Spatial modeling with system of stochastic partial differential equations

Xiangping Hu[1] and Ingelin Steinsland[2]*

To define a spatial process as the solution to a stochastic partial differential equation (SPDE) is an approach to spatial modeling that is gaining popularity. The model corresponds to a Gaussian random spatial process with Matérn covariance function. The SPDE approach allows for computational benefits and provides a framework for making valid complex models (e.g., nonstationary spatial models). Using systems of SPDEs to define spatial processes extends the class of models that can be specified as SPDEs, while the computational benefits are kept. In this study, we give an overview of the current state of spatial modeling with systems of SPDEs. Systems of SPDEs have contributed toward modeling and computational efficient inference for spatial Gaussian random field (GRF) models with oscillating covariance functions and multivariate GRF models. For multivariate GRF models special systems of SPDEs corresponding to models known from the literature are set up. Little work has been done for exploring opportunities and properties of spatial processes defined as systems of SPDEs. We also describe some of the interesting topics for further research.

How to cite this article:
WIREs Comput Stat 2016, 8:112–125. doi: 10.1002/wics.1378

Keywords: Stochastic partial differential equations, Gaussian random fields, Spatial models, Covariance function

INTRODUCTION

The foundation of spatial statistics can be summarized in Tobler's 'First Law of Geography': Things that are nearby tend to be more alike than things that are far apart.[1] Air temperature and air pressure are two examples of spatial phenomena where it is reasonable to assume that locations that are close, are more similar than locations that are further apart. When modeling and doing inference about a spatial phenomena we want to account for, utilize, and estimate this dependency related to geographical distance. We might need to estimate the temperature at a location without observations based on observations from nearby weather stations. This can be done by utilizing the spatial dependence, and we need a spatial model. A simple spatial model considers one quantity (e.g., temperature at noon today), and assumes that the dependency decreases monotonically with distance, and can be specified as a function of the absolute value of the distance (stationary and isotropic).[2] For many applications, we want to relax these assumptions and/or extend to more quantities. For example, when modeling temperature it is reasonable that the dependency in the dominating wind direction is stronger than perpendicular to it (anisotropic model), or that dependency is stronger over oceans than over land (nonstationary). Furthermore, pressure on a global scale is known from physics to be oscillating, which calls for a dependency model with a oscillating feature. In

*Correspondence to: ingelins@math.ntnu.no

[1]Department of Mathematics, University of Oslo, Oslo, Norway

[2]Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim, Norway

Conflict of interest: The authors have declared no conflicts of interest for this article.

many situations, two or more spatial quantities are of interest simultaneously. For example, temperature and precipitation are important quantities for spring flood predictions,[3,4] and for winter road maintenance.[5,6] A multivariate spatial model is needed whenever the spatial phenomena are dependent. Another common situation is that a time-series of the spatial process is of interest. For example, precipitation over time causes large floods. When there are both temporal and spatial dependency, we need a spatiotemporal model (also known as a space-time model).

Developing spatial models and methods with more flexible dependency structure (e.g., anisotropic, nonstationary, and oscillation) as well as for multivariate spatial phenomena and spatiotemporal phenomena have been areas of active research the last two decades. The main challenges can be summarized in two stages: (1) specify spatial models that are flexible, valid and identifiable from available observations and (2) find methods for inference that are computationally feasible.

Gaussian processes are popular in spatial statistics, either as a model for the phenomenon of interest directly, or as part of a hierarchical model.[2,7] A Matérn field is a special case of Gaussian spatial processes. Matérn models are popular, and Stein concluded 'use Matérn'[8] due to their predictive properties. It is known that the stationary solution x ($\mathbf{s}$) to a stochastic partial differential equation (SPDE):

$$\mathcal{L}x(\mathbf{s}) = \omega(\mathbf{s}), \tag{1}$$

is a Matérn field.[9,10] In Eq. (1), $\omega(\mathbf{s})$ is spatial Gaussian white noise and $\mathcal{L}$ is a differential operator (advanced second derivative) to be defined in next Section. A Gaussian Markov random field (GMRF) representation can be achieved by approximating the Gaussian field $x(\mathbf{s})$ using piecewise linear basis functions.[11] The Markov structure causes a sparse structure of the matrices that need to be considered, which gives computational efficiency.[12] The computational cost depends on the number of linear basis functions used, and hence it can be further reduced by using fewer basis functions (e.g., by letting each piecewise linear part be larger). Hence, the SPDE methodology can be used for approximate inference for large spatial datasets.[13]

Furthermore, the integrated nested Laplace approximation (INLA)[14] methodology can be used for efficient Bayesian inference for many Bayesian hierarchical models with a Gaussian process formulated as an SPDE. The SPDE formulation also provides frameworks for extending the Matérn model. The formulation in Eq. (1) has contributed toward formulating anisotropic and nonstationary Gaussian processes,[11,15,16] non-Gaussian models[17–19] as well as spatiotemporal models.[20,21] There have been calls for stronger interactions between physical and statistical models.[22] The deterministic version of Eq. (1) is the heat equation, and the random term can be interpreted as random forcing or random model disturbance. Whittle called this a diffusion–injection model.[23] Especially, this can be an basis for dynamical spatiotemporal models.[24]

Another extension of Eq. (1) is to introduce systems of SPDEs. An example of a 2 × 2 system is as follows:

$$\begin{pmatrix} \mathcal{L}_{11} & \mathcal{L}_{12} \\ \mathcal{L}_{21} & \mathcal{L}_{22} \end{pmatrix} \begin{pmatrix} x_1(\mathbf{s}) \\ x_2(\mathbf{s}) \end{pmatrix} = \begin{pmatrix} \omega_1(\mathbf{s}) \\ \omega_2(\mathbf{s}) \end{pmatrix}, \tag{2}$$

where $\mathcal{L}_{ij}$ are differential operators similar to the one given in Eq. (1) and $\omega_1(\mathbf{s})$ and $\omega_2(\mathbf{s})$ are independent spatial white noise processes. To this day, systems of SPDEs have been used to obtain spatial models with oscillating dependency[11,25] and spatial multivariate models.[26] As spatial modeling with systems of SPDEs is the focus of this study, we give a short overview of specific challenges and alternative approaches for spatial multivariate models.

Multivariate spatial modeling and inference has been an area of active research for many years, and it is still challenging.[2,26–31] There are two principal approaches to multivariate spatial modeling, conditional modeling and specifying the joint model directly.[2] Many models are based on combinations of univariate models,[30] and this include linear models of coregionalization,[7] convolution methods[32,33] and methods using latent dimensions.[34] The Matérn model has been extended to multivariate spatial models,[28,29] and ensuring valid models has proven to be involved. One of the main challenges in multivariate spatial modeling is specification of the cross-dependency (directly or indirectly through conditioning),[30,31] i.e., the dependency between the spatial phenomena at different locations. Many models assume (by construction) symmetry in the cross-dependency to ensure valid models (i.e., models with positive definite covariance matrices). An implication of this assumption is that the dependency between temperature in London and air pressure in Amsterdam is the same as the dependency between temperature in Amsterdam and air pressure in London. This is not always reasonable, and there has been some progress toward finding classes of valid models with asymmetric cross-dependency.[35] Systems of SPDEs can be set up to archive both symmetric and asymmetric cross-dependency.[25,26]

All the extensions of the SPDE approach mentioned above are easily done such that valid models are guaranteed. Furthermore, the Markov property, and hence the computational efficiency is also kept. Therefore, spatial modeling with SPDEs can contribute toward both the two main challenges in spatial statistics stated above simultaneously: (1) specify flexible valid models and (2) computationally feasible inference.

In the next section, we give an introduction to spatial modeling formulated using (a single) SPDE, and describe relevant computational, modeling and inference issues. Brief overviews of extensions to non-stationary and non-Gaussian models using the SPDE formulation are also included. An introduction to systems of SPDEs in spatial modeling is given in the subsequent section. This includes multivariate spatial models, models with oscillating covariance function, and connections to some common (non-SPDE) model formulations. The exploration of opportunities and properties for models formulated as a system of SPDEs has just started. Some of the interesting opportunities for future research using systems of SPDEs are also provided. The article ends with a short conclusion.

MODELING AND INFERENCE WITH SPDES

Gaussian Random Fields and Gaussian Markov Random Fields

A spatial stochastic process is defined over a continuous domain. In practice, we consider (or observe) it for a fixed number of locations $\mathbf{s} = (\mathbf{s}_1, \ldots, \mathbf{s}_n)$. We refer to the joint distribution for any fixed number of locations as a random field. A random field $x(\mathbf{s})$ is a Gaussian random field (GRF) if any collection $\{x(\mathbf{s}_i); i = 1, 2, \ldots, n\}$ for every $n \geq 1$ is jointly Gaussian distributed. GRFs are popular (as components) in spatial models.[7,24,27,36]

A GRF $x(\mathbf{s})$ can be explicitly specified through its mean function $\mu(\mathbf{s})$ and covariance function $\mathrm{cov}(x(\mathbf{s}_1), x(\mathbf{s}_2))$, $\mathbf{s}_i \in \mathbb{R}^d$. A spatial model is stationary isotropic if the covariance function between two points only depends on the absolute value of the distance $h = \| \mathbf{s}_1 - \mathbf{s}_2 \|$; $\mathrm{cov}(x(\mathbf{s}_1), x(\mathbf{s}_2)) = C(\|\mathbf{s}_1 - \mathbf{s}_2\|) = C(h)$. When doing inference for GRFs the covariance matrix Σ of the observations needs to be inverted. This has computational cost $\mathcal{O}(n^3)$ for a $n \times n$ matrix.[12]

In some situations, we have discrete spatial domains, e.g., when considering pixels in an image or data for administrative units. In these situations, GMRF models are popular.[12] A GMRF is a GRF with a Markov property. The Markov property refers to conditional independence given neighborhoods. Examples of neighborhoods are all pixels or administrative units that share a border. For a GMRF $= (x_1, x_2, \ldots, x_n)$ we let $\mathbf{x}_{(-i)}$ denote all other elements in $\mathbf{x}$ but x_i, and $\mathbf{x}_{\mathrm{ne}(i)}$ denotes the elements in the neighborhood of x_i. The Markov property implies that $\pi(x_i|\mathbf{x}_{(-i)}) = \pi(x_i|\mathbf{x}_{\mathrm{ne}(i)})$ for all elements x_i. This conditional independence structure is reflected in the precision matrix $\mathbf{Q}$ (inverse of covariance matrix, $\mathbf{Q} = \sum^{-1}$). Only entries in $\mathbf{Q}$ corresponding to conditional dependent variables are nonzero:

$$i \notin \mathrm{ne}(j) \Rightarrow Q_{ij} = 0. \tag{3}$$

This sparse structure allows for numerical linear algebra methods for sparse matrices, and therefore makes GMRFs computationally attractive. For a spatial structure (in $\mathbb{R}^2$), the computational cost of inference is $\mathcal{O}(n^{3/2})$.

Because of these computational benefits, it is appealing to model also continuous spatial phenomena as GMRFs. Traditionally, this has required both a regular discretion of the domain, and that the model has to be defined through conditional dependency instead of covariance functions.[12]

Gaussian Random Fields with SPDEs

The wish to do spatial modeling using GRFs and at the same time obtain computational benefits from Markov properties motivated the (re-)introduction of spatial modeling using SPDEs.[11,13,17] Lindgren et al.[11] showed that the SPDE approach allow us to build models using GRFs and do inference benefiting from GMRFs. Furthermore, flexible spatial models can be defined using the SPDE approach. These models are defined in terms of processes instead of the covariance functions, and therefore avoid the difficult task of ensuring positive definite covariance matrices.

The basis of the SPDE approach is the equation as follows:

$$\tau\left(\kappa^2 - \Delta\right)^{\alpha/2} x(\mathbf{s}) = \omega(\mathbf{s}). \quad \mathbf{s} \in \mathbb{R}^d, \tag{4}$$

where ω is spatial Gaussian white noise, Δ is the Laplacian

$$\Delta = \sum_{i=1}^{d} \frac{\partial^2}{\partial x_i^2}, \quad (5)$$

$\kappa > 0$ controls the spatial dependency, $\tau > 0$ is a precision parameter, and α is related to the smoothness. It has been shown that the stationary solutions $x(\mathbf{s})$ to Eq. (4) are GRFs with Matérn covariance functions for every $\alpha > d/2$,[9,10] and GMRFs when α is an integer[11] (for most of this article, we set $d = 2$). The Matérn covariance function between two locations $\mathbf{s}_1$ and $\mathbf{s}_2$ is given by the following equation:

$$C(\mathbf{s}_1, \mathbf{s}_2) = \frac{\sigma^2}{\nu 2^{\nu-1}} (\kappa \| \mathbf{s}_2 - \mathbf{s}_1 \|)^{\nu} K_\nu(\kappa \| \mathbf{s}_2 - \mathbf{s}_1 \|), \quad (6)$$

where K_ν is the modified Bessel function of the second kind and order $\nu > 0$, κ is a positive scaling parameter and σ^2 is the marginal variance. The parameters in the SPDE formulation (Eq. (4)) and the Matérn covariance function (Eq. (6)) are coupled. The smoothness parameter of the Matérn function is $\nu = \alpha - d/2$, and the marginal variance is given as follows:

$$\sigma^2 = \frac{\Gamma(\nu)}{\Gamma(\alpha)(4\pi)^{d/2}\kappa^{2\nu}\tau^2}. \quad (7)$$

Hence, a larger SPDE parameter τ gives a smaller marginal variance. The range is a quantity often used to describe spatial models, and it is interpreted as the distance at which the dependency becomes negligible. By defining range (ρ) as the distance for which the correlation is 0.1, an empirical relation between range and the SPDE parameters has been derived: $\rho = \sqrt{8\nu}/\kappa$.[11] We see that the range does not depend on the SPDE parameter τ, only on κ and ν. Larger κ implies shorter range, and larger ν (or α) implies longer range.

The Matérn family of correlation functions has been widely used in spatial statistics since there is a parameter that allows for varying the differentiability, and the exponential family is a special case of this family.[8] Indeed, Stein[8] concluded 'use the Matérn model.' Figure 1 shows a sample from a Matérn GRF (panel a) and the corresponding correlation function (panel b).

Lindgren et al.[11] showed an explicit link between GRFs and GMRFs for integer values of α by using the finite element method (FEM)[37] together with piecewise linear basis functions. The spatial domain is first discretized into nonintersecting triangles with m nodes (often refereed to as the mesh). The basis functions ψ_k are piecewise linear functions in each triangle. Basis function ψ_k has value 1 at vertex k and 0 at all other vertexes. The values in the interior of a triangle are determined by linear interpolation.

Figure 2[38] illustrates an approximation to a smooth surface using a triangulation and piecewise linear basis functions.

Mathematically, the approximation can be expressed as follows:

$$x(\mathbf{s}) = \sum_{k=1}^{m} \psi_k(\mathbf{s})\omega_k. \quad (8)$$

where $\psi_k(\mathbf{s})$ are the linear basis functions, ω_k are weights, and m is the number of vertexes in the triangulation.

The probability distribution of the spatial field $x(\mathbf{s})$ is now determined by the joint distribution of

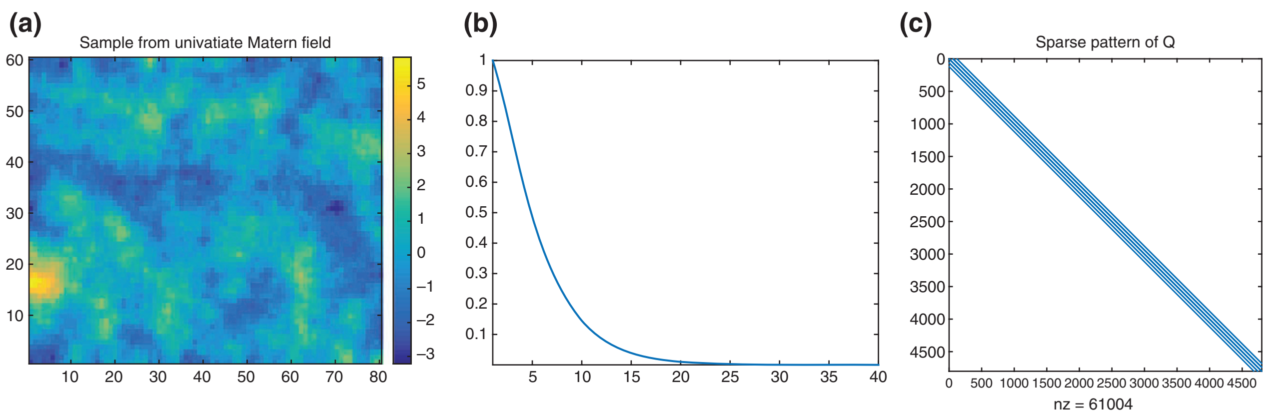

FIGURE 1 | A sample of a GRF (a) from SPDE (Eq. (4)) with $\tau = 1$; $\kappa = 0.3$, $\alpha = 2$, its correlation function corr(h) (b), and the sparse pattern of the precision matrix $\mathbf{Q}$ (c).

the weights $\boldsymbol{w} = (\omega_1, \omega_2, \ldots, \omega_m)$. We typically want to estimate the weights, and it has been shown that the weights have a GMRF structure.[11] Hence, modeling can be done using GRFs, while computations can utilize GMRFs. In Figure 1, panel c is the nonzero structure of the precision matrix **Q** from the SPDE formulation and the linear approximation based on triangulation. This precision matrix is very sparse: Only elements that corresponds to neighboring nodes in the triangulation (a vertex between them) are nonzero. Note that the corresponding covariance matrix, $\sum = \mathbf{Q}^{-1}$, is a full matrix. Working with the sparse precision matrix reduces the computational cost considerable.

It is worth to note that the computational cost depends on the size of the mesh (m), but not on the size of the dataset (n). Hence, the SPDE approach can also be used for approximate inference for large spatial datasets[13] by choosing a coarser mesh. This can be seen as a low-rank method, and it is an intuitive alternative to other low-rank methods such as fixed rank kriging[39] and predictive processes.[40] The accuracy of the approximation depends on the density of the triangulation. Generally, the accuracy increases with decreasing edge sizes of the triangles.

The SPDE can be defined on any manifold, and valid models are achieved.[11] This allows us to define models for other domains than $\mathbb{R}^2$. For example, temperature for the whole globe has been modeled using a Matérn model on the sphere.[11] Figure 3 is a sample from a GRF on the sphere (panel a), its covariance function (panel b), and the sparse pattern of precision matrix **Q** (panel c). The precision matrix is sparse, and hence fast inference is feasible.

When we consider spatial models on a finite subspace of $\mathbb{R}^2$, there will be boundary effects. The common approach is to use the Neumann boundaries (the boundaries have zero normal derivatives) and build the mesh on a larger region than the observation region. Thereby, the boundary effects are moved to regions we are not interested in.[11] In some cases, there is a physical boundary, the boundary effects are

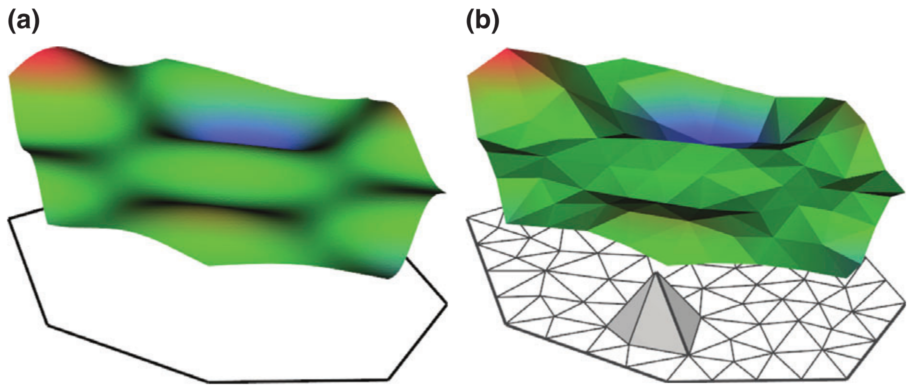

FIGURE 2 | A smooth surface (a) and its local linear approximation (b) using triangulation and piecewise linear basis functions.

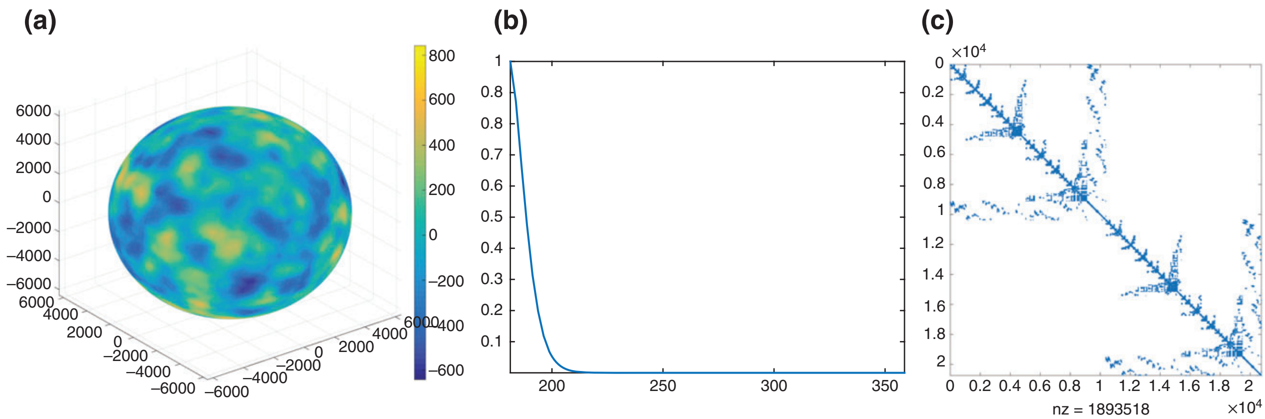

FIGURE 3 | A sample of a Matérn GRF (a) on the sphere (radius = 6738.1) from SPDE (Eq. (4)) with $\alpha = 2$; $\tau = 5$, $\kappa^2 = 360$, its correlation function corr(h) (b) and the sparse pattern (c) of the precision matrix **Q**.

desired, and then we can included the effect in the model.[41] There are no boundary issues for processes defined on the (whole) sphere.

Model Framework, Inference, and Sampling

When modeling a spatial phenomenon there are in most cases explanatory variables we want to include. Furthermore, the phenomenon is often not directly observed. Therefore, a spatial GRF is often one of several components in the spatial model. (Bayesian) hierarchical models[2,7,12,24,42] have been a popular framework for spatial modelers, and most models using the SPDE formulation have been within this framework. A hierarchical model has two stages, the data model and the process model. If a full Bayesian model is set up, a third stage is added, the parameter model. Let **y** denotes the observations, and η the spatial phenomenon of interest. Furthermore, let θ generically denote parameters. The three stages of hierarchical models are then as follows:

- Stage I: *Data model*. This model describes the probability distribution of the observations given the (true) underlying process η. We denote it $\pi(\mathbf{y}|\boldsymbol{\eta}, \boldsymbol{q}_y)$, where θ_y are related parameters. Examples of data models are Gaussian measurement uncertainty (and θ_y is the measurement variance), or Poisson models where η is the intensity (or risk).
- Stage II: *Process model*. This model describes the distribution of η, given related parameters θ_η, $\pi(\eta|\theta_\eta)$. A typically spatial model is linear and includes both explanatory variables and a spatial random field, e.g.,

$$\eta = \mathbf{X}\beta + \mathbf{x}(s), \tag{9}$$

where $\mathbf{X}\beta$ captures fixed effect with design matrix $\mathbf{X}$ and coefficients β, and $\mathbf{x}(s)$ is a spatial GRF. When using the SPDE approach $\mathbf{x}(s)$ is defined by a SPDE, e.g., Eqs (1) or (2).

- Stage III: *Parameter model*. In this stage, the prior distributions of the parameters are defined, $\pi(\beta, \theta_x, \theta_y)$. Priors for the SPDE parameters have to be set carefully, and methodology for this has been developed.[43,44]

To make inference, one has to choose which approach to take. This choice depends on personal preferences, the size of the system, and which methodology that is available for the model structure. When the GRF is defined using the SPDE approach it is the weights $\boldsymbol{w} = (\omega_1, \omega_2, \ldots, \omega_m)$ of the basis representation in Eq. (8) that are to be estimated (or integrated out). These form a GMRF, and the sparsity of the corresponding precision matrix gives large computational benefits for most inference approaches. Many hierarchical models using SPDE-based GRFs fit into the INLA framework, and full Bayesian analysis using INLA has been popular.[11,15,20,43,45] In the INLA framework, the GMRF used for the inference includes both the weights ω and the coefficients for the fixed effects β. Many spatial SPDE models fit into the INLA framework. However, some do not, and full Bayesian analysis using Markov chain Monte Carlo methods can then be used.[46] In some works, an empirical Bayes approach is taken.[16,25,26,47,48] Wallin and Bolin[19] presented a maximum likelihood estimation technique based on the Monte Carlo expectation–maximization algorithm for non-Gaussian spatial Matérn fields.

Sampling from a SPDE model can be done computationally efficient by using the GMRF representation.[12,49]

SPDE Approach and Its Extensions

The SPDE approach has been extended in several ways to accommodate a rich class of models. A more general model can be expressed as follows:

$$\tau(\mathbf{s})\left(\kappa(\mathbf{s})^2 - \nabla H(\mathbf{s})\nabla\right)^{\alpha/2} x(\mathbf{s}) = \varepsilon(\mathbf{s}), \tag{10}$$

where ∇ is the gradient $\Delta = \nabla \cdot \nabla$. This gives several opportunities to extend the Matérn model. Nonstationary random fields can be constructed by allowing the scaling parameter $\kappa(\mathbf{s})$ and the variance parameter $\tau(\mathbf{s})$ to vary in space. We let $\varepsilon(\mathbf{s})$ denote a general noise process, and reserve $\omega(s)$ for white noise. We note that the SPDE in Eq. (10) is driven by (right hand side) by $\varepsilon(\mathbf{s})$. A richer class of model can be achieved by using other noise processes than white noise.

For modeling nonstationarity, Lindgren et al.[11] use a nonparametric approach, while Ingebrigtsen et al.[15] use spatially varying explanatory variable to model the parameters in the SPDE. Spatial random fields with anisotropy can be obtained through the 2×2 matrix H, and local anisotropy[16] is obtained by letting this matrix vary through space, i.e., $H(\mathbf{s})$. However, they report unsolved identifiability issues that make the model unpractical.[16]

A rich class of spatial Gaussian models can be obtained by using spatially dependent Gaussian noise $\varepsilon(\mathbf{s})$. The nested SPDE models[17] are a generalization of the SPDE (Eq. (4));

$$\mathcal{L}_1 x(\mathbf{s}) = \mathcal{L}_2 \omega(\mathbf{s}), \tag{11}$$

where $\mathcal{L}_1$ and $\mathcal{L}_2$ are linear operators. In other words, the model is extended by allowing spatial dependent noise, $\varepsilon(\mathbf{s}) = \mathcal{L}_2\omega(\mathbf{s})$. The system where $\mathcal{L}_1$ is a product of (i.e., successive use of) the operator used in Eq. (4), e.g., $\mathcal{L}_1 = \tau(\kappa^2 - \Delta)^{\alpha/2}$, and $\mathcal{L}_2$ is (a product of) gradient based operator(s) have been studied in depth by Bolin and Lindgren.[17] This nested SPDE class of model includes both the Matérn model and different oscillating covariance function. The SPDE approach has been further extended to non-Gaussian Matérn random fields by using non-Gaussian noise.[18,19]

SPATIAL MODELING USING SYSTEMS OF SPDES

Defining spatial Gaussian processes (with mean vector zeros) as the solution to a system of SPDEs gives new opportunities for spatial modeling and inference. The attractive properties of using (one) SPDE to define a spatial process extend to systems of SPDEs. Hence, defining spatial Gaussian processes as systems of SPDEs provide computationally efficient approximations, models that can be define on any manifold, and complex valid models can be archived by changing the parameters in the SPDE formulation.

The following $p \times p$ system of SPDEs was studied by Hu et al.;[26]

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

where $\mathcal{L}_{ij} = \tau_{ij}\left(\kappa_{ij}^2 - \Delta\right)^{\alpha_{ij}/2}$, $\kappa_{ij} > 0$, $\tau_{ii} > 0$, while τ_{ij} for $i \neq j$ can take any real value. Each $\varepsilon_i(\mathbf{s})$ is Gaussian noise (white noise or spatially dependent noise), and the noise processes $\varepsilon_i(\mathbf{s})$, $i = 1, 2, \dots, p$ are independent, but not necessarily identically distributed. Δ is the Laplacian as defined in Eq. (5). Compared to the system of SPDEs set up in Eq. (2), this general system has dimension $p \times p$. Instead of assuming white independent noise $\{\omega_i(s)\}$, we allow for spatial dependency within the noise processes $\{\varepsilon_i(s)\}$. Furthermore, the matrix of operators is generally not symmetric, i.e., $\mathcal{L}_{ij} \neq \mathcal{L}_{ji}$. It has been shown that spatial models defined as the solution to this system are GRF models.[26] When finding properties of models and when comparing models it is often easier to work in the frequency domain, i.e., the Fourier transform is applied. The Fourier transform of the system of SPDEs (Eq. (12)) has been derived.[26] To achieve computational benefits, a similar approach is taken as when a model is specified with (one) SPDE, and the piecewise linear base functions are used together with Markov approximation. Therefore, the precision matrix $\mathbf{Q}$ for the GRF $\mathbf{x}(\mathbf{s}) = (x_1(\mathbf{s}), x_2(\mathbf{s}), \dots, x_p(\mathbf{s}))$ is sparse, and the sparse structure is known by construction.

To this day, systems of SPDEs have contributed to spatial modeling in two directions; multivariate spatial models and spatial models with oscillating covariance functions. Multivariate spatial models with oscillating covariance function can be modeled using a system of SPDEs with spatially dependent Gaussian noise. The next subsections describe these models. Several other interesting systems of SPDEs models have been proposed, but not further analyzed and explored through simulation studies or case studies. Some of these are summarized in the last subsection.

Multivariate Spatial Models Using Systems of SPDEs

Often more than one spatial variable is of interest, and generally a p-variate spatial field $\mathbf{x} = (x_1(\mathbf{s}), x_2(\mathbf{s}), \dots, x_p(\mathbf{s}))$ is of interest. A multivariate GRF can be fully specified by its mean functions $\mu_i(\mathbf{s})$, the marginal-covariance functions (also called direct covariance function)[30,31] $C_{ii}(\mathbf{s}_1, \mathbf{s}_2) = \mathrm{cov}(x_i(\mathbf{s}_1), x_i(\mathbf{s}_2))$ for each of the spatial variables $i = 1, \dots, p$, and its cross-covariance functions $C_{ij}(\mathbf{s}_1, \mathbf{s}_2) = \mathrm{cov}(x_i(\mathbf{s}_1), x_j(\mathbf{s}_2))$ for $i, j = 1, \dots, p$, $i \neq j$. The marginal-covariance functions describe the dependency structures for spatial fields x_i, and the cross-covariance function C_{ij} specifies the dependency between the spatial fields x_i and x_j.[30] For example, if we want to model temperature $x_1(\mathbf{s})$ and pressure $x_2(\mathbf{s})$ simultaneously, this will require a bivariate spatial model ($p = 2$). The cross-covariance $C_{12}(\mathbf{s}_1, \mathbf{s}_2)$ is the covariance between temperature at location $\mathbf{s}_1$ and pressure at location $\mathbf{s}_2$. If the cross-correlation is stationary and isotropic, it is only a function of the distance, $C_{ij}(\mathbf{s}_1, \mathbf{s}_2) = C_{ij}(\|\mathbf{s}_1 - \mathbf{s}_2\|) = C_{ij}(\mathbf{h})$. Furthermore, the cross-covariance is said to be symmetric if $C_{ij}(\mathbf{s}_1, \mathbf{s}_2) = C_{ji}(\mathbf{s}_1, \mathbf{s}_2)$. If the covariance between temperature at location $\mathbf{s}_1$ and pressure at location $\mathbf{s}_2$ differ from the covariance between pressure at location $\mathbf{s}_1$ and at temperature location $\mathbf{s}_2$ the cross-correlation is asymmetric. A major challenge in multivariate spatial statistics is to specify valid, flexible, and interpretable models, especially for the dependency between spatial fields (the cross-covariance function). Furthermore, to do inference expensive computations and implementations are often needed and parameters are often hard to estimate /identify.[28–30,34]

In order to construct a multivariate random field we can use a system of SPDEs (Eq. (12)). The model is specified through the SPDE parameters and not through the marginal-covariance and cross-covariance functions. Valid models are ensured and both symmetric and asymmetric models can be obtained. Figure 4 illustrates an example of a bivariate GRF constructed by a system of SPDEs. Figure 4 (a) and (b) shows a sample of the bivariate model, Figure 4(c) shows the correlation and cross-correlation functions, and Figure 4(d) gives the sparse pattern of the precision matrix **Q**.

To gain insight on how to model multivariate spatial models using systems of SPDEs, we start by considering some special cases. For the sake of easier presentation, we mostly set up bivariate models ($p = 2$). Because of identifiability issues, the smoothness parameters are often fixed when fitting models to data,[27] and this has also been the practice for SPDE models[15–17,20,25,26,47] Therefore, we fix the smoothness parameters to $\alpha_{ij} = 2$, $i, j = 1, 2$ for most of the presentation. Hence, the systems of SPDEs we work with for constructing bivariate spatial models are as follows:

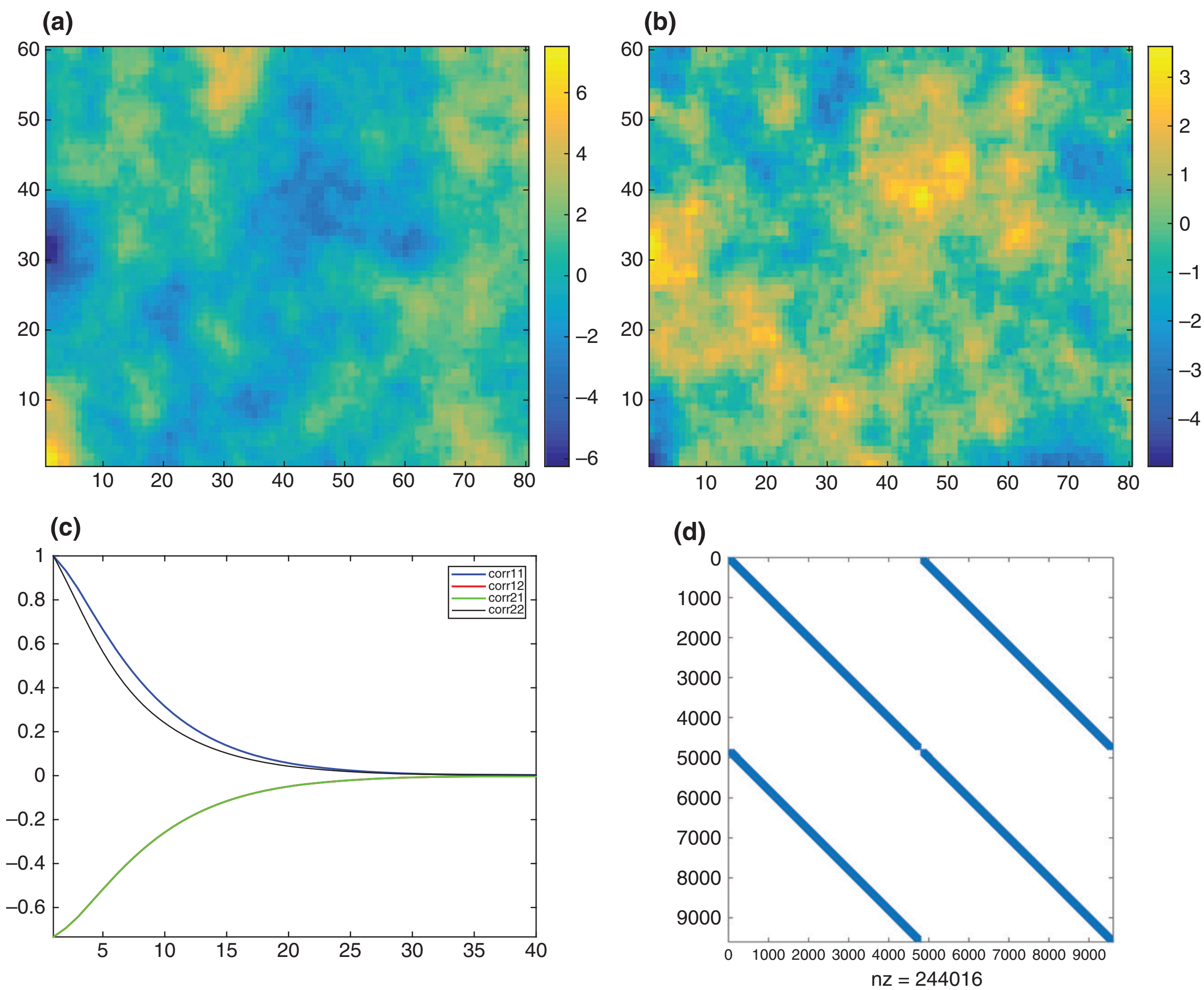

FIGURE 4 | A sample of a bivariate GRF on $\mathbb{R}^2$ (a, b) defined with the system of SPDEs in Eq. (21) with $\tau_{11} = \tau_{22} = 1$, $\tau_{21} = 0.4$, $\kappa_{11}^2 = 0.2$, $\kappa_{21}^2 = 0.5$, and $\kappa_{22}^2 = 0.4$, the corresponding correlation and cross-correlation functions (cross-correlation functions are symmetric; $\mathrm{corr}_{12}(h) = \mathrm{corr}_{21}(h)$) (c), and the sparse structure of the precision matrix (d).

$$\begin{pmatrix} \tau_{11}\left(\kappa_{11}^2-\Delta\right) & \tau_{12}\left(\kappa_{12}^2-\Delta\right) \\ \tau_{21}\left(\kappa_{21}^2-\Delta\right) & \tau_{22}\left(\kappa_{22}^2-\Delta\right) \end{pmatrix} \begin{pmatrix} x_1(\mathbf{s}) \\ x_2(\mathbf{s}) \end{pmatrix} = \begin{pmatrix} \omega_1(\mathbf{s}) \\ \omega_2(\mathbf{s}) \end{pmatrix}, \tag{13}$$

where $\omega_1(\mathbf{s})$ and $\omega_2(\mathbf{s})$ are independent white noise processes, $\kappa_{ij} > 0$, $\tau_{ii} > 0$, while τ_{ij} for $i \neq j$ can take any real value. The solution to Eq. (13) is a bivariate GRF.[26]

Independent Models

The trivial bivariate spatial model assumes that the spatial processes are independent. Assume we have two independent random fields $\mathbf{x}_1$ and $\mathbf{x}_2$ that are constructed by the SPDE approach as follows:

$$\begin{aligned} \tau_{11}\left(\kappa_{11}^2-\Delta\right)x_1(\mathbf{s}) &= \omega_1(\mathbf{s}), \\ \tau_{22}\left(\kappa_{22}^2-\Delta\right)x_2(\mathbf{s}) &= \omega_2(\mathbf{s}). \end{aligned} \tag{14}$$

In this case, $x_1(\mathbf{s})$ and $x_2(\mathbf{s})$ are two GRFs with Matérn covariance functions. We can rewrite Eq. (14) to a system of SPDEs:

$$\begin{pmatrix} \tau_{11}\left(\kappa_{11}^2-\Delta\right) & 0 \\ 0 & \tau_{22}\left(\kappa_{22}^2-\Delta\right) \end{pmatrix} \begin{pmatrix} x_1(\mathbf{s}) \\ x_2(\mathbf{s}) \end{pmatrix} = \begin{pmatrix} \omega_1(\mathbf{s}) \\ \omega_2(\mathbf{s}) \end{pmatrix}. \tag{15}$$

The fields $x_1(\mathbf{s})$ and $x_2(\mathbf{s})$ as defined in Eq. (15) have different variances and correlation range controlled by τ_{ii} and κ_{ii}, $i = 1, 2$. If we constrain $\kappa_{11} = \kappa_{22}$ the two fields are assumed to have the same range, and if in addition $\tau_{11} = \tau_{22}$ they are two independent realizations of the same process, also known as two replicates.[43]

Linear Combinations of Matérn Fields

A popular multivariate spatial model class is linear models of coregionalization (LMC).[27,50] In its simplest form, the model is constructed by linear combinations of independent spatial fields. Assuming that f_1 and f_2 are independent Matérn random fields with $\tau = 1$, and define $x_1(\mathbf{s})$ and $x_2(\mathbf{s})$ as

$$\begin{aligned} x_1(\mathbf{s}) &= a_{11}f_1(\mathbf{s}); \\ x_2(\mathbf{s}) &= a_{21}f_1(\mathbf{s}) + a_{22}f_2(\mathbf{s}); \end{aligned} \tag{16}$$

It is often reasonable to assume that the $f_i(\mathbf{s})$, $i = 1, 2$ share a common correlation function.[27] In these models, both $x_1(\mathbf{s})$ and $x_2(\mathbf{s})$ are Matérn models. This LMC model can be formulated as systems of SPDEs as follows:

$$\begin{pmatrix} \tau_{11}\left(\kappa^2-\Delta\right) & 0 \\ \tau_{21}\left(\kappa^2-\Delta\right) & \tau_{22}\left(\kappa^2-\Delta\right) \end{pmatrix} \begin{pmatrix} x_1(\mathbf{s}) \\ x_2(\mathbf{s}) \end{pmatrix} = \begin{pmatrix} \omega_1(\mathbf{s}) \\ \omega_2(\mathbf{s}) \end{pmatrix}, \tag{17}$$

with

$$a_{11} = \frac{1}{\tau_{11}}, \quad a_{21} = -\frac{\tau_{21}}{\tau_{11}\tau_{22}}, \quad a_{22} = \frac{1}{\tau_{22}}. \tag{18}$$

The system of SPDEs (Eq. (13)) where also the range parameters are constrained to be equal, $\kappa_{ij} = \kappa$, reads;

$$\begin{pmatrix} \tau_{11}\left(\kappa^2-\Delta\right) & \tau_{12}\left(\kappa^2-\Delta\right) \\ \tau_{21}\left(\kappa^2-\Delta\right) & \tau_{22}\left(\kappa^2-\Delta\right) \end{pmatrix} \begin{pmatrix} x_1(\mathbf{s}) \\ x_2(\mathbf{s}) \end{pmatrix} = \begin{pmatrix} \omega_1(\mathbf{s}) \\ \omega_2(\mathbf{s}) \end{pmatrix}. \tag{19}$$

It can be shown that both $x_1(\mathbf{s})$ and $x_2(\mathbf{s})$ are mixtures of Matérn random fields, and this model corresponds to

$$\begin{aligned} x_1(\mathbf{s}) &= \frac{\tau_{12}}{\tau_{12}\tau_{21}-\tau_{22}\tau_{11}}f_2(\mathbf{s}) - \frac{\tau_{22}}{\tau_{12}\tau_{21}-\tau_{22}\tau_{11}}f_1(\mathbf{s}); \\ x_2(\mathbf{s}) &= \frac{\tau_{21}}{\tau_{12}\tau_{21}-\tau_{22}\tau_{11}}f_1(\mathbf{s}) - \frac{\tau_{11}}{\tau_{12}\tau_{21}-\tau_{22}\tau_{11}}f_2(\mathbf{s}). \end{aligned} \tag{20}$$

where $f_1(\mathbf{s})$ and $f_2(\mathbf{s})$ are independent Matérn fields as in Eq. (16).

Triangular System of SPDEs

The LMC model (Eq. (17)) can be generalized by allowing the range parameters κ to differ between the operators, i.e.,

$$\begin{pmatrix} \tau_{11}\left(\kappa_{11}^2-\Delta\right) & 0 \\ \tau_{21}\left(\kappa_{21}^2-\Delta\right) & \tau_{22}\left(\kappa_{22}^2-\Delta\right) \end{pmatrix} \begin{pmatrix} x_1(\mathbf{s}) \\ x_2(\mathbf{s}) \end{pmatrix} = \begin{pmatrix} \omega_1(\mathbf{s}) \\ \omega_2(\mathbf{s}) \end{pmatrix}, \tag{21}$$

which correspond to setting $\tau_{12} = 0$ in Eq. (13). This gives a triangular system. As others have experienced poor identifiability of parameters for multivariate spatial models,[27,28] the analyses of Hu et al.[26,47] were restricted to this triangular system of SPDEs. Through simulation studies as well as case studies, it has been demonstrated that parameters can be identified for this triangular system. For models defined through this triangular system of SPDEs, we can first notice that $x_1(\mathbf{s})$ is a Matérn field. The process $x_2(\mathbf{s})$, can be a Matérn field (if $\kappa_{11} = \kappa_{21}$). If $\kappa_{11} \neq \kappa_{21}$ the process $x_2(\mathbf{s})$ can be seen as a nested SPDE:

$$\mathcal{L}_{22}x_2(\mathbf{s}) = \omega_2(\mathbf{s}) - \mathcal{L}_{21}x_1(\mathbf{s}). \tag{22}$$

If κ_{12} is close to κ_{11}, $\mathcal{L}_{21}x_1(\mathbf{s})$ is close to $\omega_2(\mathbf{s})$, i.e., white noise, and hence $x_2(\mathbf{s})$ is close to a Matérn field.[26] With this triangular system of SPDEs, the cross-covariance can be both positive and negative, and it is controlled by τ_{21}; a negative τ_{21} implies positive correlation between $x_1(\mathbf{s})$ and $x_2(\mathbf{s})$. The example in Figure 4 is constructed using a triangular system of SPDEs (Eq. (21)) with positive τ_{21}, i.e., a negative correlation. With a triangular system the cross-covariance functions are always symmetric.[26,47]

From a modeling point of view, one might want both fields to be marginally Matérn field. A class of cross-covariance functions that fulfill this requirement has been introduced by Gneiting et al.[28] Furthermore, it has been shown that their parsimonious Matérn model[28] can be defined as the solution to the following system of SPDEs:[26]

$$\begin{pmatrix} \tau_{11}\left(\kappa^2-\Delta\right) & 0 \\ \tau_{21}\left(\kappa^2-\Delta\right) & \tau_{22}\left(\kappa^2-\Delta\right) \end{pmatrix} \begin{pmatrix} x_1(\mathbf{s}) \\ x_2(\mathbf{s}) \end{pmatrix} = \begin{pmatrix} \omega_1(\mathbf{s}) \\ \omega_2(\mathbf{s}) \end{pmatrix}, \tag{23}$$

i.e.,, a triangular system, where all range parameters are equal ($\kappa_{ij} = \kappa$). There are also some restrictions on the $\{\tau_{ij}\}$ parameters.[26]

Spatial Oscillating Models Using Systems of SPDEs

Many nature spatial phenomena are oscillating in nature. Examples are ocean waves and pressure on a global scale. In order to construct a GRF with oscillating covariance a coupled system of SPDEs can be used

$$\begin{pmatrix} \kappa_1^2\cos(\pi\theta)-\Delta & -\kappa_2^2\sin(\pi\theta) \\ \kappa_2^2\sin(\pi\theta) & \kappa_1^2\cos(\pi\theta)-\Delta \end{pmatrix} \begin{pmatrix} x_1(\mathbf{s}) \\ x_2(\mathbf{s}) \end{pmatrix} = \begin{pmatrix} \omega_1(\mathbf{s}) \\ \omega_2(\mathbf{s}) \end{pmatrix}, \tag{24}$$

where θ is an oscillation parameter. The solution components $x_1(\mathbf{s})$ and $x_2(\mathbf{s})$ to Eq. (24) are independent GRFs with oscillating covariance functions.[11]

This is equivalent to a complex version of the basic SPDE (Eq. (4)). The idea is that we replace κ with $\kappa\exp(i\pi\theta)$, $x(\mathbf{s})$ with $x_1(\mathbf{s}) + ix_2(\mathbf{s})$ and $\omega(\mathbf{s})$ with $\omega_1(\mathbf{s}) + i\omega_2(\mathbf{s})$. It can be shown that the real and imaginary components $x_1(\mathbf{s})$ and $x_2(\mathbf{s})$ of the stationary solution to the new equation are independent and have the same oscillating covariance function.

Spatial multivariate random fields with oscillating covariance function can be constructed using a system of SPDEs.[25] The main idea is to use the GRFs with oscillating covariance functions introduced above as one or both of the noise processes. Hence, the system of SPDEs used has the form:

$$\begin{pmatrix} \tau_{11}\left(\kappa_{11}^2-\Delta\right) & 0 \\ \tau_{21}\left(\kappa_{21}^2-\Delta\right) & \tau_{22}\left(\kappa_{22}^2-\Delta\right) \end{pmatrix} \begin{pmatrix} x_1(\mathbf{s}) \\ x_2(\mathbf{s}) \end{pmatrix} = \begin{pmatrix} \varepsilon_1(\mathbf{s}) \\ \varepsilon_2(\mathbf{s}) \end{pmatrix}. \tag{25}$$

with $\varepsilon_1(\mathbf{s})$ and/or $\varepsilon_2(\mathbf{s})$ obtained from Eq. (24) (e.g., $\varepsilon_1(\mathbf{s})$ is set to be the solution $x_1(\mathbf{s})$ of Eq. (24)).

Figure 5(a) and (b) shows a sample from a bivariate GRF constructed by a system of SPDEs (25) on the sphere. The correlation and cross-correlation functions are plotted in Figure 5(c). The first field, $x_1(\mathbf{s})$, is a Matérn field, the second field $x_2(\mathbf{s})$ is a GRF with oscillating covariance function, and the two fields have a positive and symmetric cross-correlation. It is interesting to note that even though $x_2(\mathbf{s})$ has an oscillating correlation function, the cross-correlation functions do not oscillate. The sparse pattern of the precision matrix $\mathbf{Q}$ is given in Figure 5(d).

Suggested Spatial System of SPDEs Models

In this section, we review spatial models formulated as systems of SPDEs that have been proposed in the literature, but not fully tested through simulation studies or case studies. Most of these models are natural extensions of models that have been tested. Partly, they have not yet been tested because they are at the research frontier. Many of the models have more parameters than the explored less complex versions. And (informed) fear of identifiability issues might be a reason why some of them are not yet further explored.

The models presented above all have fixed smoothness using $\alpha = 2$. By using a different value for α, or letting the different operators have different values α_{ij}, this assumption can be relaxed. The bivariate spatial triangular SPDE model then reads.

$$\begin{pmatrix} \tau_{11}\left(\kappa_{11}^2-\Delta\right)^{\alpha_{11}/2} & 0 \\ \tau_{21}\left(\kappa_{21}^2-\Delta\right)^{\alpha_{21}/2} & \tau_{22}\left(\kappa_{22}^2-\Delta\right)^{\alpha_{22}/2} \end{pmatrix} \begin{pmatrix} x_1(\mathbf{s}) \\ x_2(\mathbf{s}) \end{pmatrix} = \begin{pmatrix} \omega_1(\mathbf{s}) \\ \omega_2(\mathbf{s}) \end{pmatrix}. \tag{26}$$

There are two challenges using this model. First, the smoothness is hard to estimate even for a univariate stationary model, i.e., the model corresponding to Eq. (1). Second, the properties of the model are not clear. In the triangular system above, $x_1(\mathbf{s})$ has

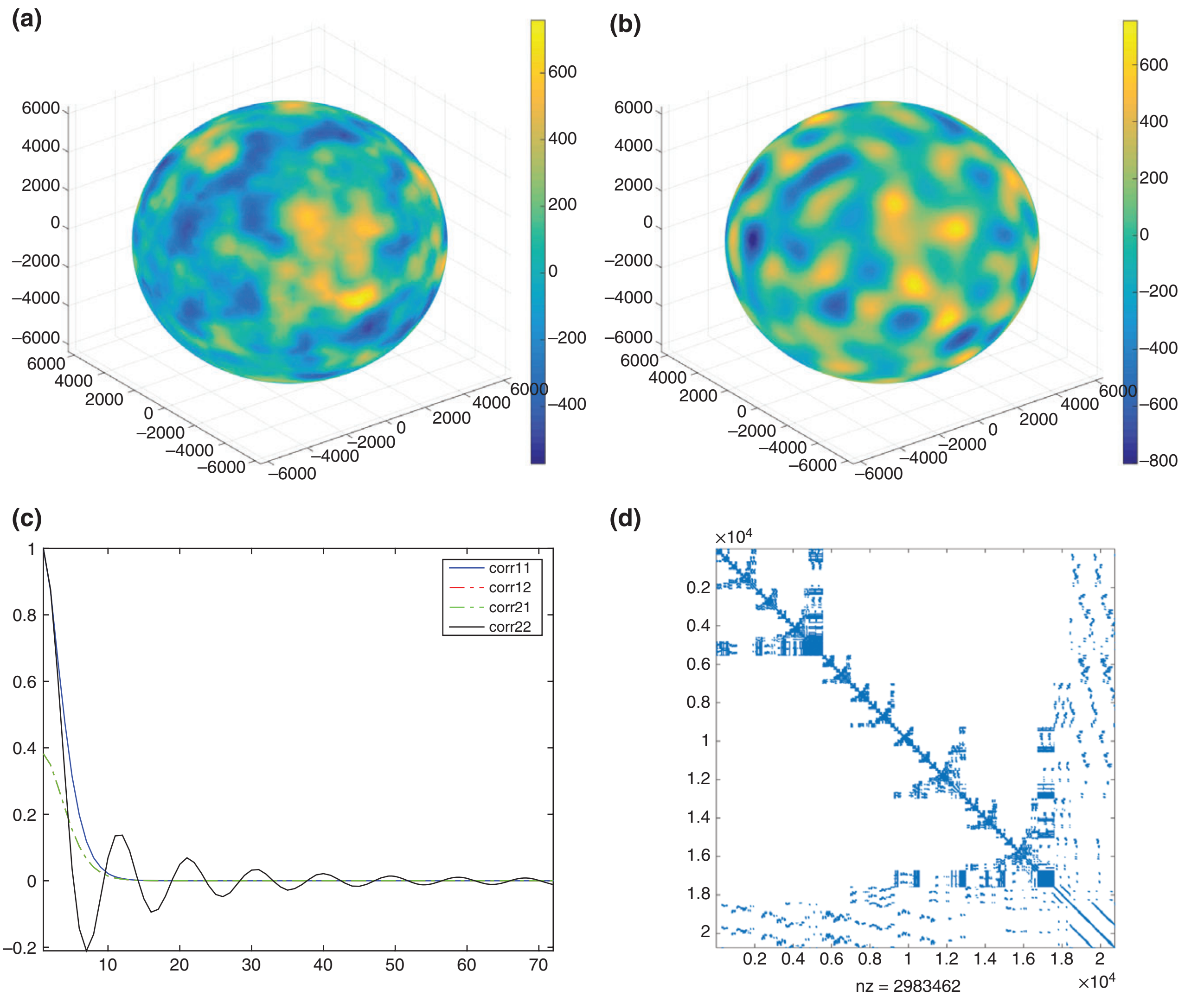

FIGURE 5 | A sample from a bivariate GRF on the sphere $\mathbb{S}^2$ (radius = 6738.1) (a, b) defined with the SPDE in Eq. (25) with τ_{11} = 5, τ_{21} = − 1, τ_{22} = 20, $\kappa_{11}^2 = 360$, $\kappa_{21}^2 = 1$, and $\kappa_{22}^2 = 500$, the corresponding correlation and cross-correlation functions (cross-correlation functions are symmetric; $\text{corr}_{12}(h)$ = $\text{corr}_{21}(h)$) (c), and the sparse structure of precision matrix (d).

properties as the corresponding univariate model. The properties of $x_2(\mathbf{s})$ however will depend on both (the smoothness) of x_1 and the operator $\mathcal{L} = \tau_{21}\left(\kappa_{21}^2 - \Delta\right)^{\alpha_{21}/2}$.

Using the triangular system (Eq. (21)) implicitly assumes a symmetric cross-correlation function.[25,26] One way to construct a bivariate GRF with a asymmetric cross-covariance functions is to use the full system of SPDEs given in Eq. (13).[26] This model has eight parameters, increasing to 12 if also smoothness parameters are allowed to vary, and identifiability might become an issue. The interpretation of the parameters becomes harder than the triangular system; The sign of the cross-correlation is no longer obvious from the values of τ_{12} and τ_{21}. However, it will be interesting to analyze these systems properties, gaining knowledge about identifiability, e.g., through doing simulation studies, and testing its flexibility for case studies.

One way of constructing an oscillating anisotropic model with the computational benefits through the SPDE approach is to use a coupled system of SPDEs:[11]

$$\begin{pmatrix} \kappa_1^2\cos(\pi\theta) - \nabla\cdot\mathbf{H}_1\nabla & -\kappa_2^2\sin(\pi\theta) + \nabla\cdot\mathbf{H}_2\nabla \\ \kappa_2^2\sin(\pi\theta) - \nabla\cdot\mathbf{H}_2\nabla & \kappa_1^2\cos(\pi\theta) - \nabla\cdot\mathbf{H}_1\nabla \end{pmatrix} \begin{pmatrix} x_1(\mathrm{s}) \\ x_2(\mathrm{s}) \end{pmatrix} = \begin{pmatrix} \omega_1(\mathrm{s}) \\ \omega_2(\mathrm{s}) \end{pmatrix}, \tag{27}$$

where ∇ is the gradient and $\mathbf{H_1}$ and $\mathbf{H_1}$ are 2×2 matrices that define the anisotropy. The solution components $x_1(\mathbf{s})$ and $x_2(\mathbf{s})$ to this system of SPDEs are independent. This model can be seen as combining the oscillating model defined by Eq. (24) with the methodology for construction anisotropic models.[16] The oscillating model defined by Eq. (24) is then a special case with $\mathbf{H}_1 = \mathbf{I}$ and $\mathbf{H}_2 = \mathbf{0}$.

One approach for specifying a p-variate GRF is to use a the triangular systems of SPDEs below:[25,26]

$$\begin{pmatrix} \mathcal{L}_{11} & & & \\ \mathcal{L}_{21} & \mathcal{L}_{22} & 0 & \\ \vdots & \vdots & \ddots & \\ \mathcal{L}_{p1} & \mathcal{L}_{p2} & \dots & \mathcal{L}_{pp} \end{pmatrix} \begin{pmatrix} x_1(\mathbf{s}) \\ x_2(\mathbf{s}) \\ \vdots \\ x_p(\mathbf{s}) \end{pmatrix} = \begin{pmatrix} \varepsilon_1(\mathbf{s}) \\ \varepsilon_2(\mathbf{s}) \\ \vdots \\ \varepsilon_p(\mathbf{s}) \end{pmatrix}, \tag{28}$$

where $\varepsilon_i(\mathbf{s})$ and $\mathcal{L}_{ij} = \tau_{ij}\left(\kappa_{ij}^2 - \Delta\right)^{\alpha_{ij}/2}$ are as in Eq. (12). This model defines multivariate GRFs with symmetric cross-correlation matrices. One can introduce oscillation into the multivariate GRFs by using noise process(es) with oscillating covariance.[25] With this triangular system of SPDEs, if only noise process $\varepsilon_i(\mathbf{s})$ has an oscillating covariance function, then the random fields $x_j(\mathbf{s})$, $j < i$ have nonoscillating covariance function and the random fields $x_j(\mathbf{s})$, $j \geq i$ have oscillating covariance functions.[25] The number of parameters grows fast with the number of fields p. The triangular SPDE system (Eq. (28)) with $p = 3$, $\alpha = 2$, and white noise has been tested through a small simulation study.[26] Inference was fast and it was possible to identify parameters (τs and κs).

RESEARCH DIRECTIONS

Using systems of SPDEs in spatial modeling is a new approach, and only some of the simplest systems of SPDEs have been explored. Therefore, there are many interesting directions for further research, including exploration and testing the models described in the subsection above. One new direction is merging systems of SPDEs with parametric[15] or nonparametric[16] nonstationary SPDE models. This could be an approach to achieve nonstationary multivariate models that are both valid statistical models, and for which computational efficient inference is available.

Another interesting direction would be combining the ideas of nested SPDEs[17] with systems of SPDEs, and let some of the operators in the system be other operators than the Laplacian operator, e.g., the gradient. The nested SPDE approach[17] can (for commuting operators) be set up as a system of SPDEs. Fore example, the model specified by Eq. (11) can be rewritten into this system of SPDEs as follows:

$$\begin{pmatrix} \mathcal{L}_1 & 0 \\ -1 & \mathcal{L}_2 \end{pmatrix} \begin{pmatrix} x_0(\mathbf{s}) \\ x(\mathbf{s}) \end{pmatrix} = \begin{pmatrix} \omega_1(\mathbf{s}) \\ 0 \end{pmatrix}. \tag{29}$$

where $\mathcal{L}_1$, $\mathcal{L}_2$, and $x(\mathbf{s})$ are as in Eq. (11). This model can be seen as a special version of a triangular system, and if both $x_0(\mathbf{s})$ and $x(\mathbf{s})$ are of interest, this can be seen as a bivariate spatial model. An interesting extension would be to also include a random noise term $\omega_2(\mathbf{s})$. Another operator that might be useful for off-diagonal elements is to use the shift operator to give a asymmetric multivariate covariance structure.[35]

For systems of SPDEs to be a powerful modeling tool, we need to be able to interpret the models. Today we only have good interpretations of the simplest models. We are neither able to fully interpret the model of the 2 × 2 system of SPDEs in Eq. (12), nor the model specified by the triangular 2 × 2 system in Eq. (21). Research that leads to better understanding of the models would therefore be very valuable. Some interpretations might come from physics, or from relations to spatial–temporal models defined through SPDEs.[24,51]

Furthermore, we need to gain understanding of identifiability issues. We do not know how much data of what kind that are required to identify parameters for different systems of SPDEs, and when replicates are required. In a Bayesian setting, one can claim that identifiability is not an issue, since proper priors will give proper posteriors. But these issues are closely related to how to set priors, and prior sensitivity.

Currently, there is no easy-to-use software available for systems of SPDEs. For systems of SPDEs to be a realistic options for applies users, this is needed. Inference for hierarchical systems of SPDEs fits the INLA framework for fast Bayesian inference. Using this connection, the modeling framework can be extended to non-Gaussian likelihood functions, e.g., Bernoulli, binominal, Poisson, or beta likelihood.

CONCLUSION

The SPDE approach to spatial modeling is a tool for solving two of the main challenges in spatial statistics, computational cost and extensions of spatial models to flexible (and) multivariate models. Systems of SPDEs open new opportunities while the computational benefits are kept. Especially for modeling multivariate spatial phenomena systems of SPDEs have proven to be powerful. But new classes of dependency structure, such as a oscillating covariance function, have also been formulated as a system of SPDEs. A flexible model class is indeed available through the system of SPDEs, and many models are yet to be explored.

A challenge in (multivariate) spatial statistics is identifiability of parameters. The SPDE approach has so far not contributed directly toward solving and understanding this challenge. Therefore, caution has to be taken when more complicated models, e.g., nonstationary multivariate models, are explored. But the system of SPDE approach enables fast sampling and inference. This makes testing of identification properties through simulation studies feasible.

FURTHER READINGS

Bolin D. Models and methods for random fields in spatial statistics with computational efficiency from Markov properties. PhD Thesis, Lund University, 2012.

Ingebrigtsen R. Bayesian spatial modelling of non-stationary processes and misaligned data utilising Markov properties for computational efficiency. PhD Thesis, Norwegian University of Science and Technology, 2015.

Fuglstad GA. Modelling spatial non-stationarity. PhD Thesis, Norwegian University of Science and Technology, 2015.

Lindgren G. *Stationary Stochastic Processes: Theory and Applications*. Boca Raton: CRC Press; 2012.

Lindgren G, Rootzén H, Sandsten M. *Stationary Stochastic Processes for Scientists and Engineers*. Boca Raton: CRC press; 2013.

REFERENCES

1. Tobler WR. A computer movie simulating urban growth in the Detroit region. *Econ Geogr* 1970, 46:234–240.
2. Wikle CK. Modern perspectives on statistics for spatio-temporal data. *Wiley Interdiscip Rev Comput Stat* 2015, 7:86–98.
3. Ahl RS, Woods SW, Zuuring HR. Hydrologic calibration and validation of swat in a snow-dominated Rocky Mountain watershed, Montana, USA. *J Am Water Resour Assoc* 2008, 44:1411–1430.
4. Troin M, Caya D. Evaluating the SWAT's snow hydrology over a Northern Quebec watershed. *Hydrol Process* 2014, 28:1858–1873.
5. Shao J, Lister PJ. An automated nowcasting model of road surface temperature and state for winter road maintenance. *J Appl Meteorol* 1996, 35:1352–1361.
6. Berrocal VJ, Raftery AE, Gneiting T, Steed RC. Probabilistic weather forecasting for winter road maintenance. *J Am Stat Assoc* 2010, 105:105.
7. Banerjee S, Carlin BP, Gelfand AE. *Hierarchical Modeling and Analysis for Spatial Data*. Boca Raton: Chapman & Hall; 2004.
8. Stein ML. *Interpolation of Spatial Data: Some Theory for Kriging*. New York: Springer-Verlag; 1999.
9. Whittle P. On stationary processes in the plane. *Biometrika* 1954, 41:434–449.
10. Whittle P. Stochastic processes in several dimensions. *Bull Int Stat Inst* 1963, 40:974–994.
11. Lindgren F, Rue H, Lindström J. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *J R Stat Soc Series B Stat Methodol* 2011, 73:423–498.
12. Rue H, Held L. *Gaussian Markov random Fields: Theory and Applications*. Boca Raton: Chapman & Hall; 2005.
13. Bolin D, Lindgren F. A comparison between Markov approximations and other methods for large spatial data sets. *Comput Stat Data Anal* 2013, 61:7–21.
14. Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J R Stat Soc Series B Stat Methodol* 2009, 71:319–392.
15. Ingebrigtsen R, Lindgren F, Steinsland I. Spatial models with explanatory variables in the dependence structure. *Spatial Stat* 2014, 8:20–38.
16. Fuglstad GA, Lindgren FK, Simpson DP, Rue H. Exploring a new class of non-stationary spatial Gaussian random fields with varying local anisotropy. *Stat Sin* 2015, 14:115–133.
17. Bolin D, Lindgren F. Spatial models generated by nested stochastic partial differential equations, with an application to global ozone mapping. *Ann Appl Stat* 2011, 5:523–550.
18. Bolin D. Spatial Matérn fields driven by non-Gaussian noise. *Scand J Stat* 2014, 41:557–579.
19. Wallin J, Bolin D. Geostatistical modelling using non-Gaussian Matérn fields. *Scand J Stat* 2015, 42:872–890.

20. Cameletti M, Lindgren F, Simpson D, Rue H. Spatio-temporal modeling of particulate matter concentration through the SPDE approach. *AStA Adv Stat Anal* 2013, 97:109–131.
21. Blangiardo M, Cameletti M. *Spatial and Spatio-Temporal Bayesian Models with R-INLA*. Chichester, UK: John Wiley & Sons; 2015.
22. Berliner LM. Physical-statistical modeling in geophysics. *J GeophysRes: Atmos* 1984–2012, 108:2003.
23. Whittle P. *Systems in Stochastic Equilibrium*. New York: John Wiley & Sons; 1986.
24. Cressie N, Wikle CK. *Statistics for Spatio-Temporal Data*, vol. 465. Hoboken, NJ: Wiley; 2011.
25. Hu X, Lindgren F, Simpson D, Rue H. Multivariate gaussian random fields with oscillating covariance functions using systems of stochastic partial differential equations. 2013. Available at: http://arxiv.org/abs/1307.1384.
26. Hu X, Lindgren F, Simpson D, Rue H. Multivariate Gaussian random fields using systems of stochastic partial differential equations. 2013. Available at: http://arxiv.org/abs/1307.1379.
27. Diggle PJ, Ribeiro PJ Jr. *Model-Based Geostatistics*. New York: Springer; 2006.
28. Gneiting T, Kleiber W, Schlather M. Matérn cross-covariance functions for multivariate random fields. *J Am Stat Assoc* 2010, 105:1167–1177.
29. Apanasovich TV, Genton MG, Sun Y. A valid Matérn class of cross-covariance functions for multivariate random fields with any number of components. *J Am Stat Assoc* 2012, 107:180–193.
30. Genton MG, Kleiber W. Cross-covariance functions for multivariate geostatistics. *Stat Sci* 2015, 30:147–163.
31. Wackernagel H. *Multivariate Geostatistics: An Introduction with Applications*. Berlin: Springer-Verlag; 2003.
32. Ver Hoef JM, Barry RP. Constructing and fitting models for cokriging and multivariable spatial prediction. *J Stat Plan Inference* 1998, 69:275–294.
33. Majumdar A, Gelfand AE. Multivariate spatial modeling for geostatistical data using convolved covariance functions. *Math Geol* 2007, 39:225–245.
34. Apanasovich TV, Genton MG. Cross-covariance functions for multivariate random fields based on latent dimensions. *Biometrika* 2010, 97:15–30.
35. Li B, Zhang H. An approach to modeling asymmetric multivariate spatial covariance structures. *J Multivar Anal* 2011, 102:1445–1453.
36. Gelfand AE, Diggle PJ, Fuentes M, Guttorp P. *Handbook of Spatial Statistics*. Boca Raton: CRC Press; 2010.
37. Bathe KJ. Finite element method. In: Wah B, ed. *Wiley Encyclopedia of Computer Science and Engineering*. Hoboken, NJ: John Wiley & Sons; 2007.
38. Hu X. Multivariate Gaussian random fields: the stochastic partial differential equation approach. PhD Thesis, Norwegian University of Science and Technology, 2013.
39. Cressie N, Johannesson G. Fixed rank kriging for very large spatial data sets. *J R Stat Soc Series B Stat Methodol* 2008, 70:209–226.
40. Banerjee S, Gelfand AE, Finley AO, Sang H. Gaussian predictive process models for large spatial data sets. *J R Stat Soc Series B Stat Methodol* 2008, 70:825–848.
41. Lindgren F, Rue H. Bayesian spatial modelling with R-inla. *J Stat Softw* 2015, 63:1–25.
42. Berliner LM. Hierarchical Bayesian time series models. In: Hanson KM, ed. *Maximum Entropy and Bayesian Methods*. Dordrecht: Springer; 1996, 15–22.
43. Ingebrigtsen R, Lindgren F, Steinsland I, Martino S. Estimation of a non-stationary model for annual precipitation in southern Norway using replicates of the spatial field. *Spatial Stat* 2015, 14:338–364.
44. Fuglstad G, Simpson D, Lindgren F, Rue H. Interpretable priors for hyperparameters for Gaussian random fields. 2015. Available at: http://arxiv.org/abs/1503.00256.
45. Blangiardo M, Cameletti M, Baio G, Rue H. Spatial and spatio-temporal models with R-inla. *Spat Spatio-temporal Epidemiol* 2013, 7:39–55.
46. Geirsson ÓP, Hrafnkelsson B, Simpson D. Computationally efficient spatial modeling of annual maximum 24-h precipitation on a fine grid. *Environmetrics* 2015, 26:339–353.
47. Hu X, Steinsland I, Simpson D, Martino S, Rue H. Spatial modelling of temperature and humidity using systems of stochastic partial differential equations. 2013. Available at: http://arxiv.org/abs/1307.1402.
48. Fuglstad G-A, Simpson DP, Lindgren FK, Rue H. Does non-stationary spatial data always require non-stationary random fields? *Spatial Statistics* 2015, 14:505–531.
49. Rue H. Fast sampling of Gaussian Markov random fields. *J R Stat Soc Series B Stat Methodol* 2001, 63:325–338.
50. Gelfand AE, Schmidt AM, Banerjee S, Sirmans CF. Nonstationary multivariate process modeling through spatially varying coregionalization. *Test* 2004, 13:263–312.
51. Sigrist F, Künsch HR, Stahel WA. Stochastic partial differential equation based modelling of large space–time data sets. *J R Stat Soc Series B Stat Methodol* 2015, 77:3–33.