ARTICLE IN PRESS

Spatial Statistics ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Spatial Statistics

journal homepage: www.elsevier.com/locate/spasta

Spatial models with explanatory variables in the dependence structure

Rikke Ingebrigtsen [a,*], Finn Lindgren [b], Ingelin Steinsland [a]

[a] *Department of Mathematical Sciences, Norwegian University of Science and Technology (NTNU), Trondheim, N-7491, Norway*

[b] *Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY, United Kingdom*

ARTICLE INFO

Article history:
Received 16 February 2013
Accepted 25 June 2013
Available online xxxx

Keywords:
Non-stationary covariance models
Gaussian random fields
Stochastic partial differential equations
Annual precipitation
Approximate Bayesian inference

ABSTRACT

Geostatistical models have traditionally been stationary. However, physical knowledge about underlying spatial processes often requires models with non-stationary dependence structures. Thus, there has been an interest in the literature to provide flexible models and computationally efficient methods for non-stationary phenomena. In this work, we demonstrate that the stochastic partial differential equation (SPDE) approach to spatial modelling provides a flexible class of non-stationary models where explanatory variables can be easily included in the dependence structure. In addition, the SPDE approach enables computationally efficient Bayesian inference with integrated nested Laplace approximations (INLA) available through the R-package `r-inla`. We illustrate the suggested modelling framework with a case study of annual precipitation in southern Norway, and compare a non-stationary model with dependence structure governed by elevation to a stationary model. Further, we use a simulation study to explore the annual precipitation models. We investigate identifiability of model parameters and whether the deviance information criterion (DIC) is able to distinguish datasets from the non-stationary and stationary models.

* Corresponding author.
E-mail addresses: rikke.ingebrigtsen@math.ntnu.no (R. Ingebrigtsen), f.lindgren@bath.ac.uk (F. Lindgren), ingelin.steinsland@math.ntnu.no (I. Steinsland).

2211-6753/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.spasta.2013.06.002

Please cite this article in press as: Ingebrigtsen, R., et al., Spatial models with explanatory variables in the dependence structure. Spatial Statistics (2013), http://dx.doi.org/10.1016/j.spasta.2013.06.002

1. Introduction

Datasets from disciplines like meteorology and environmental sciences often contain geographical information; typically coordinates for the observation locations. Using this information in the statistical analysis enables us to learn more about the underlying spatial process. The main objectives in spatial statistics are to explore the data, make predictions at locations where the process is unobserved, learn about the process or make synthetic stochastic process generators (Banerjee et al., 2003; Gelfand et al., 2010; Cressie and Wikle, 2011).

For spatial processes that are continuous in space Gaussian random fields (GRFs) are often used, either to model the spatial phenomena directly, or as part of a Bayesian hierarchical model. A continuous random field $\{x(\boldsymbol{u}) : \boldsymbol{u} \in \mathcal{D}\}$ defined on the spatial domain $\mathcal{D} \in \mathbb{R}^2$ is a GRF if all finite dimensional distributions of the field are Gaussian. More precisely, for all $n \in \mathbb{N}$ and all choices of locations $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n \in \mathcal{D}$ the vector $(x(\boldsymbol{u}_1), \ldots, x(\boldsymbol{u}_n))$ follows a multivariate Gaussian distribution. A GRF is characterised by its mean function $\mu(\boldsymbol{u})$ and covariance function $C(\cdot, \cdot)$, where for two locations $\boldsymbol{u}_i$ and $\boldsymbol{u}_j$ the covariance between them is $\mathrm{Cov}(x(\boldsymbol{u}_i), x(\boldsymbol{u}_j)) = C(\boldsymbol{u}_i, \boldsymbol{u}_j)$. If the covariance function is only a function of the relative position between two locations, i.e. $C(\boldsymbol{u}_i, \boldsymbol{u}_j) = C(\boldsymbol{u}_i - \boldsymbol{u}_j)$, the GRF is said to be stationary.

Frequently quoted in spatial statistics is Tobler's first law of geography: *everything is related to everything else, but near things are more related than distant things* (Tobler, 1970). In statistical terms this means that near things tend to be more dependent than distant things. In a GRF this is expressed through the covariance function which gives the strength of the dependence between two locations. For a stationary field this relationship is the same throughout the domain. However, many environmental processes are non-stationary by nature. Topographical variables such as mountains, distance to the coast, lakes, or meteorological variables such as wind might influence the quantitative interpretation of what *near* is, and thus, stationary GRFs are inappropriate models.

Driven by the need for more complex dependence structures for environmental phenomena, modelling frameworks of non-stationary spatial models have been an area of active research for more than two decades. Sampson and Guttorp (1992) warped the spatial domain into a space where the process is stationary. This approach has been further developed in Schmidt and O'Hagan (2003), Schmidt et al. (2011) and Bornn et al. (2012). Another popular approach has been kernel convolution methods (e.g. Higdon et al. (1999), Fuentes (2002), Paciorek and Schervish (2006), Reich et al. (2011) and Kleiber and Nychka (2012)). Also a two stage approach with a discrete GRF controlling the variance in another discrete GRF have been used (Yue and Speckman, 2010). Some works also include covariate information in the non-stationary model; Schmidt et al. (2011) adds covariate(s) as extra dimension(s) in the space to be warped to a stationary field. Reich et al. (2011) extend the work of Fuentes (2002) such that the values of the covariates (and not only the differences) can be taken into account in the covariance function.

Common for all these approaches is that computationally efficient inference is not straightforward. Often advanced Markov Chain Monte Carlo schemes have to be used (e.g. Yue and Speckman (2010), Schmidt et al. (2011) and Reich et al. (2011)), or step-wise procedures are used (e.g. Paciorek and Schervish (2006) and Kleiber and Nychka (2012)).

In this paper, we present a method for defining spatial dependence structures using explanatory variables. The stochastic partial differential equation (SPDE) approach to spatial modelling introduced by Lindgren et al. (2011) enables Gaussian random field models with spatially varying parameters, and offers a framework where dependence structures can be motivated from physical, or other, properties of the spatial process.

Lindgren et al. (2011) used an SPDE formulation to link Gaussian random fields and Gaussian Markov random fields (GMRFs). This means that modelling can be done in continuous space with GRFs, while inference can gain computational speed-up from properties of GMRFs (see e.g. Rue and Held (2005) for an introduction to GMRFs). A GMRF is simply a multivariate Gaussian vector parametrised using the precision matrix $\boldsymbol{Q}$, which is the inverse of the covariance matrix. The term is mainly used when the precision matrix contains many zero elements, which directly corresponds to conditional independence properties, or Markov structure. Such sparse precision matrices make numerical methods involving GMRFs computationally efficient.

Please cite this article in press as: Ingebrigtsen, R., et al., Spatial models with explanatory variables in the dependence structure. Spatial Statistics (2013), http://dx.doi.org/10.1016/j.spasta.2013.06.002

The SPDE formulation was motivated by computational challenges in spatial statistics. However, the SPDE framework is not limited to numerically efficient GRFs, it is also possible to modify the SPDE to introduce new classes of spatial models. In Lindgren et al. (2011) some of these opportunities are explored; defining GRFs on manifolds (e.g. the sphere), oscillating covariance function and non-stationary models where the non-stationary field is determined by low-dimensional spatial parameter fields. It is possible to use these complex spatial models without expert knowledge about SPDEs; they are implemented in the R-package `r-inla`, which makes it seamless to define SPDE models and enables Bayesian inference with integrated nested Laplace approximations (INLA) (Rue et al., 2009).

The main contribution of this paper is to demonstrate another useful application of the SPDE formulation; We introduce a parametric model with explanatory variables in the parameters that define the SPDE model. This results in a non-stationary spatial model. We find two main advantages with our approach: (1) With a parametric model for the non-stationary dependence structure knowledge about the spatial processes can be utilised, the model can be interpreted and understanding about the process can be gained, and (2) The SPDE–GMRF–INLA links enable us to sample from the models and to do fully Bayesian inference fast. This makes it computational feasible to conduct simulation studies to further explore properties of the models. We use a simulation study to find whether we are able to distinguish between a stationary and non-stationary model, and to compare estimated parameters with parameters used to simulate datasets.

We analyse a spatial dataset of annual precipitation in southern Norway to illustrate our proposed modelling framework for non-stationary GRFs. The precipitation process is driven by complex interactions between meteorological variables such as pressure, wind, temperature, and humidity, and the topography. The process is known to be non-stationary; especially for mountainous terrain (Daly et al., 1994). Therefore, we suggest an SPDE based model where both the mean and dependence are allowed to vary with elevation.

Interpolation of precipitation in Norway has shown to be a challenging task. Classical geostatistical models (i.e. stationary models) have been explored, but did not preform well. Operationally, the Norwegian Meteorological Institute uses an interpolation method based on triangulation with a term adjusting for elevation (Jansson et al., 2007), and this model do not quantify any uncertainty. Physically based reanalysis models also have challenges reconstructing the precipitation field (Orskaug et al., 2011).

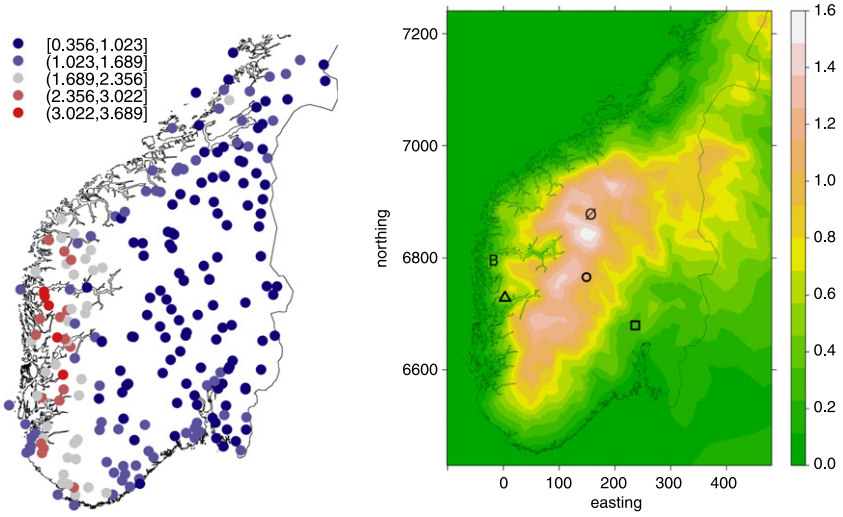
Also in statistics researchers have been interested in spatial models for precipitation, especially for precipitation accumulated over shorter time periods where zero-inflated processes must be used in the modelling (e.g. Sansó and Guenni (2000)). Both Schmidt et al. (2011) and Yue and Speckman (2010) analyse total seasonal precipitation with their respective non-stationary model.

The rest of this paper is organised as follows. Section 2 describes the annual precipitation data from southern Norway. Section 3 gives the necessary background on SPDE based spatial models, and introduces our new class of non-stationary models with explanatory variables in the dependence structure. In Section 4 we define the models used for the annual precipitation, explain the inference method (INLA) briefly, and describe how we assess the models. Results of the analysis of the annual precipitation data and a simulation study are presented in Section 5. The paper ends with a discussion in Section 6.

2. Annual precipitation in southern Norway

The motivation for studying annual precipitation comes from hydropower production. In Norway, most of the power production is hydroelectric; a renewable and clean energy source, which makes it is desirable to exploit existing power plants to their full extent. The power production is influenced by changes in the amount of precipitation from season to season, and from year to year. For optimal planning of the power production, spatial interpolation of precipitation is important, because it is used as input data in hydrological models. Most catchments are in mountain areas with rough weather conditions and poor infrastructure. This makes maintaining gauges challenging and expensive, and the areas of interest are often poorly monitored. Also, for most catchments the vast majority of the catchment is at a higher elevation than the closest precipitation gauges. Hence a good

Please cite this article in press as: Ingebrigtsen, R., et al., Spatial models with explanatory variables in the dependence structure. Spatial Statistics (2013), http://dx.doi.org/10.1016/j.spasta.2013.06.002



(a) Annual precipitation (m). (b) Smoothed elevation (km).

Fig. 1. The precipitation data are from the Norwegian Meteorological Institute and consist of observations from 233 weather stations in southern Norway. The smoothed elevation model is based on GLOBE and computed on a triangulation of the domain. The locations of the three weather stations Kvamskogen (△), Hemsedal (○), and Hønefoss (□) are indicated on the elevation map in (b). These stations will receive closer examination in Section 5.1. Also indicated on the map is Brekke (B) and Øygarden (Ø), these two weather stations have, respectively, the maximum and minimum annual normal in Norway. The coordinate reference system is UTM33 in km.

understanding of the precipitation process itself, as well as a good understanding and quantification of the interpolation uncertainty, is important.

The annual precipitation dataset we analyse in this paper was obtained from a web portal provided by the Norwegian Meteorological Institute (www.eKlima.no). Daily precipitation observations, over the one year period 2008-09-01–2009-08-31, from stations in southern Norway (up to and including the county Nord-Trøndelag) were added to obtain the annual values. Stations with incomplete records were removed. The remaining dataset consists of observations from 233 weather stations. The annual precipitation data are presented in Fig. 1a.

The data exhibit large differences in the amount of precipitation between the western and eastern parts of the country, with observations on the west coast being almost ten times as high as those from the dry inland regions. These differences are typical for the Norwegian climate, a climate characterised by extreme variations in precipitation. The wettest part of the country is the western part, where the annual normal for one of the weather stations (Brekke, see Fig. 1b) is as high as 3.573 m. The amount of precipitation in this region is among the highest in Europe. The weather station with the lowest annual normal (Øygarden, see Fig. 1b) is located in the middle of the country, and the amount of precipitation there is only 0.278 m (source: www.met.no). This large difference can be explained by the Norwegian topography and prevailing westerly winds.

Norway is a mountainous country, where the western and eastern parts of southern Norway are separated by the mountain range Langfjella. Due to the prevailing westerly winds, most of the large weather systems hitting Norway come from the west. Hence, the weather systems first reach the steep west coast of Norway, where humid air is forced to ascend because of the topography. The air will cool down and release precipitation in form of rain or snow, a phenomenon known as orographic precipitation. The eastern part of Norway is a leeward region in relation to the weather systems coming from the west, which explains the low annual normals in the "rain shadow".

Motivated by the orographic effect, we explore the use of elevation to explain the amount of precipitation. Two spatial models for the annual precipitation data are compared in this work; one where elevation is regarded as a linear effect and one where, in addition to the linear effect, we let the dependence structure of the spatial field depend on elevation. The latter model is obtained using

Please cite this article in press as: Ingebrigtsen, R., et al., Spatial models with explanatory variables in the dependence structure. Spatial Statistics (2013), http://dx.doi.org/10.1016/j.spasta.2013.06.002

the non-stationary GRF that will be introduced in Section 3.2. The two models are further described in Section 4.

A model for the topography of southern Norway is based on the 30-arc-second (1 km) gridded global digital elevation model (DEM), GLOBE (GLOBE Task Team, 1999). In Fig. 1b, the smoothed elevation model used in this paper can be seen. Fine scale details of the topography is lost using this smoothed elevation model. However, for the purpose of modelling annual precipitation in southern Norway the scale is sufficiently small to capture the orographic impact. In a more local model, and on a shorter time scale, an elevation model with higher resolution is probably of more importance.

Three weather stations have been chosen for detailed analysis; One in West-Norway: Kvamskogen–Jonshøgdi, at 455 m above sea level, where it was measured 3.101 m precipitation. One in the middle of Norway: Hemsedal–Hølto, at 648 m above sea level and annual precipitation 0.7746 m the current year. And one in East-Norway: Hønefoss–Høyby, located 140 m above sea level, which received 0.6506 m precipitation over the one year period. The three stations are indicated with different symbols on the topography map in Fig. 1b.

3. The stochastic partial differential equation approach

This section introduces the SPDE approach used in this work; Section 3.1 summarises the general methodology and Section 3.2 introduces the non-stationary GRF with explanatory variables in the covariance parameters. For further theoretical details on the approach refer to Lindgren et al. (2011), and for a more practical introductory tutorial on how to use SPDE based GRFs using the `r-inla` package, refer to Lindgren (2012).

3.1. Background

Using SPDEs in geostatistical modelling is a relatively new approach. It was introduced by Lindgren et al. (2011) and has since been extended and applied in various contexts (Bolin and Lindgren, 2011, 2013; Simpson et al., 2012a,b; Cameletti et al., 2013). However, the original idea dates back to the work of Whittle (1954, 1963), where it is shown that the solution to the following SPDE

$$(\kappa^2 - \Delta)^{\alpha/2} x(\boldsymbol{u}) = \mathcal{W}(\boldsymbol{u}), \quad \boldsymbol{u} \in \mathbb{R}^d, \alpha = \nu + d/2, \kappa > 0, \nu > 0, \tag{1}$$

is a Gaussian random field with Matérn covariance function. The innovation process $\mathcal{W}$ on the right hand side of (1) is spatial Gaussian white noise, and Δ is the Laplacian. The Matérn covariance function between locations $\boldsymbol{u}$ and $\boldsymbol{v}$ in $\mathbb{R}^d$ is given by

$$C(\boldsymbol{u}, \boldsymbol{v}) = \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)} (\kappa \|\boldsymbol{v} - \boldsymbol{u}\|)^{\nu} K_{\nu}(\kappa \|\boldsymbol{v} - \boldsymbol{u}\|), \tag{2}$$

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.

Lindgren et al. (2011) show that an approximation to the solution of the SPDE in (1) can be obtained using a finite element method (FEM), a numerical technique for solving partial differential equations (see e.g. Brenner and Scott (2008)). The approximated solution is computationally efficient for inferential purposes: instead of using a GRF with dense covariance matrix, the computations can be carried out with a GMRF with sparse precision matrix (the inverse of the covariance matrix).

The first step is to replace the infinite dimensional GRF with a finite basis function representation

$$x(\boldsymbol{u}) = \sum_{i=1}^{m} \psi_i(\boldsymbol{u}) w_i. \tag{3}$$

Here, the w_i's are weights and the ψ_i's are piece-wise linear basis functions defined on a triangulation of the domain, i.e. a subdivision into non-intersecting triangles with m nodes; see Fig. 2 for the triangulation of southern Norway that will be used for the analysis of the precipitation dataset in Section 5.1.

Please cite this article in press as: Ingebrigtsen, R., et al., Spatial models with explanatory variables in the dependence structure. Spatial Statistics (2013), http://dx.doi.org/10.1016/j.spasta.2013.06.002

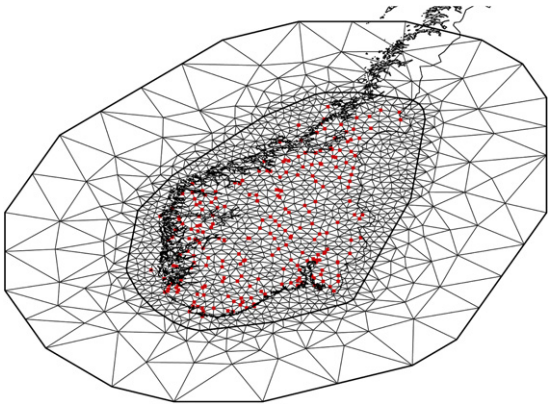


Fig. 2. Triangulation of southern Norway. The red dots indicate the observation locations of the annual precipitation data.

The basis functions $\{\psi_i\}_{i=1,\dots,m}$ are deterministic and defined for every node in the triangulation, or mesh. Lindgren et al. (2011) use piece-wise linear basis functions, which have compact support; they are equal to one in the mesh nodes and zero in all other nodes. The stochastic properties of the approximation are determined by the weights $\boldsymbol{w} = (w_1, \dots, w_m)$. These are chosen so that the representation in (3) approximates the distribution of the solution to the SPDE in (1). The distribution of the weights $\boldsymbol{w}$ is Gaussian, with Markov properties determined by the triangulation, leading to a sparse precision matrix. The elements of the precision matrix depend on α and κ, see Result 2 in Lindgren et al. (2011).

3.2. Non-stationarity with explanatory variables in the dependence structure

One of the main advantages with the SPDE approach is its flexibility. Properties of the random field are now characterised by an SPDE rather than a covariance function, which enables us to modify the SPDE instead of the covariance function to obtain GRFs with other dependence structures than the stationary Matérn covariance. The local nature of the differential operators allows local specification of parameters, and only mild regularity conditions are necessary to ensure a valid model.

In the following, we assume a two-dimensional spatial domain, i.e. $d = 2$, and fix $\nu = 1$, which implies $\alpha = 2$. The parameter ν in the Matérn covariance function determines the mean-square differentiability of the field (Diggle and Ribeiro, 2007, Chapter 3.4.1), which can influence predictions made by the model. As ν is difficult to identify from data it is common to either fix ν or to fit models for half-integer values of ν and choose its value by model selection (Diggle and Ribeiro, 2007, Chapter 5.4). For stationary models, the authors' discussion response from Lindgren et al. (2011) provides GMRFs corresponding to any ν, but no work has yet been done to carry that out for non-stationary models.

The parameter κ is a scaling parameter, linked to the range ρ by the empirically derived relationship $\rho = \sqrt{8}/\kappa$. For a GRF with Matérn covariance with parameters κ and $\nu = 1$ spatial correlation is 0.1 at distance ρ. Thus, we can think of κ as a range parameter governing the spatial dependence structure of the GRF. We add more flexibility to the dependence structure by rescaling the field $x(\boldsymbol{u})$ with a variance parameter τ, which yields the marginal variance

$$\sigma^2 = \frac{1}{4\pi\kappa^2\tau^2}. \tag{4}$$

Motivated by the annual precipitation data we introduce space-dependent covariance parameters, i.e. τ and κ are now functions of the spatial location $\boldsymbol{u}$, $\boldsymbol{u} \in \mathcal{D}$, where $\mathcal{D}$ is the spatial domain. The modified SPDE becomes

$$(\kappa(\boldsymbol{u})^2 - \Delta)(\tau(\boldsymbol{u})x(\boldsymbol{u})) = \mathcal{W}(\boldsymbol{u}), \quad \boldsymbol{u} \in \mathbb{R}^2, \tag{5}$$

Please cite this article in press as: Ingebrigtsen, R., et al., Spatial models with explanatory variables in the dependence structure. Spatial Statistics (2013), http://dx.doi.org/10.1016/j.spasta.2013.06.002

where the solutions $x(\boldsymbol{u})$ are non-stationary GRFs because τ and κ vary with location, with the consequence that the variance and correlation range vary with location.

Define $\log \tau(\boldsymbol{u})$ and $\log \kappa(\boldsymbol{u})$ as sums of basis functions,

$$\log \tau(\boldsymbol{u}) = \theta_1^\tau + \sum_{k=2}^{N} b_k^\tau(\boldsymbol{u})\theta_k^\tau, \qquad \log \kappa(\boldsymbol{u}) = \theta_1^\kappa + \sum_{k=2}^{N} b_k^\kappa(\boldsymbol{u})\theta_k^\kappa, \tag{6}$$

where the θ's are weight parameters and $b_k^\tau(\cdot)$ and $b_k^\kappa(\cdot)$ are deterministic basis functions defined on the domain $\mathcal{D}$, in practice in the nodes of the triangulation. The basis functions $b_k^\tau(\cdot)$ and $b_k^\kappa(\cdot)$ should typically be smoothly varying with respect to the triangulation to ensure that they can be adequately represented on the mesh. In Lindgren et al. (2011) the basis functions were generic low order spherical harmonics and B-splines. In this work we investigate the use of parametric models to control the SPDE parameters, using explanatory variables as basis functions.

As noted by Lindgren et al. (2011) the deformation method for non-stationary models due to Sampson and Guttorp (1992) can be reformulated as a non-stationary SPDE in the special case of deformations within a plane. However, for deformations into a higher dimensional space the two approaches differ considerably. Some deformation and SPDE models can have similar properties, but in general the two model classes are not equivalent, since the SPDE approach measures distances within or along a manifold. The non-stationary model produced by (5) and (6) leads to local changes of distance metric that cannot be represented by a simple deformation model, and hence leads to different models than the explanatory variable models from Schmidt et al. (2011). Since the non-stationary parameters control the local distance metric in the manifold they must be defined on the mesh.

From a technical point of view, having space-dependent τ and κ only changes the elements of the precision matrix of the Markov representation of the GRF. The weight vector $\boldsymbol{w}$ from the representation in (3) is $\mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}(\boldsymbol{\theta})^{-1})$, where $\boldsymbol{\theta} = (\theta_1^\tau, \ldots, \theta_N^\tau, \theta_1^\kappa, \ldots, \theta_N^\kappa)$, i.e. the elements of the precision matrix $\boldsymbol{Q}(\boldsymbol{\theta})$ are functions of the θ's in (6). Following Lindgren et al. (2011), we introduce two finite element structure matrices, $\boldsymbol{C}$ and $\boldsymbol{G}$, and write the precision matrix as

$$\boldsymbol{Q}(\boldsymbol{\theta}) = \boldsymbol{T}(\boldsymbol{K}^2\boldsymbol{C}\boldsymbol{K}^2 + \boldsymbol{K}^2\boldsymbol{G} + \boldsymbol{G}\boldsymbol{K}^2 + \boldsymbol{G}\boldsymbol{C}^{-1}\boldsymbol{G})\boldsymbol{T}, \tag{7}$$

where $\boldsymbol{T}$ and $\boldsymbol{K}$ are diagonal matrices with $T_{ii} = \tau(\boldsymbol{u}_i)$ and $K_{ii} = \kappa(\boldsymbol{u}_i)$, and i is an index over the m nodes in the triangulation. The matrix $\boldsymbol{C}$ is diagonal, with $C_{ii} = \int \psi_i(\boldsymbol{u})\, d\boldsymbol{u}$, and $\boldsymbol{G}$ is sparse positive semi-definite, with $G_{ij} = \int \nabla\psi_i(\boldsymbol{u}) \cdot \nabla\psi_j(\boldsymbol{u})\, d\boldsymbol{u}$. Recall that the ψ_i's are the basis functions from the representation in (3).

The relationship between the model parameters and the marginal variance given by (4) is valid in the stationary case but not in the non-stationary case. However, by disregarding the spatial interaction between the non-stationary parameter fields we obtain nominal approximations to the variance and correlation range of the non-stationary GRF as functions of elevation;

$$\sigma^2(\boldsymbol{u}) \approx \frac{1}{4\pi\kappa(\boldsymbol{u})^2\tau(\boldsymbol{u})^2}, \tag{8}$$

and

$$\rho(\boldsymbol{u}) \approx \frac{\sqrt{8}}{\kappa(\boldsymbol{u})}. \tag{9}$$

These approximations are valid for slowly varying $\kappa(\boldsymbol{u})$, and can be used for easy interpretation of the parameters. In Section 5.1, Fig. 5 we check how close the approximation is to the actual field variances, which can be obtained numerically (see Rue et al. (2009)). Such a comparison for the spatial range is not done, since using a single value to represent the local spatial range in a non-stationary model is inherently problematic, and (9) is used only as a convenient qualitative measure of local range.

4. Models

We suggest two Bayesian hierarchical models for the annual precipitation data: one where the spatial field is a stationary GRF and one where it is a non-stationary GRF. In Section 5.1 the two models are compared in terms of model fit and predictive performance using the data from southern Norway.

Please cite this article in press as: Ingebrigtsen, R., et al., Spatial models with explanatory variables in the dependence structure. Spatial Statistics (2013), http://dx.doi.org/10.1016/j.spasta.2013.06.002

4.1. Annual precipitation model

Let the spatial process $\{\xi(\boldsymbol{u}) : \boldsymbol{u} \in \mathcal{D}\}$, represent the true level of annual precipitation in southern Norway. We assume that this process is observed with additive measurement error at the $n = 233$ weather stations. This yields the following data model for the observations $y_1, \ldots, y_n$

$$y_i = \xi(\boldsymbol{u}_i) + \epsilon_i, \quad i = 1, \ldots, n, \tag{10}$$

where the noise terms $\epsilon_1, \ldots, \epsilon_n$ are iid $\mathcal{N}(0, 1/\tau_\epsilon)$, and independent of $\xi(\cdot)$.

Furthermore, we assume that the precipitation process can be modelled by three parts: an intercept β_0, a linear effect of elevation β_1, and a spatial field $x(\boldsymbol{u})$. The process model can be written as

$$\xi(\boldsymbol{u}) = \beta_0 + \beta_1 h(\boldsymbol{u}) + x(\boldsymbol{u}), \quad \boldsymbol{u} \in \mathcal{D}, \tag{11}$$

where $h(\boldsymbol{u})$ is the elevation at location $\boldsymbol{u}$.

4.2. Stationary and non-stationary annual precipitation models

The two precipitation models we compare in this work are obtained by using different models for the spatial field $x(\boldsymbol{u})$: (1) a stationary GRF, $x_S(\boldsymbol{u})$, and (2) a non-stationary GRF, $x_{N\text{-}S}(\boldsymbol{u})$. We refer to the models as the stationary and non-stationary model, respectively, and for later reference we write the two process models here

$$\xi_S(\boldsymbol{u}) = \beta_0 + \beta_1 h(\boldsymbol{u}) + x_S(\boldsymbol{u}), \quad \boldsymbol{u} \in \mathcal{D}, \tag{12}$$

$$\xi_{N\text{-}S}(\boldsymbol{u}) = \beta_0 + \beta_1 h(\boldsymbol{u}) + x_{N\text{-}S}(\boldsymbol{u}), \quad \boldsymbol{u} \in \mathcal{D}. \tag{13}$$

For the GRFs we use the SPDE based model as defined in Section 3.2 and (6). For the stationary x_S,

$$\log \tau = \theta_1^\tau \quad \text{and} \quad \log \kappa = \theta_1^\kappa. \tag{14}$$

And for the non-stationary $x_{N\text{-}S}$,

$$\log \tau(\boldsymbol{u}) = \theta_1^\tau + h(\boldsymbol{u})\,\theta_h^\tau \quad \text{and} \quad \log \kappa(\boldsymbol{u}) = \theta_1^\kappa + h(\boldsymbol{u})\,\theta_h^\kappa, \tag{15}$$

i.e. elevation is included as a log-linear effect on κ and τ. Note that the stationary model is equal to the non-stationary model with $\theta_h^\tau = \theta_h^\kappa = 0$.

4.3. Identifiability constraints

In both the stationary and the non-stationary model, the three components of the model are non-identifiable, since the spatial field $x(\boldsymbol{u})$ can capture both an overall mean and an elevation dependent effect. To separate the linear effect of elevation, and to make the model identifiable we require two linear orthogonality constraints for the spatial field:

$$\int_{\mathcal{D}} x(\boldsymbol{u})\, d\boldsymbol{u} = \int_{\mathcal{D}} h(\boldsymbol{u}) x(\boldsymbol{u})\, d\boldsymbol{u} = 0.$$

In the implementation, this is accomplished by discretising the integrals on the mesh into

$$\mathbf{1}^T \boldsymbol{C}\boldsymbol{w} = \boldsymbol{h}^T \boldsymbol{C}\boldsymbol{w} = 0,$$

where $\boldsymbol{C}$ is the same diagonal matrix as in (7), and $\boldsymbol{w}$ and $\boldsymbol{h}$ are the vectors of weights for the basis function representations of $x(\boldsymbol{u})$ and $h(\boldsymbol{u})$.

4.4. Prior assumptions

To complete the specification of the Bayesian hierarchical models we need to assign prior distributions to the parameters: β_0, β_1, τ_ϵ, θ_1^τ, θ_1^κ, θ_h^τ and θ_h^κ. We assume that the parameters are a priori independent. For the β's we use vague Gaussian prior distributions with zero mean and precision 0.0001. For the rest of the parameters we use informative priors; The measurement error precision τ_ϵ is assumed to have a Gamma distribution with shape parameter 2 and rate parameter 0.02.

Please cite this article in press as: Ingebrigtsen, R., et al., Spatial models with explanatory variables in the dependence structure. Spatial Statistics (2013), http://dx.doi.org/10.1016/j.spasta.2013.06.002

The four SPDE parameters are assigned Gaussian priors, we use $\theta_1^\tau \sim \mathcal{N}(4, 0.1)$, $\theta_1^\kappa \sim \mathcal{N}(-4, 0.1)$, $\theta_h^\tau \sim \mathcal{N}(0, 1)$, and $\theta_h^\kappa \sim \mathcal{N}(0, 1)$. As the prior means of θ_h^τ and θ_h^κ are set to zero, we assume that the spatial field is stationary. In the stationary case, the prior distributions for θ_1^τ and θ_1^κ yield a priori 95% credible intervals of (83, 284) km for the correlation range and (0.15, 0.54) m for the marginal standard deviation in the field. The prior variances are chosen to be low in order to limit the possible values to be physical realistic, e.g. considering the size of the spatial domain and the magnitude of the annual precipitation observations.

Note that the priors for the SPDE parameters are chosen to be informative to overcome identifiability issues. Especially the range of the field is hard to identify. We will say more about this in Section 5.2, where we present a simulation study. According to our experience, priors for the θ's in the SPDE based non-stationary GRFs should be chosen with care.

4.5. Inference

In our setting, the purpose of the spatial modelling is to learn about the process and/or to do predictions at locations without observations. We use a Bayesian framework, hence, we want to obtain posterior distributions for the parameters and/or the posterior predictive distributions at locations of interest. To obtain this we use the integrated nested Laplace approximation (INLA) methodology (Rue et al., 2009), which can be used for Bayesian inference in the class of latent Gaussian models. In this subsection, we give a brief description of how the models for annual precipitation from Section 4.2 fit into this model class, and outline the main features of the INLA methodology.

If we combine the process model in (11) with the basis function representation of the GRF in (3), we get the following linear predictor for the annual precipitation

$$\xi(\boldsymbol{u}_i) = \beta_0 + \beta_1 h(\boldsymbol{u}_i) + \sum_{j \in T_i} \psi_j(\boldsymbol{u}_i) w_j. \tag{16}$$

Here, T_i denotes the triangle containing location $\boldsymbol{u}_i$ and the sum is over the three basis functions that are non-zero at location $\boldsymbol{u}_i$. On vector form (16) becomes

$$\xi(\boldsymbol{u}_i) = [1\ h(\boldsymbol{u}_i)\ \boldsymbol{A}_i][\beta_0, \beta_1, \boldsymbol{w}]^T, \tag{17}$$

where $\boldsymbol{A}$ is an $n \times m$ matrix with elements $A_{ij} = \psi_j(\boldsymbol{u}_i)$ and $\boldsymbol{A}_i$ denotes row i of this matrix. Hence, $\xi(\boldsymbol{u}_i)$ is a linear combination of $\boldsymbol{w}$ and $\boldsymbol{\beta} = (\beta_0, \beta_1)$. The weight vector $\boldsymbol{w}$ is assigned a Gaussian prior with zero mean and sparse precision matrix as given in (7). The fixed effects $\boldsymbol{\beta}$ are assigned independent vague Gaussian priors. Thus, $\boldsymbol{x}^* = (\boldsymbol{\beta}, \boldsymbol{w})$ is jointly Gaussian given the SPDE parameters $\boldsymbol{\theta}$, and thus fits into the latent Gaussian model framework from Rue et al. (2009). The precipitation observations $\boldsymbol{y} = (y_1, \ldots, y_n)$ are conditionally independent and Gaussian given the linear predictor $\xi(\boldsymbol{u}_i)$ and the measurement error variance $1/\tau_\epsilon$, i.e. $y_i|(\xi(\boldsymbol{u}_i), \tau_\epsilon) \sim \mathcal{N}(\xi(\boldsymbol{u}_i), 1/\tau_\epsilon)$. In Rue et al. (2009) the parameters $\boldsymbol{\theta}$ and τ_ϵ are referred to as hyperparameters, denote these by $\boldsymbol{\theta}^*$, i.e. $\boldsymbol{\theta}^* = (\tau_\epsilon, \boldsymbol{\theta})$.

We would like to make inference based on the posterior density $\pi(\boldsymbol{x}^*, \boldsymbol{\theta}^*|\boldsymbol{y})$. INLA provides a recipe for computing approximations to marginal posterior densities for the hyperparameters $\pi(\theta_j^*|\boldsymbol{y})$ and for latent variables $\pi(x_i^*|\boldsymbol{y})$. Such approximations are based on use of Laplace or other related analytic approximations and on numerical integration schemes. INLA can also compute several measures used to assess models such as the deviance information criterion (DIC) (Spiegelhalter et al., 2002) and posterior predictive densities for locations of interest.

In order for the INLA methodology to work fast and efficiently, the latent Gaussian field has to satisfy some properties; First, the latent field $\boldsymbol{x}^*$, often of large dimension, admits conditional independence properties, i.e. it should be a Gaussian Markov random field (GMRF) with a sparse precision matrix $\boldsymbol{Q}$ (Rue and Held, 2005). The efficiency of INLA relies on efficient algorithms for sparse matrices. Second, because INLA needs to integrate over the hyperparameter space $\boldsymbol{\theta}^*$, the dimension of $\boldsymbol{\theta}^*$ should not be too large, say ≤ 14, due to the numerical integration scheme and optimisation methods used. For our precipitation models the latent field has a sparse precision matrix and the number of hyperparameters are three in the stationary model and five in the non-stationary model.

Please cite this article in press as: Ingebrigtsen, R., et al., Spatial models with explanatory variables in the dependence structure. Spatial Statistics (2013), http://dx.doi.org/10.1016/j.spasta.2013.06.002

The first part of the INLA methodology is to explore the (joint) posterior of the hyperparameters; $\pi(\boldsymbol{\theta}^*|\boldsymbol{y})$. The key idea is to utilise that the identity

$$\pi(\boldsymbol{\theta}^*|\boldsymbol{y}) = \frac{\pi(\boldsymbol{x}_0^*, \boldsymbol{\theta}^*|\boldsymbol{y})}{\pi(\boldsymbol{x}_0^*|\boldsymbol{\theta}^*, \boldsymbol{y})} \tag{18}$$

is valid for any value of $\boldsymbol{x}_0^*$. Following e.g. Steinsland and Jensen (2010) it can be shown that we are able to computationally efficient evaluate both $\pi(\boldsymbol{x}_0^*, \boldsymbol{\theta}^*|\boldsymbol{y})$ and $\pi(\boldsymbol{x}_0^*|\boldsymbol{\theta}^*, \boldsymbol{y})$ up to a normalising constant that does not depend on $\boldsymbol{\theta}^*$ when the likelihood is Gaussian (if the likelihood is not Gaussian the denominator in (18) needs to be approximated). Hence we are able to evaluate the unnormalised posterior $\pi(\boldsymbol{\theta}^*|\boldsymbol{y})$ for any value of $\boldsymbol{\theta}^*$ by inserting a value of $\boldsymbol{x}_0^*$ in (18). To find good evaluation points of $\pi(\boldsymbol{\theta}^*|\boldsymbol{y})$ its mode is first found by a numerical optimisation algorithm. Next, the (numerical) Hessian at the mode is used to distribute evaluation points (see Rue et al. (2009) for details).

The second part of the INLA methodology is to use the posterior for $\boldsymbol{\theta}^*$ to compute numerical approximations to the marginals of the latent field and the hyperparameters

$$\pi(x_i^*|\boldsymbol{y}) = \int \pi(x_i^*|\boldsymbol{\theta}^*, \boldsymbol{y})\pi(\boldsymbol{\theta}^*|\boldsymbol{y})d\boldsymbol{\theta}^*,$$

$$\pi(\theta_j^*|\boldsymbol{y}) = \int \pi(\boldsymbol{\theta}^*|\boldsymbol{y})d\boldsymbol{\theta}_{-j}^*.$$

Note that for our precipitation models we have a Gaussian likelihood, and the accuracy of the approximations depends only on the numerical integration scheme.

All inference presented in this paper is carried out using the `r-inla` package (available at www.r-inla.org). It makes inference from latent Gaussian models using the INLA methodology readily available, as well as it provides an user-friendly interface for defining SPDE models, see Lindgren (2012).

4.6. Model assessment

Three performance indices were chosen to evaluate predictive performance and model fit: the deviance information criterion (DIC), the continuous ranked probability score (CRPS) and the root-mean-square error (RMSE). We briefly recall these here.

The DIC is a model comparison criterion introduced by Spiegelhalter et al. (2002). The measure is widely used to assess model complexity and fit for Bayesian hierarchical models, and it is defined as

$$\text{DIC} = \bar{D} + p_D,$$

where $\bar{D}$ is the posterior mean of the deviance and p_D is the effective number of parameters. The DIC is negatively oriented, i.e. lower values indicate better model fit. As a rule of thumb, a complex model should be preferred over a less complex model if the DIC is reduced with more than ten.

To evaluate predictive performance we use RMSE and CRPS. The RMSE is simply the square-root of the mean of the squared difference between the posterior predictive mean and observed value.

CRPS is a score function on the same scale as the observations (metres in our case) and is defined as

$$\text{CRPS}(F, y) = \int_{-\infty}^{\infty} (F(u) - 1\{y \leq u\})^2 du,$$

where F is the predictive cumulative distribution and y is the observed value (Matheson and Winkler, 1976; Unger, 1985; Gneiting and Raftery, 2007). While RMSE only assesses the posterior predictive mean, the whole posterior predictive distribution is taken into consideration with CRPS. Like DIC, CRPS and RMSE are negatively oriented.

To compare the stationary and non-stationary model we have used cross-validation on the annual precipitation dataset. The 233 weather stations were randomly divided into 13 groups, 12 of them containing 18 stations and one with 17 stations. Each group was in turn left out and a stationary and non-stationary model fitted to the remaining data. The fitted models were then used to predict at the station locations left out, and CRPS and RMSE were computed using these predictive densities and precipitation observations.

Please cite this article in press as: Ingebrigtsen, R., et al., Spatial models with explanatory variables in the dependence structure. Spatial Statistics (2013), http://dx.doi.org/10.1016/j.spasta.2013.06.002

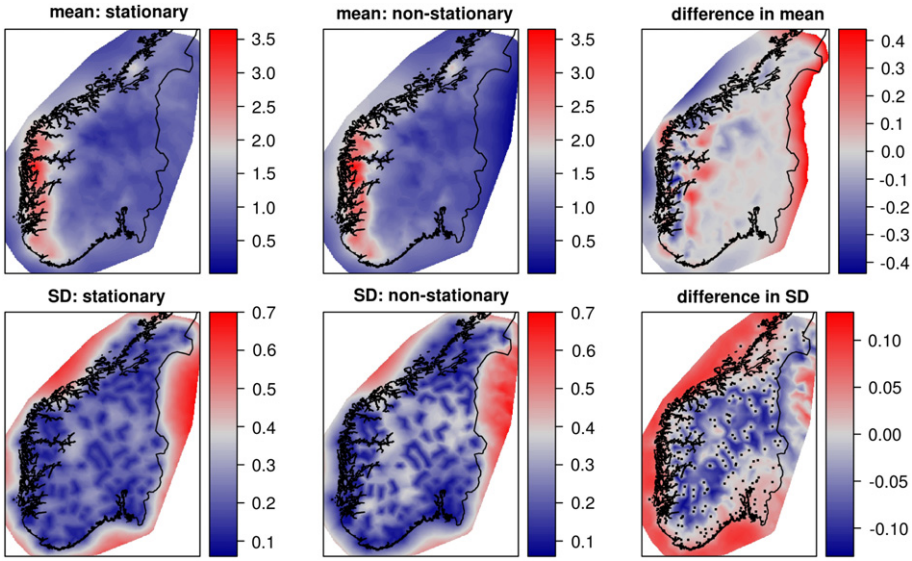


Fig. 3. Posterior mean and standard deviation of the linear predictor in (16). For comparison the difference (stationary minus non-stationary) between the two models is included in the two rightmost panels. In the lower right panel the observation locations are indicated by black dots.

5. Results

5.1. Annual precipitation in southern Norway

In this section, results for the annual precipitation data from southern Norway are presented. The dataset was described in Section 2. The two different models from Section 4.2 were fitted to the annual precipitation data: a stationary model (12), and a non-stationary model (13).

One of our main interests lies in reconstructing the annual precipitation field with uncertainty. In Fig. 3 posterior mean and standard deviation of the linear predictor defined in (16) are presented. The non-stationary model predicts more precipitation than the stationary model near the west coast, while the stationary model predicts more precipitation in the mountain range dividing East- and West-Norway. The predictions made by the two models are most similar in eastern Norway, where the landscape is relatively flat. In terms of prediction uncertainty, the stationary model is more uncertain than the non-stationary model along the coastline. In the inland the predictions made by the non-stationary model have higher uncertainty than the predictions made by the stationary model. At the observation locations the prediction uncertainty is, as expected, low and near identical using the two different models.

Posterior means and 95% credible intervals for the model parameters can be seen in Table 1. Posterior marginal densities are in Fig. 4. The linear effect of elevation (β_1) is positive, hence the precipitation increases with elevation, more in the stationary model than the non-stationary model. Also, the posterior marginal for β_1 is narrower in the non-stationary model than the stationary model.

The θ's in (14) and (15) control the dependence structure and their posterior estimates can be used to compute the spatial correlation range ρ and the marginal standard deviation in the field σ. In the stationary case, the posterior mean of the correlation range is 233 km and the posterior mean of the standard deviation is 0.738 m. In the non-stationary case, the correlation range and standard deviation vary in space approximately accordingly to (9) and (8). Maps of the approximate correlation range and standard deviation for the estimated non-stationary model can be seen in Fig. 5, both vary significantly over the domain. The marginal standard deviations shown in the map presented in the top right panel of Fig. 5 are the actual values in the model, calculated numerically from the estimated

Please cite this article in press as: Ingebrigtsen, R., et al., Spatial models with explanatory variables in the dependence structure. Spatial Statistics (2013), http://dx.doi.org/10.1016/j.spasta.2013.06.002

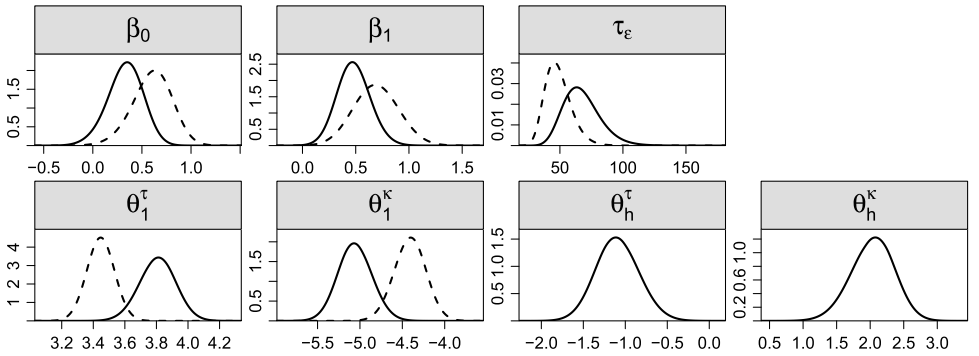


Fig. 4. Posterior marginal densities for the model parameters. The stationary model in dashed line and the non-stationary model in solid line.

Table 1
Posterior mean and 95% credible interval for the model parameters as estimated by the stationary (S) and non-stationary (N-S) models.

Parameter	Posterior mean		95% credible interval	
	S	N-S	S	N-S
β_0	0.602	0.329	(0.177, 0.975)	(−0.049, 0.665)
β_1	0.692	0.481	(0.285, 1.112)	(0.184, 0.799)
τ_ϵ	48.70	67.51	(31.57, 72.06)	(42.96, 100.5)
θ_1^τ	3.446	3.807	(3.275, 3.618)	(3.577, 4.028)
θ_1^κ	−4.410	−5.057	(−4.785, −4.049)	(−5.444, −4.652)
θ_h^τ		−1.093		(−1.587, −0.572)
θ_h^κ		2.035		(1.372, 2.642)

precision matrix, as discussed in connection with (8). How well the approximated and the actual model standard deviation agree can be seen in the bottom right panel of Fig. 5, where σ is plotted as a function of elevation together with the model standard deviations. At least for interpreting the effect of elevation the approximation can be useful. The lower panels of Fig. 5 show that both the correlation range and standard deviation decrease with increasing elevation, the range faster than the standard deviation. For southern Norway this means a larger range/standard deviation along the coastline. The coastal areas to the south-east and north in our observation area are relatively flat compared to the mountainous west coast. Hence, the area with large correlation range and standard deviation reaches further inland in these regions. Interestingly, we need an elevation of approximately 0.3 km in the non-stationary model to get the range and standard deviation of the stationary model.

The parameters θ_h^τ and θ_h^κ of the non-stationary model control the influence of elevation in the dependence structure, and if both are equal to zero, we have a stationary model. From their posterior distributions (Fig. 4 and Table 1) we find that zero is outside any reasonable credible interval, an indication that we should use a non-stationary model, and that elevation is a reasonable explanatory variable for the non-stationarity. Indeed, the DIC for the stationary model is −94.13, and −151.96 for the non-stationary model, a difference of 57.83, well above the recommended limit of ten in Spiegelhalter et al. (2002), i.e. according to DIC the non-stationary model should be preferred.

Results for the three selected weather stations – Kvamskogen, Hemsedal and Hønefoss – introduced in Section 2, can be seen in Figs. 6 and 7, and in Table 2. The three stations were left out and the stationary and non-stationary models were fitted to the remaining 230 stations. Fig. 6 show the posterior predictive densities. Both of the models underestimate the amount of precipitation at Kvamskogen, a station west in Norway at 455 m above sea level. However, the non-stationary model underestimates less than the stationary model. The best predictions are obtained at Hemsedal, a station in an inland mountain region east of the divide between East- and West-Norway. For this

Please cite this article in press as: Ingebrigtsen, R., et al., Spatial models with explanatory variables in the dependence structure. Spatial Statistics (2013), http://dx.doi.org/10.1016/j.spasta.2013.06.002

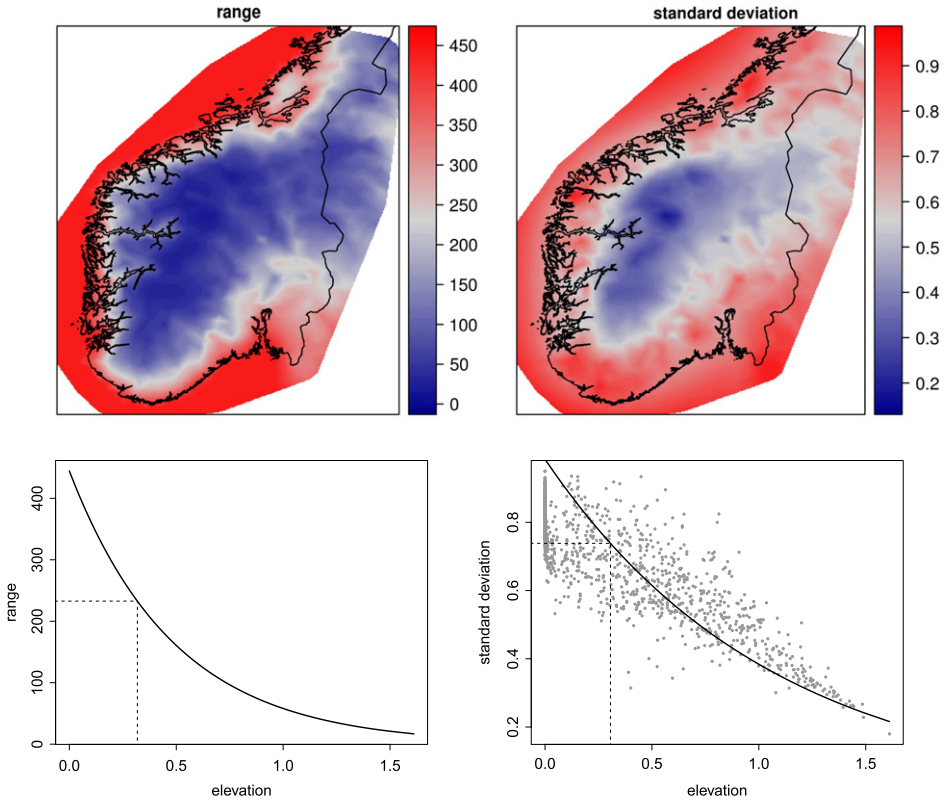


Fig. 5. Estimated correlation range ρ and standard deviation σ in the non-stationary model. The two upper panels show the spatial variation, and the two lower panels the dependence on elevation. In the lower left panel the spatial correlation in the stationary model is indicated with a dashed line. The stationary range is 233 km, which in the non-stationary case corresponds to an elevation of 0.318 km. In the same way, a dashed line at the stationary standard deviation of 0.738 m is drawn on the graph of the non-stationary standard deviation. The corresponding elevation is 0.306 km. The grey dots in the lower right panel are the marginal standard deviations obtained from the estimated precision matrix.

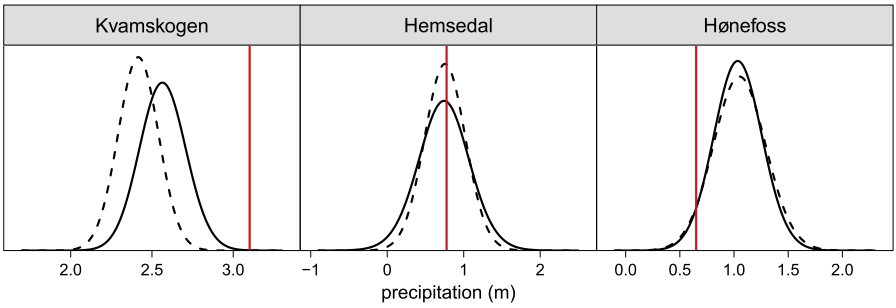


Fig. 6. Posterior predictive densities at the three selected weather stations. The stationary model in dashed line and the non-stationary model in solid line. The red vertical line is the observed amount of precipitation.

location the stationary model performs better than the non-stationary model, and the non-stationary model is slightly more uncertain. For the last station, Hønefoss, located in a flat landscape east in Norway, the amount of precipitation is overestimated by the two models. The prediction made by

Please cite this article in press as: Ingebrigtsen, R., et al., Spatial models with explanatory variables in the dependence structure. Spatial Statistics (2013), http://dx.doi.org/10.1016/j.spasta.2013.06.002

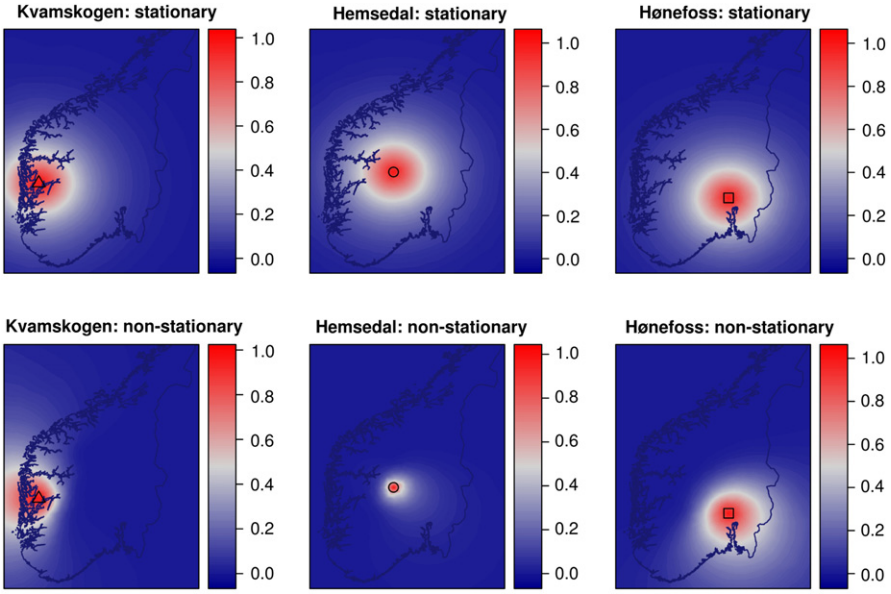


Fig. 7. Spatial correlation plots for the three stations Kvamskogen, Hemsedal and Hønefoss using the stationary model (top row) and the non-stationary model (bottom row).

Table 2
Predicted annual precipitation at the three weather stations Kvamskogen, Hemsedal, and Hønefoss with the stationary (S) and non-stationary (N-S) model. Presented are the posterior mean and selected quantiles in the posterior predictive distribution. The two bottom rows show the CRPS and the residual (observed–predicted) at the three station locations using the stationary and the non-stationary model. The underlined values indicate the model among the two with the best fit according to our chosen predictive measures.

Station (observation)	Kvamskogen (3.1010 m)		Hemsedal (0.7764 m)		Hønefoss (0.6506 m)	
	S	N-S	S	N-S	S	N-S
Posterior mean	2.4159	2.5671	0.7556	0.7412	1.0523	1.0369
Quantiles:						
0.025	2.1713	2.2887	0.2268	0.0645	0.5768	0.5981
0.500	2.4159	2.5661	0.7555	0.7416	1.0513	1.0361
0.975	2.6604	2.8511	1.2850	1.4156	1.5335	1.4809
CRPS	0.6150	0.4533	0.0637	0.0804	0.2703	0.2634
Residual	0.6851	0.5339	0.0208	0.0352	−0.4017	−0.3863

the non-stationary model is slightly better than the prediction by the stationary model. Predictive measures in terms of CRPS and residuals are found in Table 2.

The spatial correlation plots in Fig. 7 illustrate the non-stationarity imposed by the elevation. Our non-stationary model is such that the elevation in the terrain between two locations influence the correlation between these. With the stationary model, the correlation range is the same at the three stations, while with the non-stationary model the range is different at different locations and it is also clearly not symmetric. In the non-stationary model the range decreases with elevation, thus, making the low altitude areas in East- and West-Norway, separated by the mountains, less dependent than they would be in a flatter topography. For Kvamskogen we see that the non-stationary spatial correlation is very skewed with longer correlation range towards the west coast than towards the mountains in the east. Also, the most eastern location, Hønefoss, has some skewness with more correlation towards south-east/the coast, but much less than at Kvamskogen, and this is also the location among these three that have spatial correlation most similar to the stationary model.

Please cite this article in press as: Ingebrigtsen, R., et al., Spatial models with explanatory variables in the dependence structure. Spatial Statistics (2013), http://dx.doi.org/10.1016/j.spasta.2013.06.002

Table 3
Continuous ranked probability score (CRPS) and the root-mean-squared error (RMSE) for the stationary model and the non-stationary model. The results are averages based on 13-fold cross-validation. The underlined values indicate the model with the best fit according to CRPS and RMSE.

	Stationary	Non-stationary
CRPS	0.1267	0.1241
RMSE	0.2229	0.2121

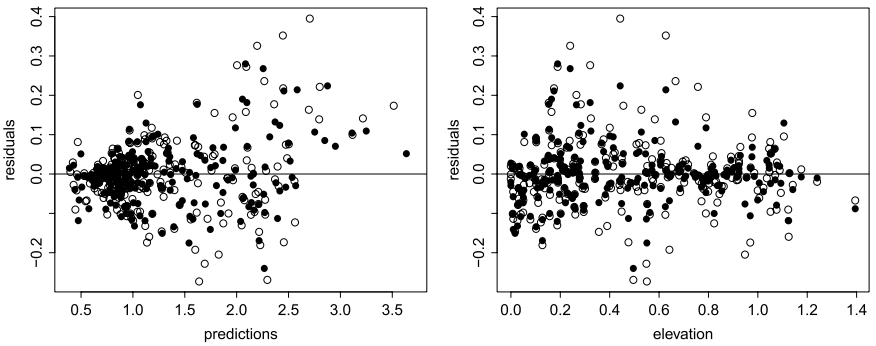


Fig. 8. Residuals from the stationary model (○) and the non-stationary model (●) plotted against the predictions made by the two models (left panel) and the elevation (right panel).

The mountain location of Hemsedal is at high elevation, with tall mountains immediately to the west. Compared to the stationary model, the non-stationary model has less spatial correlation at this location, especially westwards, which is reasonable from a meteorological point of view (recall the discussion on orographic precipitation in Section 2).

To compare model fit and predictive performance of the stationary and non-stationary model, 13-fold cross-validation was carried out. The average CRPS and RMSE for the stationary and non-stationary models are found in Table 3. The two predictive measures favour the non-stationary model. However, the difference in CRPS is 0.0026 m and in RMSE 0.0108 m, i.e. very low.

Fig. 8 is a plot of the residuals, defined as observation minus mean of the linear predictor in (16). In the left panel, the mean of the linear predictor at each station is plotted against the residuals. For both the stationary and non-stationary models the residuals are increasing with the amount of predicted precipitation. This means that both models systematically underestimate the amount of precipitation at locations where there are high amounts of precipitation, slightly more in the stationary model than in the non-stationary model. Further, the absolute values of the residuals are increasing with increasing predicted precipitation. This may indicate that the variance $1/\tau_\epsilon$ should vary with the amount of precipitation. Precipitation observations are both uncertain and have biases due to wind exposure, especially for snow (Wolff et al., 2013). Therefore, in a more realistic model, properties of the location and measurement situations should also be included. In the right panel of Fig. 8, the elevation at each station is plotted against the residuals. In this plot, the residuals seem more centred around zero and we find no apparent pattern.

5.2. Simulation study

To learn more about our model we use a simulation study. The purpose is two-fold: First, we would like to check whether we recover the parameters from simulated datasets by evaluating the coverage of posterior credible intervals. Second, we would like to evaluate how well DIC differentiate between simulated datasets from the stationary and non-stationary model.

Stationary and non-stationary models with the same prior distributions as in Section 4.4 were fitted to 100 datasets simulated from a stationary model with parameters $\theta_1^\tau = 3.4$ and $\theta_1^\kappa = -4.4$.

Please cite this article in press as: Ingebrigtsen, R., et al., Spatial models with explanatory variables in the dependence structure. Spatial Statistics (2013), http://dx.doi.org/10.1016/j.spasta.2013.06.002

Table 4
Comparison of the stationary (S) and non-stationary (N-S) models fitted to 100 simulated datasets from a stationary model. Presented are RMSE, empirical bias and 95% credible interval coverage for the model parameters.

Parameter		RMSE		Bias		Coverage	
		S	N-S	S	N-S	S	N-S
β_0	0.4	0.1603	0.1686	−0.0032	0.0026	0.97	0.97
β_1	0.6	0.1735	0.1765	−0.0102	−0.0133	0.95	0.97
τ_ϵ	50	11.6423	11.5908	4.6264	4.4430	0.96	0.96
θ_1^τ	3.4	0.0773	0.1065	0.0063	0.0350	0.97	0.97
θ_1^κ	−4.4	0.1644	0.2126	0.0699	0.1541	0.95	0.93
θ_h^τ	0.0		0.1651		−0.0234		0.97
θ_h^κ	0.0		0.6090		−0.3766		0.87

Table 5
Comparison of the stationary (S) and non-stationary (N-S) models fitted to 100 simulated datasets from a non-stationary model. Presented are RMSE, empirical bias and 95% posterior credible interval coverage for the model parameters.

Parameter		RMSE		Bias		Coverage	
		S	N-S	S	N-S	S	N-S
β_0	0.4	0.1428	0.1332	0.0056	0.0112	0.99	0.98
β_1	0.6	0.1626	0.1433	−0.0234	−0.0232	0.96	0.97
τ_ϵ	50	10.7919	13.7129	−1.1660	7.6481	0.94	0.95
θ_1^τ	3.8	0.4838	0.1652	−0.4741	−0.1358	0.01	0.84
θ_1^κ	−5.1	0.9708	0.6382	0.9529	0.6138	0.00	0.05
θ_h^τ	−1.1		0.2977		0.2200		0.84
θ_h^κ	2.0		0.8971		−0.8194		0.26

In the same way, 100 datasets from a non-stationary model with parameters $\theta_1^\tau = 3.8$, $\theta_1^\kappa = -5.1$, $\theta_h^\tau = -1.1$, and $\theta_h^\kappa = 2.0$ were simulated, and stationary and non-stationary models were fitted to the these non-stationary datasets. In both cases, $\beta_0 = 0.4$, $\beta_1 = 0.6$, and $\tau_\epsilon = 50$. The values of the parameters used for the simulations were chosen based on the data analysis in the previous subsection.

RMSE, empirical bias and 95% posterior credible interval coverage for the model parameters are presented in Tables 4 and 5. In Table 4 the simulated datasets come from the stationary model, and in Table 5 from the non-stationary model. For the stationary model there are only two parameters in the SPDE, therefore the lines for θ_h^τ and θ_h^κ are left blank.

From the results of the simulated stationary datasets (Table 4) the poorest coverage for the credible intervals is for the parameters controlling the range in the non-stationary model (θ_1^κ and θ_h^κ). For the simulated non-stationary datasets (Table 5) the coverage is as low as 0.05 for θ_1^κ and 0.26 for θ_h^κ, and with considerable empirical biases. Zhang (2004) showed that for a fixed domain and infill asymptotics there are no consistent estimators for all parameters in the Matérn class (defined in (2)). For our stationary model with $\nu = 1$, only the quantity $\kappa^2\sigma^2$ can be estimated more precisely with increasing number of (infilling) observations. From the equation for σ^2 in (4) we find that it is τ^2 that can be consistently estimated. Hence, there is limited information about κ in a single replicate over a fixed domain. Further, Zhang (2004) also demonstrated empirically that the log-likelihood function is very flat for κ for any fixed value of τ (translated to our notation). This explains our problems with estimating the parameters controlling κ (θ_1^κ and θ_h^κ). In addition, in the non-stationary parametrisation we use, neither $\tau(\boldsymbol{u})$ nor $\kappa(\boldsymbol{u})$ can be consistently estimated from a single field replicate, although for the particular parameter values used in the simulation study this effect is small for θ_1^τ and θ_h^τ. As we work in a Bayesian framework, priors also influence the posteriors, especially for parameters about which there is limited information in the data. The near non-identifiability of κ is similar to the identifiability issue discussed in Section 4.3. Loosely speaking, from a single field realisation without constraints it is not possible to distinguish between the two cases large β_0 and

Please cite this article in press as: Ingebrigtsen, R., et al., Spatial models with explanatory variables in the dependence structure. Spatial Statistics (2013), http://dx.doi.org/10.1016/j.spasta.2013.06.002

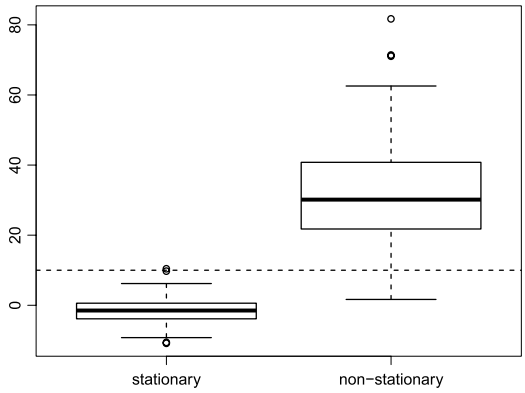


Fig. 9. Difference in DIC, the stationary model minus the non-stationary model, for the simulated datasets from the stationary and non-stationary models. The horizontal line indicates a difference of ten.

κ, and small β_0 and κ, where the latter is a nearly intrinsic model with infinite correlation range. Adding the constraints on the spatial effect does not remove this issue, and we therefore use the prior distributions to choose the case with large κ, corresponding to a short correlation length. For the parameters θ_h^τ and θ_h^κ we used Gaussian priors with mean 0 and variance 1. These are very informative priors compared to the values used in the simulation study (and obtained for the precipitation data). From Table 5 we see that both these parameters have a large bias towards the prior mean. The obvious remedy for prior influence is to choose a weaker prior, but not so weak as to enable the intrinsic situation. However, by increasing the variance of the prior, we get in to numerical problems, as the optimisation procedure in the `r-inla` software that is used to find the mode of $\pi(\boldsymbol{\theta}^* \mid \boldsymbol{y})$ (as described in Section 4.5) fails. This is due to the very flat (log) posterior, and properly overcoming this problem requires improving the numerical optimisation implementation. To develop better understanding of the identifiability issues and practical guidelines for how priors should be set is an important topic for further research.

Using difference in DIC as a model choice criterion has been questioned (e.g. Fong et al. (2010)). In our stationary versus non-stationary setting, we can use the simulation study to investigate the quality of the rule of thumb that the difference should be at least ten in favour of the more complex model for this to be chosen. In Fig. 9 box plots of the difference in DIC between stationary model fit and non-stationary model fit for the simulated datasets can be seen. Out of the 100 simulated datasets from the stationary model, the difference in DIC is greater than zero 31 times, and greater than ten only one time. From a classical hypothesis testing point of view, with the stationary model as null hypothesis, this implies a test level of $\alpha = 0.01$. For the non-stationary data, the difference is positive for all of the datasets, and greater than ten 92 times. Thus, DIC correctly selects the non-stationary model most of the times. For the models fitted to the precipitation data the difference in DIC was 57.83, which is in the upper tail of the differences for the non-stationary simulated datasets, and well above any of the simulated stationary datasets. The evidence for an elevation dependent non-stationary model for the annual precipitation in southern Norway is therefore present.

6. Discussion

The purpose of this paper was to introduce a framework for non-stationary spatial modelling by introducing explanatory variables in the dependence structure using the SPDE approach. The precipitation process is known to be non-stationary, and the topography influences both the amount of precipitation and the dependence structure. Analysing the dataset of annual precipitation in southern Norway, we have explored how our elevation driven non-stationary model can be used to learn about the process, and how predictions change, when elevation is an explanatory variable in both the mean and dependence structure in a Bayesian hierarchical model. By fitting both a stationary

Please cite this article in press as: Ingebrigtsen, R., et al., Spatial models with explanatory variables in the dependence structure. Spatial Statistics (2013), http://dx.doi.org/10.1016/j.spasta.2013.06.002

model and a non-stationary model to the data we found a difference in DIC of 57.83, which indicates that the underlying spatial process is non-stationary.

Obtaining better predictions for locations without observations, and better quantification of prediction uncertainty, was one of the main motivations for modelling the annual precipitation data with the non-stationary GRF. From Fig. 3 we have found that both the expected value and the standard deviation of the linear predictor in (16) are considerable different in some areas. For the predictive probability densities of the three stations in Fig. 6, the non-stationary model gives slightly more uncertainty at Kvamskogen and Hemsedal, while only at Kvamskogen is the posterior mean considerably shifted. To compare the predictive performance of the stationary and non-stationary models we performed 13-fold cross-validation, and the results for RMSE and CRPS in Table 3 show that the non-stationary model only performs slightly better than the stationary model (2% for CRPS and 5% for RMSE). From Table 2 and Fig. 6 we see that for Kvamskogen and Hønefoss the non-stationary model performs better, while the stationary model is better at Hemsedal.

To predict correct levels of precipitation is a challenging task, even using fine scale physical models with reanalysis forcing (Orskaug et al., 2011). There is not enough information available in the data (and explanatory variables) for a statistical model to give good predictions at all locations in such a topographical and climatic complex observation area as southern Norway. The non-stationary model were thought to solve some of the topographical challenges, but also this model smooths too much: at all stations with high predictions the actual observed values are even higher (see Fig. 8). Also Paciorek and Schervish (2006) experienced that even though the results seem more reasonable with a non-stationary model, the predictive performance did not change much compared to a stationary model when analysing meteorological data.

There are several ways to extend our precipitation model, and considering other explanatory variables is the most obvious. In preliminary work we explored other topographical explanatory variables (gradient, eastward gradient and distance from coast) and combinations of them. However, neither replacing elevation with nor including any of these, gave significant improvement to predictive measures. Thus, to keep the presentation simple we have omitted the other models from the paper.

The class of non-stationary models introduced in this paper takes explanatory variables into the dependence structure using the SPDE approach to spatial modelling. The method is related to deformation methods but differs in that the explanatory variables influence the local distance metric within the manifold. Because of this these models require that the explanatory variables going into the dependence structure are defined on and known for the entire triangulated domain mesh. In Lindgren et al. (2011) the non-stationarity stems from generic low order spherical harmonics and B-splines. We have investigated the use of an explanatory variable (elevation) controlling the dependence structure. With this parametric approach some flexibility is lost, but we gain a more direct interpretation of how elevation influences the spatial dependence. We have only used one explanatory variable in the models for the SPDE parameters, but there is no theoretical limit on the number of basis functions in the log-linear models in (6). However, as the weight parameters have to be estimated there are computational limitations (see Section 4.5). As there is limited information in one replicate of the spatial field, it is important to understand the models to be able to set required constraints and appropriate priors (see Sections 4.3 and 5.2). With our SPDE based non-stationary spatial model this is actually computationally feasible. Software for SPDE model definition and fast inference using integrated nested Laplace approximations is available (www.r-inla.org), which makes it possible to test prior sensitivity and investigate parameter identifiability using simulation studies.

Acknowledgements

We thank Håvard Rue for fruitful discussions and technical help with `r-inla`, and Jo Eidsvik and Janine B Illian for useful comments on the manuscript. Kolbjørn Engeland, Sjur Kolberg, Stian Solvang Johansen, Asgeir Petersen-Øverleir, Ole Einar Tveito, Jean-Marie Lepioufle and Matthias Mohr have contributed with interesting and informative discussions about precipitation processes, as well as the importance of, and challenges with, modelling precipitation in Norway.

This work is supported by the project Spatio-temporal Modelling and Approximate Bayesian Inference (196670/V30) founded by The Research Council of Norway.

Please cite this article in press as: Ingebrigtsen, R., et al., Spatial models with explanatory variables in the dependence structure. Spatial Statistics (2013), http://dx.doi.org/10.1016/j.spasta.2013.06.002

References

Banerjee, S., Gelfand, A., Carlin, B., 2003. Hierarchical Modeling and Analysis for Spatial Data. In: Monographs on Statistics & Applied Probability, Chapman and Hall/CRC.
Bolin, D., Lindgren, F., 2011. Spatial models generated by nested stochastic partial differential equations, with an application to global ozone mapping. Annals of Applied Statistics 5 (1), 523–550.
Bolin, D., Lindgren, F., 2013. A comparison between Markov approximations and other methods for large spatial data sets. Computational Statistics & Data Analysis 61, 7–21.
Bornn, L., Shaddick, G., Zidek, J., 2012. Modeling nonstationary processes through dimension expansion. Journal of the American Statistical Association 107 (497), 281–289.
Brenner, S., Scott, L., 2008. The Mathematical Theory of Finite Element Methods, third ed. In: Texts in Applied Mathematics, vol. 15. Springer.
Cameletti, M., Lindgren, F., Simpson, D., Rue, H., 2013. Spatio-temporal modeling of particulate matter concentration through the SPDE approach. AStA Advances in Statistical Analysis 97 (2), 109–131.
Cressie, N., Wikle, C.K., 2011. Statistics for Spatio-Temporal Data. Wiley.
Daly, C., Neilson, R., Phillips, D., 1994. A statistical-topographic model for mapping climatological precipitation over mountainous terrain. Journal of Applied Meteorology 33 (2), 140–158.
Diggle, P., Ribeiro, P., 2007. Model-Based Geostatistics. Springer.
Fong, Y., Rue, H., Wakefield, J., 2010. Bayesian inference for generalized linear mixed models. Biostatistics 11 (3), 397–412.
Fuentes, M., 2002. Spectral methods for nonstationary spatial processes. Biometrika 89 (1), 197–210.
Gelfand, A., Diggle, P., Guttorp, P., Fuentes, M., 2010. Handbook of Spatial Statistics. In: Chapman & Hall/CRC Handbooks of Modern Statistical Methods, CRC Press.
GLOBE Task Team. 1999. GLOBE: the global land one-kilometer base elevation digital elevation model. version 1.0. URL http://www.ngdc.noaa.gov/mgg/topo/globe.html.
Gneiting, T., Raftery, A.E., 2007. Strictly proper scoring rules, prediction, and estimation. Journal of the American Statistical Association 102 (477), 359–378.
Higdon, D., Swall, J., Kern, J., 1999. Non-stationary spatial modeling. Bayesian Statistics 6, 761–768.
Jansson, A., Tveito, O.E., Pirinen, P., Scharling, M., 2007. NORDGRID—a preliminary investigation on the potential for creation of a joint nordic gridded climate dataset, Tech. Rep. 3. Climate. Met.No Report.
Kleiber, W., Nychka, D., 2012. Nonstationary modeling for multivariate spatial processes. Journal of Multivariate Analysis 112, 76–91.
Lindgren, F., 2012. Continuous domain spatial models in R-INLA. The ISBA Bulletin 19 (4), 14–20. URL http://bayesian.org/bulletin.
Lindgren, F., Rue, H., Lindström, J., 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. Journal of the Royal Statistical Society, Series B: Statistical Methodology 73 (4), 423–498 (with discussion).
Matheson, J., Winkler, R., 1976. Scoring rules for continuous probability distributions. Management Science 22 (10), 1087–1096.
Orskaug, E., Scheel, I., Frigessi, A., Guttorp, P., Haugen, J., Tveito, O., Haug, O., 2011. Evaluation of a dynamic downscaling of precipitation over the Norwegian mainland. Tellus, Series A: Dynamic Meteorology and Oceanography 63 (4), 746–756.
Paciorek, C., Schervish, M., 2006. Spatial modelling using a new class of nonstationary covariance functions. Environmetrics 17 (5), 483–506.
Reich, B., Eidsvik, J., Guindani, M., Nail, A., Schmidt, A., 2011. A class of covariate-dependent spatiotemporal covariance functions for the analysis of daily ozone concentration. Annals of Applied Statistics 5 (4), 2425–2447.
Rue, H., Held, L., 2005. Gaussian Markov Random Fields: Theory and Applications. In: Monographs on Statistics and Applied Probability, vol. 104. Chapman & Hall/CRC.
Rue, H., Martino, S., Chopin, N., 2009. Approximate bayesian inference for latent Gaussian models by using integrated nested laplace approximations. Journal of the Royal Statistical Society, Series B: Statistical Methodology 71 (2), 319–392.
Sampson, P., Guttorp, P., 1992. Nonparametric estimation of nonstationary spatial covariance structure. Journal of the American Statistical Association 87 (417), 108–119.
Sansó, B., Guenni, L., 2000. A nonstationary multisite model for rainfall. Journal of the American Statistical Association 95 (452), 1089–1100.
Schmidt, A., Guttorp, P., O'Hagan, A., 2011. Considering covariates in the covariance structure of spatial processes. Environmetrics 22 (4), 487–500.
Schmidt, A., O'Hagan, A., 2003. Bayesian inference for non-stationary spatial covariance structure via spatial deformations. Journal of the Royal Statistical Society, Series B: Statistical Methodology 65 (3), 743–758.
Simpson, D., Lindgren, F., Rue, H., 2012a. In order to make spatial statistics computationally feasible, we need to forget about the covariance function. Environmetrics 23 (1), 65–74.
Simpson, D., Lindgren, F., Rue, H., 2012b. Think continuous: Markovian Gaussian models in spatial statistics. Spatial Statistics 1, 16–29.
Spiegelhalter, D., Best, N., Carlin, B., Van Der Linde, A., 2002. Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society, Series B: Statistical Methodology 64 (4), 583–616.
Steinsland, I., Jensen, H., 2010. Utilizing Gaussian Markov random field properties of Bayesian animal models. Biometrics 66 (3), 763–771.
Tobler, W.R., 1970. A computer movie simulating urban growth in the Detroit region. Economic Geography 46, 234–240.
Unger, D., 1985. A method to estimate the continuous ranked probability score. In: Proceedings of the Ninth Conference on Probability and Statistics in Atmospheric Sciences. American Meteorological Society, pp. 206–213.
Whittle, P., 1954. On stationary processes in the plane. Biometrika 41, 434–449.
Whittle, P., 1963. Stochastic processes in several dimensions. Bulletin of the International Statistical Institute 40, 974–994.
Wolff, M., Isaksen, K., Brækkan, R., Alfnes, E., Petersen-Øverleir, A., Ruud, E., 2013. Measurements of wind-induced loss of solid precipitation: description of a Norwegian field study. Hydrology Research 44 (1), 35–43.
Yue, Y., Speckman, P., 2010. Nonstationary spatial Gaussian Markov random fields. Journal of Computational and Graphical Statistics 19 (1), 96–116.
Zhang, H., 2004. Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. Journal of the American Statistical Association 99 (465), 250–261.

Please cite this article in press as: Ingebrigtsen, R., et al., Spatial models with explanatory variables in the dependence structure. Spatial Statistics (2013), http://dx.doi.org/10.1016/j.spasta.2013.06.002