# Nonparametric Approach for Spatial-Temporal Model

Tingjin Chu*
School of Statistics, Renmin University, Beijing, China, 100872
email: chutj03@ruc.edu.cn

**Abstract**

The spatial-temporal data is very common in many areas, and spatial-temporal models are often used to find the underlying relationships among various factors and predict future trend through analyzing spatial-temporal data. In this paper, a nonparametric method is proposed to model spatial-temporal error process of spatial-temporal models. More specifically, spatial-temporal error process is approximated by a Karhunen-Loève expansion, and a Newton-Raphson algorihtm is proposed to obtain parameter estimates. The proposed method is fast to compute, and its performance is illustrated in the simulation studies.

Keywords: Karhunen-Loève expansion, Spline Basis, Newton-Raphson algorihtm,

## 1 Introduction

The spatial-temporal data sets, that is, data sets collected across different locations at various time, are very common in many research areas, e.g. climate changes, real estate prices monitoring. In order to investigate the underlying relationships among various factors and predict future trend, spatial-temporal models are often employed. One challenge for spatial-temporal model is to estimate spatial-temporal covariance matrix efficiently, which is essential for parameter estimation and prediction in the later steps. Note that spatial-temporal covariance matrix is usually a positive definite matrix of order N (number of observations) whose off diagonal elements are not zero. Therefore, the computational complexity for the determinant and inversion of covariance matrix is O($N^3$), which is time consuming, if not infeasible, when sample size is large.

Many research has been done for spatial-temporal covariance structure. For example, spectral method is used to construct the stationary covariance function (Cressie and Huang, 1999). Gneiting (2002) used monotone functions to construct nonseparable stationary covariance functions. In this paper, a Karhunen-Loève expansion is used to obtain a low-rank spatial-temporal model. Through this approach, accurate estimation can be obtained, while keeping computational time at a reasonable level. The remainder of paper is organized as follows. In Section 2, low-rank spatial-temporal model is introduced and an algorithm is proposed for the model. In Section 3, simulation studies are conducted to evaluate the proposed method. A discussion is added in Section 4.

# 2 Spatial-Temporal Model

## 2.1 Model

Let $R \in \mathbb{R}^d$ and $T \in \mathbb{R}$ be the spatial and temporal domain of interest respectively, with $d \geq 1$ denoting the dimension of spatial domain. We consider the following linear spatial-temporal model $\{y(\boldsymbol{s}, t) : \boldsymbol{s} \times t \in R \times T\}$:

$$y(\boldsymbol{s}, t) = \boldsymbol{x}(\boldsymbol{s}, t)^T \boldsymbol{\beta} + \varepsilon_1(\boldsymbol{s}, t) + \varepsilon_2(\boldsymbol{s}, t), \tag{1}$$

where $\mu(\boldsymbol{s}, t)$ is an unknown spatial-tempo mean function at location $\boldsymbol{s}$ and time $t$, $\varepsilon_1(\boldsymbol{s}, t)$ is a spatial-temporal error process with mean zero and a covariance function $\gamma((\boldsymbol{s}, t), (\boldsymbol{s}', t'))$, and $\varepsilon_2(\boldsymbol{s}, t)$ is the measurement error following i.i.d $N(0, \sigma^2)$.

For model (1), in order to obtain accurate regression parameter estimates $\widehat{\boldsymbol{\beta}}$, which is often the interest of research, spatial-temporal error process $\varepsilon_1(\boldsymbol{s}, t)$ must be estimated accurately first. There are various ways to model spatial-temporal error process $\varepsilon_1(\boldsymbol{s}, t)$, and here Karhunen-Loève expansion is applied to $\varepsilon_1(\boldsymbol{s}, t)$, that is,

$$\varepsilon_1(\boldsymbol{s}, t) = \sum_{j=1}^{\infty} \xi_j(t) \varphi_j(\boldsymbol{s}), \tag{2}$$

where $\{\xi_j(t)\}_{j=1}^{\infty}$ are independent random variables with respect to $j = 1, \ldots$. Moreover, $\xi_j(t) \sim N(0, \lambda_j(t))$, with variances $\lambda_1(t) \geq \lambda_2(t) \geq \ldots \geq 0$, and $\{\varphi_j(\cdot)\}_{j=1}^{\infty}$ are orthonormal eigenfunctions over $R$ such that $\int_R \varphi_j(\boldsymbol{s}) \varphi_{j'}(\boldsymbol{s}) d\boldsymbol{s} = 1$ when $j = j'$ and zero otherwise. Furthermore, we assume that

$$\gamma((\boldsymbol{s}, t), (\boldsymbol{s}', t)) = \gamma((\boldsymbol{s}, t'), (\boldsymbol{s}', t')).$$

That is, for different time $t$ and $t'$, $\varepsilon_1(\boldsymbol{s}, t)$ and $\varepsilon_1(\boldsymbol{s}, t')$ follow the same spatial process. Based on the above assumption, it can be seen that $\lambda_j(t) = \lambda_j(t')$, and we denote $\xi_j(t) \sim N(0, \lambda_j)$. Approximating $\varepsilon_1(\boldsymbol{s}, t)$ with the first $J$ eigenfunctions, and incorporating it into model (1), we obtain

$$y(\boldsymbol{s}, t) \approx \boldsymbol{x}(\boldsymbol{s}, t)^T \boldsymbol{\beta} + \sum_{j=1}^{J} \xi_j(t) \varphi_j(\boldsymbol{s}) + \varepsilon_2(\boldsymbol{s}, t). \tag{3}$$

Suppose that we observe $y(\boldsymbol{s}_i, t_j)$, where $j = 1, \ldots, n_s$ be $n_s$ spatial locations, $j = 1, \ldots, n_t$ be $n_t$ time locations, and total number of observations is $N = n_s n_t$. Then in model 3, the spatial-temporal covariance matrix takes the form of $\boldsymbol{\Sigma} = diag\{\boldsymbol{\Sigma}(t_1), \ldots, \boldsymbol{\Sigma}(t_{n_t})\}$ with $\boldsymbol{\Sigma}(t_j) = \boldsymbol{\Phi}^T \boldsymbol{\Lambda} \boldsymbol{\Phi} + \sigma^2 \boldsymbol{I}$, where $\boldsymbol{\Lambda} = diag\{\lambda_1, \ldots, \lambda_J\}$ is an diagonal matrix of $J \times J$, and $\boldsymbol{\Phi} = \varphi_j(\boldsymbol{s}_i)_{j=1,i=1}^{j=J,i=n_s}$ is a $J \times n_s$ matrix. Then the determinant and inversion of $\boldsymbol{\Sigma}$ can obtained by following formulas, which have only the complexity of $\mathrm{O}(J^3)$ instead of $\mathrm{O}(N^3)$,

$$\boldsymbol{\Sigma}(t_j)^{-1} = (\boldsymbol{\Phi}^T \boldsymbol{\Lambda} \boldsymbol{\Phi} + \sigma^2 \boldsymbol{I})^{-1} = \sigma^{-2} \boldsymbol{I} - \sigma^{-2} \boldsymbol{\Phi}^T \{\boldsymbol{\Phi} \boldsymbol{\Phi}^T + \sigma^2 \boldsymbol{\Lambda}^{-1}\}^{-1} \boldsymbol{\Phi},$$

$$|\boldsymbol{\Sigma}(t_j)| = |\boldsymbol{\Phi}^T \boldsymbol{\Lambda} \boldsymbol{\Phi} + \sigma^2 \boldsymbol{I}| = \sigma^{2n} |\boldsymbol{\Lambda}| |\boldsymbol{\Phi} \boldsymbol{\Phi}^T / \sigma^2 + \boldsymbol{\Lambda}^{-1}|.$$

### 2.2 Algorithm

Since $\varphi_j(\boldsymbol{s})$ are unknown, basis functions are used for approximation. Let $\boldsymbol{\phi}(\boldsymbol{s}) = (\phi_1(\boldsymbol{s}), \ldots, \phi_M(\boldsymbol{s}))^T$ be $M$ basis functions, we approximate $\varphi_j(\boldsymbol{s})$ by $\varphi_j(\boldsymbol{s}) \approx \boldsymbol{b}_j^T \boldsymbol{\phi}(\boldsymbol{s})$. Therefore, model (3) is approximated by model (4)

$$y(\boldsymbol{s}, t) \approx \boldsymbol{x}(\boldsymbol{s}, t)^T \boldsymbol{\beta} + \sum_{j=1}^{J} \xi_j(t) \boldsymbol{b}_j^T \boldsymbol{\phi}(\boldsymbol{s}) + \varepsilon_2(\boldsymbol{s}, t). \tag{4}$$

When $\boldsymbol{s} \in \mathbb{R}^d$ with $d = 1$, functional data analysis can be used. Here, we use a Newton-Raphson algorihtm on a Stiefel manifold proposed by Peng and Paul (2009), and modify it for two-dimensional cases. That is, instead of one dimensional basis functions, e.g., B-spline basis functions, we use two-dimensional basis functions, e.g., orthonormalized radial basis function (Buhmann, 2003). More specifically, the radial basis function is defined as $g(c\|\boldsymbol{s} - \boldsymbol{\kappa}_m\|)$, where $g$ is a pre-specified continuous function, $\boldsymbol{\kappa}_m$ is a knot point, and $c > 0$ is a constant. Commonly used choices for $g$ include $g(h) = h^2 \log(h)$, which leads to thin-plate splines, and $g(h) = e^{-h^2}$, which results in Gaussian radial splines.

# 3 Simulation Studies

In the simulation studies, the spatial domain $R = [0, 5] \times [0, 5]$ with number of locations $n_s = 25$. At each spatial location $\boldsymbol{s}_i$, $i = 1, \ldots, n_s$, there are $n_t$ observations, for $t = 1, \ldots, n_t$. Here we choose $n_t = 20, 40, 80$ with the corresponding sample size $N = 500, 1000, 2000$. The linear regression model has seven covariates with regression coefficients $\boldsymbol{\beta} = (4, 3, 2, 1, 0, 0, 0)^T$. The covariates are generated from standard normal distributions with a cross-covariate correlation of 0.5. In addition, we standardize each covariate to have sample mean 0 and sample variance 1, and the response to have a sample mean 0. Consequently, there is no intercept in this model. For $\varepsilon_2(\boldsymbol{s}, t)$, it is a zero-mean stationary Gaussian process with an covariance function $\gamma((\boldsymbol{s}, t), (\boldsymbol{s}', t')) = \sigma_1^2 \exp(-\|\boldsymbol{s} - \boldsymbol{s}'\|/c_s - \|t - t'\|/c_t)$, with $\sigma_1^2 = 8$, $c_s = 1$ and $c_t = 5$. Moreover, $\varepsilon_2(s)$ are independently generated from $N(0, \sigma_2^2)$ with $\sigma_2 = 2$.

For each sample size $N$, we simulate 100 data sets, and for each simulated data set, $\boldsymbol{\beta}$ are obtained using proposed method (*denoted as LR*) and least square method (*denoted as LSE*). The mean and mean square error (MSE) of the resulting estimates are reported in Table 1. For both methods, MSE becomes smaller as the sample size increases. Moreover, it can be clearly seen that the proposed method performs better than LSE in terms of MSE.

# 4 Discussion

In this paper, a Karhunen-Loève expansion is used to obtain a a low-rank spatial-temporal model, whose parameter estimates are obtained through a Newton-Raphson algorihtm on a Stiefel manifold with two dimensional basis functions. A simulation study is conducted to compare the performance of proposed method and least square estimates, and it can be seen that there is significant improvement by incorporating spatial-temporal error in the estimation. Note that in the proposed method, temporal part of error process are not estimated and is treated to be i.i.d instead. If

| $N$ | Method | $\beta_1$ | $\beta_2$ | $\beta_3$ | $\beta_4$ | $\beta_5$ | $\beta_6$ | $\beta_7$ |
|---|---|---|---|---|---|---|---|---|
| 500 | LR | 4.018 | 3.013 | 1.984 | 1.009 | -0.001 | -0.001 | -0.002 |
| | MSE | 0.029 | 0.033 | 0.031 | 0.024 | 0.028 | 0.029 | 0.034 |
| | LSE | 4.010 | 3.015 | 1.988 | 1.005 | -0.011 | 0.005 | 0.001 |
| | MSE | 0.032 | 0.040 | 0.037 | 0.031 | 0.032 | 0.034 | 0.038 |
| 1000 | LR | 3.986 | 3.012 | 2.015 | 1.007 | -0.013 | -0.001 | -0.017 |
| | MSE | 0.014 | 0.012 | 0.014 | 0.012 | 0.011 | 0.013 | 0.011 |
| | LSE | 3.977 | 3.009 | 2.017 | 1.015 | -0.012 | -0.004 | -0.011 |
| | MSE | 0.019 | 0.017 | 0.018 | 0.019 | 0.014 | 0.017 | 0.013 |
| 2000 | LR | 4.005 | 2.996 | 2.000 | 0.985 | 0.001 | -0.007 | 0.013 |
| | MSE | 0.006 | 0.006 | 0.006 | 0.005 | 0.007 | 0.006 | 0.007 |
| | LSE | 4.001 | 3.002 | 2.001 | 0.985 | 0.000 | -0.008 | 0.009 |
| | MSE | 0.007 | 0.010 | 0.009 | 0.007 | 0.009 | 0.007 | 0.009 |

Table 1: Simulation results : mean, mean square error (MSE) of regression coefficients estimates for sample size $N = 500,\ 100,\ 2000$.

we can incorporate temporal part of error process into algorithm, the accuracy of parameter estimation can be further enhanced, and we will leave it to the future research.

## References

Buhmann, M. D. (2003). *Radial Basis Functions: Theory and Implementations*. Cambridge University Press, Cambridge.

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

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

Peng, J. and Paul, D. (2009). A geometric approach to maximum likelihood estimation of the functional principal components from sparse longitudinal data. *Journal of Computational and Graphical Statistics*, 18:995–1015.