View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by University of Bath Research Portal

Citation for published version:
Fuglstad, G-A, Lindgren, F, Simpson, D & Rue, H 2015, 'Exploring a New Class of Non-stationary Spatial Gaussian Random Fields with Varying Local Anisotropy', Statistica Sinica, vol. 25, no. 1, pp. 115-133.

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

Link to publication

University of Bath

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 13. May. 2019

Statistica Sinica **25** (2015), 115-133
doi:http://dx.doi.org/10.5705/ss.2013.106w

EXPLORING A NEW CLASS OF NON-STATIONARY SPATIAL GAUSSIAN RANDOM FIELDS WITH VARYING LOCAL ANISOTROPY

Geir-Arne Fuglstad[1], Finn Lindgren[2], Daniel Simpson[1] and Håvard Rue[1]

[1]*NTNU and* [2]*University of Bath*

Abstract: Gaussian random fields (GRFs) play an important part in spatial modelling, but can be computationally infeasible for general covariance structures. An efficient approach is to specify GRFs via stochastic partial differential equations (SPDEs) and derive Gaussian Markov random field (GMRF) approximations of the solutions. We consider the construction of a class of non-stationary GRFs with varying local anisotropy, where the local anisotropy is introduced by allowing the coefficients in the SPDE to vary with position. This is done by using a form of diffusion equation driven by Gaussian white noise with a spatially varying diffusion matrix. This allows for the introduction of parameters that control the GRF by parametrizing the diffusion matrix. These parameters and the GRF may be considered to be part of a hierarchical model and the parameters estimated in a Bayesian framework. The results show that the use of an SPDE with non-constant coefficients is a promising way of creating non-stationary spatial GMRFs that allow for physical interpretability of the parameters, although there are several remaining challenges that would need to be solved before these models can be put to general practical use.

Key words and phrases: Anisotropy, Bayesian, Gaussian random fields, Gaussian Markov random fields, non-stationary, spatial.

1. Introduction

Many spatial models for continuously indexed phenomena, such as temperature, precipitation and air pollution, are based on Gaussian random fields (GRFs). This is mainly due to the fact that their theoretical properties are well understood and that their distributions can be fully described by mean and covariance functions. In principle, it is enough to specify the mean at each location and the covariance between any two locations. However, specifying covariance functions is hard and specifying covariance functions that can be controlled by parameters in useful ways is even harder. This is the reason why the covariance function usually is selected from a class of known covariance functions such as the exponential covariance function, the Gaussian covariance function, or the Matérn covariance function.

But even when the covariance function is selected from one of these classes, the feasible problem sizes are severely limited by a cubic increase in computation time as a function of the number of observations and a quadratic increase in computation time as a function of the number of prediction locations. This computational challenge is usually tackled either by reducing the dimensionality of the problem (Cressie and Johannesson (2008); Banerjee et al. (2008)), by introducing sparsity in the precision matrix (Rue and Held (2005)) or the covariance matrix (Furrer, Genton, and Nychka (2006)), or by using an approximate likelihood (Stein, Chi, and Welty (2004); Fuentes (2007)). Sun, Li, and Genton (2012) offers comparisons of the advantages and challenges associated with the usual approaches to large spatial datasets.

We explore a new class of non-stationary GRFs that provide both an easy way to specify the parameters and allow for fast computations. The main computational tool used is Gaussian Markov random fields (GMRFs) (Rue and Held (2005)) with a spatial Markovian structure where each position is conditionally dependent only on positions close to itself. The strong connection between the Markovian structure and the precision matrix results in sparse precision matrices that can be exploited in computations. The main problem associated with such an approach is that GMRFs must be constructed through conditional distributions, which presents a challenge as it is generally not easy to determine whether a set of conditional distributions gives a valid joint distribution. Additionally, the conditional distributions have to be controlled by useful parameters in such a way that not only is the joint distribution valid, but also that the effect of the parameters is understood. Lastly, it is desirable that the GMRF is a consistent approximation of a GRF in the sense that when the distances between the positions decrease, the GMRF "approaches" a continuous GRF. These issues are even more challenging for non-stationary GMRFs. It is is extremely hard to specify the non-stationarity directly through conditional distributions.

There is no generally accepted way to handle non-stationary GRFs, but many approaches have been suggested. There is a large literature on methods based on the deformation method of Sampson and Guttorp (1992), where a stationary process is made non-stationary by deforming the space on which it is defined. Several Bayesian extensions of the method have been proposed (Damian, Sampson, and Guttorp (2001, 2003); Schmidt and O'Hagan (2003); Schmidt, Guttorp, and O'Hagan (2011)), but all these methods require replicated realizations which might not be available. There has been some development towards an approach for a single realization, but with a "densely" observed realization (Anderes and Stein (2008)). Other approaches use kernels which are convolved with Gaussian white noise (Higdon (1998); Paciorek and Schervish (2006)), weighted sums of stationary processes (Fuentes (2001)) and expansions into a basis such as a

wavelet basis (Nychka, Wikle, and Royle (2002)). Conceptually simpler methods have been made with "stationary windows" (Haas (1990b,a)) and with piecewise stationary Gaussian processes (Kim, Mallick, and Holmes (2005)). There has also been some progress with methods based on the spectrum of the processes (Fuentes (2001, 2002a,b)). Recently, a new type of method based on a connection between stochastic partial differential equations (SPDEs) and some classes of GRFs was proposed by Lindgren, Rue, and Lindström (2011). They use an SPDE to model the GRF and construct a GMRF approximation to the GRF for computations. An application of a non-stationary model of this type to ozone data can be found in Bolin and Lindgren (2011) and an application to precipitation data can be found in Ingebrigtsen, Lindgren, and Steinsland (2013).

This paper extends on the work of Lindgren, Rue, and Lindström (2011) and explores the possibility of constructing a non-stationary GRF by varying the local anisotropy. The interest lies both in considering the different types of structures that can be achieved, and how to parametrize the GRF and estimate the parameters in a Bayesian setting. The construction of the GRF is based on an SPDE that describes the GRF as the result of a linear filter applied to Gaussian white noise. Basically, the SPDE expresses how the smoothing of the Gaussian white noise varies at different locations. This construction bears some resemblance to the deformation method of Sampson and Guttorp (1992) in the sense that parts of the spatial variation of the linear filter can be understood as a local deformation of the space, only with an associated spatially varying variance for the Gaussian white noise. The main idea for computations is that since this filter works locally, it implies a Markovian structure on the GRF. This Markovian structure can be transferred to a GMRF which approximates the GRF, and in turn fast computations can be done with sparse matrices.

This paper presents a new type of model and the main goal is to explore what can be achieved in terms of models and inference with the model. Section 2 contains the motivation and introduction to the class of non-stationary GRFs that is studied in the other sections. The form of the SPDE that generates the class is given and it is related to more standard constructions of GMRFs. In Section 3 illustrative examples are given on both stationary and non-stationary constructions. This includes some discussion on how to control the non-stationarity of the GRF. Then in Section 4 we discuss parameter estimation for these types of models. The paper ends with discussion of extensions in Section 5, and general discussion and concluding remarks in Section 6.

2. New Class of Non-stationary GRFs

A GMRF $\boldsymbol{u}$ is usually parametrized through a mean $\boldsymbol{\mu}$ and a precision matrix $\mathbf{Q}$ such that $\boldsymbol{u} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{Q}^{-1})$. The main advantage of this formulation is that

the Markovian structure is represented in the non-zero structure of the precision matrix $\mathbf{Q}$ (Rue and Held (2005)). Off-diagonal entries are non-zero if and only if the corresponding elements of $\boldsymbol{u}$ are conditionally independent. This can be seen from the conditional properties of a GMRF,

$$\mathrm{E}(u_i|\boldsymbol{u}_{-i}) = \mu_i - \frac{1}{Q_{i,i}} \sum_{j \neq i} Q_{i,j}(u_j - \mu_j), \qquad \mathrm{Var}(u_i|\boldsymbol{u}_{-i}) = \frac{1}{Q_{i,i}},$$

where $\boldsymbol{u}_{-i}$ denotes the vector $\boldsymbol{u}$ with element i deleted. For a spatial GMRF the non-zeros of $\mathbf{Q}$ can correspond to grid-cells that are close to each other in a grid, neighbouring regions in a Besag model, and so on. However, even when this non-zero structure is determined it is not clear what values should be given to the non-zero elements of the precision matrix. This is the framework of the conditionally auto-regressive (CAR) models, whose conception predates the advances in modern computational statistics (Whittle (1954); Besag (1974)). In the multivariate Gaussian case it is clear that the requirement for a valid joint distribution is that $\mathbf{Q}$ be positive definite, not an easy condition to check.

Specification of a GMRF through its conditional properties is usually done in a somewhat ad-hoc manner. For regular grids, a process such as random walk can be constructed and the only major issue is to get the conditional variance correct as a function of step-length. For irregular grids the situation is not as clear because the conditional means and variances must depend on the varying step-lengths. In Lindgren and Rue (2008) it is demonstrated that some such constructions for second-order random walk can lead to inconsistencies as new grid points are added, and they offer a surprisingly simple construction for second-order random walk based on the SPDE

$$-\frac{\partial^2}{\partial x^2} u(x) = \sigma \mathcal{W}(x),$$

where $\sigma > 0$ and $\mathcal{W}$ is standard Gaussian white noise. If the precision matrix is chosen according to their scheme one does not have to worry about scaling as the grid is refined, as it automatically approaches the continuous second-order random walk. There is an automatic procedure to select the form of the conditional means and variances.

A one-dimensional second-order random walk is a relatively simple example of a process with the same behaviour everywhere. To approximate a two-dimensional, non-stationary GRF, a scheme would require (possibly) different anisotropy and correct conditional variance at each location. To select the precision matrix in this situation poses a large problem and there is abundant use of simple models such as a spatial moving average

$$\mathrm{E}(u_{i,j}|\boldsymbol{u}_{-\{(i,j)\}}) = \frac{1}{4}(u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1})$$

with a constant conditional variance $1/\alpha$. There are ad-hoc ways to extend such a scheme to a situation with varying step-lengths in each direction, but little theory for more irregular choices of locations.

We start with the close connection between SPDEs and some classes of GRFs as presented in Lindgren, Rue, and Lindström (2011) that is not plagued by such issues. From Whittle (1954), it is known that the SPDE

$$(\kappa^2 - \Delta)u(\boldsymbol{s}) = \mathcal{W}(\boldsymbol{s}), \qquad \boldsymbol{s} \in \mathbb{R}^2, \tag{2.1}$$

where $\kappa^2 > 0$ and $\Delta = \frac{\partial^2}{\partial s_1^2} + \frac{\partial^2}{\partial s_2^2}$ is the Laplacian, gives rise to a GRF u with the Matérn covariance function

$$r(\boldsymbol{s}) = \frac{1}{4\pi\kappa^2}(\kappa||\boldsymbol{s}||)K_1(\kappa||\boldsymbol{s}||),$$

where K_1 is the modified Bessel function of the second kind of order 1. Equation (2.1) can be extended to fractional operator orders in order to obtain other smoothness parameters in the Matérn covariance function. However, for practical applications, the true smoothness of the field is very hard to estimate from data, in particular when the model is used in combination with an observation noise model. Restricting the development to smoothness 1 in the Matérn family is therefore unlikely to be a major practical serious limitation. For practical computations the model is discretized using methods similar to those in Lindgren, Rue, and Lindström (2011), which does permit other operator orders. Integer orders are easiest but, for stationary models, fractional orders are also achievable (Lindgren, Rue, and Lindström (2011, Authors' discussion response)). For non-stationary models, techniques similar to Bolin (2013, Sec. 4.2) are possible. This means that even though we here restrict the model development to the special case in (2.1), other smoothnesses, e.g. exponential covariances, are reachable by combining the different approximation techniques.

The intriguing part, that Lindgren, Rue, and Lindström (2011) expanded upon in (2.1), is that $(\kappa^2 - \Delta)$ can be interpreted as a linear filter acting locally. This means that if the continuously indexed process u were instead represented by a GMRF $\boldsymbol{u}$ on a grid or a triangulation, with appropriate boundary conditions, one could replace this operator with a matrix, say $\mathbf{B}(\kappa^2)$, only involving neighbours of each location such that (2.1) becomes approximately

$$\mathbf{B}(\kappa^2)\boldsymbol{u} \sim \mathcal{N}(0, \mathbf{I}). \tag{2.2}$$

The matrix $\mathbf{B}(\kappa^2)$ depends on the chosen grid but, after the relationship is derived, the calculation of $\mathbf{B}(\kappa^2)$ is straightforward for any κ^2. Since $\mathbf{B}(\kappa^2)$ is sparse, the resulting precision matrix $\mathbf{Q}(\kappa^2) = \mathbf{B}(\kappa^2)^{\mathrm{T}}\mathbf{B}(\kappa^2)$ for $\boldsymbol{u}$ is also sparse. This means that by correctly discretizing the operator (or linear filter), it is

possible to devise a GMRF with approximately the same distribution as the continuously indexed GRF. And because it comes from a continuous equation one does not have to worry about changing behaviour as the grid is refined.

The class of models that are studied in this paper is the one that can be constructed from (2.1), but with anisotropy added to the Δ operator. A function $\mathbf{H}$, that gives 2×2 symmetric positive definite matrices at each position, is introduced and the operator is changed to

$$\begin{aligned}\nabla \cdot \mathbf{H}(\boldsymbol{s})\nabla = &\frac{\partial}{\partial s_1}\left(h_{11}(\boldsymbol{s})\frac{\partial}{\partial s_1}\right) + \frac{\partial}{\partial s_1}\left(h_{12}(\boldsymbol{s})\frac{\partial}{\partial s_2}\right)\\ &+ \frac{\partial}{\partial s_2}\left(h_{21}(\boldsymbol{s})\frac{\partial}{\partial s_1}\right) + \frac{\partial}{\partial s_2}\left(h_{22}(\boldsymbol{s})\frac{\partial}{\partial s_2}\right).\end{aligned}$$

This induces different strengths of local dependence in different directions, which results in a range that varies with direction at all locations. Further, it is necessary for the discretization procedure to restrict the SPDE to a bounded domain. The chosen SPDE is

$$(\kappa^2 - \nabla \cdot \mathbf{H}(\boldsymbol{s})\nabla)u(\boldsymbol{s}) = \mathcal{W}(\boldsymbol{s}), \qquad \boldsymbol{s} \in \mathcal{D} = [A_1, B_1] \times [A_2, B_2] \subset \mathbb{R}^2, \tag{2.3}$$

where the rectangular domain makes it possible to use periodic boundary conditions. Neither the rectangular shape of the domain nor the periodic boundary conditions are essential restrictions for the model, but are merely the practical restrictions we choose to work with, in order to focus on the non-stationarity itself.

When using periodic boundary conditions when approximating the likelihood of a stationary process on an unbounded domain, the parameter estimates are biased, e.g., when using the Whittle likelihood in the two-dimensional case (Dahlhaus and Künsch (1987)). However, as Lindgren, Rue, and Lindström (2011, Appendix A.4) notes for the case with Neumann boundary conditions, normal derivatives set to zero, the effect of the boundary conditions is limited to a region in the vicinity of the boundary. At a distance greater than twice the correlation range away from the boundary the bounded domain model is nearly indistinguishable from the model on an unbounded domain. Therefore, the bias due to boundary effects can be eliminated by embedding the domain of interest into a larger region, in effect moving the boundary away from where it would influence the likelihood function. For non-stationary models, defining appropriate boundary conditions becomes part of the practical model formulation itself. For simplicity we ignore this issue here, leaving boundary specification for future development, but provide some additional practical comments in Section 5.

Both for interpretation and the practical use of (2.3) it is useful to decompose $\mathbf{H}$ into scalar functions. The anisotropy due to $\mathbf{H}$ is decomposed as $\mathbf{H}(\boldsymbol{s}) =$

$\gamma \mathbf{I}_2 + \boldsymbol{v}(\boldsymbol{s})\boldsymbol{v}(\boldsymbol{s})^{\mathrm{T}}$, where γ specifies the isotropic, baseline effect, and the vector field $\boldsymbol{v}(\boldsymbol{s}) = [v_x(\boldsymbol{s}), v_y(\boldsymbol{s})]^{\mathrm{T}}$ specifies the direction and magnitude of the local, extra anisotropic effect at each location. In this way, one can, loosely speaking, think of different Matérn-like fields locally each with its own anisotropy that are combined into a full process. An example of an extreme case of a process with a strong local anisotropic effect is shown in Example 2. The example shows that there is a close connection between the vector field and the resulting covariance structure of the GRF.

The main computational challenge is to determine the appropriate discretization of the SPDE in (2.3), that is how to derive a matrix $\mathbf{B}$ such as in (2.2). The idea is to look to the field of numerics for discretization methods for differential equations, then to combine these with properties of Gaussian white noise. We use that for a Lebesgue measurable subset A of $\mathbb{R}^n$, for some $n > 0$,

$$\int_A \mathcal{W}(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s} \sim \mathcal{N}(0, |A|),$$

where $|A|$ is the Lebesgue measure of A, and for two disjoint Lebesgue measurable subsets A and B of $\mathbb{R}^n$ the integral over A and the integral over B are independent (Adler and Taylor (2007, pp. 24–25)). A matrix equation such as (2.2) was derived for the SPDE in (2.3) with a finite volume method. The derivations are quite involved and technical and are given in a supplementary document. However, when the form of the discretized SPDE has been derived as an expression of the coefficients in the SPDE and the grid, the conversion from SPDE to GMRF is automatic for any choice of coefficients and rectangular domain.

3. Examples of Models

The simplest case of (2.3) is with constant coefficients. In this case one has an isotropic model (up to boundary effects) if $\mathbf{H}$ is a constant times the identity matrix or a stationary anisotropic model (up to boundary effects) if this is not the case. In both cases it is possible to calculate an exact expression for the covariance function and the marginal variance for the corresponding SPDE solved over $\mathbb{R}^2$.

For this purpose write

$$\mathbf{H} = \begin{bmatrix} H_1 & H_2 \\ H_2 & H_3 \end{bmatrix},$$

where H_1, H_2, and H_3 are constants. This gives the SPDE

$$\left[\kappa^2 - H_1 \frac{\partial^2}{\partial x^2} - 2H_2 \frac{\partial^2}{\partial x \partial y} - H_3 \frac{\partial^2}{\partial y^2}\right] u(\boldsymbol{s}) = \mathcal{W}(\boldsymbol{s}), \qquad \boldsymbol{s} \in \mathbb{R}^2.$$

If λ_1 and λ_2 are the eigenvalues of $\mathbf{H}$, then the solution of the SPDE is actually only a rotated version of the solutions of

$$\left[\kappa^2 - \lambda_1 \frac{\partial^2}{\partial \tilde{x}^2} - \lambda_2 \frac{\partial^2}{\partial \tilde{y}^2}\right] u(\boldsymbol{s}) = \mathcal{W}(\boldsymbol{s}), \qquad \boldsymbol{s} \in \mathbb{R}^2. \tag{3.1}$$

Here the new x-axis is parallel to the eigenvector of $\mathbf{H}$ corresponding to λ_1 in the old coordinate system and the new y-axis is parallel to the eigenvector of $\mathbf{H}$ corresponding to λ_2.

The marginal variance of u is

$$\sigma_m^2 = \frac{1}{4\pi\kappa^2\sqrt{\det(\mathbf{H})}} = \frac{1}{4\pi\kappa^2\sqrt{\lambda_1\lambda_2}}.$$

A proof is given in the supplementary material. One can think of the eigenvectors of $\mathbf{H}$ as the two principal directions and λ_1 and λ_2 as a measure of the "strength" of the diffusion in these principal directions. Additionally, if $\lambda_1 = \lambda_2$, which is equivalent to $\mathbf{H}$ being equal to a constant times the identity matrix, the SPDE is rotation and translation invariant and the solution is isotropic. If $\lambda_1 \neq \lambda_2$, the SPDE is still translation invariant, but not rotation invariant, and the solutions are stationary, but not isotropic.

In our case the domain is not $\mathbb{R}^2$, but $[0, A] \times [0, B]$ with periodic boundary conditions. This means that a boundary effect is introduced and the above results are only approximately true.

3.1. Stationary models

For a constant $\mathbf{H}$ the SPDE in (2.3) becomes

$$[\kappa^2 - \nabla \cdot \mathbf{H}\nabla]u(\boldsymbol{s}) = \mathcal{W}(\boldsymbol{s}), \qquad \boldsymbol{s} \in [0, A] \times [0, B].$$

This SPDE can be rewritten as

$$[1 - \nabla \cdot \hat{\mathbf{H}}\nabla]u(\boldsymbol{s}) = \sigma\mathcal{W}(\boldsymbol{s}), \qquad \boldsymbol{s} \in [0, A] \times [0, B], \tag{3.2}$$

where $\hat{\mathbf{H}} = \mathbf{H}/\kappa^2$ and $\sigma = 1/\kappa^2$. From this form it is clear that σ is only a scale parameter and that it is enough to solve for $\sigma = 1$ and then multiply the solution with the desired value of σ. Therefore, it is the effect of $\hat{\mathbf{H}}$ that is most interesting to study.

It is useful to parametrize $\hat{\mathbf{H}}$ as $\hat{\mathbf{H}} = \gamma\mathbf{I}_2 + \beta\boldsymbol{v}(\theta)\boldsymbol{v}(\theta)^{\mathrm{T}}$, where $\boldsymbol{v}(\theta) = [\cos(\theta), \sin(\theta)]^{\mathrm{T}}$, $\gamma > 0$, and $\beta > 0$. In this parametrization one can think of γ as the coefficient of the second order derivative in the direction orthogonal to $\boldsymbol{v}(\theta)$ and $\gamma + \beta$ as the coefficient of the second order derivative in the direction $\boldsymbol{v}(\theta)$. Ignoring boundary effects, γ and $\gamma + \beta$ are the coefficients of the second order derivatives in (3.1) and θ is how much the coordinate system has been rotated in the positive direction.

Example 1 (Stationary GMRF)**.** Here we consider the effects of using a constant $\hat{\mathbf{H}}$. Use the SPDE in (3.2) with domain $[0, 20] \times [0, 20]$ and periodic boundary conditions, and discretize with a regular 200×200 grid. Two different values of $\hat{\mathbf{H}}$ are used, an isotropic case with $\hat{\mathbf{H}} = \mathbf{I}_2$ and an anisotropic case with $\gamma = 1$, $\beta = 8$, and $\theta = \pi/4$. The anisotropic case corresponds to a coefficient 9 in the x-direction and a coefficient 1 in the y-direction, and then a rotation of $\pi/4$ in the positive direction. The isotropic GMRF has marginal variances 0.0802 and the anisotropic GMRF has marginal variances 0.0263.

Figure 1 shows one realization for each of the cases. Comparing Figure 1(a) and Figure 1(b) it seems that the direction with the higher coefficient for the second-order derivative has longer range and more regular behaviour. Compared to the corresponding partial differential equation (PDE) without the white noise, this is what one would expect since large values of the coefficient penalize large values of the second order derivatives.

One expects that the correlation range increases when the coefficient is increased. This is in fact what happens. Figure 2 shows the correlation of the variable at $(9.95, 9.95)$ with every other point in the grid for the isotropic and the anisotropic case. This is sufficient to describe all the correlations since the solutions are stationary. One can see that the iso-correlation curves are close to ellipses with semi-axes along $\boldsymbol{v}(\theta)$ and the direction orthogonal to $\boldsymbol{v}(\theta)$, and that the correlation decreases most slowly and most quickly in the directions used to specify $\hat{\mathbf{H}}$, with slowest decrease along $\boldsymbol{v}(\theta)$. It is interesting that the isotropic and the non-isotropic cases have approximately the same length for the minor semi-axis of the iso-correlation curves, and that the major semi-axis is longer for the anisotropic case, lengths being connected with $\sqrt{\gamma}$ and $\sqrt{\gamma + \beta}$.

The use of 3 parameters thus allows for the creation of GMRFs that are more regular in one direction than the other. One can use the parameters γ, β, and θ to control the form of the correlation function, and σ to get the desired marginal variance.

3.2. Non-stationary models

To make the solution of the SPDE in (2.3) non-stationary, either κ^2 or $\mathbf{H}$ has to be a non-constant function. One way to achieve non-stationarity is by choosing $\mathbf{H}(\boldsymbol{s}) = \gamma \mathbf{I}_2 + \beta \boldsymbol{v}(\boldsymbol{s}) \boldsymbol{v}(\boldsymbol{s})^{\mathrm{T}}$, where $\boldsymbol{v}$ is a non-constant vector field on $[0, A] \times [0, B]$ satisfying the periodic boundary conditions and $\gamma > 0$ and $\beta > 0$ are constants.

Example 2 (Non-stationary GMRF)**.** Use the domain $[0, 20]^2$ with a 200×200 grid and periodic boundary conditions for the SPDE in (2.3). Let κ^2 be 1 and let $\mathbf{H}$ be given as $\mathbf{H}(\boldsymbol{s}) = \gamma \mathbf{I}_2 + \beta \boldsymbol{v}(\boldsymbol{s}) \boldsymbol{v}(\boldsymbol{s})^{\mathrm{T}}$, where $\boldsymbol{v}$ is a 2-dimensional vector

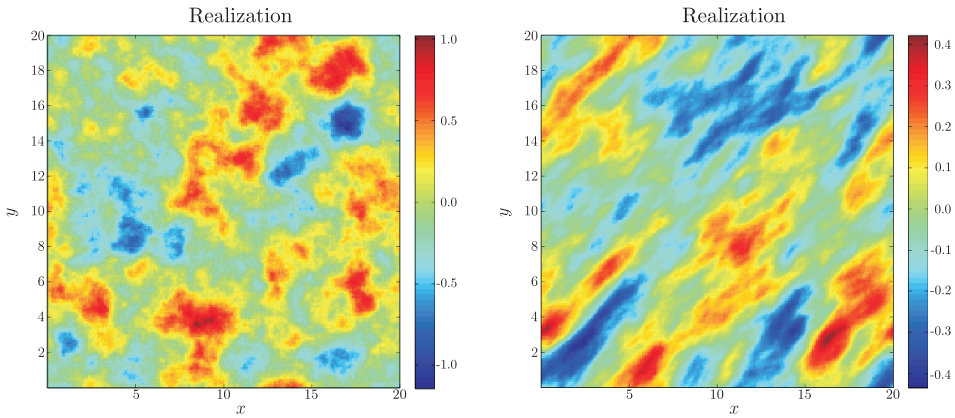


Figure 1. 1(a) Realization from the SPDE in Example 1 on $[0, 20]^2$ with a 200×200 grid and periodic boundary conditions with $\gamma = 1$, $\beta = 0$, and $\theta = 0$. 1(b) Realization from the SPDE in Example 1 on $[0, 20]^2$ with a 200×200 grid and periodic boundary conditions with $\gamma = 1$, $\beta = 8$, and $\theta = \pi/4$.

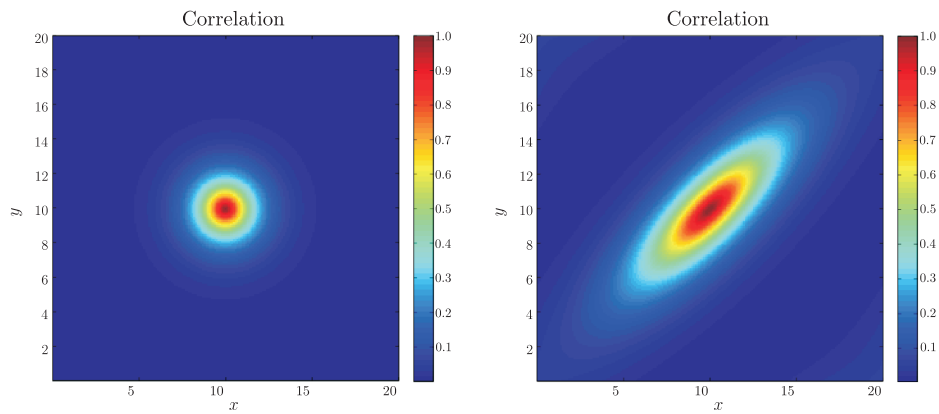
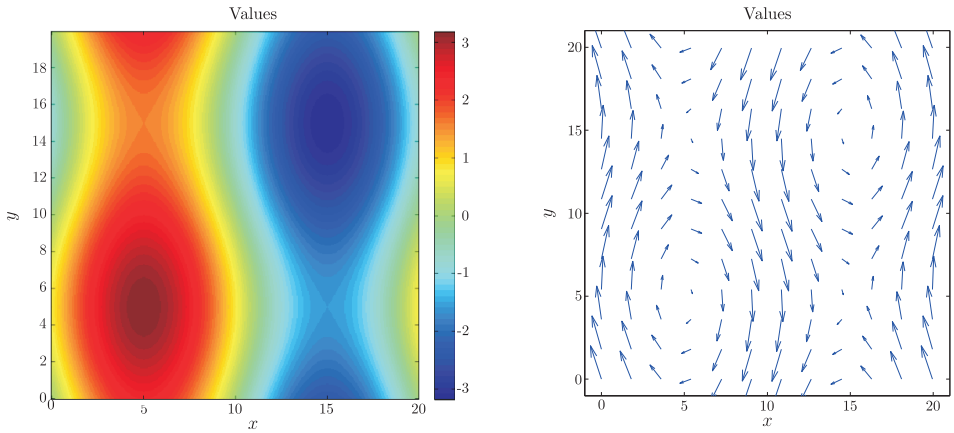


Figure 2. (a) Correlation of the centre with all other points for the solution of the SPDE in Example 1 on $[0, 20]^2$ with a 200×200 grid and periodic boundary conditions with $\gamma = 1$, $\beta = 0$, and $\theta = 0$. (b) Correlation of the centre with all other points for the SPDE in Example 1 on $[0, 20]^2$ with a 200×200 grid and periodic boundary conditions with $\gamma = 1$, $\beta = 8$, $\theta = \pi/4$.

field on $[0, 20]^2$ which satisfies the periodic boundary conditions and $\gamma > 0$ and $\beta > 0$ are constants.

To create an interesting vector field, start with the function $f : [0, 20]^2 \to \mathbb{R}$ defined by

$$f(x, y) = \left(\frac{10}{\pi}\right)\left(\frac{3}{4}\sin(\frac{2\pi x}{20}) + \frac{1}{4}\sin(\frac{2\pi y}{20})\right).$$



(a) The function used to create the vector field.

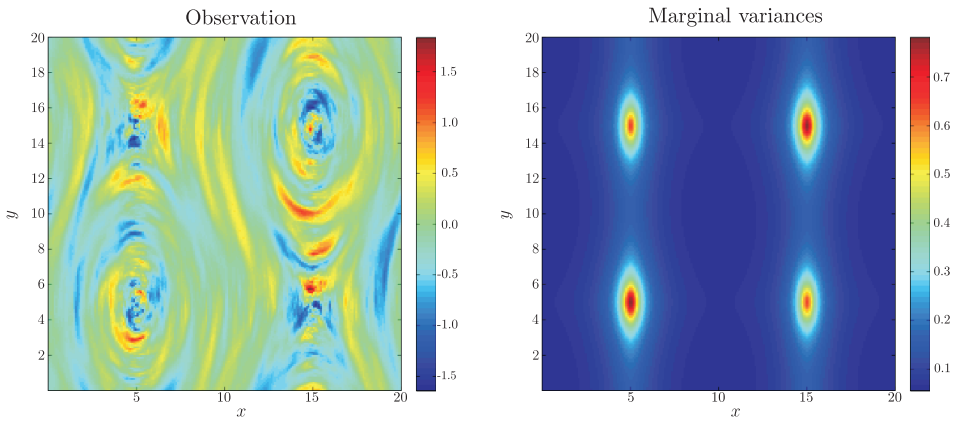
(b) The resulting vector field.

Figure 3. The gradient of the function illustrated in 3(a) is calculated and rotated 90° counter-clockwise at each point to give the vector field illustrated in 3(b).

Then calculate the gradient ∇f and let $\boldsymbol{v} : [0, 20]^2 \to \mathbb{R}^2$ be the gradient rotated 90° counter-clockwise at each point. Figure 3(a) shows the values of the function f and Figure 3(b) shows the resulting vector field $\boldsymbol{v}$. The vector field is calculated on a 400×400 regular grid, because the values between neighbouring cells in the discretization are needed.

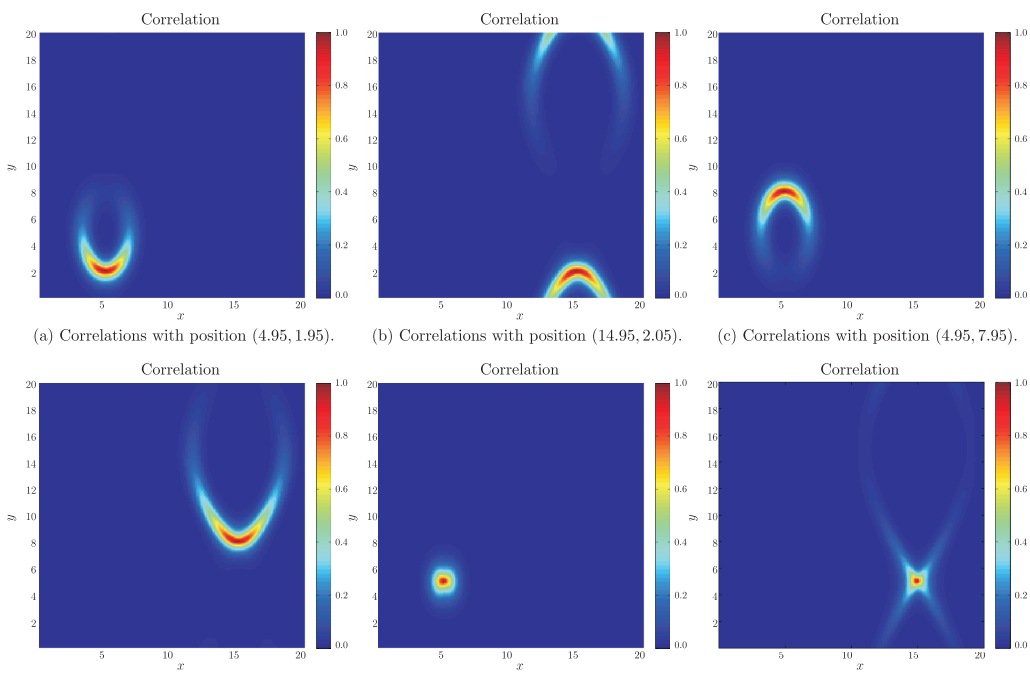
Figure 4(a) shows one realization from the resulting GMRF with $\gamma = 0.1$ and $\beta = 25$. A much higher value for β than γ is chosen to illustrate the connection between the vector field and the resulting covariance structure. From the realization it is clear that there is stronger dependence along the directions of the vector field shown in Figure 3(b) at each point than in the other directions. In addition, from Figure 4(b) it seems that positions with large values for the norm of the vector field have smaller marginal variance than positions with small values, and vice versa. This feature introduces an undesired connection between anisotropy and marginal variances. It is possible to reduce this interaction by reformulating the controlling SPDE, as discussed briefly in Section 5.

From Figure 5 one can see that the correlations depend on the direction and norm of the vector field, and that there is clearly non-stationarity. Figure 5(a) and Figure 5(c) show that the correlations with the positions $(4.95, 1.95)$ $(4.95, 7.95)$ tend to follow the vector field around the point $(5, 5)$, whereas Figure 5(b) and Figure 5(d) show that the correlations with the positions $(14.95, 1.95)$ and $(14.95, 7.95)$ tend to follow the vector field away from the point $(15, 5)$. Figure 5(e) shows that the correlations with position $(4.95, 4.95)$ and every other point are not isotropic, but concentrated close to the point itself, and Figure 5(f)



(a) One realization. (b) Marginal variances.

Figure 4. One observation and the marginal variances of the solution of the SPDE in Equation (2.3) on a 200×200 regular grid of $[0, 20]^2$ with periodic boundary conditions, $\kappa^2 \equiv 1$ and $\mathbf{H} = 0.1\mathbf{I}_2 + 25\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}$, where $\boldsymbol{v}$ is the vector field described in Example 2.



(a) Correlations with position $(4.95, 1.95)$. (b) Correlations with position $(14.95, 2.05)$. (c) Correlations with position $(4.95, 7.95)$.

(d) Correlations with position $(14.95, 7.95)$. (e) Correlations with position $(4.95, 4.95)$. (f) Correlations with position $(14.95, 4.95)$.

Figure 5. Correlations for different points with all other points for the solution of the SPDE in Example 2.

shows that the correlations with position $(14.95, 4.95)$ have high correlation along four directions that extend out from the point.

We see that allowing $\mathbf{H}$ to be non-constant can vary the dependence structure in more interesting ways than in stationary anisotropic fields. Using a vector field to control how $\mathbf{H}$ varies means that the resulting correlation structure can be partially visualized from the vector field. When $\gamma > 0$ this construction guarantees that $\mathbf{H}$ is everywhere positive definite.

4. Inference

4.1. Posterior distribution and parametrization

For inference, we introduce parameters that control the behaviour of the coefficients in (2.3) and, in turn, the behaviour of the GMRF. This is done by expanding each of the functions in a basis and using a linear combination of the basis functions weighted by parameters. For κ^2 only one parameter, say θ_1, is needed as it is assumed constant, but for the function $\mathbf{H}$ a vector of parameters $\boldsymbol{\theta}_2$ is needed. Set $\boldsymbol{\theta} = (\theta_1, \boldsymbol{\theta}_2^{\mathrm{T}})$ and give it a prior $\boldsymbol{\theta} \sim \pi(\boldsymbol{\theta})$. Then for each value of $\boldsymbol{\theta}$, a discretization, described in the supplementary document, is used to construct the GMRF $\boldsymbol{u}|\boldsymbol{\theta} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}(\boldsymbol{\theta})^{-1})$. Combine the prior of $\boldsymbol{\theta}$ with this conditional distribution to find the joint distribution of the parameters and $\boldsymbol{u}$. With a model for how an observation $\mathbf{y}$ is made from the underlying GMRF, one forms a hierarchical spatial model. The relationship between $\boldsymbol{y}$ and $\boldsymbol{u}$ is chosen to be particularly simple, namely that linear combinations of $\mathbf{u}$ are observed with Gaussian noise, $\boldsymbol{y}|\boldsymbol{u} \sim \mathcal{N}(\mathbf{A}\boldsymbol{u}, \mathbf{Q}_{\mathrm{N}}^{-1})$, where $\mathbf{Q}_{\mathrm{N}}$ is a known precision matrix.

The purpose of the hierarchical model is to do inference on $\boldsymbol{\theta}$ based on an observation of $\boldsymbol{y}$. With a Gaussian latent model it is possible to integrate out the latent field $\boldsymbol{u}$ exactly and this leads to the log-posterior

$$\begin{aligned}\log(\pi(\boldsymbol{\theta}|\boldsymbol{y})) = \mathrm{Const} + \log(\pi(\boldsymbol{\theta})) + \frac{1}{2}\log(|\mathbf{Q}(\boldsymbol{\theta})|)\\ -\frac{1}{2}\log(|\mathbf{Q}_{\mathrm{C}}(\boldsymbol{\theta})|) + \frac{1}{2}\boldsymbol{\mu}_C(\boldsymbol{\theta})^{\mathrm{T}}\mathbf{Q}_{\mathrm{C}}(\boldsymbol{\theta})\boldsymbol{\mu}_C(\boldsymbol{\theta}),\end{aligned} \tag{4.1}$$

where $\mathbf{Q}_{\mathrm{C}}(\boldsymbol{\theta}) = \mathbf{Q}(\boldsymbol{\theta}) + \mathbf{A}^{\mathrm{T}}\mathbf{Q}_{\mathrm{N}}\mathbf{A}$ and $\boldsymbol{\mu}_{\mathrm{C}}(\boldsymbol{\theta}) = \mathbf{Q}_{\mathrm{C}}(\boldsymbol{\theta})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{Q}_{\mathrm{N}}\boldsymbol{y}$. In (4.1) one sees that the posterior distribution of $\boldsymbol{\theta}$ contains terms that are hard to handle analytically. It is hard to say anything about both the determinants and the quadratic term as functions of $\boldsymbol{\theta}$. Therefore, the inference is done numerically. The model is of a form that could be handled by the INLA methodology (Rue, Martino, and Chopin (2009)), but when this was written, the R-INLA software did not have the model implemented. Instead the parameters are estimated with maximum a posteriori estimates based on the posterior density given in (4.1).

Standard deviations are estimated from the square roots of the diagonal elements of the observed information matrix.

The parametrization of $\mathbf{H}$ in the previous section employs a pre-defined vector field and a parameter β that controls the magnitude of the anisotropy due to this vector field. This is a useful representation for achieving a desired dependence structure, but in a inference setting there may not be any pre-defined vector field. Therefore, the vector field itself must be estimated. In this context the decomposition, $\mathbf{H}(\boldsymbol{s}) = \gamma \mathbf{I}_2 + \boldsymbol{v}(\boldsymbol{s})\boldsymbol{v}(\boldsymbol{s})^{\mathrm{T}}$, is more useful. For inference it is necessary to control the vector field by a finite number of parameters. The simple case of a constant matrix requires 3 parameters. Use parameters γ, v_1, and v_2 and write

$$\mathbf{H}(\boldsymbol{s}) \equiv \gamma \mathbf{I}_2 + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \begin{bmatrix} v_1 & v_2 \end{bmatrix}.$$

If $\mathbf{H}$ is not constant, it is necessary to parametrize the vector field $\boldsymbol{v}$ in some manner. Any vector field is possible for $\boldsymbol{v}$, so a basis that can generate any vector field is desirable. The Fourier basis possesses this property, but is only one of many possible choices. Let the domain be $[0, A] \times [0, B]$ and assume that $\boldsymbol{v}$ is a differentiable, periodic vector field on the domain. Then each component of the vector field can be written as a Fourier series of the form

$$\sum_{(k,l)\in\mathbb{Z}^2} C_{k,l} \exp\left[2\pi i \left(\frac{k}{A}x + \frac{l}{B}y\right)\right],$$

where i is the imaginary unit. But since the components are real-valued, each of them can also be written as a real 2-dimensional Fourier series of the form

$$A_{0,0} + \sum_{(k,l)\in E} \left[A_{k,l} \cos\left[2\pi \left(\frac{k}{A}x + \frac{l}{B}y\right)\right] + B_{k,l} \sin\left[2\pi \left(\frac{k}{A}x + \frac{l}{B}y\right)\right]\right],$$

where the set $E \subset \mathbb{Z}^2$ is given by $E = (\mathbb{N} \times \mathbb{Z}) \cup (\{0\} \times \mathbb{N})$.

Putting these Fourier series together gives

$$\begin{aligned} \boldsymbol{v}(\boldsymbol{s}) = \begin{bmatrix} A_{0,0}^{(1)} \\ A_{0,0}^{(2)} \end{bmatrix} &+ \sum_{(k,l)\in E} \begin{bmatrix} A_{k,l}^{(1)} \\ A_{k,l}^{(2)} \end{bmatrix} \cos\left[2\pi \left(\frac{k}{A}x + \frac{l}{B}y\right)\right] \\ &+ \sum_{(k,l)\in E} \begin{bmatrix} B_{k,l}^{(1)} \\ B_{k,l}^{(2)} \end{bmatrix} \sin\left[2\pi \left(\frac{k}{A}x + \frac{l}{B}y\right)\right], \end{aligned}$$

where $A_{k,l}^{(1)}$ and $B_{k,l}^{(1)}$ are the coefficients for the first component of $\boldsymbol{v}$ and $A_{k,l}^{(2)}$ and $B_{k,l}^{(2)}$ are the coefficients of the second component. This gives 2 coefficients when only the zero-frequency is included, then 18 parameters when the $(0, 1)$, $(1, -1)$,

$(1,0)$ and $(1,1)$ frequencies are included. When the number of frequencies used in each direction doubles, the number of required parameters quadruples.

4.2. Inference on simulated data

In the supplementary material, we consider data generated from known sets of parameters for models of the type

$$u(\boldsymbol{s}) - \nabla \cdot \mathbf{H}(\boldsymbol{s})\nabla u(\boldsymbol{s}) = \mathcal{W}(\boldsymbol{s}),$$

where $\mathcal{W}$ is a standard Gaussian white noise process and $\mathbf{H}(\cdot)$ is a spatially varying 2×2 matrix, with periodic boundary conditions on a rectangular domain. The matrix is parameterized as $\mathbf{H}(\boldsymbol{s}) = \gamma \boldsymbol{I} + \boldsymbol{v}(\boldsymbol{s})\boldsymbol{v}(\boldsymbol{s})^{\mathrm{T}}$. The results illustrate the ability to estimate the vector field controlling the anisotropy for four test cases.

These examples focus on simple cases where specific issues can be highlighted. The inherent challenges in estimating a spatially varying direction and strength are equally important in the more general setting where also κ and the baseline effect γ are allowed to vary. The estimation of the vector field presents an important component that must be dealt with in any inference strategy for the more general case.

5. Extensions

To make the model applicable to datasets it is necessary to also make the parameters κ and γ spatially varying functions. This results in some control also over the marginal variance and the strength of the local baseline component of the anisotropy at each location. A varying κ is discussed briefly in Section 3.2 in Lindgren, Rue, and Lindström (2011).

This comes at the cost of two more functions that must be inferred together with the vector field $\boldsymbol{v}$ that, in turn, means two more functions need to be expanded into bases. This could be done with a Fourier basis, but any basis which respects the boundary condition could in principle be used. The amount of freedom available by having four spatially varying functions comes at a price, and it would be necessary to introduce some apriori restrictions on their behaviour.

In the supplementary material we demonstrate the challenge with the non-identifiability of the sign of the vector field. It would be possible to make the situation less problematic by enforcing more structure in the estimated vector field. For example, through spline penalties which adds a preference for components without abrupt changes. Such apriori restrictions make sense both from a modelling perspective, in the sense that the properties should not change to quickly, and from a computational perspective, in the sense that it is desirable

to avoid situations as the one encountered in the previous section where the direction of the vector field flips.

The full model could be used in an application through a three-step approach. First, choose an appropriate basis to use for each function and select an appropriate prior. This means deciding how many basis elements one is willing to use from a computational point of view, and how strong the apriori penalties need to be. Second, find the maximum aposteriori estimate of the functions κ, γ, v_1, and v_2. Third, assume the maximum aposteriori estimates are the true functions and calculate the predicted values and prediction variances. Full details of such an approach are beyond the current scope. This is being studied in current work on an application to annual rainfall data in the conterminous US (Fuglstad, Simpson, Lindgren, and Rue (2013)).

Another way forward deals with the interactions of the functions κ, γ, v_1, and v_2. The functions interact in difficult ways to control marginal variance and to control anisotropy. As seen in Example 2 the vector field that controls the anisotropic behaviour is also linked to the marginal variances of the field. It would be desirable to try to separate the functions that are allowed to affect the marginal variances and the functions that are allowed to affect the correlation structure. This may present a useful feature in applications, both for interpretability and for constructing priors.

One promising way to reduce interaction is to extend ideas presented in Section 3.4 of Lindgren, Rue, and Lindström (2011), linking the use of an anisotropic Laplacian to the deformation method of Sampson and Guttorp (1992). The link is too restrictive, but the last comments about the connection to metric tensors leads to a useful way to rewrite the SPDE in (2.3). This is work in progress and involves interpreting the simple SPDE

$$[1 - \Delta]u = \mathcal{W}$$

as an SPDE on a Riemannian manifold with an inverse metric tensor defined through the strength of dependence in different directions, in a similar way as the spatially varying matrix $\mathbf{H}$. This leads to a slightly different SPDE where a separate function, that does not affect correlation structure, can be used to control marginal standard deviations. However, the separation is not perfect since the varying metric tensor gives a curved space and thus affects the marginal variances of the solution of the above SPDE, though the effect of the metric tensor on marginal standard deviations appears small.

Another issue which has not been addressed is how to define relevant boundary conditions. For rectangular domains, periodic boundary conditions are simple to implement, but a naive use of such conditions is typically inappropriate in practical applications due to the resulting spurious dependence between physically

distant locations. This problem can be partly rectified by embedding the region of interest into a larger covering domain. It is also possible to apply Neumann-type boundary conditions similar to the ones used by Lindgren, Rue, and Lindström (2011). These are easier to adapt to more general domains, but they still require a domain extension in order to remove the influence of the boundary condition on the likelihood. A more theoretically appealing, and computationally potentially less expensive, solution is to directly define the behaviour of the field along the boundary so that the models would contain stationary fields as a neutral case. Work is underway to design stochastic boundary conditions to accomplish this, and some of the solutions show potential for extension to non-stationary models.

6. Discussion

The paper explores different aspects of a new class of non-stationary GRFs based on local anisotropy. The benefit of the formulation presented is that it allows for flexible models with few requirements on the parameters. Since the GRF is based on an SPDE, there is no need to worry about how to change the discretized model in a consistent manner when the grid is refined. This is one of the more attractive features of the SPDE-based modelling.

The focus of the examples has been the matrix $\mathbf{H}$ introduced in the Laplace-operator. The examples show that a variety of different effects can be achieved by using different types of spatially varying matrices. As shown in Section 3, anisotropic fields have anisotropic Matérn-like covariance functions, through stretching and rotating the domain, and can be controlled by four parameters. It is possible to control the marginal variances, the principal directions, and the range in each of the principal directions. A spatially varying $\mathbf{H}$ gives non-stationary random fields. And by using a vector field to specify the strength and direction of extra spatial dependence in each location, there is a clear connection between the vector field and the resulting covariance structure. The covariance structure can be partially visualized from the vector field.

There are many avenues that are not explored here. The chief motivation is to explore a class of models for what can be achieved, and for the associated challenges of inference with the models. We show that a vector field constitutes a useful way to control local anisotropy in the SPDE-model of Lindgren, Rue, and Lindström (2011). What remains for a fully flexible spatial model is to allow κ and γ to be spatially varying functions, this is a simpler task than the anisotropy component since they do not require vector fields. For this more complex model there are four spatially varying functions to estimate and an expansion of each of these functions into a basis leads to many parameters. It remains to investigate appropriate choices of priors for use in applications. This question is connected with the discussion in Section 5 on other constructions of the model that separate

the functions allowed to affect marginal variances and the functions allowed to affect correlation structure.

Acknowledgements

The authors are grateful to the Guest Editors of the special issue, and to the referees for their helpful comments which improved the manuscript.

References

Adler, R. J. and Taylor, J. E. (2007). *Random Fields and Geometry.* Springer Verlag.

Anderes, E. B. and Stein, M. L. (2008). Estimating deformations of isotropic gaussian random fields on the plane. *Ann. Statist.* **31**, 719-741.

Banerjee, S., Gelfand, A. E., Finley, A. O. and Sang, H. (2008). Gaussian predictive process models for large spatial data sets. *J. Roy. Statist. Soc. Ser. B* **70**, 825-848.

Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *J. Roy. Statist. Soc. Ser. B* **2**, 192-236.

Bolin, D. and Lindgren, F. (2011). Spatial models generated by nested stochastic partial differential equations, with an application to global ozone mapping. *Ann. Appl. Statist.* **5**, 523-550.

Bolin, D. (2013). Spatial Matérn fields driven by non-Gaussian noise. *Scand. J. Statist.* In press.

Cressie, N. and Johannesson, G. (2008). Fixed rank kriging for very large spatial data sets. *J. Roy. Statist. Soc. Ser. B* **70**, 209-226.

Dahlhaus, R. and Künsch, H. (1987). Edge effects and efficient parameter estimation for stationary random fields. *Biometrika* **74**, 877-882.

Damian, D., Sampson, P. D. and Guttorp, P. (2001). Bayesian estimation of semi-parametric non-stationary spatial covariance structures. *Environmetrics* **12**, 161-178.

Damian, D., Sampson, P. D. and Guttorp, P. (2003). Variance modeling for nonstationary spatial processes with temporal replications. *J. Geophysical Research: Atmospheres* **108**.

Fuentes, M. (2001). A high frequency kriging approach for non-stationary environmental processes. *Environmetrics* **12**, 469-483.

Fuentes, M. (2002a). Interpolation of nonstationary air pollution processes: a spatial spectral approach. *Statistical Modelling* **2**, 281-298.

Fuentes, M. (2002b). Spectral methods for nonstationary spatial processes. *Biometrika* **89**, 197-210.

Fuentes, M. (2007). Approximate likelihood for large irregularly spaced spatial data. *J. Amer. Statist. Assoc.* **102**, 321-331.

Fuglstad, G.-A., Simpson, D., Lindgren, F. and Rue, H. (2013). Non-stationary spatial modelling with applications to spatial prediction of precipitation. arXiv:1306.0408. In preperation.

Furrer, Reinhard, Genton, Marc G. and Nychka, Douglas (2006). Covariance tapering for interpolation of large spatial datasets. *J. Comput. Graph. Statist.* **15**, 502-523.

Haas, T. C. (1990a). Kriging and automated variogram modeling within a moving window. *Atmospheric Environment. Part A. General Topics* **24**, 1759-1769.

Haas, T. C. (1990b). Lognormal and moving window methods of estimating acid deposition. *J. Amer. Statist. Assoc.* **85**, 950-963.

Higdon, D. (1998). A process-convolution approach to modelling temperatures in the North Atlantic Ocean. *Environ. and Ecolog. Statist.* **5**, 173-190.

Ingebrigtsen, R., Lindgren, F. and Steinsland, I. (2013). Spatial models with explanatory variables in the dependence structure of gaussian random fields based on stochastic partial differential equations. *Spatial Statistics.* In press.

Kim, H.-M., Mallick, B. K. and Holmes, C. C (2005). Analyzing nonstationary spatial data using piecewise Gaussian processes. *J. Amer. Statist. Assoc.* **100**, 653-668.

Lindgren, F. and Rue, H. (2008). On the second-order random walk model for irregular locations. *Scand. J. Statist.* **35**, 691-700.

Lindgren, F., Rue, H. and Lindström, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *J. Roy. Statist. Soc. Ser. B* **73**, 423-498.

Nychka, D., Wikle, C. and Royle, J. A. (2002). Multiresolution models for nonstationary spatial covariance functions. *Statist. Model.* **2**, 315-331.

Paciorek, C. J. and Schervish, M. J. (2006). Spatial modelling using a new class of nonstationary covariance functions. *Environmetrics* **17**, 483-506.

Rue, H. and Held, L. (2005). *Gaussian Markov Random Fields: Theory and Applications*, volume 104 of *Monographs on Statistics and Applied Probability.* Chapman & Hall, London.

Rue, H., Martino, S. and Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J. Roy. Statist. Soc. Ser. B* **71**, 319-392.

Sampson, P. D. and Guttorp, P. (1992). Nonparametric estimation of nonstationary spatial covariance structure. *J. Amer. Statist. Assoc.* **87**, 108-119.

Schmidt, A. M., Guttorp, P. and O'Hagan, A. (2011). Considering covariates in the covariance structure of spatial processes. *Environmetrics* **22**, 487-500.

Schmidt, A. M. and O'Hagan, A. (2003). Bayesian inference for non-stationary spatial covariance structure via spatial deformations. *J. Roy. Statist. Soc. Ser. B* **65**, 743-758.

Stein, M. L., Chi, Z. and Welty, L. J. (2004). Approximating likelihoods for large spatial data sets. *J. Roy. Statist. Soc. Ser. B* **66**, 275-296.

Sun, Y., Li, B. and Genton, M. G. (2012). Geostatistics for large datasets. *Advances and Challenges in Space-time Modelling of Natural Events*, 55-77. Springer Berlin Heidelberg.

Whittle, P. (1954). On stationary processes in the plane. *Biometrika* **41**, 434-449.

Department of Mathematical Sciences, NTNU, 7491 Trondheim, Norway.
E-mail: fuglstad@math.ntnu.no

Department of Mathematical Sciences, University of Bath, United Kingdom.
E-mail: f.lindgren@bath.ac.uk

Department of Mathematical Sciences, NTNU, 7491 Trondheim, Norway.
E-mail: dp.simpson@gmail.com

Department of Mathematical Sciences, NTNU, 7491 Trondheim, Norway.
E-mail: hrue@math.ntnu.no

(Received April 2013; accepted February 2014)