The Annals of Statistics
1983, Vol. 11, No. 3, 848–855

CANONICAL CORRELATIONS OF PAST AND FUTURE FOR TIME SERIES: BOUNDS AND COMPUTATION

BY NICHOLAS P. JEWELL,[1] PETER BLOOMFIELD[2] AND FLAVIO C. BARTMANN[2]

University of California, Berkeley; Princeton University; and IMECC

This paper continues an investigation into the canonical correlations and canonical components of the past and future of a stationary Gaussian time series which were introduced in Jewell and Bloomfield (1983). Bounds for the maximum canonical correlation are provided under specified conditions on the spectrum of the series. A computational scheme is described for estimating the canonical correlations and components and the procedure is illustrated on the well-known sunspot number series.

1. Introduction. In Jewell and Bloomfield (1983) (hereafter referred to as [J-B]) canonical correlations and components of the past and future of certain time series were introduced and discussed. This article is a sequel which describes some elementary bounds on the canonical correlations and a basic method for computing the correlations and respective canonical components. The well-known sunspot number series is used to provide an example of the computational procedure in the final section.

We refer to [J-B] for definitions, notation and theory concerning canonical correlations. We repeat here only the notation necessary for this paper. Let $\{x(t)\}$ be a weakly stationary Gaussian time series with zero mean and spectral measure F. We represent the process $\{x(t)\}$ in $L^2(dF)$ by a sequence of exponentials $\{e^{it\omega} : t \in \underset{\sim}{Z}\}$ on the unit circle C in $\underset{\sim}{C}$. In this spectral representation the *past* of the process, $\mathscr{P}$, is the span in $L^2(dF)$ of the exponentials $e^{it\omega}$ with $t \leq 0$. The *future beyond time s*, $\mathscr{F}_s$, is the span in $L^2(dF)$ of the exponentials $e^{it\omega}$ with $t \geq s$. $\mathscr{F}_1$ is known as the *future* of the process.

H^2 is the Hardy space of functions in $L^2(d\omega)$ on C which possess analytic extensions into the open unit disk. The Hardy space H^∞ contains those functions in $L^\infty(d\omega)$ which are also in H^2.

For reasons described in detail in [J-B] we shall restrict our attention to purely indeterministic processes. For such processes the spectral measure $dF = w\,d\omega = |h|^2\,d\omega$ where h is an outer function in H^2. The function w is the spectral density function of the process and $\bar{h}/h$ is the phase function. We write $L^2(w)$ for $L^2(dF)$.

The first canonical correlation, λ_1, of the past and future is the largest correlation between an element $f \in \mathscr{F}_1$, and $g \in \mathscr{P}$, i.e.,

$$\lambda_1 = \sup\left\{\int_C f\bar{g}w\,d\omega : f \in \mathscr{F}_1, g \in \mathscr{P}; \|f\|_{L^2(w)} = \|g\|_{L^2(w)} = 1\right\}.$$

Alternatively

$$\lambda_1 = \sup\left\{\int_C e^{i\omega}fgw\,d\omega : f, g, \text{ are in the span in } L^2(w) \text{ of the functions } 1, e^{i\omega}, \cdots,; \|f\|_{L^2(w)} = \|g\|_{L^2(w)} = 1\right\}.$$

It was also noted in [J-B] that the first canonical correlation of the past and future (if it exists) is given by the largest eigenvalue of the operator H^*H on H^2 where H is the

Received September 1982; revised January 1983.
[1] Supported in part by an NSF grant.
[2] Supported in part by DOE Contract DE-AC02-81ER10841.
AMS 1980 *subject classification.* Primary 62M15; secondary 60G25,47B35.
Key words and phrases. Time series, spectrum, prediction, canonical correlations, ARMA.

Hankel operator with symbol $\bar{h}/h$. The remaining canonical correlations (if they exist) are given by the other eigenvalues of H^*H arranged in descending order. Questions of existence are fully dealt with in [J-B]. A useful equality we shall need in Section 2 is $H^*H = I - T^*T$ where I is the identity matrix on H^2 and T is the Toeplitz operator on H^2 with symbol $\bar{h}/h$.

2. Bounds on the canonical correlations. In [J-B] it is shown that the first canonical correlation λ_1 is given by

$$\lambda_1 = d(\bar{h}/h, H^\infty) = \inf_{f \in H^\infty}\{\| \bar{h}/h - f \|_\infty\}.$$

The Helson-Szego theorem (see Helson and Szego, 1960) states that $\lambda_1 < 1$ if and only if the spectral density function w admits a representation $w = \exp(u + \tilde{v})$ with u, v real-valued functions in L^∞, $\| v \|_\infty < \pi/2$ where $\tilde{v}$ is the Hilbert transform of v. Note that this implies that $\bar{h}/h = c \exp i\{(v - \tilde{u})\}$ where c is a constant of unit modulus. The proof of the sufficiency part of the theorem is based on the fact that under the given condition on w we may take $f = kc \exp(-u - i\tilde{u})$, where k is any positive constant. We have f in H^∞ and

$$\| \bar{h}/h - f \|_\infty = \| \exp\{i(v - \tilde{u})\} - k \exp(-u - i\tilde{u}) \|_\infty = \| 1 - k \exp(-u - iv) \|_\infty.$$

It follows easily from this equality that $d(\bar{h}/h, H^\infty) < 1$. We will use the equality to obtain an upper bound for λ_1 when w satisfies the conditions of the Helson-Szego theorem.

Since $\| v \|_\infty < \pi/2$ the values of the function $\exp(-u - iv)$ lie in the sector

$$\{z : \exp(-k_1) \le | z | \le \exp(-k_2),\ | \arg z | \le \| v \|_\infty\}$$

where $k_2 \le u \le k_1$, almost everywhere. Some elementary geometry shows that if a function g has values in the sector $\{z : R_1 \le | z | \le R_2, | \arg z | \le \theta\}$, then the function Mg has values in a disk centered at 1 of radius r when

$$M = 2 \cos \theta/(R_1 + R_2) \text{ and } r^2 = (R_1^2 + R_2^2 - 2R_1R_2\cos 2\theta)/(R_1 + R_2)^2.$$

Thus $\| 1 - Mg \|_\infty^2 \le r^2$ and, without knowing any more about g, this is the best inequality we can achieve. Applying this to $\exp(-u - iv)$ we find the inequality

$$\| 1 - k \exp(-u - iv) \|_\infty \le [1 + \exp(-2(k_2 - k_1)) - 2 \exp(k_1 - k_2)\cos(2 \| v \|_\infty)]^{1/2}/[1 + \exp(k_1 - k_2)]$$

where $k = 2 \cos(\| v \|_\infty)/\{\exp(-k_1) + \exp(-k_2)\}$. We have thus established the following proposition.

PROPOSITION 1. *If w is given by $\exp(u + \tilde{v})$ where u, v are real-valued functions in L^∞ and $\| v \|_\infty < \pi/2$ then an upper bound for the maximum correlation between past and future is*

$$\frac{[1 + \| e^u \|_\infty^2 \| e^{-u} \|_\infty^2 - 2 \| e^u \|_\infty \| e^{-u} \|_\infty \cos(2 \| v \|_\infty)]^{1/2}}{1 + \| e^u \|_\infty \| e^{-u} \|_\infty}.$$

It is not clear when we would know, a priori, of such a representation for w, except in the case when we know positive constants m, M such that $m \le w \le M$. This, of course, implies that $w = \exp(u)$ for some $u \in L^\infty$. The bound in the proposition simplifies to $\lambda \le (M - m)/(M + m)$. A similar bound was derived by Bargmann and Schunemeyer (1978) for the maximum canonical correlation of two finite sets of random variables. Since, in a certain sense, the "eigenvalues" of the joint dispersion matrix of the past and future are just the values of w, this result was expected to hold. The result can also be considered as a generalization of a result of Venables (1976) to infinite dimensional spaces.

It is illuminating that the simple bound may be obtained by elementary arguments as follows. It was noted in the introduction that $\lambda_1 = \sup | \int_C fg^{i\omega} w \, d\omega |$ where f, g are in the span in $L^2(w)$ of the functions 1, $e^{i\omega}$, $e^{2i\omega}$, $\cdots$, and $\int_C |f|^2 w\, d\omega = \int_C |g|^2 w\, d\omega = 1$. We suppose that $m \le w \le M$. For f and g as above we have $\int_C e^{i\omega} fg\, d\omega = 0$ since the conditions on w imply that $f, g \in L^2$ and f, g are in the *span in* L^2 of the functions 1, $e^{i\omega}$, $e^{2i\omega}$, $\cdots$. Hence, for any constant k,

$$(*) \qquad \left| \int_C e^{i\omega} fgw\, d\omega \right| = \left| \int_C e^{i\omega} fg\{1 - k/w\} w\, d\omega \right| \le \left(\int |fg|\, w\, d\omega \right) \sup_C |1 - k/w| \le \sup |1 - k/w|.$$

It is easily seen that the best choice of the constant k to make this supremum smallest is $2Mm/(M + m)$ and, for this value of k, the bound simplifies to $(M - m)/(M + m)$. Notice that this bound is valid for *any* f and g with $\int_C |f|^2 w\, d\omega = \int_C |g|^2 w\, d\omega = 1$ and $\int_C e^{i\omega} fg\, d\omega = 0$. This is a larger class than that in which we are interested. This approach thus illustrates that the bound for the maximum correlation cannot be sharp. In fact we can be more explicit concerning this. In (*) the second inequality is sharp if and only if $|f| = |g|$ almost everywhere and the first is sharp only if fg vanishes almost everywhere on the set where $|1 - k/w|$ is not equal to its supremum. If $f, g \in H^2$ this implies that $|1 - k/w| = \sup_C |1 - k/w|$ a.e. Additionally, in this situation, sharpness of the first inequality demands that the argument of $e^{i\omega} fg$ is constant a.e. This is not possible for f, g in H^2 unless $f = g = 0$ a.e., which contradicts the condition on the norm of f and g in $L^2(w)$. Summarizing, the inequality (*) cannot be attained for f, g in the span of 1, $e^{i\omega}$, $e^{2i\omega}$, $\cdots$ in $L^2(w)$ with $\int_C |f|^2\, d\omega = \int_C |g|^2\, d\omega = 1$ where $m \le w \le M$. As an example of this phenomenon, consider a first order autoregression process $x(t) = \alpha x(t - 1) + \varepsilon(t)$ where $\varepsilon(t)$ is a zero mean white noise process with variance 1 and $0 < \alpha < 1$. The simple bound in this case is $2\alpha/(1 + \alpha^2)$ which is strictly greater than the maximum canonical correlation α.

A quite different bound may be obtained for absolutely regular processes using results of Widom (1976). Recall (see [J-B]) that a process is absolutely regular if and only if $w = |p|^2 f$ where p is a trigonometrical polynomial with all its zeros on C and $\log(f)$ has a Fourier series $\sum_{-\infty}^{\infty} f_j e^{ij\omega}$ with $\sum_{-\infty}^{\infty} |j| |f_j|^2 < \infty$. It was shown in [J-B] that a process is absolutely regular if and only if the essential correlation of the process is zero and $\sum_{j=1}^{\infty} \lambda_j^2 < \infty$ where λ_j is the jth canonical correlation of the past and future.

A result in Ibragimov and Rozanov (1978) shows that the condition on w for absolute regularity implies that the Fourier series of $\bar{h}/h$ $(= \sum_{-\infty}^{\infty} h_j e^{ij\omega})$ also has the property that $\sum_{-\infty}^{\infty} |j|\, |h_j|^2 < \infty$. Now the Fourier series of $h/\bar{h} = (\bar{h}/h)^{-1}$ is $\sum_{-\infty}^{\infty} \bar{h}_j e^{-ij\omega}$ so that this Fourier series has the same property also. Also $h/\bar{h} = h^2/|h|^2$ so that the argument of $h/\bar{h}$ is $\arg(h^2)$. Since h is analytic and non-zero on the open unit disk it follows that the change in argument of $\bar{h}/h$ as you move round C is zero. Hence $h/\bar{h}$ satisfies the conditions of Theorem 7.1 of Widom (1976). This result shows that $\det T_{h/\bar{h}} T_{\bar{h}/h} = \exp(\sum_{j=1}^{\infty} j a_j a_{-j})$ where the Fourier series of $\log(h/\bar{h})$ is $\sum_{-\infty}^{\infty} a_j e^{ij\omega}$ with the appropriate determination of $\log(h/\bar{h})$. The determinant is defined for bounded operators on a Hilbert space differing from the identity by a trace-class operator. Note that $T_{h/\bar{h}} T_{\bar{h}/h} = T^*_{\bar{h}/h} T_{\bar{h}/h} = I - H^*_{\bar{h}/h} H_{\bar{h}/h}$ (see [J-B]) and since $H_{\bar{h}/h}$ is Hilbert-Schmidt, $H^*_{\bar{h}/h} H_{\bar{h}/h}$ is trace class. The determinant of $T^*_{\bar{h}/h} T_{\bar{h}/h} = T_{h/\bar{h}} T_{\bar{h}/h}$ is given by the product of the eigenvalues of $T^*_{\bar{h}/h} T_{\bar{h}/h}$ i.e., $\det T^*_{\bar{h}/h} T_{\bar{h}/h} = \prod_{j=1}^{\infty} (1 - \lambda_j^2)$ where λ_j^2 are the eigenvalues of $H^*_{\bar{h}/h} H_{\bar{h}/h}$ and the product is taken to include the multiplicity of each λ_j^2. ($H_{\bar{h}/h}$ being Hilbert-Schmidt implies that $\sum_1^{\infty} \lambda_j^2 < \infty$ which guarantees convergence of the infinite product.) Now $\log(h/\bar{h}) = i(\widetilde{\log w})$ which implies that the Fourier coefficient of $\log(h/\bar{h})$ is the respective Fourier coefficient of $\widetilde{\log w}$ multiplied by i. Using the well-known relationships between Fourier coefficients of a function and its harmonic conjugate this implies that $a_j = -w_j$ $(j < 0)$ and $a_j = w_j$ $(j \ge 0)$ where the Fourier series of $\log w$ is $\sum_{-\infty}^{\infty} w_j e^{ij\omega}$. Thus $\det T^*_{\bar{h}/h} T_{\bar{h}/h} = \exp(-\sum_{j=1}^{\infty} j w_j w_{-j})$. Since $w_{-j} = \bar{w}_j$ we have $T^*_{\bar{h}/h} T_{\bar{h}/h} = \exp(-\sum_{j=1}^{\infty} j |w_j|^2)$. We have just established the following result.

PROPOSITION 2. *Let* $\{\lambda_j\}_{j=1}^{\infty}$ *be the canonical correlations of past and future of an absolutely regular process. Then* $\prod_{j=1}^{\infty}(1-\lambda_j^2) = \exp[-\sum_{j=1}^{\infty} j\,|w_j|^2]$, *where the Fourier series of* $\log w = \sum_{-\infty}^{\infty} w_j e^{ij\omega}$.

Thus, if λ_1^2 is the square of the maximum canonical correlation, then $1-\lambda_1^2 \geq \prod_{j=1}^{n}(1-\lambda_j)^2$. Hence $\lambda_1^2 \leq 1 - \exp[-\sum_{j=1}^{\infty} j\,|w_j|^2]$. Again we would not expect this to be a sharp upper bound. We can have equality only if $\lambda_1^2 < 1$ has multiplicity one and all the other canonical correlations are zero. Thus only ARMA (1, 1) processes give equality.

REMARKS. 1. The other canonical correlations introduced in [J-B] are those between $\mathscr{P}$ and $\mathscr{F}_n$ denoted by $\lambda_j^{(n)}, j = 1, 2, \cdots, n = 1, 2, \cdots$, if they exist. A theorem of Helson and Sarason (1967) shows that $\lambda_1^{(n)} < 1$ if and only if w admits a representation $w = |p|^2 \exp(u + \tilde{v})$ with u, v as in the Helson-Szego theorem stated above and p a trigonometric polynomial of degree less than n. If the degree of p is $k(<n)$ then the reasoning behind Proposition 1 shows that $\lambda_1^{(k)}$ is bounded above by the bound of Proposition 1. In fact this bound then holds for $\lambda_1^{(n)}$ since $\mathscr{F}_n \subseteq \mathscr{F}_k$. Concerning the bound of Proposition 2, it is easily seen that $\prod_{j=1}^{\infty}(1-\lambda_j^{(n)2}) = \det(T_{\bar{h}/h}^* T_{\bar{h}/h})_n$ where the matrix of $(T_{\bar{h}/h}^* T_{\bar{h}/h})_n$ is that of $T_{\bar{h}/h}^* T_{\bar{h}/h}$ with the first n rows and columns deleted. There doesn't seem to be any obvious way to relate this quantity to the spectral density function w.

2. It is easy to show that the Hilbert transform is bounded as an operator on $L^2(w)$ if and only if the maximum correlation between the past and the future is less than 1. See Section 5 of Helson and Szego (1960). As noted in [J-B] a theorem of Hunt, Muckenhoupt and Wheeden (1973) shows that the Hilbert transform is bounded on $L^2(w)$ if and only if

$$\sup_I \left(\frac{1}{|I|}\int_I w\,d\omega\right)^{1/2}\left(\frac{1}{|I|}\int_I w^{-1}\,d\omega\right)^{1/2} = A < \infty$$

where I ranges over all subarcs of C. Knowing the constant A and working through the proof of this theorem produces an upper bound on the norm of the Hilbert transform. By the statement at the beginning of this remark we can then produce an upper bound for λ_1. The details are exceedingly technical and will appear elsewhere.

3. Computation of the canonical correlations. It was shown in [J-B] that the canonical correlations possess a number of equivalent mathematical definitions. For instance, λ_1 is both $\|H_\phi\|$ and $d(\phi, H^\infty)$, where $\phi = \bar{h}/h$ and H_ϕ is the Hankel operator with symbol ϕ. Either characterization could be used as the basis for calculating λ_1, given the function w. We have chosen to use the former.

Since $\|H_\phi\|^2$ is the largest value in the spectrum of $H_\phi^* H_\phi$, a convergent sequence of approximations may be constructed using the power method (Riesz and Nagy, 1955, pages 230–241). To be specific, suppose that x is a eigenfunction associated with this point of the spectrum. If we start with any function x_0 satisfying $\langle x_0, x\rangle \neq 0$, and define x_n, $n > 0$, recursively by

$$\xi_n = H_\phi^* H_\phi x_{n-1}, \quad x_n = \xi_n / \|\xi_n\|$$

then $x_n \to x$ and $\|\xi_n\| \to \|H_\phi^* H_\phi\| = \lambda_1^2$.

In our case, we know that the eigenfunction x is not orthogonal to the constant function, and hence we may begin with $x_0 = 1$. For convenience we have used a two-stage version of the power method, in which x_n is defined recursively by

$$\eta_n = H_\phi x_{n-1}, \quad y_n = \eta_n/\|\eta_n\|, \quad \xi_n = H_\phi^* y_n, \quad x_n = \xi_n/\|\xi_n\|.$$

It is easily seen that the sequence $\{x_n\}$ is the same as before, and the pair (x_n, y_n) converges to a Schmidt pair (x, y) for H_ϕ, namely a pair for which

$$H_\phi x = \lambda_1 y, \quad H_\phi^* y = \lambda_1 x.$$

Furthermore, $l_n = \| \xi_n \| \to \lambda_1$.

It is a general property of the power method that the successive approximations to λ_1 converge monotonically from below. Thus each successive l_n gives a lower bound for λ_1. We may use the fact that $\lambda_1 = d(\phi, H^\infty)$ to construct corresponding upper bounds. Adamjan, Arov and Krein (1971) give the functional form of the function $F \in H^\infty$ for which $\| \phi - F \|_\infty = \lambda_1$, namely,

$$F = \phi - \lambda_1 \exp(i \arg y\bar{x}).$$

Thus

$$F_n = P\{\phi - l_n \exp(i \arg y_n \bar{x}_n)\}$$

converges to F, where P is the projection onto H^2. Hence, provided $\| F_n \|_\infty < \infty$ we have

$$u_n = \| \phi - F_n \|_\infty \geq \lambda_1 = \inf_{F \in H^\infty} \| \phi - F \|_\infty.$$

Thus we can find an interval $[l_n, u_n]$ that contains λ_1.

To implement these ideas numerically, we need to approximate two operations. In the first place, we have to factorize the spectral density function w into $| h |^2$ where h is an outer function, to obtain the function $\phi = \bar{h}/h$. Secondly, we need a mechanism for approximating the iterations of the power method.

We actually obtain ϕ directly by noting that if $w = \exp u = | h |^2$, then $h = \exp (u + i\tilde{u})/2$ and $\phi = \bar{h}/h = \exp(-i\tilde{u})$. The harmonic conjugate $\tilde{u}$, of u, is calculated as

$$\tilde{u}(\omega) = \sum_{k<0} i u_k \exp(ik\omega) - \sum_{k>0} i u_k \exp(ik\omega),$$

where u_k is the kth Fourier coefficient of u. We approximate these by

$$u_k = N^{-1} \sum_{j=1}^{N} u(\omega_j) \exp(- ik\omega_j),$$

where $\omega_j = 2\pi j/N$, for a suitably large value N.

For a more detailed discussion of the empirical factorization of an estimated spectrum, see Bhansali (1974).

It remains to describe the approximation of one step of the power method. For instance, we have to obtain $\eta_n = H_\phi x_{n-1}$, given the function x_{n-1}. Now $H_\phi x_{n-1} = (I - P)\phi x_{n-1}$, where P, as before, is the projection onto H^2. We carry out this calculation in two steps, first multiplication by ϕ and secondly the projection. Suppose that we know the values of ϕ and x_{n-1} at each $\omega_j = 2\pi j/N$. Then the product can be calculated at the same values of ω. Now the effect of the operator $(I - P)$ is to replace the Fourier coefficients with non-negative indices by zero, while leaving the Fourier coefficients with negative indices unchanged. We approximate this operation by using the discrete approximation

$$(\phi x_{n-1})_\kappa \simeq N^{-1} \sum_{j=1}^{N} \phi(\omega_j) x_{n-1}(\omega_j) \exp(-i\kappa\omega_j), \quad | \kappa | < N/2$$

$$(I - P)(\phi x_{n-1})(\omega) \simeq \sum_{\kappa<0, |\kappa|<N/2} (\phi x_{n-1})_\kappa \exp(i\kappa\omega).$$

While the resulting function, which is our approximation to $\eta_n = H_\phi x_{n-1}$, could be evaluated for any ω, we in fact only need its value at the same places, ω_j, as ϕ and x_{n-1}. Note that the next step, $y_n = \eta_n/\| \eta_n \|$, may be carried out with no further approximations, since

$$\| \eta_n \|^2 = \int | \eta_n(\omega) |^2 \, d\omega$$

is given exactly by

$$\frac{2\pi}{N} \sum | \eta_n(\omega_j) |^2.$$

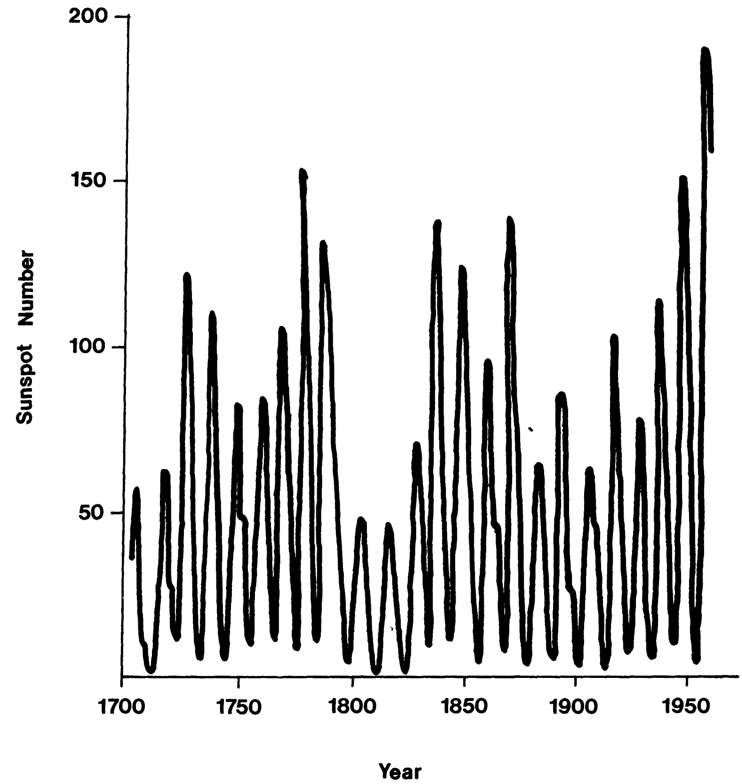

FIG. 1. *The annual sunspot numbers* 1704-1960.

TABLE 1

	Model	Source
1	$X(t) - 1.34X(t-1) + .65X(t-2) = e(t)$	Yule, Box-Jenkins, etc.
2	$X(t) - 1.62X(t-1) + X(t-2) = e(t)$	Yule
3	$X(t) - 1.30X(t-1) + .54X(t-2) + .15X(t-3)$ $-.19X(t-4) + .24X(t-5) - .14X(t-6) = e(t)$	Bailey
4	$X(t) - 1.57X(t-1) + 1.02X(t-2) - .21X(t-3) = e(t)$	Box-Jenkins
5	$X(t) - 1.42X(t-1) + .72X(t-2) = e(t) - .15e(t-1)$	Phadke and Wu
6	$X(t) - 1.25X(t-1) + 0.54X(t-2) - .19X(t-3) = e(t)$	Morris, Schaerf

The action of H_ϕ^* which is needed to compute ξ_n is approximated in a similar fashion, using the fact that $H_\phi^* y_n = P\bar{\phi} y_n$.

We also wish to find the coefficients of the canonical components. These are given by the appropriate Fourier coefficients of $\bar{x}/\bar{h}$ and $\bar{y}/h$, for the components in the past and future, respectively. For real-valued time series, where the spectral density function w is symmetric, these sequences of coefficients are reverses of each other.

The computational error introduced into these operations is that of approximating the first N Fourier coefficients of a function by discrete sums instead of integrals, and replacing the remainder by zero. We have not analyzed the magnitude of the resulting error in detail. It is clear, however, that if w is reasonably smooth then the error can be made small by choosing a sufficiently large number, N, of points on the unit circle. It should also be pointed out that in most instances we shall work with estimated spectral density functions, which are typically both smooth by construction, and computed only at finite grid of points. In these cases the statistical uncertainty in w is likely to dominate the numerical

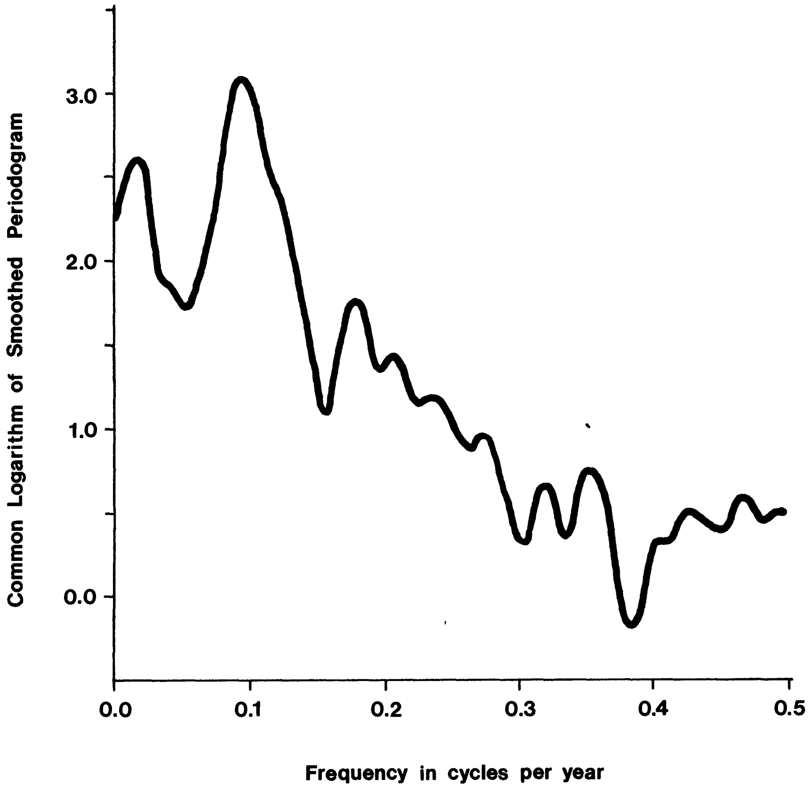

FIG. 2. *Common logarithm of smoothed periodogram of yearly sunspot numbers.*

TABLE 2

	Canonical Correlation (squared)	Associated Vector
1	.8565	$X(t)$ $-0.362X(t-1)$
2	.8518	$X(t)$ $-0.171X(t-1)$
3	.8641	$X(t)$ $-0.285X(t-1) - 0.138X(t-2)$ $+0.191X(t-3) - 0.191X(t-4) - 0.080X(t-5)$
4	.8712	$X(t)$ $-0.514X(t-1) + 0.087X(t-2)$
5	.8476	$X(t)$ $-0.296X(t-1) - 0.044X(t-2)$ $-0.067X(t-3) - 0.001X(t-4) - \ldots$
6	.8677	$X(t)$ $-0.409X(t-1) + 0.126X(t-2)$
Spec	.8923	$X(t)$ $+0.264X(t-1) - 0.170X(t-2)$ $+0.253X(t-3) - 0.231X(t-4) + \ldots$

errors introduced by our computational procedure. We have only described the computations of the first canonical correlation and components. It is easy to extend the above computation scheme to provide further canonical correlations and components.

4. An example—Sunspot numbers. Gray and Woodward (1978) discuss the well-known sunspot number series, shown in Figure 1, and tabulate the various models that have been fitted to the data. See Table 1.

We have calculated the first canonical correlation and component for each of these models, and also for the nonparametric spectrum estimate shown in Figure 2 using the methods of the previous section. The spectrum estimate was obtained by smoothing the periodgram of the data, tapered 10%, by three passes of a seven-term simple moving average. For ARMA processes the canonical correlations can be computed exactly and the canonical component coefficients can be found by solving a system of linear algebraic equations. This result is essentially in Helson and Szego (1960). A simpler proof which is more statitical in nature can be found in Bartmann (1981). See also Yaglom (1965).

Table 2 shows the results of these calculations giving the first canonical correlation squared and the associated canonical component in the past.

Model 2 requires special discussion as it does not represent a stationary time series. Its spectral density function is both unbounded and nonsummable. However, it still possesses a factorization $w = |h|^2$ where h is analytic and nonzero in the open unit disk, and this factorization may be obtained by the general method described in the previous section. The results are perhaps best interpreted as the limits of calculations for a sequence of stationary second-order models.

We note that the canonical correlations are remarkably similar, and that several of the models also give very similar canonical components. The most striking similarity is between Bailey's model and the spectrum estimate; however, this is not surprising, since the six parameters used in Bailey's model make it the closest to the nonparametric approach used in spectrum estimation. The fact that the nonparametric spectrum estimate gives the largest estimated correlation is presumably associated with extra sampling variability induced by the large effective number of parameters implied by its calculation.

REFERENCES

ADAMJAN, V. M., AROV, D. Z. and KREIN, M. G. (1971). Analytic properties of Schmidt pairs for a Hankel operator and the generalized Schur-Takagi problem. *Mat. Sb.* **86** 34–75; *Math. U.S.S.R. Sb.* 31–73.

BARGMANN, R. E. and SCHUENEMEYER, J. H. (1978). Maximum eccentricity as a union-intersection test statistic in multivariate analysis. *J. Multivariate Anal.* **8** 268–273.

BARTMANN, F. C. (1981). A new angle on past and future. Ph.D. Dissertation, Princeton University.

BHANSALI, R. J. (1974). Asymptotic properties of the Wiener-Kolmogorov predictor. I. *J. Royal Statist. Soc. Ser. B* **36** 61–73.

GRAY, H. L. and WOODWARD, W. A. (1978). New ARMA models for Wolfer's sunspot data. *Commun. Statist-Simulat., Comput.* **B7** 97–115.

HELSON, H. and SZEGO, G. (1960). A problem in prediction theory. *Ann. Mat. Pura Appl.* **51** 107–138.

HELSON, H. and SARASON, D. E. (1967). Past and future. *Math. Scand.* **21** 5–16.

HUNT, R. A., MUCKENHOUPT, B. and WHEEDEN, R. L. (1973). Weighted norm inequalities for the conjugate function and Hilbert transform. *Trans. Amer. Math. Soc.* **176** 227–251.

IBRAGIMOV, I. A. and ROZANOV, Y. A. (1978). *Gaussian Random Processes.* Springer-Verlag, New York.

JEWELL, N. P. and BLOOMFIELD, P. (1983). Canonical correlations of past and future for time series: Definitions and theory. *Ann. Statist.* **11** 837–847.

RIESZ, F. and NAGY, B. Sz. (1955). *Functional Analysis.* Ungar, New York.

VENABLES, W. (1976). Some implications of the union intersection principles for tests of sphericity. *J. Multivariate Anal.* **67** 185–190.

WIDOM, H. (1976). Asymptotic behavior of block Toeplitz matrices and determinants. II. *Advances in Math.* **21** 1–29.

YAGLOM, A. M. (1965). Stationary Gaussian processes satisfying the strong mixing condition and best predictable functionals. *Proc. Int. Research Seminar of the Statistical Laboratory, Univ. of Calif., Berkeley*, 1963, 241–252; Springer-Verlag, New York.

NICHOLAS P. JEWELL
PROGRAM IN BIOSTATISTICS
SCHOOL OF PUBLIC HEALTH
UNIVERSITY OF CALIFORNIA, BERKELEY
BERKELEY, CALIFORNIA 94720

PETER BLOOMFIELD
DEPARTMENT OF STATISTICS
NORTH CAROLINA STATE UNIVERSITY
BOX 5457
RALEIGH, NORTH CAROLINA 27650

FLAVIO C. BARTMANN
DEPARTAMENTO DE ESTATISTICA
IMECC
UNICAMP
13100 CAMPINAS-SP, BRAZIL