and what is the set of the allowed frequencies for each request.

b) Output data :
They are different for the three kinds of algorithms.
For the frequency assignment the output consists on a table that establishes a correspondence between each request and the usable frequency or hop-set and provides some information related to any relaxation of the imposed constraints.
As it concerns the up-dating program, the output consists on the list of all the available frequencies associated with a quality coefficient for any additional request or all the alternative frequency values corresposnding to all the desired changes.
For the forecasting program, the output consists on an evaluation of same performance parameters that depend on what has to be optimized.
These parameters, typically, consists of:
- Anti-interference margin against the number of users with and without frequency management
- BER against the communication range
- Processing Gain required against the number of users

c) Constraint parameters :
As said before, the relationship between input and output data is defined by a certain number of constraints that drive the assignment and that consist of:
- intermodulation products : for each request the possibility to generate IMPs with any other already assigned frequency is evaluated, and the amount of the interfering power (if any) is evaluated and compared with the target values (see para 5.4)
- harmonics : for each request the effect of the harmonics on the other nets is calculated and compared with the maximum allowed values (para. 5.5)
- minimum frequency separation : it is an input datum if the software is not dedicated and it depends on the used channelization, on the front-end characteristics of the receivers and on the out-of-band noise of the transmitters. For each request any possibile value of assignable frequencies has to maintain the minimum frequency distance from the other nets. In the frequency assignment algorithm this evaluation is performed during the construction of the compatibility matrix.
- Re-use criteria : it is present when the platform size is large enough to permit the use of the same frequency by two different nets.

In order to calculate the interfering power levels relative to any receiver, a propagation model is adopted that takes into account the environmental noise, the type of terrain and the antenna heights.

In the system operating in AM mode, moreover, a particular importance has to be given to the effects of cross-modulation.

8. CONCLUSION

The causes of radio communication channels degradation have been considered and numerical examples shown to allow us to choose appropriate resources to mitigate the cositing effects.

The identification of cositing parameters is prerequisite to the solution of the cositing and frequency management problems achieved by the use of a versatile and rigorous tool capable of handling complex and diversified situations and producing highly reliable data.

REFERENCES

1. CCIR Report 322 "World distribution and characteristics of atmospheric radio noise"

2. CCIR Recomm. 3264 "Power of radio transmitters"

3. G.C. Salisbury "Topside intermodulation interfeence aboard USS Mount Whitney, USS Blue Ridge, USS Iwo Jima, NEL Doc 206, Dec.1972

4. R.H. Bauman, C.L. Golliday, R.K. Royce, D.C. Andres, C.E. Hobbis "Adaptive cancellation of local electromagnetic interference in HF Naval Communication Systems", July 1978

5. A. Vanwelsenaers, R. Hennecart, M. Erculisse, BAMS (BE) "Frequency management for frequency hopping VHF Combat Radio Nets in on army corps" - 5th ARFA Symposium

6. "Spectrum Management and computer - Aided techniques" - CCIR 1983

7. Mathur K., Salkin M., Nishimura K. Moritos "The design of an interactive computer software system for the frequency assignment problem" IEEE transaction on EMC, vol. 26 Nr. 4 Nov. 1984

8. Trollope, L.T., Lauder W.E. "A new HF Communications Architecture for ships" IEE Conference Publication 206 - HF Communications systems and Techniques.

9. G. Cucinotta, "Sistemi di comunicazione e interferenze", Steti - Roma, Italy, 1981

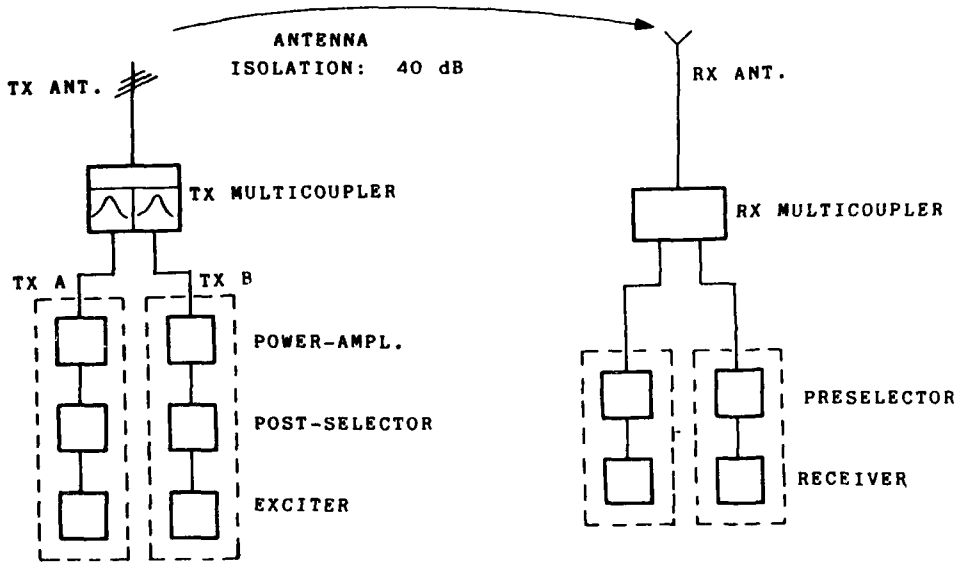


FIG. 1 - COMMUNICATION SYSTEM FOUNDED ON LINEARITY AND SELECTIVITY RESOURCES

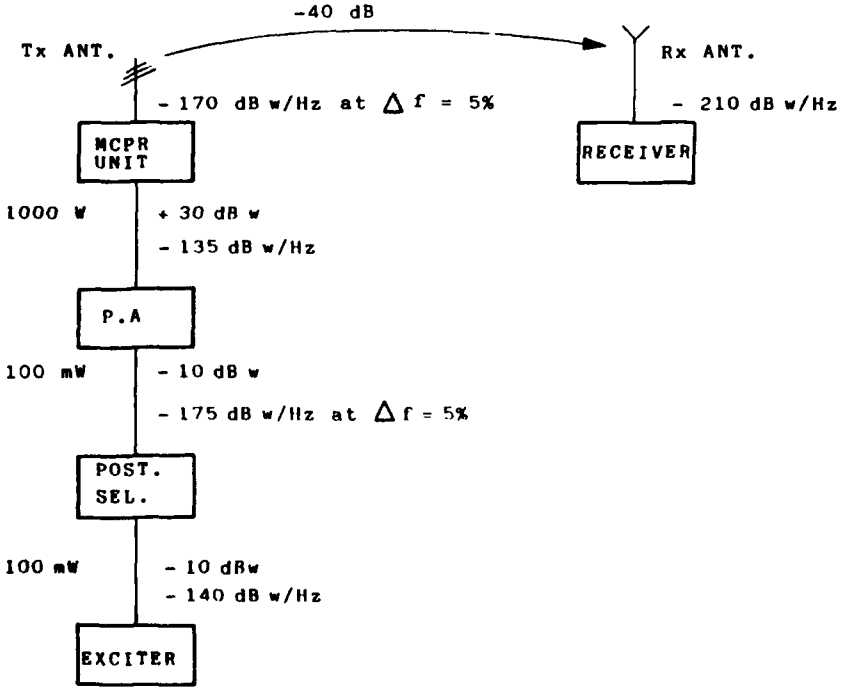
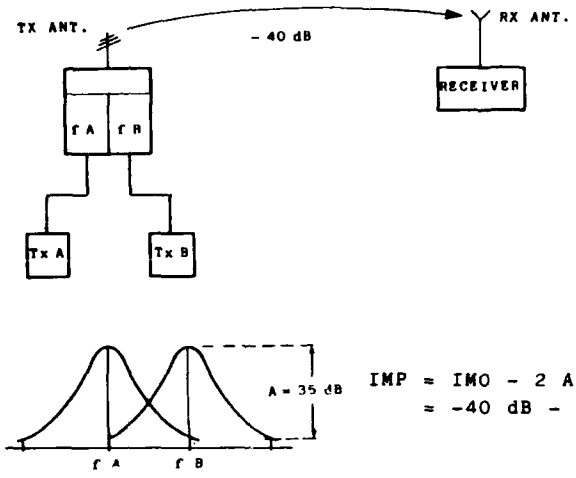


Fig. 2 - OUT-OF-BAND NOISE AT Δ f = 5%



IMP = IMO - 2 A =
= -40 dB - 70 dB = -110 dB

Fig. 3 - INTERMODULATION PRODUCTS AMONG TRANSMITTERS

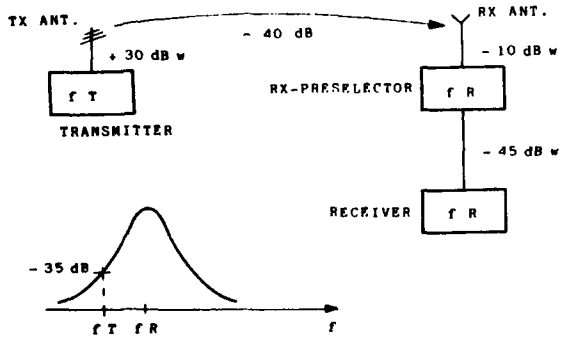


Fig. 4 - RECEIVER DESENSITIZATION

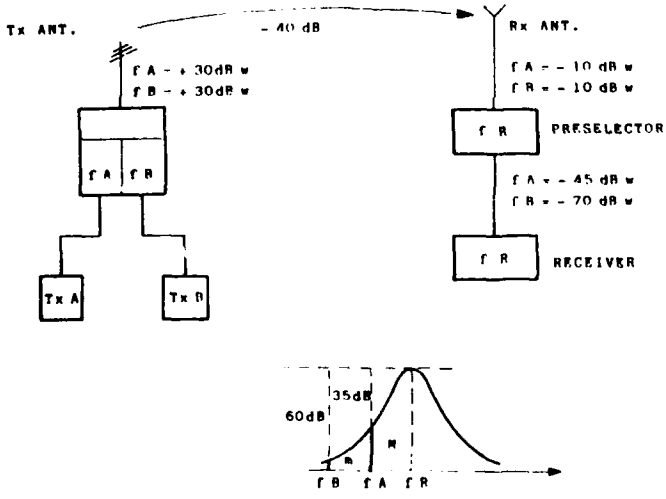


Fig. 5 - RECEIVER FRONT-END IMP

INTERMODULATION INTERFERENCES IN RADIO SYSTEMS

Mr P.S.W.Ho[1], Dr L.H.Bevan[2], Prof A.C.C.Tseung[2] and Mr W.S.Wilkinson[1]
Electrical and Electronic Engineering Department[1]/Chemical Energy Research Centre[2]
City University, Northampton Square, London EC1V 0HB, U.K.

Summary: The phenomenon of intermodulation is introduced first. The origins of intermodulation interference in a radio communication systems are considered. Mechanisms responsible for this interference, for example, ferromagnetic nonlinearity and electron tunnelling are considered, and a model based on a metal-semiconductor rectifying junction is suggested. An experimental setup for measuring intermodulation levels of specially treated test samples and the operating procedures are then described. A general principle to overcome the structural intermodulation problem is outlined. Experimental results for some chemical and metallic coatings ,and their merits are presented next. Lastly, the usefulness of monitoring the background intermodulation level of a test site is given. In conclusion, the joint effect is the first order factor in determining the degree of seriousness of the intermodulation problem and the use of high conductivity coating provides the most feasible solution.

1 Introduction

When two or more radio frequency signals f_1, f_2, ... f_n are mixed in a device with non-linear transfer characteristic, spurious frequency components, f_{im}, will be generated and are called Intermodulation Products (IMP).

f_{im} are given by :

$$f_{im} = a_1 f_1 + a_2 f_2 + \ldots\ldots a_n f_n$$

where a_1, a_2,... a_n are either zero or which may be positive or negative integers and ($a_1 + a_2 + \ldots a_n$) is the order of the intermodulation product. In general, the amplitudes of lower order IMPs are larger than the higher order IMPs. However, experimental results [1] have shown that products upto eleventh order or even higher can have adverse effects on the performance of communication systems.

2 Background for investigations

The problems of intermodulation interference in radio communication services are getting more and more serious because of the increasing demands on the available electromagnetic spectrum. As the number of channels increase, the number of IMP increases at a prodigious rate. Intermodulation frequency free planning has been an important and difficult task for communication service planners. Extensive work [2 - 4] have been done on deriving algorithms to calculate the amplitudes and frequency components of the intermodulation products for a given set of input frequencies. The time taken to carry out these calculations are very long for planning a large communication service, such as a cellular mobile radio communication network. The applications of high speed computers or parallel processing computers could reduce the time requirement.

Basically, intermodulation interference (IMI) can be divided into two types. First, those due to active devices in the communication systems, such as the non-linearity of the power amplifiers in the transmitters where common antennas are employed, and overloading of the receiver front-end. The second type is due to the passive components in the communication systems and these include cables, feeders, aerials and the supporting structures.

Effective methods have been found to solve the problems of the active type by careful shielding of equipments, using filters in the transmitter outputs and receiver inputs. However, the passive type is a much more intractable problem, and once generated, cannot be removed by filtering.

3 Causes of Passive Intermodulation Interferences

There are two major causes for the generation of passive intermodulation products, namely non-linear junction effect and B/H non-linearity which is inherent in any ferromagnetic material.

Investigations into the ferromagnetic nonlinearity effects [5,6] have identified that radio frequency connectors with nickel-plating can cause high levels of IMI and are not recommended for use in systems carrying large radio frequency currents and susceptible to intermodulation interference. Betts and Ebenezer [7] found that steel supporting structures for aerials were also a source of problems. However, ferromagnetic non-linearity of a bulk material is only a second order effect. The prime factor is the non-linear junction effect.

The non-linear current/voltage characteristics of a rusty joint (metal-oxide-metal junction) has been assumed by many researchers to be due to electrons tunnelling through the oxide layer. Tunnelling phenomena has been a hot topic among solid state physicists [8-10]. Higa [11] tried to explain the spurious signals generated on large reflector antennas by tunnelling of electrons through aluminium-aluminium oxide-aluminium junctions. Though Guenzer [12] from Naval Research Laboratory, Washington, expressed doubts on Higa's arguments. Woody et al [13] carried out a research program on Metal-Insulator-Metal (MIM) Junctions as a Surface Source of Intermodulation. They found that researchers could not agree to the correct form of the tunnelling equation and more important, the simple electron tunnelling model would not be able to predict IMP levels in real life situations due to their inherent complexities. They adapted an experimental approach and developed an empirical model based on measurements of representative MIM junctions. Their findings also confirmed that MIM junctions can generate IMPs that are much higher that IMPs generated by coaxial cables and connectors. Also they found that many junctions have very unstable IMPs which were neither predictable nor repeatable.

According to quantum theory, the probability of an electron tunnelling through a barrier is inversely proportional to the exponential of the thickness of the insulating layer. An insulating layer of thickness greater than 100 Å is impenetrable for electrons. We are not saying that electron tunnelling cannot be a candidate. For example, aluminium oxide has a film thickness of around 20 Å and tunnelling could be responsible for the observed nonlinear current/voltage characteristic of aluminium-aluminium oxide-aluminium junction.

Structural steel forms various iron oxides when exposed to the atmosphere. The thickness of these oxides are significantly larger than 100 Å. We suspect that in this case the nonlinear current/voltage characteristic might be due to junction effect instead of tunnelling; i.e. rectifying junction effect due to metal-semiconductor contacts. It has been shown that conductivity of transition metal oxides can be very high [14]. For example, at room temperature, the conductivity of FeO is about 20 Sm^{-1}, Fe_3O_4 is nearly metallic, though Fe_2O_3 has a conductivity of order only around 10^{-3} Sm^{-1}. High conductivity is dominated by nonstochiometry of the lattice structure. Conduction in undoped material will have contributions from both donors and acceptors. For cuprous oxide, the conductivity(σ) even has an oxygen pressure(P) dependence of $\sigma \sim P^n$, where n assumes values from 1/7 to 1/8.

4 Experiments

4.1 General

Various techniques have been proposed to overcome the intermodulation problem. They all employed the principle of providing an alternative radio frequency path to the rusty junction. There were particular interests in the chemical approach [15], using high dielectric materials. Their argument is based on modelling the non-linear junction as two back-to-back diodes. Putting high dielectric coating on the joint provides a shunt capacitance across it and at high frequency, this capacitance bypasses most of the induced radio frequency current. However, their results have shown that no significant improvements had been achieved. This might be due to the relatively small ratio of dielectric constants between the coating and air, hence there is always incoming wave transmitted into the corroded metal. There could also be adhesion problem of the coatings as well.

4.2 Measurement techniques and experimental procedure

The block diagram of the test setup is shown in figure(1). The two fundamental frequencies are generated by two synthesized signal generators. Power amplifiers are used to give an output of around 50 Watts for each signal. Harmonics generated in the amplifiers are removed by the high-Q cavity resonant filters connected to the outputs of the amplifiers. These signals are then fed into the combining unit by half-wavelength coaxial cables in order to avoid line resonance effects. Filter unit 3 is tuned to the IMP frequency ($2f_2-f_1$) which rejects any signals at the fundamental frequencies. A dynamic range down to -110 dBm can be measured by the spectrum analyser. The fundamental signals are terminated by 100 metres of RG214 cable which acts a linear dummy load. The intermodulation signal is terminated equally by this linear load and the spectrum analyser. With this setup, a signal to intermodulation level ratio of better than 150 dB can be measured.

Before any experiment is carried out, the test setup has to be calibrated. The filters are tuned to their respective resonant frequency and the power amplifier outputs are fixed at 50 Watts. With a clean cylindrical copper sample in the test cell, the background level of the IMP frequency was measured.

Repeatability is tested by measuring the IMP levels of the same sample for several times. We found that variations in results were normally less than 10 dB. This might be due to change in temperature, microstructures of the contact surfaces etc. However, by doing sufficient number of measurements with each sample, a statistical mean level should be achievable. Samples could then be compared with each other for IMP performance using these mean IMP levels.

5 Results

5.1 Experiment Conditions

All measurements were made under the following conditions :

f_1 = 155.2125 MHz, f_2 = 152.0875 MHz, f_{im} = 148.9625 MHz.

Power level of each input signal is 50 Watts (approximately +47 dBm).
The spectrum analyser reading error is +/- 1 dBm.

Samples were corroded by the electrolysis of artificial sea water, with an anodic potential of around 500 mV (versus hydrogen electrode) for 24 hours, unless stated otherwise.

5.2 Experimental results

Initially, the work at City University was to evaluate the use of a graphite-based coating, since graphite would not form an oxide film when it is corroded. Unfortunately the conductivity of the coating is not good enough to stop radio frequency currents from penetrating into the bulk of steel structures. Results are shown in figure(3).

After suspending the work on the graphite-based coating, we turned our attention to various metal platings. An approach to prevent intermodulation caused by the ferromagnetic nature of steel was to galvanise steel, for initial protection, then plate a layer of a more noble metal such as silver (or copper, tin, etc) on top of the zinc layer. The reason for the top layer was because the corrosion product of zinc could cause some intermodulation product while those of silver did not (or to a lesser extent because they are more conducting).

From figure(4), it can be seen that double-layer plated samples all had an initial intermodulation level of above 90 dBm, i.e. around 10 dBm lower than clean steel sample, and their intermodulation values after corrosion were also around 10 dBm better than corroded steel sample. It appears that the better intermodulation values after corrosion for the double-layer plated samples were due to their better initial intermodulation values, and the surface corrosion products of copper, zinc, silver and iron all produced similar amount of deterioration in intermodulation performance. However, the surface corrosion product could be eliminated by adding a layer of coating which does not form any solid oxides on the surface of the original high conductivity coatings (see figure(5)).

In order to make comparison with the joint effect, joint samples (see figure(2)) were made and were corroded for intermodulation measurements. By doing this, we were able to convince ourselves that the rusty joint is a first order effect in intermodulation generations (see figure(6)), although the ferromagnetic effects of steel are still important. However, according to the British Standard BS 5493 (protective coating of iron and steel structures against corrosions) and BS 729 (hot-dipping galvanised coating on iron and steel articles), it is required that for steel structures of thickness greater than 5 mm, the zinc coating must have a thickness not less than 86 microns, which is seventeen times the skin-depth for copper at the frequency that we are using.

For a new antenna tower, the structural elements are galvanised. However, the joints between structure elements could lose their initial protection during construction of the tower, and suffer more rapid deterioration giving rise to corrosion due to fretting at the joint. Moreover, a joint will tend to be a moisture trap. Corrosion at these joints can probably be assumed to be the major source of intermodulation.

Our approach to solve the intermodulation problem is to offer an alternative high conductivity path to r.f. current (see figure(7)). An initial experiment involved the covering of a corroded cylindrical steel sample with zinc coating (by electroplating) was done, and the intermodulation level of the sample was very close to the system floor level (see figure(8)).

The problem now is to develop a practical way to carry out the metallic coating of the joints. The methods being considered are metal spraying and electroless plating.

It was also found that metal-based paints would not work because the bulk conductivity of the paint is much less than the metal. An experiment was carried out using silver-based paint with volume resistivity equal to 10^{-5} m. In this case, the intermodulation level dropped from -50 dBm to -65 dBm.

Although the absolute value of -65 dBm is still high, but an improvement of 15 dBm is a convincing argument for intermodulation reduction by covering the rusty joint with a continuous layer of high conductivity material, i.e. metal.

6 Data Acquisition

It is necessary to subject the proposed coating under field trial before a definite judgement of the viability of the coating can be made. Provision has been made to allow remote data logging and data transfer from the remote test site back to the university. Moreover, the data logging system will be used to gather information about the effect of weather conditions on the variations of the background intermodulation level. This is particular useful as it will provide us with an intermodulation signature of the test site. It might even be possible to further develop the techniques and associated statistical software so that the hardware/software system can be transported to any operational site to carry out a site-intermodulation signature analysis, and the feasibility of putting up more services in that site can then be determined.

7 Conclusions

This paper describes the research programme aiming to provide a practical engineering solution of suppressing structural intermodulation interferences. The results will be useful not just to mobile radio communication services, but to satellite communication or any radio installation where intermodulation effects are prevalent due to a multiplicity of systems. A brief survey was done on the mechanism of structural intermodulation. We suggested a more probable mechanism based on metal-semiconductor rectifying effect. Further work will be done to prove it.

8 References

[1] Fudge,R.E., "The re-engineering of the VHF mobile radio services in the UK", PhD thesis, April 1984, The City University.

[2] Debney,C.W., "The measurement and prediction of intermodulation distortion in co-site radio communication systems", PhD thesis, 1983, Southampton University.

[3] Babcock,W.C., "Intermodulation interference in radio systems", The Bell System Tech. Journal, January 1953, pp 63-73.

[4] Gretsch,W.R., "The spectrum of intermodulation generated in a semiconductor diode junction", Proc. IEEE, vol.54, No.11, November 1966, pp 1528-1535.

[5] Young,C.E., "The danger of intermodulation generation by RF connector hardware containing ferromagnetic materials", Proc. 9th Annual Connector Symposium, October 1976.

[6] Bailey,G.C. and Ehrich,A.C., "A study of RF nonlinearities in nickel", Journal of Applied Physics 50, 453 (1979).

[7] Betts,J.A. and Ebenezer,D.R., "Intermodulation interference in mobile multiple transmission communication systems operating at HF (3-30 MHz)", Proceeding of IEE, Vol.120, No.11, November 1973.

[8] Stratton,R., "Volt-current characteristics for tunnelling through insulating films", J.Phy. Chem. Solids, Vol.23, pp 1177-1190, Pergamon Press 1962.

[9] Simmons,J.G., "Generalised formula for the electric tunnelling effect between similar electrodes separated by a thin insulating film", J.of Applied Phys., Vol.34, No.6, June 1963, pp 1793-1803.

[10] Fisher,J.C. and Giaever,I., "Tunnelling through thin insulating layers", J.of Applied Phys., Vol.32, No.6, June 1961, pp 172-178.

[11] Higa,W.H., "Spurious signals generated by electron tunnelling on large reflector antennas", Proc IEEE, Vol.63, No.2, February 1975, pp 306-313.

[12] Guenzer,C.S., comments on ref.[11], IEEE Proc Letters, February 1976, pp283.

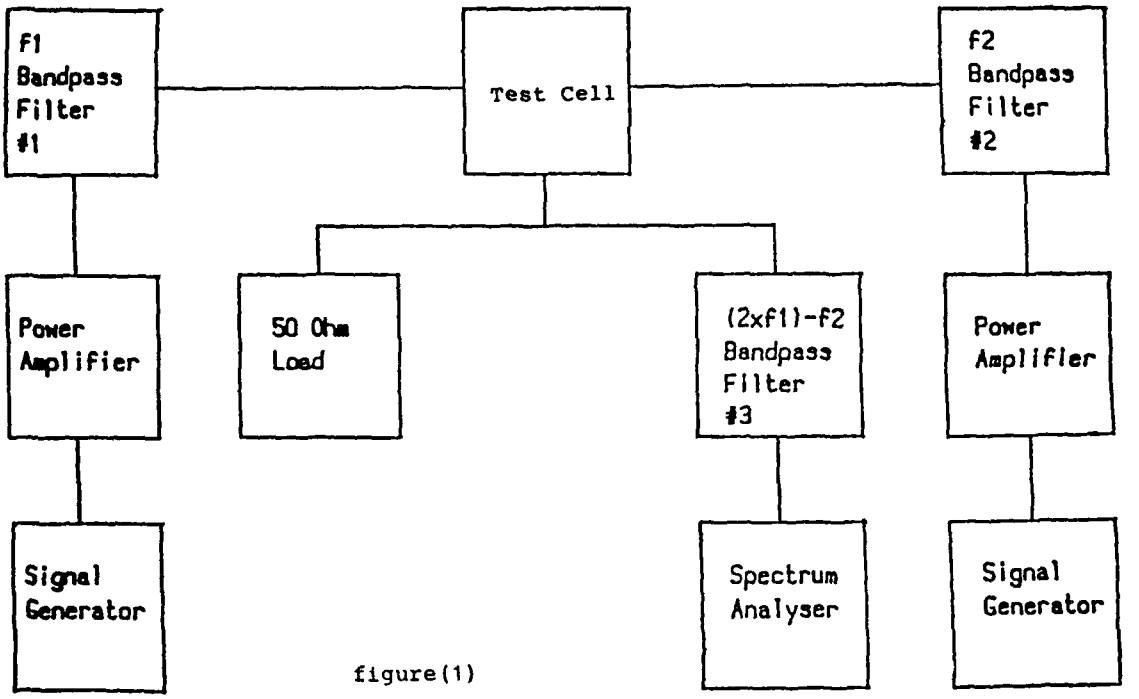
[13] Woody et al, "Metal-insulator-metal junctions as surface sources of intermodulation", February 1983, RADC-TR-83-31.

[14] Seitz,F. (ed), " Solid state physics", Vol.21, 1968, Academic Press.

[15] Copper,J.C. et al, "Chemically suppressing rusty-bolt intermodulation interference", Proc. IEEE National Symp. on EMC, April 1984, pp 233-240.

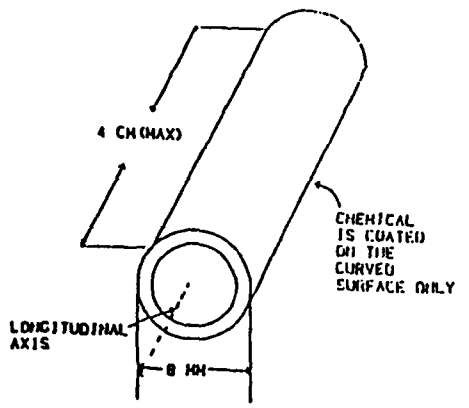
9 Acknowledgement

This study is supported jointly by Science & Engineering Research Conucil and the Home Office, UK. The authors are indebted to Mr P.M.Tomlinson and Mr P.Buley of the Directorate of Telecommunications, Home Office; also Dr R.E.Fudge, formely of the Directorate of Telecommunications, Home Office, now with British Telecom and Professor A.C.Davies of City University.

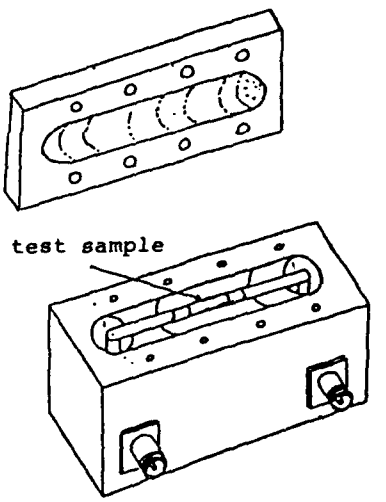


figure(1)

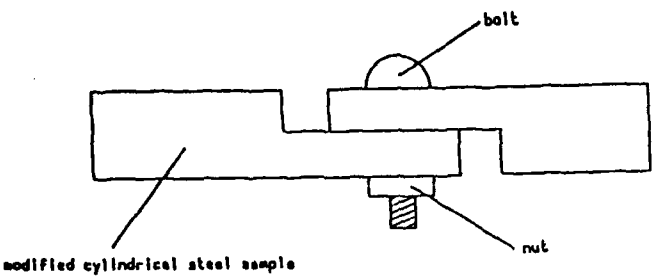
Block Diagram of Test Setup



cylindrical sample



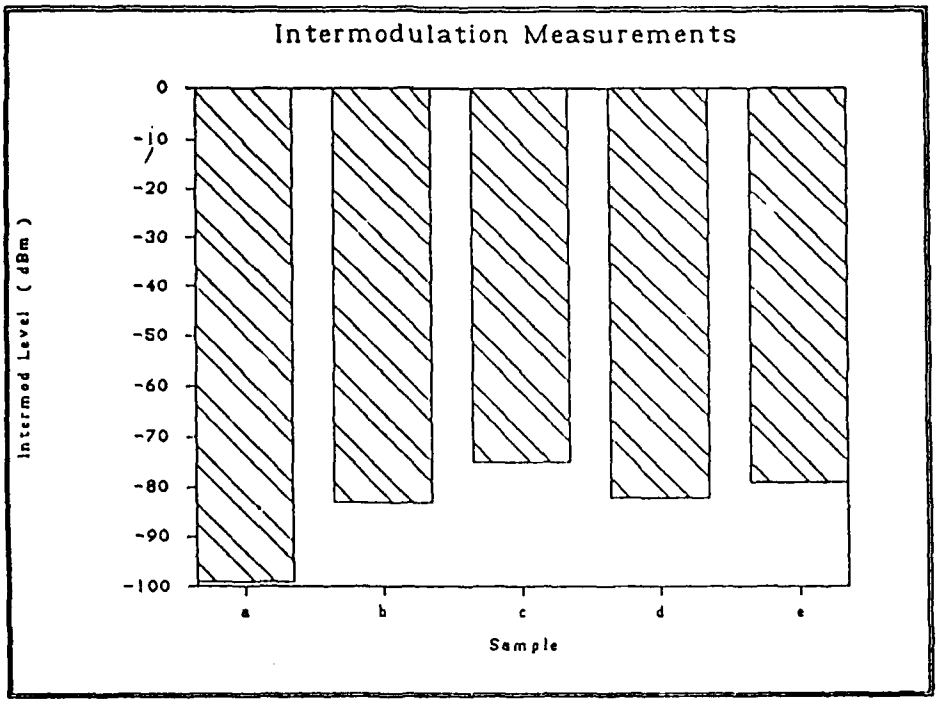
Test Cell



A Joint Sample

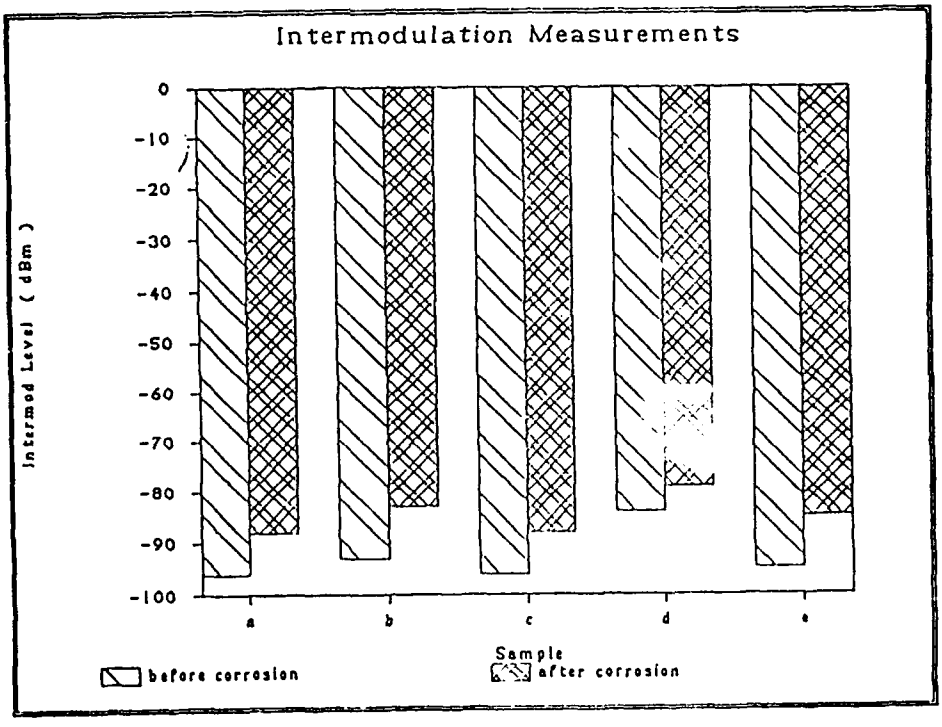
figure(2): Test Cell and Test Samples

figure(3)



a = background level
b = clean cylindrical steel sample
c = corroded sample of (b)
d = steel sample coated with graphite-based chemical
e = corroded steel sample coated with graphite-based chemical

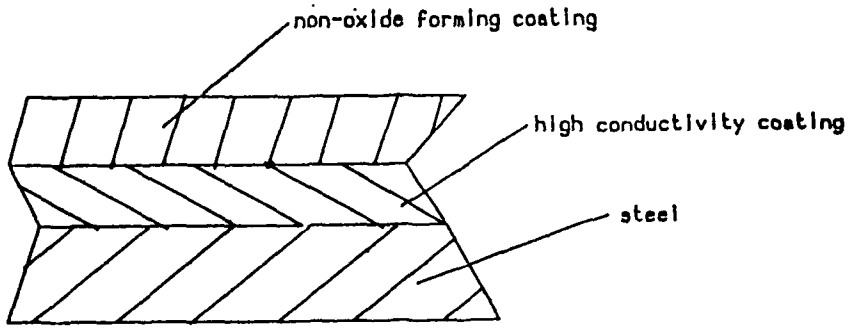
figure(4)



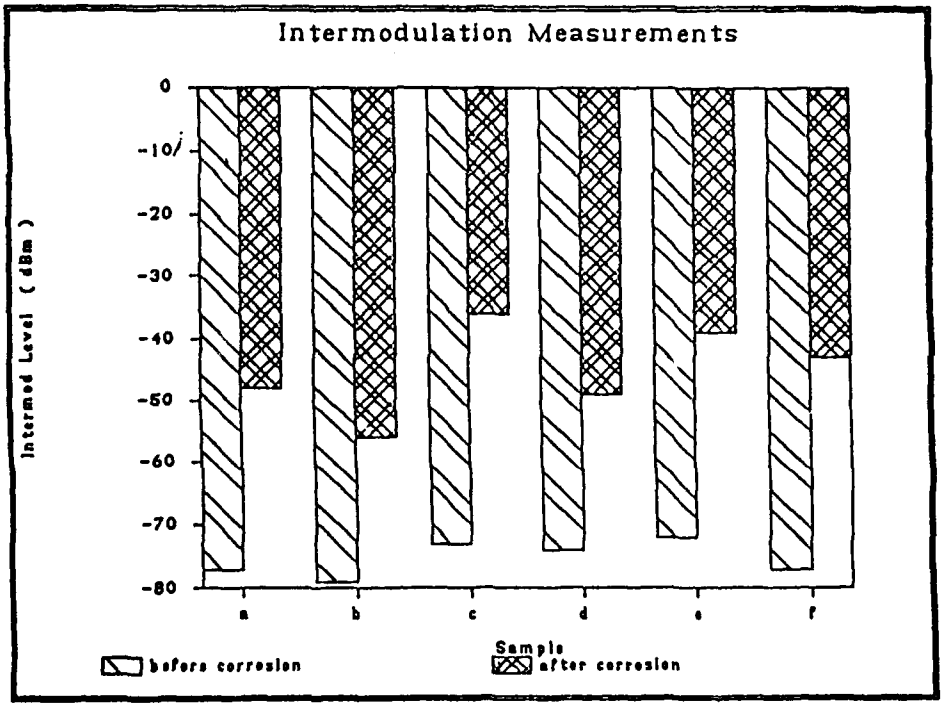
Plated cylindrical steel samples were used in this experiment

Samples (a) and (b) were corroded with concentrated hydrogen-sulphide gas.
Samples (c),(d) and (e) were corroded by the electrolysis of artificial sea water with an anodic potential of around 500 mV (versus standard hydrogen electrode) for 24 hours.

a = copper/zinc/steel
b = silver/zinc/steel
c = zinc/steel
d = silver/copper/zinc/steel
e = silver/steel

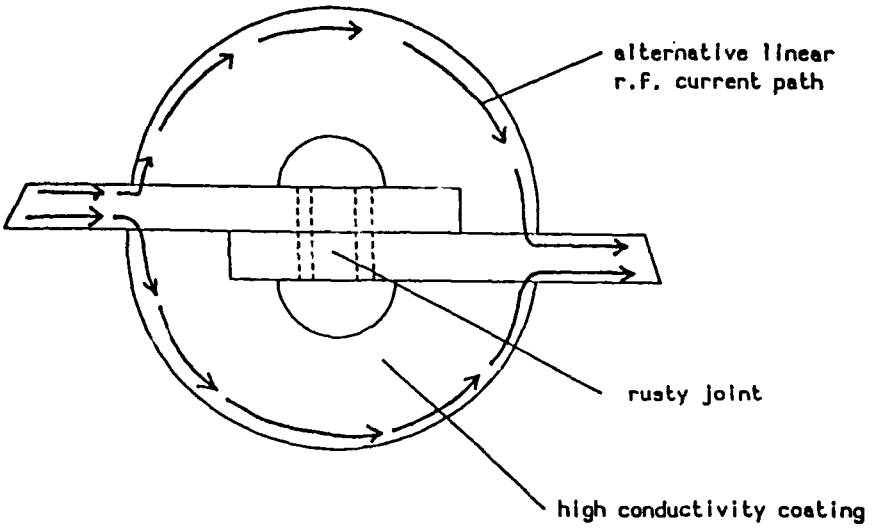


figure(5): double layer coating

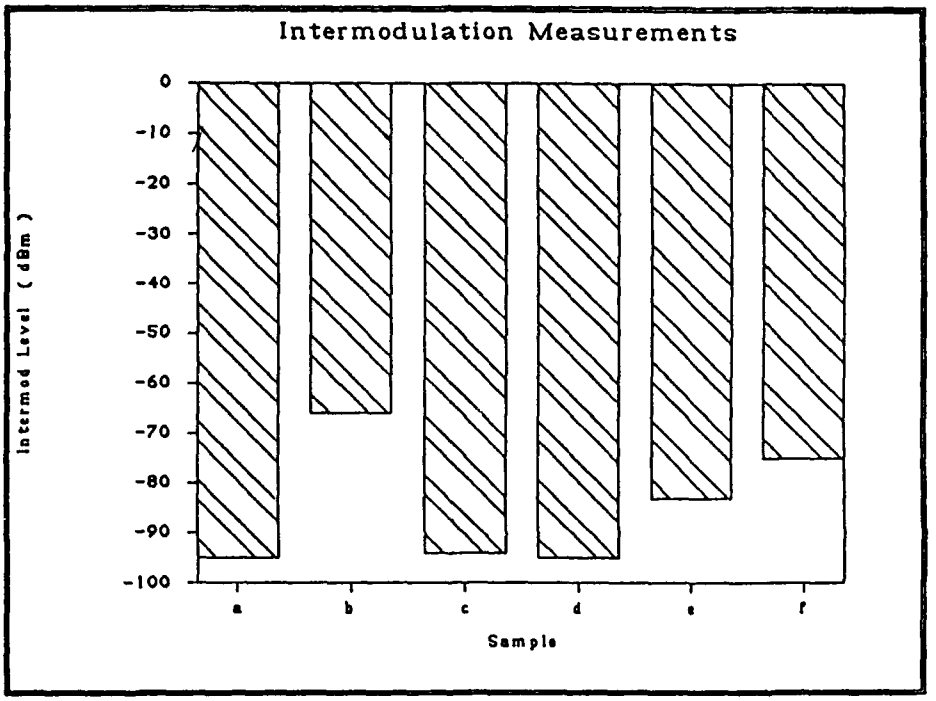


Samples used in this experiment were plain steel joints, adapted from the cylindrical samples used in previous experiments. Six clean samples were used in this experiment. Their intermodulation levels before and after corrosion were measured.

figure(6)



figure(7): Alternative r.f. current path



a = copper rod
b = corroded steel (both ends and the surface of the rod were corroded)
c = sample (b) was zinc-electroplated to a thickness of about 1 mm
d = sample (c) was re-corroded
e = clean steel sample
f = corroded steel sample (just the body, not the ends)

figure(8)

AUTOMATING THE DECONFLICTION OF JAMMING AND SPECTRUM MANAGEMENT

Samuel M. Segner
Center for Command, Control and Communications Systems
Fort Monmouth, New Jersey 07703
USA

Summary.

Powerful airborne and ground based jammers are being fielded by all services and nations as part of their intelligence/EW combat capability. For their survivability, these I/EW systems operate far from the FLOT; this creates rather large denial areas to friendly forces when they jam. Manual coordination between I/EW managers and spectrum managers is not practical to take on targets of opportunities or track the intended enemy victims when these victims counter by frequency maneuvers. Two possible architectures, one centralized, the other decentralized, are explored as is the applicability of the EMC software developed for the US Army ATFES pilot program. The proposed approach is to apply the principles of the Joint Restricted Frequency List (JRFL) and coordination via an Army derivative of the Joint Commanders EW Staff (JCEWS). The initial simplified software to demonstrate the computer aided coordination at VHF is explained.

Background (Fig 1).

This paper is a continuation of the previous paper at this meeting, Ref 1, where the overall US Army program in the Automated Battlefield Spectrum Management and Engineering (ABSME) and its test bed, the Army Automated Tactical Frequency Engineering System (ATFES) are explained.

Figure 2 outlines how the competing goals of the spectrum management community and the intelligence/electronic warfare create the need for EW deconfliction or EW fracticide resolution. Fortunately, the tool is the normal electromagnetic compatibility (EMC) analysis.

Overuse of spectrum by friendly forces limits the electronic surveillance capability of the friendly force intelligence units while activation of friendly force jammers and/or decoys degrade the performance of friendly force communications and electronic equipment.

In static situations, and when there is plenty of time to preplan the locations of possible victims from preplanned jammer or electronic surveillance sites, there is no problem demonstrating good deconfliction capability. One can make use of the types of EMC analyses available on personal computers or more sophisticated computers (such as the ATFES family) capable of running integrated EMC interference software such as the OFFSITE S/I module. The goal of automation is to provide decision aids so that within minutes and eventually seconds after electronic intelligence/surveillance systems locate and identify priority enemy victim systems, these victims can be engaged electronically with minimum impact on friendly force communications and electronic systems.

A graphic description of the friendly force (i.e., blue) jamming problem with oversimplified antenna patterns and propagation assumptions is at Figure 3. Simple cardiod antenna patterns are assumed. Propagation loss is based on square of distance and the assumption is that interference or jamming destroys the mission of an enemy (i.e., orange) or friendly force blue conventional receiver when the jammer signal level equals or exceeds the wanted signal. Hence, the EW deconfliction process' aim is to minimize the blue denial area or at least recognize it and work around it. Precise calculations can be made for any specific case with the previously mentioned software capability called OFFSITE S/I (in the case of ATFES as described in the previous paper). Area calculations can be made with another ATFES module called Line of Sight Area Coverage (LOSAC).

How real is the situation? It would be worthwhile to look at some unclassified literature in defense oriented periodicals. Figure 4 is a summary of an article published in the local Fort Monmouth newspaper. It shows substantial deployment to the European theater. QUICK FIX is airborne. TACJAM can engage three victims simultaneously. For readers of Defense Electronics, a high powered amplifier called PIRANHA, which the US Army procured for training on VHF jamming, is being advertised as a jammer. Another training-in-peacetime, jammer-in-wartime equipment mentioned in current defense oriented literature is the Heliborne Applique Communication Jammer which is advertised at 130 watts; compare this to the normal airborne communications radio rated at about 10 watts. Figure 5 is a summary of the NATO planning and shows the technology for jamming which is readily available. An extract of the Russian air defense weapons article which appeared in the same issue is at the bottom of this figure. This shows why airborne jammers face survivability issues unless they operate far from the Forward Line of Troops (FLOT).

Figure 6 summarizes the unclassified characteristics of two of the ground based jammers mentioned in Figure 4. Figure 7 summarizes the unclassified characteristics of an airborne jammer. Note that both are present in substantial numbers in a US division. Ref 2 has a fairly complete and unclassified description of the worldwide EW assets of most nations and reinforces the picture of powerful capabilities deployed among tactical forces.

Figure 8 provides some blanks which each signal officer, EW officer, air defender or artillery man can fill in with his own engineering judgement. My judgement is that ___ to ___ seconds fits the blanks; similarly, a ground based jammer closer than 10 kilometers will invite simple (i.e., without homing capability) but effective enemy artillery fire and I would rate probability-of-kill at ___ percent.

The "bottom line" of all this is that deconfliction resolution is important to military operations and, unless it is automated, it will not be responsive to the mobility of modern military forces. The alternative is to move closer to the FLOT and use the jam and skoot-before-you-are-targeted. The merit of such interim jamming is subject to effectiveness debate; the alternative here is to use coordinated multiple jammers against multiple targets, but this creates quite a control problem. Another argument against getting too close to the FLOT is the much greater target engagement area offered from a powerful rear area jammer, e.g., more than a whole division front.

Figure 9 outlines the US DOD efforts in deconfliction as part of the spectrum management contribution to the C3 countermeasures (C3CM) process. Deconfliction was empasized the the representatives of the Joint Chiefs of Staff (JCS) and the Joint Electronic Warfare Center (JEWC) who recognized that the spectrum managers already had a plateful of enigmas to solve for the electronic battlefield.

Figure 10 summarizes the essential results of the study and lists the ongoing efforts by each US service in automating spectrum management. The study called for interoperability among the automated systems, but recognized that service-unique requirements negated attempts at common hardware and software. It was felt that, whereas a data base on static civil and military facilities could be centrally maintained, it would be impractical for tactical spectrum dependent equipment, and, in light of the progress being made in computer interoperability, within echelon processing of a distributed data base was the prudent solution.

Figure 11 lists some of the ongoing groundwork as of a year ago. It should be noted that the US/NATO Joint Interoperability Tactical Command and Control System (JINTACCS) message formats have been developed for the command and EW community; there was limited participation by the spectrum management community.

Options.

There are several options for organizing the deconfliction process. One is to centralize the process in the maneuver commander's decision aid ADP systems. This has the advantage of enabling the command staff to execute risk analysis of kill vs jam. The programatic disadvantages are listed in Figure 12, but the emphasis on tight management might make this the better option for the future.

The decentralized option follows the TAFAAS approach about forming an ad-hoc C3CM or Joint Commanders Electronic Warfare Staff (JCEWS) and depends on distributed processing and interoperability to integrate the deconfliction process. A recent publication by JEWC (ref 3) reflects this approach. It will be shown that both approaches depend on interoperability and distributed processing of distributed data bases.

Going back to the five mission oriented view used by the US Army and very well explained in Ref 5, Figure 13 outlines how the decentralized option could be implemented by using remotes at the G3 EW section. Note that there is no direct electrical connection between remotes. This is a "quick-fix" to avoid the problem of different security levels which would probably still be maintained in wartime. This "quick-fix" would disappear when a multilevel secure capability is implemented in some future communications system.

Also note that each community has to be interoperable with its higher and lower echelon members and with its peers in other services and allied forces. To prevent overloading the spectrum manager of a given echelon with security problems, e.g., at corps, the ground rule being advocated is that his only connection to the intelligence/electronic warfare community should be via his counterpart at his own echelon, i.e., corps intelligence/EW.

Role and mission R&D is required to resolve interoperability between tactical computer systems whose main functions are spectrum management of friendly force communications and electronic systems and tactical computer systems whose main functions are fusing data from a variety of intelligence sources and providing tradeoff type decision aids for electronic engagement of selected targets.

Some definitions are in order with regard to the Restricted Frequency List (RFL) and are attached at Appendix A. They are summarized in Figure 14. It must be noted that the definitions originated in the context of jamming single frequency channel (SC) ground and airborne combat net radio (CNR) equipment; e.g., groundwave HF up to about 400 MHz in UHF. There is a same doubt about expanding the concept to the whole of military spectrum dependent equipment but it seems logical to this author.

Figure 15 is extracted from a Quadripartite Standardization Agreement, QSTAG 723, on Deconfliction. It has been studied by the JEWC who added that there was a time dependence or priority to be taken into account as explained in the appendix.

The figure implies that the doctrine is already in place, namely the EW community can jam any frequency and/or frequency band not on the Restricted Frequency List without further checking. If time is available, the spectrum management community should be informed. This doctrine can then be interpreted to show that the major role of the spectrum management community is to maintain the currency of the data on the Restricted Frequency List, as shown in Figure 16. This is the basis for the approach taken with the VHF deconfliction demonstration to be discussed later.

When time is available, the jamming of guarded and protected frequencies should be negotiated with the C-E section and/or the communications system spectrum managers. Jamming a TABOO frequency probably requires negotiation with higher authorities. It must be noted that the frequencies in use by selected enemy force units may be unknown during the planning stages and are found only after the combat operation starts. Alternatively, during combat operations, targets of opportunity with high priority for C3CM may be identified and targeted for electronic attack. In both cases, the enemy victims would probably respond by geographic and spectrum maneuvers and lead the jammers into a chase game.

A quick look at the centralized option is in order before going further. Figure 17 is an adaptation in Army structure of the work which the US Air Force R&D community is developing for a C3CM Battle Management Decision Aid program; the software includes a Frequency Analysis Aid for EW deconfliction. The Army ABSME program has limited itself to a study for an EW DEconfliction Management Subsystem which will be issued later in this year. Developers of US Army concepts and doctrine will be briefed on this study to determine which parts are operationally acceptable and/or should be rejected.

In this candidate, the spectrum management and the I/EW community feed information up to the maneuver commander's decision aids for tradeoff analysis of kill vs jam and then the command staff is able to analyze airborne vs ground jamming and, if airborne, use US Army assets or request US Air Force assets.

The bottom of this figure shows an interoperability need line which the author has been proposing as the way for the spectrum management community to effectively cope with enemy jamming.

One of the benefits which the spectrum management community could derive from being tied to the I/EW community is that the Army electronic surveillance assets could alert the spectrum management community of jammer activation. Figure 18 lists two of the systems (which were mentioned in Figure 4) which nominate enemy jammers as one of their targets for location and identification.

Before moving on to the actual capabilities developed, a heuristic summary is in order per Figure 19. If it is too late to standardize on exchangeable mass storage media, it is a missed opportunity as an effective alternate to electrical transmission of large files of data.

As stated in the previous paper, Ref 1, the EMC software available from EMC consultants and/or military spectrum management ADP aids are more than adequate for most cases. In the case of the ABSME effort, the OFFSITE S/I software is probably more than adequate for maneuvering force-on-force electronic engagements. Hence, the current emphasis is on interoperability between the two communities and their ADP systems.

Field Evaluation.

Obviously, there are too many missing parts to run a complete EW deconfliction resolution test bed. The philosophy adopted is outlined in Figure 20, essentially make whatever progress the state of the available systems permit.

The short term goals (to be consistent with very limited funding) was to demonstrate the ability of keeping the VHF part of the Restricted Frequency List as current as possible and sharing annotated VHF frequency assignments with the I/EW community as shown in Figure 21. When time is available in the deconfliction process, usually only during the planning stages, the ATFES EMC software capabilities, essentially OFFSITE S/I and LOSAC could be used to provide EMC analysis.

The experiment was developed for a stand alone division basis; it was assumed that the EW section and the C-E section of the G3 (operations) staff had the missions and authorities listed in Figure 21 and used interim personal computers (IPC) at a distributed command post to coordinate their decisons.

VHF DECON.

Figure 22 shows the demonstration setup. As shown in Figure 3, a blue jammer attempting to jam an organge link denies substantial area of spectrum use on the blue side. Software capabilities discussed in the previous paper, such as OFFSITE S/I or LOSAC II can make accurate EMC analyses if the characteristics of the antennas, the transmitter and the unintentional victims are known. What is missing is the ability to maintain the data base on friendly force frequency assignments as current as possible with regard to restrictions so that EW deconfliction analysis (or fracticide resolution) can be executed in the G3 (operations) part of the corps or division CEOI. Maintaining the status of each frequency assignment in the Restricted Frequency List is, therefore, the major contribution by the distributed spectrum management community.

Although the effort started with the VHF (30 to 88 MHz) band, the principles are essentially applicable to all bands. The VHF DECON demonstration was based on simple software using an off-the-shelf data base management system (ORACLE SQL) with the users having the option to call on other ATFES software for more detailed analysis.

Figure 22 outlines the types of interactions which this simple setup can demonstrate. Once these capabilities are demonstrated in the field, the principles will be expanded to all other frequency bands and spectrum usage.

The criteria for success is timeliness to execute the electronic attack and defense for the case of the targets of opportunities and unexpected activation of enemy jammers.

Conclusions.

The problem of EW deconfliction has many players and lots of studies and analysis. The spectrum management community is now working to standardize the approach among combined and joint forces. The demonstration developed by the US Army has been successsful in a wired up environment. Its ability to operate in an environment based on tactical communications among distributed organizations has yet to be proven.

The major foreseen problem is one of secure interoperability, not sophistication of EMC analysis. As summarized in Figure 24, this leads to the case of meeting international and national standardization actions, most of which are still evolving.

References.

1. S.M. Segner, USACECOM, Automating the Deconfliction of Jamming and Spectrum Management, AGARD, Oct 87, Paper #28.
2. The International Countermeasures Handbook: 12th Edition - EW Communications, Inc., Palo Alto, California.
3. JDD 2-84, April 1984, Joint Electronic Warfare Center, subj: Electronic Warfare Frequency Deconfliction in Joint Military Operations.
4. ABC Quadriprtite Standardization Agreement 623, Battlefield Spectrum Management, Deconfliction, approved May 1985.
5. J. Goghlan, Defense Electronics, Defense Electronics, June 1984, Vol 20, No 6, pg. 85.

PIG 1

AGARD CONFERENCE, 26-30 OCTOBER 1987

THEME: EFFECTS OF ELECTROMAGNETIC NOISE AND INTERFERENCE ON PERFORMANCE OF MILITARY RADIO COMMUNICATIONS SYSTEMS

PAPER #35: AUTOMATING THE DECONFLICTION OF JAMMING AND SPECTRUM MANAGEMENT

SAMUEL M. SEGNER
US ARMY COMMUNICATIONS-ELECTRONICS COMMAND
FORT MONMOUTH, NEW JERSEY 07703

A STUDY AND ANALYSIS SUBJECT OF THE US ARMY AUTOMATED BATTLEFIELD SPECTRUM MANAGEMENT AND ENGINEERING (ABSME) PROGRAM.

US ARMY COMMUNICATIONS ELECTRONICS COMMAND

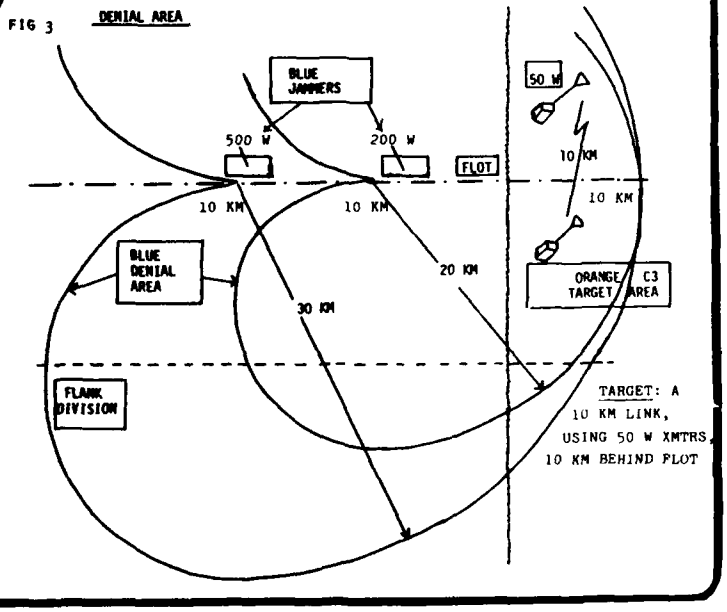
PIG 2

COMPETING GOALS FOR USE OF SPECTRUM

- FREQUENCY SPECTRUM MANAGEMENT/ENGINEERING FOR:
  - EFFICIENT/OPTIMUM USE OF THE AVAILABLE FREQUENCY SPECTRUM FOR:
    - COMMAND CONTROL COMMUNICATIONS EQUIPMENT
    - WEAPONS SYSTEMS ELECTRONICS
    - MILITARY AVIONICS
  - TAKING INTO ACCOUNT:
    - PEACETIME VS WARTIME AVAILABILITY
    - ECCM MODES OF EACH C-E EQUIPMENT:
      - POWER BOOST
      - SPREAD SPECTRUM
        - FREQUENCY HOPPING (ECM AVOIDANCE)
        - DIRECT SEQUENCE (ECM PROCESSOR)
- INTELLIGENCE/ELECTRONIC WARFARE MANAGEMENT/ENGINEERING FOR:
  - ELECTRONIC SURVEILLANCE(PASSIVE LOCATION/IDENTIFICATION)
  - ELECTRONIC ENGAGEMENT
    - JAM - DECOY

US ARMY

FIG 3 DENIAL AREA



PIG 4

I/EW SYSTEMS FOR ECM VS C3 SYSTEMS*

- TRAFFIC JAM AN/TLQ-17A(V1): MEDIUM POWER
  USAREUR: 16; FORSCOM: 08; EUSA: 03
- QUICK FIX 1B (EH-1H) AG/GLR-9 RCVR & AN/TLQ-17A(V1)
  USAREUR: 14; FORSCOM: 03; EUSA: 03
- TACJAM AN/MLQ-34: 3 SIG HI PWR & 37' MAGIC MAST
  USAREUR: 12; FORSCOM: 02
- TRAILBLAZER AN/TSQ-114A: 5 STATION INTERCEPT, 50' MAGIC MAST
  USAREUR: 03
- TEAMMATE AN/TRQ-32(V) INTERCEPT, 25' QUICK ERECT
  USAREUR: 12

*16 NOV 84, PG 6, MONMOUTH MESSAGE

US ARMY COMMUNICATIONS ELECTRONICS COMMAND

PIG 5

DEFENSE ELECTRONICS, SEP 86

PG 18: "NATO EW TRAINING PROGRAM ENTERS PHASE II"
MULTISERVICE EW SUPPORT GROUP ----------
3 AIRCRAFT ---------- EARLY 90'S

PHASE IIA: RADAR JAMMERS
HF, VHF, UHF COMM JAMMERS

PHASE III: GROUND BASED JAMMERS

PG 136: TABLE "SOVIET BASED WEAPONS"

SA 7	GRAIL	7 KM
SA 6	GAINFUL	33 KM
SA 12	GLADIATOR	100 KM

US ARMY COMMUNICATIONS ELECTRONICS COMMAND

PIG 6

AN/MLQ-34 TACJAM

Description:

The AN/TLQ-34 is a communications jammer, that can jam three targets simultaneously. Rapid set-up and tear-down, this highly mobile ground based ECM system, features quick erect mast, ground rod driver, and on-board generator.

System Organization and Operations:

2 systems per separate Armored Cavalry Regiment/Brigade
3 systems per Division
3 systems per Corps

AN/TLQ-17 TRAFFIC JAM

Description:

The AN/TLQ-17 countermeasures set is capable of disrupting hostile communications. A modern, small, highly mobile piece of equipment which uses state-of-the-art solid state micro-processor techniques. The system is deployable as a separate unit, a jeep-mounted jammer or as helicopter-borne countermeasures capability.

System Organization and Operations:

2 systems per Armored Cavalry Regiment/Brigade
3 systems per Division
9 systems per Corps

US ARMY COMMUNICATIONS ELECTRONICS COMMAND

FIG 7

EH-60 QUICK FIX

Description:

The EH-60 is a helborne system intended to provide tactical Signals Intelligence and Electronic Warfare support to the Army. The system has intercept, jamming and direction finding capabilities and interfaces with similar airborne and ground units via secure voice and data links.

System Organization and Operations:

3 aircraft per Division
4 aircraft per Armored Cavalry Regiment/Brigade

Targets:

Single channel tactical radios and communications jammers.

US ARMY COMMUNICATIONS ELECTRONICS COMMAND

FIG 8

JAMMER TRADEOFF/SURVIVABILITY ISSUES

- WITHIN _____SECONDS LOCATION IS CONFIRMED AS TARGET FOR ENEMY FIELD ARTILLERY OR AIR DEFENSE ARTILLERY.
- HI RISK ZONE [P(KILL) > ___%]
  PLOT TO 10 KM (#1, #2)
- RISK ZONE: 10 KM TO 20 KM

NOTE:

1. JAM AND SKOOT VS SUSTAINED ECM.
2. COORDINATED MULTIPLE JAM VS MULTIPLE TARGET CAPABILITY.

US ARMY COMMUNICATIONS ELECTRONICS COMMAND

FIG 9

SPECTRUM MANAGEMENT ROLE

- JTC3A WORKING GROUP STUDY REPORT, 31 MAY 85, "TACTICAL AUTOMATED FREQUENCY ALLOTMENT/ASSIGNMENT SYSTEM (TAFAAS)"

JCS-C3	ARMY (CECOM, HQ)
CINCRED-FMO	USAF (FMO, TAC)
JEWC	USN (NAVEMCEN FMO)
JTC3A	USMC (HQ)
DOD ECAC (ARMY, NAVY, AF, MC DPM)	

- ASSUME: AD HOC C3CM/CCM CELL
  - JOINT/CINC DECONFLICTION
  - CORPS/DIV DECONFLICTION

US ARMY COMMUNICATIONS ELECTRONICS COMMAND

CECOM - PROVIDING LEADERS THE DECISIVE EDGE

PIG 11

ONGOING GROUNDWORK

- CONCEPTUAL:
  - JCS PUB 24: EW IN JOINT MILITARY OPERATIONS
  - JCS PUB 28: JOINT C3CM/CCM ---------
  - TAC-TRADOC: COMPASS CALL, SEP 80
    JOINT C3CM, DEC 81, ? 86
  - DALFA: 5-80: COMPASS CALL
    1-82: EW SUPPORT OF JOINT AIRLAND OPS
    2-85: COMBINED EW OPS (NATO)
- JINTACCS MSG:
  - REQUEST/TASKING FOR EW SUPPORT
  - ALERT/INTENTION TO CONDUCT EW
  - STOP JAMMING REQUEST
  - USMCEB/NATO MSG FORMATS
- FIELD TEST:
  - COMFY FRED/SHIRE (USAF)
  - TEAM SPIRIT (CINCPAC/KOREA JFMO & PACAF S-W)
  - GREEN/BLUE FLAG (USAF)
  - COMPASS CALL (CONUS & NATO)
  - REFORGER (USAREUR VHF)

PIG 10

TACTICAL AUTOMATED FREQUENCY ALLOTMENT/ASSIGNMENT SYSTEM

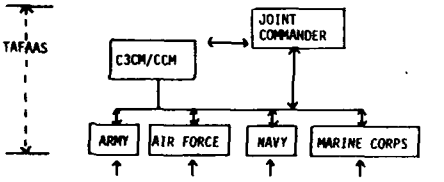
JTC3A 31 MAY 1985

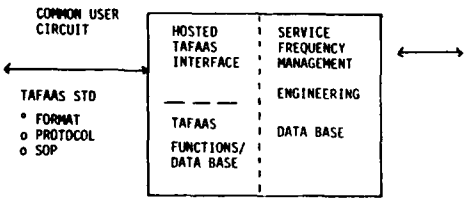
WORKING GROUP STUDY REPORT

ARMY INTEGRATED BATTLEFIELD SPECTRUM MANAGEMENT SYSTEM

USN, USMC
FREQUENCY ASSIGNMENT COMPUTER TERMINAL SYSTEM (FACTS)

USAF
SYSTEM CONFLICT ANALYSIS & MANAGEMENT SYSTEM (SCAMS)





US ARMY COMMUNICATIONS ELECTRONICS COMMAND

CECOM - PROVIDING LEADERS THE DECISIVE EDGE

PIG 12

OPTIONS FOR DECONFLICTION

1. CENTRALIZE FOR C3CM/CCM
   - KILL: ONE ON ONE VS WEAPON SURVIVABILITY.
   - JAM: ONE ON MANY VS JAMMER SURVIVABILITY.
   - DEDICATED PERSONNEL.
   - R&D FOR NEW ADP SYSTEM.
   - IMPACTS 1980-2000 EXISTING PROGRAMS.
2. DECENTRALIZE AND INTEROPERATE
   - AUGMENT BFSM EMC ANALYSES.
   - AUGMENT I/EW EMC ANALYSES.
   - EXISTING PERSONNEL.
   - RDT&E INTEROP: NETWORKING, DATA BASE MAINTENANCE AND OPS PROCEDURES.
   - RESOLVABLE IN LATE 80's

CONSENSUS - EXECUTE #2
STUDY #1 (IF #2 SHORTFALLS).

PIG 14

COMPONENTS OF RESTRICTED FREQUENCY LIST

TABOO. TABOO FREQUENCIES ARE ANY FRIENDLY FUNCTIONS/FREQUENCIES OF SUCH IMPORTANCE THAT THEY MUST NEVER BE DELIBERATELY JAMMED OR INTERFERED WITH BY FRIENDLY FORCES.

PROTECTED. PROTECTED FUNCTIONS/FREQUENCIES ARE THOSE FRIENDLY FUNCTIONS/FREQUENCIES USED FOR A PARTICULAR OPERATION, IDENTIFIED AND PROTECTED TO PREVENT THEM FROM BEING INADVERTENTLY JAMMED BY FRIENDLY FORCES WHILE ACTIVE EW OPERATIONS ARE DIRECTED AGAINST HOSTILE FORCES.

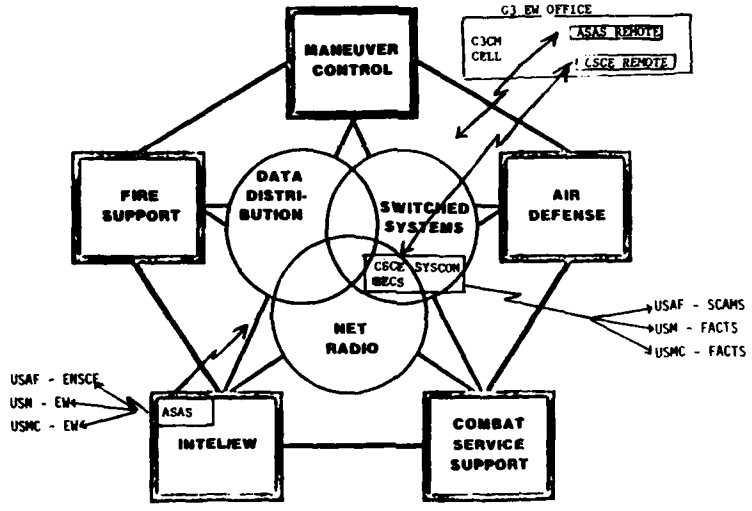
GUARDED. GUARDED FUNCTIONS/FREQUENCIES ARE ENEMY FUNCTIONS/FREQUENCIES THAT ARE CURRENTLY BEING EXPLOITED FOR COMBAT INFORMATION AND INTELLIGENCE.

US ARMY COMMUNICATIONS ELECTRONICS COMMAND

CECOM - PROVIDING LEADERS THE DECISIVE EDGE

PIG 13

ARMY TEST BED ARCHITECTURES FOR C³CM DECONFLICTION, DECENTRALIZED OPTION



CECOM - PROVIDING LEADERS THE DECISIVE EDGE

PIG 15

DYNAMICS OF RESTRICTED FREQUENCY LIST

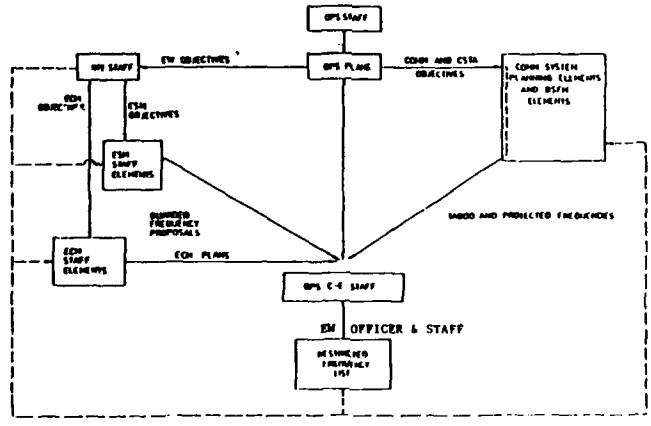


FIG 16

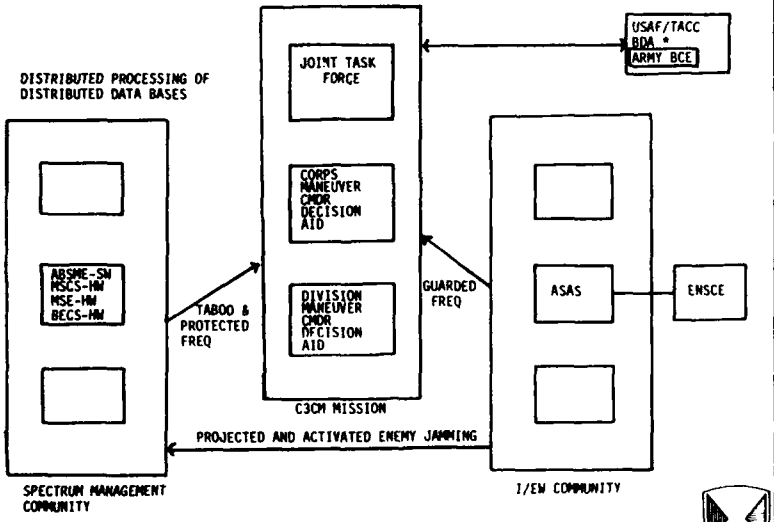
GOALS FOR AUTOMATING PROCESSING OF RESTRICTED FREQUENCY LIST

- MAXIMUM FREE PLAY TO EW BY KEEPING RFL MINIMUM AND ON TOP OF BATTLE SITUATION.
- MINIMIZE COMM DENIAL BY USE OF EMC ANALYSES S-W BASED ON DIGITIZED MAPS.
  - PATH PROFILE
  - SIGNAL PENETRATION AREAS

US ARMY COMMUNICATIONS ELECTRONICS COMMAND

FIG 17

DECONFLICTION INTEGRATED WITH KILL/JAM C3CM AIDS



*ENSCE = ENEMY SITUATION CORRELATION ELEMENT
*BDA = USAF/RADC AUTOMATED C3CM BATTLE MANAGEMENT DECISION AID

US ARMY COMMUNICATIONS ELECTRONICS COMMAND

FIG 18

AN/TSQ-114 TRAILBLAZER ANL AN/TRQ-32 TEAMMATE

AN/TSQ-114 TRAILBLAZER

Description:

Division level communications intercept and direction finding system consisting of 5 geographically dispersed components controlled through a real time data link.

System Organization and Operations:

5 stations per heavy division

Targets:

Single channel tactical radios and communications jammers.

AN/TRQ-32 TEAMMATE

Description:

The "TEAMMATE" is a mobile, multi-station communications intercept and direction finding system for the support of the Army's maneuver elements. The AN/TRQ-32 is automated and features both a quick erect antenna mast as well as a self contained generator/air conditioning system.

System Organization and Operations:

3 systems per Light Division
2 systems per Armored Cavalry Regiment

Targets:

Single channel tactical radios and communications jammers.

US ARMY COMMUNICATIONS ELECTRONICS COMMAND

FIG 19

HEURISTIC SUMMARY

- DECONFLICTION PROCESS HINGES ON COORDINATING I/EW AND BSM VIA RESTRICTED FREQUENCY LIST.
- ALTERNATE ARCHITECTURE:
  - OPTIONS OF WHERE AND DEPTH OF ANALYSES.
  - BSM MAINTAIN DISTRIBUTED DATA BASE AND DISTRIBUTION PROCESSING THEREOF.
- REQUIRED STANDARDS:
  - ELECTRONIC DATA TRANSPER.
  - MASS STORAGE MEDIA.
- INTEL PROVIDE REC ALERT.

US ARMY COMMUNICATIONS ELECTRONICS COMMAND

FIG 20

EVOLUTION OF FIELD EVALUATION/ANALYSIS

1. WATCH TOMORROW SYSTEM CONCEPT.
2. ADAPT TOMORROW TO TODAY WHENEVER POSSIBLE.
3. DON'T WAIT FOR TOMORROW -
   ("IT IS ALWAYS A DAY AWAY")
4. TEST WITH TODAY SYSTEM REALITY.
   - NETWORKING CAPABILITY
   - 1200 BPS
   - EDC
   - BULK - COMSEC

US ARMY COMMUNICATIONS ELECTRONICS COMMAND

FIG 21

CORPS PLAYERS

EWS IPC: ECM PLANNING & IDENTIPYING CONFLICT WITH RPL.
PRIORITIZE GUARDED FREQUENCY LIST.
MAINTAIN RFL IAW CORPS C-E IPC.
TASK DIVISION IPC AND/OR USAF IPC.
GRANT VIOLATIONS OF RFL.
EXTENSION: USE ATPES S-W TO ANALYZE VIOLATIONS.

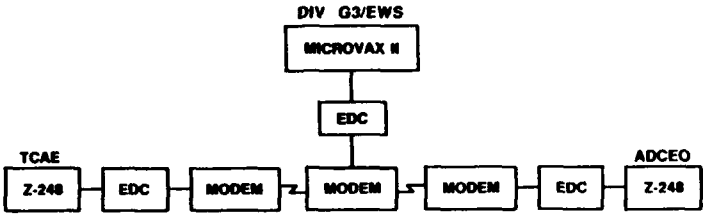
CORPS IPC: AUTHORIZE CORPS & EAC TABOO LIST.
COORDINATE AND PRIORITIZE CORPS & DIVISION PROTECTED LIST.
ANALYZES GRANTED VIOLATION OF RFL.
EXTENSION: USE ATPES TO REASSIGN.
ANALYZES ENEMY JAMMING.
EXTENSION: USE ATPES TO REASSIGN.

US ARMY COMMUNICATIONS ELECTRONICS COMMAND

PIC 22

VERY HIGH FREQUENCY DECONFLICTION VHF DECON

- REFINE A SET OF PROCEDURES FOR VHF DECONFLICTION THAT IS COMPATIBLE WITH DOCTRINE AND OPERATIONS
- DESIGN AND DEVELOP A HARDWARE/SOFTWARE PACKAGE (INCLUDING NETWORKING CAPABILITY) TO AUTOMATE RFL GENERATION AND UPDATING
- DELIVER, INSTALL, AND TEST THE SYSTEM IN FRG



PIC 23

INTERACTIONS BETWEEN BFSM AND I/EW SYSTEMS

- DECONFLICT FRIENDLY FORCE C^3 & CSTA SYSTEMS FROM FRIENDLY FORCE ISTA & JAM/DECOY SYSTEMS:
  - PLANNED MISSIONS
  - TARGETS OF OPPORTUNITY
- ALERT BFSM TO OBSERVED & CHARACTERIZED ENEMY REC:
  - OBSERVED WEAK POINTS
  - ACTIVATED JAMMERS & LOOK THRU SERVICE
- COORDINATE LOCAL MIJI ANALYSES

PIC 24

LONG TERM ISSUES

1. MESSAGE FORMAT
2. DATA BASE FORMAT
3. ACCESS CONTROL/MI ALERT
4. NETWORKING
5. ARTIFICIAL INTELLIGENCE

DETERMINING THE AZIMUTHAL DISTRIBUTION OF NOISE AND INTERFERENCE ACROSS THE HF SPECTRUM AT THE RAE RECEIVING STATION, COBBETT HILL, SURREY, ENGLAND

by

H. L. Spong
Cobbett Hill Radio Station
MOD(PE), Royal Aircraft Establishment
Farnborough, Hants, GU14 6TD, England

SUMMARY

An increasing interest in techniques for nulling out unwanted signals and interference in the HF band makes it more important to establish not only the magnitude of the signals but the direction from which they are emanating and whether they are of a spasmodic or continuous nature.

RAE has been anxious for some time to build up a data base of such information from which statistical trends may be derived and applied to research into more effective HF communications systems for the Royal Air Force.

This presentation which gives a simple outline of a measurement system built around a Multi-Vee antenna, a Rohde and Schwarz ESH3 measurement receiver and a Hewlett Packard HP9836 computer all installed at the RAE Receiving Station, Cobbett Hill, goes on to explain the various formats available on the HP9836 computer for displaying the data, for example, Signal versus Time and Signal versus Direction, etc.

Data acquired from this system has so far been based upon two measurement methods. The first takes up to five frequencies of interest across the band and measures these at the same time or times on consecutive days in an effort to determine the constancy or otherwise of the direction and magnitude of the interfering signals.

The second approach has been to evaluate the measurement system as a means of gathering data on channel occupancy; this has been looked at by measuring on five adjacent frequencies within a 3 kHz bandwidth over an interval of time and on a number of occasions, either on various or on consecutive days.

The presentation is concluded when some measured data are discussed and the respective directional characteristics of the interference is considered.

1 INTRODUCTION

Although the RAE Receiving Station at Cobbett Hill, Surrey, is located at an electrically quiet site, difficulties, in common with other European sites, are encountered on HF channels from interference and noise.

With a growing emphasis on digital modulation methods for communication purposes, RAE consider it important to acquire a bank of data from which statistical trends may be extracted for application to research into methods of combating interference, for example, by null steering antennas or implementing EDC techniques during the development of more advanced communications systems for the RAF.

To meet this need RAE has developed a versatile computer controlled HF measurement system allowing unattended measurements to be made across the HF spectrum over the whole or part of a 360 deg of azimuth. Added features are the ability to take sample readings on any reference antenna at the end of each azimuth scan for comparison purposes and at the completion of measurement to make limited data analysis.

2 MEASUREMENT SYSTEM

The measurement system can be seen by reference to the block diagram in Fig 1. At the centre of the system is a Hewlett Packard 9836 computer with a colour graphics display and one megabyte of memory which controls via suitable interface units a Rohde and Schwarz ESH3 HF measurement receiver and a Multi-Vee antenna. Measured data may be stored either on a Winchester disc or on floppy disc, while an IEEE bus allows analysed data to be output as hard copy to a printer or graphics plotter. When fully implemented it will be possible to store measured data on magnetic tape via a RS232C interface.

The software currently incorporates a number of data analysis routines which allow the operator to gain a quick assessment of the interference characteristics, however, because of the restriction in the size of memory more advanced analysis in terms of statistical computation and display has at the moment to be done by manual methods. Plans have been made to transfer data from the HP9836 via suitable modems to a larger computer such as a VAX 11/730 for further processing once the necessary hardware has been acquired.

THE MULTI-VEE ANTENNA

The RAE Multi-Vee antenna forms an integral part of the measurement system. It is based upon the well-known Sloping Vee antenna which displays good directional properties and is predominantly horizontally polarised. Fig 2 is an artist's impression of the structure which has the following physical parameters. At the top of a 16.4m wooden lattice tower is a switching unit connected to a total of 20, 100m antenna wires which radiate outwards at 18 deg intervals and slope down to ground level where each one is terminated by a 300 ohm resistor.

The switching unit which is driven from the computer contains a set of reed relays used to select the appropriate pair of wires to form the antenna during the scan. To provide an optimum antenna pattern across the band it was deemed necessary to have an apex angle of 36 deg. This meant that for any direction the antenna comprised a pair of wires 36 deg apart with a disconnected 'dormant' wire down the centre, ie wires A and C for direction B and wires B and D for direction C, etc. Modelling and antenna measurements showed that the 'dormant' wire had an insignificant effect upon the antenna pattern.

The selected pair of wires at any moment was connected via the reed relays to the primary side of a matching transformer (BALUN), the secondary of which was connected to a 50 ohm coaxial cable of known loss to feed the Rohde and Schwarz receiver housed along with the computer in the main building 250 m away.

A feature of the antenna switching unit makes use of a centre tap on the primary winding of the BALUN to make an automatic test of the integrity of the antenna wires and associated connections. This is achieved by measuring the loop resistance by virtue of the current flowing through half the BALUN primary winding, the antenna wire, the terminating resistance and the Earth return (soil resistance).

4 ANTENNA CHARACTERISTICS

Before carrying out any serious measurements it was necessary to characterise the antenna to establish whether or not it displayed the correct pattern in azimuth and elevation.

This was relatively easy to do in terms of vertical polarisation by using a well tried RAE method of flying a standard radiator (a vertical half-wavelength dipole and battery driven oscillator) suspended below a helicopter, around the antenna array whilst field strength measurements were taken. Fig 3 is an example of the result of a measurement on a 3 MHz frequency using this technique and also shows the equivalent elevation pattern obtained by measuring field strength as the helicopter climbed vertically over a fixed point at a known range.

For interest a theoretical pattern in the horizontal plane has been superimposed on the azimuthal plot to give the reader an indication of the expected directivity and front to back ratio of this type of antenna.

There are however difficulties in attempting to measure the antenna pattern using a horizontally polarised standard radiator. Attempts using a horizontal dipole were unsatisfactory due to antenna aerodynamic stability problems together with the dipole pattern effect. Doubts have also been expressed concerning the effects of airframe proximity and ground reflections. Plans to use a horizontal loop antenna are being considered, some initial tests indicate that the method shows promise and a suitable transmitter capable of driving sufficient current round the loop is being developed.

It is pointed out that the Vee antenna at low elevation angles displays a null in the direction of interest (Ref 1) for vertically polarised signals and this effect is apparent at 340 deg in Fig 3b.

Results of the vertically polarised measurements in terms of polar patterns and front to back ratio gave confidence that the antenna was performing satisfactorily such that a measurement programme could be initiated.

5 MEASUREMENT PLAN

Two types of measurement programme were considered. The first was to examine a number of frequencies across the HF spectrum in a wideband mode using a 3 kHz bandwidth for various times of the day and at the same times on consecutive days; the second was to take a number of adjacent frequencies within a 3 kHz bandwidth and measure these for a period of time daily in an endeavour to establish whether or not the system had potential from the point of view of making channel occupancy measurements. It is pointed out that measurements on adjacent frequencies cannot be made simultaneously since a complete scan of the Vee antenna and perhaps a measurement on a reference antenna would take about 2 seconds before the frequency would change.

6 MEASUREMENTS

Having checked the integrity of the antenna and the measurement system it is necessary to set up the measurement parameters which is achieved by means of an input file.

Fig 4 is an example of such a file and it can be seen that on this occasion 11 measurements (TRIALS) were called up, each lasting 10 minutes (DURATION) and separated by 6 hours (INTERVAL).

Measurements were set to commence at 1800 hours on 17 October 1986 and all 20 sectors (360 deg) were to be scanned as well as taking a measurement on the reference, in this case the WBD312 (Wide Band Dipole). The averaging period of the receiver was set to 2 seconds and the receiver settling time was 5 milliseconds.

The five frequencies chosen for the measurement are given in MHz along with the reception mode (USB) and bandwidth (2.4 kHz).

At the approach of the measurement time determined by the clock in the computer the measurement receiver performs a self-check calibration. At measurement time the system commences to take measurements making one complete scan of the Vee antenna and one sample on the reference in about 1.6 seconds before repeating on the next frequency.

Data are stored on disc as *RECORDS* and the number of readings per RECORD is a function of the number of sectors, the number of frequencies and the number of complete data sets. Data sets are a function of the number of frequencies and the averaging time. The total capacity of a floppy disc is 264 kbytes which is equivalent to storing about 6 hours of data if scanning all 20 sectors plus the reference.

At the conclusion of the set of measurements the screen will display the words 'MEASUREMENT COMPLETE' and the system will halt.

7 DATA DISPLAY

At the completion of a measurement or at a later stage the relevant file may be recalled and a basic analysis carried out by the use of a number of inbuilt routines.

Having called up the date and time of the period of interest the operator has the option of four types of data display which are described as follows:

Fig 5 shows the display of measured data in tabular form and represents the level of signal/interference in dBμV for each sector plus the reference on each frequency in turn set against the start time of each scan. It is possible to examine the tables and determine the trend on each frequency in turn against time.

A better approach is to call up the Signal versus Time routine for a particular sector of the Vee antenna or the reference. By specifying the particular time interval of interest, eg for 10 minutes from 1800 hours in Fig 6, a time amplitude plot can be produced. Up to five sectors can be overwritten on one print out, although this is satisfactory in colour the display is confusing in black and white so for clarity only results for the reference and one sector are produced here.

It is further possible to examine more closely parts of the Signal versus Time display by expanding the portion of interest.

In order to assess the relative levels of signal against frequency, it is possible to call up a Signal versus Frequency routine. It is first necessary to establish the moment in time, say, from the Signal versus Amplitude display, at which the information is needed. Figs 7 and 8 are examples of this type of format and provide as a 'snapshot' in time a histogram of the level of interference on the five frequencies for the reference and sector 11 of the Vee antenna respectively.

A more explicit form of display is the Amplitude versus Direction format, an example of which can be seen in Fig 9. This again is for a 'snapshot' in time and indicates the magnitude of the signal in dBμV against bearing. Included on the display are two scales, one showing the calibration and signal maximum of the Vee antenna and the other the level of received signal on the reference at the approximate same moment in time.

8 DISCUSSION OF RESULTS

For the purpose of this presentation two sets of results are considered, both in line with the measurement plan outlined in section 5 above.

The first method is centred on the five frequencies depicted in the input file at Fig 4.

By selecting a time of interest at say 1800 hours on 18 October 1986 and running the Signal versus Frequency routine for the reference antenna, a histogram is produced, Fig 7, which shows the amplitude of the interference measured on each frequency.

Examination of these results shows that the greatest level at that instant was +25 dBμV on the 9 MHz frequency. If we now concentrate on this frequency and run the Signal versus Direction routine, Fig 9, it is immediately apparent that this signal is emanating from a direction of approximately 180 deg, ie sector 11.

By running the Signal versus Frequency routine for sector 11 on the Vee, a second histogram is produced, Fig 8, which indicates the level of interference observed on each

of the five frequencies. Comparison of the two histograms gives an indication of the directivity of the Vee antenna at the other four frequencies, although at this stage one cannot be certain whether this apparent discrimination is entirely a function of the antenna or the fact that there was a small time difference between the two antenna measurements. This uncertainty could be resolved by running the Signal versus Time routine for each frequency on both antennas.

An example of the Signal versus Time results at 9 MHz for the reference and Vee sector 11 may be seen by reference to Fig 6 from which the signal levels shown on the histograms at 1800 hours correlate.

To establish how the direction of interference varies with time at 9 MHz, data, taken once every 6 hours from 1800 hours on 18 October 1986 to 0600 hours on 20 October 1986, has been analysed by manual means based on 60 samples of signal on each bearing every 6 hour interval. The results have been displayed for the three days in Fig 10a-c.

It will be seen that for this period of measurement at this frequency the direction of interference was remarkably constant from a bearing of about 180 deg with the variation in signal magnitude following the diurnal MUF variation.

From a study of HF predictions it is possible to speculate that the unwanted signals were emanating from an area in Southern Europe or North Africa.

It is pointed out that the system does not allow for the recording or analysis of the modulation because the scanning process allows only a brief dwell time on each sector and on each frequency.

Fig 11, which has been produced again by manual means, has been included to show the cumulative distribution of the interference on the reference antenna and sector 11 of the Vee for 1800 hours on 18 October 1986. Examination of the two histograms shows clearly the level of received signal at that period in time and gives an indication of the directional qualities of the Vee antenna.

The second type of measurement relates to a narrow band assessment; for presentation purposes the same 9 MHz channel has been selected but on a different day together with some results for a 3 MHz channel. On both occasions the receiver parameters were set to be in the CW mode having a 500 Hz bandwidth. Five frequencies were chosen separated by 500 Hz and each centred on the respective centre frequency thereby looking across a 2.5 kHz frequency window during each scan.

Fig 12a shows a set of results, manually constructed from a 10 minute measurement period at 2100 hours GMT on 1 October 1986. What is immediately apparent is that the interference is broadband in nature spanning the 2.5 kHz bandwidth and is emanating from a bearing of 072 deg, being in the direction of the European Continent. By reference to frequency prediction data for this date and time it is observed that a 9 MHz signal propagated via a 1F2 mode suggests the signal path is about 2000 km and therefore the source could be situated in Western Russia. Calculations show that the arrival angle of such a signal is about 17 deg which is well within the main lobe of the antenna.

A similar measurement at 3 MHz at 2100 hours GMT on 15 December 1986 has been analysed as shown in Fig 12b. These data show the interference to be narrow band in nature with three of the five frequencies relatively quiet. The remaining channels show interference from more than one direction and again by applying prediction information it is possible to determine approximately from where the source of the interference might be.

9 CONCLUSIONS

The measurement system has so far been used to examine only a few of the frequencies allocated to the station. It has been proved that a large amount of data can very quickly be acquired and apart from the simple inbuilt analysis routines a more powerful analysis capability is urgently required to deal with the mass of information. The versatility of the system has however been proved and much useful information can be presented.

From the data displayed in this report it has been shown that it is possible to determine the duration and direction of any particular interfering signal and from knowing the precise frequency and by reference to frequency predictions, it is possible to speculate in broad terms the approximate area from which the signal is derived.

Although the 9 MHz data discussed in this report shows interference from one main direction over several days, it must not be assumed that this condition remains constant, as has been proved from other measurements. As already stated more powerful analysis methods need implementing to determine proper trends and provide a statistical data bank.

Looking at the inband measurements, the first example considered shows a wideband interfering signal spread across the 2.5 kHz bandwidth with no noticeable clear channels, whereas the second example shows examples of clear channels within the baseband.

It has been shown that the system has great versatility for gathering data on the distribution of noise levels and, although measurements on adjacent frequencies are not made simultaneously, useful information regarding channel occupancy can be obtained.

10 REFERENCE

1 Royal Canadian Air Force, Central Experimental and Proving Establishment Report 1086. 'Sloping Vee antenna performance and design'. April 1955.

11 ACKNOWLEDGMENTS

The author wishes to acknowledge the work of Mr M.J. Maundrell, Head of Field Stations, RAE, who conceived the idea of the measurement system and was instrumental in having it built.

Copyright
©
Controller HMSO London
1987

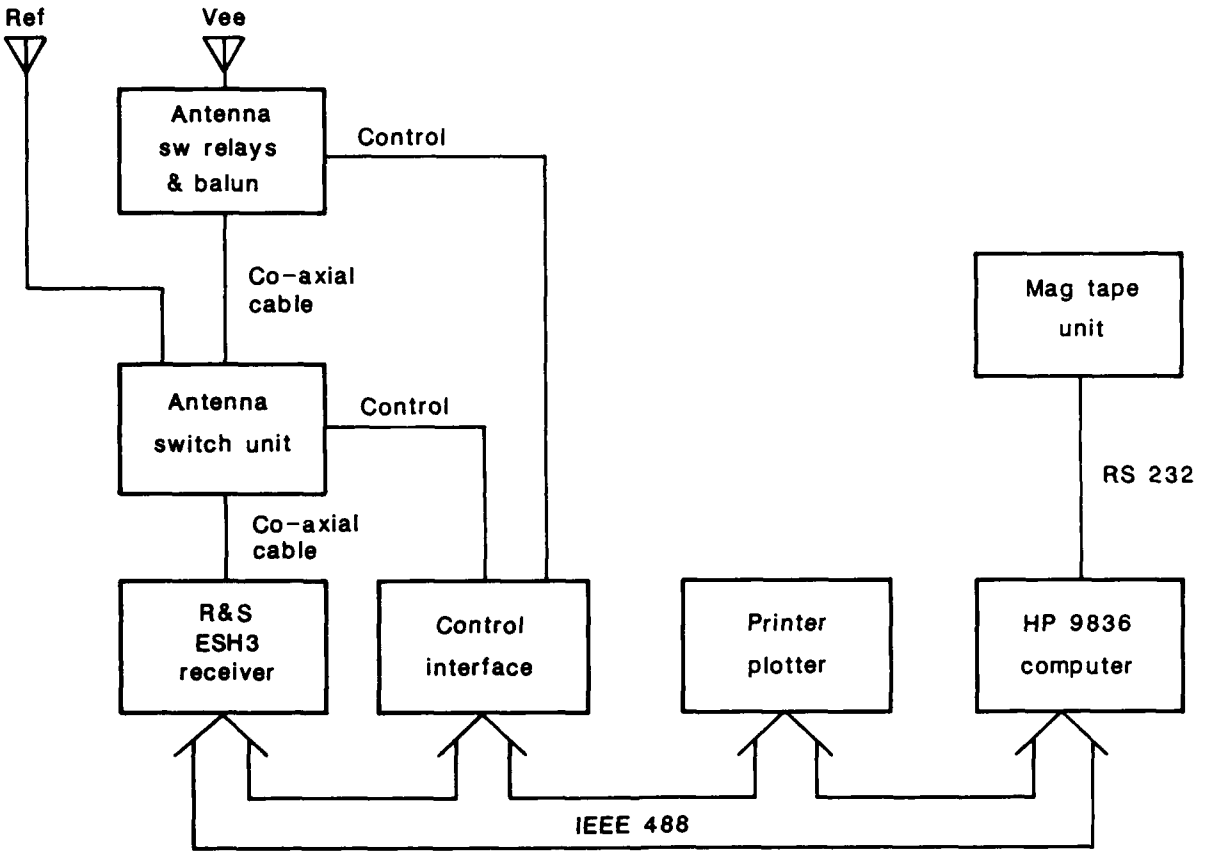


Fig 1 Block schematic of computer controlled measurement system

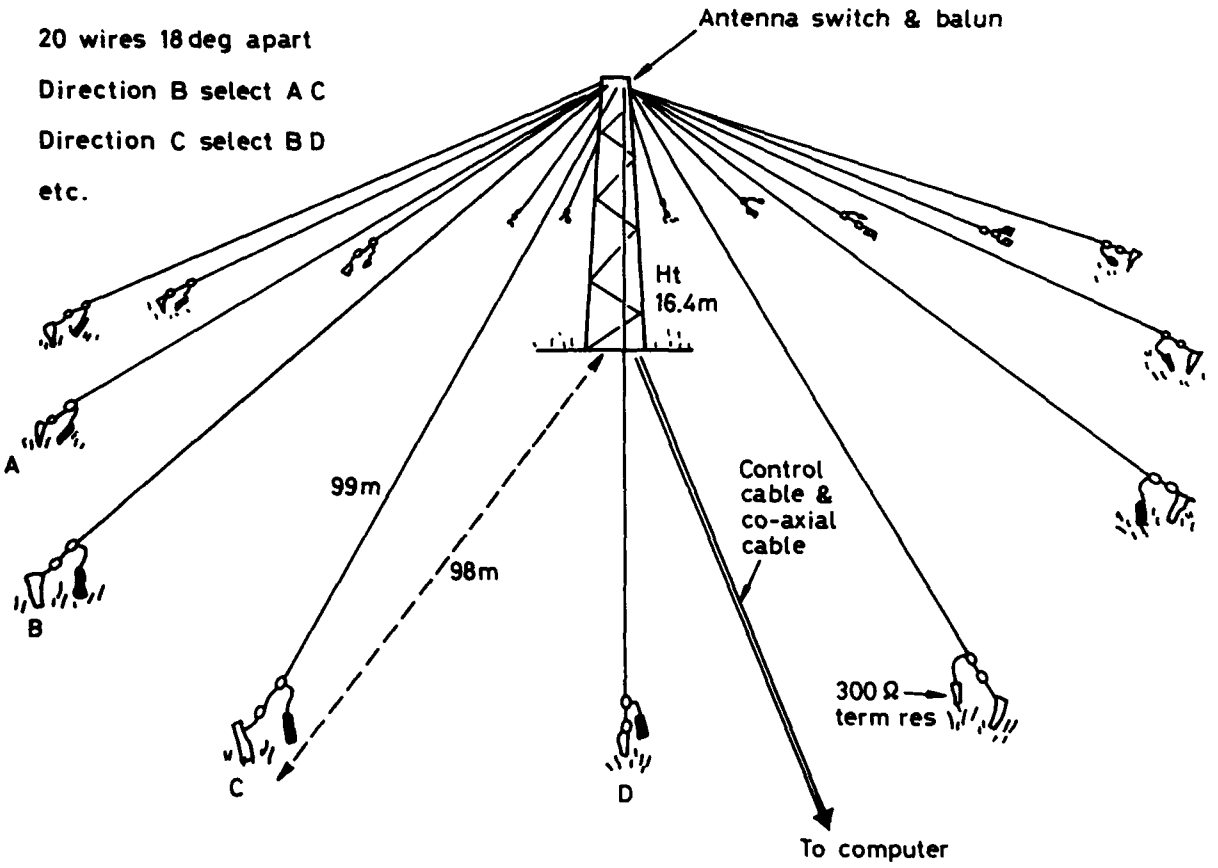


Fig 2 Sketch of multi-Vee antenna

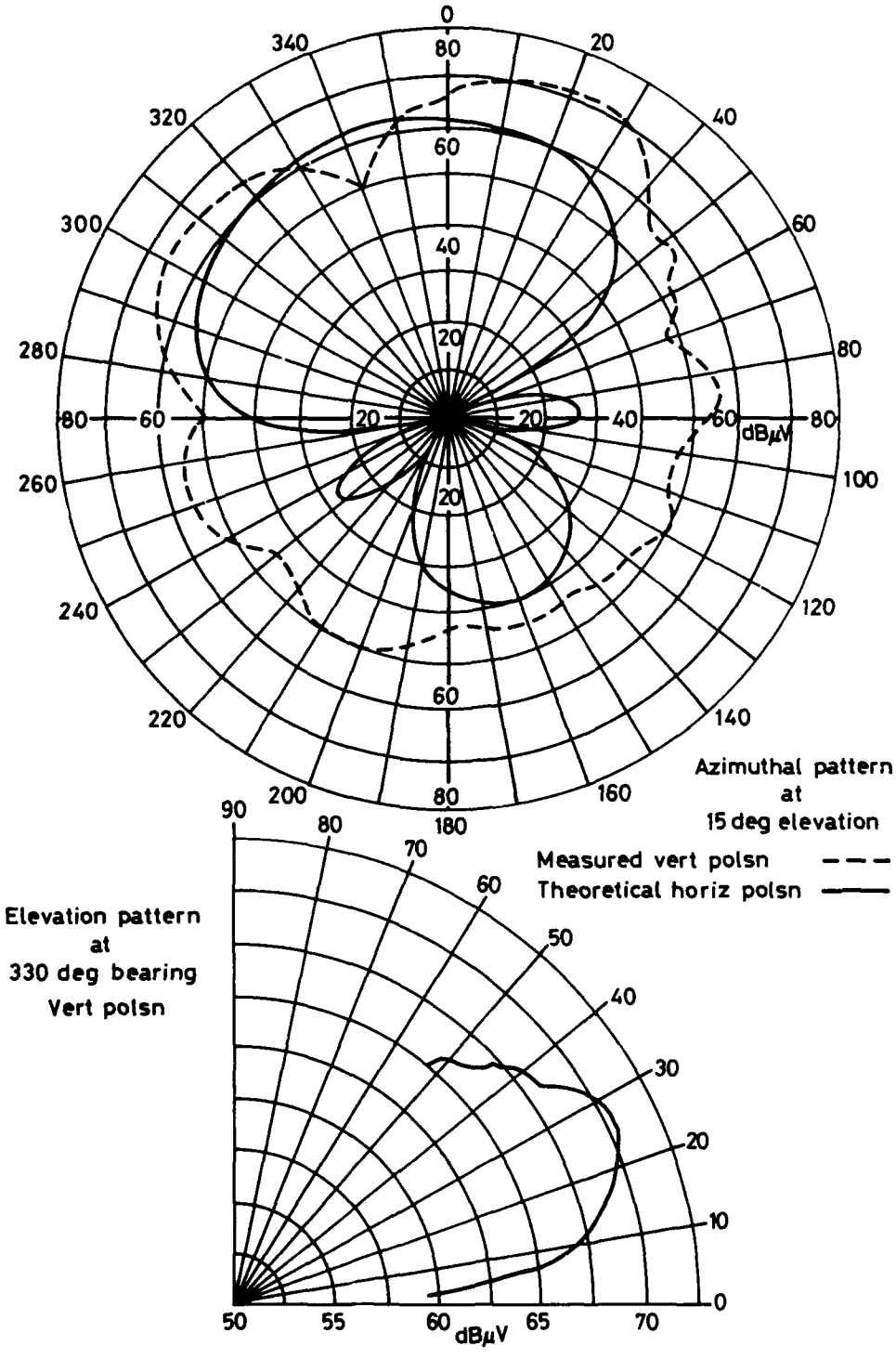


Fig 3 Multi-Vee antenna pattern at 3.0 MHz

14:03:15 17 October 1986

HF AZIMUTH OCCUPANCY MEASUREMENT SYSTEM *SCAN*

SCAN MODE: CYCLIC TRIAL No: 1

Disc ident	Xxxxxx	Scan ident (2 ch)	OC1786AI01
Start date	17 October 1986	Duration (h:mm)	0:10
Start time	18:00:00	Interval (d:hh:mm)	0:06:00
		No. of trials	11
First sector	1	Last sector	20
Reception mode	USB (A3J)	Averaging period (s)	2
Reception bandwidth	2.4 kHz	Settling time (ms)	5
Reference aerial	WBD312	Storage media	WINCHESTER
Frequency mode	DISCRETE		
Frequency (MHz)	(1) 3.0 (2) 6.0 (3) 9.0 (4) 13.0 (5) 15.0		

Fig 4 Example of an input file

Date 18:10:86 Time	Freq	Scan ident: OC1786AI05 Sectors dBμV																				
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Ref
18:00:00	1	-8	-9	-9	-8	-9	-8	-9	-7	-10	-7	-7	-7	-5	-8	-11	-9	-9	-9	-10	-10	6
18:00:02	2	-10	-8	-8	-8	-8	-10	-8	-7	-11	-10	-10	-10	-9	-9	-7	-10	-10	-11	-10	-8	8
18:00:04	3	6	2	3	6	5	5	2	11	14	19	21	19	20	15	12	7	6	3	3	5	25
18:00:06	4	-2	-8	-3	-2	-8	-10	-8	-8	-10	-9	-4	0	4	6	-2	9	-5	2	-10	-10	6
18:00:08	5	-11	-11	-10	-10	-10	-10	-10	-11	-9	-10	-10	-10	-11	-10	-9	-10	-10	-9	-9	-10	-8
18:00:10	1	-8	-8	-10	-9	-8	-7	-6	-5	-6	-6	-10	-8	-8	-7	-11	-11	-10	-9	-10	-9	8
18:00:12	2	-10	-8	-8	-7	-8	-7	-9	-7	-4	-3	-7	-3	-6	-9	-9	-10	-9	-10	-9	-9	3
18:00:14	3	-3	0	-1	-1	-2	-1	2	5	10	19	26	21	21	15	6	1	3	1	5	6	25
18:00:16	4	-10	-9	-6	-8	-2	-10	-9	-10	-8	-10	-5	1	5	9	13	12	-4	2	-6	-5	9
18:00:18	5	-10	-12	-11	-11	-9	-10	-10	-8	-11	-9	-10	-10	-9	-10	-11	-10	-10	-10	-10	-10	-7
18:00:20	1	-9	-9	-9	-8	-7	4	-9	-8	-10	-9	-10	-11	-9	-9	-11	-11	-9	-10	-9	-10	3
18:00:22	2	-9	-9	-4	-5	-7	-7	-7	-9	-2	-8	-8	-4	-2	-8	-12	-10	-6	-7	-8	-8	6
18:00:24	3	-1	3	1	2	-3	-2	0	1	12	16	20	10	16	16	7	2	4	1	2	5	19
18:00:26	4	-5	-10	-9	-9	-10	-5	-8	-8	-9	-6	-7	-9	-3	5	13	8	10	9	-9	-8	14

Fig 5 Data displayed in tabular form

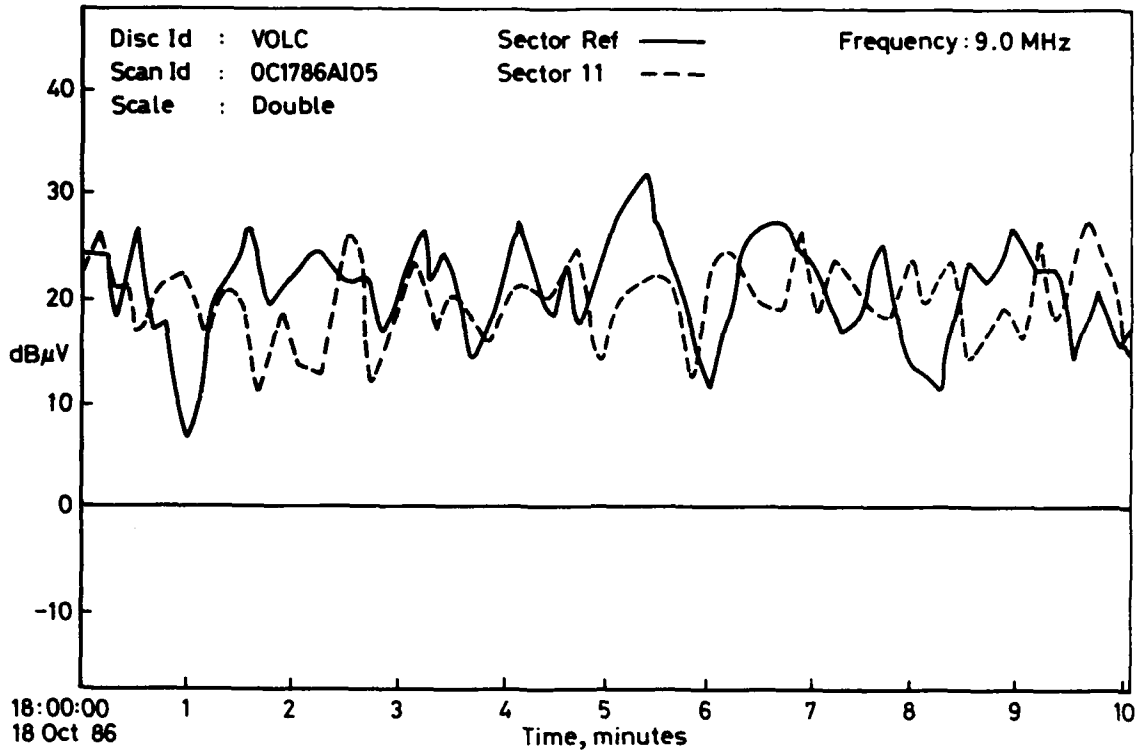


Fig 6 Signal versus time display

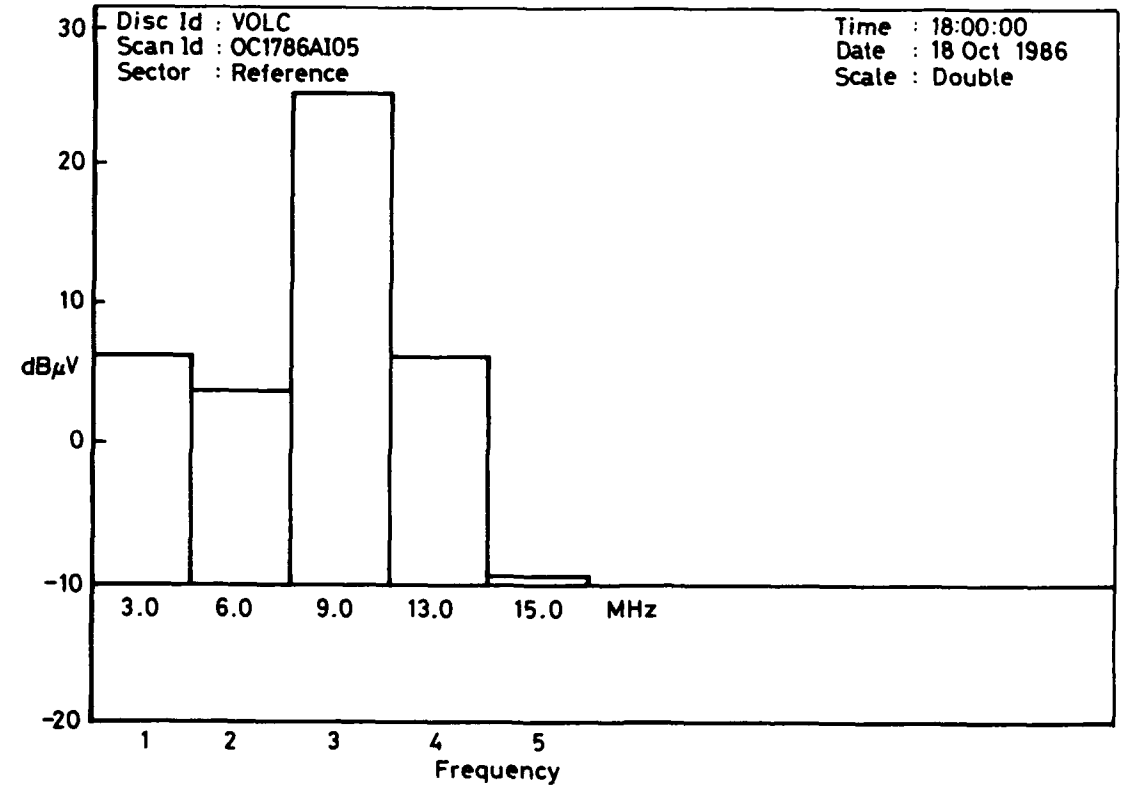


Fig 7 Signal versus frequency display (reference)

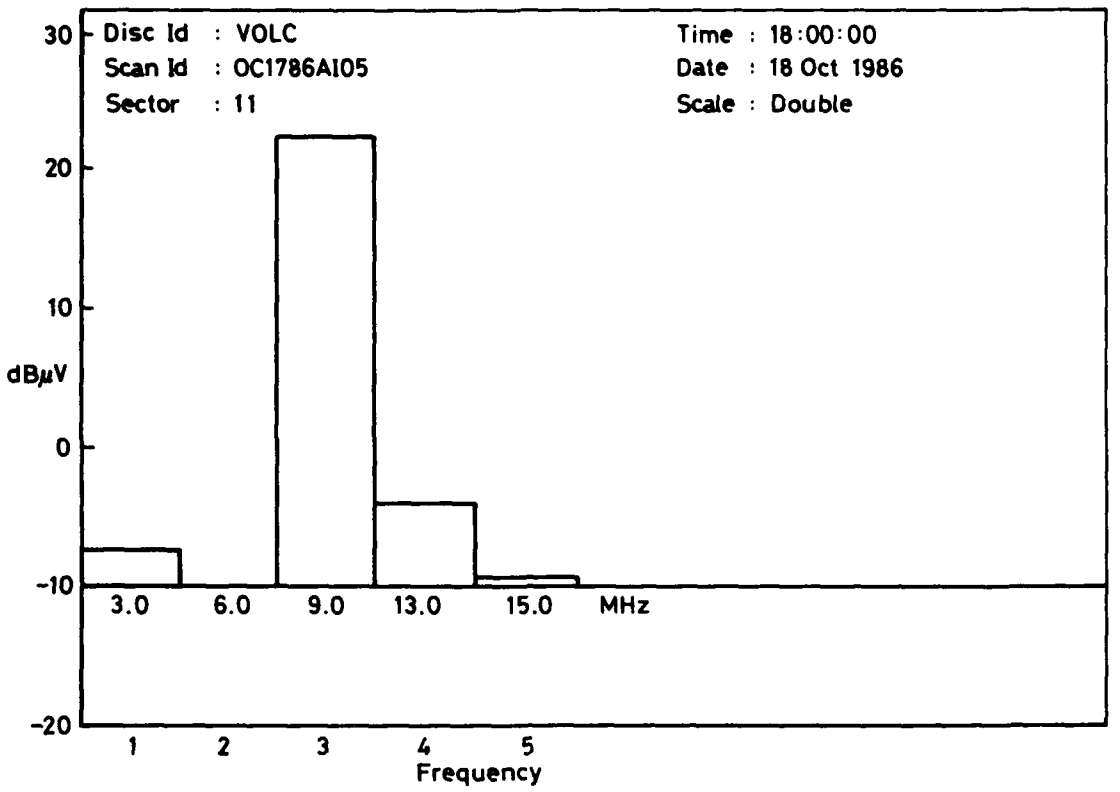


Fig 8 Signal versus frequency display (Vee sector 11)

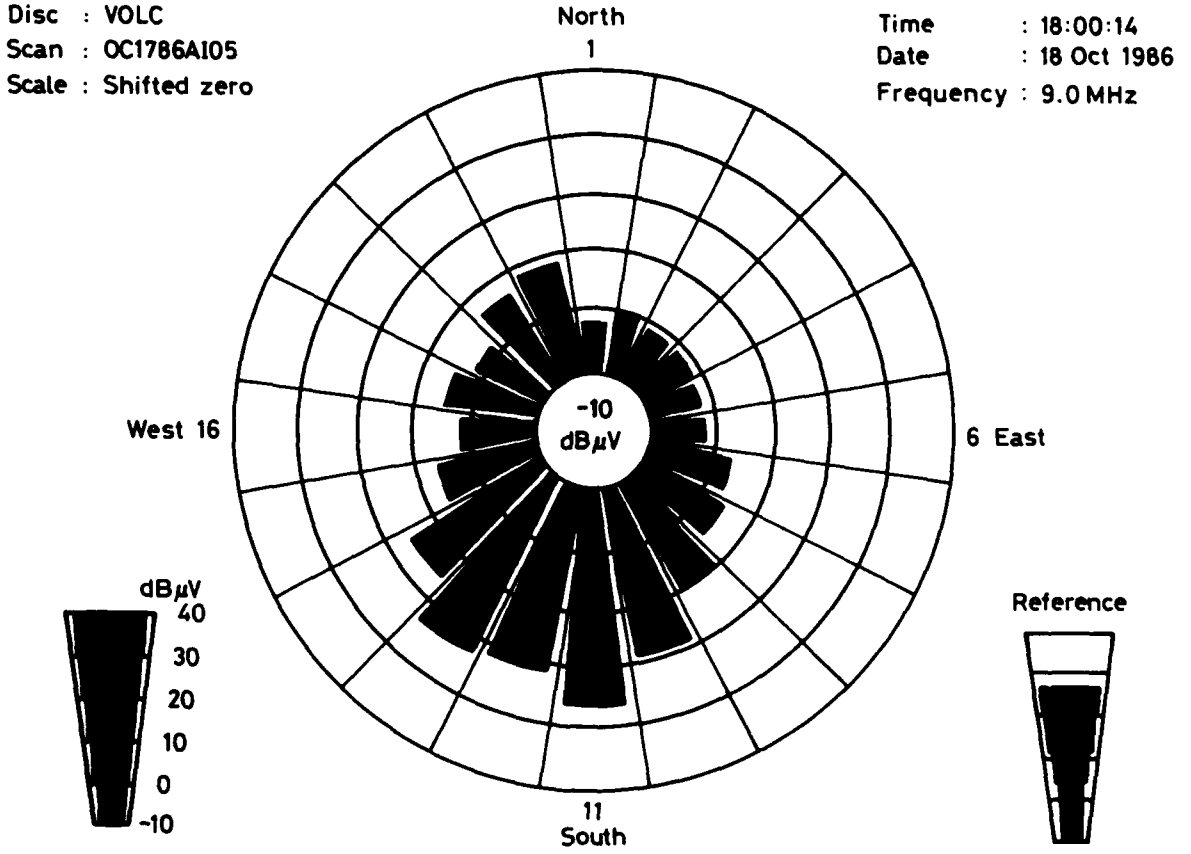


Fig 9 Signal versus direction display

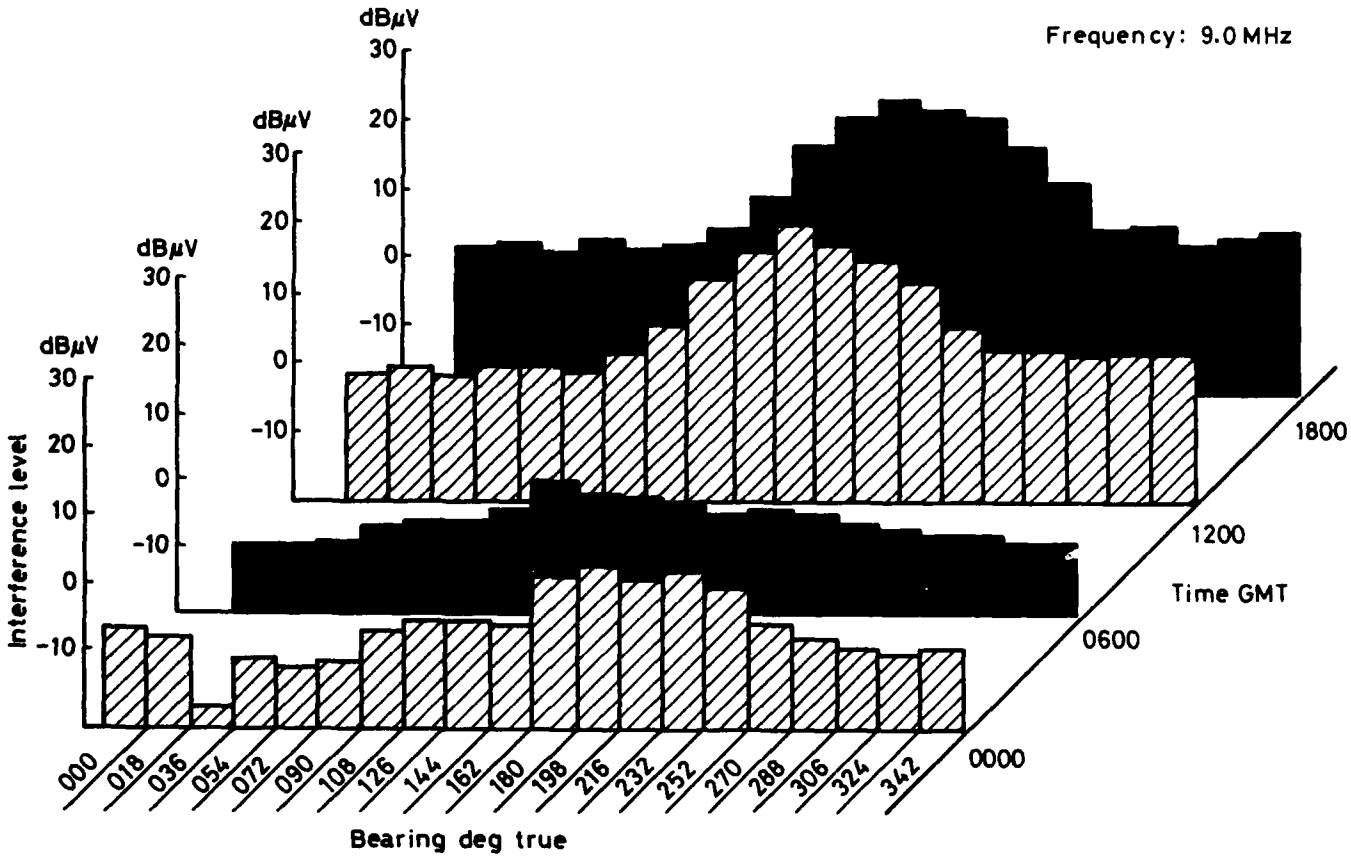


Fig 10a Interference versus bearing, 18 October 1986

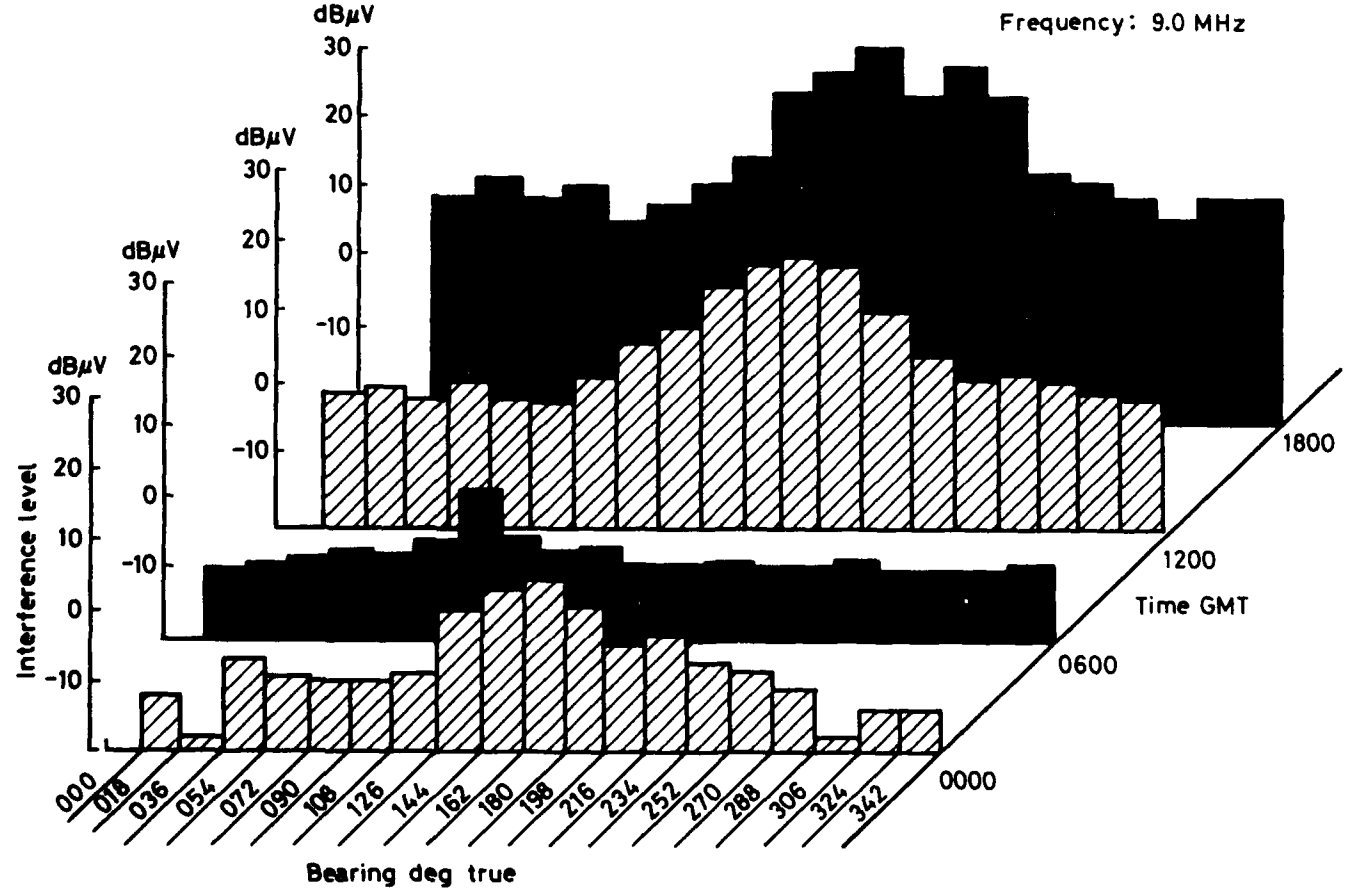


Fig 10b Interference versus bearing, 19 October 1986

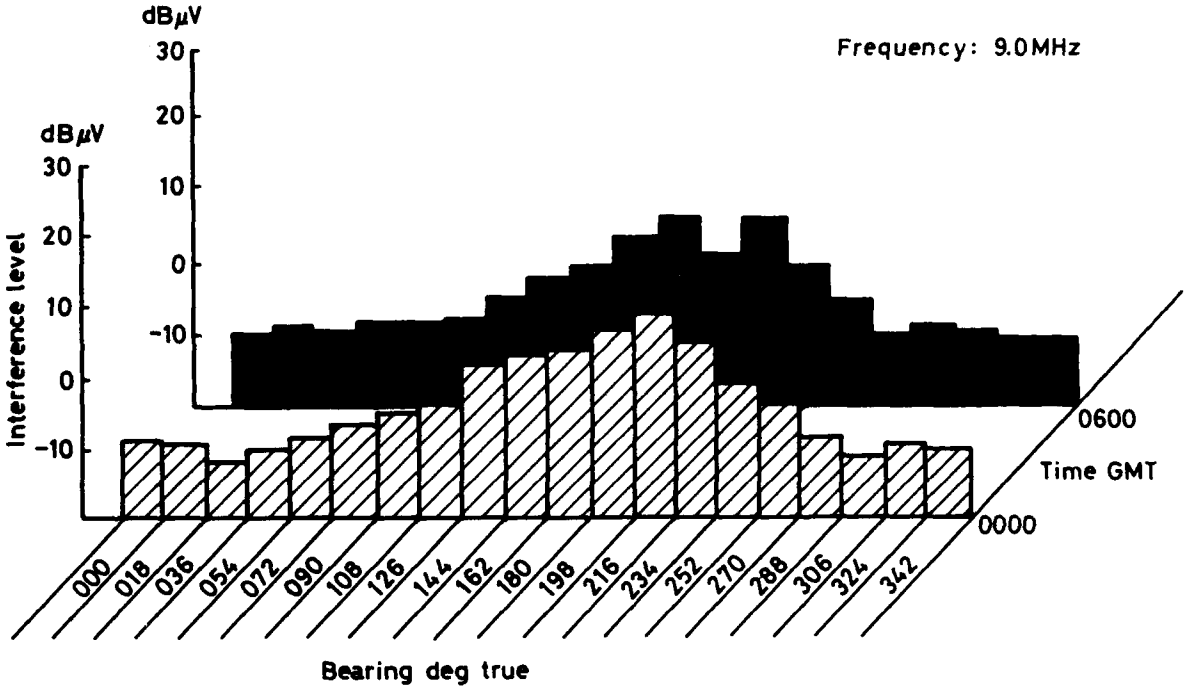


Fig 10c Interference versus bearing, 20 October 1986

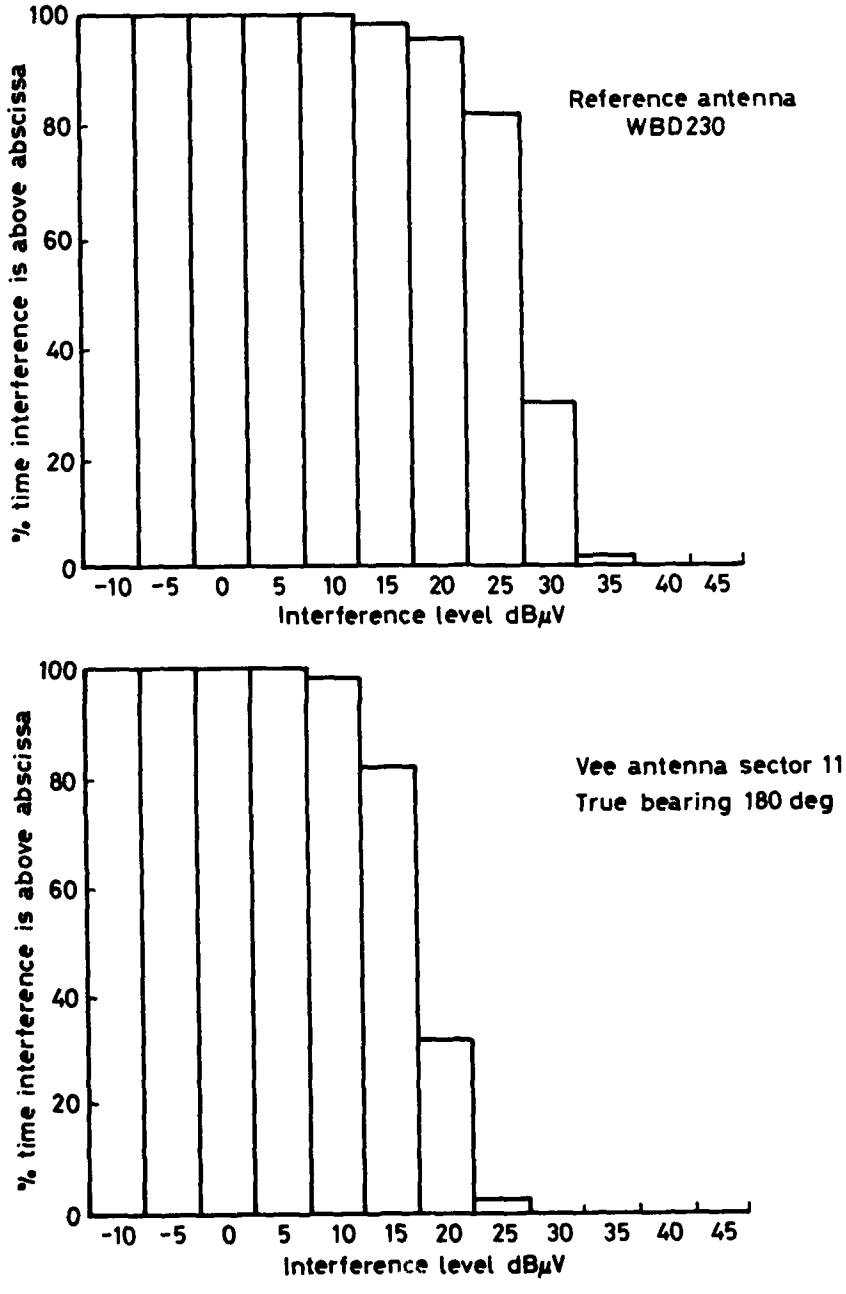


Fig 11 Cumulative distribution of interference
9.0 MHz. 1800 hours, 18 October 1986

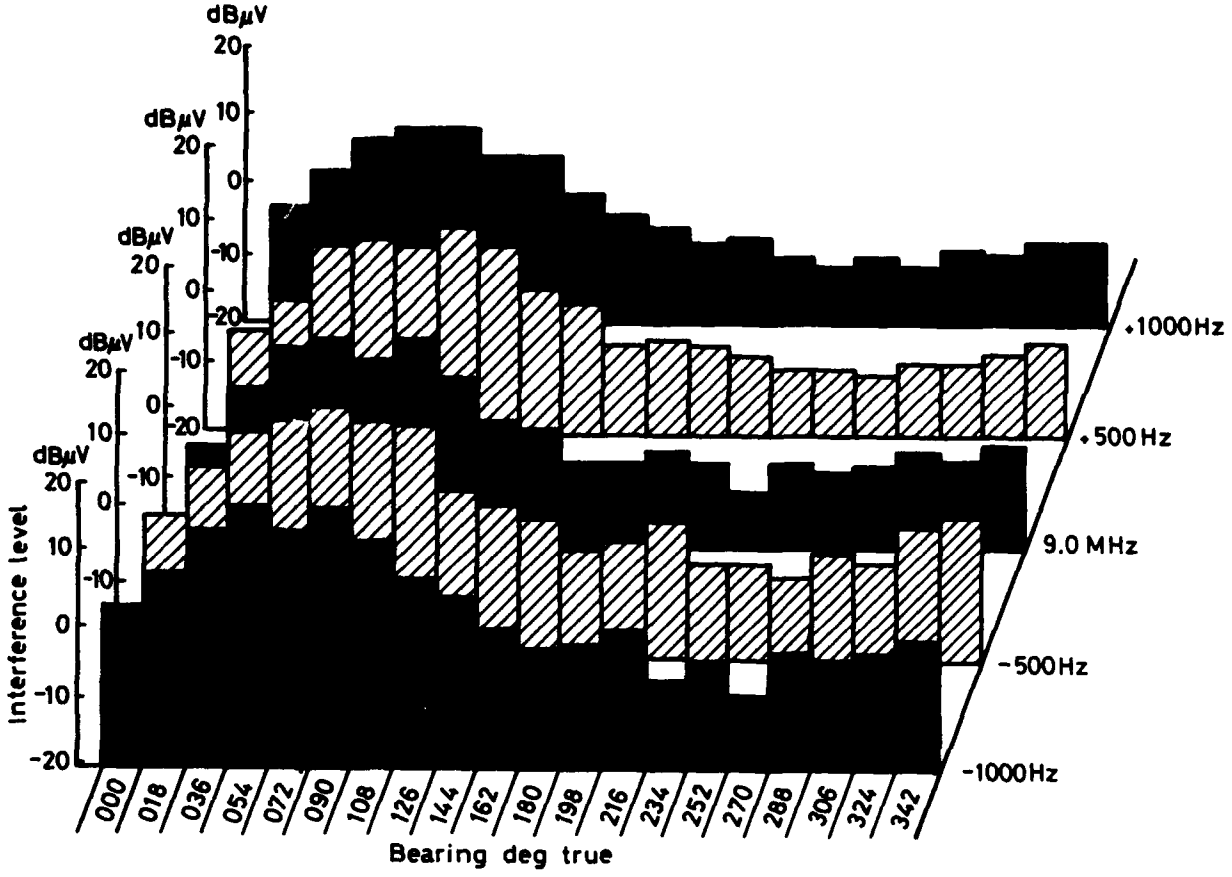


Fig 12a Interference versus bearing, 1 October 1986 at 2100 GMT

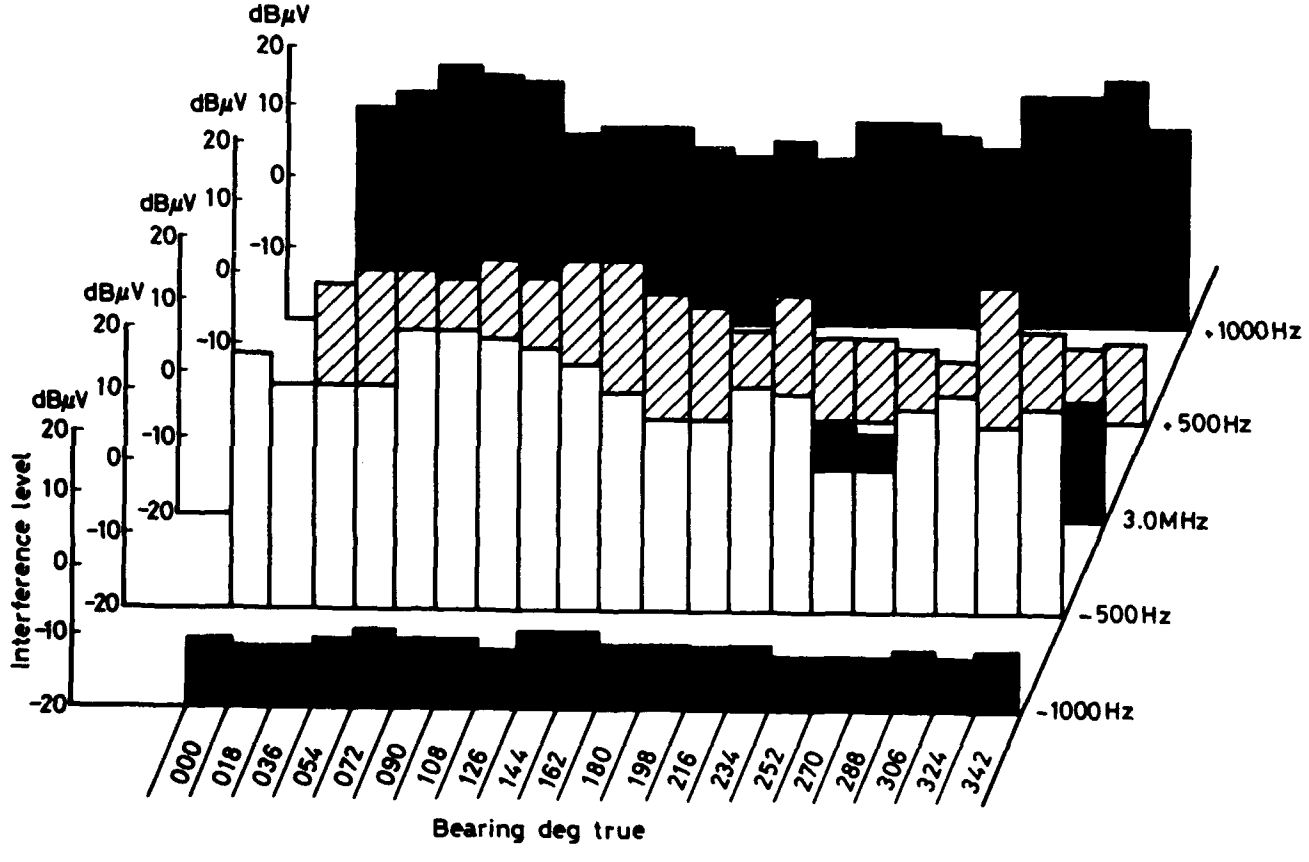


Fig 12b Interference versus bearing, 15 December 1986 at 2100 GMT

REPORT DOCUMENTATION PAGE

1. Recipient's Reference	2. Originator's Reference	3. Further Reference	4. Security Classification of Document
	AGARD-CP-420	ISBN 92-835-0487-9	UNCLASSIFIED

5. Originator Advisory Group for Aerospace Research and Development
North Atlantic Treaty Organization
7 rue Ancelle, 92200 Neuilly sur Seine, France

6. Title EFFECTS OF ELECTROMAGNETIC NOISE AND INTERFERENCE ON PERFORMANCE OF MILITARY RADIO COMMUNICATION SYSTEMS

7. Presented at the Electromagnetic Wave Propagation Panel Specialists' Meeting held in Lisbon, Portugal, 26—30 October 1987.

8. Author(s)/Editor(s)	9. Date
Various	December 1988
10. Author's/Editor's Address	**11. Pages**
Various	436

12. Distribution Statement This document is distributed in accordance with AGARD policies and regulations, which are outlined on the Outside Back Covers of all AGARD publications.

13. Keywords/Descriptors

Electromagnetic noise	Electromagnetic interference
Military communication	Simulation
Radio communication	Measurement

14. Abstract

To achieve reliable communications, a thorough knowledge of the communications channel is required. It is important for system design to characterize the channel, which in general consists of intentionally radiated signals, noise of either natural origin or radiated from man-made sources or composite variations of these signals. The various noise and interference sources existing in the communications channel impose a fundamental limit on achievable radio system performance. Communications equipment must be designed to cope with the effects of electromagnetic noise and interference.

The symposium examined how to determine the effects of these disturbances by measuring, modelling and particularly simulating the characteristics of unwanted signals.

Communication systems design and performance were also discussed since the application of radio communications systems on ships and aircraft, and the resulting mobility of the ground forces results in a very dynamic ever changing highly interactive environment. The effects of electromagnetic noise and interference on these systems performance and EMC is a most important subject for NATO.

AGARD Conference Proceedings No.420 Advisory Group for Aerospace Research and Development, NATO EFFECTS OF ELECTROMAGNETIC NOISE AND INTERFERENCE ON PERFORMANCE OF MILITARY RADIO COMMUNICATION SYSTEMS Published December 1988 436 pages To achieve reliable communications, a thorough knowledge of the communications channel is required. It is important for system design to characterize the channel, which in general consists of intentionally radiated signals, noise of either natural origin or radiated from man-made sources or composite variations of these signals. The various noise and interference sources existing in the P.T.O.	AGARD-CP-420 Electromagnetic noise Military communication Radio communication Electromagnetic interference Simulation Measurement	AGARD Conference Proceedings No.420 Advisory Group for Aerospace Research and Development, NATO EFFECTS OF ELECTROMAGNETIC NOISE AND INTERFERENCE ON PERFORMANCE OF MILITARY RADIO COMMUNICATION SYSTEMS Published December 1988 436 pages To achieve reliable communications, a thorough knowledge of the communications channel is required. It is important for system design to characterize the channel, which in general consists of intentionally radiated signals, noise of either natural origin or radiated from man-made sources or composite variations of these signals. The various noise and interference sources existing in the P.T.O.	AGARD-CP-420 Electromagnetic noise Military communication Radio communication Electromagnetic interference Simulation Measurement
AGARD Conference Proceedings No.420 Advisory Group for Aerospace Research and Development, NATO EFFECTS OF ELECTROMAGNETIC NOISE AND INTERFERENCE ON PERFORMANCE OF MILITARY RADIO COMMUNICATION SYSTEMS Published December 1988 436 pages To achieve reliable communications, a thorough knowledge of the communications channel is required. It is important for system design to characterize the channel, which in general consists of intentionally radiated signals, noise of either natural origin or radiated from man-made sources or composite variations of these signals. The various noise and interference sources existing in the P.T.O.	AGARD-CP-420 Electromagnetic noise Military communication Radio communication Electromagnetic interference Simulation Measurement	AGARD Conference Proceedings No.420 Advisory Group for Aerospace Research and Development, NATO EFFECTS OF ELECTROMAGNETIC NOISE AND INTERFERENCE ON PERFORMANCE OF MILITARY RADIO COMMUNICATION SYSTEMS Published December 1988 436 pages To achieve reliable communications, a thorough knowledge of the communications channel is required. It is important for system design to characterize the channel, which in general consists of intentionally radiated signals, noise of either natural origin or radiated from man-made sources or composite variations of these signals. The various noise and interference sources existing in the P.T.O.	AGARD-CP-420 Electromagnetic noise Military communication Radio communication Electromagnetic interference Simulation Measurement

communications channel impose a fundamental limit on achievable radio system performance. Communications equipment must be designed to cope with the effects of electromagnetic noise and interference.

The symposium examined how to determine the effects of these disturbances by measuring, modelling and particularly simulating the characteristics of unwanted signals.

Communication systems design and performance were also discussed since the application of radio communications systems on ships and aircraft, and the resulting mobility of the ground forces results in a very dynamic ever changing highly interactive environment. The effects of electromagnetic noise and interference on these systems performance and EMC is a most important subject for NATO.

Papers presented at the Electromagnetic Wave Propagation Panel Specialists' Meeting held in Lisbon, Portugal, 26—30 October 1987.

ISBN 92-835-0487-9

communications channel impose a fundamental limit on achievable radio system performance. Communications equipment must be designed to cope with the effects of electromagnetic noise and interference.

The symposium examined how to determine the effects of these disturbances by measuring, modelling and particularly simulating the characteristics of unwanted signals.

Communication systems design and performance were also discussed since the application of radio communications systems on ships and aircraft, and the resulting mobility of the ground forces results in a very dynamic ever changing highly interactive environment. The effects of electromagnetic noise and interference on these systems performance and EMC is a most important subject for NATO.

Papers presented at the Electromagnetic Wave Propagation Panel Specialists' Meeting held in Lisbon, Portugal, 26—30 October 1987.

ISBN 92-835-0487-9

communications channel impose a fundamental limit on achievable radio system performance. Communications equipment must be designed to cope with the effects of electromagnetic noise and interference.

The symposium examined how to determine the effects of these disturbances by measuring, modelling and particularly simulating the characteristics of unwanted signals.

Communication systems design and performance were also discussed since the application of radio communications systems on ships and aircraft, and the resulting mobility of the ground forces results in a very dynamic ever changing highly interactive environment. The effects of electromagnetic noise and interference on these systems performance and EMC is a most important subject for NATO.

Papers presented at the Electromagnetic Wave Propagation Panel Specialists' Meeting held in Lisbon, Portugal, 26—30 October 1987.

ISBN 92-835-0487-9

communications channel impose a fundamental limit on achievable radio system performance. Communications equipment must be designed to cope with the effects of electromagnetic noise and interference.

The symposium examined how to determine the effects of these disturbances by measuring, modelling and particularly simulating the characteristics of unwanted signals.

Communication systems design and performance were also discussed since the application of radio communications systems on ships and aircraft, and the resulting mobility of the ground forces results in a very dynamic ever changing highly interactive environment. The effects of electromagnetic noise and interference on these systems performance and EMC is a most important subject for NATO.

Papers presented at the Electromagnetic Wave Propagation Panel Specialists' Meeting held in Lisbon, Portugal, 26—30 October 1987.

ISBN 92-835-0487-9