Statistical Aspects of Estimated Principal Vectors (EOFs) Based on Small Sample Sizes

HANS VON STORCH AND GERHARD HANNOSCHÖCK

Meteorologisches Institut der Universität Hamburg, 2 Hamburg 13, FRG

(Manuscript received 24 April 1984, in final form 18 February 1985)

ABSTRACT

Statistical properties of estimated nonisotropic principal vectors [empirical orthogonal functions (EOFs)] are reviewed and discussed. The standard eigenvalue estimator is nonnormally distributed and biased: the largest one becomes overestimated, the smallest ones underestimated. Generally, the variance of the eigenvalue estimate is large. The standard eigenvalue estimator may be used to define an unbiased estimator, which, however, exhibits an increased variance. If a fixed set of EOFs is used, the EOF coefficients are not stochastically independent. The variances of the low-indexed coefficients become considerably overestimated by the respective estimated eigenvalues, those of the high-indexed coefficients underestimated. If the ratio of degrees of freedom to sample size is one-half or even less, these disadvantages are still current as is demonstrated by an example.

1. Introduction

Empirical orthogonal functions (EOFs) are widely used in atmospheric and oceanic research following Lorenz (1956) as an aid for describing climate (e.g., Kutzbach, 1967; Fechner, 1975; Barnett, 1977; Heddinghaus and Kung, 1980; Weare, 1982; Storch, 1984), and for comparing simulations of general circulation models with climate (e.g., Savijärvi, 1978; Storch, 1982; Storch and Roeckner, 1983a and b; Preisendorfer and Barnett, 1983). They are applied to construct response patterns (Hasselmann, 1979; Hannoschöck, 1984) or a spectral model (Rinne and Karhila, 1975) and are also applied as an aid for developing regression forecast techniques (Peagle and Haslam, 1982a and b).

Their algebraic and geometric properties are reported in detail by ECMWF (1977) and their ability to represent (meteorological) data is discussed by Savijärvi (1978).

Preisendorfer *et al.* (1981) give a variety of procedures to test random eigenvalues against the hypothesis of an isotropic spectrum. Gray (1981) and North *et al.* (1982) showed that in the case of small sample sizes the structure of the estimated eigenvectors becomes unstable. Kendall (1980) and Jackson (1981) report some asymptotic stochastic properties. James (1960) published a formula of the joint probability distribution of estimated eigenvalues. However, this formula seems not to be applicable in practice. Lawley (1956) derived a number of valuable asymptotic formulas for the bias and the variance of the eigenvalue estimator. Because Lawley's paper seems to be widely unknown in atmospheric research, we repeat these formulas in Sections 4 and 5. Apparently the properties of the variances and covariances of the expansion coefficients of an *a priori* fixed set of EOFs and their estimates have not been discussed so far.

The concept of EOFs is identical to the concept of principal vectors (Kendall, 1980), which is used in social and economic sciences, especially. The difference in application is that in geophysical research we deal with a larger number of degrees of freedom, whereas the effective sample size generally is small.

In the following we use the expressions "principal vectors" (PVs) for the eigenvectors of the true 2nd-moment matrix and "empirical orthogonal functions" (EOFs) for those obtained from an estimated 2nd-moment matrix. Thus, EOFs are estimated PVs.

Sometimes the view has been expressed that it should be possible to connect PVs with physical processes or modes, i.e., to identify the patterns of the PVs with spatial characteristics of the system under study. To our knowledge, such attempts have seldom been successful. This is due to the fact that PVs are constructed mathematically with the constraint of maximum convergence and orthogonality (see Section 2). If it is possible to connect PVs with physical processes or modes, these have to be orthogonal to each other, too. However, the involved processes or modes generally do not fulfil this condition. Sometimes, the first PV may be interpreted since it is not bound to orthogonality. An example in which more than the first EOF could be assigned to physical modes is given by Fraedrich and Dümmel (1983), who found that the first EOF of the vertical tropo-

A first version of this paper was presented at the II. International Meeting on Statistical Climatology, Lisboa, Portugal, 26–30 September 1983.

spheric structure may be identified with the equivalent barotropic mode and the second with the baroclinic mode. A number of physical systems whose eigenfunctions coincide with its EOFs are given by North (1984). These systems are, however, of minor importance for atmospheric and oceanic problems.

Another difficulty is that the true PVs are not available but only their estimates. It is well-known (e.g., Gray, 1981; North *et al.*, 1982), and it will be emphasized in Section 6, that one has to expect these estimates to be highly dependent on the actual sample for small and moderate sample sizes. Thus attempts to connect EOFs and processes or modes often are condemned to lead to misinterpretations.

One likely approach to overcome the latter difficulty is the use of the already mentioned procedures to test to what extent the (true) eigenvalues are "significant", and to discuss physically those EOFs belonging to eigenvalues which have turned out to be "significant" (e.g., Overland and Pease, 1982). These procedures (e.g., Overland and Preisendorfer, 1982) are based on the *a priori* idea that the large eigenvalues are simple and the small ones equal to each other. The probability distributions necessary for the performance of a statistical test are generated by Monte Carlo simulations using a white eigenvalue spectrum, i.e., an eigenvalue with multiplicity equal to the dimension involved in the problem. Thus, the rejection of the null hypothesis associated with these Monte Carlo simulations is the statistical proof that the multiplicity of the first eigenvalue is less than the problem's dimension. This is the case if the first eigenvalue is simple, but also if none but the smallest eigenvalue is simple,[1] or if the largest eigenvalue is double. Further, the rejection of this null hypothesis gives no information as to whether the eigenvalue spectrum and, much less, the PVs are estimated sufficiently well.

In view of these principal difficulties, we recommend renouncing a physical interpretation and using EOFs only for the purpose of effective data condensation as long as the data base is not large.

The objective of this paper is to collect the most important properties of PVs (Section 2), to define their estimator (Section 3), and to show that transferring the properties of the PVs to the EOFs is questionable, if the ratio of sample size to degrees of freedom is not large (Sections 4 and 6). In Section 5, two methods of constructing unbiased eigenvalue estimators are presented and discussed in terms of a Monte Carlo example.

[1] The multiplicity of an eigenvalue is the number of linearly independent eigenvectors associated with it. An eigenvalue is called simple and double, if its multiplicity is 1 and 2, respectively. Only eigenvectors belonging to simple eigenvalues are uniquely determined (besides of a multiplicative constant).

2. Principal vectors

Let $\mathbf{x} \in \mathbf{R}^n$ be a random vector. An ordered orthonormal basis of vectors $\mathbf{y}_1, \ldots, \mathbf{y}_n$ is called a system of principal vectors (PVs), if for every k with $1 \leqslant k \leqslant n$

$$E(\|\mathbf{x} - \sum_{i=1}^{k} (\mathbf{x}, \mathbf{y}_i)\mathbf{y}_i\|^2) = \min \tag{2.1}$$

is valid. In (2.1) E denotes expectation, $\|\cdot\|$ the least square norm. The linear form $a_i(\mathbf{x}) \equiv (\mathbf{x} \cdot \mathbf{y}_i)$ is called the ith generalized Fourier coefficient or principal component (of $\mathbf{x}$ on $\mathbf{y}_i$).[2]

Let $\mathbf{X} \equiv E(\mathbf{x}\mathbf{x}')$ be the matrix of the second moments of $\mathbf{x}$. Assume the order of the eigenvalues r_i of $\mathbf{X}$ as follows:

$$r_1 \geqslant r_2 \geqslant \cdots r_n.$$

(Note that $\mathbf{X}$ is symmetric and positive semidefinite.) Then the following statements are valid:

- A system of normalized eigenvectors $\mathbf{y}_i$ of $\mathbf{X}$ ordered like their corresponding eigenvalues forms a PV-system of the random variable $\mathbf{x}$.

$$E(\|\mathbf{x} - \sum_{i=1}^{k} a_i(\mathbf{x})\mathbf{y}_i\|^2) = E(\|\mathbf{x}\|^2) - \sum_{i=1}^{k} r_i = \sum_{i=k+1}^{n} r_i. \tag{2.2}$$

- Provided $E(\mathbf{x}) = 0$, the generalized Fourier coefficients have zero mean and are uncorrelated, i.e.

$$E[a_i(\mathbf{x})] = 0 \quad \text{and} \quad \operatorname{cov}[a_i(\mathbf{x}), a_j(\mathbf{x})] = r_i\delta_{ij}. \tag{2.3}$$

These properties are well-known (cf. Kendall, 1980; Jackson, 1981). A straightforward and, as far as we know, new proof is given in the Appendix.

Assume $E(\mathbf{x}) = 0$. Then Eq. (2.3) means because of $E(\|\mathbf{x}\|^2) = \operatorname{var}(\mathbf{x})$ that

$$\operatorname{var}(\mathbf{x}) = \sum_{i=1}^{n} \operatorname{var}[a_i(\mathbf{x})] = \sum_{i=1}^{n} r_i. \tag{2.4}$$

As the eigenvalues are nonnegative, the Fourier coefficient of the ith PV describes on an average $(100 * r_i)/\operatorname{var}(\mathbf{x})\%$ of the total variance. Equation (2.3) states that the expectation (first moment) of the Fourier coefficients vanishes and that their covariance matrix (second moment) is a diagonal one with the eigenvalues as diagonal elements. For this reason the first moments often are estimated beforehand and subtracted from the data.

The above used orthonormality is a question of geometry, e.g. scaling. With respect to another geometry given by the dot product $(\mathbf{x} \cdot \mathbf{y}) \equiv \mathbf{x}'N\mathbf{y}$ with

[2] The expression $a \equiv b$ means that the quantity a is defined by b. $(\mathbf{a} \cdot \mathbf{b})$ denotes the dot product of the vectors $\mathbf{a}$ and $\mathbf{b}$.

a symmetric and positive definite matrix **N**, the PVs are given by the eigenvalue equation

$$(\mathbf{X}N - rI)y = 0. \tag{2.5}$$

3. Estimation

In practice, the distribution of **x** is unknown. Therefore, one has to estimate the PVs and their eigenvalues by the eigenvectors and -values of a second-moment matrix estimated by means of, say, m samples $\mathbf{x}_1, \ldots, \mathbf{x}_m$ of the random variable **x**. These estimated PVs are called empirical orthogonal functions (EOFs). In the following, the EOFs are labelled as $\hat{\mathbf{y}}$ and the estimated eigenvalues as $\hat{r}$. Thus a caret refers to the "standard" estimator.

If the number n of components of the random vector **x** is large, it may be expensive or even practically impossible to compute the eigenvalues and eigenvectors of the sample 2nd moment matrix. But if the number m of samples is less than n, the numerical problem may be reduced by a simple algebraic trick, see Storch and Hannoschöck (1984).

The number of resulting EOFs is $\min(n, m)$. If the sample mean is subtracted, the number reduces to $\min(n, m - 1)$.

4. The estimator $\hat{r}$

The question arises as to whether the eigenvalues $\hat{r}$ of the sample second-moment matrix are satisfactory estimators of the eigenvalues r of the true second-moment matrix. If the number of samples is less than the number n of degrees of freedom, at least $n - m$ eigenvalues of the sample second-moment matrix vanish (see Storch and Hannoschöck, 1984).

Lawley (1956) derived approximate formulas for the bias and the variance of the standard eigenvalue estimator:

Theorem: Let **x** be a n-variate normally distributed random variable with $E(\mathbf{x}) = 0$ and a covariance matrix **X** whose eigenvalues are all different:

$$E(\hat{r}_i) = r_i + \frac{r_i}{m} \sum_{j=1}^{n}{}' \frac{r_j}{r_i - r_j} + \mathrm{O}\left(\frac{1}{m^2}\right) \tag{4.1}$$

$$\mathrm{var}(\hat{r}_i) = \frac{2r_i^2}{m}\left[1 - \frac{1}{m}\sum_{j=1}^{n}{}' \left(\frac{r_j}{r_i - r_j}\right)^2\right] + \mathrm{O}\left(\frac{1}{m^3}\right). \tag{4.2}$$

Here m is the number of samples available to calculate the sample covariance matrix.

In order to check the quality of formulas (4.1) and (4.2) a series of Monte Carlo simulations was done with a number of seven-dimensional examples. Without loss of generality it is assumed that the expectation vector vanishes and that the covariance matrix is diagonal:

$$E(\mathbf{x}) = 0, \quad \mathbf{X} = \mathrm{diag}(r_1, \cdots, r_7). \tag{4.3}$$

We used examples in which the relative magnitudes of the eigenvalues are comparable to what is occurring in practice. Since the results were essentially unchanged when different eigenvalue spectra $r_1, \ldots, r_7$ were used, we report only the detailed results obtained for one randomly selected example, namely

$$r_1 = 5.0, \quad r_2 = 2.5, \quad r_3 = 1.3, \quad r_4 = 0.6,$$
$$r_5 = 0.3, \quad r_6 = 0.2, \quad r_7 = 0.1 \tag{4.4}$$

and summarize results from other examples.

In Table 1, the sample means (obtained by 500 independent Monte Carlo trials) and the approximations (4.1) of the expectations of the standard eigenvalue estimators are listed.

The low-indexed sample eigenvalues are overestimated and the high-indexed underestimated according to the Monte Carlo study:

$$\left.\begin{aligned} E(\hat{r}_i) > r_i, &\quad \text{for small } i \\ E(\hat{r}_i) < r_i, &\quad \text{for large } i \end{aligned}\right\}. \tag{4.5}$$

The first eigenvalue is overestimated and the remaining six underestimated. For the largest sample size considered in Table 1, $m = 128$, the bias of the first eigenvalue is still 2%. That $\hat{r}_1 > r_1$ holds for the largest eigenvalue is reasonable: r_1 is the largest amount of the total second moment of the continuous probability space $\{\mathbf{x}\}$ expressible by just one vector and $\hat{r}_1$ the respective number of the finite $\{\mathbf{x}_1, \cdots \mathbf{x}_m\}$. Since the latter set is smaller than the former set, a bigger portion of its total second moment may be represented by just one vector, i.e. $\hat{r}_1 > r_1$.

If a flatter spectrum is considered, the biases are generally smaller and the first few eigenvalue estimators are positively biased. This behavior is not surprising. The trace of the second-moment matrix, i.e. the sum of all eigenvalues is an unbiased estimate. Thus, the positive bias of the first eigenvalue has to be balanced by negative ones of eigenvalue–estimators with higher indices. But these negative biases are bounded by the absolute value of the true eigenvalues, because the estimators are positive. Thus, if the high-indexed eigenvalues are small, one needs more of them to balance the positive biases of the large eigenvalue(s).

The approximate values (4.1) reproduce the relation (4.5) very well. For large sample sizes m, the figures coincide well with the Monte Carlo ones. For small sample sizes, however, the accuracy of (4.1) is somewhat worse. For $m = 8$ and the smallest eigenvalues, (4.1) yields negative values, which makes no sense because of the positive definiteness of the sample covariance matrix.

TABLE 1. Expectation of the standard eigenvalue estimator $\hat{r}$ and of the modified, unbiased estimators (5.2) and (5.3). Example (4.4). Monte Carlo or asymptotic estimates of the true eigenvalues listed.

Sample size m	Estimator†	MC/af‡	True eigenvalue 5.0	2.5	1.3	0.6	0.3	0.2	0.1
4	S	MC	6.84*	2.21*	0.71*	0.18*	0.0	0.0	0.0
	S	af (4.1)	7.02	2.29	0.65	0.21	0.07	−0.14	−0.10
	J	MC	5.37*	2.59*	1.24*	0.73*	0.0	0.0	0.0
	L	MC	5.78*	2.72*	1.10*	0.34*	0.0	0.0	0.0
8	S	MC	6.02*	2.40*	0.97*	0.38*	0.16*	0.06*	0.01*
	S	af (4.1)	6.01	2.40	0.97	0.40	0.19	0.03	0.00
	J	MC	5.09*	2.56*	1.26*	0.57*	0.32*	0.16*	0.05*
	L	MC	5.30*	2.60*	1.21*	0.52*	0.25*	0.10*	0.02*
16	S	MC	5.61*	2.45*	1.14*	0.50*	0.23*	0.12*	0.05*
	S	af (4.1)	5.50	2.45	1.14	0.50	0.24	0.11	0.05
	J	MC	5.07*	2.53*	1.30*	0.62*	0.30*	0.19*	0.09
	L	MC	5.19*	2.53*	1.27*	0.59*	0.29*	0.16*	0.07*
32	S	MC	5.20*	2.45*	1.22	0.55*	0.27*	0.16*	0.08*
	S	af (4.1)	5.25	2.47	1.21	0.55	0.27	0.16	0.08
	J	MC	4.93*	2.47*	1.33*	0.61*	0.30*	0.20*	0.10*
	L	MC	4.97*	2.48*	1.30	0.60*	0.30*	0.19	0.09*
64	S	MC	5.10	2.49	1.26	0.57*	0.29	0.18	0.09
	S	af (4.1)	5.13	2.49	1.26	0.58	0.29	0.18	0.09
	J	MC	4.96	2.51	1.30	0.60*	0.31*	0.20	0.10*
	L	MC	4.98	2.51*	1.29	0.60*	0.31	0.20	0.10
128	S	MC	5.09	2.49*	1.28	0.58*	0.29	0.19	0.09
	S	af (4.1)	5.06	2.49	1.28	0.59	0.29	0.19	0.09
	J	MC	5.03	2.50*	1.30	0.60*	0.30	0.20	0.10
	L	MC	5.03	2.50*	1.30	0.60*	0.30	0.20	0.10

† Either standard $\hat{r}$ (S) or Jackknife correction (J) (5.2) with $k = 4$ or Lawley correction (5.3) (L).

‡ af: asymptotic formula (4.1); MC: Monte Carlo estimate (based on 500 trials). The Monte Carlo estimates were tested by a Lilliefors-test as to whether the distributions are distributed normally. Significantly (95% level) nonnormal estimates are marked by an asterisk.

The physical significance of the bias (4.5) of the eigenvalue estimators is that the importance of the dominating (generally large-scale) structures is overestimated and that of the less weighty (generally small-scale) ones underestimated.

Furthermore, the distributions of the eigenvalue estimator are tested for normality with the Lilliefors-test (e.g., Conover, 1971). An asterisk in Table 1 indicates rejection of the null hypothesis at the 95% level of significance that the estimator is normally distributed. Note that the lack of an asterisk does not mean that the distribution is in fact normal. It is only an indication that the probability to decide erroneously "accept the alternative" is greater than 5%. Up to $m = 32$, most of the distributions of the estimators are significantly nonnormal.

In Table 2, Monte Carlo estimates of the standard deviations and approximates (4.2) of the standard eigenvalue estimator $\hat{r}$ are listed for (4.4). Especially for small sample sizes m, the Monte Carlo estimated standard deviations are quite large for the low-indexed eigenvalues, pointing at the strong variability of these estimators and of the corresponding EOFs. For small sample sizes, the approximations (4.2) do not coincide satisfactorily with the Monte Carlo numbers. Often, (4.2) yields negative variances.

5. Unbiased eigenvalue estimators

There is a possibility of using the biased standard eigenvalue estimator $\hat{r}$ given in Section 3 to define unbiased estimators. However, it has to be kept in mind that it is general experience in statistics that reduction of bias has to be paid for by increase of variance.

The first approach, known as "Jackknife", is based on an idea of Tukey (cf. Miller, 1968):

Theorem: If A is a parameter and $\hat{A}_m$ an estimator based on m samples $a_1, \ldots, a_m$ with

$$\text{bias}(\hat{A}_m) = E(\hat{A}_m - A) = \frac{c}{m} A \qquad (5.1)$$

with some constant c which may depend on A but not on m. Then the new estimator

$$\tilde{A}_m \equiv k\hat{A}_m + (k-1)\bar{A}_{m-p} \qquad (5.2)$$

is unbiased, if $m = kp$ and $\bar{A}_{m-p}$ defined as follows: Subdivide the total sample of m elements into k subgroups of length p. Delete for each j, $j = 1, \ldots,$

TABLE 2. Standard deviations of the standard eigenvalue estimator $\hat{r}$ and of the modified, unbiased estimators (5.2) and (5.3). Example (4.4). Monte Carlo or asymptotic estimates of the standard deviations of the respective eigenvalue estimator belonging to the spectrum listed.

Sample size m	Estimator*	MC/af**	Standard deviation of estimated spectrum						
			5.0	2.5	1.3	0.6	0.3	0.2	0.1
4	S	MC	3.62	1.18	0.43	0.14	0.0	0.0	0.0
	S	af (4.2)	2.99	—	—	—	—	—	—
	J	MC	3.74	1.98	1.06	0.53	0.0	0.0	0.0
	L	MC	5.64	2.26	0.81	0.23	0.0	0.0	0.0
	L	af (5.4)	4.19	1.69	0.65	0.25	0.0	0.0	0.0
8	S	MC	2.27	0.94	0.36	0.15	0.07	0.03	0.01
	S	af (4.2)	2.31	0.73	0.23	0.05	—	—	—
	J	MC	2.37	1.50	0.70	0.33	0.20	0.11	0.05
	L	MC	2.32	1.53	0.65	0.27	0.12	0.05	0.01
	L	af (5.4)	2.74	1.22	0.56	0.25	0.12	0.04	0.01
16	S	MC	1.76	0.74	0.34	0.15	0.07	0.04	0.02
	S	af (4.2)	1.70	0.72	0.34	0.15	0.05	0.01	0.02
	J	MC	1.89	1.03	0.56	0.26	0.13	0.09	0.05
	L	MC	2.12	1.01	0.49	0.22	0.11	0.06	0.02
	L	af (5.4)	1.85	0.87	0.43	0.19	0.10	0.05	0.03
32	S	MC	1.22	0.55	0.28	0.13	0.06	0.03	0.02
	S	af (4.2)	1.23	0.57	0.29	0.13	0.06	0.04	0.02
	J	MC	1.29	0.69	0.41	0.19	0.09	0.06	0.03
	L	MC	1.36	0.67	0.35	0.16	0.08	0.05	0.02
	L	af (5.4)	1.28	0.62	0.31	0.14	0.07	0.04	0.02
64	S	MC	0.89	0.43	0.22	0.10	0.05	0.03	0.02
	S	af (4.2)	0.88	0.42	0.22	0.10	0.05	0.03	0.02
	J	MC	0.91	0.49	0.26	0.12	0.06	0.05	0.02
	L	MC	0.92	0.47	0.24	0.11	0.06	0.04	0.02
	L	af (5.4)	0.89	0.44	0.23	0.10	0.05	0.03	0.02
128	S	MC	0.64	0.31	0.14	0.07	0.04	0.02	0.01
	S	af (4.2)	0.62	0.31	0.16	0.07	0.04	0.02	0.01
	J	MC	0.65	0.32	0.16	0.08	0.04	0.03	0.01
	L	MC	0.64	0.32	0.17	0.08	0.04	0.03	0.01
	L	af (5.4)	0.63	0.31	0.16	0.07	0.04	0.02	0.01

* Either standard $\hat{r}$ (S) or Jackknife correction (J) (5.2) with $k = 4$ or Lawley correction (5.3) (L).
** af: Asymptotic formulas (4.2) or (5.4); — = negative values; MC: Monte Carlo estimate (based on 500 trials).

k, the corresponding subgroup of the total data sample and estimate with the remaining $m - p$ data $\hat{A}_{m-p}$. The mean of these k numbers $\hat{A}_{m-p}$ is defined to be $\bar{A}_{m-p}$.

Equation (5.2) is obtained by some straightforward analysis solving the problem $E(\alpha\hat{A}_m + \beta\bar{A}_{m-p}) = A$ with the constraint $\alpha + \beta = 1$.

To obtain a Jackknife corrected eigenvalue estimate, totally $k + 1$ eigenvalue problems have to be solved; k to get $\bar{A}_{m-p}$ and 1 for $\hat{A}_m$.

The standard eigenvalue estimator $\hat{r}$ exhibits according to (4.1) just the form (5.1) besides an $O(1/m^2)$-term. Thus, the Jackknife should be suitable for a correction of $\hat{r}$. In order to check the Jackknife's efficiency, its performance is tested by means of two Monte Carlo experiments of 500 samples each applied to example (4.4). Once, the sample number m is varied with a fixed subgrouping $k = 4$. The results are given in Tables 1 and 2.

For all sample sizes, the bias is in part considerably reduced compared to that of the standard estimator. For $m = 4$, $\hat{r}_1$ overestimates the first eigenvalue by 37%, while the error of the Jackknife estimate is only 7%, on an average. For $m = 16$, the Jackknife yields numbers as good as the standard estimator using $m = 128$ samples. Again, the distributions are for the most part significantly nonnormal for $m \leqslant 32$.

The variance of the Jackknife (5.2) estimator is increased for all indices and sample sizes, which was to be expected in view of the reduced bias. The deterioration is more serious for small eigenvalues than for large eigenvalues, e.g., for $m = 32$, the standard deviation of (5.2) is increased by 6% for the largest eigenvalue $r_1 = 5.0$, but by 54% for $r_5 = 0.3$. Especially for small sample sizes m, e.g. $m = 4$, the substantial decrease of the bias of the largest eigenvalue is accompanied by only a small increase of the variance.

The outcome of a second series of Monte Carlo experiments using a fixed m and a varying number of subgroups k is a relative insensitivity to the selection of k.

Another proposal was given by Lawley (1956). The estimator

$$\hat{r}_i \equiv \hat{r}_i\left[1 - \frac{1}{m}\sum_{j=1}^{n}{}' \frac{\hat{r}_j}{\hat{r}_i - \hat{r}_j}\right] \tag{5.3}$$

is unbiased, but its variance is according to the following asymptotic formula increased:

$$\mathrm{var}(\hat{r}_i) = \frac{2r_i^2}{m}\left[1 + \frac{1}{m}\sum_{j=1}^{n}{}' \left(\frac{r_j}{r_i - r_j}\right)^2\right] + O\left(\frac{1}{m^3}\right). \tag{5.4}$$

Again, we studied the efficiency of (5.3) and the exactness of (5.4) with a 500-trial Monte Carlo experiment using examples (4.4). The results are included in Tables 1 and 2, too.

The success of Lawley's correction (5.3) is slightly worse than that of the Jackknife (5.2). For $m = 4$, $r_1 = 5.0$ is overestimated by 16% through (5.3) compared to 7% of the Jackknife. To obtain estimates as good as the standard $\hat{r}$ with $m = 128$, Lawley's formula needs about $m = 32$, whereas $m = 16$ are sufficient for the Jackknife.

With respect to the variance of (5.3), it may be concluded from Table 2 that formula (5.4) is not very precise for $m \leq 16$, say. According to the Monte Carlo estimates, the variance is increased compared to the standard estimator. However, the increase is greater than that of the Jackknife (5.2) for the largest, or the two largest, eigenvalues and less for the minor ones.

The combined effect of bias and variance of an estimator may be expressed in terms of the root mean square (rms) error

$$\{E[(\hat{r}_i - r_i)^2]\}^{1/2} = [\mathrm{bias}(\hat{r}_i)^2 + \mathrm{var}(\hat{r}_i)]^{1/2}. \tag{5.5}$$

This quantity is listed for the three estimators for example (4.4) with varying sample sizes in Table 3. Apparently, Lawley's formula yields an uncertainty larger than do the other two estimators, standard and Jackknife, for at least small sample sizes m and large eigenvalues. Lawley's estimate for r_1 and $m = 4$ is very poor; the rms error is 5.69, i.e. greater than the positive number $r_1 = 5$ to be estimated. Since the expectation is 5.78 (Table 1) and the standard deviation 5.64 (Table 2), the estimator (5.3) varies between 0.14 and 11.42. Apparently, this estimator is useless. On the other hand, the standard estimator and the Jackknife yield the intervals [3.22; 10.46] and [1.63; 9.11], respectively. The Jackknife is superior to the standard estimator for the largest eigenvalue, r_1, and $m \leq 8$, and inferior for the six minor eigenvalues. For large sample sizes, $m \geq 64$, say, the differences between the three methods become negligible.

Thus, we may conclude that both correction methods are appropriate methods to reduce the bias of the standard eigenvalue estimator. But this advantageous

TABLE 3. Expectation of the root mean square error (5.5) of the standard eigenvalue estimator $\hat{r}$ and of the modified, unbiased estimators (5.2) and (5.3). Example (4.4). Monte Carlo estimate of the root mean square error of the respective eigenvalue estimator belonging to the spectrum listed.

Sample size m	Estimator*	rms error of estimated spectrum						
		5.0	2.5	1.3	0.6	0.3	0.2	0.1
4	S	4.06	1.22	0.72	0.45	0.3	0.2	0.1
	J	3.76	1.98	1.06	0.57	0.3	0.2	0.1
	L	5.69	2.27	0.83	0.37	0.3	0.2	0.1
8	S	2.49	0.94	0.49	0.26	0.14	0.14	0.1
	J	2.37	1.50	0.70	0.33	0.2	0.1	0.1
	L	3.33	1.53	0.66	0.28	0.24	0.2	0.1
16	S	1.86	0.74	0.37	0.17	0.1	0.1	0.0
	J	1.89	1.03	0.56	0.26	0.14	0.1	0.0
	L	2.12	1.01	0.49	0.22	0.1	0.1	0.0
32	S	1.42	0.55	0.28	0.14	0.0	0.0	0.0
	J	1.29	0.69	0.41	0.2	0.1	0.0	0.0
	L	1.36	0.67	0.35	0.17	0.1	0.0	0.0
64	S	0.89	0.44	0.22	0.1	0.0	0.0	0.0
	J	0.91	0.49	0.26	0.1	0.0	0.0	0.0
	L	0.92	0.47	0.24	0.1	0.0	0.0	0.0
128	S	0.64	0.31	0.14	0.1	0.0	0.0	0.0
	J	0.64	0.31	0.17	0.1	0.0	0.0	0.0
	L	0.64	0.31	0.17	0.1	0.0	0.0	0.0

* Either standard $\hat{r}$ (S) or Jackknife correction (J) (5.2) with $k = 4$ or Lawley correction (5.3) (L).

aspect goes along with an increase of the variance. With respect to the largest eigenvalues, the Jackknife technique seems to be superior to the correction proposed by Lawley and the standard estimate.

6. Properties of estimated principal components of x on y

One problem occurring in atmospheric and oceanic research is the comparison of different climates, e.g. of a modeled climate with an observed one or two modeled ones (say, differing by their forcing boundary conditions). This is treated by a χ^2-test (Storch, 1982; Storch and Roeckner, 1983a; Storch, 1984; Hannoschöck, 1984; Storch and Kruse, 1985; Frankignoul, 1985) or by a nonparametric two-sample test (Preisendorfer and Barnett, 1983; Storch and Roeckner, 1983b; Storch and Kruse, 1985). In order to get a tractable low-dimensional problem, the original data consisting of sometimes hundreds of components are *a priori* expanded into a short EOF series utilizing (2.1). Often, the applied EOFs are obtained from one of the two samples to be compared. Then, instead of the original dataset, the statistical test is applied to the low-dimensional vector of generalized Fourier-coefficients ($\hat{a}_i$). For this, reliable estimates of the variances and covariances of the generalized Fourier coefficients $\hat{a}_i$ of a once fixed set of EOFs $\hat{\mathbf{y}}_1, \ldots, \hat{\mathbf{y}}_m$ are needed. The question is how to estimate these second moments.

Provided $E(\mathbf{x}) = 0$, an obvious and asymptotically reasonable (Jackson, 1981) attempt is to use the properties (2.3) valid for the principal vectors with $\hat{r}$ instead of r, i.e. to estimate the covariances and variances by

$$\widehat{\mathrm{cov}[\hat{a}_i(\mathbf{x}), \hat{a}_j(\mathbf{x})]} \equiv \hat{r}_i \delta_{ij}. \tag{6.1}$$

In the case of two dimensions

$$\mathrm{var}(\hat{a}_1) < E(\hat{r}_1), \quad \mathrm{var}(\hat{a}_2) > E(\hat{r}_2)$$

is valid and one has to be aware of sometimes large covariances of $\hat{a}_1$ and $\hat{a}_2$: We may assume without loss of generality that $\mathbf{X} = \mathrm{diag}(r_1, r_2)$ with $r_1 > r_2$. Application of (2.1), (2.5) and (4.5) yields

$$\left.\begin{aligned} \mathrm{var}(\hat{a}_1) < r_1 < E(\hat{r}_1) \\ \mathrm{var}(\hat{a}_2) > r_2 > E(\hat{r}_2) \end{aligned}\right\}. \tag{6.2}$$

In order to get an idea of the magnitude of $\mathrm{var}(\hat{a}_i) - \hat{r}_i$, a two-dimensional example is examined with $r_1 = 2$, and $r_2 = 1$. For a series of numbers m, Monte Carlo experiments were performed as follows: m samples of the two dimensional random variable denoted by x are drawn by chance. These m samples are used to calculate the sample covariance matrix $\hat{\mathbf{X}}$. The eigenvectors of $\hat{\mathbf{X}}$ are taken as a randomly selected set of EOFs, $\{\hat{\mathbf{y}}_1, \hat{\mathbf{y}}_2\}$. The variance of the corresponding generalized Fourier coefficients $\hat{a}_i = (\mathbf{x} \cdot \hat{\mathbf{y}}_i)$ is

$$\mathrm{var}(\hat{a}_i) = E(\hat{\mathbf{y}}_i' \mathbf{x}\mathbf{x}' \hat{\mathbf{y}}_i) = \hat{\mathbf{y}}_i' \mathbf{X} \hat{\mathbf{y}}_i = \sum_i \hat{y}_{ij}^2 r_i \tag{6.3}$$

with the true eigenvalues r_i. To test the accuracy of (6.1), the difference $\mathrm{var}(\hat{a}_i) - \hat{r}_i$ is calculated, finally. The whole procedure is repeated 1000 times. Sample means $\overline{\mathrm{var}(\hat{a}_i) - \hat{r}_i}$ are computed and listed in Table 4: The overestimation of $\mathrm{var}(\hat{a}_1)$ by $\hat{r}_1$ is large. For e.g. $m = 8$, 8 independent data are available to estimate the two-dimensional covariance matrix. On an average, we have: $\hat{r}_1 = 2.26$ and $\mathrm{var}(\hat{a}_1) = 1.78$. If $\hat{r}_1$ is utilized as estimator of $\mathrm{var}(\hat{a}_1)$, one has to expect an overestimation of 27%. On the other hand, $\mathrm{var}(\hat{a})_2$ is underestimated by $\hat{r}_2$ on an average by 41%. A comparison of different samples based on these variance estimators were conservative[3] for coefficients of EOF1 and radical for those of EOF2.

In Table 4 the frequency (of 1000 samples) of $\mathrm{cov}(\hat{a}_1, \hat{a}_2) > 0.25$ is listed additionally. It turns out that the covariance of the Fourier coefficients based on EOFs is not negligible especially for small sample sizes m.

A closer look at the seven-dimensional example (4.4) shows that instead of (6.2) the relations

$$E[\mathrm{var}(\hat{a}_i)] < r_i < E(\hat{r}_i) \tag{6.4.1}$$

for small indices i, and

$$E(\mathrm{var}(\hat{a}_i)) > r_i > E(\hat{r}_i) \tag{6.4.2}$$

for large indices i, are valid. As an evidence, the expectations of $\mathrm{var}(\hat{a}_i) - \hat{r}_i$, $i = 1, \ldots, 7$, were estimated by 500 independent Monte Carlo trials. The results are given in Table 5. As in the preceding two dimensional example, the respective over- and underestimations of the variances are not negligible.

Another example of this kind is given in the appendix of Storch and Roeckner (1983a).

As can be deduced from formula (2.3) and property (2.1) of PVs, instead of (6.4.1–2) is valid:

$$\mathrm{var}(\hat{a}_i) < E(\hat{r}_1), \quad \mathrm{var}(\hat{a}_n) > E(\hat{r}_n). \tag{6.5}$$

Another attempt to estimate the 2nd moment matrix could be to use the data $\mathbf{x}_1, \ldots, \mathbf{x}_m$ which have been already used for computing the EOFs and for estimating the second-moment matrix conventionally, i.e. by the sample second moments. Unfortunately, this proceeding is not successful, because the analysis of Section 2 with finite sums instead of expectations yields that the resulting estimator is just the standard estimator.

The significance of (6.4.1–2) and (6.5) is that a test procedure as outlined in the beginning of this section will be conservative for low indexed (generally large scale) EOFs and radical for high indexed (generally small and medium scale) EOFs.

[3] A conservative (radical) statistical test rejects erroneously the null hypothesis too seldom (too often).

TABLE 4. Sample means of $\text{var}(\hat{a}_i) - \hat{r}_i$ and the frequency of $\text{cov}(\hat{a}_1, \hat{a}_2) > 0.25$ for different sample sizes m to estimate the EOFs ($i = 1, 2$).

m	$\text{cov}(\hat{a}_1, \hat{a}_2) > 0.25$	$\overline{\text{var}(\hat{a}_1) - \hat{r}_1}$	$\overline{\text{var}(\hat{a}_2) - \hat{r}_2}$
4	62.4%	−0.69	0.76
6	56.9%	−0.64	0.59
8	55.3%	−0.48	0.51
12	47.3%	−0.32	0.36
50	22.2%	−0.10	0.08
100	7.7%	−0.05	0.04

7. Conclusions

Statistical properties of estimated principal vectors and eigenvalues have been studied. The following conclusions may be drawn:

1) The sample eigenvalue $\hat{r}$ is a considerably biased estimator of the true eigenvalue r. For the largest r, the bias is positive, for the smallest negative: It is of the order of $1/m$, where m is the number of independent samples. The variances of $\hat{r}$ is the order of $1/m$, too.

2) By means of correction methods (5.2) and (5.3), unbiased eigenvalue estimators are constructed. However, the decrease of the bias is accompanied by an increase of the estimator's variance. For the largest eigenvalue, at least, the Jackknife yields favorable results.

3) Estimated second moments of generalized Fourier coefficients of a fixed set of EOFs: On average, for small i (large i) the sample eigenvalue $\hat{r}_i$ will overestimate (underestimate) the variance expressed by the corresponding EOF considerably. The covariances are generally not negligible. This means that the independence of PV coefficients cannot be transferred to EOF coefficients.

Note, however, that the aforementioned problems become negligible asymptotically.

APPENDIX

Proof of (2.2–4)

a) Let $\mathbf{y}_1, \ldots, \mathbf{y}_k$ be arbitrary orthonormal vectors:

$$R(k; \mathbf{x}) \equiv \|\mathbf{x} - \sum_{i=1}^{k} a_i(\mathbf{x})\mathbf{y}_i\|^2$$

$$= (\mathbf{x} - \sum_{i=1}^{k} a_i(\mathbf{x})\mathbf{y}_i, x - \sum_{i=1}^{k} a_i(\mathbf{x})\mathbf{y}_i)$$

$$= (\mathbf{x}\cdot\mathbf{x}) - 2(\sum_{i=1}^{k} a_i(\mathbf{x})\mathbf{x}'\mathbf{y}_i) + \sum_{i,j=1}^{k} a_i(\mathbf{x})a_j(\mathbf{x})\mathbf{y}_i'\mathbf{y}_j$$

$$= \|\mathbf{x}\|^2 - 2\sum_{i=1}^{k} (\mathbf{x}'\mathbf{y}_i)(\mathbf{x}'y_i) + \sum_{i=1}^{k} a_i(\mathbf{x})^2$$

$$= \|\mathbf{x}\|^2 - 2\sum_{i=1}^{k} \mathbf{y}_i'(\mathbf{x}\mathbf{x}')y_i + \sum_{i=1}^{k} (\mathbf{x}'\mathbf{y}_i)(\mathbf{x}'\mathbf{y}_i)$$

$$= \|\mathbf{x}\|^2 - \sum_{i=1}^{k} \mathbf{y}_i'(\mathbf{x}\mathbf{x}')\mathbf{y}_i.$$

Taking expectations gives as "expected deviation":

$$E[R(k; \mathbf{x})] = \begin{cases} E(\|\mathbf{x}\|^2) - \sum_{i=1}^{k} \mathbf{y}_i' E(\mathbf{x}\mathbf{x}')\mathbf{y}_i \\ E(\|\mathbf{x}\|^2) - \sum_{i=1}^{k} \mathbf{y}_i' \mathbf{X}\mathbf{y}_i. \end{cases} \quad (*)$$

Assume (2.2) and (2.3) are already proved for the first $k-1$ principal vectors $\mathbf{y}_1, \ldots, \mathbf{y}_{k-1}$ ($1 < k < n$). The kth PV $\mathbf{y}_k$ can be constructed by adding Lagrange multipliers to (*) according to the constraints of orthonormality

$$\mathbf{y}_k'\mathbf{y}_i = 0 \quad (i = 1, \cdots, k-1); \quad \|\mathbf{y}_k\|^2 = 1.$$

Setting the derivatives with respect to y zero yields

$$0 = \frac{d}{d\mathbf{y}_k} \mathrm{E}[\mathrm{R}(k, \mathbf{x})] + \sum_{i=1}^{k-1} q_i(\mathbf{y}_k, \mathbf{y}_i) + g(\|\mathbf{y}_k\|^2 - 1)$$

$$= \frac{d}{d\mathbf{y}_k}(-\sum_{i=1}^{k} \mathbf{y}_i'\mathbf{X}\mathbf{y}_i) + \sum_{i=1}^{k-1} q_i\mathbf{y}_i + 2g\mathbf{y}_k$$

$$= -2(\mathbf{X} - gI)\mathbf{y}_k + \sum_{i=1}^{k-1} q_i\mathbf{y}_i. \quad (**)$$

Scalar multiplication by $\mathbf{y}_j$, $j < k$, from the left gives

$$0 = 2\mathbf{y}_j'(\mathbf{X} - gI)\mathbf{y}_k + q_j = 2[(X - gI)\mathbf{y}_j]'\mathbf{y}_k + q_j$$

$$= 2(q_j - g)\mathbf{y}_j'\mathbf{y}_k + q_j = q_j.$$

Thus, (**) reduces to the eigenvalue equation

$$(\mathbf{X} - gI)\mathbf{y}_k = 0$$

and (*) turns to

$$E(R(k; \mathbf{x})) = E(\|\mathbf{x}\|^2) - \sum_{i=1}^{k-1} r_i - g$$

TABLE 5. Sample means of $\text{var}(\hat{a}_i) - \hat{r}_i$ for $i = 1, \ldots, 7$ for (4.4) obtained by a series of 500 Monte Carlo trials.

Sample size	True eigenvalue r_i						
m	5.0	2.5	1.3	0.6	0.3	0.2	0.1
4	−3.53	−0.10	0.64	0.67	0.76	0.42	0.72
8	−1.89	0.13	0.48	0.41	0.37	0.33	0.32
16	−1.13	0.04	0.31	0.24	0.17	0.16	0.14
32	−0.58	0.01	0.19	0.11	0.06	0.09	0.06
64	−0.32	0.03	0.09	0.06	0.03	0.05	0.03
128	−0.12	0.01	0.04	0.03	0.02	0.02	0.01

which is minimized by $g = r_k$, being the largest residual eigenvalue. Now, (2.2) and the first part of (2.3) is proved. It is

$$E(\|\mathbf{x}\|^2) = \text{trace}(\mathbf{X}) = \sum_{i=1}^{n} r_i$$

which yields the second statement of (2.3).

b) To establish (2.4), assume $E(\mathbf{x}) = 0$. Then

$$E(a_i(\mathbf{x})) = E(\mathbf{x})'\mathbf{y}_i = 0$$

and

$$\text{cov}(a_i, a_j) = E(\mathbf{y}_i'\mathbf{x}\mathbf{x}'\mathbf{y}_j) = \mathbf{y}_i'\mathbf{X}\mathbf{y}_j = r_j\mathbf{y}_i'\mathbf{y}_j = r_j\delta_{ij}$$

which had to be shown.

REFERENCES

Barnett, T. P., 1977: The principal time and space scales of the Pacific trade windfields. *J. Atmos. Sci.*, **34,** 221–236.

Conover, W. J., 1971: *Nonparametric Statistics.* Wiley & Sons, 462 pp.

ECMWF, 1977: Workshop on the use of empirical orthogonal functions in meteorology. European Center for Medium Range Weather Forecasts, Reading, Shinfield Park, England.

Fechner, H., 1975: Darstellung des Geopotentials der 500 mb Fläche der winterlichen Nordhalbkugel durch natürliche Orthogonalfunktionen, Ber. Inst. f. Meeresk. 5, Institut für Meereskunde, Düsternbrooker Str., 2300 Kiel, FRG.

Fraedrich, K., and T. Dümmel, 1983: On single station forecasting: The geopotential height, its vertical and time structure and 500 mbar ARMA prediction. *Beitr. Phys. Atmos.* **56,** 221–239.

Frankignoul, C., 1985: Multivariate analysis of sensitivity studies with atmospheric GCM's. *Proc. 16th Int. Liège Colloq. on Ocean Hydrodynamics,* Elsevier (in press).

Gray, M., 1981: On the stability of temperature eigenvector patterns. *J. Climatol.* **1,** 273–281.

Hannoschöck, G., 1984: A multivariate signal-to-noise analysis of the response of an atmospheric circulation model to sea surface temperature anomalies, Hamburger Geophysikalische Einzelschriften (Meteorologisches Institut der Universität Hamburg, Bundesstr. 55, 2000 Hamburg 13, FRG).

Hasselmann, K., 1979: On the signal-to-noise problem in atmospheric response studies. *Meteorology of Tropical Oceans,* Roy. Meteor Soc., 251–258.

Heddinghaus, T. R., and E. C. Kung, 1980: An analysis of climatological patterns of the Northern Hemispheric circulation. *Mon. Wea. Rev.*, **108,** 1–17.

Jackson, E., 1981: Principal components and factor analysis. Part II: Additional topics related to principal components. *J. Quality Technol.*, **13,** 46–58.

James, A. T., 1960: The distribution of the latent roots of the covariance matrix. *Ann. Math. Stat.*, **31,** 151–158.

Kendall, M., 1980: *Multivariate Analysis.* Charles Griffin, 210 pp.

Kutzbach, J. E., 1967: Empirical eigenvectors of sea level pressure, surface temperature and precipitation complexes over North America. *J. Appl. Meteor.*, **6,** 791–802.

Lawley, D. N., 1956: Tests of significance for the latent roots of covariance and correlation matrices. *Biometrika,* **43,** 128–136.

Lemke, P., 1980: Applications of the inverse modelling technique to arctic and antarctic sea ice anomalies. Hamb. Geoph. Einzelschr. 49.

Lorenz, E. N., 1956: Empirical orthogonal functions and statistical weather prediction, in: Final report statistical forecasting project.

Miller, R. G., 1968; Jackknifing variances. *Ann. Math. Stat.*, **39,** 568–582.

North, G. R., 1984: Empirical orthogonal functions and normal modes. *J. Atmos. Sci.*, **41,** 879–887.

——, T. L. Bell, R. F. Cahalan and F. J. Moeng, 1982: Sampling errors in the estimation of empirical orthogonal functions. *Mon. Wea. Rev.*, **110,** 699–706.

Overland, J. E., and C. H. Pease, 1982: Cyclone climatology of the Bering Sea and its relation to sea ice extent. *Mon. Wea. Rev.*, **110,** 5–13.

——, and R. W. Preisendorfer, 1982: A significance test for principal components applied to a cyclone climatology. *Mon. Wea. Rev.*, **110,** 1–4.

Peagle, J. N., and R. B. Haslam, 1982a: Empirical orthogonal function estimates of local predictability. *J. Appl. Meteor.*, **21,** 117–126.

——, and ——, 1982b: Statistical prediction of 500 mb height field using eigenvectors. *J. Appl. Meteor.*, **21,** 127–138.

Preisendorfer, R. W., F. W. Zwiers and T. P. Barnett, 1981: Foundations of principal component selection rules. SIO Ref. Ser., 81-4, Scripps Institution of Oceanography, 200 pp.

——, and T. P. Barnett, 1983: Numerical model-reality intercomparison test using small-sample statistics. *J. Atmos. Sci.*, **40,** 1884–1896.

Rinne, J., and V. Karhila, 1975: A spectral barotropic model in horizontal orthogonal functions. *Quart. J. Roy. Meteor. Soc.*, **101,** 365–382.

Savijärvi, H., 1978: Verification and storing with empirical orthogonal functions. ECMWF Internal Rep. 18, 36 pp.

Storch, H. v., 1982: Comparison of a sequence of model generated 500 mb-topographics with climate. *Tellus,* **34,** 89–91.

——, 1984: An accidental result: The mean 1983 January 500 mb field significantly different from its 1967–81 predecessors. *Beitr. Phys. Atmos.*, **57,** 440–444.

——, and E. Roeckner, 1983a: Verification of general circulation models applied to the Hamburg University GCM. Part I: Test of individual climate states. *Mon. Wea. Rev.*, **111,** 1965–1976.

——, and ——, 1983b: On the verification of January GCM simulations. *Proc. Second Int. Meeting on Statistical Climatology,* Lisboa, Instituto de Meteorogia e Geofisica, XXX pp.

——, and G. Hannoschöck, 1984: Comments on "Empirical orthogonal function-analysis of wind vectors over the tropical Pacific region". Bull. Amer. Meteor. Soc., **65,** 162.

——, and H. A. Kruse, 1985: The significant tropospheric midlatitudinal El Niño response patterns observed in January 1983 and simulated by a GCM. *Proc. 16th Int. Liège Colloq. on Ocean Hydrodynamics.* Elsevier (in press.)

Weare, B. C., 1982: El Niño and tropical Pacific Ocean surface temperature. *J. Phys. Oceanogr.* **12,** 17–27.