T-IT/37/1//39289

Two Alternative Philosophies for Estimation of the Parameters of Time-Series

William A. Gardner

Reprinted from
IEEE TRANSACTIONS ON INFORMATION THEORY
Vol. 37, No. 1, January 1991

Two Alternative Philosophies for Estimation of the Parameters of Time-Series

William A. Gardner, *Senior Member, IEEE*

Abstract—**Two alternative philosophical frameworks for the two problems of estimating the time-invariant or time-variant autocorrelation function and its Fourier transform for stationary and cyclostationary time-series are briefly compared. One is based on the stochastic process model, and the other is based on the nonstochastic time-series model. It is then explained that results on estimator bias and variance for these two problems couched within the stochastic process framework have analogs within the nonstochastic framework. The bias and variance results for cyclostationary time-series that are available within these two frameworks are then briefly summarized.**

Index Terms—**Estimation philosophies, nonstochastic estimation, time-series analysis, ergodicity, cyclostationarity, cycloergodicity.**

I. Introduction

The study of the reliability of estimates (e.g., estimates of the autocorrelation function or its Fourier transform) obtained from single time-series can be approached from two philosophically different points of view. The orthodox point of view that comes from the conventional probabilistic theory of statistical inference is that the time-series we use to obtain our estimate in practice is but one sample path from an infinite ensemble of possible sample paths from a hypothetical stationary (or cyclostationary) stochastic process—a mathematical creation—and the estimation method (which uses a single time-series-segment of specified length) is reliable if and only if we would get approximately the same value for our estimate (with probability one) regardless of which sample path from the hypothetical stochastic process we were to use. The alternative point of view is that there exists one and only one time-series—the one we in fact use to obtain our estimate—and our estimation method is reliable if and only if we would get approximately the same value for our estimate, regardless of which time-segment of specified length from the hypothetically infinitely long time-series (also a mathematical creation) we were to use.

The orthodox point of view becomes amenable to mathematical analysis once a particular probabilistic model of the hypothetical stochastic process is adopted. Similarly, the alternative point of view becomes amenable to mathematical analysis once a particular fraction-of-time probabilistic model of the single hypothetically infinitely long time-series is adopted. These two philosophical approaches are duals when the stochastic process models are ergodic (or cycloergodic) and in a certain sense are mathematically isomorphic [1], [2], [3, Chapter 8], [4, Chapters 1, 10], [5], [6]. It thus should not be surprising that the development of reliability theory (e.g., the study of bias, variance, confidence intervals, etc.) for both parametric and nonparametric estimation can proceed within either of the two mathematical frameworks, and can for the most part be translated from one framework to the other (as explained in [4, Chapter 5]). But this does not occur much in practice. It is common for investigators to choose one philosophical viewpoint to the exclusion of the other.

A case in point is the recent paper [7], which adopts the orthodox stochastic process framework and presents results on estimation of the autocorrelation function and its Fourier transform for cyclostationary processes[1] that have analogs within the alternative framework. The purpose of this note is to clarify the relationship between the results (within the stochastic process framework) on reliability of estimates presented in [7] and the results (within the alternative framework) presented in [3], [4], [8]. In Section II, the results from [3], [4], [8], obtained within the nonstochastic framework, are briefly summarized. In Section III, the results from [3] and [7], obtained within the stochastic process framework, are briefly summarized, and then the relationship between the results in [3] and [7] and the results in [3], [4], and [8] is explained.

II. Nonstochastic Estimation [2]

A time-series $x(t)$ defined for $-\infty < t < \infty$ is said to be of *second order* if the limit (using x^* to denote the complex conjugate of x)

$$\hat{R}_x^\alpha(\tau) \triangleq \lim_{T\to\infty} \frac{1}{T}\int_{t-T/2}^{t+T/2} x(u+\tau/2)x^*(u-\tau/2)e^{-i2\pi\alpha u}\,du, \tag{1}$$

which is called the *cyclic autocorrelation function*, exists for all real τ and α, is not zero for some τ and α, and is continuous in τ at $\alpha = \tau = 0$ (and, therefore, for all α and τ).

If $\hat{R}_x^\alpha(\tau)$ is not identically zero for only $\alpha = 0$, then $x(t)$ is said to be *purely stationary*. If $\hat{R}_x^\alpha(\tau)$ is not identically zero for only $\alpha = nT_o$ for some integers n including $n \neq 0$, then $x(t)$ is said to be *purely cyclostationary* with period T_o. If $\hat{R}_x^\alpha(\tau) \neq 0$ for some values of α that are not all integer multiples of one nonzero value, then $x(t)$ is said to be *almost cyclostationary* with multiple incommensurate periods.

[1]The term *cyclostationary* is used here instead of the term *periodically correlated*, which is used in [7], because the former predates the latter (cf. [9] and references therein), is in more common use in engineering, and is more versatile (since it admits the modifiers *wide sense*, *nth order*, and *strict sense*).

[2]The material in this section is taken from [4] and [8].

Manuscript October 1989; revised May 1990.

The author is with the Department of Electrical Engineering and Computer Science, University of California, Davis, CA 95616.

IEEE Log Number 9039289.

0018-9448/91/0100-0216$01.00 © 1991 IEEE

Let us consider the normalized complex envelope

$$X_V(t,f) \triangleq \frac{1}{\sqrt{V}} \int_{t-V/2}^{t+V/2} x(u) e^{-i2\pi f u} \, du \tag{2}$$

corresponding to the spectral component of $x(t)$ with center frequency f and bandwidth on the order of $1/V$. If the limit

$$\hat{S}_x^\alpha(f) \triangleq \lim_{V\to\infty} \lim_{U\to\infty} \frac{1}{U} \cdot \int_{t-U/2}^{t+U/2} X_V(u, f+\alpha/2) X_V^*(u, f-\alpha/2) \, du \tag{3}$$

of the correlation of the normalized complex envelopes at frequencies $f+\alpha/2$ and $f-\alpha/2$ exists, and if the Fourier transform

$$F\{\hat{R}_x^\alpha(\tau)\} = \int_{-\infty}^{\infty} \hat{R}_x^\alpha(\tau) e^{-i2\pi f \tau} \, d\tau$$

exists, then we have

$$\hat{S}_x^\alpha(f) = F\{\hat{R}_x^\alpha(\tau)\}. \tag{4}$$

That is, the *spectral correlation density function* $\hat{S}_x^\alpha(f)$ and the cyclic autocorrelation function $\hat{R}_x^\alpha(\tau)$ are a Fourier transform pair. As a special case, when $\alpha = 0$, $\hat{R}_x^\alpha(\tau)$ is the conventional autocorrelation function and $\hat{S}_x^\alpha(f)$ is the conventional power spectral density function, and (4) is known as the *Wiener relation*. Consequently, for $\alpha \neq 0$, (4) is called the *cyclic Wiener relation*.

In practice the integration times T, U, V must be finite. As a result, the finite-time cyclic autocorrelation function and spectral correlation density function must be considered to be estimates of the corresponding limits. These estimates, unlike their limits, fluctuate in the time parameter t. The difference between the time-averaged (over $-\infty < t < \infty$) value (the temporal mean) and the limit of each estimate is called the *temporal bias* of the estimate, and the difference between the time-averaged squared deviation of each estimate about its temporal mean is called the *temporal variance*.

The temporal bias is completely characterized by the set of all nonzero limits $\hat{R}_x^\alpha(\tau)$ or, equivalently, $\hat{S}_x^\alpha(f)$. However, the temporal variance depends on the fourth-order cyclic moments of $x(t)$ as well as these second-order cyclic moments. However, if $x(t)$ has Gaussian fraction-of-time distributions (cf. [4]), then the temporal variance like the temporal bias is completely characterized by the set of $\hat{R}_x^\alpha(\tau)$ or the set of $\hat{S}_x^\alpha(f)$.

In [4] and [8], the general class of estimators

$$\tilde{S}_x^\alpha(t,f) \triangleq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} k_f^\alpha(t,u,v) x(t-u) x^*(t-v) \, du \, dv, \tag{5}$$

where

$$k_f^\alpha(t,u,v) = m\left(\frac{u+v}{2}, v-u\right) e^{-i2\pi f(v-u)} e^{i\pi\alpha(2t+u+v)} \tag{6}$$

for some kernel $m(\cdot,\cdot)$, is considered. For example, for $\alpha = 0$, this reduces to a general class of estimators for the power spectral density function that includes as special cases all of the commonly known estimators (e.g., continuous-time counterparts of the methods of Bartlett–Welch, Blackman–Tukey, Wiener–Daniell, etc.). Each particular estimation method is specified by a particular kernel $m(\cdot,\cdot)$. A wide variety of specific examples, including time-averaged and frequency-smoothed cyclic periodograms (the cyclic periodogram is the integrand in (3)) as well as the Fourier transform of the lag-windowed cyclic correlogram, are studied in [4] and [8]. For this general class of estimators of the spectral correlation density function $\hat{S}_x^\alpha(f)$, explicit formulas for the temporal bias and, for the Gaussian model, the temporal variance are derived and studied at length. The formulas are completely specified in terms of $\hat{S}_x^\alpha(f)$ and the double Fourier transform of $m(\cdot,\cdot)$. It is shown that, for $\Delta t \Delta f \gg 1$ and Δf small enough for a window of width Δf to resolve $\hat{S}_x^\alpha(\cdot)$, the variance is inversely proportional to $\Delta t \Delta f$. Here Δt and $1/\Delta f$ are measures of the widths of $m(\cdot,\cdot)$ in its first and second variables, respectively. For example, using (2) and (3), without the limits, as an estimate $\tilde{S}_x^\alpha(t,f)$ results in $\Delta t = U$ and $\Delta f = 1/V$.

The dependence of the bias and variance of $\tilde{S}_x^\alpha(t,f)$ on $\hat{S}_x^\beta(g)$ for values of β that differ from α by more (or less) than $\pm 1/2\Delta t$ is identified as *cycle leakage* (or inadequate *cycle resolution*) analogous to the spectral leakage and inadequate spectral resolution corresponding to the dependence on $\hat{S}_x^\beta(g)$ for values of g differing from f by more (or less) than $\pm\Delta f/2$. This leads to an explanation of how cycle leakage, cycle resolution, spectral leakage, spectral resolution, and reliability (variance) can be traded off by selection of the estimator parameters Δt and Δf and the design of window functions which make up $m(\cdot,\cdot)$.

All of the preceding is carried out in [4] not only for the auto spectral correlation density function of complex-valued time-series, but also for the cross spectral correlation density function obtained by replacing the conjugated factors x^* and X_V^* in (1), (3), and (5) with y^* and Y_V^*, respectively, for some time-series $y(t)$ other than $x(t)$.

III. Stochastic Estimation

A zero-mean stochastic process $x(t)$ defined for $-\infty < t < \infty$ is said to be of second order if the expected value

$$R_x(t,\tau) \triangleq E\{x(t+\tau/2) x^*(t-\tau/2)\} \tag{7}$$

exists for all t and τ and is not identically zero. To avoid anomalous behavior, consideration is usually restricted to second-order processes for which $R_x(t,\tau)$ is continuous.

If $R_x(t,\tau)$ is periodic in t with period, say T_o, then $x(t)$ is said to be *cyclostationary in the wide sense* (or to be *periodically correlated*). In this case, the Fourier coefficients

$$R_x^\alpha(\tau) \triangleq \frac{1}{T_o} \int_{-T_o/2}^{T_o/2} R_x(t,\tau) e^{-i2\pi\alpha t} \, dt \tag{8}$$

for $\alpha = n/T_o$ are called the *probabilistic cyclic autocorrelation functions*. Observe that if the time-series in Section II are sample paths of a stochastic process, then when the limit (1) exists (e.g., in the mean-square sense), we have

$$R_x^\alpha(\tau) = E\{\hat{R}_x^\alpha(\tau)\} \tag{9}$$

(e.g., in the mean-square sense). By analogy with (4), the Fourier transform

$$S_x^\alpha(f) \triangleq F\{R_x^\alpha(\tau)\}, \tag{10}$$

when it exists, can be considered to be a probabilistic spectral correlation density function.

If $R_x(t,\tau)$ is almost periodic in t, then (8) must be generalized to

$$R_x^\alpha(\tau) \triangleq \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} R_x(t,\tau) e^{-i2\pi\alpha t} \, dt, \tag{11}$$

but (9) still holds.

When the Fourier series associated with the periodic or almost periodic function $R_x(\cdot,\tau)$ converges (in some sense), then we can make the association

$$R_x(t,\tau)=\sum_{\alpha} R_x^{\alpha}(\tau)e^{i2\pi\alpha t}, \tag{12}$$

where α ranges over all values for which (8) or (11) is nonzero. As an aside, it is pointed out that the analogous quantity in the nonstochastic framework is, from (1),

$$\hat{R}_x(t,\tau)=\sum_{\alpha}\hat{R}_x^{\alpha}(\tau)^{i2\pi\alpha t}.$$

The preceding material (for real $x(t)$) is taken from [3], where it is also pointed out that for any finite-mean-square measurement function $z(t)$, such as

$$z(t)=x(t+\tau/2)x^*(t-\tau/2), \tag{13}$$

for which the estimate

$$\hat{m}_z^{\alpha}(T)\triangleq\frac{1}{T}\int_{t-T/2}^{t+T/2}z(u)e^{-i2\pi\alpha u}\,du \tag{14}$$

exists in the mean-square sense, the variance is given by

$$\operatorname{var}\{\hat{m}_z^{\alpha}(T)\}=\frac{1}{T^2}\int_{t-T/2}^{t+T/2}\int_{t-T/2}^{t+T/2}K_z\left(\frac{u+v}{2},u-v\right)\cdot e^{-i2\pi\alpha(u-v)}\,du\,dv, \tag{15}$$

where

$$K_z(t',\tau')=R_z(t',\tau')-m_z(t'+\tau'/2)m_z^*(t'-\tau'/2) \tag{16}$$

is the covariance and $m_z(t')=E\{z(t')\}$ is the mean of the process $z(t)$ (cf. [3, (12.193)]. Since $\hat{m}_z^{\alpha}(T)$ is simply the estimate $\hat{R}_x^{\alpha}(\tau)$ obtained from (1) with finite T, then the variance of this estimate converges to zero as $T\to\infty$ if and only if (15), with (13) substituted in, converges to zero. This is so, regardless of whether $x(t)$ is cyclostationary, almost cyclostationary, or more generally nonstationary but exhibiting cyclostationarity (in which case $R_x(t,\tau)$ contains additive periodic components in addition to other nonstationary fluctuations in t). In addition to these relatively straightforward results on mean-square cycloergodicity, analogous but more technical results on cycloergodicity with probability one are given in [6] (for discrete-time stochastic processes).

The Fourier transform of the lag-windowed cyclic correlogram is considered in [7] as an estimate of the probabilistic spectral correlation density function $S_x^{\alpha}(f)$, and for a real Gaussian cyclostationary process it is shown that the variance of this estimate converges to zero $O([\Delta t\Delta f]^{-1})$ as $\Delta t\Delta f\to\infty$ and the bias converges to zero as $\Delta f\to 0$, where Δt is the length of the data segment used to obtain the cyclic correlogram and Δf is the reciprocal of the width of the lag-window. It is also shown that if $\Delta f=1/\Delta t$ (no lag windowing), then the variance does not converge to zero as $\Delta t\to\infty$. Since the formulas for bias and (for a Gaussian process) variance derived for nonstochastic estimation in [4] and [8] are precisely the same for stochastic estimation when the process exhibits mean-square cycloergodicity of the autocorrelation [3, Section 12.7][3], then these results in [7] are analogs of the results obtained in [4] and [8]. The results in [4] and [8] apply to a more general class of estimators (which is described in Section II and includes the estimator in [7] as one special case among numerous special cases treated) and a more general class of time-series (which includes almost cyclostationary as well as cyclostationary).

Since the applicability of bias and variance analyses for nonstochastic time-series, such as those in [4] and [8], which are summarized here in Section II, to stochastic processes, such as in [7], as summarized in this section, requires cycloergodicity (or ergodicity for stationary processes), this note is concluded with a philosophical comment on cycloergodicity.

It can be argued that in all those situations for which a cycloergodic (or ergodic) stochastic process model is appropriate, the single-time-series point-of-view is also appropriate (cf. [4, Chapters 1, 10]); and in those situations where a nonergodic model is essential, the question of perfectly reliable (probabilistic variance converging to zero as averaging time approaches infinity) estimation is ill-posed. Furthermore, without the presence of at least local cyclostationarity and local cycloergodicity (or local stationarity and local ergodicity), or some other form of underlying cyclostationarity (or stationarity), such as asymptotic, in the stochastic process model, there is no sense in even studying reliability of estimates obtained from single time-series. Thus, the question of whether or not estimates with probabilistic variances that converge to zero exist (cf. [7]) is really a question of whether or not one has proposed an appropriate stochastic process model, that is, a model with appropriate cyclostationarity and cycloergodicity (or stationarity and ergodicity) properties. Nevertheless, this question of the appropriateness of a stochastic process model can be relevant in the nonstochastic framework because it is not in general obvious (although in many specific cases it is obvious) whether or not a proposed fraction-of-time probabilistic model (e.g., a set of finite-order fraction-of-time probability distributions) can indeed be obtained from some time-series. If the model is self consistent and has the cycloergodic (or ergodic) property, then there does indeed exist a time-series with that model, namely, a sample path of the stochastic process with the same model, (e.g., the same finite-order distributions).

[3]A zero-mean Gaussian process exhibits mean-square cycloergodicity of the autocorrelation if and only if (15), with $z(t)$ specified by (13), converges to zero for all α. This is guaranteed by the satisfaction of the hypothesis used in [7] to show that the variance of the estimate $S_x^{\alpha}(f)$ converges to zero (although this hypothesis is only sufficient, not necessary).

References

[1] E. M. Hofstetter, "Random processes," in *The Mathematics of Physics and Chemistry*, Vol. II, H. Margenau and G. M. Murphy, Eds. Princeton, N.J.: D. Van Nostrand Co., 1964, ch. 3.

[2] H. O. A. Wold, "On prediction in stationary time series," *Ann. Math. Stat.*, vol. 19, pp. 558–567, 1948.

[3] W. A. Gardner, *Introduction to Random Processes with Applications to Signals and Systems*, New York: Macmillan, 1985 (2nd ed., New York: McGraw-Hill, 1989).

[4] ——, *Statistical Spectral Analysis: A Nonprobabilistic Theory*. Englewood Cliffs, N.J.: Prentice-Hall, 1987.

[5] W. A. Gardner and W. A. Brown, "Fraction-of-time probability for time-series that exhibit cyclostationarity," *Signal Processing*, in press.

[6] R. A. Boyles and W. A. Gardner, "Cycloergodic properties of discrete-parameter nonstationary stochastic processes," *IEEE Trans. Inform. Theory*, vol. IT-29, no. 1, pp. 105–114, Jan. 1983.

[7] H. L. Hurd, "Nonparametric time series analysis for periodically correlated processes," *IEEE Trans. Inform. Theory*, vol. 35, no. 2, pp. 350–359, Mar. 1989.

[8] W. A. Gardner, "Measurement of spectral correlation," *IEEE Trans. Acoust. Speech and Signal Processing*, vol. ASSP-34, no. 5, pp. 1111–1123, Oct. 1986.

[9] W. A. Gardner and L. E. Franks, "Characterization of cyclostationary random signal processes," *IEEE Trans. Inform. Theory*, vol. IT-21, no. 1, pp. 4–14, Jan. 1975.