

NORGES TEKNISK-NATURVITENSKAPELIGE
UNIVERSITET

**Multivariate Gaussian Random Fields Using Systems of
Stochastic Partial Differential Equations**

by

Xiangping Hu, Daniel Simpson, Finn Lindgren and Håvard Rue

PREPRINT
STATISTICS NO. 3/2013



NORWEGIAN UNIVERSITY OF SCIENCE AND
TECHNOLOGY
TRONDHEIM, NORWAY

This preprint has URL <http://www.math.ntnu.no/preprint/statistics/2013/S3-2013.pdf>

Xiangping Hu has homepage: <http://www.math.ntnu.no/~xiangph>

E-mail: xiangph@math.ntnu.no

Address: Department of Mathematical Sciences, Norwegian University of Science and
Technology, N-7491 Trondheim, Norway.

Multivariate Gaussian Random Fields Using Systems of Stochastic Partial Differential Equations

Xiangping Hu^{*1}, Daniel Simpson¹, Finn Lindgren², and Håvard Rue¹

¹Department of Mathematical Sciences, Norwegian University of Science and Technology, N-7491 Trondheim, Norway

²Department of Mathematical Sciences, University of Bath, BA2 7AY, United Kingdom

July 4, 2013

Abstract

In this paper a new approach for constructing *multivariate* Gaussian random fields (GRFs) using systems of stochastic partial differential equations (SPDEs) has been introduced and applied to simulated data and real data. By solving a system of SPDEs, we can construct multivariate GRFs. On the theoretical side, the notorious requirement of non-negative definiteness for the covariance matrix of the GRF is satisfied since the constructed covariance matrices with this approach are automatically symmetric positive definite. Using the approximate stochastic weak solutions to the systems of SPDEs, multivariate GRFs are represented by multivariate Gaussian *Markov* random fields (GMRFs) with sparse precision matrices. Therefore, on the computational side, the sparse structures make it possible to use numerical algorithms for sparse matrices to do fast sampling from the random fields and statistical inference. Therefore, the *big-n* problem can also be partially resolved for these models. These models out-perform existing multivariate GRF models on a commonly used real dataset.

Keywords: Multivariate Gaussian random fields; Gaussian Markov random fields; covariance matrix; stochastic partial differential equations; sparse matrices

1 Introduction

Gaussian random fields (GRFs) have a dominant role in spatial modelling and there exists a wealth of applications in areas such as geostatistics, atmospheric and environmental science and many other fields (Cressie, 1993; Stein, 1999; Diggle and Ribeiro Jr, 2006). GRFs are practical since the normalizing constant can be computed explicitly. They also possess good properties since they can be explicitly specified through the mean

^{*}Corresponding author. Email: Xiangping.Hu@math.ntnu.no

function $\boldsymbol{\mu}(\cdot)$ and covariance function $\mathbf{C}(\cdot, \cdot)$. In \mathbb{R}^d , with $\mathbf{s} \in \mathbb{R}^d$, $x(\mathbf{s})$ is a continuously indexed GRF if all the finite subsets $(x(\mathbf{s}_i))_{i=1}^n$ jointly have Gaussian distributions. A GRF is said to be stationary or homogeneous if the mean function $\boldsymbol{\mu}(\cdot)$ is constant and the covariance function $\mathbf{C}(\cdot, \cdot)$ is a function only of the distance between the coordinates (Adler and Taylor, 2007). It is common to use covariance functions that are isotropic. The isotropic covariance functions are functions of just the Euclidean distance between two locations.

1.1 Matérn covariance functions and Multivariate GRFs

The Matérn family of covariance functions is a class of commonly used isotropic covariance functions introduced by Matérn (Matérn, 1986). This family of covariance functions is usually parametrized by $\sigma^2 M(\mathbf{m}, \mathbf{n} | \nu, a)$, where, $M(\mathbf{m}, \mathbf{n} | \nu, a)$ is the Matérn correlation function between locations $\mathbf{m}, \mathbf{n} \in \mathbb{R}^d$ and has the form

$$M(\mathbf{h} | \nu, a) = \frac{2^{1-\nu}}{\Gamma(\nu)} (a \|\mathbf{h}\|)^\nu K_\nu(a \|\mathbf{h}\|), \quad (1)$$

where σ^2 is the marginal variance. $\|\mathbf{h}\|$ denotes the Euclidean distance between \mathbf{m} and \mathbf{n} . K_ν is the modified Bessel function of second kind and $a > 0$ is a scaling parameter. The order ν is the smoothness parameter. It defines the critical concerns in spatial statistical modelling and simulation (Stein, 1999), such as the differentiability of the sample paths and the Hausdorff dimension (Handcock and Stein, 1993; Goff and Jordan, 1988). We follow the common practice to fix the smoothness parameter ν because it is poorly identifiable from data (Ribeiro Jr and Diggle, 2006; Lindgren et al., 2011). The range ρ connects the scaling parameter a and the smoothness parameter ν . Throughout the paper, the simple relationship $\rho = \sqrt{8\nu}/a$ is assumed from the empirically derived definition (Lindgren et al., 2011). We refer to Matérn (1986), Diggle et al. (1998), Guttorp and Gneiting (2006) and Lindgren et al. (2011) for detailed information about Matérn covariance functions. One can show that the Matérn covariance function can be reduced to the product of a polynomial and an exponential function when the smoothness parameter ν is equal to an integer m plus $\frac{1}{2}$ (Gneiting et al., 2010). The Matérn covariance function nests the popular exponential model since $M(\mathbf{h} | \frac{1}{2}, a) = \exp(-a \|\mathbf{h}\|)$.

Analogously to the univariate case, we need to specify a mean vector function $\boldsymbol{\mu}(\mathbf{s})$ and a covariance function $\mathbf{C}(\|\mathbf{h}\|)$ that assigns to each distance $\|\mathbf{h}\|$ a $p \times p$ symmetric non-negative definite covariance matrix in order to specify an isotropic p -dimensional multivariate Gaussian random field $\mathbf{x}(\mathbf{s})$. It is known that it is quite difficult to construct flexible covariance functions $\mathbf{C}(\cdot)$ that satisfy this requirement.

Gneiting et al. (2010) presented an approach for constructing multivariate random fields using matrix-valued covariance functions, where each constituent component in the matrix-valued covariance function is a Matérn covariance function. Define $\mathbf{x}(\mathbf{s}) = (x_1(\mathbf{s}), x_2(\mathbf{s}), \dots, x_p(\mathbf{s}))^T$, with $\mathbf{s} \in \mathbb{R}^d$, so that each location consists p components. T denotes the transpose of a vector or a matrix. Assume the process is second-order

stationary with mean vector zero and matrix-valued covariance function of the form

$$\mathbf{C}(\mathbf{h}) = \begin{pmatrix} C_{11}(\mathbf{h}) & C_{12}(\mathbf{h}) & \cdots & C_{1p}(\mathbf{h}) \\ C_{21}(\mathbf{h}) & C_{22}(\mathbf{h}) & \cdots & C_{2p}(\mathbf{h}) \\ \vdots & \vdots & \ddots & \vdots \\ C_{p1}(\mathbf{h}) & C_{p2}(\mathbf{h}) & \cdots & C_{pp}(\mathbf{h}) \end{pmatrix}, \quad (2)$$

where,

$$C_{ii}(\mathbf{h}) = \sigma_{ii}M(\mathbf{h}|\nu_{ii}, a_{ii}), \quad (3)$$

is the covariance function $C_{ii}(\mathbf{h}) = \mathbb{E}(x_i(\mathbf{s} + \mathbf{h})x_i(\mathbf{s}))$ within the field $x_i(\mathbf{s})$ and

$$C_{ij}(\mathbf{h}) = \rho_{ij}\sigma_i\sigma_jM(\mathbf{h}|\nu_{ij}, a_{ij}) \quad (4)$$

is the cross-covariance function $C_{ij}(\mathbf{h}) = \mathbb{E}(x_i(\mathbf{s} + \mathbf{h})x_j(\mathbf{s}))$ between the fields $x_i(\mathbf{s})$ and $x_j(\mathbf{s})$. ρ_{ij} is the co-located correlation coefficient, $\sigma_{ii} \geq 0$ is the marginal variance, and σ_i and σ_j are the corresponding standard deviations, $1 \leq i \neq j \leq p$. We use the following notations through out the paper

$$\begin{aligned} \sigma_{ii} &= \sigma_i\sigma_i, \\ \sigma_{ij} &= \sigma_i\sigma_j. \end{aligned} \quad (5)$$

Gneiting et al. (2010) presented some conditions to ensure the matrix-valued covariance function in Equation (2) is symmetric and non-negative definite, with focus on the bivariate case. They claimed that the parameters in the parametric family of matrix-valued covariance function given in Equation (2) for multivariate random fields are interpretable in terms of smoothness, correlation length, variances of the processes and co-located coefficients. It was shown that a parsimonious bivariate Matérn model, with fewer parameters than the full bivariate Matérn model, is preferred to the traditional linear model of coregionalization (LMC) (Gelfand et al., 2004). Even though the GRFs have convenient analytical properties, the covariance-based models for constructing GRFs are hindered by computational issues, or the so-called “*big-n* problem” (Banerjee et al., 2004). This is because inference with these models requires the factorization of dense $n \times n$ covariance matrices, which requires $\mathcal{O}(n^2)$ storage and $\mathcal{O}(n^3)$ floating point operations. It follows that this kind of model is not suitable for problems with realistically large amounts of data.

Statistical inference for large datasets within feasible time is still a challenge in modern spatial statistics. The size of modern datasets typically overwhelm the traditional models in spatial statistics and a great deal of effort has been expended trying to construct good models that scale well computationally. In this paper, we extend the models proposed by Lindgren et al. (2011), which exploited an explicit link between GRFs and GMRFs through stochastic partial differential equations (SPDEs). They showed that for some univariate Gaussian fields with Matérn covariance functions, it is possible to do the modelling with GRFs, but do the computations using GMRFs.

A GMRF is a discretely indexed Gaussian random field $\mathbf{x}(\mathbf{s})$ with the property that its full conditional distributions $\{\pi(x_i|\mathbf{x}_{-i}); i = 1, \dots, n\}$, only depend on the set of

neighbors ∂i to each node i . It is obvious that if $i \in \partial j$ then $j \in \partial i$ from consistency. Rather than parameterizing GMRFs through their covariance matrix, they are usually parametrized through their precision matrix (the inverse of the covariance matrix). The Markov property implies that the precision matrix \mathbf{Q} is sparse due to the fact that $Q_{ij} \neq 0 \iff i \in \partial j \cup j$ (Rue and Held, 2005). This allows for the use of numerical methods for sparse matrices for fast sampling and also for fast statistical inference (Rue, 2001; Rue and Held, 2005; Lindgren et al., 2011). The general cost of factorizing the precision matrix \mathbf{Q} for a spatial Markovian structure is $\mathcal{O}(n)$ in one dimension, $\mathcal{O}(n^{3/2})$ in two dimensions and $\mathcal{O}(n^2)$ in three dimensions (Rue and Held, 2005).

Lindgren et al. (2011) pointed out that using the link between the GRFs and GMRFs can open new doors to modelling difficult problems with simple models. The SPDE approach can be extended to Gaussian random fields on manifolds, non-stationary random fields, random fields with oscillating covariance functions and non-separable space-time models.

1.2 Matérn covariance models through SPDEs

In this paper we use the following characterization of Matérn random fields, originally due to Whittle (1954, 1963), that formed the basis for the methods of Lindgren et al. (2011). A GRF $x(\mathbf{s})$ with the Matérn covariance function can be described as a solution to the linear fractional SPDE (Whittle, 1954, 1963; Lindgren et al., 2011)

$$(\kappa^2 - \Delta)^{\alpha/2} x(\mathbf{s}) = \mathcal{W}(\mathbf{s}), \quad (6)$$

where $(\kappa^2 - \Delta)^{\alpha/2}$ is a pseudo-differential operator and $\alpha = \nu + d/2$, $\kappa > 0$, $\nu > 0$. The innovation process $\mathcal{W}(\mathbf{s})$ is a spatial standard Gaussian white noise. $\Delta = \sum_{i=1}^d \frac{\partial^2}{\partial x_i^2}$ is the Laplacian on \mathbb{R}^d . Applying the Fourier transform to the (fractional) SPDE given in \mathbb{R}^d in (6) yields

$$\left\{ \mathcal{F}(\kappa^2 - \Delta)^{\alpha/2} \phi \right\}(\mathbf{k}) = (\kappa^2 + \|\mathbf{k}\|^2)^{\alpha/2} (\mathcal{F}\phi)(\mathbf{k}), \quad (7)$$

where \mathcal{F} denotes the Fourier transform, ϕ is a smooth, rapidly decaying function in \mathbb{R}^d . See for example, Lindgren et al. (2011) for detailed description. Equation (7) is used in Section 3 for model comparison. One might think that the Matérn covariance function seems rather restrictive in statistical modelling, but it covers the most commonly used models in spatial statistics (Lindgren et al., 2011). Stein (1999, Page 14) has a practical suggestion “use the Matérn model”. For more information about the Matérn family, see, for example Diggle et al. (1998, Section 3.4.1) and Stein (1999, Section 2.6).

1.3 Outline of the paper

The rest of the paper is organized as follows. In Section 2 we discuss the construction of multivariate GRFs using systems of SPDEs, we call this approach the SPDEs approach. Additionally, a brief introduction of how to construct the multivariate GRFs using the

covariance-based models is also included in this section. Section 3 contains a detailed model comparisons between the SPDEs approach and the covariance-based models presented by Gneiting et al. (2010). Section 3.4 discusses how to sample from these models, and statistical inference for simulated data and real dataset is presented in Section 4. The paper ends with a general discussion in Section 5.

2 Model construction

In this section we discuss how to use the SPDEs approach to construct multivariate GRFs. One of the appealing properties of this approach is that the SPDE specification *automatically* constructs valid covariance functions. This means that if the solution to a system of SPDEs exists, then it will construct a matrix-valued covariance function which fulfills the symmetric non-negative definiteness property. The parameters in the parametric model from the SPDEs approach are interpretable in terms of co-located correlation coefficients, smoothness, marginal variances and correlations. And there is a correspondence to the parameters in the covariance matrix based models.

2.1 Multivariate GRFs and the SPDEs approach

A zero mean d -dimensional multivariate (or d -dimensional p -variate) GRF is a collection of continuously indexed multivariate normal random vectors

$$\mathbf{x}(\mathbf{s}) \sim MVN(\mathbf{0}, \boldsymbol{\Sigma}(\mathbf{s})),$$

where $\boldsymbol{\Sigma}(\mathbf{s})$ is a non-negative definite matrix that depends on the point $\mathbf{s} \in \mathbb{R}^d$. Define the system of SPDEs

$$\begin{pmatrix} \mathcal{L}_{11} & \mathcal{L}_{12} & \dots & \mathcal{L}_{1p} \\ \mathcal{L}_{21} & \mathcal{L}_{22} & \dots & \mathcal{L}_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{L}_{p1} & \mathcal{L}_{p2} & \dots & \mathcal{L}_{pp} \end{pmatrix} \begin{pmatrix} x_1(\mathbf{s}) \\ x_2(\mathbf{s}) \\ \vdots \\ x_p(\mathbf{s}) \end{pmatrix} = \begin{pmatrix} f_1(\mathbf{s}) \\ f_2(\mathbf{s}) \\ \vdots \\ f_p(\mathbf{s}) \end{pmatrix}, \quad (8)$$

where $\{\mathcal{L}_{ij} = b_{ij}(\kappa_{ij}^2 - \Delta)^{\alpha_{ij}/2}; \alpha_{ij} = 0 \text{ or } 2, 1 \leq i, j \leq p\}$, are differential operators and $\{f_i(\mathbf{s}); 1 \leq i, j \leq p\}$ are independent but not necessarily identically distributed noise processes. It turns out that the solution to (8) defines a multivariate GRF $\mathbf{x} = (x_1(\mathbf{s}), x_2(\mathbf{s}), \dots, x_p(\mathbf{s}))^T$. Define the operator matrix

$$\mathcal{L} = \begin{pmatrix} \mathcal{L}_{11} & \mathcal{L}_{12} & \dots & \mathcal{L}_{1p} \\ \mathcal{L}_{21} & \mathcal{L}_{22} & \dots & \mathcal{L}_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{L}_{p1} & \mathcal{L}_{p2} & \dots & \mathcal{L}_{pp} \end{pmatrix} \quad (9)$$

and

$$\mathbf{f}(\mathbf{s}) = (f_1(\mathbf{s}), f_2(\mathbf{s}), \dots, f_p(\mathbf{s}))^T,$$

and then the system of SPDEs (8) can be written in compact matrix form as

$$\mathcal{L}(\boldsymbol{\theta})\mathbf{x}(\mathbf{s}) = \mathbf{f}(\mathbf{s}), \quad (10)$$

where $\boldsymbol{\theta} = \{\boldsymbol{\alpha}, \boldsymbol{\kappa}, \mathbf{b}\}$ denotes the collection of parameters. Similarly as Equation (7) we apply Fourier transforms to Equation (10), and it yields

$$\mathcal{H}(\boldsymbol{\theta})\hat{\mathbf{x}}(\mathbf{k}) = \hat{\mathbf{f}}(\mathbf{k}), \quad (11)$$

where \mathbf{k} is the frequency, $\hat{\mathbf{x}}(\mathbf{k}) = \mathcal{F}(\mathbf{x}(\mathbf{s}))$ and $\hat{\mathbf{f}}(\mathbf{k}) = \mathcal{F}(\mathbf{f}(\mathbf{s}))$ are the Fourier transforms of the random fields and the noise processes, respectively, and $\mathcal{H}(\boldsymbol{\theta})$ is the matrix formed from the Fourier transforms of the operator matrix $\mathcal{L}(\boldsymbol{\theta})$. Each element has the form $\mathcal{H}_{ij}(\theta_{ij}) = b_{ij}(\kappa_{ij}^2 + \|\mathbf{k}\|^2)^{\alpha_{ij}/2}$ with $\theta_{ij} = \{\alpha_{ij}, \kappa_{ij}, b_{ij}\}$, $1 \leq i, j \leq p$. From Equation (11) one can find the spectral process as $\hat{\mathbf{x}}(\mathbf{k}) = \mathcal{H}^{-1}(\boldsymbol{\theta})\hat{\mathbf{f}}(\mathbf{k})$. The corresponding power spectrum is defined as $\mathbf{S}_x(\mathbf{k}) = \mathbb{E}(\hat{\mathbf{x}}\hat{\mathbf{x}}^H)$, where H denotes the Hermitian transpose of a matrix. Simple calculations using the above mentioned formulas yield

$$\mathbf{S}_x(\mathbf{k}) = \mathcal{H}^{-1}(\boldsymbol{\theta})\mathbf{S}_f(\mathbf{k})\mathcal{H}^{-H}(\boldsymbol{\theta}), \quad (12)$$

where \mathcal{H}^{-H} denotes the inverse of the Hermitian transpose of Fourier transform of the operator matrix \mathcal{L} . The Equation (12) can be written explicitly as

$$\mathbf{S}_x(\mathbf{k}) = \begin{pmatrix} S_{x_{11}}(\mathbf{k}) & S_{x_{12}}(\mathbf{k}) & \cdots & S_{x_{1p}}(\mathbf{k}) \\ S_{x_{21}}(\mathbf{k}) & S_{x_{22}}(\mathbf{k}) & \cdots & S_{x_{2p}}(\mathbf{k}) \\ \vdots & \vdots & \ddots & \vdots \\ S_{x_{p1}}(\mathbf{k}) & S_{x_{p2}}(\mathbf{k}) & \cdots & S_{x_{pp}}(\mathbf{k}) \end{pmatrix}. \quad (13)$$

Let $\mathbf{S}_f(\mathbf{k})$ denote the power spectrum matrix of the noise processes, $\mathbf{S}_f(\mathbf{k}) = \mathbb{E}(\hat{\mathbf{f}}\hat{\mathbf{f}}^H)$.

If the noise processes are mutually independent, \mathbf{S}_f is a diagonal matrix and we write it as

$$\mathbf{S}_f(\mathbf{k}) = \text{diag}(S_{f_{11}}, S_{f_{22}}, \dots, S_{f_{pp}}),$$

where $\{S_{f_i} = \mathbb{E}(\hat{f}_i\hat{f}_i^H); i = 1, \dots, p\}$. This means it is easy to obtain this matrix.

In general, with Equations (8) to (12) we can compute all the components in (13). We show how to do this for bivariate random fields in Section 3.1.

2.2 Covariance-based model for multivariate GRFs

It is also possible to construct a multivariate GRF using the covariance-based model, but the notorious non-negative definiteness restriction makes it hard. Gneiting et al. (2010) discussed one such approach in detail. The main aim of their approach is to find the proper constrains for $\nu_{ij}, a_{ij}, \rho_{ij}$ which result in valid matrix-valued covariance functions for second-order stationary multivariate GRFs with a covariance function of the form

in Equation (2), that is, a symmetric non-negative definite covariance function. Some useful theorems for constructing the covariance functions for bivariate case are presented in Gneiting et al. (2010, Section 2.2).

In this covariance-based model, the components $C_{ij}(\mathbf{h})$ in the matrix-valued covariance function $\mathbf{C}(\mathbf{h})$ given in Equation (2) are modelled directly. Gneiting et al. (2010) try to find conditions on the parameter space which result in valid multivariate Matérn models. In the bivariate case when $p = 2$, a full characterization of the parameter space is achieved. For $p \geq 3$ Gneiting et al. (2010) suggested that a parsimonious model should be used in practice. This kind of model has more restrictions on the smoothness parameter and the scale parameters, such that $a_{ij} = a$, where $a > 0$ is the common scale parameter and $\nu_{ij} = \frac{1}{2}(\nu_i + \nu_j)$ for $1 \leq i \neq j \leq p$. We refer to Gneiting et al. (2010) for more information about different kinds of conditions that yield the valid matrix-valued covariance functions.

Assume that the components of the covariance matrix are known. The power spectrum can also be obtained from in this covariance-based approach. This can in turn be used to compare parameters with the SPDEs approach. The covariance matrix of a second-order stationary multivariate Matérn GRF was given in Equation (2). The power spectrum in \mathbb{R}^d of the cross-covariance function for $x_i(\mathbf{s})$ and $x_j(\mathbf{s})$ is defined as $R_{x_{ij}}(\mathbf{k}) = \mathcal{F}(\text{Cov}(x_i(\mathbf{s}), x_j(\mathbf{s})))$, $1 \leq i, j \leq p$, which is given by

$$R_{x_{ij}}(\mathbf{k}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-i\mathbf{k}\mathbf{h}} C_{ij}(\|\mathbf{h}\|) d\mathbf{h}. \quad (14)$$

Applying the Fourier transform to the covariance matrix in Equation (2) yields

$$\mathbf{R}(\mathbf{k}) = \begin{pmatrix} R_{x_{11}}(\mathbf{k}) & R_{x_{12}}(\mathbf{k}) & \cdots & R_{x_{1p}}(\mathbf{k}) \\ R_{x_{21}}(\mathbf{k}) & R_{x_{22}}(\mathbf{k}) & \cdots & R_{x_{2p}}(\mathbf{k}) \\ \vdots & \vdots & \ddots & \vdots \\ R_{x_{p1}}(\mathbf{k}) & R_{x_{p2}}(\mathbf{k}) & \cdots & R_{x_{pp}}(\mathbf{k}) \end{pmatrix}. \quad (15)$$

The comparison between the covariance-based approach and the SPDEs approach in Section 3 is based on the following fact: if $S_{x_{ij}}(\mathbf{k}) = R_{x_{ij}}(\mathbf{k})$, for each $1 \leq i, j \leq p$, then the multivariate GRFs constructed through the SPDEs approach (8) and through the covariance-based model (2) are equivalent.

2.3 GMRF approximations to GRFs

It is known that spatial Markovian GMRFs have good computational properties since the precision matrix \mathbf{Q} is sparse. The Markov structure here means that $x_i(\mathbf{s})$ and $x_j(\mathbf{s})$ are independent conditioned on $x_{-ij}(\mathbf{s})$, when i and j are not neighbors, where $x_{-ij}(\mathbf{s})$ means for $x_{-\{i,j\}}(\mathbf{s})$. In other words, $Q_{ij} = 0$ if and only if $x_i(\mathbf{s})$ and $x_j(\mathbf{s})$ are independent conditioned $x_{-ij}(\mathbf{s})$. So it is possible to read off whether x_i and x_j are conditionally independent or not from the elements of \mathbf{Q} . Note that the precision matrix \mathbf{Q} should also be non-negative definite. The sparse structure of the precision matrix \mathbf{Q}

is crucial with our models in Bayesian inference methods on large datasets. We refer to Rue and Held (2005, Chapter 2) for more information on the theory for GMRFs.

Rue and Tjelmeland (2002) demonstrated that most of the GRFs in geostatistics can be approximated with GMRFs. Hartman and Hössjer (2008) proposed to use GMRFs instead of GRFs, from a computational point of view, when doing spatial prediction using Kriging. This approach can also be used for spatio-temporal models (Allcroft and Glasbey, 2003). In this paper we follow the approach presented in Lindgren et al. (2011) for GMRF approximations to GRFs in order to partially resolve the “*big n problem*” with our models. Lindgren et al. (2011) mainly considered univariate GRFs with Matérn covariance functions. In this case the GMRF representation was constructed explicitly through the SPDE, and the solution to the SPDE driven by the Gaussian white noise is a GRFs with Matérn covariance function. They showed how to build a GRF model with a covariance matrix Σ theoretically from SPDE, and then use the GMRF to represent the GRF, which means the precision matrix of the GMRF fulfills the condition $\mathbf{Q}^{-1} \simeq \Sigma$ on some predefined norm. Fuglstad (2011) considered a modification of the SPDE with a diffusion matrices to control the covariance structure of the GRF, and create inhomogeneous GRFs, but the GMRF representation is still the main ingredient for computations.

With our SPDEs approach we use two steps to construct the multivariate GRFs. The first step is to construct the precision matrices for the noise processes in Equation (8), and the second step is to solve the system SPDEs (8) with the constructed noise processes. We focus on the noise processes, if they are not white noise processes, generated by the SPDE given as follows

$$(\kappa_{n_i}^2 - \Delta)^{\alpha_{n_i}/2} f_i(\mathbf{s}) = \mathcal{W}_i(\mathbf{s}), \quad \alpha_{n_i} = \nu_{n_i} + d/2, \quad i = 1, 2, \dots, p, \quad (16)$$

with the requirements that $\kappa_{n_i} > 0$ and $\nu_{n_i} > 0$. κ_{n_i} and ν_{n_i} are the scaling parameter and smoothness parameter for noise process $f_i(\mathbf{s})$. $\{\mathcal{W}_i(\mathbf{s}); i = 1, 2, \dots, p\}$ are standard Gaussian white noise processes. The generated noise processes in this way are independent but not necessarily identically distributed.

The first step in constructing the GMRF representation for the noise process $f_i(\mathbf{s})$ on the triangulated lattice is to find the stochastic weak formulation of Equation (16) (Kloeden and Platen, 1999). In this paper Delaunay triangulation is chosen, and we refer to Hjelle and Dæhlen (2006) for detailed discussion about Delaunay triangulations. Denote the inner product of functions h and g as

$$\langle h, g \rangle = \int h(\mathbf{s})g(\mathbf{s})d(\mathbf{s}), \quad (17)$$

where the integral is over the region of interest. We can find the stochastic weak solution of SPDE (16) by requiring that

$$\left\{ \langle \phi_k, (\kappa_{n_i}^2 - \Delta)^{\alpha_{n_i}/2} f_i \rangle \right\}_{i=1}^M \stackrel{d}{=} \{ \langle \phi_k, \mathcal{W}_i \rangle \}_{i=1}^M. \quad (18)$$

In the second step we need to construct a finite element representation of the solution to the SPDE. We refer to Zienkiewicz et al. (2005) and Bathe (2008) for more information

about finite element methods. The finite element representation of the solution to SPDE (16) is

$$f_i(\mathbf{s}) = \sum_{k=1}^N \psi_k(\mathbf{s}) \omega_k, \quad (19)$$

where $\psi_k(\mathbf{s})$ is some chosen basis function, ω_k is some Gaussian distributed weight, and N is the number of the vertexes in the triangulation. The basis function $\psi_k(\mathbf{s})$ is chosen to be piecewise linear with value 1 at vertex k and 0 at all other vertexes. This means that a continuously indexed solution is approximated with a piecewise linear function defined through the joint distribution of $\{\omega_k; k = 1, 2, \dots, N\}$ (Lindgren et al., 2011). The chosen basis functions ensure that the local interpolation on a triangle is piecewise linear.

The third step is to choose the test functions. In this paper we follow the setting used by Lindgren et al. (2011). With $M = N$ the test functions are chosen as $\phi_k = (\kappa_{n_i}^2 - \Delta)^{1/2} \psi_k$ for $\alpha_{n_i} = 1$ and $\phi_k = \psi_k$ for $\alpha_{n_i} = 2$, which are denoted as a *least squares* and a *Galerkin* solution, respectively. When $\alpha_{n_i} \geq 3$, the approximation can be obtained by setting $\alpha_{n_i} = 2$ at the left-hand side of Equation (16) and replacing the right hand side of Equation (16) with a field generated with $\alpha_{n_i} - 2$ and let $\phi_k = \psi_k$. This iteration procedure terminates when $\alpha_{n_i} = 1$ or 2. This is the essence of the recursive Galerkin formulation. More detailed description can be found in Lindgren et al. (2011). α_{n_i} can only be integer-valued currently. When α_{n_i} is not an integer, different approximation methods must be used and this is beyond the scope of our discussion. Response to the discussion to Lindgren et al. (2011) discussed fractional α . We define the required $N \times N$ matrices \mathbf{C} , and \mathbf{G} with entries

$$C_{mn} = \langle \psi_m, \psi_n \rangle, \quad G_{mn} = \langle \nabla \psi_m, \nabla \psi_n \rangle, \quad m, n = 1, 2, \dots, N, \quad (20)$$

and

$$(\mathbf{K}_{\kappa_{n_i}^2}) = (\kappa_{n_i}^2 \mathbf{C} + \mathbf{G}). \quad (21)$$

Finally, by using these matrices given in (20) and (21) together with the Neumann boundary conditions (zero normal derivatives at the boundaries), the precision matrices \mathbf{Q}_{f_i} for noise process $f_i(\mathbf{s})$ can be obtained,

$$\begin{cases} \mathbf{Q}_{1, \kappa_{n_i}^2} &= \mathbf{K}_{\kappa_{n_i}^2}, & \text{for } \alpha_{n_i} = 1, \\ \mathbf{Q}_{2, \kappa_{n_i}^2} &= \mathbf{K}_{\kappa_{n_i}^2}^T \mathbf{C}^{-1} \mathbf{K}_{\kappa_{n_i}^2}, & \text{for } \alpha_{n_i} = 2, \\ \mathbf{Q}_{\alpha_{n_i}, \kappa_{n_i}^2} &= \mathbf{K}_{\kappa_{n_i}^2}^T \mathbf{C}^{-1} \mathbf{Q}_{\alpha_{n_i}-2, \kappa_{n_i}^2} \mathbf{C}^{-1} \mathbf{K}_{\kappa_{n_i}^2}, & \text{for } \alpha_{n_i} = 3, 4, \dots \end{cases} \quad (22)$$

As pointed out by Lindgren et al. (2011) the inverse of \mathbf{C} , i.e., \mathbf{C}^{-1} , is usually a dense matrix. This causes the precision matrix \mathbf{Q}_{f_i} to be dense and ruins all the effort we have made. So we actually used a diagonal matrix $\tilde{\mathbf{C}}$, where $\tilde{\mathbf{C}}_{ii} = \langle \psi_i, 1 \rangle$, instead of \mathbf{C} . This diagonal matrix results in a sparse precision matrix and hence sparse GMRFs models can be obtained. Using the diagonal matrix $\tilde{\mathbf{C}}_{ii}$ yields a Markov approximation to the

FEM solution. The effects of the Markov approximation have been studied by Bolin and Lindgren (2009). They claimed that the difference between the exact representation by the finite element method and the Markov approximation is negligible.

Move to our multivariate GRFs constructed by systems of SPDEs (8). In this case we take the same set of basis functions $\{\psi_k; k = 1, 2, \dots, M\}$ and construct a basis for the solution space for $(x_1, x_2, \dots, x_p)^T$ as

$$p \left\{ \begin{pmatrix} \psi_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} \psi_M \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \psi_1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ \psi_M \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \psi_1 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \psi_M \end{pmatrix} \right\}, \quad (23)$$

where there are a total of Mp basis functions $\{\psi_i\}$ which are numbered in the order listed above. The weak solution of the system of SPDEs (8) requires

$$[\langle \psi_i, \mathcal{L} \mathbf{x} \rangle]_{i=1}^{Mp} \stackrel{d}{=} [\langle \psi_i, \mathbf{f} \rangle]_{i=1}^{Mp}, \quad (24)$$

with \mathcal{L} defined in Equation (9). The finite element representation to the solution of the system of SPDEs is

$$\mathbf{x}(\mathbf{s}) = \sum_{i=1}^{Mp} \psi_i(\mathbf{s}) \omega_i \quad (25)$$

In order to find the precision matrix for the solution we need to define the following matrices,

$$\mathbf{D} = \begin{pmatrix} \tilde{\mathbf{C}} & & & \\ & \tilde{\mathbf{C}} & & \\ & & \ddots & \\ & & & \tilde{\mathbf{C}} \end{pmatrix}, \quad \mathbf{K} = \begin{pmatrix} \mathbf{K}_{11} & \mathbf{K}_{12} & \cdots & \mathbf{K}_{1p} \\ \mathbf{K}_{21} & \mathbf{K}_{22} & \cdots & \mathbf{K}_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{K}_{p1} & \mathbf{K}_{p2} & \cdots & \mathbf{K}_{pp} \end{pmatrix}, \quad (26)$$

$$\mathbf{Q}_f = \begin{pmatrix} \mathbf{Q}_{f_1} & & & \\ & \mathbf{Q}_{f_2} & & \\ & & \ddots & \\ & & & \mathbf{Q}_{f_p} \end{pmatrix},$$

where \mathbf{D} and \mathbf{Q}_f are block diagonal matrices with p blocks on the diagonal. $\{\mathbf{K}_{ij} = b_{ij}(\kappa_{ij}^2 \mathbf{C} + \mathbf{G}); i, j = 1, 2, \dots, p\}$ for $\alpha_{ij} = 2$ and $\{\mathbf{K}_{ij} = b_{ij} \mathbf{I}_{M \times M}; i, j = 1, 2, \dots, p\}$ for $\alpha_{ij} = 0$.

We summarize the results for multivariate GRFs with our systems of SPDEs approach in Main result 1.

Main result 1. *Let \mathbf{Q} be the precision matrix for the Gaussian weights ω_i in the system of SPDEs (8), then the precision matrix of the multivariate GMRf is*

$$\mathbf{Q} = \mathbf{K} \mathbf{D}^{-1} \mathbf{Q}_f \mathbf{D}^{-1} \mathbf{K} \quad (27)$$

with \mathbf{D} and \mathbf{K} defined in (26).

The form of the Main result **1** is similar to the discussion in Lindgren et al. (2011, Appendix C.4). With our approach the precision matrices for the multivariate GRFs are sparse and the smoothness of the fields is mainly controlled by the noise processes. This sparse precision matrix is used for sampling the multivariate GRFs in Section 3.4 and statistical inference with simulated data and real data in Section 4.

3 Model comparison and sampling the GRFs

In this section we focus on the construction of the bivariate GRF, i.e., $p = 2$ in \mathbb{R}^d using the SPDEs approach and then compare with the covariance-based approach presented by Gneiting et al. (2010). As discussed in Section 2.2, the SPDEs approach can construct the same multivariate GRFs as the covariance-based models if $S_{x_{ij}}(\mathbf{k}) = R_{x_{ij}}(\mathbf{k})$ for all $1 \leq i, j \leq q$. The comparison for the univariate GRF is trivial and the multivariate GRF when $p > 2$ can be done in a similar way.

3.1 Bivariate GRF with SPDEs

When $p = 2$ in the system of SPDEs (8), we can construct bivariate GRFs when $b_{12} \neq 0$ or $b_{21} \neq 0$. In this case, the system of SPDEs has the following form

$$\begin{aligned} b_{11}(\kappa_{11}^2 - \Delta)^{\alpha_{11}/2} x_1(\mathbf{s}) + b_{12}(\kappa_{12}^2 - \Delta)^{\alpha_{12}/2} x_2(\mathbf{s}) &= f_1(\mathbf{s}), \\ b_{22}(\kappa_{22}^2 - \Delta)^{\alpha_{22}/2} x_2(\mathbf{s}) + b_{21}(\kappa_{21}^2 - \Delta)^{\alpha_{21}/2} x_1(\mathbf{s}) &= f_2(\mathbf{s}), \end{aligned} \quad (28)$$

and the solution $\mathbf{x}(\mathbf{s}) = (x_1(\mathbf{s}), x_2(\mathbf{s}))^T$ to the system of equations (28) is a bivariate random field. Since it is convenient to study the properties of the bivariate GRFs in the spectral domain, the Fourier transform is applied,

$$\begin{aligned} b_{11}(\kappa_{11}^2 + \|\mathbf{k}\|^2)^{\alpha_{11}/2} \hat{x}_1(\mathbf{k}) + b_{12}(\kappa_{12}^2 + \|\mathbf{k}\|^2)^{\alpha_{12}/2} \hat{x}_2(\mathbf{k}) &= \hat{f}_1(\mathbf{k}), \\ b_{22}(\kappa_{22}^2 + \|\mathbf{k}\|^2)^{\alpha_{22}/2} \hat{x}_2(\mathbf{k}) + b_{21}(\kappa_{21}^2 + \|\mathbf{k}\|^2)^{\alpha_{21}/2} \hat{x}_1(\mathbf{k}) &= \hat{f}_2(\mathbf{k}). \end{aligned} \quad (29)$$

The matrix form of the differential operator in the spectral domain can be written as

$$\mathcal{H}(\boldsymbol{\theta}) = \begin{pmatrix} \mathcal{H}_{11}(\theta_{11}) & \mathcal{H}_{12}(\theta_{12}) \\ \mathcal{H}_{21}(\theta_{21}) & \mathcal{H}_{22}(\theta_{22}) \end{pmatrix}. \quad (30)$$

If the noise processes are mutually independent (but not necessarily identically distributed), the power spectrum matrix of the noise processes is a block diagonal matrix with the form

$$S_{\mathbf{f}}(\mathbf{k}) = \begin{pmatrix} S_{f_1}(\mathbf{k}) & 0 \\ 0 & S_{f_2}(\mathbf{k}) \end{pmatrix}, \quad (31)$$

where $S_{f_1}(\mathbf{k})$ and $S_{f_2}(\mathbf{k})$ are the power spectra in \mathbb{R}^d for the noise processes $f_1(\mathbf{s})$ and $f_2(\mathbf{s})$, respectively. If the noise processes are white, then the problems is simplified and the corresponding power spectra have the forms $S_{\mathcal{W}_1}(\mathbf{k}) = (2\pi)^{-d} \sigma_{n_1}^2$ and $S_{\mathcal{W}_2}(\mathbf{k}) =$

$(2\pi)^{-d}\sigma_{n_2}^2$, where σ_{n_1} and σ_{n_2} are the standard deviations for the white noise processes \mathcal{W}_1 and \mathcal{W}_2 , respectively. However, these two parameters are confounded with $\{b_{i,j}; i, j = 1, 2\}$ and we fix $\sigma_{n_1} = 1$ and $\sigma_{n_2} = 1$. The conclusion is also valid with other types of noise processes. Notice that we have used the new notation $\{S_{\mathcal{W}_i}; i = 1, 2\}$, because it is also possible to use more interesting noise processes which are not only the simple white noise. For instance, we can use the noise processes with Matérn covariance functions. These kinds of noise processes can be generated (independently) from the SPDEs

$$\begin{aligned}(\kappa_{n_1}^2 - \Delta)^{\alpha_{n_1}/2} f_1 &= \mathcal{W}_1, \\(\kappa_{n_2}^2 - \Delta)^{\alpha_{n_2}/2} f_2 &= \mathcal{W}_2,\end{aligned}\tag{32}$$

where \mathcal{W}_1 and \mathcal{W}_2 are standard Gaussian white noise processes and κ_{n_1} and κ_{n_2} are scaling parameters. α_{n_1} and α_{n_2} are related to smoothness parameters ν_{n_1} and ν_{n_2} for f_1 and f_2 and $\alpha_{n_i} = \nu_{n_i} + d/2$. Apply the Fourier transform to (32) and use a similar procedure as defined in Equation (10) - Equation (13), and then the power spectra for the noise processes generated from SPDEs (32) can be obtained and they have the forms

$$\begin{aligned}S_{f_1}(\mathbf{k}) &= \frac{1}{(2\pi)^d} \frac{1}{(\kappa_{n_1}^2 + \|\mathbf{k}\|^2)^{\alpha_{n_1}}}, \\S_{f_2}(\mathbf{k}) &= \frac{1}{(2\pi)^d} \frac{1}{(\kappa_{n_2}^2 + \|\mathbf{k}\|^2)^{\alpha_{n_2}}}.\end{aligned}\tag{33}$$

The power spectrum matrix of the GRFs presented in (13) can also be simplified further and has the form

$$\mathbf{S}_x(\mathbf{k}) = \begin{pmatrix} S_{x_{11}}(\mathbf{k}) & S_{x_{12}}(\mathbf{k}) \\ S_{x_{21}}(\mathbf{k}) & S_{x_{22}}(\mathbf{k}) \end{pmatrix}.\tag{34}$$

By using (12) together with the exact formula for the differential operators (30) and the spectra for the noise processes as defined in (33), we can get an exact symbolic expressions for all the elements in the power spectrum matrix (34)

$$\begin{aligned}S_{x_{11}} &= \frac{S_{f_1}|\mathcal{H}_{22}^2| + S_{f_2}|\mathcal{H}_{12}^2|}{|(\mathcal{H}_{11}\mathcal{H}_{22} - \mathcal{H}_{12}\mathcal{H}_{21})^2|}, \\S_{x_{12}} &= -\frac{\mathcal{H}_{22}S_{f_1}|\mathcal{H}_{21}^2|\mathcal{H}_{11} + \mathcal{H}_{12}S_{f_2}|\mathcal{H}_{11}^2|\mathcal{H}_{21}}{|(\mathcal{H}_{11}\mathcal{H}_{22} - \mathcal{H}_{12}\mathcal{H}_{21})^2|\mathcal{H}_{21}\mathcal{H}_{11}}, \\S_{x_{21}} &= -\frac{\mathcal{H}_{21}S_{f_1}|\mathcal{H}_{22}^2|\mathcal{H}_{12} + \mathcal{H}_{11}S_{f_2}|\mathcal{H}_{12}^2|\mathcal{H}_{22}}{|(\mathcal{H}_{11}\mathcal{H}_{22} - \mathcal{H}_{12}\mathcal{H}_{21})^2|\mathcal{H}_{22}\mathcal{H}_{12}}, \\S_{x_{22}} &= \frac{S_{f_1}|\mathcal{H}_{21}^2| + S_{f_2}|\mathcal{H}_{11}^2|}{|(\mathcal{H}_{11}\mathcal{H}_{22} - \mathcal{H}_{12}\mathcal{H}_{21})^2|}.\end{aligned}\tag{35}$$

In order to simply the problem, we make the operator matrix in the spectral domain a lower triangular matrix by setting $H_{12}(\theta_{12}, \mathbf{k}) = 0$, or equivalently, by setting $b_{12} = 0$. The system of SPDEs in the spectral domain becomes

$$b_{11}(\kappa_{11}^2 + \|\mathbf{k}\|^2)^{\alpha_{11}/2} \hat{x}_1(\mathbf{k}) = \hat{f}_1(\mathbf{k}),\tag{36}$$

$$b_{22}(\kappa_{22}^2 + \|\mathbf{k}\|^2)^{\alpha_{22}/2} \hat{x}_2(\mathbf{k}) + b_{21}(\kappa_{21}^2 + \|\mathbf{k}\|^2)^{\alpha_{21}/2} \hat{x}_1(\mathbf{k}) = \hat{f}_2(\mathbf{k}).\tag{37}$$

This means expression (35) becomes

$$\begin{aligned} S_{x_{11}} &= \frac{S_{f_1}}{|\mathcal{H}_{11}^2|}, & S_{x_{21}} &= -\frac{\bar{\mathcal{H}}_{21}S_{f_1}}{\mathcal{H}_{22}|\mathcal{H}_{11}|^2}, \\ S_{x_{12}} &= -\frac{S_{f_1}\mathcal{H}_{21}}{\mathcal{H}_{22}|\mathcal{H}_{11}|^2}, & S_{x_{22}} &= \frac{|\mathcal{H}_{21}|^2S_{f_1} + |\mathcal{H}_{11}|^2S_{f_2}}{|\mathcal{H}_{11}|^2|\mathcal{H}_{22}|^2}, \end{aligned} \quad (38)$$

where $\bar{\mathcal{H}}_{ij}$ denotes the conjugate of \mathcal{H}_{ij} . This is called the *triangular* version of SPDEs in this paper. If the operators \mathcal{H}_{12} and \mathcal{H}_{21} are real, we obtain $S_{x_{12}}(\mathbf{k}) = S_{x_{21}}(\mathbf{k})$. In other words, we have an imposed symmetry property on the cross-covariance in this case. With the triangular version of the SPDEs, under some extra conditions, the properties of the multivariate GRFs is easy to interpret. We will see more about this in Section 4. In this paper we focus on the triangular version of the SPDEs. However, the full version of the SPDEs could be handled analogously to the triangular version of the SPDEs. If there is no constraint on the operator matrix, i.e., $b_{ij} \neq 0$ for all $i, j = 1, 2$, it is called the *full* version of the SPDEs. In general, $S_{x_{12}}(\mathbf{k})$ and $S_{x_{21}}(\mathbf{k})$ are not necessarily equal. This can release the constraint shared by the Matérn model proposed by Gneiting et al. (2010) and the Linear model of coregionalization (LMC), namely that the cross-covariance has an imposed symmetry property, i.e., $C_{ij}(\mathbf{h}) = C_{ji}(\mathbf{h})$. This does in general not hold as discussion by Wackernagel (2003, Chapter 20). We refer to Goulard and Voltz (1992), Wackernagel (2003, Chapter 14) and Gelfand et al. (2004) for more information about LMC. In this paper we assume that all the operators are real which simplifies calculations and discussion. We reorganize (38) and find

$$\begin{aligned} \frac{S_{x_{12}}}{S_{x_{11}}} &= -\frac{\mathcal{H}_{21}}{\mathcal{H}_{22}}, \\ \frac{S_{x_{22}}}{S_{x_{11}}} &= \frac{|\mathcal{H}_{21}^2|}{|\mathcal{H}_{22}^2|} + \frac{|\mathcal{H}_{11}^2|}{\mathcal{H}_{22}^2} \frac{S_{f_2}}{S_{f_1}}, \\ \frac{S_{x_{22}}}{S_{x_{12}}} &= -\left(\frac{|\mathcal{H}_{11}^2|}{\mathcal{H}_{21}\mathcal{H}_{22}}\right) \frac{S_{f_2}}{S_{f_1}} - \frac{\mathcal{H}_{21}}{\mathcal{H}_{22}}. \end{aligned} \quad (39)$$

From the results given in (35) and (39), we can notice that $S_{x_{12}}$ and $S_{x_{11}}$ only depend on the power spectrum of the noise process f_1 and $S_{x_{22}}(\mathbf{k})$ depends on the noise process power spectra $S_{f_1}(\mathbf{k})$ and $S_{f_2}(\mathbf{k})$. The ratio between the power spectra $S_{x_{12}}$ and $S_{x_{11}}$ is independent of both the noise processes.

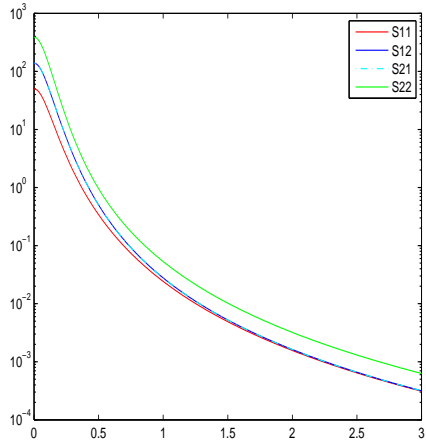
When $\{f_i(\mathbf{s}); i = 1, 2\}$ are mutually independent and generated from (32), the elements in the power spectrum matrix (34) can be written down explicitly by using (33) and (38),

Table 1: Parameter values for asymptotic behaviors of the power spectra for bivariate GRFs

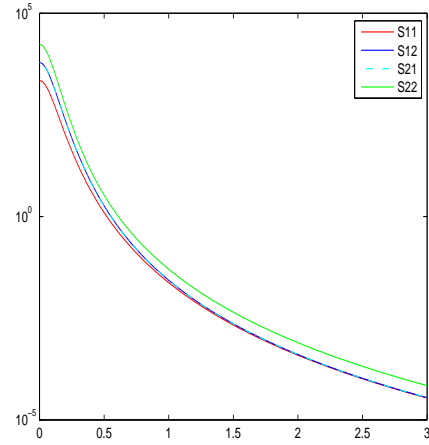
case 1			case 2		
α	κ	\mathbf{b}	α	κ	\mathbf{b}
$\alpha_{11} = 2$	$\kappa_{11} = 0.15$	$b_{11} = 1$	$\alpha_{11} = 2$	$\kappa_{11} = 0.15$	$b_{11} = 1$
$\alpha_{12} = 0$	$\kappa_{12} = 0$	$b_{12} = 0$	$\alpha_{12} = 0$	$\kappa_{12} = 0$	$b_{12} = 0$
$\alpha_{21} = 2$	$\kappa_{21} = 0.5$	$b_{21} = -1$	$\alpha_{21} = 2$	$\kappa_{21} = 0.5$	$b_{21} = -1$
$\alpha_{22} = 2$	$\kappa_{22} = 0.3$	$b_{22} = 1$	$\alpha_{22} = 2$	$\kappa_{22} = 0.3$	$b_{22} = 1$
$\alpha_{n_1} = 0$	$\kappa_{n_1} = 0.15$		$\alpha_{n_1} = 1$	$\kappa_{n_1} = 0.15$	
$\alpha_{n_1} = 0$	$\kappa_{n_2} = 0.3$		$\alpha_{n_2} = 1$	$\kappa_{n_2} = 0.3$	
case 3			case 4		
α	κ	\mathbf{b}	α	κ	\mathbf{b}
$\alpha_{11} = 2$	$\kappa_{11} = 0.15$	$b_{11} = 1$	$\alpha_{11} = 2$	$\kappa_{11} = 0.15$	$b_{11} = 1$
$\alpha_{12} = 0$	$\kappa_{12} = 0$	$b_{12} = 0$	$\alpha_{12} = 0$	$\kappa_{12} = 0$	$b_{12} = 0$
$\alpha_{21} = 2$	$\kappa_{21} = 0.15$	$b_{21} = -0.5$	$\alpha_{21} = 2$	$\kappa_{21} = 0.5$	$b_{21} = -1$
$\alpha_{22} = 2$	$\kappa_{22} = 0.3$	$b_{22} = 1$	$\alpha_{22} = 2$	$\kappa_{22} = 0.3$	$b_{22} = 1$
$\alpha_{n_1} = 0$	$\kappa_{n_1} = 0.15$		$\alpha_{n_1} = 0$	$\kappa_{n_1} = 0.15$	
$\alpha_{n_1} = 0$	$\kappa_{n_2} = 0.3$		$\alpha_{n_2} = 0$	$\kappa_{n_2} = 0.3$	

$$\begin{aligned}
S_{x_{11}}(\mathbf{k}) &= \frac{1}{(2\pi)^d (\kappa_{n_1}^2 + \|\mathbf{k}\|^2)^{\alpha_{n_1}} b_{11}^2 (\kappa_{11}^2 + \|\mathbf{k}\|^2)^{\alpha_{11}}}, \\
S_{x_{12}}(\mathbf{k}) &= -\frac{b_{21} (\kappa_{n_1}^2 + \|\mathbf{k}\|^2)^{-\alpha_{n_1}} (\kappa_{21}^2 + \|\mathbf{k}\|^2)^{\alpha_{21}/2}}{(2\pi)^d (b_{22} (\kappa_{22}^2 + \|\mathbf{k}\|^2)^{\alpha_{22}/2} b_{11}^2 (\kappa_{11}^2 + \|\mathbf{k}\|^2)^{\alpha_{11}})}, \\
S_{x_{22}}(\mathbf{k}) &= \frac{\frac{b_{21}^2 (\kappa_{n_1}^2 + \|\mathbf{k}\|^2)^{\alpha_{21}}}{(2\pi)^d (\kappa_{n_1}^2 + \|\mathbf{k}\|^2)^{\alpha_{n_1}}} + \frac{b_{11}^2 (\kappa_{11}^2 + \|\mathbf{k}\|^2)^{\alpha_{11}}}{(2\pi)^d (\kappa_{n_2}^2 + \|\mathbf{k}\|^2)^{\alpha_{n_2}}}}{(b_{11}^2 (\kappa_{11}^2 + \|\mathbf{k}\|^2)^{\alpha_{11}}) (b_{22}^2 (\kappa_{22}^2 + \|\mathbf{k}\|^2)^{\alpha_{22}})}.
\end{aligned} \tag{40}$$

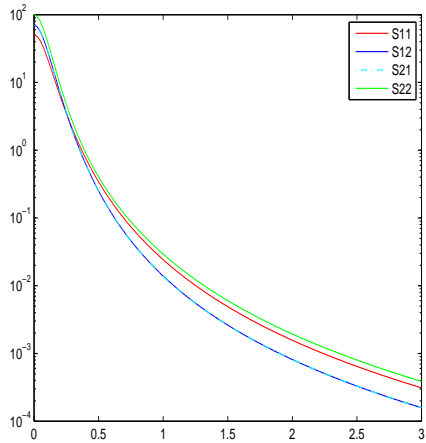
The asymptotic behavior of the power spectra for the bivariate GRF can be obtained from (40). For some selected parameter values defined in Table 1, the power spectra are shown in Figure 1. From these figure, we can notice that the parameters $\{b_{ij}; i, j = 1, 2, i > j\}$ controls the correlation between the two fields. When a smaller absolute value of b_{21} is chosen, the correlation between these two fields decreases rapidly. We can also show that the sign of $b_{21} \cdot b_{22}$ is related the sign of the correlation between the two GRFs. Obviously, b_{ij} are also related to the variance of the GRFs. The parameters κ_{ij} are related to the range of the two fields. α_{ij} and α_{n_i} are related to the smoothness of the two GRFs. The asymptotic behavior is curial when dealing with the real-world applications.



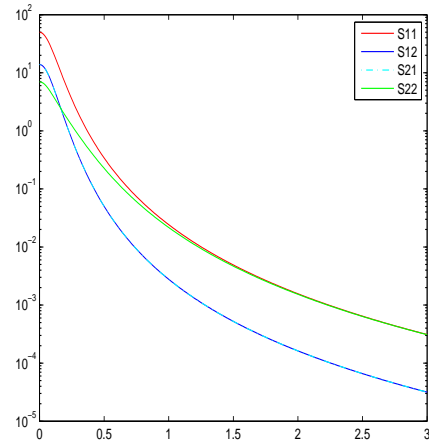
(a)



(b)



(c)



(d)

Figure 1: Asymptotic behaviors of the power spectra corresponding to case 1 (a), case 2 (b), case 3 (c) and case 4 (d) with parameter values given in Table 1.

3.2 Bivariate GRF with the covariance-based model

As presented by Gneiting et al. (2010), a multivariate GRFs can be constructed using a covariance-based model. In the bivariate setting, the matrix-valued covariance function for the bivariate GRF contains 4 elements and the corresponding power spectrum matrix becomes

$$\mathbf{R}(\mathbf{k}) = \begin{pmatrix} R_{x_{11}}(\mathbf{k}) & R_{x_{12}}(\mathbf{k}) \\ R_{x_{21}}(\mathbf{k}) & R_{x_{22}}(\mathbf{k}) \end{pmatrix}. \quad (41)$$

Applying integral given in Equation (14), together with the expression for the marginal variances in Equation (3) and cross-covariance in Equation (4), we can find closed form for the elements in the power spectrum matrix,

$$\begin{aligned} R_{x_{ii}}(\mathbf{k}) &= \frac{1}{(2\pi)^d} \left[a_{ii}^{2\nu_{ii}} (\sqrt{4\pi})^d \frac{\sigma_{ii} \Gamma(\nu_{ii} + d/2)}{(a_{ii}^2 + \|\mathbf{k}\|^2)^{\nu_{ii} + d/2} \Gamma(\nu_{ii})} \right], \\ R_{x_{ij}}(\mathbf{k}) &= \frac{1}{(2\pi)^d} \left[a_{ij}^{2\nu_{ij}} (\sqrt{4\pi})^d \frac{\rho_{ij} \sigma_i \sigma_j \Gamma(\nu_{ij} + d/2)}{(a_{ij}^2 + \|\mathbf{k}\|^2)^{\nu_{ij} + d/2} \Gamma(\nu_{ij})} \right]. \end{aligned} \quad (42)$$

Notice that if $d = \{2, 4, \dots\}$, then the expressions in (42) can be simplified even further by (recursively) using the well-known formula

$$\Gamma(\nu + 1) = \nu \Gamma(\nu).$$

Since all the components $\{C_{ij}; i, j = 1, 2\}$ in the matrix-valued covariance matrix \mathbf{C} are from Matérn family, as defined in Equation (3) and Equation (4) (Gneiting et al., 2010), the power spectra of the marginal covariance functions $\{C_{ii}; i = 1, 2\}$ and the cross covariance functions $\{C_{ij}; i, j = 1, 2, i \neq j\}$ have similar forms as indicated in (42).

3.3 Parameter matching

As mentioned in Section 2.2, the model comparison in this paper is based on the fact that if $S_{ij}(\mathbf{k}) = R_{ij}(\mathbf{k})$, for each $1 \leq i, j \leq p$, and then the multivariate GRFs constructed from the SPDEs approach (8) and from the covariance-based model (2) will be equivalent.

By comparing each element in (40) and (42), we can get the results given as follows,

$$\begin{aligned} 0 &= \rho_{12} - \rho_{21}, \\ \kappa_{11} = \kappa_{21} = \kappa_{22} = \kappa_{n_1} = \kappa_{n_2} &= a_{11} = a_{12} = a_{21} = a_{22}, \\ \alpha_{11} + \alpha_{n_1} &= \nu_{11} + \frac{d}{2}, \\ \alpha_{11} + \frac{\alpha_{22}}{2} + \alpha_{n_1} - \frac{\alpha_{21}}{2} &= \nu_{12} + \frac{d}{2}, \end{aligned} \quad (43)$$

and

$$\begin{aligned}
\frac{1}{b_{11}^2} &= \frac{(4\pi)^{d/2} a_{11}^{2\nu_{n1}} \sigma_{11} \Gamma(\nu_{11} + d/2)}{\Gamma(\nu_{11})}, \\
-\frac{b_{21}}{b_{22} b_{11}^2} &= \frac{(4\pi)^{d/2} a_{12}^{2\nu_{12}} \Gamma(\nu_{12} + d/2) \rho_{12} \sigma_1 \sigma_2}{\Gamma(\nu_{12})}, \\
\frac{b_{21}^2 + b_{11}^2}{b_{11}^2 b_{22}^2} &= \frac{(4\pi)^{d/2} \sigma_{22} a_{22}^{2\nu_{22}} \Gamma(\nu_{22} + d/2)}{\Gamma(\nu_{22})}.
\end{aligned} \tag{44}$$

More equalities can be obtained based on the relationship between $\alpha_{21} + \alpha_{n_2}$ and $\alpha_{11} + \alpha_{n_1}$. When $\alpha_{21} + \alpha_{n_2} \leq \alpha_{11} + \alpha_{n_1}$, we can get

$$\alpha_{n_2} + \alpha_{22} = \nu_{22} + \frac{d}{2}. \tag{45}$$

When $\alpha_{21} + \alpha_{n_2} \geq \alpha_{11} + \alpha_{n_1}$, we can get

$$\alpha_{11} + \alpha_{22} + \alpha_{n_1} - \alpha_{21} = \nu_{22} + \frac{d}{2}. \tag{46}$$

The Equation (45) and Equation (46) are critical when discussing about the asymptotic behaviors of bivariate GRF and the corresponding power spectra. From (43), we can notice that in order to construct equivalent models to the ones as discussed by Gneiting et al. (2010), we need to have the constraint that all the scaling parameters should be equal in addition to some constraints for the smoothness parameters given in the third and the fourth equations in (43) and one more conditional equality given in (45) or (46). From the first equation in (43), we notice the extra symmetry restriction discussed in Gneiting et al. (2010, Section 4) also must be fulfilled. Additionally, from the second equation in (44), we notice that there is one important relationship

$$\rho_{12} b_{22} b_{21} \leq 0, \tag{47}$$

between the co-located correlation coefficient ρ_{12} and b_{21} and b_{22} . It shows that the correlation ρ_{12} must have opposite sign to the product of b_{21} and b_{22} . This is an important information not only for sampling from the bivariate GRFs but also for interpreting the results from inference.

From the results given in (44), we can obtain the following results

$$\begin{aligned}
-\frac{b_{22}}{b_{21}} &= \frac{a_{11}^{\nu_{11}} \sigma_1 \Gamma(\nu_{11} + d/2) / \Gamma(\nu_{11})}{a_{12}^{\nu_{12}} \rho \sigma_2 \Gamma(\nu_{12} + d/2) / \Gamma(\nu_{12})}, \\
\frac{b_{22}^2}{b_{11}^2 + b_{21}^2} &= \frac{a_{11}^{\nu_{11}} \sigma_{11} \Gamma(\nu_{11} + d/2) / \Gamma(\nu_{11})}{a_{22}^{\nu_{22}} \sigma_{22} \Gamma(\nu_{22} + d/2) / \Gamma(\nu_{22})}.
\end{aligned} \tag{48}$$

Notice that when $a_{11} = a_{21} = a_{22}$ and $\nu_{11} = \nu_{21} = \nu_{22}$, the above results in (48) can be simplified to the following form

$$\begin{aligned}
-\frac{b_{22}}{b_{21}} &= \frac{\sigma_1}{\rho\sigma_2}, \\
\frac{b_{22}^2}{b_{11}^2 + b_{21}^2} &= \frac{\sigma_{11}}{\sigma_{22}}.
\end{aligned}
\tag{49}$$

These results show that the parameters $\{b_{ij}; i, j = 1, 2\}$ are not only connected to the correlation between these two fields, but also connected with the marginal variance of the GRF. From the results given in (43) to (49), we can notice that the parameters in the system of SPDEs have similar interpretations as the parameters from the covariance-based model.

From these results we see that it is possible to construct *multivariate* GRFs through the SPDEs approach as the covariance-based approach proposed by Gneiting et al. (2010). Additionally, there are three main advantages for our SPDEs approach. The first advantage is that our new approach does not explicitly depend on the theory of positive definite matrix. We do not need to worry about the notorious requirement of positive definite covariance matrices. The second advantage is that we can remove the symmetry property which is shared by the covariance-based approach and the LMC approach (Gelfand et al., 2004; Gneiting et al., 2010). The third advantage, which has not yet been discussed, is that we can construct multivariate GRFs on manifolds, such as the sphere \mathbb{S}^2 , by simply reinterpret the systems of SPDEs to be defined on the manifold. Lindgren et al. (2011, Section 3.1) discussed the theoretical background in the univariate setting, which is basically the same as for our multivariate setting.

Furthermore, we can actually go even further for the multivariate GRFs, such as multivariate GRFs with oscillating covariance functions and non-stationary multivariate GRFs. These kinds of multivariate GRFs are under development but are beyond the scope of this paper.

3.4 Sampling bivariate GRFs and trivariate GRFs

As presented in Section 2.3 we can construct multivariate GRFs by the SPDEs approach theoretically, but do the computations using the GMRF representation. The precision matrix \mathbf{Q} is usually sparse and the sparseness of the precision matrices enables us to apply numerical linear algebra for sparse matrices for sampling from the GRFs and for fast statistical inference. We now assume that the multivariate GMRF has mean vector $\boldsymbol{\mu}$ and precision matrix \mathbf{Q} , i.e., $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{Q}^{-1})$. When sampling from the GMRF, the forward substitution and backward substitution is applied by using the Cholesky triangle \mathbf{L} , where by definition $\mathbf{Q} = \mathbf{L}\mathbf{L}^T$. The commonly used steps for sampling a GRF are as follows.

1. Calculate the Cholesky triangle \mathbf{L} from the Cholesky factorization;
2. Sample $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$, where \mathbf{I} is the identity matrix with the same dimensions as the precision matrix \mathbf{Q} ;

Table 2: Parameters for sampling bivariate GRFs

positively correlated GRFs			negatively correlated GRFs		
α	κ	b	α	κ	b
$\alpha_{11} = 2$	$\kappa_{11} = 0.15$	$b_{11} = 1$	$\alpha_{11} = 2$	$\kappa_{11} = 0.15$	$b_{11} = 1$
$\alpha_{12} = 0$	$\kappa_{12} = 0$	$b_{12} = 0$	$\alpha_{12} = 0$	$\kappa_{12} = 0$	$b_{12} = 0$
$\alpha_{21} = 2$	$\kappa_{21} = 0.5$	$b_{21} = -1$	$\alpha_{21} = 2$	$\kappa_{21} = 0.5$	$b_{21} = 1$
$\alpha_{22} = 2$	$\kappa_{22} = 0.3$	$b_{22} = 1$	$\alpha_{22} = 2$	$\kappa_{22} = 0.3$	$b_{22} = 1$
$\alpha_{n_1} = 0$	$\kappa_{n_1} = 0.15$		$\alpha_{n_1} = 0$	$\kappa_{n_1} = 0.15$	
$\alpha_{n_2} = 0$	$\kappa_{n_2} = 0.3$		$\alpha_{n_2} = 0$	$\kappa_{n_2} = 0.3$	

- Solve the linear system of equations $L\mathbf{v} = \mathbf{z}$ for \mathbf{v} . Then \mathbf{v} has the correct precision matrix \mathbf{Q} , and $\mathbf{v} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}^{-1})$;
- Finally correct the mean by computing $\mathbf{x} = \boldsymbol{\mu} + \mathbf{v}$.

Then \mathbf{x} is a sample of the GMRF with correct mean $\boldsymbol{\mu}$ and precision matrix \mathbf{Q} . If \mathbf{Q} is a band matrix, a band-Cholesky factorization can be used with the algorithm given by Rue and Held (2005, Algorithm 2.9, Page 45). Different types of parametrization of GMRFs and their corresponding sampling procedures can be found in Rue and Held (2005). If the mean of the field is $\mathbf{0}$, the Step 4 is not needed and \mathbf{v} is a sample from the GMRF.

With white noise at the right hand of SPDEs (28), and the parameters from the differential operators given in Table 2, we can get samples for positively correlated random fields with $b_{21} < 0$ and negatively correlated random fields with $b_{21} > 0$. These two samples are shown in Figure 2 and Figure 3. Choosing the reference points in the middle of the two GRFs for these two cases, we can get the corresponding marginal correlations within each of the GRFs and the cross-correlations between the two GRFs. The results are shown in Figure 4 and Figure 6. We can see that the random fields are isotropic and have the same correlation range. This is because we have chosen almost the same parameters for our cases, except the signs of b_{21} .

From Figure 2 - Figure 6 and many more samples from the bivariate GRFs which we have simulated, we can notice that the sign of $b_{21} \cdot b_{22}$ is related to the sign of the correlation between these two GRFs, which corresponds to the comparison result in (47). The smoothness of the GRFs are related to the values α_{ij} . These results also verify the conclusion given in (43).

We can also notice that the first GRF is a Matérn random field when we use the white noise process as the driving process or under the condition that $\kappa_{n_1} = \kappa_{11}$ with the triangular systems of SPDEs. The second field, in general, is not a Matérn random field, but it can be relatively close to a Matérn random field. With additional conditions, the second random fields could also be a Matérn random field, but this is not focused in this paper. With the triangular system of SPDEs, together with some other conditions, the correlation range which was mentioned in Section 1 for the first random field can be calculated by the empirically derived formula $\rho = \sqrt{8\nu}/\kappa$ but not for the second random

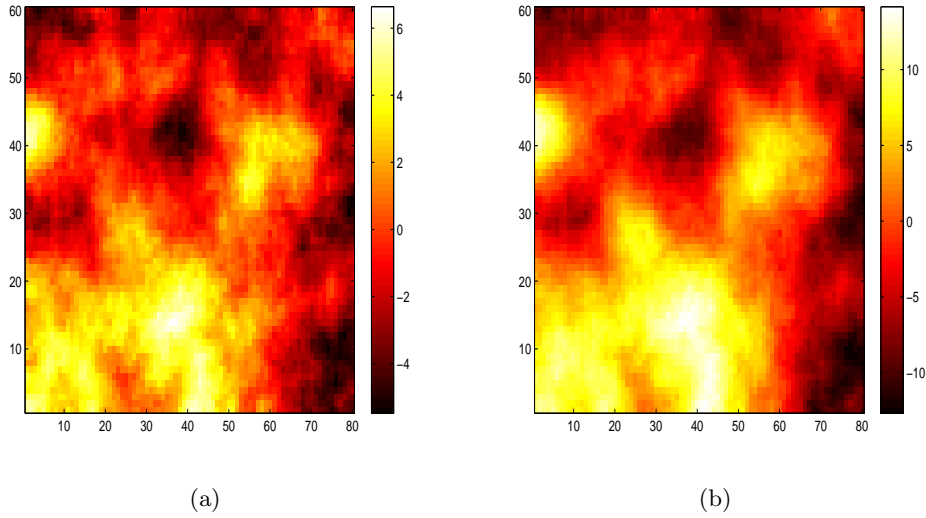


Figure 2: One simulated realization from the bivariate Gaussian Random Field with positive correlation with parameters $\alpha_{11} = 2, \alpha_{12} = 0, \alpha_{21} = 2, \alpha_{22} = 2, \alpha_{n_1} = 0, \alpha_{n_2} = 0, \kappa_{11} = 0.15, \kappa_{22} = 0.3, \kappa_{21} = 0.5, \kappa_{n_1} = 0.15, \kappa_{n_2} = 0.3, b_{11} = 1, b_{12} = 0, b_{22} = 1$ and $b_{21} = -1$.

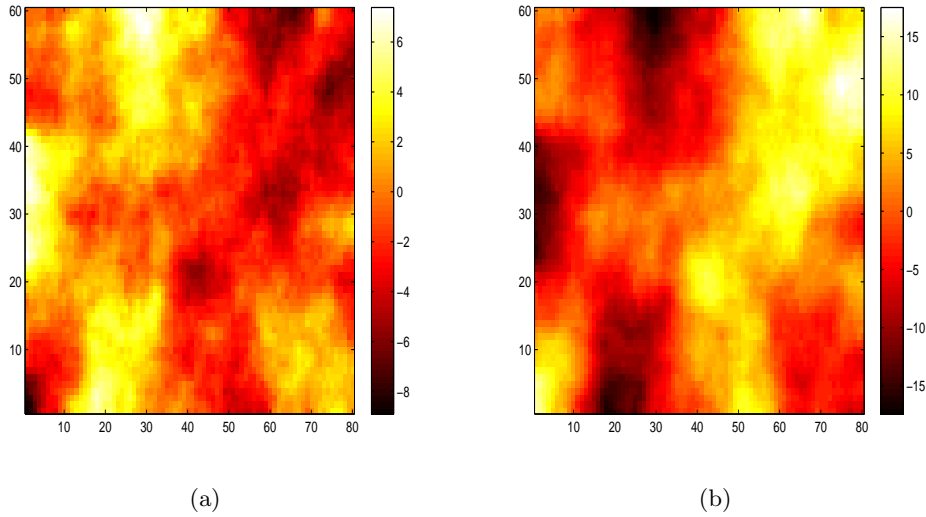


Figure 3: One simulated realization from the bivariate Gaussian Random Field with negative correlation with parameters $\alpha_{11} = 2, \alpha_{12} = 0, \alpha_{21} = 2, \alpha_{22} = 2, \alpha_{n_1} = 0, \alpha_{n_2} = 0, \kappa_{11} = 0.15, \kappa_{22} = 0.3, \kappa_{21} = 0.5, \kappa_{n_1} = 0.15, \kappa_{n_2} = 0.3, b_{11} = 1, b_{12} = 0, b_{22} = 1$ and $b_{21} = 1$.

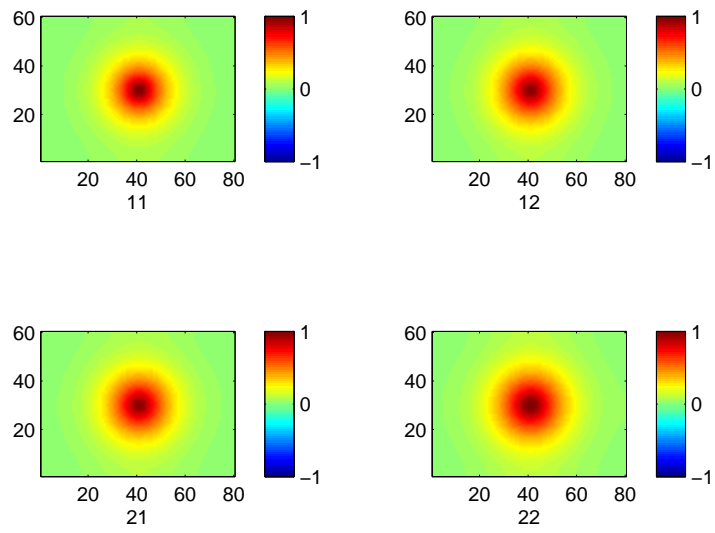


Figure 4: Marginal correlations (diagonal direction) and cross-correlations (anti-diagonal direction) with the reference points in the middle of the two GRFs for the positively correlated random fields.

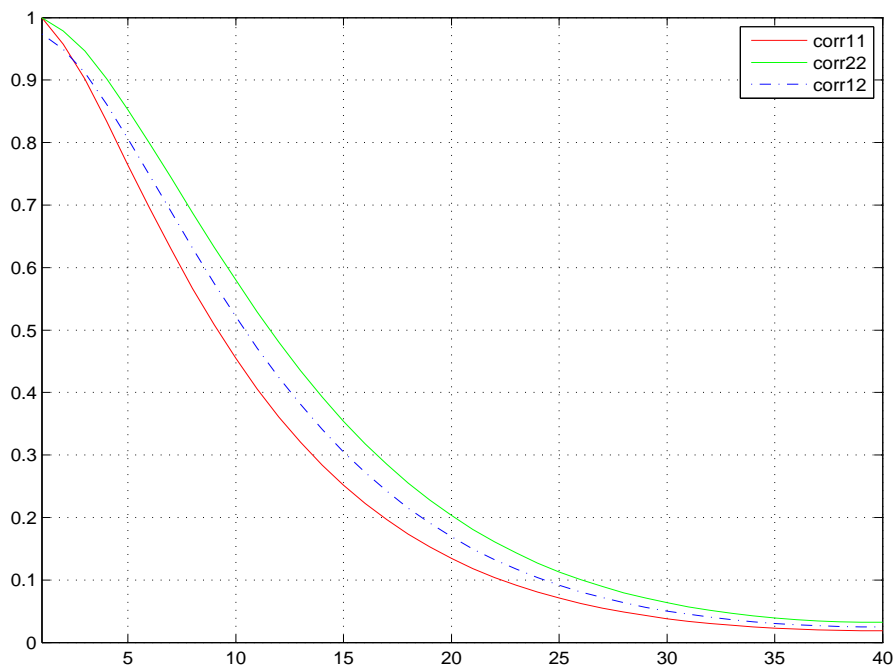


Figure 5: Marginal correlations and cross-correlations with the reference points in the middle of the two GRFs for the positively correlated random fields. 'corr11' means the marginal correlation within random field 1. 'corr22' means the marginal correlation within random field 2. 'corr12' means the cross-correlation between random fields 1 and 2.

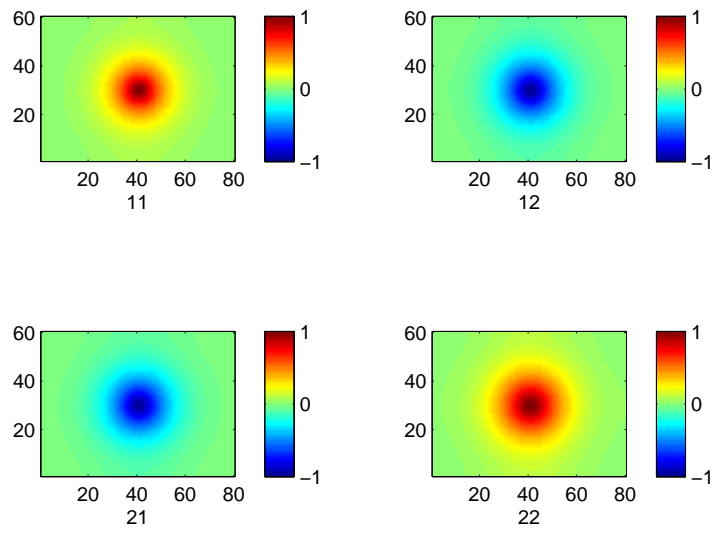


Figure 6: Marginal correlations (diagonal direction) and cross-correlations (anti-diagonal direction) with the reference points in the middle of the two GRFs for the negatively correlated random fields.

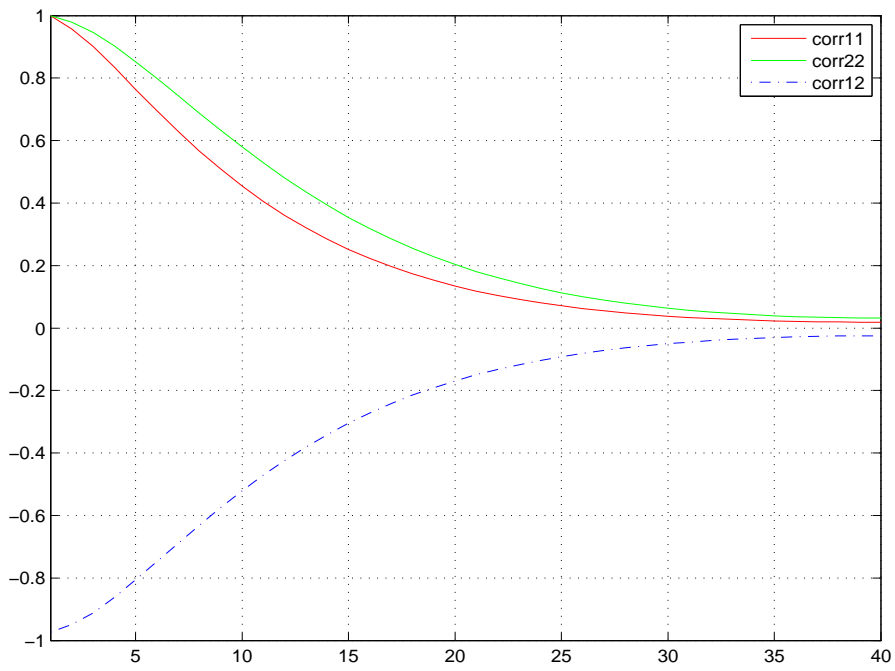


Figure 7: Marginal correlations and cross-correlations with the reference points in the middle of the two GRFs for the negatively correlated random fields. 'corr11' means the marginal correlation within random field 1. 'corr22' means the marginal correlation within random field 2. 'corr12' means the cross-correlation between random fields 1 and 2.

Table 3: Parameters for sampling trivariate GRFs

α	κ	b
$\alpha_{11} = 2$	$\kappa_{11} = 0.5$	$b_{11} = 1$
$\alpha_{12} = 0$	$\kappa_{12} = 0$	$b_{12} = 0$
$\alpha_{13} = 0$	$\kappa_{13} = 0$	$b_{13} = 0$
$\alpha_{21} = 2$	$\kappa_{21} = 0.6$	$b_{21} = 0.8$
$\alpha_{22} = 2$	$\kappa_{22} = 0.4$	$b_{22} = 1$
$\alpha_{23} = 0$	$\kappa_{23} = 0$	$b_{23} = 0$
$\alpha_{31} = 2$	$\kappa_{31} = 0.5$	$b_{31} = 1$
$\alpha_{32} = 2$	$\kappa_{32} = 1$	$b_{32} = 0.9$
$\alpha_{33} = 2$	$\kappa_{33} = 0.3$	$b_{33} = 1$
$\alpha_{n_1} = 1$	$\kappa_{n_1} = 0.5$	
$\alpha_{n_2} = 1$	$\kappa_{n_2} = 0.4$	
$\alpha_{n_3} = 1$	$\kappa_{n_3} = 0.3$	

field. In general we need to find the correlation ranges for the random fields numerically.

For trivariate GRFs the sampling procedure is exactly the same as for bivariate GRFs. We use the triangular system of the SPDEs. The true values given in Table 7 are used. One sample from the trivariate GRFs is shown in Figure 8 with the corresponding correlation functions given in Figure 9. From these figures, we notice that the trivariate random fields have similar interpretation as the bivariate random fields, but are more complicated. Since the triangular version of the system of SPDEs has been used, the sign of $b_{21} \cdot b_{22}$ is related to the sign of the correlation between the first two GRFs. But for the third fields, it is not only related to the sign of $b_{32} \cdot b_{33}$ but also the influence from the sign of $b_{21} \cdot b_{22}$. By choosing a different parametrization, we can end up with more interesting models such as the random field 1 and random field 3 are positively correlated in some locations and negatively correlated in other locations. We are not going to discuss these issues here since this is ongoing research.

4 Examples and applications

In order to illustrate how to use the SPDEs approach for constructing multivariate GRFs and the usefulness of our approach, some examples both with simulated data and real data are chosen. First, some basic theory on inference with multivariate GMRFs is given. We focus on the triangular system of SPDEs in this paper.

We use the bivariate GMRFs as an example. The multivariate GMRFs can be done analogously. Assume that we have used N triangles in the discretization for each random field $x_i(\mathbf{s})|\theta_i$, $i = 1, 2$. For the bivariate case $\mathbf{x}(\mathbf{s}) = (x(\mathbf{s})_1, x(\mathbf{s})_2)^T$ has a $2N$ -dimensional multivariate Gaussian distribution with probability density function

$$\pi(\mathbf{x}|\boldsymbol{\theta}) = \left(\frac{1}{2\pi}\right)^{2N} |\mathbf{Q}(\boldsymbol{\theta})|^{1/2} \exp\left(-\frac{1}{2}\mathbf{x}^T \mathbf{Q}(\boldsymbol{\theta}) \mathbf{x}\right), \quad (50)$$

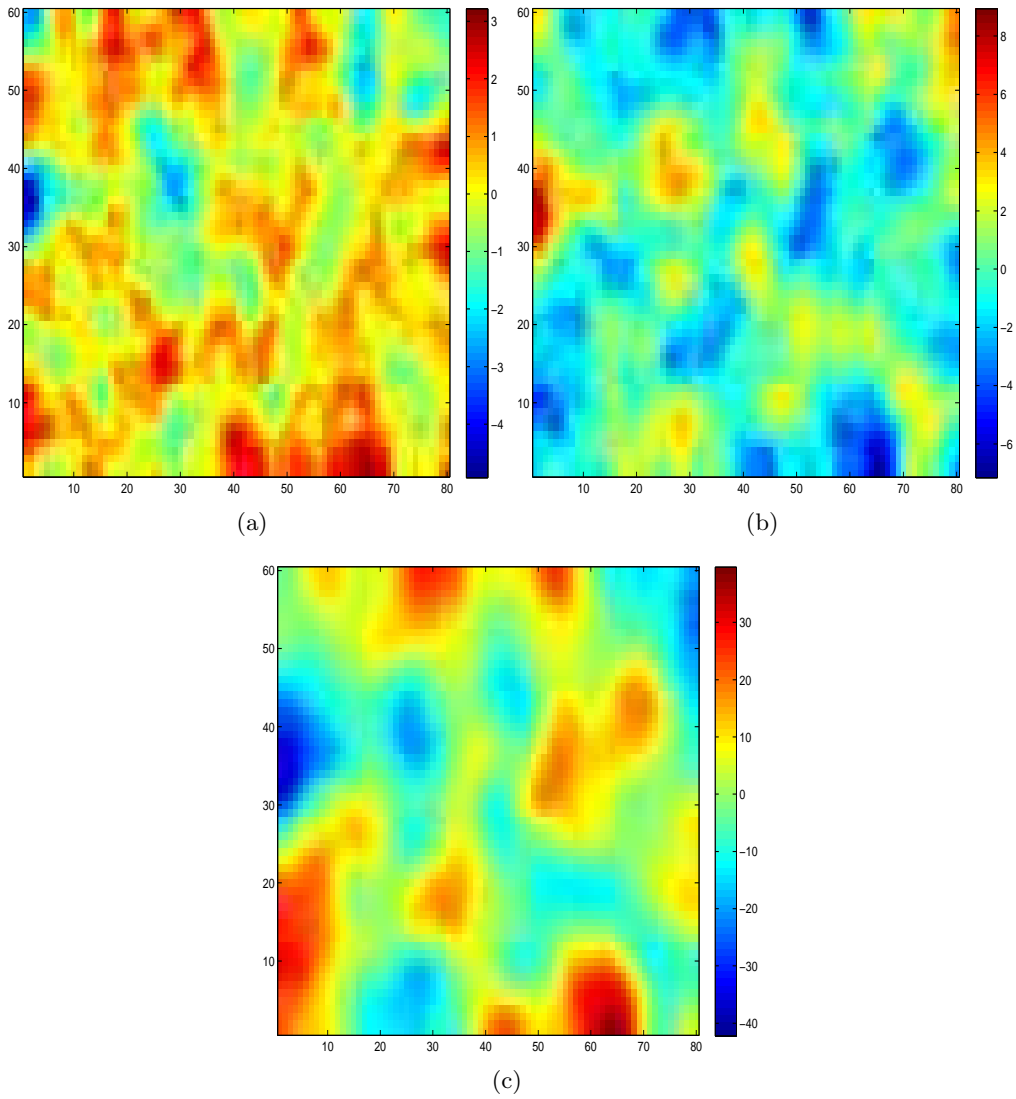


Figure 8: One simulated realization from the trivariate Gaussian Random Fields with parameters given in Table 3.

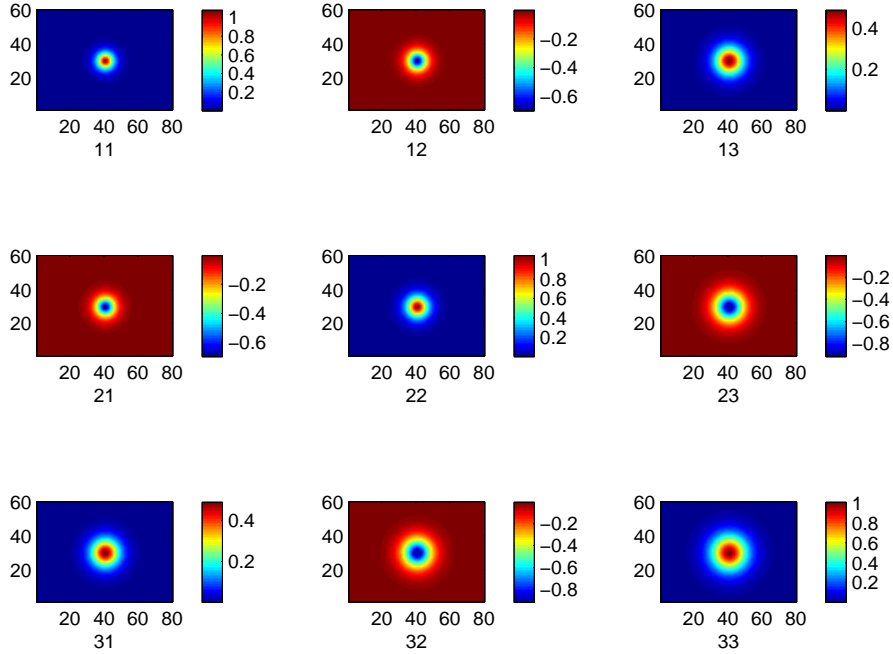


Figure 9: Marginal correlation matrices (diagonal) and cross-correlation matrices (off-diagonal) with each of the reference point in the middle of the three GRFs.

where $\mathbf{Q}(\boldsymbol{\theta})$ is the precision matrix of the bivariate GMRF with the parameters $\boldsymbol{\theta}$. Furthermore, we assume the length of the data $\mathbf{y} = (y_1, y_2)^\top$ is $t = k_1 + k_2$ where y_1 is the observation of $x_1(\mathbf{s})$ with length k_1 and y_2 is the observation of $x_2(\mathbf{s})$ with length k_2 . Then \mathbf{y} is a t -dimensional random variable with probability density function

$$\pi(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta}) = \left(\frac{1}{2\pi}\right)^t |\mathbf{Q}_n|^{1/2} \exp\left(-\frac{1}{2}(\mathbf{y} - \mathbf{A}\mathbf{x})^\top \mathbf{Q}_n (\mathbf{y} - \mathbf{A}\mathbf{x})\right), \quad (51)$$

where \mathbf{Q}_n is defined in Section 2.3 with size $t \times t$, and \mathbf{A} is a $t \times 2N$ matrix which links the sparse observations to our bivariate GMRFs. Notice that the density function $\pi(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta})$ is independent of $\boldsymbol{\theta}$. Hence we can write the probability density function as $\pi(\mathbf{y}|\mathbf{x})$. We first find the probability density function of $\mathbf{x}|\mathbf{y}, \boldsymbol{\theta}$ from (50) - (51)

$$\begin{aligned} \pi(\mathbf{x}|\mathbf{y}, \boldsymbol{\theta}) &\propto \pi(\mathbf{x}, \mathbf{y}|\boldsymbol{\theta}) \\ &= \pi(\mathbf{x}|\boldsymbol{\theta})\pi(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta}) \\ &\propto \exp\left(-\frac{1}{2}(\mathbf{x}^\top (\mathbf{Q}(\boldsymbol{\theta}) + \mathbf{A}^\top \mathbf{Q}_n \mathbf{A})\mathbf{x} - 2\mathbf{x}^\top \mathbf{A}^\top \mathbf{Q}_n \mathbf{y})\right). \end{aligned} \quad (52)$$

Let $\boldsymbol{\mu}_c(\boldsymbol{\theta}) = \mathbf{Q}_c^{-1}(\boldsymbol{\theta})\mathbf{A}^\top \mathbf{Q}_n \mathbf{y}$, and $\mathbf{Q}_c(\boldsymbol{\theta}) = \mathbf{Q}(\boldsymbol{\theta}) + \mathbf{A}^\top \mathbf{Q}_n \mathbf{A}$, and then $\pi(\mathbf{x}|\mathbf{y}, \boldsymbol{\theta})$ can be denoted as

$$\mathbf{x}|\mathbf{y}, \boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{\mu}_c(\boldsymbol{\theta}), \mathbf{Q}_c^{-1}(\boldsymbol{\theta})),$$

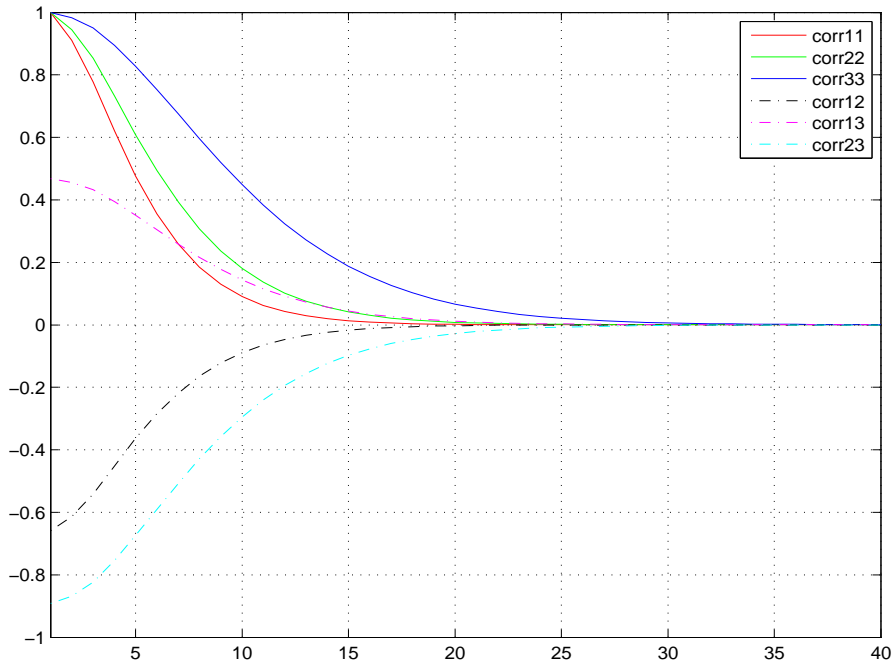


Figure 10: Marginal correlation matrices and cross-correlation matrices with each of the reference point in the middle of the three GRFs. 'corr11' means the marginal correlation within random field 1. 'corr22' means the marginal correlation within random field 2. 'corr33' means the marginal correlation within random field 3. 'corr12' means the cross-correlation between random fields 1 and 2. 'corr13' means the cross-correlation between random fields 1 and 3. 'corr23' means the cross-correlation between random fields 2 and 3.

or in the canonical parametrization as

$$\mathbf{x}|\mathbf{y}, \boldsymbol{\theta} \sim \mathcal{N}_c(\mathbf{b}_c, \mathbf{Q}_c(\boldsymbol{\theta})),$$

with $\mathbf{b}_c = \mathbf{A}^\top \mathbf{Q}_n \mathbf{y}$. Thus $\mathbf{x}|\mathbf{y}, \boldsymbol{\theta}$ is a $2N$ -dimensional multivariate Gaussian distribution. The canonical parametrization for a GMRF is useful with successive conditioning (Rue and Held, 2005). For more information about the canonical parametrization for the GMRFs, we refer to Rue and Held (2005, Chapter 2.2.3).

The probability density function (52) can be used to integrate out \mathbf{x} from the joint density of \mathbf{x} , \mathbf{y} and $\boldsymbol{\theta}$,

$$\begin{aligned} \pi(\mathbf{y}, \boldsymbol{\theta}) &= \frac{\pi(\boldsymbol{\theta}, \mathbf{x}, \mathbf{y})}{\pi(\mathbf{x}|\boldsymbol{\theta}, \mathbf{y})} \\ &= \frac{\pi(\boldsymbol{\theta})\pi(\mathbf{x}|\boldsymbol{\theta})\pi(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta})}{\pi(\mathbf{x}|\mathbf{y}, \boldsymbol{\theta})}, \end{aligned} \quad (53)$$

where $\pi(\boldsymbol{\theta})$ is the prior distribution of $\boldsymbol{\theta}$. With (53), the posterior distribution of $\boldsymbol{\theta}$ can be obtained as

$$\begin{aligned} \pi(\boldsymbol{\theta}|\mathbf{y}) &\propto \pi(\boldsymbol{\theta}) \frac{|\mathbf{Q}(\boldsymbol{\theta})|^{1/2} |\mathbf{Q}_n|^{1/2}}{|\mathbf{Q}_c(\boldsymbol{\theta})|^{1/2}} \exp\left(-\frac{1}{2} \mathbf{x}^\top \mathbf{Q}(\boldsymbol{\theta}) \mathbf{x}\right) \\ &\times \exp\left(-\frac{1}{2} (\mathbf{y} - \mathbf{A}\mathbf{x})^\top \mathbf{Q}_n(\boldsymbol{\theta}) (\mathbf{y} - \mathbf{A}\mathbf{x})\right) \\ &\times \exp\left(\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}_c(\boldsymbol{\theta}))^\top \mathbf{Q}_c(\boldsymbol{\theta}) (\mathbf{x} - \boldsymbol{\mu}_c(\boldsymbol{\theta}))\right). \end{aligned} \quad (54)$$

The quadratic terms in the exponential functions in (54) can be simplified by using $\boldsymbol{\mu}_c(\boldsymbol{\theta})$ and $\mathbf{Q}_c(\boldsymbol{\theta})$. It is also convenient to use the logarithm of the posterior distribution of $\boldsymbol{\theta}$. Reorganize (54) to get the formula which will be used for inference

$$\begin{aligned} \log(\pi(\boldsymbol{\theta}|\mathbf{y})) &= \text{Const} + \log(\pi(\boldsymbol{\theta})) + \frac{1}{2} \log(|\mathbf{Q}(\boldsymbol{\theta})|) \\ &\quad - \frac{1}{2} \log(|\mathbf{Q}_c(\boldsymbol{\theta})|) + \frac{1}{2} \boldsymbol{\mu}_c(\boldsymbol{\theta})^\top \mathbf{Q}_c(\boldsymbol{\theta}) \boldsymbol{\mu}_c(\boldsymbol{\theta}). \end{aligned} \quad (55)$$

From (55) we can see that it is difficult to handle the posterior distribution of $\boldsymbol{\theta}$ analytically since both the determinants and the quadratic terms are hard to handle. Thus numerical methods should be applied for the statistical inference in this paper.

Furthermore, even though it is not the topic for our paper, we point out that it is also possible to obtain the probability density function $\pi(\mathbf{x}|\mathbf{y})$ by integrate out $\boldsymbol{\theta}$. But this is difficult and needs to be obtained numerically using the following expression (56)

$$\pi(\mathbf{x}|\mathbf{y}) = \int_{\mathbb{R}^m} \pi(\mathbf{x}|\boldsymbol{\theta}, \mathbf{y}) \pi(\boldsymbol{\theta}|\mathbf{y}) d\boldsymbol{\theta} \quad (56)$$

where m is the dimension of $\boldsymbol{\theta}$. When the dimension of $\boldsymbol{\theta}$ is large, then this integration might be infeasible in practice.

4.1 Statistical inference with simulated data

First, we illustrate how to do the statistical inference for simulated data with the SPDEs approach with known true parameter values. These datasets both contain one realization with 2000 observations at different locations. The parameters used for generating the simulated data are presented in Table 4. To make it simpler, the nugget effects are assumed to be known, $\tau_1 = \tau_2 = 0.001$. As discussed in Section 1 it is hard to estimate the smoothness parameters, so we fix $\{\alpha_{ij}; i, j = 1, 2\}$ in the system of SPDEs to the known values. The smoothness parameters α_{n_1} and α_{n_2} for the noise processes are also fixed to the known values due to the same reason. The scaling parameters κ_{n_1} and κ_{n_2} for the noise processes are restricted with $\kappa_{n_1} = \kappa_{11}$ and $\kappa_{n_2} = \kappa_{22}$ for the simulated data. Thus in the simulated data examples, only $\boldsymbol{\theta} = \{\kappa_{11}, \kappa_{21}, \kappa_{22}, b_{11}, b_{21}, b_{22}\}$ needed to be estimated. Since κ_{11}, κ_{21} and κ_{22} have to take positive values, we a priori assign *log-normal* distributions with mean zero and large variances for each of the parameters. b_{11}, b_{21} and b_{22} are given *normal* priors with mean zero and large variances.

In the first example, the two GRFs are negatively correlated and the realizations are shown in Figure 11(a) and Figure 11(b). The corresponding estimated conditional mean for the negatively correlated GRFs are given in Figure 11(c) and Figure 11(d). We can notice that there are no large differences between the estimated conditional mean for the GRFs and the true bivariate GRFs. The estimates for the parameters are given in Table 5 with their standard derivations. From this table, we notice that the estimates for all the parameters are quite accurate with accuracy to 2 digits. The true values of the parameters are within 1 standard deviation from the estimates.

The second example with the simulated data uses two GRFs that are positively correlated, and the realizations are shown in Figure 12(a) and Figure 12(b). The corresponding estimated conditional mean for the bivariate GRFs are illustrated in Figure 12(c) and Figure 12(d). From Figure 12(a) to Figure 12(d), we can again notice that the estimated conditional mean for the positively correlated bivariate GRFs are almost the same as the true bivariate GRFs. The estimates are accurate and the true values of the parameters are within 1 standard deviation from the estimates.

The third example uses a trivariate GRF. We reuse the parameters given in Table 7 as true parameters for the simulated data. One realization from the trivariate GRF is given in Figure 8 in Section 3.4. The nugget effects are assumed to be known as $\tau_1 = \tau_2 = \tau_3 = 0.01$. In this example, we also fix $\{\alpha_{ij}; i, j = 1, 2\}$ and $\{\alpha_{n_i}, i = 1, 2, 3\}$. The scaling parameters for the noise processes are fixed with similar settings as the bivariate GRFs $\{\kappa_{n_i} = \kappa_{ii}; i = 1, 2, 3\}$. Therefore we have 12 parameters to be estimated. They are b_{ij} and $\{\kappa_{ij}; i, j = 1, 2, 3, i \geq j\}$. The estimates for these parameters are given in Table 7. We can notice that the estimates are accurate in this example since all the estimated are accurate to 2 digits, and the true values of the parameters are within 2 standard deviations from the corresponding estimates. The estimated conditional mean for the trivariate GRF are given in Figure 13(a) - Figure 13(c). Comparing the results given in Figure 13 with the corresponding fields given in Figure 8, we can see that there are no large differences between the true random fields and the estimated conditional mean. From the examples we can notice that our approach works well not only for

Table 4: Parameters for simulating the bivariate GRFs

dataset 1			dataset 2		
α	κ	b	α	κ	b
$\alpha_{11} = 2$	$\kappa_{11} = 0.3$	$b_{11} = 1$	$\alpha_{11} = 2$	$\kappa_{11} = 0.15$	$b_{11} = 1$
$\alpha_{12} = 0$	$\kappa_{12} = 0$	$b_{12} = 0$	$\alpha_{12} = 0$	$\kappa_{12} = 0$	$b_{12} = 0$
$\alpha_{21} = 2$	$\kappa_{21} = 0.5$	$b_{21} = 1$	$\alpha_{21} = 2$	$\kappa_{21} = 0.5$	$b_{21} = -1$
$\alpha_{22} = 2$	$\kappa_{22} = 0.4$	$b_{22} = 1$	$\alpha_{22} = 2$	$\kappa_{22} = 0.3$	$b_{22} = 1$
$\alpha_{n_1} = 1$	$\kappa_{n_1} = 0.3$		$\alpha_{n_1} = 0$	$\kappa_{n_1} = 0.15$	
$\alpha_{n_1} = 0$	$\kappa_{n_2} = 0.4$		$\alpha_{n_2} = 0$	$\kappa_{n_2} = 0.3$	

Table 5: Inference with simulated dataset 1

Parameters	True value	Estimated	Standard deviations
κ_{11}	0.3	0.295	0.019
κ_{21}	0.5	0.471	0.044
κ_{22}	0.4	0.380	0.020
b_{11}	1	1.009	0.069
b_{21}	1	1.032	0.064
b_{22}	1	0.997	0.059

bivariate GRFs but also for multivariate GMRFs.

4.2 Inference with real data

We illustrate how to use the SPDEs approach for multivariate data in real application in this section. The meteorological dataset used by Gneiting et al. (2010) is analyzed in this paper. This meteorological dataset contains one realization consisting of 157 observations of both temperature and pressure in the north American Pacific Northwest, and the temperature and pressure are always observed at the same locations. It contains the following data: pressure errors (in Pascal) and temperature errors (in Kelvin), measured against longitude and latitude. The errors are defined as the forecasts minus the observations. The data are valid on 18th, December of 2003 at 16:00 local time, at a forecast horizon of 48 hours. For information about the data, see, for example, Eckel and Mass

Table 6: Inference with simulated dataset 2

Parameters	True value	Estimated	Standard deviations
κ_{11}	0.15	0.139	0.124
κ_{21}	0.5	0.487	0.059
κ_{22}	0.3	0.293	0.061
b_{11}	1	0.991	0.017
b_{21}	-1	-1.002	0.033
b_{22}	1	1.011	0.018

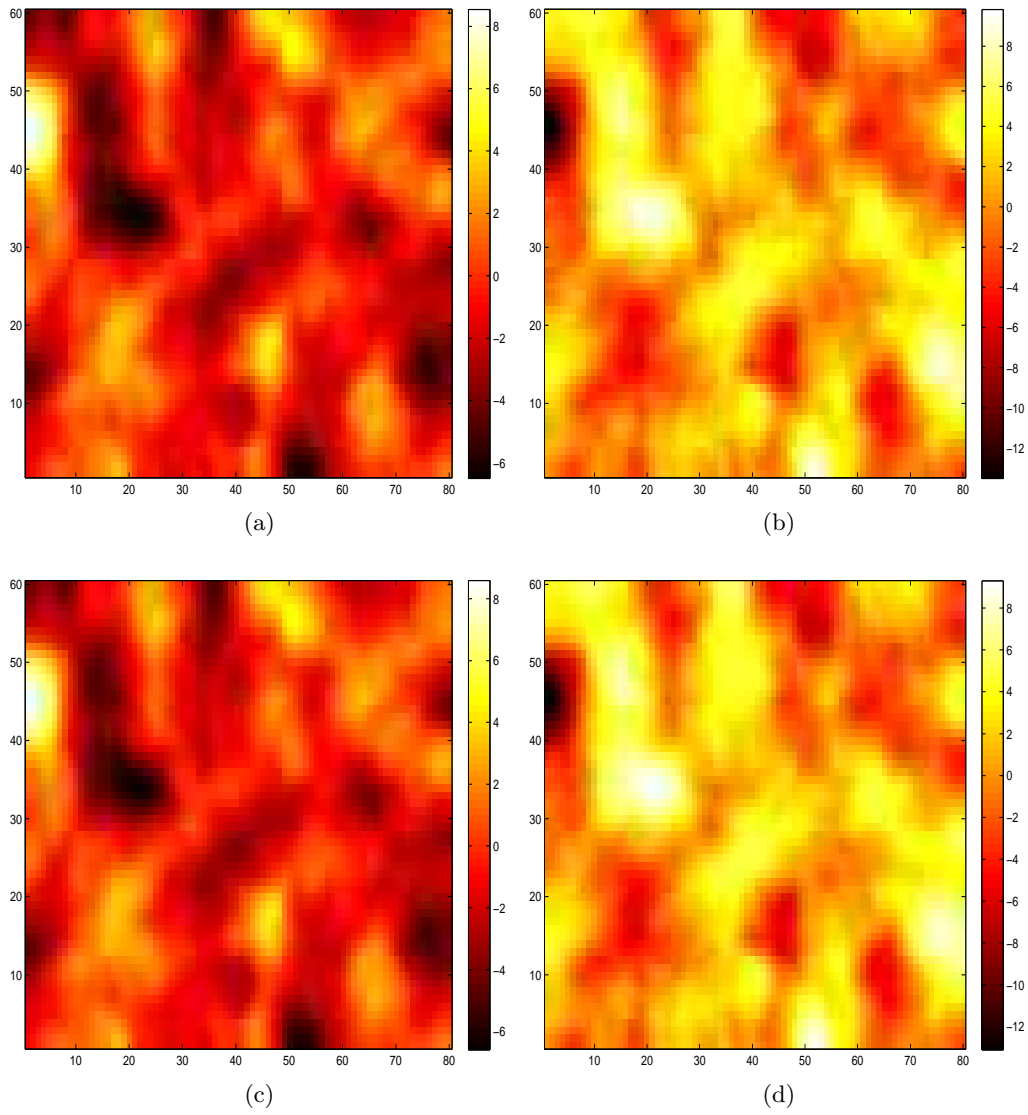


Figure 11: The true bivariate GRFs (a) - (b) and the estimated conditional mean for the bivariate GRFs (c) - (d) for simulated dataset 1.

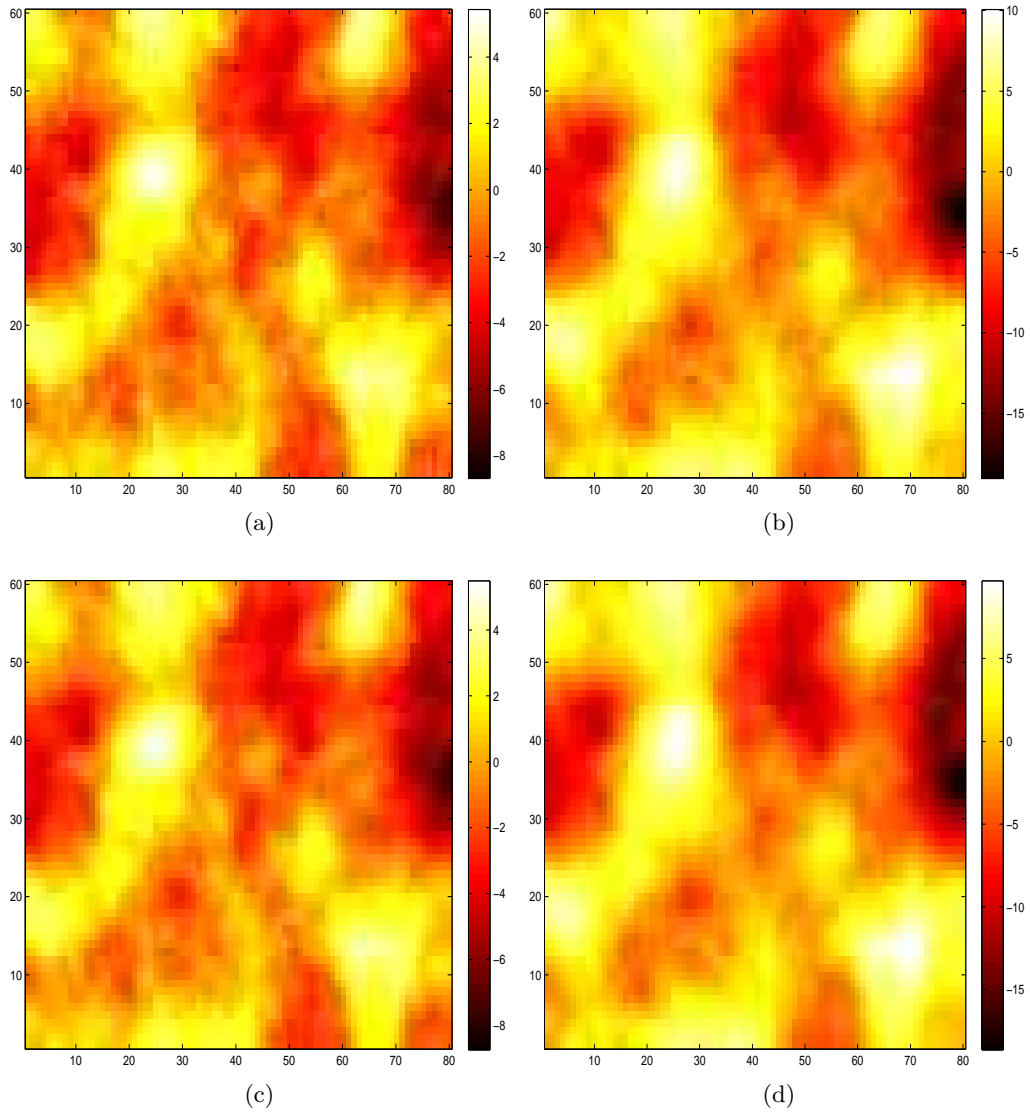


Figure 12: The true bivariate GRFs (a) - (b) and the estimated conditional mean for the bivariate GRFs (c) - (d) for simulated dataset 2.

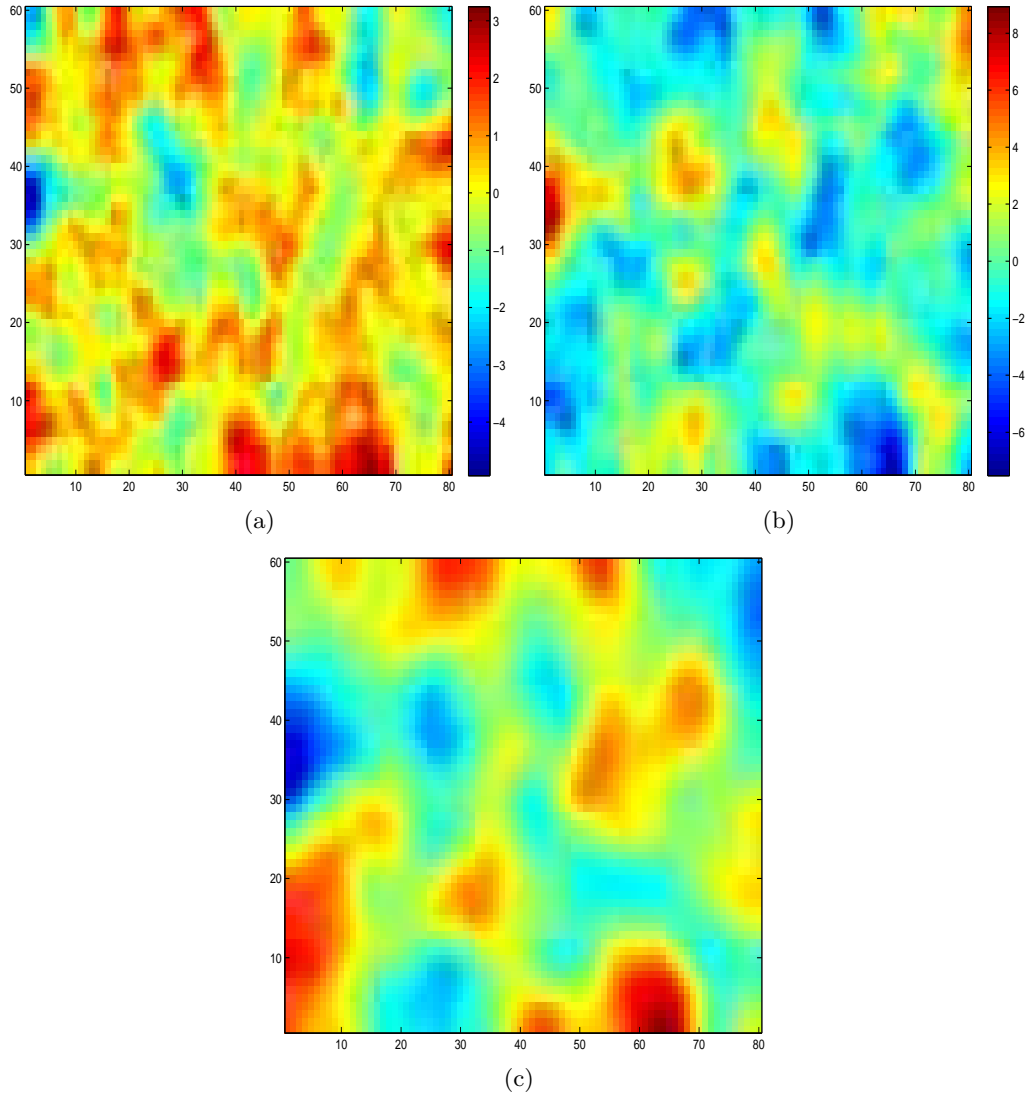


Figure 13: The estimated conditional mean for the trivariate GRFs.

Table 7: Inference for the trivariate GRF

Parameters	True value	Estimated	Standard deviations
b_{11}	1	1.002	0.014
b_{21}	0.8	0.807	0.024
b_{22}	1	0.984	0.013
b_{31}	1	0.958	0.031
b_{32}	0.9	0.878	0.026
b_{33}	1	0.986	0.012
κ_{11}	0.5	0.481	0.022
κ_{21}	0.6	0.582	0.044
κ_{22}	0.4	0.393	0.021
κ_{31}	0.5	0.564	0.050
κ_{32}	1	1.008	0.034
κ_{33}	0.3	0.286	0.012

(2005) and Gneiting et al. (2010). Gneiting et al. (2010) have chosen this dataset with the aim of doing probabilistic weather field forecasting, which is a big research area (Gel et al., 2004; Berrocal et al., 2007; Sloughter et al., 2007; Berrocal et al., 2008). One of the main focuses in this area is to fit and sample spatially correlated error fields. This aim fits our motivation well since the SPDEs approach can be applied to construct multivariate random fields in order to capture the dependence structures not only within the fields but also between the fields.

The main aim of this example is to illustrate how to fit a bivariate random field with the SPDEs approach for pressure and temperature errors data. This bivariate random field can be used to explain the features of the two random fields as in Gneiting et al. (2010). The constructed bivariate spatial random fields are used to represent the error fields for temperature and pressure. It is known that the temperature and the pressure are negatively correlated (Courtier et al., 1998; Ingleby, 2001). As pointed out by Gneiting et al. (2010), the forecast fields are usually smooth fields. The observation field for the pressure is smooth. However, the observation field for temperature is rough. So the model should set up to give the same type of behavior. Without any confusion, we will call the temperature error field the temperature and the pressure error field the pressure. In general, we need to choose the order of the fields at the modelling stage. The simple way is fitting the data with both the orders and selecting the best one using some predefined scoring rules. For more information about the scoring rules, see for example, Gneiting et al. (2005). In this paper we will set the first field $x_1(\mathbf{s})$ as the pressure and the second field $x_2(\mathbf{s})$ as the temperature.

The bivariate random field with our approach is constructed with a system of SPDEs,

$$\begin{aligned}
(\kappa_{n_1}^2 - \Delta)^{\alpha_{n_1}/2} f_1(\mathbf{s}) &= \mathcal{W}_1(\mathbf{s}), \\
(\kappa_{n_2}^2 - \Delta)^{\alpha_{n_2}/2} f_2(\mathbf{s}) &= \mathcal{W}_2(\mathbf{s}), \\
b_{11}(\kappa_{11}^2 - \Delta)^{\alpha_{11}/2} x_1(\mathbf{s}) &= f_1(\mathbf{s}), \\
b_{22}(\kappa_{22}^2 - \Delta)^{\alpha_{22}/2} x_2(\mathbf{s}) + b_{21}(\kappa_{21}^2 - \Delta)^{\alpha_{21}/2} x_1(\mathbf{s}) &= f_2(\mathbf{s}).
\end{aligned} \tag{57}$$

Since the main purpose of this example is to illustrate that we can construct the same (or similar) model as the covariance-based approach presented by Gneiting et al. (2010), only the models which can result in similar models as theirs have been chosen from the parameter matching results given in Section 3.3. We assume that $\kappa_{n_1} = \kappa_{11}$ and $\kappa_{n_2} = \kappa_{22}$. This particular setting corresponds to the first random field $x_1(\mathbf{s})$ being a Matérn random field and the second random field $x_2(\mathbf{s})$ being close to a Matérn random field. These settings make the model constructed through our SPDEs approach closer to the Gneiting et al. (2010) approach.

The results for the estimates with the SPDEs approach are given in Table 8. From the table we can notice that we can capture the negative dependence structure between temperature and pressure since $b_{21} > 0$. The estimated co-located correlation coefficient $\rho_{pt} = -0.53$ which is quite similar as the result from Gneiting et al. (2010). The standard deviations for pressure and temperature are $\sigma_P = 202.1$ Pascal and $\sigma_T = 2.76$ Celsius degrees. In order to compare the predictive perform between our approach and approach proposed by Gneiting et al. (2010), we randomly leave out 25 observations from each field and use only 132 observations for parameters estimation. The relative error \mathcal{E} has been chosen to compare the predictions and is defined as

$$\mathcal{E} = \frac{\|\hat{\mathbf{y}} - \mathbf{y}\|_2}{\|\hat{\mathbf{y}}\|_2}.$$

where $\|\cdot\|_2$ denotes the 2-norm. $\hat{\mathbf{y}}$ denotes the vector of predictions for the observations \mathbf{y} . The predictive performances for these two approaches are given in Figure 14. We can notice that the results from our SPDEs approach and the approach in Gneiting et al. (2010) are quite similar. Table 9 shows the corresponding predictive errors for these two approaches. From this table, we can notice that our model gives slightly better results than the covariance-based model but the difference is not statistically significant. Another merit with our approach which has not been discussed until now is that the SPDEs approach in general is much more computationally efficient since the precision matrix \mathbf{Q} is sparse. The conditional mean in 3D and 2D for the bivariate GRF are shown in Figure 15 and Figure 16 for our approach, respectively. The corresponding results for the covariance-based approach are shown in Figure 17 and Figure 18. As we can see from the 3D and 2D figures, the bivariate GRFs from our approach and from the covariance-based approach give quite similar results.

For the nugget effects (measurement error variances) of pressure and temperature $\boldsymbol{\tau} = (\tau_1, \tau_2)^T$, we are going to use an iterative bias correction to estimate them. This is based on formula

$$\sigma_{ij}^2 = \tau_i^2 + V_{ij}, \quad i = 1, 2, \quad j = 1, \dots, n, \tag{58}$$

Table 8: Estimated parameters for the SPDEs approach

κ_{11}	κ_{21}	κ_{22}	b_{11}	b_{21}	b_{22}
6.353×10^{-3}	0.413	2.243×10^{-2}	0.2165	1.298×10^{-4}	5.458

Table 9: Predictive relative errors for the SPDEs approach and the covariance-based models

relative errors		
Models	random field 1	random field 2
SPDEs approach	0.777	0.690
Covariance-based model	0.821	0.716

where i indicates the pressure with $i = 1$ and the temperature with $i = 2$. $\sigma_{ij}^2 = \text{Var}(y_{ij} - \hat{y}_{ij})$ is the variance which contains the nugget effects and the kriging variances. $\{\tau_i^2 = \text{Var}(y_{ij} - x_{ij}); i = 1, 2, j = 1, \dots, n\}$ are the nugget effects for the pressure when $i = 1$ or temperature when $i = 2$. $\{V_{ij} = \text{Var}(x_{ij} - \hat{y}_{ij}); i = 1, 2, j = 1, \dots, n\}$ are the kriging variances and they are from the model bias. n denotes the total number of data points in each field. y_{ij} denotes the observed value at data point for each field. \hat{y}_{ij} is the predicted values from a given model and x_{ij} denotes the true value which is unknown. See Appendix for a simple proof of (58).

One simple way is to use Equation (59) where we just subtract the kriging variance V_{ij} from the empirical variance $\hat{\sigma}_{ij}^2$ to get the estimate of nugget effect $\hat{\tau}_i^2$ for each field

$$\hat{\tau}_i^2 = \frac{1}{n} \sum_j (\hat{\sigma}_{ij}^2 - V_{ij}), \quad (59)$$

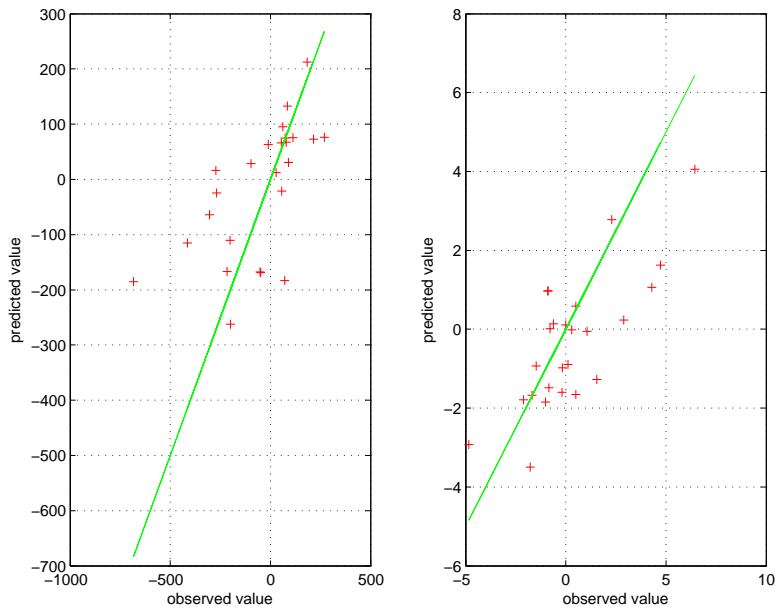
where $\hat{\sigma}_{ij}^2 = (y_{ij} - \hat{y}_{ij})^2$. But another preferable way which we have used in this paper is with the formula

$$\hat{\tau}_i^2 = \sum_j \{w_{ij} (\hat{\sigma}_{ij}^2 - V_{ij})\}, \quad (60)$$

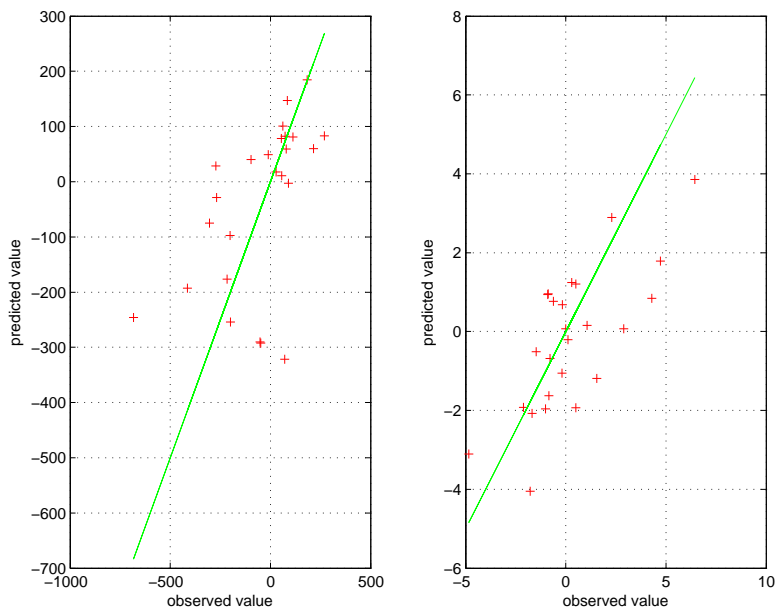
where $w_{ij} = \frac{1/(\tau_i^2 + V_{ij})}{\sum_j (1/(\tau_i^2 + V_{ij}))}$. This can give an unbiased estimator of nugget effects. See Appendix for the proof. The initial values are chosen to be similar to the results given in Gneiting et al. (2010), and the results from the bias correction approach are shown in Table 10. The convergence curves of the nugget effects for the fields are illustrated in Figure 19.

Table 10: The nugget effects for pressure and temperature

	initial values	estimated values
τ_1	68.4	81.1
τ_2	0.01	0.53



(a)



(b)

Figure 14: The predictive performances of SPDEs approach (a) and the covariance-based approach (b)

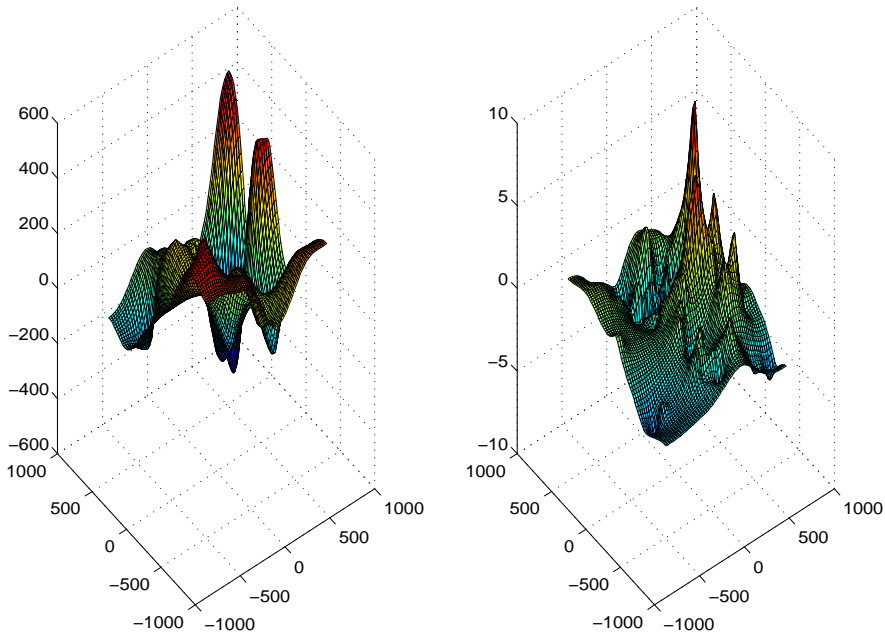


Figure 15: Estimated conditional mean for bivariate GRFs for SPDEs approach in 3D.

5 Discussion

Spatial modelling of multivariate data are of demand in many areas, such as weather forecasting (Courtier et al., 1998; Reich and Fuentes, 2007), air quality (Brown et al., 1994; Schmidt and Gelfand, 2003), economics (Gelfand et al., 2004; Sain and Cressie, 2007). The important issue for modelling spatial multivariate data is that the approach can not only model the marginal covariances within each field, but also has the ability to model the cross-covariances between the random fields. In addition, we need to solve the theoretical challenge for the positive definite constraint of the covariance functions and the computational challenge for large dataset.

The main aim of this work is to illustrated the possibility of constructing *multivariate* GRFs through the SPDEs approach. We notice that the parameters in the SPDEs approach is interpretable and can link our approach to the covariance-based approach. By using the approximate weak solution of the corresponding system of SPDEs, we can represent multivariate GRFs by GMRFs. Since the precision matrices of the GMRFs are sparse, numerical algorithms for sparse matrices can be applied, and therefore fast sampling and inference are feasible. Our approach inherited the properties from the approach discussed by Lindgren et al. (2011). There are three main advantages for our newly proposed SPDEs approach. The first advantage is that our new SPDEs approach does not depend on direct construction of positive definite matrix. The notorious requirement of positive definite covariance matrix is fulfilled automatically. The second advantage is that we can remove the symmetry property constraint shared by both the

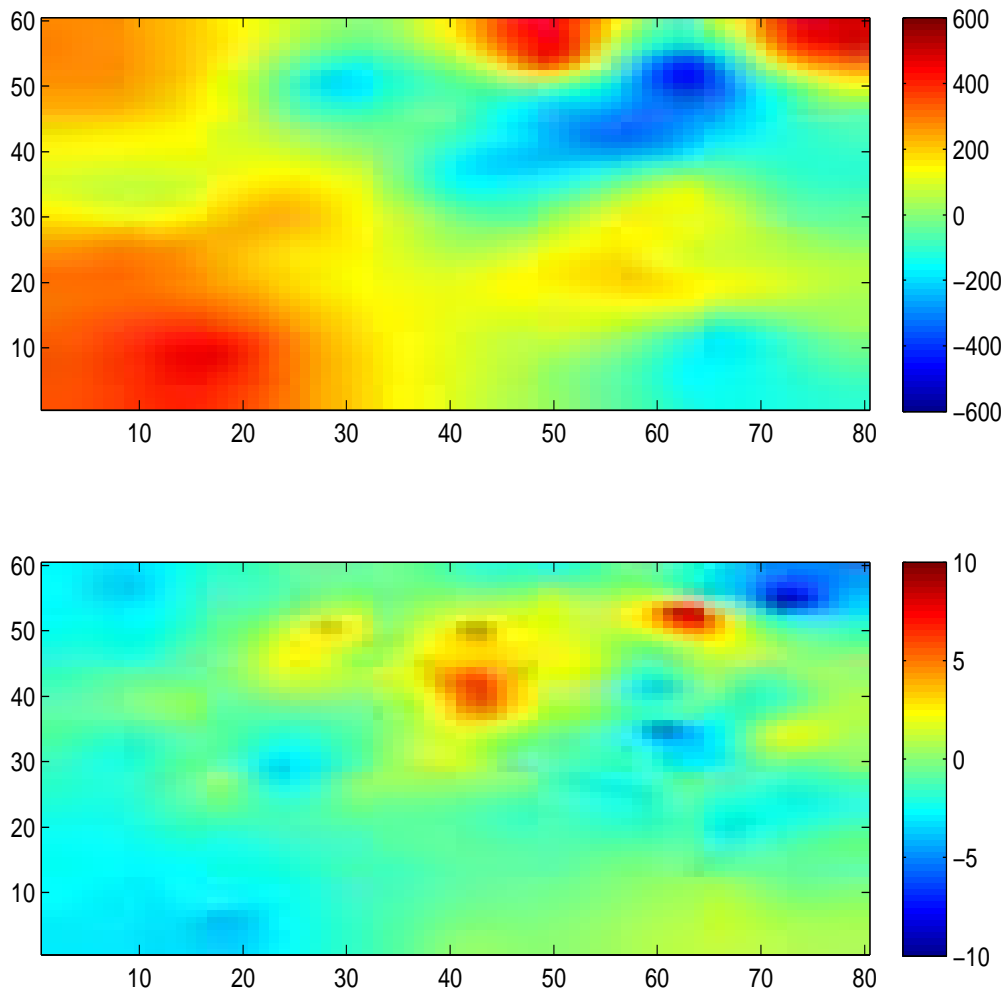


Figure 16: Estimated conditional mean for bivariate GRFs for SPDEs approach in 2D.

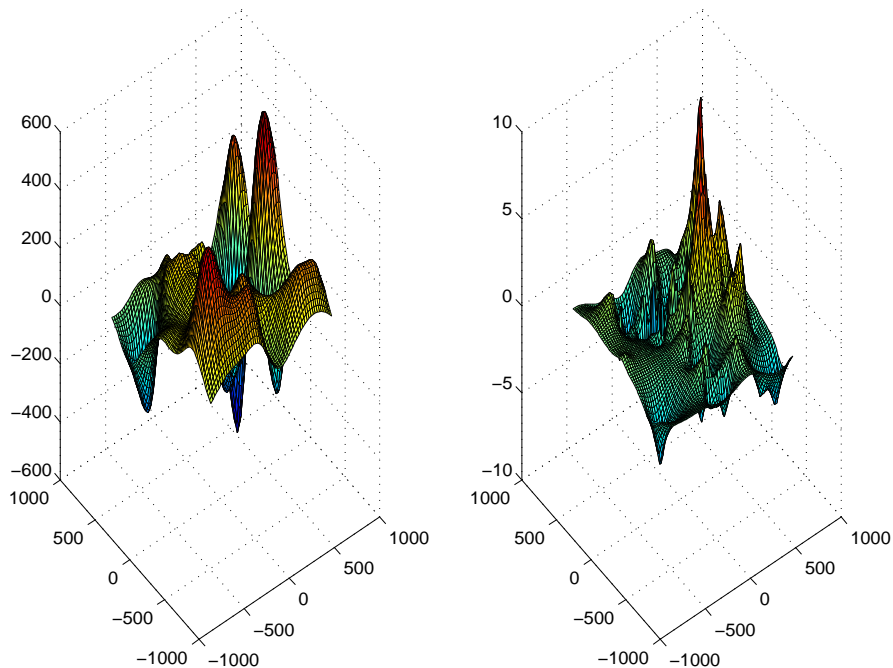


Figure 17: Estimated conditional mean for bivariate GRFs for covariance-based model in 3D.

covariance-based approach Gneiting et al. (2010) and the LMC approach (Gelfand et al., 2004; Gneiting et al., 2010). The third advantage is that we can construct the multivariate GRFs on manifolds, such as on the sphere \mathbb{S}^2 . The extension follows the discussion given by Lindgren et al. (2011). We just need to reinterpret the systems of SPDEs to be defined on the manifold.

One issue that needs to be pointed out is that we have chosen $\kappa_{n_1} = \kappa_{11}$ and $\kappa_{n_2} = \kappa_{22}$ in the model. This restriction used in this paper is to make the model closer to the models presented by Gneiting et al. (2010) and also make the inference easier. This may not be needed in other applications and κ_{n_1} and κ_{n_2} can be estimated from the data. However, κ_{n_1} and κ_{11} might be not distinguishable with the triangular systems of SPDEs. So one suggestion is that if we are use the triangular systems of SPDEs, then we fix κ_{n_1} based on some other information or set $\kappa_{n_1} = \kappa_{11}$ when doing inference.

Another issue that needed to be pointed out is that we have prosecuted the *full* version of the SPDEs approach but have not used in all the examples. This version could give us more flexibility to construct multivariate GRFs. The modelling procedures and the inference are the same as for the *triangular* version of the SPDEs which were used and discussed extensively in this paper.

We also want to point out that the proposed approach costs effort during the implementation and pre-processing stages since we need to build the system of SPDEs, discretize the fields and do the approximation to obtain a GMRF representation. But we believe, as pointed out by Lindgren et al. (2011), that “*such costs are unavoidable*”

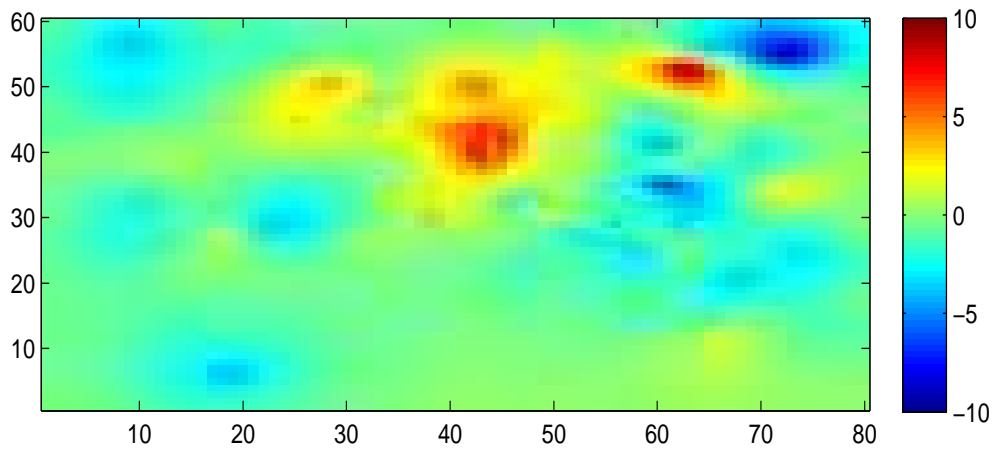
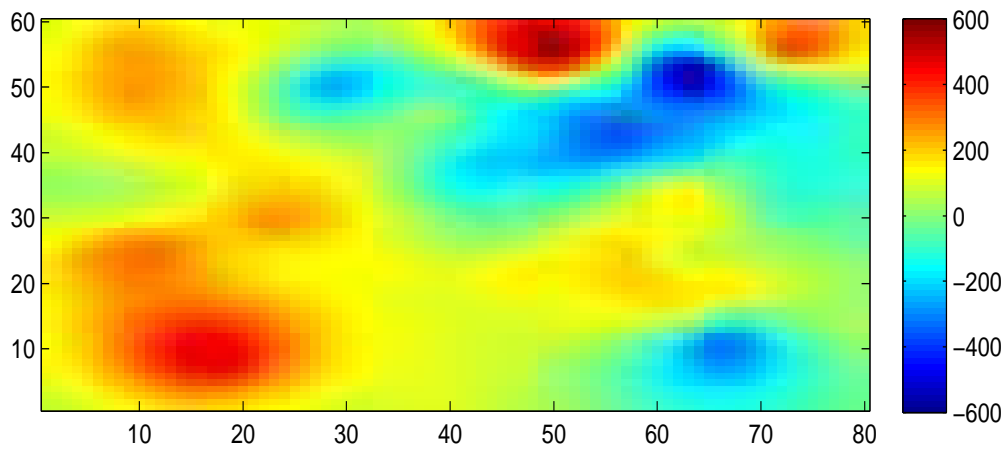


Figure 18: Estimated conditional mean for bivariate GRFs for covariance-based model in 2D.

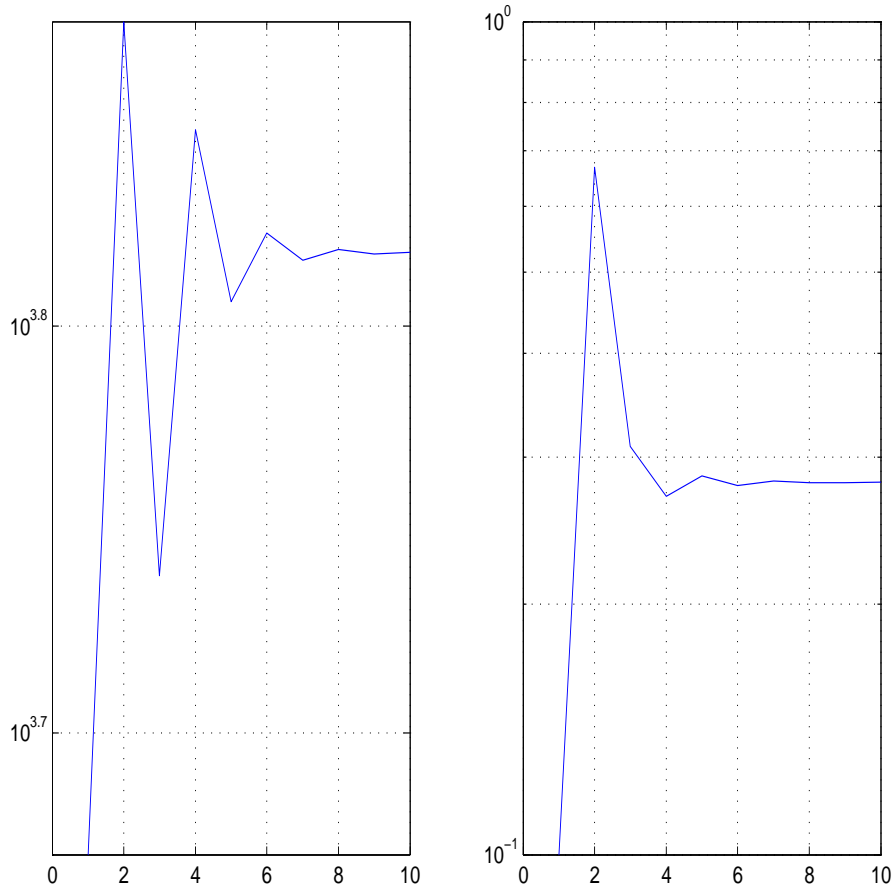


Figure 19: Bias correction iteration. The curves are for τ_1^2 and τ_2^2 respectively

when efficient computations are required”.

Similarly, as pointed out by Lindgren et al. (2011), although the SPDEs approach presented here is not generally applicable for all covariance functions, it covers many useful model in spatial statistics. And it is possible to extend our approach to construct even richer class of models by following the discussion given by, for instance, Bolin and Lindgren (2011) and Fuglstad (2011). Our approach extends the applicability of multivariate GRFs in practical applications since we can build and interpret the model using GRFs but do computation with GMRFs. It is further possible to include this approach in the integrated nested Laplace approximation (INLA) framework (Rue et al., 2009). These extensions are under study.

Appendix

In order to prove Equation (58), we need to write down the expression explicitly. Let y_{ij} denote the observed value for each field, \hat{y}_{ij} the predicated value from a given model, and x_{ij} the true value. Then we have

$$\begin{aligned}\text{Var}(y_{ij} - \hat{y}_{ij}) &= \text{Var}((y_{ij} - x_{ij}) + (x_{ij} - \hat{y}_{ij})) \\ &= \text{Var}(y_{ij} - x_{ij}) + \text{Var}(x_{ij} - \hat{y}_{ij}) \\ &\quad + 2\text{Cov}(y_{ij} - x_{ij}, x_{ij} - \hat{y}_{ij}) \\ &= \text{Var}(y_{ij} - x_{ij}) + \text{Var}(x_{ij} - \hat{y}_{ij}),\end{aligned}\tag{61}$$

since

$$\begin{aligned}\text{Cov}(y_{ij} - x_{ij}, x_{ij} - \hat{y}_{ij}) &= \mathbb{E}((y_{ij} - x_{ij})(x_{ij} - \hat{y}_{ij})) \\ &= \mathbb{E}[\mathbb{E}((y_{ij} - x_{ij})(x_{ij} - \hat{y}_{ij})|\mathbf{y})] \\ &= \mathbb{E}[(y_{ij} - \hat{y}_{ij}) \times 0] \\ &= 0.\end{aligned}\tag{62}$$

So Equation (58) is established.

Now we show that with the weights $w_{ij} = \frac{1/(\tau_i^2 + V_{ij})}{\sum_j (\tau_i^2 + V_{ij})}$, the estimator in Equation (60) is an unbiased estimator for the nugget effects. Using Equation (61), we know that $(y_{ij} - \hat{y}_{ij}) \sim \mathcal{N}(0, \tau_i^2 + V_{ij})$. Then

$$\begin{aligned}\text{Var}((y_{ij} - \hat{y}_{ij})^2) &= \mathbb{E}((y_{ij} - \hat{y}_{ij})^4) - (\mathbb{E}((y_{ij} - \hat{y}_{ij})^2))^2 \\ &= 3(\tau_i^2 + V_{ij})^2 - (\tau_i^2 + V_{ij})^2 \\ &= 2(\tau_i^2 + V_{ij})^2\end{aligned}\tag{63}$$

and thus the estimator is unbiased since the scaling constant cancels in the weight coefficient.

References

- R.J. Adler and J.E. Taylor. *Random fields and geometry*. Springer Verlag, 2007. ISBN 0387481125.
- D.J. Allcroft and C.A. Glasbey. A latent Gaussian Markov random-field model for spatiotemporal rainfall disaggregation. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 52(4):487–498, 2003. ISSN 1467-9876.
- S. Banerjee, B.P. Carlin, and A.E. Gelfand. *Hierarchical modeling and analysis for spatial data*. Chapman & Hall, 2004. ISBN 158488410X.
- K.J. Bathe. *Finite element method*. Wiley Online Library, 2008.
- V.J. Berrocal, A.E. Raftery, and T. Gneiting. Combining spatial statistical and ensemble information in probabilistic weather forecasts. *Monthly Weather Review*, 135(4):1386–1402, 2007.
- V.J. Berrocal, A.E. Raftery, and T. Gneiting. Probabilistic quantitative precipitation field forecasting using a two-stage spatial model. *The Annals of Applied Statistics*, 2(4):1170–1193, 2008.
- D. Bolin and F. Lindgren. Wavelet markov models as efficient alternatives to tapering and convolution fields. Technical report, Mathematical Statistics, Centre for Mathematical Sciences, Faculty of Engineering, Lund University, 2009.
- D. Bolin and F. Lindgren. Spatial models generated by nested stochastic partial differential equations, with an application to global ozone mapping. *The Annals of Applied Statistics*, 5(1):523–550, 2011.
- P.J. Brown, N.D. Le, and J.V. Zidek. Multivariate spatial interpolation and exposure to air pollutants. *Canadian Journal of Statistics*, 22(4):489–509, 1994.
- P. Courtier, E. Andersson, W. Heckley, D. Vasiljevic, M. Hamrud, A. Hollingsworth, F. Rabier, M. Fisher, and J. Pailleux. The ecmwf implementation of three-dimensional variational assimilation (3d-var). i: Formulation. *Quarterly Journal of the Royal Meteorological Society*, 124(550):1783–1807, 1998.
- N.A.C. Cressie. *Statistics for spatial data*, volume 298. Wiley-Interscience, 1993.
- P.J. Diggle and P.J. Ribeiro Jr. *Model-based Geostatistics*. Springer, 2006.
- P.J. Diggle, JA Tawn, and RA Moyeed. Model-based geostatistics. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 47(3):299–350, 1998. ISSN 1467-9876.
- F.A. Eckel and C.F. Mass. Aspects of effective mesoscale, short-range ensemble forecasting. *Weather and Forecasting*, 20(3):328–350, 2005.
- G.A. Fuglstad. Spatial modelling and inference with spde-based gmrf. Master’s thesis, Department of Mathematical Sciences, Norwegian University of Science and Technology, 2011.
- Y. Gel, A.E. Raftery, and T. Gneiting. Calibrated probabilistic mesoscale weather field forecasting. *Journal of the American Statistical Association*, 99(467):575–583, 2004.
- A.E. Gelfand, A.M. Schmidt, S. Banerjee, and CF Sirmans. Nonstationary multivariate process modeling through spatially varying coregionalization. *Test*, 13(2):263–312, 2004.
- T. Gneiting, A.E. Raftery, A.H. Westveld III, and T. Goldman. Calibrated probabilistic forecasting using ensemble model output statistics and minimum crps estimation. *Monthly Weather Review*, 133(5):1098–1118, 2005.
- T. Gneiting, W. Kleiber, and M. Schlather. Matérn Cross-Covariance Functions for Multivariate Random Fields. *Journal of the American Statistical Association*, 105(491):1167–1177, 2010. ISSN 0162-1459.

- J.A. Goff and T.H. Jordan. Stochastic modeling of seafloor morphology: Inversion of sea beam data for second-order statistics. *Journal of Geophysical Research*, 93(B11):13589–13, 1988. ISSN 0148-0227.
- M. Goulard and M. Voltz. Linear coregionalization model: tools for estimation and choice of cross-variogram matrix. *Mathematical Geology*, 24(3):269–286, 1992. ISSN 0882-8121.
- P. Guttorp and T. Gneiting. Studies in the history of probability and statistics XLIX On the Matern correlation family. *Biometrika*, 93(4):989, 2006. ISSN 0006-3444.
- M.S. Handcock and M.L. Stein. A Bayesian analysis of kriging. *Technometrics*, 35(4):403–410, 1993. ISSN 0040-1706.
- L. Hartman and O. Hössjer. Fast kriging of large data sets with Gaussian Markov random fields. *Computational Statistics & Data Analysis*, 52(5):2331–2349, 2008. ISSN 0167-9473.
- Ø. Hjellev and M. Dæhlen. *Triangulations and applications*. Springer Verlag, 2006.
- N.B. Ingleby. The statistical structure of forecast errors and its representation in the met. office global 3-d variational data assimilation scheme. *Quarterly Journal of the Royal Meteorological Society*, 127(571):209–231, 2001.
- P.E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*. Springer, 3rd edition, 1999. ISBN 3540540628.
- F. Lindgren, H. Rue, and J. Lindström. 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 (Statistical Methodology)*, 73(4):423–498, 2011.
- B. Matérn. *Spatial variation*. Springer-Verlag Berlin, 1986.
- B.J. Reich and M. Fuentes. A multivariate semiparametric bayesian spatial modeling framework for hurricane surface wind fields. *The Annals of Applied Statistics*, 1(1):249–264, 2007.
- P.J. Ribeiro Jr and P.J. Diggle. *MODEL BASED GEOSTATISTICS*. Springer Series in Statistics. Springer, 2006.
- H. Rue. Fast sampling of Gaussian Markov random fields. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2):325–338, 2001. ISSN 1467-9868.
- H. Rue and L. Held. *Gaussian Markov random fields: theory and applications*. Chapman & Hall, 2005. ISBN 1584884320.
- H. Rue and H. Tjelmeland. Fitting Gaussian Markov random fields to Gaussian fields. *Scandinavian Journal of Statistics*, 29(1):31–49, 2002. ISSN 1467-9469.
- H. Rue, S. Martino, and N. Chopin. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71(2):319–392, 2009. ISSN 1467-9868.
- S.R. Sain and N. Cressie. A spatial model for multivariate lattice data. *Journal of Econometrics*, 140(1):226–259, 2007.
- A.M. Schmidt and A.E. Gelfand. A bayesian coregionalization approach for multivariate pollutant data. *Journal of Geophysical Research*, 108(D24):8783, 2003.
- J.M.L. Slaughter, A.E. Raftery, T. Gneiting, and C. Fraley. Probabilistic quantitative precipitation forecasting using bayesian model averaging. *Monthly Weather Review*, 135(9):3209–3220, 2007.
- M.L. Stein. *Interpolation of Spatial Data: some theory for kriging*. Springer Verlag, 1999. ISBN 0387986294.

- H. Wackernagel. *Multivariate geostatistics: an introduction with applications*. Springer Verlag, 2003. ISBN 3540441425.
- P. Whittle. On stationary processes in the plane. *Biometrika*, 41(3-4):434–449, 1954. ISSN 0006-3444.
- P. Whittle. Stochastic processes in several dimensions. *Bull. Int. Statist. Inst.*, 40:974–994, 1963.
- O.C. Zienkiewicz, R.L. Taylor, R.L. Taylor, and JZ Zhu. *The finite element method: its basis and fundamentals*, volume 1. Butterworth-heinemann, 2005.