

NORGES TEKNISK-NATURVITENSKAPELIGE  
UNIVERSITET

**Multivariate Gaussian Random Fields with Oscillating  
Covariance Functions using Systems of Stochastic Partial  
Differential Equations**

by

Xiangping Hu, Finn Lindgren, Daniel Simpson and Håvard Rue

PREPRINT  
STATISTICS NO. 4/2013



NORWEGIAN UNIVERSITY OF SCIENCE AND  
TECHNOLOGY  
TRONDHEIM, NORWAY

This preprint has URL <http://www.math.ntnu.no/preprint/statistics/2013/S4-2013.pdf>

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

E-mail: [xiangph@math.ntnu.no](mailto:xiangph@math.ntnu.no)

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



# Multivariate Gaussian Random Fields with Oscillating Covariance Functions using Systems of Stochastic Partial Differential Equations

Xiangping Hu<sup>\*1</sup>, Finn Lindgren<sup>2</sup>, Daniel Simpson<sup>1</sup>, and Håvard Rue<sup>1</sup>

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

<sup>2</sup>Department of Mathematical Sciences, University of Bath, BA2 7AY, United Kingdom

July 4, 2013

## Abstract

In this paper we propose a new approach for constructing *multivariate* Gaussian random fields (GRFs) with oscillating covariance functions through systems of stochastic partial differential equations (SPDEs). We discuss how to build systems of SPDEs that introduces oscillation characteristics in the covariance functions of the multivariate GRFs. By choosing different parametrization of the equations, some GRFs can be made with oscillating covariance functions but other fields can have Matérn covariance functions or close to Matérn covariance functions. The multivariate GRFs constructed by solving the systems of SPDEs automatically fulfill the hard requirement of nonnegative definiteness for the covariance functions. The approximate weak solutions to the systems of SPDEs are used to represent the multivariate GRFs by multivariate Gaussian *Markov* random fields (GMRFs). Since the multivariate GMRFs have sparse precision matrices (inverse of the covariance matrices), numerical algorithms for sparse matrices can be applied to the precision matrices for sampling and inference. Thus from a computational point of view, the *big-n* problem can be partially solved with these types of models. Another advantage of the method is that the oscillation in the covariance function can be controlled directly by the parameters in the system of SPDEs. We show how to use this proposed approach with simulated data and real data examples.

**Keywords:** Multivariate Gaussian random fields, Oscillating covariance functions, Multivariate Gaussian Markov random fields, Sparse matrix, Stochastic partial differential equations

---

<sup>\*</sup>Corresponding author. Email: [Xiangping.Hu@math.ntnu.no](mailto:Xiangping.Hu@math.ntnu.no)

# 1 Introduction

Statistics for spatial data appeared from hundreds of years ago, but spatial models for this type of data appeared much later (Cressie, 1993). The spatial models have been widely used to model the spatial data in many areas. Gaussian random fields (GRFs) are some of the most commonly used models in spatial statistics. Since the normalizing constant can be computed explicitly, the GRFs are convenient to be used in many applications, such as geo-statistical data, environmental and atmospheric data, longitudinal and survival data (Cressie, 1993; Stein, 1999; Rue and Held, 2005; Gelfand et al., 2010). GRFs also have other good properties, such as the fact that a GRF can be explicitly specified through a mean  $\boldsymbol{\mu}(\cdot)$  and a covariance function  $\mathbf{C}(\cdot, \cdot)$ . Let the coordinates members of  $\mathbb{R}^d$ . If  $(x(\mathbf{s}_1), x(\mathbf{s}_2), \dots, x(\mathbf{s}_n))$  is Gaussian for every selection of points  $(\mathbf{s}_1^T, \mathbf{s}_2^T, \dots, \mathbf{s}_n^T)$  for every  $n \geq 1$ , then we call  $x(\mathbf{s})$  a continuously indexed GRF. The covariance matrix of a collection  $(x(\mathbf{s}_1), x(\mathbf{s}_2), \dots, x(\mathbf{s}_n))$  is given by  $\boldsymbol{\Sigma} = [C(\mathbf{s}_i, \mathbf{s}_j)]$ . However, there is a hard nonnegative definite requirement that must be fulfilled for the function  $\mathbf{C}(\cdot, \cdot)$ . This is one of the main concerns when we build models with GRFs using covariance function based approaches.

A widely used class of covariance functions is the Matérn family which was introduced by Matérn (Matérn, 1986). This family of covariance functions captures the most common form of the empirical behavior of stationary covariance functions, namely that the correlation between the locations  $\mathbf{m}$  and  $\mathbf{n}$  should decrease when the Euclidean distance  $\|\mathbf{h}\| = \|\mathbf{m} - \mathbf{n}\|$  increases (Diggle and Ribeiro Jr, 2006). This family of covariance functions is isotropic and usually written as  $\sigma^2 M(\mathbf{m}, \mathbf{n} | \nu, \kappa)$ , where  $M(\mathbf{m}, \mathbf{n} | \nu, \kappa)$  is the Matérn correlation function between the spatial locations  $\mathbf{m}, \mathbf{n} \in \mathbb{R}^d$ . The Matérn correlation function is a two-parameter family with the form

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

where  $\sigma^2$  is the marginal variance,  $K_\nu$  is the modified Bessel function of second kind and  $\kappa > 0$  is the scaling parameter. The order  $\nu$  is a smoothness parameter and must be positive. The smoothness parameter  $\nu$  is of critical concern in spatial statistics since it defines the differentiability of the sample paths and the Hausdorff dimension (Goff and Jordan, 1988; Handcock and Stein, 1993; Gneiting et al., 2010). It is known that the smoothness parameter  $\nu$  is poorly identifiable from the data and hence it is usually fixed (Diggle and Ribeiro Jr, 2006; Lindgren et al., 2011; Hu et al., 2012a). The Matérn family contains the commonly used models with exponential covariance function  $M(\mathbf{h} | \frac{1}{2}, \kappa) = \exp(-\kappa \|\mathbf{h}\|)$ . The Matérn covariance function is the key factor in the explicit link between the GRFs and GMRFs through the stochastic partial differential equations (SPDEs) discussed by Lindgren et al. (2011). Gneiting et al. (2010) presented one direct approach for constructing the multivariate GRFs by using matrix-valued covariance functions, and all components in the matrix-valued covariance function are Matérn covariance functions. Hu et al. (2012a) discussed the important role of Matérn covariance function for constructing stationary and isotropic multivariate GRFs with systems of SPDEs.

In spatial statistics oscillating models usually deal with ocean waves, and we usually work entirely with their spectra only rarely with covariance functions (Lindgren, 2010). For time series there are plenty of applications with oscillating models in discrete time since the  $p$ -order auto-regressive (AR( $p$ )) processes can result in oscillating models (Wei, 2006). However, it is less common for continuous time models. Lindgren et al. (2011) have discussed an approach for constructing univariate GRFs with oscillating covariance functions. In their approach they have chosen a coupled system of SPDEs to construct two independent random fields with the same precision matrix. Their discussion was focused on the case  $\alpha = 2$ . Since it is important for our approach for constructing multivariate GRFs with oscillating covariance functions, we give an overview of their approach in Section 3.1.

In this paper we focus on the methodology for constructing multivariate GRFs with oscillating covariance functions through systems of SPDEs. This work is an extension of the discussion given by Lindgren et al. (2011) for the univariate case. It is also an extension of the approach discussed by Hu et al. (2012a) to construct larger class of useful models. In our approach the GRFs are constructed by solving systems of stochastic partial differential equations. One of the main advantages of this approach is that we do not need to consider the notorious nonnegative definite requirement for the covariance functions. This requirement is fulfilled automatically because we are working on the processes directly. During the computational stage the GRFs are represented by the Gaussian *Markov* random fields (GMRFs) by GMRF approximations. A GMRF  $x(\mathbf{s})$  is a discretely indexed GRF with some Markov property. The full conditionals  $\{\pi(x_i|\mathbf{x}_{-i}); i = 1, \dots, n\}$  of a GMRF depend only on a set of neighbors of each site  $i$ . Denote the neighbors of the node  $i$  by  $\partial i$ . The Markov property implies that  $Q_{ij} \neq 0$  if and only if  $j \in \partial i \cup i$ , where  $\mathbf{Q}$  denotes the precision matrix of the GMRF. Consistency requirement implies that if  $i \in \partial j$ , then  $j \in \partial i$ . The precision matrix  $\mathbf{Q}$  for the GMRF is sparse which enables us to use numerical algorithms for sparse matrices. Thus the *big-n* problem (Banerjee et al., 2004) can be partially solved in our case with these types of models. We refer to Rue et al. (2009, Section 2.1) for a condensed overview of the theory of GMRFs. Detailed discussions about GMRFs are given in Rue and Held (2005, Chapter 2).

## 2 State-of-the-art and preliminaries

We review the state-of-the-art research on GRFs with SPDE approach in Section 2.1 and the methodologies for constructing multivariate GRFs in Section 2.2. The GMRF approximation for representing a GRF with a GMRF is introduced in Section 2.3 since it is crucial for computations.

## 2.1 GRFs through the SPDE approach

As mentioned in Section 1, Lindgren et al. (2011) proposed a novel approach for constructing GRFs by using the SPDE

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

where  $(\kappa^2 - \Delta)^{\alpha/2}$  is a pseudo-differential operator,  $\kappa$  is the scaling parameter,  $\alpha$  is related to the smoothness parameters  $\nu > 0$  and  $\alpha = \nu + d/2$ .  $\Delta$  is the Laplacian with definition

$$\Delta = \sum_{i=1}^d \frac{\partial^2}{\partial x_i^2}.$$

A range parameter  $\rho$  connects the scaling parameter  $\kappa$  and the smoothness parameter  $\nu$ . The simple and empirically derived relationship  $\rho = \sqrt{8\nu}/\kappa$  is commonly used (Lindgren et al., 2011; Hu et al., 2012a). It corresponds to correlation near 0.1 at distance  $\rho$ , with parameters  $\kappa$  and  $\nu$ .  $\mathcal{W}(\mathbf{s})$  is the innovation process which is a standard spatial Gaussian white noise. The most important result given by Whittle (1954, 1963), and used by Lindgren et al. (2011) extensively, is that the solution  $x(\mathbf{s})$  to SPDE (2) is a GRF with the Matérn covariance function given in Equation (1). We follow the terminology used by Lindgren et al. (2011) and call GRFs with Matérn covariance functions Matérn random fields. Lindgren et al. (2011) commented that their approach can be extended in many directions.

Fuglstad (2011) has extended the SPDE approach to include a diffusion matrix in Equation (2) which provides a way of controlling the covariance structures of the GRF. Using the diffusion matrix  $\mathbf{H}$  in SPDE (2), it is not only possible to construct homogeneous isotropic fields, but also anisotropic fields. Fuglstad (2011) showed that it is possible to construct inhomogeneous fields. The SPDE discussed by Fuglstad (2011) has the form

$$\kappa^2 x(\mathbf{s}) - \nabla \cdot \mathbf{H}(\mathbf{s}) \nabla x(\mathbf{s}) = \mathcal{W}(\mathbf{s}), \quad (3)$$

where  $\mathbf{H}$  is a  $2 \times 2$  matrix-valued function,  $\nabla$  is the gradient operator and  $\mathcal{W}(\mathbf{s})$  is a standard spatial Gaussian white noise process. The main contribution of Fuglstad (2011) is that he introduced the matrix  $\mathbf{H}$  to (2) to control the structure of covariance matrix. However, he focused on discussing the univariate GRFs with  $\alpha = 2$ .

Bolin and Lindgren (2011) discussed how to use nested SPDEs to construct stationary and non-stationary GRFs. The equation chosen by Bolin and Lindgren (2011) has the form

$$\mathcal{L}_1 x(\mathbf{s}) = \mathcal{L}_2 \mathcal{W}(\mathbf{s}) \quad (4)$$

for some linear differential operators  $\mathcal{L}_1$  and  $\mathcal{L}_2$ .  $\mathcal{W}(\mathbf{s})$  is a standard spatial Gaussian white noise process. The SPDE (4) may not exist in the common sense since the operator  $\mathcal{L}_2$  may contain some differentiation of the noise process  $\mathcal{W}(\mathbf{s})$ . In this case SPDE (4) can then be interpreted as the following nested system of SPDEs

$$\begin{aligned} \mathcal{L}_1 x_0(\mathbf{s}) &= \mathcal{W}(\mathbf{s}), \\ x(\mathbf{s}) &= \mathcal{L}_2 x_0(\mathbf{s}), \end{aligned} \quad (5)$$

when  $\mathcal{L}_1$  and  $\mathcal{L}_2$  are commutative operators. As pointed out by Bolin and Lindgren (2011) this interpretation not only avoids the apparent problem with the differentiation of the Gaussian white noise process, but also gives an interpretation of the consequence of the additional differential operator  $\mathcal{L}_2$ . The random field  $x(\mathbf{s})$  is obtained by applying  $\mathcal{L}_2$  to the solution of (4) with  $\mathcal{L}_2 = \mathbf{I}$ .

Fuglstad (2010) discussed the approximated solutions to a SPDE for constructing space-time GMRFs. The equation discussed by him has the form

$$\frac{\partial}{\partial t}x(s, t) - \nabla \cdot \nabla x(s, t) = \tau \mathcal{W}(s, t), \quad (s, t) \in [0, L] \times [0, T], \quad (6)$$

where,  $\nabla = \frac{\partial}{\partial x}$ ,  $\tau > 0$  is a constant and  $\mathcal{W}(s, t)$  is Gaussian space-time white noise. Fuglstad (2010) only discussed  $s \in \mathbb{R}$  and argued that this equation has real physical meaning since the prototype of (6) is the stochastic heat equation

$$\frac{\partial}{\partial x}q(s, t) + \nabla F(s, t) = f(s, t). \quad (7)$$

The heat equation relates the change of the size  $q$  in time to the spatial divergences of the flux  $F(s, t)$  and the source term  $f(s, t)$ . Fuglstad (2010) chose a finite volume method for solving Equation (6), and claimed that the finite volume method gave the correct distribution of the total energy of the solution to SPDE (6).

Lindgren et al. (2011) discussed the methodology for constructing non-stationary GRFs and non-separable space-time models. They claimed that if the parameters  $\kappa$  and  $\tau$  depend on the coordinate  $\mathbf{s}$ , then we can construct a non-stationary GRF. The SPDE then has the form

$$(\kappa^2(\mathbf{s}) - \Delta)^{\alpha/2} \{\tau(\mathbf{s})x(\mathbf{s})\} = \mathcal{W}(\mathbf{s}). \quad (8)$$

The non-separable space-time models have interaction between space and time in the covariance structure. In general it is difficult to construct this kind of model through a covariance function based approach. However, the SPDE approach can be used. One of the non-separable SPDEs which can result in this kind of model is

$$\left\{ \frac{\partial}{\partial t} + (\kappa^2 + \mathbf{m} \cdot \nabla - \nabla \cdot \mathbf{H} \nabla) \right\} x(\mathbf{s}, t) = \varepsilon(\mathbf{s}, t), \quad (9)$$

where  $\mathbf{m}$  is a transport vector,  $\mathbf{H}$  is a positive definite diffusion matrix and  $\varepsilon(\mathbf{s}, t)$  is a stochastic space-time noise process. We refer to Lindgren et al. (2011, Section 3.5) for detailed discussion on this topic.

## 2.2 Multivariate GRFs

A multivariate GRF with  $p$  components  $\mathbf{x}(\mathbf{s}) = (x_1(\mathbf{s}), x_2(\mathbf{s}), \dots, x_p(\mathbf{s}))^T$ ,  $\mathbf{s} \in \mathbb{R}^d$ , is a collection of continuously indexed multivariate normal random vectors such that

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

where  $\boldsymbol{\mu}$  is the mean of the random field and  $\boldsymbol{\Sigma}$  is the covariance matrix. Assume, at the current stage, that the process is second-order stationary with mean zero. One approach for constructing stationary and isotropic multivariate GRFs using covariance-based models was proposed by Gneiting et al. (2010). The covariance function  $\mathbf{C}$  in their approach is given by

$$\mathbf{C}(\mathbf{h}) = \begin{pmatrix} C_{11}(\mathbf{h}) & C_{12}(\mathbf{h}) & \cdots & C_{1p}(\mathbf{h}) \\ C_{21}(\mathbf{h}) & C_{22}(\mathbf{h}) & \cdots & C_{2p}(\mathbf{h}) \\ \vdots & \vdots & \ddots & \vdots \\ C_{p1}(\mathbf{h}) & C_{p2}(\mathbf{h}) & \cdots & C_{pp}(\mathbf{h}) \end{pmatrix}, \quad (10)$$

where  $\{C_{ii}(\mathbf{h}) = \sigma_{ii}M(\mathbf{h}|\nu_{ii}, \kappa_{ii}); i = 1, \dots, p\}$  are the marginal covariance functions and  $\{C_{ij}(\mathbf{h}) = \rho_{ij}\sigma_i\sigma_jM(\mathbf{h}|\nu_{ij}, \kappa_{ij}); i, j = 1, \dots, p, i \neq j\}$  are the cross-covariance functions.  $\{C_{ii}(\mathbf{h}) = \mathbb{E}(x_i(\mathbf{s}+\mathbf{h})x_i(\mathbf{s})); i = 1, 2, \dots, p\}$  give information about the covariance structures within the fields  $\{x_i(\mathbf{s})\}$ .  $\{C_{ij}(\mathbf{h}) = \mathbb{E}(x_i(\mathbf{s}+\mathbf{h})x_j(\mathbf{s})); i, j = 1, 2, \dots, p, i \neq j\}$  describes the covariance structure between fields  $\{x_i(\mathbf{s})\}$  and  $\{x_j(\mathbf{s})\}$ .  $\{\rho_{ij}\}$  are the co-located correlation coefficients.  $\{\sigma_{ii} \geq 0\}$  are the marginal variances, and  $\{\sigma_i\}$  and  $\{\sigma_j\}$  are the corresponding standard deviations. They satisfy the relationships throughout this paper  $\sigma_{ii} = \sigma_i^2, \sigma_{ij} = \sigma_i\sigma_j$ . The main difficulty in constructing useful multivariate models using this kind of approach is the nonnegative definiteness requirement for the covariance functions. Gneiting et al. (2010) proposed a way to specify valid parametric models through the covariance functions given in Equation (10) directly. Several theorems were presented in order to ensure the matrix-valued covariance function to be symmetric and nonnegative definite.

Hu et al. (2012a) proposed to use a system of SPDEs to construct a multivariate GRF. They claimed that the notorious requirement nonnegative definiteness for the covariance function is automatically fulfilled with their approach. Hu et al. (2012a) also discussed the link between the system of SPDEs approach and the covariance function based approach discussed by Gneiting et al. (2010). The system of SPDEs which has been used for constructing the multivariate GRFs by Hu et al. (2012a) is

$$\begin{pmatrix} \mathcal{L}_{11} & \mathcal{L}_{12} & \cdots & \mathcal{L}_{1p} \\ \mathcal{L}_{21} & \mathcal{L}_{22} & \cdots & \mathcal{L}_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{L}_{p1} & \mathcal{L}_{p2} & \cdots & \mathcal{L}_{pp} \end{pmatrix} \begin{pmatrix} x_1(\mathbf{s}) \\ x_2(\mathbf{s}) \\ \vdots \\ x_p(\mathbf{s}) \end{pmatrix} = \begin{pmatrix} \varepsilon_1(\mathbf{s}) \\ \varepsilon_2(\mathbf{s}) \\ \vdots \\ \varepsilon_p(\mathbf{s}) \end{pmatrix}, \quad (11)$$

where  $\{\mathcal{L}_{ij} = b_{ij}(\kappa_{ij}^2 - \Delta)^{\alpha_{ij}/2}; i = j = 1, 2, \dots, p\}$  are differential operators with  $\{\alpha_{ij} = 0 \text{ or } 2\}$ ,  $\{\kappa_{ij}\}$  and  $\{\nu_{ij}\}$  are scaling parameters and smoothness parameters.  $\{\varepsilon_i(\mathbf{s}); i = 1, 2, \dots, p\}$  are independent but not necessarily identically distributed noise processes, and  $\{b_{ij}\}$  are the parameters related to the marginal variances of the fields and the cross-covariances between the fields. Hu et al. (2012a) pointed out that the GMRF approximation can be applied to the GRFs. Hence they can use computationally efficient GMRFs for sampling and inference. However, the constructed GRFs are always stationary and isotropic, and the covariance functions are not oscillating.



### 2.3 GMRFs approximation to GRFs

Generally speaking, GRFs are commonly used in statistical modelling because of their good theoretical properties. However, the GRFs have a bottle-neck on the computational side. The computational cost for factorizing a dense covariance matrix  $\Sigma$  with dimension  $n \times n$  is  $\mathcal{O}(n^3)$ . Even though the computational power is at an all time high, it seems that in many situations it is infeasible to do the computations in reasonable time. Banerjee et al. (2004, Appendix A.5) informally call this situation “the big  $n$  problem”.

There are many different approaches trying to avoid or overcome “the big  $n$  problem”, such as covariance tapering (Furrer et al., 2006; Zhang and Du, 2008; Kaufman et al., 2008; Shaby and Ruppert, 2012), likelihood approximations (Vecchia, 1988; Stein et al., 2004), and fixed rank kriging and fixed rank filtering (Cressie and Johannesson, 2008; Cressie et al., 2010). The approach which has been chosen in this paper is based on the GMRF approximation to GRFs. The sparsity of the precision matrix  $\mathbf{Q}$  enables the numerical algorithms for sparse matrix for fast inference with large datasets (Rue, 2001; Rue and Held, 2005; Lindgren et al., 2011). The general cost for factorizing the sparse matrix  $\mathbf{Q}$  is  $\mathcal{O}(n)$ ,  $\mathcal{O}(n^{3/2})$  and  $\mathcal{O}(n^2)$  in one dimension, two dimensions and three dimensions, respectively (Rue and Held, 2005). Hartman and Hössjer (2008) proposed to use the GMRFs for GRFs for spatial prediction with Kriging, due to the pleasant computational properties of GMRFs.

In this paper we only give an overview of GMRF approximation to univariate GRF and refer to Hu et al. (2012a, Section 2.3) for detailed discussion on GMRF approximation to multivariate GRF. In order to find a GMRF approximation of a GRF on a triangulated lattice, we at first need to find the stochastic weak formulation of SPDE (2) (Kloeden and Platen, 1999). In this paper we use Delaunay triangulation. We refer to Hjelle and Dæhlen (2006) for more information about triangulations. Denote the inner product of functions  $h$  and  $g$  as

$$\langle h, g \rangle = \int h(\mathbf{s})g(\mathbf{s})d(\mathbf{s}), \quad (12)$$

where the integration is within the region of interest. The stochastic weak solution of SPDE (2) is found by requiring

$$\left\{ \langle \phi_i, (\kappa^2 - \Delta)^{\alpha/2} x \rangle; i = 1, \dots, M \right\} \stackrel{d}{=} \left\{ \langle \phi_i, \mathcal{W} \rangle; i = 1, \dots, M \right\}, \quad (13)$$

where  $M$  is the number of test functions  $\{\phi_i(\mathbf{s})\}$  and “ $\stackrel{d}{=}$ ” denotes equality in distribution.

Then we need to find the finite element representation of the solution to the SPDE. The finite element representation of the solution is

$$x(\mathbf{s}) = \sum_{i=1}^N \psi_i(\mathbf{s})\omega_i \quad (14)$$

with basis functions  $\{\psi_i(\mathbf{s}); i = 1, 2, \dots, N\}$  and Gaussian distributed weights  $\{\omega_i; i = 1, 2, \dots, N\}$ .  $N$  is the number of vertexes in the triangulation. We refer to Zienkiewicz

et al. (2005) and Brenner and Scott (2008) for more information and theoretical background of finite element methods. The approach given by Lindgren et al. (2011) for choosing the basis functions is used in this paper. With  $M = N$  we choose each basis function  $\psi_i(\mathbf{s})$  to be piecewise linear on each triangle with  $\psi_i(\mathbf{s}) = 1$  at vertex  $i$  and  $\psi_i(\mathbf{s}) = 0$  at other vertexes. This choice of basis functions means that the local interpolation on a triangle is linear. Lindgren et al. (2011) pointed out that other methods, such as kernel method, are useful in theory but not necessary in practice. When  $\alpha_{ij} = 1$  the *least squares* approximation is chosen,  $\phi_k = (\kappa^2 - \Delta)^{\frac{1}{2}} \psi_k$ . When  $\alpha_{ij} = 2$  the *Galerkin* solution is chosen,  $\phi_k = \psi_k$ . When  $\alpha_{ij} \geq 3$  the *recursive Galerkin* formulation is used. We refer to Lindgren et al. (2011, Section 2.3) for more information about the recursive Galerkin formation.

## 2.4 Outline of the paper

The structure of the rest of the paper is organized as follows. Section 3 gives the detailed discussion on how to construct multivariate GRFs with oscillating covariance functions through the systems of SPDEs approach. Examples with simulated data and real data are given in Section 4. Discussion and future work in Section 5 ends this paper.

## 3 Model formulation

GRFs with oscillating covariance functions can be used in many situations, for example, for modelling global pressure (Lindgren et al., 2011) and ocean waves (Lindgren, 2010). First, an overview for constructing the *univariate* GRFs with oscillating covariance function is given since it is needed for constructing *multivariate* GRFs with oscillating covariance functions. Next, we introduce a general approach for constructing the multivariate GRFs with oscillating covariance functions. Then explicit approach for constructing the *bivariate* GRFs is discussed. At last we discuss the procedure for sampling the multivariate GRFs with oscillating covariance functions.

### 3.1 Univariate GRFs with oscillating covariance functions

Lindgren et al. (2011, Section 3.3) discussed how to construct an univariate GRF with oscillating covariance function using a SPDE with complex number. For the case  $\alpha = 2$  the SPDE has the form

$$\{\kappa^2 \exp(i\pi\omega) - \Delta\} x(\mathbf{s}) = \mathcal{W}(\mathbf{s}), \quad (15)$$

where  $\omega \in [0, 1]$  is the oscillation parameter,  $x(\mathbf{s}) = x_1(\mathbf{s}) + ix_2(\mathbf{s})$ , and  $\mathcal{W}(\mathbf{s}) = \mathcal{W}_1(\mathbf{s}) + i\mathcal{W}_2(\mathbf{s})$ . The innovation processes  $\mathcal{W}_1(\mathbf{s})$  and  $\mathcal{W}_2(\mathbf{s})$  are independent standard Gaussian white noise processes. Lindgren et al. (2011) pointed out that the real and imaginary parts,  $x_1(\mathbf{s})$  and  $x_2(\mathbf{s})$ , of the stationary solution  $x(\mathbf{s})$  are independent with identical spectrum densities

$$R(\mathbf{k}) = \frac{1}{(2\pi)^d (\kappa^4 + 2 \cos(\pi\omega) \kappa^2 \|\mathbf{k}\|^2 + \|\mathbf{k}\|^4)}, \quad \mathbf{k} \in \mathbb{R}^d. \quad (16)$$

With this approach the common isotropic stationary Matérn random fields can be obtained by setting  $\omega = 0$ . We can notice that  $\omega = 1$  generates intrinsic stationary random fields. We refer to Rue and Held (2005, Chapter 3) for more information on the intrinsic random fields. When  $\omega \in (0, 1)$ , the constructed GRFs have covariance functions with oscillation. The oscillation is increasing with larger value of  $\omega$ . The closed form of the precision matrix  $\mathbf{Q}$  for the stationary GRFs with oscillation can be obtained from (16),

$$\mathbf{Q}(\kappa^2, \omega) = \kappa^4 \mathbf{C} + 2 \cos(\pi\omega) \kappa^2 \mathbf{G} + \mathbf{G} \mathbf{C}^{-1} \mathbf{G}. \quad (17)$$

The matrices  $\mathbf{C}$  and  $\mathbf{G}$  in Equation (17) are defined through

$$\begin{aligned} C_{ij} &= \langle \psi_i, 1 \rangle, \\ G_{ij} &= \langle \nabla \psi_i, \nabla \psi_j \rangle, \end{aligned} \quad (18)$$

with basis functions  $\{\psi_i; i = 1, 2, \dots, n\}$ . We use  $C_{ij} = \langle \psi_i, 1 \rangle$ , instead of  $C_{ij} = \langle \psi_i, \psi_j \rangle$ , in order to make the precision matrix sparse. This setting yields a Markov approximation to the FEM solution. Bolin and Lindgren (2009) studied the effects of the Markov approximation and claimed that the difference between the Markov approximation and exact FEM representation is negligible.

Lindgren et al. (2011) pointed out that this complex-valued version of SPDE (15) can be rewritten as a special case of the coupled systems of SPDEs

$$\begin{pmatrix} h_1 - \Delta & -h_2 \\ h_2 & h_1 - \Delta \end{pmatrix} \begin{pmatrix} x_1(\mathbf{s}) \\ x_2(\mathbf{s}) \end{pmatrix} = \begin{pmatrix} \mathcal{W}_1(\mathbf{s}) \\ \mathcal{W}_2(\mathbf{s}) \end{pmatrix}, \quad (19)$$

where  $h_1 = \kappa^2 \cos(\pi\omega)$  and  $h_2 = \kappa^2 \sin(\pi\omega)$ . The random fields  $x_1(\mathbf{s})$  and  $x_2(\mathbf{s})$  from Equation (19) have the same precision matrix  $\mathbf{Q}$  given in Equation (17). Lindgren et al. (2011) commented that it is surprising that these two fields from the coupled system of SPDEs (19) are always independent regardless of the choices of parameters. Additionally, the univariate GRFs with oscillating covariance functions from Equation (15) are always isotropic. However, it is possible to construct non-isotropic GRFs by slightly modifying the coupled system of SPDEs (19). We are not going to discuss this issue here and we focus only on the isotropic GRFs. We refer to Lindgren et al. (2011, Appendix C.4) for more information about the oscillating and non-isotropic cases.

### 3.2 Multivariate GRFs with oscillating covariance functions

The multivariate GRFs with oscillating covariance functions, in this paper, all have the assumption that the mean is zero, i.e.,  $\mathbf{x}(\mathbf{s}) \sim \text{MVN}(\mathbf{0}, \mathbf{\Sigma})$ . Hu et al. (2012a) proposed to construct the multivariate GRFs using the system of SPDEs given in (11). In their approach the multivariate GRFs are always isotropic and stationary. The covariance functions cannot be oscillating. They argued that, under some conditions, it is possible to construct the multivariate GRFs with Matérn covariance functions as discussed by Gneiting et al. (2010). In this section we are going to discuss how to construct multivariate GRFs where some components of the random fields have oscillating covariance

functions. The main idea is to replace the noise processes by noise processes with oscillating covariance functions. With this approach the systems of SPDEs have the same form as given in (11), but the noise processes are different. Even though the system of SPDEs in (11) is theoretically general, we recommend to use the triangular system of SPDEs in applications. The triangular system of SPDEs is

$$\begin{pmatrix} \mathcal{L}_{11} & & & \\ \mathcal{L}_{21} & \mathcal{L}_{22} & & \\ \vdots & \vdots & \ddots & \\ \mathcal{L}_{p1} & \mathcal{L}_{p2} & \dots & \mathcal{L}_{pp} \end{pmatrix} \begin{pmatrix} x_1(\mathbf{s}) \\ x_2(\mathbf{s}) \\ \vdots \\ x_p(\mathbf{s}) \end{pmatrix} = \begin{pmatrix} \varepsilon_1(\mathbf{s}) \\ \varepsilon_2(\mathbf{s}) \\ \vdots \\ \varepsilon_p(\mathbf{s}) \end{pmatrix}, \quad (20)$$

where  $\{\mathcal{L}_{ij}; i, j = 1, 2, \dots, p, i \geq j\}$  are differential operators as defined in (11),  $\{\mathcal{L}_{ij} = 0; i, j = 1, 2, \dots, p, i < j\}$  and  $\{\varepsilon_i(\mathbf{s}); i = 1, 2, \dots, p\}$  are noise processes where some of them have oscillating covariance functions. We recommend to use as fewer noise processes with oscillating covariance functions as possible. This system has many advantages, such as interpretation of the properties of the fields. For example, we know which components of the random field must have non-oscillating covariance functions and have oscillating covariance functions. However, there are some components of the random field which might have oscillating covariance functions. We divide the random fields into three categories  $\mathbf{x}_m$ ,  $\mathbf{x}_o$  and  $\mathbf{x}_p$ , where  $\mathbf{x}_m$  denotes the random fields with non-oscillating covariance functions,  $\mathbf{x}_o$  denotes the random fields with oscillating covariance functions and  $\mathbf{x}_p$  denotes the random fields with covariance functions which might be oscillating. Assume that only  $\{\varepsilon_i(\mathbf{s}); i = 1, 2, \dots, p\}$  is the noise process with oscillating covariance function and other noise processes  $\{\varepsilon_j(\mathbf{s}); j = 1, 2, \dots, p, j \neq i\}$  are noise processes with non-oscillating covariance functions, and then we can obtain the following results.

- If only the covariance function for the noise process  $\varepsilon_i(\mathbf{s})$  is oscillating, the covariances functions for all the random fields  $x_j(\mathbf{s}) (j < i)$  are non-oscillating,  $x_j(\mathbf{s}) \in \mathbf{x}_n(\mathbf{s})$ ;
- If only the covariance function for the noise process  $\varepsilon_i(\mathbf{s})$  is oscillating, the random field  $x_j(\mathbf{s}) (j = i)$  has an oscillating covariance function,  $x_j(\mathbf{s}) \in \mathbf{x}_o(\mathbf{s})$ ;
- If only the covariance function for the noise process  $\varepsilon_i(\mathbf{s})$  is oscillating, the random fields  $x_j(\mathbf{s}) (j > i)$  belong to  $\mathbf{x}_p(\mathbf{s})$ , which means that the covariance functions for these random fields might be oscillating.

This result is rather intuitive since we can obtain it by checking the system of SPDEs (20) directly. However, these results are important in the real-world application since it gives information for how to build models in a reasonable way. For instance, we can get information about how to choose the order of the random fields.

### 3.3 Bivariate GRFs with oscillating covariance functions

The methodology for constructing non-oscillating and isotropic bivariate GRFs explicitly has been studied by Gneiting et al. (2010) and Hu et al. (2012a). In this section we discuss the approach for constructing bivariate GRFs explicitly with oscillating covariance functions using systems of SPDEs. We start the investigation with random fields constructed by the full system of SPDEs

$$\begin{aligned} b_{11}(\kappa_{11}^2 - \Delta)^{\alpha_{11}/2} x_1(\mathbf{s}) + b_{12}(\kappa_{12}^2 - \Delta)^{\alpha_{12}/2} x_2(\mathbf{s}) &= \varepsilon_1(\mathbf{s}), \\ b_{22}(\kappa_{22}^2 - \Delta)^{\alpha_{22}/2} x_2(\mathbf{s}) + b_{21}(\kappa_{21}^2 - \Delta)^{\alpha_{21}/2} x_1(\mathbf{s}) &= \varepsilon_2(\mathbf{s}), \end{aligned} \quad (21)$$

where  $b_{ij}$  and  $\{\mathcal{L}_{ij}(\mathbf{s}); i, j = 1, 2\}$  are the same as in (11), and  $\{\varepsilon_i(\mathbf{s}); i = 1, 2\}$  are noise processes which can have oscillating covariance functions. By changing the properties of the noise processes we can construct more interesting random fields. Use the matrix notion and define the operator matrix as

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

and let  $\boldsymbol{\varepsilon}(\mathbf{s}) = (\varepsilon_1(\mathbf{s}), \varepsilon_2(\mathbf{s}))^H$ , where  $H$  denotes the Hermitian transpose of a vector or a matrix.  $\theta_{ij} = \{\alpha_{ij}, \kappa_{ij}, b_{ij}\}$  is defined as the collection of parameters for  $\mathcal{L}_{ij}$ . The system of equations (21) can then be written in a compact matrix form as

$$\mathcal{L}(\boldsymbol{\theta})\mathbf{x}(\mathbf{s}) = \boldsymbol{\varepsilon}(\mathbf{s}), \quad (23)$$

where  $\boldsymbol{\theta} = \{\theta_{ij}, i, j = 1, 2\}$ . With (23) we can obtain the power spectrum  $\mathbf{S}_x = \mathbb{E}(\hat{\mathbf{x}} \cdot \hat{\mathbf{x}}^H)$  by

$$\mathbf{S}_x = \mathcal{H}^H \mathbf{S}_\varepsilon \mathcal{H}^{-H}, \quad (24)$$

where  $-H$  denotes the inverse of the complex conjugate of the matrix.  $\hat{x}_{ij}$  is the Fourier transform of  $x_{ij}$ ,  $\hat{x}_{ij} = \mathcal{F}(x_{ij})$ , and  $\mathcal{H}$  is the Fourier transform of the operator matrix  $\mathcal{L}$ ,

$$\mathcal{H}(\boldsymbol{\theta}) = \begin{pmatrix} \mathcal{H}_{11}(\theta_{11}) & \mathcal{H}_{12}(\theta_{12}) \\ \mathcal{H}_{21}(\theta_{21}) & \mathcal{H}_{22}(\theta_{22}) \end{pmatrix}. \quad (25)$$

$\mathbf{S}_\varepsilon(\mathbf{k}) = \mathbb{E}(\hat{\boldsymbol{\varepsilon}} \hat{\boldsymbol{\varepsilon}}^H)$  is the power spectrum matrix for the independent noise processes

$$\mathbf{S}_\varepsilon(\mathbf{k}) = \begin{pmatrix} S_{\varepsilon_1}(\mathbf{k}) & 0 \\ 0 & S_{\varepsilon_2}(\mathbf{k}) \end{pmatrix}, \quad (26)$$

where  $\mathbf{k}$  is the frequency. Since the noise processes are mutually independent, the power spectrum matrix of noise processes is a diagonal matrix. Using Equation (24) - Equation

(25), the elements in the power spectrum matrix of the bivariate fields from the full system of SPDEs in (21) can be obtained,

$$\begin{aligned}
S_{x_{11}}(\mathbf{k}) &= \frac{S_{\varepsilon_1}|\mathcal{H}_{22}^2| + S_{\varepsilon_2}|\mathcal{H}_{12}^2|}{|(\mathcal{H}_{11}\mathcal{H}_{22} - \mathcal{H}_{12}\mathcal{H}_{21})^2|}, \\
S_{x_{12}}(\mathbf{k}) &= -\frac{\mathcal{H}_{22}S_{\varepsilon_1}|\mathcal{H}_{21}^2|\mathcal{H}_{11} + \mathcal{H}_{12}S_{\varepsilon_2}|\mathcal{H}_{11}^2|\mathcal{H}_{21}}{|(\mathcal{H}_{11}\mathcal{H}_{22} - \mathcal{H}_{12}\mathcal{H}_{21})^2|\mathcal{H}_{21}\mathcal{H}_{11}}, \\
S_{x_{21}}(\mathbf{k}) &= -\frac{\mathcal{H}_{21}S_{\varepsilon_1}|\mathcal{H}_{22}^2|\mathcal{H}_{12} + \mathcal{H}_{11}S_{\varepsilon_2}|\mathcal{H}_{12}^2|\mathcal{H}_{22}}{|(\mathcal{H}_{11}\mathcal{H}_{22} - \mathcal{H}_{12}\mathcal{H}_{21})^2|\mathcal{H}_{22}\mathcal{H}_{12}}, \\
S_{x_{22}}(\mathbf{k}) &= \frac{S_{\varepsilon_1}|\mathcal{H}_{21}^2| + S_{\varepsilon_2}|\mathcal{H}_{11}^2|}{|(\mathcal{H}_{11}\mathcal{H}_{22} - \mathcal{H}_{12}\mathcal{H}_{21})^2|}.
\end{aligned} \tag{27}$$

Define poles and zeros as the roots of the denominators and numerators of the power spectrum elements  $\{S_{x_{ij}}; i, j = 1, 2\}$ , respectively. From (27), we can see that the poles of the power spectrum, in general, are the same for both the fields, but zeros of the power spectrum will be different. It gives us a possibility to construct bivariate GRFs with oscillating covariance functions by carefully re-parametrization of system of SPDEs (21). However, we will not discuss this approach in this paper, but leave it for future research. The approach we have chosen here is to change the noise process at the right hand of system of SPDEs (21).

Theoretically, we could choose the full version of the system of SPDEs given in (21) and give more flexibility for constructing bivariate random fields. However, we choose to simplify the model. Hu et al. (2012a) used the triangular version of the SPDEs system to construct bivariate GRFs and this suggestion is followed in this paper. In the following sections, we focus on a special form of the triangular system of the SPDEs discussed in Section 3.2. In the special form the operator matrix is

$$\mathcal{L}_1(\boldsymbol{\theta}) = \begin{pmatrix} b_{11}(h_{11} - \Delta) & 0 \\ b_{21} & b_{22}(h_{22} - \Delta) \end{pmatrix}, \tag{28}$$

where the subscript “1” in  $\mathcal{L}_1$  is used to denote the first operator matrix we use for constructions. Some other operator matrices are discussed in Appendix A. We can rewrite the system of SPDEs with matrix notation as

$$\mathcal{L}_1(\boldsymbol{\theta})\mathbf{x}(\mathbf{s}) = \boldsymbol{\varepsilon}(\mathbf{s}), \tag{29}$$

and Equation (29) can be written down explicitly as

$$\begin{aligned}
b_{11}(h_{11} - \Delta)x_1(\mathbf{s}) &= \varepsilon_1(\mathbf{s}), \\
b_{21}x_1(\mathbf{s}) + b_{22}(h_{22} - \Delta)x_2(\mathbf{s}) &= \varepsilon_2(\mathbf{s}).
\end{aligned} \tag{30}$$

In this form both random fields can have oscillating covariance function. The following discussion are based on the system of SPDEs (30). Let  $\varepsilon_1(\mathbf{s})$  be a noise process with non-oscillating covariance function, such as a white noise process or noise process with Matérn covariance function, and  $\varepsilon_2(\mathbf{s})$  be a noise process with oscillating covariance

function generated from the complex-valued SPDEs (15). We can then conclude that the first field  $x_1(\mathbf{s})$  is a stationary and isotropic random field with non-oscillating covariance function, and that  $x_2(\mathbf{s})$  is a random field with oscillating covariance function. On the other hand, if  $\varepsilon_1(\mathbf{s})$  has an oscillating covariance function and  $\varepsilon_2(\mathbf{s})$  has a non-oscillating covariance function, then the covariance functions for both random fields  $x_1(\mathbf{s})$  and  $x_2(\mathbf{s})$  are oscillating given that  $b_{21} \neq 0$ .

We use the power spectra  $\{S_{x_{ii}}(\mathbf{k}); i = 1, 2\}$  of the random fields together with the cross spectrum  $S_{x_{21}}(\mathbf{k})$  to investigate the properties of the random fields,

$$\begin{aligned} S_{x_{11}}(\mathbf{k}) &= \frac{S_{\varepsilon_1}}{b_{11}^2(h_{11} + \|\mathbf{k}\|^2)^2}, \\ S_{x_{21}}(\mathbf{k}) &= -\frac{b_{21}S_{\varepsilon_1}}{b_{22}(h_{22} + \|\mathbf{k}\|^2)b_{11}^2(h_{11} + \|\mathbf{k}\|^2)^2}, \\ S_{x_{22}}(\mathbf{k}) &= \frac{b_{21}^2S_{\varepsilon_1} + b_{11}^2(h_{11} + \|\mathbf{k}\|^2)^2S_{\varepsilon_2}}{b_{11}^2(h_{11} + \|\mathbf{k}\|^2)^2b_{22}^2(h_{22} + \|\mathbf{k}\|^2)^2}. \end{aligned} \quad (31)$$

The following results can be obtained from (31).

- The marginal variance of the first random field  $x_1(\mathbf{s})$  is related only to the parameters  $b_{11}$  and  $h_{11}$ .
- The marginal variance of the second random field  $x_2(\mathbf{s})$  is related only to  $\{b_{ij}, h_{ij}; i, j = 1, 2\}$ .
- The sign of  $b_{11}$  is irrelevant to the sign of the cross-correlation between the two fields. Since the marginal variance of the first random field  $x_1(\mathbf{s})$  is only related to  $b_{11}$  and  $h_{11}$ , and there is a requirement  $h_{11} > 0$ , we can set  $b_{11} > 0$ ;
- The sign of the correlation between the two fields only depends on the sign of the product of  $b_{22}$  and  $b_{21}$ . We recommend to set  $b_{22} > 0$ , and then the sign of the correlation between the fields will be related only to the sign of  $b_{21}$ . If  $b_{21} > 0$ , the two random fields are negatively correlated. If  $b_{21} < 0$ , the random fields are positively correlated.

### 3.4 Sampling the bivariate GRFs

The common approach for sampling GRFs uses the covariance matrix  $\Sigma$  or precision matrix  $\mathbf{Q}$ . Since the bivariate GRFs from the systems of SPDEs are represented by the GMRFs, the precision matrices  $\mathbf{Q}$  are sparse. Therefore, the direct approach for sampling a (multivariate) GMRF is usually through the Cholesky triangle  $\mathbf{L}$ , where  $\mathbf{Q} = \mathbf{L}\mathbf{L}^T$ . The commonly used procedure for getting a sample from the GMRF  $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{Q}^{-1})$  is through the following steps

- I. Use the Cholesky factorization to find the Cholesky triangle  $\mathbf{L}$  of the precision matrix  $\mathbf{Q}$ . We usually do the Cholesky factorization with standard libraries.

Table 1: Parameters for sampling the bivariate GRFs

Parameters		
$\alpha$	$\kappa$	$b$
$\alpha_{11} = 2$	$h_{11} = 0.25$	$b_{11} = 0.5$
$\alpha_{12} = 0$	$h_{22} = 0.36$	$b_{12} = 0$
$\alpha_{21} = 2$	$\kappa_{n_1} = 0.5$	$b_{21} = 0.25$
$\alpha_{22} = 2$	$\kappa_{n_2} = 0.6$	$b_{22} = 1$
$\alpha_{n_1} = 2$	$\omega = 0.95$	
$\alpha_{n_2} = 2$		

- II. Get a sample  $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ .  $\mathbf{I}$  is an identity matrix and has the same dimensions as the precision matrix  $\mathbf{Q}$ .
- III. Solve a linear system of equations with Cholesky triangle  $\mathbf{L}\mathbf{v} = \mathbf{z}$ . Thus  $\mathbf{v}$  has the correct covariance matrix  $\mathbf{Q}^{-1}$  since  $\text{Cov}(\mathbf{v}) = \text{Cov}(\mathbf{L}^{-T}\mathbf{z}) = (\mathbf{L}\mathbf{L}^T)^{-1} = \mathbf{Q}^{-1}$ .
- IV. Correct the mean by  $\mathbf{x} = \boldsymbol{\mu} + \mathbf{v}$ , and then  $\mathbf{x}$  has the correct mean  $\boldsymbol{\mu}$  and covariance matrix  $\mathbf{Q}^{-1}$ ,  $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{Q}^{-1})$ .

If the precision matrix  $\mathbf{Q}$  is a band matrix, the Cholesky triangle  $\mathbf{L}$  will be also a band matrix. The corresponding algorithm for finding the Cholesky triangle when  $\mathbf{Q}$  is a band matrix can be found in Rue and Held (2005, Algorithm 2.9). We also refer to Rue and Held (2005, Chapter 2) for detailed discussion about different sampling algorithms for GMRFs with different kinds of parametrization. Hu et al. (2012b) showed that it is possible to find a sparser triangular matrix  $\tilde{\mathbf{L}}$  with incomplete orthogonal factorization for sampling the GMRF, but they pointed out that it needs longer computation time for finding the sparse triangular matrix.

In the following two examples, we choose all values of parameters to be equal. However, we set  $\varepsilon_1(\mathbf{s})$  to be a noise process with a non-oscillating covariance function and  $\varepsilon_2(\mathbf{s})$  to be a noise process with an oscillating covariance function in the first example. In the second example, we simply switch the order of the noise processes, i.e., we set  $\varepsilon_1(\mathbf{s})$  to be a noise process with an oscillating covariance function and  $\varepsilon_2(\mathbf{s})$  to be a noise process with a non-oscillating covariance function.

One sample from the GMRF in the first example is shown in Fig. 1 and one sample in the second example is shown in Fig. 4. The corresponding correlation matrices are shown in Fig. 2 and Fig. 5. In these figures, the red lines indicate that the correlation is 0. In the first example the random field  $x_1(\mathbf{s})$  has a non-oscillating covariance function and the second random field  $x_2(\mathbf{s})$  has an oscillating covariance function. In the second example both the fields have oscillating covariance functions. These two examples verify the conclusion in Section 3.3.



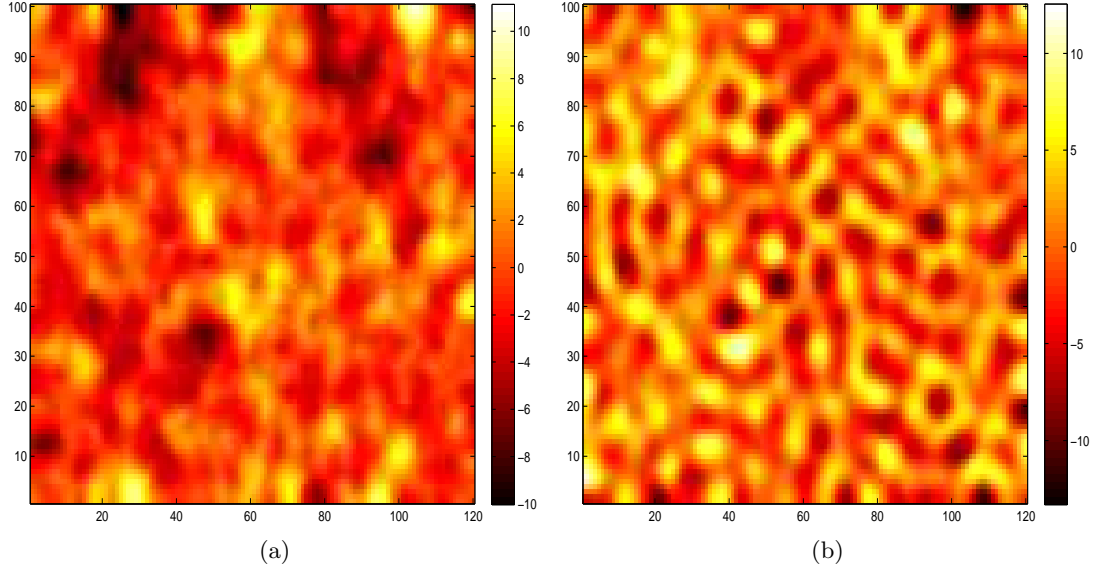


Figure 1: A realization of the bivariate random field with parameters given in Table 1. In this example the first noise process has a non-oscillating covariance function and the second noise process has an oscillating covariance function.

## 4 Inference with simulated data and real data

In this section we illustrate how to use our approach with some simulated data examples and one real data example. In the first example, the covariance function of the first random field is non-oscillating, but the second random field has an oscillating covariance function. In the second example, both the random fields have oscillating covariance functions. The third and the forth examples show that if the two fields are independent, the inferences give indications about this, no matter which field is oscillating. One real data example in the end shows that our approach can be applied in practical applications. As pointed out by Diggle et al. (1998, Chapter 5) and Lindgren et al. (2011, Section 2), the smoothness parameter  $\{\nu_{ij}; i, j = 1, 2\}$  are poorly identifiable. Therefore, we fix the values of  $\{\alpha_{ij}\}$  in the simulated data examples and in the real data example.

### 4.1 Posterior for the hyper-parameters

The first step of the inference is usually to derive the (log-) posterior distribution  $\pi(\boldsymbol{\theta}|\mathbf{y})$  of  $\boldsymbol{\theta}$ . The well known Bayesian formula (32) is at the core of Bayesian inference,

$$\begin{aligned}\pi(\mathbf{y}, \boldsymbol{\theta}) &= \frac{\pi(\boldsymbol{\theta}, \mathbf{x}, \mathbf{y})}{\pi(\mathbf{x}|\boldsymbol{\theta}, \mathbf{y})} \\ &= \frac{\pi(\boldsymbol{\theta})\pi(\mathbf{x}|\boldsymbol{\theta})\pi(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta})}{\pi(\mathbf{x}|\mathbf{y}, \boldsymbol{\theta})}.\end{aligned}\tag{32}$$

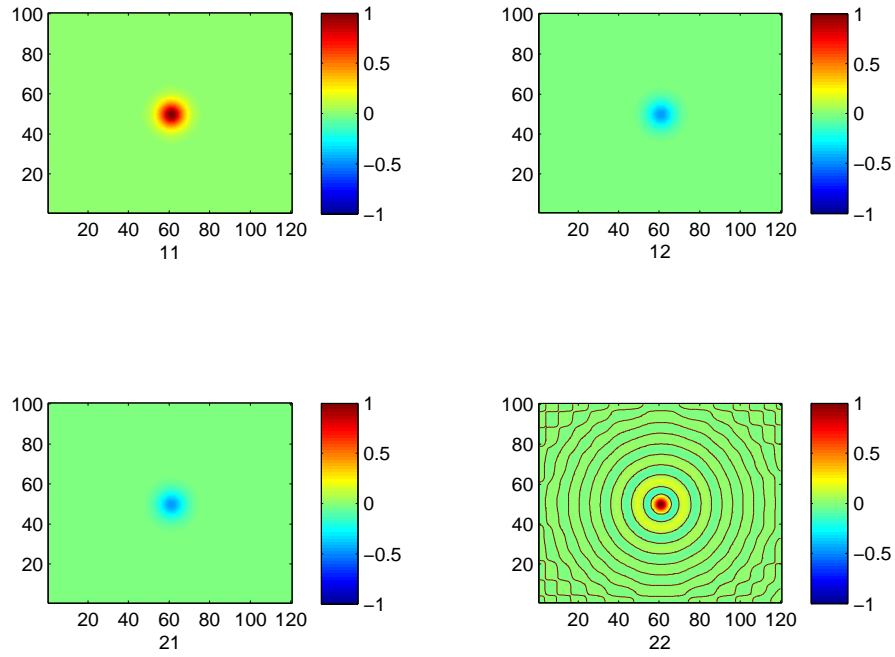


Figure 2: Correlations and cross-correlation functions for the bivariate random field with parameters given in Table 1. In this example the first noise process has a non-oscillating covariance function and the second noise process has an oscillating covariance function. We see that only the second random field has an oscillating covariance function.

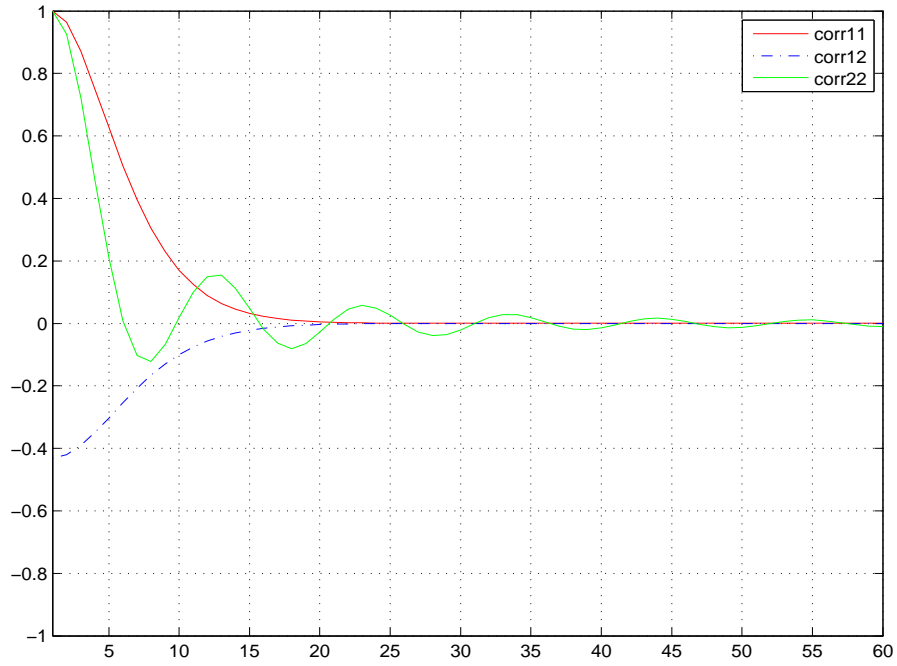


Figure 3: Correlations and cross-correlation functions for the bivariate random field with parameters given in Table 1. In this example the first noise process has a non-oscillating covariance function and the second noise process has an oscillating covariance function. We see that only the second random field has an oscillating covariance function.

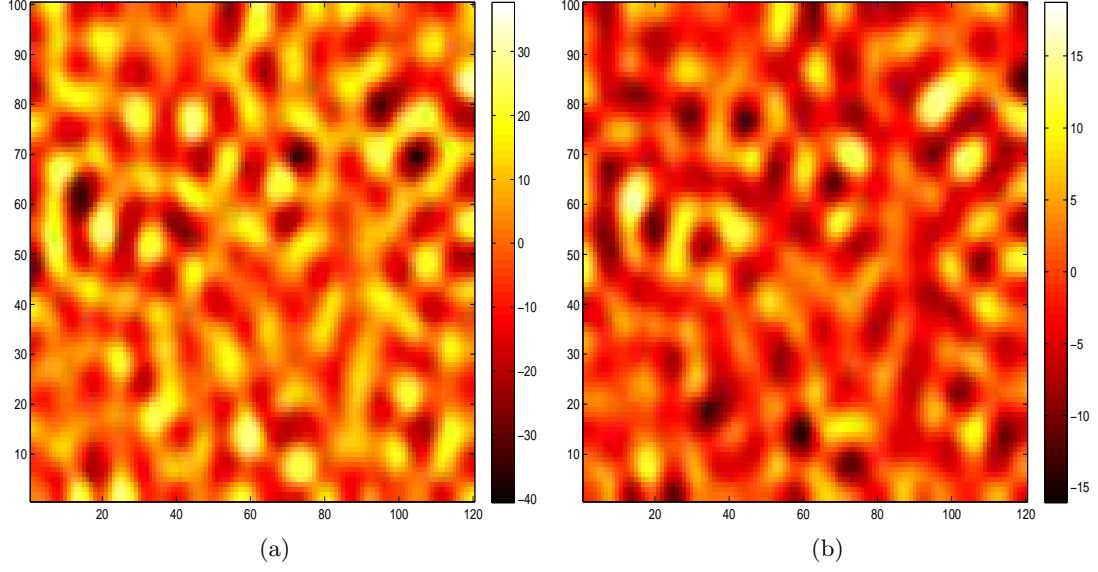


Figure 4: A realization of the bivariate random field with the parameters given in Table 1. In this example the first noise process has an oscillating covariance function and the second noise process has a non-oscillating covariance function.

where  $\pi(\boldsymbol{\theta})$  is the prior distribution of the hyper-parameters, and we return to this topic in Section 4.2,  $\pi(\mathbf{x}|\boldsymbol{\theta})$  is the density for the bivariate random fields,  $\pi(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta})$  is the density for the observations given the random field and the parameters and  $\pi(\mathbf{x}|\mathbf{y}, \boldsymbol{\theta})$  is the full conditional of the random fields given the observations and parameters.

Assume that there are  $N$  triangles in the domain for each of the random field  $\{x_i(\boldsymbol{\theta}); i = 1, 2\}$ . With the bivariate random fields  $\mathbf{x} = (x_1, x_2)^T$ ,  $2N$  triangles are used and hence the probability density of the bivariate random field has the form

$$\pi(\mathbf{x}|\boldsymbol{\theta}) = \left(\frac{1}{2\pi}\right)^{2N} |\mathbf{Q}(\boldsymbol{\theta})|^{1/2} \exp\left(-\frac{1}{2}\mathbf{x}^T \mathbf{Q}(\boldsymbol{\theta}) \mathbf{x}\right), \quad (33)$$

where  $\mathbf{Q}(\boldsymbol{\theta})$  is the precision matrix for the bivariate field. We assume that the length of the data is  $t = k_1 + k_2$ , where  $\{k_i; i = 1, 2\}$  are the length of the observations for each field. Then  $\pi(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta})$  has the form

$$\pi(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta}) = \left(\frac{1}{2\pi}\right)^t |\mathbf{Q}_n|^{1/2} \exp\left(-\frac{1}{2}(\mathbf{y} - \mathbf{A}\mathbf{x})^T \mathbf{Q}_n (\mathbf{y} - \mathbf{A}\mathbf{x})\right), \quad (34)$$

where  $\mathbf{Q}_n$  is the precision matrix for the measurement errors with dimension  $t \times 2k$ , and  $\mathbf{A}$  is a matrix with dimension  $t \times 2N$  that links the sparse observations to the dense random fields. One thing we want to point out is that the length of the observations for each field can be different and they are not necessarily observed at the same locations. We used the notation  $\pi(\mathbf{y}|\mathbf{x})$  instead of  $\pi(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta})$  since this function is independent of  $\boldsymbol{\theta}$ . The full conditional  $\pi(\mathbf{x}|\mathbf{y}, \boldsymbol{\theta})$  can be obtained,

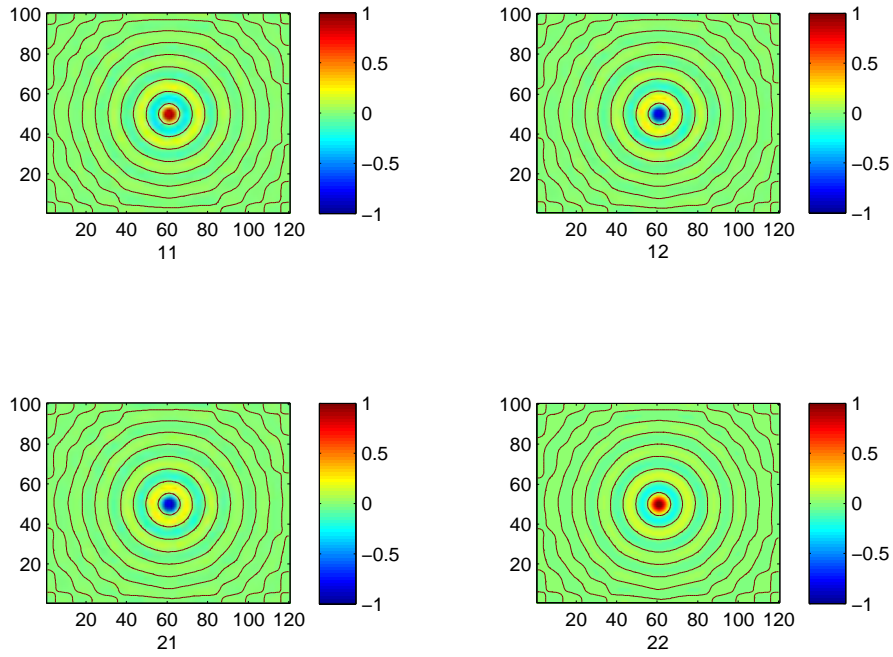


Figure 5: Correlations and cross-correlation functions for the bivariate random field with parameters given in Table 1. In this example the first noise process has an oscillating covariance function and the second noise process has a non-oscillating covariance function. We see that both the random fields have oscillating covariance functions.

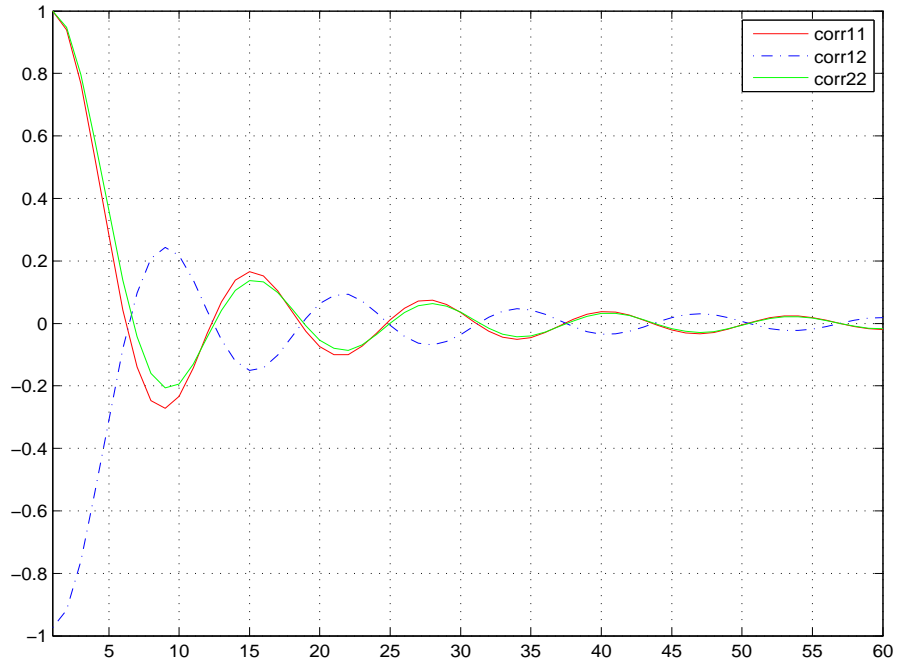


Figure 6: Correlations and cross-correlation functions for the bivariate random field with parameters given in Table 1. In this example the first noise process has an oscillating covariance function and the second noise process has a non-oscillating covariance function. We see that both the random fields have oscillating covariance functions.

$$\begin{aligned}
\pi(\mathbf{x}|\mathbf{y}, \boldsymbol{\theta}) &\propto \pi(\mathbf{x}, \mathbf{y}|\boldsymbol{\theta}) \\
&= \pi(\mathbf{x}|\boldsymbol{\theta})\pi(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta}) \\
&\propto \exp\left(-\frac{1}{2}\left[\mathbf{x}^T(\mathbf{Q}(\boldsymbol{\theta}) + \mathbf{A}^T\mathbf{Q}_n\mathbf{A})\mathbf{x} - 2\mathbf{x}^T\mathbf{A}^T\mathbf{Q}_n\mathbf{y}\right]\right).
\end{aligned} \tag{35}$$

Denote  $\boldsymbol{\mu}_c(\boldsymbol{\theta}) = \mathbf{Q}_c^{-1}(\boldsymbol{\theta})\mathbf{A}^T\mathbf{Q}_n\mathbf{y}$ , and  $\mathbf{Q}_c(\boldsymbol{\theta}) = \mathbf{Q}(\boldsymbol{\theta}) + \mathbf{A}^T\mathbf{Q}_n\mathbf{A}$ . Then we have

$$\mathbf{x}|\mathbf{y}, \boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{\mu}_c(\boldsymbol{\theta}), \mathbf{Q}_c^{-1}(\boldsymbol{\theta})). \tag{36}$$

Thus  $\mathbf{x}|\mathbf{y}, \boldsymbol{\theta}$  is a  $2N$ -dimensional multivariate Gaussian distribution. We can write (36) in the canonical form  $\mathbf{x}|\mathbf{y}, \boldsymbol{\theta} \sim \mathcal{N}_c(\mathbf{A}^T\mathbf{Q}_n\mathbf{y}, \mathbf{Q}_c(\boldsymbol{\theta}))$ . For more information about canonical form of the GMRFs, we refer to, for example, Rue and Held (2005, Chapter 2).

With a given prior  $\pi(\boldsymbol{\theta})$ , together with (33) to (36), the posterior distribution of  $\boldsymbol{\theta}$  becomes

$$\begin{aligned}
\pi(\boldsymbol{\theta}|\mathbf{y}) &\propto \pi(\boldsymbol{\theta}) \frac{|\mathbf{Q}(\boldsymbol{\theta})|^{1/2}|\mathbf{Q}_n|^{1/2}}{|\mathbf{Q}_c(\boldsymbol{\theta})|^{1/2}} \exp\left(-\frac{1}{2}\mathbf{x}^T\mathbf{Q}(\boldsymbol{\theta})\mathbf{x}\right) \\
&\times \exp\left(-\frac{1}{2}(\mathbf{y} - \mathbf{A}\mathbf{x})^T\mathbf{Q}_n(\boldsymbol{\theta})(\mathbf{y} - \mathbf{A}\mathbf{x})\right) \\
&\times \exp\left(\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_c(\boldsymbol{\theta}))^T\mathbf{Q}_c(\boldsymbol{\theta})(\mathbf{x} - \boldsymbol{\mu}_c(\boldsymbol{\theta}))\right).
\end{aligned} \tag{37}$$

And hence the logarithm of the posterior distribution is

$$\begin{aligned}
\log(\pi(\boldsymbol{\theta}|\mathbf{y})) &= \text{Const} + \log(\pi(\boldsymbol{\theta})) + \frac{1}{2}\log(|\mathbf{Q}(\boldsymbol{\theta})|) \\
&\quad - \frac{1}{2}\log(|\mathbf{Q}_c(\boldsymbol{\theta})|) + \frac{1}{2}\boldsymbol{\mu}_c(\boldsymbol{\theta})^T\mathbf{Q}_c(\boldsymbol{\theta})\boldsymbol{\mu}_c(\boldsymbol{\theta}).
\end{aligned} \tag{38}$$

## 4.2 Priors for the parameters

The prior distribution is important in Bayesian inference, and choosing the priors is an important part of inference. Two common approaches for choosing the prior distribution are the conjugate prior approach and the non-informative prior approach. There is no unique way for choosing priors. We refer to Robert (2007, Chapter 3) for detailed discussion about the prior information and prior distribution.

General speaking, it is hard to specify an informative prior for the hyper-parameters in our system of SPDEs approach. Therefore, the non-informative approach has been chosen. The following choice for the priors of the parameters are recommended with the bivariate random fields.

- $b_{11}$  and  $b_{22}$  should be positive values. So log-normal distributions are used for these two parameters. Gamma distribution can also be considered;

- Because of the requirement on the systems of SPDEs that  $\{h_{ij}; i, j = 1, 2\}$  and  $\{\kappa_{n_i}; i = 1, 2\}$  should be positive values, we can use log-normal or gamma distributions;
- $b_{21}$  is related to the sign of the correlation of the two random fields and it can be either positive or negative. Therefore, a Gaussian distribution can be used;
- The oscillation parameter  $\omega$  should fulfill the requirement  $\omega \in [0, 1]$  and hence a beta distribution can be used.

### 4.3 Inference with simulated data

Four simulated data examples are presented in this section to illustrate how to use our proposed approach. The datasets are divided into 2 groups. In the first group we use the correlated random fields given in Section 3.4. In the second group the fields are independent. We want our model to capture these features, and to return whether  $b_{21} = 0$  or not. However, if the first noise process is generated from the univariate SPDE given in Equation (2),  $\kappa_{n_1}^2$  and  $h_{11}$  are not identifiable. See Appendix B for more information. We use the setting  $\kappa_{n_1}^2 = h_{11}$  in this situation. It is our experience that  $\omega$  is likely to be in the range of  $(0.5, 1)$  if we have empirical knowledge that the random field has an oscillating covariance function, and hence we recommend to use a beta distribution with negative skew. In all of our simulated data examples, we use the following priors for the parameters (if they are needed to be estimated) following the discussion given in Section 4.2.

- $b_{11}, b_{22}, h_{11}, h_{22}, \kappa_{n_1}$  and  $\kappa_{n_2}$  have the log-normal distributions with  $\mu = 0$  and  $\sigma^2 = 100$ ;
- $b_{21}$  has a normal distribution with  $\mu = 0$  and  $\sigma^2 = 100$ ,  $b_{21} \sim \mathcal{N}(0, 100)$
- $\omega$  has a beta distribution with  $\alpha = 1$  and  $\beta = 1$ ,  $\omega \sim \text{Beta}(1, 1)$ , i.e., it is a uniform distribution.

The results for the first and second simulated datasets are given in Table 2 and Table 3, respectively. We can notice that the estimates are quite precise. Most of the true values are within 1 standard derivation away from the estimates. None of the true values are 2 standard deviations away from the estimates. The estimated conditional mean of the bivariate fields for these two datasets are given in Fig. 7 and Fig. 8. Compare with the true random fields given in Fig. 1(a) - Fig. 1(b) and Fig. 4(a) - Fig. 4(b). There is no large difference between them.

Similarly, the results for the second group are given in Table 4 and Table 5. In both examples, the estimates are precise and they are within 2 standard derivations from the true values. We can notice that if the fields are independent, i.e.,  $b_{21} = 0$ , our model captures this characteristic since  $b_{21}$  is small and 0 is within the 95% credible interval.



Table 2: Inference for the simulated dataset 1

Parameters	True values	Estimates	Standard deviations
$b_{11}$	0.5	0.495	0.013
$b_{21}$	0.25	0.248	0.017
$b_{22}$	1	1.027	0.032
$h_{11}$	0.25	0.248	0.010
$h_{22}$	0.36	0.355	0.029
$\kappa_{n_2}$	0.6	0.601	0.004
$\omega$	0.95	0.953	0.092

Table 3: Inference for the simulated dataset 2

Parameters	True values	Estimates	Standard deviations
$b_{11}$	0.5	0.497	0.014
$b_{21}$	0.25	0.234	0.012
$b_{22}$	1	0.964	0.029
$h_{11}$	0.25	0.269	0.024
$h_{22}$	0.36	0.339	0.022
$\kappa_{n_1}$	0.5	0.496	0.005
$\kappa_{n_2}$	0.6	0.636	0.049
$\omega$	0.95	0.956	0.113

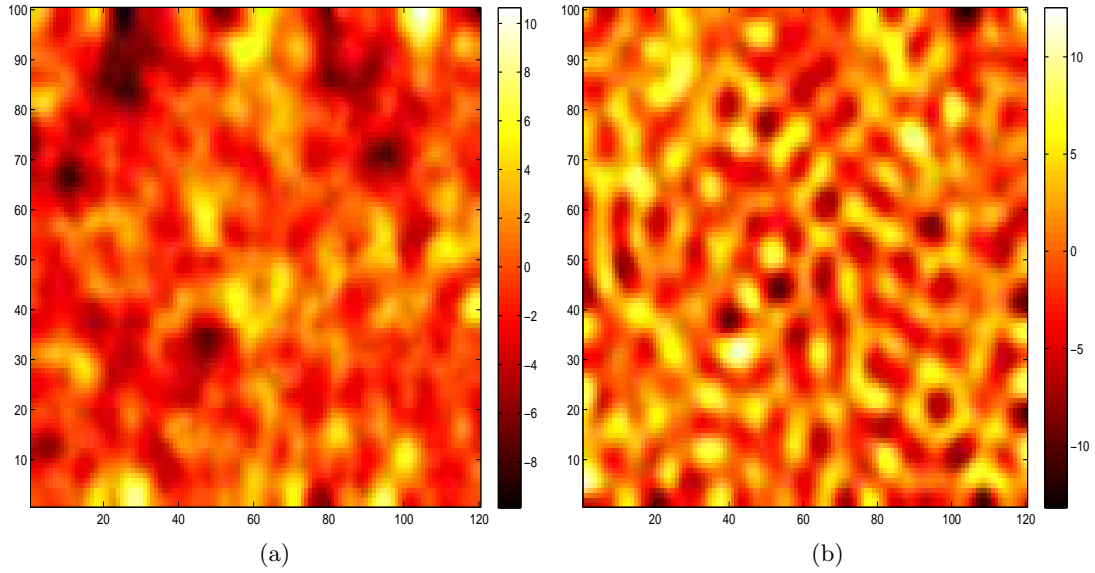


Figure 7: Estimated conditional mean of the bivariate random field for the dataset 1.

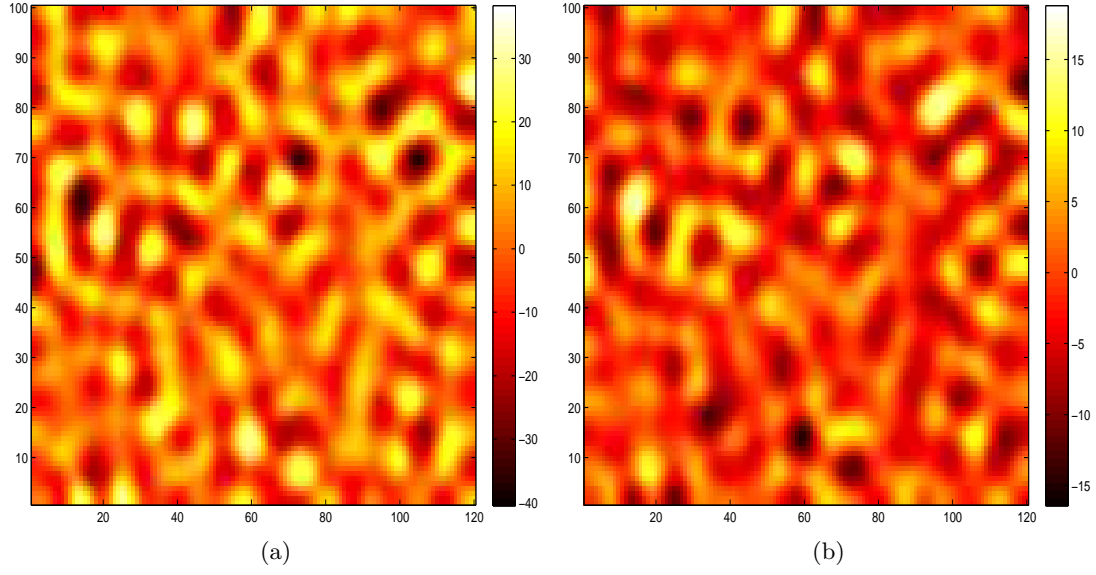


Figure 8: Estimated conditional mean of the bivariate random field for the dataset 2.

Table 4: Inference for the simulated dataset 3

Parameters	True value	Estimated	Standard deviations
$b_{11}$	0.5	0.491	0.012
$b_{21}$	0	0.012	0.010
$b_{22}$	0.3	0.301	0.010
$h_{11}$	0.25	0.247	0.009
$h_{22}$	0.36	0.374	0.033
$\kappa_{n_2}$	0.6	0.596	0.004
$\omega$	0.95	0.951	0.092

Table 5: Inference for the simulated dataset 4

Parameters	True value	Estimated	Standard deviations
$b_{11}$	0.5	0.487	0.015
$b_{21}$	0	0.001	0.002
$b_{22}$	0.3	0.308	0.009
$h_{11}$	0.25	0.284	0.026
$h_{22}$	0.36	0.359	0.122
$\kappa_{n_1}$	0.5	0.502	0.004
$\kappa_{n_2}$	0.6	0.599	0.102
$\omega$	0.95	0.949	0.107

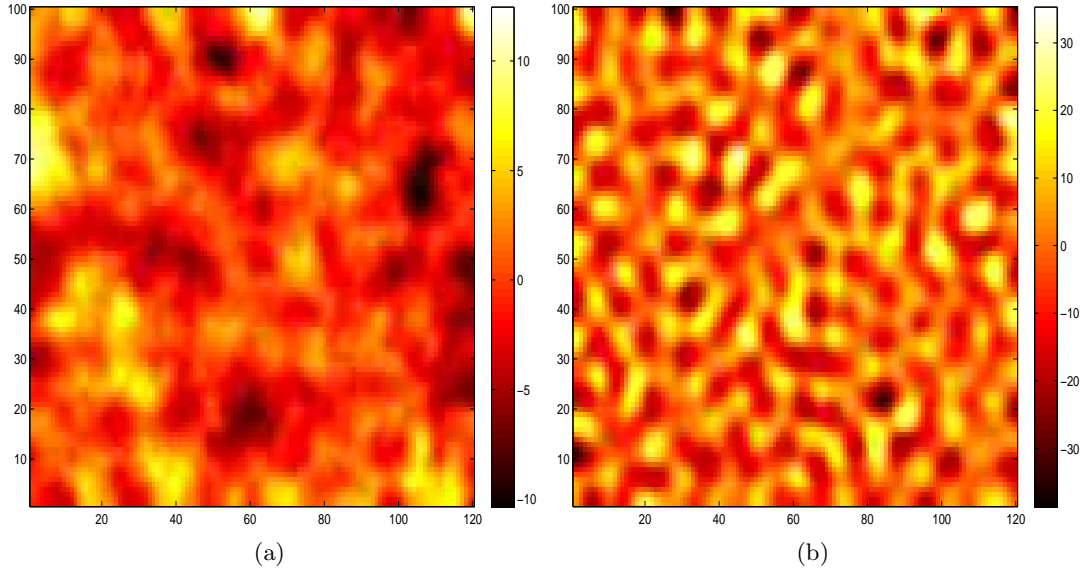


Figure 9: One realization of the bivariate random field with parameters given in Table 4. The first field random field has a non-oscillating covariance function and the second field has an oscillating covariance. The two random fields are independent.

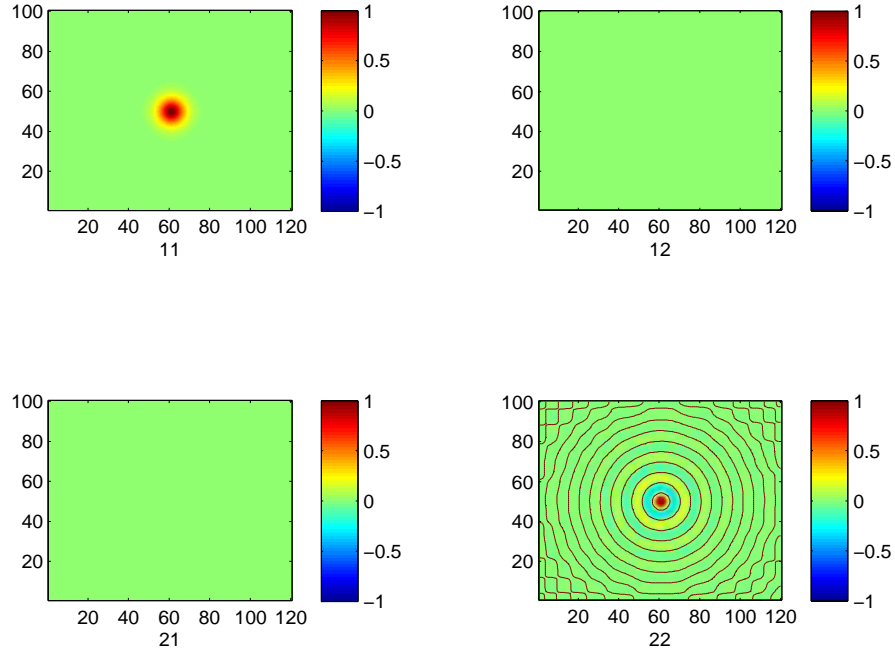


Figure 10: The correlation and cross correlation functions with parameters given in Table 4. The first field random field has a non-oscillating covariance function and the second field has an oscillating covariance. The two random fields are independent.

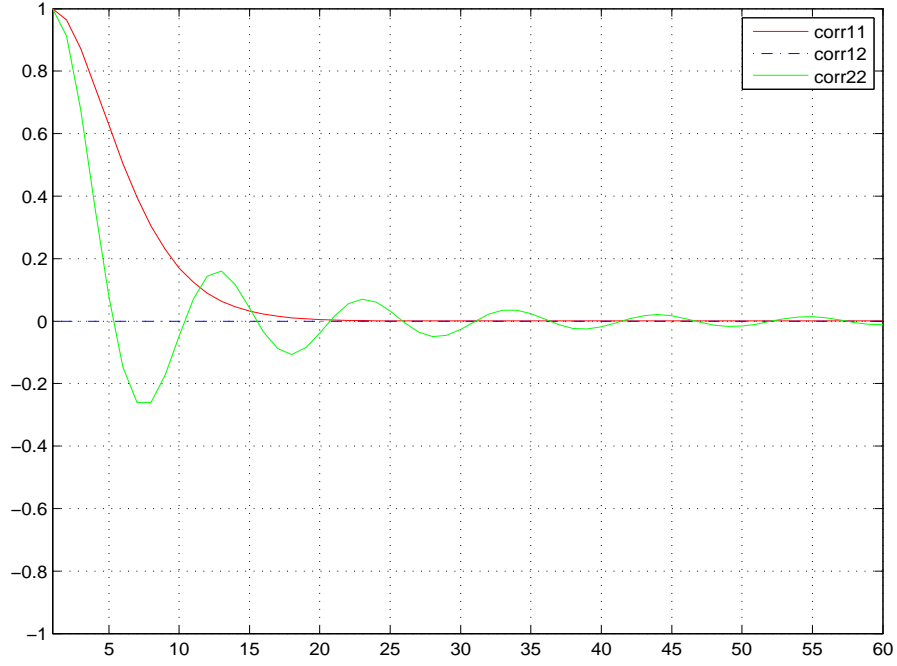


Figure 11: The correlation and cross correlation functions with parameters given in Table 4. The first field random field has a non-oscillating covariance function and the second field has an oscillating covariance. The two random fields are independent.

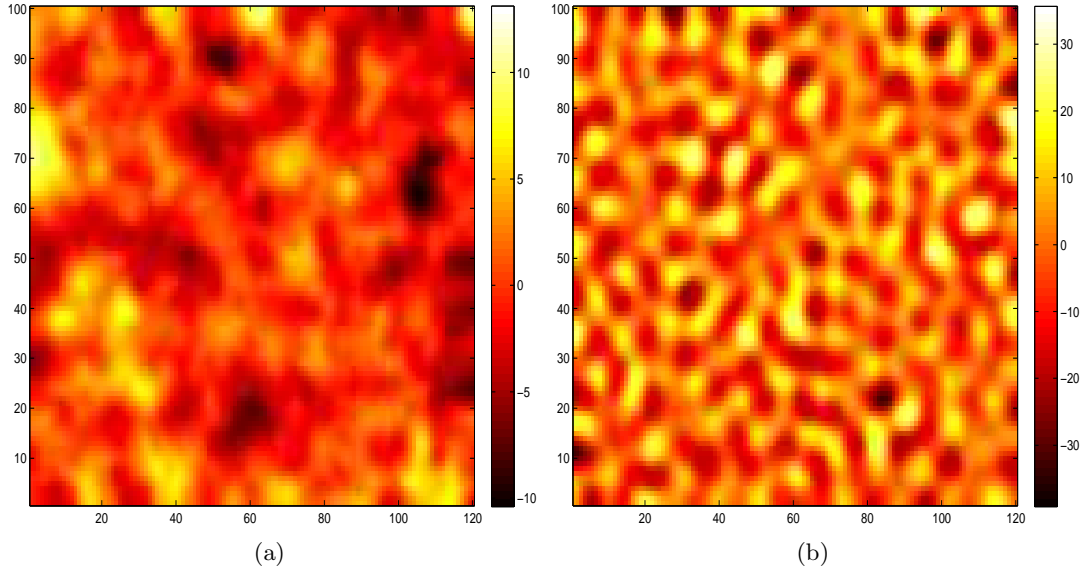


Figure 12: Estimated conditional mean of the bivariate random fields for dataset 3.

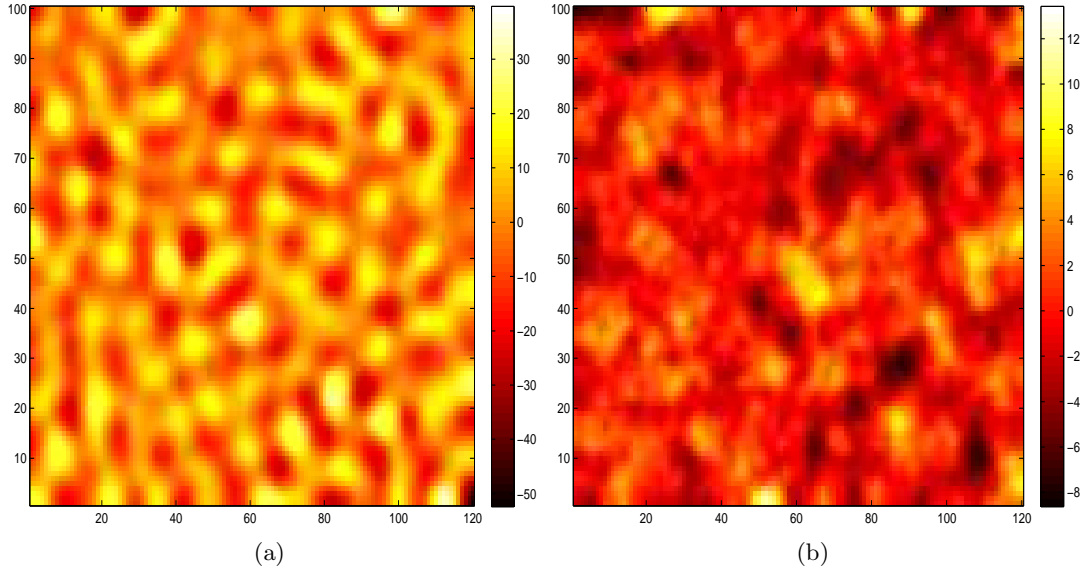


Figure 13: One realization of the bivariate random fields with parameters given in Table 5. The first random field has an oscillating covariance function and the second field has a non-oscillating covariance. The two random fields are independent.

#### 4.4 Inference with real data

In this section a dataset has been chosen in order to illustrate how to use our approach in real-world applications. This dataset is from the ERA 40 database and can be downloaded from the ERA 40 project homepage. The dataset contains the temperature and pressure on the whole globe on 4th of September, 2002. The main objective for this section is to illustrate how to use our model for a big dataset. All the results are only from the prediction point of view. The dataset contains 10368 observations both for temperature and pressure, and the observations are on the grid. The grid is constructed with the latitude and longitude. The latitudes are from  $90^\circ$  to  $-90^\circ$  and longitudes are from  $0^\circ$  to  $357.5^\circ$ , with increments of  $2.5^\circ$  for both axes. The dataset contains the temperatures in Kelvin and the mean sea level pressure in Pascal. We have subtracted the monthly mean for the temperature and pressure, respectively.

Since the dataset is on the entire globe, we need to construct our model on the sphere. Jones (1963) discussed how to construct stochastic processes on a sphere using the spectral representations for spherically symmetric and the axially symmetric cases. Another approach is to consider the sphere as a surface in  $\mathbb{R}^3$ . However, this has the disadvantage that the correlation between points are determined by the chordal distances (Lindgren et al., 2011). Gneiting (1998) pointed out that the random fields constructed on the plane were not suitable for this kind of dataset since the great circle distances in the original covariance function would not work in general. Jun and Stein (2007) discussed an approach for constructing space-time covariance functions on spheres using a sum of

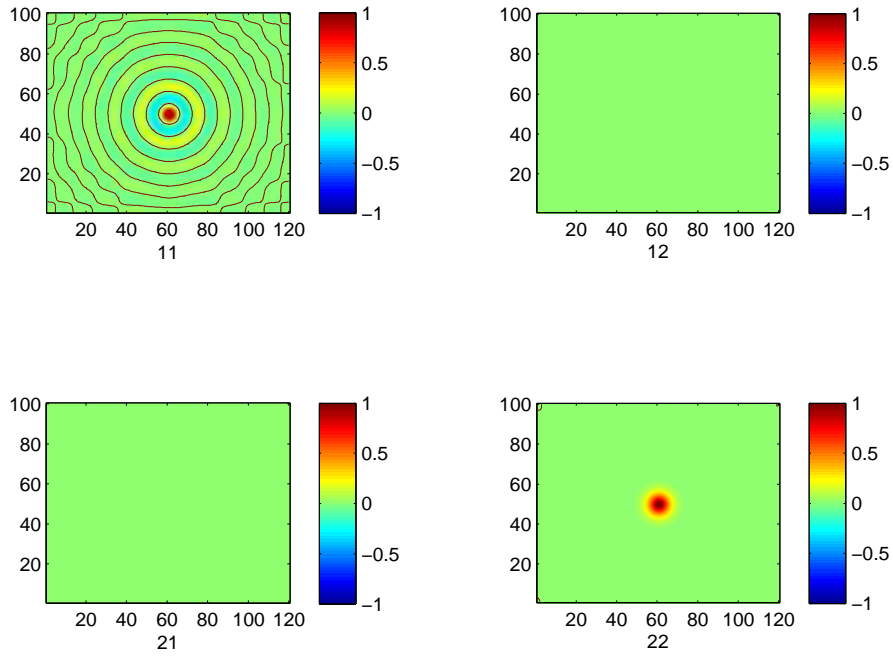


Figure 14: The correlation functions and cross correlation functions with parameters given in Table 5. The field random field has an oscillating covariance function and the second field has a non-oscillating covariance. The two random fields are independent.

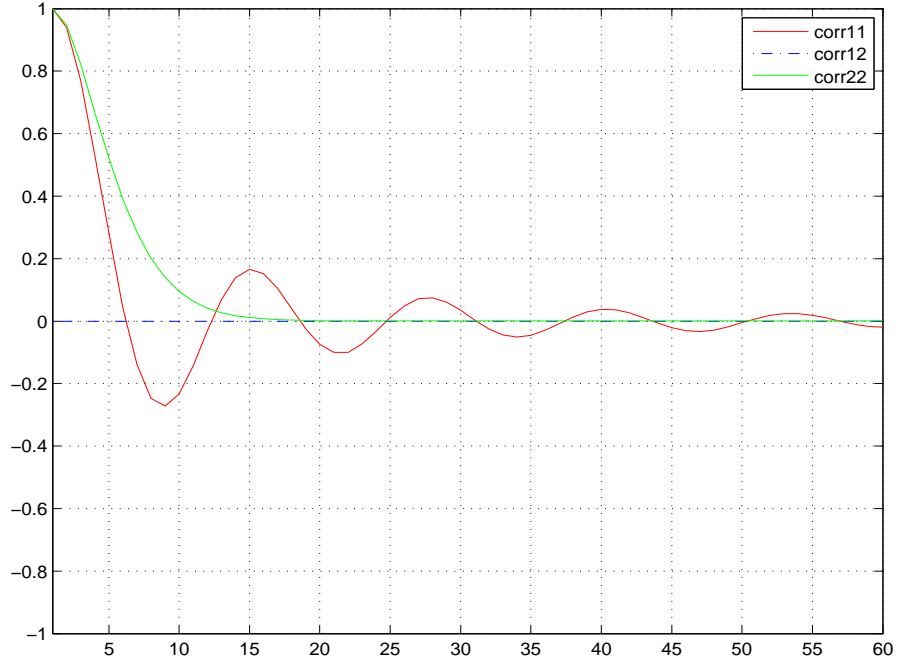


Figure 15: The correlation functions and cross correlation functions with parameters given in Table 5. The field random field has an oscillating covariance function and the second field has a non-oscillating covariance. The two random fields are independent.

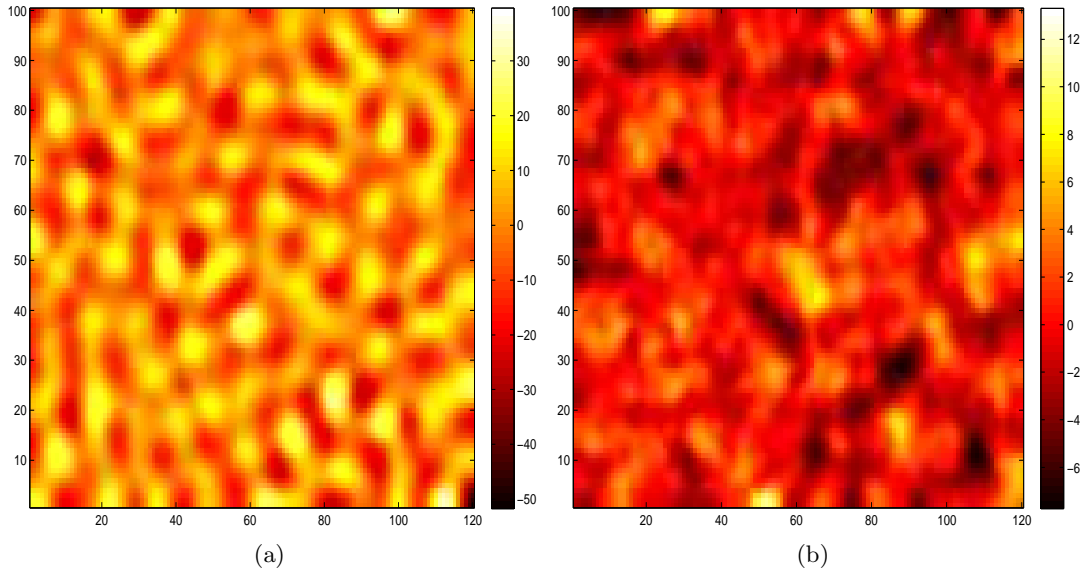


Figure 16: Estimated conditional mean of the bivariate random fields for dataset 4.

Table 6: Inference for real dataset

Parameters	Estimated
$b_{11}$	$8.952 \times 10^5$
$b_{21}$	1.546
$b_{22}$	$4.714 \times 10^1$
$h_{11}$	$1.224 \times 10^{-6}$
$h_{22}$	$8.089 \times 10^{-6}$
$\kappa_{n_2}$	$1.013 \times 10^{-3}$
$\omega$	0.819

independent processes. The main idea for their approach is to sum independent processes where each process is obtained by applying the first-order differential operations to a fully symmetric processes on sphere  $\times$  time. We refer to Jun and Stein (2007) for more information on the fully symmetric processes on sphere  $\times$  time.

In this paper we follow the approach discussed by Lindgren et al. (2011) to construct the GRFs on the sphere. They claimed that using the SPDE approach for constructing GRFs on the sphere is similar to constructing the GRFs on  $\mathbb{R}^d$ . By reinterpreting the SPDE defined on  $\mathbb{S}^2$ , the solutions of the SPDE are GRFs defined on  $\mathbb{S}^2$ . Our proposed system of SPDEs approach inherits this property. The only place has been changed is that the system of SPDEs is directly defined on  $\mathbb{S}^2$ . Another advantage of our approach is that the GMRF approximation can still be used. In other words, we can use GMRFs to represent GRFs for computation. For more information about GRFs on manifolds, we refer to Lindgren et al. (2011, Section 3.1).

Since it is known that the pressure on the globe has an oscillating covariance function, it is reasonable to set  $x_1(\mathbf{s})$  as the temperature and  $x_2(\mathbf{s})$  as the pressure, and let the second noise process  $\varepsilon_2(\mathbf{s})$  have an oscillating covariance function, but not the first noise process  $\varepsilon_1(\mathbf{s})$ . The original dataset is shown in Fig. 17(a) and Fig. 17(b), and the reconstructed temperature and pressure are shown in Fig. 18(a) and Fig. 18(b). We also give 3D images for the true datasets on the sphere in Fig. 19(a) and Fig. 19(b), and the reconstructed temperature and pressure on the sphere in Fig. 20(a) and Fig. 20(b). One thing we want to point out is that we follow the methodology given in Lindgren et al. (2011) and construct the GRF on the unit radius sphere  $\mathbb{S}^2$ . Another important point is that we set  $\kappa_{n_1}^2 = h_{11}$  to simplify the model and inference.

In order to check the predictive performance of our approach, we have divided the dataset into two subsets. We used a subset containing 5368 observations for both temperature and pressure for estimating the parameters and predict the remaining 5000 observations. The estimates are given in Table 6. From the results we notice that the model captures the empirical knowledge that the temperature and pressure are negative correlated since  $b_{21} > 0$ . The prediction for the 5000 observations are given in Fig. 21. From a prediction point of view, the model works well since most of the prediction are close to the true observed values. The correlation functions are given in Fig. 22, and we notice that the covariance function of pressure indeed has oscillation behavior.



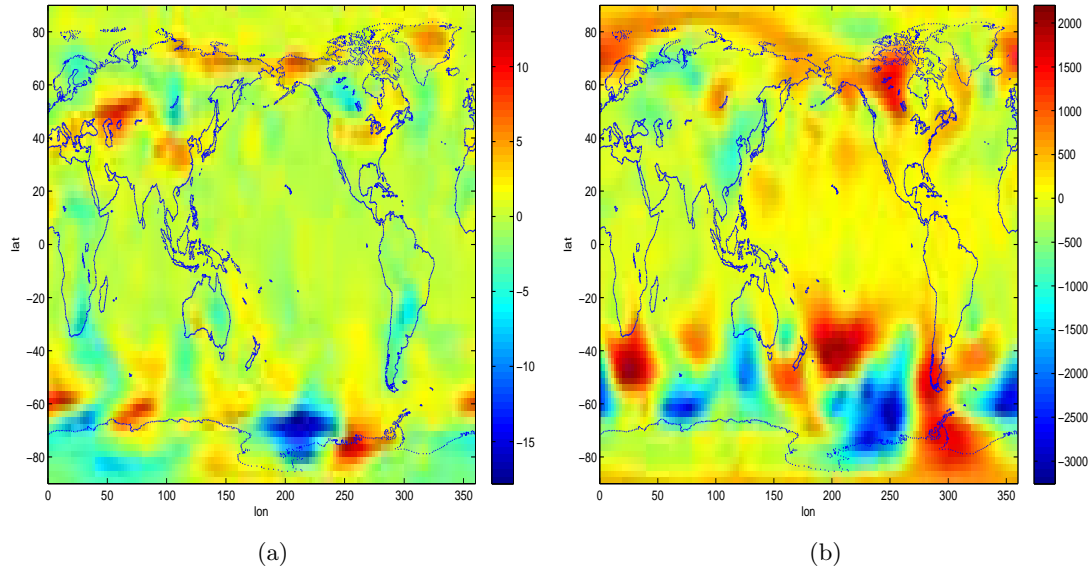


Figure 17: Dataset with temperature (a) and pressure (b) from ERA 40 database.

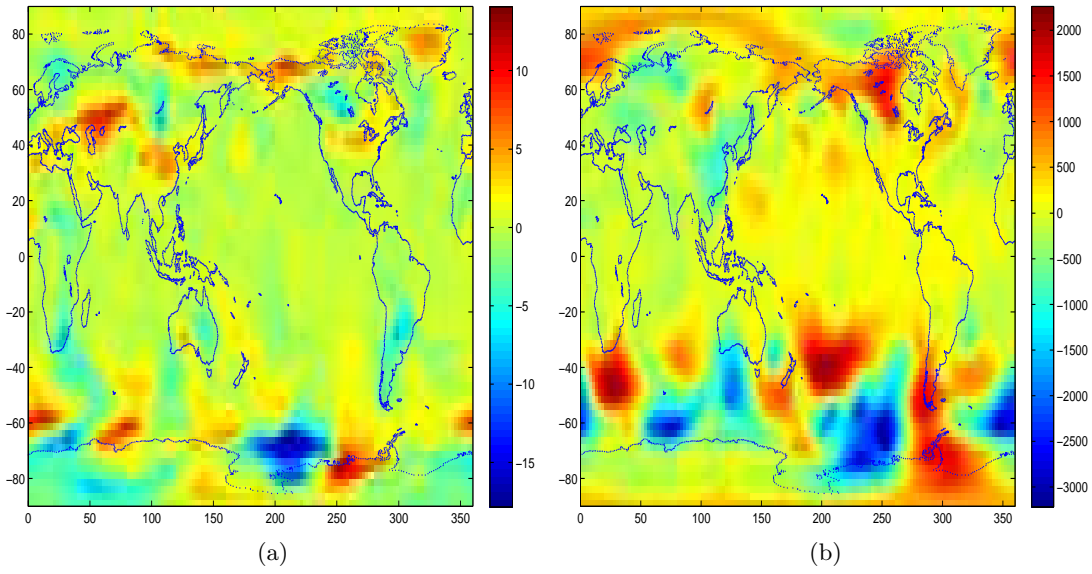


Figure 18: Reconstructed bivariate random fields for temperature (a) and pressure (b).

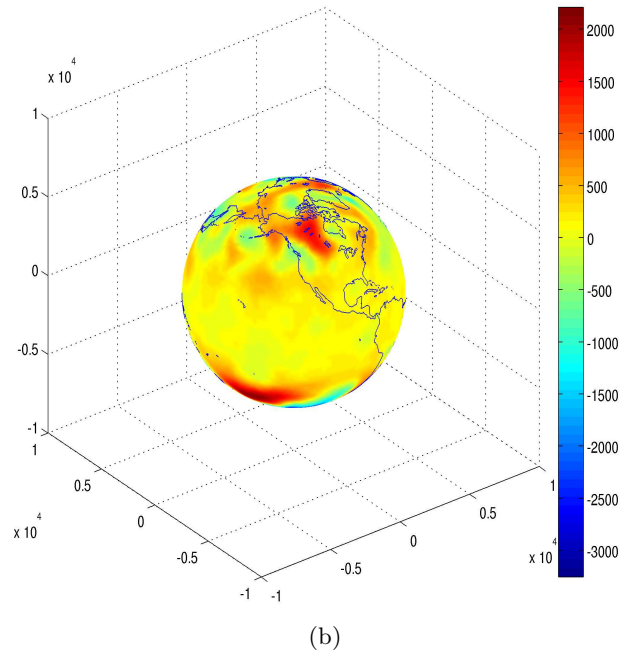
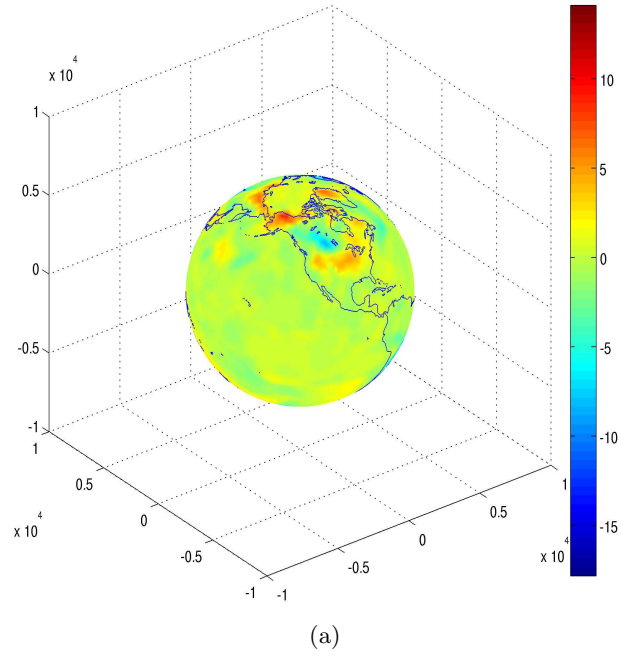
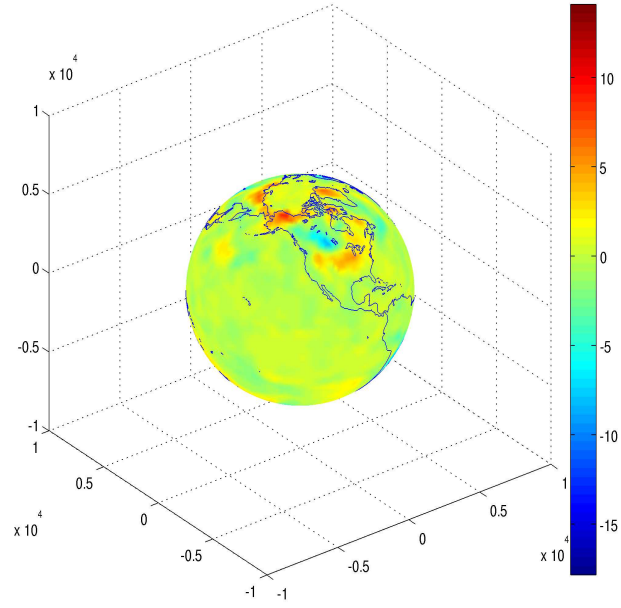
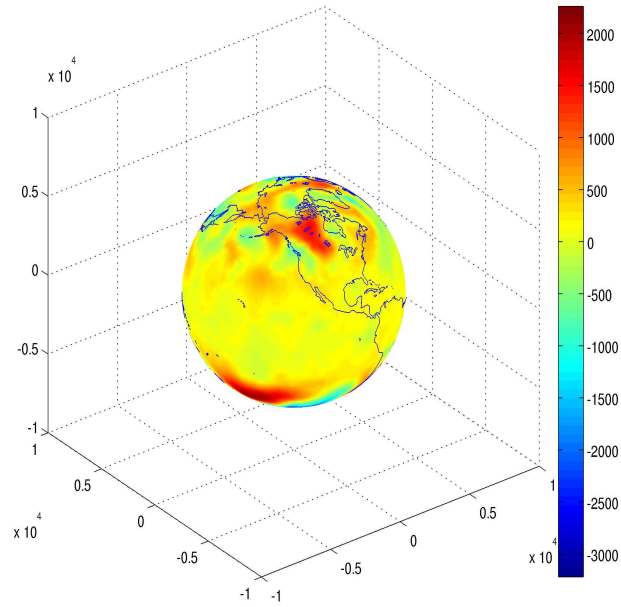


Figure 19: Dataset with temperature (a) and pressure (b) on the sphere from ERA 40 database.



(a)



(b)

Figure 20: Reconstructed bivariate random fields for temperature (a) and pressure (b) on the sphere.

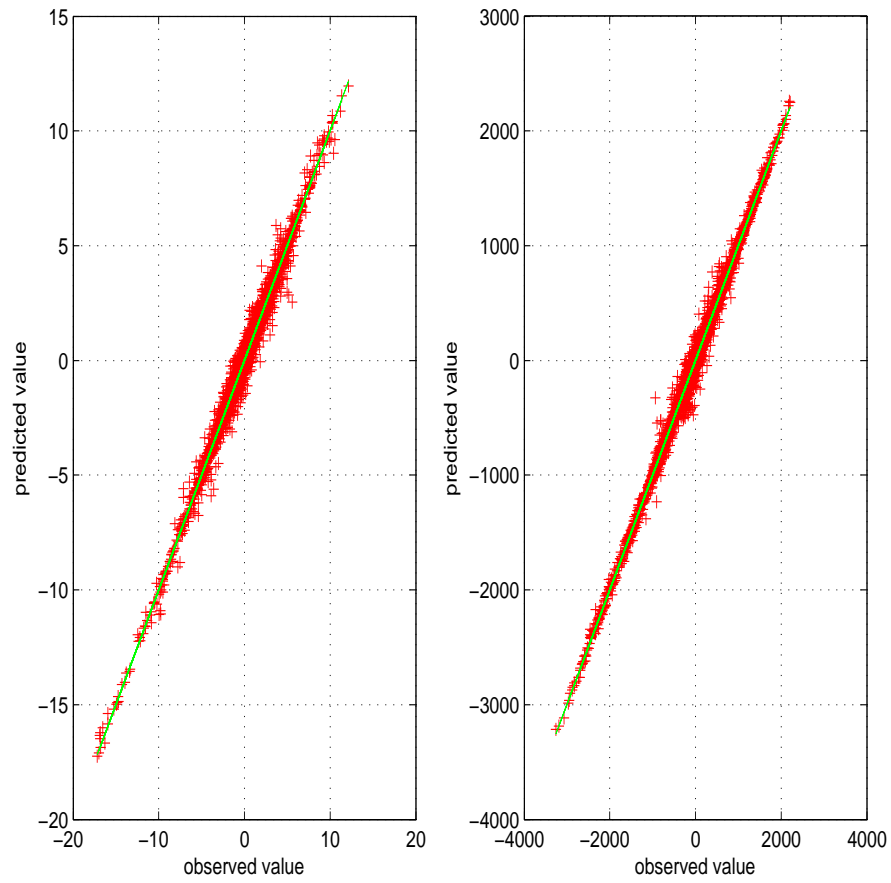


Figure 21: Prediction for the leaving out 5000 observations for temperature (left) and pressure (right)

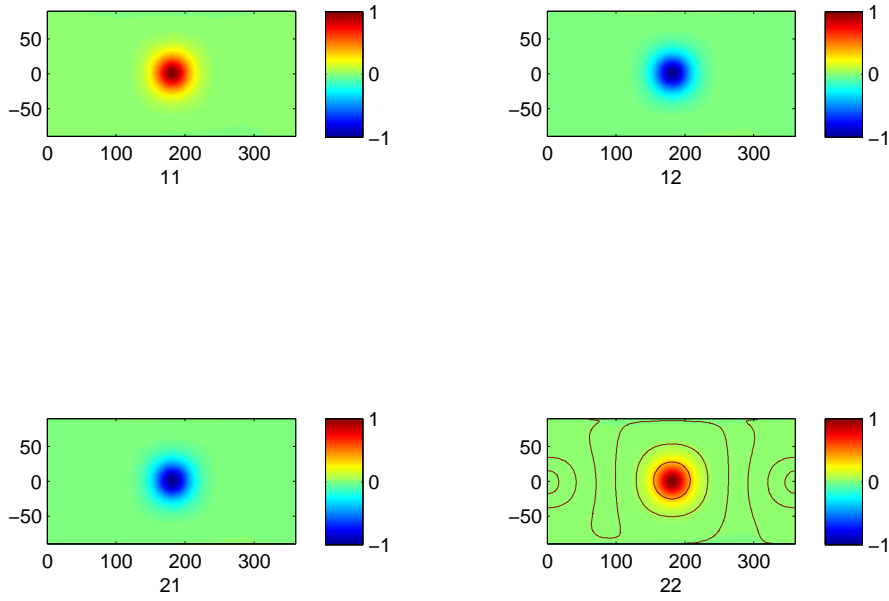


Figure 22: Correlation functions for the bivariate random fields. 11 indicates the marginal correlation of the temperature. 12 and 21 indicate the cross-correlation between the temperature and pressure. 22 indicates the marginal correlation of the pressure.

## 5 Discussion and future work

Due to the increasing importance of spatial statistics in applications, new approaches for handling different complex datasets are demanded. The methodologies for dealing with multivariate datasets appear in many areas, such as in air quality (Brown et al., 1994; Schmidt and Gelfand, 2003), weather forecasting (Courtier et al., 1998; Reich and Fuentes, 2007), and economics (Gelfand et al., 2004; Sain and Cressie, 2007). Two of the most important issues with these methodologies are how to handle large datasets and how to ensure the nonnegative definiteness constraint for the covariance function. Gneiting et al. (2010) gave some theorems in order to construct valid covariance functions for multivariate random fields. In their approach every component in the matrix-valued covariance function was a Matérn covariance function. Hu et al. (2012a) proposed to use the systems of SPDEs to construct multivariate GRFs with isotropic and non-oscillating covariance functions. The summary paper by Sun et al. (2012) discussed the approaches for how to handle large datasets. They discussed several approaches such as separable covariance structures (Genton, 2007; Fuentes, 2006), covariance tapering (Furrer et al., 2006; Zhang and Du, 2008), likelihood approximations (Vecchia, 1988; Stein et al., 2004), fixed rank kriging and fixed rank filtering (Cressie and Johannesson, 2008; Cressie et al., 2010) and Gaussian Markov random fields approximation (Rue and Tjelmeland, 2002; Rue et al., 2004; Rue and Held, 2005; Rue et al., 2009; Lindgren et al., 2011).

This paper is an extension of Lindgren et al. (2011) and Hu et al. (2012a). The

main contribution of this paper is the proposed approach for constructing multivariate random fields with oscillating covariance functions using systems of SPDEs. The main idea is to use noise processes with oscillating covariance functions in order to introduce oscillation in the covariance functions of the random fields. We recommend to use the triangular systems of SPDEs since these models have some advantages. For instance, we have fewer hyper-parameters and we can locate which random fields have non-oscillating, oscillating, possibly oscillating covariance functions. This approach can construct many models discussed by Hu et al. (2012a) if we set the oscillation parameter  $\omega = 0$ . It also inherits most of the advantages of the SPDE approach discussed by Lindgren et al. (2011) and systems of SPDEs approach discussed by Hu et al. (2012a).

The two main challenges in multivariate random fields mentioned above can be partially solved with our model. On the theoretical side, the covariance functions of the multivariate random fields fulfill the nonnegative definite constraint automatically. On the computational side, the GMRF representation makes the precision matrices to be sparse. Thus numerical algorithms for sparse matrices can be used for fast sampling and inference. Four simulated datasets and one real dataset have been used to illustrate how to use our approach in different situations. The results have illustrated the effectiveness of the proposed approach.

There are several possible extensions for further research, such as constructing non-stationary multivariate GRFs from the systems of SPDEs, and spatio-temporal models both in  $\mathbb{R}^d$  and on manifolds. It should also be possible to use the integrated nested Laplace approximation (INLA) framework (Rue et al., 2009) for doing inference for the multivariate GRFs. More applied work using the proposed approach is under development.

## Appendix A

There are different kinds of parametrization for the system of SPDEs. The main idea is to change the operators in (22) and use a different parametrization as discussed in Section 3.3. Two intuitive operator matrices are

$$\mathcal{L}_2(\boldsymbol{\theta}) = \begin{pmatrix} b_{11}(h_{11} - \Delta) & 0 \\ b_{21}(h_{21} - \Delta) & b_{22}(h_{22} - \Delta) \end{pmatrix}, \quad (39)$$

$$\mathcal{L}_3(\boldsymbol{\theta}) = \begin{pmatrix} b_{11}(h_{11} - \Delta) & 0 \\ b_{21} & b_{22} \end{pmatrix}. \quad (40)$$

With (39) and (40) the correlations between the fields will be changed. The operator matrix given in (39) introduces more flexibility since we have one more parameter  $h_{21}$  to control the range of cross-correlation. However, it might be hard to estimate all the parameters in this case. With operator matrix given in (40), we have fewer parameters in the model, but the correlation structure between the fields is simplified. The second random field has the same correlation range as the first field. These two systems can be used in different applications.

## Appendix B

When we use the triangular systems of SPDEs, we need to set the constraint  $\kappa_{n_1}^2 = h_{11}$  since when the first noise process is generated from Equation (2), they are not identifiable together. Write the system of equations for the bivariate random field given in (30) explicitly as

$$\begin{aligned} b_{11}(h_{11} - \Delta)x_1(\mathbf{s}) &= \varepsilon_1(\mathbf{s}), \\ b_{21}x_1(\mathbf{s}) + b_{22}(h_{22} - \Delta)x_2(\mathbf{s}) &= \varepsilon_2(\mathbf{s}). \end{aligned}$$

Assume the first noise process  $\varepsilon_1(\mathbf{s})$  is generated by

$$(\kappa_{n_1}^2 - \Delta)\varepsilon_1(\mathbf{s}) = \mathcal{W}(\mathbf{s}). \quad (41)$$

We can now rewrite the first equation in system (30) as a system of equations

$$\begin{aligned} b_{11}(h_{11} - \Delta)x_1(\mathbf{s}) &= \varepsilon_1(\mathbf{s}), \\ (\kappa_{n_1}^2 - \Delta)\varepsilon_1(\mathbf{s}) &= \mathcal{W}(\mathbf{s}). \end{aligned} \quad (42)$$

This system of equations can be rewritten into one equation with white noise as the driving process,

$$(h_{11} - \Delta)(\kappa_{n_1}^2 - \Delta)x_1(\mathbf{s}) = \mathcal{W}(\mathbf{s}). \quad (43)$$

It is obvious that  $\kappa_{n_1}^2$  and  $h_{11}$  are not identifiable from each other since  $(h_{11} - \Delta)$  and  $(\kappa_{n_1}^2 - \Delta)$  commute. Therefore, we suggest the constraint  $\kappa_{n_1}^2 = h_{11}$ . However, if the first noise process is oscillating and is generated from Equation (15), we don't need this constraint because they are identifiable. However, we still recommend to use this setting in order to simplify the inference.

## References

- S. Banerjee, B.P. Carlin, and A.E. Gelfand. *Hierarchical modeling and analysis for spatial data*. Chapman & Hall, 2004. ISBN 158488410X.
- 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.
- S.C. Brenner and L.R. Scott. *The mathematical theory of finite element methods*, volume 15. Springer Verlag, 2008.
- 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, J. Pailleux, D. Vasiljevic, M. Hamrud, A. Hollingsworth, F. Rabier, and M. Fisher. The ecmwf implementation of three-dimensional variational assimilation (3d-var). i: Formulation. *Quarterly Journal of the Royal Meteorological Society*, 124:1783–1807, 1998.

- N. Cressie and G. Johannesson. Fixed rank kriging for very large spatial data sets. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 70(1):209–226, 2008.
- N. Cressie, T. Shi, and E.L. Kang. Fixed rank filtering for spatio-temporal data. *Journal of Computational and Graphical Statistics*, 19(3):724–745, 2010.
- 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.
- M. Fuentes. Testing for separability of spatial-temporal covariance functions. *Journal of Statistical Planning and Inference*, 136(2):447–466, 2006.
- G.A. Fuglstad. Approximating solutions of stochastic differential equations with gaussian markov random fields. Technical report, Department of Mathematical Science, Norwegian University of Science and Technology, 2010.
- 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.
- R. Furrer, M.G. Genton, and D. Nychka. Covariance tapering for interpolation of large spatial datasets. *Journal of Computational and Graphical Statistics*, 15(3):502–523, 2006.
- 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.
- A.E. Gelfand, P.J. Diggle, M. Fuentes, and P. Guttorp. *Handbook of spatial statistics*. CRC Press, 2010.
- M.G. Genton. Separable approximations of space-time covariance matrices. *Environmetrics*, 18(7):681–695, 2007.
- T. Gneiting. Simple tests for the validity of correlation function models on the circle. *Statistics & probability letters*, 39(2):119–122, 1998.
- 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.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. Hjelle and M. Dæhlen. *Triangulations and applications*. Springer Verlag, 2006.
- X. Hu, D.P. Simpson, F. Lindgren, and H. Rue. Multivariate gaussian random fields using systems of stochastic partial differential equations. *statistical report, Norwegian University of Science and Technology*, 2012a.
- X. Hu, D.P. Simpson, and H. Rue. Specifying gaussian markov random fields with incomplete orthogonal factorization using givens rotations. Technical report, Department of Mathematical Science, norwegian University of Science and Technology, 2012b.



- R.H. Jones. Stochastic processes on a sphere. *The Annals of mathematical statistics*, 34(1): 213–218, 1963.
- M. Jun and M.L. Stein. An approach to producing space–time covariance functions on spheres. *Technometrics*, 49(4):468–479, 2007.
- C.G. Kaufman, M.J. Schervish, and D.W. Nychka. Covariance tapering for likelihood-based estimation in large spatial data sets. *Journal of the American Statistical Association*, 103 (484):1545–1555, 2008.
- 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.
- G. Lindgren. *A second course on stationary stochastic processes*. Center for Mathematical Sciences, Lund University, December 2010.
- 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.
- C. Robert. *The Bayesian choice: from decision-theoretic foundations to computational implementation*. Springer Verlag, 2007.
- 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, I. Steinsland, and S. Erland. Approximating hidden Gaussian Markov random fields. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66(4):877–892, 2004. ISSN 1467-9868.
- 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.
- B. Shaby and D. Ruppert. Tapered covariance: Bayesian estimation and asymptotics. *Journal of Computational and Graphical Statistics*, 21(2):433–452, 2012.
- M.L. Stein. *Interpolation of Spatial Data: some theory for kriging*. Springer Verlag, 1999. ISBN 0387986294.
- M.L. Stein, Z. Chi, and L.J. Welty. Approximating likelihoods for large spatial data sets. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66(2):275–296, 2004.
- Y. Sun, B. Li, and M.G. Genton. Geostatistics for large datasets. *Advances and challenges in space-time modelling of natural events*, pages 55–77, 2012.

- A.V. Vecchia. Estimation and model identification for continuous spatial processes. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 297–312, 1988.
- W.W.S. Wei. *Time series analysis: univariate and multivariate methods*. Addison-Wesley, 2006.
- 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.
- H. Zhang and J. Du. Covariance tapering in spatial statistics. *Positive definite functions: From Schoenberg to space-time challenges*, pages 181–196, 2008.
- 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.