

# AD-A257 279



## TATION PAGE

Form Approved  
OMB No. 0704-0188

2

do not average 1 hour per response, including the time for reviewing instructions, searching existing data sources, reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Avenue, Washington, DC 20540, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE September 1992		3. REPORT TYPE AND DATES COVERED	
4. TITLE AND SUBTITLE Stationary Time Series Analysis Using Information and Spectral Analysis				5. FUNDING NUMBERS DAAL03-90-G-0069	
6. AUTHOR(S) Emanuel Parzen					
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Texas A+M University College Station, TX 77843-3143					
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U. S. Army Research Office P. O. Box 12211 Research Triangle Park, NC 27709-2211				10. SPONSORING / MONITORING AGENCY REPORT NUMBER ARO 27574.11-MA	
11. SUPPLEMENTARY NOTES The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.					
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited.				12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) The technical aims of this paper are: to discuss some roles of information ideas and spectral analysis in time series analysis (sections 1 and 2); extend spectral estimation by exponential models and extend (to time series) goodness of fit tests by components (sections 3 and 4). Section 0 presents some philosophy. The practice of statistics and time series analysis can "stand on the shoulders of giants" if we develop a framework which unifies diverse methods. Information ideas are central to a unified framework since they clarify and extend methods by providing many levels of relationship between time series analysis, classical statistical methods for independent samples, and signal processing problems called inverse problems with positivity constraints.					
14. SUBJECT TERMS Entropy; Renyi information; Stationary time series; Spectral exponential models; Spectral components				15. NUMBER OF PAGES 15	
				16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED		18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED		19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	
				20. LIMITATION OF ABSTRACT UL	

149K



**STATIONARY TIME SERIES ANALYSIS  
USING INFORMATION AND SPECTRAL ANALYSIS**

Technical Report # 187

September 1992

Emanuel Parzen

Texas A&M Research Foundation

Project No. 6547

Sponsored by the U. S. Army Research Office

Professor Emanuel Parzen, Principal Investigator

Approved for public release; distribution unlimited

**STATIONARY TIME SERIES ANALYSIS  
USING INFORMATION AND SPECTRAL ANALYSIS**

EMANUEL PARZEN

Department of Statistics, Texas A&M University

*Dedicated to Maurice Priestley on his 60th Birthday*

**0. Introduction to philosophical and technical aims and means**

The technical aims of this paper are: to discuss some roles of information ideas and spectral analysis in time series analysis (sections 1 and 2); extend spectral estimation by exponential models and extend (to time series) goodness of fit tests by components (sections 3 and 4). Section 0 presents some philosophy.

We believe that a major problem of statistical theory is how to develop technology transfer from esoteric methods to exoteric methods. We define exoteric statistical methods as those that have reached the status of a *consumer product*, where the consumers are applied researchers. Esoteric methods are known mainly to experts who are researching the theory and are often alleged to be an *intellectual game*. More methods need to reach the status of consumer products (applicable methods) because computing power enables us to apply several methods to a real problem and reduces the personal investment required to learn how to apply a new method. It should now be possible to implement the growing consensus that problem solving by comparison of several methods leads to conclusions which have increased confidence.

The practice of statistics and time series analysis can "stand on the shoulders of giants" if we develop a framework which unifies diverse methods. Information ideas are central to a unified framework since they clarify and extend methods by providing many levels of relationship between time series analysis, classical statistical methods for independent samples, and signal processing problems called inverse problems with positivity constraints.

---

Research supported by U. S. Army Research Office

We believe that statisticians who work in time series analysis may find that their audience is not their fellow statisticians. While their work is appreciated by many researchers in the many fields in which time series analysis is applied and developed, they may feel not valued by the majority of statisticians (to whom the extensive literature of time series analysis seems to be disjoint from the main stream of statistical methods). My experience is that statisticians should understand that time series methods provide many of the right foundations for the successful unification of statistical methods. I feel fortunate to have studied time series analysis intensively before beginning (in 1977) my current work on non-parametric data modeling, unification of statistical methods, and *change* analysis.

Another benefit that I have derived from working in time series analysis has been the friendship of Maurice Priestley and his wife Nancy since we first met in 1958. As I express my esteem for Maurice Priestley and honor his 60th birthday, let me commend Priestley (1981) as the best book to read to learn about time series analysis in both the time and frequency domains.

### 1. Entropy, Cross-Entropy, Renyi Information

The pioneer work of Kullback (1959) made statisticians aware of the fundamental role in statistical theory played by the (Kullback-Liebler) information divergence between two probability distributions  $F$  and  $G$ ; we define it by a definition which differs from usual definitions by a factor of 2:

$$I(F; G) = (-2) \int_{-\infty}^{\infty} \log\{g(x)/f(x)\} f(x) dx$$

when  $F$  and  $G$  are continuous with probability density functions  $f(x)$  and  $g(x)$ ;

$$I(F; G) = (-2) \sum \log\{p_G(x)/p_F(x)\} p_F(x)$$

when  $F$  and  $G$  are discrete, with probability mass functions  $p_F(x)$  and  $p_G(x)$ . One reason for the importance of information divergence (among other possible definitions) is that it possesses a decomposition

$$I(F; G) = H(F; G) - H(F),$$

in terms of entropy  $H(F)$  and cross-entropy  $H(F; G)$ ; we define

$$H(F) = (-2) \int_{-\infty}^{\infty} \{\log f(x)\} f(x) dx,$$

$$H(F; G) = (-2) \int_{-\infty}^{\infty} \{\log g(x)\} f(x) dx.$$

The fundamental work of Renyi (1961) deserves more attention by statisticians; we define Renyi information of index  $\lambda$  as follows for continuous  $F$  and  $G$ :

$$IR_{\lambda}(F; G) = \frac{2}{\lambda(1+\lambda)} \log \int \left( \left\{ \frac{g(y)}{f(y)} \right\}^{1+\lambda} - (1+\lambda) \left\{ \frac{g(y)}{f(y)} - 1 \right\} \right) f(y) dy$$

for  $\lambda \neq 0, -1$ ;

$$IR_0(F; G) = 2 \int \left\{ \frac{g(y)}{f(y)} \log \frac{g(y)}{f(y)} - \frac{g(y)}{f(y)} + 1 \right\} f(y) dy$$

$$IR_{-1}(F; G) = -2 \int \left\{ \log \frac{g(y)}{f(y)} - \frac{g(y)}{f(y)} + 1 \right\} f(y) dy$$

An analogous definition holds for discrete  $F$  and  $G$ .

When we think of these formulas as relating non-negative functions  $f$  and  $g$ , we denote it by  $IR_{\lambda}(f; g)$ . This definition provides: (1) extensions to non-negative functions which are not densities, and also (2) a non-negative integrand which can provide diagnostic measures at each value of  $y$ . The above definitions hold also for multivariate  $F$  and  $G$ .

## 2. Asymptotic Information of Stationary Normal Time Series

This section discusses the elegance of formulas provided by a unification of information measures of stationary normal time series and information measures of non-negative functions which are spectral density functions.

When a time series  $\{Y(t), t = 1, 2, \dots\}$  is modeled by alternative probability measures  $P_1$  and  $P_2$  for the infinite sequence, we define asymptotic information divergence (or rate of information divergence)

$$\text{AsymIR}_{\lambda}(P_2; P_1) = \lim_{n \rightarrow \infty} (1/n) IR_{\lambda}(P_2^{(n)}; P_1^{(n)}).$$

where  $P_i^{(n)}$  is the multivariate distribution under  $P_i$  of  $Y(t), t = 1, \dots, n$ .

Let  $Y(\cdot)$  be zero mean stationary with covariance function

$$R(v) = E[Y(t)Y(t-v)].$$

and correlation function

$$\rho(v) = R(v)/R(0).$$

Parzen ((1981), (1983)) shows how empirical modeling of a time series from data can be approached by using concepts of information to measure the predictability of  $Y(t)$  from past values  $Y(t-1), \dots, Y(t-m)$ . We define the information about  $Y(t)$  in its infinite past  $Y(t-1), Y(t-2), \dots$ , and in its finite past by

$$I_\infty = I(Y|Y_{-1}, Y_{-2}, \dots) = E_{Y_{-1}, Y_{-2}, \dots} I(f_{Y|Y_{-1}, Y_{-2}, \dots}; f_Y) = \lim_{m \rightarrow \infty} I_m,$$

$$I_m = I(Y|Y_{-1}, \dots, Y_{-m}) = E_{Y_{-1}, \dots, Y_{-m}} I(f_{Y|Y_{-1}, \dots, Y_{-m}}; f_Y)$$

where  $f_{Y|Y_{-1}, \dots, Y_{-m}}$  denotes the conditional probability density of  $Y(t)$  given  $Y(t-1), \dots, Y(t-m)$ .

**An important classification of time series is by memory type: no memory, short memory, long memory according as  $I_\infty = 0$ ,  $0 < I_\infty < \infty$ ,  $I_\infty = \infty$ .** Memory type is related to the dynamic range of the spectral density function of the time series.

The spectral density function  $f(\omega)$ ,  $0 \leq \omega < 1$ , is defined as the Fourier transform of the correlation function (assuming it exists):

$$f(\omega) = \sum_{v=-\infty}^{\infty} \exp(-2\pi i v \omega) \rho(v)$$

We call a time series *bounded memory* if the spectral density is bounded above and below:

$$0 < c_1 \leq f(\omega) \leq c_2 < \infty.$$

Long memory occurs when there are zeroes or infinities in  $f(\omega)$ . Bounded memory is intuitively the same as short memory or finite memory.

Let  $P_f$  denote the probability measure on the space of infinite sequences  $R_\infty$  corresponding to a *normal* zero mean stationary time series with spectral density function  $f(\omega)$ .

An important result of Pinsker [(1964), p. 196] can be interpreted as providing a formula for asymptotic information divergence between two zero mean stationary time series with respective rational spectral density functions  $f(\omega)$  and  $g(\omega)$ . Write  $\text{AsymIR}_\lambda(f, g)$  for  $\text{AsymIR}_\lambda(P_f; P_g)$ . Adapting Pinsker (1964) one can prove that

$$\text{AsymIR}_{-1}(f, g) = \int_0^1 \{(f(\omega)/g(\omega)) - 1 - \log(f(\omega)/g(\omega))\} d\omega$$

The right hand side is called by electrical engineers the Itakura-Saito formula and plays an important role in signal processing.

Because spectral densities are even functions we can take the integral to be over  $0 \leq \omega < .5$ ; then one obtains the following important theorem.

*Theorem:* Unification of information measures of Pinsker (1964) and Itakura-Saito (1970).

$$\text{AsymIR}_{-1}(f, g) = \text{IR}_{-1}(f(\omega)/g(\omega))_{0,.5}$$

The validity of this identity of information measures can be extended to non-normal asymptotically stationary time series (Ephraim, Lev-Ari, Gray (1988)). It can be called Pinsker's information theoretic justification of the Itakura-Saito distortion measure. Statisticians might try to understand these formulas as extensions of the formula for the information divergence between two univariate normal distributions with zero means and different variances.

For bounded memory time series (and  $-1 < \lambda < 0$ ), Kazakos and Kazakos (1980) prove

$$\text{AsymIR}_\lambda(f, g) = (1/\lambda) \int_0^1 \left\{ \log(f(\omega)/g(\omega)) - (1/(1+\lambda)) \log \left\{ 1 + (1+\lambda) \left( (f(\omega)/g(\omega)) - 1 \right)_+ \right\} \right\} d\omega$$

Kazakos and Kazakos (1980) also give formulas for asymptotic information of multiple stationary time series. They illustrate how formulas for  $-1 < \lambda < 0$  can lead to formulas for  $\lambda = 0$  and  $\lambda = 1$ .

### 3. Estimation of Finite Parameter Spectral Densities

This section formulates in terms of our Renyi information notation the classic asymptotic maximum likelihood Whittle theory of time series parameter estimation.

For a random sample of a random variable with unknown probability density  $f$ , maximum likelihood estimators  $\hat{\theta}$  of the parameters of a finite parameter model  $f_\theta$  of the probability density  $f$  can be shown to be equivalent to minimizing

$$IR_{-1}(f^{\sim}, f_\theta)$$

where  $f^{\sim}$  is a raw estimator of  $f$  (initially, a symbolic sample probability density formed from the sample distribution function  $F^{\sim}$ ). A similar result, called Whittle's estimator (Whittle (1953)), holds for estimation of spectral densities of a bounded memory zero mean stationary time series for which one assumes a finite parametric model  $f_\theta(\omega)$  for the true unknown spectral density  $f(\omega)$ .

A raw fully nonparametric estimator of  $f(\omega)$  from a time series sample  $Y(t)$ ,  $t = 1, \dots, n$ , is the sample spectral density (or periodogram)

$$f^{\sim}(\omega) = \left| \sum_{t=1}^n Y(t) \exp(-2\pi i \omega t) \right|^2 \div \sum_{t=1}^n |Y(t)|^2$$

Note that  $f^{\sim}(\omega)$  is not a consistent estimator of  $f(\omega)$ ; nevertheless,

$$E[f^{\sim}(\omega)] \text{ converges to } f(\omega),$$

a fact which can be taken as the *definition of the spectral density*  $f(\omega)$ .

An estimator  $\hat{\theta}$  which is asymptotically equivalent to maximum likelihood estimator is obtained by minimizing  $\text{Asym}IR_{-1}(f^{\sim}; f_\theta) = IR_{-1}(f^{\sim}, f_\theta)_{0.5} =$

$$\int_0^1 \left\{ (f^{\sim}(\omega)/f_\theta(\omega)) - 1 - \log(f^{\sim}(\omega)/f_\theta(\omega)) \right\} d\omega$$

which can be interpreted as choosing  $\theta$  to make  $\frac{f^{\sim}(\omega)}{f_\theta(\omega)}$  as flat or constant as possible.



Whittle pioneered the representation of a parametric model for a Gaussian stationary time series with zero mean and unit variance

$$f_{\theta}(\omega) = \sigma^2 / \gamma_{\theta}(\omega)$$

where  $\gamma_{\theta}(\omega)$  is the square modulus of the transfer function of the whitening filter represented by the spectral density model  $f_{\theta}$ , and  $\sigma^2$  is defined to satisfy

$$\int_0^1 \log f_{\theta}(\omega) d\omega = \log \sigma^2 = -I_{\infty}.$$

Note that  $0 < \sigma^2 < 1$  and  $\sigma^2$  is the ratio of the conditional variance of  $Y(\cdot)$  given its infinite past and the unconditional variance of  $Y(\cdot)$ .

Minimizing  $\text{AsymIR}_{-1}(f^{\sim}, f_{\theta})$  is equivalent to minimizing

$$\left(1/\sigma^2\right) \int_0^1 \{f^{\sim}(\omega)\gamma_{\theta}(\omega)\} d\omega + \log(\sigma^2)$$

which is equivalent to minimizing over  $\theta$

$$\sigma_{\theta}^2 = \int_0^1 \gamma_{\theta}(\omega) f^{\sim}(\omega) d\omega$$

and setting

$$\sigma^{-2} = \int_0^1 \gamma_{\theta^{\sim}}(\omega) f^{\sim}(\omega) d\omega = \sigma_{\theta^{\sim}}^2.$$

The information divergence between the data and the fitted model is given by

$$\text{IR}_{-1}(f^{\sim}, f_{\theta^{\sim}}) = \log \sigma_{\theta^{\sim}}^2 - \log \sigma^{-2} = I_{\infty}^{\sim} - I_{\infty}^{\wedge}$$

defining  $-I_{\infty}^{\wedge} = \log \sigma^{-2}$ ,

$$-I_{\infty}^{\sim} = \log \sigma^{-2} = \int_0^1 \log f^{\sim}(\omega) d\omega$$

This criterion (however, corrected for bias in  $I_{\infty}^{\sim}$ ) arises from information approaches to model identification (Parzen (1983)). A model fitting criterion (but not a parameter estimation criterion) is provided by the information increment

$$\begin{aligned} & I(Y | \text{all past } Y; Y \text{ values in model } \theta) \\ &= \int_0^1 -\log \{f^{\sim}(\omega)/f_{\theta^{\sim}}(\omega)\} d\omega = \text{IR}_{-1}(f^{\sim}/f_{\theta^{\sim}})_{0,-.5} \end{aligned}$$

One can regard it as a measure of the distance of the whitening spectral density

$$f^*(\omega) = \hat{f}(\omega)/\hat{f}_{\theta^*}(\omega)$$

from a constant function; note that  $f^*(\omega)$  is constructed to integrate to 1.

Our philosophy of identification of parametric models for a stationary time series is to choose them so that by the criteria of the next section of this paper an optimal smoother of the computed function  $f^*(\omega)$  is a constant; then a "parameter-free" non-parametric estimator of the spectral density  $f(\omega)$  by a smoother of  $\hat{f}(\omega)$  is given by the parametric estimator  $\hat{f}_{\theta^*}$ . By "parameter-free" we mean that we are free to choose the parameters to make the data (raw estimator) shape up to a smooth estimator. The parameters are not regarded as having any significance or interpretation; they are merely coefficients of a representation of  $f(\omega)$ .

Portmanteau statistics to test goodness of fit of a model to the time series use sums of squares of correlations of residuals; an analogous statistic is

$$IR_1(\hat{f}/\hat{f}_{\theta^*})_{0,.5} = \log \int_0^{.5} \{\hat{f}(\omega)/\hat{f}_{\theta^*}(\omega)\}^2 d\omega$$

Goodness of fit of the model to the data (as measured by how close  $f^*(\omega)$  is to the spectral density of white noise) is the ultimate model identification criterion to decide between competing parametric models.

#### 4. Goodness of fit by components and exponential models

Our philosophy of stationary time series model identification argues that goodness of fit tests of a model should test for whiteness

$$f^*(\omega) = \hat{f}(\omega)/\hat{f}_{\theta^*}(\omega).$$

In this section we propose an analogue of the concept of *components* introduced in the classical goodness of fit theory by Durbin and Knott (1972):

$$T^*(J) = 2^{-.5} \int_0^1 f^*(\omega) J(\omega) d\omega$$

for various score functions  $J(\omega)$ , which one evaluates in practice as a sum over a dense grid of frequencies.

One usually forms a sequence of components whose score functions

$$J_0(\omega) = 1, J_1(\omega), J_2(\omega), \dots$$

are a complete orthonormal set of functions in  $L_2[0, 1]$ . Choices are: harmonics ( $\cos 2\pi j\omega, j = 0, 1, 2, \dots$ ); Legendre polynomials; Hermite polynomial functions of the standard normal quantile function  $\Phi^{-1}$ . Note that  $\int_0^1 J_j(\omega) d\omega = 0$  for  $j = 1, 2, \dots$ .

Under the assumption that  $f(\omega) = f(\omega; \theta)$  for some parameter vector  $\theta$ , the asymptotic distribution of  $T^*(J_j)$  is the same as the spectral average  $2^{-.5} \int_0^1 (1/f(\omega)) J_j(\omega) f(\omega) d\omega$ ; the latter are asymptotically normal with mean

$$\int_0^1 J_j(\omega) d\omega = 0,$$

and variance

$$(2/n) \int_0^1 (1/2f^2(\omega)) |J_j(\omega)|^2 f^2(\omega) d\omega = (1/n) \int_0^1 |J_j(\omega)|^2 d\omega = 1/n.$$

Thus properly defined components are asymptotically independent  $\text{Normal}(0, 1/n)$ .

A component based quadratic test of the goodness of fit of the model, with an asymptotic chi-square distribution, is

$$S_{k,m} = \sum_{j=k}^m |T^*(J_j)|^2.$$

These component tests have the asymptotic optimality properties of score tests if we model the true spectral density  $f(\omega)$  by an exponential model extending Bloomfield (1973).

*The main technical contribution of this paper may be the following proposal: estimate  $f(\omega)$  by assuming an exponential model of order  $m$  using score functions  $J_j(\omega)$ ,  $j = 1, \dots, m$ . The choice of score functions and criteria for determining from the data an optimal order  $\hat{m}$  require further research.*

Note that an exponential model for the spectral density provides smooth estimators of the log spectral density and therefore of cepstral correlations (the Fourier coefficients

of the log spectrum) and coefficients of the AR( $\infty$ ) and MA( $\infty$ ) representations of a time series required for prediction.

An exponential model of order  $m$ , denoted  $f_{\theta,m}$ , is defined

$$\log f_{\theta,m}(\omega) = \theta_0 + \theta_1 J_1(\omega) + \dots + \theta_m J_m(\omega)$$

The coefficient  $\theta_0$  has the interpretation  $\theta_0 = \int_0^1 \log f(\omega) d\omega = \log \sigma_\infty^2$  where  $\sigma_\infty^2$  is the infinite memory one step ahead prediction mean square error. An exponential model can be expressed

$$f_{\theta,m} = \sigma_\infty^2 \exp\left(\sum_{j=1}^m \theta_j J_j(\omega)\right)$$

Maximum likelihood estimators  $\theta^m = (\hat{\theta}_1, \dots, \hat{\theta}_m)$  of  $\theta^m = (\theta_1, \dots, \theta_m)$  are equivalent to minimizing

$$V(\theta) = \int_0^1 d\omega f^-(\omega) \exp\left(-\sum_{j=1}^m \theta_j J_j(\omega)\right)$$

and then estimating  $\sigma_\infty^2$  by  $V(\hat{\theta}^m)$ .

An estimated spectral density is given by

$$f_{\hat{\theta}^m}(\omega) = V(\hat{\theta}^m) \exp\left(\sum_{j=1}^m \hat{\theta}_j J_j(\omega)\right);$$

it satisfies

$$\int_0^1 f^{*m}(\omega) d\omega = 1,$$

defining

$$f^{*m}(\omega) = f^-(\omega) / f_{\hat{\theta}^m}(\omega)$$

The product of the Fisher score function (derivative with respect to  $\theta_j$  of the optimization criterion  $V(\theta)$ ) and  $2^{-.5}$  is denoted  $U_j(\theta)$ ; for  $j = 1, \dots, m$

$$U_j(\theta) = 2^{-.5} \int_0^1 d\omega (f^-(\omega) / f_\theta(\omega)) J_j(\omega)$$

A goodness of fit test of a model of order  $m$  is given by a score test of an order  $m$  sub-model against an order  $M$  "full" model:

$$U_j(\hat{\theta}^m) = 0, j = m + 1, \dots, M$$

An overall chi-square test uses the sum of squares of these score statistics. To compute the parameter estimators, let

$$U(\theta^m) = (U_1(\theta^m), \dots, U_m(\theta^m)).$$

An approximate Newton-Raphson iterative scheme for computing can be shown, following Bloomfield (1973), to be

$$\theta^{m(n+1)} = \theta^{m(n)} - .5U(\theta^{m(n)})$$

Note that the vector of correction terms in this iteration is the vector of score tests. *Exponential models for the spectral density use the same score statistics for iterative evaluation of estimators as are used for component tests of goodness of fit.*

An initial estimator of  $\theta_j$ , adapting Bloomfield (1973), is

$$\theta_j^{(1)} = (1/n) \sum_{t=1}^n \log f(2\pi t/n) J_j(2\pi t/n)$$

We conclude our heuristic discussion by emphasizing that the approach outlined above to goodness of fit and spectral density estimation needs further research about the problems of choosing score functions  $J_j(\omega)$  and determining an optimal order  $m$ .

## REFERENCES

- Akaike, H. (1974). "A new look at the statistical model identification," *IEEE Trans. Autom. Contr.*, AC-19, 716-723.
- Bloomfield, P. (1973). "An exponential model for the spectrum of a scalar time series." *Biometrika*, 60, No. 2, 217-226.
- Durbin, J. and Knott, S. (1972). "Components of Cramér-von Mises statistics I", *J. Roy. Statist. Soc. Ser. B*, 34, pp. 290-307.
- Itakura, F. and Saito, S. (1970). "A statistical method for estimation of speech spectral density and format frequencies," *Electron. Commun. Japan*, 53-A, 36-43.
- Jones, L. K. (1989). "Approximation theoretic derivation of logarithmic entropy principles for inverse problems and unique extension of the maximum entropy method to incorporate prior knowledge," *SIAM J. Appl. Math.*, vol. 49, pp. 650-661.
- Jones, L. K. and Byrne, C. L. (1990). "General entropy criteria for inverse problems, with applications to data compression, pattern classification, and cluster analysis," *IEEE Transactions on Information Theory*, vol. 36, no. 1, pp. 23-30.
- Kazakos, D. and Papantoni-Kazakos, P. (1980). "Spectral distance measures between Gaussian processes," *IEEE Trans. Automat. Contr.*, vol. AC-25, no. 5, pp. 950-959 .
- Kullback, S. (1959). *Information Theory and Statistics*. New York: Wiley.
- Newton, H. J. (1988). *TIMESLAB: A Time Series Analysis Laboratory*, Wadsworth: Pacific Grove, California.
- Parzen, E. (1967). *Time Series Analysis Papers*, Holden-Day, San Francisco, California.
- Parzen, E. (1958) "On asymptotically efficient consistent estimates of the spectral density function of a stationary time series" *J. Royal Statist. Soc., Ser. B.*, 20, 303-322.
- Parzen, E. (1962). "Spectral analysis of asymptotically stationary time series", *Bull. Inst. Internat. Statist.*, 39, 2, 87-103.
- Parzen, E. (1969). "Multiple time series modeling" *Multivariate Analysis - II*, edited by P. Krishnaiah, Academic Press: New York, 289-409.
- Parzen, E. (1974). "Some Recent Advances in Time Series Modeling", *IEEE Transactions on Automatic Control*, AC-19, 723-730.
- Parzen, E. (1977). "Multiple Time Series: Determining the Order of Approximating Autoregressive Schemes", *Multivariate Analysis - IV*, Edited by P. Krishnaiah, North Holland: Amsterdam, 283-295.
- Parzen, E. (1979). "Nonparametric Statistical Data Modeling", *Journal of the American Statistical Association*, (with discussion), 74. 105-131.
- Parzen, E. (1981). "Time Series Model Identification and Prediction Variance Horizon," *Proceedings of Second Tulsa Symposium on Applied Time Series Analysis*. Academic Press: New York. 425-447.
- Parzen, E. (1982). "Maximum Entropy Interpretation of Autoregressive Spectral Densities," *Statistics and Probability Letters*, 1, 2-6.
- Parzen, E. (1983). "Time Series Model Identification by Estimating Information," *Studies in Econometrics, Time Series, and Multivariate Statistics in Honor of T. W. Anderson*, ed. S. Karlin, T. Amemiya, L. Goodman, Academic Press: New York, 279-298.
- Parzen, E. (1983). "Time Series ARMA Model Identification by Estimating Information" *Proceedings of the 15th Annual Symposium on the Interface of Computer Science and*

*Statistics*, Amsterdam: North Holland.

- Parzen, E. (1983). "Time Series Model Identification by Estimating Information, Memory, and Quantiles," *Questo*, 7, 531-562.
- Parzen, E. (1986). "Quantile Spectral Analysis and Long Memory Time Series," *Journal of Applied Probability*, Vol. 23A, 41-55.
- Parzen, E. (1992) "Comparison Change Analysis". *Nonparametric Statistics and Related Topics* (ed. A. K. Saleh), Elsevier: Amsterdam, 33-15.
- Pinsker, M. S. (1964). *Information and Information Stability of Random Variables and Processes*. San Francisco, CA: Holden-Day.
- Priestley, M. B. (1981). *Spectral Analysis and Time Series*, Academic Press: London.
- Renyi, A. (1961). "On measures of entropy and information." *Proc. 4th Berkeley Symp. Math. Statist. Probability, 1960*, 1, 547-561. University of California Press: Berkeley.
- Renyi, A. (1967). "On some basic problems of statistics from the point of view of information theory," *Proc. 5th Berkeley Symp. on Math., Stat. and Probability*, 531-543.
- Shore, J. (1981). "Minimum cross-entropy spectral analysis," *IEEE Trans Acoust. Speech, Signal Processing*, vol. ASSP-29, no. 2, pp. 230-237.
- Whittle, P. (1953). "Estimating and information in stationary time series." *Ark. Math.* 2, 423-434.
- Whittle, P. (1953). "The analysis of multiple stationary time series," *J. Royl Statist. Soc.*, B. 15:125-139.

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability Codes	
Dist.	Avail and/or Special
A-1	

DTIC QUALITY INSPECTED 2