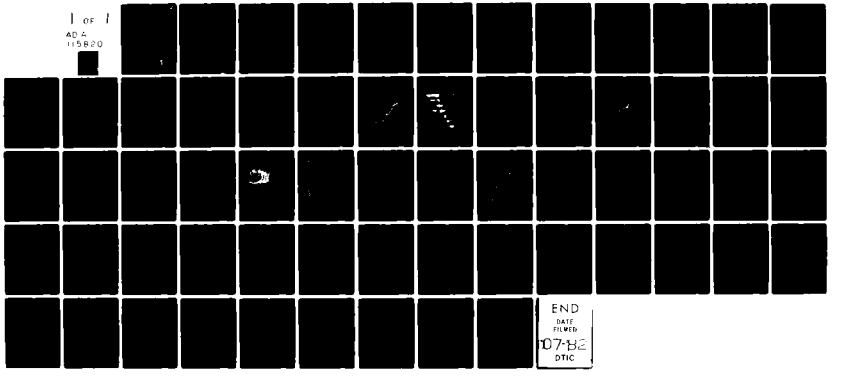
AD-A115 820 LOWELL UNIV MA CENTER FOR ATMOSPHERIC RESEARCH F/G 4/1
AUTOMATIC PROCESSING OF DIGITAL IONOGRAMS AND FULL WAVE SOLUTIO--ETC(U)
NOV 81 B W REINISCH, H E MOSES, J S TANG F19628-80-C-0113
UNCLASSIFIED ULRF-417/CAR AFGL-TR-82-0014 NL



AFGL-TR-82-0014

AUTOMATIC PROCESSING OF DIGITAL IONOGRAMS AND FULL WAVE SOLUTIONS FOR THE PROFILE INVERSION PROBLEM

B. W. Reinisch
H. E. Moses
J. S. Tang

University of Lowell
Center for Atmospheric Research
450 Aiken Street
Lowell, Massachusetts 01854

November 1981

Final Report

July 7, 1980 to August 7, 1981

AIR FORCE GEOPHYSICS LABORATORY
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
HANSCOM AFB, MASSACHUSETTS 01731

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFGL-TR-82-0014	2. GOVT ACCESSION NO. AD-A115 820	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) AUTOMATIC PROCESSING OF DIGITAL IONOGRAMS AND FULL WAVE SOLUTIONS FOR THE PROFILE INVERSION PROBLEM		5. TYPE OF REPORT & PERIOD COVERED Final 7 Jul 1980-7 Aug 1981 6. PERFORMING ORG. REPORT NUMBER ULRF-417/CAR
7. AUTHOR(s) B. W. Reinisch H. E. Moses J. S. Tang		8. CONTRACT OR GRANT NUMBER(s) F19628-80-C-0113
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Lowell, Center for Atmospheric Research, 450 Aiken Street, Lowell, Massachusetts 01854		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 62101F 464306AH
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Geophysics Laboratory Hanscom AFB, Massachusetts 01731 Contract Monitor: William N. Hall/PHY		12. REPORT DATE November 1981 13. NUMBER OF PAGES 61
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Digital Ionograms
Autoscaling of Ionograms
Spread F
Full Wave Solution
WKB Solution
Inverse Scattering
Gelfand-Levitan Algorithm

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Methods for the automatic scaling of Digisonde ionograms are described. Making use of the available ionogram information, i.e. signal amplitude, polarization and incidence angle, the bottomside ionogram scaling algorithm extracts the ordinary echo trace with good accuracy even under very disturbed ionospheric conditions. The automated parameters for some 600 ionograms from Goose Bay, Labrador, were compared with the manually scaled values. For 82% of all ionograms foF2

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. Abstract

was scaled within 0.5 MHz, and for 93% within 1 MHz in spite of the frequent occurrence of heavy spread F.

New full wave methods for the profile inversion are explored in the second part of this report. Special profiles are investigated for which the full-wave equations can be solved exactly.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

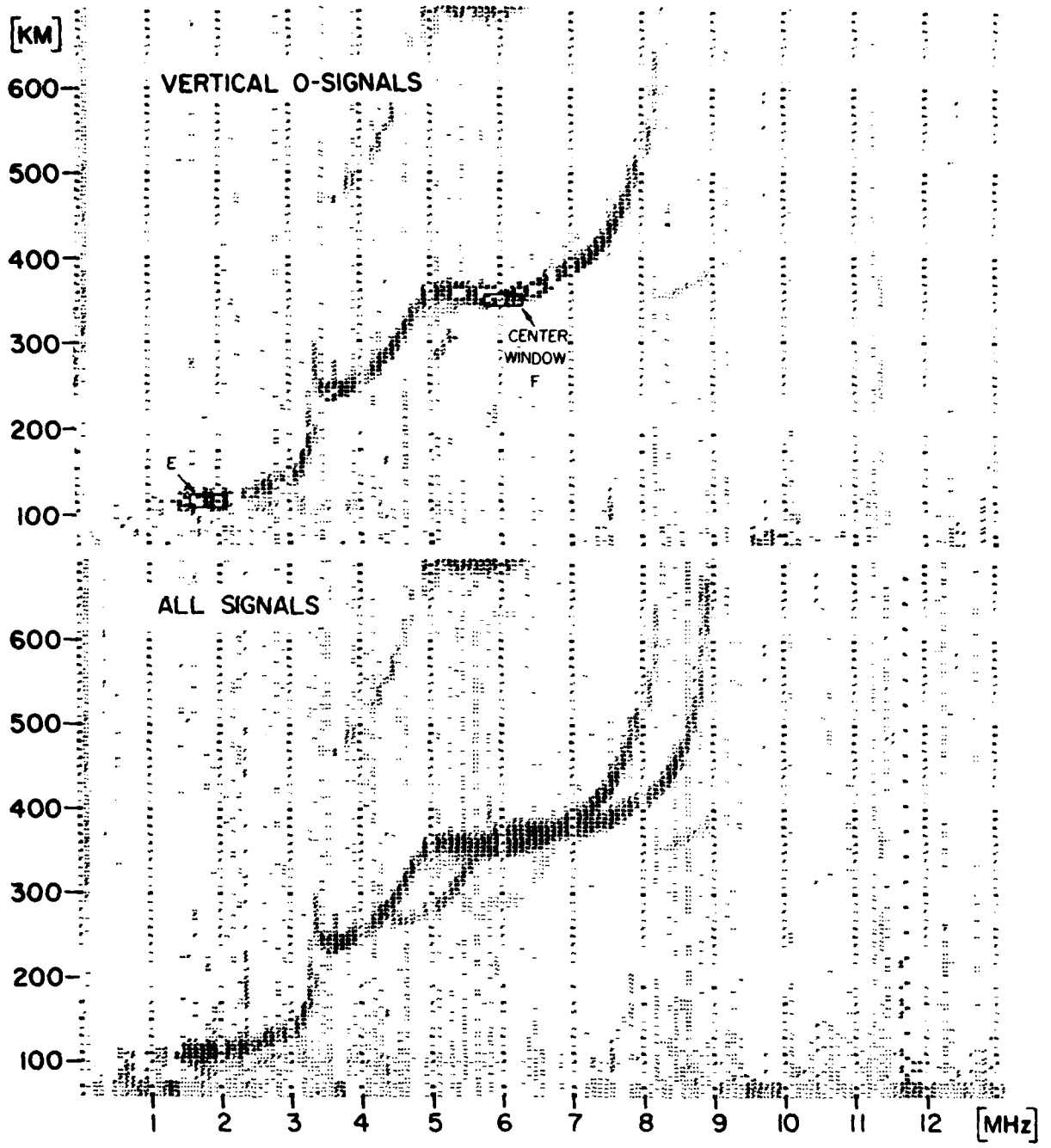
1.0 INTRODUCTION

Ionograms contain as basic information the virtual range versus frequency. The more advanced digital ionosondes also measure amplitude, phase, polarization, incidence angle and spectral signature of the echoes. This allows improved methods for the conversion of the ionogram to vertical electron density profiles. This report discusses two different efforts toward that goal. Section 2 describes the automatic scaling of Digisonde ionograms, and Section 3 investigates the possibility of solving the Schroedinger wave equation for different potential functions in an effort to compare the WKB method with the full wave solution.

2.0 AUTOMATIC PROCESSING OF DIGITAL IONOGRAMS

Automatic processing of ionograms has become possible with the availability of advanced digital sounders [Bibl and Reinisch, 1978]. Since about 1970, when the first digital ionograms were routinely recorded on magnetic tape, software and hardware techniques for the automatic extraction of the echo traces from the ionograms [Bibl et al, 1973] have been developed. The initial concept was to decide immediately after transmission of each individual frequency which of the received signals are echoes and to retain only the amplitudes and heights of the identified echoes. This approach is clearly inadequate for many scientific investigations in a disturbed ionosphere where the researchers want to see the complete range-versus-frequency display, i.e. the raw ionogram with all the nuances in the signal characteristics. Bibl and Reinisch [1978] described the on- or off-line printing of Digisonde ionograms using an optically weighted font (Optifont) to retain the digital resolution in the quasi-analog ionogram display (Figure 1), establishing the transition from the familiar analog to the digital ionograms. It may suffice to record on magnetic tape a limited number of "identified" echo points per frequency for the archiving of routine ionograms or for the monitoring of ionospheric trends [Buchau et al, 1978]. In areas with frequent occurrence of spread F, this approach is ineffective unless the range spread for each frequency is also recorded.

A different objective is the automatic calculation of the vertical electron density profile. In that case, "the" vertical echo trace must be extracted and oblique, ducted and multiple echoes disregarded; a trace must be found within the spread F signals, relying on amplitude, Doppler and incidence angle information that are contained in the digital ionograms. This is only possible by examining the ionogram in its entirety. This report describes this approach for the auto-



UNDISTURBED DAY IONOGRAM
GOOSE BAY 16 JUN 1980 17:20 AST

Figure 1

matic scaling of Digisonde ionograms recorded at the Goose Bay Ionospheric Observatory (53.3N, 60.5W geographic, 64.6N, 12.1W geomagnetic). Once the ordinary vertical-echo trace is extracted from the ionogram, the electron density profile can be calculated using the standard lamination technique [Becker, 1967] or the profile-fitting method [Huang and Reinisch, 1982], the latter optimally suited for the case of automated data. This method will be described in a later paper. The current report discusses the accuracy of the autoscaling of the F-trace during disturbed ionospheric conditions, e.g. spread F and the mid-latitude trough. Since the technique for the autoscaling of the E-trace is not quite complete, we will not describe it here.

2.1 The Ionogram Data

The routine ionograms at Goose Bay measure and record amplitude, polarization, incidence angle and Doppler frequency as a function of frequency and range. The lower part of Figure 1 shows a quiet amplitude ionogram containing all signals. Removal of the non-vertical and X-polarization signals results in the upper ionogram of Figure 1, which is much easier to scale automatically. As will be shown later, the X-trace data are not discarded; they are used for the accurate determination of foF2. For bottomside ionograms, the O-trace is generally better presented than the X-trace and our autoscaling effort concentrated therefore on the O-trace. This is in contrast to topside ionograms where O and X-polarization echoes dominate alternatingly [Reinisch and Huang, 1982; see also the Special Issue on Topside Sounding and the Ionosphere of the Proceedings of the IEEE, Vol. 57, No. 6, June 1969].

Ideal ionograms like the one in Figure 1 pose no difficulties for automatic scaling, yet they are useful to illustrate the scaling procedure. To start the scaling of the F-trace, the ionogram for h' > 160 km is surveyed and the

"center window" is determined. The search for the complete O-trace starts at the center window, which is generally the best defined part of the trace and as such can be identified most reliably.

The tape recorded Digisonde ionograms are actually composite ionograms, since each pixel, i.e. frequency-range-bin, is accompanied by a status word [Bibl and Reinisch, 1978] which gives the signal polarization (O or X), the Doppler frequency (four spectral lines), and the approximate incidence angles (three directions). The current report discusses the off-line processing of ionograms for which only the data stored on tape are available. The processing algorithm is tailored toward future on-line application in which case a larger data set could be used, like separate O and X ionograms.

2.2 Bottomside Ionogram Scaling Algorithm-BISA

Efficient software was developed that determines the O-echo trace h'(f) in about five seconds of CPU time on the University of Lowell Cyber 71 computer or the AFGL CDC 6600 machine. If the F-region scaling and the electron density profile inversion routine are added, a total of about ten seconds of CPU time is required.

Figures 2a and 2b show the flow chart and block diagram for the main BISA program. The ionogram data are entered from magnetic tape and unpacked in dual arrays of 128 amplitudes and 128 status words for each ionogram frequency. Each functional unit of the processing sequence is briefly described in the following paragraphs.

Noise Elimination

For each sounding frequency an individual noise threshold is determined since it is not white noise but man-made interferers that generally obscure the ionospheric echoes. The typical signature of interference is a fairly constant

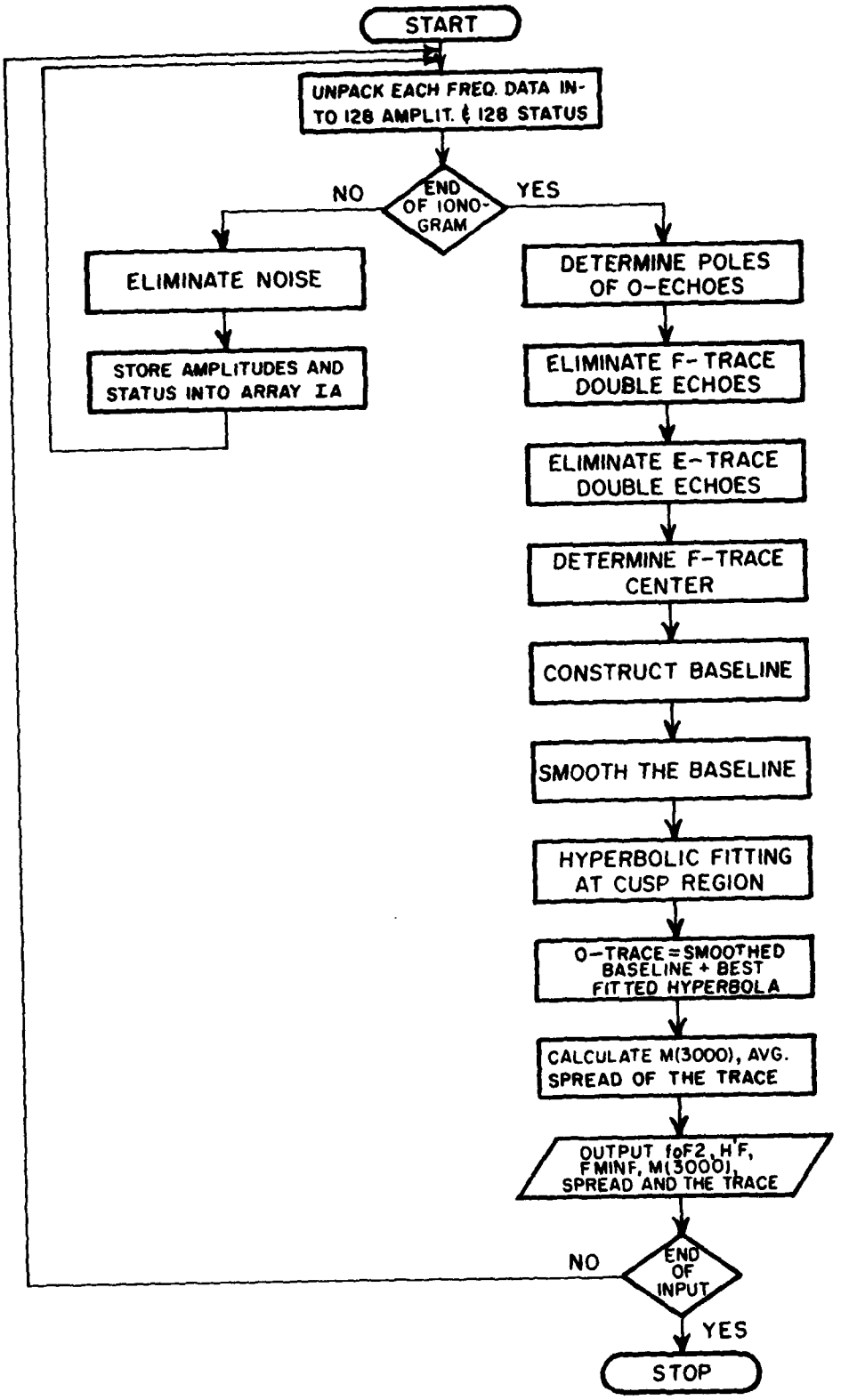


Figure 2a. BISA Flow Chart

BLOCK DIAGRAM FOR THE EXTRACTION OF THE F-TRACE

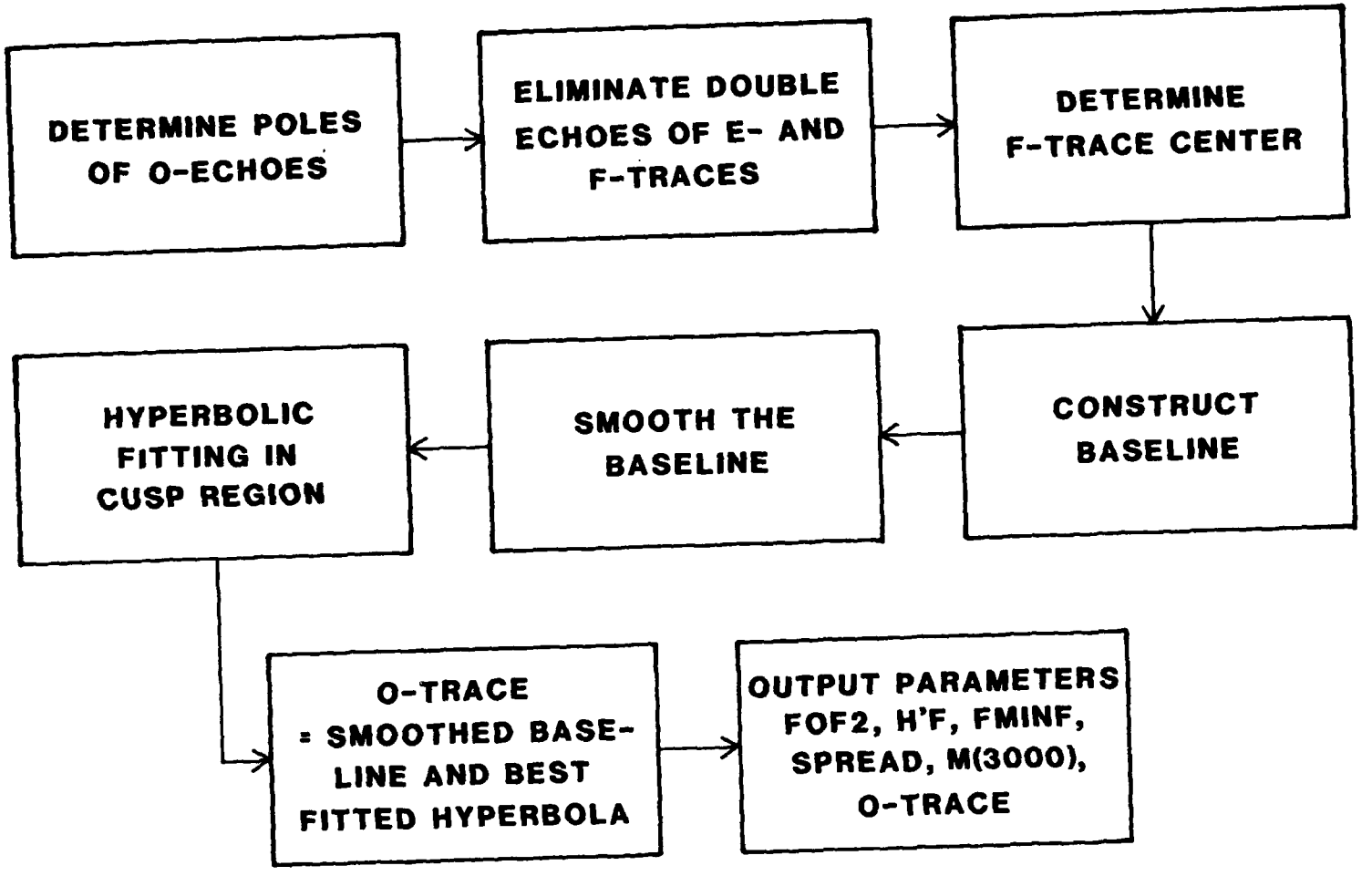


Figure 2b

distribution of amplitudes over the range bins, while the echo amplitudes generally occupy only two or three range bins. The most probable amplitude found over all range bins for a given frequency is used to determine the noise threshold. The half-point above the most probable amplitude on the amplitude probability curve is defined as signal threshold. During severe spread F conditions the most probable amplitude could be the signal level, setting too high a threshold. To avoid this problem two probability curves are formed, one for the lower 64 and one for the higher 64 range bins. The distribution that peaks at a lower amplitude sets the threshold [Reinisch and Smith, 1976].

O-Echo Poles

In an effort to minimize the required CPU time the program uses for certain tasks selected data points, called the poles. For each frequency, the position of the first amplitude maximum of an O-echo between two zero amplitudes is defined as pole. The number of poles per frequency is not limited.

Elimination of E and F Region Double Echoes

Before beginning the trace identification, all data points that are double echoes are eliminated. An F-region pole at twice the range (±20 km) of a lower pole is checked by comparing the average amplitudes over nine range bins of the lower and upper echoes. If the upper average amplitude is smaller the pole and all poles above it are removed.

The same method is applied to E-region echoes, except that a 5-bin window is used. Elimination of double E or Es echoes is important since they could otherwise be mistaken for F echoes.

F-Trace Center

During disturbed ionospheric conditions, it is often difficult to identify the main overhead echo trace. At Goose

Bay this is especially true when the mid-latitude F-region trough moves over the station. The direction finding capability of the Digisonde identifies some of the oblique echoes as such, but not all of them. By finding the frequency-range window with maximum signal energy one can be reasonably sure that the trace identification starts at a good point on the main echo trace.

For each range the pole amplitudes for all frequencies from the first ionogram frequency to 10 MHz are added. Grouping three height ranges together, a search is made for the three largest peaks in this sum array. A 5-frequency by 3-range bin window is slid along each of the three maximum ranges, and the sum over the amplitudes of the 15 pixels in each window position is formed. The window position with the maximum average amplitude is designated as F-trace center.

Coarse Baseline

As a first approximation of the echo trace the baseline is constructed. A 5-frequency by 35-range bin window is placed over the trace center, and the median range h_m of all poles within this window is determined. The lowest poles in the interval $h_m - 5 \leq h \leq h_m + 20$ bins (bin spacing is 5 km) are selected as baseline points. The window is shifted one step toward higher frequency and the next baseline point is determined. In each window the range difference D between the highest and lowest pole is determined. If D is larger than 20 the upper height of the window is increased to $h_m + D$. If no pole is found for a frequency the range of the previous frequency is substituted. The process stops when no pole is found in five consecutive frequencies. It is assumed that the end of the trace is found and the baseline point is given the range 128.

Starting again at the center, the process is repeated in the direction of lower frequencies. If not terminated by the lack of proper poles, the process stops at f =

foE - 0.3 MHz. Predicted values for foE are tabulated as function of time of day and season.

Smoothing of Baseline

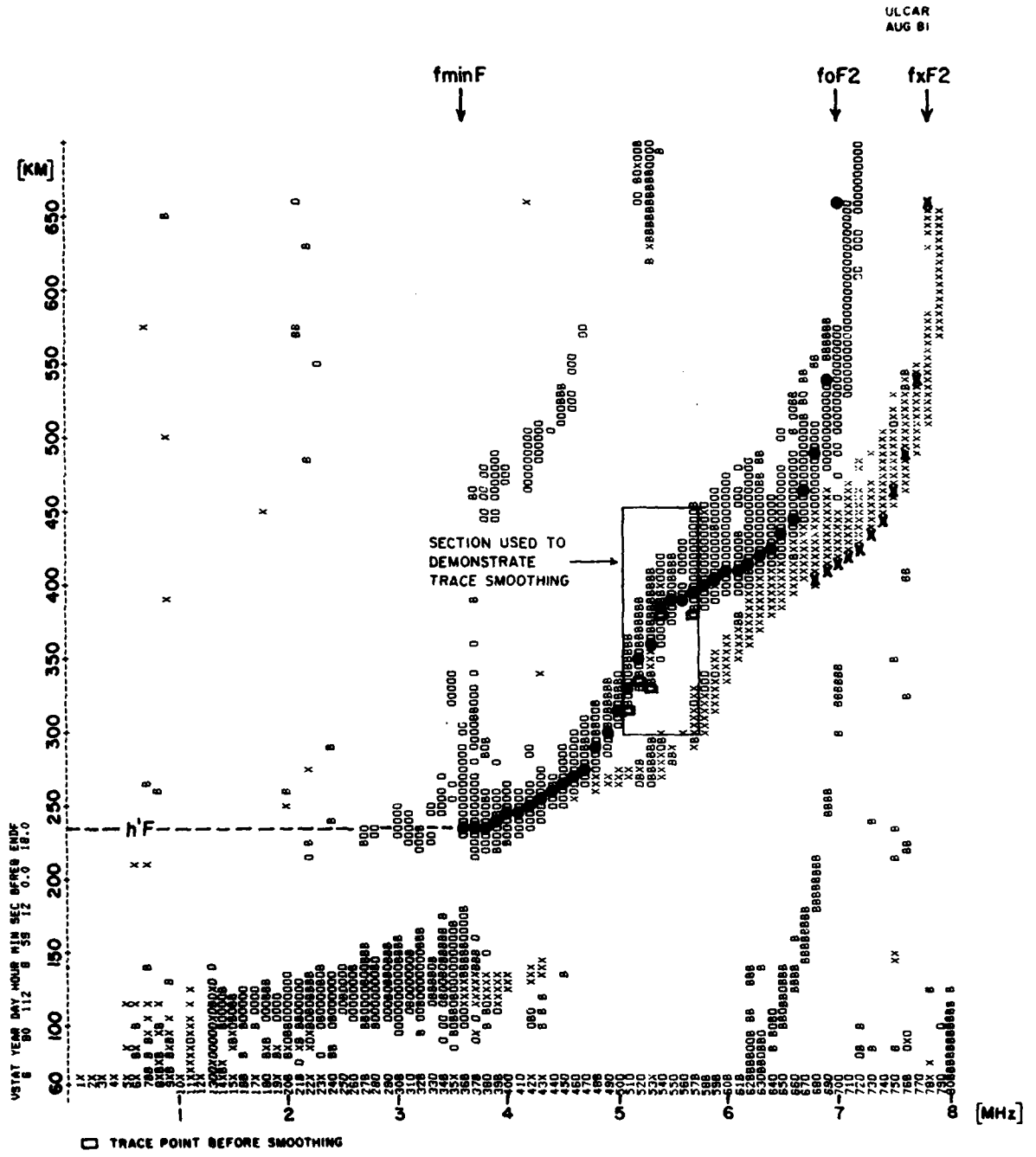
The sequence of heights for the baseline points shows some discontinuities that are caused by high interference on some frequencies, and blank pixels that were originally occupied by oblique or X-polarization echoes. A refined linear smoothing technique is applied to adjust the heights of the baseline points.

Consider the framed trace section in Figure 3. The O-trace points are marked by small circles. In the framed section the original baseline points are shown as small rectangles if they differ from the final trace points. Smoothing is performed over six frequencies, as illustrated in Figure 4. The shaded areas show the amplitudes of the O-echoes, the white areas the amplitude of oblique or X-echoes. The height for frequency 5.3 is to be found; the heights for f = 5.4, 5.5 and 5.6 MHz have already been adjusted. The smoothing starts at the trace center and proceeds to the lower and to the upper frequency end, respectively.

A straight line is fitted to any three of the six frequencies, resulting in a total of 20 lines. The six pixel amplitudes for each line are added and the line with the maximum amplitude determines the height of the frequency under investigation. Only three of the 20 possible lines are shown in Figure 4. The heavy line is the one with the largest amplitude sum setting the height for f = 5.3 MHz to 360 km, as marked by the X.

Determination of each line is based on the min-max approach. The equal-error or Chebyshev [Scheid, 1968, p. 269] line is determined which misses three given data points by equal amounts with alternating signs (see the deivations $-\Delta$, $+\Delta$ and $-\Delta$ in Figure 4). There exists exactly one such line

$$h = c + df \qquad (2.1)$$



GOOSE BAY JAN 9 1980 10:00 AST

Figure 3

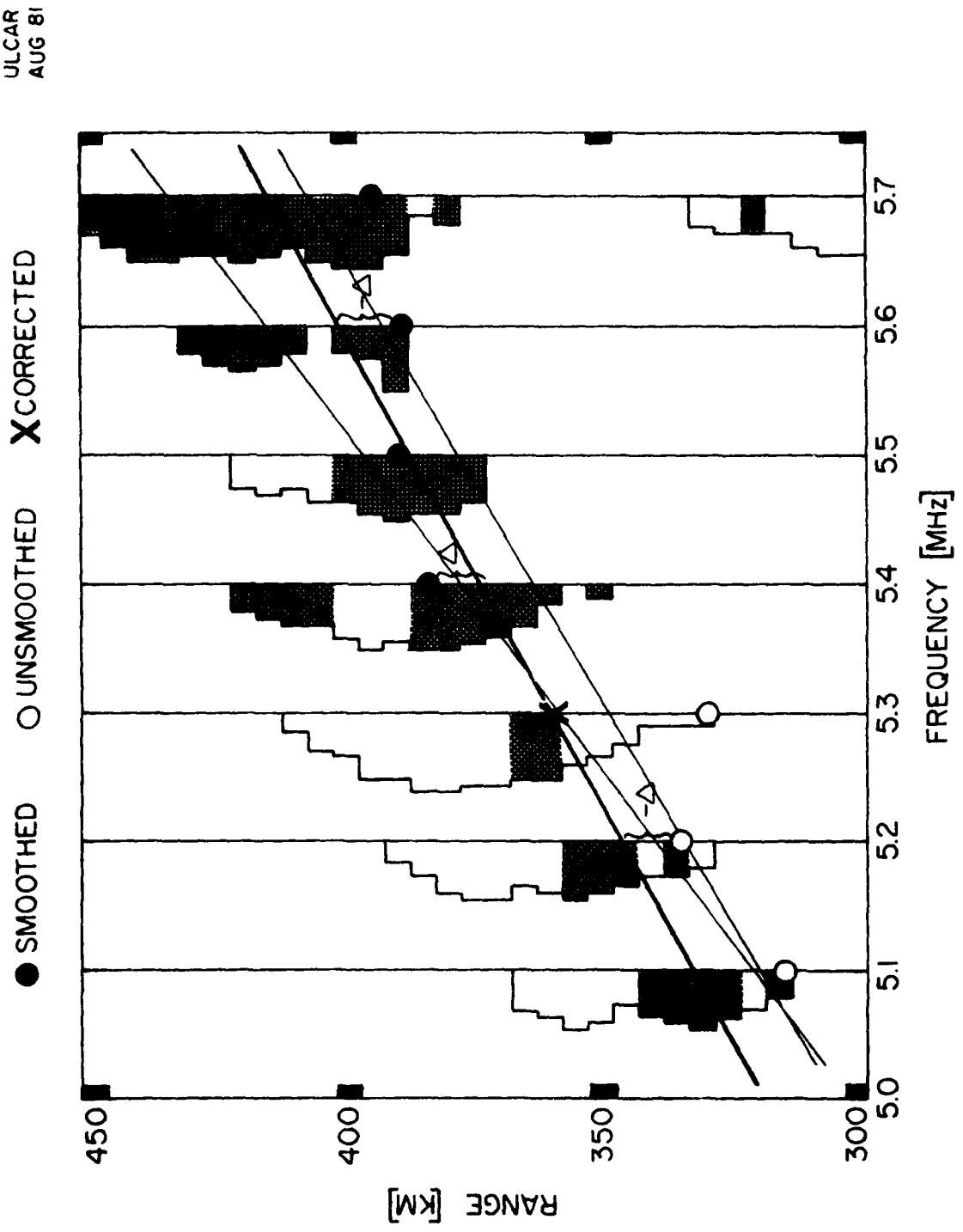

ULCAR
AUG 81
● SMOOTHED
○ UNSMOOTHED
X CORRECTED
RANGE [KM]
450
400
350
300
5.0
5.1
5.2
5.3
5.4
5.5
5.6
5.7
FREQUENCY [MHz]
TRACE SMOOTHING (LEFT TO THE TRACE CENTER)

Figure 4

which is defined by the following three equations:

$$c + df_1 - h_1 = \Delta \tag{2.2}$$

$$c + df_2 - h_2 = -\Delta \tag{2.3}$$

$$c + df_3 - h_3 = \Delta \tag{2.4}$$

The solutions for c, d and Δ are:

$$c = h_1 + \Delta - df_1 \tag{2.5}$$

$$d = \frac{h_3 - h_1}{f_3 - f_1} \tag{2.6}$$

$$\Delta = \frac{1}{2}\,[h_2 - h_1 - d(f_2 - f_1)]. \tag{2.7}$$

Hyperbolic Trace Fitting in Cusp Region

If the ionogram data have polarization tagging, some curve fitting can be performed in the h'-f domain. The cusp region of the ionogram trace can be described with reasonable accuracy by a polynomial or a rational function. A simple rational function like

$$h'_x = h_0 + \frac{1}{a + bf}, \quad b < 0 \tag{2.8}$$

has a singularity at f = - a/b which serves well to fit the trace near the critical frequency. The same function transposed to lower frequencies by half the gyrofrequency fH describes the corresponding O-trace:

$$h'_o = h_0 + \frac{1}{a + bf^*} \; ; \; f^* = f + \frac{1}{2}\, fH_{300} \tag{2.9}$$

A reference height of 300 km is used for the calculation of fH. Fitting simultaneously two curves to the cusp data makes the method very robust. Even if one of the traces is absent, the critical frequency is determined from the other trace. The hyperbolas with the maximum sum of amplitudes in narrow strips around them are selected to represent the O and X

traces near the critical frequencies. The optimum positioning of the hyperbolas is illustrated in Figure 5 for a Goose Bay ionogram with spread F. The range interval over which the amplitudes are summed is given for each frequency by the ranges of the preceding and the following frequency as shown in Figure 5. The Digisonde ionogram used in this figure does not show the signal amplitudes but instead the polarization tags O and X, and the letter B indicating oblique echoes. Only O-amplitudes contribute to the O-hyperbola's amplitude summation and X-amplitudes to the X-hyperbola. The hyperbolic traces for the cusp region that the program determined for the ionogram in Figure 5 are marked by heavy dots (O-trace) and X's (X-trace). The hyperbolic section of the trace ends where the X-hyperbola intersects the baseline (3.5 MHz), and the baseline then defines the trace toward lower frequencies.

Equations 2.8 and 2.9 have three coefficients h_0, a, and b, which are determined according to the method of min-max rational functions [Scheid, 1968, p. 289]. The baseline is used as a guide in finding the best fit. Starting at the high frequency end of the baseline, a hyperbola is fitted to the last three baseline points by minimizing the distance x from the hyperbola (see Figure 6). The hyperbola is determined in such a way that the distances are x, -x and x, where x is a positive or negative number. The range deviations of the three points (f_i, h'_i) from the curve are

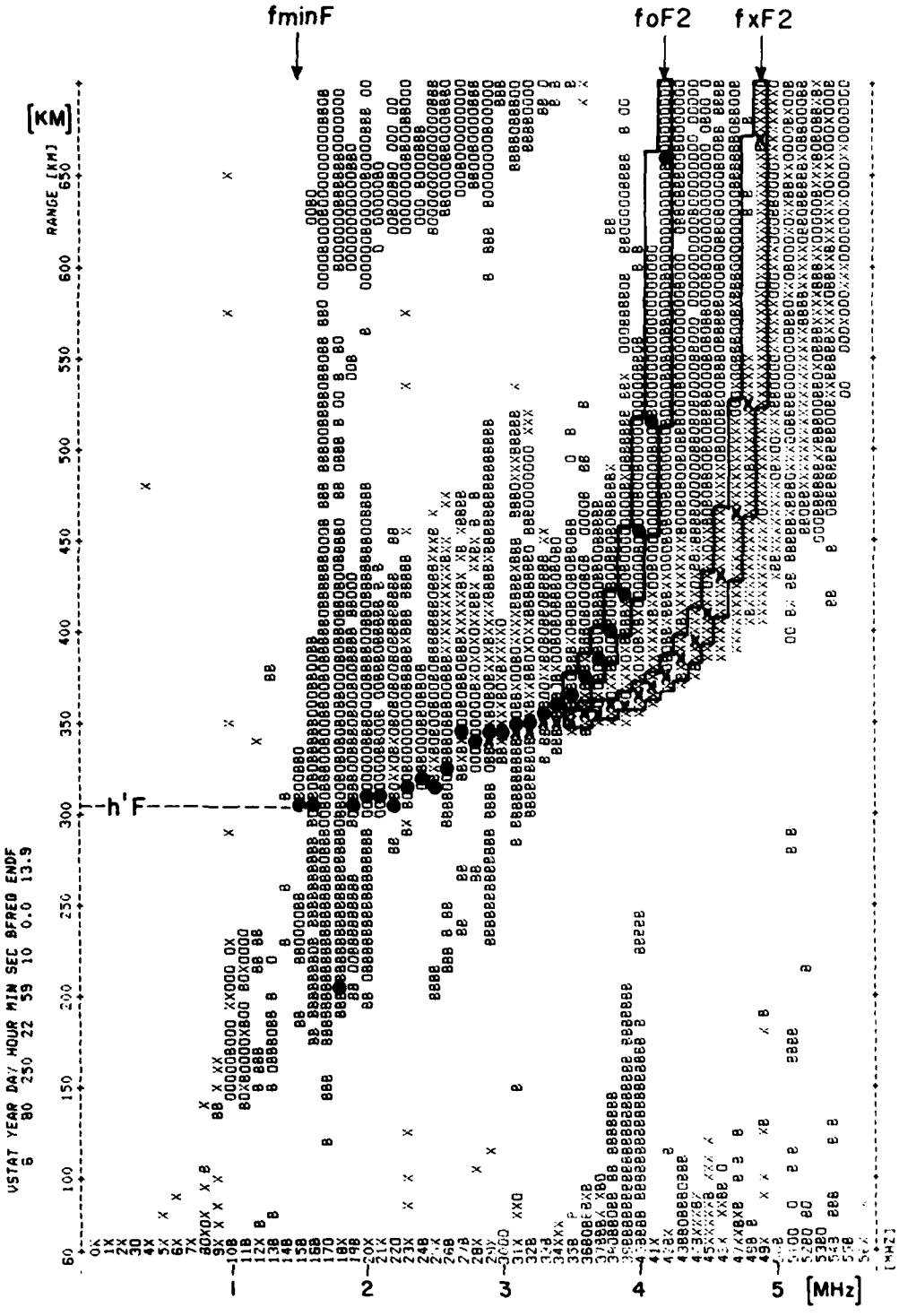
$$h'_i - \{h_0 + \frac{1}{a + bf_i}\} = (-1)^{i-1} x \qquad i = 1, 2, 3. \qquad (2.10)$$

These equations can be written in the form (deleting the primes on the heights):

$$(h_1 - h_0 - x)a + (h_1 - h_0 - x)f_1 b - 1 = 0 \qquad (2.11)$$

$$(h_2 - h_0 + x)a + (h_2 - h_0 + x)f_2 b - 1 = 0 \qquad (2.12)$$

$$(h_3 - h_0 - x)a + (h_3 - h_0 - x)f_3 b - 1 = 0 \qquad (2.13)$$



GOOSE BAY 6 SEP 1980 23.00 AST

Figure 5

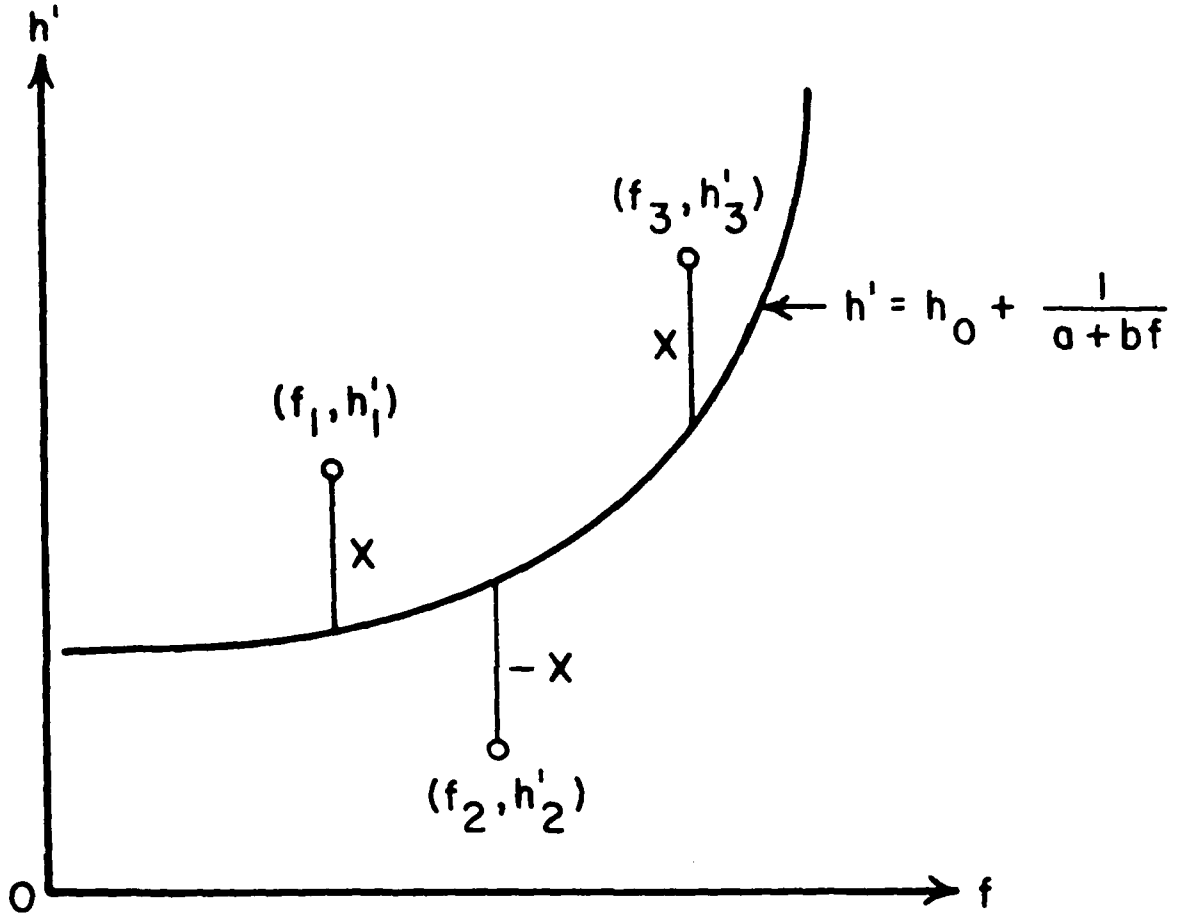


Figure 6. Min-Max Fitting

From (2.11) and (2.12) one obtains

$$a = \frac{1}{h_1-h_0-x} - \frac{f_1}{f_2-f_1}\left(\frac{1}{h_1-h_0-x} - \frac{1}{h_2-h_0+x}\right) \tag{2.14}$$

$$b = -\frac{1}{f_2-f_1}\left(\frac{1}{h_1-h_0-x} - \frac{1}{h_2-h_0+x}\right) \tag{2.15}$$

Elimination of a and b leads to a quadratic equation for the determination of x [Scheid, 1968, p. 289]:

$$\begin{vmatrix} h_1 - h_0 - x & (h_1 - h_0 - x) f_1 & 1 \\ h_2 - h_0 + x & (h_2 - h_0 + x) f_2 & 1 \\ h_3 - h_0 - x & (h_3 - h_0 - x) f_3 & 1 \end{vmatrix} = 0 \tag{2.16}$$

or:

$$Ax^2 + Bx + C = 0 \tag{2.17}$$

The coefficients A, B and C are:

$$A = 2\,(f_1 - f_3) \tag{2.18}$$

$$B = B_0 + Ah_0 \tag{2.19}$$

$$C = C_0 - C_1 h_0 \tag{2.20}$$

where

$$B_0 = -f_1(h_3+2h_1-h_2) - f_2(h_3-h_1) - f_3(h_2-2h_3-h_1) \tag{2.21}$$

$$C_0 = f_1h_1(h_3-h_2) + f_2h_2(h_1-h_3) + f_3h_3(h_2-h_1) \tag{2.22}$$

$$C_1 = f_1(h_3-h_2) + f_2(h_1-h_3) + f_3(h_2-h_1) \tag{2.23}$$

Selecting the smaller of the two solutions for equation 2.17 we get:

$$x = -\frac{B}{2A}\left[1 - \sqrt{1 - \frac{4AC}{B^2}}\right] \tag{2.24}$$

The coefficients B and C are functions of the base height h_0 (equation 2.8), which is not known. Inspection of equation (2.24) shows that $|x|$ will be a minimum if $4AC/B^2=0$. This may not be possible to achieve, but one can try to find the minimum value of $|AC/B^2|$. This is done by setting the derivative with respect to h_0 equal to zero:

$$\frac{d}{dh_0}\left(\frac{AC}{B^2}\right) = 0 \tag{2.25}$$

resulting in

$$h_0 = \frac{C_1 B_0 + 2 AC_0}{AC_1} \tag{2.26}$$

Once h_0 and x are calculated from (2.26) and (2.24) the coefficients a and b can be calculated specifying the X-hyperbola. The corresponding O-hyperbola is simply determined by replacing f by f^*. Finally the O and X sums are determined.

It is of course unlikely that the hyperbola fitted to the last three points (highest frequency) of the baseline will be a good approximation for the X-trace in the cusp region. So the process is repeated for the next three points and so on, down to frequencies at the center window, and then again for non-abutting points. If we number the baseline data points from 1 to N, we can write the triple point sets to which hyperbolas are fitted in the following way:

GROUP 1			GROUP 2			GROUP 3		
P_1	P_2	P_3	P_1	P_3	P_5	P_1	P_4	P_7
P_2	P_3	P_4	P_2	P_4	P_6	P_2	P_5	P_8
-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-
P_{N-2}	P_{N-1}	P_N	P_{N-4}	P_{N-2}	P_N	P_{N-6}	P_{N-3}	P_N

The total number of trial hyperbolas is close to 2N requiring precious CPU time. A factor of two in time could be gained by

deleting Group 1 of the guiding points. From all these hyperbolas, the one with the maximum O and X amplitude sum is selected for the trace representation.

O-Echo Trace

Combining the O-hyperbola with the smoothed baseline gives the echo trace for the O-polarization echoes. If the Digisonde outputs complete O and X ionograms, as can be done with the new Digisonde 256, one can devise a complete O and X trace. This would be of interest mainly for stations who want to study the ionization valley between E and F-region.

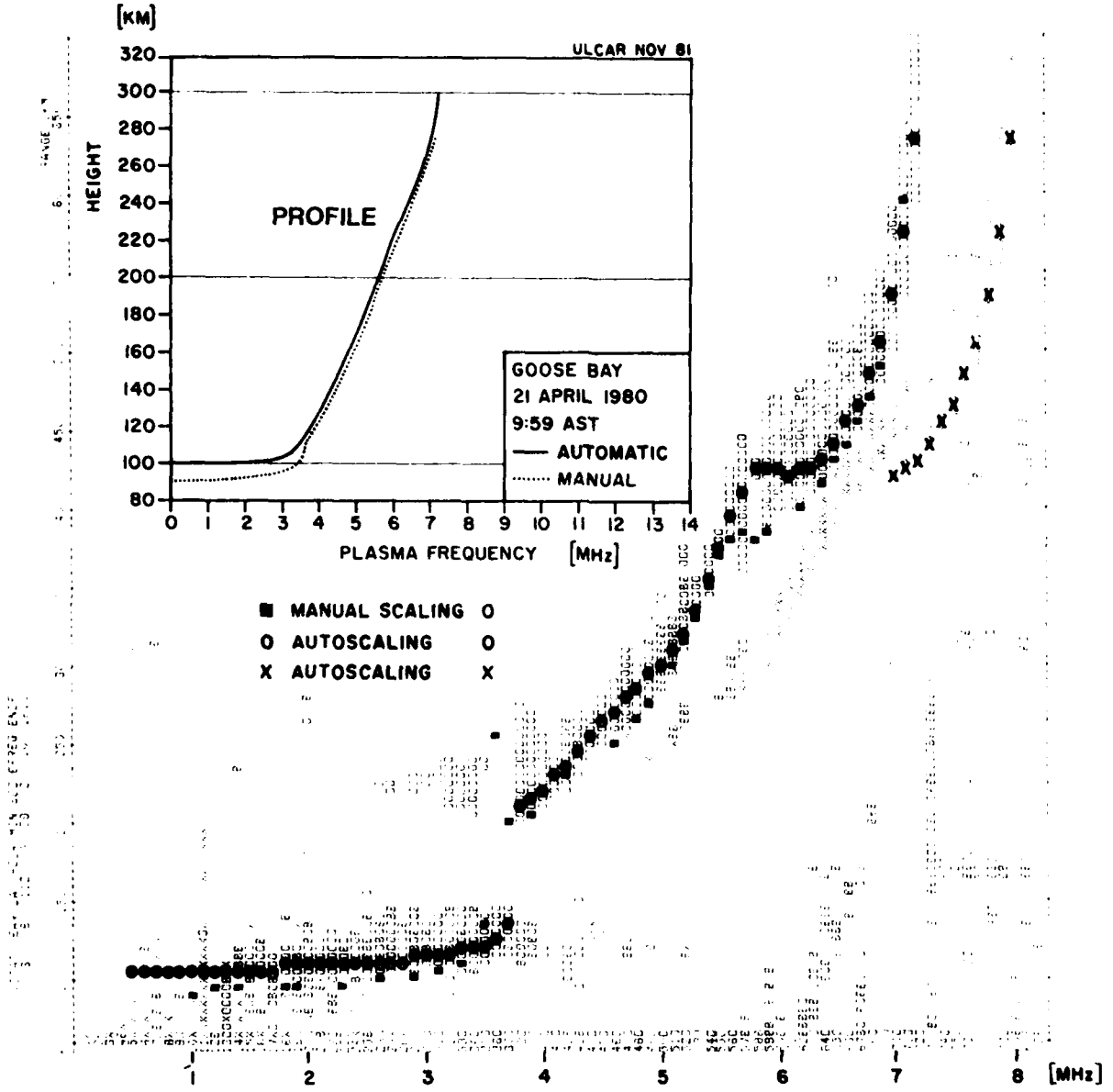
Figures 7 and 8 illustrate the performance of the scaling algorithm for Goose Bay ionograms with light spread. The preliminary results for the E-region scaling are included in these figures, together with the corresponding electron density profiles.

Parameter Extraction

At the present time, the program determines the following F-region parameters: critical frequency of the F-region, foF2, minimum height h'F, the M(3000) factor and the spread F figure. The foF2 value is obtained in the hyperbolic fitting procedure as -a/b. The smallest range value in the F-trace is used as h'F. The M(3000) factor is obtained from the transformed oblique ionogram which is calculated from the vertical ionogram trace by multiplying each frequency with the transmission factor M(h') [Smith et al, 1979, p. 12]:

$$f_{ob} = M(h') \cdot f_v \qquad (2.27)$$

The transmission curve M(h') is calculated by fitting a polynomial to the URSI specified data set [URSI Handbook of Ionogram Interpretation and Reduction, Second Edition, November 1972, p. 21; World Data Center A Report UAG-23]. The MUF(3000), obtained as the highest frequency in the oblique ionogram, is divided by foF2 to find the M(3000) factor.



SPRING DAY IONOGRAM

GOOSE BAY 21 APRIL 1980 9:59 AST

Figure 7

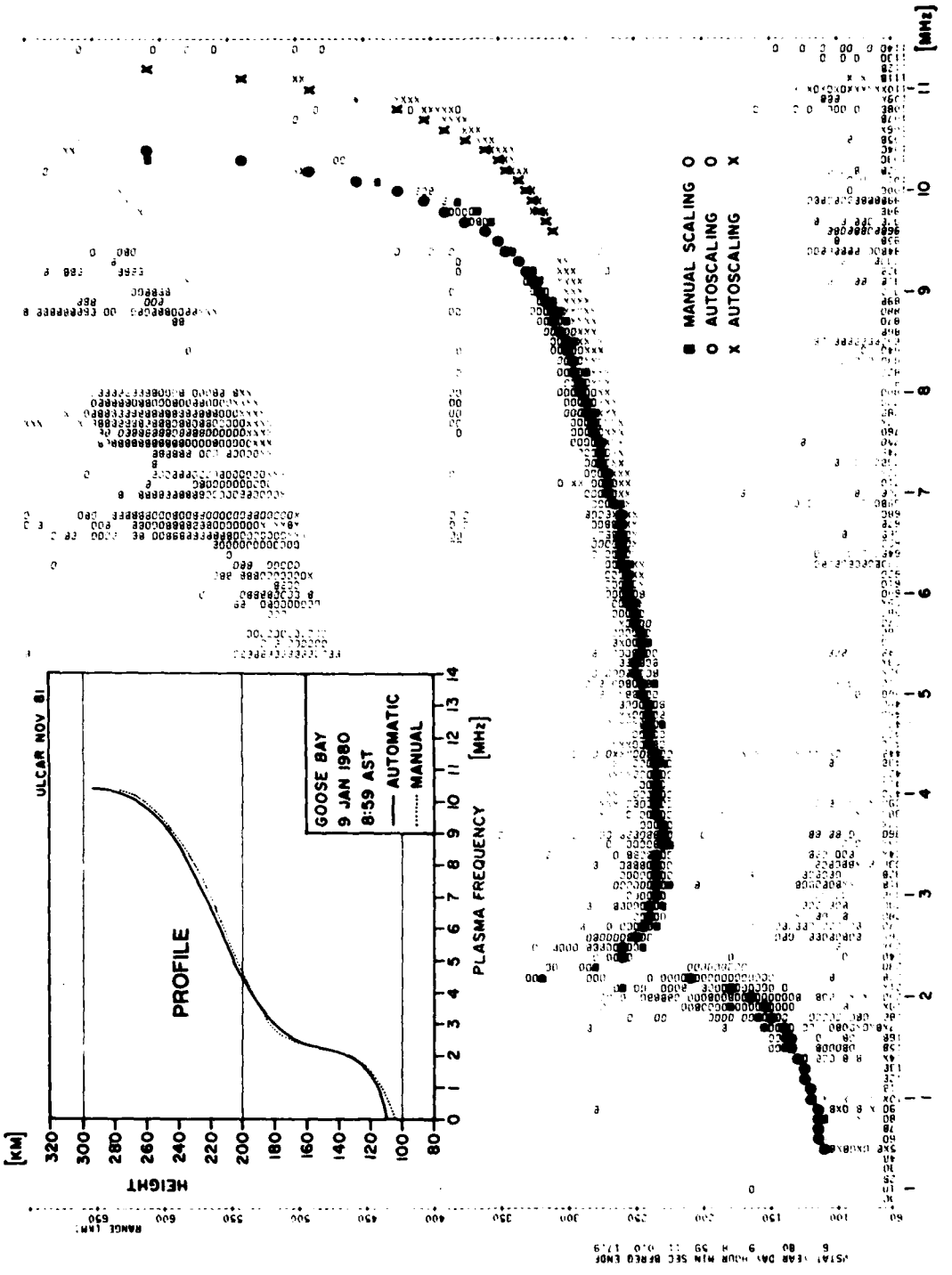

AUTOMATIC ELECTRON DENSITY PROFILE
GOOSE BAY 9 JAN 1980 8:59 AST
MANUAL SCALING O
AUTOSCALING O
AUTOSCALING X
[MHz]
PLASMA FREQUENCY
HEIGHT
[KM]
PROFILE
ULCAR NOV 81
GOOSE BAY
9 JAN 1980
8:59 AST
AUTOMATIC
MANUAL

Figure 8

For each ionogram an average spread F figure SF is determined by calculating the mean range spread:

$$SF = \frac{1}{foF2-fminF} \sum_{fminF}^{foF2} \Delta h(f). \qquad (2.28)$$

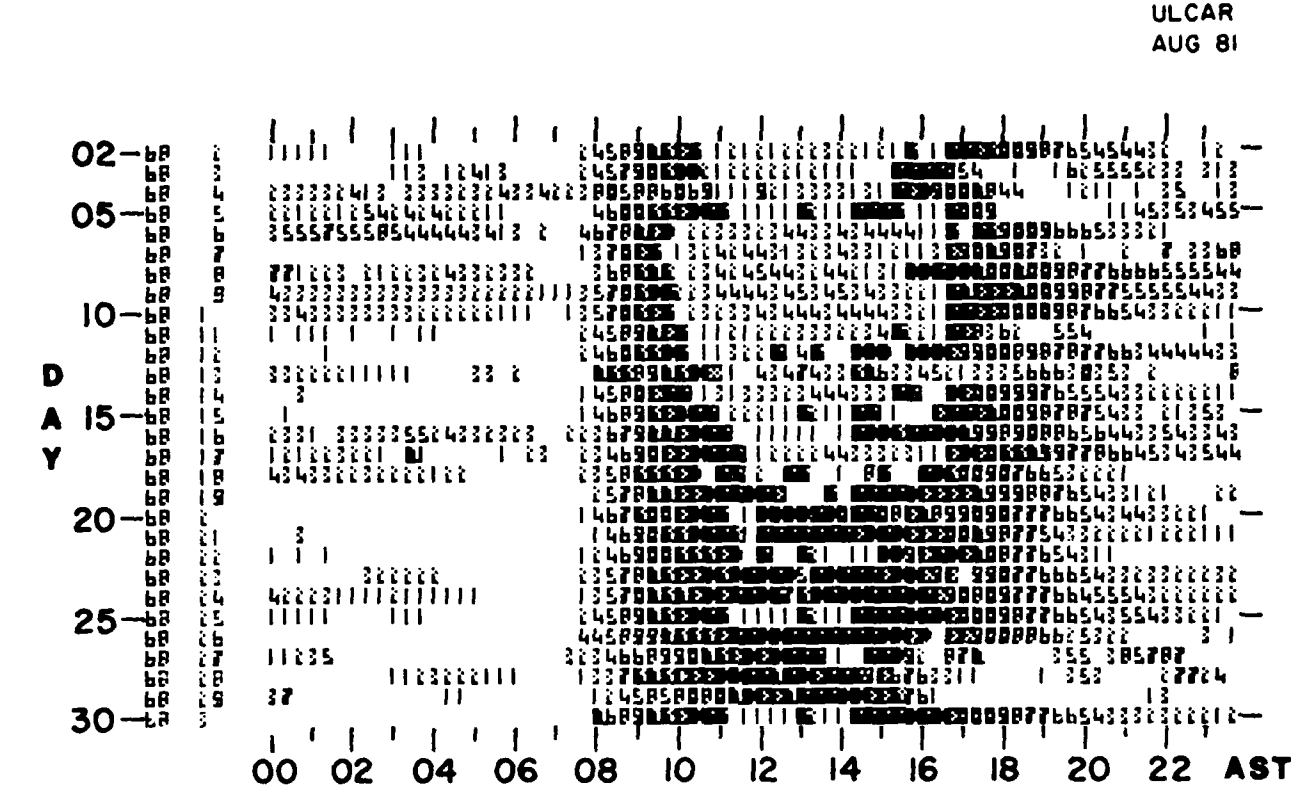
The upper limit of the range spread $\Delta h(f)$ is defined by a drop in average echo amplitude (averaged over four range bins or 20 km) by 12 dB or more below the average trace amplitude.

2.3 Performance Evaluation of the Ionogram Scaling Algorithm

It is always easy to automatically scale a few selected ionograms by tailoring the scaling algorithms to these specific ionogram. The real test is the application of the algorithm to a large number of ionograms covering day, night and twilight as well as quiet and disturbed conditions. To test the performance of the BISA program one month of ionograms, i.e. approximately 2000 ionograms, for January 1980 from Goose Bay, Labrador, were processed and the parameters of the approximately 700 hourly ionograms were compared with the manually scaled values.

2.3.1 Results of the Autoscaling

Figures 9, 10 and 11 show the diurnal and day to day variations of foF2, h'F, and the spread F figure for January 1980 using the Optifont to visually enhance the variations. With three ionograms per hour the 72 daily values are printed in one line with consecutive days following each other. The foF2 values in Figure 9 are printed in increments of 0.5 MHz although the original scaling resolution is 0.1 MHz. From a pre-sunrise minimum below 4 MHz at 07 AST, foF2 increases rapidly after sunrise. When foF2 increases from 11.5 to 12 MHz the Optifont numbers go from 15 to 0, so the 12 MHz contour line in Figure 9 is indicated as the sharp break in in-



foF2

GOOSE BAY, LABRADOR, JANUARY 1980

$foF2 = 4 + \frac{1}{2}x$ MHz

Figure 9

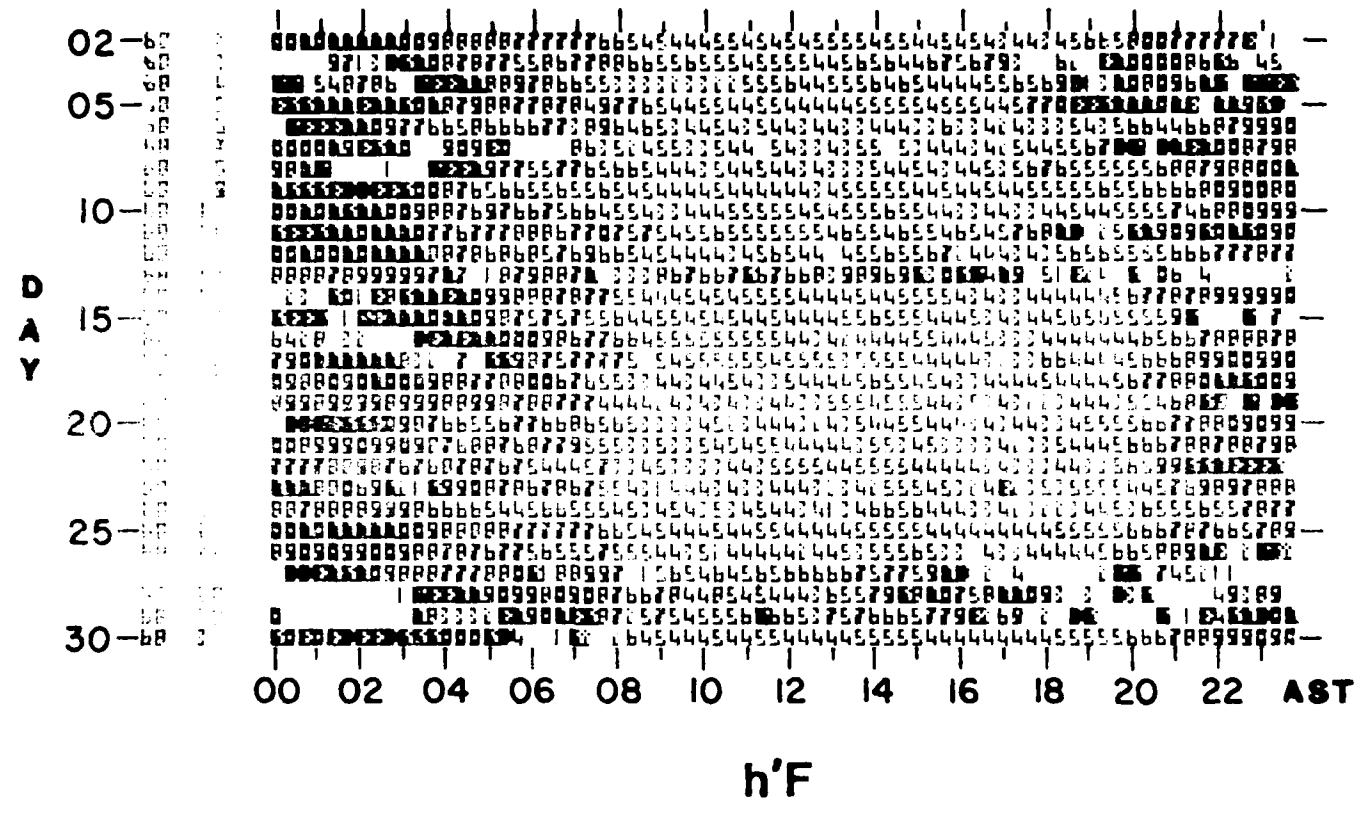
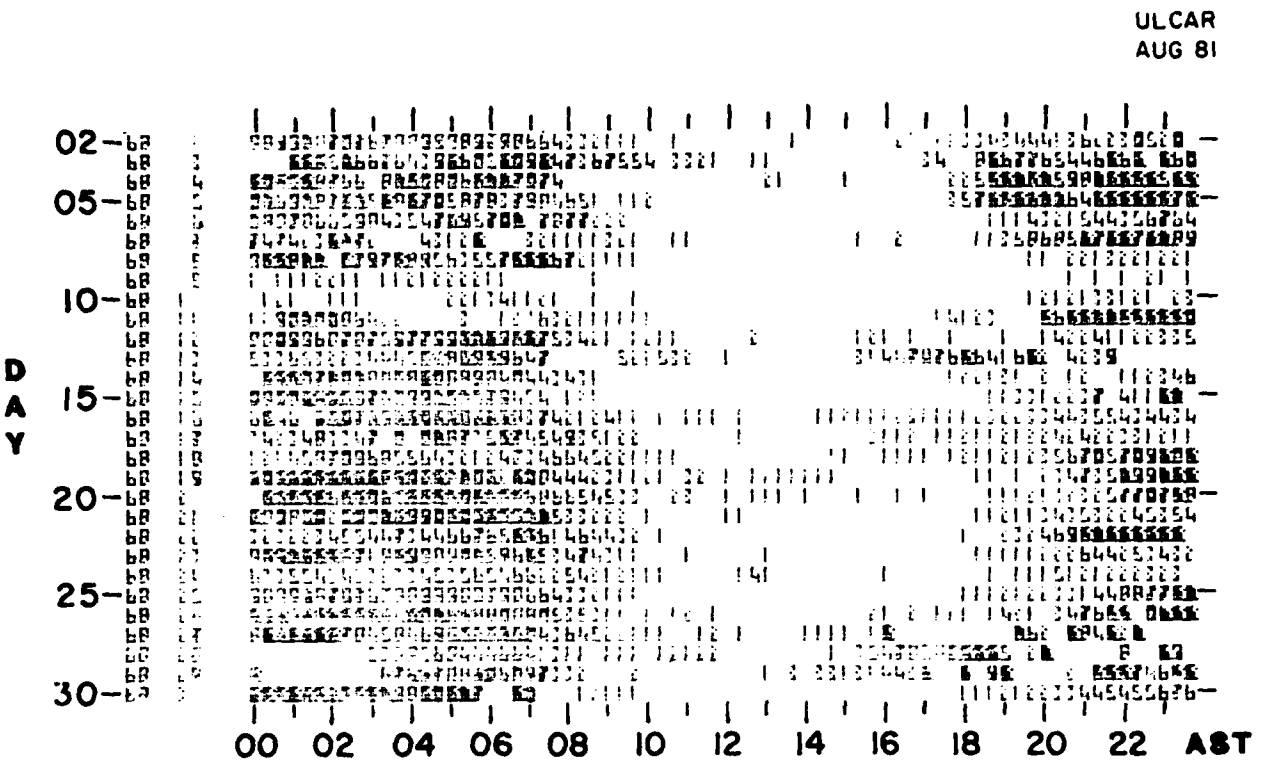


Figure 10

GOOSE BAY, LABRADOR, JANUARY 1980

h'F = 175 + 10x [km]



SPREAD F

GOOSE BAY, LABRADOR, JANUARY 1980

AVERAGE RANGE SPREAD OF F TRACE IN 5 KM INCREMENTS

$$\sigma = \frac{\int_{fminF}^{foF2} \delta h'_f}{foF2 - fminF}$$

Figure 11

tensity. The highest foF2 values were observed around January when foF2 reached values between 14.5 and 15 MHz. Later in the month foF2 stayed at or below 12 MHz. The occurrence of the F-layer trough becomes visible as a sudden decrease in foF2 during the early evening hours on days 3, 5, 7, 11, 127, 28 and 29. Days 12 and 13 show irregular foF2 values indicating ionospherically disturbed conditions. Indeed, the magnetic A-index was above 45 on these days, while A < 25 on the other days.

Figure 10 displays the h'F values in 10 km increments using the same format as the previous figure. The height variations during the day hours are small except during the disturbed days. The disturbed days and the occurrence of the trough are even better indicated by the spread F figures displayed in Figure 11. The numbers printed give the average range spread in multiples of 5 km.

The difficulty of the scaling task is illustrated by the sequence of ionograms in Figure 12 recorded in Goose Bay on 7 January 1980 between 1819 and 2039 AST. It is the time when the F-region trough moves over the station causing spread F and oblique echoes. The results of the autoscaling, indicated by the small filled circles, are in good agreement with manual interpretation. The amplitude ionograms in Figure 12 contain the O and X-echoes as well as oblique returns. If reflections from the trough walls or other irregularities, like at 1959 AST, produce several retardation cusps in the ionogram, the program most likely selects the one with the strongest amplitudes. A manual evaluator may scale foF2 at 1959 AST as 2.7 or 3.2 MHz, while the program scaled 4.3 MHz and jumped to the lower value at 2019 AST. These differences in "interpretation" are the main contributor to errors in foF2 above 0.5 MHz.

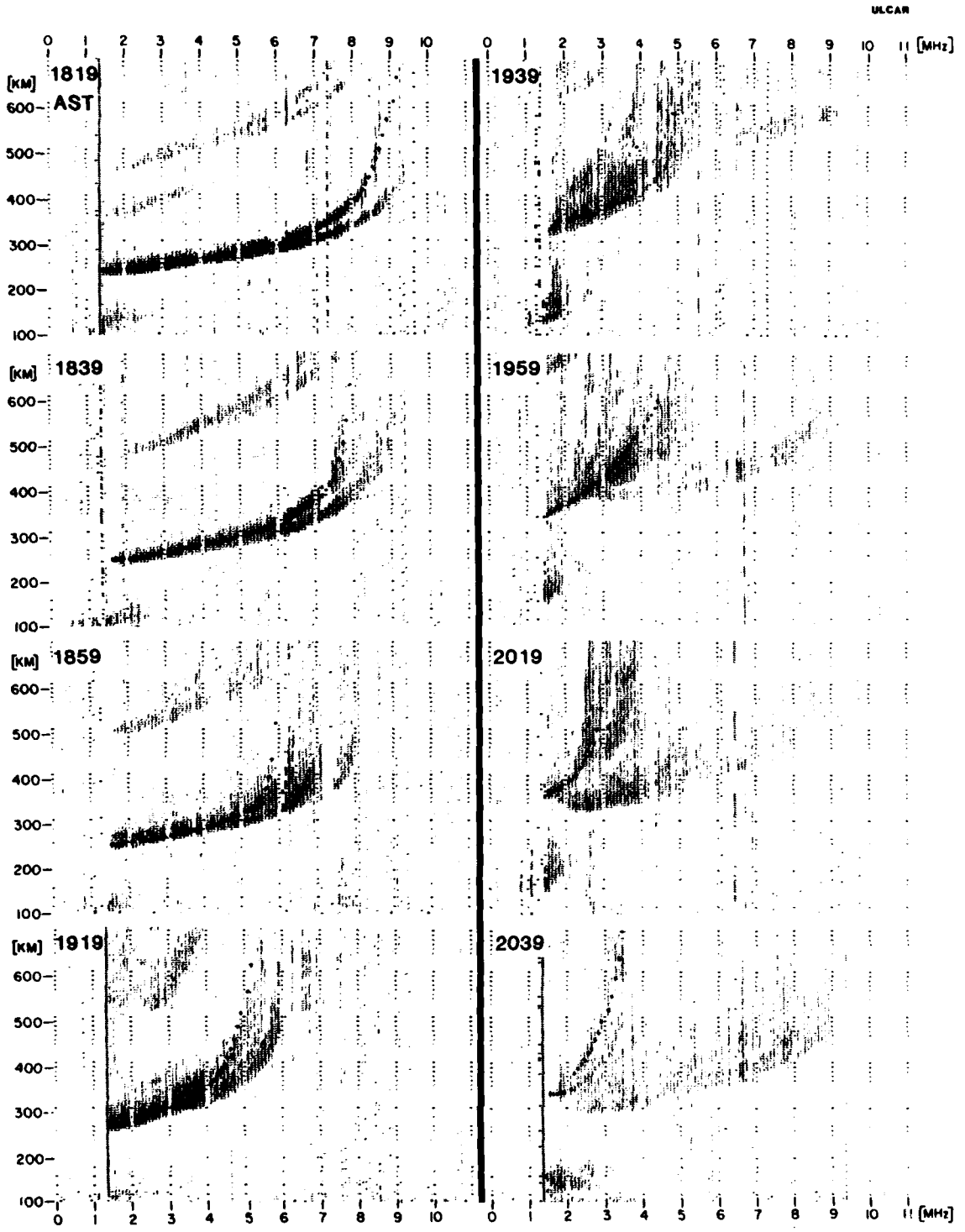

GOOSE BAY 7 JAN 1980
ULCAR
1819 AST
1839
1859
1919
1939
1959
2019
2039
[KM]
[MHz]
AUTOSCALING DURING TROUGH CONDITIONS

Figure 12

2.3.2 Comparing Manual and Autoscaling

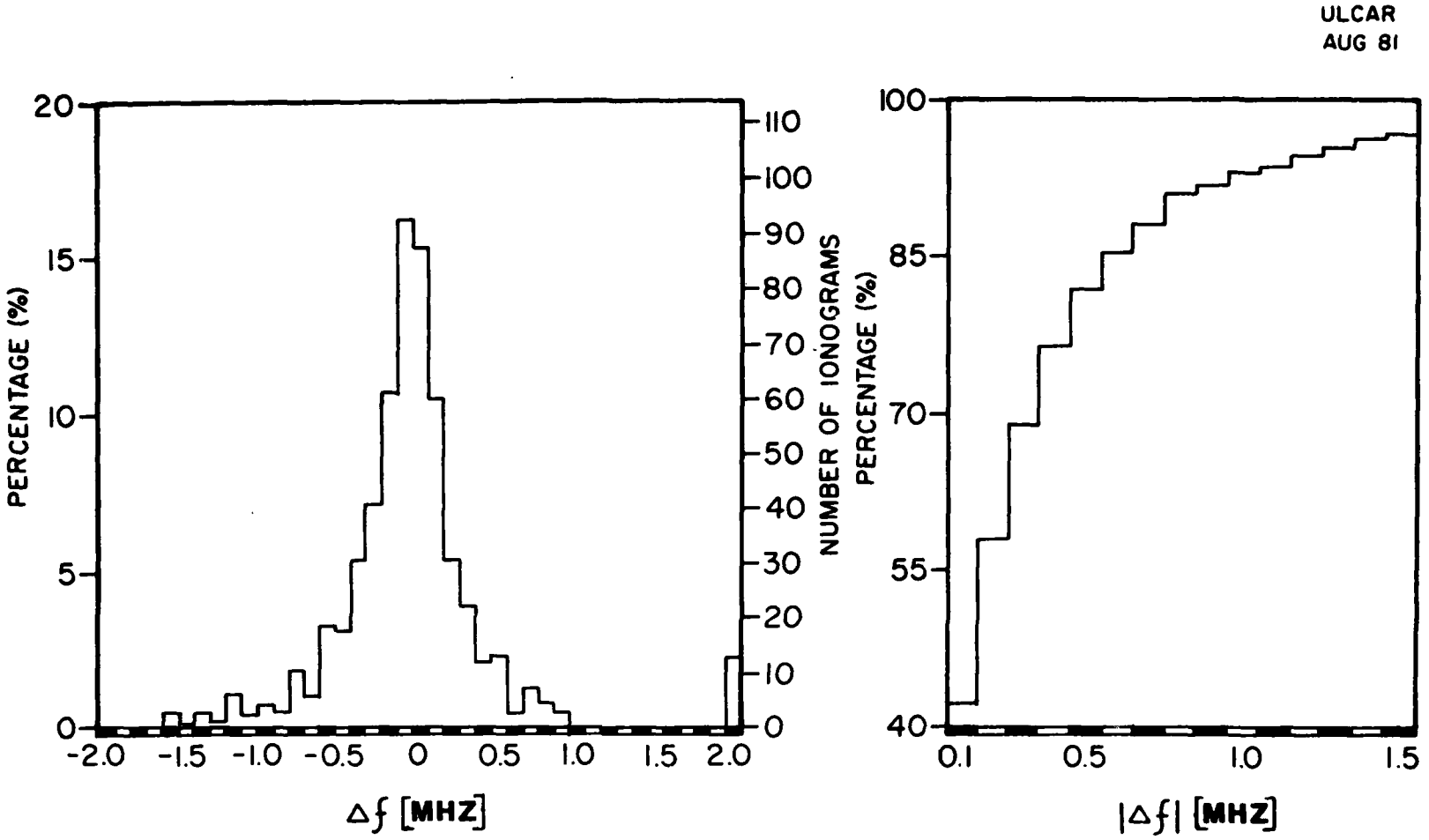
Since manually scaled parameters for the hourly ionograms were available, they were used as reference against which the autoscaled values are compared. For a total of about 570 ionograms the parameters foF2, h'F and fminF were investigated for their accuracy. Ionograms with technical errors were removed from the data base.

Figure 13 shows the error distribution function of foF2 where the error is defined as the difference between the manual and autoscaled value. The error distribution is Gaussian in shape; 82% of all ionograms have errors of less than 0.5 MHz, and 93% fall within the 1 MHz limit. If the analysis is limited to ionograms without spread F the statistics are only slightly improved as can be seen in Figure 14. This demonstrates BISA's success in the scaling of disturbed ionograms. The diurnal variation in the autoscaling of foF2 can be checked in Figure 15. It is apparent that the scaling accuracy is higher for the day ionograms than for the night measurements. This is expected because of the complicated signature of the night ionograms at Goose Bay.

Comparison of the median foF2 values shows very good agreement between manual and autoscaling (Figure 16). Two larger deviations occur at 1700 and 1900 AST which are caused by differences in interpreting the ionograms in the presence of oblique signals. The manual median values are based on the URSI conventions regarding qualifying letters, no qualifications were used for the autovalues. This explains some of the smaller deviations.

The error distribution functions for fminF and h'F (Figures 17 and 18) verify the excellent performance of BISA.

Figure 13



ERROR DISTRIBUTION OF FOF2 (MANUAL FOF2-BISA FOF2)

(USING ALL 577 IONOGRAMS, JANUARY 1980, GOOSE BAY, LABRADOR)

ULCAR
AUG 81

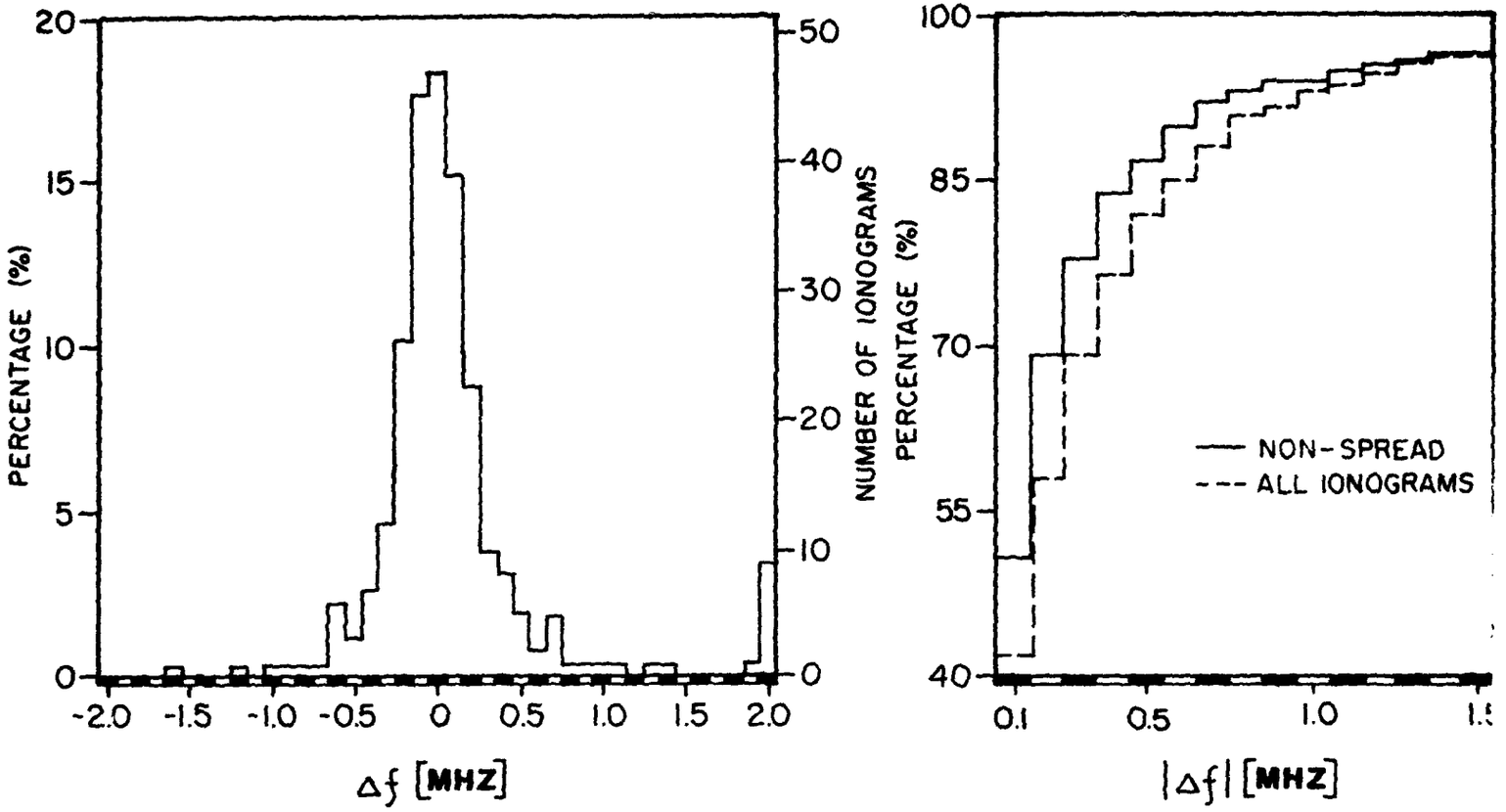


Figure 14

ERROR DISTRIBUTION OF FOF2 (MANUAL FOF2-BISA FOF2)

(USING 256 NON-SPREAD IONOGRAMS, JANUARY 1980, GOOSE BAY, LABRADOR)

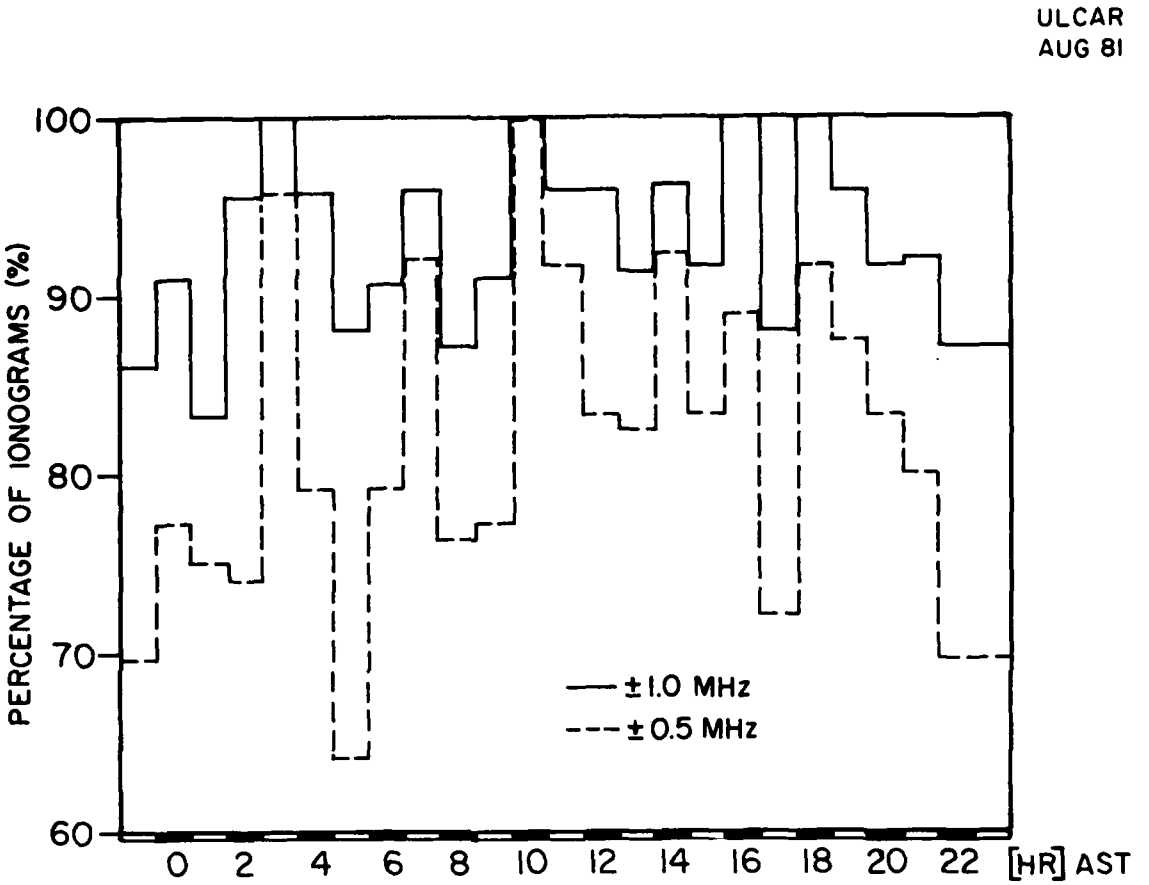


Figure 15

DIURNAL VARIATION OF ERROR (MANUAL FOF2-BISA FOF2)

(USING 577 IONOGRAMS, JANUARY 1980, GOOSE BAY, LABRADOR)

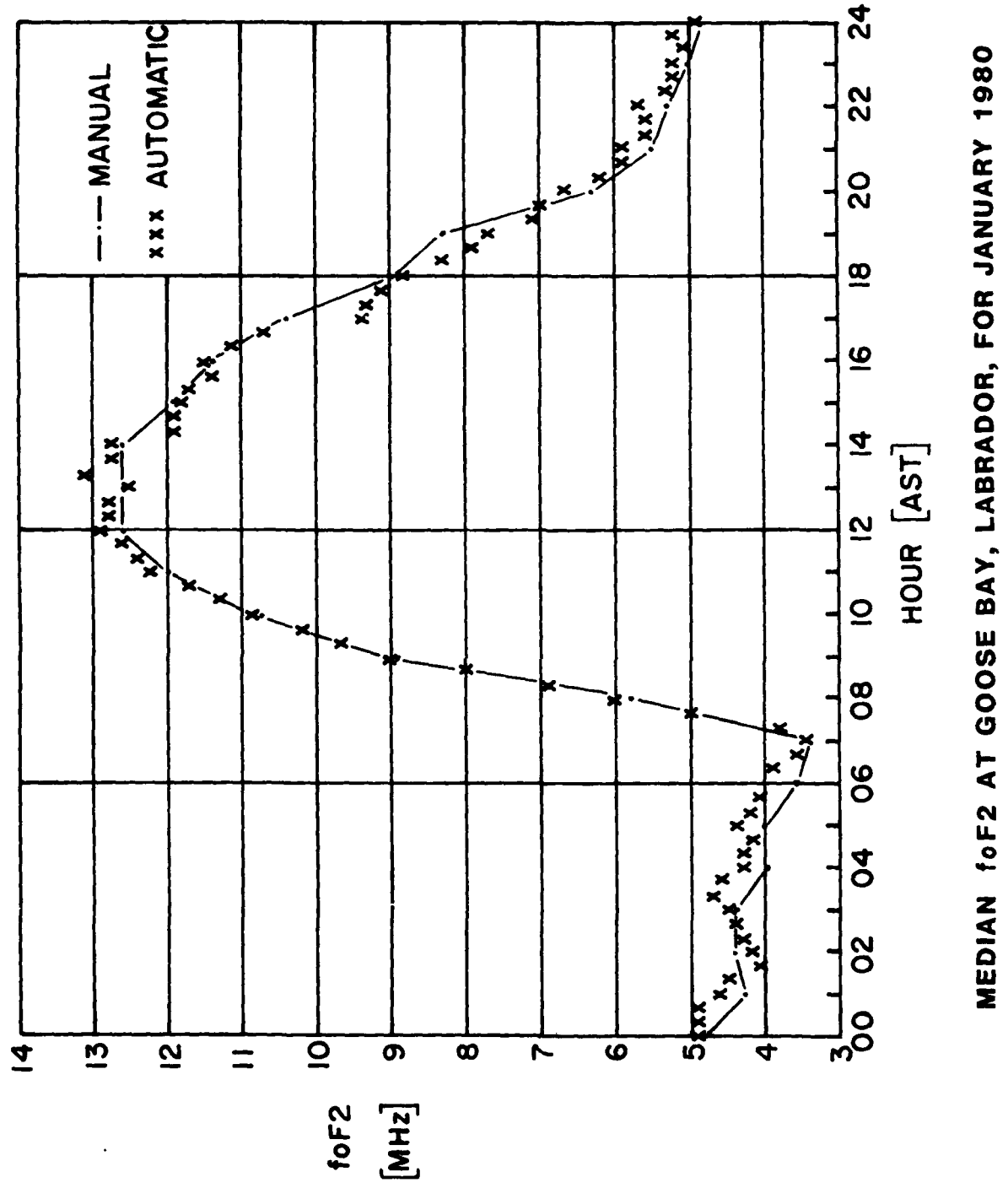

MEDIAN foF2 AT GOOSE BAY, LABRADOR, FOR JANUARY 1980
— · — MANUAL
xxx AUTOMATIC
foF2 [MHz]
HOUR [AST]
14
13
12
11
10
9
8
7
6
5
4
3
00
02
04
06
08
10
12
14
16
18
20
22
24

Figure 16

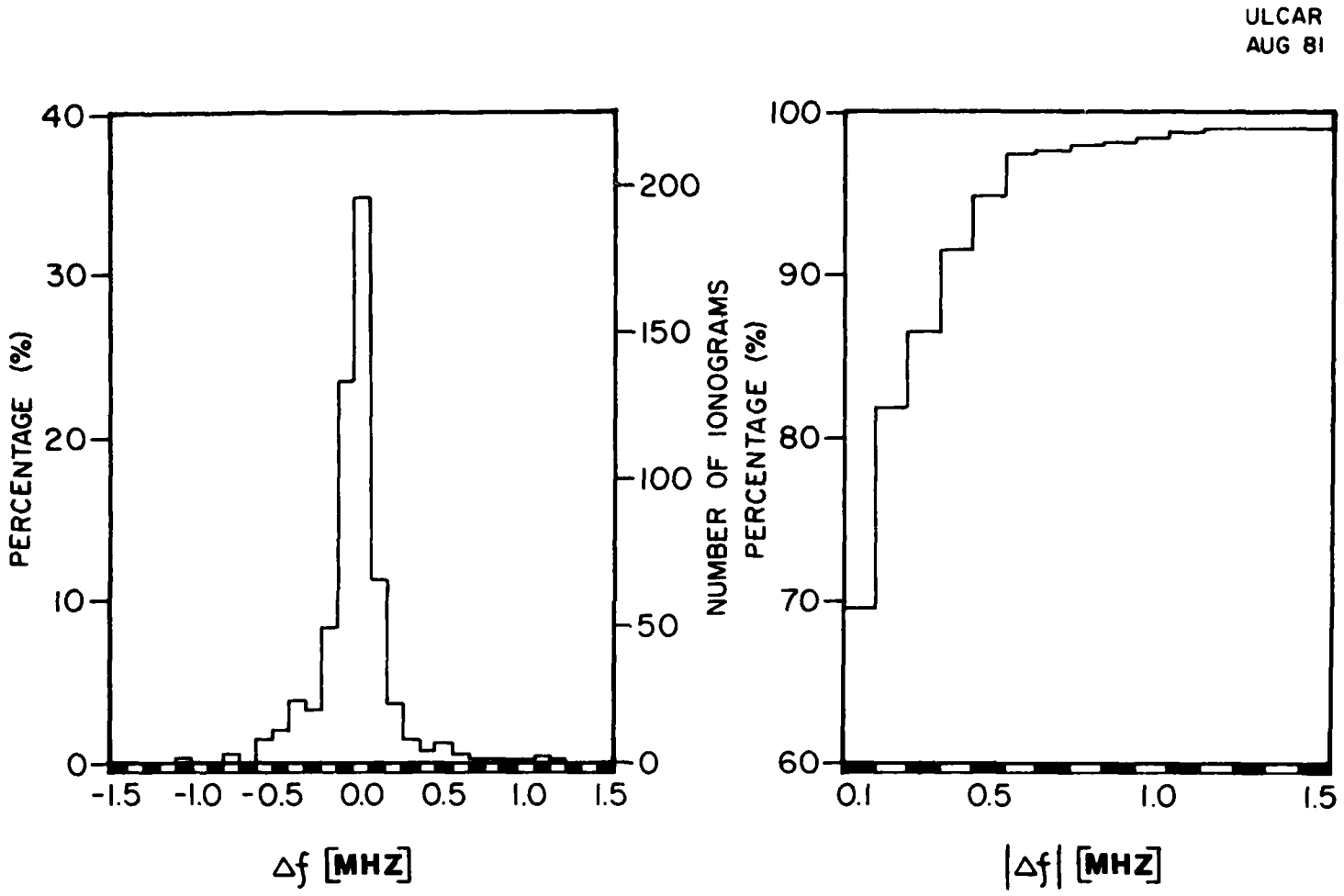


Figure 17

ERROR DISTRIBUTION OF FMINF (MANUAL FMINF-BISA FMINF)

(USING 569 IONOGRAMS, JANUARY 1980, GOOSE BAY, LABRADOR)

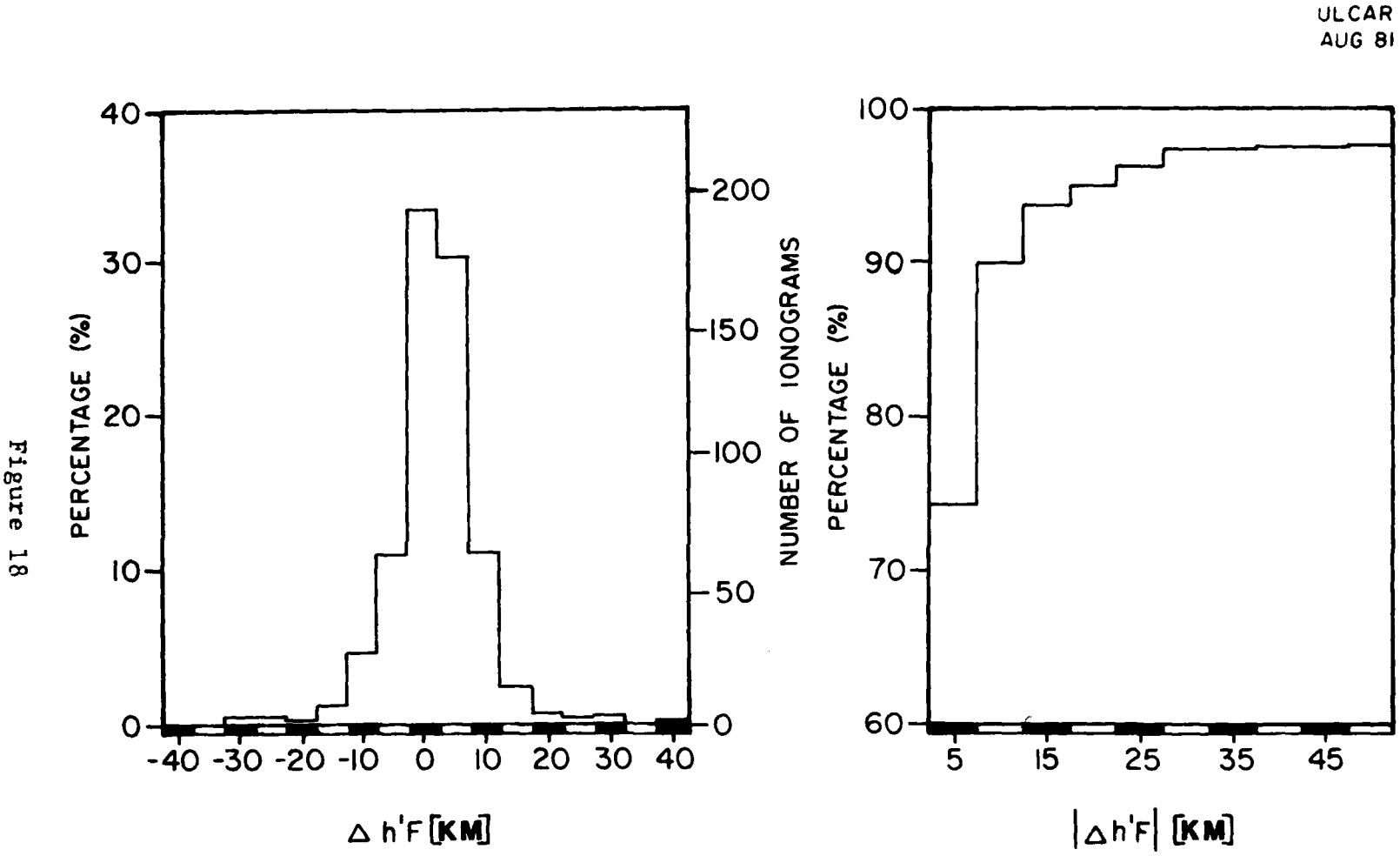


Figure 18

ERROR DISTRIBUTION OF Δ h'F (MANUAL h'F -BISA h'F)

(USING 580 IONOGRAMS, JANUARY 1980, GOOSE BAY, LABRADOR)

2.4 Summary and Outlook

By comparing the autoscaled ionospheric parameters with manually scaled values for a whole winter month at Goose Bay, it could be shown that the BISA program can be successfully employed even at stations with an unusual high rate of ionospheric disturbances. The program must now be completed by adding the E-region scaling algorithm and by specifying some qualifying parameters. The basic concepts for these tasks have been developed, so that the program can be completed within the next year. Solutions for the complementary task of converting the autotrace into an electron density profile are based on the profile-fitting method [Huang and Reinisch, 1982]. We originally developed and applied this method to autoscaled topside ionograms. The necessary adjustments for bottomside profiles will be made in the next year.

Efficiency in the computer coding, one of the achieved design goals, makes it possible to implement the software in an on-line ionogram processor, the so-called real-time ionogram scaler (RIS). The RIS will be based on the 8086 microprocessor and uses Fortran 86 as source language.

3.0 COMPARISON OF FULL WAVE THEORY AND WKB APPROXIMATION FOR DETERMINING IONOSPHERIC DENSITY PROFILES FROM REFLECTION COEFFICIENTS

3.1 The Full Wave Theory and the WKB Approximation. Time-Delay From the Phase of the Reflection Coefficient. [Cohen et al, 1982]

3.1.1 The Basic Wave Equation

Let $E(z,t)$ be a component of the electric field perpendicular to the z-axis or vertical axis along which the electron density is stratified. Then with

$$E_k(z,t) = e^{-i\omega t} E_k(z) \ , \qquad (\omega = ck) \tag{3.1}$$

the equation for $E_k(z,t)$ is [Budden, 1966, page 129]

$$\frac{d^2}{dz^2} E_k(z) + k^2 n^2 E_k(z) = 0 \ . \tag{3.2}$$

In (3.1) ω is the angular frequency, $k = \frac{\omega}{c}$ is the wave number in free space and n is the index of refraction given by

$$n^2 = 1 - \frac{N(z)4\pi e^2}{mc^2k^2} \tag{3.3}$$

where m, e, c are the mass of the electron, the charge of the electron and velocity of light in Gaussian cgs units.

In our discussion we have ignored electron collisions and the effect of the earth's magnetic field. This physical (as opposed to mathematical) approximation is often assumed in ionospheric sounding methods. At the magnetic equator, the earth's magnetic field is parallel to the earth's surface and hence the polarization of the transmitted wave can

be chosen so that the magnetic field does not affect the scattered wave. In other situations, the effect of the magnetic field may also be eliminated [Budden, 1966]. The effect of collisions is negligible at higher altitudes ($z > 100$ km). One of the ultimate objectives of the present train of research is to determine whether these physical approximations lead to errors greater than those made by the mathematical approximations. In any case, we shall assume as is customary that the use of Eq. (3.3) for the index of refraction is a good approximation for use in ionospheric sounding in certain geographical regions.

Equation (3.2) can also be written

$$\frac{d^2}{dz^2} E_k(z) + (k^2 - V(z))\, E_k(z) = 0 \qquad (3.4)$$

where $V(z)$ is given by

$$V(z) = K\, N(z), \text{ with } K = \frac{4\pi e^2}{mc^2} . \qquad (3.4a)$$

Equation (3.4) is the one-dimensional Schroedinger equation which has been exhaustively studied. The potential $V(z)$ is essentially the number density $N(z)$.

In the direct problem of reflecting electromagnetic waves from the ionosphere we assume $V(z)$ (or $N(z)$) is very small for $z < z_0$ and $z > z_1$.

We look for solutions of (3.4) which behave like

$$E_k(z) = e^{ikz} + b(k)\, e^{-ikz} \text{ for } z < z_0$$

$$= t(k)\, e^{ikz} \text{ for } z > z_1 . \qquad (3.5)$$

The quantities $b(k)$ and $t(k)$ are called the reflection and transmission coefficient respectively.

For a wave with wave number k, we have

$$\Gamma(z,t) = e^{ik(z-ct)} + b(k)\, e^{-ik(z+ct)}, \quad z < z_0$$

$$= t(k)\, e^{ik(z-ct)}, \quad z > z_1 \,. \tag{3.6}$$

The boundary conditions (3.6) are interpreted to mean that a plane wave with wave number $k = \frac{2\pi}{\lambda}$ moves initially toward the scattering potential V(z) and then is partially reflected. A transmitted wave on the other side of the potential propagates in the Z-direction.

Actually, one never sends an infinite plane wave toward the potential (or ionosphere). Instead one sends a pulse containing several wave lengths. This pulse can be represented by

$$E(z,t) = \int_{-\infty}^{+\infty} A(k)\, e^{-i\omega t}\, F_k(z)\, dk \,, \tag{3.7}$$

and is thus a superposition of the plane waves. The amplitude factor A(k) has its peak value near or at the value of $k = k_0$, where k_0 is the mean wave number which appears in the pulse. One sends a pulse

$$F_{\text{incident}}(z,t) = \int_{-\infty}^{+\infty} A(k)\, e^{ik(z-ct)}\, dk \tag{3.8}$$

and gets back a reflected pulse

$$F_{\text{reflected}}(z,t) = \int_{-\infty}^{+\infty} A(k)\, b(k)\, e^{-ik(z+ct)}\, dk. \tag{3.9}$$

Since $F_{\text{incident}}(z,t)$ and $F_{\text{reflected}}(z,t)$ are of finite extent we can ask for the time for the reflected wave to return to the transmitter. This time will depend on k_0. It is given by

$$T(k_0) = \frac{1}{c}\left[\frac{\partial\phi(k)}{\partial k}\right] \text{ evaluated at } k = k_0 \tag{3.10}$$

where $\phi(k)$ is the phase of $b(k)$; i.e.

$$b(k) = |b(k)|\ e^{i\phi(k)} \ . \tag{3.11}$$

The virtual height h_v is given by

$$h_v = \frac{1}{2}\ cT(k_0) \ . \tag{3.11a}$$

It should be remembered that the theory thus far is exact. (In using (3.10) for the time delay it is convenient to think of the transmitter as being located at $z = 0$.)

3.1.2 The Approximate Inverse Problem

The direct problem is, for our purposes, the following: Given $V(z)$ (or equivalently $N(z)$), find $T(k)$.

The inverse problem is: Given $T(k)$ for all k, find $V(z)$ (or $N(z)$).

The approximate solution is obtained from the WKB approximation. Assume $E_k(z)$ has the form

$$E_k(z) = A\ e^{iF(z)} \ . \tag{3.12}$$

Substitute (3.12) into equation (3.2) to obtain

$$\left(\frac{dF}{dz}\right)^2 = k^2\ n^2 + i\ \frac{d^2F}{dz^2} \tag{3.13}$$

which becomes a Riccati equation if $y = dF/dz$. An iteration procedure can be started by assuming on the right of Eq. (3.13) that k^2n^2 is large compared to $\frac{d^2F}{dz^2}$, i.e. n/λ is large or n varies slowly compared to the wave length [see Kay, 1954 for a more careful analysis of domain of validity]. The first iteration gives $F' = \pm kn$ while the second yields

$$F' = \pm kn\ \left(1 + \frac{iF''}{k^2n^2}\right)^{1/2} \simeq \pm kn\ \left(1 + \frac{iF''}{2k^2n^2}\right) = \pm kn + \frac{i\ n'}{2n}$$

which gives F by a integration. Since Eq. (3.13) is non-linear, the two particular solutions of (3.13) cannot be superimposed to find a general solution. However, the solution F can be substituted into Eq. (3.12) to obtain particular solutions

$$E_k(z) = A\,[n(z)]^{-1/2} \exp\left[\pm\, i\, k \int_{z_0}^{z} n(z')\, dz'\right] \qquad (3.14)$$

to Eq. (3.2). It should be remembered that n(z) is a function of k (see Eq. (3.3)). We have suppressed this dependence on the notation for simplicity. If one particular solution to Eq. (3.2) is known, it can be used to give a first order linear equation which can be solved by quadratures thereby giving the general solution. Alternatively, two distinct particular solutions can be added to give the general solution of Eq. (3.2) which is (within this approximation)

$$E_k(z) = A(k)\,[n(z)]^{-1/2} \exp\left[+\, i\, k \int_{z_0}^{z} n(z')\, dz'\right].$$

$$+\, B(k)\,[n(z)]^{-1/2} \exp\left[-\, i\, k \int_{z_0}^{z} n(z')\, dz'\right]. \qquad (3.15)$$

Using the first boundary condition of Eq. (3.6) and noting $n(z) = 1$ for $z = z_0$, one finds

$$E_k(z) = e^{ikz} + b(k)\, e^{-ikz} \quad (z < z_0)$$

$$= A(k)\, e^{-ikz_0}\, e^{ikz} + B(k)\, e^{ikz_0}\, e^{-ikz} .$$

Hence $A(k) = e^{ikz_0}$

$$B(k) = b(k)\, e^{-ikz_0} . \qquad (3.16)$$

Let z_k be the value of z for which $n(z) = 0$, i.e. from Eq. (3.3) and (3.4a)

$$V(z_k) = k^2 . \qquad (3.17)$$

Our picture, the usual one, is that $V(z) = 0$ for $z < z_0$ and increases monotonically toward a maximum as z increases above z_0. As k^2 varies from 0 to the maximum of $V(z_k)$, a z_k is defined by (3.17).

From (3.15), $E_k(z_k) \to \infty$ unless

$$A(k) \exp[i k \int_{z_0}^{z_k} n(z')dz'] + B(k) \exp[- i k \int_{z_0}^{z_k} n(z')\, dz'] = 0. \tag{3.18}$$

Thus

$$B(k) = - A(k) \exp[2ik \int_{z_0}^{z_k} n(z')\, dz']$$

$$= - e^{ikz_0} \exp[2ik \int_{z_0}^{z_k} n(z')\, dz'] . \tag{3.19}$$

Or from (3.16)

$$b(k) = - e^{2ikz_0} \exp[2ik \int_{z_0}^{z_k} n(z')\, dz'] . \tag{3.20}$$

The phase $\phi(k)$ of Eq. (3.11) is

$$\phi(k) = \pi + 2\, k\, [z_0 + \int_{z_0}^{z_k} n(z')\, dz'] . \tag{3.21}$$

$$T(k) = \frac{2}{c} [z_0 + \int_{z_0}^{z_k} n(z')\, dz'] + \frac{2k}{c} \int_{z_0}^{z_k} \frac{\partial}{\partial k} n(z')\, dz'$$

$$+ \frac{2k}{c} n(z_k) \frac{\partial z_k}{\partial k} . \tag{3.22}$$

But $n(z_k) = 0$ by definition (see Eqs. (3.3) and (3.4)), also

$$\frac{\partial n(z)}{\partial k} = \frac{\partial}{\partial k}\frac{1}{k}\sqrt{k^2-V(z)}$$

$$= -\frac{1}{k^2}\sqrt{k^2-V(z)} + \frac{1}{\sqrt{k^2-V(z)}}$$

$$= -\frac{1}{k}n(z) + \frac{1}{kn(z)} . \qquad (3.23)$$

Thus finally

$$T(k) = \frac{2}{c}\left[z_0 + \int_{z_0}^{z_k}\frac{dz'}{n(z')}\right]$$

or

$$T(k) = \frac{2}{c}\left[z_0 + k\int_{z_0}^{z_k}\frac{dz'}{\sqrt{k^2-V(z')}}\right] . \qquad (3.24)$$

This is a well-known integral equation for V(z). The first term on the right $\frac{2z_0}{c}$ represents the time for the signal to reach z_0 and return to the transmitter at z = 0. If k_1 is the lowest value of k for which there is a reflection, then approximately

$$z_0 = \frac{cT(k_1)}{2} . \qquad (3.25)$$

Now that z_0 is determined approximately, one can find V(z) (or N(z)) using Abel's method of inversion [Faddeev, 1958]. For each value of k one gets T(k) experimentally. One then can find V(z) ($z_0 < z < z_k$) using Abel's method. This is the usual procedure. Note if $k^2 > V_{max}(z)$, the method fails.

3.1.3 The Exact Inverse Problem

The exact or full-wave theory version of the inverse problem is developed in Kay, 1954 and Kay and Moses, 1956.

Let b(k) be the reflection coefficient as in Eq. (3.6). The potential can be recovered from a knowledge of b(k) through the use of the Gelfand-Levitan algorithm. To be specific, let us define R(z) by

$$R(z) = (2\pi)^{-1} \int_{-\infty}^{+\infty} b(k)\, e^{-ikz}\, dk. \tag{3.26}$$

The reflection coefficient b(k) satisfies the following conditions [see Kay, 1954; Kay and Moses, 1956; Faddeev, 1958].

$$b(-k) = b^*(k),\ b(k) \text{ analytic in the upper half-plane,}$$
$$b(0) = -1,\ R(z) \equiv 0 \text{ for } z < 2z_0. \tag{3.27}$$

Let us define the Gelfand-Levitan kernel K(z,y) by

$$K(z,y) = 0 \text{ for either } y > z, \text{ or } z < z_0. \tag{3.28}$$

For $y < z$ and at the same time $z > z_0$ we require K(z,y) to satisfy the Gelfand-Levitan equation:

$$K(z,y) = -R(z+y) - \eta(z+y-2z_0) \int_{2z_0-y}^{z} K(z,u)\, R(u+y)\, du \tag{3.29}$$

where $\eta(x)$ is the Heaviside function defined by $\eta(x) = 1$ for $x > 0$, $\eta(x) = 0$ for $x < 0$.

Having found the Gelfand-Levitan kernel K(z,y), V(z) is given by the simple expression

$$V(z) = 2\,\frac{d}{dz}\,K(z,z)\ . \tag{3.30}$$

However, the electric field $E_k(z)$ can also be obtained using

$$E_k(z) = e^{ikz} + b(k)\, e^{-ikz} + \int_{2z_0-z}^{z} K(z,u)\, [e^{iku} + b(k)\, e^{-iku}]\, du. \tag{3.31}$$

Thus to find V(z) using the algorithm, we must obtain the reflection coefficient b(k) from its phase.

This is accomplished using the analytic properties of b(k) and, hence, log b(k) and a generalized form of the Hilbert transform. To be specific, (d/dk) $\phi(k)$ is found from the time delay in accordance with Eq. (3.10). The phase $\phi(k)$ is obtained by integration with the boundary condition $\phi(0) = \pi$. Let v(k) be defined by

$$v(k) = \phi(k) - 2kz_0 \tag{3.32}$$

and w(k) by

$$w(k) = \log |b(k)| \tag{3.33}$$

then

$$w(k) = \frac{P}{\pi} \int_{-\infty}^{+\infty} \frac{kv(k')}{k'(k'-k)}\, dk' \ . \tag{3.34}$$

In Eq. (3.34), the symbol P means the principal part of the integral.

It should be mentioned that there are variational principles available [Moses, 1977] which enable one to obtain V(z) when b(k) is known. These principles have an upper bound built into them.

One also has available a generalization of the Gelfand-Levitan algorithm. If, for a given reflection coefficient $b_0(k)$, one knows the electron density $V_0(z)$, one can obtain $V(z) - V_0(z)$ in terms of $b(k) - b_0(k)$ [Pechenick and Cohen, 1981]. One can view this generalization as offering at least two options. One may regard $V(z) - V_0(z)$ as the error in density due to an error $b(k) - b_0(k)$ in the reflection coefficient. Or one may think of $V_0(z)$ as being the density associated with a time delay leading to $b_0(k)$ having been obtained from a model or a previous calculation (even using the WKB method). Then V(x) is obtained as a relatively small change due to the change in the reflection coefficient. The variational principle can also be used to obtain $V(z) - V_0(z)$ from $b(k) - b_0(k)$ together with a bound on this difference.

3.2 Comparison of the WKB Method with the Full-Wave Method for Profiles for Which the Full-Wave Equation can be Solved for Exactly.

3.2.1 General Outline of Research Procedure

The use of the WKB approximation for the calculation of electron density profiles is a well-established procedure. As mentioned in the previous section, the WKB approximation is not exact. As part of our research effort we are considering potentials V(z) (which, as seen above, is equivalent to considering number densities N(z)), for which the Schroedinger equation can be solved for exactly and compare the exact time delays (or equivalently virtual heights) with the time delays which would be calculated from the potential using the WKB approximation discussed in Section 3.1. A comparison of the time-delays will give us some idea of how good the WKB approximation is. Surprisingly, perhaps, our work in this area will probably be one of the few to compare exact time-delays with WKB time delays, even though the WKB time-delay has been used for almost sixty years. The reason that this comparison was not made earlier is that the Schroedinger equation had not been solved exactly for potentials V(x) for which the WKB approximation might be expected to give an accurate result. However, in more recent years it is possible to construct potentials from a given reflection coefficient for which the Schroedinger equation can be solved exactly and in terms of elementary functions. Thus having specified a reflection b(k) and hence as part of the reflection coefficient, its phase and thus the exact time-delay, we compute the potential. Having the potential, it is then an easy matter to compute the WKB time-delay. The exact time-delay is then compared with the WKB time-delay.

Using inverse methods it is possible to give a class of potentials which are smooth and continuous and which would seem to model a single layer ionosphere insofar as the

general shape is concerned. Since these potentials are obtained from given reflection coefficients, the phase and hence the time-delay are known exactly. From such examples one can compute the time-delay using the WKB method. The WKB time-delay bears no relation to the exact time-delays. In the simplest case, which we discuss below, the scaling is inappropriate for the WKB method to be valid despite the apparent smoothness. Other examples remain to be examined for scaling.

3.2.2 General Properties of the WKB Time-Delay. Failure of the WKB Method. Tentative Conclusions.

The formula for the WKB time-delay in terms of a potential is identical to the formula for obtaining the height of a hill (essentially V(x)) from the time required for a ball rolled up the hill with a given velocity to come down again to the point at which the ball started rolling. Generally, the higher the initial velocity is, the greater the time of return will be because the ball will have to roll higher up the hill. In wave terms, the WKB time-delay will generally increase with increasing k. Our examples of potentials given below will show that the WKB behavior for the time-delay will be as expected. However, the <u>exact</u> time-delay will <u>decrease</u> with increasing k no matter how slowly growing the potential of the given class is. In the simplest case, discussed in some detail below; the WKB time-delay bears little relationship to the exact, full-wave time-delay.

Before we give the examples, we can draw some conclusions from the comparison of full-wave and WKB theory. It is mathematically possible to have layers for which the WKB time-delay bears no relation to the full-wave time delay. In practice one usually observes time-delays which increase with k in analogy to the WKB behavior. While these time-delays may

arise from potentials for which the WKB approximation is reasonably good, from the inverse point of view one will not know that the observed time-delay (which coincides with the exact time delay for the potential) can be used in a WKB approximation without comparing the potential from the WKB approximation with that of the full-wave theory. Since it is necessary to use the full wave theory for the comparison, and since in principle and probably in practice (when the numerical methods are fully established), it is as easy to use the full-wave inversion method as the WKB inversion, one should use the exact method in the first place to prevent error. The WKB method may also be used afterward if one chooses, but only to see how well it predicts the exact potential.

3.3 Examples

Our examples arise through the choice of a suitable set of reflection coefficients b(k). Let us consider the Gelfand-Levitan equations (3.26) and (3.29). It is easy to see that near $z = z_0$, V(z) is given by

$$V(z) = 2 \frac{d}{dz} (2\pi)^{-1} \int_{-\infty}^{+\infty} b(k)\, e^{-2ikz}\, dk. \tag{3.35}$$

The class of reflection coefficients we shall consider are dependent on two parameters, a positive constant a, and a positive integer n:

$$b(k) = - \frac{(ia)^n}{(k+ia)^n} . \tag{3.36}$$

We have picked b(k) so that it satisfies

$$b(0) = - 1. \tag{3.37}$$

This condition is necessary for V(z) to be positive. All the corresponding potentials $V_n(z)$ are zero for negative values of z. We thus are taking z_0 to be zero and regarding the

transmitter to be at the bottom end of the layer. [The generalization to positive z_0 simply adds a time-delay to free-space propagation from the transmitter at z = 0 to the bottom of the layer at $z = z_0$.] It is also seen that (in the vicinity of z = 0) $V_n(z)$ is of the form

$$V_n(z) = C_n z^{n-2} \quad , \; z \text{ greater than or equal to zero,} \quad (3.38)$$

for n greater than or equal to 2 where C_n is a real positive constant. For n = 2 we see from (3.35) that the potential has a jump at z = 0 from zero for z less than zero to C_2 for z greater than but near zero. For n = 3 the potential is continuous at zero and $V_3(0) = 0$. For n = 4, V_n and the first derivative of V_n are zero at z = 0 and are continuous there. As one takes higher values of n an increasing number of derivatives are continuous and zero at z = 0. Thus one can make the potential increase from zero to finite values as slowly as one wishes. High values of n correspond to increasingly smoother variations of V_n from zero to positive values. In regions sufficiently near z = 0, where V_n varies in as smooth a manner as one wishes, one might expect the best chances of the WKB time-delay to be valid. However, for all values of k and n the exact time delay <u>decreases</u>, as we shall show shortly.

A possible explanation for the failure of the WKB method is the fact that in addition to requiring that the index of refraction vary slowly and not be too small, the value of k should not be small [Budden, 1966]. For low value of n, k is small compared to the scale a, but for high values of n, k can be made increasingly large and the WKB procedure goes through formally but still gives the wrong character for the time-delay as compared with the exact time-delay. These results suggest more strongly than ever that the WKB inverse method requires a closer scrutiny even where it is believed to be valid.

The time-delay $T_n(k)$ arising from the reflection coefficient of Eq. (3.36) is readily found to be

$$T_n(k) = \frac{na}{k^2+a^2} . \qquad (3.39)$$

The exact potential $V_n(z)$ have been computed and sketched by Pechenick and Cohen [1981]. As n increases the potentials, which may be considered to be a two-parameter family of functions depending on n and a, are all positive and, for n equal to or greater than 3 start at zero for z = 0 and have a suitable number of derivatives vanishing at z = 0. The potentials increase monotonically, reach a maximum and then decay exponentially for large z. For a given value of a, higher values of n correspond to flatter potentials, i.e. the effective base of the potential increases with respect to the height.

In Figure 19, we have plotted $V_3(x)$. To the eye it looks like a "normal" potential. One would think that the WKB calculation of the time delay for that portion of the potential for z between 0 and the value of z corresponding to the maximum would be a good approximation. In Figure 20, the exact and WKB values of the time-delay are plotted as a function of k. The discrepancy between the two time delays is astonishing! The WKB time-delay does not even approach the values of the exact one for most values of k. Thus, even for potentials which look as though WKB time delay ought to be valid, this is not generally the case.

It should be noted that $V_3(x)$ does not model the ionosphere at all well because of the length to height ratio, which indeed precludes the use of the WKB method. A better exact model, which we are trying to obtain, will be more useful for comparing the WKB and exact methods.

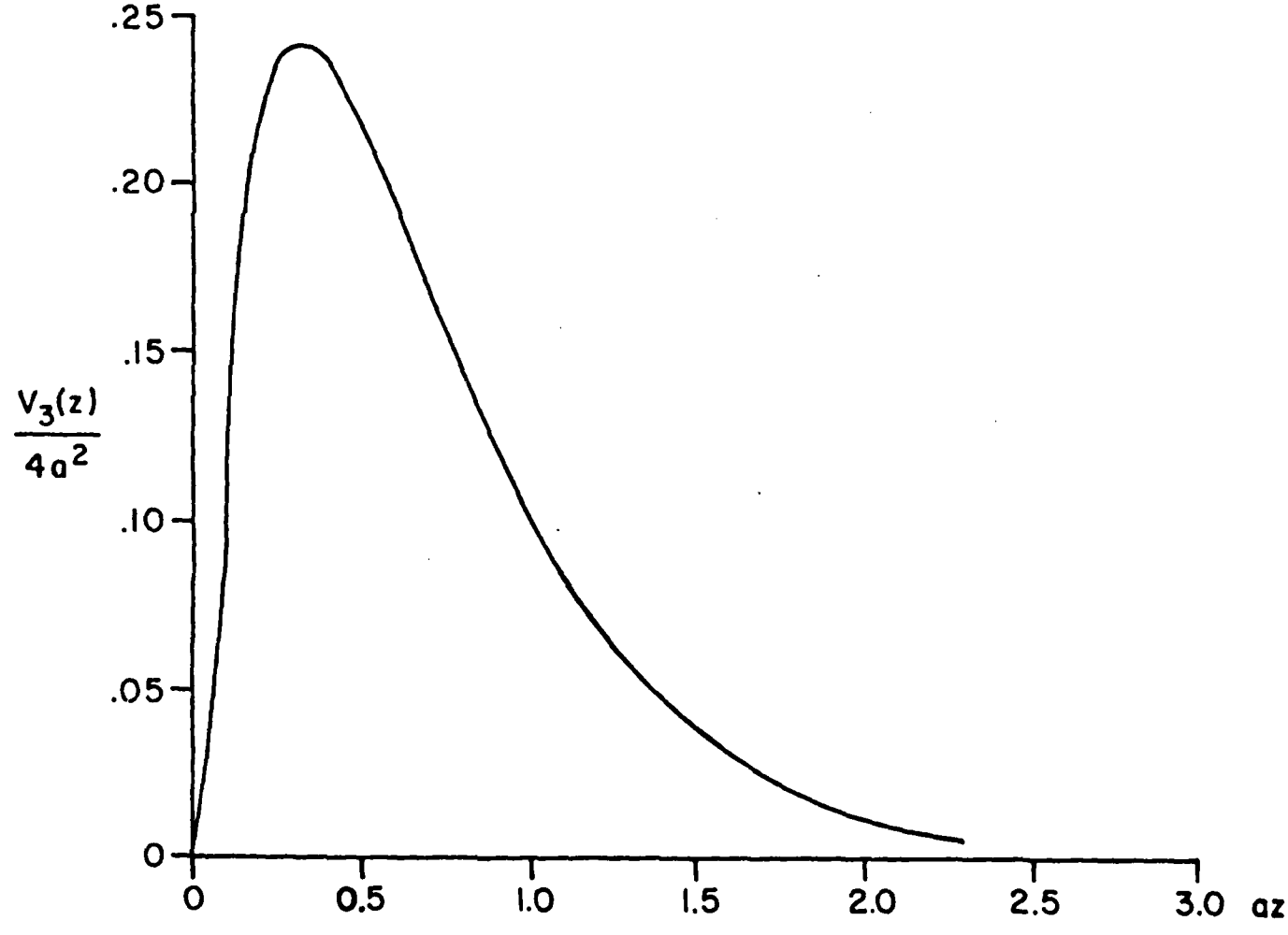


Figure 19. Plot of $V_3(z)$

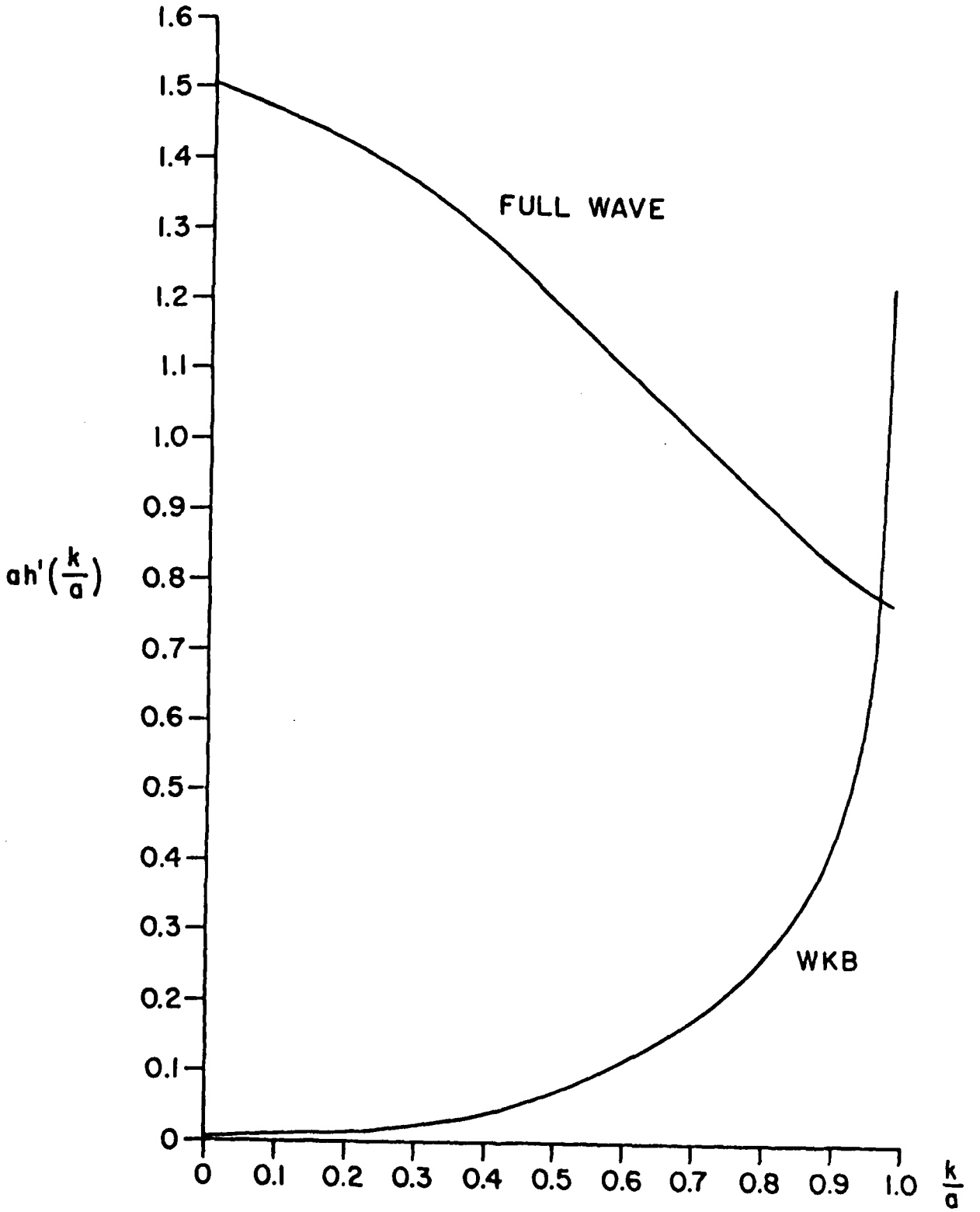


Figure 20. Full Wave and WKB Virtual Heights

3.4 Future Research Directions

Having seen that potentials exist to which the WKB approximation can not be applied, we intend to introduce potentials for which the WKB time-delay has some of the properties of the exact time-delay to find the domain where WKB has meaning.

In addition to obtaining potentials for which the WKB time-delay makes some sense, we shall give examples for even odder behavior of the exact time-delay than those of the present report. They will include potentials for which the time-delay is zero for all k and potentials for which the time-delay is negative: causality seems to be violated. The potentials for these unexpected time-delays look "normal".

The result of these investigations appear to indicate that in many circumstances time-delay is not a particularly useful notion. More useful will be the shape of the returned wave for which the shape of the pulse is known. By deconvolution techniques the reflection coefficient b(k) can be found and one can then use the full-wave solutions. These possibilities will also be examined from a numerical point of view.

4.0 REFERENCES

Becker, W., "On the Manual and Digital Computer Methods Used at Lindau for the Conversion of Multifrequency Ionograms to Electron Density-Height Profiles," Radio Sci., 2, (10), 1205-1232 (1967).

Bibl, K., J. A. Patenaude, G. M. Polucci and B. W. Reinisch, "Digital Data Generation and Processing in Airborne Ionospheric Research," Final Report, AFCRL-TR-73-0203, LTIRF-353/IP (1973).

Bibl, K., and B. W. Reinisch, "The Universal Digital Ionosonde," Radio Sci., 13, (3), 519-530 (1978).

Buchau, J., W. N. Hall, B. W. Reinisch and S. Smith, "Remote Ionospheric Monitoring," in "Effect of the Ionosphere on Space and Terrestrial Systems," edited by John M. Goodman, U.S. Government Printing Office, Washington, D. C., pp. 401-410 (1978).

Budden, K. B., Radio Waves in the Ionosphere, Cambridge University Press (1966).

Courant, R. and D. Hilbert, Methods of Mathematical Physics, Vol. 1, Interscience Division, John Wiley and Sons, New York (1953).

Faddeev, L. D., "The Inverse Problem in the Quantum Theory of Scattering," Dokl. Akad. Nauk. USSR, 121, 66 (1958).

Huang, X. and B. W. Reinisch, "Automatic Calculation of Electron Density Profiles from Digital Ionograms. Part II. True Height Inversion of Topside Ionograms with the Profile-Fitting Method," to be published in Radio Science (1982).

Kanal, M. and H. E. Moses, "A Variational Principle for the Gelfand-Levitan Equation and the Korteweg-deVries Equation," J. Math. Phys., 18, 2445 (1977).

Kay, I., "The Inverse Scattering Problem," Report No. EM-74 of the Institute of Mathematical Sciences of New York University (1954).

Kay, I. and H. E. Moses, "The Determination of the Scattering Potential from the Spectral Measure Function," Nuovo Cimento, X, 3, 276 (1956).

Moses, H. E., "A Generalization of the Gelfand-Levitan Equation for the One-Dimensional Schroedinger Equation," J. Math. Phys., 18, 2243 (1977).

Pechenick, K. R. and J. M. Cohen, "Inverse Scattering with Coinciding-Pole Reflection Coefficients," Physics Department, University of Pennsylvania, preprint (1981).

Reinisch, B. W. and X. Huang, "Automatic Calculation of Electron Density Profiles from Digital Ionograms. Part I. Automatic O and X Trace Identification for Topside Ionograms," to be published in Radio Science (1982).

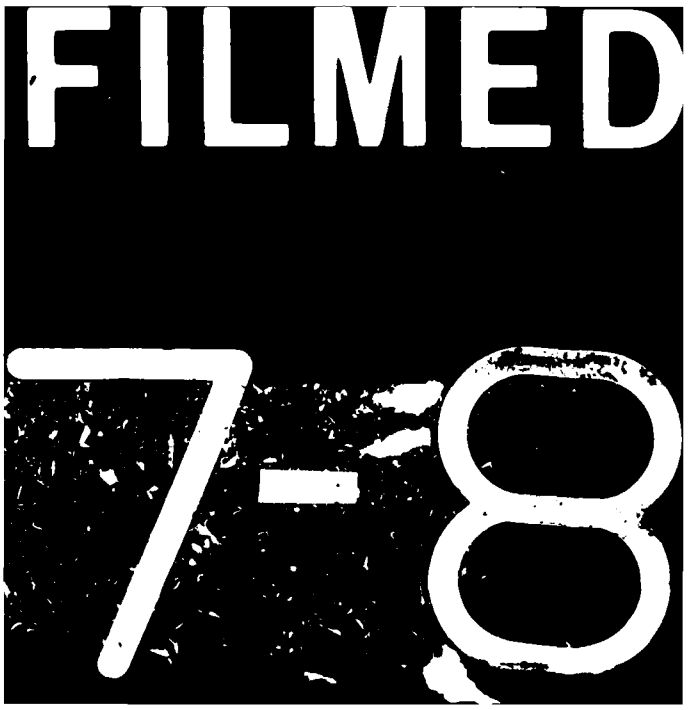
Reinisch, B. W. and S. Smith, "Geomonitor-Digital Real Time Processor for Geophysical Data," AFGL-TR-76-0292 (1976).

Scheid, F., "Numerical Analysis," Schaum's Outline Series, McGraw-Hill (1968).

Smith, S., B. W. Reinisch, J. S. Tang and K. Bibl, "Automatic Ionospheric Parameter Extraction from Digital Ionogram Data," Interim Report, AFGL-TR-80-0102, ULRF-404/CAR (1979).

ACKNOWLEDGEMENT

Parts of the autoscaling algorithm are based on the research by Mr. Huang Xueqin, visiting scholar at ULCAR. Messrs. William N. Hall and Jurgen Buchau of AFGL gave valuable directions for the automatic scaling of ionograms. The authors appreciate the illuminating discussions with Prof. J. M. Cohen of the University of Pennsylvania and Dr. I. Kay (particularly for Section 3.2) of the Institute for Defense Analysis.


FILMED
7-8