Statistica Sinica **30** (2020), 719-739
doi:https://doi.org/10.5705/ss.202017.0385

NONSEPARABLE, SPACE-TIME COVARIANCE FUNCTIONS WITH DYNAMICAL COMPACT SUPPORTS

Emilio Porcu, Moreno Bevilacqua and Marc G. Genton

Trinity College Dublin, Universidad de Valparaiso and King Abdullah University of Science and Technology

Abstract: This study provides new classes of nonseparable space-time covariance functions with spatial (or temporal) margins that belong to the generalized Wendland class of compactly supported covariance functions. An interesting feature of our covariances, from a computational viewpoint, is that the compact support is a decreasing function of the temporal (spatial) lag. We provide conditions for the validity of the proposed class, and analyze the problem of differentiability at the origin for the temporal (spatial) margin. A simulation study explores the finite-sample properties and the computational burden associated with the maximum likelihood estimation of the covariance parameters. Finally, we apply the proposed covariance models to Irish wind speed data, and compare the results with those of Gneiting–Matérn models in terms of fitting, prediction efficiency, and computational burden. The necessary and sufficient conditions, together with other results on dynamically varying compact supports, are provided in the online Supplementary Material.

Key words and phrases: Generalized Wendland covariance function, geostatistics, kriging, random field, sparse matrices.

1. Introduction

There is growing interest in space-time modeling using covariance functions; refer to Gneiting (2002a), Stein (2005), Zastavnyi and Porcu (2011), Gneiting, Genton and Guttorp (2007), and Schlather (2010). Typically, data observed over space and time are modeled as a realization of a stationary Gaussian random field that has a covariance function that is spatially isotropic and temporally symmetric (Gneiting (2002a)). Specifically, for a stationary random field $Z(\mathbf{x}, t)$, with $\mathbf{x}$ a point of $\mathbb{R}^d$ and t denoting time, spatial isotropy is coupled with temporal symmetry through a continuous function, $\phi : [0, \infty) \times [0, \infty) \to \mathbb{R}$, such that

$$\operatorname{cov}\left\{Z(\mathbf{x}, t), Z(\mathbf{x} + \mathbf{h}, t + u)\right\} = \phi\left(\|\mathbf{h}\|, |u|\right), \tag{1.1}$$

where $(\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}$ denotes a space-time lag vector and $\phi(0, 0) = \sigma^2$ is the variance of Z. For the remainder of this paper, we use r for $\|\mathbf{h}\|$ and u for $|u|$.

In addition, the margins $\phi(r, 0)$ and $\phi(0, u)$ are called, respectively, spatial and temporal margins. A covariance function ϕ is called separable if ϕ factors into the product of a purely spatial and a purely temporal covariance function.

A popular example of nonseparable space-time covariance functions of the type in (1.1) is the Gneiting class (Gneiting (2002a); Zastavnyi and Porcu (2011)). We define it here as

$$\phi(r, u) = \frac{\sigma^2}{\psi(u^2)^{d/2}} g\left(\frac{r^2}{\psi(u^2)}\right), \qquad (r, u) \in [0, \infty) \times [0, \infty), \tag{1.2}$$

where g is completely monotonic on the positive real line, such that it is infinitely differentiable on $(0, \infty)$ and $(-1)^k g^{(k)}(t) \geq 0$, for $t \geq 0$. The function ψ is strictly positive and has a completely monotonic derivative. Additionally, with no loss of generality, we assume that $g(0) = \psi(0) = 1$, such that $\phi(0, 0) = \sigma^2$. Sufficient conditions for the validity of this class were provided by Gneiting (2002a). Then, Zastavnyi and Porcu (2011) found the necessary conditions, and relaxed the hypothesis on the function ψ. A subclass of the Gneiting class in (1.2) that has become especially popular is expressed as

$$\phi(r, u) = \frac{\sigma^2}{\psi(u^2)^{d/2}} \mathcal{M}_\mu\left(\frac{r}{\psi(u^2)}\right), \qquad (r, u) \in [0, \infty) \times [0, \infty), \tag{1.3}$$

where

$$\mathcal{M}_\mu(r) = \frac{2^{1-\mu}}{\Gamma(\mu)} r^\mu \mathcal{K}_\mu(r), \qquad r \geq 0,$$

with $\mu > 0$ and $\mathcal{K}_\mu$, a modified Bessel function of the second kind of order μ, is the so-called Matérn class (Stein (1999)). Hence, the class (1.3) is termed the Gneiting–Matérn class. The parameter μ characterizes the differentiability at the origin and, thus, the differentiability of the sample paths of a Gaussian field in $\mathbb{R}^d$ with a Matérn covariance function. In particular, for a positive integer k, the sample paths are k-times differentiable, in any direction, if and only if $\mu > k$.

A spatial covariance function is called compactly supported if it vanishes after a given spatial distance. There is a large body of literature on compactly supported covariance functions in many branches of probability theory, geostatistics, and approximation theory. For further details refer to Golubov (1981), Wendland (1995), Schaback and Wu (1995), Wu (1995), Buhmann (2000), Gneiting (2002b), Zastavnyi and Trigub (2002), Zastavnyi (2006), Schaback (2011), Zhu (2012), Hubbert (2012), Porcu and Zastavnyi (2014), and Chernih, Sloan and Womersley (2014), as well as to the more recent results in Bevilacqua et al. (2018) and the review by Porcu, Zastavnyi and Bevilacqua (2018).

Compactly supported covariance functions are used for computationally efficient spatial predictions (Furrer, Genton and Nychka (2006) and the references therein) and estimations (Kaufman, Schervish and Nychka (2008)) in the covariance tapering technique and for fast and exact simulations, as well as being appealing to practitioners (Gneiting (2002b)). The recent work by Bevilacqua et al. (2018) revealed importance of such functions for kriging predictions, showing that the generalized Wendland class (Zastavnyi and Trigub (2002); Gneiting (2002a)) is compatible with the Matérn class. This implies that, under fixed domain asymptotics, and under some specific conditions on the parameters indexing the covariance functions, the misspecified linear unbiased predictor from the generalized Wendland class is asymptotically as efficient as the true simple kriging predictor using a Matérn class. Thus, a kriging prediction can be performed using a compactly supported function, without any loss of asymptotic prediction efficiency.

The problem of constructing nonseparable compactly supported space-time covariance functions is almost unexplored. A mathematical formulation of the problem is provided by Zastavnyi and Porcu (2011), who suggested replacing the function g in (1.2) with another function that has compact support. However, they could not find a solution to the problem, a characterization of which remains elusive. The present study challenges this problem. Specifically, we show how to generate covariance functions of the type in (1.2) by replacing the function g with another function that has compact support. Further, we replace the Matérn function used in the Gneiting–Matérn class in (1.3) with generalized Wendland functions that are compactly supported. In addition, the latter functions have the same properties as the Matérn class in terms of differentiability at the origin (Bevilacqua et al. (2018)).

A simulation study explores the finite sample properties of the maximum likelihood (ML) estimation of the covariance parameters. Finally, we apply our models to Irish wind speed data, and compare the results with those of Gneiting models in terms of fitting, prediction efficiency based on predictive scores and computational burden.

The remainder of the paper proceeds as follows. Section 2.1 presents the necessary background and introduces the generalized Wendland class. Section 2.2 provides the results fot the proposed classes of space-time covariance functions. Section 2.3 discusses examples and parameterization. Section 2.4 provides conditions to improve the differentiability of the temporal margins of the proposed classes. Section 3 explores our findings using a simulation and real data. Section 4 concludes the paper.

The online Supplementary Material provides several more technical results: First, we generalize the results in Section 2 to wider classes of functions with compact support. Then, a Fourier analysis and completely monotone functions are used to explore the necessary and sufficient conditions. The online Supplementary Material also provides the figures referred in the paper.

2. Compactly Supported Space-time Covariance Functions

2.1. Background material

For neatness of exposition, several preliminaries are needed. The space-time covariance function in (1.1) is positive-semidefinite. That is, for any finite collection $\{(\mathbf{x}_k, t_k)\}_{k=1}^N \subset \mathbb{R}^d \times \mathbb{R}$, and for any system of constants $\{c_k\}_{k=1}^N \subset \mathbb{R}$, we have

$$\sum_{k=1}^N \sum_{h=1}^N c_k c_h \phi(\|\mathbf{x}_k - \mathbf{x}_h\|, |t_k - t_h|) \geq 0.$$

In what follows, we propose a class of candidate functions with compact support that can be used to replace the function g in (1.2), while preserving the positive-definiteness. We introduce the generalized Wendland class (Gneiting (2002b); Zastavnyi and Trigub (2002)) $\varphi_{\nu,\kappa} : [0,\infty) \to \mathbb{R}$, defined as

$$\varphi_{\nu,\kappa}(r) = \frac{1}{B(2\kappa+1,\nu)} \int_r^\infty (t^2 - r^2)^\kappa \varphi_{\nu-1,0}(t)\,\mathrm{d}t, \qquad r \geq 0, \tag{2.1}$$

where $\kappa > 0$, and B denotes the beta function,

$$B(2\kappa+1,\nu) = \frac{\Gamma(2\kappa+1)\Gamma(\nu)}{\Gamma(2\kappa+\nu+1)}.$$

Here, $\varphi_{\nu,0}$ denotes the Askey family of functions (Askey (1973)), defined by

$$\varphi_{\nu,0}(r) := (1-r)_+^\nu, \qquad \nu > 0, \tag{2.2}$$

where $(\cdot)_+$ denotes the positive part. Let d be a positive integer. The function $\varphi_{\nu,0}(r)$ is positive-definite in $\mathbb{R}^d$ if and only if $\nu \geq (d+1)/2$ (Golubov (1981)). Zastavnyi and Trigub (2002) show that $\varphi_{\nu,\kappa}$ is positive-definite in $\mathbb{R}^d$ if and only if $\nu \geq (d+1)/2 + \kappa$. Additionally, $\varphi_{\nu,\kappa}(\cdot/b)$ is compactly supported over the ball of $\mathbb{R}^d$, with radius $b > 0$. Closed-form solutions of the integral in (2.1) can be obtained when $\kappa = k$, a nonnegative integer. In this case,

$$\varphi_{\nu,k}(r) = \varphi_{\nu+k,0}(r) P_k(r), \qquad r \geq 0,$$

where P_k is a polynomial of order k. See the first column of Table 1 for examples with $k = 0, 1, 2, 3$. These functions, termed (original) Wendland functions, were

Table 1. Generalized Wendland correlation $\varphi_{\nu,\kappa}(r)$ and Matérn correlation $\mathcal{M}_\mu(r)$ with increasing smoothness parameters κ and μ. $SP(k)$ means that the sample paths of the associated Gaussian field are k-times differentiable. Taken from Bevilacqua et al. (2018).

κ	$\varphi_{\nu,\kappa}(r)$	μ	$\mathcal{M}_\mu(r)$	$SP(k)$
0	$(1-r)_+^{\nu}$	0.5	e^{-r}	0
1	$(1-r)_+^{\nu+1}(1+r(\nu+1))$	1.5	$e^{-r}(1+r)$	1
2	$(1-r)_+^{\nu+2}(1+r(\nu+2)$ $+r^2(\nu^2+4\nu+3)(1/3))$	2.5	$e^{-r}(1+r+r^2/3)$	2
3	$(1-r)_+^{\mu+3}\big(1+r(\nu+3)$ $+r^2(2\nu^2+12\nu+15)(1/5)$ $+r^3(\nu^3+9\nu^2+23\nu+15)(1/15)\big)$	3.5	$e^{-r}(1+r/2+r^2(6/15)$ $+r^3/15)$	3

originally proposed by Wendland (1995). Other closed-form solutions of the integral (2.1) can be obtained when $\kappa = k + 1/2$, using the results in Schaback (2011). Hubbert (2012) showed some other closed-forms based on hypergeometric functions. Finally, Chernih, Sloan and Womersley (2014) showed that, for κ tending to infinity, a rescaled version of the model in (2.1) converges to a Gaussian model. As noted by Gneiting (2002b), the generalized Wendland and Matérn models exhibit the same behavior at the origin when the smoothness parameters of the two covariance models are related by the equation $\nu = \kappa + 1/2$. This fact is depicted in Table 1, where specific cases of Wendland functions are compared with the Matérn covariance in terms of the sample path differentiability of the associated Gaussian random field. Generalized Wendland functions include many other popular classes of covariance functions with compact support; for a recent review refer to Porcu, Zastavnyi and Bevilacqua (2018).

We finish this section with a new definition that leads to the results provided in the subsequent section. Let ϕ be a space-time covariance function defined in (1.1). We call a temporally dynamical radius, ψ, the continuous mapping from $[0,\infty)$ to $(0,\infty)$, such that, for each $u_o \in [0,\infty)$, the margin $\phi(\cdot, u_o)$ is compactly supported on the interval $[0, \psi(u_o))$. Clearly, both Askey and generalized Wendland classes are special cases of dynamical compact support, when $\psi \equiv b > 0$ is the constant function.

2.2. Space-time Gneiting–Wendland functions with dynamical compact support

The following results are based on a constructive criterion provided by Porcu and Zastavnyi (2012).

Lemma 1. *Let d be a positive integer. Let $(\Omega, \mathcal{F}, P)$ be a measure space with P a positive measure. Let $H(\cdot;\cdot) : \Omega \times [0,\infty) \to \mathbb{R}$ and $F(\cdot;\cdot) : [0,\infty) \times \Omega \to \mathbb{R}$, such that*

1. *$H(\xi;\cdot)$ is a temporal covariance function, for all $\xi \in \Omega$;*
2. *$F(\cdot;\xi)$ is an isotropic spatial covariance function in $\mathbb{R}^d$, for all $\xi \in \Omega$;*
3. *$H(\cdot;u)F(r;\cdot) \in L_1(\Omega, \mathcal{F}, P)$, for any $r, u \geq 0$.*

Then, the mapping

$$\phi(r,u) = \sigma^2 \frac{\int_\Omega F(r;\xi)H(\xi;u)P(\mathrm{d}\xi)}{\int_\Omega F(0;\xi)H(\xi;0)P(\mathrm{d}\xi)}, \qquad (r,u) \in [0,\infty) \times [0,\infty), \tag{2.3}$$

with $\sigma^2 > 0$, defines a space-time covariance function in $\mathbb{R}^d \times \mathbb{R}$ that is isotropic in the spatial argument and symmetric in time.

An intuitive way to understand the formal statement in Lemma 1 is to consider the integral in (2.3) as a scale mixture of a spatial and a temporal covariance. Conditions 1 and 2 are needed for well-defined spatial and temporal covariances. Condition 3 ensures the integral in (2.3) is well defined. Note that the denominator in (2.3) is a normalization constant, such that $\phi(0,0) = \sigma^2$.

Theorem 1. *Let d be a positive integer. Let $\varphi_{\nu,0}$ be the Askey function in (2.2). Let ψ be a continuous and positive function on the positive real line, with $\psi(0) = 1$ and such that $1/\psi$ is increasing and concave on the positive real line, with $\lim_{t\to\infty} \psi(t) = 0$. Then, the mapping*

$$\phi(r,u) = \psi(u)^\alpha \varphi_{\nu,0}\left(\frac{r}{\psi(u)}\right), \qquad (r,u) \in [0,\infty) \times [0,\infty), \tag{2.4}$$

defines a space-time covariance function in $\mathbb{R}^d \times \mathbb{R}$, provided that $\nu \geq (d+5)/2$ and $\alpha \geq (d+3)/2$.

Proof. The proof is an application of Lemma 1. Specifically, we use the scale mixture argument of (2.3) under appropriate choices of the functions H and F.

We now proceed formally and consider the mapping $F(r;\xi) = \varphi_{n,0}(r/\xi)$. Golubov (1981) shows that $F(\cdot;\xi)$ is an isotropic spatial covariance function in $\mathbb{R}^d$, for any $\xi > 0$, provided that $n \geq (d+1)/2$. Thus, Condition 2 in Lemma 1 is satisfied. For the choice of the function H, we consider the mapping

$$H(\xi;u) = H_{n,\gamma}(\xi;u) = \xi^n \varphi_{\gamma,0}\left(1 - \frac{\xi}{\psi(u)}\right)_+, \quad \xi > 0, u \geq 0, \quad \gamma \geq 1, n > 0.$$

Given the properties of the function ψ, we have that $H_{n,\gamma}$ is positive, decreasing and convex, with $\lim_{t\to\infty} H_{n,\gamma}(\xi;t) = 0$, for any $\xi > 0$. Thus, we can invoke the Pólya criterion (Pólya (1949)) to show that $H_{n,\gamma}(\xi;u)$ is a covariance function in $\mathbb{R}$. Therefore, Condition 1 in Lemma 1 is satisfied. Finally, note that Condition 3 of Lemma 1 holds trivially. Thus, we can now apply the scale mixture in (2.3), with $\Omega = [0,\infty)$ and P the Lebesgue measure:

$$\begin{aligned}\phi(r,u) &= \int_{(0,\infty)} F(r;\xi)H(\xi;u)\mathrm{d}\xi \\ &= \int_{(0,\infty)} \Big(1-\frac{r}{\xi}\Big)_+^n \xi^n \Big(1-\frac{\xi}{\psi(u)}\Big)_+^\gamma \mathrm{d}\xi \\ &= \frac{1}{\psi(u)^\gamma}\int_r^{\psi(u)} (\xi-r)^n(\psi(u)-\xi)^\gamma\,\mathrm{d}\xi \\ &= \frac{1}{\psi(u)^\gamma}\int_0^{\psi(u)-r} t^n(\psi(u)-r-t)^\gamma\,\mathrm{d}t \\ &= \frac{1}{\psi(u)^\gamma}\int_0^1 (\psi(u)-r)^{n+\gamma+1}v^n(1-v)^\gamma\,\mathrm{d}v \\ &= B(n+1,\gamma+1)\psi(u)^{n+1}\left(1-\frac{r}{\psi(u)}\right)^{n+\gamma+1} \\ &= B(n+1,\gamma+1)\psi(u)^{n+1}\,\varphi_{n+\gamma+1,0}\left(\frac{r}{\psi(u)}\right), \end{aligned} \tag{2.5}$$

where B denotes the beta function. The third line in the chain of equalities is justified by the fact that, by definition, ϕ is identically equal to zero whenever $r > \psi(u)$. We now let $\alpha = n+1$ and $\nu = n+\gamma+1$. Thus, (2.4) and (2.5) agree modulo a positive factor, that is the normalization constant. This fact completes the proof. The conditions on α and ν are easily verified from the previous identities.

Theorem 2. *Let d be a positive integer and $\kappa > 0$. Let $\varphi_{\nu,\kappa}$ be the generalized Wendland class of functions in (2.1). Let ψ be a continuous and positive function on the positive real line, with $\psi(0) = 1$ and such that $1/\psi(\cdot)$ is increasing and concave on the positive real line, with $\lim_{t\to\infty}\psi(t) = 0$. Then, the mapping ϕ, defined as*

$$\phi(r,u) = \psi(u)^\alpha \varphi_{\nu,\kappa}\left(\frac{r}{\psi(u)}\right), \qquad (r,u) \in [0,\infty)\times[0,\infty), \tag{2.6}$$

defines a space-time covariance function in $\mathbb{R}^d\times\mathbb{R}$, provided that $\nu \geq (d+5)/2+\kappa$ and $\alpha \geq (d+3)/2+2\kappa$.

Proof. We give a constructive proof by applying Lemma 1 for specific choices of the functions H and F in the scale mixture (2.3). For the choice of the function F, let $\kappa > 0$, $n \geq (d+1)/2 + \kappa$, and $F(r;\xi) = \varphi_{n,\kappa}(r/\xi)$, with $\varphi_{n,\kappa}$ defined as in (2.1). Clearly, Condition 1 in Lemma 1 is satisfied. Furthermore, Zastavnyi and Trigub (2002) show that $\varphi_{n,\kappa}$ can be rewritten as:

$$\varphi_{n,\kappa}(w) = \frac{1}{B(n, 2\kappa+1)} \int_w^1 (1-t)^{n-1}(t^2 - w^2)^\kappa \, \mathrm{d}t\,, \qquad w \geq 0.$$

In particular, following Daley, Porcu and Bevilacqua (2015), we have that, for $0 < y < \xi \leq 1$,

$$\begin{aligned}\varphi_{n,\kappa}\Big(\frac{y}{\xi}\Big) &= \frac{1}{B(n,2\kappa+1)} \int_{y/\xi}^1 (1-t)^{n-1}\Big(t^2 - \frac{y^2}{\xi^2}\Big)^\kappa \mathrm{d}t \\ &= \frac{1}{B(n,2\kappa+1)} \int_y^\xi \Big(1 - \frac{v}{\xi}\Big)^{n-1}(v^2-y^2)^\kappa \,\frac{\mathrm{d}v}{\xi^{2\kappa+1}}\,.\end{aligned}$$

We now choose the function

$$H(\xi;u) = H_{n,\kappa,\gamma}(\xi;u) = \xi^{n+2\kappa}\left(1 - \frac{\xi}{\psi(u)}\right)_+^\gamma, \quad \xi > 0, u \geq 0, \quad \gamma \geq 1, n > 0,$$

with κ positive and ψ as stated. Again, it is easy to show that $H_{n,\kappa,\gamma}(\xi;\cdot)$ is positive, decreasing, and convex, with $\lim_{t\to\infty} H_{n,\kappa,\gamma}(\xi;t) = 0$, for all $\xi > 0$. Thus, Condition 2 of Lemma 1 is satisfied. Condition 3 holds trivially. We can thus apply the scale mixture argument in (2.3), with $\Omega = [0,\infty)$ and P being the Lebesgue measure. We write ψ for $\psi(u)$, and have

$$\begin{aligned}&\int_0^\infty \varphi_{n,\kappa}\Big(\frac{r}{\xi}\Big) H_{n,\kappa,\gamma}(\xi;u)\mathrm{d}\xi = \int_r^\psi \varphi_{n,\kappa}\Big(\frac{r}{\xi}\Big)\xi^{n+2\kappa}\left(1-\frac{\xi}{\psi}\right)^\gamma \mathrm{d}\xi \\ &= \frac{1}{B(n,2\kappa+1)}\int_r^\psi \xi^{n+2\kappa}\left(1-\frac{\xi}{\psi}\right)^\gamma \mathrm{d}\xi \int_r^\xi \left(1-\frac{v}{\xi}\right)^{n-1}(v^2-r^2)^\kappa\,\frac{\mathrm{d}v}{\xi^{2\kappa+1}} \\ &= \frac{\psi^{-\gamma}}{B(n,2\kappa+1)}\int_r^\psi (v^2-r^2)^\kappa\,\mathrm{d}v \int_v^\psi (\psi-\xi)^\gamma(\xi-v)^{n-1}\,\mathrm{d}\xi \\ &= \frac{\psi^{-\gamma}}{B(n,2\kappa+1)} B(n,\gamma+1)\int_r^\psi (\psi-v)^{n+\gamma}(v^2-r^2)^\kappa\,\mathrm{d}v \\ &= \psi^{n+2\kappa+1}\frac{B(n,\gamma+1)}{B(n,2\kappa+1)}\int_{r/\psi}^1 (1-t)^{n+\gamma}\left(t^2-\frac{r^2}{\psi^2}\right)^\kappa \mathrm{d}t \\ &= \frac{B(n,\gamma+1)}{B(n,2\kappa+1)}\,B(n+\gamma+1,2\kappa+1)\ \psi^{n+2\kappa+1}\varphi_{n+\gamma+1,\kappa}\Big(\frac{r}{\psi}\Big) \\ &= B(n+2\kappa+1,\gamma+1)\ \psi^{n+2\kappa+1}\varphi_{n+\gamma+1,\kappa}\Big(\frac{r}{\psi}\Big),\end{aligned}$$

where B is the beta function, as before. We now let $\nu = n + \gamma + 1$ and $\alpha =$

$n+2\kappa+1$. Rescaling at the origin and using the same arguments as in Theorem 1, we easily arrive at the assertion.

Using the arguments in Gneiting (2002b), it can be shown that for any increasing sequence $\{c_n\}_{n\geq 0}$, we have $\varphi_{c_n,\kappa}\{r/(\psi(u)c_n)\} \to \mathcal{M}_{1/2+\kappa}\{r/\psi(u)\}$, with the convergence being uniform on any bounded set. Thus, the class in (2.6) converges to the Gneiting–Matérn class, and when $u=0$, the smoothness parameters of the two covariance models are related by the equation $\mu=\kappa+1/2$ (see Table 1).

2.3. Examples and parameterizations

Several examples from the mappings ψ that satisfy the requirements in Theorems 1 and 2 can be found in Table 1 in Porcu and Schilling (2011). A notable example comes from the choice

$$\psi(t;\delta,\beta) = (1+t^{\delta})^{-\beta/\delta}, \qquad t \geq 0, \tag{2.7}$$

for $0<\delta\leq 1$ and $0\leq\beta\leq\delta$. In particular, in the following sections, we work with the special case $\psi(\cdot;\beta) := \psi(\cdot;1,\beta)$, valid for $\beta\in[0,1]$.

For $\kappa = k$, a nonnegative integer, we find that the classes in (2.6) can be written as

$$\phi(r,u) = \psi(u)^{\alpha}\varphi_{\nu+k,0}\left(\frac{r}{\psi(u)}\right)P_k\left(\frac{r}{\psi(u)}\right), \qquad r,u\geq 0,$$

where the constraints on α and ν are specified in Theorems 1 and 2, and P_k is a polynomial of degree k. In particular, we use $k=0,1,2$ for ease of illustration. Using the first three entries in Table 1 together with (2.4) and (2.6), we obtain

$$\begin{aligned}
\phi(r,u) &= \psi(u)^{\alpha}\left(1-\frac{r}{\psi(u)}\right)_+^{\nu},\\
\phi(r,u) &= \psi(u)^{\alpha}\left(1-\frac{r}{\psi(u)}\right)_+^{\nu+1}\left(1+(\nu+1)\frac{r}{\psi(u)}\right),\\
\phi(r,u) &= \psi(u)^{\alpha}\left(1-\frac{r}{\psi(u)}\right)_+^{\nu+2}\left(1+(\nu+2)\frac{r}{\psi(u)}+\frac{1}{3}\left((\nu+2)^2-1\right)\left(\frac{r}{\psi(u)}\right)^2\right),
\end{aligned}\tag{2.8}$$

where α and ν must be determined according to Theorems 1 (for the first example) and 2 (for the other two examples). For geostatistical applications, it is useful to consider rescaled versions $\phi(r/b,u/a)$, because these enable us to consider the marginal spatial compact support $b>0$, the dynamical compact support $b\psi(u/a)$, and the temporal scale parameter $a>0$. In many instances, a reparameterization of the proposed covariance models is useful. For instance, using (2.7) in the construction (2.8), and replacing $\beta\alpha$ with $\tau > 0$, we obtain the space-time correlation functions:

$$\phi(r,u) = \frac{1}{(1+u/a)^{\tau}}\varphi_{\nu,k}\left(\frac{r}{b(1+u/a)^{-\beta}}\right), \qquad r,u \geq 0, \tag{2.9}$$

where $\tau \geq 2.5 + 2k$ and $k = 0, 1, 2$. If we fix τ, we obtain a parametric family with an easily interpretable space-time nonseparable parameter $0 \leq \beta \leq 1$, which includes as special case a separable covariance, obtained when $\beta = 0$.

Figure S1 (see the Supplementary Material) shows a contour plot of the nonseparable space-time correlation functions with dynamical compact support in (2.9). Specifically, we fix $b = 0.15$, $a = 0.2$, $\nu = 3.5 + \kappa$, and $\tau = 2.5 + 2\kappa$, for $\kappa = 0, 1$ and $\beta = 0, 0.5, 1$. As β increases, the rate of decay of the dynamical compact support becomes more severe. Therefore, this parameter affects the dynamical compact support, that is, the sparsness of the associated correlation matrix.

Figure S2 (see the Supplementary Material) shows a simulation on a regular grid of 12,544 sites over a unit square and over temporal instants $u = 1, 1.5$, obtained using a Cholesky decomposition, of a space-time Gaussian field with correlation (2.9) (top), fixing $\kappa = 1$, $\tau = 6.5$, $\nu = 4.5$, $b = 0.15$, $a = 0.2$, and $\beta = 0.5$. The same figure depicts a realization of a space-time Gaussian random field with a covariance function from the Gneiting–Matérn class:

$$\phi(r,u) = \frac{1}{(1+u/0.2)^{6.5}}\mathcal{M}_{1.5}\left(\frac{r}{0.0226(1+u/0.2)^{1/4}}\right), \qquad r,u \geq 0. \tag{2.10}$$

The two simulations share the same Gaussian realization after using the Cholesky decomposition method. The two covariance models have the same marginal temporal correlation and the spatial scale parameter in the Gneiting–Matérn model is chosen such that the marginal spatial correlation is lower than 0.01 when $r > 0.15$; that is, it is greater than the marginal compact support of the generalized Wendland model. It is apparent from Figure S2 that the two simulations look very similar.

Remark 1. The members of the classes in Theorems 1 and 2 are dynamically compactly supported in space. Thus, they are computationally suitable covariance models for space-time data with a relatively large number of location sites with respect to the temporal instants.

Note that the constructions in Theorems 1 and 2 can be interchanged, yielding space-time covariances that are compactly supported over time, and have a compact support that evolves dynamically with spatial distance. We omit such a specification of the mathematical conditions because the analogue specification is literal. Then, for instance, an analogue version of the model in (2.9) is

$$\phi(r,u) = \frac{1}{(1+r/b)^{\tau}} \varphi_{\nu,\kappa}\left(\frac{u}{a(1+r/b)^{-\beta}}\right), \qquad r,u \geq 0. \tag{2.11}$$

In this model, the parameter a is the marginal temporal compact support, and the decreasing dynamical compact support is given by $a\psi(r/b)$.

This kind of model is computationally more suitable for space-time data with a relatively large number of temporal instants with respect to the sites, as in the Irish wind speed data in Section 3.2.

2.4. Improving temporal differentiability at the origin

The ψ functions used for Theorems 1 and 2 are, by construction, nondifferentiable at the origin. This implies that we can govern the degree of differentiability in the spatial component, but not in the temporal one. This issue is studied in detail in the Supplementary Material, where we show the necessary conditions using a Fourier analysis that preserves the positive-definiteness of the constructions proposed in Theorems and 1 and 2.

Having a model that allows for different degrees of temporal differentiability at the origin is important for attaining greater flexibility in analyses of space-time data sets. In addition, differentiability at the origin has a crucial impact on spatial and temporal predictions (Stein (1999)). Because we are approximating the Gneiting–Matérn class with a compactly supported structure, it is important that we attain the same level of differentiability for both spatial and temporal margins.

The sufficient conditions that allow us to improve the differentiability of the temporal margin can be improved upon, based on the following facts. The function

$$\varpi_{\tau,\lambda}(r) := \varphi_{\tau,0}(r^{\lambda}) = (1-r^{\lambda})_{+}^{\tau}, \qquad r \geq 0, \quad \lambda \in (0,2), \quad \tau > 0, \tag{2.12}$$

has attracted the interest of several mathematicians in the past; refer to Gneiting (2001), and the references therein. In particular, the univariate case, $d = 1$, has an interesting history; again refer to Gneiting (2000). We have that $\varpi_{\nu,2}$ is not positive-definite on $\mathbb{R}$, regardless of the value of ν. Kuttner (1944) showed that there exists a function $\kappa_1(\lambda)$, for $\lambda \in (0,2)$, such that $\varpi_{\tau,\lambda}(r)$ is positive-definite on $\mathbb{R}$ if and only if $\tau > \kappa_1(\lambda)$. The function $\kappa_1(\lambda)$ is continuous and strictly increasing, with $\lim_{\lambda\to 0}\kappa_1(\lambda) > 0$, $\kappa_1(1) = 1$, $\lim_{\lambda\to 2}\kappa_1(\lambda) = \infty$, and $\kappa_1(\lambda) > \lambda$ if $\lambda \neq 1$.

Table 2. Lower bounds for $\kappa_1(\lambda)$ for given values of λ. Taken from Gneiting (2000).

λ	1.05	1.15	1.25	1.45	1.55	1.75	1.95
$\kappa_1(\lambda)$	1.0507	1.1572	1.2706	1.5247	1.7234	2.3462	3.9084

We now apply our results to Theorem 2 and consider the function

$$\phi(r,u) = (1+u^{\lambda})^{-\alpha}\varphi_{\nu(\tau),\kappa}\left(\frac{r}{(1+u^{\lambda})}\right), \qquad (r,u) \in [0,\infty) \times [0,\infty),$$

where ν is a function of τ, as described in (2.12). The same scale mixture arguments as in the proof of Theorem 2 apply (see Lemma 1); hence we omit them here. We have that, for a given $d \in \mathbb{N}$, ϕ is positive-definite on $\mathbb{R}^d \times \mathbb{R}$, provided that $\alpha \geq (d+3)/2$ and

$$\nu \geq \frac{(d+3)}{2} + \kappa + \tau, \qquad \tau \geq \kappa_1(\lambda), \quad \lambda \in (0,2).$$

Table 2, taken from Gneiting (2000), allows to obtain the corresponding values for a given $\lambda \in (0,2)$.

3. Numerical Results

We start by describing the performance of the ML estimation of the parameters of the Gneiting–Wendland model. Then, we compare the Gneiting–Matérn model with the proposed Gneiting–Wendland model from a modeling, prediction performance, and computational point of view when used as space-time covariance models for the Irish wind speed data.

3.1. Simulation studies

Following Remark 1, we consider two possible scenarios:

1. A data set with many spatial sites and few temporal observations. Specifically, we have $\mathbf{x}_i$, for $i = 1, 2, \ldots, 60$ sites, uniformly distributed on the unit square, and $u = 0, 0.25, \ldots, 2.25$ temporal instants;

2. A data set with few spatial sites and many temporal observations, that is $\mathbf{x}_i$, for $i = 1, 2, \ldots, 10$ sites, uniformly distributed on the unit square, and $u = 0, 0.25, \ldots, 14.75$ temporal instants.

For both scenarios, the total number of observations is kept relatively small (600 observations) in order to make the ML estimation feasible. Under Scenario 1, we simulate 1,000 zero-mean space-time Gaussian random fields, with covariance given by (2.9), setting $k = 0, 1, 2$ in order to consider different levels of differen-

Table 3. Top: Bias and standard deviation (SD) for the ML estimation of the spatial and temporal scales and variance for the Gneiting–Wendland model in Equation (2.9), for $\kappa = 0, 1, 2$ and $\beta = 0, 0.5, 1$ under Scenario 1. Bottom: Scenario 2.

		a		b		σ^2	
κ	β	bias	SD	bias	SD	bias	SD
0	0	−0.00016	0.02168	0.00076	0.04889	0.00024	0.06181
	0.5	−0.00018	0.02145	0.00092	0.04796	0.00021	0.06164
	1	−0.00017	0.02121	0.00105	0.04690	0.00019	0.06156
1	0	0.00036	0.01342	−0.01943	0.12685	0.00027	0.06156
	0.5	0.00038	0.01342	−0.01841	0.12194	0.00028	0.06156
	1	0.00042	0.01342	−0.01596	0.11485	0.00034	0.06156
2	0	0.00054	0.01225	0.01040	0.18746	0.00034	0.06132
	0.5	0.00053	0.01225	0.00910	0.18185	0.00035	0.06132
	1	0.00054	0.01225	0.00636	0.17438	0.00037	0.06140
0	0	−0.00228	0.05000	0.00209	0.06812	0.00038	0.06419
	0.5	−0.00158	0.04743	0.00251	0.06745	0.00043	0.06411
	1	−0.00140	0.04506	0.00257	0.06626	0.00037	0.00409
1	0	−0.01809	0.11091	0.00064	0.03768	0.00084	0.06395
	0.5	−0.01763	0.10266	0.00086	0.03768	0.00040	0.06496
	1	−0.01741	0.09644	0.00103	0.03782	0.00095	0.06496
2	0	−0.01131	0.16199	0.00021	0.03674	0.00040	0.06380
	0.5	−0.01388	0.15556	0.00033	0.03688	0.00095	0.06372
	1	−0.01692	0.15063	0.00040	0.03688	0.00098	0.06372

tiability in the spatial covariance margin. Then, following Theorems 2.1 and 2.2, we fix $\tau = 2.5 + 2\kappa$ and $\nu = 3.5 + \kappa$.

We set $\sigma^2 = 1$, $b = 0.15$, and $a = 0.2$ and fix $\beta = 0, 0.5, 1$. For each simulation, we use the ML to estimate the parameters σ^2, a, and b. Table 3 reports the bias and variance associated with the ML estimations of σ^2, a, and b, for $k = 0.1, 2$ and $\beta = 0, 0.5, 1$.

Similarly, under Scenario 2, we simulate 1,000 zero-mean space-time Gaussian random fields, with covariance given by (2.11), with $k = 0, 1, 2$, fixing $\tau = 2.5 + 2k$ and $\nu = 3.5 + k$. We set $\sigma^2 = 1$, $a = 0.75$, and $b = 0.2$ and consider $\beta = 0, 0.5, 1$. For each simulation we use the ML to estimate the parameters σ^2, a, and b. The spatial and temporal scale parameters in both scenarios are chosen to attain a small dependence in space and time. Table 3 reports the bias and standard deviation (SD) associated with the ML estimations of σ^2, a, and b, for $k = 0, 1, 2$ and $\beta = 0, 0.5, 1$. Overall, the bias is negligible, and increasing β does not affect the bias or the SD of the ML estimation. Under Scenario 1, the SD of the spatial marginal compact support b increases considerably with k.

Similarly, under Scenario 2, the SD of the temporal marginal compact support a increases with k.

The bottleneck when evaluating a Gaussian likelihood is the computation of the inverse and the determinant of the covariance matrix, both of which can be obtained from its Cholesky decomposition. Some computational gains can be achieved using specific algorithms for sparse matrices in our models. The sparsity of the covariance matrix changes at each iteration of the maximization algorithm. In our implementation, a Gaussian likelihood optimization is performed by exploiting algorithms for sparse matrices, as implemented in the R package *spam* (Furrer and Sain (2010)) using the maximization algorithm of Nelder and Mead (1965), and implemented in the *optim* function of the R package (R Development Core Team (2016)).

Substantial further computational gains are achieved when performing a kriging prediction, because in this case, the sparsity of the covariance matrix is fixed. More details are given in the next section.

3.2. Irish wind speed data

The main goal of this section is to compare the Gneiting–Matérn model with the proposed Gneiting–Wendland model from a modeling, prediction performance, and computational point of view. To do so, we employ the models as space-time covariance models which we apply to Irish wind speed data (Haslett and Raftery (1989)).

We consider daily wind speeds collected over 18 years, from 1961 to 1978, at 12 sites in Ireland. Following Gneiting, Genton and Guttorp (2007), we omit the Rosslare station, consider a square root transformation of the data, and remove the seasonal component. The latter is estimated by calculating the average of the square roots of the daily means over all years and stations, and regressing this on a set of annual harmonics. The resulting transformed data, $\{z(\mathbf{x}_i, t_j)\}$, for $i = 1, \ldots, 11, j = 1, \ldots, 6{,}574$, are assumed to be a realization from a zero-mean space-time Gaussian random field. Because we perform an ML estimation, we focus on a subset of the data for computational reasons. Specifically, we focus on $\boldsymbol{z} = \{z(\mathbf{x}_i, t_j)\}$, for $i = 1, \ldots, 11, j = 366, \ldots, 910$. Thus, we have $11 \times 545 = 5{,}995$ observations, and an ML estimation is still feasible.

Figure S3 (Supplementary Material) shows that the empirical temporal marginal semivariogram attains the sill at a temporal distance of approximately three days. Thus, following Remark 1, a nonseparable, temporally compactly supported covariance model, as defined in (2.11), seems to be a natural choice for this kind of data. We compare the following space-time covariance models: the Gneiting–Matérn model

$$C_M(r,u;\boldsymbol{\theta}_M)=\frac{\sigma_M^2}{\psi(r/a_M)^{\tau_M}}\mathcal{M}_\mu\left(\frac{u}{b_M\psi(r/a_M)^{\beta_M/2}}\right),\quad \mu=0.5,1.5,2.5,\quad (3.1)$$

and our Gneiting–Wendland model

$$C_W(r,u;\boldsymbol{\theta}_W)=\frac{\sigma_W^2}{\psi(r/a_W)^{\tau_W}}\varphi_{\nu,k}\left(\frac{u}{b_W\psi(r/a_W)^{-\beta_W}}\right),\qquad k=0,1,2,\qquad (3.2)$$

where $\psi(r)=1+r$, for $r\geq 0$, and $\boldsymbol{\theta}_M=(\sigma_M^2,a_M,b_M,\beta_M)^\top$ and $\boldsymbol{\theta}_W=(\sigma_W^2,a_W,b_W)^\top$.

For the model in (3.2), based on the choices of $k=0,1,2$, we fix $\tau_W=2.5+2k$ and $\nu=3.5+k$, according to Theorem 2, such that positive-definiteness is attained. Then, for each k, we deliberately choose β_W equal to $0,0.5,1$ in order to increase the sparsity of the associated covariance matrix. Then we estimate $\boldsymbol{\theta}_W$ using the ML. Similarly, for model (3.1) we consider the cases $\mu=0.5,1.5,2.5$ fixing $\tau_M=2.5+2(\mu-0.5)$ and we estimate $\boldsymbol{\theta}_M$ using ML.

This setting makes the models defined in (3.1) and (3.2) comparable, because they share the same spatial margin. In addition, the temporal margins are of the Matérn and generalized Wendland types respectively, with the same level of differentiability at the origin for $\mu=0.5,1.5,2.5$ and $k=0,1,2$, respectively.

Table 4 (top) reports the ML estimation of $\boldsymbol{\theta}_W$ for each $k=0,1,2$ and for each $\beta_W=0,0.5,1$, with the associated loglikelihood. Table 4 (bottom) reports the ML estimation of $\boldsymbol{\theta}_M$ for each $\mu=0.5,1.5,2.5$, with the associated loglikelihood.

A comparison of the two models in terms of the loglikelihood shows that the best models are obtained when $k=0$ and $\mu=0.5$, that is, when the temporal margin is not differentiable at the origin for both cases. For the Gneiting–Wendland model, the best fit is obtained for $\beta_W=0$. Increasing this parameter leads to a small loss in terms of fitting and, at the same time, a decreasing number of nonzero values in the associated covariance matrix. Overall, the estimations of the spatial scale and the variance parameters are very similar, as expected, for $k=0,1,2$ and $\mu=0.5,1.5,2.5$. A graphical comparison between the empirical and estimated temporal semivariograms using model (3.1), when $\mu=0.5$ and model (3.2), when $k=0$ and $\beta_W=0$, is provided in Figure S3 in the Supplementary Material.

In order to compare the covariance models in (3.1) and (3.2) from prediction performance and computational viewpoints, we use three predictive scores, as described in Gneiting and Raftery (2007) and Zhang and Wang (2010). Let $\hat{Z}(\mathbf{x}_i, t_j)$ be the best linear prediction of Z at the space-time location $(\mathbf{x}_i, t_j)$, based on all data except $z(\mathbf{x}_i, t_j)$. The first prediction score is the root mean squared error (RMSE), defined as

$$\text{RMSE} = \left[\frac{1}{545 \times 11} \sum_{i=1}^{11} \sum_{j=366}^{910} \Big(z(\mathbf{x}_i, t_j) - \hat{Z}(\mathbf{x}_i, t_j) \Big)^2 \right]^{1/2}. \tag{3.3}$$

The logarithmic score is defined as

$$\log S = \frac{1}{545 \times 11} \sum_{i=1}^{11} \sum_{j=366}^{910} \left[\frac{1}{2} \log(2\pi\sigma(\mathbf{x}_i, t_j)) + \frac{1}{2} (Y(\mathbf{x}_i, t_j))^2 \right], \tag{3.4}$$

where $Y(\mathbf{x}_i, t_j) = (z(\mathbf{x}_i, t_j) - \hat{Z}(\mathbf{x}_i, t_j))/\sigma(\mathbf{x}_i, t_j)$, and $\{\sigma(\mathbf{x}_i, t_j)\}^2$ is the prediction variance associated with $\hat{Z}(\mathbf{x}_i, t_j)$. Finally, we consider the continuous ranked probability score (CRPS), which can be expressed in the Gaussian case as

$$\begin{aligned} \text{CRPS} = \frac{1}{545 \times 11} \sum_{i=1}^{11} \sum_{j=366}^{910} \sigma(\mathbf{x}_i, t_j) \Bigg(& Y(\mathbf{x}_i, t_j) \left[2F\{Y(\mathbf{x}_i, t_j)\} - 1 \right] \\ & + 2F\{Y(\mathbf{x}_i, t_j)\} - \frac{1}{\sqrt{\pi}} \Bigg), \end{aligned} \tag{3.5}$$

where F is the Gaussian cumulative distribution. In Table 4, the RMSE, logS, and CRPS are shown for each covariance model. Comparing the covariances in (3.1) and (3.2), for $\mu = 0.5, 1.5, 2.5$ and $k = 0, 1, 2$, respectively, we find a very small loss of prediction efficiency for the compactly supported models. For instance, when $\mu = 0.5$ and $k = 0$ and $\beta_W = 0$, the associated RMSE is 0.2174 and 0.2198, respectively.

The three prediction scores can be computed efficiently without calculating all of the drop-one predictions in (3.3), (3.4), and (3.5) (Zhang and Wang (2010)). This efficient computation depends on the inverse of the covariance matrix. Let $\Sigma(\hat{\boldsymbol{\theta}})$ be the estimated covariance matrix associated with one of the covariance models considered in (3.1) or (3.2). Then, for instance, the RMSE can be written as RMSE$= (\boldsymbol{f}^\top \boldsymbol{f}/5995)^{1/2}$, where $\boldsymbol{f} = D\Sigma(\hat{\boldsymbol{\theta}})^{-1}\boldsymbol{z}$ and $D = (\text{diag}(\Sigma(\hat{\boldsymbol{\theta}})^{-1}))^{-1}$.

As outlined in Furrer, Genton and Nychka (2006), an efficient computation of the inverse of a (possibly large) symmetric positive-definite matrix, with a given Cholesky matrix factorization, requires the solution of two triangular linear systems using back substitution. In our implementation, for covariance models obtained from (3.2), the solution can be obtained using a Cholesky factorization

Table 4. Top: ML estimation of σ_W^2 a_W, and b_W for the covariance model in Equation (3.2), for $\kappa = 0, 1, 2$ and $\beta_W = 0, 0.5, 1$. RMSE, logS, and CRPS computed using the ML estimated covariance matrix, the percentage of nonzero values associated, and time (in seconds) needed to compute the inverse, respectively. Bottom: ML estimation of σ_M^2 a_M, b_M, and β_M for the covariance model in equation (3.1), for $\mu = 0, 1, 2$.

	β_W	a_W	b_W	σ_W^2	Loglik	RMSE	logS	CRPS	%	Time
$\kappa = 0$	0	1,313.13	4.64	0.325	−691.23	0.2198	−0.1212	0.4399	1.64	4.5
	0.5	1,274.87	3.95	0.323	−724.74	0.2210	−0.1140	0.4396	1.28	3.8
	1	1,342.21	3.12	0.335	−788.79	0.2234	−0.1020	0.4419	0.95	3.6
$\kappa = 1$	0	2,451.25	3.21	0.319	−765.29	0.2234	−0.1020	0.4419	1.28	3.7
	0.5	2,500.77	2.88	0.324	−795.82	0.2240	−0.1036	0.4487	0.09	2.8
	1	2,648.16	2.56	0.338	−829.81	0.2246	−0.1018	0.4504	0.09	2.8
$\kappa = 2$	0	3,586.95	3.33	0.319	−773.65	0.2235	−0.1065	0.4492	1.28	3.7
	0.5	3,637.85	3.09	0.323	−795.05	0.2239	−0.1043	0.4494	1.20	3.5
	1	3,768.07	2.86	0.332	−818.62	0.2244	−0.1028	0.4503	0.09	2.8
	β_M	a_M	b_M	σ_M^2	Loglik	RMSE	logS	CRPS	%	Time
$\mu = 0.5$	0.54	1,374.01	1.322	0.333	−634.44	0.2174	−0.1343	0.4375	100.00	109
$\kappa = 1.5$	1.0	2,498.58	0.528	0.326	−702.77	0.2205	−0.1196	0.4437	100.00	393
$\kappa = 2.5$	1.0	3,604.60	0.368	0.323	−724.16	0.2214	−0.1153	0.4454	100.00	441

and the block sparse Cholesky algorithm of Ng and Peyton (1993), implemented in the *spam* package (Furrer and Sain (2010)). In Table 4, for a given percentage of nonzero values in the covariance matrix, we report the total time (in seconds) needed to compute the Cholesky factor and the inverse using back substitution for the Gneiting–Wendland models. In Table 4 (bottom), we show the total time needed to compute the Cholesky factor using classical Cholesky decomposition, and the inverse using back substitution for the Gneiting–Matérn models. The time in seconds is expressed in terms of elapsed time, using the function *system.time* of the R software on a laptop with a 2.4 GHz processor and 16 GB of memory. As expected, the computational gains obtained using the Gneiting–Wendland models are significant. For instance, computing the inverse is approximately 30 times faster than when using the Gneiting–Matérn model for comparing the cases $k = 0$, $\beta_W = 1$, and $\mu = 0.5$, and approximately 157 times faster for comparing the cases $k = 2$, $\beta_W = 1$, and $\mu = 2.5$. Similar computational gains can be achieved when computing classical space-time kriging and when performing a simulation using a Cholesky decomposition.

In conclusion, we have shown that our models allow for a substantial computational gain at the expense of a very small loss in terms of fitting and prediction performance.

4. Conclusion

As outlined in Bevilacqua et al. (2018), the sizes of data sets associated with spatially or spatio-temporally correlated random processes have steadily increased, making straightforward statistical tools computationally too expensive. The use of covariance functions with an inherent or induced compact support, leading to sparse matrices, is an accessible and scalable approach. The nonseparable compactly supported space-time covariance models introduced in this paper have a spatial (temporal) marginal covariance of the generalized Wendland type and a dynamical decreasing compact support. This is an appealing feature from a computational viewpoint, particularly when dealing with data sets with a large number of sites (temporal instants) and a relatively small number of temporal instants (sites).

The recent work of Bevilacqua et al. (2018) highlights the importance of our covariance models, with its dynamical compact support, for prediction purposes. In fact, Bevilacqua et al. (2018) showed that under specific conditions, Matérn and generalized Wendland covariance models are compatible; that is, the induced Gaussian measures are equivalent. This implies that, under fixed domain asymptotics, a missspecified linear unbiased predictor with a generalized Wendland model is asymptotically as efficient as a true simple kriging predictor using a Matérn model. To some extent, this applies to our space-time dynamical support. However, caution is needed because of the lack of a solid asymptotic framework that would allow us to merge the fixed domain asymptotic in space with the increasing domain in time.

Finally, the construction of nonseparable covariance models with marginal covariances of the generalized Wendland type and with dynamical decreasing compact support is very challenging, from a theoretical point of view. This topic is left for future research.

Supplementary Material

The online Supplementary Material integrates the main results provided in the manuscript. Specifically, Section 2 provides generalizations of Theorems 1 and 2 to a broad class of functions, called multiply monotonic functions. Section 3 explores the necessary and sufficient conditions in a general framework using a Fourier analysis. The Supplementary Material also provides the figures discussed in the paper.

Acknowledgements

The research work conducted by Moreno Bevilacqua was supported in part by FONDECYT grant 1160280 Chile. Emilio Porcu was partly supported by FONDECYT 1130647 Chile. The research work conducted by Moreno Bevilacqua and Emilio Porcu was supported in part by Millennium Science Initiative of the Chilean Ministry of Economy, Development, and Tourism, grant 'Millenium Nucleus Center for the Discovery of Structures in Complex Data'. The work of Marc G. Genton was supported by King Abdullah University of Science and Technology (KAUST).

References

Askey, R. (1973). Radial characteristic functions. *Technical Report No. 1262. - Mathematics Research Center*, University of Wisconsin-Madison.

Buhmann, M. (2000). A new class of radial basis functions with compact support. *Mathematics of Computation* **70**, 233, 307–318.

Bevilacqua, M., Faouzi, T., Furrer, R. and Porcu, E. (update). Estimation and prediction using generalized Wendland functions under fixed domain asymptotics. *The Annals of Statistics*. update.

Chernih, A., Sloan. I. H. and Womersley, R. S. (2014). Wendland functions with increasing smoothness converge to a Gaussian. *Advances in Computational Mathematics* **142**, 17–33.

Daley, D. J., Porcu, E. and Bevilacqua, M. (2015). Classes of compactly supported correlation functions for multivariate random fields. *Stochastic Environmental Research Risk Assessment* **29**, 1249–1263.

Furrer, R., Genton, M. G. and Nychka, D. (2006). Covariance tapering for interpolation of large spatial datasets. *Journal of Computational and Graphical Statistics* **15**, 502–523.

Furrer, R. and Sain, S. R. (2010). spam: A sparse matrix R package with emphasis on MCMC methods for Gaussian Markov random fields. *Journal of Statistical Software* **36**, 1–25.

Gneiting, T. (1999). On the derivatives of radial positive definite functions. *Journal of Mathematical Analysis and Applications* **236**, 86–93.

Gneiting, T. (2000). Kuttner's problem and a Pólya type criterion for characteristic functions. *Proceedings of the American Mathematical Society* **128**, 1721–1728.

Gneiting, T. (2001). Criteria of the Pólya type for radial positive definite functions. *Proceedings of the American Mathematical Society* **129**, 8, 2309–2318.

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

Gneiting, T. (2002b) Compactly supported correlation functions. *Journal of Multivariate Analysis* **83**, 493–508.

Gneiting, T., Genton, M. G. and Guttorp, P. (2007). Geostatistical space-time models, stationarity, separability and full symmetry. In Finkenstadt, B., Held, L. and Isham, V. (eds.). *Statistical Methods for Space-time Systems* (Edited by B. Finkenstadt, L. Held and V. Isham), 151–175. Chapman & Hall/CRC, Boca Raton.

Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association* **102**, 359–378.

Golubov, B. I. (1981). On Abel-Poisson type and Riesz means. *Analysis Mathematica* **7**, 161–184.

Haslett, J. and Raftery, A. (1989). space-time modelling with long-memory dependence: Assessing Ireland's wind power resource. *Applied Statistics* **38**, 1–50.

Hubbert, S. (2012). Closed form representations for a class of compactly supported radial basis functions. *Advances in Computational Mathematics* **36**, 115–136.

Kaufman, C., Schervish, M. and Nychka, D. (2008). Covariance tapering for likelihood-based estimation in large spatial datasets. *Journal of the American Statistical Association* **103**, 1545–1569.

Kuttner, B. (1944). On the Riesz means of a Fourier series (II). *Journal of the London Mathematical Society* **19**, 77–84.

Ng, E. G. and Peyton, B. W. (1993). Block sparse Cholesky algorithms on advanced uniprocessor computers *SIAM Journal on Scientific Computing* **14**, 1064–8275.

Nelder, J. A. and Mead, R (1965). A simplex algorithm for function minimization *Computer Journal* **7**, 308–313.

Pólya, G. (1949). Remarks on characteristic functions. *Proceedings of the [First] Berkeley Symposium on Mathematical Statistics and Probability*, 115–123, University of California Press, Berkeley, Calif.

Porcu, E. and Schilling, R. (2011). From Schoenberg to Pick–Nevanlinna: Toward a complete picture of the variogram class. *Bernoulli* **17**, 441–455.

Porcu, E. and Zastavnyi, V. P. (2012). Characterization theorems for some classes of covariance functions associated to vector valued random fields. *Journal of Multivariate Analysis* **102**, 1293–1301.

Porcu, E. and Zastavnyi, V. P. (2014). generalized Askey functions and their walks through dimensions. *Exposithiones Matematicae* **32**, 190–198.

Porcu, E., Zastavnyi, V. P. and Bevilacqua, M. (update). Buhmann covariance functions, their compact supports, and their smoothness. *Dolomite Research Notes on Approximation*, update.

R Development Core Team (2016). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria.

Schaback, R. (2011). The missing Wendland functions. *Advances in Computational Mathematics* **34**, 67–81.

Schaback, R. and Wu, Z. (1995). Operators on radial functions, *Journal of Computational and Applied Mathematics* **73**, 257–270.

Schlather, M. (2010). Covariance functions generated by Gaussian scale mixtures. *Bernoulli* **16**, 3, 780–797.

Stein, M. L. (1999). *Statistical Interpolation of Spatial Data: Some Theory for Kriging.* Springer, New York.

Stein, M. L. (2005). Space-time covariance functions. *Journal of the American Statistical Association* **100**, 310–321.

Wendland, H. (1995). Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Advances in Computational Mathematics* **4**, 389–396.

Wu, Z. (1995). Compactly supported positive definite radial functions. *Advances in Computational Mathematics* **4**, 283–292.

Zastavnyi, V. P. (2006). On some properties of Buhmann functions. *Ukrainian Mathematical Journal* **58**, 1184–12–9. Translated from Ukrainskyi Matematychnyi Zhurnal, 2006, **58**, 1045–1067.

Zastavnyi, V. P. and Porcu, E. (2011). Characterization theorems for the Gneiting class of covariance functions. *Bernoulli* **17**, 456–465.

Zastavnyi, V. P. and Trigub, R. M. (2002). Positive-definite splines of special form. *Matamatich-eskiuı Sbornik* **193**, 1771–1800.

Zhang, H. and Wang, Y. (2010). Kriging and cross-validation for massive spatial data. *Environmetrics* **21**, 290–304.

Zhu, S. (2012). Compactly supported radial basis functions: how and why, *Oxford Preprint*.

School of Computer Science and Statistics, Trinity College Dublin, College Green, Dublin 2, Ireland.

E-mail: porcue@tcd.ie

Department of Statistics, Universidad de Valparaiso, 2360102 Valparaiso, Chile.
Millennium Nucleus Center for the Discovery of Structures in Complex Data, Chile.

E-mail: moreno.bevilacqua@uv.cl

Statistics Program, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia.

E-mail: marc.genton@kaust.edu.sa

(Received September 2017; accepted June 2018)