Time-Trend Estimation for a Geographic Region

Paul Switzer [*] Knut Sølna [†]

Abstract

We discuss a framework for estimation of temporal trend for an evolving spatial field. The spatial field is regularly sampled in time at arbitrary monitoring locations whose position may change over time. The estimation of time-trend and the quantification of estimation error derives from a probabilistic model. We illustrate with an example involving the surface temperature field in the steppe region of eastern-Europe.

Keywords: Monitoring data, Multiple time series, Regional temperature trend, Space-time modeling

1 INTRODUCTION

The problem of identifying a temporal trend in a space-time data set has been accentuated by the global warming discussion. In this paper we propose a framework for estimation of such a trend and a quantification of the associated estimation uncertainty. Historical surface temperature data will serve the purpose of motivating our modeling assumptions and illustrating the approach.

More precisely we will consider a spatial field on a fixed region which is evolving in time and for which we have discrete observations at regular time intervals. Our objective is to estimate a temporal trend in the evolution of the field, a trend which may vary with location. We will also estimate the actual change in the regional average value of the evolving field over fixed time intervals. The problem of a sampling configuration changing with time will be explicitly considered.

The estimation relies on building a probabilistic model for the time evolving field. The temporal trend, which may vary with location, is considered as a realization of a spatially homogeneous stochastic field and overlays a residual

[*]Sequoia Hall, Stanford University, Stanford CA 94305. Supported by SIMS (Societal Institute for Mathematical Sciences) through a co-operative agreement with the U.S. Environmental Protection Agency.

[†]Supported by The Norwegian Council for Sciences and Research.

component assumed to be stationary in space and time. In our example involving surface temperature data we chose a climatically homogeneous region wherein the spatial variability of the time-trend is relatively small. This region is the steppe of eastern-Europe. Having defined the space-time variability model, the estimators are the unbiased linear combinations of the observations which minimize estimation error variance. The resulting error variance will depend on the space-time data configuration.

Note that even though the quantities to be identified are modeled as random we will use the estimation- rather than prediction-terminology throughout.

An important ingredient in the analysis is time-differencing of the original data and direct modeling of time differences. For our example the one-year temperature differences are the basic modeling element. The time-differenced field is then regarded as stationary in space and time, and the problem of estimating trend reduces to one of estimating a time-average.

The emphasis of this study is to derive a framework for trustworthy estimation of the precision associated with the estimate of temporal trend of an evolving surface temperature field. Statistical heterogeneity derives both from the heterogeneity of variability of the underlying temperature field, due to regional as well as seasonal effects, and from the varying density and geographic configuration of the monitoring network. Due to the strong seasonal heterogeneity, different models will be derived for different seasons of the year.

The paper is organized as follows. The next section describes some statistics of the data used in the example and serves to motivate the model and estimators being derived in the following sections. The statistical model employed in the analysis is described in Section 3. Estimators based on this model are derived in Sections 4 and 5. The estimators are the unbiased linear estimators minimizing error variance and are based on the yearly temperature differences. In Sections 6 and 7 various aspects of the structure of the sampling pattern are discussed. We consider the case with a complete rectangular data panel in Section 7, that is a fixed number of spatial sampling stations with no missing observations in time. This enables us to derive simpler expressions for the estimator and their standard errors as functions of number of years of observation. Section 8 uses temperature data from the steppe region of eastern-Europe to illustrate the approach. The illustration involves a monitoring network which is changing over time. Section 9 contains some concluding remarks.

2 THE DATA

We chose the steppe region of eastern-Europe as a test bed for our analysis and statistical modeling. The data are monthly averaged surface temperatures at 24 monitoring sites. A map of the steppe region is shown in Figure 1. The monitoring site locations within the rectangular study region, outlined by the dashed line, are shown on the map. The period of the data covers years 1951-1990. In addition to monthly data we created an annual surface temperature time series at each monitoring site by averaging the monthly data for each year. The raw data were obtained from Goddard Institute for Space Studies.

The monitoring stations were not all simultaneously active during the study period. Figure 2 displays the missing pattern for the annually averaged data. If at least one month's observation is missing the average observation is taken to be a missing data item. It is seen that a number of stations ceased to provide information after 1970. Hence, the reality of a monitoring network changing with time is certainly present in our test region.

Figure 3 shows three yearly time series corresponding to data from the most westerly station. The solid lines correspond to annual observations for January and July, whereas the dotted line is the annually averaged time series. These time series show that there is a substantially larger inter-annual variability in the winter than the summer months. Furthermore, the 40-year time-trend is small relative to the inter-annual variability of the data for this site. The correlation coefficient between the January and July observations, computed based on these 40 years, was .2. This suggests that the correlation between seasons is small.

We converted all time series to their annually differenced form. This removes difficulties associated with choosing a baseline year for an evolving network and also the problem of having to deal with a spatially varying temperature mean field. The three time series presented in Figure 4 display the annually differenced July data at the three monitoring stations marked with plus signs in Figure 1. The dashed line corresponds to the most westerly station. We see strong similarities between time series for nearby stations. A quantification of such spatial structure in the data will obviously be an important ingredient in the estimation schemes to be derived. Moreover, Figure 4 also suggests weak inter-annual correlation.

Figure 5 shows annual time series which are averages taken over differences at all monitoring sites. Separate annual series are shown for July, January and year averages. What is remarkable, is that the amplitude of the inter-annual variability is not much less than the corresponding amplitude for a time series from a single station as seen, for example, in Figure 4. This reinforces the suspicion that the spatial correlation structure is strong.

There will certainly be errors associated with the above data set. Among these are instrumental errors which can be systematic or random. The random instrumental errors associated with monthly averaged temperature recordings are likely to be small in relation to other sources of variability. Systematic instrumental errors arise from calibration problems and placement. Note, however, that a systematic error present over the temporal span of a temperature time series is being removed when forming the differences.

3 THE MODEL

The goal of our analysis is to develop a statistical framework for quantifying uncertainty in estimates of regional temperature time-trend as well as the uncertainty of estimates of actual regional temperature changes over fixed time periods. The proposed statistical framework will be able to handle a monitoring network evolving in time.

The salient features of the steppe data are the strong spatial dependence and the weak inter-annual correlation of the temperatures. Moreover, the variability structure of the data shows a strong seasonal dependence. For this reason, the temperature fields for different months of the year are modeled separately. The quantity of prime interest, the temperature time-trend, is small compared to the inherent variability of the data.

The space-time field for monthly averaged surface temperatures will be considered as a realization of a stochastic field. Models of the following type are developed separately for each month to allow for seasonally-dependent statistics; dependence on 'month' is suppressed in the notation

$$T(x,t) \quad = \quad \mu_0(x) + b(x)\,t + \delta(x,t), \qquad (3.1)$$

where

$T(\cdot,\cdot)$ surface temperature,

x location index, taking values in a specified region D,

t year index,

$\mu_0(\cdot)$ baseline temperature field, non-random,

$b(\cdot)$ temperature time-trend, a stochastic field second-order stationary and isotropic in space,

$\delta(\cdot,\cdot)$ space-time temperature residual, stochastic field with mean zero, second-order stationary in space and time and isotropic in space.

We are assuming a prior geographical stratification into a set of regions $\{D\}$ which are more climatically homogeneous than the globe as a whole. The purpose is to make the intraregion variation of the $b(\cdot)$ field relatively small, and correspondingly for the probabilistic structure of the residual field $\delta(\cdot,\cdot)$. Note that $b(\cdot)$ is an unknown spatially varying field and that we want to incorporate uncertainty of this b field into the expressions for estimation errors. Therefore, we model b as a stochastic field. The statistical modeling would be specific to each geographic region D. Thus the stationarity statements above refer to the domain $D \times E$, with E a time interval containing the observations and in which time-stationarity is assumed to hold.

We base the analysis on the year-to-year temperature differences rather than on the temperatures themselves. This immediately reduces the question of time-trend to one of estimation of an average level. Furthermore, heterogeneity among different stations within the region is partially mitigated; there is no need to choose a baseline time period or to estimate the baseline field $\mu_0(\cdot)$. The differences are defined by

$$\begin{aligned} d(x,t) &= T(x,t) - T(x,t-1) \\ &= b(x) + R(x,t) \end{aligned} \tag{3.2}$$

where

$$R(x,t) = \delta(x,t) - \delta(x,t-1). \tag{3.3}$$

In the next section we will derive estimators for the regional temperature time-trend and also actual temperature changes for specified subregions and time periods. We use unbiased linear combinations of the observations that, given the model, minimize the variance of the estimation errors. In doing so, it is in general advantageous to express error variances in terms of semivariograms rather than covariance functions because covariance estimation require preliminary estimation of mean values. The semivariogram of the $b(\cdot)$ field is defined by

$$1/2\, E[(b(x) - b(x+\Delta x))^2] = \gamma_b(|\Delta x|). \tag{3.4}$$

The semivariogram of the space-time temperature residual will be parameterized as

$$\begin{aligned} 1/2\, E[(\delta(x,t) - \delta(x+\Delta x, t+\Delta t))^2] &= \gamma_\delta(|\Delta x|, |\Delta t|) \\ &= \sigma_\delta^2 \{1 - \rho_S(|\Delta x|)\, \rho_T(|\Delta t|)\}. \end{aligned} \tag{3.5}$$

The assumption that the correlation structure of the temperature residuals factors into spatial and temporal components, denoted $\rho_S(\cdot)$ and $\rho_T(\cdot)$ respectively, simplifies the estimation problem. Proceeding without this assumption would entail more general covariance expressions which are more difficult to estimate from spatially sparse data. In Cressie (1992) this factorization assumption is denoted separability. Moreover, assume that that the spatial cross-covariance of the b and δ fields is time invariant. It then follows that the $b(\cdot)$ and $R(\cdot,\cdot)$ fields are uncorrelated.

The semivariogram of the difference field is thus found as

$$\begin{aligned} 1/2\,E[(d(x,t) - d(x+\Delta x, t+\Delta t))^2] &= \gamma_d(|\Delta x|, |\Delta t|) \\ &= \gamma_b(|\Delta x|) + \gamma_R(|\Delta x|, |\Delta t|). \end{aligned} \tag{3.6}$$

The semivariogram $\gamma_R(\cdot,\cdot)$ can, under the parameterization defined in (3.5), be expressed as

$$\begin{aligned} \gamma_R(|\Delta x|, |\Delta t|) &= 1/2\,E[(R(x,t) - R(x+\Delta x, t+\Delta t))^2] \\ &= \gamma_T(|\Delta t|) + \gamma_S(|\Delta x|)\,h(|\Delta t|), \end{aligned} \tag{3.7}$$

where

$$\begin{aligned} \gamma_T(|\Delta t|) &= 1/2\,E[(R(x,t) - R(x, t+\Delta t))^2] = \sigma_R^2\,(1 - h(|\Delta t|)), \\ \gamma_S(|\Delta x|) &= 1/2\,E[(R(x,t) - R(x+\Delta x, t))^2] = \sigma_R^2\,(1 - \rho_S(|\Delta x|)) \\ &= (\sigma_R/\sigma_\delta)^2\,\gamma_\delta(|\Delta x|, 0), \\ \sigma_R^2 &= E[R(x,t)^2] = 2\,\sigma_\delta^2\,(1 - \rho_T(1)) \\ &= 2\,\gamma_\delta(0,1), \\ h(|\Delta t|) &= \{-\rho_T(|\Delta t - 1|) + 2\,\rho_T(|\Delta t|) - \rho_T(|\Delta t + 1|)\}/\{2\,(1 - \rho_T(1))\} \\ &= \{\gamma_\delta(0, |\Delta t - 1|) - 2\,\gamma_\delta(0, |\Delta t|) + \gamma_\delta(0, |\Delta t + 1|)\}/\{2\,\gamma_\delta(0,1)\}. \end{aligned} \tag{3.8}$$

From (3.6)-(3.8) it is seen that knowledge of the semivariogram-components involved in the problem implies knowledge also of $h(\cdot)$, from which the covariance structure of the $R(\cdot,\cdot)$ field can be deduced. This is a further consequence of the factorization assumption (3.5). The covariance structure of the $b(\cdot)$ field, however, cannot be inferred from the semivariogram-components.

We now review briefly some of the related literature. An early paper by Rodriguez-Iturbe and Mejia (1974) on the design of rainfall networks focused on the problem of estimation of areal mean values, both for single time events and for long-term averages. However, like the example which we use, they considered space-time autocorrelation to be in factorable form. The methodological paper of Stein (1986) also addresses the problem of estimating changes over time of a spatial field in a more limited framework. However, like us, he uses the time-differenced data statistics directly for the estimation of variograms. Stein uses only concurrent data for interpolation of the time-differenced field and is not involved with problems of missing data. On the other hand, Bilonick's (1988) paper on monthly acid deposition maps uses space-time variogram modeling which does not operate directly on time-differenced data and therefore carries along the additional variability of the original spatial mean field. The paper by Rouhani (1990) and Rouhani et al. (1992) consider separately estimated time-series models at each of the monitoring stations and decompose temporal variability on several time scales. The principal focus of these papers is not on regional estimation or spatial modeling but rather on the characterization of the sampling sites. Also, the Handcock and Wallis (1994) analysis of meteorological space-time fields used models which operate on the direct data rather than on the differenced data. However, by using multivariate normality they were able to incorporate uncertainty in variability parameters via a Bayesian calculation. Observed changes in the spatial mean field were gauged against the posterior uncertainty in the estimate of the mean field at a fixed time. Finally, the recent work of Brown, Le and Zidek (1994) employs a Bayesian approach to incorporate uncertainty in parameter estimates of the space-time models of autocovariance and provides a method for propagating this uncertainty to error estimates for the interpolated field. Their work is also based on multivariate normal modeling.

4 ESTIMATION OF TIME-TREND

Consider the problem of estimating the temperature time-trend averaged over some subregion $A \subset D$ (the subregion A could for example be the whole region D or a single point in D). The estimate of the averaged time-trend, m_A, will be the unbiased linear combination of the available time differenced monitoring data in the region D which minimizes the variance of the estimation error. In accordance with the model assumptions given in the previous section the difference field $d(\cdot,\cdot)$ will be regarded as stationary over the time span of the data. The regional time-trend and its estimator are defined by

$$m_A = \int_A b(x)\, d\mu(x), \tag{4.1}$$

$$\hat{m}_A = \sum_{i=1}^{n} d_i\, \lambda_i = \boldsymbol{\lambda}^T \boldsymbol{d}, \tag{4.2}$$

where

$A \subset D$ the subregion of interest, with $\mu(A) = 1$.

n number of observed temperature differences in D, indexed by $i = 1, \cdots, n$,

$d_i = T(x_i, t_i) - T(x_i, t_i - 1)$ observed temperature difference,

x_i location of difference observation i,

t_i time of difference observation i,

λ_i estimating coefficient for difference observation i,

$\boldsymbol{\lambda} = \langle \lambda_1, \ldots, \lambda_n \rangle^T$ vector of coefficients,

$\boldsymbol{d} = \langle d_1, \cdots, d_n \rangle^T$ vector of observed temperature differences.

Under the stochastic model of Section 3, the unbiasedness criterion $E[\hat{m}_A] = E[m_A]$ entails

$$\sum_{i=1}^{n} \lambda_i \quad = \quad 1. \tag{4.3}$$

Using the fact that the λ_i coefficients sum to unity, we can express the variance of the estimation error for the time-trend as $Var[m_A - \hat{m}_A] = \boldsymbol{\lambda}^T \boldsymbol{C}_A \boldsymbol{\lambda}$, where $\boldsymbol{C}_A$ is the relevant covariance matrix. The entries of the covariance matrix can be found as

$$\begin{aligned}
&\boldsymbol{C}_A(i,j) \\
&\quad = \quad E[\{d_i - \int_A b(x)\, d\mu(x)\}\, \{d_j - \int_A b(x)\, d\mu(x)\}] \\
&\quad = \quad E[\{b(x_i) - \int_A b(x)\, d\mu(x)\}\, \{b(x_j) - \int_A b(x)\, d\mu(x)\}] \\
&\qquad\quad + E[R(x_i, t_i)\, R(x_j, t_j)] \qquad\qquad (4.4) \\
&\quad = \quad -\gamma_b(|x_i - x_j|) + \int_A \gamma_b(|x_i - x|)\, d\mu(x) + \int_A \gamma_b(|x_j - x|)\, d\mu(x) \\
&\qquad\quad - \int_A \int_A \gamma_b(|x - y|)\, d\mu(x)\, d\mu(y) - \gamma_R(|x_i - x_j|, |t_i - t_j|) + \sigma_R^2,
\end{aligned}$$

Note that the terms $\int_A \gamma_b(|x_i - x|)\, d\mu(x)$ show that the centered $b(\cdot)$ field, $b(x) - \int_A b(y)\, d\mu(y)$, is not a second order stationary field. Using this decomposition we can express the estimation error variance as

$$\begin{aligned}
&Var[m_A - \hat{m}_A] \\
&= \quad -\sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i\, \lambda_j\, \gamma_b(|x_i - x_j|) + 2 \sum_{i=1}^{n} \lambda_i \int_A \gamma_b(|x_i - x|)\, d\mu(x) \\
&\qquad - \int_A \int_A \gamma_b(|x - y|)\, d\mu(x)\, d\mu(y) - \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i\, \lambda_j\, \gamma_R(|x_i - x_j|, |t_i - t_j|) + \sigma_R^2 \\
&= \quad -\sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i\, \lambda_j\, \gamma_b(|x_i - x_j|) + 2 \sum_{i=1}^{n} \lambda_i \int_A \gamma_b(|x_i - x|)\, d\mu(x) \\
&\qquad - \int_A \int_A \gamma_b(|x - y|)\, d\mu(x)\, d\mu(y) \qquad\qquad (4.5) \\
&\qquad + \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i\, \lambda_j\, (\sigma_R^2 - \gamma_S(|x_i - x_j|))\, h(|t_i - t_j|).
\end{aligned}$$

The error variance of the estimate of the time-trend of the whole region, m_D, is obtained by substituting D for A in the above expression.

If we want to estimate the time-trend at a certain location x, we let A shrink to the point x such that the surface measure, $d\mu(x)$, becomes the point mass measure. In this case the expression for the error variance turns into

$$\begin{aligned} Var[m_x - \hat{m}_x] &= -\sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j \gamma_b(|x_i - x_j|) + 2\sum_{i=1}^{n} \lambda_i \gamma_b(|x_i - x|) \\ &\quad + \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j (\sigma_R^2 - \gamma_S(|x_i - x_j|))\, h(|t_i - t_j|). \end{aligned} \tag{4.6}$$

The 'optimal' coefficients, which minimize the variance of the estimation error subject to the unbiasedness constraint, can be found using Lagrange multipliers. The error variance expressions given above can be written in the general quadratic form

$$f(\boldsymbol{\lambda}) = \boldsymbol{\lambda}^T \boldsymbol{C} \boldsymbol{\lambda} - 2\boldsymbol{\lambda}^T \boldsymbol{v} + constant, \tag{4.7}$$

where $\boldsymbol{C}$ and $\boldsymbol{v}$ are generic terms for a covariance matrix and vector respectively. If there is no linear dependence between the observations, the covariance matrix will be positive definite. In the nonsingular case, the minimizing $\boldsymbol{\lambda}$ is

$$\boldsymbol{\lambda} = \boldsymbol{C}^{-1}\{\boldsymbol{v} + \mathbf{1}(1 - \mathbf{1}^T \boldsymbol{C}^{-1} \boldsymbol{v})/(\mathbf{1}^T \boldsymbol{C}^{-1} \mathbf{1})\}, \tag{4.8}$$

where $\mathbf{1}$ is the vector all of whose entries are one. Specific examples of estimators of the above type, implicitly defined by an error variance expression, are found in (4.5) and (4.6), and also in the next subsections.

5 ESTIMATION OF REGIONAL TEMPERATURE CHANGE

Next we will consider estimating the average annual temperature change over a given period of time averaged over a geographic area. As mentioned, each month is treated separately and the dependence on month is suppressed in the notation. The change per year from year t_a to year t_b averaged over the subregion A is denoted

$$\begin{aligned} M_A(t_a, t_b) &= \int_A \{T(x, t_b) - T(x, t_a)\}\, d\mu(x)\, m^{-1} \\ &= \int_A b(x)\, d\mu(x) + \sum_{t=t_a+1}^{t_b} \int_A R(x,t)\, d\mu(x)\, m^{-1} \end{aligned} \tag{5.1}$$

with $m = t_b - t_a$. As in the previous section, the estimator will be the unbiased linear combination of the observed year-to-year temperature differences, which minimizes the variance of the estimation error. The estimator will take the form

$$\hat{M}_A(t_a, t_b) \quad = \quad \sum_{i=1}^{n} d_i \, \lambda_i = \boldsymbol{\lambda}^T \, \boldsymbol{d}. \tag{5.2}$$

The unbiasedness criterion forces the sum of the λ_i's to be unity as before. Given this constraint, the variance of the estimation error for the regional temperature change can be expressed as

$$Var[M_A(t_a, t_b) - \hat{M}_A(t_a, t_b)] \quad = \quad \boldsymbol{\lambda}^T \, \boldsymbol{C}_A[t_a, t_b] \, \boldsymbol{\lambda},$$

where $\boldsymbol{C}_A[t_a, t_b]$ is a covariance matrix with entries

$$\begin{aligned}
&\boldsymbol{C}_A[t_a, t_b](i, j) \\
&= \quad E[\{b(x_i) - \int_A b(x) \, d\mu(x)\} \, \{b(x_j) - \int_A b(x) \, d\mu(x)\}] \\
&\quad + E[\{R(x_i, t_i) - \sum_{t=t_a+1}^{t_b} \int_A R(x, t) \, d\mu(x) \, m^{-1}\} \cdot \\
&\quad \{R(x_j, t_j) - \sum_{t=t_a+1}^{t_b} \int_A R(x, t) \, d\mu(x) \, m^{-1}\}] \\
&= \quad -\gamma_b(|x_i - x_j|) + \int_A \gamma_b(|x_i - x|) \, d\mu(x) + \int_A \gamma_b(|x_j - x|) \, d\mu(x) \\
&\quad - \int_A \int_A \gamma_b(|x - y|) \, d\mu(x) \, d\mu(y) - \gamma_R(|x_i - x_j|, |t_i - t_j|) \\
&\quad + \int_{\mathcal{A}} \gamma_R(|x_i - x_\eta|, |t_i - t_\eta|) \, d\mu(\eta) + \int_{\mathcal{A}} \gamma_R(|x_j - x_\eta|, |t_j - t_\eta|) \, d\mu(\eta) \\
&\quad - \int_{\mathcal{A}} \int_{\mathcal{A}} \gamma_R(|x_\eta - x_{\acute{\eta}}|, |t_\eta - t_{\acute{\eta}}|) \, d\mu(\eta) \, d\mu(\acute{\eta}),
\end{aligned} \tag{5.3}$$

with
$\eta = (x, t)$ reference in the space-time volume,

$\mathcal{A} = A \oplus \langle t_a + 1, \ldots, t_b \rangle$ surface in the space-time volume, with $\mu(\mathcal{A}) = 1$.

Using this decomposition we can express the estimation error variance in terms of the semivariograms

$$
\begin{aligned}
&Var[M_A(t_a,t_b) - \hat{M}_A(t_a,t_b)] \\
&= \quad -\sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i\,\lambda_j\,\gamma_b(|x_i - x_j|) + 2\sum_{i=1}^{n}\lambda_i \int_A \gamma_b(|x_i - x|)\,d\mu(x) \\
&\quad -\int_A\int_A \gamma_b(|x-y|)\,d\mu(x)\,d\mu(y) - \sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\,\lambda_j\,\gamma_R(|x_i - x_j|, |t_i - t_j|) \\
&\quad + 2\sum_{i=1}^{n}\lambda_i \int_{\mathcal{A}} \gamma_R(|x_i - x_\eta|, |t_i - t_\eta|)\,d\mu(\eta) \\
&\quad -\int_{\mathcal{A}}\int_{\mathcal{A}} \gamma_R(|x_\eta - x_{\acute{\eta}}|, |t_\eta - t_{\acute{\eta}}|)\,d\mu(\eta)\,d\mu(\acute{\eta})
\end{aligned} \tag{5.4}
$$

and thereby get an expression for the unbiased minimum error variance estimator. The error variance of the $M_D(t_a, t_b)$ estimate, the actual temperature change over the whole region, is obtained by letting $A = D$. If we restrict the observations to be in the (t_a, t_b) time window and constrain the estimation coefficients to sum to $1/m$ for each of the m years, the expression for the error variance does not involve the temporal component of the semivariogram $\gamma_T(\cdot)$. In this case, $\gamma_R(|\Delta x|, |\Delta t|)$ can be replaced by $\gamma_S(|\Delta x|)\,h(|\Delta t|)$ in (5.4).

To estimate the temperature change per year at a certain point, $\{T(x,t_b) - T(x,t_a)\}\,m^{-1}$, we let A shrink to the point x. Doing this, the expression for the error variance becomes

$$
\begin{aligned}
&Var[M_x(t_a,t_b) - \hat{M}_x(t_a,t_b)] \\
&= \quad -\sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i\,\lambda_j\,\gamma_b(|x_i - x_j|) + 2\sum_{i=1}^{n}\lambda_i\,\gamma_b(|x_i - x|) \\
&\quad -\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\,\lambda_j\,\gamma_R(|x_i - x_j|, |t_i - t_j|) + 2\sum_{i=1}^{n}\sum_{t=t_a+1}^{t_b}\lambda_i\,\gamma_R(|x_i - x|, |t_i - t|)\,m^{-1} \\
&\quad \sigma_R^2 \sum_{t=t_a+1}^{t_b}\sum_{\tau=t_a+1}^{t_b}(1 - h(|t-\tau|))\,m^{-2}..
\end{aligned}
$$

The estimating coefficients based on the above error variance expressions are found as in (4.8).

6 MISSING DATA

Hitherto the estimate has been based on the one-year temperature differences. However, monitoring data typically have missing values and cannot be represented in terms of one-year differences only. For example, if a time series of yearly temperature observations at a certain spatial sampling station has a missing data-item, two one-year differences will be left undefined when we do the differencing. Therefore, if the i'th observed temperature corresponds to location x_i and time t_i, and if the previous observation at the same location x_i occurred at time $t_i - k_i$, we define d_i as

$$d_i = T(x_i, t_i) - T(x_i, t_i - k_i). \tag{6.1}$$

We will show how to modify the error variance expressions for the estimates discussed above according to this generalization in the definition of d_i.

As in Section 4 we want to estimate the regional time-trend, but based on general multi-year differences. The time-trend and its estimator are defined as in (4.1) and (4.2). However, the unbiasedness criterion translates into

$$\sum_{i=1}^{n} \lambda_i \, k_i \quad = \quad 1.$$

To simplify the derivation of the error variance expression, we introduce a transformation matrix $\boldsymbol{P}$ such that

$$\boldsymbol{d} \quad = \quad \boldsymbol{P}^T \, \tilde{\boldsymbol{d}} \tag{6.2}$$

where $\tilde{\boldsymbol{d}}$ is the vector of one-year differences assuming that there are no missing data in the temperature time series at the different spatial sampling stations. The i'th row of the transformation matrix $\boldsymbol{P}^T$ will have k_i unit entries in the positions corresponding to the one-year differences comprising the multi-year difference d_i. For example, if we had observed the temperature field at two stations over three years, and the first station had a missing observation in the second year, the matrix $\boldsymbol{P}^T$ would read

$$\boldsymbol{P}^T \quad = \quad \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{6.3}$$

assuming the one-year temperature differences in $\tilde{\boldsymbol{d}}$ are ordered with all spatial sampling stations in year one first and then according to year.

The estimate can now be written as

$$\begin{aligned} \hat{m}_A \quad &= \quad \boldsymbol{\lambda}^T \, \boldsymbol{d} = \boldsymbol{\lambda}^T \, \boldsymbol{P}^T \, \tilde{\boldsymbol{d}} \\ &\equiv \quad \tilde{\boldsymbol{\lambda}}^T \, \tilde{\boldsymbol{d}}, \end{aligned}$$

and we get the desired expressions for the error variance simply by using the covariance matrix obtained by the transformation $\boldsymbol{P}$

$$\begin{aligned} Var[m_A - \hat{m}_A] &= \tilde{\boldsymbol{\lambda}}^T \boldsymbol{C}_A \tilde{\boldsymbol{\lambda}} \\ &= \boldsymbol{\lambda}^T \boldsymbol{P}^T \boldsymbol{C}_A \boldsymbol{P} \boldsymbol{\lambda}. \end{aligned} \tag{6.4}$$

Note that $\sum_{i=1}^n \tilde{\lambda}_i = 1$, and that the entries of the relevant covariance matrix are found by summing appropriate elements of $\boldsymbol{C}_A$, defined as in (4.4).

The same procedure can be used in order to obtain the appropriate covariance matrix in the case that there are missing data and we are estimating the regional temperature change over a specified time period.

7 RECTANGULAR DATA PANEL

In this section we will assume a rectangular data panel, that is the index set $\{i\}$ of the data is representable by $\{i\} = \{j\} \times \{k\}$, where $\{j\}$ represents uniformly spaced observation times and $\{k\}$ represents the monitoring locations having an arbitrary spatial distribution. The structure inherent in this pattern makes the problem somewhat more amenable for analytic examination than the general case. In order to elucidate some qualitative features of the estimation error variance, and since the rectangular data panel formulation is of interest in its own right, we will study it in some detail. In this context we examine how the error variance depends upon the number of years of observation and on the spatial sampling configuration. Furthermore, we compare the trend estimate based on the direct data, the temperatures, with that based on the differences. We will assume that there are no missing data in the data panel, however, cases with missing data can in general be handled by modifying solutions associated with structured systems; see the Appendix.

7.1 Regional Time-Trend Estimation

Let the observed differences be indexed as indicated above. Assume that we are estimating the time-trend in a region A in the same way as above. In this case we can write the error variance as a tensor product,

$$\begin{aligned} Var[m_A - \hat{m}_A] &= Var[\sum_{i=1}^n \lambda_i\, b(x_i) - \int_A b(x)\, d\mu(x)] + Var[\sum_{i=1}^n \lambda_i\, R(x_i, t_i)] \\ &= \boldsymbol{\lambda}^T \{(\mathbf{1}\,\mathbf{1}^T) \oplus \boldsymbol{C}_b\}\, \boldsymbol{\lambda} + \boldsymbol{\lambda}^T \boldsymbol{H}_m \oplus \boldsymbol{C}_R\, \boldsymbol{\lambda}, \end{aligned} \tag{7.1}$$

where

$$
\begin{aligned}
\boldsymbol{H}_m(i,j) &= h(|i-j|) \quad 1 \le i,j \le m, \\
\boldsymbol{C}_R(i,j) &= E[R(x_i,t)\,R(x_j,t)] \\
&= \sigma_R^2 - \gamma_S(|x_i - x_j|) \quad 1 \le i,j \le l, \\
\boldsymbol{C}_b(i,j) &= E[\{b(x_i) - \int_A b(x)\,d\mu(x)\}\,\{b(x_j) - \int_A b(x)\,d\mu(x)\}] \\
&= -\gamma_b(|x_i - x_j|) + \int_A \gamma_b(|x_i - x|)\,d\mu(x) + \int_A \gamma_b(|x_j - x|)\,d\mu(x) \\
&\quad - \int_A \int_A \gamma_b(|x-y|)\,d\mu(x)\,d\mu(y) \quad 1 \le i,j \le l,
\end{aligned} \tag{7.2}
$$

with

m number of observation years,

l number of spatial sampling stations,

$n = m \cdot l$ total number of observed differences.

It is assumed that the observations are ordered according to spatial location and then according to year, hence $\{x_i\}_{i=1}^{l}$ specify the locations of the spatial sampling stations. The tensor form is a consequence of our assumption that the covariance of the temperature residuals factors into spatial and temporal components as specified in (3.5).

Suppose that we choose the coefficients as

$$
\boldsymbol{\lambda} = \boldsymbol{\lambda}_t \oplus \boldsymbol{\lambda}_x \tag{7.3}
$$

with $\sum_{j=1}^{m} \lambda_{t,j} = 1$ and $\sum_{k=1}^{l} \lambda_{x,k} = 1$ (as shown below, if the intraregion variation of the centered $b(\cdot)$ field is negligible, that is $b(x) \approx b\ \forall x \in A$, the 'optimal' coefficients are of this form). Such a factorization implies that the coefficients at each station are found as $\boldsymbol{\lambda}_t$ scaled by $\lambda_{x,i}$, with $\lambda_{x,i}$ being the appropriate coefficient at station i when $m = 1$. The expression for the error variance can, given this factorization, be written

$$
\begin{aligned}
&Var[m_A - \hat{m}_A] \\
&\quad = Var[\sum_{k=1}^{l} \lambda_{x,k}\,b(x_k) - \int_A b(x)\,d\mu(x)] + Var[\sum_{k=1}^{l} \lambda_{x,k}\,R(x_k,t)]\,(\boldsymbol{\lambda}_t^T\,\boldsymbol{H}_m\,\boldsymbol{\lambda}_t) \\
&\quad = \boldsymbol{\lambda}_x^T\,\boldsymbol{C}_b\,\boldsymbol{\lambda}_x + \boldsymbol{\lambda}_x^T\,\boldsymbol{C}_R\,\boldsymbol{\lambda}_x\,(\boldsymbol{\lambda}_t^T\,\boldsymbol{H}_m\,\boldsymbol{\lambda}_t).
\end{aligned} \tag{7.4}
$$

For the minimizing coefficients we have

$$\begin{aligned} \lim_{m\to\infty} \boldsymbol{\lambda}_t^T \boldsymbol{H}_m \boldsymbol{\lambda}_t &= \lim_{m\to\infty} \sigma_R^{-2}\, Var[\sum_{j=1}^m \lambda_{t,j}\, R(x,j)] \qquad (7.5) \\ \leq \lim_{m\to\infty} \sigma_R^{-2}\, Var[\sum_{j=1}^m R(x,j)/m] &= \lim_{m\to\infty} \sigma_R^{-2}\, Var[(\delta(x,m)-\delta(x,0))/m] \\ = 0. \end{aligned}$$

Consequently, the attainable asymptotic precision is found as

$$\lim_{m\to\infty} Var[m_A - \hat{m}_A] \;=\; Var[\sum_{k=1}^l \lambda_{k,x}\, b(x_k) - \int_A b(x)\, d\mu(x)]$$

with $\boldsymbol{\lambda}_x$ minimizing the limiting variance expression. Hence, for a fixed set of stations, the variability of the centered $b(\cdot)$ field limits the precision of the regional time-trend estimate. This will also be the limiting error variance if we observe temperature changes $T(x_i,t_b) - T(x_i,t_a)$ with $t_b - t_a \geq t$ for any $t \geq 1$ at each station i, but that we have an otherwise arbitrary missing pattern, since we then effectively sample the $b(\cdot)$ field directly.

If we use only the net temperature change between the beginning and end of the monitoring period, we reduce the data-set to one observation from each station: $T(x_i,t_b) - T(x_i,t_a) \;\; 1 \leq i \leq l$. Equivalently, we force $\boldsymbol{\lambda}_t = \mathbf{1}\, m^{-1}$, and the coefficients do not change with time. In this case the error variance becomes

$$Var[m_A - \hat{m}_A] \;=\; \boldsymbol{\lambda}_x^T \boldsymbol{C}_b\, \boldsymbol{\lambda}_x + \boldsymbol{\lambda}_x^T \boldsymbol{C}_R\, \boldsymbol{\lambda}_x\, \mathcal{H}_m\, m^{-2}$$

where

$$\mathcal{H}_m = \sum_{i=1}^m \sum_{j=1}^m h(|i-j|).$$

Hence, the 'cost' in terms of increase in the mean square estimation error associated with the constraint $\boldsymbol{\lambda}_t = \mathbf{1}\, m^{-1}$ is determined by the factor

$$(\mathcal{H}_m\, m^{-2}) \,/\, (\boldsymbol{\lambda}_t^T\, \boldsymbol{H}_m\, \boldsymbol{\lambda}_t) \;=\; (\mathbf{1}^T \boldsymbol{H}_m \mathbf{1}\, m^{-2})\, (\mathbf{1}^T \boldsymbol{H}_m^{-1} \mathbf{1})$$

with $\boldsymbol{\lambda}_t$ denoting the 'optimal' temporal coefficients given the factorization in (7.3). If we assume the temperature residuals in different years are uncorrelated, then the matrix $\boldsymbol{H}_m$ will be tridiagonal with 1's on the diagonal and $-1/2$'s on the first off diagonals. In this case, $\mathcal{H}_m = 1$, and the above factor is found to be $(m+1)(m+2)\,/\,(6\,m) \sim m\,/\,6$. Hence, the 'cost' factor grows linearly with the number of years of observation. Moreover, in this specific case, the temporal coefficients, $\boldsymbol{\lambda}_t$ are found as

$$\lambda_{t,i} \;=\; i\,\{(m+1)-i\}\,6\,/\,\{m\,(m+1)\,(m+2)\}, \qquad (7.6)$$

and the set of coefficients has an inverted parabolic shape.

As mentioned, the 'optimal' coefficients can be written as a tensor product if $b(x) = b\ \forall x \in A$. This follows directly from the expression for the coefficients

$$\begin{aligned} \boldsymbol{\lambda} &= \boldsymbol{H}_m^{-1} \oplus \boldsymbol{C}_R^{-1}\,\mathbf{1}/\{\mathbf{1}^T\,\boldsymbol{H}_m^{-1} \oplus \boldsymbol{C}_R^{-1}\,\mathbf{1}\} \\ &= \{\boldsymbol{H}_m^{-1}\,\mathbf{1}/(\mathbf{1}^T\,\boldsymbol{H}_m^{-1}\,\mathbf{1})\} \oplus \{\boldsymbol{C}_R^{-1}\,\mathbf{1}/(\mathbf{1}^T\,\boldsymbol{C}_R^{-1}\,\mathbf{1})\} \\ &= \boldsymbol{\lambda}_t \oplus \boldsymbol{\lambda}_x, \end{aligned} \qquad (7.7)$$

where $\boldsymbol{\lambda}_t$ are the 'optimal' coefficients we obtain if we have data from a single station only. Similarly, $\boldsymbol{\lambda}_x$ are the 'optimal' coefficients obtained if data are available at all stations for a single year. The computational burden of computing 'optimal' coefficients is therefore significantly smaller than in the general case.

Suppose we are interested in solving the 'forward problem', that is simulating a realization of the temperature difference field from a given model. If we simulate the field on a rectangular data panel, the tensor product form of the covariance matrix can be used in order to 'factor' the simulation of the field

$$\begin{aligned} \boldsymbol{C}_A &= \boldsymbol{H}_m \oplus \boldsymbol{C}_R \\ &= (\boldsymbol{L}_H\,\boldsymbol{L}_H^T) \oplus (\boldsymbol{L}_C\,\boldsymbol{L}_C^T) \\ &= (\boldsymbol{L}_H \oplus \boldsymbol{L}_C)\,(\boldsymbol{L}_H^T \oplus \boldsymbol{L}_C^T) \\ &= \boldsymbol{L}\,\boldsymbol{L}^T \end{aligned}$$

with $\boldsymbol{L}$ denoting the Cholesky triangle. Hence, we need only decompose small subsystems in order to solve the overall task. The above decomposition implies that we can simulate each year independently and then combine the years as if we were combining independent random variables in order to simulate a time series of differences. If there is an exponential correlation structure in the time dimension, $h(\Delta t) = \exp(-a|\Delta t|)$, the implied Markov property simplifies the calculations even further. Note that these results generalize to higher dimensions, for instance, if the correlation structure in the space dimensions factors and the spatial data configuration is a rectangular grid.

The analysis of space-time data is often performed by considering the lumped time-series obtained by spatially averaging the data. For the considered problem, it follows from the above that this actually is 'optimal' in the case $b(x) \equiv b$ provided the spatial average is taken as the one defined by the spatial coefficients $\boldsymbol{\lambda}_x$. In fact, letting y_i denote the spatial average in year i, the 'optimal' temporal coefficients , $\boldsymbol{\lambda}_t$, solve the univariate least squares regression problem

$$\boldsymbol{y} \;=\; \mathbf{1}\, b + \boldsymbol{\varepsilon} \tag{7.8}$$

with b being the time-trend parameter to be estimated, y_i the spatial average of the differences in year i and $\boldsymbol{\varepsilon}$ a zero mean noise vector with correlation matrix $\boldsymbol{H}_m$, (assuming $b(x) \equiv b$).

Consider now the problem of basing the time-trend estimate on the temperatures themselves rather than the time-differenced temperatures. If we formulate the linear regression problem in terms of all the temperatures in the rectangular data panel the regression formulation is (retaining the previous notation)

$$\begin{aligned} \boldsymbol{y} &= \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ \vdots & \vdots \\ 1 & m \end{bmatrix} \oplus \mathbf{1} \begin{bmatrix} \mu_0 \\ b \end{bmatrix} + \boldsymbol{\varepsilon} \\ &= \boldsymbol{X} \oplus \mathbf{1} \begin{bmatrix} \mu_0 \\ b \end{bmatrix} + \boldsymbol{\varepsilon} \end{aligned} \tag{7.9}$$

where, in this case, the y_i's represent temperature observations and are ordered in $\boldsymbol{y}$ according to station number and then year. Furthermore, we have assumed a constant baseline temperature, $\mu_0(x) \equiv \mu_0$, which is estimated simultaneously in this formulation. The covariance matrix of $\boldsymbol{\varepsilon}$ becomes that of the temperature residuals: $(\sigma_\delta/\sigma_R)^2 \boldsymbol{H} \oplus \boldsymbol{C}_R$, where $\boldsymbol{H}$ is an $(m+1) \times (m+1)$ matrix defined by $\boldsymbol{H}(i,j) = \rho_T(|i-j|)$. The associated least squares coefficients can be expressed as

$$\begin{aligned} \boldsymbol{\Lambda} &= \{\boldsymbol{H}^{-1}\, \boldsymbol{X}\, (\boldsymbol{X}^T\, \boldsymbol{H}^{-1}\, \boldsymbol{X})^{-1}\} \oplus \{\boldsymbol{C}_R^{-1}\, \mathbf{1}\, (\mathbf{1}^T\, \boldsymbol{C}_R^{-1}\, \mathbf{1})^{-1}\} \\ &= \boldsymbol{\Lambda}_t \oplus \boldsymbol{\lambda}_x \end{aligned} \tag{7.10}$$

with $\boldsymbol{\Lambda}_t$ now being a $(m+1) \times 2$ matrix. Hence, the coefficient of each parameter estimate can be found as the tensor product of a set of space coefficients with a set of temporal coefficients. The spatial coefficients are seen to be the same as in the case with estimation based on time-differenced temperatures, whereas the temporal coefficients solve the following least squares regression problem

$$\boldsymbol{y} \;=\; \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ \vdots & \vdots \\ 1 & m \end{bmatrix} \begin{bmatrix} \mu_0 \\ b \end{bmatrix} + \varepsilon \tag{7.11}$$

with y_i the spatially averaged temperature in year $i-1$, and the covariance matrix of ε being $\boldsymbol{H}$. If $\boldsymbol{H}$ is diagonal the temporal coefficients of the b estimate are

$$\boldsymbol{\Lambda}_t(i,2) = 12\,\{(i-1) - m/2\}\,/\,\{m\,(m+1)\,(m+2)\},$$

which translates into the coefficients given in (7.6) when the estimate is expressed in terms of the one-year temperature differences. Hence, the time-trend estimates based on the differences and the temperatures respectively coincide. This will be the case quite generally. For a time-trend estimator based on the temperatures to be unbiased, the associated coefficients, Λ_i, of the temperatures must satisfy

$$\sum_{i=1}^{n} \Lambda_i t_i \;=\; 1,$$

where t_i is observation time in years relative to a baseline year. Furthermore, since the baseline temperature field under the model assumption is an unknown parameter field, the weights associated with each spatial sampling location must sum to zero. It follows that in the least squares formulation we minimize over the same linear subspace of the temperatures, whether we base the estimate on the temperatures or on the differences. Hence the estimates coincide. Note that if we consider the baseline temperature field, $\mu_0(x)$, as a constant parameter or as a stochastic field with known structure the estimates will in general not coincide. However, typically for space-time problems, the cost of the additional constraint associated with a priori forming the differences will be a low price to pay to avoid having to deal with the baseline field. In the above example, we assumed a constant baseline temperature, but the time-trend estimates coincided due to symmetry in the temporal pattern of observations at each spatial sampling station.

7.2 Regional Temperature Change Estimation

Using a rectangular data panel with l monitoring sites and m observation times we now estimate the actual regional temperature change over the period corresponding to the time span of the data panel. Also in this problem the estimation error variance can be specified in terms of a tensor product

$$
\begin{aligned}
& Var[M_A(t_a,t_b) - \hat{M}_A(t_a,t_b)] \\
= \;& Var[\sum_{i=1}^{n} \lambda_i \, b(x_i) - \int_A b(x)\, d\mu(x)] \\
& + Var[\sum_{i=1}^{n} \lambda_i \, R(x_i,t_i) - \sum_{t=t_a+1}^{t_b} \int_A R(x,t)\, d\mu(x)\, m^{-1}] \qquad (7.12) \\
= \;& \boldsymbol{\lambda}^T \{(\mathbf{1}\,\mathbf{1}^T) \oplus \boldsymbol{C}_b\} \,\boldsymbol{\lambda} + \boldsymbol{\lambda}^T \boldsymbol{H}_m \oplus \boldsymbol{C}_R \,\boldsymbol{\lambda} - 2\boldsymbol{\lambda}^T \,(\boldsymbol{H}_m \,\mathbf{1}\, m^{-1}) \oplus \boldsymbol{\Gamma} + \Lambda
\end{aligned}
$$

where $m = t_b - t_a$ and $\boldsymbol{C}_b$, $\boldsymbol{H}_m$ and $\boldsymbol{C}_R$ are defined as in (7.2). Furthermore, $\boldsymbol{\Gamma}$ is a $l \times 1$ vector and Λ a scalar, defined by

$$
\begin{aligned}
\Gamma(i) &= \int_A E[R(x_i,t)\, R(x,t)]\, d\mu(x), \\
\Lambda &= \int_A \int_A E[R(x,t)\, R(y,t)]\, d\mu(x)\, d\mu(y)\, \mathcal{H}_m \, m^{-2}.
\end{aligned}
$$

If we reduce the data to the net total temperature change, $T(x_i,t_b) - T(x_i,t_a)$, i.e. forcing $\boldsymbol{\lambda}_t = \mathbf{1}\, m^{-1}$, the error variance becomes

$$
\begin{aligned}
& Var[M_A(t_a,t_b) - \hat{M}_A(t_a,t_b)] \\
= \;& Var[\sum_{k=1}^{l} \lambda_{x,k} \, b(x_k) - \int_A b(x)\, d\mu(x)] \\
& + Var[\sum_{k=1}^{l} \lambda_{x,k} \, R(x_k,t) - \int_A R(x,t)\, d\mu(x)]\, \mathcal{H}_m \, m^{-2}.
\end{aligned}
$$

Thus, the time-dependent part of the error variance decays like $\mathcal{H}_m \, m^{-2}$ with respect to the number of years of observation. By the argument in (7.5) $\lim_{m\to\infty} \mathcal{H}_m \, m^{-2} = 0$, hence

$$
\lim_{t_b - t_a \to \infty} Var[M_A(t_a,t_b) - \hat{M}_A(t_a,t_b)] \;=\; Var[\sum_{k=1}^{l} \lambda_{x,k} \, b(x_k) - \int_A b(x)\, d\mu(x)]
$$

with $\boldsymbol{\lambda}_x$ minimizing the limiting variance expression.

In the case that the temperature time-trend is constant in A, $b(\cdot) \equiv b$, the 'optimal' coefficients can be expressed as $\boldsymbol{\lambda} = m^{-1}\, \mathbf{1} \oplus \tilde{\boldsymbol{\lambda}}_x + \nu \, \boldsymbol{\lambda}_t \oplus \boldsymbol{\lambda}_x$, where $\tilde{\boldsymbol{\lambda}}_x = \boldsymbol{C}_R^{-1} \boldsymbol{\Gamma}$ and $\nu = 1 - \mathbf{1}^T \tilde{\boldsymbol{\lambda}}_x$. Furthermore, with $\boldsymbol{\lambda}_t$ and $\boldsymbol{\lambda}_x$ being defined as in (7.7), and $\tilde{\boldsymbol{\lambda}}_x$ being the 'unconstrained minimizing coefficients' obtained if there are temperature differences available at all the stations for a single year only. Substituting the expression for the 'optimal' coefficients in the expression for the error variance we get

$$\begin{aligned} & Var[M_A(t_a,t_b) - \hat{M}_A(t_a,t_b)] \\ = \; & \{\int_A \int_A \sigma_R^2 \, \rho_S(|x-y|) \, d\mu(x) \, d\mu(y) - \boldsymbol{\Gamma}^T \boldsymbol{C}_R^{-1} \boldsymbol{\Gamma}\} \, \mathcal{H}_m \, m^{-2} \\ & + \nu^2 \{\boldsymbol{\lambda}_x^T \boldsymbol{C}_R \, \boldsymbol{\lambda}_x^T \, (\boldsymbol{\lambda}_t^T \boldsymbol{H}_m \, \boldsymbol{\lambda}_t)\}, \end{aligned}$$

which displays the time-span dependence of the components of the error variance.

8 AN ILLUSTRATION

We now return to the data set from the eastern-European steppe region introduced in Section 2 to illustrate calculations based on the model of Section 3. Model parameters are identified in Section 8.1, the estimates of time-trend and regional temperature change are presented in Section 8.2 along with a discussion of precision and sensitivity. In accordance with the discussion in Sections 4 and 5 the following analysis will be based on the annually differenced time series.

8.1 Parameter Estimation

First we focus on the inter-station or spatial correlation structure described by the spatial semivariogram functions $\gamma_b(\cdot)$ and $\gamma_S(\cdot)$; defined in (3.6) - (3.8). Aspects of this spatial structure are exhibited in Figure 6. Each dot in the semivariogram scatter corresponds to a station pair, the argument being the distance between the two locations. The vertical coordinate is calculated simply as

$$V_{i,j} \quad = \quad 1/2 \sum_{t \in S_{i,j}} \{d(x_i,t) - d(x_j,t)\}^2 / |S_{i,j}|,$$

where $S_{i,j}$ is the set of time indices for which differences are available at both the station located at x_i and the one located at x_j. The set has cardinality $|S_{i,j}|$. For each station pair this is a time-average of inter-station squared differences when the data are the time differenced temperatures.

Note that $V_{i,j}$ will be an unbiased estimator for $\gamma_b(|\Delta x|) + \gamma_S(|\Delta x|)$, with $\Delta x = x_i - x_j$. Figure 6 shows the semivariogram scatter plots for January, July and the annually averaged data. The solid lines in the plots are the fitted model-semivariograms. In all cases a linear model semivariogram was chosen.

It is seen that during summer there is less spatial variability in the temperature field. The spatial correlation structure was checked with respect to time-homogeneity and isotropy, which are both small effects in these data. More elaborate models and estimation schemes for spatial variation, capturing such effects as location dependent second order moments and anisotropy, are possible, see for instance Sampson and Guttorp (1992). However, for the present example, taking note of the fact that we will be estimating regionally averaged quantities, the above parameterization is believed to be adequate.

Next we turn to estimation of the semivariogram for the spatial modulation of the trend field, $b(\cdot)$, which together with the above estimate define the spatial semivariogram for the residual field $R(\cdot,\cdot)$. Figure 7 is being used for exploring the structure of the $\gamma_b(\cdot)$ semivariogram. The plots correspond to data in January, July and the annually averaged data set respectively. Rather than squaring spatial differences as above, for each station pair we formed products of spatial differences between the stations when the differences were separated in time by at least t_{min} years. Since the data analysis suggests that samples from the residual field $R(\cdot,\cdot)$ are uncorrelated when they are sufficiently separated in time the product will be an unbiased estimator for $\gamma_b(|\Delta x|)$, with $\Delta x = x_i - x_j$, if t_{min} is chosen large enough. In this example t_{min} was chosen to be four. Each station pair dot in the semivariogram scatter is obtained by using the distance between the stations as argument, and its vertical coordinate is calculated as

$$\bar{V}_{i,j} \quad = \quad 1/2 \sum_{(t,\bar{t}) \in \bar{S}_{i,j}} \{d(x_i,t) - d(x_j,t)\}\{d(x_i,\bar{t}) - d(x_j,\bar{t})\}/|\bar{S}_{i,j}|,$$

where $\bar{S}_{i,j}$ is the set of pairs $(t,\bar{t})$ such that: $t,\bar{t} \in S_{i,j}$ and $|t-\bar{t}| \geq t_{min}$. As above $S_{i,j}$ is the set of time indices for which differences are available at both stations. For each station pair, this is a time-average of products of inter-station differences that are separated by at least t_{min} years when the data are the time differenced temperatures. Since the estimate involves products, the scatter also takes on negative values. The plot shows that $\gamma_b(\cdot)$ is relatively small in magnitude compared to the corresponding $\gamma_R(\cdot)$ semivariogram. Assuming a linear semivariogram model also for the $b(\cdot)$ field, it is seen from the error variance expression (7.4), that a statement about smallness of $\gamma_b(\cdot)$ must be made relative to the value of $\Delta = \sigma_R^2 \; (\boldsymbol{\lambda}_t^T \, \boldsymbol{H}_m \, \boldsymbol{\lambda}_t) = \sigma_R^2 \; (\mathbf{1}^T \, \boldsymbol{H}_m^{-1} \, \mathbf{1})^{-1}$. For the problem at hand, $\Delta \approx .0001$ for the annually averaged data set. The average value of the scatter in Figure 7 concerning the annually averaged data is $-.001$. The noise in the scatter is large for an accurate estimate of the $\gamma_b(\cdot)$ semivariogram. However, for purposes of this illustration we take $\gamma_b(\cdot) \equiv 0$. Hence, the $\gamma_S(\cdot)$ semivariograms are specified by the solid lines of Figure 6.

As shown by the previous plot the spatial structure in the temperature field is strong. However, the analysis of Section 2 suggests that the temporal structure, exhibited by the function $h(\cdot)$ defined in (3.8), is rather weak. In order to estimate $h(\cdot)$, the autocorrelations of the centered time series of differences at each station were computed. Centering is with reference to the 'average' time series derived by averaging the time series at the monitoring sites having no missing data. The dotted lines in the plots of Figure 8 correspond to temporal autocorrelations for each of the 27 monitoring stations, and the solid lines to the simple mean over these. If $\gamma_b(\cdot) \equiv 0$, as we indeed have assumed, the solid line will be an unbiased estimator of $h(\cdot)$. The similarity of the structure in the plot with the pattern $\langle h(0)\ h(1)\ h(2)\ h(3) \cdots \rangle = \langle 1\ -.5\ 0\ 0 \cdots \rangle$, which corresponds to the $\delta(\cdot,\cdot)$ residuals in different years being uncorrelated, is rather striking. This particular pattern will also constitute our model assumption with respect to temporal structure.

Recall the structure of the semivariogram for the temperature differences

$$\begin{aligned} \gamma_d(|\Delta x|, |\Delta t|) &= \gamma_b(|\Delta x|) + \gamma_T(|\Delta t|) + \gamma_S(|\Delta x|)\, h(|\Delta t|), \\ \gamma_T(|\Delta t|) &= \sigma_R^2\ (1 - h(|\Delta t|)). \end{aligned}$$

Having estimated $\gamma_b(\cdot)$, $\gamma_S(\cdot)$ and $h(\cdot)$, we need only an estimate of $\gamma_T(1) = \sigma_R^2\ (1-h(1))$ in order to obtain an estimate of the semivariogram for the temperature differences. The quantity $\gamma_T(1)$ together with the estimate of $h(\cdot)$ define our estimate of the temporal semivariogram component. The estimate of $\gamma_T(1)$ is defined by

$$\hat{\gamma}_T(1) = 1/2 \sum_{(t,x_i)\in S} \{d(x_i,t) - d(x_i,t-1)\}^2/|S|,$$

where S is the set of time-location pairs for which temperature differences are available in consecutive years. We have thereby formed the second order differences of the temperature time series and thus removed the effect of the linear time-trend in order to enable estimation of the variance of the model residual. The numerical values of the σ_R estimates were 3.9, 1.6 and .8 degrees Celsius, for the data in January, July and the annually averaged data respectively. The variance of the annually averaged data is about 1/14 the value obtained averaging those of January and July, which suggests that the temperature field has a weak temporal correlation structure even on a month to month basis.

8.2 Numerical Results

Armed with the parameter estimates of the previous section, we are now able to calculate actual temperature time-trend and change estimates for the steppe region.

We first estimated, for each month, the annual regional temperature changes, $\hat{M}_D(t, t-1)$, as defined in (5.2) for years $1952 - 1990$. Due to the strong spatial structure the regional time series so defined has qualitatively the same behavior as the corresponding time-series at each station, though with a smaller variability. In the month July these estimates were in the range -3 to 3 degrees Celsius, and the associated estimated standard errors were approximately .1 degrees. Hence, we are able to estimate the regionally averaged temperature changes with high degree of accuracy.

The regional temperature time-trend estimate and the associated standard errors were calculated separately for the months January and July as well as for the annual average. The results are .019 degrees Celsius $+/-$ .033 degrees Celsius for January, $-.013$ degrees Celsius $+/-$ .013 degrees Celsius for July and .012 degrees Celsius $+/-$ .007 degrees Celsius for the annually averaged data. In view of the error estimates, we see that a warming trend cannot easily be discerned from a 'non-warming' scenario.

If we had complete data for these 24 stations over a period of ≈ 90 years, the estimate of the annually averaged temperature time-trend would have a standard error $\approx .002$ degrees Celsius for the eastern-European steppe region. Note that this is under the assumption that $\gamma_b(\cdot) \equiv 0$. When this is not the case, the error term associated with interpolating a spatially varying trend field will dominate when sufficiently many years of data are being used in the estimation.

The estimates of uncertainty presented above are estimates given the model. It is therefore of interest to examine the sensitivity of the uncertainty of the temperature time-trend estimate to the model parameters. Specifically, we examine how the error variance expression for the time-trend in the rectangular data-panel case depends on the model parameters. The expression for the error variance is given in (7.4) and we take $\gamma_b(\cdot) \equiv 0$.

Consider first the sensitivity of the error variance with respect to the spatial semivariogram, $\gamma_S(\cdot)$, as defined in (3.8). We will assume that the semivariogram $\gamma_S(\cdot)$ is approximately linear for short distances and examine the sensitivity with respect to the slope of the variogram. The factor $\sigma_R^{-2}\, \boldsymbol{\lambda}_x^T \boldsymbol{C}_R \boldsymbol{\lambda}_x = \sigma_R^{-2}\, (\mathbf{1}^T \boldsymbol{C}_R^{-1}\, \mathbf{1})^{-1}$ scales the error variance and the slope parameter is the only parameter involved in its definition. Hence, this factor embodies the sensitivity to the slope parameter. Define the relative slope to be the magnitude of the semivariogram at 1000 km divided by the variance. When we increased the relative slope parameter of the semivariogram $\gamma_S(\cdot)$ from .25 to .75, the square root of the factor $\sigma_R^{-2}\, \boldsymbol{\lambda}_x^T \boldsymbol{C}_R \boldsymbol{\lambda}_x$ decreased approximately linearly from .96 to .78. For these data, the relative slope was approximately .5. The sensitivity of the factor $\sigma_R^{-2}\, \boldsymbol{\lambda}_x^T \boldsymbol{C}_R \boldsymbol{\lambda}_x$ to the slope estimate, given a linear model over the chosen region, is thus rather modest.

At this point it is of interest to examine how an alternative design for the monitoring network affects the standard error of the temperature time-trend. The alternative design involves a subset of five monitoring stations, chosen to be dispersed in the region. These stations are those marked by 'crosses' in Figure 1. Changing the monitoring network only affects the factor $\sigma_R^{-2}\,\boldsymbol{\lambda}_x^T\,\boldsymbol{C}_R\,\boldsymbol{\lambda}_x$ in the error variance expression. The value of the square root of this factor was .87 in the case with the alternative network design and .86 in the case with all stations available.

Evidently, the estimate of the variance of the $R(\cdot,\cdot)$ field, σ_R^2, also scales the error variance estimate, and is thus an important parameter. However, with temperature residuals in different years being uncorrelated the precision of this parameter estimate will be high.

Finally, in order to examine the sensitivity with respect to temporal correlation structure, assume that the temperature residuals, $\delta(\cdot,\cdot)$, are exponentially correlated in time. That is $\rho_T(i) = \rho^i$, with the model assumption corresponding to $\rho = 0$. The factor $\boldsymbol{\lambda}_t^T\,\boldsymbol{H}_m\,\boldsymbol{\lambda}_t = (\boldsymbol{1}^T\,\boldsymbol{H}_m^{-1}\,\boldsymbol{1})^{-1}$ scales the error variance, and under the model ρ is the only parameter involved in its definition, as is seen from (3.8) and (7.2). In Figure 9 the numerical value of this factor has been plotted on a square root scale as a function of number of years of observation and for different values of ρ. We took $\rho = 0$, $1/3$ and $2/3$, corresponding to $h(1) = -1/2$, $-1/3$ and $-1/6$, respectively. In the example considered, with 39 years of data, increasing the value of ρ from 0 to $1/3$ approximately doubles the standard error.

9 CONCLUSION

This paper sets forth a statistical framework for the analysis of time trends of a spatial field observed periodically at fixed monitoring locations. We derive estimators for a temporally and regionally averaged trend which allow for evolution of the monitoring network and incomplete data, and which take advantage of the spatial-temporal statistics of the field. A probabilistic model leads to estimates of statistical precision.

An illustration using temperature data from the steppe region of eastern-Europe showed strong spatial structure and weak inter-annual temporal structure. Furthermore, a strong seasonal dependence was evident in the spatial structure. Spatial heterogeneity in the time trend could not be detected. We emphasize the assessment of the uncertainty associated with trend estimates. Based on 40 years of temperature observations, we estimated the regional temperature time-trend in eastern-Europe to be .012 degrees Celsius per year and the associated standard error to be .007 degrees Celsius.

References

Bilonick, R. A. (1988), "Monthly Hydrogen Ion Deposition Maps for the Northeastern U.S. from July 1982 to September 1984," *Atmospheric Environment*, Vol. 22, No. 9, 1909-1924.

Brown, P. J., Nhu, D. L., and Zidek J. V. (1994), "Multivariate spatial interpolation and exposure to air pollutants," *The Canadian Journal of Statistics*, Vol. 22, No. 4, 489-509.

Cressie, N. A. C. (1993), "Statistics for spatial data", New York: John Wiley.

Dietrich, C. R. (1993), "Computationally Efficient Cholesky Factorization of a Covariance Matrix with Block Toeplitz Structure," *Journal of Statistical Computation and Simulation*, 45, 203-218.

Handcock, M. S., and Wallis, J. R. (1994), "An Approach to Statistical Spatial-Temporal Modeling of Meteorological Fields," *Journal of the American Statistical Association*, Vol. 89, No. 426, 368-378.

Rodriguez-Iturbe, I., and Mejida, J. M. (1974), "The Design of Rainfall Networks in Time and Space," *Water Resources Research*, Vol. 10, No. 4, 713-728.

Rouhani, S., Ebrahimpour, R. M., Imran, Y., and Gianella, E. (1992), "Multivariate Geostatistical Trend Detection and Network Evaluation of Space-Time Acid Deposition Data - I. Methodology," *Atmospheric Environment*, Vol. 26A, No. 14, 2603-2614.

Rouhani, S., and Wackernagel, H. (1990), "Multivariate Geostatistical Approach to Space-Time Data Analysis," *Water Resources Research*, Vol. 26, No. 4, 585-591.

Sampson, P. D., and Guttorp, P. (1992), "Nonparametric Estimation of Nonstationary Spatial Covariance Structure," *Journal of the American Statistical Association*, Vol. 87, No. 417, 108-119.

Stein, M. (1986). "A Simple Model for Spatial-Temporal Processes," *Water Resources Research*, Vol. 22, No. 13, 2107-2110.

APPENDIX: INCOMPLETE DATA PANEL

The calculation of coefficients in the case of a rectangular data panel with missing data involves solving a system of the form

$$\boldsymbol{P}^T \boldsymbol{C} \boldsymbol{P} \boldsymbol{x} = b, \tag{A.1}$$

with $\boldsymbol{C}$ defined as in (4.4) and $\boldsymbol{P}$ having a structure as explained in (6.3). The transformation matrix $\boldsymbol{P}$ will be the identity if there are no missing data. When the correlation structure factors into spatial and temporal components $\boldsymbol{C}$ will be a structured matrix, a block Toeplitz matrix as indicated in (7.1).

Such structure calls for a fast linear systems solver. As shown in (7.7) the tensor product form of $\boldsymbol{C}$ can be exploited when solving systems involving this matrix. Furthermore, the temporal correlation matrix is Toeplitz and hence can be solved by a version of the fast algorithms making use of this structure,

The system at hand is not block Toeplitz, but if the set of missing data is a small subset of the data in the data panel, the transformed matrix will in a sense be 'close' to a structured matrix. In this appendix we illustrate how to exploit structure of the original system in the case with missing data using a simple Schur complement like approach. Assume thus that the system defined by the matrix $\boldsymbol{C}$ can be solved in a particularly efficient manner, see for instance Dietrich (1993) , and denote $\boldsymbol{V} = \boldsymbol{C}^{-1}$. For ease of exposition, assume that $\boldsymbol{P}$ has the form

$$\boldsymbol{P} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{P}_0 \end{bmatrix}. \tag{A.2}$$

That is, all the missing data are associated with the last set of coefficients in the full coefficient vector.
The corresponding decompositions of $\boldsymbol{b}$ and $\boldsymbol{C}^{-1}$ are denoted

$$\boldsymbol{b}^T = \begin{bmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 \end{bmatrix}^T,$$

$$\boldsymbol{V} = \begin{bmatrix} \boldsymbol{V}_{11} & \boldsymbol{V}_{12} \\ \boldsymbol{V}_{12}^T & \boldsymbol{V}_{22} \end{bmatrix}. \tag{A.3}$$

As will be seen, only $\boldsymbol{V_{22}}$ needs to be explicitly formed. The system (A.1) can now be written

$$\begin{aligned} \boldsymbol{x}_1 &= \boldsymbol{V}_{11}\,\boldsymbol{b}_1 + \boldsymbol{V}_{12}\,\tilde{\boldsymbol{b}}_2 \\ \boldsymbol{P}_0\,\boldsymbol{x}_2 &= \boldsymbol{V}_{12}^T\,\boldsymbol{b}_1 + \boldsymbol{V}_{22}\,\tilde{\boldsymbol{b}}_2, \end{aligned}$$

where the auxiliary variable $\tilde{\boldsymbol{b_2}}$, which satisfies $\boldsymbol{b}_2 = \boldsymbol{P}_0^T\,\tilde{\boldsymbol{b}}_2$, has been introduced. After elimination of $\tilde{\boldsymbol{b}}_2$ the solution can be found as

$$\begin{aligned} \boldsymbol{x}_1 &= [\boldsymbol{V}_{11} - \boldsymbol{V}_{12}\,\boldsymbol{V}_{22}^{-1}\,\boldsymbol{V}_{12}^T]\,\boldsymbol{b}_1 + \boldsymbol{V}_{12}\boldsymbol{V}_{22}^{-1}\,\boldsymbol{P}_0\,\boldsymbol{x}_2 \\ \boldsymbol{x}_2 &= \tilde{\boldsymbol{V}}_{22}\,\boldsymbol{P}_0^T\,\boldsymbol{V}_{22}^{-1}\,\boldsymbol{V}_{12}^T\,\boldsymbol{b}_1 + \tilde{\boldsymbol{V}}_{22}\,\boldsymbol{b}_2 \end{aligned} \tag{A.4}$$

with $\tilde{\boldsymbol{V}}_{22} = [\boldsymbol{P}_0^T\,\boldsymbol{V}_{22}^{-1}\,\boldsymbol{P}_0]^{-1}$. Thus we need to solve systems defined by the linear operators $\boldsymbol{V}_{22}$ and $\boldsymbol{P}_0^T\,\boldsymbol{V}_{22}^{-1}\,\boldsymbol{P}_0$ in addition to the original operator $\boldsymbol{C}$.

If the relative number of missing data is small this added computational cost might be small compared to the reduction in cost resulting from having obtained a full system which is structured. Note that if we do not introduce multi-year increments, but rather deal with the problem of missing data by discarding the one-year increments not being defined, the solution can be written as

$$\boldsymbol{x} \quad = \quad [\boldsymbol{V}_{11} - \boldsymbol{V}_{12}\,\boldsymbol{V}_{22}^{-1}\,\boldsymbol{V}_{12}^{T}]\,\boldsymbol{b} \tag{A.5}$$

with $\boldsymbol{C}^{-1}$ partitioned as in (A.3) and $\boldsymbol{V}_{22}$ being the block associated with the missing data. In this case $\boldsymbol{x} = \boldsymbol{x}_1$. This corresponds to applying the Schur complement of $\boldsymbol{V}_{22}$ to $\boldsymbol{b}$.