Electronic Journal of Statistics
Vol. 18 (2024) 2771–2797
ISSN: 1935-7524
https://doi.org/10.1214/24-EJS2270

Extending the generalized Wendland covariance model

Moreno Bevilacqua[1,2], **Xavier Emery**[3,4] **and Tarik Faouzi**[5]

[1] *Facultad de Ingenierìa y Ciencias, Universidad Adolfo Ibañez,*
e-mail: moreno.bevilacqua@uai.cl

[2] *Dipartimento di Scienze Ambientali, Informatica e Statistica,*
Ca' Foscari University of Venice,
e-mail: moreno.bevilacqua@uai.cl

[3] *Department of Mining Engineering, Universidad de Chile,*
e-mail: xemery@ing.uchile.cl

[4] *Advanced Mining Technology Center, Universidad de Chile,*
e-mail: xemery@ing.uchile.cl

[5] *Departamento de Matemática y Ciencia de la Computación,*
Universidad de Santiago de Chile,
e-mail: tarik.faouzi@usach.cl

Abstract: The generalized Wendland covariance model is a flexible compactly supported covariance model that allows for a continuous parameterization of smoothness of the underlying Gaussian random field, and includes the celebrated Matérn as a special limit case. However, the generalized Wendland model does not cover the full range of validity of the smoothness parameter of the Matérn model. In this paper, we provide new necessary and sufficient conditions of validity of the generalized Wendland model that allows to fill this gap. The effectiveness of our proposal is illustrated through a simulation study and a re-analysis of a large geo-referenced dataset of yearly total precipitation anomalies.

Keywords and phrases: Infill asymptotics, Matérn model, microergodic parameter, sparse matrices.

Received February 2024.

Contents

1. Introduction

Gaussian processes or Gaussian random fields (RFs) are the mathematical foundation for the statistical analyses of geo-referenced data, which allows to describe their marginal behavior and to assess their spatial dependence structure. They are a cornerstone in spatial statistics analysis [15, 49, 26, 3] as well as other disciplines such as machine learning [57], numerical analysis [55], and computer experiments [46], just to mention a few.

The finite-dimensional distributions of a Gaussian RF are characterized by a mean value and a covariance function. In particular, a necessary and sufficient requirement for a given function to be the covariance function of a Gaussian RF is that it is positive semidefinite. Such a requirement is traditionally ensured by selecting a parametric family of covariance functions. If the covariance depends exclusively on the distance between any pair of points located over the domain, it is called isotropic, a popular assumption in spatial statistics. There is a rich catalog of available parametric isotropic covariance functions (see for instance [12] or [3]), among which the Matérn ($\mathcal{MT}$ hereafter) model [40, 31, 49] is by far the most popular. It has played a central role in spatial statistics for decades and, more recently, in other disciplines such as numerical analysis, approximation theory, and machine learning (see [44] for an exhaustive recent review of the $\mathcal{MT}$ model).

A key benefit of the $\mathcal{MT}$ model is that it allows parameterizing in a continuous fashion the differentiability of the sample paths of the associated Gaussian RF, through a positive smoothness parameter $\nu > 0$. The greater is the smoothness parameter, the higher is the level of differentiability of the sample paths. In particular, one can parameterize the fractal dimension of the sample paths, a measure of roughness for non-differentiable RFs, when $0 < \nu < 1$ [29]. Several interesting special cases arise, such as the exponential covariance and, up to a suitable rescaling, the Gaussian covariance when $\nu \to \infty$. Additionally, the $\mathcal{MT}$ model is associated with a class of stochastic partial differential equations [56] that has inspired a fertile body of literature on the approximation of continuously indexed Gaussian RFs through Gaussian Markov RFs [38].

From a computational perspective, a drawback of the $\mathcal{MT}$ model is to be globally supported, *i.e.*, the covariance function does not vanish in the domain of reference. This implies that, for a given set of n spatial points, the associated covariance matrix is dense. Computing the maximum likelihood estimator and/or the optimal predictor can be prohibitive when n is large.

Several approaches have been proposed in the recent years to deal with this issue (see [32] and references therein for an extensive review). One of these approaches considers sparse matrices, *i.e.*, matrices in which most of the elements are zero. In general, the sparseness of a matrix is a desirable feature from a computational viewpoint, since sparse matrix algorithms [19, 18] can be exploited to speed-up the computation associated with estimation, prediction and/or simulation of a Gaussian RF.

A possible solution of the aforementioned problem, followed by [38] among others, is to focus on the inverse covariance matrix, *i.e.* the precision matrix, that can be generally well approximated with a sparse matrix, at least for dense covariance matrices. Another solution is to focus on the so-called Vecchia's method [53] and its extensions [17, 30, 35], which imply a sparse approximation of the Cholesky factor of the precision matrix. These two approaches have proven to be effective solutions at least when the underlying covariance function is the $\mathcal{MT}$ model.

The goal of the covariance tapering approach [23, 36, 50] is to obtain sparse covariance matrices. This is achieved by multiplying the $\mathcal{MT}$ model with a taper function, that is, a correlation function being additionally compactly supported over a ball with given radius. Thus, the resulting covariance tapered matrix is sparse, with the level of sparseness depending on the radius of compact support.

However, as shown in [6], a better alternative approach is to use flexible compactly supported models that leads to sparse covariance matrices. Probably the most famous case is the generalized Wendland ($\mathcal{GW}$ hereafter) covariance model [27, 59, 10], which allows for the parameterization of the differentiability of the sample paths of the underlying Gaussian RF, through a nonnegative smoothness parameter $\kappa \geq 0$, in the same fashion as the $\mathcal{MT}$ model.

Several connections between the $\mathcal{MT}$ and $\mathcal{GW}$ models have been studied. For instance, [6] shows that, under specific conditions, $\mathcal{MT}$ and $\mathcal{GW}$ covariance models lead to equivalent Gaussian measures. One consequence of this result is that, when the true covariance function belongs to the $\mathcal{MT}$ family, prediction can be performed with the $\mathcal{GW}$ without any loss of prediction efficiency under fixed domain asymptotics. In addition, [4] shows that the $\mathcal{MT}$ model is actually a special limit case of a compactly supported reparameterization of the $\mathcal{GW}$ model. In particular, the reparameterized $\mathcal{GW}$ model has an additional parameter that, for given smoothness and spatial dependence parameters, allows switching from the world of flexible compactly supported covariance functions to the world of flexible globally supported covariance functions.

However, as argued in Section 3 hereinafter, it turns out that the $\mathcal{GW}$ model does not cover the full range of validity of the smoothness parameter of the Matérn model. Specifically, the $\mathcal{GW}$ model cannot attain the $\mathcal{MT}$ model when $0 < \nu < 0.5$ and, as a consequence, it is not able to fully parameterize the fractal dimension of the associated Gaussian RF as for the $\mathcal{MT}$ model [29].

The goal of this paper is to fill this gap. In particular, we provide new necessary and sufficient conditions of validity for the $\mathcal{GW}$ model, extending the parametric space of the smoothness parameter from $\kappa \geq 0$ to $\kappa > -0.5$. As a consequence of our result, the $\mathcal{GW}$ model is able to cover the full range of validity of the smoothness parameter of the $\mathcal{MT}$ model, $\nu > 0$. In particular, the results in [6] and [4] can be applied also to the extended $\mathcal{GW}$ model. This implies that statistical analysis of continuous spatial data displaying very rough sample paths using a flexible compactly supported correlation model is now feasible (see the real data example in Section 5).

The remainder of the paper is organized as follows. In Section 2 we review the celebrated $\mathcal{MT}$ model, while in Section 3 we review the $\mathcal{GW}$ model and provide new necessary and sufficient validity conditions that allow the $\mathcal{GW}$ model to cover the full range of validity of the smoothness parameter of the $\mathcal{MT}$ model. In Section 4, we report a simulation study that explores the finite sample properties of the weighted composite likelihood estimation and maximum likelihood estimation methods of the covariance parameters of the extended reparameterized $\mathcal{GW}$ model under both increasing and fixed domain asymptotics settings. In Section 5, we compare the extended reparameterized $\mathcal{GW}$ model with the $\mathcal{MT}$ model in a re-analysis of a large spatial geo-referenced data set of yearly total precipitation anomalies. Concluding remarks are consigned in Section 6.

The extended $\mathcal{GW}$ covariance model has been implemented in the `GeoModels` package [8] for the open-source R statistical environment.

2. The Matérn correlation model

We denote $\{Z(\boldsymbol{s}) : \boldsymbol{s} \in D\}$ a stationary Gaussian RF defined on a set D of $\mathbb{R}^d$. To simplify notation, we focus on stationary zero mean and unit variance Gaussian RF such that $\mathbb{E}(Z(\boldsymbol{s})) = 0$, $\mathbb{V}(Z(\boldsymbol{s})) = 1$ with stationary correlation function $C : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$. We consider the class Φ_d of continuous mappings $\phi : [0, \infty) \to \mathbb{R}$ with $\phi(0) = 1$, such that

$$\operatorname{corr}(Z(\boldsymbol{s}), Z(\boldsymbol{s}')) = C(\boldsymbol{s}, \boldsymbol{s}') = \phi(\|\boldsymbol{s}' - \boldsymbol{s}\|),$$

with $\boldsymbol{s}, \boldsymbol{s}' \in D$, and $\|\cdot\|$ denoting the Euclidean norm. Gaussian RFs with such covariance functions are called weakly stationary and isotropic.

[48] characterized the class Φ_d as scale mixtures of the characteristic functions of random vectors uniformly distributed on the spherical shell of $\mathbb{R}^d$, with any nonnegative measure F on $[0, \infty)$:

$$\phi(x) = \int_0^\infty \Omega_d(xr) F(\mathrm{d}r), \qquad x \geq 0,$$

with $\Omega_d(x) = \Gamma(d/2)(2/x)^{d/2-1} J_{d/2-1}(x)$ for $x > 0$ and 1 for $x = 0$, Γ the gamma function, and J_ν the Bessel function of the first kind of order ν. The class Φ_d is nested, with the inclusion relation $\Phi_1 \supset \Phi_2 \supset \ldots \supset \Phi_\infty$ being strict, and where $\Phi_\infty := \bigcap_{d\geq 1} \Phi_d$ is the class of continuous mappings ϕ, the radial version of which is positive semidefinite on any d-dimensional Euclidean space.

Fourier transforms of radial versions of members of Φ_d that are absolutely integrable, for a given d, have a simple expression, as reported in Yaglom [58] or Stein [49]. For an absolutely integrable member ϕ of the class Φ_d, we define its isotropic spectral density as

$$\widehat{\phi}(z) = \frac{z^{1-d/2}}{(2\pi)^{\frac{d}{2}}} \int_0^\infty u^{d/2} J_{d/2-1}(uz)\phi(u)\mathrm{d}u, \qquad z \geq 0. \tag{2.1}$$

The $\mathcal{MT}$ correlation function is defined as

$$\mathcal{MT}_{\nu,\alpha}(x) = \begin{cases} 1 & \text{if } x = 0, \\ \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{x}{\alpha}\right)^\nu \mathcal{K}_\nu\left(\frac{x}{\alpha}\right), & \text{if } x > 0, \end{cases} \tag{2.2}$$

with $\mathcal{K}_\nu$ the modified Bessel function of the second kind of order ν, and where $\alpha, \nu > 0$ are necessary and sufficient condition for $\mathcal{MT}$ to belong to the class Φ_∞ [49]. The isotropic spectral density, $\widehat{\mathcal{MT}}_{\alpha,\nu}$, has a simple expression:

$$\widehat{\mathcal{MT}}_{\nu,\alpha}(z) = \frac{\Gamma(\nu + d/2)}{\pi^{d/2}\Gamma(\nu)} \frac{\alpha^d}{(1+\alpha^2 z^2)^{\nu+d/2}}, \qquad z \geq 0. \tag{2.3}$$

The importance of the $\mathcal{MT}$ model stems from the parameter ν that controls the differentiability (in the mean square sense) of the associated Gaussian RF and its sample paths. Specifically, for any integer $k = 0, 1, \ldots$, the sample paths of a Gaussian RF field with $\mathcal{MT}$ covariance function are k-times differentiable, in any direction, if and only if $\nu > k$. In particular the $\mathcal{MT}$ model is able to parameterize the fractal dimension of the sample paths, a measure of roughness for non-differentiable Gaussian RFs (see [29] for a formal definition of the fractal dimension) which equals d if $\nu > 1$ and $d + 1 - \nu$ if $0 < \nu < 1$, with smaller values of ν indicating rougher sample paths.

A rescaled version of the Matérn covariance converges to the Gaussian (squared exponential) covariance as $\nu \to \infty$, that is

$$\mathcal{MT}_{\nu,\alpha/(2\sqrt{\nu})}(x) \xrightarrow[\nu\to\infty]{} \exp(-x^2/\alpha^2), \qquad x \geq 0,$$

with the convergence being uniform on any compact set of $\mathbb{R}^d$. For this reason $\mathcal{MT}_{\nu,\alpha/(2\sqrt{\nu})}$ is sometimes adopted as parameterization, in particular in the machine learning community [57].

When $\nu = k+1/2$, for k a nonnegative integer, the $\mathcal{MT}$ covariance simplifies into the product of an exponential covariance with a polynomial of order k. For instance, $\mathcal{MT}_{1/2,\alpha}(x) = \exp(-x/\alpha)$ and, in general,

$$\mathcal{MT}_{k+1/2,\alpha}(x) = \exp(-x/\alpha) \sum_{i=0}^{k} \frac{(k+i)!}{2k!} \binom{k}{i} (2x/\alpha)^{k-i}, \quad k = 0, 1, \ldots.$$

Table 1 summarizes the cases $\mathcal{MT}_{\nu,\alpha}(x)$ for $\nu = 0.5, 1.5, 2.5, 3.5$.

Table 1
The $\mathcal{MT}_{\nu,a}$ model for $\nu = 0.5, 1.5, 2.5, 3.5$. $SP(k)$ means that the sample paths of the associated Gaussian RF are k times differentiable.

ν	$\mathcal{MT}_{\nu,a}(x)$	$SP(k)$
0.5	$e^{-\frac{x}{a}}$	0
1.5	$e^{-\frac{x}{a}}\left(1+\frac{x}{a}\right)$	1
2.5	$e^{-\frac{x}{a}}\left(1+\frac{x}{a}+\frac{x^2}{3a^2}\right)$	2
3.5	$e^{-\frac{x}{a}}\left(1+\frac{x}{2a}+\frac{6x^2}{15a^2}+\frac{x^3}{15a^3}\right)$	3

3. The generalized Wendland correlation model and its extension

3.1. The generalized Wendland correlation model

The $\mathcal{GW}$ family of correlation functions [27, 59, 10, 6, 4] allows, as in the $\mathcal{MT}$ case, for a continuous parameterization of smoothness of the underlying Gaussian RF, being additionally compactly supported. For $\kappa \geq 0$ and $\beta > 0$, it can be defined in terms of the Gauss hypergeometric function ${}_2F_1$ [1] as:

$$\mathcal{GW}_{\kappa,\mu,\beta}(x) := \begin{cases} M\left(1-\frac{x^2}{\beta^2}\right)^{\kappa+\mu} {}_2F_1\left(\frac{\mu}{2}, \frac{\mu+1}{2}; \kappa+\mu+1; 1-\frac{x^2}{\beta^2}\right), & 0 \leq x < \beta, \\ 0, & x \geq \beta, \end{cases} \tag{3.1}$$

with $M = \frac{\Gamma(\kappa)\Gamma(2\kappa+\mu+1)}{\Gamma(2\kappa)\Gamma(\kappa+\mu+1)2^{\mu+1}}$ or, equivalently, through an integral representation:

$$\mathcal{GW}_{\kappa,\mu,\beta}(x) := \begin{cases} \frac{1}{B(2\kappa,\mu+1)} \int_{x/\beta}^1 u(u^2-(x/\beta)^2)^{\kappa-1}(1-u)^\mu \, du, & 0 \leq x < \beta, \\ 0, & x \geq \beta, \end{cases} \tag{3.2}$$

where $B(\cdot,\cdot)$ is the Beta function. For a given smoothness parameter $\kappa \geq 0$,

$$\mu \geq (d+1)/2 + \kappa \tag{3.3}$$

is a necessary and sufficient condition for $\mathcal{GW}_{\kappa,\mu,\beta}$ to belong to the class Φ_d.

The associated spectral density is given by [11]:

$$\widehat{\mathcal{GW}}_{\kappa,\mu,\beta}(z) = L\beta^d {}_1F_2\Big(\frac{d+1}{2}+\kappa; \frac{d+1+\mu}{2}+\kappa, \frac{d+\mu}{2}+1+\kappa; -\frac{(z\beta)^2}{4}\Big), \quad z \geq 0, \tag{3.4}$$

where ${}_1F_2$ is a special case of the generalized hypergeometric function [1], and L is a normalization constant:

$$L = \frac{2^{-d}\pi^{-\frac{d}{2}}\Gamma(\mu+2\kappa+1)\Gamma(2\kappa+d)\Gamma(\kappa)}{\Gamma(\kappa+d/2)\Gamma(\mu+2\kappa+d+1)\Gamma(2\kappa)}.$$

When computing (3.1) or (3.2), a numerical integration or efficient evaluation of the hypergeometric function is obviously feasible, but could be cumbersome to (spatial) statisticians used to handle closed-form parametric correlation models.

However, similarly to the Matérn, in some special cases, the computation can be considerably simplified. An important example is when κ is a nonnegative integer, as exemplified next.

Example 1 (original Wendland functions, see [54]). *If $\kappa = k$ is a nonnegative integer, then*

$$\mathcal{GW}_{k,\mu,\beta}(x) = \left(1 - \frac{x}{\beta}\right)_+^{\mu+k} P_k(x;\mu,\beta), \quad k = 0, 1, 2, \dots \tag{3.5}$$

where $P_k(x,\mu,\beta)$ is a polynomial of degree k that can be expressed as:

$$P_k(x;\mu,\beta) = L_k \sum_{j=0}^{k} c_{j,k}(\mu) \left(\frac{x}{\beta}\right)^{k-j} \left(1 - \frac{x}{\beta}\right)^{j}, \tag{3.6}$$

where $c_{j,k}(\mu) = \frac{(k+j)!}{2^j j!(k-j)!} \frac{\Gamma(2k+\mu+1)}{\Gamma(k+j+\mu+1)}$ and $L_k = \frac{2^k k!}{(2k)!}$.

The proof of this example derives from Theorem 4.1 in [33] (see also Appendix A.1 for a shorter proof). Note that when the smoothness parameter is an integer, then the sample paths of the associated Gaussian RF are k times differentiable. Figure 1 illustrates the cases $k = 0$ and $k = 1$ for $\mu = m, m+2, m+4, m+8$ with $m = 0.5(d+1) + k$ and $d = 2$. Table 2 summarizes the cases $\mathcal{GW}_{k,\mu,\beta}(x)$ for $k = 0, 1, 2, 3$.

TABLE 2

The $\mathcal{GW}_{\kappa,\mu,\beta}$ model for $\kappa = 0, 1, 2, 3$. $SP(k)$ means that the sample paths of the associated Gaussian RF are k times differentiable.

κ	$\mathcal{GW}_{\kappa,\mu,\beta}(x)$	$SP(k)$
0	$\left(1-\frac{x}{\beta}\right)_+^{\mu}$	0
1	$\left(1-\frac{x}{\beta}\right)_+^{\mu+1}\left(1+\frac{x}{\beta}(\mu+1)\right)$	1
2	$\left(1-\frac{x}{\beta}\right)_+^{\mu+2}\left(1+\frac{x}{\beta}(\mu+2)+\left(\frac{x}{\beta}\right)^2(\mu^2+4\mu+3)\frac{1}{3}\right)$	2
3	$\left(1-\frac{x}{\beta}\right)_+^{\mu+3}\left(1+\frac{x}{\beta}(\mu+3)+\left(\frac{x}{\beta}\right)^2(2\mu^2+12\mu+15)\frac{1}{5}+\left(\frac{x}{\beta}\right)^3(\mu^3+9\mu^2+23\mu+15)\frac{1}{15}\right)$	3

Another example is when κ is a half-integer and μ is an integer. In this case the $\mathcal{GW}$ model can be expressed as a combination of Legendre polynomials of a certain degree. The following example considers the case $\kappa = 0.5$.

Example 2 (missing Wendland functions, see [47]). *If $\kappa = 0.5$ and μ is a positive integer, then*

$$\mathcal{GW}_{\frac{1}{2},\mu,\beta}(x) = \begin{cases} \frac{M 2^{\mu+1}\Gamma(\mu+\frac{3}{2})}{\sqrt{\pi}\Gamma(\mu+2)} z^{-\mu-1}\Big[\frac{\mu}{2}\left(P_{\mu-1}(z) - zP_{\mu}(z)\right)\log\left(\frac{z+1}{z-1}\right) + zP_{\mu-1}(z) \\ +\mu\sum_{p=1}^{\mu-1}\frac{1}{p}P_{p-1}(z)(zP_{\mu-p}(z) - P_{\mu-1-p}(z))\Big], \quad 0 \le x < \beta, \\ 0, \quad x \ge \beta, \end{cases} \tag{3.7}$$

where $z = \left(1 - \frac{x^2}{\beta^2}\right)^{-\frac{1}{2}}$ and P_μ is the Legendre polynomial of degree μ [1].

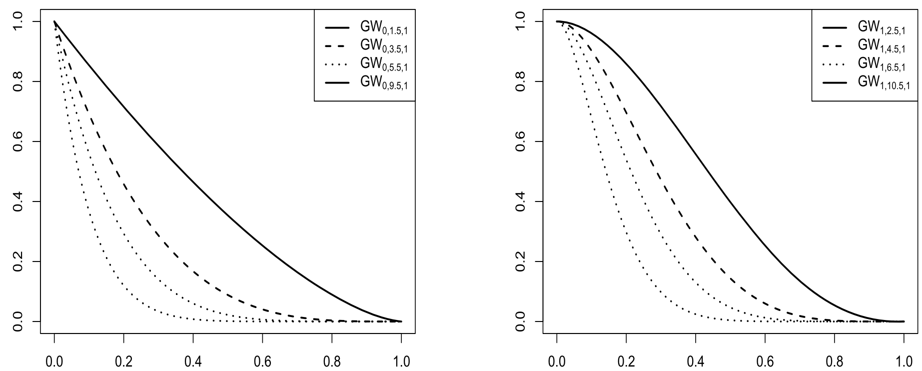


Fig 1. *Examples of the* $\mathcal{GW}_{k,\mu,\beta}$ *model with* $\beta = 1$ *and* $\kappa = 0$ *(left) or* $\kappa = 1$ *(right) for* $\mu = m, m+2, m+4, m+8$ *with* $m = 0.5(d+1) + \kappa$ *and* $d = 2$.

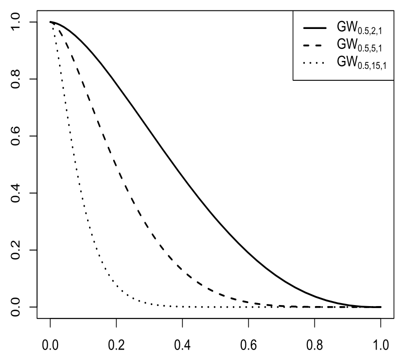


Fig 2. *Examples of the* $\mathcal{GW}_{\frac{1}{2},\mu,1}$ *model for* $\mu = 2, 5, 15$ *(from top to bottom).*

The proof of this result is given in Appendix A.2, together with a general closed-form expression of $\mathcal{GW}_{\kappa,\mu,\beta}$ for any half-integer $\kappa = 0.5, 1.5, 2.5, \ldots$ and positive integer μ. The case $\mathcal{GW}_{\frac{1}{2},\mu,1}$ for $\mu = 2, 5, 15$ is depicted in Figure 2.

3.2. Properties and reparameterization

As in the Matérn case, the $\mathcal{GW}$ model allows parameterizing in a continuous fashion the mean-square differentiability of the underlying Gaussian RF and its associated sample paths, through the smoothness parameter κ.

Specifically, for any integer $k = 1, 2, \ldots$, the sample paths of the $\mathcal{GW}_{\kappa,\mu,\beta}$ model are k times differentiable, in any direction, if and only if $\kappa > k - 0.5$ and for $0 \le \kappa \le 0.5$ they are not differentiable. In addition, it can be shown that the fractal dimension of the associated sample paths equals d if $\kappa > 0.5$ and $d + 1 - \kappa$ if $0 \le \kappa < 0.5$. Since in general the fractal dimension belongs to the interval $[d, d+1]$ [29], this implies that the $\mathcal{GW}$ model, unlike the $\mathcal{MT}$ model, is not able to entirely cover the full range of the fractal dimension.

The connections between the $\mathcal{MT}$ and the $\mathcal{GW}$ families have been investigated by [6] and [4]. In particular, [6] established that two Gaussian measures with respective covariance functions $\sigma_0^2\mathcal{MT}_{\nu,\alpha}$ and $\sigma_1^2\mathcal{GW}_{\kappa,\mu,\beta}$ are equivalent (for a formal definition of equivalent Gaussian measures, see for instance [49]) if $\mu > d + \frac{1}{2}$, $d = 1, 2, 3$ and

$$\frac{\sigma_0^2}{\alpha^{2\nu}} = \left(\frac{\Gamma(2\kappa+\mu+1)}{\Gamma(\mu)}\right)\frac{\sigma_1^2}{\beta^{1+2\kappa}}, \quad \nu = \kappa + 0.5. \tag{3.8}$$

One remarkable implication of the former result is that, when the true covariance belongs to the $\mathcal{MT}$ family, asymptotic efficiency prediction and asymptotically correct estimation of mean square error can be achieved using a $\mathcal{GW}$ model provided that the conditions in (3.8) are fulfilled. It is important to stress that the right condition in (3.8) shows that the equivalence is valid only for $\nu \geq 0.5$, *i.e.* it does not cover the full parametric space of the $\mathcal{MT}$ smoothness parameter.

In addition, [4] considered a compactly supported reparameterization of the $\mathcal{GW}$ model, that is

$$\mathcal{RGW}_{\kappa,\mu,\beta}(x) := \mathcal{GW}_{\kappa,\mu,\delta(\kappa,\mu,\beta)}(x), \tag{3.9}$$

where $\delta(\kappa,\mu,\beta) := \beta\left(\frac{\Gamma(\mu+2\kappa+1)}{\Gamma(\mu)}\right)^{\frac{1}{1+2\kappa}}$ is the compact support and proved that

$$\lim_{\mu\to\infty}\mathcal{RGW}_{\kappa,\mu,\beta}(x) = \mathcal{MT}_{\kappa+1/2,\beta}(x), \quad \kappa \geq 0, \tag{3.10}$$

with uniform convergence for $x \in (0,\infty)$. As a consequence, the $\mathcal{RGW}$ model can be viewed as a generalization of the Matérn model with an additional parameter that, for given smoothness and spatial dependence parameters, allows switching from the world of flexible compactly supported covariance functions to the world of flexible globally supported covariance functions. However, it can be appreciated from (3.10) that the $\mathcal{RGW}$ model cannot attain the $\mathcal{MT}_{\nu,\alpha}$ model when $0 < \nu < 0.5$.

To sum up, the $(\mathcal{R})\mathcal{GW}$ model does not cover the full range of validity of the smoothness parameter of the Matérn model and, in particular, it is not able to fully parameterize the fractal dimension of the associated Gaussian RF as in the Matérn case. To fill this gap, we now enlarge the conditions of validity of the $\mathcal{GW}$ model.

3.3. Extension of parameter validity conditions

The following theorem extends the $(\mathcal{R})\mathcal{GW}$ model for negative values of the smoothness parameter. Specifically, it provides necessary and sufficient conditions for the $(\mathcal{R})\mathcal{GW}$ model to belong to the class Φ_d, when $-0.5 < \kappa < 0$. The proof is deferred to the Appendix.

Theorem 1. *Let $-0.5 < \kappa < 0$. Then the $(\mathcal{R})\mathcal{GW}_{\kappa,\mu,\beta}$ correlation model belongs to the class Φ_d, if and only if*

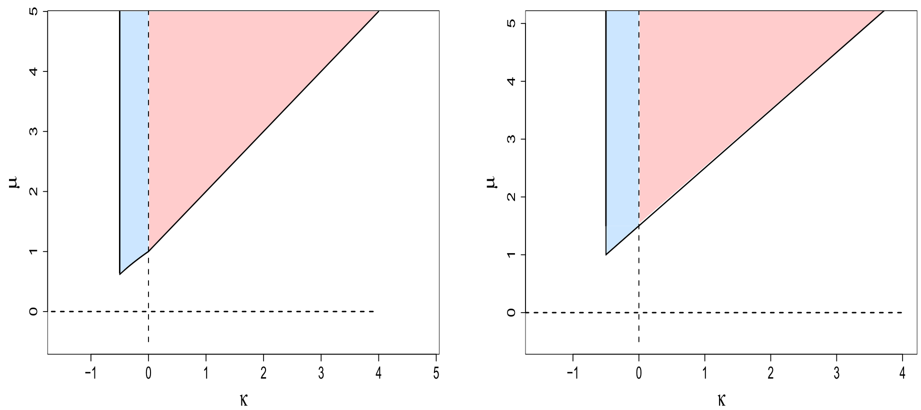


FIG 3. (κ, μ) *region of necessary and sufficient conditions for* $(\mathcal{R})\mathcal{GW}_{\kappa,\mu,\beta}$ *to belong to* Φ_d, *when* $d = 1$ *(left part) and* $d = 2$ *(right part). The blue region is obtained by applying Theorem 1.*

1. $\mu \geq (d+1)/2 + \kappa$ *for* $d \geq 2$
2. $\mu \geq 0.5(\sqrt{8\kappa + 9} - 1)$ *for* $d = 1$.

Recalling that for the case $\kappa \geq 0$ the condition $\mu \geq (d+1)/2 + \kappa$ is necessary and sufficient for $(\mathcal{R})\mathcal{GW}_{\kappa,\mu,\beta}$ to belong to the class Φ_d, we now have a more global picture of the validity conditions of the $(\mathcal{R})\mathcal{GW}$ model. Figure 3 depicts the (κ, μ) region of necessary and sufficient validity conditions of the $(\mathcal{R})\mathcal{GW}_{\kappa,\mu,\beta}$ model for the case $d = 1, 2$ (from left to right), highlighting the region provided by Theorem 1 (blue color).

More importantly, using the results in [6], it can be shown that the equivalence condition between the $\mathcal{MT}$ and $\mathcal{GW}$ models given in (3.8) can be extended to the case $-0.5 < \kappa < 0$, that is, it is valid for the full parametric space of the $\mathcal{MT}$ smoothness parameter. Similarly, using the new condition, it can be shown that

$$\lim_{\mu \to \infty} \mathcal{RGW}_{\kappa,\mu,\beta}(x) = \mathcal{MT}_{\kappa+1/2,\beta}(x), \quad \kappa > -\frac{1}{2},$$

that is, the uniform convergence of the $\mathcal{RGW}_{\kappa,\mu,\beta}$ model to the $\mathcal{MT}_{\kappa+1/2,\beta}$ model when $\mu \to \infty$ is valid for the full parametric space of the smoothness parameter. Finally, using the proposed extension, the $(\mathcal{R})\mathcal{GW}$ is now able to fully parameterize the fractal dimension index.

Figure 4 shows some examples of the $\mathcal{RGW}_{\kappa,\mu,\beta}$ correlation model when $-0.5 < \kappa < 0$ and they are compared with the $\mathcal{RGW}_{0,\mu,\beta}$, that is using the lower bound of the smoothness parameter known so far. In addition, Figure 5 depicts two realizations on a fine grid on a unit square, of a zero mean and unit variance Gaussian RF with correlation functions $\mathcal{RGW}_{-0.1,1.4,0.1}$ and $\mathcal{RGW}_{-0.4,1.1,0.1}$, respectively, obtained with the Cholesky decomposition of the covariance matrix approach. As expected, one can appreciate the roughness of the sample paths when decreasing the smoothness parameter.

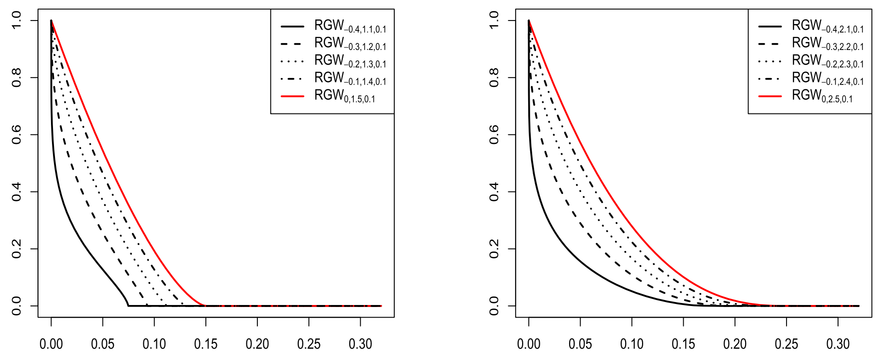


FIG 4. *Examples of extended $\mathcal{RGW}$ correlation model. Left: the $\mathcal{RGW}_{\kappa,1.5+\kappa,0.1}$ model for $\kappa = -0.1, -0.2, -0.3, -0.4$ (black lines) from top to bottom. The red line is the $\mathcal{RGW}_{0,1.5,0.1}$ model. Right: the $\mathcal{RGW}_{\kappa,2.5+\kappa,0.1}$ model for $\kappa = -0.1$, -0.2, -0.3, -0.4 (black lines) from top to bottom. The red line is the $\mathcal{RGW}_{0,2.5,0.1}$ model.*

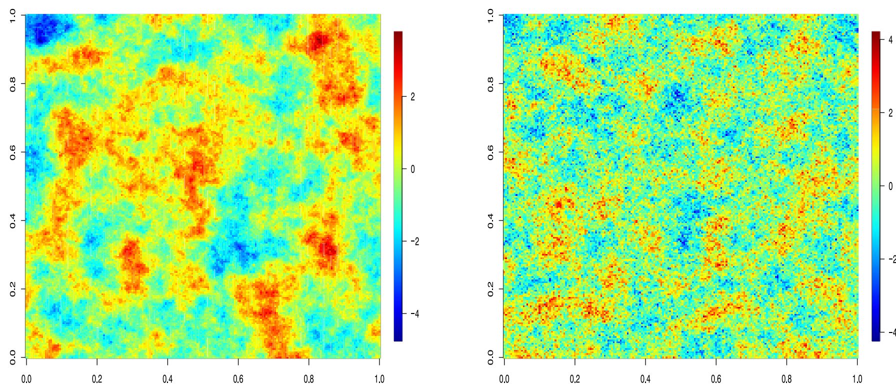


FIG 5. *Two realizations on a unit square from a zero mean unit variance Gaussian RF with covariance functions $\mathcal{RGW}_{-0.1,1.4,0.1}$ and $\mathcal{RGW}_{-0.4,1.1,0.1}$ (from left to right).*

4. A simulation study

In this section, we perform a simulation study on the $\mathcal{GW}$ family focusing on the proposed new range of the smoothness parameter, *i.e.* $-0.5 < \kappa < 0$, under both increasing and fixed domain asymptotics. In particular, we consider the $\mathcal{RGW}$ parameterization (3.9) that includes the $\mathcal{MT}$ as a special case.

In the $\mathcal{RGW}_{\kappa,\mu,\beta}$ model, κ is the smoothness parameter describing the behavior near the origin of the correlation model, β is a spatial dependence parameter, $\delta(\kappa, \mu, \beta)$ as defined in (3.9) is the compact support parameter, and μ is an additional parameter. Small values of μ lead to a compactly supported correlation model, while values μ becoming infinitely large lead to the $\mathcal{MT}$ globally supported correlation model.

As outlined in [4], the μ parameter can be estimated with the goal of looking for an improvement of the $\mathcal{MT}$ family from a modeling viewpoint, or can be

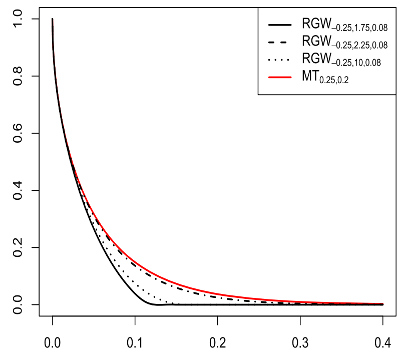


Fig 6. *The covariance models in the simulation study: 1) $\sigma^2\mathcal{RGW}_{-0.25,1.75,\beta}$; 2) $\sigma^2\mathcal{RGW}_{-0.25,2.25,\beta}$; 3) $\sigma^2\mathcal{RGW}_{-0.25,3.25,\beta}$; and 4) $\sigma^2\mathcal{MT}_{0.25,\beta}$ (in red color) with $\sigma^2 = 1$ and $\beta = 0.08$.*

arbitrary fixed with the goal to seek highly sparse matrices to reduce the computational complexity, in particular for prediction. In our simulation study and in the application, we adopt the second approach, which is clearly more suitable for large datasets.

4.1. Increasing domain asymptotics

We first assume a scenario where all the parameters can be consistently estimated, *i.e.* an increasing domain scenario. Under this setting, asymptotic results such as consistency and asymptotic normality associated with the maximum likelihood estimator or other estimation methods such as composite likelihood are well established [39, 7].

One interesting scenario is when the true correlation model is globally supported, *i.e.* it is of $\mathcal{MT}$ type but, for computational reason, a misspecified $\mathcal{RGW}$ model is considered in the estimation and prediction step fixing small values of μ.

With this goal in mind, we simulate 500 realizations of a zero mean Gaussian RF with covariance model $\sigma^2\mathcal{MT}_{\kappa+0.5,\beta}$ with $\sigma^2 = 1$, $\beta = 0.08$ and $\kappa = -0.25$, observed at $n = 1500$ locations uniformly distributed in the unit square. Then, for each simulated dataset, we estimate the parameters σ^2, β and κ using the following covariance models: 1) $\sigma^2\mathcal{RGW}_{\kappa,1.75,\beta}$; 2) $\sigma^2\mathcal{RGW}_{\kappa,2.25,\beta}$; 3) $\sigma^2\mathcal{RGW}_{\kappa,3.25,\beta}$ and the true model 4) $\sigma^2\mathcal{MT}_{\kappa+0.5,\beta} \equiv \sigma^2\mathcal{RGW}_{\kappa,\infty,\beta}$. Note that, in the covariance models 1), 2), 3) and 4), the μ parameter is increasing and for the case 1) it is set very close to its lower bound.

Figure 6 depicts the four covariance models considered in the simulation, where the true model $\sigma^2\mathcal{MT}_{0.25,0.08}$ is highlighted with a red color. It can be appreciated that, for the models 1), 2) and 3), the compact support increases when increasing μ.

The fraction of zero values in the associated covariance matrix (computed as the ratio between the number of zero entries and the total number of entries in

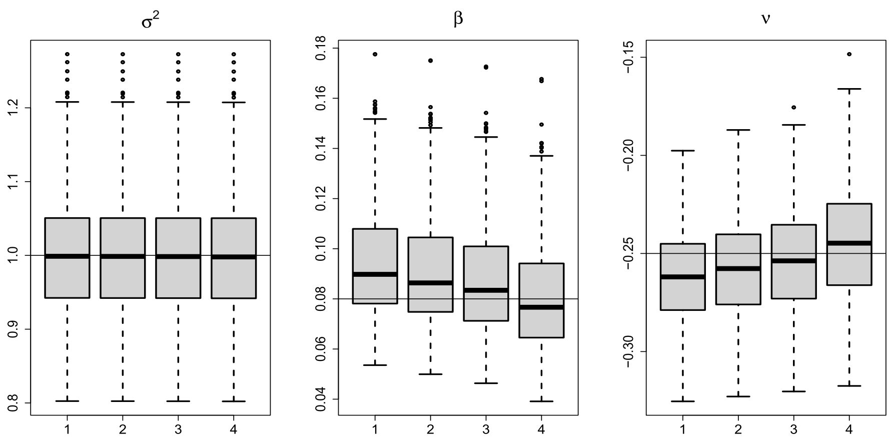


FIG 7. *Boxplots of the NNWCL estimates of σ^2, β and κ (from left to right) for the covariance models 1), 2), 3) and 4), where model 4) is true covariance model $\sigma^2\mathcal{MT}_{\kappa+0.5,\beta}$ and models 1), 2), 3) are the misspecified models $\sigma^2\mathcal{RGW}_{\kappa,\mu,\beta}$ with increasing values of $\mu = 1.75$, 2.25, 3.25. The true parameters values are $\sigma^2 = 1$, $\beta = 0.08$ and $\kappa = -0.25$.*

the covariance matrix, that is 1500^2) is given by 0.96, 0.93, 0.85 for models 1), 2) and 3), respectively, and it is clearly 0 for model 4).

In principle, the estimation of the parameters can be performed using maximum likelihood (ML). However, as outlined in [4], we point out that ML estimation can partially take advantage of the computational benefits associated with the $\mathcal{RGW}$ model because the compact support $\delta(\kappa, \mu, \beta)$ depends on β, μ and κ. Even when considering a fixed μ, the covariance matrix can be highly or slightly sparse, depending on the value of β and κ in the optimization process. An alternative strategy is to use estimation methods with a good balance between statistical efficiency and computational complexity that do not require any restrictions on the covariance model, such as composite likelihood methods [20, 7, 9], multi-resolution approximation methods [34] or, more generally, using Vecchia's approximations [35].

In this simulation study we consider the nearest neighbors weighted pairwise conditional composite likelihood (NNWCL) proposed in [9] where the number of neighbors is set to 5. Figure 7 depicts the boxplots of the NNWCL estimates of σ^2, β and κ (from left to right) for the covariance models 1), 2), 3) and 4). For the variance parameter σ^2 there are no relevant differences between the distribution of the estimates for the four covariance models. However, for β and κ estimation, it can be appreciated a slight bias when using the misspecified models 1), 2) and 3) that tends to decrease when increasing μ, as expected.

More importantly we evaluate the predictive performances of the four covariance models using cross-validation. In particular, for each simulated dataset $k = 1, \ldots, 500$, we randomly choose 90% of the spatial locations as a training subset and we use the remaining 10% as a testing subset for predictions, and we repeat this resampling approach 100 times. NNWCL estimation is per-

TABLE 3
Prediction performances of the covariance models 1) $\sigma^2\mathcal{RGW}_{\kappa,1.75,\beta}$; 2) $\sigma^2\mathcal{RGW}_{\kappa,2.25,\beta}$; 3) $\sigma^2\mathcal{RGW}_{\kappa,3.25,\beta}$ and 4) $\sigma^2\mathcal{RGW}_{\kappa,\infty,\beta} \equiv \sigma^2\mathcal{MT}_{\kappa+0.5,\beta}$.

Model	1	2	3	4
RMSE	0.6926	0.6913	0.6912	0.6912
LSCORE	1.2393	1.2308	1.2303	1.2293
CRPS	0.6884	0.6878	0.6875	0.6875

formed using the training subset and then NNWCL estimates are used to compute three prediction scores [28] using the testing subset for each covariance models. Specifically, for each $j-th$ left-out testing subset of the k-th dataset $(z^L_{j,k}(\boldsymbol{s}_1),\dots,z^L_{j,k}(\boldsymbol{s}_M))$, for $j=1,\dots,100$ we compute

1. the root mean squared error

$$\overline{\mathrm{RMSE}}_{j,k} = \left[\frac{1}{M}\sum_{i=1}^{M}\left(z^L_{j,k}(\boldsymbol{s}_i) - \widehat{Z}^L_{j,k}(\boldsymbol{s}_i)\right)^2\right]^{\frac{1}{2}}$$

2. the logarithmic score

$$\overline{\log S}_{j,k} = \frac{1}{M}\sum_{i=1}^{M}\left[\frac{1}{2}\log\{2\pi\sigma^L_{j,k}(\boldsymbol{s}_i)\} + \frac{1}{2}\{g^L_{j,k}(\boldsymbol{s}_i)\}^2\right], \tag{4.1}$$

3. the continuous ranked probability

$$\overline{\mathrm{CRPS}}_{j,k} = \frac{\sum_{i=1}^{M}\sigma^L_{j,k}(\boldsymbol{s}_i)(g^L_{j,k}(\boldsymbol{s}_i)(2\Phi(g^L_{j,k}(\boldsymbol{s}_i))-1) + 2\Phi(g^L_{j,k}(\boldsymbol{s}_i)) - \frac{1}{\sqrt{\pi}})}{M}, \tag{4.2}$$

where $\widehat{Z}^L_{j,k}(\boldsymbol{s}_i)$ is the optimal linear predictor at location $\boldsymbol{s}_i$, $\sigma^L_{j,k}(\boldsymbol{s}_i)$ is the corresponding standard error deviation, and $g^L_{j,k}(\boldsymbol{s}_i) = (z^L_{j,k}(\boldsymbol{s}_i) - \widehat{Z}^L_{j,k}(\boldsymbol{s}_i))/\sigma^L_{j,k}(\boldsymbol{s}_i)$. Table 3 reports the overall means RMSE $= \sum_{k=1}^{500}\sum_{j=1}^{100}\overline{\mathrm{RMSE}}_{j,k}/50,000$, $\log \mathrm{S} = \sum_{k=1}^{500}\sum_{j=1}^{100}\overline{\log S}_{j,k}/50,000$ and CRPS $= \sum_{k=1}^{500}\sum_{j=1}^{100}\overline{\mathrm{CRPS}}_{j,k}/50,000$ for each of the four covariance models. As expected, the best prediction performances are achieved with model 4) (*i.e.* the true $\mathcal{MT}$ model). However, it can be appreciated that the prediction performances of the $\mathcal{RGW}$ models are very similar, in particular when increasing the fixed μ parameter.

By taking advantage of the sparsity of the covariance matrices associated with the $\sigma^2\mathcal{RGW}_{\kappa,\mu,\beta}$ model for $\mu = 1.75,\ 2.25,\ 3.25$, the computation of the Cholesky factor needed for the computation of the optimal linear predictor can be speeded-up by a factor of approximately 27, 11 and 8, respectively, in this numerical experiment (we use the R package `spam` [25] for the computation of the Cholesky factor for sparse matrices) with respect to the $\mathcal{MT}$ model. The comparison involves the times expressed in seconds computed 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.

These numerical results show that the prediction of a RF with a globally supported $\sigma^2\mathcal{MT}_{\kappa+0.5,\beta}$ covariance model can be performed using a more computationally convenient $\sigma^2\mathcal{RGW}_{\kappa,\mu,\beta}$ covariance model (fixing an arbitrary small value of μ) without a significant loss of prediction efficiency.

4.2. Fixed domain asymptotics

Under fixed domain asymptotics, the $(\mathcal{R})\mathcal{GW}$ model parameters cannot be estimated consistently, as in $\mathcal{MT}$ case [60]. Only the microergodic parameters (see [41] and [2] for a formal definition) can be estimated consistently. If we assume μ fixed and known and $d = 1, 2, 3$, then [6] showed that the microergodic parameter of the $\sigma^2(\mathcal{R})\mathcal{GW}_{\kappa,\mu,\beta}$ model is given by $\sigma^2/\beta^{1+2\kappa}$ for $\kappa \geq 0$. In addition, they studied the asymptotic distribution of the ML estimator of this parameter (assuming a known smoothness parameter κ). These results can easily be generalized to the proposed smoothness parameter extension $\kappa > -0.5$.

If we assume μ unknown, then we need to distinguish between the parameterizations $\sigma^2\mathcal{GW}_{\kappa,\mu,\beta}$ and $\sigma^2\mathcal{RGW}_{\kappa,\mu,\beta}$. In the former case, the associated microergodic parameter is given by

$$m(\sigma^2, \beta, \mu, \kappa) := \frac{\sigma^2}{\beta^{2\kappa+1}} \left(\frac{\Gamma(2\kappa + \mu + 1)}{\Gamma(\mu)} \right).$$

It can be obtained from the following result that establishes the equivalence of two zero mean Gaussian measures with two different $\mathcal{GW}$ covariance models. We omit the proof since it uses the same arguments as in [6].

Theorem 2. *For a given $\kappa \geq -0.5$, consider two zero mean Gaussian measures with covariance function $\sigma_0^2\mathcal{GW}_{\kappa,\mu_0,\beta_0}$ and $\sigma_1^2\mathcal{GW}_{\kappa,\mu_1,\beta_1}$ and let $\mu_i > d+\kappa+1/2$. For any bounded set $D \subset \mathbb{R}^d$, $d = 1, 2, 3$, the two measures are equivalent on the paths of $\{Z(\boldsymbol{s}) : \boldsymbol{s} \in D\}$, if, and only if,*

$$m(\sigma_0^2, \beta_0, \mu_0, \kappa) = m(\sigma_1^2, \beta_1, \mu_1, \kappa). \tag{4.3}$$

Note that Theorem 2 takes into account the proposed extension of the $\mathcal{GW}$ model, that is $\kappa \geq -0.5$. The condition (4.3) fixes the condition in Theorem 3 of [5], which is partially wrong since it works only when $\kappa = 0$. Applying Theorem 2 to the second parameterization $\sigma^2\mathcal{RGW}_{\kappa,\mu,\beta}$, that is replacing β_i with $\delta(\kappa, \mu_i, \beta_i)$, $i = 1, 2$, in (4.3), it turns out that the microergodic parameter is given by $\sigma^2/\beta^{1+2\kappa}$. This implies that, under the parameterization $\sigma^2\mathcal{RGW}_{\kappa,\mu,\beta}$, the parameters $\sigma^2/\beta^{1+2\kappa}$ and μ are microergodic, while for the original parameterization $\sigma^2\mathcal{GW}_{\kappa,\mu,\beta}$ the parameter $m(\sigma^2, \beta, \mu, \kappa)$ is microergodic.

As an example, we simulated 500 realizations observed at $n = 100$ locations uniformly distributed in the unit square from a zero mean Gaussian RF with covariance model $\sigma^2\mathcal{RGW}_{\kappa,\mu,\beta}$ by setting $\kappa = -0.25$ $\mu = 2.25$, $\beta = 0.6$ and

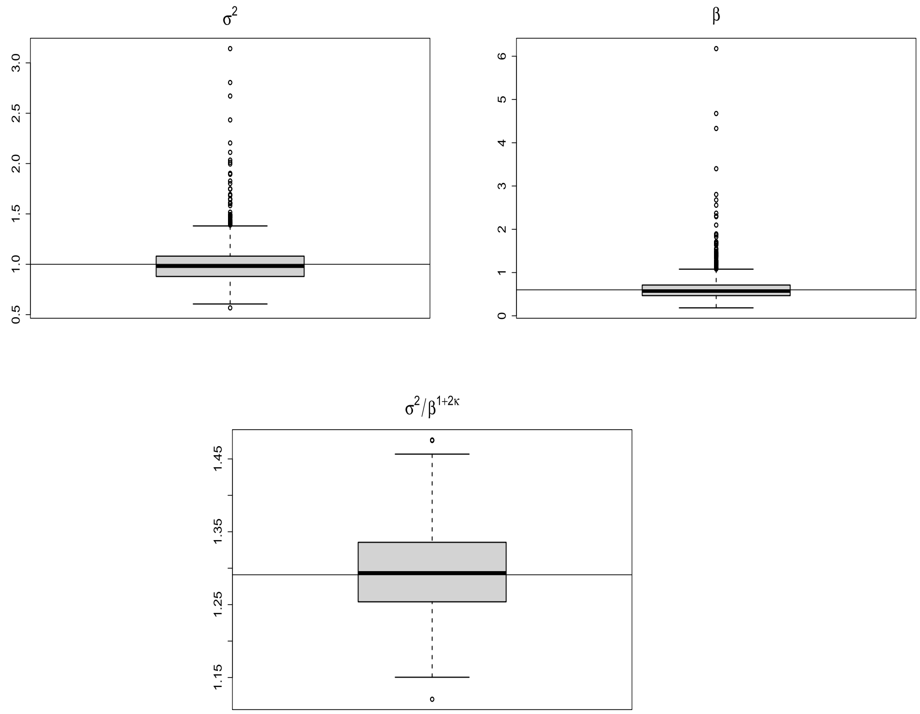


Fig 8. *Boxplots of the ML estimates of σ^2 (top left), β (top right) and of the microergodic parameter $\sigma^2/\beta^{1+2\kappa}$ (bottom) for the covariance model $\sigma^2\mathcal{RGW}_{\kappa,2.25,\beta}$. The true parameters values are $\sigma^2 = 1$, $\beta = 0.6$. The smoothness parameter is $\kappa = -0.25$*

then we estimated σ^2 and β with ML assuming κ and μ fixed and known. The microergodic parameter in this case is given by $\sigma^2/\beta^{1+2\kappa}$ and the associated ML estimator $\hat{\sigma}^2_{ML}/\hat{\beta}^{1+2\kappa}_{ML}$ is consistent and asymptotically Gaussian [6]. Note that we choose $\beta = 0.6$ such that the compact support $\delta(-0.25, 2.25, 0.6) = 1.21$ is large with respect to the spatial extent of the sampling region and, as a consequence, we expect that the fixed domain asymptotics results provide an accurate description of the behavior of the ML estimate of the microergodic parameter. Figure 8 displays the boxplots of the ML estimates of the individual parameter σ^2 and β and of the microergodic parameter $\sigma^2/\beta^{1+2\kappa}$. The empirical distributions of the individual parameters estimates are not Gaussian and show a high variability. However, the distribution of the microergodic parameter estimate is clearly Gaussian, as expected.

5. Application to yearly total precipitation anomalies

We now consider a dataset of yearly total precipitation anomalies registered at $7,352$ location sites in the USA since 1895 to 1997. The yearly totals have been standardized by the long-run mean and standard deviation for each station from

Table 4

NNWCL estimates for parameters of covariance model 1) $\sigma^2\mathcal{RGW}_{\kappa,1.75,\beta}$; 2) $\sigma^2\mathcal{RGW}_{\kappa,2.25,\beta}$; 3) $\sigma^2\mathcal{RGW}_{\kappa,3.25,\beta}$; 4) $\sigma^2\mathcal{MT}_{\kappa+0.5,\beta} \equiv \sigma^2\mathcal{RGW}_{\kappa,\infty,\beta}$ with associated composite likelihood information criteria (CLIC). Prediction performance measures RMSE, LSCORE, and CRPS, as well as the estimated compact support $\delta(\hat{\kappa},\mu,\hat{\beta})$, the percentage of zeros in the estimated covariance matrix and the computational time (in seconds) to perform the associated Cholesky decomposition, are also reported.

	$\hat{\sigma}^2$	$\hat{\beta}$	$\hat{\kappa}$	CLIC	RMSE	LSCORE	CRPS	$\delta(\hat{\kappa},\mu,\hat{\beta})$	%	TIME
1)$\sigma^2\mathcal{RGW}_{\kappa,1.75,\beta}$	0.7865 (0.087)	417.66 (90.95)	-0.2524 (0.015)	$457,292$	0.4665	0.6356	0.9328	633.80	0.87	3.17
2)$\sigma^2\mathcal{RGW}_{\kappa,2.25,\beta}$	0.7864 (0.094)	407.5245 (107.57)	-0.2503 (0.015)	$457,703$	0.4661	0.6347	0.9324	821.08	0.80	8.92
3)$\sigma^2\mathcal{RGW}_{\kappa,3.25,\beta}$	0.7863 (0.104)	397.1918 (112.51)	-0.2482 (0.016)	$458,086$	0.4662	0.6349	0.9322	1196.34	0.65	17.23
4)$\sigma^2\mathcal{MT}_{\kappa+0.5,\beta}$	0.7860 (0.103)	376.07 (129.93)	-0.2426 (0.018)	$458,175$	0.4663	0.6352	0.9317	∞	0	64.81

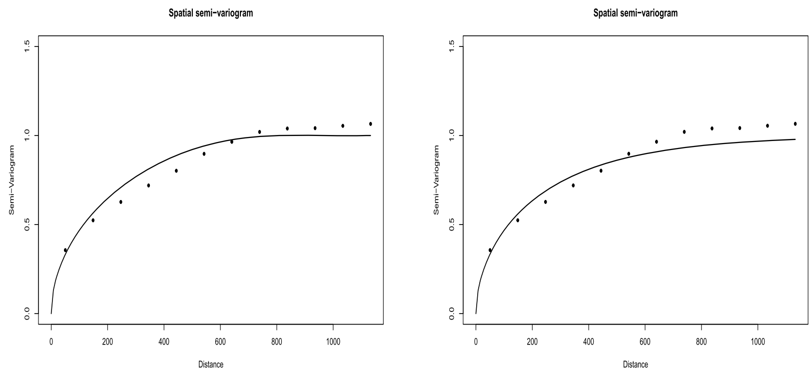
1962. A previous analysis in [36] adapted a zero-mean Gaussian random field with an exponential covariance model, while [4] considered an additional nugget effect by fixing the smoothness parameter $\kappa = 0.5$ for the $\mathcal{MT}$ model and $\kappa = 0$ for the $\mathcal{RGW}$ model.

Here we present an improved analysis by estimating the smoothness parameter of a zero mean Gaussian RF with $\mathcal{RGW}$ covariance model and $\mathcal{MT}$ as a special limit case. Specifically, as in Section 4.1, we consider a $\mathcal{RGW}$ covariance model fixing increasing values of the μ parameters that is: 1) $\sigma^2\mathcal{RGW}_{\kappa,1.75,\beta}$; 2) $\sigma^2\mathcal{RGW}_{\kappa,2.25,\beta}$; 3) $\sigma^2\mathcal{RGW}_{\kappa,3.25,\beta}$; 4) $\sigma^2\mathcal{MT}_{\kappa+0.5,\beta} \equiv \sigma^2\mathcal{RGW}_{\kappa,\infty,\beta}$. The nugget effect has not been considered because, when estimating the smoothness parameter, its estimates is very close to zero. The estimation is performed using the NNWCL method, as in Section 4.1, using 30 neighbors.

Table 4 depicts the NNWCL estimates of σ^2, β and κ with associated standard error (computed using parametric bootstrap) for the covariance models 1), 2), 3) and 4). It is important to note that the estimate of the smoothness parameter κ is negative, that is the extension of the $\mathcal{RGW}$ model proposed in this paper is crucial when analyzing this dataset. For each model the value of the composite likelihood information criteria (CLIC) [52] is indicated. Following this model selection criteria, the $\mathcal{RGW}$ models with fixed μ are preferred to the $\mathcal{MT}$ model.

Table 4 also reports the estimated compact support $\delta(\hat{\kappa},\mu,\hat{\beta})$ that increases when increasing μ, as expected, and the percentage of zero entries in the estimated covariance matrix, for each model. The computational gains that can be achieved using the $\mathcal{RGW}$ model with small values of μ when computing the optimal linear kriging are considerable. In particular, Table 4 shows the time needed for the computation of the Cholesky factor of the estimated covariance matrix. It can be appreciated that the computation of the Cholesky factor is speeded up to a factor of 21, 17 and 3, approximately, for the models 1), 2), 3) with respect to the $\mathcal{MT}$ model 4).

Finally, to compare the models in terms of prediction performance, we used leave-one-out cross-validation as described in [61]. In particular the authors show that RMSE, LSCORE and CRPS leave-one-out cross-validation can be computed in just one step by using the estimated covariance matrix. The pre-



(a) (b)

Fig 9. *Empirical semivariogram of precipitation anomalies data versus the estimated semivariogram using a) the $\sigma^2\mathcal{RGW}_{\kappa,2.25,\beta}$ covariance model, and b) the $\sigma^2\mathcal{MT}_{\kappa+0.5,\beta}$ covariance model.*

diction scores are reported in Table 4 for each covariance model. The performance prediction of the compactly supported models 1), 2), 3) are very similar to the $\mathcal{MT}$ model 4). In particular, the model with lower RMSE and LSCORE is the compactly supported model 2) $\sigma^2\mathcal{RGW}_{\kappa,2.25,\beta}$. These results shows that, accounting for the prediction performance and the computational complexity, the best model is $\sigma^2\mathcal{RGW}_{\kappa,2.25,\beta}$.

Figure 9 compares the empirical semivariogram of the yearly total precipitation anomalies data with the estimated semivariogram using the covariance model $\sigma^2\mathcal{RGW}_{\kappa,2.25,\beta}$ and the $\sigma^2\mathcal{MT}_{\kappa+0.5,\beta}$ covariance model.

6. Conclusions

The (reparameterized) generalized Wendland correlation model can be viewed as a generalization of the Matérn model. In this paper, we have provided new necessary and sufficient conditions for this class of model, extending the parametric space of the smoothness parameter to the interval $(-0.5, 0)$. The proposed new conditions allows the generalized Wendland model to cover the full range of validity of the smoothness parameter of the Matérn model.

As a result, the generalized Wendland model can be used as a more flexible alternative of the Matérn model or as a computational convenient approximation of the Matérn model, even when the sample paths are very rough. Some numerical evidences have been provided in a simulation study and in a real data application.

Although the generalized Wendland model involves the computation of a Gauss hypergeometric function or the evaluation of a specific integral, in some special cases (Section 3 and Appendix A.1 to A.2) the computation can be considerably simplified. This can be very useful for practitioners.

More generally, the numerical computation of the generalized Wendland model can be performed through efficient implementation of the Gauss hypergeometric function as in the R package `GeoModels` [8] or using adaptive integration as in the R package `GeneralizedWendland` [22]. As an alternative, an approximation of the generalized Wendland model can be obtained using a computationally efficient polynomial approximation [24]. Finally, the proposed result can be applied to other classes of models, such as the space-time Wendland model proposed in [43] and the multivariate Wendland model in [16].

Appendix A

A.1. Closed-form expression of the ordinary Wendland functions

Let $\kappa \geq 0$, $\beta > 0$. Using formula 9.6.5 of [37], the generalized Wendland covariance (3.1) can be rewritten as:

$$\begin{aligned}\mathcal{GW}_{\kappa,\mu,\beta}(x) = M \left(1 - \frac{x^2}{\beta^2}\right)_+^{\kappa+\mu} \left(\frac{\beta + x}{2\beta}\right)^{-\mu} \\ \times {}_2F_1\left(\mu, -\kappa; \kappa + \mu + 1; \frac{(\beta - x)_+}{\beta + x}\right), \quad x \geq 0.\end{aligned}$$

If, furthermore, $\kappa = k \in \mathbb{N}$, then the hypergeometric function in the above expression is a terminating series. The generalized Wendland covariance then reads as

$$\begin{aligned}\mathcal{GW}_{k,\mu,\beta}(x) &= M \left(1 - \frac{x^2}{\beta^2}\right)_+^{k+\mu} \left(\frac{\beta + x}{2\beta}\right)^{-\mu} \sum_{n=0}^{k} \frac{(\mu)_n (-k)_n}{(k + \mu + 1)_n \, n!} \left(\frac{(\beta - x)_+}{\beta + x}\right)^n \\ &= M 2^\mu \sum_{n=0}^{k} \frac{(\mu)_n (-k)_n}{(k + \mu + 1)_n \, n!} \left(1 - \frac{x}{\beta}\right)_+^{n+k+\mu} \left(1 + \frac{x}{\beta}\right)^{k-n}, \\ &= \left(1 - \frac{x}{\beta}\right)_+^{k+\mu} P_k(x; \mu, \beta), \quad x \geq 0,\end{aligned}$$

with

$$P_k(x; \mu, \beta) = M 2^\mu \sum_{n=0}^{k} \frac{(\mu)_n (-k)_n}{(k + \mu + 1)_n \, n!} \left(1 - \frac{x}{\beta}\right)_+^{n} \left(1 + \frac{x}{\beta}\right)^{k-n}, \quad x \geq 0.$$

$P_k(\cdot; \mu, \beta)$ is a polynomial of degree k whose coefficients depend on k, μ and β. If, furthermore, μ is an integer, $\mathcal{GW}_{k,\mu,\beta}$ is a polynomial of x of degree $\mu + 2k$ in the interval $[0, \beta]$ (Figure 1). The expression of $\mathcal{GW}_{k,\mu,\beta}$ can be shown to be equivalent to the expression in (3.5) obtained using the results in [33], but our proof is much more straightforward.

A.2. Closed-form expression of the missing Wendland functions

Using formula 7.3.1.100 in [45], one can rewrite the generalized Wendland function at $x \in (0, \beta)$ as

$$\begin{aligned}\mathcal{GW}_{\kappa,\mu,\beta}(x) &= M\left(1-\frac{x^2}{\beta^2}\right)^{\kappa+\mu} {}_2F_1\left(\frac{\mu}{2}, \frac{\mu+1}{2}; \kappa+\mu+1; \left(1-\frac{x^2}{\beta^2}\right)\right) \\ &= M2^{\kappa+\mu}\Gamma(\kappa+\mu+1)\left(1-\frac{x^2}{\beta^2}\right)^{\frac{\kappa+\mu}{2}}\left(\frac{x}{\beta}\right)^{\kappa} P_{-\kappa-1}^{-\kappa-\mu}\left(\frac{\beta}{x}\right),\end{aligned}$$

where P_κ^λ stands for the associated Legendre function of the first kind of degree κ and order λ. This function can be transformed into an associated Legendre function of the second kind $Q_{-\lambda-\frac{1}{2}}^{-\kappa-\frac{1}{2}}$ using Whipple's formula [42, formulae 14.3.10, 14.9.14 and 14.9.17], leading to:

$$\begin{aligned}\mathcal{GW}_{\kappa,\mu,\beta}(x) = \frac{M2^{\kappa+\mu+\frac{1}{2}}\Gamma(\kappa+\mu+1)}{\sqrt{\pi}\Gamma(2\kappa+\mu+1)}\left(1-\frac{x^2}{\beta^2}\right)^{\frac{2\kappa+2\mu-1}{4}}\left(\frac{x}{\beta}\right)^{\kappa+\frac{1}{2}} \\ \times e^{-\mathrm{i}(\kappa+\frac{1}{2})\pi} Q_{\kappa+\mu-\frac{1}{2}}^{\kappa+\frac{1}{2}}\left[\left(1-\frac{x^2}{\beta^2}\right)^{-\frac{1}{2}}\right], \quad 0<x<\beta,\end{aligned} \tag{A.1}$$

with i the imaginary unit.

From the following identities [45, p. 775–777], valid for $z > 1$ and any nonnegative integer ν:

$$Q_\nu^0(z) = \frac{1}{2}P_\nu(z)\log\left(\frac{z+1}{z-1}\right) - \sum_{p=1}^{\nu}\frac{1}{p}P_{p-1}(z)P_{\nu-p}(z),$$

where P_ν is the Legendre polynomial of degree ν, and

$$\sqrt{z^2-1}Q_{\nu+1}^{\lambda+1}(z) = (\nu+1-\lambda)zQ_{\nu+1}^{\lambda}(z) - (\nu+1+\lambda)Q_\nu^\lambda(z),$$

one can obtain a closed-form expression of Q_ν^λ on $(1, +\infty)$ for every nonnegative integers λ and ν such that $\nu \geq \lambda$, therefore a closed-form expression of the generalized Wendland covariance (A.1) for κ equal to a positive half-integer and μ equal to a positive integer. For instance, for $\lambda = 1$ and ν a positive integer, one has:

$$\begin{aligned}\sqrt{z^2-1}Q_\nu^1(z) &= \frac{\nu}{2}\left(zP_\nu(z) - P_{\nu-1}(z)\right)\log\left(\frac{z+1}{z-1}\right) - zP_{\nu-1}(z) \\ &\quad - \nu\sum_{p=1}^{\nu-1}\frac{1}{p}P_{p-1}(z)(zP_{\nu-p}(z) - P_{\nu-1-p}(z)),\end{aligned}$$

which gives, for $0 < x < \beta$ and μ a positive integer,

$$\mathcal{GW}_{\frac{1}{2},\mu,\beta}(x) = \frac{M 2^{\mu+1}\Gamma(\mu+\frac{3}{2})}{\sqrt{\pi}\Gamma(\mu+2)} z^{-\mu-1}\Big[\frac{\mu}{2}\left(P_{\mu-1}(z) - zP_\mu(z)\right)\log\left(\frac{z+1}{z-1}\right) + zP_{\mu-1}(z) + \mu\sum_{p=1}^{\mu-1}\frac{1}{p}P_{p-1}(z)(zP_{\mu-p}(z) - P_{\mu-1-p}(z))\Big], \tag{A.2}$$

with $z = \left(1-\frac{x^2}{\beta^2}\right)^{-\frac{1}{2}}$ and $P_\mu(z) = 2^{-\mu}\sum_{p=0}^{\lfloor\frac{\mu}{2}\rfloor}\frac{(-1)^p(2\mu-2p)!}{p!(\mu-p)!(\mu-2p)!}z^{\mu-2p}$, $z\in\mathbb{R}$, is the Legendre polynomial of degree μ.

In the general case, for any $z\in\mathbb{R}$ and any positive integers λ and ν such that $\nu\geq\lambda$, one finds [51, eqs. 2.7, 2.31 and 4.10]

$$\begin{aligned}Q_\nu^\lambda(z) &= \frac{1}{2}P_\nu^\lambda(z)\log\left(\frac{z+1}{z-1}\right)\\ &\quad-\frac{(\nu+\lambda)!}{2(\nu-\lambda)!}\sum_{p=0}^{\lambda-1}\frac{(-1)^p(2p+1)}{(\nu-p)(p+\nu+1)}(P_p^{-\lambda}(-z) - (-1)^\nu P_p^{-\lambda}(z))\\ &\quad-\sum_{p=0}^{\nu-\lambda-1}\frac{(1-(-1)^{p+\nu+\lambda})(2p+2\lambda+1)}{2(\nu-\lambda-p)(p+\nu+\lambda+1)}\left[1+\frac{p!(\nu+\lambda)!}{(p+2\lambda)!(\nu-\lambda)!}\right]P_{p+\lambda}^\lambda(z),\end{aligned} \tag{A.3}$$

with

$$P_\tau^\lambda(z) = \frac{\Gamma(\tau+\lambda+1)}{\Gamma(\tau-\lambda+1)}\left(\frac{z+1}{z-1}\right)^{\frac{\lambda}{2}}\sum_{p=0}^{\tau}\frac{(-1)^{p+\tau}(p+\tau)!(z+1)^p}{p!(\tau-p)!\Gamma(p+\lambda+1)2^p}, \quad 0\leq\lambda\leq\tau\in\mathbb{N},$$

and

$$P_p^{-\lambda}(z) = \left(\frac{z-1}{z+1}\right)^{\frac{\lambda}{2}}\sum_{q=0}^{p}\frac{(q+p)!(z-1)^q}{q!(q+\lambda)!(p-q)!2^q}, \quad \lambda > p\in\mathbb{N}.$$

Plugging (A.3) into (A.1) gives a closed-form expression of the missing Wendland functions for any positive integer μ and half-integer κ.

Although the associated Legendre function of the second kind in (A.1) tends to infinity as x tends to 0 or β, the value of $\mathcal{GW}_{\kappa,\mu,\beta}(x)$ remains finite, due to the power terms $(\frac{x}{\beta})^{\kappa+1/2}$ and $(1-\frac{x^2}{\beta^2})^{(2\kappa+2\mu-1)/4}$, which make $\mathcal{GW}_{\kappa,\mu,\beta}(x)$ tend to 1 as x tends to zero and to 0 as x tends to β.

A.3. Proof of Theorem 1

Let d be a positive integer, and let β, χ, γ and δ be positive real numbers such that:

1. $\delta > d/2$;

2. $2(\chi-\delta)(\gamma-\delta) \geq \delta$;
3. $2(\chi+\gamma) \geq 6\delta+1$.

Consider the four-parameter Gauss hypergeometric model proposed in [21]:

$$\mathcal{GH}_{\delta,\chi,\gamma,\beta}(x) = \frac{\Gamma(\chi-d/2)\Gamma(\gamma-d/2)}{\Gamma(\chi-\delta+\gamma-d/2)\Gamma(\delta-d/2)}\left(1-\frac{x^2}{\beta^2}\right)_+^{\chi-\delta+\gamma-d/2-1} \times {}_2F_1\left(\chi-\delta;\gamma-\delta;\chi-\delta+\gamma-d/2;\left(1-\frac{x^2}{\beta^2}\right)_+\right), \quad x \geq 0. \quad \text{(A.4)}$$

Based on Theorem 4.2 in [13], and Theorem 3 and Eq. 44 in [59], it can be shown that, under conditions a), b) and c), $\mathcal{GH}_{\delta,\chi,\gamma,\beta}$ belongs to the class Φ_d, and that its spectral density is given by

$$\widehat{\mathcal{GH}}_{\delta,\chi,\gamma,\beta}(z) = Ka^d{}_1F_2\left(\delta;\chi,\gamma;-\frac{(z\beta)^2}{4}\right), \quad z \geq 0, \quad \text{(A.5)}$$

where $K = \frac{\Gamma(\delta)\Gamma(\chi-d/2)\Gamma(\gamma-d/2)}{2^d\pi^{\frac{d}{2}}\Gamma(\delta-d/2)\Gamma(\chi)\Gamma(\gamma)}$ is a normalization constant.

Using the properties of the gamma function, it is straightforward to show that a special case of (A.5) is the spectral density of a reparameterized $\mathcal{GW}$ model, that is

$$\widehat{\mathcal{GH}}_{\delta,\chi,\chi+0.5,\beta}(z) = \widehat{\mathcal{GW}}_{\delta-(d+1)/2,2(\chi-\delta),\beta}(z), \quad z \geq 0.$$

Now we focus on the restriction $d/2 < \delta < d/2+1/2$. If we use the standard parameterization of the $\mathcal{GW}$ model, *i.e.*, if we set $\kappa = \delta-(d+1)/2$ and $\mu = 2(\chi-\kappa)-(d+1)$, it can be appreciated that $-1/2 < \kappa < 0$, that is, we obtain an extension of the parametric space associated with the smoothness parameter of the generalized Wendland model $\mathcal{GW}_{\kappa,\mu,a}$ including the case $-1/2 < \kappa < 0$.

The sufficient conditions, under the restriction $-1/2 < \kappa < 0$, can be obtained putting together the above conditions 1), 2) and 3), by setting $\delta = \kappa+(d+1)/2$, $\chi = (\mu+2\kappa+(d+1))/2$ and $\gamma = \chi+0.5$. Solving the system of inequalities, we obtain the following sufficient conditions:

1. $\mu \geq (d+1)/2+\kappa$ when $d \geq 2$
2. $\mu \geq 0.5(\sqrt{8\kappa+9}-1)$ when $d = 1$.

for $\mathcal{GW}_{\kappa,\mu,\beta}$ to belong to the class Φ_d when $-1/2 < \kappa < 0$.

Necessary conditions can be obtained following Cho and Yun [14, Theorem 6.1], who show that if ${}_1F_2(a;b;c;-\frac{x^2}{4}) \geq 0$, then $b > a$, $c > a$, $b+c \geq 3a+0.5$. Taking into account the spectral density (A.5) and setting $\delta = \kappa+(d+1)/2$, $\chi = (\mu+2\kappa+(d+1))/2$ and $\gamma = \chi+0.5$, we find the solution of the system inequalities (under the restriction $-1/2 < \kappa < 0$), which is given by $\mu \geq (d+1)/2+\kappa$ when $d \geq 1$.

Finally, noting that, when $d = 1$, $(d+1)/2+\kappa < 0.5(\sqrt{8\kappa+9}-1)$ for $-1/2 < \kappa < 0$, we put together the necessary and sufficient conditions for the cases $d = 1$ and $d \geq 2$, obtaining the main result.

Funding

Moreno Bevilacqua and Xavier Emery acknowledge the funding of the National Agency for Research and Development of Chile, through grants ANID Fondecyt 1240308 (MB), ANID PIA ANILLOS ACT210096 (MB), ANID project Data Observatory Foundation DO210001 (MB), MATH-AMSUD1167 22-MATH-06 (AMSUD220041) (MB), ANID Fondecyt 1210050 (XE) and ANID PIA AFB230001 (XE).

References

[1] Abramowitz, M. and Stegun, I. A., editors (1970). *Handbook of Mathematical Functions*. Dover, New York.

[2] Bachoc, F. (2021). *Asymptotic Analysis of Maximum Likelihood Estimation of Covariance Parameters for Gaussian Processes: An Introduction with Proofs*, pages 283–303. Springer International Publishing, Cham. MR4299285

[3] Banerjee, S., Carlin, B. P., and Gelfand, A. E. (2014). *Hierarchical Modeling and Analysis for Spatial Data (2nd ed.)*. Chapman and Hall/CRC Press, Boca Raton: FL. MR3362184

[4] Bevilacqua, M., Caamaño-Carrillo, C., and Porcu, E. (2022). Unifying compactly supported and Matérn covariance functions in spatial statistics. *Journal of Multivariate Analysis*, 189:104949. MR4375672

[5] Bevilacqua, M. and Faouzi, T. (2019). Estimation and prediction of Gaussian processes using generalized Cauchy covariance model under fixed domain asymptotics. *Electronic Journal of Statistics*, 13(2):3025–3048. MR4010591

[6] Bevilacqua, M., Faouzi, T., Furrer, R., and Porcu, E. (2019). Estimation and prediction using Generalized Wendland covariance functions under fixed domain asymptotics. *Annals of Statistics*, 47(2):828–856. MR3909952

[7] Bevilacqua, M. and Gaetan, C. (2014). Comparing composite likelihood methods based on pairs for spatial Gaussian random fields. *Statistics and Computing*, pages 1–16. MR3375623

[8] Bevilacqua, M., Morales-Oñate, V., and Caamaño-Carrillo, C. (2024). *GeoModels: Procedures for Gaussian and Non Gaussian Geostatistical (Large) Data Analysis*. R package version 2.0.4.

[9] Caamaño-Carrillo, C., Bevilacqua, M., López, C., and Morales-Oñate, V. (2024). Nearest neighbors weighted composite likelihood based on pairs for (non-)Gaussian massive spatial data with an application to Tukey-hh random fields estimation. *Computational Statistics and Data Analysis*, 191:107887. MR4667247

[10] Chernih, A. and Hubbert, S. (2014). Closed form representations and properties of the generalised Wendland functions. *Journal of Approximation Theory*, 40:17–33. MR3128401

[11] Chernih, A., Sloan, I. H., and Womersley, R. S. (2014). Wendland functions

with increasing smoothness converge to a Gaussian. *Advances in Computational Mathematics*, 40(1):185–200. MR3158020

[12] Chilès, J.-P. and Delfiner, P. (2012). *Geostatistics: Modeling Spatial Uncertainty, Second Edition*. John Wiley and Sons, New York. MR2850475

[13] Cho, Y.-K., Chung, S.-Y., and Yun, H. (2020). Rational extension of the Newton diagram for the positivity of ${}_1F_2$ hypergeometric functions and Askey–Szegö problem. *Constructive Approximation*, 51(1):49–72. MR4045736

[14] Cho, Y.-K. and Yun, H. (2018). Newton diagram of positivity for ${}_1F_2$ generalized hypergeometric functions. *Integral Transforms and Special Functions*, 29(7):527–542. MR3806555

[15] Cressie, N. (1993). *Statistics for Spatial Data*. Wiley, New York, revised edition. MR1239641

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

[17] Datta, A., Banerjee, S., Finley, A. O., and Gelfand, A. E. (2016). Hierarchical nearest-neighbor Gaussian process models for large geostatistical datasets. *Journal of the American Statistical Association*, 111(514):800–812. MR3538706

[18] Davis, T., editor (2006). *Direct Methods for Sparse Linear Systems*. Society for Industrial and Applied Mathematics, Philadelphia. MR2270673

[19] Duff, I. S., Erisman, A. M., and Reid, J. K., editors (1989). *Direct Methods for Sparse Matrices*. Clarendon, New York. MR1029273

[20] Eidsvik, J., Shaby, B., Reich, B., Wheeler, M., and Niemi, J. (2014). Estimation and prediction in spatial models with block composite likelihoods. *Journal of Computational and Graphical Statistics*, 23(2):295–315. MR3215812

[21] Emery, X. and Alegría, A. (2022). The Gauss hypergeometric covariance kernel for modeling second-order stationary random fields in Euclidean spaces: Its compact support, properties and spectral representation. *Stochastic Environmental Research and Risk Assessment*, 36:2819–2834.

[22] Fischer, T. C., Furrer, R., and Stocker, J. (2022). *GeneralizedWendland: Fully Parameterized Generalized Wendland Covariance Function*. R package version 0.5-2.

[23] 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. MR2291261

[24] Furrer, R. and Hediger, M. (2023). Asymptotic analysis of ML-covariance parameter estimators based on covariance approximations. *Electronic Journal of Statistics*, 17(2):3050–3102. MR4667731

[25] 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.

[26] Gaetan, C. and Guyon, X. (2009). *Spatial Statistics and Modeling*. Springer. MR2569034

[27] Gneiting, T. (2002). Compactly supported correlation functions. *Journal of Multivariate Analysis*, 83:493–508. MR1945966

[28] Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102:359–378. MR2345548

[29] Gneiting, T., Sevcikova, H., and Percival, D. B. (2011). Estimators of fractal dimension: Assessing the roughness of time series and spatial data. *Statistical Science*, 27:247–277. MR2963995

[30] Guinness, J. (2018). Permutation and grouping methods for sharpening Gaussian process approximations. *Technometrics*, 60(4):415–429. MR3878098

[31] Guttorp, P. and Gneiting, T. (2006). Studies in the history of probability and statistics XLIX: On the Matérn correlation family. *Biometrika*, 93(4):989–995. MR2285084

[32] Heaton, M. J., Datta, A., Finley, A. O., Furrer, R., Guinness, J., Guhaniyogi, R., Gerber, F., Gramacy, R. B., Hammerling, D., Katzfuss, M., Lindgren, F., Nychka, D. W., Sun, F., and Zammit-Mangion, A. (2019). A case study competition among methods for analyzing large spatial data. *Journal of Agricultural, Biological and Environmental Statistics*, 24(3):398–425.

[33] Hubbert, S. (2012). Closed form representations for a class of compactly supported radial basis functions. *Advances in Computational Mathematics*, 36(1):115–136. MR2886183

[34] Katzfuss, M. (2017). A multi-resolution approximation for massive spatial datasets. *Journal of the American Statistical Association*, 112(517):201–214. MR3646566

[35] Katzfuss, M. and Guinness, J. (2021). A general framework for Vecchia approximations of Gaussian processes. *Statistical Science*, 36:124–141. MR4194207

[36] Kaufman, C. G., Schervish, M. J., and Nychka, D. W. (2008). Covariance tapering for likelihood-based estimation in large spatial data sets. *Journal of the American Statistical Association*, 103:1545–1555. MR2504203

[37] Lebedev, N. (1965). *Special Functions and their Applications*. Prentice-Hall, London. MR0174795

[38] Lindgren, F., Rue, H., and Lindström, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B*, 73:423–498. MR2853727

[39] Mardia, K. V. and Marshall, J. (1984). Maximum likelihood estimation of models for residual covariance in spatial regression. *Biometrika*, 71:135–146. MR0738334

[40] Matèrn, B. (1986). *Spatial Variation: Stochastic Models and their Applications to Some Problems in Forest Surveys and Other Sampling Investigations*. Springer, Heidelberg, 2nd edition. MR0169346

[41] Matheron, G. (1989). *Estimating and Choosing*. Springer, New York. MR0979761

[42] Olver, F. W., Lozier, D. M., Boisvert, R. F., and Clark, C. W. (2010). *NIST Handbook of Mathematical Functions*. Cambridge University Press, Cambridge. MR2723248

[43] Porcu, E., Bevilacqua, M., and Genton, M. G. (2020). Nonseparable, space-time covariance functions with dynamical compact supports. *Statistica Sinica*, 30(2):719–739. MR4214159

[44] Porcu, E., Bevilacqua, M., Schaback, R., and Oates, C. J. (2024). The Matérn model: A journey through statistics, numerical analysis and machine learning. *Statistical Science*, 39(3):469–492. MR4766860

[45] Prudnikov, A., Brychkov, Y. A., and Marichev, O. (1990). *Integrals and Series: More Special Functions*, volume 3. Gordon and Breach Science Publishers. MR1054647

[46] Santner, T., Williams, B. J., and Notz, W. I. (2003). *The Design and Analysis of Computer Experiments*. Springer, New York. MR2160708

[47] Schaback, R. (2011). The missing Wendland functions. *Advances in Computational Mathematics*, 34(1):67–81. MR2783302

[48] Schoenberg, I. J. (1938). Metric spaces and completely monotone functions. *Annals of Mathematics*, 39:811–841. MR1503439

[49] Stein, M. L. (1999). *Interpolation of Spatial Data. Some Theory of Kriging*. Springer, New York. MR1697409

[50] Stein, M. L. (2013). Statistical properties of covariance tapers. *Journal of Computational and Graphical Statistics*, 22(4):866–885. MR3173747

[51] Szmytkowski, R. (2009). On the derivative of the associated Legendre function of the first kind of integer degree with respect to its order (with applications to the construction of the associated Legendre function of the second kind of integer degree and order). *Journal of Mathematical Chemistry*, 46:231–260. MR2598493

[52] Varin, C. and Vidoni, P. (2005). A note on composite likelihood inference and model selection. *Biometrika*, 92(3):519–528. MR2202643

[53] Vecchia, A. V. (1988). Estimation and model identification for continuous spatial processes. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 297–312. MR0964183

[54] Wendland, H. (1995). Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Advances in Computational Mathematics*, 4:389–396. MR1366510

[55] Wendland, H. (2004). *Scattered Data Approximation*. Cambridge. MR2131724

[56] Whittle, P. (1954). On stationary processes in the plane. *Biometrika*, 41(3-4):434–449. MR0067450

[57] Williams, C. K. and Rasmussen, C. E. (2006). *Gaussian processes for machine learning*. MIT press Cambridge, MA. MR2514435

[58] Yaglom, A. M. (1987). *Correlation Theory of Stationary and Related Random Functions. Volume I: Basic Results*. Springer, New York. MR0893393

[59] Zastavnyi, V. P. (2006). On some properties of Buhmann functions. *Ukrainian Mathematical Journal*, 58(8):1184–1208. MR2345078

[60] Zhang, H. (2004). Inconsistent estimation and asymptotically equivalent

interpolations in model-based geostatistics. *Journal of the American Statistical Association*, 99:250–261. MR2054303

[61] Zhang, H. and Wang, Y. (2010). Kriging and cross-validation for massive spatial data. *Environmetrics*, 21(3-4):290–304. MR2842244