An Efficient Approach for Statistical Calculations with Globally Gridded Filtered Time Series

X. H. GAO AND J. L. STANFORD

Physics Department, Iowa State University, Ames, Iowa

9 February 1987 and 1 February 1988

1. Introduction

In atmospheric observational studies many statistical concepts and calculations are involved, for example, variance, spectrum, correlation, EOF (Empirical Orthogonal Function), EEOF (Extended Empirical Orthogonal Function), etc. Some statistical calculations are quite simple, while others may be lengthy and expensive. In many cases, use of the frequency domain can result in improved efficiency for statistical calculations with time-filtered data. Only the Fourier coefficients within the filtered frequency band are significantly nonzero; moreover, these coefficients contain all the information of interest. This paper discusses several statistical calculations based on Fourier coefficients and presents the formulas needed to effect these calculations:

1) A method is given which allows an estimate of the degrees of temporal freedom of two correlated time series, utilizing the frequency spectra of these two time series.

2) A simple way to perform seasonal analyses is proposed which employs a half-year summer/winter projection operator in the frequency domain.

3) A modified lag-correlation calculation is suggested from which lag correlations may be easily obtained in the frequency domain.

4) A spectral approach is presented for EOF and EEOF analyses which reduces the size of the matrix to be solved in the eigenproblem. No matter how many grid points are covered by the analyzed area or how many lag steps are involved, the size of the matrix to be solved is always $s \times s$ where s is the number of Fourier coefficients in the filter bandpass, resulting in a significant reduction in computation time.

Some of these calculation methods have been used in the study of low frequency oscillations in the large-scale stratospheric temperature field (Gao and Stanford 1988).

Corresponding author address: Prof. John L. Stanford, Dept. of Physics, Iowa State University, 12 Physics Building, Ames, Iowa 50011.

2. Estimation of confidence level for correlations

a. General concepts

Observational data are often in the form of time series $T(r, t)$, with $t = 1, 2, \cdots, N$. Here N is the length of the time series, and r is an index (or indices) other than t, such as a grid point index. Consider an N-dimensional space spanned by a set of N orthonormal basis vectors $\mathbf{e}_t$. Then the time series $T(r, t)$ can be represented by an N-dimensional vector $\mathbf{T}(r)$ having projections $T(r, t)$ on axes $\mathbf{e}_t$:

$$\mathbf{T}(r) = \sum_t T(r, t)\mathbf{e}_t \tag{1}$$

If the time mean is removed from the data, as is usually done, then the variance of the time series $T(r, t)$ is

$$\sigma_r^2 \equiv \sum_t T^2(r, t) = \mathbf{T}(r)\cdot\mathbf{T}(r) = |\mathbf{T}(r)|^2, \tag{2}$$

the square of the modulus of vector $\mathbf{T}(r)$.

The correlation between these two time series is defined as

$$R_{rq} \equiv \sum_t T(r, t)\cdot T(q, t)/\sigma_r\sigma_q = \frac{\mathbf{T}(r)}{\sigma_r}\cdot\frac{\mathbf{T}(q)}{\sigma_q}, \tag{3}$$

the scalar product of two unit vectors, $\mathbf{T}(r)/\sigma_r$ and $\mathbf{T}(q)/\sigma_q$.

b. Confidence level from m-dimensional probability

To estimate the confidence level for the correlation, we start from the calculation of a quantity $P(R_c, m)$ which is the probability of obtaining a correlation $R_{ij} \geqslant R_c$, in m-dimensional space. Here R_{ij} is the correlation between two completely random time series $T(i, t)$ and $T(j, t)$ each of length m. The calculation can be visualized geometrically as the calculation of $P(R_c, m)$ for all dot products $R_{ij} = \mathbf{T}(i)\cdot\mathbf{T}(j)/\sigma_i\sigma_j \geqslant R_c$. Since $\mathbf{T}(i)$, $\mathbf{T}(j)$ are completely random, their distributions in m-dimensional space must be spherically symmetric. If $\mathbf{T}(i)/\sigma_i$ is considered fixed along some direction in the m-dimensional space, say $\mathbf{T}(i)$ is parallel to $\mathbf{x}_1$, the distribution of $\mathbf{T}(j)$ with respect to $\mathbf{T}(i)$ will also be seen to be spherically symmetric. Therefore,

$$R_{ij} = \frac{\mathbf{T}(i)}{\sigma_i} \cdot \frac{\mathbf{T}(j)}{\sigma_j} \geqslant R_c \tag{4}$$

gives a 'super cone' in the m-dimensional space. With this geometrical visualization it can be seen that the probability of obtaining a random correlation value greater than R_c is

$$P(R_c, m) = V_c(R_c, m)/V_s(m), \tag{5}$$

where $V_s(m)$ is the volume of the m-dimensional sphere of radius R, and $V_c(R_c, m)$ is the volume of the common part of this sphere and the cone defined by (4), as shown in Fig. 1 for $m = 3$.

1) CALCULATION OF $V_s(m)$

Suppose $A_m = \beta_m r^{m-1}$ is the "surface area" of an m-dimensional sphere of radius r. Here β_m is a function of m, whose form is to be determined. Then

$$V_s(m) = \int_0^R A_m dr = \frac{\beta_m}{m} R^m. \tag{6}$$

Using

$$\int_0^\infty A_m(r) e^{-r^2} dr = \prod_{i=1}^m \int_{-\infty}^\infty e^{-x_i^2} dx_i = \pi^{m/2},$$

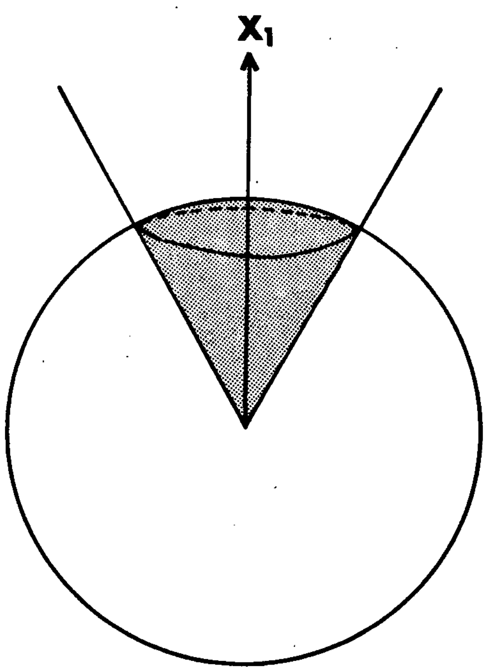

FIG. 1. Intersection of the m-dimensional sphere and cone, illustrated for $m = 3$.

β_m is determined by

$$\beta_m = \pi^{m/2}\left[\int_0^\infty r^{m-1} \cdot e^{-r^2} dr\right]^{-1}$$

$$= (2\pi)^{m/2}[(m-2)!!]^{-1} \begin{cases} \sqrt{2/\pi}, & \text{for } m \text{ odd.} \\ 1, & \text{for } m \text{ even.} \end{cases} \tag{7}$$

As a check, $\beta_3 = 4\pi$ in agreement with the surface area of a three-dimensional sphere. In general

$$V_s(m) = \frac{1}{m} \beta_m R^m = \frac{1}{m} A_m R$$

$$A_m(R) = \beta_m R^{m-1}$$

$$\beta_m = (2\pi)^{m/2}[(m-2)!!]^{-1} \begin{cases} \sqrt{2/\pi}, & \text{for } m \text{ odd.} \\ 1, & \text{for } m \text{ even.} \end{cases}$$

2) CALCULATION OF $V_c(R_c, m)$

The m-dimensional space cone (see Fig. 2) is given by the equation

$$\sum_{i=2}^m x_i x_i \leqslant k x_1, \tag{8}$$

with $\tan\phi = k = (1 - R_c^2)^{1/2}/R_c$, $x_0 = R_c R$. The volume V_c, defined as that common to the cone and the sphere, is

$$V_c(R_c, m) = \int_0^{x_0} dx \left[\beta_{m-1}(xk)^{m-1} \frac{1}{m-1}\right]$$

$$+ \int_{x_0}^R dx \frac{\beta_{m-1}}{m-1} [(x+R)(R-x)]^{(m-1)/2}. \tag{9}$$

Letting $x/R = \cos\alpha$,

$$V_c(R_c, m) = \frac{\beta_{m-1}}{\beta_m} V_s(m) \int_0^\phi \sin^{m-2}\alpha d\alpha. \tag{10}$$

Using the results obtained in (10),

$$P(R_c, m) = \frac{\beta_{m-1}}{\beta_m} \int_0^\phi \sin^{m-2}\alpha d\alpha, \tag{11}$$

where $P(R_c, m)$ can be easily obtained by numerical calculation. For correlations between two time series, the 5% confidence level can be obtained by $P(R_c, m) = 0.05$. Here m is the degrees of temporal freedom which we will discuss next.

c. Estimation of degrees of temporal freedom

The degrees of temporal freedom of two time series involved in a correlation calculation are usually much less than their length N, especially when bandpass filtering has been employed. To estimate the degrees of temporal freedom, the following procedure is usually

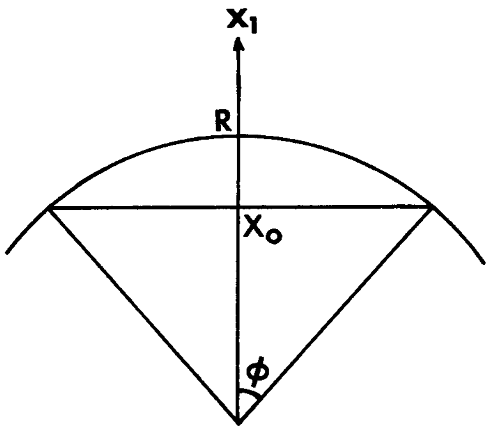

FIG. 2. Illustration of the integration limits for $V_c(R_c, m)$, Eq. (9).

used. The degrees of temporal freedom (DOTF) are estimated by

$$\text{DOTF} \approx L/\tau, \qquad (12)$$

where $L = N\Delta t$ is the length of the time series, and

$$\tau = [1 + 2\sum_{i=1}^{N} \gamma_r(i)\gamma_q(i)]\Delta t. \qquad (13)$$

Here the γ are autocorrelations at lags $i\Delta t$ (Livezey and Chen 1983). Equation (13) is not suitable for calculation in the frequency domain. One should find a formula to estimate τ in the frequency domain. Equation (13) is a modified form of the original formula given by Davis (1976) for two time series of infinite length,

$$\tau = \sum_{i=-\infty}^{\infty} \gamma_r(i)\gamma_q(i)\Delta t. \qquad (14)$$

If use is made of the Fourier expansion, with t varying from $-\infty$ to $+\infty$, the time series can be extended periodically to infinity. Then, using (12),

$$\text{DOTF} = \left[\frac{N^2}{2}\sum_{n=1}^{N/2} I_n(r)I_n(q)/\sigma_r^2\sigma_q^2\right]^{-1}, \qquad (15)$$

where $I_n(r) = \frac{1}{2}[C_n^2(r) + S_n^2(r)]$ is the power at frequency ω_n, C_n and S_n being the Fourier cosine and sine coefficients, except $I_{N/2}(r) = C_{N/2}^2(r)$. Here we have used $\gamma_r(i) = N\sum_{n=1}^{N/2} I_n(r)\cos\omega_n i/\sigma_r^2$. Equation (15) provides a method of estimating DOTF from the spectra of the time series.

In the band-limited white-noise case,

$$I_n(r) \approx I_n(q) \approx \frac{\sigma^2}{n_f N}, \qquad (16)$$

with $\sigma = \sigma_r = \sigma_q$. Here n_f is the number of frequencies ω_n covered by the filter band. Therefore,

$$\text{DOTF} \approx 2n_f. \qquad (17)$$

Here $2n_f$ is the number of Fourier sine and cosine coefficients in the bandpass of the filter.

Madden (1976) has used spectral density to estimate the effective degrees of freedom for the time average of a single time series. Equation (15) can be considered as a generalization of his method.

3. Combination of frequency band-pass filter with seasonal projection

Correlations between time series are common calculations in statistical studies of atmospheric motions. In many cases the time series need to be filtered to enhance signals in a certain frequency band. It is also often of interest to investigate the seasonal dependence of correlations between these time series.

A straightforward way to accomplish this is to filter the datasets, then select out the segments of the time series for the seasons of interest, and finally perform the correlations in the ordinary way. This involves passing a digital filtering program over the data time series (not necessarily trivial, especially for low-frequency signals) or reconstructing the needed time series from the relevant Fourier coefficients (which also can require a significant amount of computing time for long record lengths). In addition, the ordinary method of correlation calculations in the time domain is itself often quite lengthy.

The following provides a different technique for calculating correlations of bandpass filtered, seasonally projected time series. It is based on Fourier coefficients and takes advantage of a seasonal projection operator, both of which enhance the efficiency of calculation.

Let $|T_i\rangle$ and $|T_j\rangle$ be two time series each an exact integral number of years in length, and let Γ be a projection operator. For example, a winter or summer half-year projection would be obtained by $\Gamma|T\rangle$. As discussed before, a time series can be represented by an N-dimensional vector, where N is the length of the time series. The projection operator Γ is represented by a Hermitian matrix and satisfies the relations

$$\Gamma = \Gamma^\dagger, \quad \Gamma^2 = \Gamma. \qquad (18)$$

Here $\Gamma^\dagger$ indicates the transpose, complex conjugate of matrix Γ. The correlation between $\Gamma|T_i\rangle$ and $\Gamma|T_j\rangle$ can be obtained by

$$\begin{aligned} R_{ij} &= \langle T_i|\Gamma^\dagger\Gamma|T_j\rangle \cdot [\langle T_i|\Gamma\Gamma|T_i\rangle]^{-1/2} \\ &\quad \times [\langle T_j|\Gamma\Gamma|T_j\rangle]^{-1/2} \\ &= \langle T_i|\Gamma|T_j\rangle / \sqrt{\langle T_i|\Gamma|T_i\rangle\langle T_j|\Gamma|T_j\rangle}. \qquad (19) \end{aligned}$$

If the time series are bandpass filtered, the advantage of calculating R_{ij} in the frequency domain becomes apparent: only the components of the vector $\Gamma|T_j\rangle$ within the bandpass will contribute to R_{ij}, and usually the number of these components is much less than the length of the time series, N. This is to be contrasted with the usual method of calculation of correlations in the time domain which would have N terms.

As an example, suppose $|T_i\rangle$ and $|T_j\rangle$ are four years of 3-day mean temperatures, so the length of these time series is $N = 488$. For one 40–50 day bandpass filter, only 15 frequencies (30 Fourier coefficients) were used. In the time domain, if the data start at the beginning of summer, the summer (winter) half-year projection operator can be written as

$$\Gamma = \frac{1}{2}[1 \underset{(-)}{+} \theta(\sin\Omega t)]$$
$$= \frac{1}{2} \underset{(-)}{+} \frac{2}{\pi}\sum_{m=0}^{\infty}\frac{1}{2m+1}\sin(2m+1)\Omega t, \quad (20)$$

where the θ function is $1(-1)$ when its argument is $+(-)$, and Ω is the frequency of annual cycle. Since the time series $|T\rangle$ are bandpass filtered, they only have the spectral components within the filter bandpass. Therefore, in correlation calculations, only the first few terms in (20) make a contribution. For example, in the case of the 40–50 day band, the only contribution comes from the first three terms.

Operated on by Γ, the data series $|T\rangle$ becomes $\Gamma|T\rangle$, and the in-band coefficients are easily shown to be

$$S_l' = \frac{1}{2}S_l \pm \left(\frac{1}{\pi}C_{l-Y} - \frac{1}{\pi}C_{l+Y}\right) \pm \frac{1}{3\pi}(C_{l-3Y} - C_{l+3Y})$$
$$C_l' = \frac{1}{2}C_l \pm \left(\frac{1}{\pi}S_{l+Y} - \frac{1}{\pi}S_{l-Y}\right) \pm \frac{1}{3\pi}(S_{l+3Y} - S_{l-3Y}). \quad (21)$$

Here Y is the number of years covered by the dataset. The correlation function can then be calculated in the frequency domain and that calculation is much more efficient than usual calculations in the time domain.

4. The modification of lag correlation and EOF and EEOF analyses

a. *Lag correlation modification*

Lag correlations are very important in investigating the propagation properties of disturbance features. In extended empirical orthogonal function (EEOF) analysis the **R**-matrix to be solved for eigenvectors consists of correlations and lag correlations. For a large field with fine mesh, the **R**-matrix requires a large number of correlation and lag correlation computations, the number being $\frac{1}{2}\{I(J+1)[I(J+1)-1]\}$, where I is the number of grid points and J is the number of lag steps. For bandpass filtered time series, the degrees of temporal freedom are much less than the data length. Once the Fourier coefficients of the time series are obtained, calculation of the **R**-matrix in the frequency domain has obvious advantages. The **R**-matrix calculation can be accomplished with a modification of the lag correlation, together with certain approximations, as shown below.

The lag correlation between two data series $T(r, t)$ and $T(q, t)$ is

$$R_{rq}(\tau) = \frac{1}{\sigma_r'\sigma_q'}\sum_{t=1}^{N-\tau} T(r, t)T(q, t+\tau),$$

where

$$(\sigma_r')^2 = \sum_{t=1}^{N-\tau} T^2(r, t), \quad (\sigma_q')^2 = \sum_{\tau+1}^{N} T^2(q, t).$$

If $T(q, t)$ is extrapolated for an extra period τ by using the Fourier expansion of $T(q, t)$, then $T(q, N+s) = T(q, s)$, $s \leqslant \tau$. The $R_{rq}(\tau)$ is modified to

$$\bar{R}_{rq}(\tau) = \frac{1}{\sigma_r\sigma_q}\sum_{t=1}^{N} T(r, t)T(q, t+\tau)$$
$$= \frac{\mathbf{T}(r)}{\sigma_r}\cdot\frac{\mathbf{T}(q, \tau)}{\sigma_q}. \quad (22)$$

The Fourier coefficients of $T(q, t+\tau)$ are obtained by the following technique. Consider one frequency component, ω_n, for which the Fourier term is

$$S_n \sin\omega_n(t+\tau) + C_n \cos\omega_n(t+\tau)$$
$$= (S_n \cos\omega_n\tau - C_n \sin\omega_n\tau)\sin\omega_n t$$
$$+ (S_n \sin\omega_n\tau + C_n \cos\omega_n\tau)\cos\omega_n t. \quad (23)$$

The Fourier coefficients of $T(q, t+\tau)$ are seen to be

$$\begin{bmatrix} S_n(q, \tau) \\ C_n(q, \tau) \end{bmatrix} = \begin{bmatrix} \cos\omega_n\tau, & -\sin\omega_n\tau \\ \sin\omega_n\tau, & \cos\omega_n\tau \end{bmatrix}\begin{bmatrix} S_n(q) \\ C_n(q) \end{bmatrix}$$

Here $\omega_n\tau$ represents the angle rotated in the ω_n plane spanned by $\sum_t (\sin\omega_n t)\mathbf{e}_t$ and $\sum_t (\cos\omega_n t)\mathbf{e}_t$. Then the lag correlation

$$\bar{R}_{rq}(\tau) = \frac{\mathbf{T}(r)}{\sigma_r}\cdot\frac{\mathbf{T}(q, \tau)}{\sigma_q} \quad (24)$$

can be calculated in the frequency domain.

For $N \gg \tau$, $R_{rq}(\tau) \approx \bar{R}_{rq}(\tau)$, with relative error of about τ/N. Here N is the data record length.

b. *EOF and EEOF analyses*

EOF and EEOF analyses are mathematically eigenproblems, both related to solving a matrix equation of the form

$$\mathbf{R}|\mathbf{E}\rangle = \lambda|\mathbf{E}\rangle. \quad (25)$$

Here **R** is an $M \times M$ matrix, where M is the total number of time series involved in the analysis, λ is an eigenvalue, and $|\mathbf{E}\rangle$ an eigenvector. For example, the temperature field on a global $5° \times 5°$ grid contains $M = 72 \times 35 + 2 = 2522$ separate time series. So if one wants to perform an EOF analysis on these data, the matrix **R** to be solved is of size 2522×2522, too large even for present super computers. For EEOF analyses the situation is even worse, the matrix size being $(J + 1)$ times larger, when J steps of lag correlations (or covariances) are used. (It should be noted that high efficiencies can be obtained in the time domain using techniques such as those employed by Rasmusson et al. (1981) or Anderson and Rosen (1983). Here we discuss a technique using the frequency domain.) For many cases, the computational work can be reduced significantly, as detailed below.

We inspect two approaches to this eigenproblem. Suppose the data to be analyzed consist of M time series $T(r, t)$, each with N time steps: $r = 1, 2, \cdots, M$; $t = 1, 2, \cdots, N$. The data can be treated as M vectors in an N-dimensional space. They also can be treated as N vectors in a M-dimensional space, every vector representing a "pattern" describing the state of the field at time t. Let this state be denoted by

$$\mathbf{S}(t) = (T(1, t), T(2, t), \cdots, T(M, t))^{\mathrm{T}}, \quad t = 1, 2, \cdots, N. \tag{26}$$

where the superscript T denotes the transposed matrix. The eigenproblem related to EOF analyses is equivalent to the following geometrical problem: Find a unit vector ζ in the M-dimensional space which maximizes (minimizes) the sum of the squared projections of $\mathbf{S}(t)$ on ζ.

$$\sum_{t=1}^{N} (\zeta \cdot \mathbf{S}(t))(\mathbf{S}(t) \cdot \zeta) = \zeta \cdot [\sum_t \mathbf{S}(t)\mathbf{S}(t)] \cdot \zeta$$
$$\zeta \cdot \zeta = 1. \tag{27}$$

One approach to solving this problem is a variational method

$$\delta\{\zeta \cdot [\sum_t \mathbf{S}(t)\mathbf{S}(t)] \cdot \zeta\}$$
$$= \delta\zeta \cdot [\sum \mathbf{SS}] \cdot \zeta + \zeta \cdot [\sum \mathbf{SS}] \cdot \delta\zeta = 0.$$
$$\zeta \cdot \delta\zeta = 0. \tag{28}$$

The problem becomes an eigenproblem

$$[\sum_t \mathbf{SS}] \cdot \zeta = \lambda\zeta. \tag{29}$$

In matrix equation form (29) is the same as (25) and the **R**-matrix is the representation of $\sum_t \mathbf{SS}$. This is the method used in the usual EOF and EEOF analyses and requires the solution of the eigenproblem of the $M \times M$ matrix R. Frequently, when large area and high resolution are involved, the size of R becomes too large to handle.

Now, we search for another approach to solve the same problem. Let

$$\mathbf{S}(C_n) = [C_n(1), C_n(2), \cdots, C_n(M)]^{\mathrm{T}} \tag{30}$$

These are actually patterns of Fourier coefficients. It will be clear that for filtered data $\sum_t \mathbf{SS}$ can be more easily obtained in the frequency domain:

$$\sum_t \mathbf{SS} = \sum_{\Delta\omega} [\mathbf{S}(S_n)\mathbf{S}(S_n) + \mathbf{S}(C_n)\mathbf{S}(C_n)] \cdot \frac{N}{2}. \tag{31}$$

In many cases the bandpass of the filter is relatively narrow. In such situations, the number s of Fourier coefficients which lie in the bandpass is much less than N, the length of the time series. For example, as noted earlier, for four years of 3-day mean temperature data, $N = 488$, but for 40–50 day bandpassed signals only about 30 Fourier coefficients are involved. All of the analysis can be done by using only those coefficients.

Returning to the geometrical problem, there will now be only s vectors in M-dimensional space. In consequence, in the huge M-dimensional space, the Fourier coefficient vectors of interest here, $\mathbf{S}(S_n)$ and $\mathbf{S}(C_n)$, only span an s-dimensional subspace, with $s \ll M$. If, instead, the problem is solved in the s-dimensional subspace, a more efficient computation results.

Let $\mathbf{S}_f$ $(f = 1, 2, \cdots, s)$ be the s Fourier coefficient vectors. The unit vector ζ to be found in (27) can be decomposed into a linear superposition of $\mathbf{S}_f$,

$$\zeta = \sum_{f=1}^{s} \alpha_f \mathbf{S}_f. \tag{32}$$

Then the problem consists of maximizing

$$V(\alpha_f) \equiv \sum_{i=1}^{s} (\sum_f \alpha_f \mathbf{S}_f \cdot \mathbf{S}_i)^2 \tag{33}$$

with the restriction $\sum_{f,i} \alpha_i \alpha_f \mathbf{S}_f \cdot \mathbf{S}_i = 1$. This can be solved by the Lagrangian method. Let

$$G_{ij} \equiv \mathbf{S}_i \cdot \mathbf{S}_j = G_{ji}. \tag{34}$$

Then

$$\frac{\partial V(\alpha_f)}{\partial \alpha_j} = 0, \quad \text{yielding} \quad \sum_{f,i} \alpha_f G_{fi} G_{ij} = 0.$$

$$\frac{\partial}{\partial \alpha_j} \sum_{i,f} \alpha_i \alpha_f G_{fi} = 0, \quad \text{yielding} \quad \sum_f \alpha_f G_{fi} = 0.$$

Thus the problem becomes

$$GG|\alpha\rangle = \lambda G|\alpha\rangle \tag{35}$$

or

$$G|\alpha\rangle = \lambda|\alpha\rangle. \tag{36}$$

This is also an eigenproblem, but the matrix G is much smaller in size than that of R. Using the same example mentioned before, matrix **G** is on the order of 30×30 for the 40–50 day features, compared with matrix **R** of size 2522×2522 for global $5° \times 5°$ grids.

Once (36) is solved, (32) gives the desired eigenvectors (the EOFs or EEOFs), and the time variation of the coefficients of the EOF patterns can also be easily obtained. It is worth noting that the idea used in the modified lag correlation can also be used to simplify the calculation of the matrix **G** in EEOF analysis.

Acknowledgments. This paper is based upon work supported jointly by the National Science Foundation and the National Aeronautics and Space Administration under Grants ATM8402901 and ATM8603943.

REFERENCES

Anderson, J. R., and R. D. Rosen, 1983: The latitude-height structure of 40–50 day variations in atmospheric angular momentum. *J. Atmos. Sci.,* **40,** 1584–1591.

Davis, R. E., 1976: Predictability of sea surface temperature and sea level pressure anomalies over the North Pacific Ocean. *J. Phys. Oceanogr.,* **6,** 249–466.

Gao, X. H., and J. L. Stanford, 1988: Possible feedback path for low frequency atmospheric oscillations. *J. Atmos. Sci.,* submitted.

Livezey, R. E., and W. Y. Chen, 1983: Statistical field significance and its determination by Monte Carlo techniques. *Mon. Wea. Rev.,* **111,** 45–59.

Madden, R. A., 1976: Estimate of the natural variability of time-averaged sea-level pressure. *Mon. Wea. Rev.,* **104,** 942–952.

Rasmusson, E. M., P. A. Arkin, W. Y. Chen and J. B. Jalickee, 1981: Biennial variations in surface temperature over the United States as revealed by singular decomposition. *Mon. Wea. Rev.,* **109,** 587–598.