COMPUTERS AND BIOMEDICAL RESEARCH 19, 497–507 (1986)

Spectral Methods for Principal Components Analysis of Event-Related Brain Potentials

ROBERT R. RAWLINGS,* JOHN W. ROHRBAUGH,† HENRI BEGLEITER,‡ AND MICHAEL J. ECKARDT†

*Division of Biometry and Epidemiology and †Section of Clinical Brain Research, Laboratory of Clinical Studies, National Institute on Alcohol Abuse and Alcoholism, 5600 Fishers Lane, Rockville, Maryland, 20857, and ‡Department of Psychiatry, Downstate Medical Center, State University of New York, Brooklyn, New York, 11203

Received June 13, 1985

Principal components analysis has been a widely used method for the analysis of event-related, electrical brain potentials (ERPs). Recent emphasis has been placed on measuring the topography of ERPs, as derived from the instantaneous measurements from multiple locations, and on defining diagnostic differences in ERPs among various clinical populations. One goal of the present paper is to discuss inherent difficulties in utilizing PCA as an analytical technique in multiple location and multiple group studies. Another goal is to demonstrate the utility of spectral analysis and its equivalency to PCA when the signal imbedded in stationary noise model is used. Spectral analysis readily permits analysis of multiple lead and multiple group studies. © 1986 Academic Press, Inc.

The method of principal components analysis (PCA) has been widely used for the analysis of event-related, electrical brain potentials (ERPs) (*1*, *2*). ERPs consist of characteristic changes in voltage over time with respect to a specified stimulus. They show great variation depending on stimulus, task, and subject variables. When recorded from the scalp, the ERP signal is embedded in unrelated noise in the form of spontaneous electroencephalographic (EEG) activity. The signal-to-noise ratio is unfavorable and must be improved by some statistical procedure, usually by averaging over an ensemble of trails. Although the signal-to-noise ratio is improved in proportion to the number of sweeps averaged, some unaveraged noise is in practice, invariably present in the composite wave form. In ERP research the ERP is considered to be a deterministic signal while the spontaneous EEG activity process is considered to be a stochastic process. In this article we adhere to the designation of this stochastic process as noise. The term noise is utilized to emphasize the deviation from the deterministic ERP signal. In any case the information contained in the EEG noise as well as that contained in the ERP signal is utilized in the statistical procedures outlined here. It should be noted that, physiologically, the ERP and EEG components may share common origins. Both are believed to arise from syn-

0010-4809/86 $3.00

chronized graded potentials which are conducted by volume to the scalp. It is thought that the observed EEG potentials arise from the dendrites aligned perpendicular to the surface of the cerebral cortex. The precise origins of the ERP signals are not presently known but there is some evidence that the P300 and later components of the ERP arise from both cortical and subcortical regions of the brain (*3, 4*). Some attempts at developing models to explain the generation of EEGs and ERPs are contained in Freeman (*5*) and Nunez (*6*).

One of the objectives of PCA is the decomposition of the complex ERP waveform into a relatively few, and hence more manageable "components." Consequently, experimentally induced changes in these components seem easier to interpret than in the original ERP waveform. Although PCA has been used successfully as a descriptive technique, attempts to relate these derived components to latencies of some of the major original waveform peaks have relied on arbitrary methodology and often appear to result in subjective interpretation. Moreover, the appropriateness of PCA has been questioned on the basis of mathematical (*7*) and empirical (*8–10*) issues.

Recent emphasis has been placed on measuring the topography of ERPs, as derived from the instantaneous measurements from multiple locations, and on defining diagnostic differences in ERPs among various clinical populations. One goal of the present paper is to discuss inherent difficulties in utilizing PCA as an analytical technique in multiple location and multiple group studies. Another goal is to demonstrate the utility of spectral analysis and its equivalency to PCA when the signal imbedded in stationary noise model (SSN) is used. Spectral analysis readily permits analysis of multiple lead and multiple group studies. Moreover, rigorous methods of statistical inference are available for the SSN model. It should be pointed out that while there has been considerable research using "spectral methods" (*11–14*) the mathematical methods discussed in the present article are more general. In particular, our proposed methods are especially powerful for statistical inference in multiple lead problems and simultaneously utilize ERP phase, ERP magnitude, and EEG power spectra, EEG cross-spectra information in the rigorous analyses of multiple time series. The earlier research using spectral methods concentrates primarily on the estimation of power spectra and uses these estimates along with the usual real-valued multivariate methods for statistical inferences.

Principal Components Analysis

The rationale for utilizing PCA on time series data (i.e., voltage varying over time) is a mathematical result known as the Karhunen–Loeve expansion (*15*). Let $x(t) = \mu(t) + \eta(t)$ denote a stochastic process on the interval $[0,T]$ and let $EX(t) = \mu(t)$ where E denotes the mathematical expectation. The term $\eta(t)$ can be considered a noise process such that $E\eta(t) = 0$. Only the discrete time case for $t = t_1, \ldots, t_n$ will be discussed. The stochastic process is an n-dimensional vector $x = (x(t_1), \ldots, x(t_n))'$ where $'$ denotes the transpose. The mean is the vector $\mu = (\mu(t_1), \ldots, \mu(t_n))'$ and the covariance matrix is $\Sigma = E(x -$

$\mu)(x - \mu)'$. The eigenvalues and eigenvectors of Σ are $\{(\lambda_i, \Phi_i); i = 1, \ldots n\}$. The random process $\eta(t) = x(t) - \mu(t)$ can be expressed as the expansion $\sum_{i=1}^{n} a_i\Phi_i$ where the principal components are given by $a_i = \eta'\Phi_i$. It should be observed that the $\{\Phi_i, i = 1, \ldots, n\}$ are deterministic while the a_i are random variables with $E(a_i) = 0$. The process η can be approximated by using the expansion $\boldsymbol{\eta} = \sum_{i=1}^{k} a_i\Phi_i$ and the following mean square error can be obtained $E(\eta - \boldsymbol{\eta})'(\eta - \boldsymbol{\eta}) = E\left(\sum_{i=k+1}^{n} a_i\Phi_i\right)'E\left(\sum_{i=k+1}^{n} a_i\Phi_i\right) = E\left(\sum_{i=k+1}^{n} a_i^2\right) = E\left(\sum_{i=k+1}^{n} \Phi_i'\eta\eta'\Phi_i\right) = \sum_{i=k+1}^{n} \Phi_i'\Sigma\Phi_i = \sum_{i=k+1}^{n} \lambda_i$ from $\Sigma\Phi_i = \lambda_i\Phi_i$. Since $\mu = \sum_{i=1}^{n} (\mu'\Phi_i)\Phi_i = \sum_{i=1}^{k}(\mu'\Phi_i)\Phi_i + \sum_{i=k+1}^{n} (\mu'\Phi_i)\Phi_i$ it is seen that by utilizing only the first k principal components μ can be approximated with an error of $\sum_{i=k+1}^{n} (\mu'\Phi_i)\Phi_i$ (i.e., one obtains the projection of the vector μ onto the subspace spanned by the first k eigenvectors). From the properties of eigenvectors, the expression $\Sigma = \Phi\Lambda\Phi'$ can be derived where Φ is the matrix composed of the columns Φ_i and Λ is the diagonal matrix of eigenvalues λ_i. The above expression also can be written as $\Sigma = (\Phi\Lambda^{1/2})(\Phi\Lambda^{1/2})' = E(\eta\eta')$. Define the random vector $F' = (f_1, \ldots, f_n)$ with the properties $E(FF') = I$ and $E(F) = 0$. Then the above matrix decomposition can be written as $\Sigma = E(\Phi\Lambda^{1/2}F)(\Phi\Lambda^{1/2}F)' = E(\eta\eta')$. From the previous Karhunen–Loeve expansion, $\eta = \sum_{i=1}^{n} a_i\Phi_i = \Phi a$ where $a = (a_1, \ldots, a_n)'$. The vector $F = \Lambda^{-1/2}a = \Lambda^{-1/2}\Phi'\eta$ so that the new variables are scaled principal components. These F values are called factor scores. η can be expressed as $\Phi\Lambda^{1/2}F$ and the matrix $\Phi\Lambda^{1/2}$ is referred to as the factor loading matrix. For a given orthogonal matrix B, $\eta = (\Phi\Lambda^{1/2}B)(B'F)$ so that $B'\Lambda^{-1/2}\Phi'\eta = B'F$. Hence, a rotated factor loading matrix results in factor scores for rotated factors. The mean square error and the signal error described above remain the same under orthogonal rotations.

If we wish to study k groups in terms of common principal components, then we require that $\Sigma_1 = \ldots = \Sigma_k = \Sigma$, so that a set of n common principal components can be obtained. The signals $\mu_1, \ldots, \mu_k$ can be decomposed into the following forms $\mu_i = \sum_{j=1}^{n} (\mu_i'\Phi_j)\Phi_j$ for $i = 1, \ldots, k$. If it were known, somehow, that the k signals made no significant contributions to the subspace spanned by the last l eigenvectors then we could restrict our analyses to the first $n - l$ eigenvectors. However, since we have assumed no relationship among the signals and the noise there is no justification for assuming that the μ_i will lie in the subspace spanned by the first $n - l$ eigenvectors. It is customary in ERP research to restrict analysis to those eigenvectors with corresponding eigenvalues larger than some chosen value. While this procedure will result in an acceptable mean square error for the noise, it will not necessarily result in an acceptable error for the signal.

The decomposition of $\mu = \sum_{i=1}^{n} (\mu'\Phi_i)\Phi_i$ by the Karhunen–Loeve expansion provides an orthogonal decomposition of a signal in relation to a specific stochastic noise process η. For another stochastic noise process η_1, an orthogonal decomposition of μ in terms of a different set of orthogonal functions will generally be obtained. Thus, it is not generally possible to compare the ERPs from several leads on the basis of common components. Also, comparing the ERPs from the same lead under different experimental conditions is difficult because it must be assumed that η and η_1 are not different if we wish to compare common components. When comparing q leads, if one assumes that $\Sigma_1 = \ldots = \Sigma_q = \Sigma$ then it is possible to obtain a single set of eigenvalues and eigenvectors for all the leads. However, this requires the assumption that the noise processes for each lead are the same. Consequently, PCA does not provide a methodology which can practically handle multiple electrodes with different noise processes, thereby enabling a component to be described in terms of scalp distributions.

Statistical Properties of PCA

The above discussion has been concerned with the probability model of the stochastic process η. Since, in general, population parameters are not known, statistical estimates must be determined. For the single population problem the parameters $\mu, (\lambda_i, \Phi_i)$, $i = 1, \ldots, n$ must be estimated from a sample of N observation vectors $x_1, \ldots, x_N$. If the observation vectors x_i are obtained from a multivariate normal distribution then the usual maximum likelihood estimates are $\hat{\mu} = 1/N \sum_{i=1}^{N} x_i$ and $(1/(N-1)) \sum_{i=1}^{N} (x_i - \hat{\mu})(x_i - \hat{\mu})' = \hat{\Sigma}$. The estimates $\{(\hat{\lambda}_i, \hat{\Phi}_i), i = 1, \ldots n\}$ are obtained from $\hat{\Sigma}$. If rotated factors are obtained for analysis then it should be noted that the rotation matrix B is random because it is sample dependent. From the factor analysis literature there are various "rules-of-thumb" in terms of the ratio N/n for specifying a sample size large enough to provide stable estimates for factor loading matrices. Because of the large number of time points and the small number of subjects in the typical ERP study, the ratio N/n will usually not be large enough to obtain stable estimates of the factor loading matrix. This implies that the interpretation of the components is not very reliable. In particular, there can be considerable variation in the amplitudes and latencies of waveform peaks from sample to sample or study to study. Because of the limited sample size a subset of the eigenvectors is all that is estimable resulting in, at best, a subspace S spanned by this subset of eigenvectors. Consequently, only the projection of the signal onto S can be studied. As discussed above, since we have not assumed a relationship between μ and η, there is no particular reason to assume that μ will lie in S.

From above it is seen that $x = \sum_{i=1}^{n} (\mu'\Phi_i)\Phi_i + \sum_{i=1}^{n} a_i\Phi_i = \sum_{i=1}^{n}(a_i + \mu'\Phi_i)\Phi_i = \sum_{i=1}^{n} y_i\Phi_i$ where the y_i are independent with $E(y_i) = \mu'\Phi_i = v_i$. Using the N observation vectors we can test the hypothesis $\{v_i = 0; i = 1, \ldots, n\}$ by performing independent t tests. However, as discussed above, we can only examine a subset of the v_i.

In the k-group problem we can estimate the parameters for each group in a manner similar as the above. In order to be able to compare identical principal components in the k groups it is often assumed in ERP studies that $\Sigma_1 = \ldots = \Sigma_k = \Sigma$ so that the common principal components can be obtained from Σ. If the observation vectors $\{x_{ij}: i = 1, \ldots, k; j = 1, \ldots, N_i\}$ are obtained from k multivariate normal distributions then the usual maximum likelihood estimates are $\hat{\mu}_i = (1/N_i) \sum_{j=1}^{N_i} x_{ij}$, $i = 1, \ldots, k$ and $\hat{\Sigma} = (1/(N_1 + \ldots + N_k - k)) \sum_{i=1}^{k} \sum_{j=1}^{N_i} (x_{ij} - \hat{\mu}_i)(x_{ij} - \hat{\mu}_i)'$. The estimates $\{\hat{\lambda}_l, \Phi_l; l = 1, \ldots n\}$ are obtained from $\hat{\Sigma}$. From above, for observation x in group i, it is seen that $x = \sum_{j=1}^{n} (\mu_i'\Phi_j)\Phi_j + \sum_{j=1}^{n} a_j\Phi_j = \sum_{j=1}^{n} (a_j + \mu_i'\Phi_j)\Phi_j = \sum_{j=1}^{n} y_{ij}\Phi_j$ where y_{ij} are independent with $E(y_{ij}) = \mu_i'\Phi_j = v_{ij}$. Using the $N = N_1 + \ldots + N_k$ observation vectors we can test the hypotheses $\{v_{ij} = \ldots = v_{kj}; j = 1, \ldots, n\}$ by performing analyses of variance. Usually, because of the limited sample sizes involved only a limited number of principal components can be examined. A question that frequently arises is whether or not one should use unrotated factors or rotated factors when comparing groups. If the group differences are examined with a multivariate analysis of variance then the results will be the same. This result follows from the fact that the MANOVA tests are invariant with respect to orthogonal transformations (*16*).

Rather than merely assuming that $\Sigma_1 = \ldots = \Sigma_k = \Sigma$ it would be better to test this hypothesis as in Anderson (*16*). Also, it would be instructive to examine the principal components in the k groups as in (*17*) and to test for common principal components (*18, 19*). Because of the small sample sizes it would usually not be possible to perform these between-groups comparisons in ERP studies.

Another aspect of PCA signal decomposition should be discussed at this point. If it is assumed that a particular signal consists of the sum of several component waveforms from physiological generators then the PCA analysis will not necessarily provide estimates of the component waveforms. The eigenvectors are usually assumed to represent the component waveforms. These eigenvectors are derived from the noise covariance matrix, and we have not assumed any relationship between the signal and noise. In Van Rotterdam (*7*) it is shown mathematically that there is no justification for assuming that the orthogonal principal components correspond to independent physiological generators. In (*7*) the derived principal components, from nonorthogonal components, are essentially mathematical artifacts. In Woods and McCarthy (*8*) an empirical investigation is conducted using PCA and nonorthogonal components. Even though the derived components do not appear graphically to differ greatly from the true component waveforms, the resultant analyses of variance were considerably different from the true results. Since, the same Σ can be obtained from many diverse models we should not expect PCA to provide a unique solution. One of the reasons for utilizing PCA methodology in ERP studies is to obtain orthogonal variables rather that correlated variables so that more reliable statistical inference can be obtained from the multivariate vectors of large dimension. However, because of the small sample sizes typically encountered, it is not possible to use all of the eigenvectors. In order to resolve the problem of reliably estimating the large number of parameters another approach is to restrict the class of noise processes to those with specially structured covariance matrices containing fewer parameters. A commonly used covariance matrix model is the Toeplitz matrix form which arises when considering stationary time series. In the next section, it will be described how the use of a stationary noise model will allow orthogonal components which are deterministic (i.e., not sample dependent), and will allow rigorous methods of statistical inference for multiple lead and multiple population problems. It will also be shown how the principal components analysis of Toeplitz matrices is related to the spectral analysis of stationary time series.

Spectral Analysis

Consider the p-dimensional multivariate time series $\mathbf{x}_{jl}(t) = \boldsymbol{\mu}_j(t) + \boldsymbol{\eta}_{jl}(t)$ where $t = 0, 1, \ldots, T-1; j = 1, \ldots, q; l = 1, \ldots, N_j$ such that q is the number of groups and N_j denotes the number of independent time series in the jth group. $\boldsymbol{\eta}_{jl}(t)$ are zero-mean stationary Gaussian processes with the cross-correlation matrix $R_j(t-u) = [\gamma_j^{mn}(t-u); m, n = 1, \ldots, p]$ and $\boldsymbol{\mu}_j(t)$ denotes the event-related potential for group j. The spectral density matrix $F_j(\lambda)$ for group j is defined by the representation $R_j(t-u) = \int_{-\pi}^{\pi} e^{i\lambda(t-u)}F_j(\lambda)d\lambda$. The multivariate finite Fourier transform (FFT) of $\mathbf{x}(t)$ is defined by $\hat{\mathbf{x}}(k) = \sum_{t=0}^{T-1} \mathbf{x}(t)e^{-i\lambda_k t}$ where $\lambda_k = 2\pi k/T$ for $k = 0, \ldots, T-1$. When the FFT is applied to the above time series model, $\hat{\mathbf{x}}_{jl}(k) = \hat{\boldsymbol{\mu}}_j(k) + \hat{\boldsymbol{\eta}}_{jl}(k)$ is obtained. From Hannan (*20*), it is known that $\hat{\boldsymbol{\eta}}_{jl}(k)$ has a multivariate complex normal distribution, $N(0,2\pi TF_j(\lambda))$ and for large T the vectors $\hat{\boldsymbol{\eta}}_{jl}(k)$ and $\hat{\boldsymbol{\eta}}_{jl}(k')$ are independent for $k \neq k'$. By performing the inverse FFT, the following representation of $\boldsymbol{\mu}(t) = 1/T \sum_{k=0}^{T-1} \hat{\boldsymbol{\mu}}(k)e^{i\lambda_k t}$ can be obtained in terms of the complex trigonometric terms $e^{i\lambda_k t}$.

The sampling interval is determined by the choice of the highest frequencies to be included in the analyses (Nyquist frequency). If there is little power at frequencies beyond the Nyquist frequency then the effects of aliasing can be avoided. Otherwise, the power at the higher frequencies should be removed by preceding A-D conversion by analog filtering. Empirically, we have found that a sampling interval of 7 msec is satisfactory with a corresponding Nyquist frequency of 71 Hz. The length of the sample, T, in PCA analysis is usually determined by the "components" of interest (e.g., P300). Much of current interest in ERP research is with the P300 and later components (*21*). Reasonably large values of T are necessary since the P300 component can have peak latencies of 450 msec or more, and is followed by later components. Empirically, sample lengths of $T = 100$ (700 msec) have been found to be satisfactory with the lowest nonzero frequency available being $1/T$ cycles/unit time (1.42 Hz). Otherwise, one may be required to "taper" the data to reduce the bias produced by using a finite T. In Shumway (*22*) it is shown that the spectral density matrices can be approximated with a bias of only $0(T^{-1})$. While this sample length is large enough for good statistical estimates it is not so large that a change in stationarity is a major problem (the issue of stationarity will be discussed later). The choice of the number of replications, N_j, is determined, as usual, by the types of statistical analyses to be performed (e.g., discriminant analysis or multivariate analysis of variance at a specific frequency). In this case, however, we must consider complex multivariate normal distributions rather than real multivariate normal distributions. If only small numbers of replications are available then it is possible that statistical estimates and tests can be improved by frequency averaging as described in Shumway (*22*).

The relationship of the above representation for $\boldsymbol{\mu}(t)$ to the principal components analysis of the Toeplitz matrix for the noise process $\boldsymbol{\eta}(t)$ will be discussed for the univariate case. In Brillinger (*23*) it is shown that the eigenvectors for the Toeplitz matrix are approximately

$$\left\{\frac{1}{\sqrt{T}}e^{i\lambda_k 0}, \frac{1}{\sqrt{T}}e^{i\lambda_k 1}, \ldots, \frac{1}{\sqrt{T}}e^{i\lambda_k(T-1)}\right\}' = \frac{1}{\sqrt{T}}\psi_k \quad \text{for } k = 0, \ldots, T-1.$$

The principal components are given by $a_k = (1/\sqrt{T})\psi_k^*\eta$ where $\eta = (\eta(0), \ldots, \eta(t-1))'$ and $\sqrt{T}(a_0, \ldots, a_{T-1})' = \sqrt{T}a$ is the FFT for η. An approximation to η in terms of principal components of the Toeplitz matrix is $(1/\sqrt{T}) \sum_{k=0}^{T-1} a_k\psi_k = (1/\sqrt{T})\psi a$. It should be observed that the eigenvectors for this special covariance structure are not sample dependent. It is noted in (*23*) that the Cramer representation, $x(t) = \int_0^{2\pi} e^{i\lambda t}dz_x(\lambda)$, is the limit, as T increases, of the principal component representation of $\eta(t)$. It is also shown that eigenvalues of the Toeplitz matrix are approximately $2\pi f(2\pi k/T)$ for $k = 0, \ldots, T-1$, where $f(\lambda)$ is the spectral density matrix of $\eta(t)$. The mean μ, has the representation $(1/\sqrt{T}) \sum_{i=0}^{T-1} b_i\psi_i$ in terms of the complex trigonometric functions, where $\sqrt{T}(b_0, \ldots, b_{T-1})'$ is the FFT for μ. Using the result $b_k = b^*_{T-k}$, it can be shown that μ has a representation in terms of real trigonometric functions. Let T be odd and let $b_j = c_j + id_j$ where c_j and d_j are the real and imaginary components of b_j. Consider the expression

$$\frac{1}{\sqrt{T}} \sum_{j=0}^{T-1} b_j\psi_j = \frac{1}{\sqrt{T}} b_0\psi_0 + \frac{1}{\sqrt{T}} \sum_{j=1}^{(T-1)/2} (b_j\psi_j + b_j^*\psi_j^*).$$

It can be shown that $b_je^{i\lambda_jt} + b_j^*e^{-i\lambda_jt} = 2(c_j\cos \lambda_jt - d_j\sin \lambda_jt)$. The Polar representation of a complex number results in $c_j = \gamma_j\cos \Theta_j$, $d_j = \lambda_j\sin \Theta_j$ where $\gamma_j = \sqrt{c_j^2 + d_j^2}$ and $d_j/c_j = \tan \Theta_j$ so that $\Theta_j = \tan^{-1} (d_j/c_j)$. We obtain $b_je^{i\lambda t} + b_j^* e^{-i\lambda_jt} = 2(\gamma_j\cos \Theta_j\cos \lambda_jt - \gamma_j\sin \Theta_j\sin \lambda_jt)$. Let $\Phi_j = -\Theta_j$ so that $2(\gamma_j\cos \Phi_j\cos \lambda_jt + \sin \Phi_j\sin \lambda_jt) = 2\gamma_j\cos (\lambda_jt + \Phi_j)$ is obtained. Hence, μ can be represented as a sum of cosine functions with different amplitudes, phases, and frequencies, or we can say that μ can be represented by a sum of harmonic oscillators. A similar result holds in the case that T is even.

It should be noted that in the stationary noise case it is not necessary to compute the eigenvalues and eigenvectors by the usual Principal components methods since these estimates can be obtained by utilizing the simpler FFT methods. In any case, the usual PCA methods will not be correct unless the Toeplitz matrix form of the covariance matrix is taken into consideration. An estimate of the covariance matrix can be obtained by utilizing the FFT of the spectral density function of the noise (*23*).

An important question is whether the signal imbedded in stationary noise model, SSN, can provide useful results in ERP studies. In Rawlings (*24*) discriminant analyses were performed on both the SSN model, and the signal imbedded in nonstationary noise model, SNN. These analyses were performed specifically to determine if there was any evidence of deterioration of performance due to nonstationarity. While there was indication that the SNN model performed better in some cases, it was also evident that the SSN model performed well relative to the SNN model. More importantly, for the multiple lead (four scalp electrodes) problem, for which the SNN model cannot practically be used, the spectral discriminant analysis obtained the best overall nonerror rate (81%) using the leaving-one-out estimate. The model $\hat{\mathbf{x}}_{jl}(k) = \hat{\boldsymbol{\mu}}_j(k) + \hat{\boldsymbol{\eta}}_{jl}(k)$ discussed at the beginning of this section, is the complex analog of the general linear model, and Goodman (*25*) and Khatri (*26*) have derived the usual likelihood ratio tests. Shumway (*22*) discussed the use of the likelihood ratio tests, $L(\lambda_k)$ and $L'(\lambda_k)$ for testing the equality of group means at each frequency, as well as testing the equality of group spectral matrices at each frequency, respectively. Rawlings *et al.* (*24*) used the likelihood ratio test, $L''(\lambda_k)$, for testing the homogeneity of complex multivariate normal populations, while Giri (*27*) derived the complex analog of the multivariate T^2 test. The tests $L(\lambda_k)$, $L'(\lambda_k)$ and $L''(\lambda_k)$ allow one to perform multivariate analysis of variance, and linear or quadratic discriminant analysis. The Giri (*27*) results provide for mean comparisons among correlated variates at each frequency (i.e., provides for a test of equal power of a signal at several leads at each frequency). There is considerable evidence in ERP research that stationarity of EEG activity varies over time as a function of cerebral state. The result of this is that the spectral density matrices will be different under different cerebral states. These spectral matrix differences can be tested using $L'(\lambda_k)$ and the spectral quadratic discriminant functions as constructed in (*24*) take efficient advantage of these EEG cerebral state differences. The SSN model has been utilized successfully by Woestenberg *et al.* (*28*) in developing a statistical Wiener filter for a single lead in ERP experiments. This type of filter can be easily generalized to multiple leads by using the results of Giri (*27*) so that a filter for the entire ERP scalp distribution can be developed. Also, Woestenberg *et al.* (*29*) utilized spectral regression methods and the SSN model to adjust for eye movement evoked potentials in EEGs.

Conclusions

While the Karhunen–Loeve method provides an orthonormal expansion for a general stochastic process in the case where the parameters of the process are known, it is not generally possible to obtain a complete orthonormal expansion when utilizing the small sample sizes usually encountered in ERP studies because of the large number of parameters involved. Hence, one cannot represent arbitrary ERP waveforms in terms of the derived orthonormal basis. In the case of multiple leads, each lead has its own principal components expansion and each expansion will generally be in terms of different orthonormal bases. Consequently, unless one assumes that the covariance matrices are the same for each lead, it is not possible to compare identical components across leads. Similar problems are encountered when studying multiple populations with quadratic discriminant analysis.

If instead of considering a general noise process, a stationary noise process is used the difficulties described above can be overcome. Principal components analysis and spectral analysis have been shown to be equivalent in this case. We have also discussed some of the methods of statistical inference which are available for this model. We have available the multivariate general linear model for complex variables, so that it is possible to analyze rigorously any ERP experiments designed according to the multivariate general linear model format. We have also discussed literature in which the SSN model has performed well and we are certain that this model can be usefully applied in many more complex studies. By utilizing the SSN model we can test scientifically useful multivariate hypotheses not meaningful in the case of a general noise problem. Perhaps when more is understood about the physiological mechanism generating the ERP and noise we can restrict the noise process to a class of nonstationary processes with few parameters (e.g., nonstationary autoregressive), and we can restrict the ERP to a functional form with few parameters (e.g., exponentially damped sine waves).

REFERENCES

1. HORST, R. L., AND DONCHIN, E. Beyond Averaging. II. Single-trial classification of exogenous event-related potentials using stepwise discriminant analysis. *Electroenceph. Clin. Neurophysiol.* **48**, 113–126 (1980).
2. SQUIRES, K. C., DONCHIN, E., HERNING, R., AND MCCARTHY, G. On the influence of task relevance and stimulus probability in event-related-potential components. *Electroenceph. Clin. Neurophysiol.* **42**, 1–14 (1977).
3. WOOD, C. C., ALLISON, T., GOFF, W. R., WILLIAMSON, P., AND SPENCER, D. On the neural origin of P300 in man. *Prog. Brain Res.* **54**, 51–56 (1980).
4. HALGREN, E., SQUIRES, N., WILSON, C., ROHRBAUGH, J., BABB, T., AND CRANDALL, P. Endogenous potentials generated in the human hippocampal formation and amygdala by infrequent events. *Science* **210**, 803–805 (1980).
5. FREEMAN, W. J. Linear analysis of the dynamics of neural masses. *Annu. Rev Biophys. Bioeng.* **1**, 225–256 (1972).
6. NUNEZ, P. L. The brain wave equation: A model for the EEG. *Math. Biosci.* **21**, 279–297 (1974).
7. VON ROTTERDAM, A. Limitations and difficulties in signal processing by means of the principal-components analysis. *IEEE Trans. Biomed. Eng.* **BME-17**, 268–269 (1970).
8. WOOD, C. W., AND MCCARTHY, G. Principal component analysis of event-related potentials: Simulation studies demonstrate misallocation of variance across components. *Electroenceph. Clin. Neurophysiol.* **59**, 249–260 (1984).
9. ROSLER, F., AND MANZEY, D. Principal components and varimax-rotated components in event-related potential research: Some remarks on their interpretation. *Biol. Psychol.* **13**, 3–26 (1981).
10. WASTELL, D. On the correlated nature of evoked brain activity: Biophysical and statistical considerations. *Biol. Psychol.* **13**, 51–69 (1981).
11. WALTER D. O. Spectral analysis for electroencephalograms: Mathematical determination of physiological relationships from records of limited duration. *Exp. Neurol.* **8**, 155–181 (1963).
12. BURCH, N. R. Period analysis of the electroencephalograms on a general-purpose digital computer. *Ann. N.Y. Acad. Sci.* **115**, 827–843 (1964).
13. LARSEN, L. E., AND WALTER, D. O. On automatic methods of sleep staging by EEG spectra. *Electroenceph. Clin. Neurophysiol.* **28**, 459–467 (1970).
14. DUMERMUTH, G. Numerical spectral analysis of the electroencephalogram. *In* "Handbook of EEG Clin. Neurophysiol." (A. Rémond, Ed.), Vol. 5A. Elsevier, Amsterdam, 1973.
15. FUKUNAGA, K. "Introduction to Statistical Pattern Recognition." Academic Press, New York, 1972.
16. ANDERSON, T. W. "An Introduction to Multivariate Statistical Analysis." Wiley, New York, 1958.
17. KRZANOWSKI, W. J. Between-groups comparison of principal components. *J. Amer. Stat. Assoc.* **74**, 703–707 (1979).
18. KRZANOWSKI, W. J. Principal component analysis in the presence of group structure. *Appl. Stat.* **33**, 164–168 (1984).
19. FLURY, B. N. Common principal components in K groups. *J. Amer. Stat. Assoc.* **79**, 892–898 (1984).
20. HANNAN, E. J. "Multiple Time Series." Wiley, New York, 1970.
21. H. BEGLEITER (Ed.) "Evoked Brain Potentials and Behavior. Plenum, New York, 1979.
22. SHUMWAY, R. H. Discriminant analysis for time series. *In* "Handbook of Statistics," Vol. 2, "Classification, Pattern Recognition, and Reduction of Dimensionality." North-Holland, Amsterdam, 1982.
23. BRILLINGER, D. "Time Series Data Analysis and Theory." Holden-Day, San Francisco, 1981.
24. RAWLINGS, R., ECKARDT, M., AND BEGLEITER, H. Multivariate time series discrimination in the spectral domain. *Comput. Biomed. Res.* **17**, 352–361 (1984).

25. GOODMAN, N. R. Statistical analysis based on a certain multivariate complex distribution (an introduction). *Ann. Math. Stat.* **34,** 152–177 (1963).
26. KHATRI, C. G. Classical statistical analysis based on a certain multivariate complex Gaussian distribution. *Ann. Math. Stat.* **36,** 98–114 (1965).
27. GIRI, N. On the complex analogues of T^2 and R^2 tests. *Ann. Math. Stat.* **36,** 664–670 (1965).
28. WOESTENBURG, J. C., VERBATEN, M., SJOUW, W., AND SLANGEN, J. A statistical Wiener filter using complex analysis of variance. *Biol. Psychol.* **13,** 215–225 (1981).
29. WOESTENBURG, J., VERBATEN, M., AND SLANGEN, J. The removal of the eye-movement artifact from the EEG by regression analysis in the frequency domain. *Biol. Psychol.* **16,** 127–147 (1983).