ARL-TR-86-22 Copy No. 6

A TEST FOR ALIASING USING BISPECTRAL ANALYSIS

Melvin J. Hinich
Murray A. Wolinsky

APPLIED RESEARCH LABORATORIES
THE UNIVERSITY OF TEXAS AT AUSTIN
POST OFFICE BOX 8029, AUSTIN, TEXAS 78713-8029

11 August 1986

Technical Report

Approved for public release;
distribution unlimited.

Prepared for:

**OFFICE OF NAVAL RESEARCH
DEPARTMENT OF THE NAVY
ARLINGTON, VA 22217**

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER

2. GOVT ACCESSION NO. AD-A176138

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
A TEST FOR ALIASING USING BISPECTRAL ANALYSIS

5. TYPE OF REPORT & PERIOD COVERED
technical report

6. PERFORMING ORG. REPORT NUMBER
ARL-TR-86-22

7. AUTHOR(s)
Melvin J. Hinich
Murray A. Wolinsky

8. CONTRACT OR GRANT NUMBER(s)
N00014-86-K-0237

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Applied Research Laboratories
The University of Texas at Austin
Austin, Texas 78713-8029

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS
Office of Naval Research
Department of the Navy
Arlington, Virginia 22217

12. REPORT DATE
11 August 1986

13. NUMBER OF PAGES
13

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)
aliasing
hypothesis test
bispectrum
bicovariance function
principal domain
support set
bandlimited

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)
Aliasing is the term used for a signal confounding problem that arises when a continuous-time signal is sampled at a rate slower than twice the highest frequency component of a Fourier series representation of the signal. Aliasing can be especially serious for social science time series applications since the sampling designs used to contruct most social science data bases are fixed by considerations other than the nature of the continuous-time mechanisms that generate the observed processes. Once a sampling process is used to collect

cont'd

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. (cont'd)

data, it would be valuable to be able to test the observations for the presence of a significant amount of aliasing. We will show that the overlooked property of the principal domain of a discrete-time bandlimited stationary signal can be used to motivate an amended version of the Hinich bispectrum test for Gaussianity (Hinich, 1982) as a test for aliasing.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

I. INTRODUCTION

The two sinusoids $x(t)=\cos(2\pi f_0 t)$ and $y(t)=\cos(4\pi f_0 t)$ yield identical sampled values if they are sampled at times $t=n/f_0$, where n is an integer. Thus these two functions and the constant function with value one cannot be identified from a discrete-time sample of a time series which is sampled at rate f_0. This is a simple example of the aliasing problem that can arise when a time series is sampled at equal time points. If the time between successive observations is τ (for a sampling rate of $1/\tau$) then each frequency component of the series for frequencies $-1/2\tau \leq f \leq 1/2\tau$ can be confounded by components whose frequencies are $f+n/\tau$ where n is a signed integer. This confounding by aliases will not occur if the underlying continuous-time signal is bandlimited at frequency f_0 and the sampling rate is greater than or equal to the Nyquist frequency $2f_0$, i.e., if $1/\tau \geq 2f_0$. Aliasing is usually avoided in engineering and physical science applications by filtering the signal to eliminate its energy above a certain cutoff frequency and then sampling the filtered signal at twice the cutoff frequency. This type of sampling design is usually impossible for most social science applications since the sampling rate is determined according to logistic and cost considerations without any regard for the spectral characteristics of the underlying process.

Once a sampling process is used to collect data, it would be valuable to be able to test the observations for the presence of a significant amount of aliasing. We will show that an overlooked property of the principal domain of a discrete-time bandlimited stationary signal can be used to motivate an amended version of the Hinich bispectrum test for Gaussianity (Hinich, 1982) as a test for aliasing.

The bispectrum of a discrete-time signal is a periodic function in two frequency indices. There is a surprisingly persistent confusion between the statistics and engineering literature as to the triangular form of the principal domain of the bispectrum of a time series that is a discrete-time sample of a continuous-time bandlimited signal. Huber et al. (1971), Kim and Powers (1979), and Matsuoka and Ulrych (1984) have equilateral triangular domains that are a subset of the triangle given by Brillinger and Rosenblatt (1967), Lii and Rosenblatt (1982), Subba Rao and Gabr (1980), Hinich (1982), and Subba Rao (1983).

We will now demonstrate that the set of positive support of the continuous-time bispectrum is a proper subset of the principal domain of the bispectrum of the sampled process. Huber et al. and others in the engineering literature are really referring to the support set rather than the principal domain, and that support set is the one of substantive interest. The extra triangle in Brillinger and Rosenblatt (1967) is due to the frequencies that sum to one (2π in angular frequency units). We will show that the bispectrum must be zero in that extra triangle if the sampling rate equals or exceeds the Nyquist rate. This result will be explained after some basic definitions are reviewed.

II. CONTINUOUS-TIME BISPECTRA

To review the mathematics of bispectra, let $x(t)$ denote a real zero mean stationary continuous-time stochastic process. Assume that all expected values, sums, and integrals used below exist. The bicovariance function of the process is $c(u,v)=Ex(t)x(t+u)x(t+v)$, which does not depend on t since the process is stationary. Its Fourier transform

$$B(f,g) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} c(u,v)\, \exp[-i2\pi(fu+gv)]\, du\,dv \tag{1}$$

is called its bispectrum. Although this two-frequency index notation is standard, it hides the three-frequency interaction that is so important for applications of bispectral estimation. To help the exposition, we will now switch to the three-index notation used by Brillinger and Rosenblatt (1967), namely, $B(f,g,h)$, where $h=-f-g$. To motivate this notation, consider the following redefinition of the bispectrum as a Fourier transform of $c(t+u,t+v)=c(u,v)$ for all t:

$$\begin{aligned} B(f,g,h) &= \iint c(t+u,t+v)\, \exp\{-i2\pi[f(t+u)+g(t+v)+ht]\}du\,dv \\ &= \iint c(u,v)\, \exp\{-i2\pi[fu+gv+(f+g+h)t]\}du\,dv. \end{aligned} \tag{2}$$

In order for $B(f,g,h)$ to equal $B(f,g)$ regardless of t, then $f+g+h=0$. Note that the right-hand side of expression (2) is invariant to permutations of the frequency indices f, g, and $h=-f-g$. Thus the bispectrum's symmetry lines are $f=g$, $f=h(2f=-g)$, and $g=h(2g=-f)$. Another symmetry holds since $c(u,v)$ is real; namely $B(-f,-g,-h)=B^*(f,g,h)$, where * denotes complex conjugation. This skew symmetry yields another three symmetry lines: $f=-g$, $f=-h(g=0)$, and $g=-h(f=0)$ (see Fig. 1). Thus the cone $C= f,g:0\leq f,\ g\leq f$ is a principal domain of this continuous time bispectrum in the (f,g) plane.

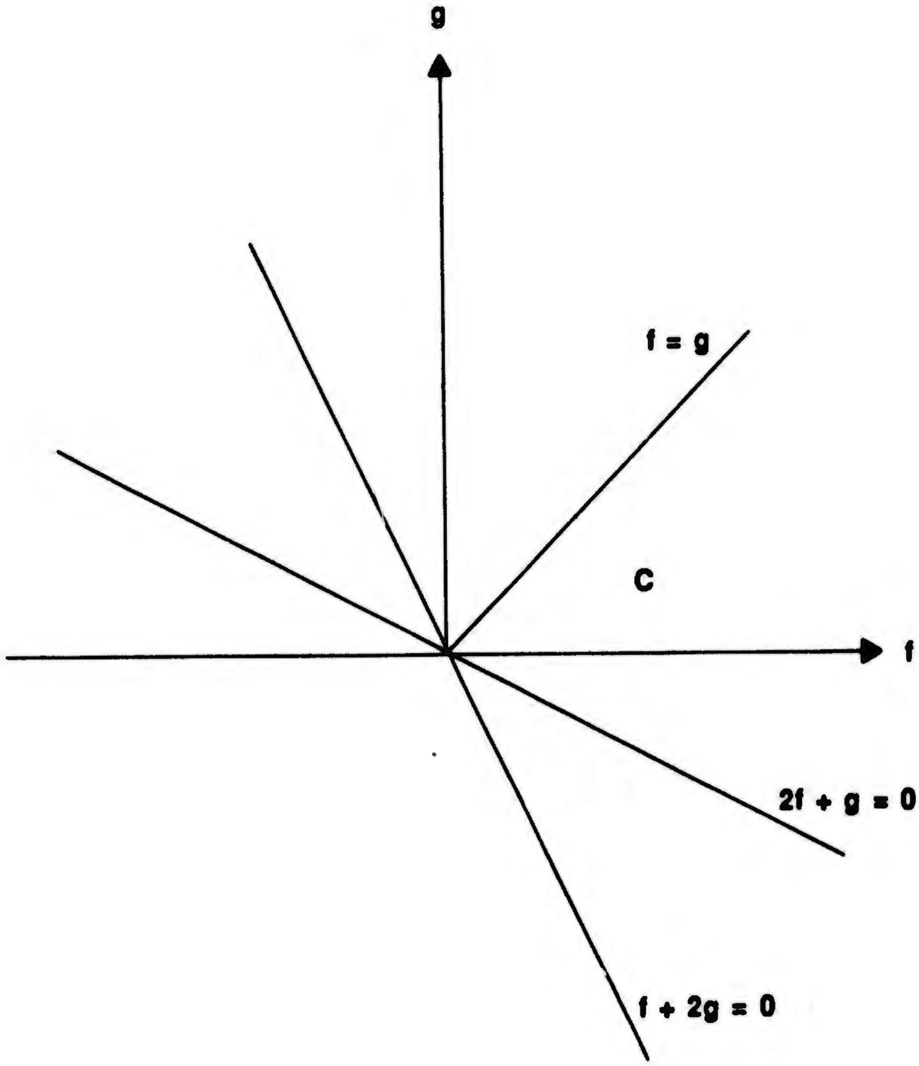

FIGURE 1
SYMMETRIES OF BISPECTRUM B(f,g)

ARL:UT
AS-86-443
MJH - GA
8 - 6 - 86

III. THE PRINCIPAL DOMAIN FOR DISCRETE-TIME

Now suppose that the process is bandlimited at frequency f_0. Then there is no variance in the process for frequencies beyond f_0 and thus the bispectrum cuts off at $f=(+/-)f_0$, $g=(+/-)f_0$, and $f+g=(+/-)f_0$. Then the continuous-time support set is the equilateral triangle $\{f,g:0\leq f\leq f_0,\ g\leq f,\ f+g=f_0\}$. But the discrete-time principal domain is a larger triangle if the process is sampled at the Nyquist frequency $2f_0$.

The principal domain can be derived from Eq. (2) in a straightforward manner (see Rosenblatt, 1983, for a general treatment of polyspectra). Consider the discrete-time sequence $\{x(n\tau)\}$ where $\tau=1/(2f_0)$. The bicovariance function of this sampled version of $\{x(t)\}$ is really an array $\{c(j\tau,k\tau):j,k=0,(+/-)1,(+/-)2,\ldots\}$. Then the sampled-data bispectrum is defined, analogous with Eq. (2), to be given by the Fourier transform in three indices:

$$_{\tau}B(f,g,h) = \Sigma_j\ \Sigma_k\ c(j\tau,k\tau)\ \exp\{-i2\pi[(fj\tau+gk+(f+g+h)\tau]\}, \qquad (3)$$

where now f+g+h is not just constrained to be zero, but can be equal to n/τ for any signed integer n. Thus the sampling introduces an infinite set of parallel symmetry lines $2f+g=n/\tau$, $2f-g=n/\tau$, $f+2g=n/\tau$, and $f-2g=n/\tau$. The cone C is only cut by the symmetries $2f+g=n/\tau$, and thus the principal domain of $_{\tau}B$ is the triangle $\{f,g:0\leq f\leq 1/2\tau,\ g\leq f,\ 2f+g=1/\tau]\}$ in the cone C. We will now give an expression for B in terms of the underlying bispectrum for frequencies in the odd triangle $OT=\{f,g:g\leq f,\ 1/2\tau\leq f+g\ \leq(1/\tau)-f\}$ adjoining the equilateral triangle $ET=\{f,g:g\leq f,\ 0\leq f+g\leq 1/2\tau\}$ (Fig. 2).

The discrete-time bispectrum $_{\tau}B$ is a periodic function of period $1/\tau$ in each of its three indices. A special case of a formula in Brillinger and Rosenblatt (1967, p. 190) is that for f, g, and h in the principal domain of $_{\tau}B$,

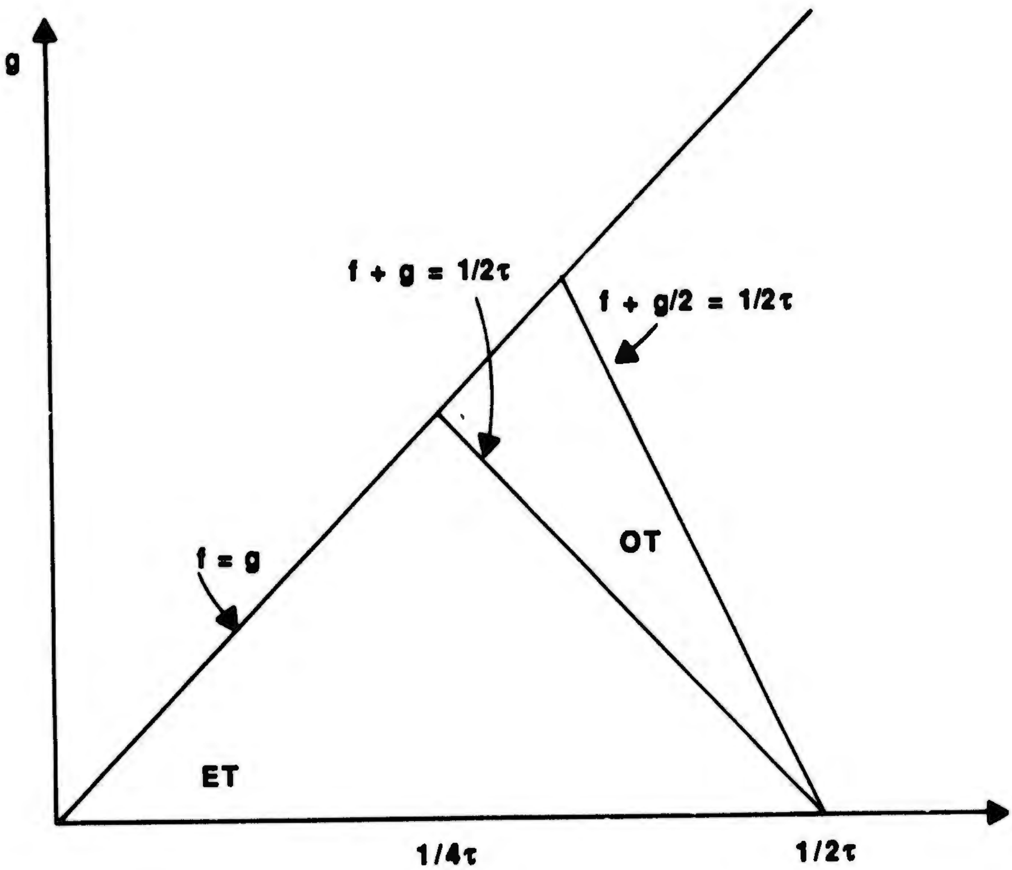

FIGURE 2
DISCRETE-TIME PRINCIPAL DOMAIN

ARL:UT
AS-86-444
MJH - GA
8 - 6 - 86

$$ {}_{\tau}B(f,g,h) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} B(f+k/\tau,\ g+m/\tau,\ h+n/\tau), \text{ with} $$

$$ f+g+h+(k+m+n)/\tau = 0, \qquad (4) $$

where the signed integers are restricted to keep the indices in B's principal domain. For example, if $h=-f-g$, then $k+m+n=0$. If $h=(1/\tau)-f-g$, then $k+m+n=-1$. But B is bandlimited at f_0 and so the sum is restricted to the k, m, and n such that $|f+k/\tau|\le f_0$, $|g+m/\tau|\le f_0$, and $|h+n/\tau|\le f_0$.

Now consider the case when there is no aliasing which is the case if $\tau=1/2f_0$. If f and g are in OT, then $f+g>1/2\tau=f_0$, $f<f_0$, and $g<f_0$. Thus the key term in the sum to consider is $B(f,g,h-1/\tau)$, where $h=-f-g+1/\tau$. But this is $B(f,g,-f-g)$ which is zero since $f+g>f_0$. All the other terms are zero for a similar reason. Thus ${}_{\tau}B$ is zero for (f,g) in OT.

A. Testing for Aliasing

Let $S(f_j,g_k)$ denote a consistent estimator of $B(f_j,g_k)$ for a grid of equally spaced bifrequencies (f_j,g_k) in OT for a sample of size N of $\{x(n\tau)\}$. This estimator can be computed by smoothing the sample bicovariance (Subba Rao, 1983), smoothing the sample bispectrum in the bifrequency domain (Hinich, 1982), or by dividing the sample into pieces and averaging the piecewise sample bispectra and then doing bifrequency smoothing (Lii and Rosenblatt, 1982). Under some mixing conditions such as the ones given by Brillinger (1975) or Rosenblatt (1985), for large N the distribution of $2|S(f_j,g_k)|^2$ is approximately a central chi-square with two degrees of freedom if ${}_{\tau}B(f_j,g_k)=0$. The estimators for the different bifrequencies in the domain are asymptotically independent (see the above references for exact statements of the asymptotic properties alluded to here). The statistic in the Hinich test for Gaussianity is the sum of these approximately chi-square statistics over the points in the principal domain grid. Estimates of the power of this test for several types of nonlinear and non-Gaussian models is given by Ashley, Hinich, and Patterson (1986). The test is applied to stock price data by Hinich and Patterson (1985).

The obvious modification of this test for the problem on hand is to sum the chi-square statistics only over the triangle OT. Under the null hypothesis that there is no aliasing for a sampling rate $1/\tau$, the distribution of the sum is approximately central chi-square with 2K degrees of freedom, where K is the number of grid points in OT. Since Ashley et al. show that the approximation used in the Hinich test is good for samples as small as N=256, there is no reason to doubt its application to this restricted sum test for aliasing. If one does not want to use the large sample properties of the estimated bispectrum, then the Subba Rao and Gabr (1980) test can be modified in a similar way.

In a sense the test has already been applied. The results in Hinich and Patterson (1985) show that many daily stock series have large peaks in various parts of the principal domain, including OT. Since stock prices change by the minute, it is not surprising that high frequency components are aliased. All statistically significant terms in OT verify aliasing for the sampling interval of one day.

B. Analysis of Ten Stock Price Series

To provide an example of our method with new data, we compute our chi-squared alias statistic for rates of return from ten randomly selected stocks using the same sample period of N = 1000 consecutive days for each of these ten stocks. We compute the bispectrum for points in the principal domain, averaging in the frequency domain over a square whose sides have a resolution bandwidth of 0.03 1/day. In other words, 900 raw sample bispectrum values are averaged to produce an estimate of the bispectrum for the square. The spectrum for each stock is averaged using a truncated cosine kernel whose bandwidth is 0.237 1/day. There are K=24 center points in the alias triangle OT. We obtained similar results for our aliasing test using different smoothing parameters for the bispectral and spectral estimates.

The results are presented in Table I. The values for the chi-square statistics with 48 degrees of freedom are given in column 1. Under the

TABLE I

ALIASING TEST STATISTICS

	χ^2_{48} [1]	Z [2]
American Airlines	74.6	5.4
Alberto Culver	95.2	7.0
CBS	43.9	2.5
Cambell Soup	103.5	7.5
El Paso Natural Gas	291.2	17.3
Swift & Co.	74.7	5.4
Federated Department Stores	77.4	5.6
Northern Gas	76.4	5.5
Indianapolis Power and Light	92.7	6.8
Merrill Lynch	54.6	3.6

1. Test statistic is approximately a central chi-squared variate with 48 degrees of freedom under H_0.
2. Test statistic is a normal approximation of chi squared with large d.f's. It is N(0,1) under H_0.

null hypothesis that there is no aliasing, the statistic is approximately chi-squared with 48 degrees of freedom. If there is aliasing the distribution is approximately a noncentral chi squared with 48 degrees of freedom and a positive noncentrality parameter.

The second column presents the Gaussian approximation to a chi-squared random variable with large degrees of freedom. This statistic is approximately N(0,1) under the null hypothesis. Each and every stock series has a test statistic that is not consistent with the null hypothesis. The results overwhelmingly indicate that the data are aliased for the sampling interval of one day. The results suggest that the underlying continuous time mechanism that generates the price series has a spectrum whose bandwidth far exceeds the 0.5 1/day folding frequency for the measured rates of return.

For each series analyzed, the bispectrum in the equilateral triangle is similar in form to that of the OT triangle. The statistically significant bispectral values are concentrated at certain bifrequency pairs as is the case for the bispectra presented in the Hinich-Patterson JBES paper. The results are consistent with a hypothesis that the generating mechanism is nonlinear. The aliasing test statistics indicate that a higher sampling rate must be used to identify the structure of the nonlinear mechanism.

REFERENCES

Ashley, R., M. J. Hinich, and D. Patterson (1986), "A Diagnostic Test for Nonlinear Serial Dependence in Time Series Fitting Errors," Journal of Time Series Analysis 7(3), 165-178.

Brillinger, D. (1975), Time Series, Data Analysis and Theory (Holt, Rinehart, and Winston, New York).

Brillinger, D., and M. Rosenblatt (1967), "Asymptotic Theory of kth Order Spectra," in Spectral Analysis of Time Series, B. Harris, Ed. (John Wiley, New York), 153-188.

Brillinger, D, and M. Rosenblatt (1967), "Computation and Interpretation of kth Order Spectra", in Spectral Analysis of Time Series, B. Harris, Ed. (John Wiley, New York), 189-232.

Hinich, M. J. (1982), "Testing for Gaussianity and Linearity of a Stationary Time Series," J. Time Series Analysis 3(3), 169-176.

Hinich, M. J., and D. M. Patterson (1985), "Evidence of Nonlinearity in Daily Stock Returns," J. Business and Economic Statistics 3(1), 69-77.

Huber, P., B. Kleiner, T. Gasser, and G. Dummermuth (1971), "Statistical Methods for Investigating Phase Relations," IEEE Trans. Audio and Electroacoustics AU 19-20, 78-86.

Kim, Y. C., and E. Powers (1979), "Digital Bispectral Analysis and Its Applications to Nonlinear Wave Interactions," IEEE Trans. Plasma Sci. PS-7, 120-131.

Lii, K. S., and M. Rosenblatt (1982), "Deconvolution and Estimation of Transfer Function Phase and Coefficients for NonGaussian Linear Processes," Ann. Statist. 10(4), 1195-1208.

Matsuoka, T., and T. J. Ulrych (1984), "Phase Estimation Using the Bispectrum," Proceedings IEEE 72(10), 1403-1411.

Rosenblatt, M. (1983), "Cumulants and Cumulant Spectra," in Handbook of Statistics, Vol. 3, D. Brillinger and P. Krishnaiah, Eds. (North-Holland, Amsterdam), Chap. 17.

Rosenblatt, M. (1985), Stationary Sequences and Random Fields (Birkhauser, Boston).

Subba Rao, T. (1983), "The Bispectral Analysis of Nonlinear Stationary Time Series with Reference to Bilinear Time Series Models," in Handbook of Statistics Vol. 3, D. Brillinger and P. Krishaiah, Eds. (North-Holland, Amsterdam), Chap. 14.

Subba Rao, T., and M. Gabr (1980), "A Test for Linearity of Stationary Time Series," J. Time Series Analysis 1, 145-158.

11 August 1986

DISTRIBUTION LIST FOR
ARL-TR-86-22
UNDER CONTRACT N00014-86-K-0237

Copy No.	
	Office of Naval Research Arlington, VA 22217-5000
1 - 3	Attn: Mathematical Sciences Division (Code 411)
	Commander Naval Ocean Systems Center San Diego, CA 92152
4	Attn: D. Gingras (Code 733)
5 - 16	Commanding Officer and Director Defense Technical Information Center Cameron Station, Building 5 5010 Duke Street Alexandria, VA 22314
17	Glen E. Ellis, ARL:UT
18	Marshall E. Frazer, ARL:UT
19	Melvin J. Hinich, ARL:UT
20	Joseph F. Willman, ARL:UT
21	Murray A. Wolinsky, ARL:UT
22	Library, ARL:UT
23 - 27	Reserve, ARL:UT