Spatiotemporal Model Identification for a Class of Continuous Space-Time Processes

Molly I. Hartfield
Radian International
Austin, TX 78720-1088
(molly_hartfield@radian.com)

Richard F. Gunst
Department of Statisitcal Science
Southern Methodist University
Dallas, TX 75275-0332
(rgunst@mail.smu.edu)

ABSTRACT

Environmental data are routinely collected at irregularly spaced monitoring stations and at intermittent times, times which may differ by location. Dynamic temporal models of climatological variates, models for contaminant migration, and models for numerous other environmental processes require increasingly more complex stochastic structure than classical linear and generalized linear models. In this paper a class of continuous-time, continuous-space statistical models is introduced to accommodate many of these more complex environmental processes. The models incorporate temporal and spatial variability in a cohesive manner by assuming that temporal change follows a stochastic differential equation with possibly temporally and spatially correlated errors. A wide range of ARIMA temporal models and geostatistical spatial models are included in the class of models investigated. Techniques for identifying the structure of the temporal and spatial components of this class of models are detailed. The model identification techniques are based on transformations of data to remove large-scale structure, thereby facilitating graphical identification of small-scale temporal and spatial structure.

KEY WORDS: ARIMA, Geostatistics, Intrinsic Random functions, Kriging, REML

1. INTRODUCTION

Spatiotemporal data often are collected across a geographic region at multiple time points. Such data may be collected in order to construct regional maps of an attribute of interest at specified time points, to characterize changes in the attribute over time, or to predict values of the attribute at future times across the region. A critical feature of many spatiotemporal applications is that measurements of the attribute of interest are taken at irregular spatial locations and irregular temporal intervals. Continuous-time, continuous-space models and corresponding analytic methods are needed to adequately model these space-time processes. In this article, a new class of spatiotemporal models which naturally includes space and time through the application of stochastic differential equations is introduced and methods for model identification are discussed.

Rouhani and Myers (1990) and Rouhani et al. (1992) discuss some of the problems associated with the analysis of space-time geohydrological data. The temporal and spatial sampling scales differ, complicating the definition of common scales of variation and often leading to a disparity in the amount of data available in the two domains. Because time series methods are based on the notion of ordered data, spatial data do not easily conform to time series methods. Conversely, adding time as a dimension to a spatial statistical analysis often does not take advantage of the ordered structure of temporal data.

Bennet (1974), Cliff and Ord (1974), Martin and Oeppen (1975), Ali (1979), Pfeifer and Deutsch (1980a, 1980b, 1981) and DeGooijer and Anderson (1985) all investigate spatiotemporal models based upon the general form of a discrete-time ARIMA model. Rouhani and Wackernagel (1990) and Rouhani et al. (1992) model time series at various sampling locations as separate but correlated one-dimensional random processes. Bilonick (1985) and Rouhani and Hall (1989) apply three-dimensional kriging to spatiotemporal data. Haas (1995) conducts three-dimensional "moving-cylinder" spatiotemporal kriging. Oehlert (1993), using a similar idea, tiles

the region of interest with small rectangles, and within each tile fits a separate time-trend model. Switzer (1989) and Loader and Switzer (1992) use repeat measurements in time to increase the specificity among estimated spatial covariances. In the same spirit, but using different methods, Guttorp, Sampson and Newman (1992) take advantage of repeat measurements in time to estimate the covariance structure of a nonstationary spatial process.

Stein (1986), Haslett and Raftery (1989), Host, Omre and Switzer (1995), Solna and Switzer (1996) and Huang and Cressie (1996) all propose spatiotemporal models which can be seen to fit within a general linear model (GLM) framework. Not all of the models in these papers contain deterministic components, and the stochastic components of the models, which reflect both spatial and temporal variability, are defined and estimated somewhat differently by the various authors. Given the GLM framework, however, once the spatiotemporal covariance matrix has been specified, the subsequent prediction and estimation methods are equivalent to best linear unbiased prediction and estimation (e.g., Christensen 1991). All these authors assume models with separable covariance structures (e.g., Bogaert 1996, Rodriguez-Iturbe and Mejia 1974): the spatiotemporal correlation function is the product of a spatial correlation function ρ_S and a temporal correlation function ρ_T.

In an effort to take advantage of the ordered structure of the temporal domain without oversimplifying or de-emphasizing the spatial aspects of the process, a class of models which is based on stochastic differential equations in time is introduced in Section 2. These models do not require constant sampling intervals in space or in time. Temporal and spatial model identification procedures are discussed in Sections 3-6. Concluding remarks are made in Section 7.

2. A CLASS OF SPATIOTEMPORAL MODELS

Let Z(s,t) be a random variable that is observable at location **s** at time t, where **s** and t vary over continuous index sets $\mathbf{S} \subset \mathbb{R}^k$ (k = 1, 2, or 3) and $\mathbf{T} \subset \mathbb{R}$, respectively. The proposed class of models, which is based on the assumption of an underlying d^{th}-order stochastic differential equation in time, can be expressed as

$$Z(\mathbf{s},t) = \mu_\alpha(\mathbf{s},t) + g_\beta\{\mathbf{Y}(\mathbf{s}), t\} + W_\xi^d(\mathbf{s},t) + e_\sigma(\mathbf{s},t), \tag{1}$$

where μ_α is a nonstochastic function of explanatory variables, location, and time; g_β is a nonstochastic function of time and a stochastic function of independent spatial processes $Y_1(\mathbf{s}), \ldots, Y_k(\mathbf{s})$; W_ξ^d is a continuous-time, zero-mean, spatiotemporal ARIMA(p, d, q) error process; and e_σ is a zero-mean spatiotemporal random field representing microscale variability and measurement errors. The model component $\mu_\alpha(\mathbf{s},t) + g_\beta\{\mathbf{Y}(\mathbf{s}), t\}$ is assumed to be d times differentiable with respect to time. The zero-mean, second-order stationary spatial processes $Y_i(\mathbf{s})$ are mutually independent with

$$\mathrm{Cov}[Y_i(\mathbf{s}), Y_j(\mathbf{s}+\mathbf{g})] = \begin{cases} \mathrm{Cov}_{Y_i}(\|\mathbf{g}\|) & i = j \\ 0 & i \neq j \end{cases}.$$

For q = 0, p = 0, 1, and d = 0, 1, 2, the W_ξ^d process has the following representations

$$W_\xi^d = \begin{cases} a_\xi(\mathbf{s},t) & d = 0 \\ \int_{t_0}^{t} a_\xi(\mathbf{s},u)\,du & d = 1 \\ \int_{t_0}^{t}\int_{t_0}^{v} a_\xi(\mathbf{s},u)\,du\,dv & d = 2 \end{cases}$$

with

$$a_\xi(\mathbf{s},t) = \begin{cases} b_\omega(\mathbf{s},t) & p = 0 \\ \int_0^t \phi^{t-u} b_\omega(\mathbf{s},u)\,du & p = 1 \end{cases}$$

and zero-mean, second-order spatially correlated disturbances b_ω having

$$\mathrm{Cov}[b_\omega(\mathbf{s},t), b_\omega(\mathbf{s}+\mathbf{g}, t+h)] = \begin{cases} \mathrm{Cov}_b(\|\mathbf{g}\|) & h = 0 \\ 0 & h \neq 0 \end{cases}.$$

The errors are white-noise with

$$\mathrm{Cov}(e_\sigma(\mathbf{s},t), e_\sigma(\mathbf{s}+\mathbf{g}, t+h)) = \begin{cases} \sigma_{ee} & \mathbf{g} = \mathbf{0}, h = 0 \\ 0 & \text{otherwise} \end{cases} .$$

Representations of W_ξ^d and $a_\xi(\mathbf{s},t)$ are readily expressed for other choices of p, d, and q but the above ones illustrate the richness of the model possibilities and will be the focus of this paper. Similarly, one could expand the model even further by adding a term like g_β for spatial large-scale structure with random temporal coefficients but that too is beyond the scope of this work.

Consistent with the approach taken in the spatial statistics literature (Journel and Huijbregts 1978, Cressie 1993), the spatiotemporal stochastic process (1) permits different scales of variation. The *large-scale* variation includes the nonstochastic mean structure $\mu_\alpha(\mathbf{s},t)$ and the component $g_\beta\{\mathbf{Y}(\mathbf{s}), t\}$ which is nonstochastic with respect to time but may be a stochastic function of location. The *small-scale* variation $W_\xi^d(\mathbf{s}, t)$ often represents smooth (L_2-continuous) stochastic fluctuations about the large-scale structure. Small-scale variation generally represents local or regional stochastic variation in space and short-term temporal variation. This type of variation can include *microscale* spatial variation -- varaiation at a scale smaller than the smallest distance between locations -- in which case would contain a discontinuous jump at the origin, the *nugget effect*. The error variation $e_\sigma(\mathbf{s}, t)$ represents measurement imprecision.

An alternative representation of the class of models given by (1) is in terms of the assumed underlying stochastic differential equation (SDE). The underlying SDE with respect to time can be written as:

$$Z^{(d)}(\mathbf{s}, t) = f(\mathbf{s}, t) + a_\xi(\mathbf{s}, t), \tag{2}$$

where $Z^{(d)}(\mathbf{s}, t) = \frac{d^d Z(\mathbf{s},t)}{(dt)^d}$ and $f(\mathbf{s}, t) = \frac{d^d[\mu_\alpha(\mathbf{s},u)+g_\beta\{y(\mathbf{s}),u\}]}{dt^d}$.

Equation (2) is of the same form as the differential equations associated with stochastic growth curve models in Seber and Wild (1989, Chapter 7) except that in model (2) a spatial index has been included and in the stochastic growth curve context d is constrained to be 1. Model (1)

is the solution to the spatiotemporal SDE given by (2), plus a microscale variation term $e_\sigma(\mathbf{s}, t)$. Often, growth is modeled in terms of the relative rate of change rather than in terms of the magnitude of change. In such a case, the spatiotemporal SDE (2) can be written in terms of the logs of the process as:

$$[log\{Z(t)\}]^{(d)} = f(\mathbf{s}, t) + a_\gamma(\mathbf{s}, t).$$

Solna and Switzer's (1996) model can be expressed in terms of the class of models proposed here by writing it as

$$Z(\mathbf{s},t) = \mu(\mathbf{s}) + Y(\mathbf{s})t + b_\omega(\mathbf{s},t),$$

where $\mu(\mathbf{s},t) = \mu(\mathbf{s})$ and $g_\beta\{\mathbf{Y}(\mathbf{s}), t\} = Y(\mathbf{s})t$, the latter being a linear time trend with location-specific slopes generated by a spatial stochastic process. The small-scale component $W_\xi^d(\mathbf{s}, t) = b_\omega(\mathbf{s},t)$, where $b_\omega(\mathbf{s},t)$ is assumed to be a separable spatiotemporal process, white noise with respect to time for a fixed location and intrinsically stationary with respect to space for a fixed time.

The Huang and Cressie (1996) and Haslett and Raftery (1989) models also can be seen to fall within the class of models represented by (1). The Huang and Cressie model is

$$Z(\mathbf{s},t_j) = \alpha_1 a(\mathbf{s},t_{j-1}) + \alpha_2 a(\mathbf{s},t_{j-2}) + \ldots + \alpha_p a(\mathbf{s},t_{j-p}) + b(\mathbf{s},t_j),$$

where the $\mathbf{b}_t = (\, b(\mathbf{s}_1,t) \quad b(\mathbf{s}_2,t) \quad \ldots \quad b(\mathbf{s}_m,t)\,)'$ are independent, identically distributed $N(\mathbf{0},\Sigma_\mathbf{b})$ and $\Sigma_\mathbf{b} = Cov[b(\mathbf{s}_i,t),b(\mathbf{s}_j,t)]$ is the matrix of second-order spatial stationary covariances, assumed constant over time. The Haslett and Raftery model is

$$Z(\mathbf{s},t) = \mu(\mathbf{s}) + \nabla^{-d}\phi(B)^{-1}\theta(B)b(\mathbf{s},t),$$

where $\mu(\mathbf{s})$ is a location-specific mean, the $\nabla^{-d}\phi(B)^{-1}\theta(B)b(\mathbf{s},t)$ term is of the form of an ARIMA(p,d,q) model, and the $\mathbf{b}_t$ are independent, identically distributed $N(\mathbf{0},\Sigma_\mathbf{b})$. The Haslett

and Raftery model generalizes the Huang and Cressie model from a spatially correlated autoregressive model to a spatially correlated ARIMA spatiotemporal model with a possibly nonstationary mean. Further, the Haslett and Raftery model allows for the possibility that d is a fractional value, allowing for long-memory temporal processes to be accommodated (e.g., Hosking 1981).

3. MODEL IDENTIFICATION

The issue of whether to include large-scale variation terms in geostatistical universal kriging models has been debated by practitioners (e.g., Journel 1985, Journel & Rossi 1989). Cressie (1993) stresses the lack of uniqueness in the specification of deterministic and stochastic components of spatial statistical models. The specification often is decided based on how much of the variability to attribute to large-scale effects so that the covariances appear to be spatially stationary. Similarly, in the time series context decomposing the model into large-scale and small-scale variation also is a challenging practical question, despite the existence of the Wold decomposition theorem (e.g., Brockwell and Davis 1991, Section 5.7), since the form of the decomposition often is not clear in practice.

In the spatiotemporal setting, specifying a decomposition becomes even more difficult because of the ambiguity that is introduced with regard to scales of variability. For example, some effects may be considered large-scale with respect to time but small-scale with respect to space. Although data generally provide some direction with regard to an appropriate decomposition, alternative models which describe the data equally well are common, as pointed out by Cressie (1993) in a spatial context.

The class of models (1) emphasizes temporal components in order to permit the modeling of dynamic environmental systems. The lack of uniqueness of spatiotemporal model specification and

the broad scope of this class of models introduce many complex issues regarding model specification and parameter estimation. Because these issues are so broad, the scope of the current investigation is limited to model specification and identification. In order to better focus on some of the key model specification issues, attention is restricted to small-scale ARIMA$(p, d, 0)$ components $W_\xi^d(\mathbf{s}, t)$ with $p = 0$ or 1 and $d = 0$, 1, or 2. Even this restriction will allow the nonuniqueness issues to be clearly illustrated.

One approach that has been found useful in the specification and parameter estimation of the class of models represented by (1) consists of the following steps.

1. Characterize the large-scale spatial and temporal structure $\mu_\alpha(\mathbf{s}, t) + g_\beta\{\mathbf{Y}(\mathbf{s}), t\}$ of the model using exploratory data analysis.
2. Specify the temporal structure of $W_\xi^d(\mathbf{s}, t)$ through comparisons of theoretical and empirical primary increment semivariograms.
3. Specify the spatial structure of $W_\xi^d(\mathbf{s}, t)$ through the choice of a generalized covariance or semivariogram model.
4. Estimate the parameters associated with $W_\xi^d(\mathbf{s}, t)$ and $e_\sigma(\mathbf{s}, t)$.
5. Specify the functional form and estimate the corresponding spatial and temporal parameters of $\mu_\alpha(\mathbf{s}, t) + g_\beta\{\mathbf{Y}(\mathbf{s}), t\}$.

Two principles guided the formulation of these steps. First, the more extensive the scientific information that is available about the behavior of the process, the richer should be the specification of the large-scale components because they represent structural components of the process. Second, since the emphasis in this work is on modeling dynamic environmental processes, the temporal components of the model are allowed to capture as much of the stochastic variability as possible prior to the specification and estimation of the spatial processes.

Thus the sequence of steps focuses on large-scale processes before specifying small-scale processes, and it introduces temporal components and fitting prior to the specifying and fitting of spatial processes. In the next two sections, the critical first three steps are discussed in more detail. Estimation of parameters and prediction using the fitted model is discussed in a companion paper, Hartfield and Gunst (2000).

4. CHARACTERIZING THE LARGE-SCALE STRUCTURE

As in any statistical analysis, exploratory data analysis (EDA) should precede formal modeling. Multivariate graphical displays can assist in delineating temporal and spatial model structures. Plots of $z(\mathbf{s},t)$ versus time, without regard for location, averaged across location, and separately by location may reveal a general temporal pattern, any consistent deviations from that pattern, and whether the pattern changes systematically as one moves across the region. Two- and three-dimensional plots of $z(\mathbf{s},t)$ versus location coordinates without regard for time, averaged across time, and separately by time can reveal general spatial patterns, variability in such patterns across time and whether the spatial pattern changes systematically in time.

This information may suggest whether certain effects should be modeled as deterministic or stochastic functions of space or time, the large-scale structure of model (1). For example, if the observations are increasing at the same rate across time and across all locations, this may suggest that a constant linear trend term should be included in the model; that is, $\mu(\mathbf{s},t)$ may be specified as $\mu(\mathbf{s},t) = \alpha_1 + \alpha_2 t$. On the other hand, if the observations appear to increase linearly in time but at different rates for different locations, this may suggest that a linear trend term should be included in the model, but with location-dependent coefficients $Y_1(\mathbf{s})$ and $Y_2(\mathbf{s})$.

Figure 1 is a plot of smoothed annual temperature anomalies for 138 stations in the contiguous United States from 1920 to 1980 (Gunst, Basu,and Brunell 1993). The actual

anomalies (differences in temperatures from a baseline) have been smoothed using a locally linear loess smoother in order to better highlight the temporal trends. As a function of time, the gross trends appear to be largely quadratic with location-dependent coefficients. It is not clear whether the apparent trends are due to large-scale structural trends or to autocorrelation. The techniques discussed in the next section will help clarify which model components are contributing to the quadratic features apparent in Figure 1.

[Insert Figure 1]

Figure 2 is consists of smoothed contour plots of anomalies versus latitude and longitude for 9 of the 61 years, 1960-1969. These plots are representative of the plots across the entire time span. The gross spatial gradients appear to be quadratic, with time-dependent parameters. The diversity of quadratic features suggests that each year is distinctively different from the prior and subsequent ones and that the features may not have a smooth representation across time.

[Insert Figure 2]

A model that incorporates the large-scale features observed in Figures 1 and 2 is (1) with $\mu_\alpha(\mathbf{s},t) = \mathbf{x}_\mathbf{s}'\alpha_t$ and $g_\beta\{\mathbf{Y}(\mathbf{s}), t\} = Y_1(\mathbf{s})t + Y_2(\mathbf{s})t^2$. In the definition of the mean μ, $\mathbf{x}_\mathbf{s}$ designates the vector (1, lat(**s**), long(**s**), lat(**s**)*long(**s**), lat(**s**)^2, long(**s**)^2) of components of a quadratic function of latitude and longitude. In this article, the means $\mu_\alpha(\mathbf{s},t)$ are assumed to have fixed regression coefficients for a given year. It is possible to model the spatial gradient coefficients α_t in a hierarchical framework (e.g., Royle and Berliner 1999) or as a stationary temporal random process. To do so would require proper accounting for nonperiodic but regularly occurring events like El Niño - Southern Oscillation (ENSO) events and nonregular events like volcanic eruptions, both of which affect temperature change. Neither of these options will be pursued in this work but they are natural extensions of it.

It is well known that estimation of the parameters of the mean structure prior to modeling the stochastic components of the model can lead to biased estimators of the spatial random field parameters (e.g., Cressie 1987). Gunst (1995) documents the substantial bias that can occur when spatial covariances are estimated from nonfiltered time series. Similar concerns in the estimation of linear models led to a number of alternatives to least squares and maximum likelihood approaches; a notable alternative being restricted (residual) maximum likelihood (e.g., Searle, Casella, McCulloch 1992). In the geostatistical literature intrinsic random function kriging (Matheron 1973) is a widely cited alternative. In its simplest form, intrinsic random function kriging is equivalent to modeling increments (contrasts) of the data. In the next two sections, this approach is shown to greatly aid the identification of the temporal and spatial small-scale components of (1).

5. SPECIFYING THE TEMPORAL STRUCTURE OF $\mathbf{W}_{\xi}^{d}$

In this section attention is concentrated on the use of *primary temporal increments* for filtering polynomial time trends such as those noted in Figure 1. This type of filtering aids in the specification of the degree of time trends in the large-scale component of model (1), aids in the specification of reasonable choices for p and d in the small-scale component of the model, and helps distinguish whether variability is due to the large-scale or to the small-scale components. In the next section, this type of temporal filtering is shown to greatly aid in the proper specification of spatial covariance functions.

Although the functional form of t in $\mu_{\alpha}(\mathbf{s},t)$ and $g_{\beta}\{\mathbf{Y}(\mathbf{s}),t\}$ in model (1) can be much more general, for ease of exposition in this article the large-scale temporal component in model (1) is restricted to a polynomial: $\mu_{\alpha}(\mathbf{s},t) = \sum_{r=0}^{2}\alpha_r(\mathbf{s})t^r$ and $g_{\beta}\{\mathbf{Y}(\mathbf{s}),t\} = \sum_{r=0}^{2}Y_r(\mathbf{s})t^r$, where some of the

$\alpha_r(\mathbf{s})$ and the $Y(\mathbf{s})$ terms may be identically zero. The postulated temporal trend observed in Figure 1 is included in this class of large-scale structures. In the context of filtering spatial data to determine its stochastic structure, Cressie (1987) proposes the use of primary spatial increments to filter polynomial gradients. Cressie (1988) uses temporal increments in a similar fashion to identify the degree of nonstationarity in a time series. The latter application provides a method for filtering temporal data in such a way that polynomial large-scale variation and nonstationary small-scale effects are eliminated.

To calculate primary temporal increments, define time lags $h_{ij} = |t_j - t_i|$, $t_i, t_j \in T$. Partition the time lags into bins B_1 through B_N, where $B_k = \{h_{ij} : (b_k - \tau) < h_{ij} \leq (b_k + \tau)\}$ and τ is the tolerance around the nominal lag b_k. The nominal lag corresponding to bin B_1 is b_1 and the remaining nominal lags ordinarily are multiples of b_1. For a fixed location $\mathbf{s}$, time t, and separation distance h, define primary temporal increments of order zero, one and two (higher-order filters are similarly defined) as:

$$
\begin{aligned}
P_0(\mathbf{s},t,h) &= Z(\mathbf{s},t+h) - Z(\mathbf{s},t) \\
P_1(\mathbf{s},t,h) &= h_1 Z(\mathbf{s},t+h_1+h) - (h+h_1)Z(\mathbf{s},t+h_1) + hZ(\mathbf{s},t) \\
&= h_1\{Z(\mathbf{s},t+h_1+h) - Z(\mathbf{s},t+h_1)\} - h\{Z(\mathbf{s},t+h_1) - Z(\mathbf{s},t)\} \\
P_2(\mathbf{s},t,h) &= h_1(h_1+h_2)\{h_2 Z(\mathbf{s},t+h_1+h_2+h) - (h+h_2)Z(\mathbf{s},t+h_1+h_2) + hZ(\mathbf{s},t+h_1)\} \\
&\quad - h(h+h_2)\{h_1 Z(\mathbf{s},t+h_1+h_2) - (h_1+h_2)Z(\mathbf{s},t+h_1) + h_2 Z(\mathbf{s},t)\} \\
&= h_1(h_1+h_2)[h_2\{Z(\mathbf{s},t+h_1+h_2+h) - Z(\mathbf{s},t+h_1+h_2)\} \\
&\quad - h\{Z(\mathbf{s},t+h_1+h_2) - Z(\mathbf{s},t+h_1)\}] \\
&\quad - h(h+h_2)[h_1\{Z(\mathbf{s},t+h_1+h_2) - Z(\mathbf{s},t+h_1)\} - h_2\{Z(\mathbf{s},t+h_1) - Z(\mathbf{s},t)\}],
\end{aligned}
\tag{3}
$$

respectively, where $h_1, h_2 \in B_1$. These reduce to Cressie's (1988) primary increments for equally spaced data in time and $h_1 = h_2 = 1$. In the large-scale temporal structure, $P_0(\mathbf{s},t,h)$ filters a constant mean, $P_1(\mathbf{s},t,h)$ filters a constant mean and a linear trend, and $P_2(\mathbf{s},t,h)$ filters a constant

mean and linear and quadratic trends. Not only do these primary temporal increments filter the large-scale polynomial temporal effects, $\mu_\alpha(\mathbf{s},t) + g_\beta\{\mathbf{y}(\mathbf{s}),t\}$, but they also accommodate certain forms of nonstationarity among the small-scale temporal effects, $W_\xi^d(\mathbf{s},t)$. As demonstrated below, plots of estimated semivariograms of primary temporal increments can reveal information about both $W_\xi^d(\mathbf{s},t)$ and the form of $\mu_\alpha(\mathbf{s},t) + g_\beta\{\mathbf{y}(\mathbf{s}),t\}$ with respect to time.

Define the temporal semivariogram $\delta(t,h)$ as:

$$\delta(t,h) = \frac{1}{2}E\left[\left\{W_\xi^d(\mathbf{s},t)+e_\sigma(\mathbf{s},t)\right\} - \left\{W_\xi^d(\mathbf{s},t+h)+e_\sigma(\mathbf{s},t+h)\right\}\right]^2. \quad (4)$$

Table 1 shows $\delta(t,h)$ for the six combinations of p and d considered in this work. In Table 1, $\sigma_{bb} = Cov_b(0)$ and $\lambda = -log(\phi)$, where ϕ is the autoregressive parameter when $p = 1$. As the table indicates, for models with stationary small-scale variation, $\delta(t,h)$ follows the form of a white noise or an exponential semivariogram model corresponding to the uncorrelated ($p = 0$) or autoregressive ($p = 1$) cases, respectively. Notice that $\delta(t,h)$ has an asymptote in both instances. For a stochastic walk, $\delta(t,h)$ is linear for large h. For a stochastic slope, $\delta(t,h)$ depends not only on the separation distance h, but also on t, and is cubic with respect to h. For the case $d = 2$ and $p = 1$, the expression for $\delta(t,h)$ is an approximation, but the impact of this approximation on the subsequent work has been found to be negligible.

[Insert Table 1]

Although $\delta(t,h)$ contains key information about the temporal structure of $W_\xi^d(\mathbf{s},t)$, if the large-scale structure of model (1) is not constant, $\delta(t,h)$ cannot be estimated directly. The expected values of squared primary temporal increments are shown below to be functions of $\delta(t,h)$. Hence, the information about the temporal structure of $W_\xi^d(\mathbf{s},t)$ can be assessed through averages of squared primary temporal increments.

Define the following variance functions of the primary increments:

$$\gamma_0(\mathbf{s},t,h) = \frac{1}{2}\mathrm{Var}[P_0(\mathbf{s},t,h)] \qquad \gamma_1(\mathbf{s},t,h) = \frac{\mathrm{Var}[P_1(\mathbf{s},t,h)]}{2hh_1(h_1+h)}$$
$$\gamma_2(\mathbf{s},t,h) = \frac{\mathrm{Var}[P_2(\mathbf{s},t,h)]}{2h_1h_2h(h_1+h_2)(h_2+h)^2(h_1+h_2+h)}. \tag{5}$$

Observe that γ_0 is a temporal semivariogram and that γ_1 and γ_2 can be considered temporal increment semivariograms. Estimators of these temporal increment semivariograms are

$$\hat{\gamma}_0(\mathbf{s},b_k) = \frac{1}{|N_0(\mathbf{s},b_k)|} \sum_{N_0(\mathbf{s},b_k)} \frac{\{P_0(\mathbf{s},t_i,h)\}^2}{2},$$

$$\hat{\gamma}_1(\mathbf{s},b_k) = \frac{1}{|N_1(\mathbf{s},b_k)|} \sum_{N_1(\mathbf{s},b_k)} \frac{\{P_1(\mathbf{s},t_i,h)\}^2}{2hh_1(h+h_1)}, \tag{6}$$

$$\hat{\gamma}_2(\mathbf{s},b_k) = \frac{1}{|N_2(\mathbf{s},b_k)|} \sum_{N_2(\mathbf{s},b_k)} \frac{\{P_2(\mathbf{s},t_i,h)\}^2}{2h_1h_2h(h_1+h_2)(h_2+h)^2(h_1+h_2+h)},$$

where $P_0(\mathbf{s},t_i,h)$, $P_1(\mathbf{s},t_i,h)$, $P_2(\mathbf{s},t_i,h)$ are primary increments calculated as in (3), $N_0(\mathbf{s},b_k)=\{(t_i,t_i+h) : h \in B_k\}$, $N_1(\mathbf{s},b_k)=\{(t_i, t_i \pm h_1, t_i \pm h_1 \pm h) : h_1 \in B_1, h \in B_k\}$, and $N_2(\mathbf{s},b_k)=\{(t_i, t_i \pm h_1, t_i \pm h_1 \pm h_2, t_i \pm h_1 \pm h_2 \pm h) : h_1,h_2 \in B_1, h \in B_k\}$. The $\pm$ in N_1 and N_2 indicate that increments are calculated both prior to $(-)$ and following $(+)$ t_i.

Using the nominal separation distance b_1 for the first bin in $\hat{\gamma}_1$ and nominal distances b_1 and $2b_1$ for the first two bins in $\hat{\gamma}_2$, the expectations of the individual terms being averaged in (6) are:

$$E[\hat{\gamma}_0(\mathbf{s},t,h)] = \tfrac{1}{2}\left[\sum_{r=1}^{2} \alpha_r(\mathbf{s})\{(t+h)^r - t^r\}\right]^2 + \tfrac{1}{2}\sum_{r=1}^{2} \sigma^2_{Y_r}(\mathbf{s})\{(t+h)^r - t^r\}^2 + \delta(t,h)$$

$$E[\hat{\gamma}_1(\mathbf{s},t,h)] = \tfrac{b_1h(b_1+h)}{2}\left\{\alpha_2^2(\mathbf{s}) + \sigma^2_{Y_2}(\mathbf{s})\right\} + \left[\tfrac{\delta(t,b_1)}{b_1} + \tfrac{\delta\{(t+b_1),h\}}{h} - \tfrac{\delta\{t,(b_1+h)\}}{b_1+h}\right] \tag{7}$$

$$E[\hat{\gamma}_2(\mathbf{s},t,h)] = \frac{1}{(b_1+h)}\left[h\left\{\frac{\delta(t,b_1)}{b_1}\right\} + (2b_1+h)\left\{\frac{\delta\{(t+b_1),b_1\}}{b_1}\right\} - (b_1+h)\left\{\frac{\delta(t,2b_1)}{2b_1}\right\}\right.$$
$$+ b_1\left\{\frac{\delta\{(t+b_1+h),h\}}{h}\right\} - 2b_1\left\{\frac{\delta\{(t+b_1),(b_1+h)\}}{b_1+h}\right\}$$
$$\left. + b_1\left\{\frac{\delta(t,(2b_1+h))}{2b_1+h}\right\}\right]$$

where $\sigma^2_{Y_r}(\mathbf{s}) = \mathrm{Var}[Y_r(\mathbf{s})]$.

If the large-scale temporal structure is a polynomial of order zero, then $E[\hat{\gamma}_0(\mathbf{s},t,h)] = \delta(t,h)$, which Table 1 indicates is constant or asymptotes if the process is stationary ($d = 0$), exhibits a positive linear trend if the process is a stochastic walk ($d = 1$), or exhibits a positive quadratic or cubic trend if the process is a stochastic slope ($d = 2$). On the other hand, if the large-scale temporal structure is a polynomial of order one or two with coefficients of sufficiently large magnitude, the first equation in (7) indicates that $E[\hat{\gamma}_0(\mathbf{s},t,h)]$ is at least a quadratic function with respect to h, regardless of whether the coefficients are deterministic or spatially stochastic, and regardless of the structure of the small-scale variation. Similar considerations can lead to characterizations of the expected behavior of the first and second primary increments using the expressions in (7) and Table 1.

Valuable supplementary statistics to the primary increments are the linvariogram and the quadvariogram (Cressie 1987):

$$L(\mathbf{s},b_h) = \frac{1}{h}\sum_{j=1}^{h-1}\hat{\gamma}_1(\mathbf{s},b_j) \qquad h = 2, 3, \ldots$$
$$Q(\mathbf{s},b_h) = \frac{6}{h^3}\sum_{j=1}^{h-2}\frac{(h-1-j)\,\hat{\gamma}_2(\mathbf{s},b_j)}{(j+1)} \qquad h = 3, 4, \ldots \tag{8}$$

If linear trends dominate the large-scale structure of the model, the linvariogram and the first primary increments will asymptote to the same value provided d = 0 or 1. If d = 2, the linvariogram and the first primary increment will diverge for large lags h. The quadvariogram and

the second primary increments asymptote to the same value for any of the models discussed in this paper. An examination of the behavior of the semivariogram, the first primary increments and linvariogram, and the second primary increments and quadvariogram thus can greatly aid in the identification of the temporal components of model (1).

Figure 3 contains graphs of the order 0, 1, and 2 primary increment averages for the first 30 lags of the temperature anomaly data. In this figure, the increment averages for each station at each lag are averaged across stations. One-year bin widths were used in the calculations. Smoothing of the order 0 primary increments indicate either a linearly or an exponentially increasing trend. The linear trend fits the overall plot well but does not conform well to the curvature apparent in the first few lags. The exponential trend captures the curvature in the first few lags and the overall trend but does not closely fit the last few plotted values. Note that any trend, including both a linear or an exponential trend, in the order 0 increments indicates either that the large-scale temporal trend is not constant across time or that the small-scale temporal variability is not white noise.

[Insert Figure 3]

The order 1 increment averages and the linvariogram values in Figure 3 asymptote to the same value, suggesting that the large-scale structure is at most linear (Cressie 1988). This also suggests, along with the trend in the average order 0 increment values, that the small-scale structure could be an AR(1), an ARIMA (0,1,0), or an ARIMA (1,1,0) process. From Table 1, these three small-scale processes are the only ones consistent with the order 0 increment plot. In addition, these three processes would all be filtered by the order 1 increments.

Figure 4 provides another perspective on the specification of the nonstationary temporal structure using smoothed temporal increments. Figure 1 is reproduced in Figure 4 to allow comparisons with the smoothed increment values. All the increment plots remove most of the

quadratic features of the smoothed anomalies; however, there is still prominent curvature in the order 0 increments graphed in Figure 4. Plots of the smoothed order 1 and order 2 increments are flat except for the first few lags. This apparent curvature is due to spatial effects, as is shown in the next section. For most of the lags, the smoothed order 1 increments are relatively constant when compared to the large curvature in Figure 1. This suggests, as did Figure 3, that the quadratic features observed in the smoothed anomalies are due to the small-scale component of model (1) and not to large-scale quadratic temporal trends.

[Insert Figure 4]

The theoretical temporal semivariogram functions and the expectations in equation (7) can further help distinguish the small-scale structure. Note that equation (7) indicates that linear or quadratic large-scale structure should be evident by quadratic curvature in the order 0 increment plot. The lack of such curvature further suggests that the quadratic features in the anomaly plot are due to small-scale structure.

Figure 5 shows typical theoretical increment semivariograms for several reasonable choices of small-scale components for the temperature anomaly data. Figures 5(a) and 5(b) show the theoretical order 0 increment plot for AR(1) and ARIMA(0,1,0) small-scale components. Neither of these curves conclusively captures the shape of the order 0 plot in Figure 3. Figures 5 (c) and (d) depict first-order increment semivariograms and linvariograms for models with AR(1) components. Note that Figure 5(c), a model with no large-scale trend and a small-scale AR(1) component, is similar to the corresponding plot in Figure 3. Adding a quadratic trend component (recall that any linear component is filtered by the first-order increments) causes the first-order increment semivariogram and the linvariogram to be rotated from a horizontal asymptote. The two actually diverge if additional lags are plotted. This further confirms that the curvature in Figure 1 is not due to a quadratic large-scale component. Figures 5 (e) and (f) demonstrate that an

ARIMA(0,1,0) without a quadratic large-scale component behaves similarly to the plot in Figure 3. An ARIMA (1,1,0) does not.

[Insert Figure 5]

From this analysis of the temporal increments, it seems reasonable to model the temporal component of the temperature anomaly data as an AR(1) or an ARIMA (0,1,0) process. Additional evaluations of the individual series for all 138 stations indicates that 96 are selected to be white noise or AR(1) series using an AIC criterion. Differenced series are not white noise: only 6 are identified as white noise or AR(1) series using the AIC criterion. Hence, the AR(1) will be used for the small-scale temporal structure. It is again important to emphasize that this analysis does not uniquely determine that the only possible small-scale temporal component is an AR(1); nevertheless, it clearly demonstrates that an AR(1) is a reasonable choice. While not claiming to definitively model temperature change, other studies have also used AR(1) models as a reasonable temporal error structure for individual stations; e.g., Wigley and Jones (1981).

6. SPECIFYING THE SPATIAL STRUCTURE OF W_{ξ}^{d}

One approach to the estimation of the small-scale spatial structure is to repeat the primary increment approach for fixed times on the spatial data values. This is the intrinsic random function (IRF) kriging approach. However, the primary temporal increments have already filtered any large-scale temporal structure and are readily available for further use. One adaptation makes their use even more effective.

In the previous section, no accommodation was made for the large-scale spatial structure of model (1). In Section 3, it was suggested that quadratic gradients characterized the large-scale spatial structure. It is well-known that such large-scale gradients bias parameter estimates for the

small-scale spatial structure (Cressie 1987). This is clearly evident in spatial semivariogram plots. Figure 6 displays spatial semivariogram plots, averaged across the 61 years, for bins of 100 km. The positive bias caused by the spatial gradients causes the semivariogram to increase in an approximate linear fashion. When spatial gradients are removed by averaging semivariograms for each year of the least squares residuals from quadratic fits in latitude and longitude, the semivariogram plot reaches a sill at approximately 750 km. Hence, the residual semivariogram indicates that a well-defined generalized spatial covariance function (Cressie 1987, Kitanidis 1993) exists for the temperature anomalies.

[Insert Figure 6]

To further clarify the small-scale temporal structure in the temperature anomalies, Figures 3 and 4 were reconstructed from the spatial residuals for each year. Figures 7 and 8 display the results. The main qualitative difference between Figures 3 and 7 is that the order 0 plot in Figure 7 more clearly indicates an AR(1) small-scale error component than does the order 0 plot in Figure 3. The order 0 plot is much smoother and closer to an exponential shape than a linear one. Although the scale of the residual axis in Figure 8 is reduced by a factor of 4 over that of the anomaly axis in Figure 4, the residual anomalies from the quadratic spatial fits still exhibit curvature in time; however, the order 1 smoothed temporal increment plot is relatively constant. In addition, the order 1 smoothed increment plot in Figure 7 does not have the initial curvature apparent in the first few lags of the corresponding plot in Figure 3. Both Figures 7 and 8 confirm the conclusions drawn from Figures 3 and 4 but they do so with greater clarity. It is recommended, therefore, that temporal increment plots be made from spatial increments (residuals).

[Insert Figures 7 and 8]

Once the temporal structure of the model has been identified, it is known what order of primary increments will filter the temporal effects due to $\mu_\alpha(\mathbf{s},t) + g_\beta\{\mathbf{Y}(\mathbf{s}),t\}$. The spatial structure of $W_\xi^d(\mathbf{s},t)$ then can be determined by estimated *spatial* semivariograms of the selected primary temporal increments. Spatial semivariograms $\gamma_k(g) = \text{var}\{P_k(\mathbf{s}+\mathbf{g},t,h) - P_k(\mathbf{s},t,h)\}/2$ of primary temporal increments of the chosen order k are a multiple of the spatial semivariogram values among the $W_\xi^d(\mathbf{s},t)$ terms. To clarify this point, let

$$\widehat{\gamma}_k(g) = \frac{1}{2n_k c_k' c_k |M_k(\mathbf{g})|} \sum_{M_k(\mathbf{g})} \{P_k(\mathbf{s}_i+\mathbf{g},t,h) - P_k(\mathbf{s}_i,t,h)\}^2, \tag{9}$$

where n_k is the scale factor in the denominator of $\gamma_k(\mathbf{s},t,h)$ in (5), $c_k' c_k$ is the squared norm of the coefficient vector for the primary temporal increments of order k (see below),

$M_0(\mathbf{g})=\{(\mathbf{s}_i,\mathbf{s}_i+\mathbf{g},t,t+h) : g - \eta < ||\mathbf{g}|| < g + \eta; h \in B_1\}$,

$M_1(\mathbf{g})=\{(\mathbf{s}_i,\mathbf{s}_i+\mathbf{g},t,t+h_1,t+h_1+h) : g_k - \eta < ||\mathbf{g}|| < g + \eta;\ h_1,\ h \in B_1\}$, and

$M_2(\mathbf{g})=\{(\mathbf{s}_i,\mathbf{s}_i+\mathbf{g},t,t+h_1,t+h_1+h_2,t+h_1+h_2+h) : g - \eta < ||\mathbf{g}|| < g + \eta;\ h_1,\ h_2,\ h \in B_1\}$, g is the nominal spatial separation distance, and η is the bin tolerance. Write the kth primary temporal increment as $P_k(\mathbf{s},t,h) = c_k' z(\mathbf{s})$, where $z(\mathbf{s}) = \{(z(\mathbf{s},t_1), z(\mathbf{s},t_2), \ldots, z(\mathbf{s},t_n)\}'$. For example, the general forms of the coefficient vector c_k for equally spaced temporal data with $h_1 = h_2 = h = 1$ are $c_0 = (0'\ 1\ -1\ 0')'$, $c_1 = (\ 0'\ 1\ -2\ 1\ 0')'$, and $c_2 = (0'\ -2\ 6\ -6\ 2\ 0')'$. If the spatial semivariogram of the small-scale process $W_\xi^d(\mathbf{s},t)$ is written as $\gamma_w(g) = \text{var}\{W_\xi^d(\mathbf{s}+\mathbf{g},t) - W_\xi^d(\mathbf{s},t)\}/2$, then $E\{\widehat{\gamma}_k(g)/c_k' c_k\} = \sigma_{ee} + \gamma_w(g)q_k(t,h)$, where $q_k(t,h)$ is a function of the temporal expressions in Table 1.

The importance of this result is that the shape of the spatial semivariogram can be discerned from a plot of the spatial semivariogram values of the temporal increments. As a function of the nominal spatial separation distance g, the expected values of the estimators (9) are products of a

constant term in time and the spatial semivariogram $\gamma_w(g)$. Hence, a plot of $\hat{\gamma}_k(g)$ versus g will reveal the covariance structure of the $b_\omega(s,t)$ terms through the shape of the semivariogram, allowing for the specification of a parametric semivariogram or covariance model for $b_\omega(s,t)$.

Figure 9 displays the semivariograms of the original anomalies and the gradient-removed residuals alongside the comparable plots for the first-order temporal increments and the gradient-removed increments. For this data set, both sets of graphs exhibit similar shapes, although the order 1 temporal increment plots are smoother and noticeably smaller in magnitude because of the removal of temporal variation from the increments. Increment plots often shown greater smoothness in simulations. The shape of the gradient-removed semivariogram plots in Figure 9 suggests that a spherical semivariogram model would appropriately characterize the spatial dependence.

[Insert Figure 9]

7. CONCLUDING REMARKS

This work presents a general class of spatiotemporal models which are applicable to a wide range of environmental studies for which data are collected across a geographic region at multiple points in time. The primary focus of the models and model identification techniques formulated in this paper are on characterizing changes in time across the region of interest. Systematic steps for model identification are presented which can be applied to any of the specific models within the general class of models. The information gained from these steps can be scientifically informative. For example, knowing that the rate of change is well-modeled as a stochastic walk process may provide evidence in support of a scientific theory about a growth process or migratory behavior across a region or may suggest an alternative reality than was originally postulated about the sites of interest . Knowing that the value of the process is systematically different as one moves across

the region or that the process is changing at different rates in different areas of the region may provide support for decisions that are to be made or actions that are to be taken in different areas of the region.

The methods proposed here allow the data themselves to suggest an appropriate model but these methods are not so rigid that scientific information about the process of interest cannot be incorporated. Setting the models in the context of stochastic differential equations helps to draw the connection between the statistical modeling approach and a more deterministic approach, which often involves describing growth or migration in terms of a (deterministic) differential equation. If scientific information is available about the differential equation that should describe the process, it may be possible to incorporate that information into the large-scale temporal component of the model. More generally, when information about the way in which the process or the rate of change differs spatially across the region is available, that also may be incorporated into the large-scale temporal component of the model.

Finally, in addition to providing statistical justifications for the proposed models and identification methods, the procedures proposed in this paper can be reasonably implemented in practice. Proposed model identification techniques are primarily graphic, with theoretical justifications for why these graphics should direct investigators to meaningful modeling decisions. The use of graphics also allows scientific judgment to be incorporated into the model identification procedure.

REFERENCES

Ali, M. M. (1979). "Analysis of stationary spatial-temporal processes: estimation and prediction," *Biometrika*, 66, 513-518.

Bennet, R. J. (1974). "The representation and identification of spatio-temporal systems: an example of population diffusion in north-west England," *Transactions and Papers of the Institute of British Geographers*, 66, 73-94.

Bilonick, R. A. (1985). "The space-time distribution of sulfate deposition in the northeastern United States," *Atmospheric Environment*, 19, 1829-1845.

Bogaert, P. (1996). "Comparison of kriging techniques in a space-time context," *Mathematical Geology*, 28, 73-86.

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

Christensen, R. C. (1991). *Linear Models for Multivariate, Time Series and Spatial Data*, New York: Springer-Verlag.

Cliff, A. D. and Ord, J. K. (1974). "Space-time modeling with an application to regional forecasting," *Transactions and Papers of the Institute of British Geographers*, 66, 119-128.

Cressie, N. (1987). "A nonparametric view of generalized covariances for kriging," *Mathematical Geology*, 19, 425-449.

----- (1988). "A graphical procedure for determining nonstationarity in time series," *Journal of the American Statistical Association*, 83, 1108-1116; corrigenda, 85, 272.

----- (1993). *Statistics for Spatial Data*. New York: John Wiley & Sons.

DeGooijer, J. G. and Anderson, O. D. (1985). "Moments of the sampled space-time autocovariance and autocorrelation function," *Biometrika*, 72, 689-693.

Gunst, R. F. (1995) "Estimating spatial correlations from spatial-temporal meteorological data," *Journal of Climate*, 8, 2454-2470.

Gunst, R. F., Basu, S., and Brunell, R. (1993). "Defining and estimating global mean temperature anomalies," *Journal of Climate*, 6, 1368-1374.

Guttorp, P., Sampson, P. D. and Newman, K. (1992). "Nonparametric estimation of spatial covariance with application to monitoring network evaluation," in *Statistics in the Environmental and Earth Sciences*, eds. A. T. Walder and P. Guttorp, London: Charles W. Griffin.

Haas, T. C. (1995). "Local prediction of spatio-temporal processes with an application to wet sulfate deposition," *Journal of the American Statistical Association*, 90, 853-861.

Hartfield, M. I. and Gunst, R. F. (1999). "Parameter estimation for a class of spatio-temporal models," Technical Report SMU/DS/TR-291, Department of Statistical Science, Southern Methodist University, Dallas, Texas.

Haslett, J. and Raftery, A. E. (1989). "Space-time modeling with long-memory dependence: assessing Ireland's wind power resource," *Applied Statistics*, 38, 1-50.

Hosking, J. R. M. (1981). "Fractional differencing," *Biometrika*, 68, 165-176.

Host, G., Omre, H. and Switzer, P. (1995) "Spatial interpolation errors for monitoring data," *Journal of the American Statistical Association*, 90, 853-861.

Huang, H. and Cressie, N. (1996). "Spatio-temporal prediction of snow water equivalent using the kalman filter," *Computational Statistics and Data Analysis*.

Journel, A. G. (1985). "The deterministic side of geostatistics," *Journal of International Association for Mathematical Geology*, 17, 1-14.

Journel, A. G. and Huijbregts, C. 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.

Kitanidis, P. K. (1993). "Generalized covariance functions in estimation," *Mathematical Geology*, 25, 525-540.

Loader, C. and Switzer, P. (1992). "Spatial covariance estimation for monitoring data," in *Statistics in the Environmental and Earth Sciences*, eds. A.T. Walder and P. Guttorp, London: Charles W. Griffin, pp. 52-59.

Martin R. L. and Oeppen, J. E. (1975). "The identification of regional forecasting models using space-time correlation functions," *Transactions and Papers of the Institute of British Geographers*, 66, 95-118.

Matheron, G. (1973). "The intrinsic random functions and their applications," *Advances in Applied Probability*, 5, 439-468.

Oehlert, G. W. (1993). "Regional trends in wet sulfate deposition," *Journal of the American Statistical Association*, 88, 390-399.

Pfeifer, P. E. and Deutsch S. J. (1980a). "A three-stage iterative approach for space-time modeling," *Technometrics*, 22, 35-47.

----- (1980b). "Identification and estimation of first order space-time ARMA models," *Technometrics*, 22, 397-408.

----- (1981). "Variance of the sample space-time autocorrelation function," *Journal of the Royal Statistical Society*, Ser. B, 43, 28-33.

Rodrigues-Iturbe, I. and Mejia, J. M. (1974). "The design of rainfall networks in time and space," *Water Resources Research*, 10, 713-728.

Rouhani, S. Ebrahimpour, M. R., Yaqub, I. and Gianella, E. (1992). "Multivariate geostatistical trend detection and network evaluation of space-time acid deposition data--I. methodology," *Atmospheric Environment*, 26A, 2603-2614.

Rouhani, S. and Hall, T. J. (1989). "Space-time kriging of groundwater data," in Vol. 2 of *Geostatistics*, ed. M. Armstrong, New York: Kluwer Academic Publishers.

Rouhani, S. and Myers, D. E. (1990). "Problems in space-time kriging of geohydrological data," *Mathematical Geology*, 22, 611-623.

Rouhani, S. and Wackernagel, H. (1990). "Multivariate geostatistical approach to space-time data analysis," *Water Resources Research*, 26, 585-591.

Royle, J. A. and Berliner, M. (1999). "A hierarchical approach to multivariate spatial modeling and prediction," *Journal of Agricultural, Biological, and Environmental Statistics*, 4, 29-56.

Searle, S. R., Casella, G., and McCulloch, C. E. (1992). *Variance Componenst*. New York: John Wiley & Sons.

Seber, G. A. F. and Wild, C. J. (1989). *Nonlinear Regression*, New York: John Wiley & Sons.

Solna, K. and Switzer P. (1996). "Time trend estimation for a geographic region." *Journal of the American Statistical Association*, 91, 577-589.

Stein, M. (1986). "A simple model for spatio-temporal processes," *Water Resources Research*, 22, 2107-2110.

Switzer, P. (1989). "Nonstationary spatial covariance estimated from monitoring data," from Vol. 1 of *Geostatistics*, ed. M. Armstrong, New York: Kluwer Academic Publishers, pp. 639-651.

Wigley, T. M. L. and Jones, P. (1995). "Detecting CO_2-induced Climatic Change" *Nature*, 292, 205-208.

Table 1. Expressions for $\delta(t,h)$ for Six Small-Scale Variation Structures

Small-Scale Process	*Parameters*
Stationary	
$\sigma_{bb} + \sigma_{ee}$	$d = 0\ p = 0$
$\frac{\sigma_{bb}}{2\lambda}(1 - e^{-h\lambda}) + \sigma_{ee}$	$d = 0\ p = 1$
Stochastic Walk	
$\frac{\sigma_{bb}}{2} h + \sigma_{ee}$	$d = 1\ p = 0$
$\frac{\sigma_{bb}}{2\lambda^3}\{\lambda h - (1\text{-}e^{-h\lambda})\} + \sigma_{ee}$	$d = 1\ p = 1$
Stochastic Slope	
$\frac{\sigma_{bb}}{2}\left(\frac{h^3}{3} + th^2\right) + \sigma_{ee}$	$d = 2\ p = 0$
$\frac{\sigma_{bb}}{2\lambda^5}\left\{\lambda^3\left(\frac{h^3}{3} + th^2\right) - \frac{1}{2}\lambda^2 h^2 - \lambda h + \left(1 - e^{-h\lambda}\right)\right\} + \sigma_{ee}^2$	$d = 2\ p = 1$

FIGURE CAPTIONS

Figure 1. Smoothed Annual Temperature Anomalies

Figure 2. Smoothed Annual Temperature Anomalies, by Year

Figure 3. Temporal Primary Increments Averaged Across Years, No Spatial Gradient Removal

Figure 4. Smoohted Annual Temperature Anomalies and Primary Increments

Figure 5. Effects of Model Assumptions on Increment Semivariograms

Figure 6. Average Spatial Semivariograms

Figure 7. Temporal Primary Incremetns Averaged Across years, Spatial Gradients Removed

Figure 8. Smoothed Residual Primary Increments

Figure 9. Anomaly and First-Order Temporal Increment Spatial Semivariograms

Figure 1. Smoothed Annual Temperature Anomalies

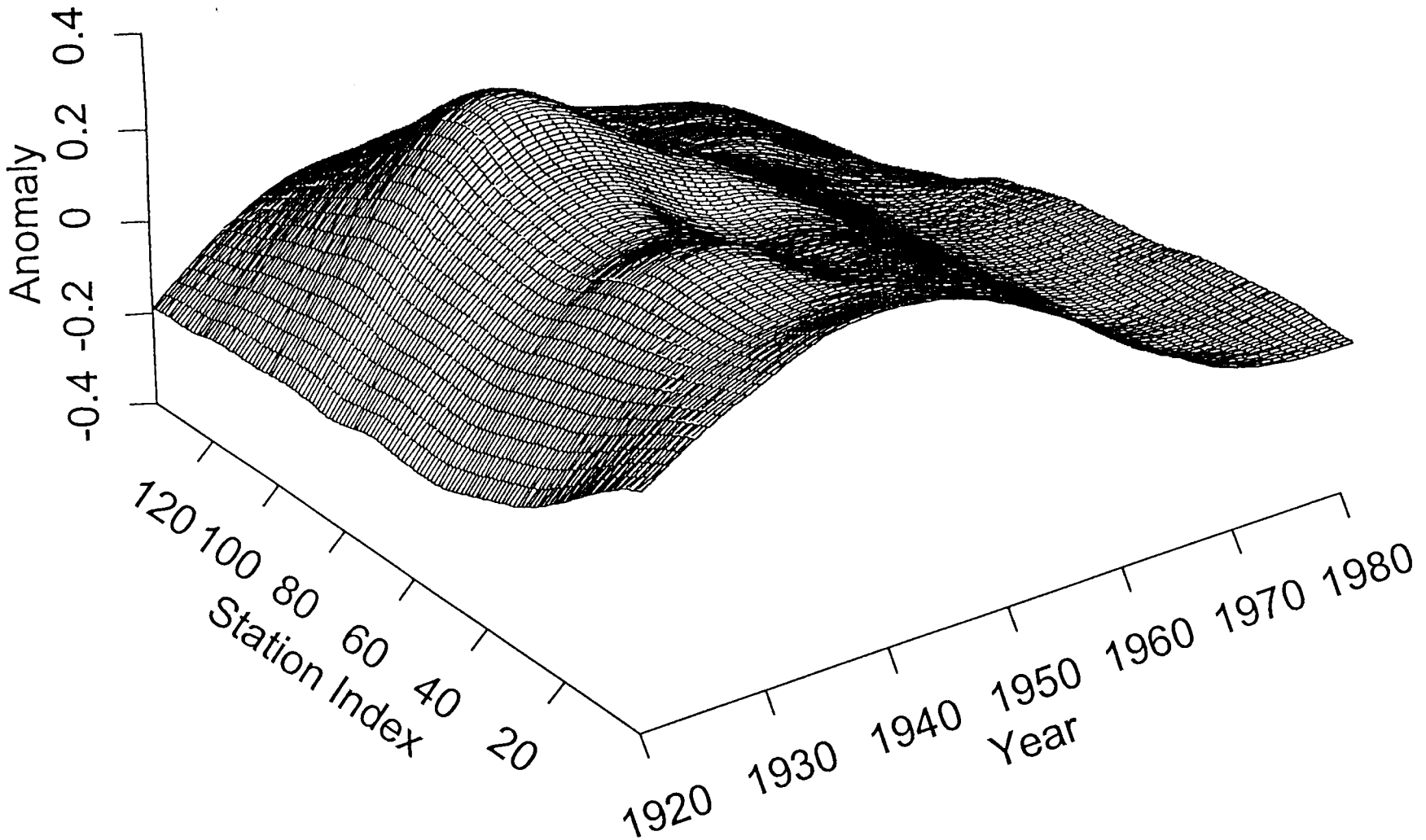


Figure 2. Smoothed Annual Temperature Anomalies, by Year

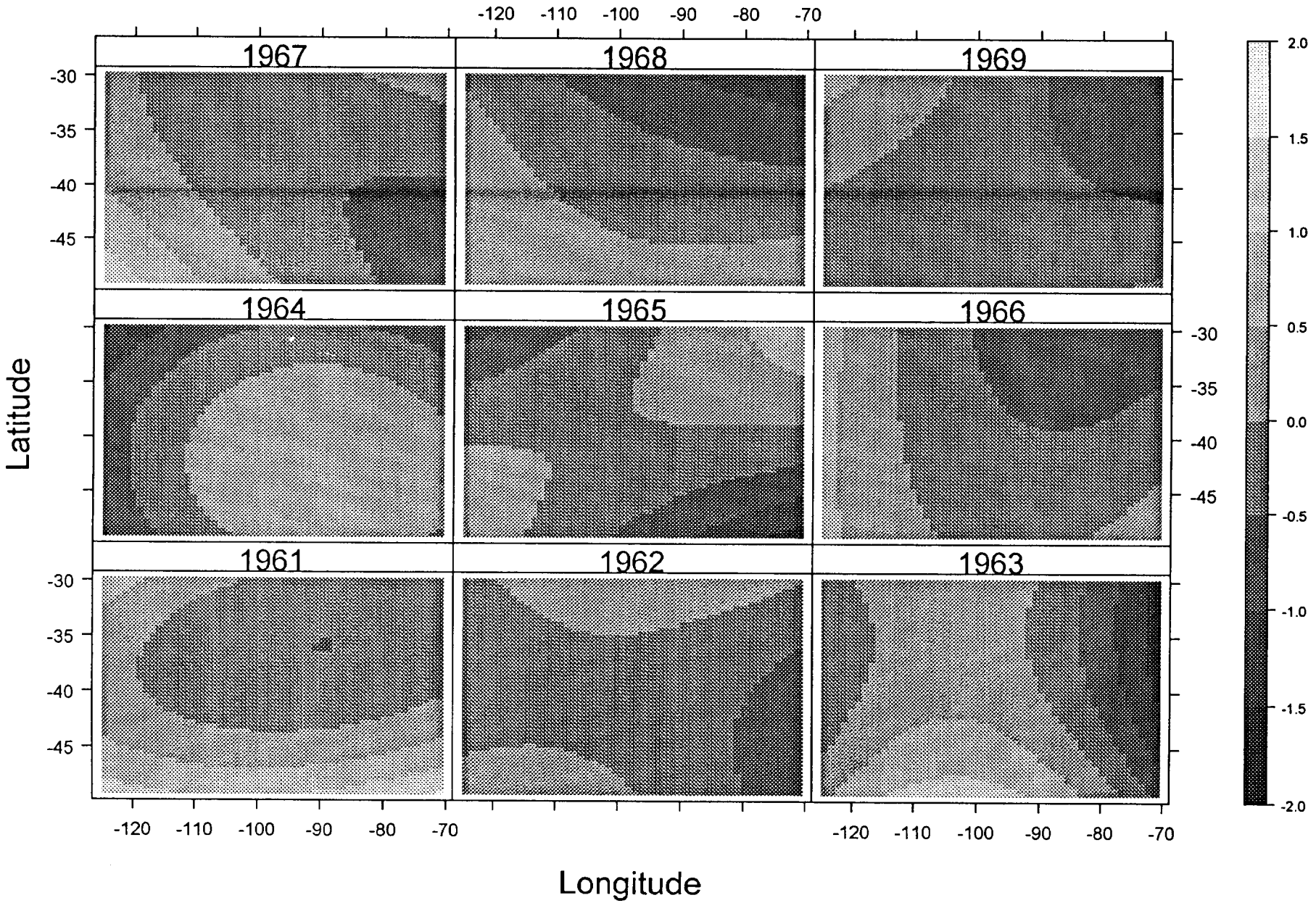
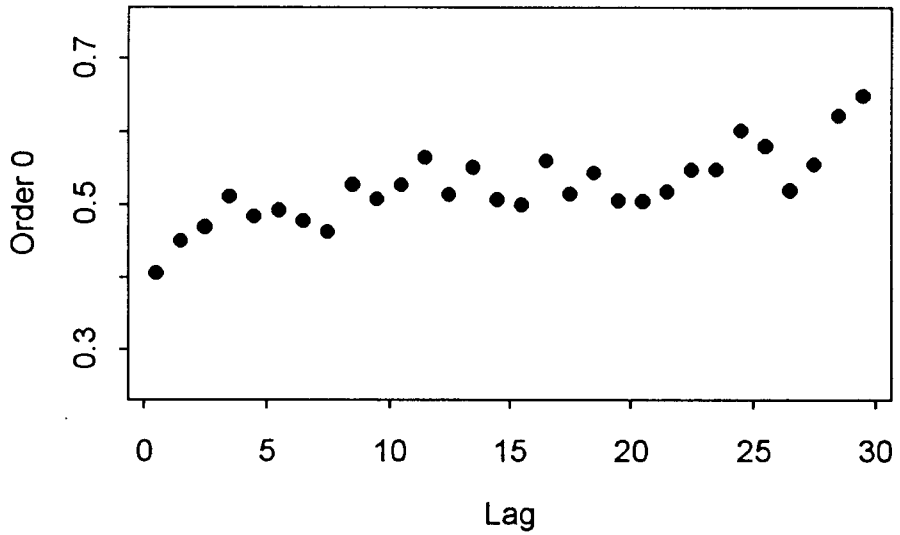
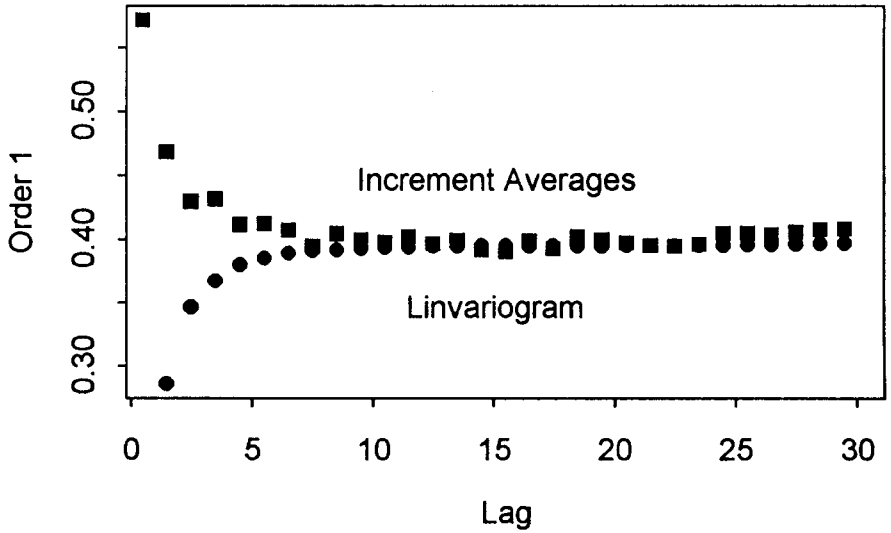


Figure 3. Temporal Primary Increments Averaged Across Years
No Spatial Gradient Removal





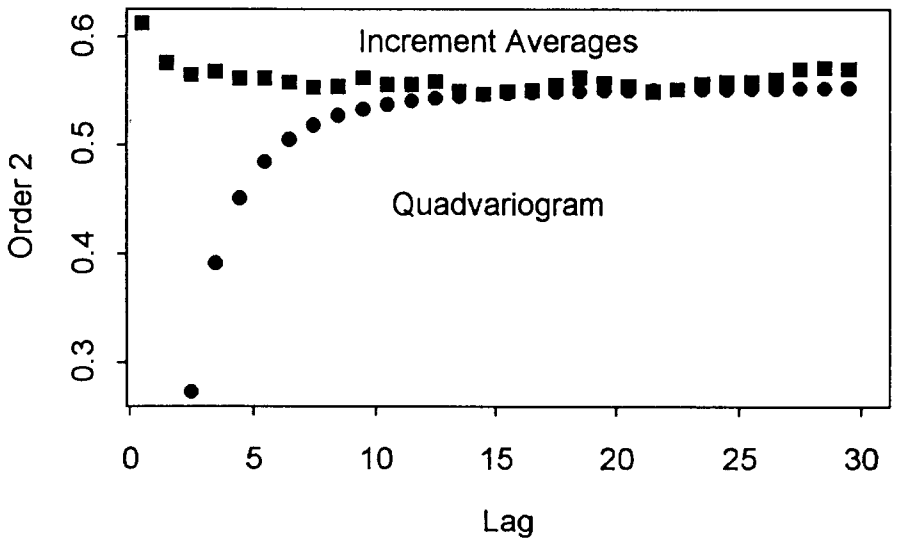


Figure 4. Smoothed Annual Temperature Anomalies and Primary Increments

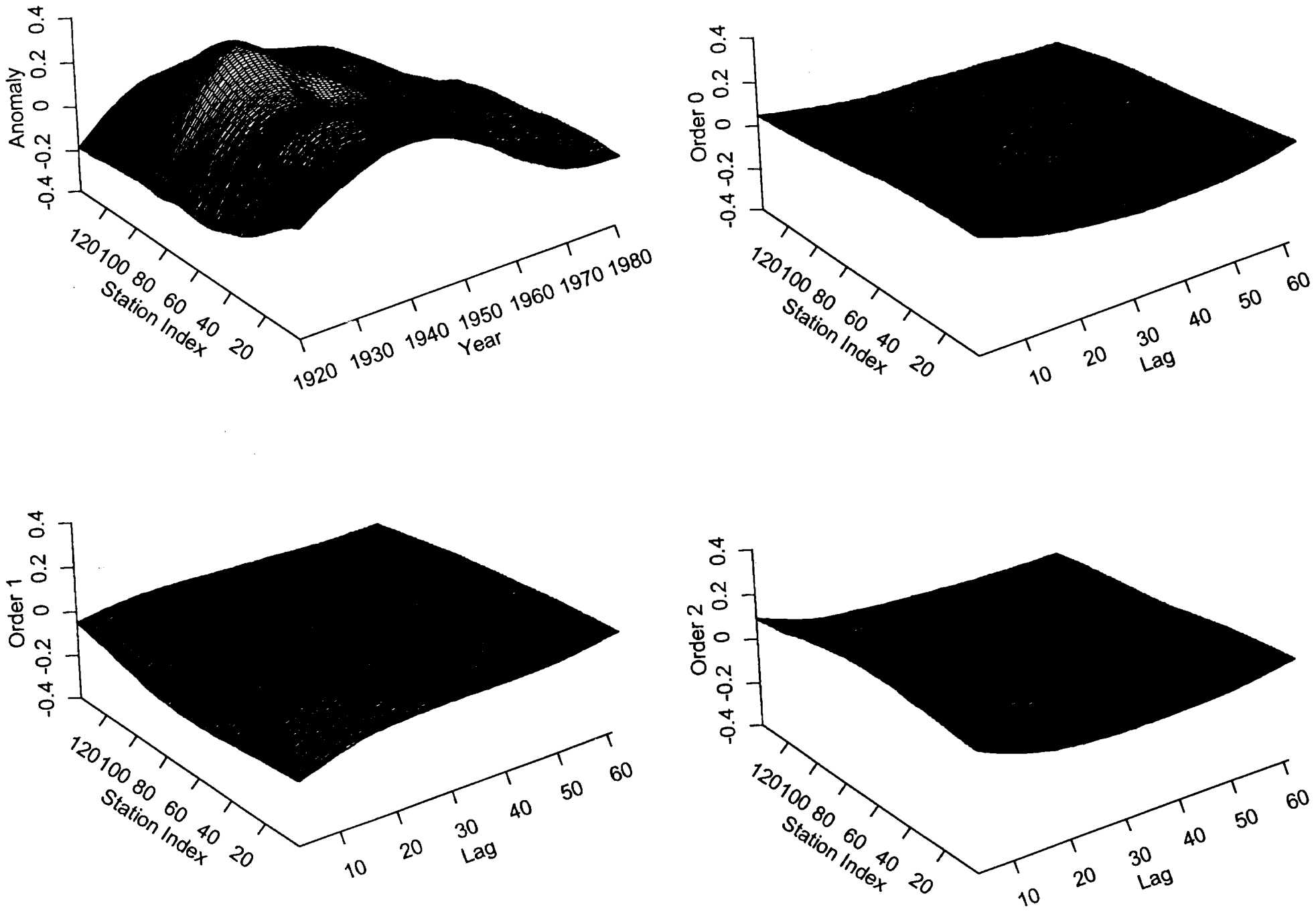


Figure 5. Effects of Model Assumptions on Increment Semivariograms

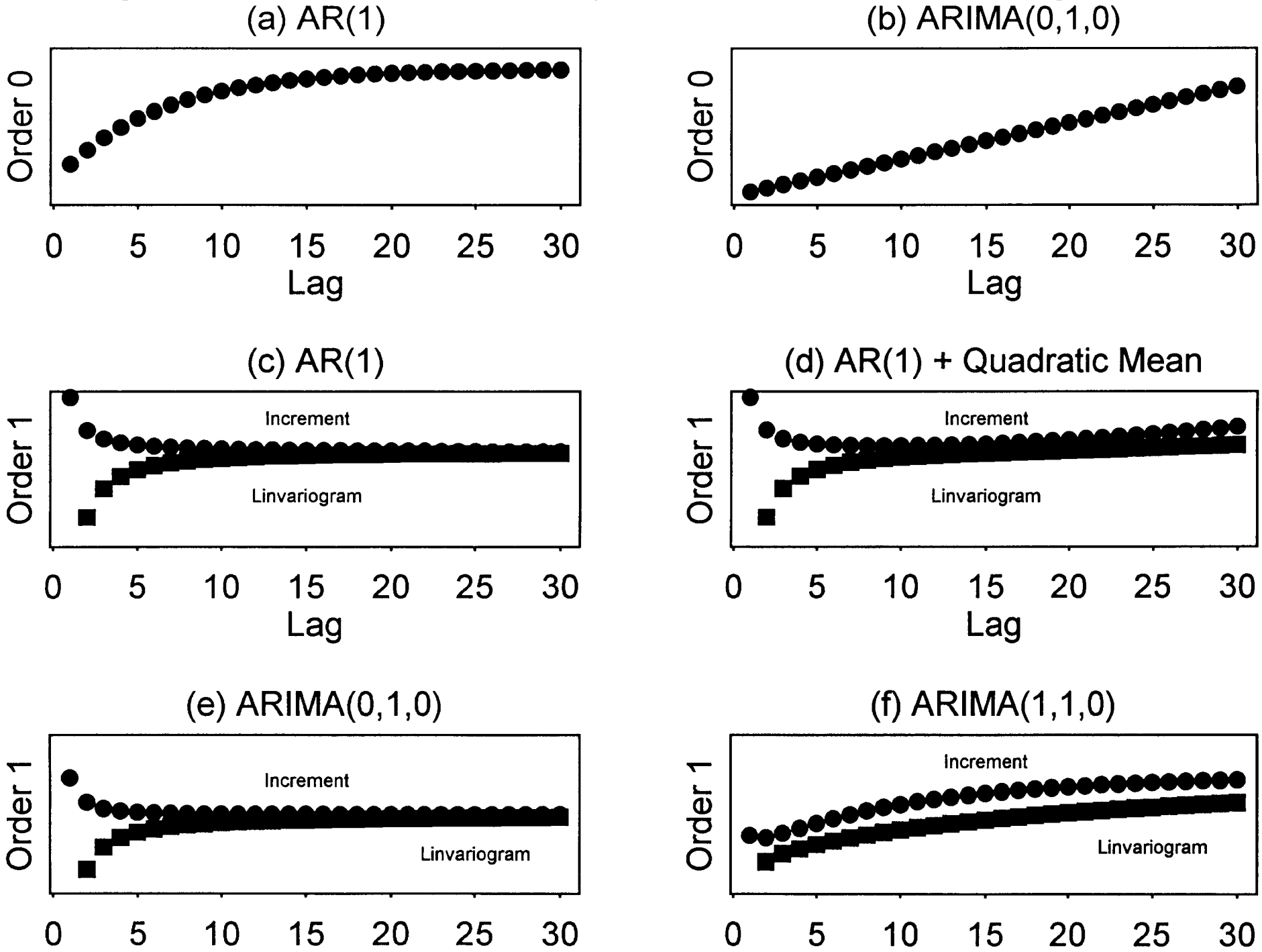


Figure 6. Average Spatial Semivariograms.

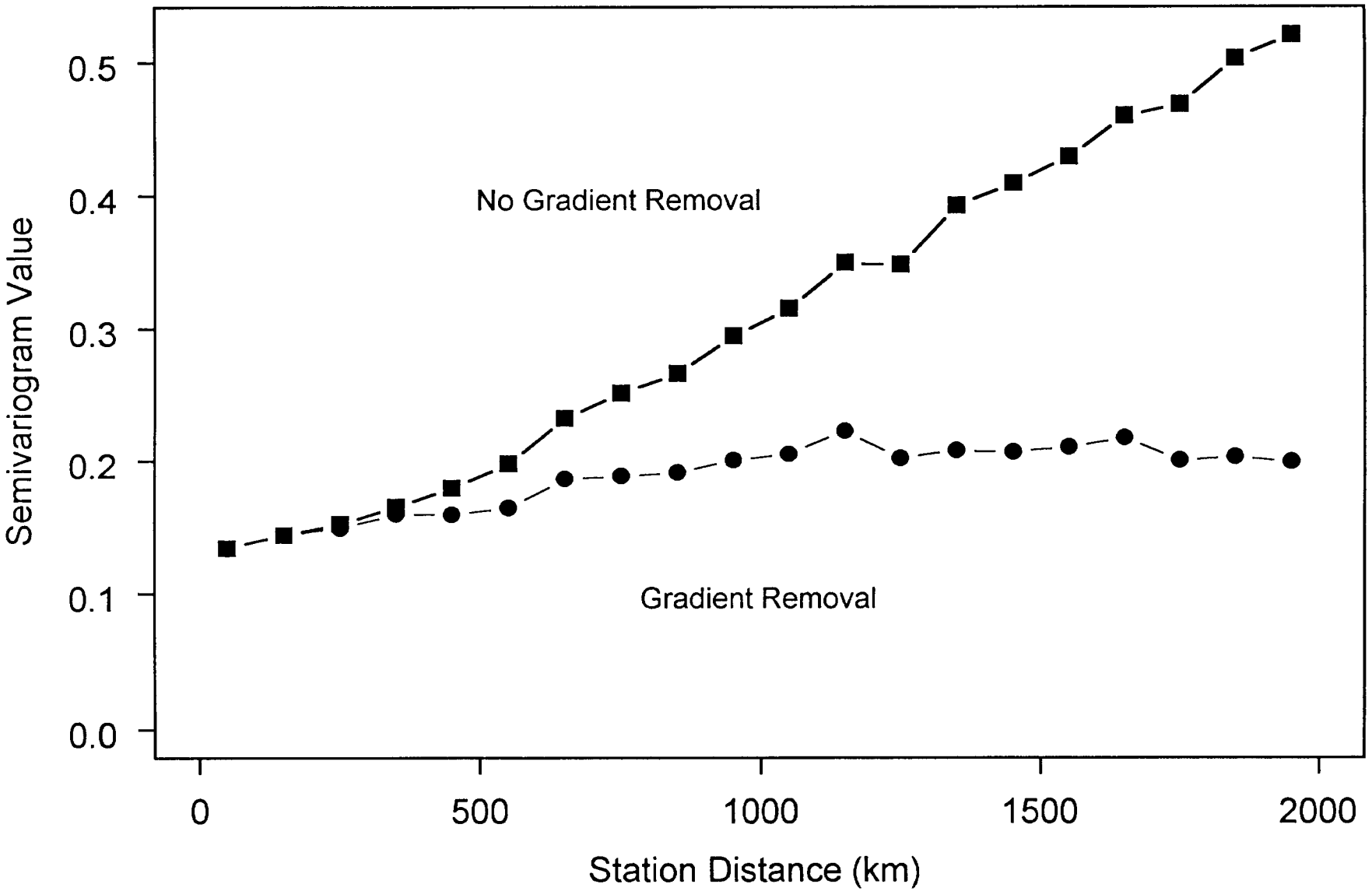
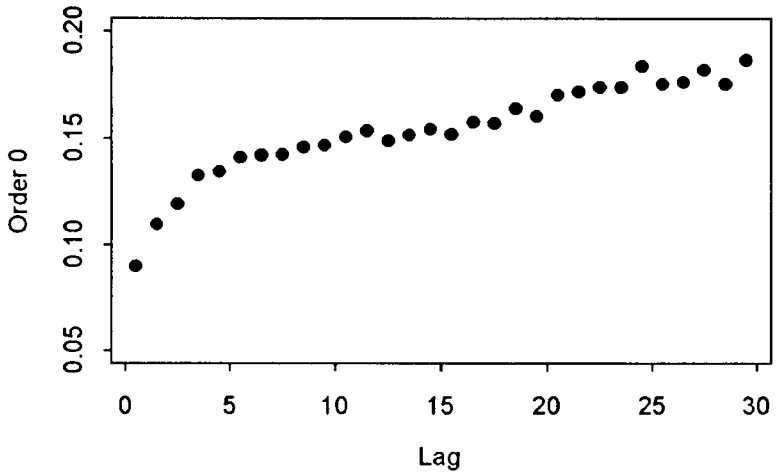
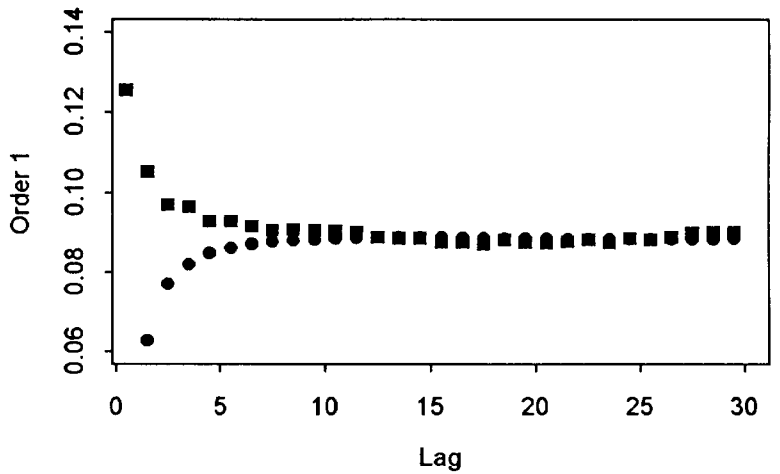


Figure 7. Temporal Primary Increments Averaged Across Years
Spatial Gradients Removed





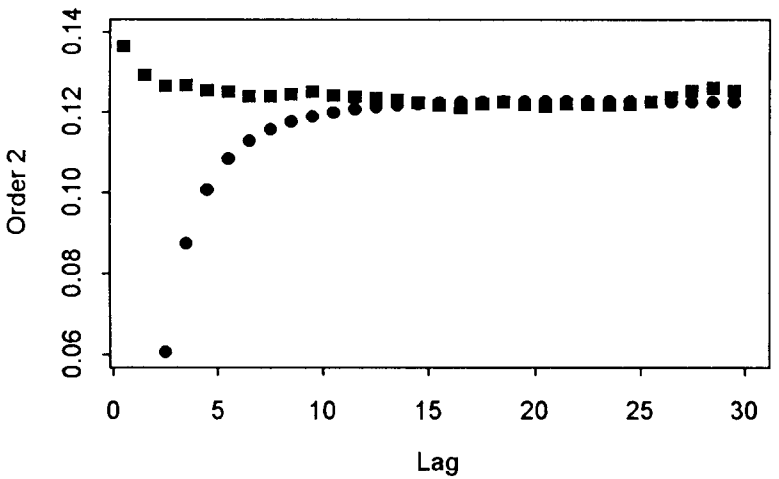


Figure 8. Smoothed Residual Primary Increments

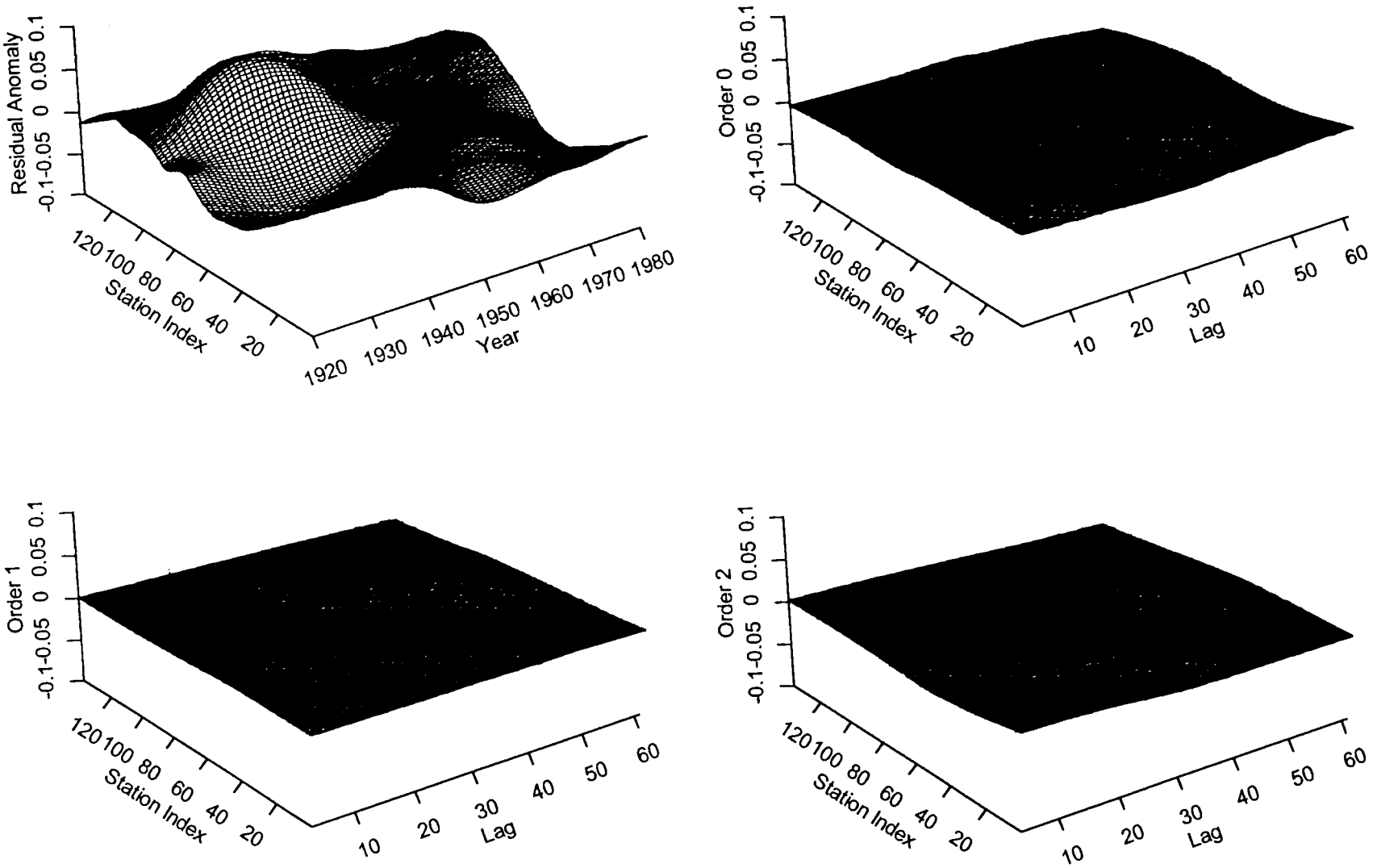


Figure 9. Anomaly and First-Order Temporal Increment Spatial Semivariograms.