Statistica Sinica Preprint No: SS-2020-0010

Title	Space-Time Estimation and Prediction under Infill Asymptotics with Compactly Supported Covariance Functions
Manuscript ID	SS-2020-0010
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202020.0010
Complete List of Authors	Tarik Faouzi, Emilio Porcu and Moreno Bevilacqua
Corresponding Author	Emilio Porcu
E-mail	georgepolya01@gmail.com
Notice: Accepted version subject to English editing.	

Space-Time Estimation and Prediction under fixed-domain Asymptotics with Compactly Supported Covariance Functions

Tarik Faouzi[1], Emilio Porcu[2], and Moreno Bevilacqua[3].

Abstract

We study estimation and prediction of Gaussian processes with space-time covariance models belonging to the *Dynamical Generalized Wendland* family ($\mathcal{DGW}$) (Porcu et al., 2020, Statistica Sinica), under fixed-domain asymptotics. Such a class is nonseparable, has dynamical compact supports, and parameterizes differentiability at the origin similarly to the space-time Matérn class (Ip and Li, 2017, Statistica Sinica).

The results of the paper are classified into two parts. In the first part, we establish strong consistency and asymptotic normality for the maximum likelihood estimator of the microergodic parameter associated to the $\mathcal{DGW}$ covariance model, under fixed-domain asymptotics. The last part focuses on optimal kriging prediction under the $\mathcal{DGW}$ model and asymptotically correct estimation of mean square error using a misspecified model. Theoretical results are in turn based on equivalence of Gaussian measures under some given families of space-time covariance functions, where both space or time are compact. Such technical results are provided in the Online Supplement to this paper.

Keywords: fixed-domain asymptotics; Microergodic parameter; Maximum likelihood, Space-Time Generalized Wendland family.

[1] Department of Statistics, Universidad del BioBio, 4081112 Concepcion, Chile
E-mail:tfaouzi@ubiobio.cl

[2] School of Computer Science and Statistics, Trinity College Dublin.
and Millennium Nucleus Center for the Discovery of Structures in Complex Data, Chile
E-mail: georgepolya01@gmail.com, porcue@tcd.ie

[3] Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibáñez, Viña del Mar, Chile.
Millennium Nucleus Center for the Discovery of Structures in Complex Data, Chile
E-mail: moreno.bevilacqua@uai.cl

1. Introduction

1.1 Context and State of the Art

This paper is concerned with fixed-domain asymptotics for estimation and kriging prediction of Gaussian random fields defined over product spaces $D \times \mathcal{T}$, where D is a subset of $\mathbb{R}^d$ (d is a positive integer) and $\mathcal{T}$ is a compact interval of the real line. The most notable application refers to D as spatial domain and $\mathcal{T}$ as time. Although we focus on the space-time case, our results can be analogously applied to the anisotropic spatial case where the rate of decay in correlation in one coordinate is different from that of the remaining d coordinates. Here we assume that the process is observed at n (possibly unevenly spaced) locations and repeatedly over m time points.

There might be other choices for space-time asymptotics: for instance, for the temporal part one might consider an increasing asymptotic framework while keeping a fixed-domain approach for the spatial part. We are not aware of any contribution of this type, and such a setting looks challenging. Alternatively, one might consider both space and time under an increasing domain fashion. In this case, the results of Mardia and Marshall (1984) on maximum likelihood estimation (ML throughout) would apply, and space-time asymptotics becomes a straightforward extension of the results obtained in the spatial case.

Instead, there is a lack of general results for the case of fixed-domain asymptotics. Some results have been given for specific classes of covariance functions. For instance, Zhang (2004), Wang and Loh (2011) and Kaufman and Shaby (2013) have studied the asymptotic properties of the ML estimation of the microergodic parameter of Matérn covariance model. Additionally, Stein (1999) and Kaufman and Shaby (2013) have studied the asymptotic effect of the misspecified kriging prediction on the prediction variance, under the Matérn covariance model. Recently, Bevilacqua et al. (2019) have considered a fixed-domain asymptotic framework for Gaussian random fields defined over a compact set of $\mathbb{R}^d$ under the Generalized Wendland ($\mathcal{GW}$ throughout) class of compactly supported correlation functions (Zastavnyi and Trigub, 2002). Bevilacqua and Faouzi (2019) have explored a similar problem using the generalized Cauchy class, which allows for decouping of fractal dimensions with the Hurst effect.

The literature on space-time fixed-domain asymptotics is sparse, with the notable exception of Ip and Li (2017), who do asymptotic analysis on the basis of a class of space-time covariance functions, proposed by Fuentes et al. (2008), having both spatial and temporal margins belonging to the Matérn family (Stein, 1999). Through the paper we call this family the

Dynamical Matérn ($\mathcal{DM}$) family of space-time covariance functions. A recent paper by Porcu et al. (2020) gives a thorough review of space-time covariance functions.

1.2 Our Contribution

This paper considers a class of nonseparable space-time covariance functions proposed by Porcu et al. (2020). The members of this class are dynamically compactly supported, meaning that for any fixed temporal lag, the spatial margin is dynamically compactly supported, that is, there is a decreasing and continuous function, h, such that for every fixed temporal lag, t_o, the spatial margin of the space-time covariance function is compactly supported over a ball with radius $h(t_o)$ embedded in $\mathbb{R}^d$. Specifically, the spatial margin belongs to the $\mathcal{GW}$ class. For the remainder of the paper, we call this class Dynamical $\mathcal{GW}$ class and use the acronym $\mathcal{DGW}$ for it.

We study the problem of ML estimation of the $\mathcal{DGW}$ class defined over the product space $\mathcal{D} \times \mathcal{T}$, under fixed-domain asymptotics. Further, we study the problem of kriging prediction under the same asymptotic framework. Since results on fixed-domain asymptotics largely rely on the equivalence of Gaussian measures (Skorokhod and Yadrenko, 1973), we derive conditions for the equivalence of Gaussian measures under, either, two $\mathcal{DGW}$ families with different parameters, or under a $\mathcal{DGW}$ and a $\mathcal{DM}$ family. Such conditions are provided in Section B of the Online Supplement (OS throughout). We explore the implications of these results in terms of consistency and asymptotic distribution of the ML estimator for the microergodic parameter. Finally, we assess the consequences of previous results in terms of efficiency of the misspecified best linear unbiased predictors.

The plan of the paper is the following: Section 2 contains the necessary mathematical notation, a description of the covariance functions used in this paper, while some background material on equivalence of Gaussian measures is deferred to A in OS. Section 3 provides preliminary results related to the space-time Fourier transforms of both $\mathcal{DM}$ and $\mathcal{DGW}$ models. We also find conditions for equivalence of Gaussian measures under both models (see B in OS) and these results are the basis for Section 4, which studies the problem of consistency and asymptotic normality for the ML estimators of the parameters indexing the $\mathcal{DGW}$ family. The problem of misspecified kriging predictions under the $\mathcal{DGW}$ is then explored in Section 5. A short conclusion finishes the paper. Technical proofs are deferred to Section C in OS.

2. Background Material

2.1 Preliminaries and Notation

We denote by $Z = \{Z(\boldsymbol{s},t),(\boldsymbol{s},t) \in D \times \mathcal{T}\}$ a zero mean Gaussian random field with index set on $D \times \mathcal{T}$, with stationary covariance function $C : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ that is spatially isotropic and temporally symmetric. That is, there exists a continuous function $K : [0,\infty)^2 \to \mathbb{R}$ such that $K(0,0) = 1$ and $C(\mathbf{h},u) = \sigma^2 K(\|\mathbf{h}\|,|u|)$, $(\mathbf{h},u) \in \mathbb{R}^d \times \mathbb{R}$, where σ^2 denotes the variance parameter. Here, $\|\cdot\|$ denotes the Euclidean norm. We denote by $\Phi_{d,T}$ the set of such functions. For the reminder of the paper, we use r for $\|\mathbf{h}\|$ and t for $|u|$. Additionally, we denote with Φ_d the family of spatially isotropic covariance functions defined on $\mathbb{R}^d$. The classes Φ_d and $\Phi_{d,T}$ are nested, with the inclusion relations

$$\Phi_1 \supset \Phi_2 \supset \ldots \supset \Phi_\infty \quad \text{and} \quad \Phi_{1,T} \supset \Phi_{2,T} \supset \ldots \supset \Phi_{\infty,T}$$

being strict, where $\Phi_\infty := \bigcap_{d\geq 1} \Phi_d$ and $\Phi_{\infty,T} := \bigcap_{d\geq 1} \Phi_{d,T}$. There is a rich mathematical theory for both classes Φ_d and $\Phi_{d,T}$. For a recent account on the class Φ_d, the reader is referred to Daley and Porcu (2013), while Porcu et al. (2006) have provided extensive material for the class $\Phi_{d,T}$.

In particular, the results in Porcu et al. (2006) (see also Gneiting and Guttorp, 2010) show that a continuous function ϕ with $\phi(0,0) = 1$ belongs to the class $\Phi_{d,T}$ if and only if there exists a probability measure F, defined on the positive quadrant of $\mathbb{R}^2$ such that

$$K(r,t) = \int_0^\infty \int_0^\infty \Omega_d(r\xi_1)\cos(t\xi_2)F(\mathrm{d}(\xi_1,\xi_2)), \qquad t \geq 0, r \geq 0,$$

where $\Omega_d(t) = t^{-(d-2)/2}J_{(d-2)/2}(t)$ and J_ν is the Bessel function of the first kind of order $\nu > 0$ (Abramowitz and Stegun, 1970). Classical Fourier inversion arguments show that, if K is absolutely integrable, then $K \in \Phi_{d,T}$ if and only if the function $f : [0,\infty)^2 \to \mathbb{R}$, defined by

$$f(z,\tau) = \frac{1}{(2\pi)^{(d+1)/2}} \int_0^\infty \int_0^\infty \Omega_d(z\xi_1)\cos(\tau\xi_2)\phi(\xi_1,\xi_2)\xi_1^{d-1}\mathrm{d}\xi_1\mathrm{d}\xi_2 \qquad (2.1)$$

is nonnegative and integrable. The function f is called *isotropic spectral density* throughout.

2.2 The Matérn and the Generalized Wendland classes of covariance functions

The Matérn class (Matérn, 1986; Handcock and Stein, 1993) of continuous functions $K_{\mathcal{M}}(r;\alpha,\nu)$, $r \geq 0$, $\alpha,\nu > 0$, is defined as

$$K_{\mathcal{M}}(r;\alpha,\nu) = \frac{2^{1-\nu}}{\Gamma(\nu)}\left(\frac{r}{\alpha}\right)^\nu \mathcal{K}_\nu\left(\frac{r}{\alpha}\right), \qquad (2.2)$$

where $\mathcal{K}_\nu$ is a modified Bessel function of the second kind of order ν (Abramowitz and Stegun, 1970). $K_{\mathcal{M}}(\cdot;\alpha,\nu)$ belongs to the class Φ_∞.

We now introduce the $\mathcal{GW}$ class $K_{\mathcal{GW}}(\cdot;\beta,\mu,\kappa):[0,\infty)\to\mathbb{R}$, defined as (Gneiting, 2002; Zastavnyi, 2002)

$$K_{\mathcal{GW}}(r;\beta,\mu,\kappa) := \begin{cases} \frac{1}{B(2\kappa,\mu+1)}\int_{r/\beta}^{1} u(u^2-(r/\beta)^2)^{\kappa-1}(1-u)^{\mu}\,\mathrm{d}u, & 0\le r<\beta,\\ 0, & r\ge\beta, \end{cases} \tag{2.3}$$

where $\kappa>0$, $\beta>0$ is the compact support parameter and B denotes the Beta function. For $\kappa=0$ the $\mathcal{GW}$ class is defined as (Askey, 1973):

$$K_{\mathcal{GW}}(r;\beta,\mu,0) := \begin{cases} (1-r/\beta)^{\mu}, & 0\le r<\beta,\\ 0, & r\ge\beta. \end{cases} \tag{2.4}$$

Closed-form solutions of the integral in (2.3) can be obtained when $\kappa=k$, a nonnegative integer. In this case, (2.3) can be factorized as

$$K_{\mathcal{GW}}(r;\beta,\mu,k) = K_{\mathcal{GW}}(r;\beta,\mu+k,0)P_k(r), \qquad r\ge 0,$$

where P_k is a polynomial of order k.

The $\mathcal{GW}$ class belongs to the class Φ_d, for a fixed $d\in\mathbb{N}$, provided $\mu\ge(d+1)/2+\kappa$.

Both $K_{\mathcal{M}}$ and $K_{\mathcal{GW}}$ are flexible models, as they allow for parameterizing in a continuous fashion the mean square and sample path differentiability of a Gaussian random field with these covariance functions. In particular, for the $\mathcal{M}$ case, given a positive integer k, the sample paths are k times differentiable, in any direction, if and only if $\nu>k$. Similarly , for the $\mathcal{GW}$

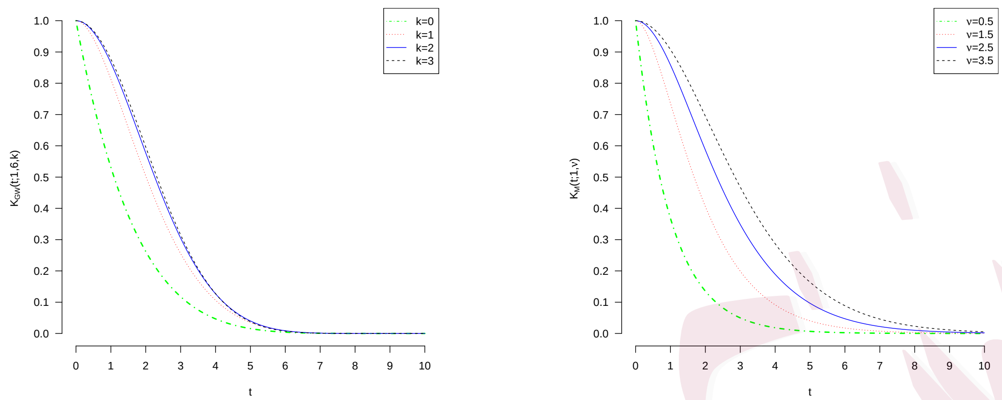

Figure 1: Left: $K_{\mathcal{GW}}(t;10,6,k)$ for $k = 0,1,2,3$, Right: $K_{\mathcal{M}}(t;1,\nu)$ for $\nu = 0.5, 1.5, 2.5, 3.5$

case, the sample paths are k times differentiable, in any direction, if and only if $\kappa > k - 0.5$. Figure 1 depicts $K_{\mathcal{GW}}(t;10,6,k)$ for $k = 0,1,2,3$ and $K_{\mathcal{M}}(t;1,\nu)$ for $\nu = 0.5, 1.5, 2.5, 3.5$.

2.3 The $\mathcal{DM}$ and $\mathcal{DGW}$ Families of Space-Time Covariance Functions

The $\mathcal{DM}$ family of space-time covariance functions has been introduced by Fuentes et al. (2008): the motivation for such a proposal is to provide a class of space-time covariance functions having spatial or temporal margins of the Matérn type. That is, either the spatial margin $C(\cdot,0)$ or the temporal margin $C(\mathbf{0},\cdot)$ are proportional to the class $K_{\mathcal{M}}(\|\cdot\|;\alpha,\nu)$ as being defined at (2.2). The $\mathcal{DM}$ family is the building block for the *tour de force* in Ip and Li (2017), that has largely inspired our work. To introduce the $\mathcal{DM}$ family, we follow a different path than Fuentes et al. (2008): let $\boldsymbol{\theta} = (\nu, \zeta, \upsilon, \epsilon)^\top$ with $\top$ denoting the transpose of a vector. We assume ν, υ, ζ positive, and $\epsilon \in [0,1]$. We define the parameter ℓ, which depends on $\boldsymbol{\theta}$, through

$$\ell(\boldsymbol{\theta}) = \frac{\zeta^{2\nu-d}\upsilon^{2\nu-1}\Gamma(\nu)}{\Gamma(\nu - \frac{d+1}{2})}\mathbf{1}_{\{\epsilon=0\}} + \frac{\zeta^{2\nu-d}\upsilon^{2\nu-1}\Gamma(\nu)^2}{\Gamma(\nu-\frac{d}{2})\Gamma(\nu-\frac{1}{2})}\mathbf{1}_{\{\epsilon=1\}} + x\mathbf{1}_{\{\epsilon\in(0,1)\}}, \qquad (2.5)$$

with $\mathbf{1}_A$ being the indicator function of any Borel set of the real line. Here, Γ denotes the Gamma function (Gradshteyn and Ryzhik, 2007), and x is a positive constant that is kept fixed and is disregarded for the rest of our exposition.

We define the $\mathcal{DM}$ class, $K_{\mathcal{DM}}(\cdot,\cdot;\boldsymbol{\theta}) : [0,\infty^2) \to \mathbb{R}$, through the identity

$$K_{\mathcal{DM}}(r,t;\boldsymbol{\theta}) = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}ut} g_{\boldsymbol{\theta}}(r,u)\mathrm{d}u, \qquad (r,t) \in [0,\infty), \qquad (2.6)$$

where i is the imaginary unit. Here, the function $g_{\boldsymbol{\theta}}$ is defined as

$$g_{\boldsymbol{\theta}}(r,u) = \frac{\ell(\boldsymbol{\theta})\pi^{d/2}}{2^{\nu-d/2-1}\Gamma(\nu)}\left(\frac{r}{a(u)}\right)^{\nu-d/2}\left(\upsilon^2 + \epsilon u^2\right)^{-\nu}\mathcal{K}_{\nu-d/2}\left(a(u)r\right),$$

with $a(u) = \sqrt{\zeta^2(\upsilon^2+u^2)/\upsilon^2+\epsilon u^2}$, $u \in \mathbb{R}$. In Equation (2.6), the parameter ζ^{-1} (spatial range) explains the rate of decay of the spatial correlation, υ^{-1} (temporal range) explains the rate of decay for the temporal correlation, and ℓ is a scale parameter of the associated random field. The parameter ϵ allows to switch from separability (when $\epsilon = 0$) to different levels of nonseparability. The arguments in Fuentes et al. (2008) show that, $K_{\mathcal{DM}}(\cdot,\cdot;\boldsymbol{\theta})$ is a member of the class $\Phi_{\infty,T}$. Also, Fuentes et al. (2008) show that some special cases admit partial Fourier transforms that admit closed forms of the Matérn type.

We now follow Porcu et al. (2020) to introduce the $\mathcal{DGW}$ class of space-time covariance functions. Let $\mu, \beta > 0$, $\delta \in (0, 2]$, $\gamma > 0$ and $\kappa \geq 0$. Let $\boldsymbol{\chi} = (\mu, \kappa, \beta, \delta, \lambda, \gamma)^\intercal$.

The range of the parameter λ is specified below. Let us consider the function

$$h_{\delta,\gamma}(t) = \left(1 + \left(\frac{t}{\gamma}\right)^{\delta}\right)^{-1}, \quad t \geq 0. \tag{2.7}$$

We define the $\mathcal{DGW}$ class, denoted $K_{\mathcal{DGW}}(\cdot, \cdot; \boldsymbol{\chi})$ (Porcu et al., 2020) as follows:

$$K_{\mathcal{DGW}}(r, t; \boldsymbol{\chi}) = [h_{\delta,\gamma}(t)]^{\lambda} K_{\mathcal{GW}}(r; \beta h_{\delta,\gamma}(t), \mu, \kappa), \qquad r, t \geq 0. \tag{2.8}$$

According to Theorem 1 in Porcu et al. (2020) (see also Table 1 therein) that $K_{\mathcal{DGW}}(\cdot, \cdot; \boldsymbol{\chi})$ belongs to class $\Phi_{d,T}$ for some integer d, provided

$$\mu \geq (d+3)/2 + \kappa + \alpha, \qquad \text{and} \qquad \lambda > \max((d+3)/2, 2\kappa + 3). \tag{2.9}$$

The constant α is positive and bigger than a lower bound $\kappa_1(\delta)$ that is specified in Table 1 of Porcu et al. (2020). Here, α is fixed and does not enter into the parameter $\boldsymbol{\chi}$. For the remainder of the paper, we suppose that α is always bigger than the lower bound $\kappa_1(\delta)$. As for parameters interpretation, we note that β is the spatial compact support when $t = 0$, that is, $K_{\mathcal{DGW}}(\cdot, 0; \boldsymbol{\chi}) = K_{\mathcal{GW}}(\cdot; \beta, \mu, \kappa)$, with $K_{\mathcal{GW}}$ as in (2.3), is compactly supported over a ball with radius β embedded in $\mathbb{R}^d$. The parameter κ determines the differentiability at the origin for the spatial margin $K_{\mathcal{DGW}}(\cdot, 0; \boldsymbol{\chi})$. The parameter $\gamma > 0$ is the temporal scale, and the parameter δ indexes fractal dimension for the temporal sample paths. Finally, the function $h_{\delta,\gamma}$ is the temporal radius, as for every $t_o > 0$, the margin $K_{\mathcal{DGW}}(\cdot, t_o; \boldsymbol{\chi})$ is compactly supported over a ball with radius $\beta h_{\delta,\gamma}(t_o)$ embedded in $\mathbb{R}^d$.

For the remainder of the paper, we use $f_{\mathcal{DM}}(\cdot, \cdot; \boldsymbol{\theta})$ and $f_{\mathcal{DGW}}(\cdot, \cdot; \boldsymbol{\chi})$ for the Fourier transforms of, respectively, $K_{\mathcal{DM}}(\cdot, \cdot; \boldsymbol{\theta})$ and $K_{\mathcal{DGW}}(\cdot, \cdot; \boldsymbol{\chi})$, that are uniquely determined according to Equation (2.1).

3. Preliminary Results

3.1 Fourier Transforms and Tails for the $\mathcal{DGW}$ and $\mathcal{DM}$ classes

For d a positive integer and $\kappa \geq 0$, we define $\eta := (d+1)/2 + \kappa$. Next result describes the behaviour of the isotropic spectral density associated with the $K_{\mathcal{DGW}}$, $f_{\mathcal{DGW}}(\cdot, \cdot; \boldsymbol{\chi})$, defined at (2.8), and determined according to (2.1). Some further notation is needed. For given functions $g_1(x)$ and $g_2(x)$, we write $g_1(x) \asymp g_2(x)$ to mean that there exist constants c and C such that $0 < c < C < \infty$ and $c|g_2(x)| \leq |g_1(x)| \leq C|g_2(x)|$ for all x.

Note that $f \sim g$ means that the function f is asymptotically equal to the function g. We consider the function ${}_1F_2$ defined as:

$$ {}_1F_2(a; b, c; z) = \sum_{k=0}^{\infty} \frac{(a)_k z^k}{(b)_k (c)_k k!}, \qquad z \in \mathbb{R}, $$

which is a special case of the generalized hypergeometric functions ${}_qF_p$ (Abramowitz and Stegun, 1970), with $(q)_k = \Gamma(q+k)/\Gamma(q)$ for $k \in \mathbb{N} \cup \{0\}$, being the Pochhammer symbol. Finally, for a complex number z, we use $\Im(z)$ to denote its immaginary part. We are ready to provide the first result of the paper.

Theorem 1. *Let $\mathcal{DGW}$ be the class of functions $K_{\mathcal{DGW}}(\cdot,\cdot;\boldsymbol{\chi})$ defined at Equation (2.8), and let $f_{\mathcal{DGW}}(\cdot,\cdot;\boldsymbol{\chi})$ be the spectral density associated with $K_{\mathcal{DGW}}(\cdot,\cdot;\boldsymbol{\chi})$ and determined according to Equation (2.1). Let $\boldsymbol{\varsigma} := (\mu, \kappa, \eta, d)^\top$. Let*

$$ \varrho_{\lambda,\eta} = \frac{2^{\delta}(d+\lambda-2\eta)\Gamma(\frac{\delta+1}{2})\Gamma(\frac{\delta+2}{2})\sin(\frac{\pi\delta}{2})}{\gamma^{\delta}\pi^{\frac{3}{2}}}, $$

$$ \varrho_{\lambda,\eta+1} = \frac{2^{\delta}(d+\lambda-2\eta-2)\Gamma(\frac{\delta+1}{2})\Gamma(\frac{\delta+2}{2})\sin(\frac{\pi\delta}{2})}{\gamma^{\delta}\pi^{\frac{3}{2}}}, $$

$$ c_3^{\boldsymbol{\varsigma}} = \frac{\Gamma(\mu+2\eta)}{\Gamma(\mu)}, \qquad c_4^{\boldsymbol{\varsigma}} = \frac{\Gamma(\mu+2\eta)}{\Gamma(\eta)2^{\eta-1}}, \qquad c_5^{\boldsymbol{\varsigma}} = \frac{\pi}{2}(\mu+\eta) $$

and

$$ L^{\boldsymbol{\varsigma}} = \frac{2^{-\kappa-d+1}\Gamma(\kappa)\pi^{-\frac{d}{2}}\Gamma(\mu+1)\Gamma(2\kappa+d)}{B(2\kappa,\eta)\Gamma(\kappa+\frac{d}{2})\Gamma(\mu+2\eta)}. $$

Then, for $\kappa \geq 0$, $\beta > 0$, $\delta \in (0,2)$, $\lambda > 2\kappa+3$ and $\mu \geq \eta+1+\alpha$, we have

1. $f_{\mathcal{DGW}}(z,\tau;\boldsymbol{\chi}) = -\beta^d \gamma^{3/2}\tau^{1/2}\sqrt{2}\pi^{-3/2} L^{\varsigma} \times$
$$\times \Im\left(\int_0^\infty \frac{\mathcal{K}_{1/2}(\gamma t \tau)\, {}_1F_2\left(\eta; \eta+\frac{\mu}{2}, \eta+\frac{\mu}{2}+\frac{1}{2}; -\frac{\left(z\beta(1+e^{i\pi\delta/2}t^\delta)^{-1}\right)^2}{4}\right)}{(1+e^{i\pi\delta/2}t^\delta)^{d+\lambda}} t^{1/2}\mathrm{d}t \right);$$

2. For $\tau, z \to \infty$,
$$f_{\mathcal{DGW}}(z,\tau;\boldsymbol{\chi}) = \beta^{-(1+2\kappa)} L^{\varsigma} c_3^{\varsigma} z^{-2\eta} \times$$
$$\left(\left[\varrho_{\lambda,\eta}\tau^{-(1+\delta)} - \mathcal{O}\left(\tau^{-(1+2\delta)}\right)\right] + \left[\varrho_{\lambda,\eta+1}\tau^{-(1+\delta)} - \mathcal{O}\left(\tau^{-(1+2\delta)}\right)\right]\mathcal{O}(z^{-2}) \right) +$$
$$+ \left[\varrho_{\lambda,0}\tau^{-(1+\delta)} - \mathcal{O}\left(\tau^{-(1+2\delta)}\right)\right]\mathcal{O}(z^{-(\mu+\eta)}); \tag{3.1}$$

3. For $z \to \infty$, $\tau \to \infty$, $f_{\mathcal{DGW}}(z,\tau;\boldsymbol{\chi}) \asymp z^{-2\eta}\tau^{-1-\delta}$.

The proof of this result is deferred to Section C in OS.

To describe the asymptotic behaviour of the spectral density associated with the $\mathcal{DM}$ class, a result from Ip and Li (2017) is needed.

Theorem 2. *Let $f_{\mathcal{DM}}(\cdot,\cdot;\boldsymbol{\theta})$ be the spectral density function associated with the $\mathcal{DM}$ class in Equation (2.6), and being uniquely determined according to (2.1). Then, for $\upsilon > 0$ and $\epsilon \in [0,1]$, we have*

1. $f_{\mathcal{DM}}(z,\tau;\boldsymbol{\theta}) = \ell(\boldsymbol{\theta})(\zeta^2\upsilon^2 + \upsilon^2 z^2 + \zeta^2\tau^2 + \epsilon^2 z^2\tau^2)^{-\nu}$;

2. As $z, \tau \to \infty$ and $\epsilon \in (0,1]$,
$$\frac{1}{f_{\mathcal{DM}}(z,\tau;\boldsymbol{\theta})} \sim \ell^{-1}(\boldsymbol{\theta})(\epsilon z\tau)^{2\nu}\left(1 + \frac{\nu\zeta^2\upsilon^2}{\epsilon^2 z^2\tau^2} + \frac{\nu\upsilon^2}{\epsilon^2\tau^2} + \frac{\nu\zeta^2}{\epsilon^2 z^2} + \mathcal{O}(\tau^{-4}z^{-4})\right);$$

3. As $z, \tau \to \infty$ and $\epsilon = 0$,

$$\frac{1}{f_{\mathcal{DM}}(z,\tau;\boldsymbol{\theta})} \sim \ell^{-1}(\boldsymbol{\theta})(\upsilon^2 z^2+\zeta^2\tau^2)^{\nu}\left(1+\nu\frac{\zeta^2\upsilon^2}{\zeta^2\tau^2+\upsilon^2 z^2}+\mathcal{O}\left((\zeta^2\tau^2+\upsilon^2 z^2)^{-2}\right)\right).$$

The following section relates technical results that provide the crux for the proofs of the main results in the manuscript.

4. Maximum likelihood estimation for $\mathcal{DGW}$ Classes

The following results refer to Gaussian random fields with covariance function $C(r,t) = \sigma^2 K(r,t;\boldsymbol{\tau})$, where $\boldsymbol{\tau}$ is a given parameter vector, and where $K \in \Phi_{d,T}$ for a given d. Following the arguments in Zhang (2004), an immediate consequence of Theorem 2 in OS is that for fixed κ, δ, μ and λ the parameters σ^2, β and γ cannot be estimated consistently. Instead, we show here that the microergodic parameter $\sigma^2/(\gamma^{\delta}\beta^{2\kappa+1})$ is consistently estimable. We shall then assess the asymptotic distribution of the ML estimator. Let $D \times \mathcal{T}$ be a bounded subset of $\mathbb{R}^d \times \mathbb{R}$ and let $\boldsymbol{Z}_{nm} = (Z(\boldsymbol{s}_1,t_1),\ldots,Z(\boldsymbol{s}_n,t_m))^{\top}$ be a finite realization of a zero mean stationary Gaussian random field $Z(\boldsymbol{s},t)$, $(\boldsymbol{s},t) \in D\times\mathcal{T}$, with a given parametric covariance function $\sigma^2 K(r,t;\boldsymbol{\tau})$. Here we consider the $\mathcal{DGW}$ covariance model, that is $\sigma^2 K_{\mathcal{DGW}}(r,t;\boldsymbol{\chi})$, where $K_{\mathcal{DGW}}$ and the corresponding parameters have been defined at (2.8). At the same time, in the current exposition $\boldsymbol{\chi}$ will not contain the parameters that are fixed, but only those that are to be estimated through ML. Specifically κ, δ, λ and μ are assumed known and fixed, that is we assume $\boldsymbol{\chi} = (\beta, \gamma)^\top$, being the spatial and temporal scale parameters. Then, the Gaussian log-likelihood function is defined as:

$$\mathcal{L}_{nm}(\sigma^2, \beta, \gamma) = -\frac{1}{2}\left(nm\log(2\pi\sigma^2) + \log\left(|R_{nm}(\beta,\gamma)|\right) + \frac{1}{\sigma^2}\boldsymbol{Z}_{nm}^\top R_{nm}(\beta,\gamma)^{-1}\boldsymbol{Z}_{nm}\right), \tag{4.1}$$

where $R_{nm}(\beta, \gamma) = [K(\|\boldsymbol{s}_i - \boldsymbol{s}_j\|, |t_l - t_k|; \beta, \gamma)]_{i,j=1;,l,k=1}^{n;m}$ is the correlation matrix. Let $\hat{\sigma}^2_{nm}$ be the ML estimator of the variance parameter obtained by maximizing $\mathcal{L}_{nm}(\sigma^2, \beta, \gamma)$ with respect to σ^2, and given by

$$\hat{\sigma}^2_{nm}(\beta, \gamma) = \frac{1}{nm}\boldsymbol{Z}_{nm}^\top R_{nm}(\beta,\gamma)^{-1}\boldsymbol{Z}_{nm}. \tag{4.2}$$

We now establish strong consistency and asymptotic distribution of the random variable $\hat{\sigma}^2_{nm}(\beta,\gamma)/(\gamma^\delta\beta^{2\kappa+1})$ *i.e.*, the ML estimator of the microergodic parameter.

Theorem 3. *Let $Z(\boldsymbol{s}, t)$, $(\boldsymbol{s}, t) \in D \times \mathcal{T} \subset \mathbb{R}^d \times \mathbb{R}$, $d = 1, 2$, be a zero mean Gaussian random field with covariance model $\sigma^2 K_{\mathcal{DGW}}(\cdot, \cdot; \boldsymbol{\chi})$ and let $\boldsymbol{\chi} = (\beta_0, \gamma_0)^\top$, with $\lambda > 2\kappa + 3$ and $\mu > \eta + 1 + \alpha$. For κ, δ, λ and μ fixed and known and arbitrary β and γ , we have, as $n, m \to \infty$,*

1. $\frac{\hat{\sigma}^2_{nm}(\beta,\gamma)}{\gamma^\delta\beta^{2\kappa+1}} \xrightarrow{a.s.} \frac{\sigma_0^2}{\gamma_0^\delta\beta_0^{2\kappa+1}}$, *and*

2. $\sqrt{n \times m}\left(\frac{\hat{\sigma}^2_{nm}(\beta,\gamma)}{\gamma^\delta\beta^{2\kappa+1}} - \frac{\sigma_0^2}{\gamma_0^\delta\beta_0^{2\kappa+1}}\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, 2\left(\frac{\sigma_0^2}{\gamma_0^\delta\beta_0^{2\kappa+1}}\right)^2\right)$.

The proof is deferred to Section C in OS. The second point of Theorem 3 provides the asymptotic distribution of the microergodic parameter for arbitrary dependence parameters β and γ. Nevertheless, in practical applications both parameters must be estimated. In principle the asymptotic distribution of the random variable $\frac{\hat{\sigma}^2_{nm}(\hat{\beta},\hat{\gamma})}{\hat{\gamma}^{\delta}\hat{\beta}^{2\kappa+1}}$, with $\hat{\boldsymbol{\chi}} = (\hat{\beta},\hat{\gamma})^{\intercal}$, can be obtained following the arguments in Kaufman and Shaby (2013) or Bevilacqua et al. (2019). However, to establish strong consistency and asymptotic distribution of the sequence of random variables $\frac{\hat{\sigma}^2_{nm}(\hat{\beta},\hat{\gamma})}{\hat{\gamma}^{\delta}\hat{\beta}^{2\kappa+1}}$, we need to prove the monotone behaviour of $\frac{\hat{\sigma}^2_{nm}(\beta,\gamma)}{\gamma^{\delta}\beta^{2\kappa+1}}$ when viewed as a function of $(\beta,\gamma) \in I \times J$, with $I \times J$ is a product of bounded intervals. Unfortunately, we have not been able to inspect such a monotonicity property.

In the following, to assess the quality of the approximation of Theorem 3 (Point 2) we consider a simulation study that takes into account the case when γ and β are arbitrary, and we also explore the case when both are estimated with ML.

Specifically, we consider 500 simulations, using Cholesky decomposition, of a Gaussian random field with $\mathcal{DGW}$ space-time covariance function observed in $[0,1]^2 \times [0,1]$. In particular, we consider x^2 location sites uniformly distributed in $[0,1]^2$ with $x = 6, 8, 10, 12, 14$ and $0, 0.1, \ldots, 0.9, 1$ temporal instants, that is, we consider $n = 36, 64, 100, 144, 196$, and $m = 11$. The increasing total number of space-time observations in the three dimensional unit cube is $n \times m = 396, 704, 1100, 1584, 2156$ respectively.

For each simulation, we consider $\kappa = 0, 1$, $\delta = 1.75$, $\lambda = 5$ and $\mu = 5.5 + \kappa$ as known and fixed, and we set $\sigma_0^2 = 1$, $\beta_0 = 1$, $\gamma_0 = 3$. We estimate the microergodic parameter with

$$\frac{\hat{\sigma}_i^2(x_i, y_i)}{x_i^{2\kappa+1} y_i^{\delta}} = \frac{\boldsymbol{Z}_i^{\top} R_{nm}(x_i, y_i)^{-1} \boldsymbol{Z}_i}{nm x_i^{2\kappa+1} y_i^{\delta}},$$

where $x_i = \beta_0$, $y_i = \gamma_0$ for the case with arbitrary dependence parameters (here we set them equal to the true dependence parameters) and $x_i = \hat{\beta}_i$, $y_i = \hat{\gamma}_i$ for the case of estimated parameters with ML. Here $\boldsymbol{Z}_i$ is the data vector of simulation i.

For the first case ML estimation is namely obtained using (4.2) and for the second case ML estimation is obtained trough the maximization, with respect of β and γ, of the log profile likelihood $\mathcal{L}_{nm}(\hat{\sigma}_{nm}^2(\beta, \gamma), \beta, \gamma)$.

Using the asymptotic distributions stated in Theorem 3, Table 1 compares the sample quantiles of order $0.05, 0.25, 0.5, 0.75, 0.95$, mean and variance of

$$\sqrt{\frac{n \times m}{2}} \Big(\frac{\hat{\sigma}_i^2(x_i, y_i) \beta_0^{2\kappa+1} \gamma_0^{\delta}}{\sigma_0^2 x_i^{2\kappa+1} y_i^{\delta}} - 1 \Big),$$

when $x_i = \beta_0$ and $y_i = \gamma_0$ with the associated theoretical values of the standard Gaussian distribution. In the same table we also explore the case $x_i = \widehat{\beta}_i$ and $y_i = \widehat{\gamma}_i$.

As expected, the best approximation is achieved overall when using the true dependence parameters, *i.e.*, $x_i = \beta_0$, $y_i = \gamma_0$ and in the case of $x_i = \widehat{\beta}_i$, $y_i = \widehat{\gamma}_i$ the asymptotic distribution seems a satisfactory approximation of the sample distribution, visually improving when increasing n. Note that the variance increases when the smoothness parameter κ increases. This pattern in well known in the purely spatial case when estimating the microergodic parameter of the $\mathcal{GW}$ or Matérn covariance models. In addition, when $x_i = \beta_0$ and $y_i = \gamma_0$ the sample quantiles does not depends on κ, as expected. We repeat this numerical experiment by considering arbitrary dependence parameters sufficiently "far" from the true values parameters and we observe that in this case the convergence can be very slow as observed also in Kaufman and Shaby (2013) and Bevilacqua et al. (2019).

5. Prediction using the $\mathcal{DGW}$ model

We now consider kriging prediction, under fixed domain asymptotics, of a Gaussian random field at an unknown space-time location $(\boldsymbol{s}_0, t_0) \in \mathcal{D} \times \mathcal{T}$, using the $\mathcal{DGW}$ model $\sigma^2 K_{\mathcal{DGW}}(r, t; \boldsymbol{\chi})$. Specifically, we focus on two properties:

(A) asymptotic efficient prediction and

(B) asymptotically correct estimation of the prediction variance.

Table 1: For $\kappa = 0, 1$ and $\delta = 1.75$, sample quantiles, mean and variance of $\sqrt{\frac{n\times m}{2}}\left(\frac{\hat{\sigma}_i^2(x_i,y_i)\beta_0^{2\kappa+1}\gamma_0^\delta}{\sigma_0^2 x_i^{2\kappa+1}y_i^\delta} - 1\right)$, $i = 1,\dots,500$, for $x_i = \widehat{\beta}_i, \beta_0$ and $y_i = \widehat{\gamma}_i, \gamma_0$ when $\beta_0 = 1$ and $\gamma_0 = 3$ with $n\times m = 396, 704, 1100, 1584, 2156$, compared with the associated theoretical values of the standard Gaussian distribution.

$\kappa = 0$	(x, y)	$n \times m$	5%	25%	50%	75%	95%	Mean	Var
	$(\widehat{\beta},\widehat{\gamma})$	396	-1.962	-0.904	0.060	0.836	2.142	0.029	1.534
		704	-1.889	-0.759	0.010	0.836	2.031	0.030	1.386
		1100	-1.728	-0.741	0.068	0.852	1.868	0.070	1.278
		1584	-1.642	-0.738	0.008	0.704	1.717	0.017	1.141
		2156	-1.643	-0.720	-0.009	0.639	1.669	-0.094	1.119
	(β_0, γ_0)	396	-1.535	-0.705	-0.014	0.720	1.788	0.022	1.061
		704	-1.662	-0.733	0.032	0.704	1.758	0.001	1.060
		1100	-1.675	-0.700	0.032	0.709	1.682	0.021	1.052
		1584	-1.634	-0.646	0.014	0.717	1.601	0.005	1.017
		2156	-1.645	-0.648	-0.094	0.659	1.660	-0.079	1.012
$\kappa = 1$	(x, y)	$n \times m$	5%	25%	50%	75%	95%	Mean	Var
	$(\widehat{\beta},\widehat{\gamma})$	396	-2.179	-0.971	-0.110	0.733	2.462	-0.036	1.839
		704	-2.039	-0.806	0.041	0.877	1.938	0.015	1.510
		1100	-1.939	-0.782	0.104	0.800	1.850	0.002	1.382
		1584	-1.683	-0.735	-0.030	0.653	1.977	-0.002	1.270
		2156	-1.693	-0.720	-0.009	0.679	1.723	-0.096	1.194
	(β_0, γ_0)	396	-1.535	-0.705	-0.014	0.720	1.788	0.022	1.061
		704	-1.662	-0.733	0.032	0.704	1.758	0.001	1.060
		1100	-1.675	-0.700	0.032	0.709	1.682	0.021	1.052
		1584	-1.634	-0.646	0.014	0.717	1.601	0.005	1.017
		2156	-1.645	-0.648	-0.094	0.659	1.660	-0.079	1.012
$N(0, 1)$			-1.645	-0.674	0	0.674	1.645	0	1

Stein (1988) shows that both asymptotic properties hold when the Gaussian measures are equivalent (see Section A in OS). Let $P(\sigma_i^2 K_{\mathcal{DGW}}(\boldsymbol{\chi}_i))$, $i = 0, 1$ be two probability zero mean Gaussian measures with $\mathcal{DGW}$ classes as in (2.8), with $\boldsymbol{\chi}_i = (\beta_i, \gamma_i)^\top$ (See A in OS for details).

Under $P(\sigma_0^2 K_{\mathcal{DGW}}(\boldsymbol{\chi}_0))$, and using Theorem 2 in OS, properties **(A)** and **(B)** hold provided

$$\frac{\sigma_0^2}{\gamma_0^\delta \beta_0^{2\kappa+1}} = \frac{\sigma_1^2}{\gamma_1^\delta \beta_1^{2\kappa+1}},$$

and $\mu > \eta + 1 + \alpha$, $\delta > (d+1)/2$ and $d = 1, 2$. Similarly, let $P(\sigma_0^2 K_{\mathcal{DM}}(\boldsymbol{\theta}))$ and $P(\sigma_1^2 K_{\mathcal{DGW}}(\boldsymbol{\chi}))$ be two zero mean Gaussian measures under $\mathcal{DM}$ and $\mathcal{DGW}$ models, respectively. Under $P(\sigma_0^2 K_{\mathcal{DM}}(\boldsymbol{\theta}))$, properties **(A)** and **(B)** hold when $\mu > \eta + 1 + \alpha$, Point *2* of Theorem 3 in OS holds and $d = 1, 2$.

Actually, Stein (1993) gives a substantially weaker condition for asymptotic efficiency prediction based on the asymptotic behaviour of the ratio of the spectral densities. Let

$$\widehat{Z}_{nm}(\boldsymbol{\chi}_1) = \boldsymbol{c}_{nm}(\boldsymbol{\chi}_1)^\top R_{nm}(\boldsymbol{\chi}_1)^{-1} \boldsymbol{Z}_{nm} \tag{5.1}$$

be the best linear unbiased predictor at an unknown location $(\boldsymbol{s}_0, t_0) \in \mathcal{D} \times \mathcal{T}$, under the misspecified model $P(\sigma_1^2 K_{\mathcal{DGW}}(\boldsymbol{\chi}_1))$, where $\boldsymbol{c}_{nm}(\boldsymbol{\chi}_1) = [K_{\mathcal{DWG}}(\boldsymbol{s}_0 - \boldsymbol{s}_i, t_0 - t_j; \boldsymbol{\chi}_1)]_{i=1,j=1}^{n,m}$ and $R_{nm}(\boldsymbol{\chi}_1) = [K_{\mathcal{DWG}}(\boldsymbol{s}_i - \boldsymbol{s}_j, t_i - t_j; \boldsymbol{\chi}_1)]_{i=1,j=1}^{n,m}$ is the correlation matrix. If the correct model is $P(\sigma_0^2 K_{\mathcal{DGW}}(r, t; \boldsymbol{\chi}_0))$, then the mean squared error of the kriging predictor is given by:

$$\begin{aligned}&\mathrm{Var}_{\boldsymbol{\chi}_0}\left[\widehat{Z}_{nm}(\boldsymbol{\chi}_1)-Z(\boldsymbol{s}_0,t_0)\right] \qquad (5.2)\\ =\;& \sigma_0^2\Big(1-2\boldsymbol{c}_{nm}(\boldsymbol{\chi}_1)^\top R_{nm}(\boldsymbol{\chi}_1)^{-1}\boldsymbol{c}_{nm}(\boldsymbol{\chi}_0)+\boldsymbol{c}_{nm}(\boldsymbol{\chi}_1)^\top R_{nm}(\boldsymbol{\chi}_1)^{-1}R_{nm}(\boldsymbol{\chi}_0)R_{nm}(\boldsymbol{\chi}_1)^{-1}\boldsymbol{c}_{nm}(\boldsymbol{\chi}_0)\Big).\end{aligned}$$

If $\beta_0=\beta$ and $\gamma_0=\gamma$, *i.e.*, true and misspecified models coincide, this expression simplifies to

$$\mathrm{Var}_{\boldsymbol{\chi}_0}\left[\widehat{Z}_{nm}(\boldsymbol{\chi}_0)-Z(\boldsymbol{s}_0,t_0)\right]=\sigma_0^2\big(1-\boldsymbol{c}_{nm}(\boldsymbol{\chi}_0)^\top R_{nm}(\boldsymbol{\chi}_0)^{-1}\boldsymbol{c}_{nm}(\boldsymbol{\chi}_0)\big). \qquad (5.3)$$

Similarly, $\mathrm{Var}_{\boldsymbol{\theta}_0}\left[\widehat{Z}_{nm}(\boldsymbol{\chi}_1)-Z(\boldsymbol{s}_0,t_0)\right]$ and $\mathrm{Var}_{\boldsymbol{\theta}_0}\left[\widehat{Z}_{nm}(\boldsymbol{\theta})-Z(\boldsymbol{s}_0,t_0)\right]$ can be defined under $P(\sigma_0^2K_{\mathcal{DM}}(\boldsymbol{\theta}_0)$. Here $\widehat{Z}_{nm}(\boldsymbol{\theta})$ is the best linear unbiased predictor using the $\mathcal{DM}$ model and recall that $\boldsymbol{\theta}=(\nu,\zeta,\upsilon,\epsilon)^\top$ is the set of correlation parameters. The following results are an application of Theorems 1 and 2 of Stein (1993).

Theorem 4. *Let $P(\sigma_i^2K_{\mathcal{DGW}}(\boldsymbol{\chi}_i))$, $i=0,1$ and $P(\sigma_0^2K_{\mathcal{DM}}(\boldsymbol{\theta}_0))$ be three Gaussian probability measures on $\mathcal{D}\times\mathcal{T}\subset\mathbb{R}^d\times\mathbb{R}$ and let $\mu_i>\eta_i+1+\alpha_i$. Then, for all $(\boldsymbol{s}_0,t_0)\in\mathcal{D}\times\mathcal{T}$:*

1. *Under $P(\sigma_0^2K_{\mathcal{DGW}}(\boldsymbol{\chi}_0))$, as $n\to\infty$,*

$$\frac{\mathrm{Var}_{\boldsymbol{\chi}_0}\big[\widehat{Z}_{nm}(\boldsymbol{\chi}_1)-Z(\boldsymbol{s}_0,t_0)\big]}{\mathrm{Var}_{\boldsymbol{\chi}_0}\big[\widehat{Z}_{nm}(\boldsymbol{\chi}_0)-Z(\boldsymbol{s}_0,t_0)\big]}\longrightarrow 1, \qquad (5.4)$$

for any fixed $\beta_1>0$ and $\gamma_1>0$.

2. *Under* $P(\sigma_0^2 K_{\mathcal{DM}}(\boldsymbol{\theta}_0))$*, if* $\nu = \eta$ *as* $n \to \infty$,

$$\frac{\mathrm{Var}_{\boldsymbol{\theta}_0}\left[\widehat{Z}_{nm}(\boldsymbol{\chi}_1) - Z(\boldsymbol{s}_0, t_0)\right]}{\mathrm{Var}_{\boldsymbol{\theta}_0}\left[\widehat{Z}_{nm}(\boldsymbol{\theta}) - Z(\boldsymbol{s}_0, t_0)\right]} \longrightarrow 1, \tag{5.5}$$

   for any fixed $\beta_1 > 0$*,* $\gamma_1 > 0$ *and* $\boldsymbol{\theta} = (\nu, \zeta, \upsilon, \epsilon)$.

3. *Under* $P(\sigma_0^2 K_{\mathcal{DGW}}(\boldsymbol{\chi}_0))$*, if* $\frac{\sigma_0^2 \beta_0^{-(2\kappa+1)}}{\gamma_0^\delta} = \frac{\sigma_1^2 \beta_1^{-(2\kappa+1)}}{\gamma_1^\delta}$*, then*

$$\frac{\mathrm{Var}_{\boldsymbol{\chi}_1}\left[\widehat{Z}_{nm}(\boldsymbol{\chi}_1) - Z(\boldsymbol{s}_0, t_0)\right]}{\mathrm{Var}_{\boldsymbol{\chi}_0}\left[\widehat{Z}_{nm}(\boldsymbol{\chi}_1) - Z(\boldsymbol{s}_0, t_0)\right]} \longrightarrow 1. \tag{5.6}$$

4. *Under* $P(\sigma_0^2 K_{\mathcal{DM}}(\boldsymbol{\theta}_0))$*, for* $\epsilon \in (0,1]$*, if* $\sigma_1^2 \varrho_{\lambda,\eta} c_3^\varsigma \beta^{-2\eta} = \ell(\boldsymbol{\theta}_0)\epsilon^{-2\nu}$*,* $\nu = \eta$ *and* $1 + 2\kappa = \delta$*, then as* $n \to \infty$,

$$\frac{\mathrm{Var}_{\boldsymbol{\chi}_1}\left[\widehat{Z}_{nm}(\boldsymbol{\chi}_1) - Z(\boldsymbol{s}_0, t_0)\right]}{\mathrm{Var}_{\boldsymbol{\theta}_0}\left[\widehat{Z}_{nm}(\boldsymbol{\chi}_1) - Z(\boldsymbol{s}_0, t_0)\right]} \longrightarrow 1. \tag{5.7}$$

As an illustration of the results in Theorem 4, we perform a small numerical experiment and in particular we focus on Points 1 and 3. Let us define the ratios (5.4) and (5.6) as $U_1(\beta_1, \gamma_1)$ and U_2, respectively. We randomly select $n_j = 36, 64, 100, 144, 196$, $(j = 1, \ldots 500)$ location sites without replacement from a fine regular grid on the unit square and wee keep this location sites fixed across the eleven temporal instants $0, 0.1, \ldots, 1$. We then compute the ratios U_{1j} and U_{2j}, $j = 1, \ldots, 500$, using the closed form expressions in Equation (5.2) and (5.3), when predicting the space-time location site $\boldsymbol{s}_0 = (0.53, 0.53)$ and $t_0 = 0.6$. Specifically, for $\kappa = 0, 1$ we set $\mu = 5.5 + \kappa$,

$\delta = 1.75$, $\lambda = 5$, as in the numerical experiment of Section 4. The parameters of the correct model $\boldsymbol{\chi}_0 = (\beta_0, \gamma_0)^\intercal$ are fixed as ,$\beta_0 = 1$, $\gamma_0 = 3$, with additionally $\sigma_0^2 = 1$, and the parameters of the misspecified model $\boldsymbol{\chi}_1 = (\beta_1, \gamma_1)^\intercal$ are fixed as $\sigma_1^2 = 1.25$, $\gamma_1 = 3.05$ and the spatial parameter is obtained using the equivalence condition, that is $\beta_1 = \beta_0((\sigma_0^2/\sigma_1^2)(\gamma_0/\gamma_1)^\delta)^{-(1+2\kappa)}$ (see also A in OS). This gives $\beta_1 = 1.21436$ for $\kappa = 0$ and $\beta_1 = 1.066881$ for $\kappa = 1$.

Table 2 reports the overall mean $\bar{U}_1 = \sum_{j=1}^{500} U_{1j}/500$ and $\bar{U}_2 = \sum_{j=1}^{500} U_{2j}/500$ when increasing the number of spatiotemporal sites $n \times m = 396, 704, 1100, 1584, 2196$. It can be appreciated that the convergence to 1 of $\bar{U}_1$ is much faster than $\bar{U}_2$. This results are consistent to the purely spatial case in Bevilacqua et al. (2019). Additionally there is no significative differences between the cases $\kappa = 0, 1$.

Table 2: $\bar{U}_1 = \sum_{j=1}^{500} U_{1j}/500$, $\bar{U}_2 = \sum_{j=1}^{500} U_{2j}/500$ when increasing the number of spacetime locations for $\kappa = 0, 1$.

$m \times n$	κ=0		κ=1	
	$\bar{U}_1$	$\bar{U}_2$	$\bar{U}_1$	$\bar{U}_2$
396	1.00249	1.05611	1.00104	1.05730
704	1.00104	1.04349	1.00035	1.04338
1100	1.00048	1.03826	1.00013	1.03781
1584	1.00022	1.03513	1.00005	1.03500
2156	1.00012	1.03354	1.00002	1.03337

Conclusions

There is a clear lack of general results on the asymptotic properties of ML estimator under fixed-domain asymptotics, in particular in the space-time setting. This is reflected in the literature, where the results are sparse and are estabilished for *ad hoc* families of covariance functions. Similarly, our paper has established some results holding for the $\mathcal{DGW}$ family under fixed-domain asymptotics.

Future works might be devoted on how to evade from the current asymptotics setting and consider a more realistic setting for the temporal component. A promising solution might be to embed time into the circle, so that the associated Gaussian random field becomes periodic.

Acknowledgement

The authos are very grateful to Donald Richards for the thorough revision and the very helpful discussions. Partial support was provided by

FONDECYT grant 1130647, Chile for Emilio Porcu. Partial support was provided by FONDECYT grant 1200068, Chile for Moreno Bevilacqua and by Millennium Science Initiative of the Ministry of Economy, Development, and Tourism, grant "Millenium Nucleus Center for the Discovery of Structures in Complex Data" for Moreno Bevilacqua and Emilio Porcu. Moreno Bevilacqua also acknowledges support from the MATH-AMSUD program, grant 20-MATH-03. The research work conducted by Tarik Faouzi was supported in part by grant DIUBB 2020525 IF/R Chile.

References

Abramowitz, M. and I. A. Stegun (1970). *Handbook of Mathematical Functions*. New York: Dover.

Askey, R. (1973). Radial characteristic functions. *Technical report, Research Center, University of Wisconsin*.

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

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

Daley, D. J. and E. Porcu (2013). Dimension walks and Schoenberg spectral measures. *Proc. Amer. Math. Society 141*, 1813–1824.

Fuentes, M., L. Chen, and J. M. Davis (2008). A class of nonseparable and nonstationary spatial temporal covariance functions. *Environmetrics: The official journal of the International Environmetrics Society 19*(5), 487–507.

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

Gneiting, T. and P. Guttorp (2010). Continuous parameter spatio-temporal processes. *Handbook of Spatial Statistics 97*, 427–436.

Gradshteyn, I. S. and I. M. Ryzhik (2007). *Tables of Integrals, Series, and Products* (seventh ed.). Academic Press, Amsterdam.

Handcock, M. S. and M. L. Stein (1993). A bayesian analysis of kriging. *Technometrics 35*(4), 403–410.

Ip, R. H. and W. Li (2017). On some Matérn covariance functions for spatio-temporal random fields. *Statistica Sinica 27*, 805–822.

Kaufman, C. G. and B. A. Shaby (2013). The role of the range parameter for estimation and prediction in geostatistics. *Biometrika 100*, 473–484.

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

Matérn, B. (1986). Spatial variation, volume 36 of lecture notes in statistics.

Porcu, E., M. Bevilacqua, and M. G. Genton (2020). Nonseparable, space-time covariance functions with dynamical compact supports. *Statistica Sinica 30*, to appear.

Porcu, E., R. Furrer, and D. Nychka (2020). 30 years of space-time covariance functions. *Wiley Interdiscip. Rev. Comput. Stat.*, –.

Porcu, E., P. Gregori, and J. Mateu (2006). Nonseparable stationary anisotropic space time covariance functions. *Stochastic Environmental Research and Risk Assessment 21*, 113–122.

Skorokhod, A. and M. Yadrenko (1973). On absolute continuity of measures corresponding to homogeneous Gaussian fields. *Theory of Probability and Its Applications 18*, 27–40.

Stein, M. (1988). Asymptotically efficient prediction of a random field with a misspecified covariance function. *The Annals of Statistics 16*, 55–63.

Stein, M. L. (1993). A simple condition for asymptotic optimality of linear predictions of random fields. *Statistic and Probability Letters 17*, 399–404.

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

Wang, D. and W.-L. Loh (2011). On fixed-domain asymptotics and covariance tapering in gaussian random field models. *Electronic Journal of Statistics 5*, 238–269.

Zastavnyi, V. (2002). Positive-definite radial functions and splines. *Doklady Mathematics 66*, 213–216.

Zastavnyi, V. and R. Trigub (2002). Positive definite splines of special form. *English transl. in Sb. Math. 193*, 1771–1800.

Zhang, H. (2004). Inconsistent estimation and asymptotically equivalent interpolations in model-based geostatistics. *Journal of the American Statistical Association 99*, 250–261.