A Reparametrization Approach for Dynamic Space-Time Models

Hyeyoung Lee* and Sujit K. Ghosh*

Institute of Statistics Mimeo Series #2587

Abstract

Researchers in diverse areas such as environmental and health sciences are increasingly working with data collected across space and time. The space-time processes that are generally used in practice are often complicated in the sense that the auto-dependence structure across space and time is non-trivial, often non-separable and non-stationary in space and time. Moreover, the dimension of such data sets across both space and time can be very large leading to computational difficulties due to numerical instabilities. Hence, space-time modeling is a challenging task and in particular parameter estimation based on complex models can be problematic due to the such curse of dimensionality. We propose a novel reparametrization approach to fit dynamic space-time models which allows the use of a very general form for the spatial covariance function. Our modeling contribution is to present an unconstrained reparametrization for a covariance function within the dynamic space-time models. A major benefit of the proposed unconstrained reparametrization method is that we are able to implement the modeling of a very high dimensional covariance matrix that automatically maintains the positive definiteness constraint. We demonstrate the applicability of our proposed reparametrized dynamic space-time models for a large data set of total nitrate concentrations.
Keywords: Computational efficiency, Dynamic models, Reparametrization, Spatial models.

*Department of Statistics, North Carolina State University, Raleigh, NC 27695, U.S.A.

1 Introduction

Researchers in diverse areas such as environmental and health sciences are increasingly faced with working with data that are observed over space and time. In recent years, there has been widespread attention in the statistical literature given to modeling space-time data (Gelfand et al., 1998; Kyriakidis and Journel, 1999; Cressie and Huang, 1999; ; Gneiting, 2002; Stein, 2003; Huang and Hsu, 2004).

Dynamic linear models (DLM), or state-space models have been widely used to analyze multidimensional time series such as spatial time series. Many researchers have employed a DLM framework in environmental problems, in particular, temporally rich in data. A Bayesian approach has been commonly taken in the DLM framework for analyzing space-time data. Examples include Tonellato (1997), Sansó and Guenni (1999), Wikle et al. (1999), Stroud et al. (2001), Huerta et al. (2004), and Banerjee et al. (2005) among others. Huang and Cressie (1996) and Mardia et al. (1998) presented a non-Bayesian approaches as well.

In this article, we take a Bayesian approach and use Markov chain Monte Carlo (MCMC) simulation to obtain the inference based on the posterior distribution of the parameters. Recent developments in MCMC computing allow fully Bayesian analyses of complex multilevel models for dynamic space-time data. Pole et al. (1994) and West and Harrison (1997) are good references on the DLM from a Bayesian point of view.

In general, space-time modeling is a challenging task which requires the manipulation of large data sets and considers both spatial and temporal correlations. Moreover, space-time processes are often complicated in that the dependence structure across space and time is non-trivial, often non-separable and non-stationary in space and/or time. Several modeling strategies have been proposed to address this problem. Wikle and Cressie (1999) developed a space-time Kalman-filter that achieves dimension reduction by decomposing the state-process into sets of basis functions and time series. To avoid a high computational load they used an empirical Bayesian method for the estimation of model parameters rather than a fully Bayesian hierarchical approach. However, the additional variability in estimating parameters is ignored in such a method. Stroud et al. (2001) specified very simple random walk dynamics, while Xu and Wikle (2005) discussed efficient estimation approaches via the expectation-maximization (EM) algorithm for the parameter and covariance matrices with high dimensionality in dynamic space-time

models. Thus, it appears that most of the methods available currently in the literature are limited either by restrictive assumptions on the covariance functions (e.g., stationary, isotropic, separable etc.) or if such restrictive assumptions are not made the methods suffer from curse of dimensionality due high dimensional matrix inversions. We propose a method that avoids such curse of dimensionality without making strong restrictive assumptions.

This article is organized as follows. First, we present a brief overview of the dynamic space-time models in Section 2.1 and related inference based on standard method in Section 2.2/ In Section 2.3, we propose a reparametrization approach to fit dynamic space-time models that are based on a very general class of spatial covariance function. These models are not limited by the stationarity and isotropy restrictions for the covariance function. One of the main contribution of this article is to present an unconstrained reparametrization method to be used for a covariance function within dynamic space-time modeling framework. Using this unconstrained reparametrization method, we are able to implement the models to fit high-dimensional data without many restrictive assumptions. We demonstrate the applicability of our proposed reparametrized dynamic space-time models for a large data set of total nitrate concentrations in Section 3. Section 4 presents some general discussions and scope for future research.

2 Dynamic Space-Time Models (DSTM)

Dynamic linear model (West and Harrison, 1997) is probably one of the most widely known and used subclass in dynamic models. The term *dynamic* is related to changes in time series processes due to the passage of time. DLM can be seen as a generalization of traditional regression models that allows changes in parameter values over time. DLM provide a very flexible framework that permits smooth and abrupt changes in the time series generating process. We adapt a DLM framework to space-time data and call the model as dynamic space-time model (DSTM). A DSTM framework views the data as arising from spatial time series.

2.1 Overview

Suppose the process $\{Z(s,t); s \in \mathbb{R}^2, t \in [0,\infty)\}$ is observed on a finite number of sites labeled as $s_1, ..., s_n$ at each time t, where $t = t_1, t_2, \cdots, t_m$.

Assuming that t_j's are observed over equal lags, we wil relabel $t_j = j$ for $j = 1, \ldots, m$. Consider the $n \times 1$ vector time series $\boldsymbol{Z}_t = (Z(s_1, t), ..., Z(s_n, t))^T$ at time t. For each t, DSTM is usually formed by an observation equation and an evolution equation. An observation equation describes the relationship between the observation ($\boldsymbol{Z}_t$) and the regressors (X_t) that takes the form of a multivariate regression process,

$$\boldsymbol{Z}_t = X_t \boldsymbol{\beta}_t + \boldsymbol{\nu}_t, \quad \boldsymbol{\nu}_t \sim N_n(\mathbf{0}, \Sigma_t^{\nu}) \tag{1}$$

where X_t is an $n \times p$ observed design matrix and $\boldsymbol{\beta}_t$ is a $p \times 1$ vector of regression coefficients or state parameters. An evolution equation describes the dynamics of the vector of regression coefficients or state parameters $\boldsymbol{\beta}_t$ through time,

$$\boldsymbol{\beta}_t = G_t \boldsymbol{\beta}_{t-1} + \boldsymbol{\omega}_t, \quad \boldsymbol{\omega}_t \sim N_p(\mathbf{0}, \Sigma_t^{\omega}) \tag{2}$$

where G_t is a $p \times p$ evolution matrix. There are several ways to model the G_t's. The most common assumption is that the G_t's are structurally known, possibly up to some finite number of parameters. In this article, we do not make any structural assumption about the G_t's, but we assume that $G_t = G$ for all t and that G follows a matrix-valued normal distribution with mean G_0 and variance-covariance parameters Ω_0 and Σ_0^G. That is, $G \sim MN_{p \times p}(G_0, \Omega_0, \Sigma_0^G)$. We also assume that the $\boldsymbol{\nu}_t$ and $\boldsymbol{\omega}_t$ error vectors are independent and have multivariate normal distributions with mean $\mathbf{0}$ and variance-covariance matrices Σ_t^{ν} and Σ_t^{ω}, respectively. The model is completed with a normal prior for the initial state, $\boldsymbol{\beta}_1 \sim N(\boldsymbol{\beta}_0, \Sigma_0^{\omega})$, where $\boldsymbol{\beta}_0$ is assumed known.

Equivalently, the model can be written using hierarchical specifications as follows:

$$\begin{aligned}
\boldsymbol{Z}_t | \boldsymbol{\beta}_t &\sim N_n(X_t \boldsymbol{\beta}_t, \Sigma_t^{\nu}), \\
\boldsymbol{\beta}_t | \boldsymbol{\beta}_{t-1}, G &\sim N_p(G \boldsymbol{\beta}_{t-1}, \Sigma_t^{\omega}), \\
G &\sim MN_{p \times p}(G_0, \Omega_0, \Sigma_0^G), \\
\boldsymbol{\beta}_1 &\sim N(\boldsymbol{\beta}_0, \Sigma_0^{\omega}),
\end{aligned}$$

where the matrix valued normal distribution $MN_{p \times p}(G_0, \Omega_0, \Sigma_0^G)$ has the probability density function given by

$$p(G) = (2\pi)^{-p^2/2} |\Omega_0|^{-p/2} |\Sigma_0^G|^{-p/2} \exp\left(-\frac{1}{2} \mathrm{tr}\left[{\Omega_0}^{-1} (G - G_0) {\Sigma_0^G}^{-1} (G - G_0)^T \right] \right).$$

This Bayesian hierarchical approach not only helps organize our thinking about the model, but also fully accounts for all sources of uncertainty without making substantial structural assumptions such as spatial stationarity, isotropy, etc.

In order to keep our illustrations simple, first we consider the following simplified DSTM. Specifically, first we assume that $\Sigma_t^\nu = \Sigma^\nu$ and $\Sigma_t^\omega = \Sigma^\omega$ in equations (1) and (2). That is, the conditional variance-covariance matrices Σ^ν and Σ^ω do not change over time (and hence are static). Later we discuss how to relax these assumptions. However notice that the marginal variances, $Var[\boldsymbol{Z}_t]$'s do change over time. Specifically,

$$\begin{aligned} Var[\boldsymbol{Z}_t] &= X_t Var[\boldsymbol{\beta}_t] X_t^T + \Sigma^\nu \\ Var[\boldsymbol{\beta}_t] &= G Var[\boldsymbol{\beta}_{t-1}] G^T + \Sigma^\omega. \end{aligned}$$

Using the precceding recursive relations we can easily see that $Var[\boldsymbol{Z}_t]$ does evolve over time even when Σ_t^μ is assumed to be contant over time. Then our simplified DSTM can be written as,

$$\boldsymbol{Z}_t = X_t \boldsymbol{\beta}_t + \boldsymbol{\nu}_t, \;\; \boldsymbol{\nu}_t \sim N_n(\mathbf{0}, \Sigma^\nu), \tag{3}$$

$$\boldsymbol{\beta}_t = G\boldsymbol{\beta}_{t-1} + \boldsymbol{\omega}_t, \;\; \boldsymbol{\omega}_t \sim N_p(\mathbf{0}, \Sigma^\omega), \tag{4}$$

and $\boldsymbol{\beta}_1 \sim N(\boldsymbol{\beta}_0, \Sigma_0^\omega)$. Here, Σ^ν and Σ^ω are unstructured variance-covariance matrices and for the ease of exposition, we initially assume that $\boldsymbol{Z}_t$ and $\boldsymbol{X}_t$ are observed at all time points t. Later we discuss how to relax this assumption if some observations are missing.

The Bayesian model is completed with the specification of a prior distribution for the parameters. These include the data-model covariance matrix Σ^ν and the error covariance matrix Σ^ω. Prior specifications for Σ^ν and Σ^ω can be tricky as these matrices are usually high-dimensional (e.g., Σ^ν is $n \times n$) and they need to be positive definite (pd). It is customary to use the inverse Wishart distributions to model such covariance matrices. Note that our models require no restrictive assumptions such as stationarity and isotropy as Σ^ν is left unstructured. However, if such assumptions are deemed necessary, we can easily incorporate them in our modeling framework.

2.2 Updating Equations

Suppose that the time at origin t=0 represents the current time, and that existing information available is denoted by the D_0, initial information set.

Generally, we have the information set at time t, D_t. As time evolves, so does the information. Observing the values of $\boldsymbol{Z}_t$ at time t implies that D_t includes both the previous information set D_{t-1} and the observations $\boldsymbol{Z}_t$, that is, $D_t = \{\boldsymbol{Z}_t, D_{t-1}\}$.

Traditionally, inference for dynamic models is made sequentially by obtaining the prior predictive and updated distributions for the state parameters $\boldsymbol{\beta}_t$ for each time t. The prior predictive distributions are respectively obtained by

$$\begin{aligned} p(\boldsymbol{\beta}_t|D_{t-1}) &= \int p(\boldsymbol{\beta}_t|\boldsymbol{\beta}_{t-1})p(\boldsymbol{\beta}_{t-1}|D_{t-1})d\boldsymbol{\beta}_{t-1} \\ p(\boldsymbol{Z}_t|D_{t-1}) &= \int p(\boldsymbol{Z}_t|\boldsymbol{\beta}_t)p(\boldsymbol{\beta}_t|D_{t-1})d\boldsymbol{\beta}_t \end{aligned}$$

and the updated distribution is obtained by Bayes' theorem as

$$p(\boldsymbol{\beta}_t|D_t) \propto p(\boldsymbol{Z}_t|\boldsymbol{\beta}_t)p(\boldsymbol{\beta}_t|D_{t-1})$$

where D_t represents all the available information upto time t.

The above updating scheme in the dynamic space-time model may not be easy to implement when the G matrix is completely unknown and the matrices Σ^ν and Σ^ω are completely unstructured, even if we are abale to derive the analytical form of the full conditional distributions required for the above updating equations. The main limitation of these types of multivariate updating schemes are the problems associated with high dimensional matrix inversions. Moreover these type of multivariate updating schemes are not generally applicable when some observations are missing or censored, as it is generally very hard to sample from truncated multivariate distributions. In order to avoid such numerical instabilities and to accelerate model fitting, we propose a method via reparametrization of the model which leads to an univariate scheme for the aforementioned DSTM. As by-products of using this reparametrization method, we show that it is possible to obtain several extensions of the usual DSTM (e.g., relaxing the Gaussian assumption, allowing nonstationary and nonsperable covariances etc.).

2.3 A Reparametrization Method

In this section, we describe a reparametrization method for the covariance matrices Σ^ν and Σ^ω. Modeling a covariance matrix Σ is difficult because

(i) it is a high dimensional parameter and (ii) it is restricted to be positive definite. Pourahmadi (1999) introduced an unconstrained parameterization procedure to model a temporal covariance matrix. The Cholesky decomposition of the inverse of a covariance matrix is used to associate a unique unit lower triangular and a unique diagonal matrix with each covariance matrix. The entries of the lower triangular matrix and the log of the diagonal matrix are completely unconstrained and have interpretations similar to regression coefficients and prediction variances, respectively, when regressing a measurement on its predecessors. Using the Cholesky decomposition and the ensuing unconstrained and statistically meaningful reparameterization, Daniels and Pourahmadi (2002) provides a convenient and intuitive framework for developing conditionally conjugate prior distributions for covariance matrices, and show their connections with generalized inverse Wishart priors. However, to the best of our knowledge this type of reparameterization of covariance matrices has been used only to model temporal processes, taking advantage of the natural ordering of time. We extend these methodologies to spatial and temporal processes and show several extensions.

For our DSTM framework, we define two lower triangular matrices T^ν and T^ω and two diagonal matrices D^ν and D^ω such that $T^\nu \Sigma^\nu T^{\nu T} = D^\nu$ and $T^\omega \Sigma^\omega T^{\omega T} = D^\omega$. Such decompositions of positive definite matrices are unique. More precisely, let T^ν and T^ω be the lower triangular matrices with 1's as their diagonal entries and $-\phi_{ii'}$, $i > i'$ and $-\psi_{kk'}$, $k > k'$ as their lower triangular entries, respectively. Also, let D^ν and D^ω be diagonal matrices with entries $\sigma_1^{\nu 2}, \cdots, \sigma_n^{\nu 2}$ and $\sigma_1^{\omega 2}, \cdots, \sigma_p^{\omega 2}$, respectively. We now re-express the equations (3) and (4) using the entries of their lower triangular and diagonal matrices.

Let $Z_{it} = Z(s_i, t)$, $i = 1, \cdots, n$, $t = 1, \cdots, m$ and $X_{itk} = X_k(s_i, t)$, $k = 1, \cdots, p$. Notice that $\boldsymbol{Z}_t = (Z_{1t}, \cdots, Z_{nt})^T$ and $\{X_t\}_{n \times p} = ((X_{itk}))_{1 \le i \le n, 1 \le k \le p}$, which appear in (3). Then we can represent the DSTM defined by (3) and (4) as follows. The observation equation can be written as,

$$Z_{it} = \sum_{k=1}^{p} \beta_{kt} X_{itk} + \sum_{i'=1}^{i-1} \phi_{ii'} Z_{i't} + \nu_{it}, \tag{5}$$

$$Z_{1t} = \sum_{k=1}^{p} \beta_{kt} X_{1tk} + \nu_{1t}, \tag{6}$$

where $i = 2, \cdots, n$, $t = 1, \cdots, m$, $E[\nu_{it}] = 0$, $E[\nu_{it}^2] = \sigma_i^{\nu 2}$, and $E[\nu_{it}\nu_{i't}] = 0$, $i \neq i'$. Notice that $\boldsymbol{\nu}_t = (\nu_{1t}, \cdots, \nu_{nt})^T$ as in (3). The evolution equation can

now be written as,

$$\beta_{kt} = \sum_{k'=1}^{p} \beta_{k't-1} g_{kk'} + \sum_{k'=1}^{k-1} \psi_{kk'} \beta_{k't} + \omega_{kt}, \quad (7)$$

$$\beta_{1t} = \sum_{k'=1}^{p} \beta_{k't-1} g_{1k'} + \omega_{1t}, \quad (8)$$

where $k = 2, \cdots, p$, $t = 2, \cdots, m$, $E[\omega_{kt}] = 0$, $E[\omega_{kt}^2] = \sigma_k^{\omega 2}$, and $E[\omega_{kt}\omega_{k't}] = 0$, $k \neq k'$, and initial state equation can be written as,

$$\beta_{k1} = \beta_{k0} + \sum_{k'=1}^{k-1} \psi_{kk'} \beta_{k'1} + \omega_{k1}, \quad (9)$$

where $k = 2, \cdots, p$. Notice that $\boldsymbol{\omega}_t = (\omega_{1t}, \cdots, \omega_{nt})^T$ as in (4). The model is completed with

$$\beta_{11} = \beta_{10} + \omega_{11}. \quad (10)$$

Here, notice that no structural constraints are required for the elements of T^ν, T^ω, D^ν, and D^ω (i.e., $\phi_{ii'}, \psi_{kk'} \in \mathbb{R}$ and $\sigma_i^{\nu 2}, \sigma_k^{\omega 2} \in (0, \infty)$). In particular, for our applications we may specify a prior distribution for these parameters as,

$$\begin{aligned}
\phi_{ii'} &\sim N(\phi_0, \sigma_\phi^2), \ \ 1 \leq i' < i \leq n, \\
\psi_{kk'} &\sim N(\psi_0, \sigma_\psi^2), \ \ 1 \leq k' < k \leq p, \\
\sigma_{\nu i}^2 &\sim IG(a_\nu, b_\nu), \ \ i = 1, \cdots, n, \\
\sigma_{\omega k}^2 &\sim IG(a_\omega, b_\omega), \ \ k = 1, \cdots, p, \\
g_{kk'} &\sim N(g_0, \sigma_g^2), \ \ k, k' \in 1, \cdots, p,
\end{aligned}$$

where ϕ_0, σ_ϕ^2, ψ_0, σ_ψ^2, a_ν, b_ν, a_ω, b_ω, g_0 and σ_g^2 are all known values that can be used to quantify the prior information if available, otherwise we can use values that would generate a set of vague priors. Here $N(a, b)$ denotes a normal distribution with mean a and variance b, and $IG(a, b)$ denotes an inverse gamma distribution with mean $\frac{b}{a-1}$ for $a > 1$ and variance $\frac{b^2}{(a-1)^2(a-2)}$ for $a > 2$. For instance, we choose these values in such a way that these will have minimal impact on the posterior inference of the parameters. Other prior distributions can also be adapted for our framework very easily.

The advantages of using our proposed reparametrized DSTM (RDSTM) can be summarized as follows:

i) Numerical stability: RDSTM avoid numerical instabilities caused by multivariate updating scheme for DLM when the dimensions are very large.

ii) Routine handling of missing data: RDSTM allow missing data to be imputed from its full conditional distribution.

iii) Implementation using WinBUGS: RDSTM can be implemented using WinBUGS, while general DLM cannot.

We can extend the above framework to allow $\phi_{ii'}$ and $\psi_{kk'}$ to vary with time t and thus allowing Σ_t^ν to change with time t. Also letting $\xi_i(t) = \log \sigma_i^{\nu 2}(t)$ and $\eta_k(t) = \log \sigma_k^{\omega 2}(t)$, we can replace the equations (5)-(10) with time varying coefficients. However for such extensions we need to use some autoregressive models for these time varying coefficients. Although these extensions still keep all the updating equations univariate within the MCMC methodology, but it would require much more time to process a single cycle of the Gibbs sampler as each such cyle will consist of a large number of univariate sampling. Neverthless, the method will not suffer from numerical instabilities and curse of dimensionality as compared to multivariate updating schemes.

In next section, we demonstrate the applicability of our RDSTM to total nitrate concentration data. We also provide practical benefits of the estimates obtained from RDSTM in the context of atmospheric sciences, in addition to computational efficiency.

3 Application to Total Nitrate Concentration Data

3.1 Motivation

The release of different types of hazardous emissions has been instrumental in polluting the atmosphere over the last few decades, and hence atmospheric deposition has become a major topic of concern in environmental studies. The U.S. Environmental Protection Agency (EPA) established the Clean Air Status and Trends Network (CASTNET) to monitor air pollutant emissions and pollutant deposition (National Research Council, 2004). One of the goals of the scientists working with environmental data is to improve

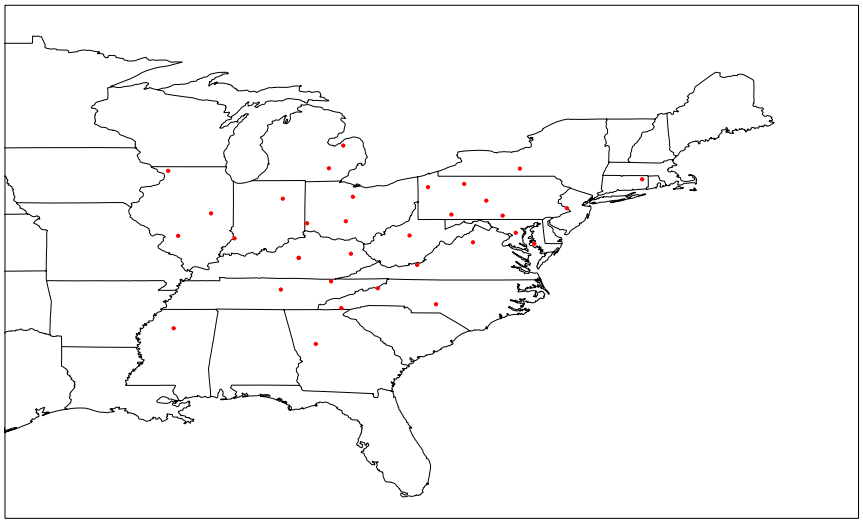

Figure 1: Locations of Stations

the accuracy of total nitrate prediction and finding the relationships of total nitrate with other predictors variables. To this end, one would like to explore the empirical relationship that may exist between the observed values of total nitrate and other observed variables. To estimate this empirical relationship in the observed data, we employ RDSTM proposed in Section 2. We use total nitrate as the response variable and consider the following variables as explanatory or predictor variables within a regression model: sulfate (SO_4), ammonium (NH_4), ozone (O_3), relative humidity (RH), wind speed (WS), and precipitation (P). We expect these variables to be reasonably good predictors of total nitrate. RDSTM allows us to estimate dynamic relationships that vary with weeks or months between total nitrate and the chemical species and meteorological variables.

3.2 The CASTNET Data

All of the data for this study were obtained from the U.S. EPA CASTNET sites. A complete description of this network can be found at:
`http://www.epa.gov/castnet`. Figure 1 shows the locations of the stations used in this study. We used $n = 33$ stations in the eastern U. S. These sites were selected on the basis of the extent of NO_X $(NO_2 + NO)$ emissions, where NO represents nitrous oxide. Note that all the CASTNET sites are located in rural locations.

Our data were collected between January, 1997 and July, 2004, and a total of $m = 394$ weeks are available during this period. Thus, the data appears to be rich on temporal scale than the spatial resolutions, and hence we consider DSTM which are ideal for modeling spatial times data. The chemical species used in this study were nitric acid (HNO_3) $(\mu mol/m^3)$, nitrate (NO_3^-) $(\mu mol/m^3)$, sulfate (SO_4) $(\mu mol/m^3)$, ammonium (NH_4) $(\mu mol/m^3)$ and ozone (O_3) (ppb). Nitric acid and nitrate were summed to get total nitrate $(\mu mol/m^3)$. Residual ammonium was also used, which is calculated as $(NH_4 - 2 \times SO_4)$. In the statistical analysis one could use ammonium, sulfate and residual ammonium. However, since residual ammonium is a known linear combination of the other two, only one of the other two should be used. By using sulfate and residual ammonium, one is prevented from having two ammonium variables in the statistical analysis, which also makes sense thermodynamically.

Meteorological variables are observed on site at each of the CASTNET stations. In this study we used relative humidity (RH) (%), wind speed (WS) (m/s) and precipitation (P) $(mm/week)$. All of the meteorological variables were measured hourly. With the exception of ozone, which was measured hourly, the chemical species were averaged over a week from Tuesday to Tuesday. To conform to this weekly pattern, RH and WS variables were averaged over the same period and the daily maximum O_3 values were averaged over each week. Precipitation was summed over the same period. A more detail study of this data has been conducted by these authors along with collaborators from EPA and results of such data analysis will be presented elsewhere. In this article, we present the data analysis only as an illustration of the proposed reparametrization method.

3.3 Model Fitting

Our results from the RDSTM (equation (5)-(10)) were obtained numerically using a Markov chain Monte Carlo (MCMC) procedure via the `WinBUGS` software. Usually a complete data set (i.e., no missing data) is required to fit the multivariate version of DSTM, however we have have missing values in the data. Our proposed RDSTM overcomes this limitation of a regular DSTM and performs imputations using Gibbs sampling. Gibbs sampling provides a natural solution by imputing values for the missing data at each iteration, sampling from their conditional posterior predictive distribution given the available data. Regression coefficients are then updated conditional on the imputed values. We assumed that each of the standardized covariates, when missing, follows a standard normal distribution, that is, $X_{itk}^{miss} \sim N(0, 1)$.

We analyzed the data using vague priors (i.e., proper priors with large variance) on parameters to have minimal impacts on the posterior inference. We assigned independent $N(0, 10^3)$ priors to $\phi_{ii'}$, $\psi_{kk'}$ and $g_{kk'}$, and independent $G(10^3, 10^3)$ priors to $1/\sigma_{\nu i}^2$ and $1/\sigma_{\omega k}^2$. Some sensitivity analysis were also performed to see the impacts of prior specifications and it was observed that the posterior inference was insensitive to sucg prior specifications as long as the priors were kept relatively vague. We obtained 10,000 iterates using a single chain from the MCMC sampler. The first 5000 iterates were discarded as a part of the Markov chain burn-in period, and all the posterior summaries reported were based on Monte Carlo estimates from the remaining 5000 iterates. The number of burn-in and final MCMC sample sizes were chosen using trace plots for parameters by diagnosing for their convergence performances to stationary region. We examined trace plots of the sampled values versus iteration to look for evidence of when the simulation appears to have stabilized to a stationary distribution.

3.4 Results

For every covariate, the RDSTM provides posterior estimates of the dynamic regression coefficient ($\boldsymbol{\beta}_t$) that changes with each week from January, 1997 to July, 2004, $t = 1, \ldots, 394$. This is in sharp contrast to the (static) linear regression coefficients. It is of interest to know how the covariates are dynamically related to total nitrate. RDSTM can provide the dynamic nature of the regression coefficient (see Figure 2) which provides a more informative relation between the total nitrate and other predictors.

We first checked how many weeks in each month have a significant positive/negative effect on total nitrate. For illustrative purpose, we used 95% equal-tail credible intervals to see whether the dynamic regression coefficients $\boldsymbol{\beta}_t$ are significant in the sense that these intervals do not contain the number zero. More formal methods, such as the use of Bayes factors and posterior predictive p-value can also be used for this purpose. To this end, we counted two separate numbers of significant weeks in each month for all the covariates. The first count provides the number of weeks which have significant positive coefficients (i.e., the lower limit of 95% interval is positive) and the other the significant negative coefficients (i.e., the upper limit is negative). The left-side plot in Figure 2 summarizes this result for SO_4. The right-side plot in Figure 2 presents the box plots for the posterior medians of significant coefficients for SO_4 in each month. Here, '* ' indicates the regression coefficient from linear regression models. These plots explain how strongly SO_4 is related to total nitrate across months and thus provides a better interpretation of the regression coefficients. We also computed the mean and the standard deviation of these posterior medians to summarize our findings numerically. This is given in Table 1 for SO_4. Here, N_{sig} represents the number of significant weeks, and N_{tot} represents the total number of weeks. Using these dynamic regression coefficients, we can find which covariate has the biggest effect on the response variable, total nitrate at what time of the year. In this article, we only show the results for SO_4. More detailed analysis that includes results based on other explanatory variables will be reported elsewhere.

In the left plot of Figure 2, sulfate seems to have a positive relationship with total nitrate uniformly over all months. This follows from the fact that in rural areas power plants are the main source of both SO_X (SO_3 plus SO_2) and NO_X (NO plus NO_2). Also, both pollutants would build up during stagnant meteorological conditions and be diluted during periods of high winds. In addition to the counts of significant weeks, the plot on the right-side of Figure 2 shows that the relationship and associated uncertainties between total nitrate and sulfate that varies with months. Sulfate seems to have a stronger relationship with total nitrate along with higher uncertainty during winter months. During the winter sulfate levels are low, and hence we have high level of residual ammonium. This results in increase of total nitrate because the residual ammonium reacts with nitrate. For linear regression models, the regression coefficients for June through September are far below from those of RDSTM.

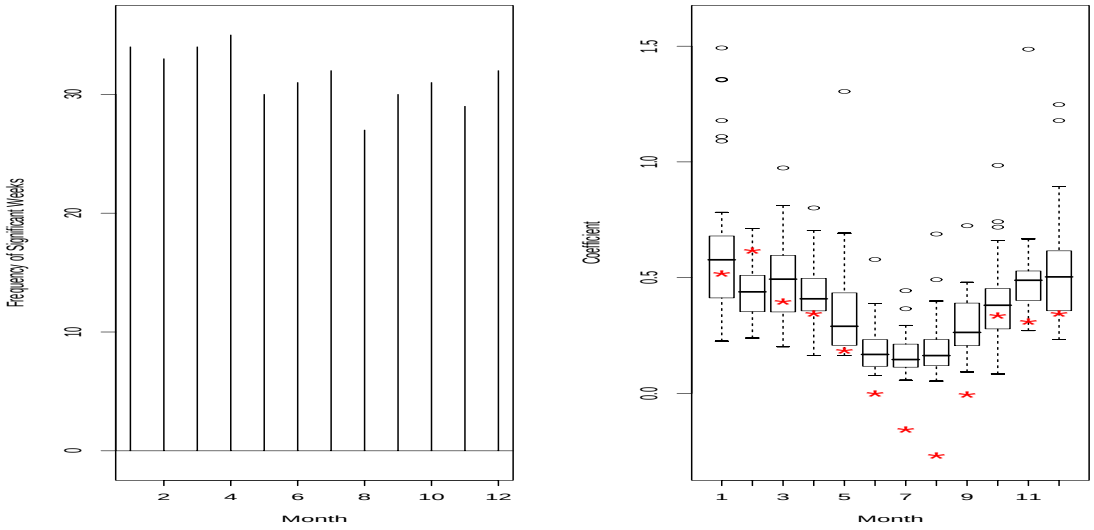

Figure 2: Frequencies and Box Plots for Posterior Medians of Significant Weeks in Each Month for SO_4

Month	N_{sig}	N_{tot}	Mean	S.E.
January	34	34	0.646	0.322
February	33	33	0.441	0.116
March	34	35	0.496	0.167
April	35	35	0.425	0.134
May	30	34	0.354	0.224
June	31	35	0.194	0.106
July	32	35	0.169	0.085
August	27	31	0.202	0.136
September	30	30	0.290	0.137
October	31	31	0.405	0.194
November	29	29	0.502	0.212
December	32	32	0.528	0.236

Table 1: Mean and S.D. of Posterior Medians of Significant Regression Coefficients for SO_4

Table 1 shows the mean and the standard deviation of the posterior medians of the significant regression coefficients. The mean ranges from 0.169 (July) to 0.646 (January) and such differences are significant as evident from non-overlapping box plots (e.g., compare the box plots of June-September to the rest of the box plots). The above analysis and related results clearly show the advantages of the RDSTM over more traditional static models in terms of being much more informative in extracting the relationship between the response (e.g. total nitrate) and other explanatory variables (e.g., SO_4). As the models are not limited by stationarity and other restrictive assumptions, our estimates are robust against such model misspecifications.

4 Discussions and Future Research

The illustration is the previous section clearly demonstrates that our proposed reprametrized method can be used fit relatively high-dimensional (33 $\times$ 394=13,002) data without making any restrictive assumptions about the spatial covariance functions. Also it demonstrates that data irregularities like missing or censored observations can be handled easily. As a part of future research the results from RDSTM could be used to diagnose the inherent problems numerical models like Community Multiscale Air Quality (CMAQ) might have with the simulations of total nitrate. As a first step in this effort, we have applied RDSTM to atmospheric measurements and obtained a dynamic relationship between total nitrate and the covariates over time. In the future, RDSTM may be applied to obtain predictive values of TNO_3 based on atmospheric data at grid locations used in CMAQ. By using model comparison methods (Fuentes and Raftery, 2005) to compare the predictions with CMAQ model values, if needed, improvements can be done to CMAQ in order to obtain more realistic predictions.

The model proposed by Wikle and Cressie (1999) closely resembles our DSTM. They introduced a dimension reduced approach to space-time Kalman filtering (STKF). The model can be written as,

$$\begin{aligned} \boldsymbol{Z}_t &= \Phi \boldsymbol{a}_t + \boldsymbol{v}_t^*, \quad \boldsymbol{v}_t^* \sim N(\boldsymbol{0}, R + V), \\ \boldsymbol{a}_t &= H \boldsymbol{a}_{t-1} + \boldsymbol{\eta}_t^*, \quad \boldsymbol{\eta}_t^* \sim N(\boldsymbol{0}, JQJ^T). \end{aligned}$$

This model has the same form as our DSTM except that Φ does not change over time and the variance matrices $R+V$ and JQJ^T are also assumed to be known (or estimated externally). Wikle and Cressie (1999) used an empirical

Bayesian technique for computing efficiency, because a fully Bayesian hierarchical approach requires highly intensive computational resources for Kalman filtering. But doing so, some statistical precision (and hence efficiency) is lost which can be obtained with the fully Bayesian approach. However, our RDSTM achieve both computing efficiency by using the reparametrized univariate scheme and statistical efficiency by using the fully hierarchical approach. Further they used standard method of moments to estimate R, V, Q, and B for computational efficiency and then pluuged in such estimated values as known quantities within the Kalman filter. This again leads to underestimation of uncertainty. Moreover, the use of maximum likelihood type procedures are not also possible to implement due to numerical instabilities. As we mentioned earlier, our RDSTM approach are not limited by such approximations due to high dimensionality. In addition, it is not clear if the positive definiteness is ensured for the estimated covariance matrices obtained by such moment based methods. In contrast, our RDSTM guarantees that the covariance matrices are positive definite and the estimates are robust to spatial covariance misspecifications. Finally, it is well-known that a fully hierarchical method of estimation provides the true measure of uncertainty as compared to two-stage methods that uses plugged-in estimates from the first stage.

As a final remark our proposed RDSTM also can be extended to develop a very flexible class of space-time models that are not subject to structural restrictions. Several possible extensions are possible and we list a few of those as follows:

i) Dynamic modeling of covariance function Σ^{ν} and Σ^{ω}: we can allow Σ^{ν} and Σ^{ω} to depend on t, that is, $\phi_{ii'}$'s, $\psi_{kk'}$'s, $\sigma_i^{\nu 2}$'s, and $\sigma_k^{\omega 2}$'s depend on t and thus making these parameters dynamic as well.

ii) Relaxing the Gaussian assumption on distributions for $\boldsymbol{\nu}_t$'s and $\boldsymbol{\omega}_t$'s: we can assume other than Gaussian distributions such as t-distributions for $\boldsymbol{\nu}_t$'s and $\boldsymbol{\omega}_t$'s if we suspect the presence of outliers.

iii) Extension of the first-order Markovian assumption: we can assume $\boldsymbol{\theta}_t$ is generated by a higher order Markovian process rather than a first-order in evolution equations in (7) and (8).

Acknowledgements

We are indebted to the contribution of Dr. Jerry M. Davis at NC State University and Dr. David M. Holland from EPA who provided the data sets used as an illustration in this article. We are also thankful to Dr. Prakash Bhave and Dr. Davis for providing many insights on chemistry and meteorology of chemical species used for this study. The research work of both authors was partially funded by an EPA cooperative grant (Co-Op 533246).

References

Banerjee, S., Gamerman, D., and Gelfand, A. E.(2005). Spatial process modelling for univariate and multivariate dynamic spatial data, *Environmetrics*, **16**, 465-479.

Brown, P. E., Karesen, K. F., Roberts, G. O., and Tonellato S. (2000). Blur-generated non-separable space-time models, *Journal of the Royal Statistical Society, Series B*, **62**, 847-860.

Cressie, N. A. C. and Huang, H.-C. (1999). Classes of nonseparable, spatiotemporal stationary covariance functions. *Journal of the American Statistical Association*, **94**, 1330-1340.

Daniels, M. J. and Pourahmadi M. (2002). Bayesian analysis of covariance matrices and dynamic models for longitudinal data. *Biometrika*, **89**, 553-566.

Fuentes, M. and Raftery, A.E. (2005). Model evaluation and spatial interpolation by Bayesian combination of observations with outputs from numerical models. *Biometrics*, **66**, 36-45.

Gelfand, A. E., Ghosh, S. K., Knight, J. R., and Sirmans, C. F. (1998). Spatio-Temporal Modeling of Residential Sales Data. *Journal of Business & Economic Statistics*, **16**, 312-321.

Gneiting, T. (2002). Nonseparable, stationary covariance functions for space-time data. *Journal of the American Statistical Association*, **97**, 590-600.

Huang, H.-C. and Cressie, N. A. C. (1996). Spatio-temporal prediction of snow water equivalent using the Kalman filter. *Computational Statistics and Data Analysis*, **22**, 159-175.

Huang, H.-C. and Hsu, N.-J. (2004). Modeling transport effects on ground-level ozone using a non-stationary space-time model. *Environmetrics*, **15**, 251-268.

Huerta, G., Sanso, B., and Stroud, J. R. (2004). A spatio-temporal model for Mexico city ozone levels. *Journal of the Royal Statistical Society*, Series C, **53**, 231-248.

Kyriakidis, P. C. and Journel, A. G. (1999). Geostatistical space-time models: a review. *Mathematical Geology*, **31**(6), 651-684.

Mardia, K. V., Goodall, C., Redfern, E. J., and Alonso, F. J. (1998). The kriged Kalman filter (with discussion). *Test*, **7**, 217-285.

National Research Council (2004). *Air Quality Management in the United States*, The National Academic Press.

Pole, A., West M. and Harrison, P. J. (1994). *Applied Bayesian forecasting and times series analysis*, Chapman and Hall: New York.

Pourahmadi M. (1999). Joint mean-covariance models with applications to longitudinal data: unconstrained parameterisation. *Biometrika*, **86**, 677-690.

Sansó, B. and Guenni, L. (1999). Venezuelan rainfall data analyzed using a Bayesian space-time model. *Applied Statistics*, **48**, 345-362.

Stein, M. L. (2003). Space-time covariance functions. *Journal of the American Statistical Association*, **100**, 310-321.

Stroud, J. R., Müller, P., and Sanso, B. (2001). Dynamic models for spatio-temporal data. *Journal of Royal Statistical Society, Series B*, **63**, 673-689.

Tonellato, S. (1997). Bayesian dynamic linear models for spatial time series, Technical report (Rapporto di riceria 5/1997), Dipartimento di Statistica, Universita CaFoscari di Venezia, Venice, Italy.

West M. and Harrison, P. J. (1997). *Bayesian Forecasting and Dynamic Models*, Springer: New York, 2nd ed.

Wikle, C., Berliner, M., and Cressie, N. (1999). Hierarchical Bayesian space-time models. *Environmental and Ecological Statistics*, **5**, 117-154.

Wikle, C. and Cressie, N. (1999). A dimension reduced approach to space-time kalman filtering. *Biometrika*, **86**, 815-829.

Xu, K. and Wikle, C. (2005). Estimation of Parameterized Spatio-Temporal Dynamic Models. *Ecological and Environmental Statistics*, to appear.