Spectral Approach to Optimal Estimation of the Global Average Temperature

SAMUEL S. P. SHEN

Mathematics Department, University of Alberta, Edmonton, Alberta, Canada

GERALD R. NORTH AND KWANG-Y. KIM

Climate System Research Program, Texas A&M University, College Station, Texas

(Manuscript received 22 September 1993, in final form 8 June 1994)

ABSTRACT

Making use of EOF analysis and statistical optimal averaging techniques, the problem of random sampling error in estimating the global average temperature by a network of surface stations has been investigated. The EOF representation makes it unnecessary to use simplified empirical models of the correlation structure of temperature anomalies. If an adjustable weight is assigned to each station according to the criterion of minimum mean-square error, a formula for this error can be derived that consists of a sum of contributions from successive EOF modes. The EOFs were calculated from both observed data and a noise-forced EBM for the problem of one-year and five-year averages. The mean square statistical sampling error depends on the spatial distribution of the stations, length of the averaging interval, and the choice of the weight for each station data stream. Examples used here include four symmetric configurations of 4×4, 6×4, 9×7, and 20×10 stations and the Angell–Korshover configuration. Comparisons with the 100-yr U.K. dataset show that correlations for the time series of the global temperature anomaly average between the full dataset and this study's sparse configurations are rather high. For example, the 63-station Angell–Korshover network with uniform weighting explains 92.7% of the total variance, whereas the same network with optimal weighting can lead to 97.8% explained total variance of the U.K. dataset.

1. Introduction

A number of investigations have recently focused on the problem of global-scale climate change. In particular, the issue of whether the earth's surface temperature has warmed in the last century is of interest because of the possibility of warming forced by increases in the so-called greenhouse gases (e.g., Folland et al. 1992). Empirical investigations from several laboratories throughout the world have indicated that the global average temperature may have increased by as much as 0.6°C over the last century (Hansen and Lebedeff 1987; Jones et al. 1986a; Jones et al. 1986b,c; Jones and Wigley 1990; Jones et al. 1991; Vinnikov et al. 1990; Folland et al. 1992; Karl et al. 1993). The importance of the outcome suggests further scrutiny of the methodology of such analyses. In particular, questions have been raised about the adequacy of various methods of removing some rather obvious biases from the record, such as those attributable to the systematically changing urban heat island effects or by changes in the routine practices of sampling the sea surface temperature by buckets or by engine intake water. Investigation of such contributions to the error budget requires a careful review of the instrument histories contributing to the record. Another potentially large error is that of statistical sampling, although it can be minimized by optimal averaging schemes (e.g., Kagan 1979). Yet the broad spectrum of spatial scales of temperature anomalies gives rise to the question of whether a relatively sparse network of arbitrarily distributed stations omits nonnegligible contributions to the variance at fine scales. The problem as such has been a concern for a long time. Recent studies include those by Trenberth and Olson (1992) and Madden et al. (1993). Both of them rely heavily on Monte Carlo techniques based upon GCM studies using perpetual month simulations.

The purpose of this paper, combining the use of an optimal averaging technique and EOF analysis, is to evaluate the statistical sampling error incurred in estimating the global temperature by a sparse network of a finite number of stations whose data streams are averaged through a time interval τ, for example, one or five years. In particular, we set up a minimal mean-square-error criterion, which is a function of the network configuration and the weight assigned to each station for a specified averaging time interval. Our present investigation is an outgrowth of an earlier study

Corresponding author address: Gerald R. North, Climate System Research Program, Texas A & M University, College Station, Texas 77843.
E-mail: North@csrp.tamu.edu

for comparing area-averaged rain rate designs (North and Nakamoto 1989).

Past work using statistical optimal averaging schemes has required an empirical, rather simplified, and sometimes inappropriate homogeneity or isotropy assumption on the correlation field, such as that given by Eq. (10) in Vinnikov et al. (1990). Our mean-square error minimization method takes a spectral approach and does not require explicit simplified models of the correlation field. By spectral approach, we mean that the correlation structure is expressed in terms of a sum of EOF basis functions. EOFs sometimes have a dynamical interpretation, but more directly they provide information about how variance is distributed with spatial scale even for nonhomogeneous fields. As will be shown in this paper, the mean-square sampling error can be expressed in terms of a sum of contributions from successive EOF modes [see Eq. (31)] and the error variance contributions due to temperature anomalies at different spatial scales become explicit. In addition, the spectral approach has some computational advantage. When considering different averaging time interval length τ—for example, one year, two years, or five years—the EOFs and their eigenvalues, which depend on τ, can be readily recalculated. But the empirical and analytic expressions of the correlation fields for different τ values, which are required by the classical optimal averaging scheme, are not easily parameterized. The EOFs, by definition, take into account the second moment (variance and covariance) properties of any nonhomogeneous field.

In section 2, we express the dependence of the mean-square error on the network configuration, the weights, and the correlation structure of the temperature anomaly. The scheme to obtain minimum mean-square error—that is, the scheme of optimal weighting—and the EOFs are introduced in section 3. Several examples are described in section 4, and conclusions are given in section 5.

2. Weighted estimator

We wish to measure the globally averaged temperature $\bar{T}_\tau(t)$ time averaged through the interval τ centered at t. The true value of this quantity is

$$\bar{T}_\tau(t) = \frac{1}{4\pi}\int_{4\pi} T_\tau(\hat{\mathbf{r}}, t)\, d\Omega, \tag{1}$$

where

$$T_\tau(\hat{\mathbf{r}}, t) = \frac{1}{\tau}\int_{t-\tau/2}^{t+\tau/2} T(\hat{\mathbf{r}}, t')\, dt' \tag{2}$$

is the τ-length averaging, $\hat{\mathbf{r}}$ is a unit vector pointing from the center of the earth to a point in question on the spherical surface of the earth, $T(\hat{\mathbf{r}}, t)$ is the temperature at $\hat{\mathbf{r}}$ and t, $d\Omega$ is the element of solid angle, and 4π is used to denote that the integral is to be taken over the entire sphere.

Departures of the τ-average from the ensemble average are due to natural variability. Such a departure is called an anomaly:

$$\Delta T_\tau(\hat{\mathbf{r}}, t) \equiv T_\tau(\hat{\mathbf{r}}, t) - \langle T_\tau(\hat{\mathbf{r}}, t)\rangle, \tag{3}$$

where $\langle T_\tau(\hat{\mathbf{r}}, t)\rangle$ denotes the ensemble average of $T_\tau(\hat{\mathbf{r}}, t)$. Similarly, we can define an anomaly of the global average temperature:

$$\Delta\bar{T}_\tau(t) \equiv \bar{T}_\tau(t) - \langle \bar{T}_\tau(t)\rangle. \tag{4}$$

By definition

$$\langle \Delta\bar{T}_\tau(t)\rangle = 0. \tag{5}$$

In what follows we will deal with the anomaly field and its global average. However, to keep the notation simple, we drop the prefix Δ.

The global average temperature anomaly may be estimated from the data streams collected from a given network of N_{net} stations $\{\hat{\mathbf{r}}_j, j = 1, 2, 3, \ldots, N_{\text{net}}\}$ by the estimator

$$\hat{\bar{T}}_\tau(t) \equiv \sum_{j=1}^{N_{\text{net}}} w_j T_\tau(\hat{\mathbf{r}}_j, t), \tag{6}$$

where w_j is the weight assigned to the jth station. If we assume that $\langle T_\tau(\hat{\mathbf{r}}, t)\rangle$ hardly changes in space, then the no-bias constraint on the weights is

$$\sum_{j=1}^{N_{\text{net}}} w_j = 1. \tag{7}$$

This constraint is not strictly necessary, since we are dealing with an anomaly field. However, we enforce it as a safety measure, since we do not have perfect knowledge of the ensemble average. Of course, if this condition is removed, the resulting mean-square error would be smaller. Hence, our estimate may be considered as the upper bound of the mean-square error.

We may write the estimator of the global average anomaly defined by (6) into an integral form

$$\hat{\bar{T}}_\tau(t) = \frac{1}{4\pi}\int_{4\pi} w_{\{\text{net}\}}(\hat{\mathbf{r}}) T_\tau(\hat{\mathbf{r}}, t)\, d\Omega, \tag{8}$$

where

$$w_{\{\text{net}\}}(\hat{\mathbf{r}}) \equiv 4\pi \sum_{j=1}^{N_{\text{net}}} w_j \delta(\hat{\mathbf{r}} - \hat{\mathbf{r}}_j). \tag{9}$$

The subscript $\{\text{net}\}$ is used to emphasize that the function is dependent on the configuration of stations and $\delta(\cdot)$ is the Dirac delta function.

Now we may form the mean-square error

$$\epsilon^2 = \langle (\bar{T}_\tau - \hat{\bar{T}}_\tau)^2\rangle = \left\langle \left(\frac{1}{4\pi}\int_{4\pi} d\Omega\, [1 - w_{\{\text{net}\}}(\hat{\mathbf{r}})]\, T_\tau(\hat{\mathbf{r}}, t)\right)^2 \right\rangle, \tag{10}$$

where ϵ^2 is a function of the weight vector $\mathbf{w} = \{w_j\}$. A natural question is what is the size of ϵ^2 and how does it vary for different choices of the station network $\{\hat{\mathbf{r}}_j\}$ and of the weight vector $\mathbf{w}$. In particular, one might wish to know the best choice of $\mathbf{w}$, in the sense of minimum ϵ^2, for a given configuration $\{\hat{\mathbf{r}}_j\}$, since the latter, representing the locations of historical stations, is often beyond the control of the investigator. Another meaningful problem is to find the network $\{\hat{\mathbf{r}}_j\}$ such that ϵ^2 is least for uniform $\mathbf{w} = \{w_i = 1/N_{\text{net}}\}$. Both problems have been studied for the case of a planet with rotationally invariant statistics (North et al. 1992; Hardin et al. 1992; Hardin and Upson 1994). The Angell–Korshover network of stations is an example of a configuration $\{\hat{\mathbf{r}}_i\}$ in which a set of stations was found by trial and error that in some practical sense minimized ϵ^2 (Angell and Korshover 1983; Angell 1992). In the present study we apply the spectral technique to the case where the statistics are not homogeneous but rather are the actual nonhomogeneous statistical characteristics of the anomaly of the real earth's surface air temperature field.

Expanding the formula for the mean-square error leads to

$$\epsilon^2 = \langle (\bar{T}_\tau - \hat{\bar{T}}_\tau)(\bar{T}_\tau - \hat{\bar{T}}_\tau) \rangle$$
$$= \int_{4\pi} d\Omega' \int_{4\pi} d\Omega'' \left[\frac{1}{(4\pi)^2} - \frac{2}{4\pi} \sum_{i=1}^{N_{\text{net}}} w_i \delta(\hat{\mathbf{r}}' - \hat{\mathbf{r}}_i) + \sum_{i,j=1}^{N_{\text{net}}} w_i w_j \delta(\hat{\mathbf{r}}' - \hat{\mathbf{r}}_i)\delta(\hat{\mathbf{r}}'' - \hat{\mathbf{r}}_j) \right] \rho_\tau(\hat{\mathbf{r}}', \hat{\mathbf{r}}''), \quad (11)$$

where we have introduced the temporally smoothed covariance

$$\rho_\tau(\hat{\mathbf{r}}', \hat{\mathbf{r}}'') \equiv \langle T_\tau(\hat{\mathbf{r}}', t) T_\tau(\hat{\mathbf{r}}'', t) \rangle. \quad (12)$$

In practice, we form the covariance matrix by cross product between the expansion coefficients of $T(\hat{\mathbf{r}}, t)$. Then, the eigenvectors (vectors of spherical harmonic expansion coefficients) of this covariance matrix are translated into the eigenfunctions in real space (see Kim and North 1993). All second moment information about the anomaly field averaged over the interval τ is contained in this temporally smoothed covariance function. The eigenfunctions of this covariance, when viewed as a kernel of a linear integral operator, are very useful in a number of estimation problems (see, e.g., North et al. 1995; North and Kim 1995; Kim and North 1992, 1993). The eigenfunctions of the covariance kernel are the EOFs for the smoothing interval τ (also called Karhunen–Loève functions) and their eigenvalues are the individual component contributions to the total variance of the anomaly field (see, e.g., North et al. 1982). These eigenfunctions and their eigenvalues in general depend on τ but we will not clutter the notation further with this index.

3. Minimizing the mean-square error

Next we take the variation of the mean-squared error with respect to the weights $\{w_i\}$, holding the station locations fixed, all subject to the no-bias constraint (7). This is accomplished with the method of Lagrange multipliers (e.g., Arfken 1985). We simply extremize

$$J[\mathbf{w}] = \epsilon^2[\mathbf{w}] - 2\Lambda \left[\sum_{j=1}^{N_{\text{net}}} w_j - 1 \right], \quad (13)$$

where 2Λ is a Lagrange multiplier and is always regarded as being a constant in the computational procedures following. This philosophy is the same as the method of optimal statistical averaging (e.g., Kagan 1979). Taking the partial derivatives of $J[\mathbf{w}]$ with respect to the weights and the Lagrange multiplier and setting them individually to zero,

$$\frac{\partial J}{\partial w_i} = 0, \quad i = 1, \ldots, N_{\text{net}}, \quad (14)$$

$$\frac{\partial J}{\partial \Lambda} = 0. \quad (15)$$

Inserting the expression for ϵ^2 results in

$$\sum_{j=1}^{N_{\text{net}}} w_j \rho_\tau(\hat{\mathbf{r}}_i, \hat{\mathbf{r}}_j) - \Lambda = \frac{1}{4\pi} \int_{4\pi} \rho_\tau(\hat{\mathbf{r}}, \hat{\mathbf{r}}_i)\, d\Omega, \quad i = 1, \ldots, N_{\text{net}}, \quad (16)$$

$$\sum_{i=1}^{N_{\text{net}}} w_i = 1. \quad (17)$$

The above system is for $N_{\text{net}} + 1$ unknowns, the N_{net} weights $\{w_i\}$, and the Lagrange multiplier Λ.

To proceed we introduce the EOFs. These are the eigenfunctions of the τ-averaged covariance kernel,

$$\rho_\tau(\hat{\mathbf{r}}, \hat{\mathbf{r}}') = \langle \bar{T}_\tau(\hat{\mathbf{r}}, t)\, \bar{T}_\tau(\hat{\mathbf{r}}', t) \rangle, \quad (18)$$

and the eigenfunctions $\psi_n(\hat{\mathbf{r}})$ are defined by

$$\int_{4\pi} \rho_\tau(\hat{\mathbf{r}}, \hat{\mathbf{r}}') \psi_n(\hat{\mathbf{r}}')\, d\Omega' = \lambda_n \psi_n(\hat{\mathbf{r}}). \quad (19)$$

These EOFs have the orthogonality and completeness properties (Arfken 1985)

$$\int_{4\pi} \psi_n(\hat{\mathbf{r}})\, \psi_m(\hat{\mathbf{r}})\, d\Omega = \delta_{nm}, \quad (20)$$

$$\sum_{n=1}^{\infty} \psi_n(\hat{\mathbf{r}})\, \psi_n(\hat{\mathbf{r}}') = \delta(\hat{\mathbf{r}} - \hat{\mathbf{r}}'), \quad (21)$$

where δ_{mn} is the Kronecker delta. The above two properties allow us to express

$$\rho_\tau(\hat{\mathbf{r}}, \hat{\mathbf{r}}') = \sum_{n=1}^{\infty} \lambda_n \psi_n(\hat{\mathbf{r}}) \psi_n(\hat{\mathbf{r}}'). \quad (22)$$

This is where, in principle, the EOFs pay off, since they conveniently express the spatial correlation structure over the earth for the averaging period τ. From the above formulas, the linear equations (16) become

$$\sum_{j=1}^{N_{\text{net}}} \alpha_{ij} w_j - \Lambda = \beta_i, \quad i = 1, 2, \cdots, N_{\text{net}}, \quad (23)$$

where

$$\alpha_{ij} = \sum_{n=1}^{\infty} \lambda_n \psi_n(\hat{\mathbf{r}}_i)\psi_n(\hat{\mathbf{r}}_j) \quad (24)$$

$$\beta_i = \frac{1}{4\pi} \sum_{n=1}^{\infty} \lambda_n \psi_n(\hat{\mathbf{r}}_i)\bar{\psi}_n \quad (25)$$

and

$$\bar{\psi}_n \equiv \frac{1}{4\pi} \int_{4\pi} \psi_n(\hat{\mathbf{r}})\, d\Omega. \quad (26)$$

Up to this point, we have converted the problem (16) of averaging the unknown covariance function $\rho_\tau(\hat{\mathbf{r}}, \hat{\mathbf{r}}_i)$ for all stations ($i = 1, 2, \ldots, N_{\text{net}}$) to the problem (26) of averaging the EOFs. The EOF approach automatically takes care of the nonhomogeneity property of temperature anomalies and $\bar{\psi}_n$ can be computed following the method of Kim and North (1993). The EOF $\psi_n(\hat{\mathbf{r}})$ is expanded into a series of spherical harmonics:

$$\psi_n(\hat{\mathbf{r}}) = \sum_{j=1}^{\infty} a_j^{(n)} Y_j(\hat{\mathbf{r}}), \quad n = 1, 2, \cdots, \quad (27)$$

where Y_j are the usual spherical harmonics but with rearranged indices. Then the average of the EOFs are

$$\bar{\psi}_n = a_1^{(n)}, \quad n = 1, 2, 3, \cdots. \quad (28)$$

For details of this procedure, see section 2 of Kim and North (1991).

We now form the $N_{\text{net}} + 1$ dimensional vector of unknowns $\{v_i\} = \{w_1, \cdots, w_{N\text{net}}, \Lambda\}$ and the $N_{\text{net}} + 1$ dimensional vector of the rhs: $\{b_i\} = \{\beta_1, \cdots, \beta_{N\text{net}}, 1\}$, whence the v_j are the solutions of the linear system

$$\sum_{j=1}^{N_{\text{net}}+1} M_{ij} v_j = b_i, \quad (29)$$

where

$$M_{ij} = \begin{cases} \alpha_{ij}, & i, j \leq N_{\text{net}}, \\ -1, & i \leq N_{\text{net}}, \quad j = N_{\text{net}} + 1, \\ 1, & i = N_{\text{net}} + 1, \quad j \leq N_{\text{net}}, \\ 0, & i, j = N_{\text{net}} + 1. \end{cases} \quad (30)$$

We are now in a position to find the optimal weights by simply inverting $\mathbf{M}$. We call the optimal weights $\{w_i^{\text{opt}};\ i = 1, \cdots, N_{\text{net}}\}$. These optimal weights are used to reevaluate the optimal mean-square error that can be written as the sum of the contributions of the EOFs in successive order

$$\epsilon_{\text{opt}}^2 = \sum_{n=1}^{\infty} \lambda_n [\bar{\psi}_n - \sum_{i}^{N_{\text{net}}} w_i^{\text{opt}} \psi_n(\hat{\mathbf{r}}_i)]^2. \quad (31)$$

It is worthy of note that the quantity in brackets will vanish for sufficiently large N_{net} provided the configuration is such that it "covers" the globe in the sense of a Riemann sum.

It is interesting to ask about the convergence of the sum over n in (31). This gives us some idea about the number of EOF modes needed to get a good answer. We address this in the next section, which covers numerical examples.

4. Numerical examples

In this section we explore the random sampling error budget for a few network configurations. We use modified data (smoothed to T11 and T15) obtained from the U.K. dataset (Jones et al. 1986b,c) for our examples. Starting with such a gridded dataset entails some error being introduced into our procedure from the outset since smoothing has already been applied in putting the data into gridded form. In the present application we consider this to be unimportant. We simply are considering the U.K. data to be an example of a dataset with nearly the same statistical properties as the real temperature anomalies. The original U.K. data are presented on a $5° \times 5°$ latitude–longitude grid. For computational convenience we further smoothed it by expanding it into spherical harmonics truncated at T11. Since, from (31), the effects of this truncation are also potentially important in what follows, we used a T15 truncation to check the sensitivity to truncation level. We also considered two time averaging intervals, $\tau = 1$, 5 years.

Two sets of EOFs were used as a measure of sensitivity to their shapes and eigenvalue spectra. The first set was computed directly from τ-length averages of the U.K. data; the second set was for τ-length averages based upon a noise-forced energy balance model. Both sets of EOFs have been described in detail by Kim and North (1993).

In all of the examples we used five different network configurations:

(i) 4×4 uniform (in latitude and longitude),
(ii) 6×4 uniform,
(iii) 9×7 uniform,
(iv) Angell–Korshover grid (63 stations), and
(v) 20×10 uniform stations.

Figure 1 shows the convergence of ϵ^2 [(°C)^2] as a function of truncation level in (31) using the data-derived EOFs all for $\tau = 1$ year; that is, the ordinate is

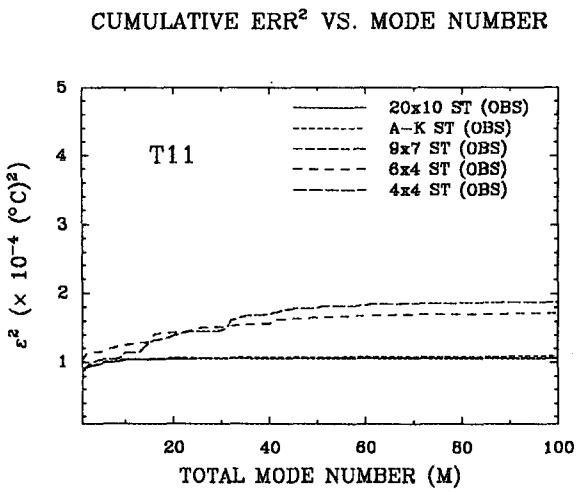


FIG. 1. Cumulative error squared [(°C)²] versus truncation level of the formula for the mean square error (32) [also see (31)] for the case of one year averages, EOFs computed from U.K. data, and EOFs and data are truncated at T11.

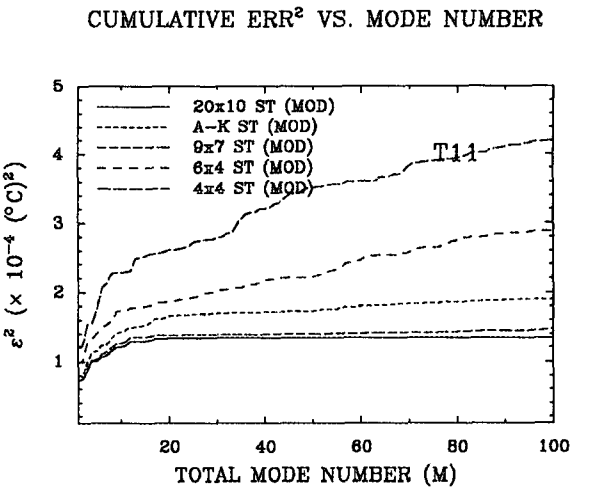


FIG. 2. Same as Fig. 1 except EOFs are computed from the noise-forced energy balance model of Kim and North (1992, 1993).

$$\epsilon_M^2 = \sum_{n=1}^{M} \lambda_n [\bar{\psi}_n - \sum_{i}^{N_{\text{net}}} w_i^{\text{opt}} \psi_n(\hat{\mathbf{r}}_i)]^2 \qquad (32)$$

and the abscissa is the total mode number M. Equation (32) is an ensemble average estimate of the error squared, and actual error for a particular realization can be several times larger than (32). In all five cases it can be seen that the convergence is essentially complete after about 50 EOF modes (cf. T11, which contains 144 spherical harmonic modes). For the densest network we find standard errors (ϵ) of about 0.01°C, while those for the coarsest network are only about 50% larger. For comparison we show in Fig. 2 the corresponding graphs for the EBM-derived EOFs. We can see that the convergence is slower and the asymptotic mean-square errors are about twice as large (standard errors about 1.41 times as large).

To see whether our results are sensitive to the truncation level of our spherical harmonic expansion, we repeated the procedure with a T15 truncation. The resulting convergence graph is shown in Fig. 3, which almost coincides with Fig. 1. We conclude that for this problem the T11 truncation is ample.

As an additional measure of the error we took the 100 years of U.K. monthly data and sampled it with our five configurations defined earlier in this section. Then the global average from the sparse station sampling is compared with that from the full dataset using correlation as an index of similarity. We use the zeroth-order component of spherical harmonic expansion as the mean of the full data. We used three weighting schemes to check for the sensitivity to weighting of station data. The first is uniform weighting (each station gets the same weight: $1/N_{\text{net}}$), the second is weighting by area (stations representing larger areas get proportionally larger weights), and finally our optimal weighting (using both data-derived and model-derived EOFs) as derived in the last section. The area weight for the station at (θ_i, ϕ_i) of the 4×4 network, for example, is the area of the earth surface region $(\theta_i - 45°, \theta_i + 45°) \times (\phi_i - 22.5°, \phi_i + 22.5°)$ $(i = 1, 2, \cdots, 16)$. The area weights for other networks of uniform grid are computed in the similar way. The area weight for a station of the Angell–Korshover network is computed according to the number of $5° \times 5°$ boxes associated with the station.

In computing sample correlations we used the 100 years of data at our disposal. Of course, these are not statistically independent, since there is serial correlation in the data. The serial correlation in the data increases the uncertainty in estimating the true correlation with

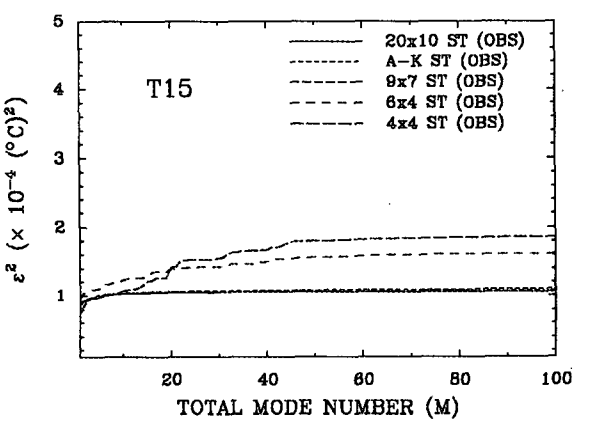


FIG. 3. Same as Fig. 1 except all calculations are based upon spherical harmonic truncations at T15.

TABLE 1. Sample correlations r of one year global average estimates based upon the U.K. 100 year dataset with the various network configurations and for different weighting schemes.

config.	uniform	area wt.	opt wt (data)	opt wt (model)
4 × 4	0.8813	0.9120	0.9168	0.9180
6 × 4	0.9041	0.9522	0.9634	0.9508
9 × 7	0.9763	0.9971	0.9949	0.9967
A–K	0.9630	0.9697	0.9893	0.9906
20 × 10	0.9759	0.9999+	0.9999+	0.9989

TABLE 3. Sample correlations of five year global average estimates based upon the U.K. 100 year dataset with the various network configurations for different weighting schemes. The last column shows the percent error for the optimally weighted case as in Table 2.

config.	uniform	area wt.	opt wt (data)	pct samp error
4 × 4	0.8770	0.9101	0.9008	43%
6 × 4	0.9053	0.9528	0.9673	25
9 × 7	0.9769	0.9967	0.9960	9.0
A–K	0.9644	0.9717	0.9857	17
20 × 10	0.9762	0.9999+	0.9999+	0.06

what is effectively less than 100 independent samples. Nevertheless, the sample correlation r, defined as the correlation coefficient between the time series constructed from the full U.K. data and that from sparse network sampling, are still indicative of the degree of correspondence between the coefficient of the zeroth-order spherical harmonic in the T11 truncation and the estimate based upon the networks of finitely many stations. The results are shown in Table 1.

From Table 1 we can see that the A–K network of 63 stations leads to a sample correlation $r = 0.963$, giving an r^2 of 0.922; in other words only ~8% of the variance in the record is due to sampling error. However, if we define the percent sampling error by $(1 - r^2)^{1/2} \times 100\%$, the same figures say that for an individual measurement, the sampling error is about 27% of the standard deviation of the natural fluctuation. On the other hand, optimal weighting brings the sample correlation to $r = 0.989$, which corresponds to the sampling error being 15% of the climatological standard deviation. We show the percentage sampling error $[(1 - r^2)^{1/2} \times 100\%]$ in Table 2 for all the five cases considered.

Some numerical results presented in Tables 1 and 2 may be puzzling, such as the fact that for the 9 × 7 grid, the area weighting leads to less error than optimal weighting. This is actually not surprising since the error here is defined according to the sample correlation r and the time series constructed is only a single realization. This sort of error fluctuation is not unusual. Nonetheless, the mean-square error is still smaller when optimal weights are applied (see Figs. 1–3).

Table 2 shows another interesting feature. The results in the last two columns are remarkably similar, in spite of the fact that the data-derived and the model-derived EOFs differ considerably (Kim and North 1993). On the other hand the data-derived and the model-derived eigenvalues are rather similar. We believe this similarity of the error assessments indicates that error budget calculations (and perhaps many other estimation theory calculations) can be conducted with rather crude EOFs as long as the eigenvalues are accurately calculated. This is an important result if true, since EOFs in general are rather difficult to compute from short data records (e.g., North et al. 1982; Kim and North 1993).

To get an idea how much the error is lowered by taking a longer averaging interval τ, we show in Table 3 the results for 5-year averages using the observational EOFs truncated at spherical harmonic resolution T11.

Some improvement is expected by the use of longer time averaging, since the correlation lengths are somewhat larger over oceans for the longer time period (Kim and North 1991, 1992). On comparing Table 2 with Table 3, however, we see only mixed improvement. We suspect the sampling errors (only 20 ensemble members for the 5-year averages) in our estimates of the correlations are to blame for the ambiguity.

Finally, we show the corresponding result for the 1-year averaging but using the T15 spherical harmonic truncation (Table 4). These results correspond to the convergence diagram in Fig. 3.

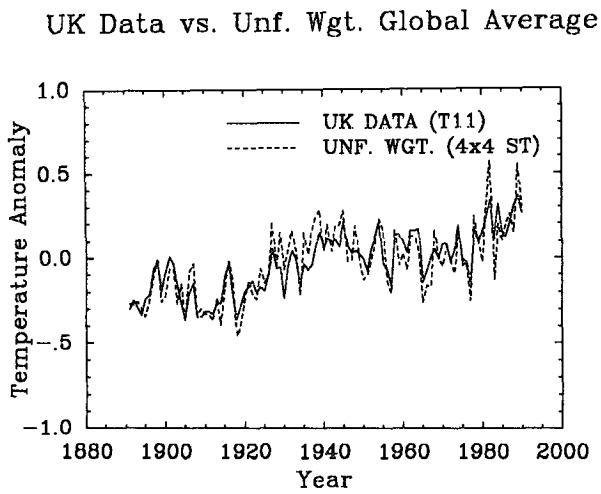
As an additional means of displaying the close correspondence between the global averages of the annual mean at the 5° × 5° grid level (smoothed to T11) and the sample estimates by sparse network stations, we show some time series in Figs. 4 and 5. In Fig. 4 we

TABLE 2. Percent random sampling error $\sqrt{1 - r^2} \times 100\%$ compared to the natural variability of annual averages based upon data in Table 1.

config.	uniform	area wt.	opt wt (data)	opt wt (model)
4 × 4	47%	41%	40%	40%
6 × 4	43	31	27	31
9 × 7	22	7.6	10	8.1
A–K	27	24	15	14
20 × 10	22	1.7	1.0	4.7

TABLE 4. Sample correlations of one year global average estimates based upon the U.K. 100 year dataset with the various network configurations for different weighting schemes based upon the T15 spherical harmonic truncation in computing the EOFs from observational data.

config.	uniform	area wt.	opt wt (data)	pct samp error
4 × 4	0.8602	0.9040	0.9330	36%
6 × 4	0.9082	0.9566	0.9771	21
9 × 7	0.9765	0.9931	0.9928	12
A–K	0.9578	0.9688	0.9701	24
20 × 10	0.9775	0.9999+	0.9999	2.6



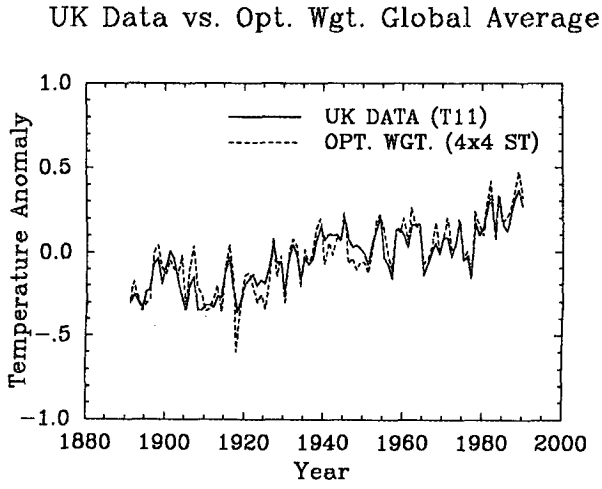
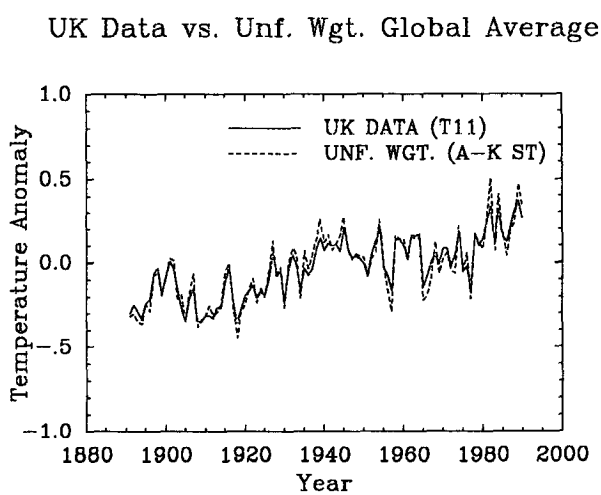


FIG. 4. Time series of global average temperatures based upon U.K. data (solid line) smoothed to T11 along with (a) a uniformly weighted 16-station symmetric configuration of stations. (b) Same but for optimally weighted station data.

show the fit for the 16-station case first (a) for the uniform weighting of 1-year averaged station data and (b) for the optimally weighted station data. From Table 1, we see that the sample correlations for the 16-station network are 0.8813 and 0.9168 for uniform weighting and optimal weighting, respectively. Hence the discrepancy between the two time series shown in Fig. 4 is noticeable. The sample correlations for the A–K network configuration are large (0.9630 for uniform weighting and 0.9893 for optimal weighting), and one expects a very fine agreement between the two time series as shown in Fig. 5.

5. Conclusions

In this paper we have developed a simple spectral formalism based upon earlier work (North and Nakamoto 1989; North et al. 1992; Hardin et al. 1992; Hardin and Upson 1994) to calculate the mean-square error of the time mean global average surface temperature field. The error is a function of the weights associated with each station. These weights were adjusted to give a minimum mean-square error. We developed the mean-square error formula into a sum of contributions from EOF modes, each term contributing successively less to the total mean square error. In this way the convergence of the formula could be explicitly examined. It was found that when the EOFs derived from the data were used, only about 50 modes were necessary to approximate the mean-square error. To get an idea how sensitive the mean-square error is to the choice of EOF basis, we used a set of EOFs derived from a noise-forced energy balance model (Kim and North 1993). The theoretical mean-square error was about twice as large when the model-derived EOFs were used in the calculation. Getting a rough idea about



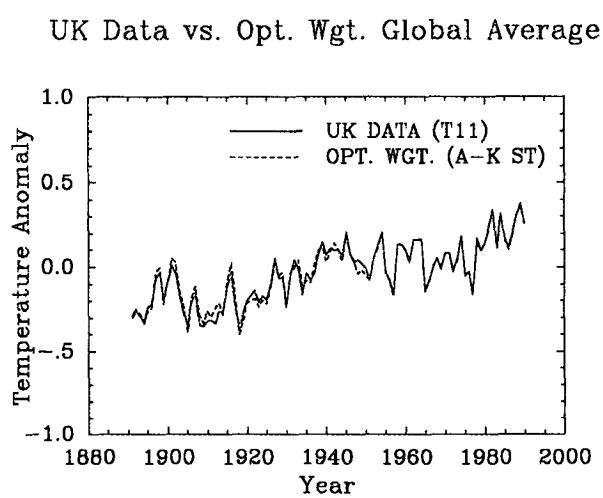


FIG. 5. Same as Fig. 4 except for the Angell–Korshover network.

these sensitivities is important since we anticipate having no choice but to use model-derived EOFs in some estimation theory problems in the future.

We calculated the 100-year sample correlation between the sparse station configuration estimates of the global averaged temperatures with the exact values defined as the coefficient of the zeroth-order spherical harmonic in T11 or T15 truncation. The latter estimate is exact only in the context of the spherical harmonic expansion. Casting the observed data into the expansion coefficients involves some errors, which we neglect in this study. Then, we found the magnitude of sampling error relative to the standard deviation of natural variability. This normalized error was found to be close to 25% for the Angell–Korshover network of 63 stations but is nearly halved if optimal weighting coefficients were applied to the stations. This suggests that when the number of well-dispersed stations is below about 100, it is advantageous to make use of optimal weighting. We note of course that the correlation to the U.K. data at fine scale is not necessarily the same as with perfect data. In fact, we correlated the coefficients of the zeroth harmonics in T11 truncation derived from the U.K. data with the Hansen–Lebedeff data and found the correlation coefficient to be only 0.93! This is presumably because the two groups used different means of removing biases, etc. (We do not mean to imply that one is in any sense better than the other.)

The illustrations drawn here were mainly pedagogical, in that some important features were ignored in order to isolate and study the effects of sampling error alone. Another limitation in our study is that we used the U.K. dataset, which had been interpolated onto a $5° \times 5°$ grid; we further smoothed the field by expanding into spherical harmonics and truncating at T11 (in keeping with our previous EBM work). These procedures have a tendency to filter out high wavenumber variance that might contribute to the mean square error at high EOF index. We think this is not a serious shortcoming since the series for the mean-square error seemed to converge in about 50 terms, whereas, the convergence might have been expected to be slower if this high wavenumber variance were important. The reason for this lack of variability at high wavenumbers has been remarked upon by many authors and is equivalent to the fact that the autocorrelation length for annual average data is about 1500 km or larger (Hansen and Lebedeff 1988; North 1984; North et al. 1992). As a test of sensitivity to level of truncation we repeated much of our procedure at the T15 level and found little difference.

One additional very interesting problem remains, that of the seasonal dependence of the global mean temperature. A study analogous to the one presented here would be very interesting for several reasons. First, the time series would have to be treated as a cyclostationary process rather than a stationary one. The effects of the cyclostationarity are strong since it is well known that both the variance and the local correlation lengths are strongly dependent on season (e.g., Plantico et al. 1990; Karl et al. 1993). Furthermore, there is some evidence that the warming signal and certainly the natural variability are dependent on season. We are presently engaged in a study of this phenomenon by means of cyclostationary EOFs.

In spite of the limitations just alluded to we feel that the magnitude of the sampling error is probably small compared to the other errors in the budget once the number of well-distributed stations is above about 60 and the optimal weighting is applied.

Acknowledgments. The authors gratefully acknowledge support from a grant from the U.S. Department of Energy "Quantitative Links" program. Shen wishes to thank the Environment Canada for a subvention grant by the Atmospheric Environment Service.

REFERENCES

Angell, J. K., 1992: Changes in tropospheric and stratospheric global temperatures, 1958–1988. *Greenhouse-Gas-Induced Climatic Change: A Critical Appraisal of Simulations and Observations,* M. Schlesinger, Ed., Elsevier, 231–247.

——, and J. Korshover, 1983: Global temperature variations in the troposphere and stratosphere, 1958–1982. *Mon. Wea. Rev.,* **111,** 901–921.

Arfken, G., 1985: *Mathematical Methods for Physicists.* 3d ed. Academic Press, 985 pp.

Folland, C. K., T. R. Karl, N. Nicholis, B. S. Nyenzi, D. E. Parker, and K. Ya. Vinnikov, 1992: Observed climate variability and change. *Climate Change, The IPCC Scientific Assessment,* J. Houghton, B. Callander, and S. Varney, Eds., Cambridge University Press, 135–170.

Hansen, J., and S. Lebedeff, 1987: Global trends of measured surface air temperature. *J. Geophys. Res.,* **92,** 13 345–13 372.

——, and ——, 1988: Global surface air temperatures: Update through 1987. *Geophys. Res. Lett.,* **15,** 323–326.

Hardin, J. W., and R. B. Upson, 1994: Estimation of the global average temperature with optimally weighted point gauges. *J. Geophys. Res.,* **98,** 23 275–23 282.

——, G. R. North, and S. S. Shen, 1992: Minimum error estimates of global mean temperature through optimal arrangement of gauges. *Environmetrics,* **3,** 15–27.

Jones, P. D., and T. M. L. Wigley, 1990: Satellite data under scrutiny. *Nature,* **344,** 711.

——, ——, and P. B. Wright, 1986a: Global temperature variations between 1861 and 1984. *Nature,* **322,** 430–434.

——, S. C. B. Raper, R. S. Bradley, H. F. Diaz, P. M. Kelly, and T. M. L. Wigley, 1986b: Northern Hemisphere surface air temperature variation: 1851–1984. *J. Climate Appl. Meteor.,* **25,** 161–179.

——, ——, ——; ——, ——, and ——, 1986c: Southern Hemisphere surface air temperature variations, 1851–1984. *J. Climate Appl. Meteor.,* **25,** 1213–1230.

——, T. M. L. Wigley, and G. Farmer, 1991: Marine and land temperature data sets: A comparison and a look at recent trends. *Greenhouse-Gas-Induced Climatic Change: A Critical Appraisal of Simulations and Observations,* M. Schlesinger, Ed., Elsevier, 153–172.

Kagan, R. L., 1979: Averaging meteorological fields (in Russian). *Gidrometeoizdat,* 212 pp.

Karl, T. R., R. G. Quale, and P. Y. Groisman, 1993: Detecting climate variations and change: New challenges for observing and data management systems. *J. Climate,* **6,** 1481–1494.

Kim, K.-Y., and G. R. North, 1991: Surface temperature fluctuations in a stochastic climate model. *J. Geophys. Res.*, **96,** 18 573–18 580.

——, and ——, 1992: Seasonal cycle and second-moment statistics of a simple coupled climate system. *J. Geophys. Res.*, **97,** 20 437–20 448.

——, and ——, 1993: EOF analysis of surface temperature field in a stochastic climate model. *J. Climate,* **6,** 1681–1690.

Madden, R. A., D. J. Shea, G. W. Branstator, J. J. Tribbia, and R. O. Weber, 1993: The effects of imperfect spatial and temporal sampling on estimates of the global mean temperature: Experiments with model data. *J. Climate,* **6,** 1057–1066.

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

——, and S. Nakamoto, 1989: Formalism for comparing rain estimation designs. *J. Atmos. Oceanic Technol.*, **6,** 985–992.

——, and K.-Y. Kim, 1995: Detection of forced climate signals. Part II: Simulation results. *J. Climate,* in press.

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

——, S. S. Shen, and J. W. Hardin, 1992: Estimation of the global mean temperature with point gauges. *Environmetrics,* **3,** 1–14.

——, K.-Y. Kim, S. S. Shen, and J. W. Hardin, 1995: Detection of forced climate signals. Part I: Filter theory. *J. Climate,* in press.

Plantico, M. S., T. R. Karl, G. Kukla, and J. Gavin, 1990: Are recent changes of temperature, cloudiness, sunshine, and precipitation across the United States related to rising levels of anthropogenic greenhouse gases? *J. Geophys. Res.*, **95,** 16 617–16 632.

Trenberth, K. E., and J. G. Olson, 1992: Representativeness of a 63-station network for depicting climate changes. *Greenhouse-Gas-Induced Climatic Change: A Critical Appraisal of Simulations and Observations,* M. Schlesinger, Ed., Elsevier, 249–259.

Vinnikov, K. Ya., P. Ya. Groisman, and K. M. Lugina, 1990: Empirical data on contemporary global climate changes (temperature and precipitation). *J. Climate,* **3,** 662–677.