Classes of Nonseparable, Spatio-Temporal Stationary Covariance Functions

Article *in* Journal of the American Statistical Association · January 1999

DOI: 10.1080/01621459.1999.10473885 · Source: CiteSeer

Hsin-Cheng Huang
National Chiao Tung University

Classes of Nonseparable, Spatio-temporal Stationary Covariance Functions

Noel CRESSIE, and Hsin-Cheng HUANG

Suppose that a random process $Z(s;t)$, indexed in space and time, has a spatio-temporal stationary covariance $C(h;u)$, where $h \in \mathbb{R}^d$ $(d \geq 1)$ is a spatial lag and $u \in \mathbb{R}$ is a temporal lag. Separable spatio-temporal covariances have the property that they can be written as a product of a purely spatial covariance and a purely temporal covariance. Their ease of definition is counterbalanced by the rather limited class of random processes to which they correspond. In this article, we derive a new approach that allows one to obtain many classes of nonseparable, spatio-temporal stationary covariance functions and we fit several such to spatio-temporal data on wind speed over a region in the tropical western Pacific ocean.

1. INTRODUCTION

Let $\{Z(\boldsymbol{s};t) : \boldsymbol{s} \in D \subset \mathbb{R}^d;\ t \in [0,\infty)\}$ denote a spatio-temporal random process that is observed at N space-time coordinates $(\boldsymbol{s}_1;t_1),\ldots,(\boldsymbol{s}_N;t_N)$. Optimal prediction (in space and time) of the unobserved parts of the process, based on the observations

$$\boldsymbol{Z} \equiv (Z(\boldsymbol{s}_1;t_1),\ldots,Z(\boldsymbol{s}_N;t_N))',$$

is often the ultimate goal but, in order to achieve this goal, a model is needed for how various parts of the process covary in space and time.

For example, $\boldsymbol{Z}$ might be the wind speed measured every 6 hours at n monitoring sites distributed throughout a region of interest (Section 4). Thus, between November 1992 and February 1993, there are on the order of $N = 480n$ observations for the spatio-temporal process representing wind speed. Although wind speed is potentially observable at any space-time coordinate $(\boldsymbol{s}_0;t_0)$, where $\boldsymbol{s}_0$ may not be a monitoring site and t_0 may be a time in the middle of a 6-hour period, the uncertainty associated with the unobserved parts of the process can be expressed probabilistically by modeling the wind speed to be a *random* process in space and time. Further, one might assume certain functional forms for the first and second moments (mean, variance, and covariance) of the random process.

In all that is to follow, we assume that the spatio-temporal process $Z(\cdot;\cdot)$ satisfies the regularity condition, $\mathrm{var}(Z(\boldsymbol{s};t)) < \infty$, for all $\boldsymbol{s} \in D$, $t \geq 0$. Then we can define the mean function as,

$$\mu(\boldsymbol{s};t) \equiv E(Z(\boldsymbol{s};t))$$

and the covariance function as,

$$K(\boldsymbol{s},\boldsymbol{r};t,q) \equiv \mathrm{cov}(Z(\boldsymbol{s};t),Z(\boldsymbol{r};q)); \quad \boldsymbol{s},\boldsymbol{r} \in D,\ t > 0,\ q > 0.$$

Noel Cressie is Professor of Statistics and Director, Program in Spatial Statistics and Environmental Sciences, Department of Statistics, The Ohio State University, Columbus, OH 43210-1247. Hsin-Cheng Huang is Assistant Research Fellow, Institute of Statistical Science, Academia Sinica, Taipei 115, Taiwan. This research was supported by the Office of Naval Research under grant N00014-93-1-0001 and the U.S. Environmental Protection Agency under cooperative agreement CR 822919-01-0, with Iowa State University. The authors would like to thank anonymous referees for comments that improved the presentation, and Chris Wikle for his help in obtaining the wind-speed data.

Furthermore, the optimal (minimum mean squared prediction error) linear predictor (e.g., Toutenburg, 1982, p. 14) of $Z(\boldsymbol{s}_0;t_0)$ is

$$Z^*(\boldsymbol{s}_0;t_0) = \mu(\boldsymbol{s}_0;t_0) + \boldsymbol{c}(\boldsymbol{s}_0;\boldsymbol{t}_0)'\Sigma^{-1}(\boldsymbol{Z} - \boldsymbol{\mu}), \quad (1)$$

where $\Sigma \equiv \mathrm{cov}(\boldsymbol{Z})$; $\boldsymbol{c}(\boldsymbol{s}_0;t_0) \equiv \mathrm{cov}(Z(\boldsymbol{s}_0;t_0),\boldsymbol{Z})$, and $\boldsymbol{\mu} \equiv E(\boldsymbol{Z})$; the minimum mean squared prediction error (MSPE) is $\boldsymbol{c}(\boldsymbol{s}_0;t_0)'\Sigma^{-1}\boldsymbol{c}(\boldsymbol{s}_0;t_0)$.

In the rest of this article, we shall assume that the covariance function is in fact stationary in space and time, namely

$$K(\boldsymbol{s},\boldsymbol{r};t,q) = C(\boldsymbol{s}-\boldsymbol{r};t-q), \quad (2)$$

for certain functions C. This assumption is often made so that the covariance function can be estimated from data.

Now, the function C has to satisfy a positive-definiteness condition in order to be a valid covariance function. That is, for any $(\boldsymbol{r}_1;q_1),\ldots,(\boldsymbol{r}_m;q_m)$, any real $a_1,\ldots,a_m$, and any positive integer m, C must satisfy

$$\sum_{i=1}^{m}\sum_{j=1}^{m} a_i a_j C(\boldsymbol{r}_i - \boldsymbol{r}_j; q_i - q_j) \geq 0. \quad (3)$$

Then and only then is (1) a valid, statistically optimal, spatio-temporal predictor of $Z(\boldsymbol{s}_0;t_0)$ with nonnegative MSPE. We further assume that C is continuous, although this assumption will be relaxed in Section 5. For continuous functions, positive-definiteness is equivalent to the process having a spectral distribution function (e.g., Matern, 1960, p. 12).

To ensure positive-definiteness, one often specifies the covariance function C to belong to a parametric family whose members are known to be positive-definite. That is, one assumes

$$\mathrm{cov}(Z(\boldsymbol{s};t),Z(\boldsymbol{s}+\boldsymbol{h};t+u)) = C^0(\boldsymbol{h};u|\boldsymbol{\theta}), \quad (4)$$

where C^0 satisfies (3) for all $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^p$.

Our goal in this article is to introduce new parametric families C^0 defined in (4) that will increase substantially the choices a modeler has for valid (i.e., positive-definite) spatio-temporal stationary covariances. One commonly used class (e.g., Rodriguez-Iturbe and Mejia, 1974)

Journal of the American Statistical Association
, Vol. , No. , Theory and Methods

Figure 1. Contour plot of $C(h;u) \equiv \exp\{-||h|| - |u|\}$, versus $||h||$ and $|u|$. The horizontal axis represents the modulus of the spatial lag and the vertical axis represents the temporal lag.

consists of separable covariances,

$$C^0(\boldsymbol{h}; u|\boldsymbol{\theta}) = C^1(\boldsymbol{h}|\boldsymbol{\theta}_1)C^2(u|\boldsymbol{\theta}_2), \tag{5}$$

where C^1 is a positive-definite function in $\mathbb{R}^d$, C^2 is a positive-definite function in $\mathbb{R}^1$, and $\boldsymbol{\theta}' = (\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')$. Valid spatial covariance models and valid temporal covariance models are readily available (e.g., Matern, 1960; Cressie, 1993, Sections 2.3 and 2.5) and hence they can be combined in product form via (5) to give valid spatio-temporal covariance models. A simple example of a separable model (5) is: $C^1(\boldsymbol{h}) = \exp(-\theta_1||\boldsymbol{h}||)$; $\theta_1 > 0$, and $C^2(u) = \exp(-\theta_2|u|)$; $\theta_2 > 0$, and hence

$$C^0(\boldsymbol{h}; u|\boldsymbol{\theta}) = \exp(-\theta_1||\boldsymbol{h}|| - \theta_2|u|).$$

The contour plot of the spatio-temporal covariance function for $\theta_1 = \theta_2 = 1$ is given in Figure 1. Separable models are often chosen for convenience rather than for their ability to fit the data well; at least they are guaranteed to satisfy (3) and hence are valid.

However, the class (5) is severely limited, since it does not model space-time interaction. Notice that, for any two fixed spatial lags $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$,

$$C^0(\boldsymbol{h}_1; u) \propto C^0(\boldsymbol{h}_2; u); \quad u \in \mathbb{R}.$$

Thus, for two spatial locations, the cross-covariance function between the time series at each location always has the same shape, regardless of the relative displacement of the locations. An analogous result holds for any pair of time points and the cross-covariance function of the two spatial processes.

Another type of separability involves adding spatial and temporal covariances; that is, $C^0(\boldsymbol{h}; u|\boldsymbol{\theta}) = C^1(\boldsymbol{h}|\boldsymbol{\theta}_1) + C^2(u|\boldsymbol{\theta}_2)$. For this model, covariance matrices of certain configurations of spatio-temporal data are singular (Myers and Journel, 1990; Rouhani and Myers, 1990), which is unsatisfactory when using (1) for optimal prediction.

Nonseparable stationary covariance functions that model space-time interactions are in great demand. Using simple stochastic partial differential equations over space and time, Jones and Zhang (1997) have developed a four-parameter family of spectral densities that implicitly yield such stationary covariance functions, although not in closed form.

In this article, a new and simple methodology is given for developing whole classes of nonseparable spatio-temporal stationary covariance functions, in closed form. In Section 2, we derive a theoretical result that shows how positive-definiteness in $\mathbb{R}^{d+1}$ can be obtained from positive-definiteness in $\mathbb{R}^d$. This result is used in Section 3 to define various classes of valid spatio-temporal stationary covariance models, including the separable models as a special case. Several are fitted to spatio-temporal data on wind speed over a region in the tropical western Pacific ocean; see Section 4. Finally, a short discussion of our approach is given in Section 5.

2. THEORETICAL RESULTS ON POSITIVE-DEFINITENESS

Consider the stationary spatio-temporal covariance function C given by (2). Assume that C is continuous and that its spectral distribution function possesses a spectral density $g(\boldsymbol{\omega}; \tau) \geq 0$. That is, by Bochner's Theorem (Bochner, 1955),

$$C(\boldsymbol{h}; u) = \int\int e^{i\boldsymbol{h}'\boldsymbol{\omega} + iu\tau} g(\boldsymbol{\omega}; \tau) d\boldsymbol{\omega} d\tau.$$

If, in addition, $C(\cdot;\cdot)$ is integrable, then

$$\begin{aligned} g(\boldsymbol{\omega}; \tau) &= (2\pi)^{-d-1} \int\int e^{-i\boldsymbol{h}'\boldsymbol{\omega} - iu\tau} C(\boldsymbol{h}; u) d\boldsymbol{h} du \\ &= (2\pi)^{-1} \int e^{-iu\tau} h(\boldsymbol{\omega}; u) du, \end{aligned} \tag{6}$$

where

$$\begin{aligned} h(\boldsymbol{\omega}; u) &\equiv (2\pi)^{-d} \int e^{-i\boldsymbol{h}'\boldsymbol{\omega}} C(\boldsymbol{h}; u) d\boldsymbol{h} \\ &= \int e^{iu\tau} g(\boldsymbol{\omega}; \tau) d\tau. \end{aligned}$$

The construction of C, or equivalently of g, in this article proceeds by specifying appropriate models for $h(\boldsymbol{\omega}; u)$. We assume that

$$h(\boldsymbol{\omega}; u) = \rho(\boldsymbol{\omega}; u) k(\boldsymbol{\omega}), \tag{7}$$

where the following two conditions are satisfied:

(C1) For each $\boldsymbol{\omega} \in \mathbb{R}^d$, $\rho(\boldsymbol{\omega}; \cdot)$ is a continuous autocorrelation function, $\int \rho(\boldsymbol{\omega}; u) du < \infty$, and $k(\boldsymbol{\omega}) > 0$.

(C2) $\int k(\boldsymbol{\omega}) d\boldsymbol{\omega} < \infty$.

Then (6) becomes

$$g(\boldsymbol{\omega}; \tau) \equiv (2\pi)^{-1} k(\boldsymbol{\omega}) \int e^{-iu\tau} \rho(\boldsymbol{\omega}; u) du > 0,$$

by (C1). Furthermore,

$$\int\int g(\boldsymbol{\omega}; \tau) d\tau d\boldsymbol{\omega} = \int k(\boldsymbol{\omega}) d\boldsymbol{\omega} < \infty,$$

by (C2). Therefore, assuming $h(\boldsymbol{\omega}; u)$ is given by (7) such that conditions (C1) and (C2) are satisfied, we see that

$$C(\boldsymbol{h}; u) \equiv \int e^{i\boldsymbol{h}'\boldsymbol{\omega}} \rho(\boldsymbol{\omega}; u) k(\boldsymbol{\omega}) d\boldsymbol{\omega} \tag{8}$$

is a valid (i.e., positive-definite) continuous spatio-temporal stationary covariance function on $\mathbb{R}^d \times \mathbb{R}$.

It is not hard to see that any continuous, integrable, spatio-temporal stationary covariance function can be written as in (7) with conditions (C1) and (C2) satisfied. Simply define

$$\rho(\boldsymbol{\omega}; u) \equiv \frac{h(\boldsymbol{\omega}; u)}{\int g(\boldsymbol{\omega}; \tau) d\tau},$$

and

$$k(\boldsymbol{\omega}) \equiv \int g(\boldsymbol{\omega}; \tau) d\tau$$

in (7).

Also notice that the covariance functions defined by (8) are generally not separable. However, the separable covariances arise as a special case: If the autocorrelation function ρ in (7) is purely a function of u, then (8) can be written in separable form. To sum up, our goal in this article is to find functions $h(\boldsymbol{\omega}; u)$ given by (7) that satisfy (C1) and (C2), *and for which the integral in (8) can be evaluated.* There are many new classes that can be defined in this manner, as the next section illustrates.

3. CLASSES OF CONTINUOUS SPATIO-TEMPORAL STATIONARY COVARIANCE MODELS

In this section, we give some parametric families of continuous spatio-temporal covariance functions $C(\boldsymbol{h}; u)$. Based on the results in Section 2, we have only to look for functions $\rho(\boldsymbol{\omega}; u)k(\boldsymbol{\omega})$ that satisfy the two conditions (C1) and (C2) and for which the integral in (8) can be evaluated. Then $C(\boldsymbol{h}; u)$ defined by (8) is a continuous spatio-temporal covariance function with corresponding spectral density, $g(\boldsymbol{\omega}; \tau) = (2\pi)^{-1} k(\boldsymbol{\omega}) \int e^{-iu\tau} \rho(\boldsymbol{\omega}; u) du$.

To construct the families of nonseparable spatio-temporal stationary covariances that follow, we used covariance functions and spectral density functions given in Matern (1960, Chapter 2). Through these examples, it will be seen generally how other closed-form Fourier transform pairs could be used to do the same.

Example 1. Let

$$\rho(\boldsymbol{\omega}; u) = \exp\{-\|\boldsymbol{\omega}\|^2 u^2/4\}$$

and

$$k(\boldsymbol{\omega}) = \exp\{-c_0\|\boldsymbol{\omega}\|^2/4\}; \quad c_0 > 0.$$

It is clear that both conditions (C1) and (C2) are satisfied. Therefore, from (8) and Matern (1960, p. 17),

$$C(\boldsymbol{h}; u) \propto \frac{1}{(u^2 + c_0)^{d/2}} \exp\left\{-\frac{\|\boldsymbol{h}\|^2}{u^2 + c_0}\right\}$$

is a continuous spatio-temporal covariance function in $\mathbb{R}^d \times \mathbb{R}$. So, a three-parameter spatio-temporal stationary covariance family can be given as,

$$C^0(\boldsymbol{h}; u|\boldsymbol{\theta}) = \frac{\sigma^2}{(a^2u^2 + 1)^{d/2}} \exp\left\{-\frac{b^2\|\boldsymbol{h}\|^2}{a^2u^2 + 1}\right\},$$

where $\boldsymbol{\theta} = (a, b, \sigma^2)'$, $a \geq 0$ is the scaling parameter of time, $b \geq 0$ is the scaling parameter of space, and $\sigma^2 = \boldsymbol{c}^0(\boldsymbol{0}; 0|\boldsymbol{\theta}) > 0$. Notice that, because of redundancy in the parameters a, b, and c_0, without loss of generality, we have put $c_0 = 1$. The contour plot of the spatio-temporal covariance function for $a = b = \sigma^2 = 1$ and $d = 2$ is shown in Figure 2.

Figure 2. Contour plot of $C(h; u) \equiv \frac{1}{u^2+1} \exp\left\{-\frac{\|h^2\|}{u^2+1}\right\}$, *versus* $\|h\|$ *and* $|u|$, *where* $h = (h_1, h_2)$. *The horizontal axis represents the modulus of the spatial lag and the vertical axis represents the temporal lag.*

Example 2. Let

$$\rho(\boldsymbol{\omega}; u) = \exp\{-\|\boldsymbol{\omega}\|^2 |u|/4\}$$

and

$$k(\boldsymbol{\omega}) = \exp\{-c_0\|\boldsymbol{\omega}\|^2/4\}; \quad c_0 > 0.$$

It is clear that both conditions (C1) and (C2) are satisfied. Therefore, from (8) and Matern (1960, p. 17),

$$C(\boldsymbol{h}; u) \propto \frac{1}{(|u| + c_0)^{d/2}} \exp\left\{-\frac{\|\boldsymbol{h}\|^2}{|u| + c_0}\right\}$$

is a continuous spatio-temporal covariance function in $\mathbb{R}^d \times \mathbb{R}$. So, a three-parameter spatio-temporal stationary covariance family can be given as,

$$C^0(\boldsymbol{h}; u|\boldsymbol{\theta}) = \frac{\sigma^2}{(a|u| + 1)^{d/2}} \exp\left\{-\frac{b^2\|\boldsymbol{h}\|^2}{a|u| + 1}\right\},$$

where $\boldsymbol{\theta} = (a, b, \sigma^2)'$, $a \geq 0$ is the scaling parameter of time, $b \geq 0$ is the scaling parameter of space, and $\sigma^2 = C^0(\boldsymbol{0}; 0|\boldsymbol{\theta}) > 0$. Again, without loss of generality, we have put $c_0 = 1$.

Example 3. Let

$$\rho(\boldsymbol{\omega}; u) = \exp\{-\|\boldsymbol{\omega}\| u^2\}$$

and

$$k(\boldsymbol{\omega}) = \exp\{-c_0\|\boldsymbol{\omega}\|\}; \quad c_0 > 0.$$

It is clear that both conditions (C1) and (C2) are satisfied. Therefore, from (8) and Matern (1960, p. 18),

$$C(\boldsymbol{h}; u) \propto \frac{1}{(u^2 + c_0)^d} \left\{1 + \frac{\|\boldsymbol{h}\|^2}{(u^2 + c_0)^2}\right\}^{-(d+1)/2}$$

is a continuous spatio-temporal covariance function in $\mathbb{R}^d \times \mathbb{R}$. So, a three-parameter spatio-temporal stationary covariance family can be given as,

$$C^0(\boldsymbol{h}; u|\boldsymbol{\theta}) = \frac{\sigma^2(a^2u^2 + 1)}{\{(a^2u^2 + 1)^2 + b^2\|\boldsymbol{h}\|^2\}^{(d+1)/2}},$$

where $\boldsymbol{\theta} = (a, b, \sigma^2)'$, $a \geq 0$ is the scaling parameter of time, $b \geq 0$ is the scaling parameter of space, and $\sigma^2 = C^0(\boldsymbol{0}; 0|\boldsymbol{\theta}) > 0$. Again, without loss of generality, we have put $c_0 = 1$.

Example 4. Let

$$\rho(\boldsymbol{\omega}; u) = \exp\{-\|\boldsymbol{\omega}\||u|\}$$

and

$$k(\boldsymbol{\omega}) = \exp\{-c_0\|\boldsymbol{\omega}\|\}; \quad c_0 > 0.$$

It is clear that both conditions (C1) and (C2) are satisfied. Therefore, from (8) and Matern (1960, p. 18),

$$C(\boldsymbol{h}; u) \propto \frac{1}{(|u| + c_0)^d} \left\{1 + \frac{\|\boldsymbol{h}\|^2}{(|u| + c_0)^2}\right\}^{-(d+1)/2}$$

is a continuous spatio-temporal covariance function in $\mathbb{R}^d \times \mathbb{R}$. So, a three-parameter spatio-temporal stationary covariance family can be given as,

$$C^0(\boldsymbol{h}; u|\boldsymbol{\theta}) = \frac{\sigma^2(a|u|+1)}{\{(a|u|+1)^2 + b^2||\boldsymbol{h}||^2\}^{(d+1)/2}},$$

where $\boldsymbol{\theta} = (a, b, \sigma^2)'$, $a \geq 0$ is the scaling parameter of time, $b \geq 0$ is the scaling parameter of space, and $\sigma^2 = C^0(\boldsymbol{0}; 0|\boldsymbol{\theta}) > 0$. Again, without loss of generality, we have put $c_0 = 1$.

Example 5. Let

$$\rho(\boldsymbol{\omega}; u) = \frac{c_0^{d/2}}{(u^2+c_0)^{d/2}} \exp\left\{-\frac{||\boldsymbol{\omega}||^2}{4(u^2+c_0)} + \frac{||\boldsymbol{\omega}||^2}{4c_0}\right\}$$

and

$$k(\boldsymbol{\omega}) = \exp\left\{-\frac{||\boldsymbol{\omega}||^2}{4c_0}\right\}; \quad c_0 > 0.$$

Since for each $\boldsymbol{\omega} \in \mathbb{R}^d$, $\rho(\boldsymbol{\omega}; u)$ is decreasing and convex for $u \in (0, \infty)$, it follows that condition (C1) is satisfied. Also, condition (C2) is clearly satisfied. Therefore, from (8) and Matern (1960, p. 17), the function, $\exp\{-(u^2+c_0)||\boldsymbol{h}||^2\}$, is a valid spatio-temporal covariance function in $\mathbb{R}^d \times \mathbb{R}$. Because the product of two valid covariance functions is a valid covariance function, we obtain

$$C(\boldsymbol{h}; u) \propto \exp\{-(u^2+c_0)||\boldsymbol{h}||^2 - a_0u^2\}; \quad a_0 > 0, \ c_0 > 0.$$

So, a four-parameter spatio-temporal stationary covariance family can be given as,

$$C^0(\boldsymbol{h}; u|\boldsymbol{\theta}) = \sigma^2 \exp\{-a^2u^2 - b^2||\boldsymbol{h}||^2 - cu^2||\boldsymbol{h}||^2\},$$

where $\boldsymbol{\theta} = (a, b, c, \sigma^2)'$, $a \geq 0$ is the scaling parameter of time, $b \geq 0$ is the scaling parameter of space, $c \geq 0$, and $\sigma^2 = C^0(\boldsymbol{0}; 0|\boldsymbol{\theta}) > 0$. The contour plots of the spatio-temporal covariance function for $a = b = \sigma^2 = 1$, $d = 2$, and $c = 0, 1, 5, 10$ are shown in Figure 3. Notice that a separable covariance function is obtained when $c = 0$; see Figure 3 (a).

Example 6. Let

$$\rho(\boldsymbol{\omega}; u) = \frac{c_0^{d/2}}{(|u|+c_0)^{d/2}} \exp\left\{-\frac{||\boldsymbol{\omega}||^2}{4(|u|+c_0)} + \frac{||\boldsymbol{\omega}||^2}{4c_0}\right\}$$

and

$$k(\boldsymbol{\omega}) = \exp\left\{-\frac{||\boldsymbol{\omega}||^2}{4c_0}\right\}; \quad c_0 > 0.$$

Since for each $\boldsymbol{\omega} \in \mathbb{R}^d$, $\rho(\boldsymbol{\omega}; u)$ is decreasing and convex for $u \in (0, \infty)$, it follows that condition (C1) is satisfied. Also, condition (C2) is clearly satisfied. Therefore, from (8) and Matern (1960, p. 17), the function, $\exp\{-(|u|+c_0)||\boldsymbol{h}||^2\}$, is a valid spatio-temporal covariance function in $\mathbb{R}^d \times \mathbb{R}$. Again, because the product of two valid covariance functions is a valid covariance function, we obtain

$$C(\boldsymbol{h}; u) \propto \exp\{-(|u|+c_0)||\boldsymbol{h}||^2 - a_0|u|\}; \ a_0 > 0, \ c_0 > 0.$$

So, a four-parameter spatio-temporal stationary covariance family can be given as,

$$C^0(\boldsymbol{h}; u|\boldsymbol{\theta}) = \sigma^2 \exp\{-a|u| - b^2||\boldsymbol{h}||^2 - c|u|||\boldsymbol{h}||^2\},$$

where where $\boldsymbol{\theta} = (a, b, c, \sigma^2)'$, $a \geq 0$ is the scaling parameter of time, $b \geq 0$ is the scaling parameter of space, $c \geq 0$, and $\sigma^2 = C^0(\boldsymbol{0}; 0|\boldsymbol{\theta}) > 0$. Notice that a separable covariance function is obtained when $c = 0$.

Example 7. Let

$$\begin{aligned}\rho(\boldsymbol{\omega}; u) &= \{u^2 + 1 + (u^2+c)||\boldsymbol{\omega}||^2\}^{-\nu-d/2} \\ &\quad \times \{1 + c||\boldsymbol{\omega}||^2\}^{\nu+d/2}; \quad c > 0, \ \nu > 0,\end{aligned}$$

and

$$k(\boldsymbol{\omega}) = \{1 + c||\boldsymbol{\omega}||^2\}^{-\nu-d/2}; \quad c > 0, \ \nu > 0.$$

Since for each $\boldsymbol{\omega} \in \mathbb{R}^d$, $\rho(\boldsymbol{\omega}; u)$ is decreasing and convex for $u \in (0, \infty)$, it follows that condition (C1) is satisfied. Also, condition (C2) is clearly satisfied. Therefore, from (8), and Matern (1960, p. 18) or, more explicitly, from Handcock and Wallis (1994, p. 370),

$$C(\boldsymbol{h}; u) \propto \begin{cases} \dfrac{1}{(u^2+1)^\nu (u^2+c)^{d/2}} \left\{\left(\dfrac{u^2+1}{u^2+c}\right)^{1/2} ||\boldsymbol{h}||\right\}^\nu \\ \times K_\nu\left(\left(\dfrac{u^2+1}{u^2+c}\right)^{1/2} ||\boldsymbol{h}||\right); & \text{if } ||\boldsymbol{h}|| > 0, \\ \dfrac{1}{(u^2+1)^\nu (u^2+c)^{d/2}}; & \text{if } ||\boldsymbol{h}|| = 0 \end{cases}$$

is a continuous spatio-temporal covariance function in $\mathbb{R}^d \times \mathbb{R}$, where K_ν is the modified Bessel function of the second kind of order ν (see, e.g., Abramowitz and Stegun, 1972, pp. 374ff.). So, a five-parameter spatio-temporal covariance family can be given as,

$$C^0(\boldsymbol{h}; u|\boldsymbol{\theta}) = \begin{cases} \dfrac{\sigma^2(2c^{d/2})}{(a^2u^2+1)^\nu (a^2u^2+c)^{d/2}\Gamma(\nu)} \left\{\dfrac{b}{2}\left(\dfrac{a^2u^2+1}{a^2u^2+c}\right)^{1/2} ||\boldsymbol{h}||\right\}^\nu \\ \times K_\nu\left(b\left(\dfrac{a^2u^2+1}{a^2u^2+c}\right)^{1/2} ||\boldsymbol{h}||\right); & \text{if } ||\boldsymbol{h}|| > 0, \\ \dfrac{\sigma^2 c^{d/2}}{(a^2u^2+1)^\nu (a^2u^2+c)^{d/2}}; & \text{if } ||\boldsymbol{h}|| = 0, \end{cases}$$

where $\boldsymbol{\theta} = (a, b, c, \nu, \sigma^2)'$, $a \geq 0$ is the scaling parameter of time, $b \geq 0$ is the scaling parameter of space, $c > 0$,

Figure 3. Contour plot of $C(h; u) \equiv \exp\left\{-u^2 - ||h||^2 - cu^2||h||^2\right\}$, *versus* $||h||$ *and* $|u|$, *where* $h = (h_1, h_2)$. *(a)* $c = 0$; *(b)* $c = 1$; *(c)* $c = 5$; *(d)* $c = 10$. *The horizontal axis represents the modulus of the spatial lag and the vertical axis represents the temporal lag.*

Figure 4. Contour plot of $C(h; u) \equiv \dfrac{4}{(u^2+1)^{1/2}(u^2+4)} \exp\left\{-\left(\frac{u^2+1}{u^2+4}\right)^{1/2} ||h||\right\}$, *versus* $||h||$ *and* $|u|$, *where* $h = (h_1, h_2)$. *The horizontal axis represents the modulus of the spatial lag and the vertical axis represents the temporal lag.*

$\nu > 0$, and $\sigma^2 = C^0(\mathbf{0}, 0|\boldsymbol{\theta}) > 0$. Notice that a separable covariance function is obtained when $c = 1$.

In particular, for $\nu = \frac{1}{2}$, we obtain a four-parameter spatio-temporal covariance family given as

$$C^0(\boldsymbol{h}; u|\boldsymbol{\theta}) = \frac{\sigma^2 c^{d/2}}{(a^2u^2+1)^{1/2}(a^2u^2+c)^{d/2}} \exp\left\{-b\left(\frac{a^2u^2+1}{a^2u^2+c}\right)^{1/2}||\boldsymbol{h}||\right\},$$

where $\boldsymbol{\theta} = (a, b, c, \sigma^2)'$. The contour plot of the spatio-temporal covariance function for $\nu = 1/2$, $a = b = \sigma^2 = 1$, $c = 4$, and $d = 2$ is shown in Figure 4.

4. APPLICATION TO WIND-SPEED DATA

In this section, we apply the new classes of spatio-temporal stationary covariance functions to the problem of mapping the east-west component of the wind speed over a region in the tropical western Pacific ocean. The data used in this article are collected on a regular grid of 17 x 17 sites with grid spacing of about 210km. Observations were taken every six hours from November 1992 through February 1993. That is, there are 289 spatial locations and 480 time points.

We first do some exploratory data analysis. Figure 5 shows the wind-speed fields (in m/s) at each of the first 15 time points, as well as the wind-speed field averaged over the 480 time points, where a positive value represents an east wind and a negative value represents a west wind. We can see strong spatio-temporal dependence from the first 15 wind-speed fields. We can also see that the mean wind-speed field is relatively flat. Figure 6 shows the time series plots of wind speed (in m/s) for sites at locations $(5, 5)$, $(5, 13)$, $(13, 5)$, $(13, 13)$, where we have used Cartesian coordinates based on the grid $\{1, ..., 17\} \times \{1, ..., 17\}$. A plot of the sample standard deviation versus the sample mean of wind speed (over time) obtained from each of the 289 sites is shown in Figure 7. No specific pattern (e.g., increasing pattern) is seen in this figure, indicating homoskedasticity. Therefore, from Figure 5, Figure 6, and Figure 7, a spatio-temporal stationarity assumption of the wind-speed field seems reasonable.

Let $Z(\boldsymbol{s}_i; t)$ be the observed east-west component of the wind speed (in m/s) for time t at site i; $t = 1, \ldots, 480$, $i = 1, \ldots, 289$. The empirical spatio-temporal variogram estimator is given by

$$2\hat{\gamma}(\boldsymbol{h}(l); u) \equiv \frac{1}{|N(\boldsymbol{h}(l); u)|} \sum_{(i,j,t,t') \in N(\boldsymbol{h}(l);u)} \big(Z(\boldsymbol{s}_i; t) - Z(\boldsymbol{s}_j; t')\big)^2,$$

where

$$N(\boldsymbol{h}(l); u) \equiv \big\{(i, j, t, t') : \boldsymbol{s}_i - \boldsymbol{s}_j \in \mathrm{Tol}(\boldsymbol{h}(l)); \; |t - t'| = u, \; i, j = 1, \ldots, 289\big\},$$

$\mathrm{Tol}(\boldsymbol{h}(l))$ is some specified "tolerance" region around $\boldsymbol{h}(l)$, and $|N(\boldsymbol{h}(l); u)|$ is the number of distinct elements in $N(\boldsymbol{h}(l); u)$; $l = 1, \ldots, L$, $u = 0, 1, \ldots, U$. The parameter $\boldsymbol{\theta}$ in a parametric spatio-temporal covariance function $C^0(\boldsymbol{h}; u|\boldsymbol{\theta})$ can then be estimated by fitting $\{2\hat{\gamma}(\boldsymbol{h}(l); u)\}$ to the spatio-temporal variogram,

$$\begin{aligned} 2\gamma(\boldsymbol{h}; u|\boldsymbol{\theta}) &\equiv \mathrm{var}\big(Z(\boldsymbol{s}+\boldsymbol{h}; t+u) - Z(\boldsymbol{s}; u)\big) \\ &= C^0(\mathbf{0}; 0|\boldsymbol{\theta}) - C^0(\boldsymbol{h}; u|\boldsymbol{\theta}); \quad \boldsymbol{h} \in \mathbb{R}^d, \; u \in \mathbb{R}. \end{aligned}$$

We use the weighted-least-squares method (e.g., Cressie, 1993, p. 96-97) to estimate $\boldsymbol{\theta}$, by minimizing

$$W(\boldsymbol{\theta}) \equiv \sum_{l=1}^{L} \sum_{u=0}^{U} |N(\boldsymbol{h}(l); u)| \left\{ \frac{\hat{\gamma}(\boldsymbol{h}(l); u)}{\gamma(\boldsymbol{h}(l); u|\boldsymbol{\theta})} - 1 \right\}^2 \qquad (9)$$

over all possible $\boldsymbol{\theta}$. Zimmerman and Zimmerman (1991) performed simulation experiments to compare (9) with likelihood-based methods and found that weighted least squares is sometimes the best fitting procedure and never does badly.

The estimated variogram based on 27 spatial lags, up to half the maximum possible distance, and 51 time lags is shown in Figure 8. From Figure 8, we can see a clear nonseparable feature for the spatio-temporal covariance, since the covariance (or variogram) for a larger spatial lag is almost constant, while the covariance for a smaller spatial lag is not. Recall that for a separable spatio-temporal stationary covariance function, we have $C^0(\boldsymbol{h}_1; u) \propto C^0(\boldsymbol{h}_2; u)$; $u \in \mathbb{R}$. Therefore, it is not appropriate to fit a separable spatio-temporal covariance function to the wind-speed data.

To select an appropriate spatio-temporal covariance function among the seven classes of models, we first look at the empirical spatial variogram for each time lag. We can see that the empirical spatial variogram for smaller time lags is clearly concave in $||\boldsymbol{h}||$, which is satisfied only by the spatio-temporal variograms of Example 2, Example 4, and Example 6. Therefore, we consider three spatio-temporal variogram models based on these three examples with the function $\tau^2 I(\boldsymbol{h} = \mathbf{0}, u = 0)$ added to each variogram to account for the nugget effect (see Section 5). We also consider the addition of a purely spatial variogram $2\alpha_1||\boldsymbol{h}||^{\alpha_2}$, which is needed to account for the empirical spatial variogram having almost the same shape for any larger time lag. Then the three spatio-temporal semivariogram models are given as follows:

Figure 5. Wind-speed fields (in m/s) for the first 15 time points and the mean wind-speed field (over time) on $17 \times 17 = 289$ wind-speed sites. Note that a positive value represents an east wind and a negative value represents a west wind.

Figure 6. Time series plots of wind speed (in m/s) for sites $(5, 5)$, $(5, 13)$, $(13, 5)$, $(13, 13)$ located on the grid $\{1, \ldots, 17\} \times \{1, \ldots, 17\}$.

Figure 7. Sample standard deviation versus sample mean of wind speed (in m/s). Means and standard deviations are calculated over time at each observation location.

Figure 8. Empirical spatio-temporal variogram evaluated at spatio-temporal lags $\{h(1), \ldots, h(27)\} \times \{0, 1, \ldots, 50\}$.

Model I. Based on Example 2, define the semivariogram model,

$$\gamma_1(\boldsymbol{h}; u|\boldsymbol{\theta}) \equiv \begin{cases} 0; & \text{if } u = \|\boldsymbol{h}\| = 0, \\ \sigma^2\left\{1 - \dfrac{1}{a|u|+1}\exp\left(-\dfrac{b^2\|\boldsymbol{h}\|^2}{a|u|+1}\right)\right\} & \\ +\tau^2 + \alpha_1\|\boldsymbol{h}\|^{\alpha_2}; & \text{otherwise.} \end{cases}$$

Using the weighted-least-squares criterion (9) for estimating $\boldsymbol{\theta} \equiv (a, b, \alpha_1, \alpha_2, \sigma^2, \tau^2)'$, we obtain $a = 0.399$, $b = 0.00235$, $\alpha_1 = 3.33 \times 10^{-6}$, $\alpha_2 = 1.999$, $\sigma^2 = 5.895$, and $\tau^2 = 0.164$. The weighted-least-squares value is $W(\boldsymbol{\theta}) = 2.01172 \times 10^6$.

Model II. Based on Example 4, define the semivariogram model,

$$\gamma_2(\boldsymbol{h}; u|\boldsymbol{\theta}) \equiv \begin{cases} 0; & \text{if } u = \|\boldsymbol{h}\| = 0, \\ \sigma^2\left\{1 - \dfrac{a|u|+1}{\left((a|u|+1)^2 + b^2\|\boldsymbol{h}\|^2\right)^{3/2}}\right\} & \\ +\tau^2 + \alpha_1\|\boldsymbol{h}\|^{\alpha_2}; & \text{otherwise.} \end{cases}$$

Using the weighted-least-squares criterion (9) for estimating $\boldsymbol{\theta} \equiv (a, b, \alpha_1, \alpha_2, \sigma^2, \tau^2)'$, we obtain $a = 0.1381$, $b = 0.00249$, $\alpha_1 = 3.53 \times 10^{-6}$, $\alpha_2 = 1.999$, $\sigma^2 = 5.861$, and $\tau^2 = 0$. The weighted-least-squares value is $W(\boldsymbol{\theta}) = 1.94985 \times 10^6$.

Model III. Based on Example 6, define the semivariogram model,

$$\gamma_3(\boldsymbol{h}; u|\boldsymbol{\theta}) \equiv \begin{cases} 0; & \text{if } u = \|\boldsymbol{h}\| = 0, \\ \sigma^2\{1 - \exp(-a|u| - b^2\|\boldsymbol{h}\|^2 - c|u|\|\boldsymbol{h}\|^2)\} & \\ +\tau^2 + \alpha_1\|\boldsymbol{h}\|^{\alpha_2}; & \text{otherwise.} \end{cases}$$

Using the weighted-least-squares criterion (9) for estimating $\boldsymbol{\theta} \equiv (a, b, c, \alpha_1, \alpha_2, \sigma^2, \tau^2)'$, we obtain $a = 0.186$, $b = 0.00238$, $c = 0$, $\alpha_1 = 3.65 \times 10^{-6}$, $\alpha_2 = 1.999$, $\sigma^2 = 5.313$, and $\tau^2 = 0.275$. The weighted-least-squares value is $W(\boldsymbol{\theta}) = 1.88228 \times 10^6$.

Table 1 displays comparable parameter estimates for the three models. The contour plots of the fitted spatio-temporal variogram functions for Model I, Model II, and Model III, with respect to the spatial lag $\|\boldsymbol{h}\|$ and the temporal lag u, are shown in Figure 9 (a), (b), and (c), respectively. Based on the smallest weighted-least-squares value of $W(\boldsymbol{\theta})$, Model III provides the closest fit. We conclude that, for the wind-speed data, the nonseparable empirical variogram is better fitted by a purely spatial variogram plus a spatio-temporal variogram with spatio-temporal interaction parameter $c = 0$. The three-dimensional plot of the fitted spatio-temporal variogram function based on Model III is displayed in Figure 10.

Table 1. Comparable parameter estimates for Model I, II, and III.

	Model I	Model II	Model III
α_1	3.33×10^{-6}	3.53×10^{-6}	3.65×10^{-6}
α_2	1.999	1.999	1.999
σ^2	5.895	5.861	5.313
τ^2	0.164	0	0.275
$W(\theta)$	2.01172×10^6	1.94985×10^6	1.88228×10^6

5. DISCUSSION

All spatio-temporal stationary covariances constructed according to the approach given in Section 2 are continuous. A discontinuity at the origin ($\boldsymbol{h} = \boldsymbol{0}, u = 0$) is allowed by adding the function $\tau^2 I(\boldsymbol{h} = \boldsymbol{0}, u = 0)$, to $C^0(\boldsymbol{h}; u|\boldsymbol{\theta})$ in (4), which is sometimes called a nugget effect in the geostatistics literature (e.g., Cressie, 1993, p. 59). In terms of the original process, this discontinuous component corresponds to an additive white-noise process. Further, notice that all the examples in Section 3 give stationary covariance functions that depend on spatial lag $\boldsymbol{h}$ through its modulus $\|\boldsymbol{h}\|$. This (spatial) isotropy can be relaxed by replacing $\|\boldsymbol{h}\|$ with $\|\boldsymbol{A}\boldsymbol{h}\|$, for any nonsingular matrix $\boldsymbol{A}$, in which case the model is sometimes referred to as (spatial) geometrically anisotropic (e.g., Cressie, 1993, p. 64).

It is obvious from Section 2 that, at any fixed temporal lag u, it is not enough to fit a positive-definite function of $\boldsymbol{h}$ and hope that the resulting fit, $C^1(\boldsymbol{h}|\boldsymbol{\theta}(u))$, is positive-definite in *both* $\boldsymbol{h}$ and u. In spite of the attractiveness of such a fitting procedure, it does not generally lead to a valid spatio-temporal covariance function. However, there is a circumstance where it does. Suppose that we can write the spectral density $g^1(\boldsymbol{\omega}|\boldsymbol{\theta}(u))$ of $C^1(\boldsymbol{h}|\boldsymbol{\theta}(u))$ as,

$$g^1(\boldsymbol{\omega}|\boldsymbol{\theta}(u)) = K_{\boldsymbol{\omega}}(u); \quad u \in (-\infty, \infty), \tag{10}$$

where $K_{\boldsymbol{\omega}}(\cdot)$ is a valid positive covariance function in $\mathbb{R}^1$ for each $\boldsymbol{\omega} \in \mathbb{R}^d$. Now, if we can write $K_{\boldsymbol{\omega}}(u) = \rho(\boldsymbol{\omega}; u)k(\boldsymbol{\omega})$, where $\rho(\boldsymbol{\omega}; \cdot)$ is a positive correlation function and $\int \rho(\boldsymbol{\omega}; u)du < \infty$ for each $\boldsymbol{\omega} \in \mathbb{R}^d$, we see that equation (10) is a special case of equation (7), where (C1) is satisfied. Hence, if the integrability condition (C2) is also satisfied, the integral on the right hand side of (8) is $C^1(\boldsymbol{h}|\boldsymbol{\theta}(u))$, which from Section 2 is a valid spatio-temporal covariance function.

Finally, although our results have been presented in a spatio-temporal context, they also allow construction of valid covariance models in $\mathbb{R}^{d+1}$ based on spatial covariance models in $\mathbb{R}^d$ and $\mathbb{R}^1$. For example, putting $d = 2$ and $u = h_3$ in (8) yields a valid stationary covariance model, $C(h_1, h_2, h_3)$, in $\mathbb{R}^3$.

Figure 9. Contour plot of the weighted-least-squares fitted spatio-temporal variogram, versus $\|h\|$ and $|u|$, where $h = (h_1, h_2)$. (a) Model I; (b) Model II; (c) Model III.

Figure 10. Three-dimensional plot of the weighted-least-squares fitted spatio-temporal variogram of Model III, versus $\|h\|$ and $|u|$, where $h = (h_1, h_2)$.

[Received , 1998. Revised , 1999.]

REFERENCES

Abramowitz, M., and Stegun, I. A., (1972), *Handbook of Mathematical Functions* Washington, DC.: National Bureau of Standards.

Bochner, S., (1955), *Harmonic Analysis and the Theory of Probability* Berkeley and Los Angeles: University of California Press.

Cressie, N., (1993), *Statistics for Spatial Data, rev. edn.* New York: Wiley.

Handcock, M. S., and Wallis, J. R., (1994), "An approach to statistical spatial-temporal modeling of meteorological fields," *Journal of the American Statistical Association*, 89, pp. 368–378.

Jones, R. H., and Zhang, Y., (1997), "Models for continuous stationary space-time processes," in *Modelling Longitudinal and Spatially Correlated Data*, eds. T. G. Gregoire, D. R. Brillinger, P. J. Diggle, E. Russek-Cohen, W. G. Warren, and R. D. Wolfinger, Lecture Notes in Statistics, No. 122, New York: Springer, pp. 289–298.

Matern, B., (1960), *Spatial Variation*. Meddelanden fran Statens Skogsforskningsinstitut, 49, No. 5. [Second edition (1986), Lecture Notes in Statistics, No. 36, New york: Springer.]

Myers, D. E., and Journel, A. G., (1990), "Variograms with zonal anisotropies and noninvertible kriging systems," *Mathematical Geology*, 22, pp. 779–785.

Rodriguez-Iturbe, I., and Mejia, J. M., (1974), "The design of rainfall networks in time and space," *Water Resources Research*, 10, pp. 713–729.

Rouhani, S., and Myers, D. E., (1990), "Problems in space-time kriging of geohydrological data," *Mathematical Geology*, 22, pp. 611–623.

Toutenburg, H., (1982), *Prior Information in Linear Models*, New york: Wiley.

Zimmerman, D. L., and Zimmerman, M. B., (1991), "A comparison of spatial semivariogram estimators and corresponding ordinary kriging predictors," *Technometrics*, 33, pp. 77-91.