Estimating Spatial Correlations From Spatial-Temporal Meteorological Data

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

ABSTRACT

Meteorological data are collected over space and time. Techniques for modeling the temporal characteristics of meteorological data are well known and accepted. While methods for accommodating the spatial character of meteorological variables have been available for over 30 years, they are far less frequently used. In part, this is because many theoretical and computational issues remain to be resolved concerning the application of spatial modeling methods to meteorological data. One of these issues is the fitting of spatial correlation functions. In this paper, it is shown that temporal correlations can seriously bias fitted covariance functions that are used in optimal interpolation and optimal spatial averaging (kriging) methods. A result of this bias is that the fitting of correlation functions to temporal correlations can result in overestimates of spatial correlations. In contrast, the fitting of structure function (semivariogram) models from data for fixed time periods does not suffer from the temporal biases of correlation function fitting and should be the preferred method for estimating spatial correlations. Semivariogram estimation and structure function modeling must, however, accommodate anisotropic trends in the data.

1. Introduction

Spatial modeling has a long history, especially in the meteorological and geoscience literature. Almost simultaneously Matheron (1962) and Gandin (1963) proposed methods that have been termed, respectively, (point or block) *kriging* and *optimum interpolation*. Kagan (1979) extended optimal interpolation to *optimal spatial averaging*. Apart from the methods used to estimate spatial correlations that are embedded in these methods, the statistical technique of *best linear unbiased prediction* in a more general setting is attributed to Goldberger (1962). All of these derivations were predated by Wold (1938), Kolmogorov (1941), and Wiener (1949) in a temporal setting. Cressie (1990) provides a history of the subject.

The seminal contributions of Matheron and Gandin in a spatial setting center around the estimation of spatial correlations. Both proposed similar estimation methods based on the calculation and fitting of the same quantity, termed the *variogram* by Matheron and the *structure function* by Gandin. Gandin also proposed fitting correlation functions. The fitting of models to semivariograms (half the variogram values) predominates in the geoscience and the statistical literature (e.g., Cressie 1991; Journel and Huijbregts 1978), whereas the fitting of correlation functions pervades the meteorological literature (e.g.,

Thiebaux and Pedder 1987; see also Section 2). Each is appropriate and effective when used as warranted by the statistical properties of the data being analyzed.

A concern with the use of correlation function fitting with meteorological data is that the common occurrence of temporal trends and autocorrelations can induce substantial bias in the estimation of spatial correlations. Semivariogram model fitting can be more difficult in many respects than correlation function fitting, so the former cannot be discounted. Nevertheless, semivariogram models are fit to squared differences in meteorological variates measured at a specific time. The proper application of semivariogram fitting can therefore ameliorate all or most of the temporal effects. Consequently, the potential for temporal bias in the estimated spatial correlations from semivariogram model fits is reduced, if not altogether eliminated.

The objective of this paper is twofold. First, the common but often inappropriate use of models fit to correlations calculated from time series is shown to lead to substantial bias in the estimation of spatial correlations. Prewhitening the individual time series does not overcome this bias. Second, the fitting of models to sample semivariograms largely overcomes the temporal bias but must be done with suitable accommodation of anisotropic behavior. One method that is easy to implement is to detrend the data with a suitable trend model and then to calculate and model the residual semivariograms. Satisfactory detrending of the meteorological data often enables one to fit isotropic semivariogram models.

In Section 2 of this paper, the aliasing of temporal and spatial effects in correlations calculated from time series data is discussed. In particular, it is shown that temporal correlations overestimate spatial correlations when even modest temporal effects exist in the time series. In contrast, sample semivariogram calculations, detailed in Section 3, do not suffer from similar temporal biases because they are calculated at fixed points in time and they involve differences in spatial variates. Ensemble averages of the sample semivariograms can then provide stable estimates of spatial variability and isotropic semivariogram modeling, the topic of Section 4. Parameters of the semivariogram models are shown to provide explicit information on the strength and extent of spatial correlation. In Section 5, suggestions are made for semivariogram modeling of anisotropic meteorological data, with emphasis on trend removal. Estimation of spatial correlations from fitted semivariogram models is the focal point of Section 6. Concluding remarks are made in Section 7.

2. The Aliasing of Temporal and Spatial Effects in Time Series Correlations

Arguably, the fitting of a smooth curve to spatial correlations (Pearson's r) plotted as a function of station separation distance is simpler and more readily understood by modelers than is the estimation and fitting of semivariograms. Figure 1 is a plot of correlations calculated from annual temperature anomalies (Jones et al. 1991) for 138 stations in the continental United States whose latitudes are between 30° and 50° N. These 138 stations have complete monthly data over the period 1920 to 1980. Figure 1 is typical of correlation plots to which curves are fit for the purpose of estimating spatial correlations. The exponential curves proposed by Gandin (1963) are often used to fit such

data (e.g., Julian and Thiebaux 1975, Table 1; Thiebaux and Pedder 1987, Section 4.3). Recent applications include Vinnikov, Groisman, and Lugina 1990, Brown and Eischeid (1992) and Briffa and Jones (1993).

Two features of the correlation plot are noteworthy. First, as the separation distances decrease, the correlations appear to approach their maximum value, 1. Second, as the separation distances increase, the correlations weaken and become negative. The exponential correlation functions used to fit these correlations are partially justified because they can accommodate the "negative lobe" of the correlation plot.

The fitting of correlation functions to plots such as Figure 1 is appropriate only when the meteorological data do not exhibit deterministic or autocorrelated temporal behavior (e.g., Thiebaux and Pedder 1987, Sections 4.2, 4.3). As obvious as this restriction may seem, it is not addressed in many applications. The potential for incurring bias in the estimation of *spatial* correlations from *temporal* data is readily demonstrated using the two time series plotted in Figure 2. Plotted in the figure are annual anomalies for two stations in the southwestern United States, San Diego and Los Angeles California. The calculated correlation between the anomalies is 0.86 over the period 1900-1980 and 0.80 over the period 1920-1980, the time period used in the correlation plot in Figure 1. Since the separation distance between San Diego and Los Angeles is 181 km, one would conclude from a correlation function fit to Figure 1 that there is a strong spatial correlation between the station anomalies. The calculated correlations appear to confirm the expectation -- however, these "spatial" correlations are biased upward because of the temporal components of the series.

The time series plots in Figure 2 suggest the presence of both linear and periodic temporal components. Consider the following model for the two time series $\{\mathbf{y}_{1t}\}$ and $\{\mathbf{y}_{2t}\}$, t = 1, 2, ... , n:

$$\begin{aligned} \mathbf{y}_{it} &= \mathbf{x}_{it} + \mathbf{u}_{it}(\mathbf{s}_i) \\ \mathbf{x}_{it} &= \alpha_i + \beta_i t + \phi_i \mathbf{x}_{i,t-1} + \mathbf{a}_{it}. \end{aligned} \tag{2.1}$$

This model was chosen for investigation because the anomaly data in Figure 2 are well represented by such a representation (see below). In this model, the series $\{\mathbf{y}_{1t}\}$ and $\{\mathbf{y}_{2t}\}$ represent meteorological time series from two stations situated at locations $\mathbf{s}_1$ and $\mathbf{s}_2$, respectively, where $\mathbf{s}_i$ can be a two- or three-dimensional location vector. These series are composed of a temporal component $\{\mathbf{x}_{it}\}$ and a spatial component $\{\mathbf{u}_{it}(\mathbf{s}_i)\}$. The temporal component is assumed to be a first-order autoregressive process with deterministic linear trend. The temporal innovation series $\{\mathbf{a}_{1t}\}$ and $\{\mathbf{a}_{2t}\}$ are assumed to be a bivariate white noise process with variances σ_{1t}^2 and σ_{2t}^2 and lag-zero temporal cross-correlation ρ_t. The spatial error series $\{\mathbf{u}_{1t}\}$ and $\{\mathbf{u}_{2t}\}$ are also assumed to be a bivariate white noise process with variances σ_{1s}^2 and σ_{2s}^2 and spatial correlation ρ_s. If more than two stations were being modeled in this example, the spatial correlations would change with the distances (and directions) between the pairs of stations. Since only two stations are being modeled here, the spatial correlation ρ_s is modeled as a fixed constant value for all t = 1, 2, ... , n. Model (2.1) is simultaneously simple and complex enough to illustrate the effects that various temporal and spatial components of a meteorological time series can have on the calculated value of a correlation coefficient.

The effects of the temporal and the spatial components of model (2.1) on calculated correlations can be more readily appreciated if some of the model parameters are set to common values:

$$\sigma_{1t}^2 = \sigma_{2t}^2 = \sigma_{tt},\ \sigma_{1s}^2 = \sigma_{2s}^2 = \sigma_{ss},\ \alpha_1 = \alpha_2 = \alpha,\ \beta_1 = \beta_2 = \beta,\ \phi_1 = \phi_2 = \phi. \tag{2.2}$$

These common values are used merely for simplicity of exposition. Results corresponding to those discussed below can be derived for any model parameter values of interest.

The correlations plotted in Figure 1 are calculated from the formula for Pearson's r:

$$\mathbf{r}_{ij} = \frac{\hat{\sigma}_{ij}}{\left(\hat{\sigma}_{ii} \times \hat{\sigma}_{jj}\right)^{1/2}} \tag{2.3}$$

where $\hat{\sigma}_{ij}$ is the sample covariance calculated from the time series for stations i and j, and $\hat{\sigma}_{ii}$ and $\hat{\sigma}_{jj}$ are the corresponding sample variances. An adaptation of Theorem 11.2.1 of Brockwell and Davis (1991) can be used to show that $\mathbf{r}_{ij}$ is a consistent (but not unbiased) estimator of ρ_{ij}, where ρ_{ij} is $\mathbf{r}_{ij}$ with each of the sample moments replaced by its mathematical expectation. For the two series defined by (2.1) with the common parameter values in (2.2), tedious but straightforward derivations result in the limiting value:

$$\rho_{12} = \frac{\beta^2 s_{tt} + (1-\phi^2)^{-1}\rho_t\sigma_{tt} + \rho_s\sigma_{ss}}{\beta^2 s_{tt} + (1-\phi^2)^{-1}\sigma_{tt} + \sigma_{ss}}$$
$$s_{tt} = (n-1)^{-1}\sum_{t=1}^{n}(t-\bar{t})^2. \tag{2.4}$$

One can calculate the value of (2.4) for any model parameters of interest and assess the effects of the temporal and spatial components of model (2.1) on the correlation between the two time series. Table 1 has one such set of calculations for ρ_{12} when $\sigma_{tt} = \sigma_{ss}$ and $\beta = 0$; i.e., for

$$\rho_{12} = \frac{(1-\phi^2)^{-1}\rho_t + \rho_s}{(1-\phi^2)^{-1} + 1}\ . \tag{2.5}$$

In (2.5) ρ_{12} is only a function of the autocorrelation parameter ϕ, the temporal lag-zero cross-correlation ρ_t, and the spatial correlation ρ_s. For strong autocorrelation, say $\phi = 0.9$, the calculated values in Table 1 clearly indicate that $\mathbf{r}_{12}$ is estimating a correlation ρ_{12} that is dominated by the temporal lag-zero cross-correlation. When the autocorrelation is weaker, say $\phi = 0.3$, $\mathbf{r}_{12}$ estimates a combination of the temporal and spatial correlations. It is clear from (2.4) and (2.5) that $\mathbf{r}_{12}$ estimates the spatial correlation only when there is no temporal component to the data.

As mentioned above, the series plotted in Figure 2 suggest the presence of both a linear trend and temporal autocorrelation. Indeed, least squares fits to the San Diego and Los Angeles series, using year as the predictor, have respective slopes $\hat{\beta}$ of 0.029 and

0.021 for the period 1920-1980. First-order autoregressive models fit to the respective least squares residuals (anomaly - linear fit) have autoregressive parameters estimates $\hat{\phi}$ of 0.29 and 0.30, estimated innovation variances $\hat{\sigma}_{tt}$ of 0.33 and 0.36, and lag-zero estimated cross correlation $\hat{\rho}_t$ of 0.78. The first-order model for the residuals was selected separately for each series using Akaike's information criterion, AIC. Brockwell and Davis (1992, Chapter 7) recommend separate prewhitening of each time series prior to estimating the lag-zero cross-correlations; hence, the cross correlations were obtained after each series was prewhitened by removing the linear trend and the autoregressive component. Only the lag-zero cross-correlation was statistically significant between the prewhitened series. Similar results are obtained if one models the linear and autoregressive components simultaneously using transfer function models.

Table 2 contains calculations of (2.4) using temporal parameter values comparable to those of the estimated San Diego and Los Angeles series: $\beta = 0.025$, $\phi = 0.35$, $\rho_t = 0.8$, and $\sigma_{tt} = 0.3$. The spatial variance was obtained from the semivariogram model fits in Section 4: $\sigma_{ss} = 0.16$. As a function of the spatial correlation ρ_s, the values of ρ_{12} in Table 2 change very little, implying that the value of the calculated correlation from the anomaly series is dominated by the temporal components of the model, not by the true spatial correlation. In fact, as indicated by the values in Table 2, a small spatial correlation of approximately 0.4 could lead to a ρ_{12} correlation between the anomaly series similar to the actual calculated value of 0.80.

The foregoing analyses are intended to stress that correlations calculated from time series data need not be indicative of the magnitude of spatial correlations. Indeed, a correlation near 1 that is calculated from time series for two stations may be heavily biased by the temporal characteristics of the time series and may not reflect the true magnitude of the spatial correlation between the stations. This conclusion might appear to be counterintuitive, especially when viewing plots such as Figure 1. Two additional observations make this conclusion more apparent. First, there is a high proportion of stations in relative close proximity in Figure 1 that do not have anomaly correlations close to 1. Sixty-six percent of the 65 station pairs whose separation distances are less than 100 km have anomaly correlations less than 0.8; 68% of the 19 stations within 50 km of one another have anomaly correlations less than 0.8.

Second, a well-documented property of spatial variability is the so-called *nugget effect* (e.g., Journel and Huijbregts 1978, Sections II.A.3, III.A.3; Cressie 1991, Sections 2.3.1, 3.2.1). This effect exists because of measurement errors and random local variation at distances less than the closest observations in a data set. Figure 3 is a "box-and-whisker" plot (e.g., Mason, Gunst, and Hess 1989, Section 4.3) of the absolute differences in annual temperature anomalies for U.S. station pairs for the year 1980 (other years are similar). Only station pairs that are in relative close proximity to one another are included in the plot; specifically, those whose station separation distances are no greater than 500 km. The box limits are plotted at the lower and upper quartiles of the respective anomaly differences. The box limits thus include 50% of the anomaly differences. The whiskers (vertical dashed lines) extend to the largest and smallest anomaly differences that do not exceed robust three-sigma estimates of uncertainty. Individual anomaly differences that do exceed these limits are shown as horizontal lines beyond the whiskers. Finally, the notched indentations in the centers of the box locate the calculated medians and approximate (due

to the correlations among the anomaly differences) 95% confidence limits on the true medians.

As station separation distances are reduced, one might expect the anomaly differences to be approaching zero because of the anticipated high correlations among the stations. If so, one would expect that the medians of the boxes in Figure 3 would approach zero and the widths of the boxes would be getting smaller. This is not the case. Box plots have been constructed for each year from 1880 to 1991 and they all exhibit the same general pattern as is shown in Figure 3. Station temperature measurements, no matter how closely they are taken, are subject to measurement errors, instrument differences, local topographical, and random sources of variation that prevent the measurements from being exactly alike. This has long been recognized in meteorology (e.g., Buell 1972) and is referred to as the *nugget effect* in geoscience applications. In addition, it is important to note that the station separation distances for data used in global temperature anomaly calculations are typically no closer than tens of kilometers. Only 65 of the 9,453 station pairs of U.S. stations having complete monthly data for all months between 1920 and 1980 are within 100 km of one another. The characteristics noted of the boxes in Figure 3 are not just a feature of these temperature anomalies. They are common in spatial data. Cressie's (1984) square-root difference cloud of coal ash data is another example.

It must be stressed that the characteristics of spatial data noted in these analyses, while common in spatial data, do not necessarily exist in all meteorological measurements. That they are present in temperature anomaly data mandates that alternative spatial correlation estimation techniques be considered, if for no other reason than to confirm that traditional spatial correlation methods applied to temperature anomaly data are not seriously biased. Semivariogram (structure function) calculations and modeling can be used for this purpose.

3. Sample Semivariogram Calculations

Semivariograms do not require data to be indexed by time. Rather, semivariogram calculations are based explicitly on the distances and directions between spatial locations. Cressie (1991, Section 2.4) details the rationale for calculating semivariogram values rather than directly calculating spatial covariances (covariograms) or correlations. Fuller (1976, Section 6.2) details a similar rationale for temporal modeling. In this section, discussion is restricted to isotropic sample semivariogram calculations. Semivariogram calculations and model fitting for anisotropic meteorological data are discussed in Section 5.

The semivariogram for two random spatial variables is defined as

$$\gamma(s_1, s_2) = \mathrm{var}\{z(s_1) - z(s_2)\} / 2 \,, \qquad (3.1)$$

where $z(s_i)$ denotes an element of a random field of meteorological variables (see Journal and Huijbregts (1978) for detailed discussions of spatial random fields). In general, the stochastic variation of random fields can depend on both the distances and the directions between locations. The variation of many meteorological variables, at least in sufficiently

restricted regions or after detrending, can be satisfactorily modeled as a function of the only the distance $\mathbf{d} = |\mathbf{s}_1 - \mathbf{s}_2|$ between locations; i.e., they need not be modeled as a function of their specific locations nor the directions between the locations. This simplification is the most widely used in practice.

Figure 4 is a plot of 1980 annual temperature anomalies for the 138 U.S. meteorological stations discussed in the last section. Superimposed over the plotted anomalies is a nonparametric perspective plot (S-Plus 1991). The fitted surface is included in Figure 4 only to make the overall trends in the anomalies clearer. There appears to be an increasing east-west trend in the anomalies, as well as perhaps a south-north trend. Of primary importance in this plot, however, is the relatively constant *variation* in the anomalies as a function of location.

Figure 5 confirms these impressions. In the upper plots in Figure 5, the anomalies are seen to be roughly linearly decreasing with latitude and quadratically varying with longitude. If one were to attempt to estimate spatial (random) variability from the data, the estimates would be biased by the trends (see Section 5). Estimates of spatial variability must be made from detrended data, similar to recommendations made for the estimation of variation in time series models (e.g., Brockwell and Davis 1992, Chapter 7) . Katz(1988) discusses the need to detrend meteorological data when drawing inferences on climate variation.

To detrend the anomaly data in Figure 5, a five-term quadratic regression model in latitude and longitude was fit to the anomalies. The lower two plots in Figure 5 show the residuals from the fit as a function of latitude and longitude. The trends in the anomalies are largely removed. Note, however, that the *variation* of the anomalies remains about the same magnitude as on the original plots and that it is relatively homogeneous across the two plots. Trend removal is one method that can be used to deal with anisotropy in semivariogram modeling and for removing apparent inhomogeneity in spatial variances. For the remainder of this section, the residuals of this fit to the anomalies will be used to estimate the spatial variability of the anomalies.

For isotropic meteorological data, the "classical" method of moments variogram estimator (Matheron 1962) is the sample semivariogram

$$\gamma(\mathbf{d}) = \frac{1}{2\mathbf{n}(\mathbf{d})} \sum_{i,j}^{\mathbf{n}(\mathbf{d})} \{z(\mathbf{s}_i) - z(\mathbf{s}_j)\}^2 , \qquad (3.2)$$

or half the average of the squared differences of the $\mathbf{n(d)}$ meteorological values or anomalies whose locations are separated by a distance $\mathbf{d}$. When the locations are spaced on a regular lattice, the $\mathbf{n(d)}$ pairs of locations that are separated by each distance $\mathbf{d}$ are readily determined. When locations are irregularly spaced as with many meteorological data, the station pairs must be binned into groups of distances (e.g., increments of 100 km). Further details on such binning can be found in Cressie (1991, Section 2.4) or Journel and Huijbregts (1978, Section III.C.4). The latter also recommend (Section III.B.7) that at least 30 location pairs be in each bin and that the largest binned distance that is modeled be no greater than half of the maximum distance among the locations. For the U.S. temperature anomaly data, this effectively limits the range for which semivariogram values should be calculated to 2,000 km.

Cressie and Hawkins (1980) propose robust estimators of the semivariogram based on averages of the square roots of the absolute differences. Other robust methods can be found in Cressie (1991, Section 2.4). Basu and Gunst (1993) show that extreme anomaly values can cause spikes in semivariogram plots, spikes that induce computational problems when semivariogram models are fit to the sample semivariogram values. They further show that robust methods only partially ameliorate the effects of the extreme values on the semivariogram calculations; i.e., spikes due to such *influential observations* are not completely removed. They introduce influence diagnostics for the identification of extreme values and show that the removal of the influential observations can be more effective than robust methods for smoothing the sample semivariogram.

Figure 6 displays individual sample semivariogram plots for residuals from quadratic least squares fits to temperature anomalies for each of the 30 years 1951-1980. The residual semivariogram values were calculated for bins of length 100 km up to a maximum distance of 2,000 km. In other words, 20 bins were used, with station pairs whose (great circle) separation distances were less than or equal to 100 km placed in the first bin and similarly for the other 19 bins. It is important to recognize the need to restrict the distance over which semivariogram values are calculated. Apart from the admonitions of Journel and Huijbregts (1978), restricted ranges are necessary in global modeling because of the likelihood that temporal influences on meteorological variables are not the same for all stations over, say, a latitude band. Although different detrending fits could be used for regions widely separated, different local and regional conditions make it unlikely that one semivariogram calculation would adequately represent the spatial variation across a latitude band. Preliminary investigations of United States and European station data support this concern.

From Figure 6 it is apparent that there is variation across the individual residual sample semivariogram plots, as one might expect from any statistics calculated for each of 30 consecutive years. General characteristics of the semivariogram plots include the following:

(a) the semivariogram values initially increase until separation distances reach approximately 750-1,000 km, the approximate *range* of spatial variation;
(b) the semivariogram values are relatively constant (at the *sill* value) for separation distances larger than the range; and
(c) the semivariogram values approach a positive -- not zero -- *nugget* value for small separation distances.

The nugget, sill, and range characteristics are extensively discussed in Cressie (1991) and Journel and Huijbregts (1978). The nugget effect was noted previously in the discussion of the box plots in Figure 3. It occurs because the values of the anomalies (and hence the residuals) are not identical for small station separation distances.

The dark squares in Figure 6 are the ensemble averages of the 30 residual semivariogram values for each bin. The averages suggest a smooth change in spatial variability with separation distance rather than the somewhat jagged appearance of many of the individual plots. In addition, the nugget, sill, and range are better visualized from the plot of average residual semivariogram values than from most of the individual plots. The plot of average residual semivariogram values can thus be used to typify the spatial variation in anomalies over this 30-year period. Note that since separate semivariogram

values are calculated separately for each year, temporal influences that affect correlation calculations do not bias the individual semivariogram values or the ensemble averages.

The numerical values of the nugget, the sill, and the range quantify the key characteristics of the spatial variation of the anomalies. Numerical values for these quantities are readily determined from fits to the average semivariogram values in Figure 6. This is the topic of the next section.

4. Semivariogram Model Fitting

One of the primary reasons for fitting models to sample semivariograms, rather than using the sample values themselves, is that kriging and optimal spatial averaging require the use of estimated spatial covariances between each pair of stations. These estimates are elements of a spatial covariance matrix used in generalized least squares fitting of the anomaly values (Goldberger 1962; Stein and Corsten 1991). The spatial covariance matrix must be positive definite. This in turn implies that the corresponding spatial semivariogram matrix must be conditionally negative definite (Armstrong and Diamond 1984). Spatial semivariogram matrices formed from sample semivariogram values often are not conditionally negative definite and cannot be guaranteed to be so *apriori*. By properly selecting semivariogram models that are fit to sample semivariogram values, spatial semivariogram matrices formed from the fitted models *apriori* can be guaranteed to be conditionally negative definite. Moreover, spatial models fit to semivariograms enable key spatial characteristics such as the nugget, the sill, and the range to be quantified through the use of these quantities as parameters of the models.

Two widely used spatial semivariogram models are the Gaussian and the spherical models:

Gaussian:

$$\gamma(\mathbf{d}) = \begin{cases} 0 & \mathbf{d} = 0 \\ \theta_1 + (\theta_2 - \theta_1)[1 - \exp\{-(\mathbf{d}/\theta_3)^2] & \mathbf{d} > 0 \end{cases} \tag{4.1}$$

Spherical:

$$\gamma(\mathbf{d}) = \begin{cases} 0 & \mathbf{d} = 0 \\ \theta_1 + (\theta_2 - \theta_1)\{1.5(\mathbf{d}/\theta_3) - 0.5(\mathbf{d}/\theta_3)^3\} & 0 < \mathbf{d} \le \theta_3 \\ \theta_2 & \mathbf{d} > \theta_3 \end{cases} \tag{4.2}$$

Both of these semivariogram models have nugget θ_1, sill θ_2, and range θ_3 parameters. Two key differences exist between these models. First, the spherical model reaches its sill when the distance between two locations equals the range parameter value. The Gaussian model reaches its sill only asymptotically. For this reason, Journal and Huijbregts (1978, Section III.B.2) define $\sqrt{3}\theta_3$ to be the "practical" range, the distance at which the semivariogram reaches approximately 95% of the sill value. Second, the spherical model is approximately linear for small separation distances whereas the Gaussian is initially convex; both are concave for large separation distances. A practical implication of their behavior near the origin is that the Gaussian model often has a larger estimated nugget

than a spherical model fit to the same data. On the other hand, nonlinear curve fitting algorithms often do not converge for spherical models when the first few semivariogram values are not approximately linear.

Figure 7 shows the fits of these two semivariogram models to the averages of the residual semivariogram values. The model parameter estimates were calculated using a Levenberg-Marquardt nonlinear least squares curve fitting algorithm (Press et al. 1989). For the spherical model, the nugget, sill, and range parameter estimates were 0.079, 0.163, and 1,210 while for the Gaussian model they were 0.091, 0.163, and 591. Note that the sill values were virtually identical and that the nugget value for the Gaussian model is slightly larger than that for the spherical. The common sill estimate is evident in Figure 7, as is the slight convexity of the Gaussian model as the separation distance approaches zero. The major difference in the fit is for the range parameters, where the spherical model estimate of the range is over twice as large as that for the Gaussian model. However, the Gaussian model range parameter is not a true range estimate since the Gaussian has an asymptotic sill. Note that the "practical sill" of $\sqrt{3}\hat{\theta}_3 = 1{,}024$ for the Gaussian model is not too different from the estimated range for the spherical model.

The sill estimates translate to an overall spatial standard deviation estimate of 0.40 for individual anomaly values. Similarly, the nugget estimates indicate that the microscale and measurement error variation component of the overall anomaly variability has a standard deviation estimate of 0.28 from the spherical fit and 0.30 from the Gaussian fit. The anomalies are approximately uncorrelated at distances beyond the range estimates.

Spherical model fits to sample semivariograms are often preferred because of their linear behavior for small separation distances and consequent smaller nugget estimates than those from Gaussian fits. As will be seen in Section 6, both provide similar estimates of spatial correlation for U.S. temperature anomalies.

5. Anisotropic Semivariogram Modeling

Meteorological data often are not isotropic. The anomalies plotted in Figures 4 and 5 indicate different trends in the north-south and the east-west directions. Anisotropy can occur either in trend (drift) behavior or in the stochastic variability of the data. There are several methods currently available for dealing with anisotropy, some appropriate for trend and others for spatial variability. Some of the methods believed to be most germane to the fitting of semivariogram models will be discussed in this section; however, much research remains to be conducted on quantifying anisotropic characteristics of meteorological data.

One very important method for accommodating anisotropy is the fitting of trend models and the subsequent detrending of the data. Figure 8 contains plots of annual semivariogram values for the original U.S. temperature anomalies for the years 1951-1980. Note that the semivariogram values in Figure 8 are much larger than those of the residuals in Figure 6 because the apparent spatial variability in the anomalies consists of both trend effects and stochastic variability. The averages of the semivariogram values over the 30 years are plotted in Figure 9. These averages do not show a clear range or the reaching of a sill. Nonlinear fitting of the average semivariogram values by a spherical semivariogram model failed due to the lack of a sill in the averages; consequently, the

spatial variability was modeled using Gaussian and linear semivariogram models. The linear semivariogram model is defined by

Linear:

$$\gamma(\mathbf{d}) = \begin{cases} 0 & \mathbf{d} = 0 \\ \theta_1 + (\theta_2 - \theta_1)(\mathbf{d}/\theta_3) & 0 < \mathbf{d} \le \theta_3 \\ \theta_2 & \mathbf{d} > \theta_3 \end{cases} \tag{5.1}$$

The nugget, sill, and range estimates for the Gaussian fit were 0.102, 0.495, and 1,558 while those for the linear fit were 0.058, 0.427, and 2,090. Note that the sill and range estimates for these two semivariogram model fits are much larger than those of the residual semivariogram fits. In fact, the range parameter estimate is larger than the maximum range (2,000 km) of the data that are fit by the semivariogram models.

The semivariogram models fit in Figure 9 contain the effects of anisotropy through the differences in the anomaly trends for the two directions noted in Figure 5. The fitting of the quadratic models in latitude and longitude to annual temperature anomalies in the last section was done in order to detrend the anomalies and remove the anisotropic effects so that the spatial variability could be correctly modeled. The residual plots in Figure 5 show that the anisotropic trends were removed but that the spatial variability evident in the anomalies remained in the residuals. The removal of the trends enabled the spatial variability of the anomalies to be fit using isotropic Gaussian and spherical semivariogram models, both of which produced very similar fits. These fitted semivariogram models can be used with universal kriging or optimal averaging methods to provide improved estimates of regional mean temperature anomalies. This latter topic is not the focus of this paper but it is the ultimate application of the semivariogram model fitting.

A confirmation of the removal of anisotropic behavior in the residuals was made by calculating and fitting semivariogram values in north-south (i.e., $90^{\circ} \pm 45^{\circ}$) and east-west (i.e., $0^{\circ} \pm 45^{\circ}$) directions. Again, 100 km bins were used in each direction. The ensemble averages for each direction overlapped those of Figure 6. The fitted parameter values for the two directions, listed in Table 3, are quite close to those for the isotropic fit. These features of the directional semivariograms confirm that the trend removal eliminated large-scale anisotropic spatial variability from the anomaly values.

A second method for accommodating spatial anisotropy is to sufficiently restrict the regions in which isotropic semivariogram models are fit (e.g., Journel and Rossi 1989; Haas 1990). For example, in some years much of the eastern portion of the United States suffers from little anisotropy in the temperature anomalies. One could conceivably partition a large region such as the continental United States into smaller regions, each of which exhibits isotropic behavior, perhaps after the separate detrending. Semivariogram models then could be fit to the individual regions and regional estimates of mean temperature anomalies could be calculated. These regional estimates could then be combined in an optimal weighted average to produce an estimated mean anomaly for the United States. This possible option is not deemed practical for many types of meteorological data because of the requirement for a large number of data values in each region and because of the likelihood that the anisotropy might not simply consist of different forms of isotropic behavior in each of several regions.

Two general patterns of anisotropic spatial variability have been recognized in previous studies: geometric and zonal anisotropy (Journel and Huijbregts 1978, Section III.B.4). Geometric anisotropy occurs when a linear transformation of the coordinates results in isotropic spatial variation in each of two or more directions. Separate isotropic semivariograms are fit in each of the isotropic directions and then recombined into one anisotropic semivariogram. Zonal anisotropy occurs when the spatial variation consists of several component semivariograms, similar to a variance component model, each of which may have a distinct geometric anisotropy. Journel and Huijbregts (1978, Section IV.D) contains examples of the fitting of each type of anisotropic semivariogram. Cressie (1989) also has an example of geometric anisotropy.

Thiebaux (1977) fits anisotropic correlation functions to 500 mbar winter height and wind data. Assuming temporal and spatial homogeneity of variances and correlations, Thiebaux formed anisotropic correlation functions by taking products of isotropic latitudinal and longitudinal correlation functions: $\mathbf{R}(\tau,\sigma) = \mathbf{R}_{\phi}(\tau,\sigma)\mathbf{R}_{\lambda}(\tau,\sigma)$. Justification for this product form was based on separability assumptions along meridonal and zonal directions.

Due to the widespread availability of regression algorithms, detrending followed by isotropic semivariogram modeling is an especially attractive approach to the spatial modeling of anisotropic meteorological data. To the extent that such an approach can be shown to be generally applicable, it is preferable to the other approaches. The residual anomalies for each of the years 1951 to 1980 were examined for evidence of anisotropy; none were found. Hence, the fitting of the individual and average residual semivariogram values by isotropic semivariogram models satisfactorily characterizes the spatial variability in the temperature anomalies. Thus, the other alternatives mentioned in this section were not deemed necessary.

6. Estimating Spatial Correlations

Once sample semivariogram models have been fit to spatial data, spatial correlations can be calculated. Correlation curves, similar to those commonly used to fit the correlations in Figure 1, are readily obtainable. Under spatial second-order stationarity assumptions, taking the mathematical expectation of both sides of (2.1) leads to the following relationship between semivariogram values and spatial correlations:

$$\gamma(\mathbf{d}) = \sigma_{zz}\{1 - \rho(\mathbf{d})\} \ . \qquad (5.2)$$

In this expression, σ_{zz} is the variance (i.e., the sill) of the meteorological variates and $\rho(\mathbf{d})$ is the spatial correlation. Second-order stationarity is not necessary for a derivation similar to (5.2), it is used here only for convenience. The importance of (5.2) is that once a semivariogram model has been fit to the sample semivariogram values, the spatial correlation can be estimated for locations any distance **d** apart by solving (5.2):

$$\hat{\rho}(\mathbf{d}) = 1 - \hat{\gamma}(\mathbf{d}) / \hat{\sigma}_{zz} \ . \qquad (5.3)$$

Figure 10 is a spatial correlation plot from the fits to the anomaly semivariogram values and to the residual semivariogram values. There are major differences in the estimated spatial correlations from the fits to the two sets of variogram values. The spatial correlations estimated from the fits to the anomaly values are uniformly larger than those fit to the residual semivariogram values; moreover, the two correlation curves for the anomaly data are not in close agreement. Both the lack of close agreement and the larger apparent spatial correlations for the anomalies are due to the trend effects and not to the true spatial correlation effects.

The spatial correlation curves for the residuals approach approximately 0.4 while those of the anomalies approach values in excess of 0.8. This is due to the smaller spatial variance of the detrended anomalies and the fact that local variation and measurement errors constitute a larger portion of the spatial variability of the residuals than of the anomalies. Plots similar to Figure 1 leave the impression that spatial correlations should approach 1 as the separation distance is decreased. This simply is not the case, as the box plots in Figure 3 suggest and the correlation curves in Figure 10 confirm. In terms of the estimated model parameters, it is readily shown that the maximum value of $\hat{\rho}(\mathbf{d})$ is $\hat{\rho}(\mathbf{0}) = \mathbf{1} - \hat{\theta}_1 / \hat{\theta}_2$. For the Gaussian fit to the residual semivariogram values this maximum value is 0.44 and for the spherical model it is 0.52. The corresponding estimates from the anomaly fits are 0.79 for the Gaussian fit and 0.86 for the linear fit. The presence of trend in the anomalies has a dramatic effect on the estimates of spatial correlation.

Recall again that the station separation distance for San Diego and Los Angeles is 181 km. Inserting this distance into both the spherical and the Gaussian fitted models for the anomaly residuals results in an estimated spatial correlation of 0.40. With this estimated spatial correlation, the estimated value of ρ_{12} from Table 2 is 0.80, the same as the actual calculated value from the two time series. Thus, this relatively small estimated spatial correlation for these two stations is entirely consistent with the very large apparent spatial correlation calculated from the time series.

6. Concluding Remarks

This paper builds on much previous work on spatial modeling in both the meteorological and the geoscience literature. An essential ingredient in spatial modeling is the proper estimation of spatial correlations. The primary goal of this paper is to emphasize statistically valid methods for estimating spatial correlation. Statistical models can include a wide range of trend and stochastic components, but failure to account for any of them can seriously bias the estimation of spatial correlations and, ultimately, defeat the goal of optimal interpolation or averaging.

There is a vast literature on semivariogram and structure function modeling, but much more research is needed. The specific spatial models that are most appropriate for modeling many meteorological variables, including temperature anomalies, awaits further substantive analysis. Application of these methods to regions of the globe with sparse data requires some knowledge of the models that have been found useful in data-rich regions of the globe. The investigation of joint spatial/temporal models is still in its infancy.

Acknowledgments: Appreciation is expressed to the editor and the referees for substantive comments on an earlier version of this paper. Collaborative efforts with Sabyasachi Basu and Wayne A. Woodward on spatial-temporal issues is also gratefully acknowledged. Support for this work was provided by Contract DE FG03 93ER61645 from the U.S. Department of Energy.

References

Armstrong, M. and Diamond, P. (1984). "Testing Variograms for Positive-Definiteness," *Mathematical Geology*, 16, 407-421.

Basu, S. And Gunst, R.F. (1993). "The Effects of Influential Observations on classical and Robust Semivariogram Estimation," Technical Report No. SMU/DS/TR- , Department of Statistical Science, Dallas, TX: Southern Methodist University.

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.

Brockwell, P.J. and Davis, R.A. (1991). *Time Series: Theory and Methods*. New York: Springer-Verlag.

Brown, T.J. and Eischeid, J.K. (1992). "An Examination of Spatial Statistical Techniques for Interpolation of Gridded Climate Data," *Proceedings of the 12th Conference on Probability and Statistics in the Atmospheric Sciences,* J39-J42.

Buell, C.E. (1972), "Correlation Functions for Wind and Geopotential on Isobaric Surfaces," *Journal of Applied Meteorology*, 11, 51-59.

Cressie, N. (1984). "Toward Resistant Geostatistics," in G. Verly, M. David, A.G. Journel, and A. Marechal (eds.), *Geostatistics for Natural Resource Characterization*. Dordrecht, Holland: D. Reidel Publishing Co.

Cressie, N. (1985). "Fitting Variogram Models by Weighted Least Squares," *Mathematical Geology*, 17, 563-586.

Cressie, N. (1989). "Geostatistics," *The American Statistician*, 43, 197-202. Comment and Reply, *The American Statistician*, 44, 255-258.

Cressie, N. (1990). "The Origins of Kriging," *Mathematical Geology*, 22.

Cressie, N. (1991). *Statistics for Spatial Data.* New York: John Wiley and Sons, Inc.

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

Gandin, L.S. (1963). *Objective Analysis of Meteorological Fields.* Leningrad: Gidrometeorologicheskoe Izdatel'stvo. Translated by the Israel Program for Scientific Translation, Jerusalem, 1965.

Goldberger, A.S. (1962). "Best Linear Unbiased Prediction in the Generalized Linear Regression Model," *Journal of the American Statistical Association*, 57, 369-375.

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

Jones, P.D., Raper, S.C.B., Cherry, B.S.G., Goodess, C.M., Wigley, T.M.L., Santer, B., Kelly, P.M., Bradley, R.S., and Diaz, H.F. (1991). "An Updated Global Grid Point Surface Air Temperature Anomaly Data Set: 1851-1990," Technical Report ORNL/CDIAC-37, NDP-020/R1, Oak Ridge, TN: Oak Ridge National Laboratory.

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

Journel, A.G. and Rossi, M.E. (1989). "When Do We Need a Trend Model in Kriging," *Mathematical Geology*, 21, 715-739.

Julian, P.R. and Thiebaux, H.J. (1975). "On Some Properties of Correlation Functions Used in Optimum Interpolation Schemes," *Monthly Weather Review*, 103, 605-616.

Kagan, R.L. (1979). *Averaging Meteorological Fields*. Leningrad: Gidrometeoizdat. 212 pp. (in Russian).

Katz, R.W. (1988). "Statistical Procedures for Making Inferences about Climate Variability, *Journal of Climate*, 1, 1057-1064.

Kolmogorov, A.N. (1941). "The Local Structure of Turbulence in an Incompressible Fluid at Very Large Reynolds Numbers," *Doklady Akademii Nauk SSSR*, 30, 301-305.

Matheron, G. (1962). *Traite de Geostatistique Appliquee, Tome I. Memoires du Bureau de Recherches Geologiques et Minieres, No. 14*, Paris: Editions Technip.

Matheron, G. (1963). "Principles of Geostatistics," *Economic Geology*, 58, 1246-1266.

Press, W.H., Flannery, B.P., Teukolsky, S.A., and Vetterling, W.T. (1989). *Numerical Recipes: The Art of Scientific Computing (FORTRAN Version)*. Cambridge: Cambridge University Press. 702 pp.

S-Plus (1991). *S-Plus User's Manual, Volume 1, Version 3.0*. Seattle, WA: Statistical Sciences, Inc.

Stein, A. and Corsten, L.C.A. (1991). "Universal Kriging and Cokriging as a Regression Procedure," *Biometrics*, 47, 575-587.

Thiebaux, H.J. (1977). "Extending Estimation Accuracy with Anisotropic Interpolation," *Monthly Weather Review*, 105, 691-699.

Thiebaux, H.J. and Pedder, M.A. (1987). *Spatial Objective Analysis: With Applications in Atmospheric Science*. New York: Academic Press.

Vinnikov, K.Y., Groisman, P.Y., and Lugina, K.M. (1990). "Empirical Data on Contemporary Global Climate Changes (Temperature and Precipitation)," *Journal of Climate*, 3, 662-667.

Wiener, N. (1949). *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. Cambridge, MA: MIT Press.

Wold, H. (1938). *A Study in the Analysis of Stationary Time Series*. Uppsala: Almqvist and Wiksells.

List of Tables

Table 1. Temporal and Spatial Effects on the Correlation Between Two Time Series.
Table 2. Correlation Values Using Estimates From the Anomalies in Figure 2.
Table 3. Comparison of Estimated Semivariogram Parameters.

List of Figures

Figure 1. U.S. Station-Pair Correlations: Annual Anomalies 1920-1980.
Figure 2. Annual Temperature Anomalies for San Diego and Los Angeles.
Figure 3. U.S. Station-Pair 1980 Annual Anomaly Differences.
Figure 4. Smoothed Fit to 1980 U.S. Annual Anomalies.
Figure 5. Trend Removal in 1980 U.S. Annual Anomalies.
Figure 6. Semivariogram Values: U.S. Annual Anomaly Residuals, 1951-1980.
Figure 7. Semivariogram Model Fits: U.S. Annual Anomaly Residuals, 1951-1980.
Figure 8. Semivariogram Values: U.S. Annual Anomalies, 1951-1980.
Figure 9. Semivariogram Model Fits: U.S. Annual Anomalies, 1951-1980.
Figure 10. Estimated Spatial Correlations: U.S. Annual Anomalies, 1951-1980.

Table 1. Temporal and Spatial Effects on the Correlation Between Two Time Series.

$\phi = .3$

ρ_t/ρ_s	.1	.3	.5	.7	.9
.1	.10	.20	.29	.39	.48
.3	.20	.30	.40	.49	.59
.5	.31	.41	.50	.60	.69
.7	.41	.51	.60	.70	.80
.9	.52	.61	.71	.80	.90

$\phi = .5$

ρ_t/ρ_s	.1	.3	.5	.7	.9
.1	.10	.19	.27	.36	.44
.3	.21	.30	.39	.47	.56
.5	.33	.41	.50	.59	.67
.7	.44	.53	.61	.70	.79
.9	.56	.64	.73	.81	.90

$\phi = .7$

ρ_t/ρ_s	.1	.3	.5	.7	.9
.1	.10	.17	.24	.30	.37
.3	.23	.30	.37	.44	.50
.5	.36	.43	.50	.57	.64
.7	.50	.56	.63	.70	.77
.9	.63	.70	.76	.83	.90

$\phi = .9$

ρ_t/ρ_s	.1	.3	.5	.7	.9
.1	.10	.13	.16	.20	.23
.3	.27	.30	.33	.36	.40
.5	.44	.47	.50	.53	.56
.7	.60	.64	.67	.70	.73
.9	.77	.80	.84	.87	.90

Table 2. Correlation Values Using Estimates From the Anomalies in Figure 2.

ρ_s								
.1	**.2**	**.3**	**.4**	**.5**	**.6**	**.7**	**.8**	**.9**
.75	.77	.79	.80	.82	.84	.86	.88	.90

Table 3. Comparison of Estimated Semivariogram Parameters.

Gaussian Model			Spherical Model		
Nugget	Sill	Range	Nugget	Sill	Range
Isotropic Semivariogram Values					
.091	.163	591	.079	.163	1,210
North-South Semivariogram Values					
.088	.156	602	.089	.158	1,168
East-West Semivariogram Values					
.094	.167	579	.086	.167	935

Fig. 1. U.S. Station-Pair Correlations: Annual Anomalies 1920-1980.

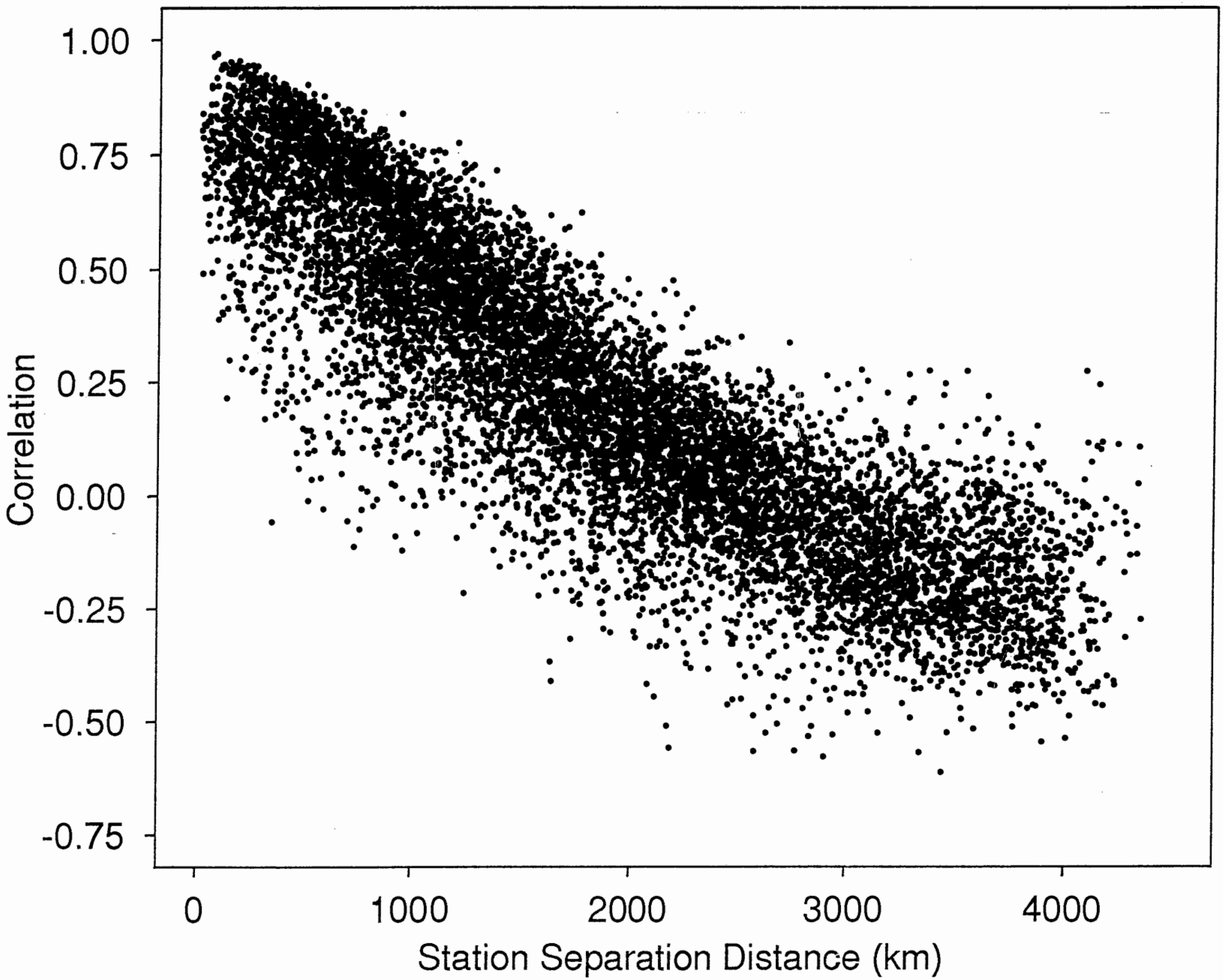

Fig. 2. Annual Temperature Anomalies for San Diego and Los Angeles.

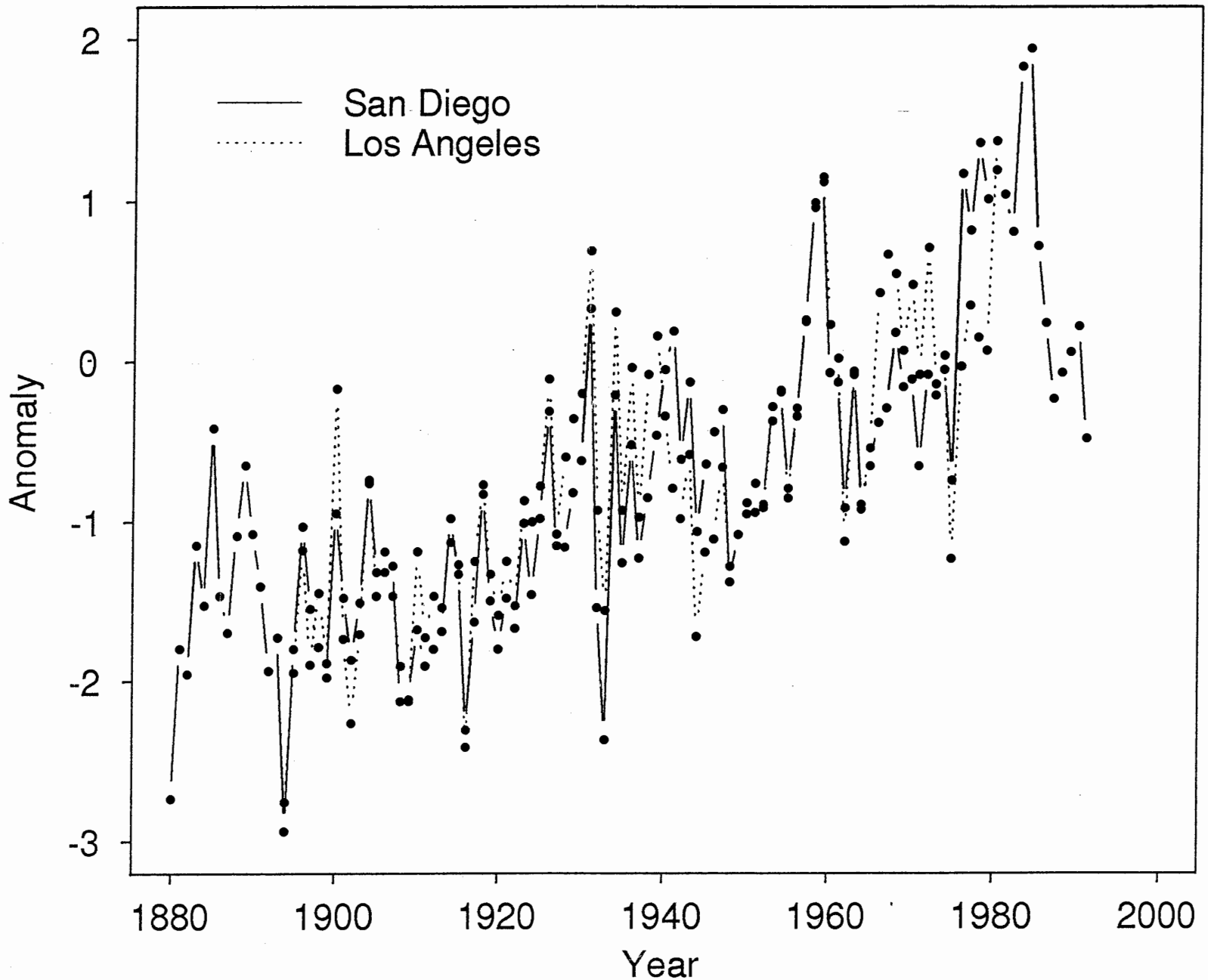

Fig. 3. U.S. Station-Pair 1980 Annual Anomaly Differences.

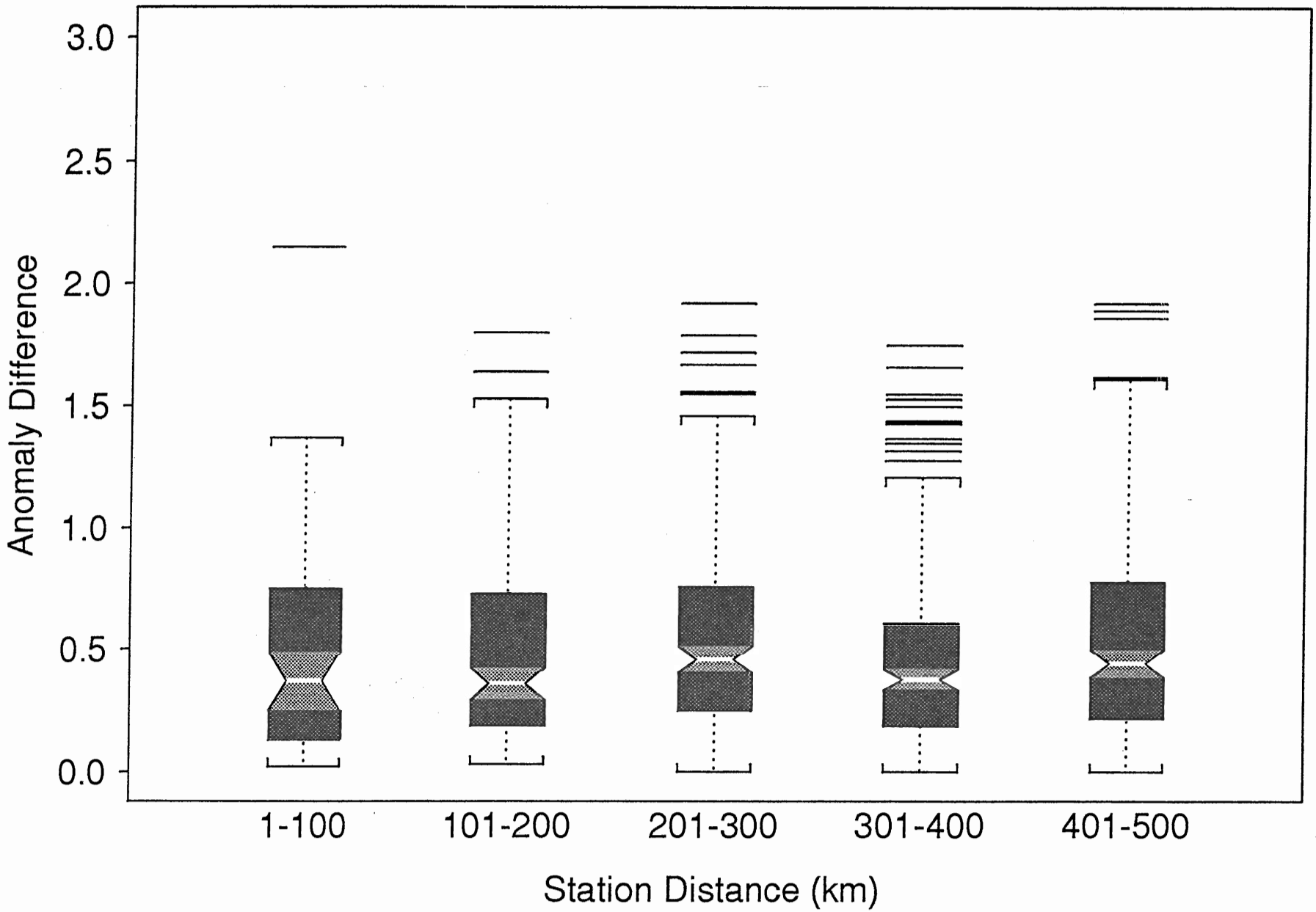

Fig. 4. Smoothed Fit to 1980 U.S. Annual Anomalies.

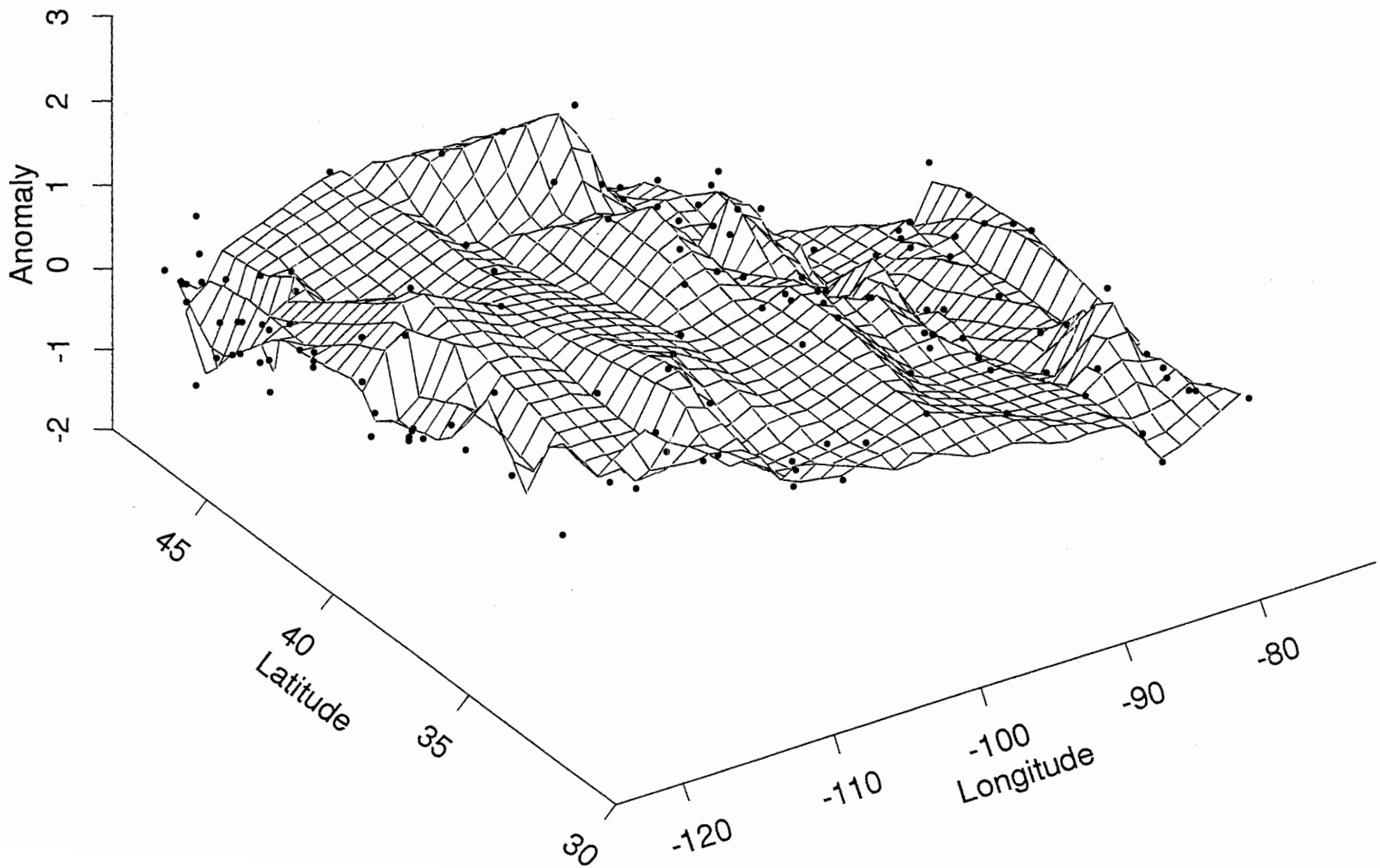

Fig. 5. Trend Removal in 1980 U.S. Annual Anomalies.

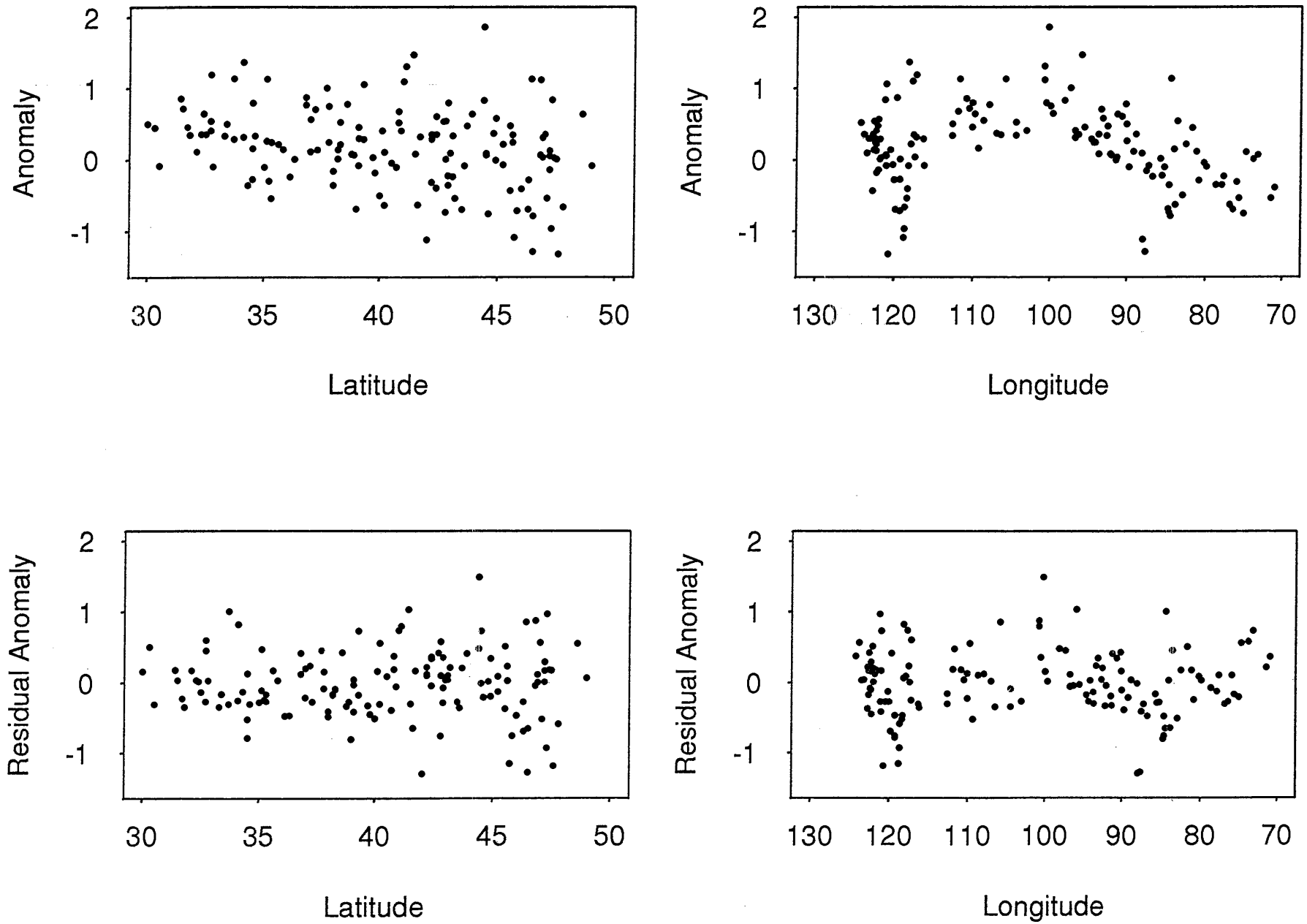

Fig. 6. Semivariogram Values: U.S. Annual Anomaly Residuals, 1951-1980.

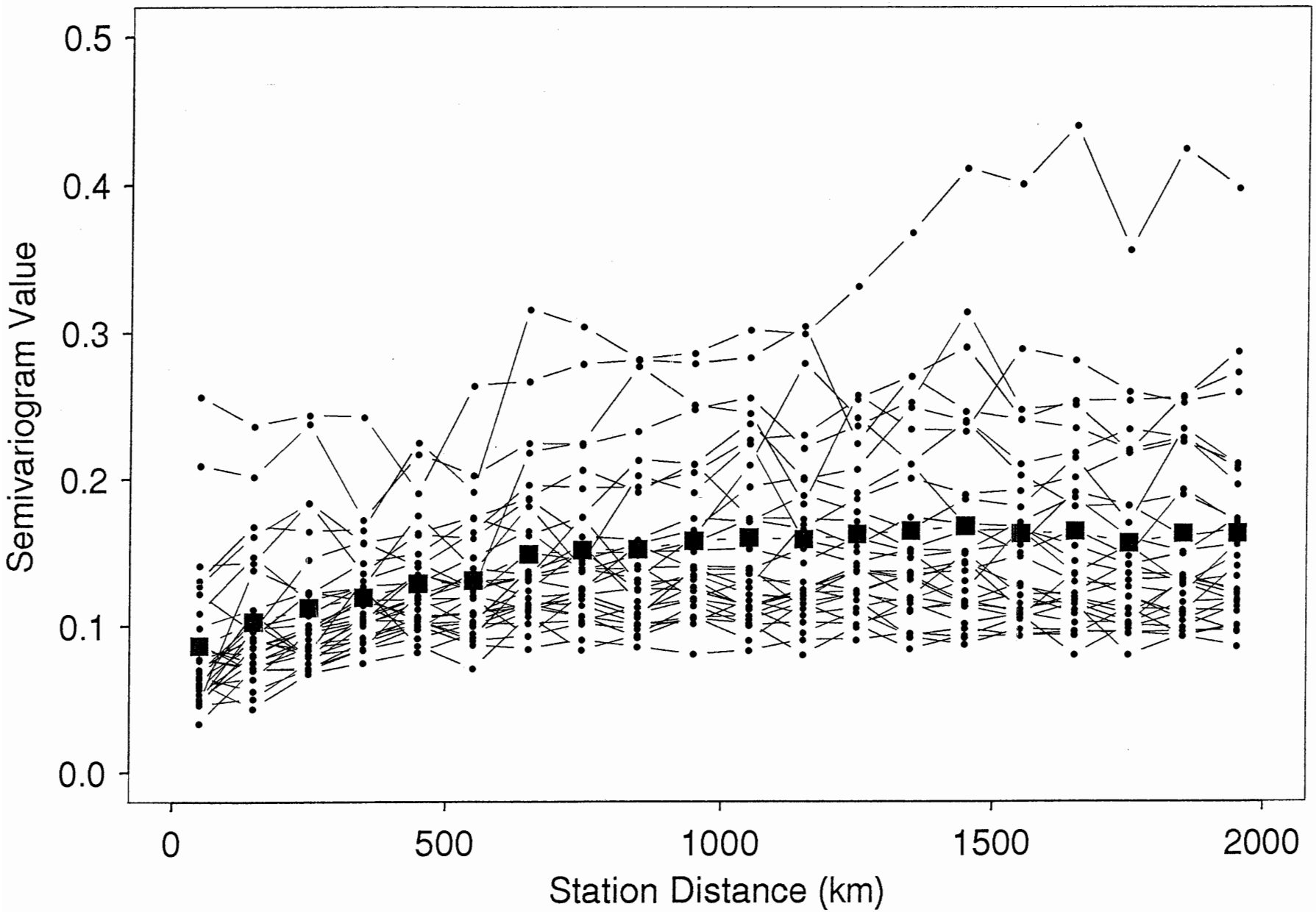

Fig. 7. Semivariogram Model Fits: U.S. Annual Anomaly Residuals, 1951-1980.

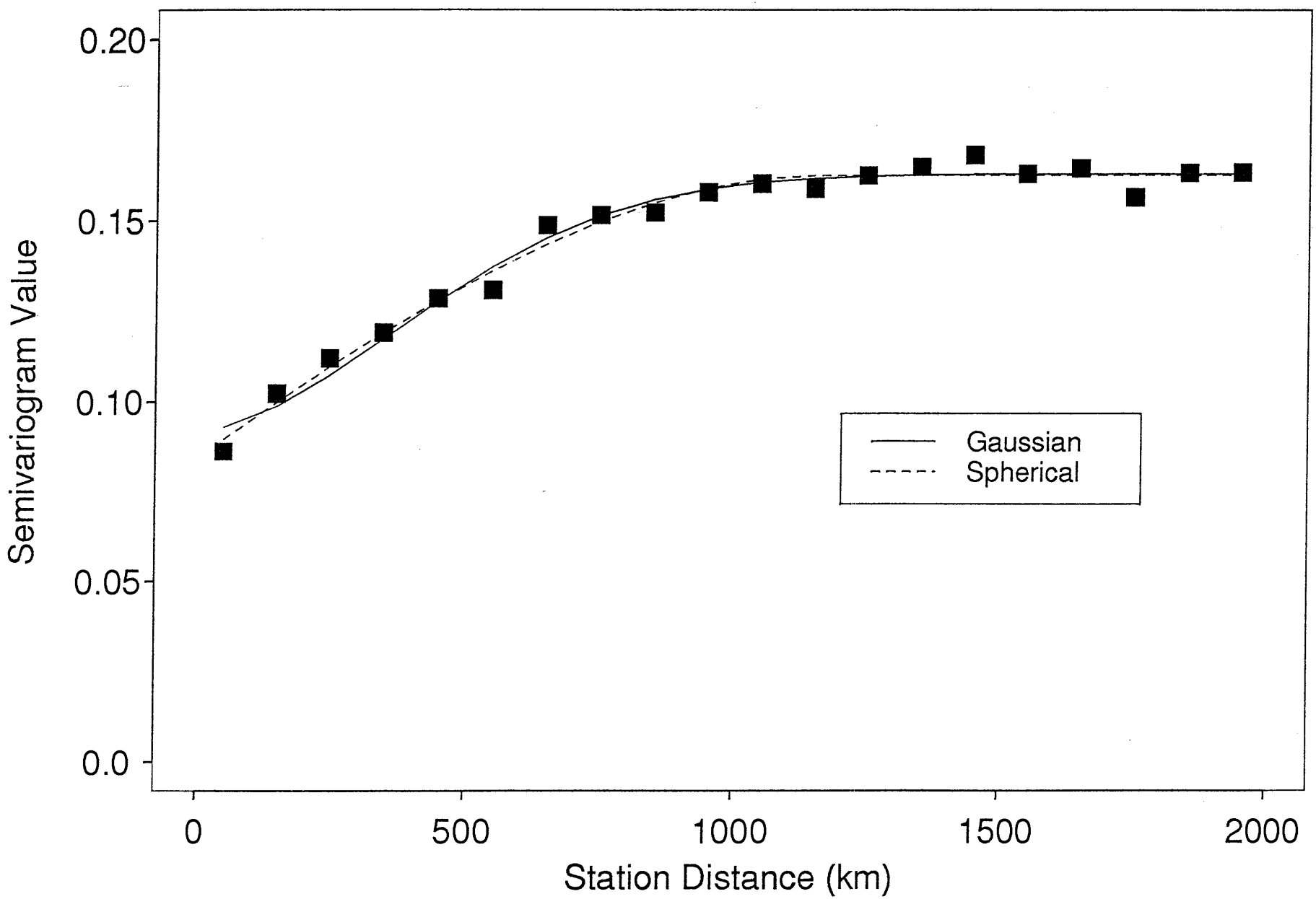

Fig. 8. Semivariogram Values: U.S. Annual Anomalies, 1951-1980.

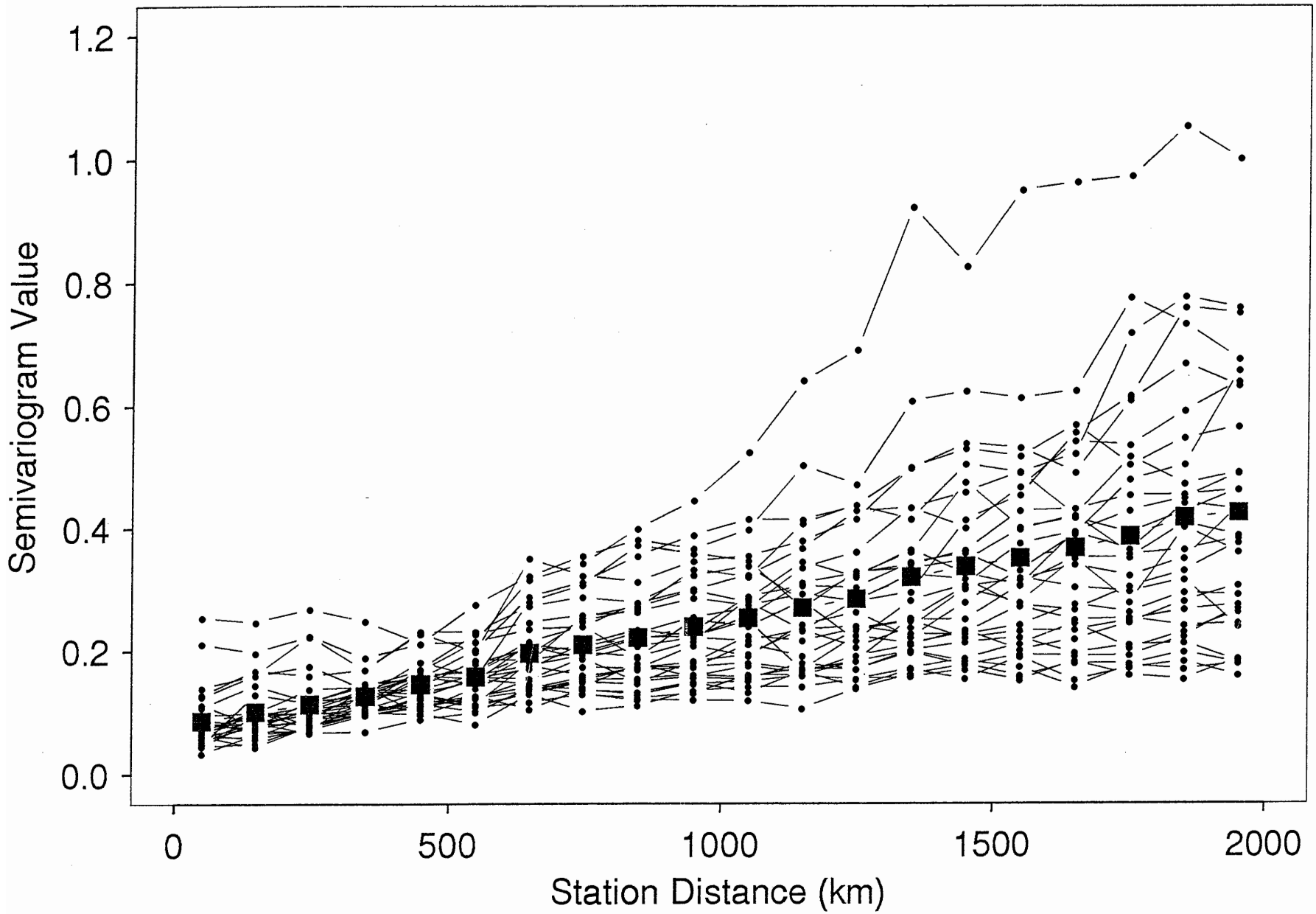

Fig. 9. Semivariogram Model Fits: U.S. Annual Anomalies, 1951-1980.

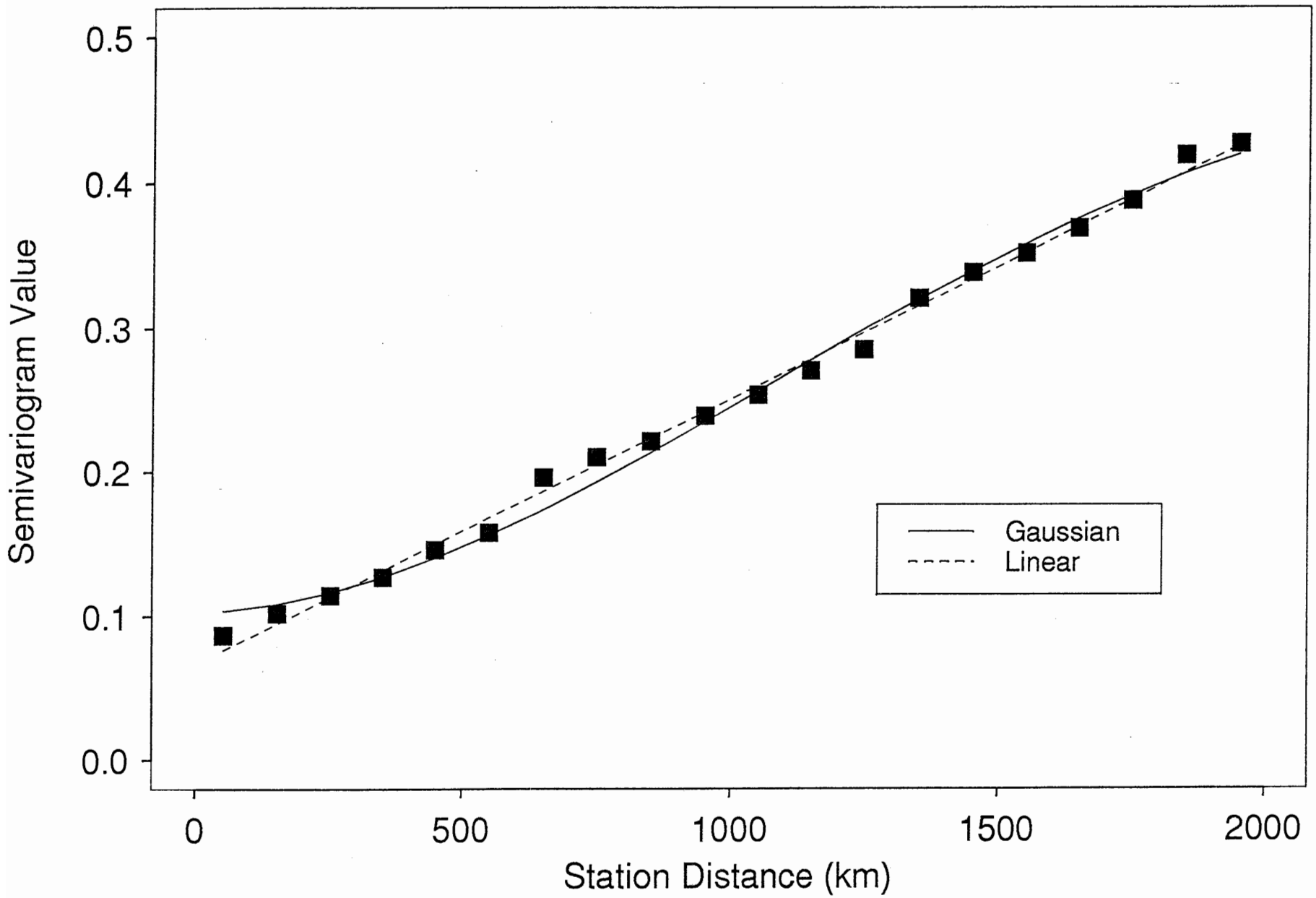

Fig. 10. Estimated Spatial Correlations: U.S. Annual Anomalies, 1951-1980.