Spatial Process Modelling for Univariate and Multivariate Dynamic Spatial Data

Sudipto Banerjee

Division of Biostatistics, University of Minnesota, Minneapolis, Minnesota, U.S.A.

Dani Gamerman

Instituto de Matemática, Universidade Federal do Rio de Janeiro, Brazil.

Alan E. Gelfand

Institute of Statistics and Decision Sciences, Duke University, Durham, North Carolina, USA.

ABSTRACT: There is a considerable literature in spatiotemporal modelling. The approach adopted here applies to the setting where space is viewed as continuous, but time is taken to be discrete. We view the data as a time series of spatial processes and work in the setting of dynamic models, achieving a class of dynamic models for such data. We seek rich, flexible, easy-to-specify, easy-to-interpret, computationally tractable specifications which allow very general mean structures and also non-stationary association structures.

Our modelling contributions are as follows. In the case where univariate data is collected at the spatial locations, we propose the use of a spatiotemporally varying coefficients form. In the case where multivariate data is collected at the locations, we need to capture associations among measurements at a given location and time as well as dependence across space and time. We propose the use of suitable multivariate spatial process models developed through coregionalization.

We adopt a Bayesian inference framework. The resulting posterior and predictive inference enables summaries in the form of tables and maps, which help to reveal the nature of the spatiotemporal behaviour as well as the associated uncertainty. We illuminate various computational issues and then apply our models to the analysis of climate data obtained from the National Center for Atmospheric Research to analyze precipitation and temperature measurements obtained in Colorado in 1997.

Keywords: Bayesian inference; coregionalization; dynamic models; multivariate spatial processes; non-stationarity; spatially varying coefficients.
1 Introduction

Dramatic increases in computational capability have encouraged increased amounts of spatiotemporal data collection in many fields. For example, the application we investigate here involves climate data, in particular temperature and rainfall collected at several locations in a given region over a given period of time. That is, space is viewed as continuous and time is taken to be discrete. Spatiotemporal models for such data, by now, have a substantial literature. We offer only a brief review.

Early approaches include the STARMA (Pfeifer and Deutsch, 1980, a, b) and STARMAX (Stoffer, 1986) models. They add spatial covariance structure to standard time series models. More recently, Handcock and Wallis (1994) employ stationary Gaussian process models with an AR(1) model for the time series at each location to study global warming. Carroll et al. (1997) again use these processes, assuming a separable form for the space-time covariance function to study ground level ozone. See also Brown et al. (2000) for a modelling approach motivated by a stochastic differential equation which results in a nonseparable spatiotemporal covariance function.

Capturing nonstationary spatial structure within the modelling of spatiotemporal data has been developed in Guttrop et al. (1994), extending the deformation approach of Sampson and Guttorp (1992). Kernel convolution was employed by Higdon (1999) to achieve nonstationarity. Golan Kibria et al. (2002) achieve nonstationarity in a spatiotemporal setting through hierarchical modelling, allowing the spatial covariance structure for the data to vary about a stationary form using an Inverse-Wishart distribution.

There is also much recent spatiotemporal modelling, which employs Markov random field structure in the form of conditionally autoregressive (CAR) specifications. See, for example, Waller et al. (1997), who study disease mapping and Gelfand et al. (1998), who look at single family home sales. Pace et al. (2000) work with simultaneous autoregressive (SAR) models extending them to allow temporal neighbours as well as spatial neighbours. Eckert et al. (2003) attempts a survey in the context of real estate applications.

The approach taken in this paper is to view the data as arising from a time series of spatial processes. In particular, we work in the setting of dynamic models (West and Harrison, 1997), describing the temporal evolution in a latent space. We achieve a class of dynamic models for spatiotemporal data.
Here, there is a growing literature. Non Bayesian approaches include Huang and Cressie (1996), Wilde and Cressie (1999) and Mardia et al. (1998). Bayesian approaches include Tonellato (1997), Sansó and Guenni (1999), Stroud et al. (2001) and Huerta et al. (2003). The paper by Stroud et al. is attractive in being applicable to any data set which is continuous in space and discrete in time and allows straightforward computation using Kalman filtering.

Our contributions are as follows. First, we consider the case where univariate data is collected at each location. Here, we seek rich, flexible, easy-to-specify, easy-to-interpret, computationally tractable specifications, which allow very general (nonlinear) mean structure, and also non-stationary association structure. In particular, we propose a suitable extension of the spatially varying coefficients approach presented in Gelfand et al. (2003a) to accommodate temporal dependence. See also Gamerman, Moreira and Rue (2003), in this regard. We then turn to the case of multivariate data collection at the locations. Here it is evident that we need to enable dependence among measurements at a given location and time. Then, the dynamic spatial model will introduce dependence across space and time. Here we employ suitable extension of the multivariate model of coregionalization as described in Gelfand et al. (2003b).

We adopt a Bayesian inference framework. The resulting posterior and predictive inference enable summaries in the form of tables and maps which help to reveal the nature of the spatiotemporal behaviour as well as the associated uncertainty. The models are fitted through Markov Chain Monte Carlo methods and we offer some discussion of the computational details and the run time constraints imposed by a large number of locations and time points.

We demonstrate the performance of our modelling approach using a simulated data set. We then turn to a real weather data set obtained from the National Center for Atmospheric Research (NCAR), Boulder, CO. The data provide monthly (maximum) precipitation and temperature measurements obtained at 50 sites, over the twelve months in 1997. Also supplied is the elevation at each site. The univariate model attempts to explain temperature given precipitation. In the resulting dynamic model, the coefficient of precipitation is allowed to be spatiotemporally varying. In the absence of any other covariate information, it is expected that temperature will affect precipitation differently across the state. With time slices of the estimated coefficient surface, we can see this both dynamically and statically.

The multivariate model attempts to explain precipitation and temperature jointly, given elevation. We anticipate dependence between precipitation and temperature within and across sites and time periods.
Moreover, the variability in precipitation and the decay in spatial association between precipitation measurements, with regard to distance, will not be the same as those for temperature, so a suitably rich bivariate process is required. The coregionalization model enables this and allows us to infer about all types of variability and association.

The plan of the paper is as follows. Section 2.1 offers a brief review of the dynamic linear model framework. Section 2.2 adapts this framework to the univariate spatiotemporal process setting, with spatiotemporally varying coefficients. Section 3.1 provides a simulation example, while Section 3.2 provides the analysis of the temperature data given precipitation. Section 4 takes up the multivariate process setting, which Section 5 illustrates, first with a simulation example and then with a bivariate precipitation and temperature model. Section 6 provides a summary, discussion and extensions.

2 Univariate spatiotemporal dynamic models

2.1 Review of dynamic linear models

Dynamic linear models, often referred to as state-space models in the time-series literature, offer a versatile framework for fitting several time-varying models (West and Harrison, 1997). We brieﬂy outline the general dynamic linear modelling framework. Thus, let \( \mathbf{Y}_t \) be a \( m \times 1 \) vector of observables at time \( t \). \( \mathbf{Y}_t \) is related to a \( p \times 1 \) vector, \( \mathbf{\theta}_t \), called the state vector, through a measurement equation. In general, the elements of \( \mathbf{\theta}_t \) are not observable, but are generated by a first-order Markovian process, resulting in a transition equation. Therefore, we can describe the above framework as,

\[
\mathbf{Y}_t = \mathbf{F}_t \mathbf{\theta}_t + \mathbf{\epsilon}_t; \quad \mathbf{\epsilon}_t \sim N(\mathbf{0}, \mathbf{\Sigma}_\epsilon^t),
\]

\[
\mathbf{\theta}_t = \mathbf{G}_t \mathbf{\theta}_{t-1} + \mathbf{\eta}_t; \quad \mathbf{\eta}_t \sim N(\mathbf{0}, \mathbf{\Sigma}_\eta^t),
\]

where \( \mathbf{F}_t \) and \( \mathbf{G}_t \) are \( m \times p \) and \( p \times p \) matrices respectively. The first equation is the measurement equation, where \( \mathbf{\epsilon}_t \) is a \( m \times 1 \) vector of serially uncorrelated Gaussian variables with mean \( \mathbf{0} \) and an \( m \times m \) covariance matrix, \( \mathbf{\Sigma}_\epsilon^t \). The second equation is the transition equation with \( \mathbf{\eta}_t \) being a \( p \times 1 \) vector of serially uncorrelated zero-centered Gaussian disturbances and \( \mathbf{\Sigma}_\eta^t \) the corresponding \( p \times p \) covariance matrix. Note that, under
(1), the association structure can be computed explicitly across time, e.g., $\text{Cov}(\theta_t, \theta_{t-1}) = G_t \text{Var}(\theta_{t-1})$ and $\text{Cov}(Y_t, Y_{t-1}) = F_t G_t \text{Var}(\theta_{t-1}) F_t^T$.

$F_t$ (in the measurement equation) and $G_t$ (in the transition equation) are referred to as system matrices which may change over time. $F_t$ and $G_t$ may involve unknown parameters but, given these parameters, temporal evolution is in a pre-determined manner. The matrix $F_t$ is usually specified by the design of the problem at hand, while $G_t$ is specified through modelling assumptions; for example, $G_t = I_p$, the $p \times p$ identity matrix would render a nonstationary multivariate $AR(1)$ process for $\theta_t$. As a result, the system is linear, and for any time point $t$, $Y_t$ can be expressed as a linear combination of the present $\epsilon_t$ and the present and past $\eta_t$’s.

### 2.2 Dynamic spatiotemporal models

In this section we adapt the above dynamic modelling framework to univariate spatiotemporal models with spatially varying coefficients. For this we consider a collection of sites $S = \{s_1, \ldots, s_N\}$, and time-points $T = \{t_1, \ldots, t_N\}$, yielding observations $Y(s,t)$, and covariate vectors $x(s,t)$, for every $(s,t) \in S \times T$.

The response, $Y(s,t)$, is first modelled through a measurement equation, which incorporates the measurement error, $\epsilon(s,t)$, as serially and spatially uncorrelated zero-centered Gaussian disturbances. The transition equation now involves the regression parameters (slopes) of the covariates. The slope vector, say $\tilde{\beta}(s,t)$, is decomposed into a purely temporal component, $\beta_t$, and a spatiotemporal component, $\beta(s,t)$. Both these are generated through transition equations, capturing their Markovian dependence in time. While the transition equation of the purely temporal component is as in usual state-space modelling, the spatiotemporal component is generated by a multivariate Gaussian spatial process. Thus, we may write down the spatiotemporal modelling framework as,

$$Y(s,t) = \mu(s,t) + \epsilon(s,t); \epsilon(s,t) \overset{\text{iid}}{\sim} N\left(0, \sigma^2 \right).$$  \hspace{1cm} (2)

$$\mu(s,t) = x^T(s,t) \tilde{\beta}(s,t).$$

$$\tilde{\beta}(s,t) = \beta_t + \beta(s,t)$$  \hspace{1cm} (3)

$$\beta_t = \beta_{t-1} + \eta_t, \, \eta_t \overset{\text{iid}}{\sim} N_p\left(0, \Sigma_{\eta} \right).$$

$$\beta(s,t) = \beta(s, t-1) + w(s,t).$$
In (3), \( w(s,t) = A v(s,t) \), with \( v(s,t) = (v_1(s,t), ..., v_p(s,t))^T \). The \( v_l(s,t) \) are serially independent replications of a Gaussian processes with unit variance and correlation function \( \rho_l(\phi_l, \cdot) \) (henceforth denoted by \( GP(0, \rho_l(\phi_l, \cdot)) \), for \( l = 1, 2, ..., p \) and independent across \( l \). In the current work, we assume that \( A \) does not depend upon \( (s,t) \). Nevertheless, this still allows flexible modelling for the spatial covariance structure, as we discuss below.

Moreover, allowing a spatially-varying coefficient \( \beta(s,t) \) to be associated with \( x(s,t) \) provides an arbitrarily rich explanatory relationship for the \( x' \)s with regard to the \( Y' \)s. In this sense, our proposed mean structure is more flexible than that in any previous work while retaining straightforward and natural specification. By comparison, in Stroud et al. (2001), at a given \( t \), a locally weighted mixture of linear regressions is proposed and only the purely temporal component of \( \tilde{\beta}(s,t) \) is used. Such a specification requires both number of basis functions and number of mixture components.

Returning to our specification, note that if \( v_l(\cdot,t) \overset{\text{iid}}{\sim} GP(0, \rho(\phi_l, \cdot)) \), then they are serially independent and \textit{identical} Gaussian processes with a common correlation function and decay parameter and we immediately obtain the separable (intrinsic) model, implying the \( p \)-dimensional Gaussian process, \( w(\cdot,t) \overset{\text{iid}}{\sim} GP(0, \rho(\phi_l, \cdot) \Sigma_w) \). Here, \( \Sigma_w = AA^T \), which clearly shows that \( A \) can, without losing generality, be taken as the lower-triangular Cholesky square-root of \( \Sigma_w \). Separable modelling of spatial covariances has received much attention in the literature (see, e.g., Mardia and Goodall (1993), Le and Zidek (1992), and Banerjee and Gelfand (2002)). While allowing nice closed form results for model fitting, separable models impose a common correlation structure on the different components of \( w(\cdot,t) \), and hence of \( \beta(s,t) \). In this paper, we work with models where this assumption is relaxed.

Allowing different correlation functions and decay parameters for the \( v_l(s,t) \), i.e., \( v_l(\cdot,t) \overset{\text{iid}}{\sim} GP(0, \rho_l(\phi_l, \cdot)) \), so that they are still serially independent but no longer identical, leads to the linear model of coregionalization (Wackernagel, 1996; Gelfand et al., 2003). Consequently, the multivariate Gaussian process for \( w(\cdot,t) \) is given by, \( w(\cdot,t) \overset{\text{iid}}{\sim} GP(0, \sum_{l=1}^p \rho_l(\cdot) \Sigma_w, l) \), where \( \Sigma_w, l = a_l a_l^T \) and \( A = [a_1, a_2, ..., a_p] \). With such a specification, \( \Sigma_w, l \) are all singular matrices with rank one. However, we have \( \sum_{l=1}^p \Sigma_w, l = AA^T \), which we denote by \( \Sigma_w \), as a positive definite matrix. This means that, as in the intrinsic setting, we may take \( A \) to be lower-triangular, making it the Cholesky square root of \( \Sigma_w \).

Following Section 2.1, we can compute the general association structure for the \( Y' \)s under (2) and (3). For instance, we have \( \text{Cov}(Y(s,t),Y(s',t-1)) = x^T (s,t) \Sigma_{\bar{\beta}(s,t), \bar{\beta}(s',t-1)} x(s,t-1) \), where \( \Sigma_{\bar{\beta}(s,t), \bar{\beta}(s',t-1)} = \)
\( (t - 1) \left( \Sigma_{\eta} + \sum_{i=1}^{p} \beta_i (s - s'; \phi_i) a_i a_i^T \right) \). Furthermore, \( \text{var} \left( Y (s, t) \right) = x^T (s, t) \left[ \Sigma_{\eta} + \Sigma_{w} \right] x (s, t) \), which results in \( \text{corr} (Y (s, t), Y (s', t - 1)) \) to be \( O (1) \) as \( t \to \infty \).

A Bayesian hierarchical model for (2) and (3) may be completed by prior specifications as,

\[
\beta_0 \sim N (m_0, C_0) \text{ and } \beta (\cdot, 0) \equiv 0, \tag{4} \\
\Sigma_{\eta} \sim \text{IW} (a_{\eta}, B_{\eta}), \quad \Sigma_{w} \sim \text{IW} (a_{w}, B_{w}) \text{ and } \sigma^2 \sim \text{IG} (a_c, b_c), \\
m_0 \sim N (0, \Sigma_0); \quad \Sigma_0 = 10^5 \times I_p
\]

where \( B_{\eta} \) and \( B_{w} \) are \( p \times p \) precision (hyperparameter) matrices for the inverted Wishart distribution.

Consider now data, in the form \( Y (s_i, t_j) \) with \( i = 1, 2, ..., N_s \) and \( j = 1, 2, ..., N_t \). Let us collect, for each time point, the observations on all the sites. That is, we form, \( Y_t = (Y (s_1, t) , ..., Y (s_{N_s}, t))^T \) and the \( N_s \times N_t p \) block diagonal matrix \( F_t = (x^T (s_1, t), x^T (s_2, t), ..., x^T (s_{N_s}, t)) \) for \( t = t_1, ..., N_t \). Analogously we form the \( N_t \times 1 \) vector \( \theta_t = \left( \beta^T (s_1, t), ..., \beta^T (s_{N_s}, t) \right)^T \) and the \( N_s \times 1 \) measurement error vector \( \varepsilon_t \). Then we write the data equation for a dynamic spatial model as,

\[
Y_t = F_t \theta_t + \varepsilon_t; \quad t = 1, 2, ..., N_t; \quad \varepsilon_t \sim N (0, \sigma^2 I_{N_s}) \tag{5} \\
\theta_t = 1_{N_s} \otimes \beta_t + \beta^*_t,
\]

where \( \beta^*_t = (\beta (s_1, t), ..., \beta (s_{N_s}, t))^T \).

\[
\beta_t = \beta_{t-1} + \eta_t, \quad \eta_t \sim N_p (0, \Sigma_{\eta});
\]

and, with \( w_t = (w^T (s_1, t), ..., w^T (s_{N_s}, t))^T \),

\[
\beta^*_t = \beta^*_{t-1} + w_t, \quad w_t \sim N \left( 0, \sum_{l=1}^{p} (R_l (\phi_l) \otimes \Sigma_{w,l}) \right).
\]

With the prior specifications in (4), we can design a Gibbs sampler with Gaussian full conditionals for the temporal coefficients \( \{ \beta_t \} \), the spatiotemporal coefficients \( \{ \beta^*_t \} \), inverted Wishart for \( \Sigma_{\eta} \), and Metropolis steps for \( \phi \) and the elements of \( \Sigma_{w,l} \). A natural choice would seem to be a Metropolis update of \( \Sigma_{w} = \sum_{l=1}^{p} \Sigma_{w,l} \) using inverted-Wishart proposals. In practice, however, this does not seem to work well. A major difficulty is tuning the proposal to deliver a healthy acceptance rate. Wishart proposals are not tuned well.
by their scale parameter, which is a matrix, and are clumsy in moving around the parameter space, often leading to extremely autocorrelated chains. Gelfand et al. (2003b) have investigated this route and report similar problems with slow-moving chains. Consequently, they convert the joint distribution of \( w \) into a sequence of conditional distributions (also see, Royle and Berliner, 1996). This approach is very difficult to implement under (2) and (3).

Here, we outline a different strategy to update the chains. Our approach is to reparametrize the model in terms of the square root matrix, \( A \), of \( \Sigma_w \) and update the elements of the lower triangular matrix \( A \). To be precise, consider the full conditional distribution,

\[
 f (\Sigma_w | \gamma, \phi_1, \phi_2) \propto f (\Sigma_w | a_{\gamma}, B_{\gamma}) \times \frac{1}{\prod_{i=1}^{p} R_i (\phi_i) \otimes \Sigma_{w,i}} \exp \left( -\frac{1}{2} \beta^{T} \left( J^{-1} \otimes \left( \sum_{i=1}^{p} R_i (\phi_i) \otimes \Sigma_{w,i} \right)^{-1} \right) \beta^{T} \right) .
\]

The one to one relationship between elements of \( \Sigma_w \) and the Cholesky square-root \( A \) is well known (see, e.g., Harville, 1997, page 235). So, we reparametrize the above full conditional as,

\[
 f (A | \gamma, \phi_1, \phi_2) \propto f (h (A) | a_{\gamma}, B_{\gamma}) \frac{\partial h}{\partial a_{i,j}} \times \frac{1}{\prod_{i=1}^{p} R_i (\phi_i) \otimes (a_i a_i^T)} \exp \left( -\frac{1}{2} \beta^{T} \left( J^{-1} \otimes \left( \sum_{i=1}^{p} R_i (\phi_i) \otimes (a_i a_i^T) \right)^{-1} \right) \beta^{T} \right) .
\]

Here, \( h \) is the function taking the elements of \( A \), say \( a_{i,j} \), to those of the symmetric p.d. matrix \( \Sigma_w \). In the 2 \( \times \) 2 case we have,

\[
h (a_{11}, a_{21}, a_{22}) = (a_{11}^2, a_{11} a_{21}, a_{21}^2 + a_{22}^2) ,
\]

and the Jacobian is \( 4a_{11}^2 a_{22} \). Now, the elements of \( A \) are updated with univariate random-walk Metropolis proposals - log-normal (Gamma works well too) for \( a_{11} \) and \( a_{22} \), and normal for \( a_{21} \). Additional computational burden is wrought; now the likelihood needs to be computed for each of the three updates, but the chains are much better tuned (by controlling the scale of the univariate proposals) to move around the parameter space, thereby leading to better convergence behaviour.
3 Illustrating univariate spatiotemporal dynamic models

3.1 Simulation example - no covariates

We present a simple simulation case, generating a Gaussian random field with a covariance structure specified through the exponential family of covariance functions, \( \rho(\phi, d) = \sigma^2 \exp(-\phi d) \), where \( d \in \mathbb{R}^+ \). In particular, \( d \) becomes the Euclidean distance between pairs of sites. The field is generated on a randomly sampled set of \( N_t = 100 \) points within a 10x10 square. That is, the \( x \) and \( y \) coordinates lie between 0 and 10. From each of these sites we next draw \( N_t = 10 \) observations, pretending that they are observed at different time-points and are dynamically generated by the model in Section 2.1. For this example, in (2) and (3) we take \( p = 1 \), only an intercept term in \( \mu(s,t) \); i.e., \( \mu(s,t) = \beta_0(s,t) = \beta_0 + \beta_1(s,t) \) with \( \beta_0 = \beta_{0,t-1} + \eta_t \); \( \eta_t \sim \mathcal{N}(0, \sigma_n^2) \) and \( \beta_0(s,t) = \beta_0(s,t-1) + w(s,t) \); \( w(z,t) \sim \mathcal{GP}(0, \sigma_w^2 \rho(\phi, \cdot)) \). The true values for the simulation are, \( \beta_0 \equiv 10 \), (i.e. \( \beta_0 \sim \mathcal{N}(m_0, C_0) \) with \( m_0 = 10.0, C_0 = 0 \), \( \sigma_n^2 = \sigma_w^2 = \sigma_z^2 = 1.0; \phi = 0.50 \). The maximum distance among the generated sites was approximately 13km. \( \phi = 0.50 \) means the effective isotropic range is 6kms.

We analyse the data with a vague Normal prior on \( \beta_0 \), independent \( IG(2, 0.001) \) priors on \( \sigma_n^2, \sigma_w^2 \) and \( \sigma_z^2 \), and a \( G(0.1, 0.1) \) prior on \( \phi \). The time-varying intercepts, \( \{\beta_{0,t}\} \), along with the true values are displayed in Figure 1, while the remaining parameter estimates are presented in Table 1. The results are encouraging, with the true values being included in their respective 95% credible intervals despite the short length of the time series. So we turn to a “real data” example.

3.2 Modelling temperature given precipitation

Our spatial domain, shown in Figure 2 along with elevation (in 100m units) contours, provides a sample of 50 locations (indicated by “+”) in Boulder, Colorado. Each site provides information on monthly maximum temperature, and monthly mean precipitation. We denote the temperature summary in location \( s \) at time \( t \), by \( Y(s,t) \), and the precipitation by \( x(s,t) \). Forming a covariate vector \( \mathbf{x}^T(s,t) = (1, x(s,t)) \), we analyse the data using a coregionalized dynamic model, as outlined in Section 2.2. As a result, we have an intercept process \( \hat{\beta}_0(s,t) \) and a slope process \( \hat{\beta}_1(s,t) \), and the two processes are dependent.
We analysed the data with vague (independent) Normal priors on $\beta_0$, an $IW (2,0.01I_2)$ prior on the $2 \times 2$ $\Sigma_\eta$ and $\Sigma_{w_0}$ matrices, independent $IG (2,0.1)$ priors on $\sigma_\eta^2$, and independent $G (0.1,0.1)$ priors on $\phi_1$ and $\phi_2$. Three parallel MCMC chains were run for 5000 iterations each. Convergence was diagnosed using the CODA package in R, by monitoring mixing of the chains, the G-R statistic, auto-correlations and cross-correlations. Satisfactory convergence was diagnosed within 3000 iterations. A post burn-in sample of size 6000 (2000 from each chain) were used for the posterior summaries.

Figure 3 displays the time-varying intercepts and slopes (coefficient of precipitation). As expected, the intercept is higher in the summer months and lower in the winter months- highest in July, lowest in December. In fact, the gradual increase from January to July, and the subsequent decrease towards December is evident from the plot. Thus, seasonal pattern is retrieved although no such structure is imposed. Precipitation seems to have a negative impact on temperature, although this seems to be significant only in the months of January, March, May, June, November and December.

Table 2 displays the medians and credible intervals for variances and the correlations in the $\Sigma_\eta$ matrix. The corresponding results for the elements of $\Sigma_w$ are given in Table 3. A significant negative correlation is seen between the intercept and the slope processes. This is expected and justifies our use of dependent processes. Next, in Table 4, we provide the measurement error variances for temperature along with the estimates of the spatial correlation parameters for the intercept and slope process. Also presented are the ranges implied by $\phi_1$ and $\phi_2$ for the marginal intercept process, $w_1 (s)$, and the marginal slope process, $w_2 (s)$. The first range is computed by solving for the distance $d$, $\rho_1 (\phi_1, d) = 0.05$, while the second range is obtained by solving $(a^2_1 \rho_2 (\phi_2, d) + a^2_2 \rho_2 ((\phi_2, d)) / (a^2_1 + a^2_2) = 0.05$. The ranges are presented in units of 100-kms with the maximum observed distance between our sites being approximately 742 kms.

Finally, in Figures 4 and 5, we display the time-sliced image-contour plots for the intercept and slope processes. For both the processes, the spatial variation is better captured in the central and western edges of the domain. For the intercept process, all the months display similar spatial pattern, with denser contour variations towards the west than the east. For the slope process, the result is similar, but now the spatial pattern seems to be more pronounced in the months with more extreme weather - in the winter months of November through January, and the summer months of June through August.
4 Multivariate dynamic spatial models

This section looks at multivariate spatiotemporal data, with each pair \((s, t)\) yielding a vector of \(m\) observations, say \(Y(s, t) = (Y_1(s, t), ..., Y_m(s, t))^T\), and a set of \(p_k\) terms \((p_k - 1\) covariates, plus an intercept term\) corresponding to \(Y_k(s, t)\), forming the vector \(x_k(s, t)\), \(k = 1, 2, ..., m\). As in Section 2, \(s \in \{s_1, ..., s_N_s\}\) and \(t \in \{t_1, t_2, ..., t_N_t\}\) and the modelling proceeds analogous to the univariate case, with obvious changes in the dimensions. Thus the measurement equation assumes,

\[
Y(s, t) = \mu(s, t) + \epsilon(s, t) \quad \text{and} \quad \epsilon(s, t) \sim N(0, \Sigma_c); \quad \Sigma_c = \sigma^2 I_m
\]

\[
\mu(s, t) = X(s, t) \beta(s, t),
\]

where \(\beta(s, t) = (\beta_1^T(s, t), ..., \beta_m^T(s, t))^T\) is an \(\sum_{k=1}^m p_k \times 1\) vector, with each \(\beta_k(s, t)\) a \(p_k \times 1\) vector, and \(X(s, t) = \text{Diag}(x_1^T(s, t), ..., x_m^T(s, t))\), where each \(x_k(s, t)\) is a \(p_k \times 1\) vector, for \(k = 1, 2, ..., m\). Thus, \(X(s, t)\) is an \(m \times \sum_{k=1}^m p_k\) matrix. The transition equation is,

\[
\hat{\beta}(s, t) = \hat{\beta}(s, t) + \beta(s, t)
\]

\[
\beta_{t-1} = \beta_{t-1} + \eta_t \sim N_p(0, \Sigma_\eta).
\]

\[
\beta(s, t) = \beta(s, t - 1) + w(s, t),
\]

where \(\Sigma_\eta\) is a \(\sum_{k=1}^m p_k \times \sum_{k=1}^m p_k\) covariance matrix, \(w(s, t)\) is a \(\sum_{k=1}^m p_k\) dimensional spatial process.

Note that, under the above formulation, \(\beta(s, t) = (\beta_1^T(s, t), ..., \beta_m^T(s, t))^T\) is a \(\sum_{k=1}^m p_k\) dimensional spatial process. Recognizing the computational costs associated with the above generality, we work with a spatiotemporally varying intercept for each component of \(Y(s, t)\) and thus set all the terms in \(\beta_k(s, t)\) to 0, except the first term, which we label \(\beta_{0}^T(s, t)\), for \(k = 1, 2, ..., m\). Recalling that the first column of the covariate matrix corresponds to the intercept, and forming an \(m\) dimensional spatial process, \(\beta_0(s, t) = (\beta_{10}(s, t), ..., \beta_{m0}(s, t))^T\), we can now rewrite the multivariate model more simply as,

\[
Y(s, t) = X(s, t) \beta_t + \beta_0(s, t) + \epsilon(s, t) \quad \text{and} \quad \epsilon(s, t) \sim N(0, \Sigma_c).
\]

\[
\beta_t = \beta_{t-1} + \eta_t \sim N_p(0, \Sigma_\eta);
\]

\[
\beta_0(s, t) = \beta_0(s, t - 1) + w_0(s, t),
\]

where \(\Sigma_c = \text{Diag}(\sigma^2 I_k)\) is a \(m\) by \(m\) covariance matrix, \(w_0(s, t) = A_0 v_0(s, t)\) is an \(m\)-dimensional spatial process, with \(A_0\), an \(m \times m\) matrix, and \(v_0(s, t) = (v_{10}(s, t), ..., v_{m0}(s, t))^T\) is an
\( m \times 1 \) vector, with each \( v_0(\cdot, t) \overset{\text{iid}}{\sim} GP(0, \rho_1(\phi_t)) \). Writing \( \Sigma_{\mathbf{w}_0} = A_0 A_0^T \), the hierarchy is completed by assigning the priors,

\[
\begin{align*}
\beta_{t=0} &\sim N(\mathbf{m}_0, C_0) \quad \text{and} \quad \beta_0(\cdot, 0) \equiv 0. \\
\Sigma_{\eta} &\sim IW(\alpha_{\eta}, B_{\eta}) \quad \Sigma_{\mathbf{w}_\gamma} \sim IW(\alpha_{\gamma}, B_{\gamma}) \quad \text{and} \quad \sigma_{\mathbf{k}_k}^2 \sim IG(a_k, b_k), \quad k = 1, 2, ..., m.
\end{align*}
\]  

(6)

Note that here, for instance, \( \Sigma_{Y[i, t], Y[j, t-1]} = (t - 1) \left\{ \sum_{i=1}^{m} \rho_1(s - s', \phi_t) \mathbf{a}_0 \mathbf{a}_0^T + X(s, t) \Sigma_{\eta} X^T(s', t-1) \right\} \), where \( \mathbf{a}_0 \) is the \( l \)th column of \( A_0 \). As a result, nonstationarity enters only through the dynamic updating of \( \beta_t \). One might want a richer cross-covariance structure. Allowing \( A_0 \) to vary with \( s \) would increase the flexibility of the class of models.

Consider now data, in the form \((Y(s_t, t_j))\) with \( i = 1, 2, ..., N_i \) and \( j = 1, 2, ..., N_t \). Let us collect, for each time point, the observations on all the sites. That is, we form, \( Y_t = \left(Y^T(s_1, t), ..., Y^T(s_{N_i}, t)\right)^T \) and the \( mN \times p \) matrix \( X_t = \left(X^T(s_1, t), X^T(s_2, t), ..., X^T(s_{N_i}, t)\right)^T \) for \( t = t_1, ..., t_{N_t} \). Analogously, we form \( \beta_{0t} = \left(\beta_0^T(s_1, t), ..., \beta_0^T(s_{N_i}, t)\right)^T \) and \( \epsilon_t \). Then, we write down the measurement equation at time \( t \) as,

\[ Y_t = X_t \beta_t + \beta_{0t} + \epsilon_t; \quad t = 1, 2, ..., N_t; \quad \epsilon_t \overset{\text{iid}}{\sim} N(0, I_{N_i} \otimes \Sigma_{\epsilon}). \]  

(7)

The transition equations for \( \beta_t \) and \( \beta_{0t} \) are,

\[
\begin{align*}
\beta_t &= \beta_{t-1} + \eta_t, \quad \eta_t \overset{\text{iid}}{\sim} N_p(0, \Sigma_{\eta}); \quad \beta_0 \overset{\text{iid}}{\sim} N_m(\mathbf{m}_0, C_0) \\
\beta_{0t} &= \beta_{0, t-1} + \mathbf{w}_0(t); \quad \mathbf{w}_0, t \overset{\text{iid}}{\sim} N_m \left(0, \sum_{k=1}^{m} R_k(\phi_k) \otimes \Sigma_{\mathbf{w}_0,k} \right).
\end{align*}
\]

with \( \Sigma_{\mathbf{w}_0} = \sum_{k=1}^{m} \Sigma_{\mathbf{w}_0,k} \), exactly as in Section 2.2.

Updating presents no new problems compared with the univariate setting, except that we now have to deal with higher dimensional objects, thereby incurring greater computational burden. Again, a Gibbs sampler is easily designed for the above model, resulting in Gaussian kernels for \( \mathbf{m}_0, \beta_t \)'s and \( \beta_{0t} \)'s, inverted Wishart for \( \Sigma_{\eta} \) (we take \( C_0 = 0 \)), and finally, Metropolis steps for the \( \phi_k \)'s and the elements of \( \Sigma_{\mathbf{w}_0} \) - the latter being transformed into the square root space of \( A_0 \), as in Section 2.2.
5 Illustrating the multivariate models

5.1 A simulation example: "intercept only" model

We illustrate our multivariate model fitting with a simulation example. We generate a sample of \( N_s = 30 \) sites on a 10x10 grid. Thus, the \( x \) and \( y \) coordinates lie between 0 and 10. From each of these sites we drew \( N_t = 8 \) bivariate observations \( Y(s,t) \) with the following specifications,

\[
\mathbf{m}_0 = (5.0, 10.0)^T ; \quad (C_0)_{ij} = 0; \quad \sigma_{1e}^2 = 0.25; \quad \sigma_{2e}^2 = 0.75, \\
\Sigma_{\eta} = \begin{pmatrix} 1.0 & 0 \\ 0 & 2.0 \end{pmatrix}; \quad \Sigma_{w_0} = \begin{pmatrix} 1.0 & 0.70 \\ 0.70 & 1.0 \end{pmatrix};
\]

\( \phi_1 = 0.50; \phi_2 = 1.50. \rho_l(\phi_l; d) = \exp(-\phi_l d), \ l = 1, 2. \)

The maximum distance among the generated sites was approximately 13kms. \( \phi_1 \) implies a range of 6kms, while \( \phi_2 \) implies a range of 4.7 kms.

We analysed this data using our multivariate dynamic model with a flat prior on \( \mathbf{m}_0 \), independent inverted Wishart priors on \( \Sigma_{\eta} \) and \( \Sigma_{w_0} \), independent \( IG(2, 0.01) \) priors on \( \sigma_{k_\infty}^2 \), and \( G(0.1, 0.1) \) priors on \( \phi_k \)'s, \( k = 1, 2 \). Figure 6 shows the time-varying parameters for the two variables. Tables 5 through 7 present the estimated values of the various variance and dependence parameters. Given the small sample size, we see that the estimates are all reasonably close to the true values that generated the data. In fact, the true values are all included in the reported credible intervals and, therefore, we again turn to the Colorado data set for a real illustration.

5.2 Colorado - elevation data

Here we present an analysis of sampled temperature data (maximum monthly temperature) and sampled precipitation data (maximum monthly precipitation), collected over 50 sites across Colorado, from January through December in 1997. Elevation information is available (as a covariate) for each of the 50 sites. We consider the intercept and elevation (in kms) as the only components in the regression.
We present analysis of the data with vague (independent) Normal priors on $\beta_0$, an $IW (2, 0.01I_4)$ prior on the $4 \times 4 \Sigma_y$ matrix, an $IW (2, 0.01I_2)$ prior on $\Sigma_w$, independent $IG (2, 0.1)$ priors on $\sigma^2$, and independent $G (0.1, 0.1)$ priors on $\phi_1$ and $\phi_2$. Three parallel MCMC chains were run for 5000 iterations each. Convergence was diagnosed using methods similar to those in Section 3.

Figures 7 and 8 plot the time-varying parameters for temperature and precipitation respectively. For temperature, the intercept is higher in the summer months and lower in the winter months - highest in July, lowest in December, as in Section 3.2. As expected, elevation has a negative impact on temperature - higher altitudes result in lower temperatures. This effect is more pronounced in the summer months (most significant in July) and more moderate in the winters. The relationship between precipitation and elevation reveals a slight negative effect in the summer months, but a positive effect in the winter months.

Table 8 displays the credible intervals for the variances and correlations in the $\Sigma_y$ matrix. The corresponding results for the elements of $\Sigma_w$ are given in Table 9. Note the significant negative correlation between the temperature process and the precipitation process, which seems to be consistent with our earlier finding (in Section 3.2) that precipitation tended to have a negative regression coefficient with temperature. Finally, in Table 10, we provide the two measurement error variances for temperature and precipitation and the parameter estimates of the respective correlation functions, together with the effective ranges they imply, expressed in units of 100-kms. We see a larger range for the temperature process suggesting a slightly stronger spatial dependence than for precipitation.

Figures 9 and 10 display the time-sliced image plots (with overlaid contours) of the posterior means of the spatial residual surfaces for temperature and precipitation respectively. Here also the spatial story seems to be more pronounced for temperature than for precipitation. For the former, the shades are more starkly different for the different months, especially between the summer and winter months, revealing a seasonal effect for the pattern of spatial variation. Such an effect is not as prominent for precipitation.

6 Conclusions and extensions

We have shown how to accommodate both univariate and multivariate dynamic spatial data within a hierarchical modelling framework. The novel feature of the univariate modelling is the incorporation of spa-
tiotemporally varying intercept and slope coefficients in the transition stage of the model. A bivariate spatiotemporal process is required and is handled through a coregionalization approach. The extension of the modelling for the slope from a constant $\beta_t$ regardless of $s$ to an entire process $\beta_t(s,t)$ for each $t$ considerably increases the flexibility of the specification.

The multivariate case also benefits from the use of coregionalization. For the precipitation and temperature data we expect spatial and temporal dependence not only within the precipitation measurements, not only within the temperature measurements, but also between precipitation and temperature measurements. We would also expect different spatial ranges for each of the associated processes. Modelling through coregionalization enables this.

Spatial prediction at any time point is straightforward. For instance, in the univariate case, the predictive distribution for $Y(s_0,t)$ at a new location $s_0$ with covariate vector $x(s_0,t)$ can be simulated through (2) and (3). A posterior draw from $\beta(s_0,t)$, along with $x(s_0,t)$ provides a posterior draw of $\mu(s_0,t)$. Adding Gaussian random error with variance which is a posterior draw of $\sigma^2$, yields a value of $Y(s_0,t)$ from its predictive distribution. We have employed such simulation extensively to obtain a grid of spatial predictions. Such grids have been used to create the image-contour plots in Figures 4, 5, 9 and 10.

Temporal prediction can also be handled, although the variance calculations in Section 2.2 show that it will be explosive. For instance, the transition equations in (3) enable prediction of $\beta(s,T+1)$. With $x(s,T+1)$ we obtain a prediction of $\mu(s,T+1)$. Again, with a Gaussian random error term having a variance which is a posterior draw for $\sigma^2$, we simulate a predictive realization for $Y(s,T+1)$. Prediction of $Y(s,T+1)$ can be performed by combining the above evolution result for $\beta(s,t)$. Similar calculations lead to prediction $h$ time steps ahead (of $Y(s,T+h)$) by repeated use of evolution equation for $\beta(s,t)$.

If there is concern regarding Gaussian assumption for $Y(s,t)$ we can employ suitable Box-Cox transformation. If the first stage model is clearly not Gaussian, e.g., a count or a binary outcome, we could introduce a generalized linear model structure in place of, say, (2). Now, $\mu(s,t)$ would be the mean of $Y(s,t)$ on a suitably transformed scale and $\epsilon(s,t)$ would disappear. However, computation will become more demanding since now neither the temporal coefficients $\beta_t$ nor the spatiotemporal coefficients $\beta_{t}^s$ will have Gaussian full conditional distributions. Adaptive rejection sampling or Metropolis steps need to be incorporated to update these parameters.
Although the methods discussed here (especially the multivariate methods) may not be immediately accessible in standard software, the likelihood is easily cast into a general object-oriented template, which renders itself to easy coding in object-oriented languages such as C++ or Java. Object-oriented libraries are particularly efficient in carrying out the matrix decompositions and determinant computations involved in model fitting. Nevertheless, computing time becomes prohibitive for very large data sets, and one has to resort to more sophisticated parallel algorithms.

Acknowledgements

The research of the first author was supported in part by a grant from the Office of the Vice President of the University of Minnesota. The research of the second author was supported in part by a grant from CNPq and also by a grant from the Research Foundation at the University of Connecticut. The research of the third author was supported in part by NSF DMS 9971206 and by NIH NIEHS 5 R01-E50700. The seminal development of the work took place when both the second and third authors were at the University of Connecticut, the former as a Distinguished Visiting Research Professor. The authors thank Doug Nychka for providing the data set and for several helpful comments.

REFERENCES


Table 1: Parameter estimates for the simulation example in Section 3.1.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>median (2.5%, 97.5%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2_\eta$</td>
<td>0.777 (0.327, 2.351)</td>
</tr>
<tr>
<td>$\sigma^2_w$</td>
<td>0.987 (0.823, 1.175)</td>
</tr>
<tr>
<td>$\sigma^2_\varepsilon$</td>
<td>0.976 (0.859, 1.129)</td>
</tr>
<tr>
<td>$\phi$</td>
<td>0.526 (0.336, 1.076)</td>
</tr>
</tbody>
</table>

Table 2: Estimates for the variances and the correlation from $\Sigma_\eta$ in Section 3.2.

<table>
<thead>
<tr>
<th>$\Sigma_\eta$</th>
<th>median (2.5%, 97.5%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Sigma_\eta [1,1]$</td>
<td>0.296 (0.130, 0.621)</td>
</tr>
<tr>
<td>$\Sigma_\eta [2,2]$</td>
<td>0.786 (0.198, 1.952)</td>
</tr>
<tr>
<td>$\Sigma_\eta [1,2]/\sqrt{\Sigma_\eta [1,1] \Sigma_\eta [2,2]}$</td>
<td>-0.562 (-0.807, -0.137)</td>
</tr>
</tbody>
</table>

Table 3: Estimates for the variances and correlation from $\Sigma_w$ in Section 3.2.

<table>
<thead>
<tr>
<th>$\Sigma_w$</th>
<th>median (2.5%, 97.5%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Sigma_w [1,1]$</td>
<td>0.017 (0.016, 0.019)</td>
</tr>
<tr>
<td>$\Sigma_w [2,2]$</td>
<td>0.026 (0.0065, 0.108)</td>
</tr>
<tr>
<td>$\Sigma_w [1,2]/\sqrt{\Sigma_w [1,1] \Sigma_w [2,2]}$</td>
<td>-0.704 (-0.843, -0.545)</td>
</tr>
</tbody>
</table>

Table 4: Nugget effects and spatial correlation parameters for example in Section 3.2.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>median (2.5%, 97.5%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2_\varepsilon$</td>
<td>0.134 (0.106, 0.185)</td>
</tr>
<tr>
<td>$\phi_1$</td>
<td>1.09 (0.58, 2.04)</td>
</tr>
<tr>
<td>$\phi_2$</td>
<td>0.58 (0.37, 1.97)</td>
</tr>
<tr>
<td>Range for intercept process</td>
<td>2.75 (1.47, 5.17)</td>
</tr>
<tr>
<td>Range for slope process</td>
<td>4.68 (1.60, 6.21)</td>
</tr>
</tbody>
</table>

20
Table 5: Estimates for the variances and correlation from $\Sigma_\eta$ in Section 5.1.

<table>
<thead>
<tr>
<th>$\Sigma_\eta$</th>
<th>median (25%, 97.5%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Sigma_\eta[1,1]$</td>
<td>0.656 (0.245, 2.150)</td>
</tr>
<tr>
<td>$\Sigma_\eta[2,2]$</td>
<td>1.309 (0.488, 3.537)</td>
</tr>
<tr>
<td>$\Sigma_\eta[1,2] / \sqrt{\Sigma_\eta[1,1] \Sigma_\eta[2,2]}$</td>
<td>0.398 (-0.331, 0.825)</td>
</tr>
</tbody>
</table>

Table 6: Estimates for the variances and correlation from $\Sigma_w$ in Section 5.1

<table>
<thead>
<tr>
<th>$\Sigma_w$</th>
<th>median (25%, 97.5%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Sigma_w[1,1]$</td>
<td>0.838 (0.632, 1.121)</td>
</tr>
<tr>
<td>$\Sigma_w[2,2]$</td>
<td>1.179 (0.876, 1.605)</td>
</tr>
<tr>
<td>$\Sigma_w[1,2] / \sqrt{\Sigma_w[1,1] \Sigma_w[2,2]}$</td>
<td>0.752 (0.625, 0.833)</td>
</tr>
</tbody>
</table>

Table 7: Nugget effects and spatial correlation parameters for example in Section 5.1.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>median (25%, 97.5%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2_\ell$</td>
<td>0.375 (0.244, 0.566)</td>
</tr>
<tr>
<td>$\sigma^2_\phi$</td>
<td>0.485 (0.280, 0.731)</td>
</tr>
<tr>
<td>$\phi_1$</td>
<td>0.48 (0.27, 1.76)</td>
</tr>
<tr>
<td>$\phi_2$</td>
<td>1.58 (0.84, 2.64)</td>
</tr>
</tbody>
</table>

Table 8: Estimates of the variances and correlations from $\Sigma_\eta$ in Section 5.2

<table>
<thead>
<tr>
<th>$\Sigma_\eta$</th>
<th>median (25%, 97.5%)</th>
<th>$\Sigma_\eta$</th>
<th>median (25%, 97.5%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Sigma_\eta[1,1]$</td>
<td>0.482 (0.215, 1.045)</td>
<td>$\Sigma_\eta[3,3]$</td>
<td>0.295 (0.029, 1.503)</td>
</tr>
<tr>
<td>$\Sigma_\eta[2,2]$</td>
<td>0.014 (0.006, 0.029)</td>
<td>$\Sigma_\eta[4,4]$</td>
<td>2.378 (1.012, 7.020)</td>
</tr>
<tr>
<td>$\Sigma_\eta[1,2] / \sqrt{\Sigma_\eta[1,1] \Sigma_\eta[2,2]}$</td>
<td>-0.492 (-0.830, -0.049)</td>
<td>$\Sigma_\eta[2,3] / \sqrt{\Sigma_\eta[2,2] \Sigma_\eta[3,3]}$</td>
<td>-0.560 (-0.910, 0.624)</td>
</tr>
<tr>
<td>$\Sigma_\eta[1,3] / \sqrt{\Sigma_\eta[1,1] \Sigma_\eta[3,3]}$</td>
<td>-0.502 (-0.773, 0.048)</td>
<td>$\Sigma_\eta[2,4] / \sqrt{\Sigma_\eta[2,2] \Sigma_\eta[4,4]}$</td>
<td>0.229 (-0.434 0.619)</td>
</tr>
<tr>
<td>$\Sigma_\eta[1,4] / \sqrt{\Sigma_\eta[1,1] \Sigma_\eta[4,4]}$</td>
<td>-0.220 (-0.651, 0.537)</td>
<td>$\Sigma_\eta[3,4] / \sqrt{\Sigma_\eta[3,3] \Sigma_\eta[4,4]}$</td>
<td>-0.651 (-0.797, -0.219)</td>
</tr>
</tbody>
</table>

Table 9: Estimates of the variances and correlations of $\Sigma_w$ from Section 5.2

<table>
<thead>
<tr>
<th>$\Sigma_w$</th>
<th>median (25%, 97.5%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Sigma_w[1,1]$</td>
<td>0.012 (0.011, 0.014)</td>
</tr>
<tr>
<td>$\Sigma_w[2,2]$</td>
<td>2.239 (1.638, 3.007)</td>
</tr>
<tr>
<td>$\Sigma_w[1,2] / \sqrt{\Sigma_w[1,1] \Sigma_w[2,2]}$</td>
<td>-0.164 (-0.303, -0.036)</td>
</tr>
</tbody>
</table>
Table 10: Estimates of the nugget effects and spatial correlation parameters from Section 5.2

<table>
<thead>
<tr>
<th>Parameters</th>
<th>median (2.5%, 97.5%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2_{n}$</td>
<td>0.0026 (0.0017, 0.0043)</td>
</tr>
<tr>
<td>$\sigma^2_{e}$</td>
<td>6.924 (5.951, 8.274)</td>
</tr>
<tr>
<td>$\phi_1$</td>
<td>0.77 (0.36, 2.47)</td>
</tr>
<tr>
<td>$\phi_2$</td>
<td>1.56 (0.41, 3.29)</td>
</tr>
<tr>
<td>Range (temperature)</td>
<td>3.89 (1.22, 8.33)</td>
</tr>
<tr>
<td>Range (precipitation)</td>
<td>1.97 (0.91, 7.32)</td>
</tr>
</tbody>
</table>
Figure 1: Posterior distributions for the time-varying parameters in the univariate simulation example in Section 3.1. The solid line represents the medians, the short-dashed lines are the upper and lower credible intervals, and the long-dashed line shows the true values.
Figure 2: Map of the region in Boulder, CO, that forms our spatial domain. The data for the illustrations come from 50 locations, marked by “+”, in this region.
Figure 3: Posterior distributions for the time-varying parameters in the temperature given precipitation example in Section 3.2. The top graph corresponds to the intercept, while the lower one is the coefficient of precipitation. Solid lines represent the medians and the dashed lines correspond to the upper and lower credible intervals.
Figure 4: Time-sliced image-contour plots displaying the posterior mean surface of the spatial residuals corresponding to the intercept process in the temperature given precipitation model in Section 3.2.
Figure 5: Time-sliced image-contour plots displaying the posterior mean surface of the spatial residuals corresponding to the slope process in the temperature given precipitation model in Section 3.2.
Figure 6: Posterior distributions for the two sets of time-varying parameters in the multivariate simulation example in Section 5.1. The solid line represents the medians, the short-dashed lines are the upper and lower credible intervals, and the long-dashed line shows the true values.
Figure 7: Posterior distributions for the time-varying parameters for temperature in the multivariate example in Section 5.2. The top graph corresponds to the intercept, while the lower one is the coefficient of elevation. Solid lines represent the medians and the dashed lines correspond to the upper and lower credible intervals.
Figure 8: Posterior distributions for the time-varying parameters for precipitation in the multivariate example in Section 5.2. The top graph corresponds to the intercept, while the lower one is the coefficient of elevation. Solid lines represent the medians and the dashed lines correspond to the upper and lower credible intervals.
Figure 9: Time-sliced image-contour maps showing the posterior mean temperature residuals corresponding to the analysis in Section 5.2.
Figure 10. Time-sliced image-contour plots displaying the posterior mean surface of the precipitation residuals for the analysis in Section 5.2.