Spatial–Temporal Analysis of Temperature Using Smoothing Spline ANOVA

ZHEN LUO

Department of Statistics, The Pennsylvania State University, University Park, Pennsylvania

GRACE WAHBA

Department of Statistics, University of Wisconsin—Madison, Madison, Wisconsin

DONALD R. JOHNSON

Space Science and Engineering Center, University of Wisconsin—Madison, Madison, Wisconsin

(Manuscript received 3 February 1997, in final form 2 May 1997)

ABSTRACT

A new method, smoothing spline ANOVA, for combining station records of surface air temperature to get the estimates of regional averages as well as gridpoint values is proposed. This method is closely related to the optimal interpolation (also optimal averaging) method. It may be viewed as a generalization of these methods from spatial interpolation methods to a method interpolating in both spatial and temporal directions. The connection of this method to the commonly used anomaly approach is discussed in the context of correcting biases resulting from incomplete sampling. A main strength of this new method is its ability to borrow information across both space and time just like optimal interpolation does across space. This increases not only the accuracy of estimates but also the ability to correct various biases resulting from incomplete sampling. Some of these biases are ignored by the anomaly approach.

1. Introduction

A number of studies (Vinnikov et al. 1980; Yamamoto and Hoshiai 1980; Jones et al. 1982; Jones et al. 1986; Hansen and Lebedeff 1987; Vinnikov et al. 1990) have used surface air temperature data at meteorological stations around the world to estimate changes over a large area (regional, zonal, or global averages, etc.). The problem addressed in these studies may be, directly or indirectly, viewed as estimating the surface air temperature as a function of location and time based on noisy data scattered in both spatial and temporal domains. Various averages may be calculated after an estimate of the function is obtained. In the context of numerical weather prediction, Lorenc (1986) reviewed a few methods for getting such estimates, which include optimal interpolation, smoothing splines, kriging, and Kalman filters, and pointed out that they are formally equivalent to each other. The purpose of this article is to introduce a new estimating method [smoothing spline ANOVA; see Gu and Wahba (1993a,b) and Luo (1996a,b)], which differs from the methods used in the previously cited studies in that it accomplishes spatial and temporal analyses at the same time. In other words, this is not just another estimating method in the spatial domain. This new method is capable of borrowing information across both spatial and temporal domains, hence making the estimates more accurate and increasing our ability to correct various biases resulting from incomplete sampling. Borrowing information across both space and time is certainly not a new idea. In numerical weather predication and many other fields, people have borrowed information across time, usually from the past, through dynamical models. [See Lorenc (1986), Derber and Rosati (1989), Thiebaux (1991), and Wahba et al. (1995), and references therein.] However, there have not been many studies that borrow information across both space and time empirically. A. Kaplan et al. (1996, manuscript submitted to *J. Geophys. Res.*) is one example. In their approach, an empirical model (a Markov model) is substituted for the dynamical model (and hence the model itself has to be estimated too). Our approach here represents a more direct generalization of optimal interpolation type methods in spatial domain to one interpolating in both spatial and temporal domains. The thin plate spline in Wahba and Wendelberger (1980), used as a spatial estimating method on rectangular domains, can also accommodate time in a very different manner than proposed here. The direct methods

Corresponding author address: Dr. Zhen Luo, Dept. of Statistics, The Pennsylvania State University, 326 Thomas Building, University Park, PA 16802-2111.
E-mail: zhen@stat.psu.edu

of this paper result in a global space and time fit that is easily decomposed into various averages and anomalies of meteorological interest.

To calculate a global average or other regional averages, the simple average of available station records is obviously biased toward the area concentrated with more stations. More sophisticated methods are needed. Vinnikov et al. (1980) subjectively contoured the station data to get gridpoint estimates, then averaged them with cosine weighting to account for the change of grid density along meridians of latitude. Jones et al. (1982) did a similar computation except with an objective method for getting gridpoint estimates (nearest neighbor inverse distance weighted average). Yamamoto and Hoshiai (1980) used ''optimal interpolation'' to estimate gridpoint values from station data. Jones et al. (1986) divided the globe into 36 × 36 boxes and within each box the inverse distance weighted average was used to estimate the gridpoint value corresponding to that box. Hansen and Lebedeff (1987) divided the globe into a number of equal-area small boxes and computed a mean value for each small box using a distance-weighted average. Then a hierarchical average of box mean values (from small boxes to bigger boxes, then to latitude bands, to hemispheres, with different weighting schemes at different levels) is used as an estimate of the global average. Vinnikov et al. (1990) used a statistically optimum averaging method to compute different regional means directly without computing gridpoint values. In section 2, we will describe the smoothing spline method to get both gridpoint estimates and regional average estimates. This method is, as pointed out by many authors (e.g., Lorenc 1986; Kimeldorf and Wahba 1971), formally equivalent to the optimal interpolation method. We will demonstrate in this article that statistically optimum averaging used by Vinnikov et al. (1990) is also formally equivalent to the estimate based on the smoothing spline method.

To compare global averages or other regional averages across time (the crudest way to look at the large area temperature change), another kind of bias exists due to the incompleteness of sampling over time; that is, the time period during which a station has records varies from station to station. Hence, the changes across years in the simple averages of each year's records are confounded with the changes across the locations of the stations used in each year's calculation. The way most previous studies have chosen to correct this bias is through the use of anomalies that are defined as the differences of raw records and the average over a prespecified reference period, usually a period with good sampling coverage. Hansen and Lebedeff (1987) used a different scheme that, in some sense, is like an anomaly approach with a varying reference period. In section 3a, we will show that while the anomaly approach is satisfactory in general, there are some significant biases resulting from incomplete sampling that the anomaly approach cannot correct. We will show in section 3b that the smoothing spline ANOVA approach as described there can correct such biases directly using raw station data.

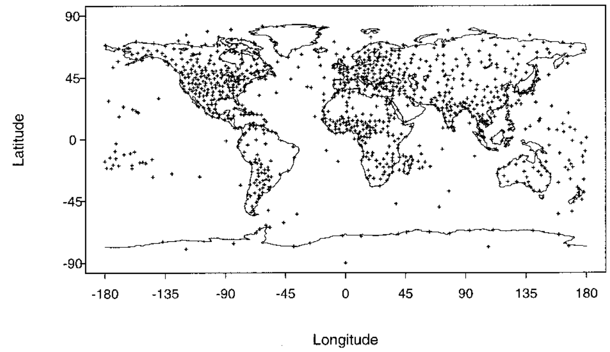

FIG. 1. The distribution of the 1000 stations used.

We chose the dataset prepared by Jones et al. (1991) to apply this new method. The data were obtained from http://cdiac.ESD.ORNL.GOV/ftp/. It is a combination of four files: ndp020r1/jonesnh.dat, ndp020r1/jonessh.dat, ndp032/ndp032.tm1, and ndp032/ndp032.tm2. This dataset is assembled from different sources of monthly temperature records at about 2000 stations distributed across the world over the period from 1851 through 1991. There are only a few stations with records dating back that far. Most stations started recording in this century. The stations are concentrated heavily in Europe and North America. Some cleaning and homogenizing to the original data have been done by Jones et al. (1991).

We did not redo the whole analysis of this dataset using our new method. Instead a subset of this dataset was chosen to illustrate our method. Only Northern Hemisphere winter mean temperatures, defined as the average of December, January, and February temperatures, are considered. The word ''winter'' throughout this article always refers to the winter in the Northern Hemisphere. The most recent 30-yr period (1961–90) is chosen. Instead of using all the stations in this dataset, we selected 1000 stations mainly due to the limit of our computing capacity (also see the discussion in section 2). These 1000 stations are chosen deliberately so that they cover the sphere as uniformly as possible. Hence, most stations left out are in Europe and North America while almost all the stations in other regions are included. Note that this selection of stations only mitigates the problem of nonuniformness of station distribution, it does not eliminate the problem. The distribution of these 1000 stations is shown in Fig. 1. To have a graphical idea of the incomplete sampling coverage, *a plot of each record's year versus latitude with the longitude nested within each year is given in Fig. 2*. Having longitude nested within each year gives a better idea of the data density. There are clearly fewer stations toward the later part of the period. The records for the Antarctic region end in 1988.

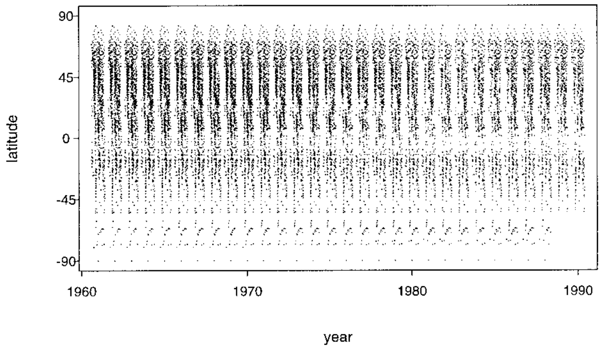

FIG. 2. The pattern of incompleteness in the data. The x axis is the longitude nested within each year.

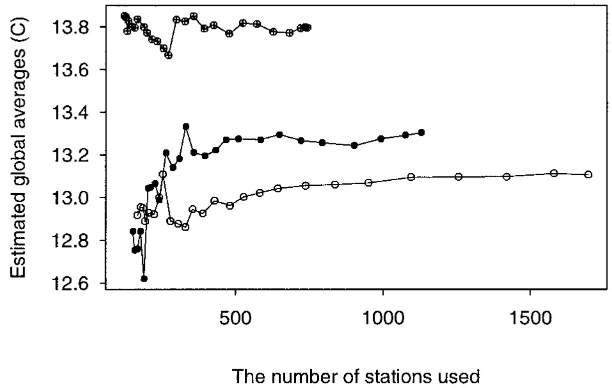

FIG. 3. Estimated global averages versus the number of stations used. Circles: estimates for 1970; solid circles: estimates for 1980; crossed circles: estimates for 1990.

2. Smoothing spline estimates for one time data

Suppose that we want to estimate the temperature field at one time, for example, winter mean temperature field in one particular year, based on noisy data at some locations:

$$y_i = f(P_i) + \epsilon_i, \quad \text{for} \quad i = 1, 2, \ldots, n. \tag{1}$$

The symbol P_i is a point on the sphere [i.e., $P_i = (\lambda_i, \phi_i)$, a latitude–longitude pair] and the sphere will be denoted by $\mathcal{S}$. The symbol ϵ_i represents a ''noise'' term that contains not only the measurement error in the record y_i, but also purely local variability that is of much smaller scale than the resolution of any model to be fitted.

A smoothing spline estimate of f, denoted by f_θ to emphasize its dependence on a smoothing parameter θ, may be defined as the function that minimizes

$$\sum_{i=1}^{n} [y_i - f_\theta(P_i)]^2 + \frac{1}{\theta}\int_{\mathcal{S}} [\Delta f_\theta(P)]^2 \, dP, \tag{2}$$

where θ is a regularization parameter (also called a smoothing parameter) that controls the trade-off between the closeness of f_θ to the data $\{y_i\}$, and the smoothness of f_θ. The smaller θ is, the more local features are smoothed out in f_θ. The second term in (2), usually called a penalty term, represents a weak constraint on f_θ. Peixoto and Oort (1992, p 84) have used a Poisson equation as a strong constraint to extend a temperature field from observation points to other points. It is a common practice in numerical weather prediction to use theoretical models as either strong constraints or weak constraints when estimating meteorological fields. Equation (2) represents the simplest example of a weak constraint. Higher powers of Δ or some other forms of penalty may be used. More work may be done in this direction to incorporate prior and other information into the penalty term. The details of computing f_θ are given in appendix A.

Based on the computed f_θ, various area averages may be calculated. We note that this scheme to get area averages leads to the same answers that Vinnikov et al. (1990) called the statistically optimum averaging [also called Best Linear Unbiased Prediction (BLUP) in statistical literature] does, if appropriate penalty form and smoothing parameter are chosen, just as the gridpoint estimates by the smoothing spline method are the same as those from optimal statistical interpolation. See appendix B for the demonstration of this formal equivalence.

Many authors have chosen to estimate the field on grid points first and then to average gridpoint estimates to get area means. A reason for this is to avoid the bias that may result from the unbalanced station distribution. Figure 3 shows the smoothing spline estimates of the global average of winter temperature versus the number of stations used in calculating these estimates, for three arbitrarily chosen years. (Stations are chosen as uniformly distributed as possible by requiring a minimal distance between any two stations. The smaller this distance is, the more stations included.) The estimated global averages tend to stabilize after a certain number of stations are included in the analysis. The initial variation of the estimates results from the change of covered area, and after a certain point, the extra stations added in the area already covered have little influence on the estimate of global average. This shows that the smoothing spline method (also ''statistically optimum averaging'') can avoid the bias resulting from unbalanced station distribution without referring to gridpoint estimates. Based on Fig. 3, it seems that 1000 stations are about adequate to estimate the global average based on this particular dataset. That is, the estimates based on the entire dataset will not differ much from the estimates based on the subset of 1000 stations.

Even though the smoothing spline estimates of global averages are not affected by the unbalanced station distribution, we may still want to average only over the area where data are available, since global averages involve the extrapolation to the region where there is no data, (e.g., for this dataset, to the west coast of the Americas). We define the region covered by this particular dataset to be the area where there is at least one station within 500 km. This region is shown in Fig. 5 as the colored region. Since most of this region is land, we will refer to the average over this region as ''land average.'' ''Global average'' will refer to the average over the whole sphere, both land and sea regions. The choice of 500 km is subjective to some extent. We note that Jones (1994) used a 5° × 5° grid, which corresponds to about 550 km × 550 km at equatorial region, and a grid box with at least one station within it was counted as covered by the data. The land average in this article is calculated by the cosine-weighted average (accounting for the change of latitude) of gridpoint values on a 200 × 100 grid falling in the land region defined above. This is just a simple numerical way to integrate a function over an irregular region.

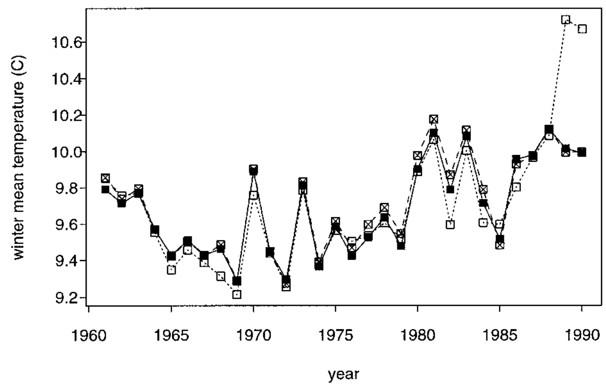

FIG. 4. Estimated land average winter temperatures by three methods. Squares: estimates by the smoothing spline method of section 2 applied on each year's temperature records; crossed squares: estimates by the smoothing spline method of section 2 applied on each year's anomalies; solid squares: estimates by the smoothing spline ANOVA method of section 3b.

3. Smoothing spline estimates for multiple time data

Suppose that we are interested in the time evolution of the temperature field. That is, we want to estimate the temperature field as a function of year and location, based on scattered noisy data:

$$y_i = f(t_i, P_i) + \epsilon_i, \qquad i = 1, 2, \ldots, n, \tag{3}$$

where $t_i \in \{1, 2, \ldots, n_t\}$ denotes the year and $P_i \in \mathcal{S}$ denotes the location of the ith data point. The value of n is the total number of observations used, n_t is the total number of years, and n_s is the total number of stations included in the study. One obvious way to estimate f is to ignore the time dependence and to estimate $f(t, P)$ as a function of P separately for each year. Besides the loss of efficiency (since the information could have been borrowed across years), this approach has a problem of biases resulting from the incomplete sampling. This problem and how the anomaly approach deals with it will be discussed first, together with the problem that the anomaly approach faces itself. After that we will describe our smoothing spline ANOVA approach.

a. Biases from incomplete sampling and anomalies

Applying the smoothing spline method of section 2 to each year's records in the period of 1961–90, we get a sequence of land averages of winter temperature, shown in Fig. 4 as squares. The last 2 yr have outstanding high values in this sequence. If we concluded that the dramatic increase of winter temperature occurred in the last 2 yr of the 1980s, then we had been misled by the bias resulting from the incomplete sampling, or the spatial sampling difference across years. The records for the Antarctic region end in 1988 (see Fig. 2). Obviously this abrupt increase of winter temperature in the last 2 yr is mainly because of the lack of data in the Antarctic region where it is much colder than most other regions of the world.

In order to correct the bias resulting from the spatial sampling difference, many previous studies have chosen anomalies, instead of raw temperature records, for comparison. An anomaly is defined as a difference between a temperature record and the average temperature over a specified reference period. Choosing 1961–90 as the reference period, a sequence of average temperature anomalies by the same smoothing spline method is shown in Fig. 4 as crossed squares. In this sequence, the outlying feature of the last 2 yr disappears.

The reason for the effectiveness of using anomalies to correct the bias resulting from the spatial sampling difference can easily be explained by the following data decomposition.

The temperature field as a function of year (t) and location (P) can always be written as a sum of its component functions

$$\begin{aligned} f(t, P) &= d_1 + d_2\phi(t) + g_1(t) + g_2(P) \\ &\quad + g_{\phi,2}(P)\phi(t) + g_{12}(t, P), \end{aligned} \tag{4}$$

where $t \in \{1, 2, \ldots, n_t\}$ and $P =$ (latitude, longitude) $\in \mathcal{S}$, and ϕ is a known linear function $\phi(t) = t - [(n_t + 1)/2]$. The following conditions on the component functions guarantee that representation (4) is unique:

$$\begin{cases} \sum_{t=1}^{n_t} g_1(t) = \sum_{t=1}^{n_t} g_1(t)\phi(t) = 0 \\ \sum_{t=1}^{n_t} g_{12}(t, P) = \sum_{t=1}^{n_t} g_{12}(t, P)\phi(t) = 0 \\ \int_{\mathcal{S}} g_2(P)\, dP = \int_{\mathcal{S}} g_{\phi,2}(P)\, dP = \int_{\mathcal{S}} g_{12}(t, P)\, dP = 0, \end{cases} \tag{5}$$

for any t and P. Because of these conditions, the component functions in (4) and their combinations are of clearly defined climatological meanings. For example, d_1 is the grand average temperature over both year and location; d_2 is the linear trend coefficient of global averages; $d_1 + d_2\phi + g_1$ is the global average temperature history; g_2, $g_{\phi,2}$, and g_{12} represent spatial variations about d_1, d_2 and g_1, respectively; $d_1 + g_2(P)$ is the average winter temperature at location P; and $d_2 + g_{\phi,2}(P)$ is the linear trend coefficient of winter temperatures at location P.

An observation is

$$y(t, P) = d_1 + d_2\phi(t) + g_1(t) + g_2(P) + g_{\phi,2}(P)\phi(t) + g_{12}(t, P) + \epsilon(t, P).$$

Considering (5), the station mean over the same period is

$$\overline{y}(P) := \frac{1}{n_t}\sum_{t=1}^{n_t} y(t, P) \simeq d_1 + g_2(P).$$

The approximate equality here is due to the fact that the records of some years may be missing and $\Sigma_t\, \epsilon(t, P)$ is only approximately zero. Therefore the anomaly is

$$y(t, P) - \overline{y}(P) \simeq d_2\phi(t) + g_1(t) + g_{\phi,2}(P)\phi(t) + g_{12}(t, P) + \epsilon(t, P).$$

Now it is clear that the differences in $g_2(P)$ resulting from different station sets across years do not affect the anomaly. However, the differences in the last two terms in (4) still do. The most suitable case in which using anomalies will eliminate any bias resulting from spatial sampling difference is when we are certain that the last two terms in (4) are not significant; that is, we know in advance that an additive model:

$$y(t, P) = d_1 + d_2\phi(t) + g_1(t) + g_2(P) + \epsilon(t, P), \quad (6)$$

is adequate. This is not the case here, however, since we know that not only is there spatial variation in the average temperature [($g_2(P)$)], but also in the temperature change trend over years. Some locations may have an increase, others may have a smaller increase or even a decrease. The pattern of such variation is actually one important aspect of the climate we would like to extract from the data. See, for example, Fig. 2 of Hegerl et al. (1996). This makes the last two terms in (4) nonnegligible when considering the bias resulting from spatial sampling differences.

Having pointed out this limitation of the anomaly approach, we would also like to make clear that it is true that the spatial variation in $g_2(P)$ is the most prominent spatial variation among the three terms in (4) involving P. The spatial variation in the average temperatures (in a range from about $-40°$C to about 40°C is much larger than the spatial variation in the year by year changes of temperature [$(g_{\phi,2}(P)\phi(t) + g_{12}(t, P)$], which is just a few degrees (°C). Therefore, the anomaly approach has eliminated the major part of biases resulting from spatial differences. This is probably one of the reasons for its satisfactory use so far.

The approach to be described in the following section is very different from the previous approaches in that we fit raw temperature records instead of anomalies. By choosing appropriate parameters, this approach has the ability to correct the bias resulting from the locational difference in all the three terms of (4) involving P. The estimates of land average winter temperature using this approach are shown in Fig. 4 as solid squares. This sequence is similar to the one obtained by the anomaly approach. This is obviously evidence that our approach has a similar ability to correct the bias resulting from spatial sampling difference as the anomaly approach does. But they also have some differences. Since we do not single out $g_2(P)$, it is reasonable to expect that our approach will correct the bias resulting from all three terms in (4) involving P.

b. *Smoothing spline model for multiple years*

For the data in (3), a smoothing spline ANOVA estimate, f_θ, is defined as the minimizer of

$$\sum_{i=1}^{n} [y_i - f(t_i, P_i)]^2 + \frac{1}{\theta_1} J_1(g_1) + \frac{1}{\theta_2} J_2(g_2) + \frac{1}{\theta_3} J_3(g_{\phi,2}) + \frac{1}{\theta_4} J_4(g_{12}), \quad (7)$$

where f has a representation as in (4), and $J_1(g_1) = \sum_{t=1}^{n_t-2} [g_1(t + 2) - 2g_1(t + d) + g_1(t)]^2$, $J_2(g_2) = \int_S (\Delta g_2)^2\, dP$, J_3 is the same as J_2, and J_4 is derived from J_1 and J_2 as the norm of the corresponding tensor-product space. See Gu and Wahba (1993a,b) and Luo (1996a,b) for more details. Other forms of penalty and penalty terms with some additional parameters may also be used depending on the situation. The details of computing a smoothing spline estimate are given in appendix C.

How to choose smoothing parameters, θ's in (7), is a very crucial issue here, because the choice affects the smoothing spline estimate to a great extent. For example, if we choose θ_3 and θ_4 to be very small, we will essentially make $g_{\phi,2}$ and g_{12} disappear in our model and adopt an additive model (6); hence, the results will be very close to those from the anomaly approach. If we choose θ_2, θ_3, and θ_4 to be very small, then we will essentially get an estimate assuming no g_2, $g_{\phi,2}$, and g_{12} components, that is, no spatial variation at all. In that case, we will be very close to the naive approach of averaging each winter's raw temperatures separately.

There are basically two types of techniques for choosing smoothing parameters. One consists of the so-called objective or data-driven methods such as cross validation (CV), generalized cross validation (GCV), and generalized maximum likelihood (GML) estimations (see Wahba 1990, chap. 4). The other consists of "subjective" methods. We may compute for each choice of smoothing parameters an estimate of the standard deviation of the observation, then compare it with our prior knowledge about the size of such observation "error." We may also use the past data to estimate these param-

TABLE 1. RGCV scores for the 1000 station dataset. Here, log10(θ_1) and log10(θ_2) are fixed at $-.1$ and 4.5, respectively, and (*) indicates a local minimum.

log10(θ_4)	log10(θ_3)				
	1.5	1.25	1	.75	.5
4.4		0.63452	0.63617		0.64697
4.1	0.62752	0.62737(*)	0.62747	0.62909	0.63298
3.8		0.63958	0.63905		0.64201

eters. This is exactly Vinnikov et al.'s (1990) approach for deciding both their smoothing parameter and covariance function. In general, these subjective criteria rarely give us a precise choice of smoothing parameters, but still, they are very important in guiding us and are even sufficient for our needs in many applications. It is important to keep these criteria in mind even when we use data-driven criteria since so-called objective methods may give us misleading results also, not to mention that some important information is very hard to be formulated into objective criteria.

In our particular application, we decide to use a *subjective* method to choose θ_1 and θ_2 and an *objective* method to choose θ_3 and θ_4. The main reason is because of the large computational demand of choosing all four θ's by an objective method. Another reason is that we have a relatively clear idea about how much smoothing should be done to g_1 and g_2. As a matter of fact, we want little smoothing done to them. A way to relate this information to a smoothing parameter is through the concept of ''degrees of freedom.'' A commonly used definition of the degrees of freedom in a smoothing spline estimate is $tr[\mathbf{A}(\theta)]$ (see Wahba 1990), where $\mathbf{A}(\theta)$ is the influence matrix defined by $[f_\theta(x_1), f_\theta(x_2), \ldots, f_\theta(x_n)]^T = \mathbf{A}(\theta)(y_1, y_2, \ldots, y_n)^T$, that is, the coefficient matrix of the linear dependence of the estimated values on the observed values. This concept can be readily generalized to each component of f_θ. See appendix C for more details. The maximum degrees of freedom is $(n_t - 2)$ for g_1 and n_s for g_2, where n_t is the total number of years, and n_s is the total number of stations used in the analysis. To choose θ_1 and θ_2 in such a way that little smoothing is done to g_1 and g_2, we just choose them so that their corresponding degrees of freedom are close to their maximum values.

A commonly used data-driven method is to choose θ's that minimize the GCV score, which is defined as

$$V(\theta) = \frac{\frac{1}{n}\|\mathbf{y} - \hat{\mathbf{f}}\|^2}{\left[\frac{1}{n} tr(\mathbf{I} - \mathbf{A}(\theta))\right]^2},$$

where $\hat{\mathbf{f}} = [f_\theta(x_1), f_\theta(x_2), \ldots, f_\theta(x_n)]^T$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)^T$. The numerator $\|\mathbf{y} - \hat{\mathbf{f}}\|^2$, the residual sum of squares, can easily be computed after we get the estimate of the function. However, when the data size is very large as in our case here, computing $tr[\mathbf{I} - \mathbf{A}(\theta)]$ exactly can be very expensive; hence, we use an approximation to the GCV score called ''randomized GCV'' (RGCV) score:

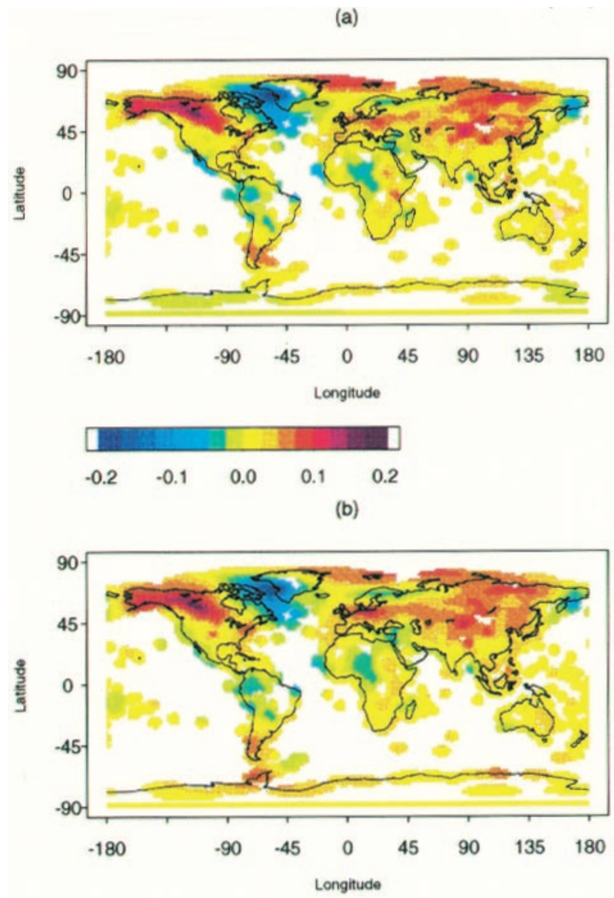

FIG. 5. Estimates of local winter temperature trend $[d_2 + g_{\phi,2}(P)]$. (a) estimates by the smoothing spline ANOVA method using the original data (truth for the simulated data); (b) estimates by the smoothing spline ANOVA method using the simulated data.

$$\mathrm{RGCV}(\theta) := \frac{\frac{1}{n}\|\mathbf{y} - \hat{\mathbf{f}}\|^2}{\left[\frac{1}{n}\boldsymbol{\xi}^T[\boldsymbol{\xi} - \hat{f}(\xi)]\right]^2},$$

where $\boldsymbol{\xi}$ is a standard multivariate normal random vector with the same length as the data vector, and $\hat{\mathbf{f}}(\xi)$ is the smoothing spline estimate when the data vector $\mathbf{y}$ is substituted by $\boldsymbol{\xi}$ [Girard (1989) and Wahba et al. (1995)]. In order to minimize the variation induced by $\boldsymbol{\xi}$, it is better to use the same $\boldsymbol{\xi}$ for all choices of θ.

Using the approach described above, θ_1 and θ_2 were chosen as $10^{-0.1}$ and $10^{4.5}$, which correspond to 27.8 degrees of freedom for g_1, and 989.8 for g_2. With θ_1 and θ_2 fixed, we choose θ_3 and θ_4 according to the RGCV criterion by a crude grid search. We first set some preliminary limits for them by the tool of the degrees of freedom of their corresponding components. For θ_3, the limits are $10^{.5}$ and $10^{1.5}$ corresponding to 565.4 and 890.5 degrees of freedom (the maximum is 1000), respectively. For θ_4, the limits are $10^{2.8}$ and $10^{4.4}$ corresponding to 7052.6 and 17 138.2 degrees of freedom (the maximum is 28 000, but the total number of observations is 20 910), respectively. Part of the search results are given in Table 1. A minimum in RGCV gives us a choice of $\theta_3 = 10^{1.25}$ and $\theta_4 = 10^{4.1}$, which correspond to 831.1 degrees of freedom for $g_{\phi,2}$ and 14 860.5 degrees of freedom for g_{12}, respectively.

The estimated standard deviation of ϵ, $\hat{\sigma}$, by the formula of Wahba (1990, section 4.7)

$$\hat{\sigma}^2 = \frac{\|\mathbf{y} - \hat{\mathbf{f}}\|^2}{tr(\mathbf{I} - \mathbf{A}(\theta))} \simeq \frac{\|\mathbf{y} - \hat{\mathbf{f}}\|^2}{\xi(\xi - \hat{\mathbf{f}}(\xi))}, \qquad (8)$$

is 0.49°C, which is a little bit larger than what a typical measurement error of mean temperature is expected to be. But this is still reasonable considering the fact that here ϵ contains not just the measurement error.

The estimates of land average winter temperatures [defined as $\int_{\text{land}} f_\theta(t, P)\, dP$] are shown in Fig. 4 as solid squares. There exists an overall cooling trend in the early 1960s and an overall warming trend from the 1970s on. The linear trend of land averages over these 30 yr is about 0.01587°C yr^{-1}. For comparison, the trend estimated by the anomaly approach of about 0.01591°C yr^{-1} is in very close agreement.

Estimated local winter mean temperature, that is, $d_1 + g_2(P)$, is shown in Fig. 7 (solid contour lines). This is the familiar pattern of (Northern Hemisphere) winter mean temperature across the world. Estimated local winter temperature (linear) trend, that is, $d_2 + g_{\phi,2}(P)$, is shown in Fig. 5a. We see that most of the European area has a warming trend (positive coefficient) except the eastern Mediterranean region and a large area of the North Atlantic, including Greenland. A cooling trend has been observed in part of Africa and America also. Strong warming trends have been noticed in parts of Siberia and North America.

The whole history of these 30-yr winter temperature anomalies [i.e., $d_2\phi(t) + g_1(t) + g_{\phi,2}(P)\phi(t) + d_{12}(t, P)$] based on our smoothing spline estimates is made into a movie that can be accessed at http://www.stat.psu.edu/~zhen.

As a by-product, we may check the residuals from smoothing spline fits to identify outliers [see Knight (1980) for a comparison with other approaches]. We plotted the residuals against the year and other variables and actually identified a few obvious outliers. They were suspected to be typos during data transcription. Details may be found in Luo (1996a), section 3.3.3. As of 17 June, 1996, the last time we visited the database at CDIAC, these possible typos were still there. We note, however, that this is not the latest database and not the one used in the latest IPCC Report; (see Nicholls et al. (1996). Also, the values at these few data points do not change the final results significantly.

c. Discussions

For the particular dataset used in section 3b, the differences between our smoothing spline estimate of the temperature field and the estimate based on the anomaly approach are not significant, as seen in Fig. 4. This does not mean, however, that these two methods will not give significantly different estimates in other situations. Besides the reason given in section 3a, there is another reason why in this particular example two approaches result in similar estimates. That is due to the fact that there is very little correlation between different years' winter temperatures as can be shown by looking at their time series' autocorrelation plot. Hence, the information borrowed across time is mainly in the 30-yr average winter temperature and 30-yr linear trend. In its own way, the anomaly approach has already borrowed a major part of this information through using the station mean temperature in the definition of each year's anomaly. Based on this analysis, we would expect more varied results in the situation where strong correlations across time do exist, for example, as pointed out by a referee, in daily or hourly temperatures. Another situation where we would expect more varied results is when the spatial variation of linear trend plays a bigger role in the total spatial variation, for instance, when a smaller spatial domain is considered and the mean temperature's variation does not dominate the total spatial variation.

The model described in section 3 can easily be extended to the situation where, besides year and location, we want to include other variables (e.g., season) into our model. For example, a model of monthly temperature similar to (4) may be defined as follows. Monthly temperature f as a function of year, month, and location, denoted by t, m, and P, respectively, can be written as a sum of several component functions:

$$\begin{aligned} f(t, m, P) &= d_1 + d_2\phi(t) + g_1(t) \\ &\quad + g_2(P) + g_{\phi,2}(P)\phi(t) + g_{12}(t, P) \\ &\quad + g_3(m) + g_{\phi,3}(m)\phi(t) + g_{13}(t, m) \\ &\quad + g_{23}(m, P) + g_{\phi,23}(m, P)\phi(t) + g_{123}(t, m, P), \end{aligned} \qquad (9)$$

where these components satisfy some side conditions similar to those in (5).

A smoothing spline ANOVA estimate can be defined as the minimizer of

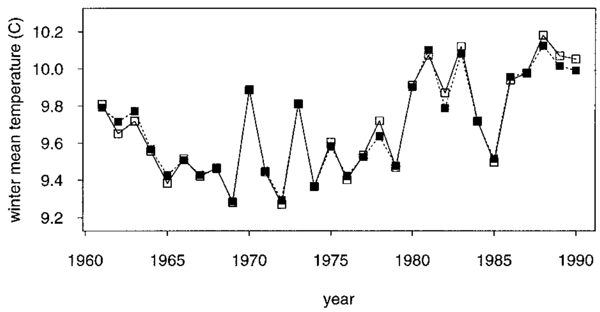

FIG. 6. Estimates of land average winter temperatures. Solid squares: estimates by the smoothing spline ANOVA method using the original data (truth for the simulated data); squares estimates by the smoothing spline ANOVA method using the simulated data.

$$\sum_{i=1}^{n} [y_i - f(t_i, m_i, P_i)]^2 + \frac{1}{\theta_1}J_1(g_1) + \frac{1}{\theta_2}J_2(g_2) + \frac{1}{\theta_3}J_3(g_{\phi,2}) + \frac{1}{\theta_4}J_4(g_{12}) + \frac{1}{\theta_5}J_5(g_3) + \frac{1}{\theta_6}J_6(g_{\phi,3}) + \frac{1}{\theta_7}J_7(g_{13}) + \frac{1}{\theta_8}J_8(g_{23}) + \frac{1}{\theta_9}J_9(g_{\phi,23}) + \frac{1}{\theta_{10}}J_{10}(g_{123}), \quad (10)$$

where J_1J_4 are the same as in (7), J_5 and J_6 are the same and may be defined as

$$J(g) := \sum_{m=1}^{12} [g(m + 2) - 2g(m + 1) + g(m)]^2,$$

with $g(13) := g(1)$ and $g(14) := g(2)$. This form of penalty is chosen because of the periodic nature of the variable month. The rest of the J's are defined through the tensor-product structure of their corresponding function spaces. See Luo (1996a) for more details. A vertical spatial coordinate can be similarly included.

4. Simulation

In order to assess the accuracy of the smoothing spline estimates described in section 3b, a small experiment is done. Using the residuals of the fit in section 3b, an estimate of the standard deviation of ϵ in (3) was calculated for each station. Pretending the estimates obtained in section 3b and these standard deviations to be the truth, a new dataset was generated by (3) with ϵ generated by a pseudonormal random variable generator. Then the same smoothing spline estimating method was applied to this dataset. The first two smoothing parameters (θ_1 and θ_2) were fixed at the same values as in section 3b. The other two smoothing parameters were chosen according to the RGCV criterion as $\log_{10}(\theta_3) = .625$ and $\log_{10}(\theta_4) = 3.5$.

The sequence of estimated land averages $[\int_{\text{land}} f(t, P)\, dP]$ is shown in Fig. 6 together with the "truth." The estimated local winter mean temperature $[d_1 + g_2(P)]$ and the corresponding "truth" are shown together in Fig. 7. The estimated local winter temperature trend $[d_2 + g_{\phi,2}(P)]$ is shown in Fig. 5b. It should be compared with Fig. 5a. In most areas (mainly the areas where sufficient data exist), the agreement between the truth (Fig. 5a) and the estimate (Fig. 5b) is quite good. But in areas where no data or very scattered data exist, there are some discrepancies. This suggests a way to get a kind of confidence interval for the estimates obtained in section 3b., that is, to repeat this experiment many times and use the variation in the estimates of these experiments to estimate the variation in the estimate of section 3b. This is called a parametric bootstrap method. See Efron and Tibshirani (1993) for more details.

Acknowledgments. The first author's research is supported in part by a startup fund provided by the Pennsylvania State University. Work of the second author is

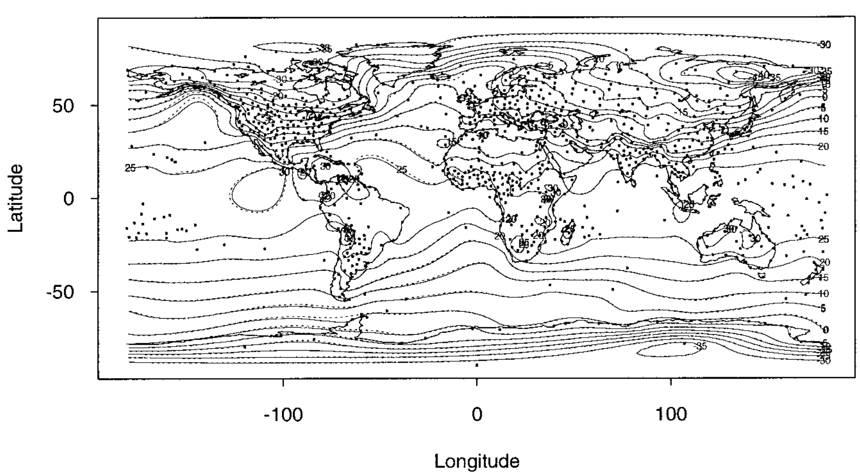

FIG. 7. Estimates of local winter mean temperature $[d_1 + g_2(P)]$. Solid contour lines: estimates using the original data (truth for the simulated data); dashed contour lines: estimates using the simulated data.

supported by NSF Grant DMS-9121003 and NASA Grant NAGW-2961. The second author would like to thank Jon Helgeland for originally suggesting the application of smoothing spline ANOVA to climate data. The authors thank two referees for their helpful comments that have improved the paper significantly.

APPENDIX A

Computational Details of Smoothing Spline Estimates for One-Time Data

It can be proved (Wahba 1981) that the minimizer of (2) has a representation

$$f_\theta(P) = d + \theta \sum_{i=1}^{n} c_i R(P, P_i), \quad \text{(A1)}$$

where d and c are given by

$$\begin{cases} d = 1^T(\theta \mathbf{Q} + \mathbf{I})^{-1} y / 1^T(\theta Q + I)^{-1} 1 \\ c = (\theta \mathbf{Q} + \mathbf{I})^{-1}(y - d1), \end{cases} \quad \text{(A2)}$$

where Q is an n by n matrix with its (i, j)-th element $R(P_i, P_j)$. Here, R is a nonnegative definite function uniquely defined by $\int (\Delta f)^2\, dP$. We will use

$$R(P, P') = \frac{1}{2\pi}\left[\frac{1}{2} q_2(z) - \frac{1}{6}\right], \quad \text{(A3)}$$

where $z = \cos(\gamma(P, P'))$, $\gamma(P, P')$ is the angle between P and P', and

$$q_2(z) = \frac{1}{2}\left\{ \ln\left(1 + \sqrt{\frac{2}{1-z}}\right)\left[12\left(\frac{1-z}{2}\right)^2 - 4\left(\frac{1-z}{2}\right)\right] - 12\left(\frac{1-z}{2}\right)^{3/2} + 6\left(\frac{1-z}{2}\right) + 1 \right\}$$

[see Wahba 1981, (3.3) and (3.4)]. This R does not correspond exactly to $\int_{\mathcal{S}} (\Delta f)^2\, dP$, but a norm topologically equivalent to it. In other words, we are not computing the minimizer of (2), but (2) with a slightly changed penalty term. The reason is computational since R corresponding to $\int_{\mathcal{S}} (\Delta f)^2\, dP$ is too expensive to compute. By the results of Stein (1990), these changes will not make the results much different when sufficient data are available.

It is not difficult to verify that

$$\int_{\mathcal{S}} R(P, P')\, dP = 0, \quad \text{for any } P' \in \mathcal{S};$$

hence $\int_{\mathcal{S}} f_\theta(P)\, dP / \int_{\mathcal{S}} 1\, dP = d$. That is, d is the global average of f_θ. We can also integrate f_θ (numerically) over a region to get an estimate of the average temperature in that region.

APPENDIX B

The Equivalence of Smoothing Spline Averaging Method to the "Statistically Optimum Averaging" Method

Considering the model in (1), Vinnikov et al. (1990) assume that f is a random field over the sphere with a constant mean, say C, and a covariance function $R(P, P')$. They also assume that $\{\epsilon_i\}$ are independent random variables such that $E(\epsilon_i) = 0$ and $\mathrm{Var}(\epsilon_i) = \sigma^2$. The $\{\epsilon_i\}$ are assumed to be independent of f as well. Then the mean squared error of estimating the average of f over a region $\mathcal{K} \subset \mathcal{S}$, $\int_{\mathcal{K}} f\, dP/b$, where $b := \int_{\mathcal{K}} 1\, dP$ by a linear combination $\Sigma_{i=1}^{n} p_i y_i$ of observed data is

$$\begin{aligned} \mathrm{MSE} &= E\left(\int_{\mathcal{K}} f\, dP/b - \sum_{i=1}^{n} p_i y_i\right)^2 \\ &= \left[\mathrm{Var}\left(\int_{\mathcal{K}} f\, dP/b\right) + \sum_i \sum_j p_i p_j R(P_i, P_j) + \sum_{i=1}^{n} p_i^2 \sigma^2 - 2\sum_{i=1}^{n} p_i \Omega_i\right] + \left[C^2\left(1 - \sum_{i=1}^{n} p_i\right)^2\right] \\ &= [\text{variance}] + [\text{bias}], \end{aligned} \quad \text{(B1)}$$

where

$$\Omega_i = \mathrm{Cov}\left[\int_{\mathcal{K}} f\, dP/b,\ f(P_i)\right] = \int_{\mathcal{K}} R(P, P_i)\, dP/b.$$

Restricting estimators to unbiased ones, that is, requiring that the "bias" term in (B1) equals zero, coefficients $\{p_i\}$ must satisfy

$$\sum_{i=1}^{n} p_i = 1. \quad \text{(B2)}$$

With the restriction (B2), the minimizer of the mean square error (MSE) is

$$p = (Q + \sigma^2 I)^{-1}\Omega + (Q + \sigma^2 I)^{-1} 1 \times \frac{1 - 1^T(Q + \sigma^2 I)^{-1}\Omega}{1^T(Q + \sigma^2 I)^{-1} 1};$$

hence the estimate of the average of f over region $\mathcal{K}$ is

$$\begin{aligned} y^T p &= y^T(Q + \sigma^2 I)^{-1}\Omega + y^T(Q + \sigma^2 I)^{-1} 1 \times \frac{1 - 1^T(Q + \sigma^2 I)^{-1}\Omega}{1^T(Q + \sigma^2 I)^{-1} 1} \\ &= \frac{y^T(Q + \sigma^2 I)^{-1} 1}{1^T(Q + \sigma^2 I)^{-1} 1} + \left(y - \frac{y^T(Q + \sigma^2 I)^{-1} 1}{1^T(Q + \sigma^2 I)^{-1} 1} 1\right)^T \times (Q + \sigma^2 I)^{-1}\Omega, \end{aligned}$$

which is exactly $\int_{\mathcal{K}} f_\theta\, dP / \int_{\mathcal{K}} 1\, dP$ with $\theta = 1/\sigma^2$ [(considering (A1) and (A2)]. The smoothing spline approach gives not only the same average estimates of *statistically* *optimum averaging* when parameters are matched appropriately, but also the same grid point estimates of *statistical optimal* interpolation. See, for example, Lorenc (1986) or Wahba (1990).

Vinnikov et al. (1990) used a prescribed $R(P, P')$ and empirically estimated σ^2. In smoothing spline estimates, σ^2 is assumed unknown and to be estimated simultaneously based on the same data. The function R is decided by the form of the penalty term, but other R's, including those estimated from current or past data, may also be used, which makes the smoothing spline approach even closer to the statistical optimal averaging approach.

APPENDIX C

Computational Details of Smoothing Spline Estimates for Multiple Time Data

It can be shown that the minimizer of (7) has a representation similar to (A1):

$$f_\theta(t, P) = d_0 + d_1\phi(t) + \sum_{i=1}^{n} c_i \sum_{\alpha=1}^{4} \theta_\alpha R_\alpha[(t_i, P_i); (t, P)], \quad \text{(C1)}$$

where each R_α is a nonnegative definite function decided by the choice of J_α and given in Table C1 where R_{time} and R_{space} will be described below and $d := (d_0, d_1)^T$ and $c := (c_1, \ldots, c_m)^T$ are the solution of

$$\begin{cases} 0 = S^T c \\ (Q_\theta + I)c = (y - Sd), \end{cases} \quad \text{(C2)}$$

where $y := (y_1, \ldots, y_n)^T$, $\mathbf{S}$ is a $n \times 2$ matrix with ith row $[1, \phi(t_i)]$, $Q_\theta = \Sigma_{\alpha=1}^4 \theta_\alpha Q_\alpha$, and $Q_\alpha = [R_\alpha(x_i, x_j)]_{i,j=1,2,\ldots,n}$. Here, $x_i = (t_i, P_i)$, which is a shorthand for the year, latitude, and longitude of the ith data point. In this article, the nonnegative definite function R_{time} is defined as follows. Let $\mathbf{L}$ be a $(n_t - 2) \times n_t$ matrix:

$$\begin{pmatrix} 1 & -2 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1 & -2 & \cdots & 0 & 0 & 0 \\ & & & \cdots & \cdots & & & \\ 0 & 0 & 0 & 0 & \cdots & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 1 & -2 & 1 \end{pmatrix}. \quad \text{(C3)}$$

Thus, $\mathbf{J}_1(g) = \mathbf{g}^T\mathbf{L}^T\mathbf{L}\mathbf{g}$, where $\mathbf{g} = [g(1), g(2), \ldots, g(n_t)]^T$. Then $R_{\text{time}}(j, j')$ is the jj'th entry of $(L^TL)^\dagger$ where † denotes the Moore–Penrose generalized inverse. In this article R_{space} is given by (A3). It would be possible to substitute other nonnegative definite functions.

Since considering time and space simultaneously means that we are dealing with a huge data size, efficient algorithms are needed to solve (C2). Luo (1996b) developed some algorithms that can handle a few thousand stations and as many years as practically needed in this application (there are only about 150 yr of historical instrumental records of temperature). After c and d are obtained, the components of f_θ and various regional averages can easily be computed.

TABLE C1. The nonnegative definite functions in the representation (C1) of a smoothing spline estimate.

α	$R\alpha$
1	$R_1(t, P; t', P') = R_{\text{time}}(t, t')$
2	$R_2(t, p; t', P') = R_{\text{space}}(P, P')$
3	$R_3(t, P; t', P') = \phi(t)\phi(t')R_{\text{space}}(P, P')$
4	$R_4(t, P; t', P') = R_{\text{time}}(t, t')\,R_{\text{space}}(P, P')$

For choosing smoothing parameters θ's, the degrees of freedom for the component of f_θ mentioned in section 3b are defined as $tr(S_\alpha)$, where $S_\alpha(\theta_\alpha) = (Q_\alpha + (1/\theta_\alpha)I)^{-1}Q_\alpha$, for $\alpha = 1, 2, 3, 4$. Let Q_{time} be a $n_t \times n_t$ matrix with its (i, j) entry being $R_{\text{time}}(i, j)$, and Q_{space} be a $n_s \times n_s$ matrix with its (i, j) entry being $R_{\text{space}}(P_i, P_j)$. Let $\lambda_1^t, \lambda_2^t, \ldots, \lambda_n^t$ be all the eigenvalues of Q_{time}, and $\lambda_1^s, \lambda_2^s, \ldots, \lambda_n^s$ be all the eigenvalues of Q_{space}, then

$$tr(S_1) = \sum_{i=1}^{n_t} \frac{n_s\lambda_i^t}{n_s\lambda_i^t + 1/\theta_1}, \qquad tr(S_2) = \sum_{j=1}^{n_s} \frac{n_t\lambda_j^s}{n_t\lambda_j^s + 1/\theta_2},$$

$$tr(S_3) = \sum_{j=1}^{n_s} \frac{\|\phi\|^2\lambda_j^s}{\|\phi\|^2\lambda_j^s + 1/\theta_3}, \quad \text{and}$$

$$tr(S_4) = \sum_{i=1}^{n_t}\sum_{j=1}^{n_s} \frac{\lambda_i^t\lambda_j^s}{\lambda_i^t\lambda_j^s + 1/\theta_4},$$

where

$$\|\phi\|^2 = \sum_{i=1}^{n_t} \phi(t)^2.$$

REFERENCES

Derber, J., and A. Rosati, 1989: A global oceanic data assimilation system. *J. Phys. Oceanogr.*, **19,** 1333–1347.

Diaz, H. F., R. S. Bradley, and J. K. Eischeid, 1989: Precipitation fluctuations over global land areas since the late 1800's. *J. Geophys. Res.*, **94,** 1195–1210.

Efron, B., and R. J. Tibshirani, 1993: *An Introduction to the Bootstrap.* Chapman and Hall, 436 pp.

Girard, D., 1989: A fast 'Monte Carlo cross-validation' procedure for large least squares problems with noisy data. *Numer. Math.*, **56,** 1–23.

Gu, C., and G. Wahba, 1993a: Semiparametric analysis of variance with tensor product thin plate splines. *J. Roy. Stat. Soc. Ser. B,* **55,** 353–368.

——, and ——, 1993b: Smoothing spline ANOVA with component-wise Bayesian confidence intervals. *J. Comput. Graphical Stat.*, **2,** 97–117.

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

Hegerl, G. C., H. von Storch, K. Hasselmann, B. D. Santer, U. Cubasch, and P. D. Jones, 1996: Detecting greenhouse gas induced climate change with an optimal fingerprint method. *J. Climate*, **9,** 2281–2306.

Jones, P. D., 1994: Hemispheric surface air temperature variations: A reanalysis and update to 1993. *J. Climate,* **7,** 1794–1802.
——, T. M. L. Wigley, and P. M. Kelly, 1982: Variations in surface air temperatures: Part 1. Northern Hemisphere, 1881–1980. *Mon. Wea. Rev.,* **110,** 59–72.
——, S. C. B. Raper, R. S. Bradley, H. F. Diaz, P. M. Kelly, and T. M. L. Wigley, 1986: Northern Hemisphere surface air temperature variations: 1851–1984. *J. Climate Appl. Meteor.,* **25,** 161–179.
——, ——, B. S. G. Cherry, C. M. Goodess, T. M. L. Wigley, B. Santer, P. M. Kelly, R. S. Bradley, and H. F. Diaz, 1991: An updated global grid point surface air temperature anomaly data set: 1851–1988. Environmental Sciences Division Publ. 3520, U.S. Department of Energy, 411 pp. [Available from the National Technical Information Service, U.S. Dept. of Commerce, 5285 Port Royal Rd., Springfield, VA 22161.]
Karl, T. R., R. W. Knight, and J. R. Christy, 1994: Global and hemispheric temperature trends: Uncertainties related to inadequate spatial sampling. *J. Climate*, **7,** 1144–1163.
Kimeldorf, G., and G. Wahba, 1971: Some results on Tchebycheffian spline functions. *J. Math. Anal. Appl.,* **33,** 82–95.
Knight, R. W., 1980: A comparison of some methods for flagging erroneous observations in certain types of meteorological data. Department of Statistics, University of Wisconsin—Madison. Tech. Rep. 610, 16 pp. [Available from Department of Statistics, University of Wisconsin—Madison, 1210 W. Dayton St., Madison, WI 53706.]
Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.,* **112,** 1177–1194.
Luo, Z., 1996a: Backfitting in smoothing spline ANOVA with application to historical global temperature data (thesis). Department of Statistics, University of Wisconsin–Madison, Tech. Rep. 964. [Available on-line from http://www.stat.psu.edu/~zhen.]
——, 1996b: Backfitting in smoothing spline ANOVA. Department of Statistics. The Pennsylvania State University, Tech. Rep. 96-05. [Available on-line from http://www.stat.psu.edu/~zhen.]
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.
Nicholls, N., G. V. Gruza, J. Jouzel, T. R. Karl, L. A. Ogallo, and D. E. Parker, 1996: Observed climate variability and change. *Climate Change 1995: The Science of Climate Change,* J. T. Houghton, L. G. Meira Filho, B. A. Callander, N. Harris, A. Kattenberg, and K. Maskell, Eds., Cambridge University Press, 134–192.
Peixoto, J. P., and A. H. Oort, 1992: *Physics of Climate.* American Institute of Physics. 520 pp.
Stein, M., 1990: Uniform asymptotic optimality of linear predictions of a random field using an incorrect second-order structure. *Ann. Stat.,* **18,** 850–872.
Thiebaux, H. J., 1991: Statistics and the environment: The analysis of large-scale earth-oriented systems. *Environmetrics*, **2,** 5–24.
Vinnikov, K. Ya., G. V. Gruza, V. F. Zakharov, A. A. Kirillov, N. P. Kovyneva, and E. Ya. Ran'kova, 1980. Contemporary variations of the Northern Hemisphere climate (in Russian). *Meteor. Gidrol.*, **6,** 5–17.
——, P. Ya. Groisman, and K. M. Lugina, 1990: Empirical data on contemporary global climate changes (temperature and precipitation). *J. Climate,* **3,** 662–677.
Wahba, G., 1981: Spline interpolation and smoothing on the sphere. *SIAM J. Sci. Stat. Comput.,* **2,** 5–16; Erratum, **3,** 385–386.
——, 1990: *Spline Models for Observational Data.* CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. 59, Society of Industrial and Applied Mathematics, 169 pp.
——, and J. Wendelberger, 1980: Some new mathematical methods for variational objective analysis using splines and gross validation. *Mon. Wea. Rev.,* **108,** 1122–1143.
——, D. R. Johnson, F. Gao, and J. Gong, 1995: Adaptive tuning of numerical weather prediction models: Randomized GCV in three-and four-dimensional data assimilation. *Mon. Wea. Rev.,* **123,** 3358–3369.
Yamamoto, R., and M. Hoshiai, 1980: Fluctuations of the Northern Hemisphere mean surface air temperature during recent 100 years, estimated by optimum interpolation. *J. Meteor. Soc. Japan,* **58,** 187–193.