THE WIENER FILTER FOR LOCALLY STATIONARY STOCHASTIC PROCESSES IS RARELY LOCALLY STATIONARY

Patrik Wahlberg and Peter J. Schreier

School of Electrical Engineering and Computer Science
The University of Newcastle
Callaghan, NSW 2308, Australia
firstname.lastname@newcastle.edu.au
web: www.peter-schreier.com

ABSTRACT

The Wiener filter (i.e., linear minimum mean squared error filter) for wide-sense stationary stochastic processes is translation-invariant, i.e., its impulse response, like the covariance function, is only a function of the time-shift. We investigate whether there is a generalization of this result to continuous-time stochastic processes that are locally stationary in Silverman's sense: Is the optimal filter for locally stationary processes locally stationary itself? The answer is surprisingly negative: Even though the optimal filter can be locally stationary in special cases, it rarely is, even when the covariance functions have Gaussian shape.

1. INTRODUCTION

We consider complex-valued zero-mean second-order continuous-time stochastic processes $z(t)$ that are locally stationary (LS) in Silverman's sense [9, 10]. This means that the covariance function has the structure $r_{zz}(t,s) = E(z(t)z^*(s)) = f((t+s)/2)\rho(t-s)$. Thus f describes the global time behavior and ρ describes the local time behavior of the process. Locally stationary processes are a generalization of wide-sense stationary (WSS) processes where $f \equiv 1$. That is, the covariance function of WSS processes only depends on the time-shift $t-s$.

We investigate linear minimum mean squared error (LMMSE) filtering of a corrupted stochastic process, where we linearly estimate a signal $x(t)$ from a measurement process $z(t)$. The LMMSE criterion leads to the well-known family of integral equations [11]

$$r_{xz}(t,s) = \int_{\mathbb{R}} h(t,u) r_{zz}(u,s)du, \tag{1}$$

indexed by $t,s \in \mathbb{R}$, where $r_{xz}(t,s) = E(x(t)z^*(s))$ is the cross-covariance function between signal and measurement, and $h(t,u)$ is the impulse response (filter kernel) of the possibly nonstationary optimum filter. An important special case is $z(t) = x(t) + n(t)$, where $n(t)$ is a noise process that is uncorrelated with $x(t)$. Then $r_{zz}(t,s) = r_{xx}(t,s) + r_{nn}(t,s)$ and $r_{xz}(t,s) = r_{xx}(t,s)$. If $z(t)$ is WSS, and $z(t)$ and $x(t)$ are jointly WSS, then $h(t,s) = h_0(t-s)$ and (1) amounts to the convolution equation

$$r_{xz}(t) = (h_0 * r_{zz})(t),$$

which can be solved by means of Fourier transformation.

This work was supported by the Australian Research Council (ARC) under the Discovery Project DP0664365.

WSS stochastic processes are very popular models of physical phenomena, mainly because of the huge arsenal of useful mathematical results available for them. However, in practice the WSS assumption is only approximately and never exactly satisfied, so more flexible tools are needed. The class of LS stochastic processes is one such class where the type of nonstationarity has a manageable form. Many other extensions of WSS processes exist in the literature, for example the class of harmonizable processes [5]. There are also several concepts that are distinct from Silverman's LS processes but still bear the name "locally stationary process," e.g. by Dahlhaus [2], and Mallat, Papanicolaou and Zhang [6].

This piece of work is based on the following question: In the WSS case the optimal filter inherits its translation-invariant form $h(t,s) = h_0(t-s)$ from $r_{xz}(t-s)$ and $r_{zz}(t-s)$. Is there a generalization of this result to LS processes? That is, if $r_{zz}(t,s) = f_1((t+s)/2)\rho_1(t-s)$ and $r_{xz}(t,s) = f_2((t+s)/2)\rho_2(t-s)$, does the optimal filter kernel have the same LS structure $h(t,s) = f((t+s)/2)\rho(t-s)$?

In general, equation (1) is difficult to solve, so we introduce the restricting assumption that f_1 and f_2 are Gaussians and ρ_1 and ρ_2 modulated Gaussians. This assumption admits explicit computations in (1) and still comprises interesting cases where $z(t)$ has bandpass character, for a range of possible combinations of bandwidth and duration. These combinations are restricted due to the Cauchy-Schwarz inequality [12]

$$|f_1(0)|^2|\rho_1(2\tau)|^2 \le f_1(\tau)f_1(-\tau), \tag{2}$$

where $\tau = t-s$ is the time-shift parameter. In fact, if f_1 is a function of short duration then $z(t)$ has short duration. The inequality (2) implies that ρ_1 has short duration as well, which means that a WSS process with covariance ρ_1 is wideband. Since $z(t)$ can be approximated locally by a multiple of this WSS process, $z(t)$ will be wideband as well.

However, the basic question whether there is an easy generalization of the WSS results to the LS case turns out to have a surprisingly negative answer: The optimal filter $h(t,s)$ has the LS structure only for very special parameter combinations for the Gaussians f_1 and f_2 and modulated Gaussians ρ_1 and ρ_2. In particular, there is a solution with the LS structure only if the modulated Gaussians ρ_1 and ρ_2 have identical modulation parameter (center frequency)—which implies that ρ has the same modulation parameter—in striking contrast to the WSS case where the optimal filter adapts its modulation parameter to an arbitrary difference of the modulation parameters of ρ_1 and ρ_2.

For simplicity, we assume that $z(t)$ and $x(t)$ are jointly proper [7, 8] (which is also called circular or circularly symmetric), i.e.,

$$E(z(t)z(s)) \equiv 0 \text{ and } E(x(t)z(s)) \equiv 0.$$

If they are improper, then a natural problem is to extend our results to take into acccount the complementary covariance functions $E(z(t)z(s))$ and $E(x(t)z(s))$. Then the optimum filter must include a conjugate-linear term in addition to the linear estimator.

The organization of the paper follows. First we introduce some notation and background in Section 2. Then we introduce LS processes in Section 3, and discuss the general LMMSE filtering problem in Section 4. We present our main result, the fact that LMMSE filtering of LS processes is rarely LS, in Section 5.

2. PRELIMINARIES

The translation operator is denoted by

$$(T_x f)(t) = f(t-x), \quad x \in \mathbb{R},$$

the modulation operator by

$$(M_\xi f)(t) = e^{2\pi j \xi t} f(t), \quad \xi \in \mathbb{R},$$

and the dilation (or scaling) operator by

$$(D_\alpha f)(t) = |\alpha|^{1/2} f(\alpha t), \quad \alpha \in \mathbb{R} \setminus 0.$$

Gaussian functions are denoted

$$g_a(x) = \exp(-2\pi a x^2), \quad x \in \mathbb{R}, \quad a \geq 0.$$

A function $f : \mathbb{R}^2 \mapsto \mathbb{C}$ is *nonnegative definite* [1], $f \in \mathrm{NND}(\mathbb{R}^2)$, if f is continuous and

$$\sum_{j,k=1}^{n} f(t_j, t_k) z_j z_k^* \geq 0, \quad \{t_j\}_{j=1}^n \subset \mathbb{R}, \quad \{z_j\}_{j=1}^n \subset \mathbb{C}, \quad n > 0. \tag{3}$$

A function $f : \mathbb{R} \mapsto \mathbb{C}$ of one real variable is nonnegative definite, denoted $f \in \mathrm{NND}(\mathbb{R})$, if $f_0 \in \mathrm{NND}(\mathbb{R}^2)$ where $f_0(t,s) = f(t-s)$. We have

$$|f(x)| \leq f(0), \quad x \in \mathbb{R}, \quad f \in \mathrm{NND}(\mathbb{R}). \tag{4}$$

A function f is $\mathrm{NND}(\mathbb{R}^2)$ if and only if it is the covariance function of a mean-square continuous stochastic process [5]. The following lemma is well known [5] and will be needed in Section 5.

Lemma 1 *If $f, g \in \mathrm{NND}(\mathbb{R}^2)$ then $f \cdot g \in \mathrm{NND}(\mathbb{R}^2)$.*

3. LOCALLY STATIONARY PROCESSES

Definition 2 *A* locally stationary (LS) process *[9, 10] is a complex-valued stochastic process $z(t)$ whose covariance function r_{zz} has the form*

$$r_{zz}(t,s) = f\left(\frac{t+s}{2}\right) \rho(t-s). \tag{5}$$

We normalize $\rho(0) = 1$ without loss of generality. With the tensor product notation $(f \otimes g)(t,s) = f(t)g(s)$ and the isometric (but not orthogonal) coordinate transformation on $\mathbb{R}^2$

$$\kappa(x,y) = (x + y/2, x - y/2)$$
$$\kappa^{-1}(x,y) = \left(\frac{x+y}{2}, x-y\right),$$

we may write (5) shorter as

$$r_{zz} = (f \otimes \rho) \circ \kappa^{-1}. \tag{6}$$

Note that $f(t) = E|z(t)|^2 \geq 0$ for all t. Continuity of f and ρ is equivalent to mean-square continuity of the process z [5]. When f is constant, we recover WSS processes as a special case of LS processes. Another extreme is $\rho = \delta_0$, where δ_0 denotes the Dirac delta distribution, which allows any nonnegative function as f and then r_{zz} is the covariance of time-variable white noise. The fact that $f(t)$ describes the signal power $E|z(t)|^2$ justifies the name *global* time variable for t in the reformulation of (5)

$$r_{zz}(t + \tau/2, t - \tau/2) = f(t)\rho(\tau). \tag{7}$$

If the global time variable is frozen at $t = t_0$ then $f(t_0)\rho(\tau)$ is the covariance function of a WSS process that approximates $z(t)$ around $t = t_0$ [12], which justifies the term *local* time variable for the time shift τ [4].

Suppose that the LS process $z(t)$ is *harmonizable* [5], which means that the covariance function has the Fourier representation

$$r_{zz}(t,s) = \iint_{\mathbb{R}^2} e^{2\pi j (t\xi - s\eta)} m_{zz}(d\xi, d\eta),$$

where m_{zz} is the spectral covariance function of two frequency variables. Then it can be shown that

$$m_{zz} = (\mathscr{F}\rho \otimes \mathscr{F}f) \circ \kappa^{-1},$$

where $\mathscr{F}$ denotes Fourier transformation [9, 12]. Thus m_{zz} has the same structure as r_{zz}, and a comparison with (6) reveals that m_{zz} is built from the Fourier transforms of the components f, ρ of r_{zz}, with an interchange of roles. This means that the global time variable corresponds to the local frequency variable, and the local time variable corresponds to the global frequency variable. Moreover, we obtain a time-frequency description of LS processes. In fact, the Wigner–Ville spectrum [3], defined by

$$W(t,\xi) = \int_{\mathbb{R}} r_{zz}(t + \tau/2, t - \tau/2) e^{-2\pi j \tau \xi} d\tau,$$

is then separable (rank-one), or more precisely, $W(t,\xi) = f(t)(\mathscr{F}\rho)(\xi)$.

4. OPTIMAL NONCAUSAL FILTERING

Suppose that $z(t)$ is a stochastic process that is a measurement of a corrupted signal $x(t)$. An estimator of $x(t)$ from $z(t)$ using a linear filter defined by a kernel h is given by

$$\widehat{x}(t) = \int_{\mathbb{R}} h(t,s) z(s) ds. \tag{8}$$

The optimum filter problem [11, 14, 15] consists of finding a filter kernel h that minimizes the MSE $E|\widehat{x}(t)-x(t)|^2$ for each $t \in \mathbb{R}$. We do *not* impose that the filter kernel (time-varying impulse response) $h(t,s)$ be causal, where causality means that $h(t,s)=0$ for $t<s$. We make the following assumptions:

$$\begin{aligned} &\text{(i)} \quad h \in L^2(\mathbb{R}^2); \\ &\text{(ii)} \quad \int_{\mathbb{R}} r_{zz}(s,s)ds < \infty; \\ &\text{(iii)} \quad z(t) \quad \text{is mean square continuous.} \end{aligned} \tag{9}$$

Then $E|\widehat{x}(t)-x(t)|^2$ is minimized for all $t \in \mathbb{R}$ if and only if the family of integral equations

$$r_{xz}(t,s) = \int_{\mathbb{R}} h(t,u) r_{zz}(u,s)du, \quad \text{a.e. } t,s \in \mathbb{R}, \tag{10}$$

is satisfied [11,14]. Here $r_{xz}(t,s)=E(x(t)z^*(s))$ is the cross-covariance function between x and z.

If z is WSS and x and z are jointly WSS then the estimator (8) is replaced by

$$\widehat{x}(t) = \int_{\mathbb{R}} h(t-s)z(s)ds = (h*z)(t) \tag{11}$$

where now the filter kernel h is defined on $\mathbb{R}$. Under the assumptions (9), where (i) is modified as $h \in L^2(\mathbb{R})$, (10) simplifies to the convolution equation

$$r_{xz}(t) = (h*r_{zz})(t), \quad t \in \mathbb{R}.$$

If we assume that $r_{xz} \in L^2(\mathbb{R})$ and that there exist $a,b>0$ such that $0<a \le \mathscr{F}r_{zz}(\xi) \le b < \infty$ when $\xi \in \operatorname{supp}(\mathscr{F}r_{xz})$, then the equation may be solved by Fourier transformation $\mathscr{F}$, which yields

$$\mathscr{F}h(\xi) = \frac{\mathscr{F}r_{xz}(\xi)}{\mathscr{F}r_{zz}(\xi)} \in L^2(\mathbb{R}).$$

This filter is commonly called the Wiener filter, even though Wiener's original work [15] concerns the problem when h is causal, i.e., $h(t)=0$ for $t<0$.

5. THE WIENER FILTER FOR LS PROCESSES IS RARELY LS

In the WSS case the optimal filter kernel h has the same structure as the covariance functions: It depends on the difference of the arguments only. It is natural to pose the question if the Wiener filter for LS processes also has the LS kernel structure. That is, we solve the family of integral equations (10) when r_{xz} and r_{zz} have the LS structure, and we ask whether there exist $f,\rho,g,\gamma,h,\lambda$ such that

$$(h \otimes \lambda) \circ \kappa^{-1} \in \mathrm{NND}(\mathbb{R}^2)$$

(to guarantee that r_{zz} is a covariance function) and

$$\begin{aligned} &f((t+s)/2)\rho(t-s) \\ &= \int_{\mathbb{R}} g((t+u)/2)\gamma(t-u)h((u+s)/2)\lambda(u-s)du, \; t,s \in \mathbb{R}. \end{aligned} \tag{12}$$

The equation (12) is invariant under the operations

$$\begin{aligned} &\text{(i)} \quad \text{Translation} \\ &\qquad (f,g,h) \longmapsto (T_a f, T_a g, T_a h) \quad \text{for} \quad a \in \mathbb{R}; \\ &\text{(ii)} \quad \text{Modulation} \\ &\qquad (\rho,\gamma,\lambda) \longmapsto (M_a\rho, M_a\gamma, M_a\lambda) \quad \text{for} \quad a \in \mathbb{R}; \\ &\text{(iii)} \quad \text{Dilation} \\ &\qquad (f,g,h,\rho,\gamma,\lambda) \longmapsto (D_a f, D_a g, D_a h, D_a\rho, D_a\gamma, D_a\lambda) \\ &\qquad\quad \text{for} \quad a>0. \end{aligned} \tag{13}$$

It seems difficult to find solutions to (12) without further assumptions. Therefore we simplify the problem to the tractable case when all functions $f,\rho,g,\gamma,h,\lambda$ are Gaussians, and ρ, γ, λ are possibly modulated. Equation (12) holds when the parameters of the Gaussians fit together as described in the following result. Due to the invariances (13) (i)–(iii) we may restrict our attention to the case when λ is not modulated and has a convenient dilation factor, and h is centered at the origin.

Proposition 3 *Suppose*

$$r_{xz}(t,s) = \int_{\mathbb{R}} g_b \otimes M_\alpha g_{c/4} \circ \kappa^{-1}(t,u) g_a \otimes g_{1/4} \circ \kappa^{-1}(u,s)du, \tag{14}$$

where $t,s \in \mathbb{R}$, $g_a(x)=e^{-2\pi a x^2}$, $0 \le a \le 1$, $b,c \ge 0$, $\alpha \in \mathbb{R}$ *and* $C>0$ *is a constant. Then there exist* $d,e \ge 0$ *and* $\beta \in \mathbb{R}$ *such that*

$$r_{xz}(t,s) = C g_{4d} \otimes M_\beta g_e \circ \kappa^{-1}(t,s), \quad C>0, \tag{15}$$

if and only if either of the following two conditions is satisfied.

$$\begin{aligned} &\textit{(i)} \quad \textit{WSS case:} \\ &\qquad a=b=d=0, \quad 0 \le e < 1/4, \\ &\qquad c=4e/(1-4e), \quad \beta \in \mathbb{R}, \quad \textit{and} \\ &\qquad \alpha = \beta/(1-4e). \end{aligned} \tag{16}$$

$$\begin{aligned} &\textit{(ii)} \quad \textit{LS case:} \\ &\qquad \alpha=\beta=0, \quad b>0, \quad c=a/b, \quad \textit{and} \\ &\qquad d = \frac{2a+b+a^2/b}{4(1+a+b+a/b)} \\ &\qquad e = \frac{2a+ab+a/b}{4(1+a+b+a/b)}. \end{aligned} \tag{17}$$

Proof: We have

$$\begin{aligned} r_{zz}(u,s) &= g_a\left(\frac{u+s}{2}\right) g_{1/4}(u-s) \\ &= \exp\left(-2\pi\left[\frac{a}{2}(u^2+s^2) + \frac{1-a}{4}(u-s)^2\right]\right) \\ &= \exp\left(-\pi a(u^2+s^2)\right)\exp\left(-\pi(1-a)(u-s)^2/2\right) \\ &\in \mathrm{NND}(\mathbb{R}^2) \end{aligned}$$

if and only if $a \le 1$. This statement follows from the observation

$$\exp\left(-\pi a(u^2+s^2)\right) \in \mathrm{NND}(\mathbb{R}^2),$$

Lemma 1, and the fact that $g_p \in \mathrm{NND}(\mathbb{R})$ if and only if $p \geq 0$, since otherwise (4) is violated. The assumption $0 \leq a \leq 1$ is thus justified. The expression (15) is

$$C \exp(-2\pi[(d+e)(t^2+s^2)+2(d-e)ts+j\beta(s-t)]). \quad (18)$$

Some calculations transform the right hand side of (12) into

$$\begin{aligned} &C' \exp\left\{-2\pi\left[\frac{t^2}{4}\left(b+c-\frac{(b-c)^2}{D}\right)\right.\right.\\ &\left.\left.\quad+\frac{s^2}{4}\left(1+a-\frac{(1-a)^2}{D}\right)\right]\right\} \\ &\times \exp\left\{-2\pi\left[\frac{ts}{2D}(b-c)(1-a)\right.\right.\\ &\left.\left.\quad+j\frac{\alpha}{D}\left(s(1-a)-t(b-c+D)\right)\right]\right\}, \end{aligned} \quad (19)$$

where $C' > 0$ and $D = 1+a+b+c$. If we compare with (18) and force equality, then either $\alpha = \beta = 0$ or $1-a = b-c+D$, which is equivalent to $b = -a$.

In the case $b = -a$ we have $a = b = 0$, and comparison of (19) and (18) yields

$$\begin{aligned} d+e &= \frac{c}{4(1+c)} \\ d-e &= -\frac{c}{4(1+c)}, \end{aligned}$$

i.e. $d = 0$ and $e = c/(4(1+c))$, and $\beta = \alpha/(1+c)$. Thus we have case (i).

If $b \neq -a$ then $\alpha = \beta = 0$. Again by comparison of (19) and (18), we obtain

$$\begin{aligned} 4(d+e) &= b+c-(b-c)^2/D \\ &= 1+a-(1-a)^2/D \end{aligned} \quad (20)$$

which implies $a = bc$, and

$$4(d-e) = (b-c)(1-a)/D. \quad (21)$$

The result (ii) follows as the solution to (20) and (21). □

We draw the following conclusions from this result. If $a = b = d = 0$ then (14) is a convolution equation for the optimal filter for WSS processes. If e in the range $0 \leq e < 1/4$ and $\beta \in \mathbb{R}$ are given, then $c = 4e/(1-4e)$ and $\alpha = \beta/(1-4e)$ solves the equation. The solution is a modulated Gaussian.

If $a > 0$ then $\beta = \alpha = 0$ must hold, and for given $d, e > 0$ there exists a solution

$$h = g_b \otimes g_{a/(4b)} \circ \kappa^{-1}$$

if and only if there exists $b > 0$ such that (17) holds. Alternatively stated, h of LS type exists if and only if there exists a common positive solution $b = x$ to the system of polynomial equations

$$\begin{cases} (1-4d)x^2 + 2x(a-2d(1+a)) + a(a-4d) = & 0 \\ (a-4e)x^2 + 2x(a-2e(1+a)) + a(1-4e) = & 0. \end{cases}$$

If $a = b = 1/n$, then $c = 1$, $d = 1/(2(n+1)) \longrightarrow 0$ as $n \longrightarrow \infty$ and

$$e = \frac{2n+1+n^2}{4(2n+2n^2)} \longrightarrow \frac{1}{8} \quad \text{as} \quad n \longrightarrow \infty$$

so

$$\lim_{n \longrightarrow \infty} \frac{4e}{1-4e} = c$$

and we get case (i) asymptotically.

It is rather surprising that there are so few solutions in the case $a, d, e > 0$, even in the tractable case when all functions are Gaussians. For a solution to exist, the given parameters $a, d, e > 0$ must be related in a special way, in the sense that there exists $b > 0$ such that (17) holds. In particular, there are only "modulation-free" ($\alpha = \beta = 0$) solutions. It seems to be an unusual coincidence that an optimal filter kernel h has the LS structure when r_{zz} and r_{xz} have the LS structure.

Remark 4 *If $a = 1$ in Proposition 3 then $c = 1/b$ and $d = e = 1/4$. This means that if $r_{zz} = r_{xz} = g_1 \otimes g_{1/4} \circ \kappa^{-1}$, then the filter kernel $h = C_b g_b \otimes g_{b^{-1}/4} \circ \kappa^{-1}$ solves* (14) *for any $b > 0$, where C_b is a constant that depends on b.*

6. CONCLUSIONS

Locally stationary processes are a generalization of WSS processes where the covariance function allows a factorization into two factors, where one controls the global time behavior and the other the local time behavior. In this paper, we have studied Wiener (LMMSE) filters for LS processes. Because LS processes are more difficult to handle than WSS processes, we had to restrict our analysis to covariance and cross-covariance functions parameterized by Gaussians to reach closed-form solutions. Our main finding is that the filter kernel rarely inherits the LS structure from the covariance functions. This is a surprisingly negative result. However, if we relax the requirement that the filter kernel have the LS structure, then it is possible to find exact solutions for the filter kernel as series expansions involving Hermite functions in certain cases. This is explored in the journal version [13] of this paper.

REFERENCES

[1] C. Berg, J. P. R. Christensen, and P. Ressel. *Harmonic Analysis on Semigroups*. Springer-Verlag, 1984.

[2] R. Dahlhaus. Fitting time series models to nonstationary processes. *Ann. Statist.*, 25:1–37, 1997.

[3] P. Flandrin. *Time-Frequency/Time-Scale Analysis*. Academic Press, 1999.

[4] A. Hanssen and L. L. Scharf. A theory of polyspectra for nonstationary stochastic processes. *IEEE Trans. Signal Process.*, 51(5):1243–52, 2003.

[5] M. Loève. *Probability Theory*. Van Nostrand, 1963.

[6] S. G. Mallat, G. Papanicolaou, and Z. Zhang. Adaptive covariance estimation of locally stationary processes. *Ann. Statist.*, 26(1):1–47, 1998.

[7] F. D. Neeser and J. L. Massey. Proper complex random processes with applications to information theory. *IEEE Trans. Inform. Theory*, 39(4):1293–1302, 1993.

[8] P. J. Schreier and L. L. Scharf. Second-order analysis of improper complex random vectors and processes. *IEEE Trans. Signal Process.*, 51(3):714–25, 2003.

[9] R. A. Silverman. Locally stationary random processes. *IRE Trans. Inf. Theory*, 3:182–187, 1957.

[10] R. A. Silverman. A matching theorem for locally stationary random processes. *Comm. Pure and Appl. Math.*, 12:373–383, 1959.

[11] H. L. V. Trees. *Detection, Estimation, and Modulation Theory*. John Wiley, 1971.

[12] P. Wahlberg and P. J. Schreier. Locally stationary harmonizable complex improper stochastic processes. Submitted, 2008.

[13] P. Wahlberg and P. J. Schreier. On Wiener filtering of locally stationary stochastic processes. Submitted, 2008.

[14] L. A. Weinstein and V. D. Zubakov. *Extraction of Signals from Noise*. Prentice–Hall, 1962.

[15] N. Wiener. *Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications*. MIT Press, 1949.