

SPATIO-TEMPORAL MODELS FOR SOME DATA SETS IN CONTINUOUS SPACE AND DISCRETE TIME

Samuel Seth Demel and Juan Du

Kansas State University

Abstract: Space time data sets are often collected at monitored discrete time lags, which are usually viewed as a component of time series. Valid and practical covariance structures are needed to model these types of data sets in various disciplines, such as environmental science, climatology, and agriculture. In this paper we propose several classes of spatio-temporal functions whose discrete temporal margins are some celebrated autoregressive and moving average (ARMA) models, and obtain necessary and sufficient conditions for them to be valid covariance functions. The possibility of taking advantage of well-established time series and spatial statistics tools makes it relatively easy to identify and fit the proposed model in practice. A spatio-temporal model with moving average type of temporal margin is fitted to Kansas daily precipitation to illustrate the application of the proposed model comparing with some popular spatio-temporal models in literature.

Key words and phrases: Autoregressive and moving average process, Fourier transform, Matérn covariance function, spatio-temporal covariance function.

1. Introduction

With many phenomena in nature observed across space and through time, the efficient statistical modeling techniques are needed to capture the spatio-temporal variability exhibited in these data sets for various application purposes. Examples in meteorology, climatology and ecology can be found in Haslett and Raftery (1989), Handcock and Wallis (1994), Sahu and Mardia (2005), Wikle and Royle (2005), among others. Often times appealing to computational convenience, one adopts separable models by simply multiplying spatial and temporal covariance functions together, which does not take into account the possible space-time interaction. However, many environmental and geophysical processes are dependent on the space-time interaction, which motivates a sequence of work in non-separable model development. See e.g., Cressie and Huang (1999), Gneiting (2002), Ma (2003), de Luna and Genton (2005), Stein (2005a,b), Gneiting, Genton, and Guttorp (2007). Most of traditional effort focuses on building space-time covariance models under the framework of continuous time and mention a need to contend with the fact that time data are usually measured at discrete

time points and normally viewed as realization of time series. Especially in environmental and agricultural research, the data are often recorded at regular time intervals and at irregular stations. The common practice is to start with simple discrete time series analysis, then combine with spatial information for model exploration. In order to carry the initial time series result over for an efficient space-time modeling, we need to develop sensible spatio-temporal covariance structure over discrete time. Some approaches have been attempted along this line, such as multivariate time series technique, autoregressive equation and spectral representation (see Bras and Rodríguez-Iturbe (1985), Storvik, Frigessi, and Hirst (2002) and Stein (2005b)). However the results are limited and most of the models do not provide explicit close-form of the covariance functions.

Since space-time data are often collected at discrete time lags, it is natural to assume the underlying space-time process $Z(s, t)$ is defined on $\mathbb{R}^d \times \mathbb{Z}$ and its covariance function is given by

$$C(s_1, s_2, t_1, t_2) = \text{Cov}(Z(s_1, t_1), Z(s_2, t_2)), \quad (s_1, t_1), (s_2, t_2) \in \mathbb{R}^d \times \mathbb{Z}. \quad (1.1)$$

The process is said to be stationary in space and time, if $EZ(s, t)$ is a constant and covariance function depends on s_1, s_2 and t_1, t_2 only through $s_1 - s_2$ and $t_1 - t_2$ with simplified notation $C(s_1 - s_2, t_1 - t_2)$. Similarly $C(s_1, s_2, t_1 - t_2)$ represents temporally stationary covariance. Following Ma (2005a), the spatial margin of the spatio-temporal covariance is defined as $C(s_1, s_2, t, t)$, i.e., the spatial covariance at a fixed time. The temporal margin looks at the temporal covariance of the process at a fixed location and is given by $C(s, s, t_1, t_2)$. To model most spatio-temporal data the usual starting point is to break the problem into these two parts, the time series part and the spatial part. There is a wealth of knowledge in the data exploration of these two types of processes, why not construct a model that takes advantage of this front end analysis efficiently. The well-established techniques of time series and spatial statistics are valuable tools for model selection and determination. Then these tools can aid in tackling the challenges of selecting and fitting complex spatio-temporal models.

The goal of this paper is to construct covariance functions to model the dependence structure of continuous space and discrete time data using an intuitive approach which utilizes existing time series and spatial statistics tools to facilitate model selection and construct models that are relatively easy to apply in practice. The sufficient or sufficient and necessary conditions are provided for the validity of the models. A spatio-temporal process where the temporal margin has a moving-average-type model is first studied in Section 2. The model is then extended to include some autoregressive-type models in Section 3. Invited by the fact that the autoregressive and moving average (ARMA) processes can be easily interpreted and techniques for estimating parameters are well-established,

we try to utilize these techniques to determine the model and find starting values to fit the overall space-time covariance function presented. Finally in Section 4, the proposed model is applied to daily precipitation of Kansas and the fitting is compared to the classical model introduced by Gneiting (2002) and Gneiting, Genton, and Guttorp (2007).

2. Moving-average-type Temporal Margin

As we know, the moving average models in time series are building blocks of more complex model structures. We begin our exploration of discrete temporal margins with this type of structure. In the case of the first-order moving average as temporal margin, we need to investigate the permissibility of real-valued functions $g_0(\mathbf{s}_1, \mathbf{s}_2)$ and $g_1(\mathbf{s}_1, \mathbf{s}_2)$, $\mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^d$, such that the following spatio-temporal function function

$$C(\mathbf{s}_1, \mathbf{s}_2, t) = \begin{cases} g_0(\mathbf{s}_1, \mathbf{s}_2), & t = 0, \\ g_1(\mathbf{s}_1, \mathbf{s}_2), & t = \pm 1, \mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^d, \\ 0, & \text{otherwise,} \end{cases} \quad (2.1)$$

in the domain of $\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{Z}$ is a covariance function. It is clear that its temporal margin at a fixed location $\mathbf{s}$, is a first-order moving average model, i.e. MA(1), provided that $|g_1(\mathbf{s}, \mathbf{s})| \leq 1/2g_0(\mathbf{s}, \mathbf{s})$, $\mathbf{s} \in \mathbb{R}^d$. It is often not easy to justify nonnegative definiteness of a proposed space-time function to be covariance function. Especially when (2.1) is not necessarily stationary in space, the well-known Bochner's Theorem can not be directly applied. However the following reformulation of (2.1) marks a promising avenue:

$$C(\mathbf{s}_1, \mathbf{s}_2, t) = \frac{g_0(\mathbf{s}_1, \mathbf{s}_2) + 2ag_1(\mathbf{s}_1, \mathbf{s}_2)}{2} \cdot \begin{cases} 1, & t = 0, \\ \frac{1}{2a}, & t = \pm 1, \\ 0, & t = \pm 2, \dots, \end{cases} \\ + \frac{g_0(\mathbf{s}_1, \mathbf{s}_2) - 2ag_1(\mathbf{s}_1, \mathbf{s}_2)}{2} \cdot \begin{cases} 1, & t = 0, \\ -\frac{1}{2a}, & t = \pm 1, \\ 0, & t = \pm 2, \dots, \end{cases} \quad (2.2)$$

which is a product-sum of purely spatial function and correlation functions of MA(1) models given a is a constant not less than 1. Based on this decomposition and structure of separable model in each summand, it is not hard to see the condition that $g_0(\mathbf{s}_1, \mathbf{s}_2) \pm 2ag_1(\mathbf{s}_1, \mathbf{s}_2)$ are spatial covariance functions, entails the validity of (2.1) as a space-time covariance function on $\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{Z}$. When $a = 1$, it turns out that this condition is also the necessary condition as shown in the following theorem, the proof is postponed to the Appendix.

Theorem 1. *The function (2.1) is a spatio-temporal covariance function on $\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{Z}$ if and only if both*

$$C_+(s_1, s_2) = g_0(s_1, s_2) + 2g_1(s_1, s_2), \quad s_1, s_2 \in \mathbb{R}^d,$$

and

$$C_-(s_1, s_2) = g_0(s_1, s_2) - 2g_1(s_1, s_2), \quad s_1, s_2 \in \mathbb{R}^d,$$

are spatial covariance functions on $\mathbb{R}^d$.

Benefiting from this theorem, we only need to check the validity of two purely spatial functions in order to verify that of the spatio-temporal function (2.1).

Example 1. The function

$$C(s_1, s_2, t) = \begin{cases} \exp(-\|s_1 - s_2\|^2), & t = 0, \\ \frac{1}{2} \exp(-\|s_1 + s_2\|^2), & t = \pm 1, \quad s_1, s_2 \in \mathbb{R}^d, \\ 0, & \text{elsewhere,} \end{cases}$$

is a covariance function on $\mathbb{R}^d \times \mathbb{Z}$, since both $\exp(-\|s_1 - s_2\|^2) + \exp(-\|s_1 + s_2\|^2)$ and $\exp(-\|s_1 - s_2\|^2) - \exp(-\|s_1 + s_2\|^2)$, $s_1, s_2 \in \mathbb{R}^d$, are spatial covariance functions on $\mathbb{R}^d$, where $\|s\| = \left(\sum_{k=1}^d s_k^2\right)^{1/2}$ is the usual Euclidean norm of $s \in \mathbb{R}^d$.

The stationary version of Theorem 1 can be obtained by replacing $g_0(s_1, s_2)$ and $g_1(s_1, s_2)$ with $h_0(s)$ and $h_1(s)$ respectively, where $h_0(s)$ is a stationary spatial covariance function. Now that the basic structure of a spatio-temporal covariance function with a moving-average-type temporal margin has been established, we now consider the spatial component to be more specific. Using Theorem 1, we impose the commonly used Matérn type spatial margin and give minimum conditions to create a valid covariance function. The Matérn spatial covariance model $(\alpha\|s\|)^\nu K_\nu(\alpha\|s\|)$, $s \in \mathbb{R}^d$, was proposed in Matérn (1960) in the general form, where α is a positive constant, and $K_\nu(x)$ stands for the modified Bessel functions of the second kind of order ν (Abramowitz and Stegun (1970)). Note that when $\nu = 1/2$ the Matérn reduces to the classic exponential spatial covariance function. Parameter ν is a smoothness parameter that controls the degree of the differentiability of the underlying process. We refer to Stein (1999) for the elaboration on the importance and flexibility of the Matérn model. The following theorem determines a spatio-temporal covariance function with spatial margin being a linear combination of Matérn spatial covariance models and temporal margin being a first-order moving average.

Theorem 2. Let ν , α_k , and β_k ($k = 1, 2$) be constants with $\nu > 0$, $0 < \alpha_1 < \alpha_2$ and $-1/2 \leq \beta_1 < \beta_2 < 1/2$. A necessary and sufficient condition for the function

$$C(\mathbf{s}, t) = \begin{cases} c(\alpha_1 \|\mathbf{s}\|)^\nu K_\nu(\alpha_1 \|\mathbf{s}\|) + (1-c)(\alpha_2 \|\mathbf{s}\|)^\nu K_\nu(\alpha_2 \|\mathbf{s}\|), & t = 0, \\ c(\alpha_1 \|\mathbf{s}\|)^\nu K_\nu(\alpha_1 \|\mathbf{s}\|)\beta_1 + (1-c)(\alpha_2 \|\mathbf{s}\|)^\nu K_\nu(\alpha_2 \|\mathbf{s}\|)\beta_2, & t = \pm 1, \mathbf{s} \in \mathbb{R}^d, \\ 0, & \text{otherwise,} \end{cases} \quad (2.3)$$

to be a stationary covariance function on $\mathbb{R}^d \times \mathbb{Z}$ is that the constant c satisfies

$$\left\{ 1 - \frac{\alpha_2^d(1-2\beta_1)}{\alpha_1^d(1-2\beta_2)} \right\}^{-1} \leq c \leq \left\{ 1 - \frac{\alpha_1^{2\nu}(1+2\beta_1)}{\alpha_2^{2\nu}(1+2\beta_2)} \right\}^{-1}. \quad (2.4)$$

The proof of this theorem is based on Theorem 1 and Bochner's theorem, the detail is given in the appendix. There are several remarks on Theorem 2.

Remark.

1. The range of β_1 and β_2 is suggested to ensure the validity of the MA(1) temporal margin $C(0, t) = 1_{\{t=0\}} + (c\beta_1 + (1-c)\beta_2)1_{\{t=\pm 1\}}$, $t \in \mathbb{Z}$, when c equals 0 or 1, where 1 represents an indicator function.
2. If $-1/2 \leq \beta_1 < \beta_2 < 1/2$ and c lies in $[0, 1]$, the proposed model (2.3) is valid for any $\alpha_i > 0$, $i = 1, 2$, by the proof of the theorem in appendix.
3. When β_2 approaches $1/2$ from its left-hand side, the lower bound of c tends to zero so that (2.3) reduces to a separable spatio-temporal model.
4. Since the interval $[0, 1]$ is only a subset of c 's permissible domain (2.4) and β_1 or β_2 may be negative, the function (2.3) is flexible to represent spatio-temporal positive and negative correlations.

Specially, taking $\nu = 1/2$ in (2.3) with the constant omitted, yields the case with linear combination of exponential type of spatial margin:

Corollary 1. Let α_k , and β_k ($k = 1, 2$) be assumed as in Theorem 2, the function

$$C(\mathbf{s}, t) = \begin{cases} c \exp(-\alpha_1 \|\mathbf{s}\|) + (1-c) \exp(-\alpha_2 \|\mathbf{s}\|), & t = 0, \\ c \exp(-\alpha_1 \|\mathbf{s}\|)\beta_1 + (1-c) \exp(-\alpha_2 \|\mathbf{s}\|)\beta_2, & t = \pm 1, \mathbf{s} \in \mathbb{R}^d, \\ 0, & \text{otherwise,} \end{cases} \quad (2.5)$$

is a stationary correlation function on $\mathbb{R}^d \times \mathbb{Z}$ if and only if the constant c satisfies

$$\left\{ 1 - \frac{\alpha_2^d(1-2\beta_1)}{\alpha_1^d(1-2\beta_2)} \right\}^{-1} \leq c \leq \left\{ 1 - \frac{\alpha_1(1+2\beta_1)}{\alpha_2(1+2\beta_2)} \right\}^{-1}. \quad (2.6)$$

It is worth noting that the bounds in (2.6) depend on the quotient α_1/α_2 only, instead of α_1 and α_2 individually. Therefore for constant c satisfying condition (2.6) and any $u \geq 0$, $C(u\mathbf{s}, t)$ is space-time covariance function, so is

$\int_0^\infty C(us, t) d\mu(u)$, given μ is a non-negative measure. Suppose φ is a completely monotone function on $[0, \infty)$, which admits a Laplace transform of a finite non-negative measure by Bernstein's theorem (Widder (1941, p.160)), we can generate a large semiparametric class of spatio-temporal covariance functions from Corollary 1.

Theorem 3. *Let $\varphi(x)$ be a completely monotone function on $[0, \infty)$. If c satisfies inequality (2.6), then*

$$C(s; t) = \begin{cases} c\varphi(\alpha_1\|s\|) + (1-c)\varphi(\alpha_2\|s\|), & t = 0, \\ c\varphi(\alpha_1\|s\|)\beta_1 + (1-c)\varphi(\alpha_2\|s\|)\beta_2, & t = \pm 1, s \in \mathbb{R}^d, \\ 0, & \text{otherwise,} \end{cases} \quad (2.7)$$

is a stationary covariance function on $\mathbb{R}^d \times \mathbb{Z}$.

Gneiting (2002) gives a table of commonly used completely monotone functions for covariance modeling. For example, if you choose $\varphi(x) = \exp\{-cx^\gamma\}$, $c > 0, 0 < \gamma \leq 1$, the resulting space-time function possesses a spatial margin which is a linear combination of powered exponentials.

3. ARMA-type Temporal Margin

In this section we consider modeling of spatial-temporal data with some other types of stationary time series margins. The product-sum format in previous section is carried on in the hope of utilizing Bochner's Theorem based on product measure. The following theorem provides a sufficient and necessary condition for the proposed space-time function to be valid covariance function with some autoregressive and moving average margins.

Theorem 4. *Let $\nu, \alpha_1, \alpha_2, \beta_1$, and β_2 be constants with $\nu > 0, 0 < \alpha_1 < \alpha_2$ and $-1 < \beta_1 < \beta_2 < 1$. A necessary and sufficient condition for the function*

$$C(s; t) = c(\alpha_1\|s\|)^\nu K_\nu(\alpha_1\|s\|)\beta_1^{|t|} + (1-c)(\alpha_2\|s\|)^\nu K_\nu(\alpha_2\|s\|)\beta_2^{|t|}, \quad s \in \mathbb{R}^d, t \in \mathbb{Z}, \quad (3.1)$$

to be a stationary covariance function on $\mathbb{R}^d \times \mathbb{Z}$ is that the constant c satisfies

$$\left\{ 1 - \frac{\alpha_2^d (1 - \beta_1)(1 + \beta_2)}{\alpha_1^d (1 + \beta_1)(1 - \beta_2)} \right\}^{-1} \leq c \leq \left\{ 1 - \frac{\alpha_1^{2\nu} (1 - \beta_2)(1 + \beta_1)}{\alpha_2^{2\nu} (1 + \beta_2)(1 - \beta_1)} \right\}^{-1}. \quad (3.2)$$

The proof of the theorem is given in the appendix.

In this model the spatial margin is a linear combination of two Matérn covariance functions. The temporal margin of (3.1) is $C(0; t) = c\beta_1^{|t|} + (1-c)\beta_2^{|t|}$, $t \in \mathbb{Z}$, with c confined in (3.2) and is a linear combination of correlation functions of two first-order autoregressive time series, which includes families of correlation

functions of stationary AR(1), AR(2), and ARMA (2, 1) time series. Roughly speaking, $\alpha_i, i = 1, 2$ and ν can be looked at as the scaling parameter and smoothness parameter for the spatial component. The β_i 's are the parameters for time series component and c plays as a balancing parameter based on strength of space and time interaction. Moreover, (3.1) is more than a mixture of two separable space-time covariance functions, because the permissible domain of c in (3.1) contains the interval $[0, 1]$ as a subset. For other values of c , (3.1) can be flexible to model negatively correlated structure. While if c lies in $[0, 1]$, we only need $\alpha_i \geq 0$ to ensure the validity of the proposed model. As Corollary 1 in previous section, a special case of Theorem 4 when $\nu = 1/2$, produces a spatio-temporal covariance function with an exponential type of spatial margin and a ARMA-type temporal margin.

One may ask whether we can always expand the proposed model to continuous process which covers the discrete one as constrained version on discrete domain. More precisely, the embedding question is: can we embed the covariance function (3.1) into a covariance function whose space-time domain is $\mathbb{R}^d \times \mathbb{R}$? To this end, let's compare (3.1) with the following covariance function whose temporal domain is $\mathbb{R}$ and the base in exponential expressions is positive

$$C(s; t) = c(\alpha_1 \|s\|)^\nu K_\nu(\alpha_1 \|s\|) \beta_1^{|t|} + (1-c)(\alpha_2 \|s\|)^\nu K_\nu(\alpha_2 \|s\|) \beta_2^{|t|}, \quad s \in \mathbb{R}^d, t \in \mathbb{R}, \quad (3.3)$$

where $0 < \beta_1 < \beta_2 < 1$ are defined as in Theorem 4. We can obtain the permissible range of c based on Theorem 3 of Ma (2005b) as following

$$\left\{ 1 - \frac{\alpha_2^d \ln \beta_2}{\alpha_1^d \ln \beta_1} \right\}^{-1} \leq c \leq \left\{ 1 - \frac{\alpha_1^{2\nu} \ln \beta_1}{\alpha_2^{2\nu} \ln \beta_2} \right\}^{-1}. \quad (3.4)$$

It is straightforward to verify the permissible interval of c in (3.2) is contained by that in (3.4). This means that a stationary random field on $\mathbb{R}^d \times \mathbb{Z}$ with covariance (3.1) can be embedded into a stationary random field on $\mathbb{R}^d \times \mathbb{R}$ with covariance (3.3) when β_1 and β_2 are both positive. However, this is not allowable if β_1 or β_2 is negative, in which case (3.3) would not be real-valued. It is unclear whether (3.1) can be embedded into a real-valued, stationary covariance function on $\mathbb{R}^d \times \mathbb{R}$ when β_1 or β_2 is negative.

When applying the proposed parametric models in this and previous sections, we can use time series techniques to fit time series for individual location developing ARMA order and starting values for β_1 , β_2 , and c , so that the final parameter estimation can be achieved by maximum likelihood estimation or Cressie (1993) weighted least square estimation (see Eq. (22) of Gneiting (2002)). For the spatial aspect we can use procedures in spatial statistics to find starting values for α_1 and α_2 . The advantage here is that we can employ some handy time series

general techniques, such as ACF, PACF or likelihood-based criteria like AIC or BIC, to determine the temporal model patterns and orders, since the temporal margin is treated as ordinary time series. Notice that this step provides an initial idea of what the marginal time series looks like for model selection, the choice of appropriate models will eventually be justified by the final space-time fitting criteria, which are often not quite sensitive to very mild difference of marginal time series choices. The simplicity is also of concern in final model selection. Hence the proposed models along with this stepwise estimation procedure is relatively convenient under growing demand on the statistical techniques for the model determination given all the different theoretical models developed in continuous space and time in literature.

Our proposed model presented can serve as an attempt in seeking of more straightforward approach to studying spatial-temporal data where at each location the temporal process can be modeled with some ARMA-type covariance structure. Estimation for the data application in next section was done using Cressie (1993) weighted least squares and techniques introduced by Gneiting (2002). Expanding these techniques to the general ARMA(p, q) would require careful expansion of the theorems presented herein and the computation should still be manageable if Cressie weighted least squares method is employed. We will consider more complex temporal margins in our future work.

4. Kansas Daily Precipitation Data

This section explores Kansas daily precipitation to illustrate the application of the proposed model. There are numerous issues that must be dealt with when it comes to the raw daily weather data, one being missing observations at certain locations and time points. This missing information can occur for many reasons; loss in a station funding, data not entered, and much more. Using space-time modeling to fill in these gaps is one way of solving this problem.

The data source is the National Oceanic and Atmospheric Administration (NOAA) and the weather stations across Kansas are shown by gray points in Figure 1. To normalize the data and provide stability, each county was aggregated by taking a daily average across the county's weather stations. The resulting dataset is 105 time series, one for each county, of daily precipitation data in millimeters over 8,030 daily time points from January 1, 1990 to December 31, 2011. The dataset was then split into the first fifteen years to fit the model and the last seven years to test the model predictive capabilities. To preform comparisons we follow the similar approach as in Gneiting (2002) and Gneiting, Genton, and Guttorp (2007). For time series margin the seasonal trend was fit and removed using annual harmonic regression, also the spatial trend was extracted by removing the station specific means.

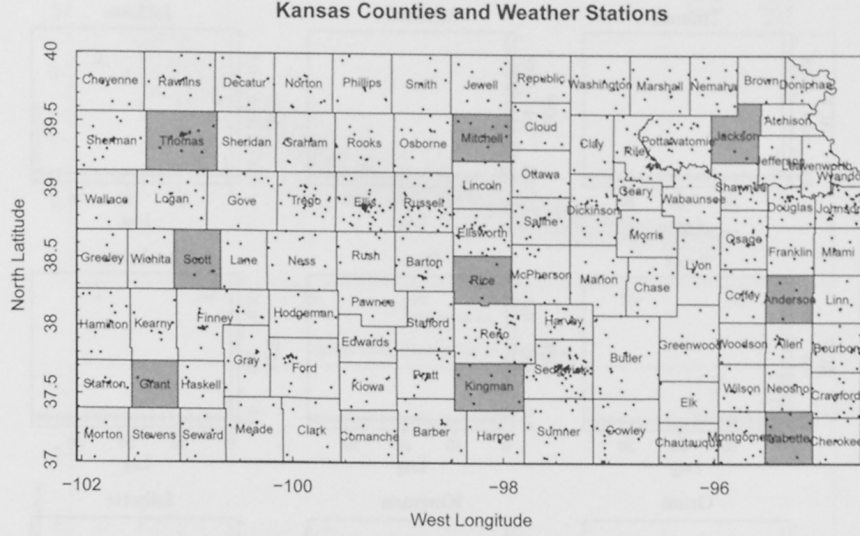


Figure 1. Kansas Counties Map: gray points are stations, orange highlighted counties sampled to display ACFs and AICs.

We begin the modeling of the cleaned data set with an exploratory analysis of temporal and spatial margins individually. Time series analysis has been familiar and often times is routine step to practitioners. Examining the ACF and AIC of the precipitation time series in all Kansas counties, we note that the majority of the counties chooses a MA(1) for a temporal margin based on these commonly used model selection criteria. To illustrate this result, referencing the orange highlighted counties of Figure 1 evenly located across Kansas, the ACF of those counties are given in Figure 2 and all of them indicates similar time series pattern of MA(1). So does the comparison of AICs for different competitive temporal marginal models. Hence it is determined that the temporal margin can be modeled with a MA(1) process. The left panel of Figure 3 shows the fitted empirical time correlations for two days of lag. Exploring the spatial component results in an exponential type structure. The right panel of Figure 3 shows the empirical spatial correlation with the fitted mixed exponential model. Suggested by these exploratory analysis, it seems reasonable to choose the space-time model with MA(1) temporal margin in Corollary 1. Incorporating nugget effect gives the proposed model,

$$C_{MA(1)}(\mathbf{s}; t) = \begin{cases} (1 - \eta) \{c \exp(-\alpha_1 \|\mathbf{s}\|) + (1 - c) \exp(-\alpha_2 \|\mathbf{s}\|)\} + \eta I_{\mathbf{s}=0, t=0}, \\ (1 - \eta) \{c \exp(-\alpha_1 \|\mathbf{s}\|) \beta_1 + (1 - c) \exp(-\alpha_2 \|\mathbf{s}\|) \beta_2\} \\ \quad + \eta \{c \beta_1 + (1 - c) \beta_2\} I_{\mathbf{s}=0}, & t = \pm 1, \mathbf{s} \in \mathbb{R}^d, \\ 0, & \text{otherwise.} \end{cases} \quad (4.1)$$

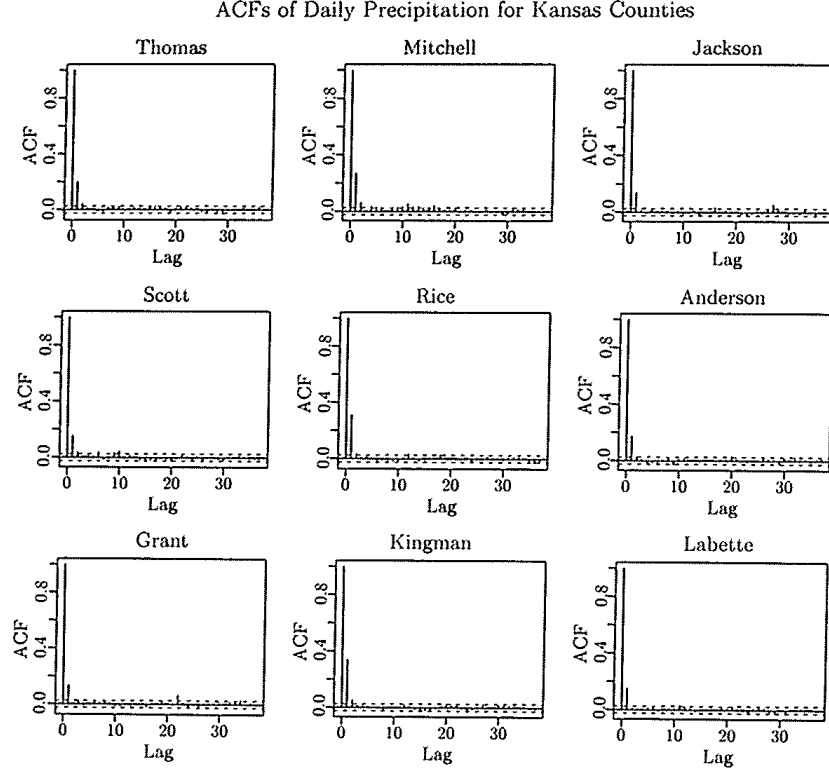


Figure 2. ACFs of nine chosen counties over Kansas

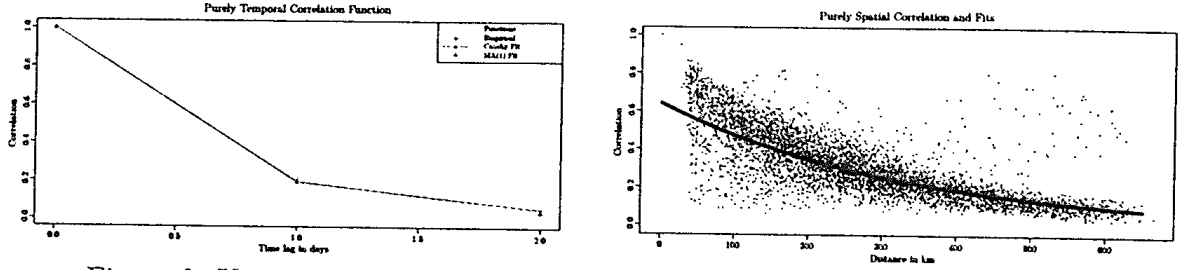


Figure 3. Kansas precipitation marginal plots: (left) time and (right) space.

Fitting space and time independent of each other allows for reasonable starting parameter values when fitting the overall space-time model using Cressie (1993) weighted least squares procedure. The estimated parameter values are $\eta = 0.322$, $\alpha_2 = 0.003$, $\alpha_1 = 0.009$, $c = 0.230$, $\beta_1 = -0.495$, and $\beta_2 = 0.495$.

To study the performance of the proposed model (4.1) comparisons are made with celebrated models in Gneiting (2002). Gneiting's separable model

$$C_{G.SEP}(s; t) = \{(1 - \eta) \exp(-c\|s\|) + \eta I_{\{s=0\}}\} \cdot \{(1 + a|t|^{2\alpha})^{-1}\}, \quad s \in \mathbb{R}^d, t \in \mathbb{R}, \quad (4.2)$$

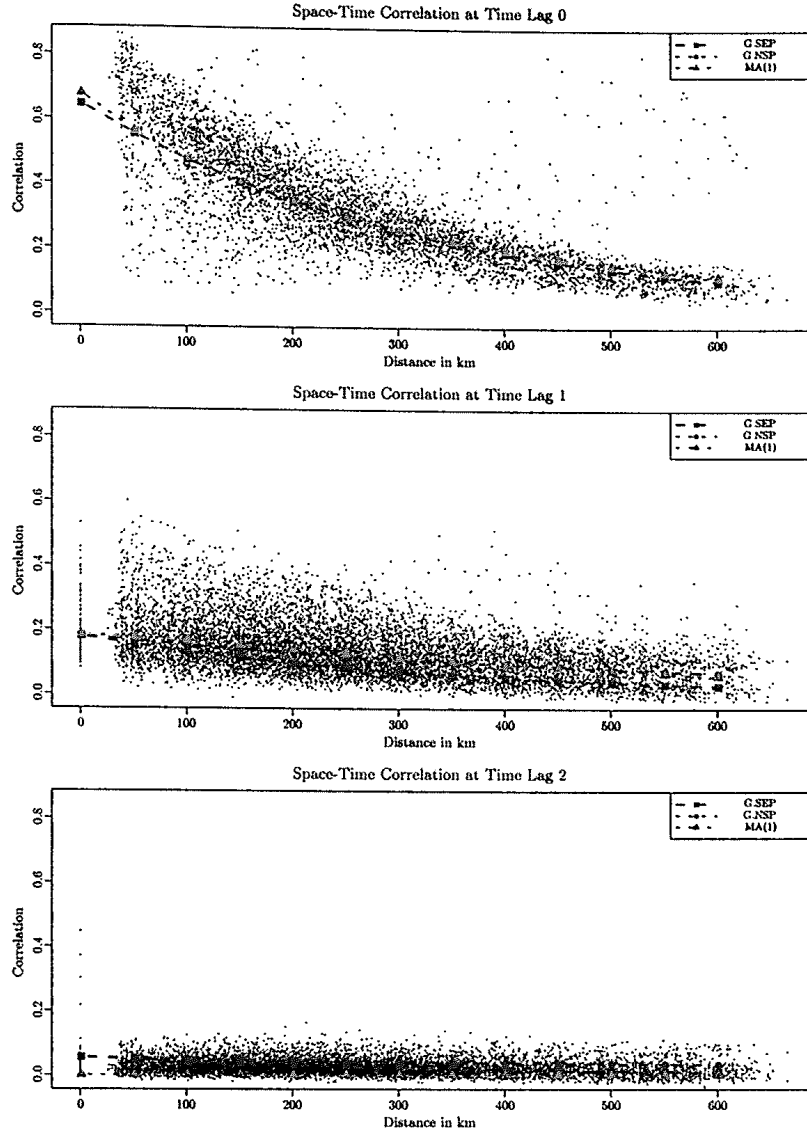


Figure 4. Empirical precipitation space-time correlations and fitted models based on 0, 1, 2 days of lag. Gneiting's separable model denoted by G.SEP (red), non separable model denoted by G.NSP (blue), and proposed model denoted by MA(1) (green).

and Gneiting's non separable model

$$C_{G.NSEP}(s; t) = \frac{1 - \eta}{1 + a|t|^{2\alpha}} \left\{ \exp \left(-\frac{c\|s\|}{(1 + a|t|^{2\alpha})^{\beta/2}} \right) + \frac{\eta}{1 - \eta} I_{\{s=0\}} \right\}, \quad s \in \mathbb{R}^d, t \in \mathbb{R}, \quad (4.3)$$

are fitted using similar procedures. Figure 4 shows the empirical space-time correlations for 0, 1, and 2 days of lag as well as the fitted space-time models. Gneiting's separable model is represented by the blue line, Gneiting's non-separable model is denoted by the red line, and the proposed model is the green

Table 1. Kansas Precipitation RMSE Statistics

Measure	G.SEP	G.NSEP	$C_{MA(1)}(s; t)$
AVG. RMSE	6.899	6.879	6.861
STD. DEV.	1.863	1.833	1.817
95% DATA.I.	(6.54, 7.24)	(6.52, 7.23)	(6.51, 7.21)
Low Count	29 (28%)	11 (10%)	65 (62%)

Table 2. Kansas Precipitation CRPS Statistics

Measure	G.SEP	G.NSEP	$C_{MA(1)}(s; t)$
AVG. CRPS	9.720	9.644	9.580
STD. DEV.	3.913	3.887	3.840
95% DATA.I.	(8.96, 10.48)	(8.89, 10.40)	(8.84, 10.32)
Low Count	0 (0%)	16 (15%)	89 (85%)

line. Notice that after two days of lag the space-time correlation is dying out. This fact is used to do one day ahead prediction based on two days of lag. As in Gneiting, Genton, and Guttorp (2007), the root mean-square error (RMSE) and continuous rank probability score (CRPS) are examined to compare the prediction performance of the models. Since the RMSE is calculated for each of 105 counties and it is meaningful to consider the consistency of the superior performance. Table 1 gives the average RMSE, standard deviation over all counties. The 95% data interval gives empirical interval that covers the majority of RMSE across counties. The “low count” tells how many counties have the lowest RMSE per model. Although all the models have a similar RMSE around 6.8 meaning on the average there is only 6.8 millimeters of error in predicting the next day’s precipitation, the proposed model denoted by $C_{MA(1)}$ does have the lowest average RMSE, variability and 62% of counties, are best fitted by the proposed model in prediction. Using the same techniques in Gneiting, Genton, and Guttorp (2007), we use the CRPS to compare predictive distributions. Table 2 shows the mean CRPS of all counties and again the proposed model has the lowest value. Also the model with MA(1) temporal structure has the most counties with the lowest CRPS with 89 out of 105 counties given by the Low Count.

Based on this analysis, the proposed model (4.1) preforms slightly better than Gneiting’s models when the temporal margin of the space-time process can be modeled with a MA(1) structure. It seems that taking into account the discrete nature of the times series does help to improve the predictability of the model. What’s more, the straightforward structure of the model gives intuitive meaning for each component, which eases the cumbersome task of determining the appropriateness of the model.

Acknowledgement

The authors would like to thank Professor Tilmann Gneiting for sharing the data information in Gneiting (2002) and two anonymous referees for their valuable comments and suggestions which helped to improve this paper.

Appendix

Proof of Theorem 1. In view of the product-sum decomposition (2.2) of $C(\mathbf{s}_1, \mathbf{s}_2, t)$ with $a = 1$, the sufficient part can be proved using the additive and multiplicative properties of covariance functions (see e.g., Schabenberger and Gotway (2005), p.34-44).

Conversely, suppose the function (2.1) is a spatio-temporal covariance function on $\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{Z}$, then $C(\mathbf{s}_1, \mathbf{s}_2, t)$ is nonnegative definite. For arbitrary n locations and m integer time points at each location, we formulate nm pairs $\mathbf{s}_i$ and t_j , and the corresponding coefficients are chosen as the products $a_i b_j$, $i = 1, \dots, n, j = 1, \dots, m$. We have

$$\begin{aligned} & \sum_{i=1}^n \sum_{i'=1}^n a_i a_{i'} C_+(\mathbf{s}_i, \mathbf{s}_{i'}) \sum_{j=1}^m \sum_{j'=1}^m b_j b_{j'} \rho_+(t_j - t_{j'}) \\ & + \sum_{i=1}^n \sum_{i'=1}^n a_i a_{i'} C_-(\mathbf{s}_i, \mathbf{s}_{i'}) \sum_{j=1}^m \sum_{j'=1}^m b_j b_{j'} \rho_-(t_j - t_{j'}) \geq 0, \end{aligned} \quad (\text{A.1})$$

where $\rho_+(t) = I_{\{t=0\}} + (1/2)I_{\{t=\pm 1\}}$ and $\rho_-(t) = I_{\{t=0\}} - (1/2)I_{\{t=\pm 1\}}$, $t \in \mathbb{Z}$, two MA(1) correlation functions. Note that only C_+ and C_- are of our interest, we can choose some special values for t_i and b_i to meet our need. To this end, in (A.1) let $t_j = j$, $b_j = \cos(j\omega)$ first, then let $t_j = j$, $b_j = \sin(j\omega)$, $j = 1, \dots, m$ and $\omega \in \mathbb{R}$. Adding these two inequalities together with m divided gives

$$\begin{aligned} & \left\{ 1 + \left(1 - \frac{1}{m} \right) \cos \omega \right\} \sum_{i=1}^n \sum_{i'=1}^n a_i a_{i'} C_+(\mathbf{s}_i, \mathbf{s}_{i'}) \\ & + \left\{ 1 - \left(1 - \frac{1}{m} \right) \cos \omega \right\} \sum_{i=1}^n \sum_{i'=1}^n a_i a_{i'} C_-(\mathbf{s}_i, \mathbf{s}_{i'}) \geq 0. \end{aligned}$$

When m tends to infinity, we obtain

$$(1 + \cos \omega) \sum_{i=1}^n \sum_{i'=1}^n a_i a_{i'} C_+(\mathbf{s}_i, \mathbf{s}_{i'}) + (1 - \cos \omega) \sum_{i=1}^n \sum_{i'=1}^n a_i a_{i'} C_-(\mathbf{s}_i, \mathbf{s}_{i'}) \geq 0. \quad (\text{A.2})$$

Taking $\omega = 0$ in (A.2) implies

$$\sum_{i=1}^n \sum_{i'=1}^n a_i a_{i'} C_+(\mathbf{s}_i, \mathbf{s}_{i'}) \geq 0,$$

which means that $C_+(s_1, s_2)$ is nonnegative definite. Similarly, taking $\omega = \pi$ in (A.2) yields that $C_-(s_1, s_2)$ is nonnegative definite. This concludes the proof.

Proof of Theorem 2. With Theorem 1, it is equivalent to show that inequalities (2.4) are necessary and sufficient condition for $g_0(s) \pm 2g_1(s)$ to be nonnegative definite. Under scenario of this theorem,

$$g_0(s) \pm 2g_1(s) = c(\alpha_1 \|s\|)^\nu K_\nu(\alpha_1 \|s\|)(1 \pm 2\beta_1) \\ + (1 - c)(\alpha_2 \|s\|)^\nu K_\nu(\alpha_2 \|s\|)(1 \pm 2\beta_2), \quad s \in \mathbb{R}^d.$$

Based on the spectral density of Matérn class of function (see Eq. (32) of Stein (1999)), the Fourier transforms of $g_0(s) + 2g_1(s)$ and $g_0(s) - 2g_1(s)$ can be readily found to be positively proportional to

$$f_1(u) = c\alpha_1^{2\nu}(\|u\|^2 + \alpha_1^2)^{-\nu-d/2}(1 + 2\beta_1) + (1 - c)\alpha_2^{2\nu}(\|u\|^2 + \alpha_2^2)^{-\nu-d/2}(1 + 2\beta_2), \\ \text{and} \\ f_2(u) = c\alpha_1^{2\nu}(\|u\|^2 + \alpha_1^2)^{-\nu-d/2}(1 - 2\beta_1) + (1 - c)\alpha_2^{2\nu}(\|u\|^2 + \alpha_2^2)^{-\nu-d/2}(1 - 2\beta_2), \\ u \in \mathbb{R}^d,$$

respectively. Hence, it is reduced to show that inequalities (2.4) are necessary and sufficient for $f_1(u) \geq 0$ and $f_2(u) \geq 0$, $u \in \mathbb{R}^d$ by Bochner's Theorem.

Suppose that $f_1(u) \geq 0$ and $f_2(u) \geq 0$ hold for every $u \in \mathbb{R}^d$. Since $0 < \alpha_1 < \alpha_2$ and $0 \leq 1 + 2\beta_1 \leq 1 + 2\beta_2$, the inequalities (2.4) follows from

$$0 \leq \lim_{u \rightarrow \infty} (\|u\|^2 + \alpha_1^2)^{\nu+d/2} f_1(u) = c\alpha_1^{2\nu}(1 + 2\beta_1) + (1 - c)\alpha_2^{2\nu}(1 + 2\beta_2) \quad (\text{A.3})$$

and

$$0 \leq f_2(0) = c\alpha_1^{-d}(1 - 2\beta_1) + (1 - c)\alpha_2^{-d}(1 - 2\beta_2). \quad (\text{A.4})$$

On the other hand, we are going to show that, under inequalities (2.4), $f_1(u) \geq 0$ and $f_2(u) \geq 0$ hold for every $u \in \mathbb{R}^d$. While this is obviously true if $0 \leq c \leq 1$, it remains to consider the cases

$$\text{Case I: } \left\{ 1 - \frac{\alpha_2^d(1 - 2\beta_1)}{\alpha_1^d(1 - 2\beta_2)} \right\}^{-1} \leq c \leq 0, \text{ and Case II: } 1 \leq c \leq \left\{ 1 - \frac{\alpha_1^{2\nu}(1 + 2\beta_1)}{\alpha_2^{2\nu}(1 + 2\beta_2)} \right\}^{-1}.$$

Case I: $1 - c$ is positive and (A.4) holds. Since $0 < \alpha_1 < \alpha_2$, we have

$$\left(\frac{\alpha_2^2}{\|u\|^2 + \alpha_2^2} \right)^{\nu+d/2} \geq \left(\frac{\alpha_1^2}{\|u\|^2 + \alpha_1^2} \right)^{\nu+d/2}.$$

Therefore

$$f_2(u) \geq \{c\alpha_1^{-d}(1 - 2\beta_1) + (1 - c)\alpha_2^{-d}(1 - 2\beta_2)\} \left(\frac{\alpha_1^2}{\|u\|^2 + \alpha_1^2} \right)^{\nu+d/2} \geq 0,$$

and with $0 \leq 1 + 2\beta_1 < 1 + 2\beta_2$,

$$f_1(\mathbf{u}) \geq \{c\alpha_1^{-d} + (1-c)\alpha_2^{-d}\} \left(\frac{\alpha_1^2}{\|\mathbf{u}\|^2 + \alpha_1^2} \right)^{\nu+d/2} (1 + 2\beta_1) \geq 0,$$

where the last inequality is obtained from

$$c \geq \left\{ 1 - \frac{\alpha_2^d(1 - 2\beta_1)}{\alpha_1^d(1 - 2\beta_2)} \right\}^{-1} \geq \left(1 - \frac{\alpha_2^d}{\alpha_1^d} \right)^{-1}.$$

Case II: In this case, c is positive and (A.3) is valid. Thus

$$f_1(\mathbf{u}) \geq \{c\alpha_1^{2\nu}(1 + 2\beta_1) + (1-c)\alpha_2^{2\nu}(1 + 2\beta_2)\}(\|\mathbf{u}\|^2 + \alpha_2^2)^{-\nu-d/2} \geq 0,$$

and

$$f_2(\mathbf{u}) \geq \{c\alpha_1^{2\nu} + (1-c)\alpha_2^{2\nu}\}(\|\mathbf{u}\|^2 + \alpha_2^2)^{-\nu-d/2}(1 - 2\beta_2) \geq 0, \quad \mathbf{u} \in \mathbb{R}^d,$$

where the last inequality is due to

$$c \leq \left\{ 1 - \frac{\alpha_1^{2\nu}(1 + 2\beta_1)}{\alpha_2^{2\nu}(1 + 2\beta_2)} \right\}^{-1} \leq \left(1 - \frac{\alpha_1^{2\nu}}{\alpha_2^{2\nu}} \right)^{-1}.$$

The proof is completed.

Proof of Theorem 4. The Fourier transform of (3.1) is positively proportional to

$$\begin{aligned} f(\mathbf{u}; \omega) &= c\alpha_1^{2\nu}(\|\mathbf{u}\|^2 + \alpha_1^2)^{-\nu-d/2} \frac{1 - \beta_1^2}{1 + \beta_1^2 - 2\beta_1 \cos \omega} \\ &\quad + (1-c)\alpha_2^{2\nu}(\|\mathbf{u}\|^2 + \alpha_2^2)^{-\nu-d/2} \frac{1 - \beta_2^2}{1 + \beta_2^2 - 2\beta_2 \cos \omega} \\ &= \{h_1(\mathbf{u}) - 2h_2(\mathbf{u}) \cos \omega\} \prod_{k=1}^2 (1 + \beta_k^2 - 2\beta_k \cos \omega)^{-1}, \quad \mathbf{u} \in \mathbb{R}^d, \omega \in [-\pi, \pi], \end{aligned}$$

where

$$\begin{aligned} h_1(\mathbf{u}) &= c\alpha_1^{2\nu}(\|\mathbf{u}\|^2 + \alpha_1^2)^{-\nu-d/2}(1 - \beta_1^2)(1 + \beta_2^2) \\ &\quad + (1-c)\alpha_2^{2\nu}(\|\mathbf{u}\|^2 + \alpha_2^2)^{-\nu-d/2}(1 - \beta_2^2)(1 + \beta_1^2), \end{aligned}$$

and

$$\begin{aligned} h_2(\mathbf{u}) &= c\alpha_1^{2\nu}(1 - \beta_1^2)\beta_2(\|\mathbf{u}\|^2 + \alpha_1^2)^{-\nu-d/2} \\ &\quad + (1-c)\alpha_2^{2\nu}(1 - \beta_2^2)\beta_1(\|\mathbf{u}\|^2 + \alpha_2^2)^{-\nu-d/2}, \quad \mathbf{u} \in \mathbb{R}^d. \end{aligned}$$

By Bochner's theorem, (3.1) is a stationary covariance function on $\mathbb{R}^d \times \mathbb{Z}$ if and only if its Fourier transform, $f(\mathbf{u}; \omega)$, is nonnegative, or equivalently,

$$h_1(\mathbf{u}) - 2h_2(\mathbf{u}) \cos \omega \geq 0, \quad \mathbf{u} \in \mathbb{R}^d, \omega \in [-\pi, \pi]. \quad (\text{A.5})$$

Moreover, inequality (A.5) holds for all $\mathbf{u} \in \mathbb{R}^d$ and $\omega \in [-\pi, \pi]$ if and only if

$$h_1(\mathbf{u}) - 2h_2(\mathbf{u}) \geq 0, \quad h_1(\mathbf{u}) + 2h_2(\mathbf{u}) \geq 0, \quad \mathbf{u} \in \mathbb{R}^d. \quad (\text{A.6})$$

In fact, on one hand we obtain (A.6) from (A.5) by simply taking $\omega = 0$ and π in (A.5), and on the other hand, inequalities (A.6) imply

$$h_1(\mathbf{u}) \geq 2|h_2(\mathbf{u})| \geq 2h_2(\mathbf{u}) \cos \omega, \quad \mathbf{u} \in \mathbb{R}^d, \omega \in [-\pi, \pi].$$

Notice that inequalities (A.6) are equivalent to

$$c\alpha_1^{2\nu}(\|\mathbf{u}\|^2 + \alpha_1^2)^{-\nu - \frac{d}{2}}(1 + \beta_1)(1 - \beta_2) + (1 - c)\alpha_2^{2\nu}(\|\mathbf{u}\|^2 + \alpha_2^2)^{-\nu - \frac{d}{2}}(1 + \beta_2)(1 - \beta_1) \geq 0, \quad (\text{A.7})$$

and

$$c\alpha_1^{2\nu}(\|\mathbf{u}\|^2 + \alpha_1^2)^{-\nu - \frac{d}{2}}(1 - \beta_1)(1 + \beta_2) + (1 - c)\alpha_2^{2\nu}(\|\mathbf{u}\|^2 + \alpha_2^2)^{-\nu - \frac{d}{2}}(1 - \beta_2)(1 + \beta_1) \geq 0. \quad (\text{A.8})$$

Hence, it suffices to show that inequalities (3.2) are necessary and sufficient for (A.7) and (A.8) to hold.

Suppose that (A.7) and (A.8) hold for any $\mathbf{u} \in \mathbb{R}^d$, then inequalities (3.2) follow by multiplying $(\|\mathbf{u}\|^2 + \alpha_1^2)^{\nu + d/2}$ on both sides of (A.7) and then letting $\mathbf{u}$ tend to infinity, i.e.

$$c\alpha_1^{2\nu}(1 + \beta_1)(1 - \beta_2) + (1 - c)\alpha_2^{2\nu}(1 + \beta_2)(1 - \beta_1) \geq 0. \quad (\text{A.9})$$

as well as substituting $\mathbf{u} = \mathbf{0}$ in (A.8), which gives

$$c\alpha_1^{-d}(1 - \beta_1)(1 + \beta_2) + (1 - c)\alpha_2^{-d}(1 - \beta_2)(1 + \beta_1) \geq 0. \quad (\text{A.10})$$

Conversely we are going to show that, under inequalities (3.2), (A.7) and (A.8) hold for any $\mathbf{u} \in \mathbb{R}^d$. While this is obviously true if $0 \leq c \leq 1$, it remains to consider the cases

$$\text{Case I: } \left\{ 1 - \frac{\alpha_2^d(1 - \beta_1)(1 + \beta_2)}{\alpha_1^d(1 + \beta_1)(1 - \beta_2)} \right\}^{-1} \leq c \leq 0;$$

$$\text{Case II: } 1 \leq c \leq \left(1 - \frac{\alpha_1^{2\nu}(1 - \beta_2)(1 + \beta_1)}{\alpha_2^{2\nu}(1 + \beta_2)(1 - \beta_1)} \right)^{-1}.$$

In Case I: $1 - c$ is positive and (A.10) is valid. Similar to the proof of Theorem 2, with $0 < \alpha_1 < \alpha_2$ the inequality in (A.8) holds, because the left-hand side (LHS) of (A.8) is greater than or equal to the following non-negative quantity

$$\{c\alpha_1^{-d}(1 - \beta_1)(1 + \beta_2) + (1 - c)\alpha_2^{-d}(1 - \beta_2)(1 + \beta_1)\} \left(\frac{\alpha_1^2}{\|u\|^2 + \alpha_1^2} \right)^{\nu+d/2}.$$

With $(1 + \beta_1)(1 - \beta_2) < (1 - \beta_1)(1 + \beta_2)$, inequality (A.7) follows because the LHS of (A.7) is greater than or equal to

$$\{c\alpha_1^{-d} + (1 - c)\alpha_2^{-d}\} \left(\frac{\alpha_1^2}{\|u\|^2 + \alpha_1^2} \right)^{\nu+d/2} (1 - \beta_2)(1 + \beta_1),$$

which is non-negative by the first inequality in (3.2).

In Case II: c is positive and (A.9) holds. We can deduce inequality (A.7) in the similar way as we did for the Case I by using the second inequality in (3.2). This completes the proof of the theorem.

References

- Abramowitz, M. and Stegun, I. (1970), *Handbook of Mathematical Functions*, Dover, New York.
- Bras, R. L. and Rodríguez-Iturbe, I. (1985). *Random Functions and Hydrology*. Addison-Wesley.
- Cressie, N. (1993). *Statistics for Spatial Data*. Revised edition. Wiley, New York.
- Cressie, N. and Huang, H. C. (1999). Classes of non-separable, spatio-temporal stationary covariance functions. *J. Amer. Statist. Assoc.* **94**, 1330-1340.
- de Luna, X. and Genton, M. G. (2005). Predictive spatio-temporal models for spatially sparse environmental data. *Statist. Sinica* **15**, 547-568.
- Gneiting, T. (2002). Nonseparable, stationary covariance functions for space-time data. *J. Amer. Statist. Assoc.* **97**, 590-600.
- Gneiting, T., Genton, M. and Guttorp, P. (2007). Geostatistical space-time models, stationarity, separability, and full symmetry. In *Statistical Methods for Spatio-Temporal Systems*. (Edited by B. Finkenstädt, L. Held and V. Isham), 151-174. Chapman and Hall/CRC, Boca Raton.
- Haslett, J. and Raftery, A. (1989). Space-time modeling with long-memory dependence: Assessing Ireland's wind power resource. *J. Roy. Statist. Soc. Ser. C* (Applied Statistics) **38**, 1-50.
- Handcock, M. S. and Wallis, J. R. (1994). An approach to statistical spatial temporal modeling of meteorological fields (with discussion), *J. Amer. Statist. Assoc.* **89**, 368-390.
- Ma, C. (2003). Spatio-temporal stationary covariance models. *J. Multivariate Anal.* **86**, 97-107.
- Ma, C. (2005a). Semi-parametric spatio-temporal covariance models with the ARMA temporal margin. *Ann. Inst. Math. Stat.* **37**, 221-233.
- Ma, C. (2005b). Spatio-temporal variograms and covariance models. *Adv. Appl. Prob.*, **37**, 706-725.

- Matérn, B. (1960). *Spatial Variation: Stochastic Models and Their Application to Some Problems in Forest Surveys and Other Sampling Investigations*. Meddelanden Fran Statens Skogsforskningsinstitut, Band 49, Stockholm.
- Sahu, S. K. and Mardia, K. V. (2005). A Bayesian Kriged-Kalman model for short-term forecasting of air pollution levels. *J. Roy. Statist. Soc. Ser. C* 54, 223-244.
- Schabenberger, O. and Gotway, C. (2005). *Statistical Methods for Spatial Data Analysis*. Chapman & Hall/CRC.
- Stein, M. L. (1999). *Interpolation of Spatial Data: Some Theory for Kriging*. Springer.
- Stein, M. L. (2005a). Space-time covariance functions. *J. Amer. Statist. Assoc.* 100, 310-321.
- Stein, M. L. (2005b). Statistical methods for regular monitoring data. *J. Royal Stat. Society Ser. B* 67, 667-687.
- Storvik, G., Frigessi, A. and Hirst, D. (2002). Stationary space-time Gaussian fields and their time autoregressive representation. *Statist. Model.* 2, 139-161.
- Widder, D. V. (1941). *The Laplace Transform*, Princeton University Press, Princeton.
- Wikle, C. K. and Royle, J. A. (2005). Dynamic design of ecological monitoring networks for non-Gaussian spatio-temporal data. *Environmetrics* 16, 507-522.

Statistics Department, Kansas State University, Manhattan, KS 66506, USA.

E-mail: s.s.demel56@gmail.com

Statistics Department, Kansas State University, Manhattan, KS 66506, USA.

E-mail: dujuan@k-state.edu

(Received August 2013; accepted February 2014)