Spatial Correlations in Station Temperature Anomalies

Richard F. Gunst
Department of Statistical Science
Southern Methodist University
Dallas, TX 75275-0332

Abstract

The statistical validity of uncertainty estimates calculated from station temperature anomalies and of inferences drawn on trends in temperature anomalies require that spatial correlations among the anomalies be properly characterized in the statistical methods that are used. Adequate characterization of these spatial correlations necessitates the explicit use of the strength of anomaly dependence as a function of station separation distances in the statistical modeling of anomalies. In this paper the estimation of spatial correlations for station temperature anomalies is discussed. Sample semivariograms are used to characterize anomaly correlations as a function of station separation distances. The nugget, sill, and range parameters of the models identify key features of the correlation structure. Application of this modeling technique is demonstrated by fitting Gaussian semivariogram models to 1990 station anomaly data for the contiguous United States and for Europe.

1. Introduction

Statistical analyses of instrumental temperature anomaly data generally involves estimation or inference procedures that are based on an assumption that the individual station anomalies are statistically independent. It is well recognized that this assumption is unreasonable for station anomaly data because of their spatial character:

anomalies from stations in close proximity to one another can be expected to be highly correlated while those from stations more distant from one another can be expected to be less correlated. The importance of recognizing the existence of and for properly accounting for spatial correlations is clearly evident in the need for statistically valid measures of uncertainty for mean anomaly estimates. Among the consequences of ignoring spatial or temporal correlations among anomalies are that the usual estimators of mean anomalies are inefficient and estimators of uncertainty can be badly biased.

Standard deviation estimates have often been used to measure the uncertainty in estimated temperature anomalies, even though the the spatial and temporal correlations of the anomalies render their use questionable. Jones and Briffa (1992), in an excellent survey of twentieth-century instrumental temperature variation patterns, use sample standard deviations to compare monthly variation in hemispheric temperature anomaly averages from 1901 to 1990. The standard deviations calculated for each hemisphere and for the globe using land-only station anomalies and using land and marine data were approximately 0.20°C. However, these sample standard deviations do not account for temporal correlations that induce trends in the anomaly series. While the relative magnitudes of the calculated standard deviations may be comparable -- and comparison of the standard deviations was a primary purpose of the calculations -- the standard deviations themselves are likely to be biased as estimates of the true annual variations.

Hansen and Lebedeff (1987) also calculate standard deviations to provide uncertainty estimates. Theirs are calculated from a 100-year simulation of a general circulation model in order to place approximate error bars on their annual hemispheric and global anomaly series. The standard deviations calculated for the global averages decreased from approximately 0.07 in the 1980s to 0.02 in the 1960s. As the authors point out, these estimates from the general circulation model do not account for a number of sources of error and are used only as a "nominal measure of the expected error."

Wigley, Jones, Kelly, and Raper (1989; see also Wigley and Jones 1981) use standard deviations of the annual anomalies to compare anomalies year-by-year. Then to assess the possibility that changes on anomalies are greenhouse-induced, they compare decadal averages. In this comparison, they assume that annual temperature anomalies can be well represented by a first-order autoregressive model and adjust the estimated standard error of the decadal averages:

$$\hat{se}(\hat{a}_d) = \{(1+r)/(1-r)\}^{1/2}\, s/n^{1/2} \tag{1.1}$$

where $\hat{a}_d$ is a decadal average anomaly, s is the standard deviation of the annual anomalies over a 1951 to 1980 reference period, r is the lag-1 estimated autocorrelation calculated over the entire record, and $n = 10$. Using the estimated values of $r = 0.7$ and $s = 0.063$, an annual estimated standard error using (1.1) with $n = 1$ would be $\hat{se}(\hat{a}_t) =$ 0.15°C. This estimate is smaller than those of Jones and Briffa, presumably because the standard deviation s is only calculated over the reference period and because temporal correlation is partially accounted for in the autoregressive model that leads to (1.1).

Solow (1988) used a robust method to estimate the uncertainty in southern hemisphere annual temperature anomalies. The surprising result of this investigation was that robust estimates of variability were essentially constant over the 128 years of data. Since there are many more stations in the late 20th century than in the late 19th century, it was expected that variability estimates would be smaller in the later period of the record. The most likely explanation offered for the constancy of the uncertainty estimates across the 128 years was the lack of any accounting for spatial correlations among the stations.

In addition to estimates of uncertainty, other traditional statistics also are used to characterize statistical properties of anomalies. Hansen and Lebedeff (1987), Briffa and Jones (1992), and others correlate annual temperature series for individual stations or for box or grid averages in order to describe spatial correlations. Their calculations use ordinary correlation (Pearson's r) estimates for each pair of stations. These calculations do not account for the temporal correlations in the individual anomaly series. Moreover, these calculations implicitly assume that the spatial correlations between pairs of stations are constant over the entire time span of the data.

The need for improved spatial and temporal modeling arises not only from the violation of the key assumption of independent anomalies, but also from data analytic difficulties that can occur from the use of the traditional statistics when the independence assumption is violated. For example, Hansen and Lebedeff's correlations for stations in the southern hemisphere do not approach unity as station distances approach zero. Briffa and Jones model correlations between grid box average anomalies using a simple exponential fit; however, they must ignore zero and negative correlations in the fitting procedure used. Both of these papers provide important first approximations to spatial modeling. They also highlight the need for more sophisticated spatial modeling techniques.

The focus of this paper is on spatial modeling of temperature anomalies. In Section 2, some of the consequences of improper spatial modeling are detailed for the estimation of means and variances. In particular, it is shown that the usual estimator of the standard error of an average is seriously biased when anomalies are spatially correlated. In Section 3 the spatial dependence of anomalies is characterized through the use of sample semivariograms. The fitting of sample semivariograms using Gaussian semivariogram models is discussed in Section 4, along with the subsequent calculation of spatial correlations from the fitted models. The modeling of anomaly correlations is illustrated for 1990 station temperature anomalies from the contiguous United States and from Europe. Concluding remarks are made in Section 5.

2. Consequences of Spatial Correlation

To illustrate the effects that spatial correlations can have on statistical estimates of spatial means and standard deviations, fix time (t) and select a spatial region (e.g., a grid box) in which error-free temperature anomalies can be considered constant at all locations in the region. Measured anomalies a_t are not error free. A simple statistical model that can be used to describe the measured anomalies is $a_t(p_i) = \delta_t + e_i$, where $a_t(p_i)$ is the anomaly recorded at time t for spatial location p_i (e.g., p_i could represent the latitude and the longitude of the ith anomaly), and e_i represents a mean-zero error. The mean anomaly δ_t is the constant value across the region and is the quantity to be estimated.

Under the conditions just described, it is well known (e.g., Aitken 1935; Bibby and Toutenberg 1977, Section 3.4; Theil 1971, Section 6.1) that the optimal (minimum variance linear unbiased; minimum variance unbiased if the errors are normally distributed) estimator of the mean δ_t is

$$\hat{\delta}_t = \sum_i \sum_j w_{ij}\, a_t(p_i) \;/\; \sum_i \sum_j w_{ij} \tag{2.1}$$

where the weights w_{ij} are elements of the inverse of the covariance matrix of the errors; i.e., $W = [w_{ij}] = \Sigma^{-1}$ and $\Sigma = [\sigma_{ij}]$ with $\sigma_{ij} = \mathrm{cov}\{e_i,e_j\}$. An alternative expression for this estimator is $\hat{\delta}_t = \sum_i \omega_i\, a_t(p_i)$ where $\omega_i = \sum_j w_{ij} \,/\, \sum_i \sum_j w_{ij}$ and $\sum_i \omega_i = 1$.

The estimator (2.1) reduces to two well-known estimators under appropriate conditions on the the errors e_i. If the errors have a common variance, $\sigma_{ii} = \sigma^2$ for all i, and are uncorrelated, $\sigma_{ij} = 0$ for $i \neq j$, then $w_{ii} = \sigma^{-2}$, $w_{ij} = 0$ $(i \neq j)$ and $\hat{\delta}_t = \bar{a}_t$, the average of the anomalies. If the errors have different variances in the region but are uncorrelated,

$$\hat{\delta}_t = \sum_i w_i \, a_t(p_i) \; / \; \sum_i w_i \tag{2.2}$$

where $w_i = w_{ii} = \sigma_{ii}^{-1}$.

The ordinary average can be considered a weighted estimator with all weights equal. From (2.1), the ordinary average is optimal as an estimator of a spatial mean only if the variances of the anomalies for all stations in a region can be considered constant. Inverse- (i.e., with $w_i = d_i^{-1}$) and linear- (i.e., with $w_i = r - d_i$ for a region of radius r) distance-weighted estimators are optimal only if the variances of the anomalies have a reciprocal relationship to the distance weights. Also, the optimality of both the ordinary average and distance-weighted averages requires that anomalies in a region be uncorrelated. If anomalies are correlated, neither ordinary averages nor distance-weighted averages are optimal for estimating the spatial mean anomaly for the region.

Major consequences of violating the assumptions imposed on the optimality of (2.1) or its special cases are that the resulting estimators of the mean anomaly are inefficient relative to the optimal estimator and that standard deviations are biased estimators of uncertainty. Under the constant mean model $a_t(p_i) = \delta_t + e_i$ with the e_i having mean 0, all weighted estimators of the form (2.2) are unbiased estimators of the mean δ_t if the weights do not depend on the values of the anomalies. Both the ordinary average and distance-weighted averages satisfy this requirement. The variances of the estimators, however, are only correct if the weights are the optimal weights for the errors. In particular, if the errors are correlated the usual sample variance is an unbiased estimator of a weighted average of the true anomaly variances. In addition, $s/n^{1/2}$ is not the correct standard error estimator for either the average anomaly $\bar{a}$ or for a distance-weighted average. While not the focus of this paper, $s/n^{1/2}$ also is not the correct estimator for temporally correlated anomalies (e.g., Fuller 1976, Chapter 6).

As a specific illustration of the effect that spatially correlated errors can have on mean and variance estimation, suppose that in a region the correlation between the anomalies from any pair of stations is a constant value, $\rho > 0$, and the anomalies from all the stations have the same variance σ^2; i.e., suppose $\mathrm{var}\{a_t(p_i)\} = \sigma^2$ and $\mathrm{cov}\{a_t(p_i),a_t(p_j)\} = \rho\sigma^2$. Then $\overline{a}_t$ is an unbiased estimator of δ_t, but the standard error of $\overline{a}_t$ is

$$\mathrm{se}(\overline{a}_t) = (\sigma/n^{1/2}) \bullet \{1+(n-1)\rho\}^{1/2}, \tag{2.3}$$

not $\sigma/n^{1/2}$. The optimal estimator (2.1) has standard error

$$\mathrm{se}(\hat{\delta}_t) = (\sigma/n^{1/2})/\{1+(n-1)\rho\}^{1/2}. \tag{2.4}$$

In addition, the expected value of the sample variance s^2 is $\sigma^2(1-\rho)$, not σ^2. Thus, if $\rho > 0$, use of the sample variance in standard error calculations will lead to an overly optimistic estimate of the uncertainty of unweighted averages. In other words, $s/n^{1/2}$ will tend to be too small as an estimator of the standard error of the average anomaly both because $\sigma/n^{1/2}$ is smaller than the correct standard error (2.3) and because the sample variance is biased downward by a factor of $(1-\rho)$. Whether $s/n^{1/2}$ is an overestimate or underestimate of the standard error of distance-weighted averages depends on the weights actually used; nevertheless, it will in general be a biased estimator.

Similar difficulties arise with other assumptions about the spatial correlations. If the errors are assumed to have a first-order autoregressive structure, $e_i = \rho e_{i-1} + v_i$ with $\mathrm{var}(v_i) = \sigma^2$, then (ignoring terms of smaller order than $n^{-1/2}$ and high-order terms in ρ) the standard error of the sample mean is approximately (e.g., Johnston 1963, Section 7.3)

$$\mathrm{se}(\overline{a}_t) = (\sigma_e/n^{1/2})\{(1-\rho)/(1+\rho)\}^{1/2}, \tag{2.5}$$

where $\sigma_e = \mathrm{var}(e_i)^{1/2} = \sigma/(1-\rho^2)^{1/2}$. The expected value of the sample variance is approximately $\sigma_e^2(1+\rho)/(1-\rho)$.

The above examples suffice to demonstrate the theoretical concerns that arise when spatial correlations are ignored. The degree to which the spatial correlations must be accounted for in the actual calculation of spatial mean temperature anomalies is currently unknown. The need for efforts to address this issue can be effectively seen by demonstrating the existence of strong spatial correlations in anomaly data. In the next section, the existence of strong spatial correlations is shown for data from stations in the contiguous United States.

3. Sample Semivariograms for Temperature Anomaly Data

Figure 1 displays the locations of the 100 temperature stations operating during 1990 that were located between 30° and 50° north latitude and between 70° and 130° west longitude, roughly within the contiguous United States. All but two of these stations have monthly average temperatures for all 12 months of 1990. Superimposed on the map are 5° latitude by 10° longitude grid boxes. It is clear from the figure that the stations are irregularly located both within the United States and within the grid boxes. It is this irregularity that has caused the concern over how to estimate spatial mean anomalies and how to properly quantify uncertainty in the estimates.

Great circle distances were calculated between all pairs of the 100 U.S. stations. Anomalies were calculated using 1951 to 1980 as the reference period. Figure 2 displays scatterplots of the January anomaly pairs for stations within 1,000 km of one another, grouped according to the distance separating them: distance $\leq$ 250 km, 250 km $<$ distance $\leq$ 500 km, 500 km $<$ distance $\leq$ 750 km, and 750 km $<$ distance $\leq$ 1,000 km. This grouping was chosen simply for illustrative purposes. For each pair of stations, the station with the smaller World Meteorological Organization (WMO) identification number was assigned to the horizontal axis; the station with the larger WMO number was assigned to the vertical axis. The weakening of the correlational bonds between the station anomalies as a function of separation distance is clear from a comparison of the four plots. As a simple descriptive indicator of the relative strengths of the dependence between station anomalies, the (Pearson's r) correlation for the pairs of stations within 250 km is 0.82, for those between 250 and 500 km is 0.71, for those between 500 and 750 km is 0.56, and for those between 750 and 1,000 km is 0.34. Figure 3 displays a similar plot for the May anomaly pairs. The weakening of the correlational bonds with the distance separating the stations is again clear from the plots and from the calculated correlations: 0.57, 0.53, 0.29, and 0.02, respectively. The stronger correlation patterns for January than for May is also clear from a comparison of Figures 2 and 3.

As descriptive statistics, correlations are useful for providing a rough indication of the dependence between station anomalies. They are less satisfactory for inferential purposes because there is no natural way to assign stations as the first and second members of station pairs (or to vertical and horizontal axes in plots such as Figures 2 and 3). Also, sample covariances and correlations can have substantial biases when stations are spatially or temporally autocorrelated (e.g., Fuller 1976, Section 6.2). For modeling and inferential purposes, semivariogram calculations are preferable for describing spatial dependence (e.g., Cressie 1989, Section 2.4).

A theoretical semivariogram is half the variance of the difference between two anomalies: $\gamma(p_1,p_2) = \mathrm{var}\{a_t(p_1) - a_t(p_2)\}/2$. Under appropriate conditions, semivariograms are related to spatial correlations, when both exist. To establish this relationship, let $c(p_1,p_2) = \mathrm{cov}\{a_t(p_1),a_t(p_2)\}$. Then the spatial correlation between two anomalies is $\rho(p_1,p_2) = c(p_1,p_2)/c(0)$, where $c(0) = \mathrm{var}\{a_t(p_i)\}$ is assumed constant across station locations in a region. Under spatial second-order stationarity assumptions on the anomalies, the semivariogram, the covariance, and the correlation for anomalies at two locations are functions only of the distance, denoted $p_1 - p_2$, between the locations. These quantities then can be related as follows:

$$\gamma(p_1 - p_2) = c(0) - c(p_1 - p_2) = c(0)\{1 - \rho(p_1 - p_2)\} \ . \tag{3.1}$$

From this expression, the semivariogram value is seen to be small when the spatial correlation $\rho(p_1 - p_2)$ is close to 1 and relatively large, close to the spatial variance $c(0)$, when the spatial correlation is close to 0. When the spatial correlation is zero, the semivariogram value equals $c(0)$, the variance of an individual anomaly.

The calculation of semivariogram values can be performed on the individual squared anomaly differences $\{a_t(p_i) - a_t(p_j)\}^2$ or on spatially grouped squared anomaly differences. Method of moment or robust semivariogram estimators can be calculated (e.g., Cressie 1989, Section 2.4). In part due to the effect of outliers on semivariogram calculations, squared anomaly differences are usually grouped or "binned" based on the distances separating the stations. If stations are binned by the distance separating them, decisions must be made on the number of bins to use and on the distances/directions between the stations that will be placed in each bin. For the purposes of this paper, a very simple approach was taken to these alternatives. The 4,950 pairs of 1990 U.S. station anomalies were isotropically (without regard to direction) binned into 100 km "lags"; i.e., all stations whose separation distance was less than or equal to 100 km formed the first lag, those between 100 and 200 km formed the second lag, etc. This choice of the number and size of the bins was somewhat arbitrary, but was based on a number of analyses using different bin sizes.

Figure 4 shows plots of the 1990 U.S. monthly sample semivariograms, where each semivariogram plot is calculated using the 10 bins of stations whose distances are within 1,000 km of one another. For lag m, containing the n_m pairs of stations whose distances are between $100(m-1)$ km and $100m$ km, the sample semivariogram value is calculated by taking half the average of the squared anomaly differences for all pairs of stations in the mth bin:

$$\hat{\gamma}_m = \sum_{\text{bin } m} \{a_t(p_i) - a_t(p_j)\}^2 \; / \; (2n_m) \; . \qquad (3.2)$$

The limit of 1,000 km initially was chosen as a compromise between the desire for sensitivity to local change and the desire for large-area representativeness. It is currently unknown over how great a distance spatial correlation modeling of temperature anomalies can be considered appropriate. Ideally, spatial correlation modeling would be conducted over small regions, thereby allowing the maximum sensitivity to local conditions. There are, however, insufficient data over much of the globe to perform spatial correlation modeling over small regions. Hence, distances of the order of 1,000 km or more are likely to be used in any comprehensive modeling efforts of the entire globe. On the other hand, erratic sample semivariogram values (outliers) sometimes occur due to the occasional presence of very large or small average anomalies (relative to averages for neighboring bins) for lags representing stations that are great distances apart. Use of 30 to 40 lags did result in both semivariogram outliers and in the loss of sensitivity to changes in the curvature of the semivariogram plots. Cressie and Hawkins (1980) robust semivariogram calculations initially were examined as an alternative to the averaging in (3.2). This method did not result in substantive changes in the results presented below and did not overcome the lack of sensitivity to changes in the curvature in the variograms when large numbers of lags were included. Further examination of this approach is, however, warranted.

The semivariogram plots shown in Figure 4 all initially are small and increase at different rates. From equation (3.1), the plots for which the semivariograms rise sharply suggest that the spatial correlations decrease relatively rapidly across the distance plotted. Those for which the semivariogram plots are fairly flat suggest slowly changing spatial correlations, but this may be deceiving due to the differences in the sills (the plateaus at the upper ends of the plots -- see Section 4). While the spatial correlations thus appear to change most rapidly in the winter months and most slowly in the summer months, it is difficult to draw substantive conclusions about spatial correlations due to the differences in the sills.

Rather than discussing further the sample semivariogram and its implications for estimates of spatial correlations, statistical models of variogram change are now fit to the sample data. The rationale and procedures for doing so are the topic of the next section.

4. Semivariogram Modeling

Statistical models are fit to sample semivariograms for three important reasons. First, the fitted semivariogram models are smoothed versions of the sample semivariograms for which the parameter estimates in the fitted model describe key characteristics of the spatial correlation structure. Second, fitted semivariogram models can be used in spatial kriging models to predict anomalies for any location in a region, to estimate the mean anomaly for the region, and to provide estimates of uncertainty for the predictions and mean estimates. Third, sample semivariograms usually cannot be used with kriging models because the associated semivariogram matrices are not negative definite.

The use of kriging models requires that a semivariogram matrix be negative definite or, equivalently, that the covariance matrix of the anomalies be positive definite (e.g., Journel and Huijbregts 1978, Section II.A.3). In addition to kriging modeling, this requirement is also a standard assumption for most statistical modeling techniques. It is equivalent in univariate modeling that estimates of variances be nonnegative. Sample semivariogram matrices calculated directly from differences in squared anomalies generally are not negative definite. On the other hand, semivariogram matrices formed from the fitting of appropriately selected semivariogram models can be guaranteed to be negative definite by appropriate choice of the semivariogram model.

Based on comparisons of the fits of several different variogram models to the sample semivariograms in Figure 4, Gaussian semivariogram models were selected. The Gaussian model was selected because it can capture an S-shaped curvature that is apparent in many of the sample semivariograms that were plotted as part of this study. In many of the semivariogram plots the curvature is initially concave, then convex, and finally flat. Other semivariogram models investigated did not fit these sample semivariograms as well as did the Gaussian model.

The Gaussian semivariogram is a valid semivariogram in that the fitted semivariogram will result in a semivariogram matrix that is negative definite. The form of the Gaussian semivariogram model is

$$\gamma(\mathrm{h}) = \theta_1 + (\theta_2 - \theta_1)[1 - \exp\{-(\mathrm{h}/\theta_3)^2\}] , \tag{4.1}$$

where h is the great-circle distance between two station locations, θ_1 is the "nugget" parameter, θ_2 is the "sill" parameter, and θ_3 is the "range" parameter. Figure 5 is a schematic diagram of a Gaussian semivariogram model.

The nugget in the Gaussian model describes the common phenomenon of a semivariogram being zero at the origin, but not approaching zero because of measurement errors and "microscale" variation as distances between stations become small (e.g., Cressie 1989, Sections 2.3 and 3.2). An estimated nugget is the estimate of the component of the variance of a temperature anomaly that is due to measurement errors and local effects. These measurement errors and local effects occur for individual stations and are not directly estimable, other than by the nugget parameter in the model, because the data base does not have replicate temperature measurements for each station each month.

The sill is the maximum of the semivariogram. It occurs when the distance between stations is so large that the spatial correlation between them is zero. At distances for which this occurs, $\gamma(\mathrm{h}) = c(0) = \theta_2$, and the sample semivariogram plateaus. The sill is the total variance of a temperature anomaly that is due to measurement errors, local effects, regional effects (e.g. due to changes in terrain that cause anomalies to vary), and all other sources of variability.

The range is the distance where the sill is effectively reached; i.e., the distance at which there is effectively no spatial correlation between stations. For some semivariogram models, the sill actually does reach a maximum at the value of the range parameter. For others, including the Gaussian, a maximum is never actually reached but an effective sill can be defined where any remaining increases in the semivariogram can be ignored. For Gaussian models, Journel and Huijbregts (1978, p. 165) choose the range for a Gaussian model to be $\theta_3\sqrt{3}$.

Gaussian models were fit to the sample semivariograms in Figure 4 using a Gauss-Newton nonlinear least squares algorithm (PROC NLIN, SAS 1989). There are a number of computational difficulties (e.g., Dietrich 1991, Warnes and Ripley 1987) which were encountered in the fitting of the Gaussian semivariogram model, but they are beyond the scope of this paper and will be detailed in forthcoming work. A summary of the fitting procedure is as follows.

Initial values for the parameters were calculated using (a) the median of the first three semivariogram values for the nugget, (b) the median of the last four semivariogram values for the sill, and (c) using the lag m for which $\hat{\gamma}_m$ was closest to $\hat{\theta}_1 + (\hat{\theta}_2 - \hat{\theta}_1)\{1 - \exp(-1)\}$, which is obtained from (4.1) by setting $h = \theta_3$ and inserting the initial estimates of the nugget and the sill. Fully iterated estimates of the semivariogram parameters using these initial values compared well to those selected from an extensive grid search of possible initial values. The advantage of using these initial values is the reduction in computations that will be possible when fits are obtained to each month of each year for many regions of the globe.

Once initial estimates were obtained, one step of the Gauss-Newton iteration algorithm was used to compute final estimates. As is common with nonlinear fitting, fully iterated estimates of the model parameters tend to be highly correlated due to collinearities among the first derivatives of the nonlinear function. These collinearities tend to force the estimated to have unreasonable magnitudes; e.g., very large sills and ranges. One-step iterates are a compromise. They usually fit the sample semivariograms as well as the fully iterated estimates but do not suffer from the problems of collinear first derivatives as do the fully iterated estimates. One-step nonlinear iterate estimates have been used in other settings and shown to possess good statistical properties (e.g., Pregibon 1981, Wellman and Gunst 1991).

The ability to achieve satisfactory fits with the one-step Gauss-Newton algorithm is illustrated by Figure 6, which shows all the monthly fits superimposed on the sample semivariograms. From these model fits, the spatial correlation plots in Figure 7 were obtained by inverting equation (3.1). Figure 7 clearly shows the strong spatial correlations in the station anomalies for each month of the year.

Figure 8 shows Gaussian semivariogram fits to sample semivariograms having 20 lags, a maximum range of 2,000 km. These calculations were performed because of a concern that the 1,000 km range might not be sufficient to capture the full extent of the spatial correlations. It is clear from Figure 8 that some of the sample semivariograms (January-April, July, September, December) in Figure 4 indeed had not reached their sills within the maximum range of the plot, 1,000 km, while others had (May, June, August, October, November). With the greater range in Figure 8, the spatial correlations for the months that had not reached their sills within 1,000 km are greater in Figure 9 than in Figure 7. Moreover, the estimated spatial correlations shown in Figure 9 are more seasonally consistent than those in Figure 7, with the strongest spatial correlations occurring in the winter and early spring and the weakest ones in summer and early fall, which is in general agreement with the findings of Briffa and Jones (1993) for the northern hemisphere. Although some of the semivariograms in Figure 8 do not reach their sills within 2,000 km, further discussion of this issue will not be pursued here. The 2,000 km limit suffices to distinguish general patterns in the spatial correlations and is reasonably consistent values used in other studies (e.g., Briffa and Jones 1993; Haas 1990).

The estimated Gaussian model parameters in Table 1 provide further insight into characteristics of the spatial correlations. The largest estimated local variability, as estimated by the nuggets, occurs from late fall through the middle of spring, November through March. The greatest total variability, as measured by the estimated sills, occurs over roughly this same period, from December through April. The greatest spread of strong spatial correlations, as measured by the ranges, occurs from December through April, although the ranges are nearly as large in July and September. The estimation of the range fo a semivariogram model is the most problematic of the semivariogram model parameters, as noted by Mardia and Watkins (1989) and Dietrich (1991), so the range values in Table 1 should be regarded as preliminary estimates.

To assess the potential for the use of Gaussian semivariogram models for global fitting, a similar analysis was performed on stations located in Europe from 10° west to 40° east longitude and from 40° to 60° north latitude. Table 2 shows the estimated model parameters. Plots of the sample semivariograms and the fitted Gaussian semivariograms (not shown) reveals that the Gaussian model well captures the trends in the sample semivariograms and that the overall trends are generally similar to those of the U. S. The nugget values in Table 2 appear to be more seasonally uniform than those for the U. S. in Table 1 and of slightly smaller magnitudes. On the other hand, the ranges appear to be less seasonally uniform, with the largest values occurring in the spring and summer months. The lack of clear seasonality in the parameter estimates may be due to how large a geographic area has been modeled or to the need to model European semivariograms anisotropically. This question requires further investigation.

5. Concluding Remarks

The spatial correlation modeling presented in this paper is an initial effort to provide a structured method for fitting spatially correlated temperature anomalies. A number of topics must be investigated before the method presented in this paper reliably can be used to estimate regional and global mean temperature anomalies. For example, only isotropic semivariogram estimation was used in the model fits reported in the last section. Investigations are needed to determine whether these results can be improved using anisotropic semivariograms and, if so, how the pattern of anisotropy changes throughout the globe. As highlighted by the fits to the European data, alternatives to the Gaussian semivariogram model may be needed for other regions of the globe, although the present work is suggestive that the Gaussian model may be useful over large portions of the globe. It is noteworthy that the exponential and spherical semivariogram models that are used so prevalently in mining and the geosciences were not as effective as the Gaussian semivariogram model in the fitting of sample semivariograms for temperature anomalies reported in Section 4.

The issue of lag distance and the number of lags to use also awaits careful scrutiny. The lag length of 100 km and the maximum distance of 2,000 km that were used in this work may need modification for different parts of the globe, as in Briffa and Jones (1993), or as more knowledge is gained about the relative roles of local sensitivity and global representativeness. Lag selection and model estimation procedures that are not labor or computationally prohibitive are needed for the fitting of semivariograms and, ultimately, for kriging model fits to monthly anomalies in each region of the globe over the many years of instrument data that are currently available.

The analyses reported in this paper used data from the United States and Europe. These regions were chosen because they have sufficient data to investigate modeling strategies. Careful attention must be given to the modeling of regions of the globe where data are more sparse. If the modeling of many regions of the globe consistently result in the use of the same model, it may be appropriate to apply a common model throughout the globe. Likewise if parameter estimates for the models are seen to vary smoothly for contiguous regions of the globe, it may be possible to infer parameter values for regions of the globe for which little or no information is available on temperature anomalies.

The method presented in this paper represents a major advancement over current methods of quantifying spatial correlation patterns of temperature anomalies. Semivariogram model fitting has been widely used in the mining industry and geoscience research for over 40 years, but has not been widely applied outside of these fields. Once the semivariogram issues mentioned above are fully addressed, kriging methods can be used to estimate grid-box, regional, hemispheric, and global mean temperature anomalies. These methods can also provide valid estimates of uncertainty.

Acknowledgements

This research was supported by Contract DE FG05 90ER61015 from the U.S. Department of Energy. Appreciation is expressed to P. D. Jones and T. M. L. Wigley, who graciously supplied the raw station temperature data used in this work. Insightful comments on an earlier draft of this paper by P. D. Jones are also gratefully acknowledged.

References

Aitken, A.C. (1935). "On Least Squares and Linear Combination of Observations," *Proceedings of the Royal Society of Edinburgh*, 55, 42-48.

Bibby, J. and Toutenberg, H. (1977). *Prediction and Improved Estimation in Linear Models.* New York: John Wiley and Sons, Inc. pp. 188.

Briffa, K.R. and Jones, P.D. (1993). ""Global Surface Air Temperature Variations During the Twentieth Century: Part 2, Implications for Large-Scale High-Frequency Palaeoclimatic Studies," *The Holocene*, 3, 77-88.

Cressie, N.A.C. (1989). *Statistics for Spatial Data.* New York: John Wiley and Sons, Inc. pp. 900.

Cressie, N.A.C. and Hawkins, D. M. (1980). "Robust Estimation of the Variogram: I," *Mathematical Geology*, 12, 115-125.

Dietrich, C.R. (1991). "Modality of the Restricted Maximum Likelihood for Spatial Gaussian Random Fields," *Biometrika*, 78, 833-839.

Fuller, W.A. (1976). *Introduction to Statistical Time Series.* New York: John Wiley and Sons, Inc. pp. 470.

Haas, T.C. (1990). "Lognormal and Moving Window Methods of Estimating Acid Deposition," *Journal of the American Statistical Association,* 85, 950-963.

Hansen, J. and Lebedeff, S. (1987). "Global Trends of Measured Surface Air Temperature," *Journal of Geophysical Research*, 92, 13345-13372.

Johnston, J. (1963). *Econometric Methods.* New York: McGraw-Hill Book Company. pp. 300.

Jones, P.D. and Briffa, K.R. (1992). ""Global Surface Air Temperature Variations During the Twentieth Century: Part 1, Spatial, Temporal and Seasonal Details," *The Holocene*, 2, 165-179.

Journel, A.G. and Huijbregts, Ch.J. (1978). *Mining Geostatistics.* New York: Academic Press. pp. 600.

Mardia, K.V. and Watkins, A.J. (1989). "On Multimodality of the Likelihood in the Spatial Linear Model," *Biometrika*, 76, 289-295.

Pregibon, D. (1981). "Logistic Regression Diagnostics," *Annals of Statistics*, 9, 705-724.

SAS (1989). *SAS/STAT User's Guide, Fourth Edition.* Cary, NC: SAS Institute, Inc. pp. 1685.

Solow, A.R. (1989). "Detecting Changes Through Time in the Variance of a Long-Term

Hemispheric Temperature Record: An Application of Robust Locally Weighted Regression," *Journal of Climate*, 1, 290-296.

Theil, H. (1971). *Principles of Econometrics.* New York: John Wiley and Sons, Inc. pp. 736.

Warnes, J.J. and Ripley, B.D. (1987), "Problems with Likelihood Estimation of Covariance Functions of Spatial Gaussian Processes," *Biometrika*, 74, 640-642.

Wellman, J.M. and Gunst, R.F. (1991). "Influence Diagnostics for Measurement Error Models," *Biometrika*, 373-380.

Wigley, T.M.L. and Jones, P.D. (1981). "Detecting CO2-Induced Climatic Change," *Nature*, 292, 205-208.

Wigley, T.M.L., Jones, P.D., Kelly, P.M., and Raper, S.C.D. (1989). "Statistical Significance of Global Warming," In *Proceedings of the Thirteenth Annual Climate Diagnostics Workshop*, pp. A.1 to A.8. Washington, D.C.: U.S. Department of Commerce.

List of Tables

List of Figures

Table 1. Gaussian Semivariogram Parameter Estimates: 1990 U.S. Anomalies.

Month	Nugget	Sill	Range
January	.75	9.99	20.63
February	.67	7.29	23.00
March	.50	2.68	16.22
April	.55	4.20	25.39
May	.18	.75	7.53
June	.32	.76	10.95
July	.33	1.31	18.09
August	.18	.54	8.05
September	.36	2.56	15.70
October	.21	1.51	8.22
November	.44	1.27	5.85
December	.49	15.72	21.59

Table 2. Gaussian Semivariogram Parameter Estimates: 1990 Europe Anomalies.

Month	Nugget	Sill	Range
January	.86	5.20	11.19
February	.44	10.12	17.97
March	.21	3.88	15.06
April	.10	1.56	10.92
May	.18	4.74	25.40
June	.14	1.33	22.60
July	.11	1.72	14.83
August	.32	2.46	21.09
September	.17	2.04	16.51
October	.10	.90	9.20
November	.44	1.16	9.26
December	.24	1.00	13.42

Fig. 1. U.S. Station Locations: 1990

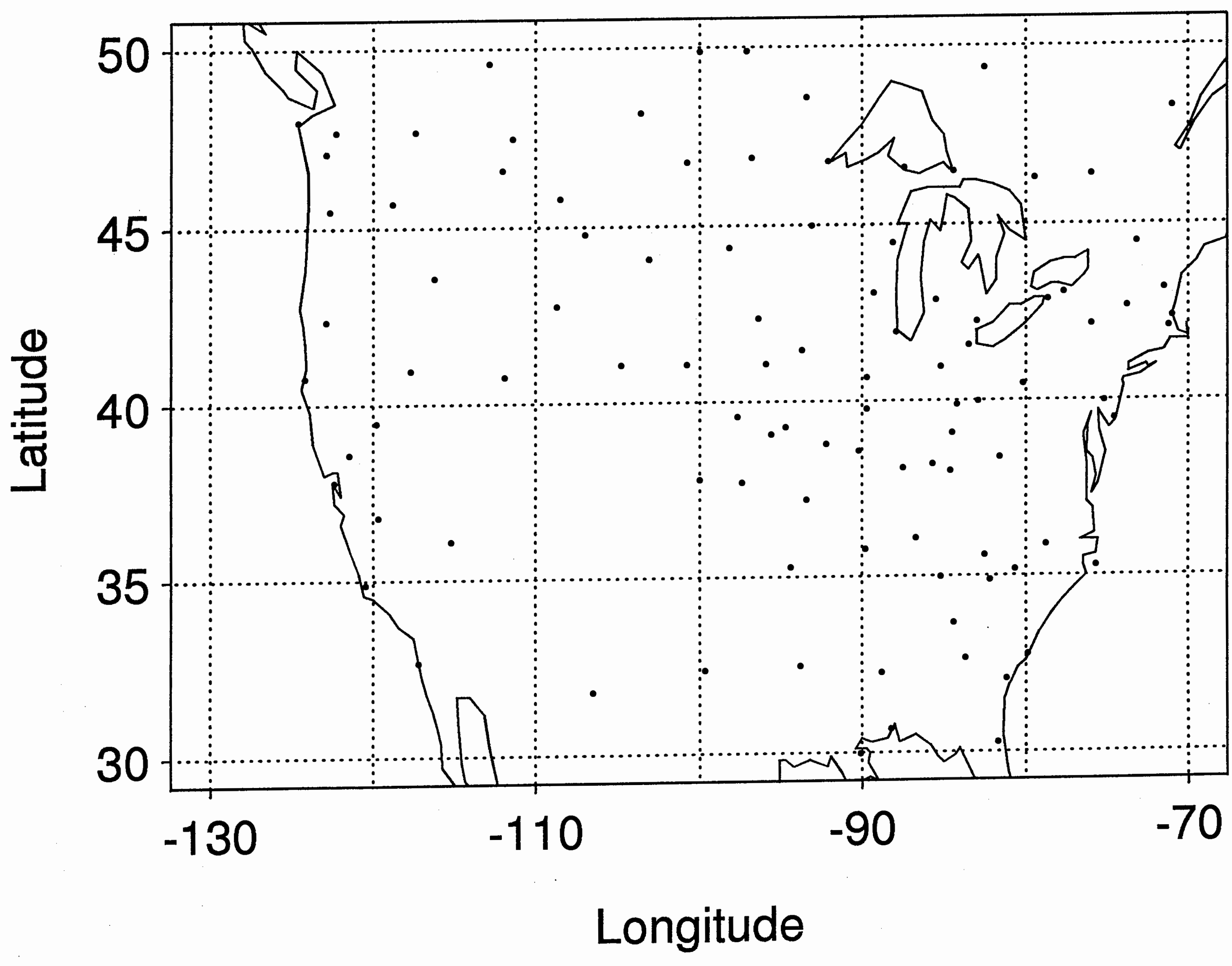

Fig. 2. Anomaly Pairs, January 1990

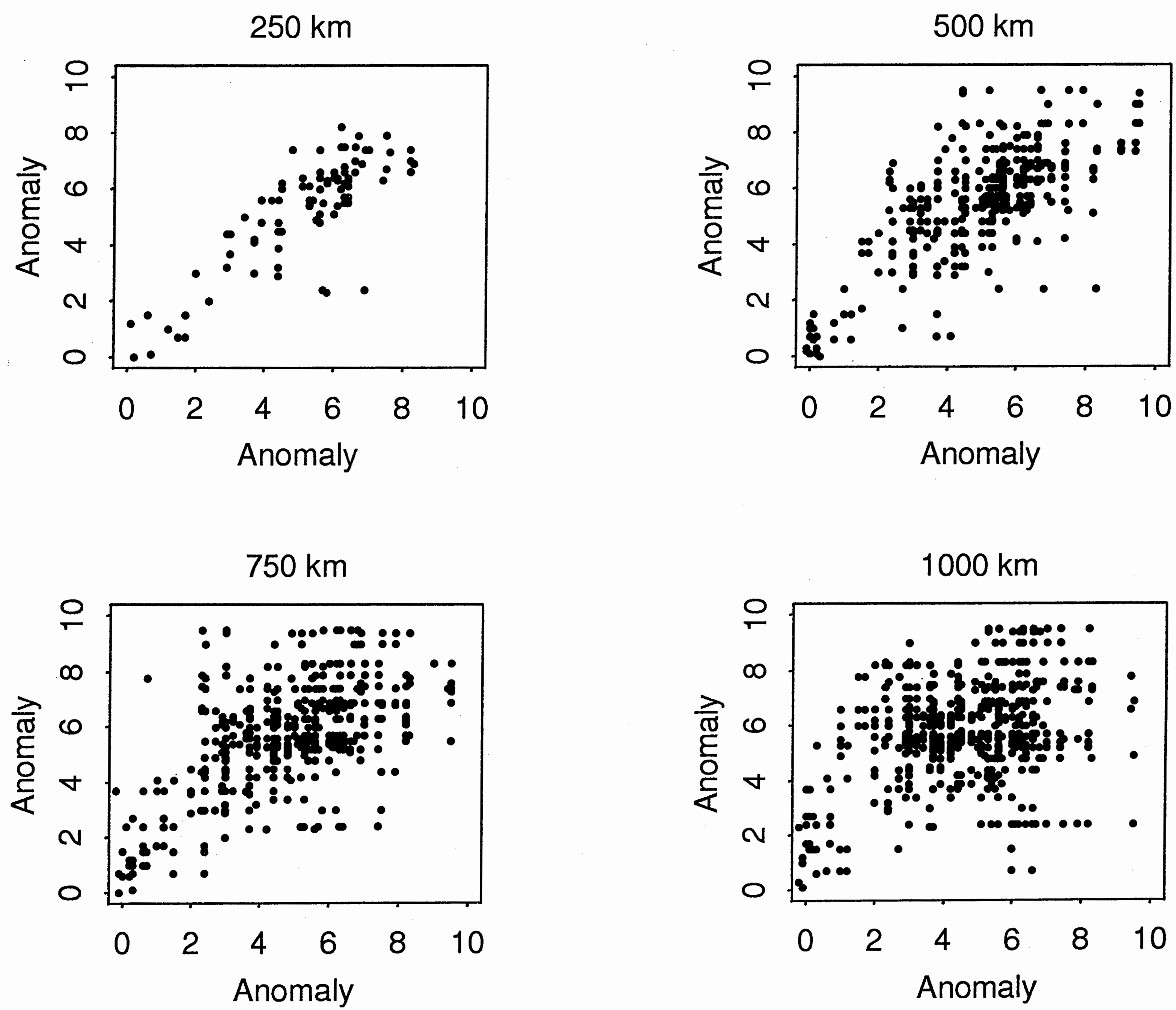

Fig. 3. Anomaly Pairs, May 1990

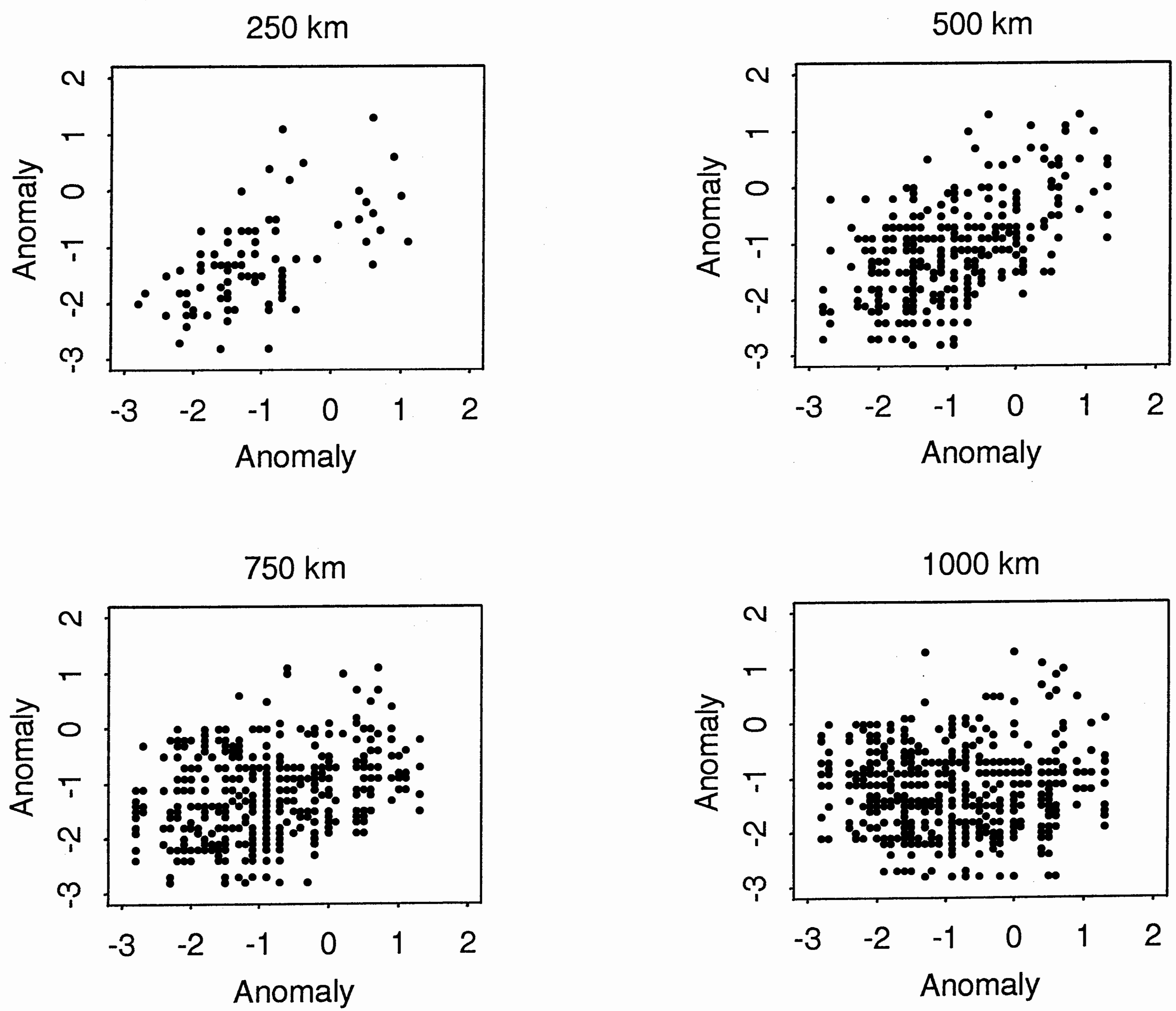

Fig. 4. Sample Semivariograms, 1990

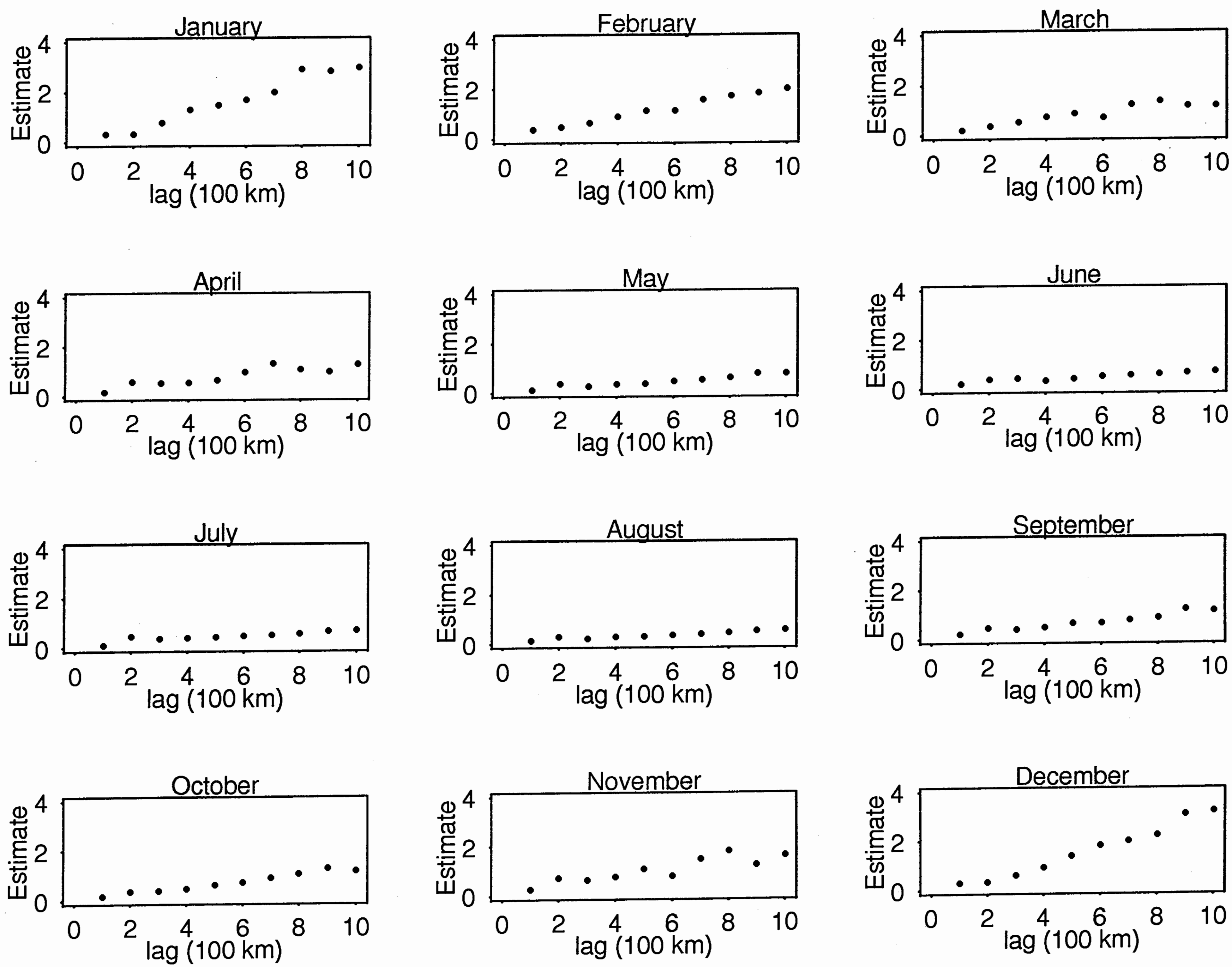

Fig. 5. Gaussian Semivariogram Model

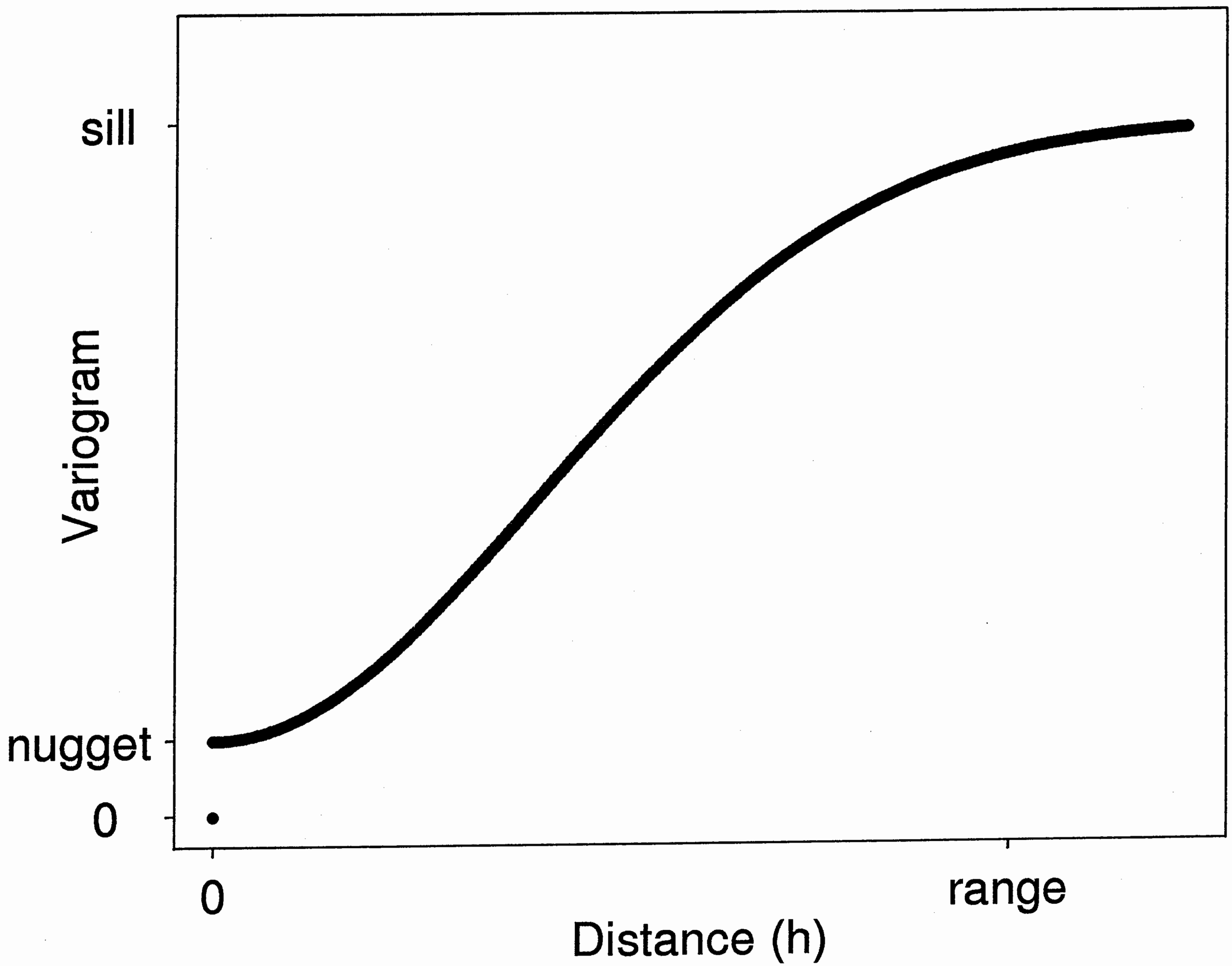

Fig. 6. One-Step Estimated Gaussian Semivariograms, U.S. 1990

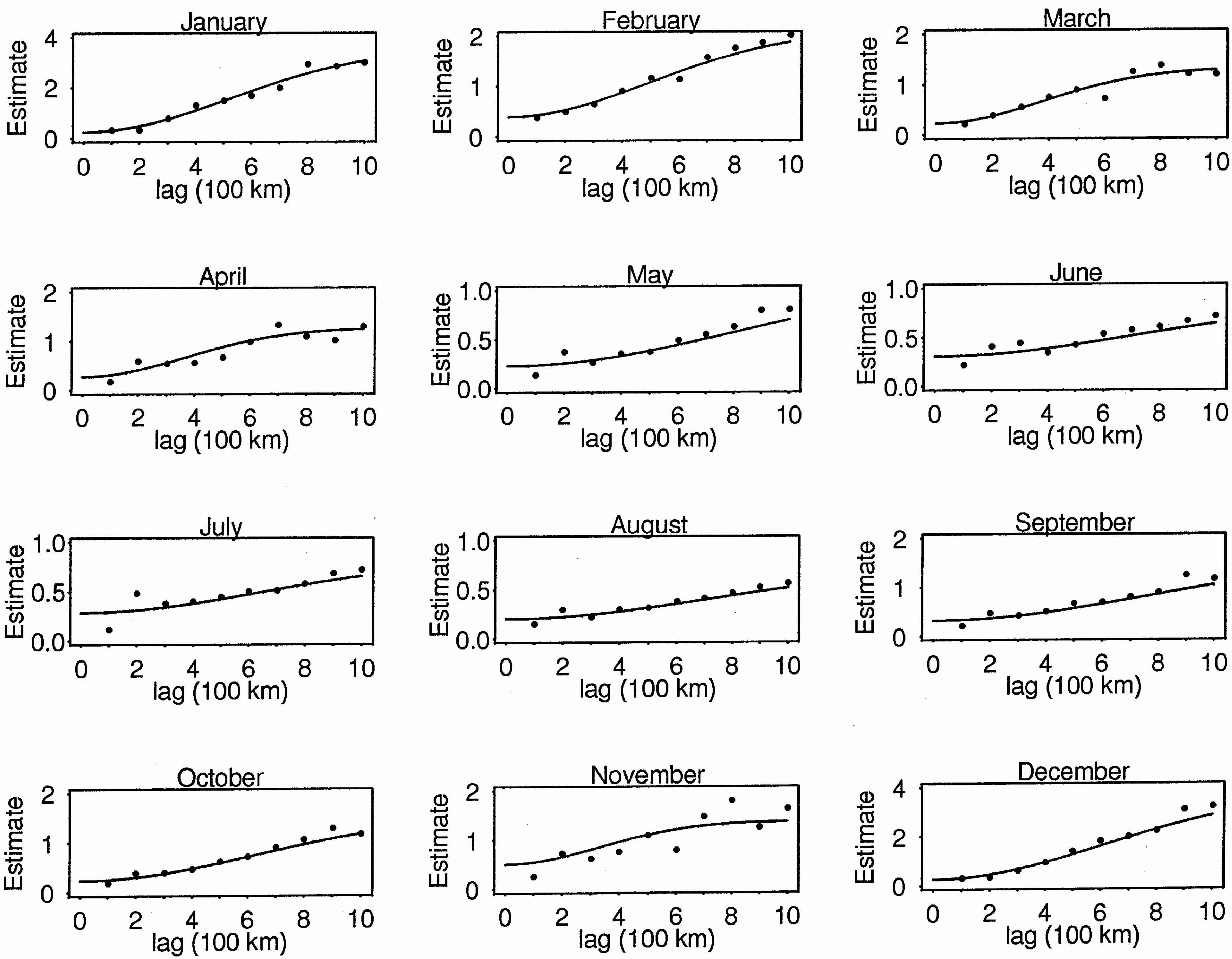

Fig. 7. One-Step Estimated Spatial Correlations, U.S. 1990

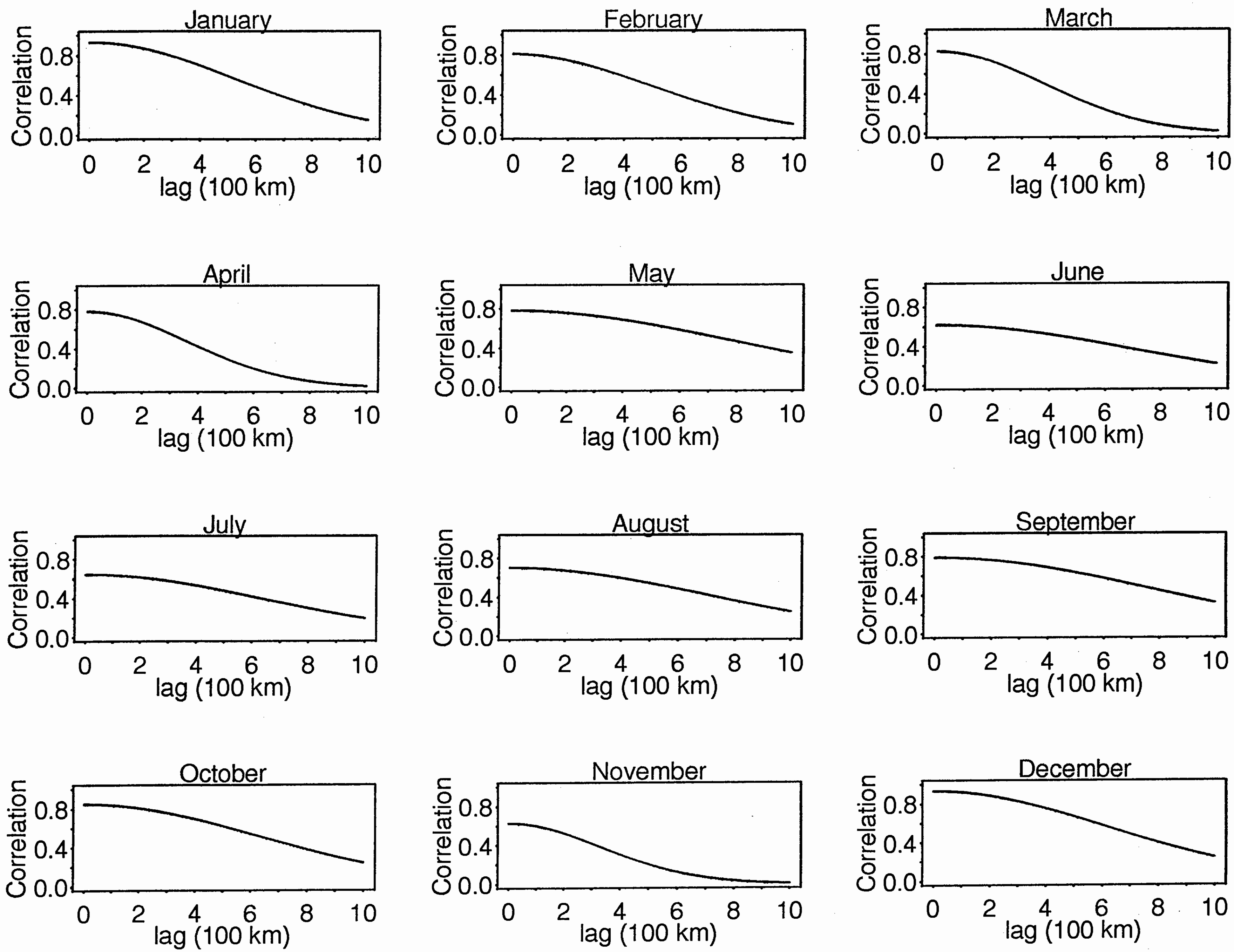

Fig. 8. One-Step Estimated U.S. 1990 Gaussian Semivariograms, 20 Lags

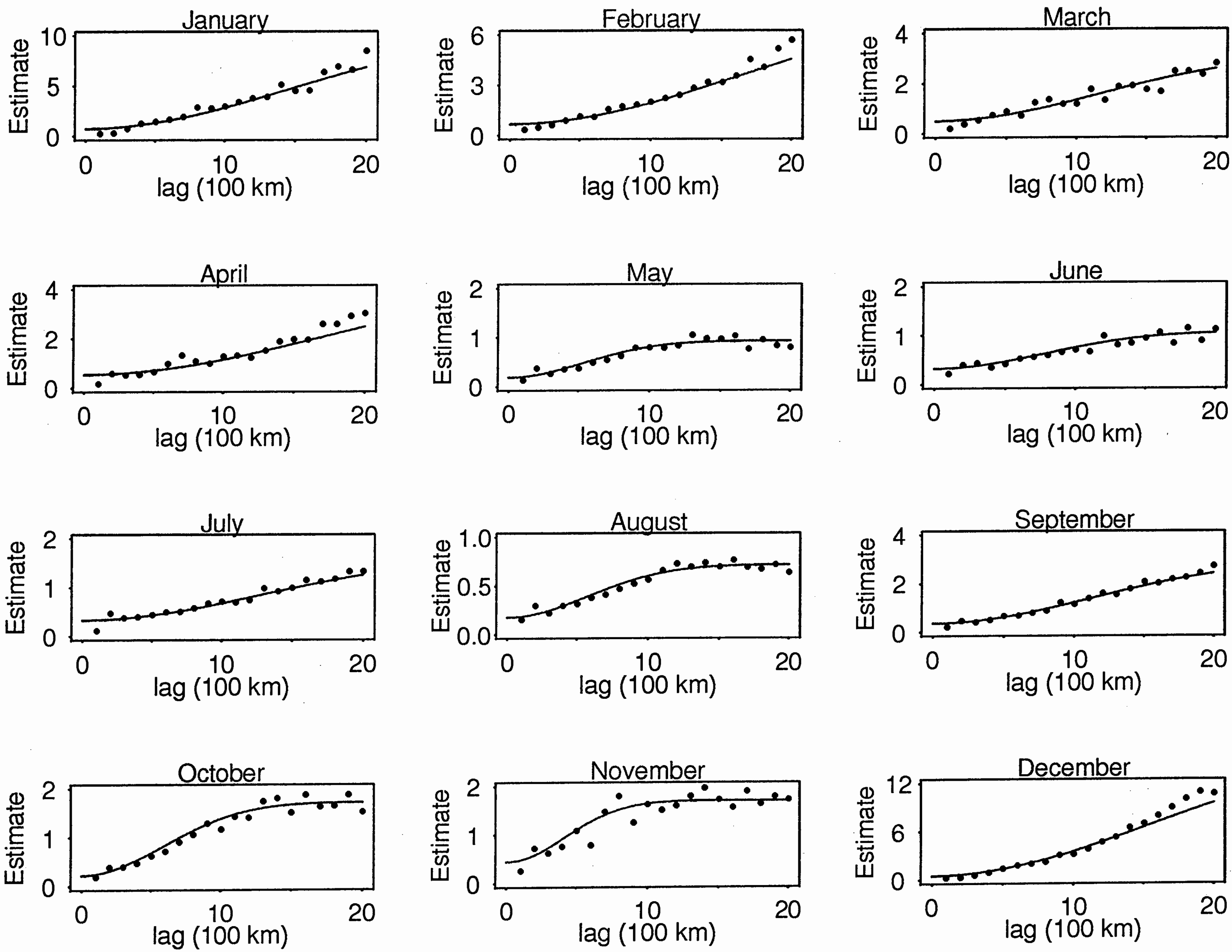

Fig. 9. One-Step Estimated U.S. 1990 Spatial Correlations, 20 Lags