Deformed SPDE models with an application to spatial modeling of significant wave height

Anders Hildeman [*], David Bolin, Igor Rychlik

Department of Space-, Earth-, and Environmental Sciences, Chalmers University of Technology, Sweden
Computer, Electrical and Mathematical Science and Engineering Division, King Abdullah University of Science and Technology, Saudi Arabia
Department of Mathematical Sciences, Chalmers University of Technology, Sweden
University of Gothenburg, Sweden

ARTICLE INFO

ABSTRACT

A non-stationary Gaussian random field model is developed based on a combination of the stochastic partial differential equation (SPDE) approach and the classical deformation method. With the deformation method, a stationary field is defined on a domain which is deformed so that the field becomes non-stationary. We show that if the stationary field is a Matérn field defined as a solution to a fractional SPDE, the resulting non-stationary model can be represented as the solution to another fractional SPDE on the deformed domain. By defining the model in this way, the computational advantages of the SPDE approach can be combined with the deformation method's more intuitive parameterization of non-stationarity. In particular it allows for independent control over the non-stationary practical correlation range and the variance, which has not been possible with previously proposed non-stationary SPDE models.

The model is tested on spatial data of significant wave height, a characteristic of ocean surface conditions which is important when estimating the wear and risks associated with a planned journey of a ship. The model parameters are estimated to data from the north Atlantic using a maximum likelihood approach. The fitted model is used to compute wave height exceedance probabilities and the distribution of accumulated fatigue damage for ships traveling a popular shipping route. The model results

* Corresponding author at: Department of Space-, Earth-, and Environmental Sciences, Chalmers University of Technology, Sweden.
E-mail address: hildeman@chalmers.se (A. Hildeman).

agree well with the data, indicating that the model could be used for route optimization in naval logistics.

1. Introduction

The probability distribution of ocean waves at a location in space and time is often referred to as the *sea state* of that location. The ability to model the sea state provides important and sometimes necessary information for risk assessment and prediction in naval logistics and marine operations. One of the most important parameters for characterizing the sea state is the *significant wave height*. The significant wave height was traditionally defined as the average wave height among the one third highest individual waves (Munk, 1944). This definition was intended to mathematically express the average wave-height estimated by a "trained observer" such as a sailor or fisherman. A more common definition today, that gives values of similar magnitude but is easier to work with mathematically, is that the significant wave height is four times the standard deviation of the sea elevation for a fixed point in space and time. Even though the significant wave height only provides limited information about the sea state it is sufficient for many situations and applications. The safety of a naval structure may depend on extreme and rare events such as loads which exceed the strength of components, or on everyday load variability that may cause changes in the properties of material, e.g., cracking (fatigue) or other types of aging processes. Typically such loads can be approximated using significant wave height (Mao et al., 2010; Lewis, 1988).

The sea state is rarely constant over large spatial regions or for any longer time periods. Instead the significant wave height can be regarded as the parameter that describes the evolution in time and space of the local wave energy (Baxevani et al., 2009). This means that the significant wave height is a varying function in space and time. As such, and with all uncertainties and lack of information associated with predicting future sea states, the significant wave height over a larger region of space and/or time has to be modeled as a random process (in space and/or time).

For small regions of the north Atlantic, the sea state during a time interval of about an hour can be considered as constant and distributed according to a log-Gaussian distribution (Jasper, 1956; Ochi, 1978). Also for other ocean regions similar transformations to Gaussianity are often possible. For applications only concerned with the sea state at one point in space during a shorter interval of time, this is sufficient. However, many applications are concerned with quantities acquired over larger regions or paths on the spatial domain and possibly during a longer time duration. Examples of such are spatial interpolation/data assimilation of sea states from measurements, or risks and wear analysis of ships given a planned route. When more than one point should be considered simultaneously it is often necessary to account for the dependency among these points, i.e., a spatial/temporal model is required. Moreover, for larger spatial regions this model has to be non-stationary (Baxevani et al., 2009; Ailliot et al., 2010).

Ailliot et al. (2010) considered non-stationary marginal standardizations and then assumed a stationary Gaussian random field model in space. Baxevani et al. (2009, 2011) developed a non-stationary spatial model based on the logarithm of significant wave height being a Gaussian random field. The covariance model was defined as a non-stationary process-convolution (Higdon, 1998) using a Gaussian kernel. Hence, the non-stationarity was characterized by one spatially-varying parameter controlling the correlation length between point pairs. The choice of a Gaussian kernel is limiting the types of possible spatial correlations that can be modeled. Moreover, it yields infinitely mean-square differentiable realizations of the significant wave height, which is theoretically unrealistic since it implies that observing a small region of the ocean thoroughly would give perfect information about the values over the whole ocean (Stein, 1999). This issue was noted by Paciorek and Schervish (2006) who instead introduced a non-stationary Matérn model based on process-convolution. This model can be employed to allow for control of the smoothness, i.e., differentiability. However, the non-stationary process-convolution model is not theoretically

suitable for modeling the significant wave height, as well as many other spatial variables. The reason being that the covariance between a pair of points are only dependent on the local spectral densities at those two specific points. In most variables studied in spatial statistics, shared information between a pair of points has to flow through the spatial domain, i.e., a spatially continuous Markov property. This is clearly true for significant wave height as well, where wave energy travels in space and time and is generated by wind friction over large regions of water. Strong correlation between two points in space indicates that they are both affected by the same wind/wave systems. This suggests that correlation between a pair of points should be dependent, not only on the parameter values at the two points themselves, but on the parameter values at all points corresponding to possible paths between the two points.

Besides non-stationary process-convolution there are other methods for acquiring non-stationarity. Sampson (2010) divides them into three categories, viz., smoothing-based methods, basis function methods, and deformation methods. Smoothing methods acquire non-stationarity by introducing a convolution over a set of independent random fields. Compared to the process-convolution methods the kernel itself is stationary but the parameters of the independent random fields are indexed by points in space, yielding a non-stationary smoothing over the set of independent random fields. The basis function methods decompose the random field into a, possibly infinite, sum of spatially varying functions with associated random coefficients. The spatial modeling is then transformed into modeling of the dependencies between the set of random coefficients. The deformation methods (Sampson and Guttorp, 1992) allow for non-stationarity by assuming that the spatial domain can be deformed using a bijective mapping into a *deformed space*. If the random field is mapped to this *deformed space* it will be stationary. The problem of non-stationarity is hence divided into that of finding a suiting bijective mapping and a suiting stationary model. The deformation method has the attractive continuous Markov property that the process-convolution methods were lacking (as long as the stationary covariance function has this property). The problem with the deformation method is parameterizing this deformation in an adequate way that is flexible enough while avoiding *folding*, see Sampson and Guttorp (1992).

A promising new method for modeling non-stationarity is the non-stationary stochastic partial differential equation (SPDE) models of Fuglstad et al. (2015b,a) and Bolin and Lindgren (2011). These are closely connected to the deformation method but does not require explicit parameterization of the deformation. More importantly, compared to the non-stationary models listed above, these models can handle complicated spatial domains, such as coast lines, while having beneficial computational properties that allow for high resolution modeling of large regions of open ocean. However, a problem with these models is that, for the non-stationary case, the spatially varying correlation parameters and marginal parameters become intertwined and their interpretation is hence lost. This means that one cannot change the marginal variance without affecting the correlation structure and vice versa, leading to some possible problems, viz.,

- The clear interpretation of each parameter is lost since they jointly represent both the spatial correlation behavior and the marginal variances.
- All parameters have to be fitted jointly. This increases the computational complexity of parameter estimation, as compared to estimating the marginal variance first and then the spatial correlation.
- The identifiability of the model is lost since more than one set of parameters can correspond to exactly the same model.

Our contribution is deriving a non-stationary SPDE-model which solves the issues of Fuglstad et al. (2015b,a) by combining the deformation method (Sampson and Guttorp, 1992) and the Matérn SPDE-approach (Lindgren et al., 2011). The model maintains all the computational advantages of Fuglstad et al. (2015b). Furthermore, due to our derived link between the deformation method and the non-stationary SPDE, it is possible to calculate properties of the non-stationary random field that have only been possible for stationary fields before. As an example, in this work we derive an analytical formula for exceedance probabilities when traveling a specified path over the spatial domain. We use our results to model significant wave height in the north Atlantic spatially and to analyze the risks and wear of a ship traveling a transatlantic route.

In this paper we will denote a point in space by s, possibly by an index if several points are present simultaneously. Likewise, a point in time will be denoted by t and possibly an index if necessary. The spatial domain will be denoted as $\mathcal{G}$, or sometimes when needed as $\mathcal{D}$. The vector "dot-product" is denoted as $\cdot$, the gradient operator of a function in several dimensions as ∇, the Laplacian as $\Delta = \nabla \cdot \nabla$, the identity matrix as I, and the domain (of an operator) as $\mathscr{D}$.

The structure of the paper is as follows. In Section 2 the data and the problems of the case study are presented. This section also act as a motivation to why a spatial model of significant wave height is needed. Also, an upper bound for exceedance probabilities of significant wave height when traveling a shipping route is given. In Section 3 the proposed model is derived. A theorem proving the relationship between the model and a deformation of a Gaussian Matérn random field is presented. Also, a semi-explicit formula for exceedance probabilities of a given curve using the proposed model is presented. Section 4 introduces how the model can practically be parameterized and estimated given data. The results of the case study are presented in Section 5. This include a comparison between the non-stationary model and a stationary anisotropic Matérn model as well as the performance when analyzing accumulated fatigue damage and exceedance probabilities for a transatlantic route in the north Atlantic. Finally, Section 6 concludes with a discussion of the results and potential further applications and extensions of the model. The manuscript also contains four appendices which present further details on the model and the estimation procedure as well as all proofs.

2. Case study: spatial distribution of significant wave height

We want to model the significant wave height, commonly denoted as H_s, spatially over a large region of open ocean. Getting consistent high quality measurements of H_s over a large spatial region, with a good spatial resolution, is not easy. In fact, this is one of the reasons for developing a spatial model of H_s, to interpolate measured values to the whole of a spatial domain. However, in this paper we are mainly concerned with the effect that encountered sea states will have on ships sailing a specified route. Although forecasting methods for wind and waves are constantly evolving, predicting the sea state more than two weeks in advance is practically impossible. At the same time, applications such as policy making, ship design, and business strategy for naval logistics require decisions to be made further into the future than two weeks. Hence, making use of the spatial probability distribution of sea states together with modern ship models can yield important information for the applications mentioned above.

To be able to fit and compare models with data, we will use data from the ERA-Interim global atmospheric reanalysis (Dee et al., 2011), performed by the European Centre for Medium-Range Weather Forecasts (ECMWF). The ERA-Interim dataset provides hindcasted data of several important atmospheric and oceanic variables on a global longitude–latitude grid. The spatial resolution of the dataset is approximately 80 km and data is available from 1979 to present. The variable in the dataset we will work with is *significant wave height of wind and ground swells*, which is available at a temporal resolution of 6 h. The values produced for a fixed point in time are based on measurements from a temporal window of $\pm$ 6 h. Hence, there is a 3 h overlap in between data from consecutive points in time which will cause a smoothing effect. We will not model the temporal evolution and therefore we want to approximate data from different points in time as independent. Therefore, we choose to thin the data to a temporal resolution of 24 h (only keeping every fourth time stamp).

Even though we do not believe that the ERA-Interim dataset represents the true values of H_s perfectly, it is probably the most accurate representation thereof available. In previous work (Baxevani et al., 2005) we have used satellite measurements to estimate a spatial log-normal Gaussian field with Gaussian correlations for H_s variability. When ECMWF data have been available we have re-estimated that model at several locations using the ERA-Interim dataset. The two datasets have given similar results which indicates that the ERA-Interim dataset performs well for our purposes. We will interpret the values of the dataset as point observations at the grid locations.

We choose the north Atlantic as our spatial domain since this region provides some of the most important trading routes of naval logistics (both historically and at present) and is known for its unpredictable storms. Furthermore, the region is known to produce data for which the bulk of
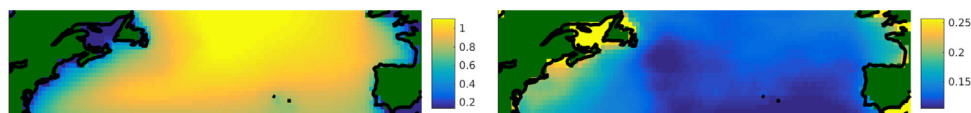
Fig. 1. Sample mean (left) and sample variance (right) of $\log H_s$.

the pointwise marginal distributions for any chosen month are well approximated by log-Gaussian distributions (Jasper, 1956; Ochi, 1978). We also restrict the analysis to the month of April, but using data from all of the 38 years (1979–2017). The reason for this restriction being that data from different months will be distributed differently due to the effects of the annual cycle, which we do not model in this work. The month of April was chosen since, among all the months of the year, for this month the tails of the logarithmized marginal data seemed to fit a normal distribution the best. However, any other month could have been chosen instead since they all showed good agreement with normal distributions. From here on, we will refer to the logarithm of H_s as X. Our model, as well as former models (Baxevani et al., 2005, 2009, 2011), are based on modeling X as a Gaussian random field.

The mean and variance of the logarithmized data can be seen in Fig. 1. Clearly the wave height is diminishing near the coasts while variance increases. Non-stationarity in the correlation structure can be observed in the left column of Fig. 4 portraying the empirical correlation function of $\log$-H_s between three different reference points and all other points on the observational domain. Apparently, the point close to the coast of USA is showing an anisotropic pattern with the principal axis on the diagonal. Contrary to this, the spatial correlation of the mid Atlantic and at the coast of Europe have the principal axis in the east–west direction. It should be noted that the data is portrayed in the longitude–latitude coordinate system and other projections would yield different non-stationarity.

2.1. Fatigue damage accumulated by a vessel

Mao et al. (2010) presented a formula approximating the fatigue damage accumulated by a ship sailing under a stationary sea state. The formula is a function of the wave period and the significant wave height and is based on the narrow-band approximation (Rychlik, 1993). According to this formula, the expected instantaneous fatigue damage to a ship, en route, as a function of the significant wave height, zero crossing period, ship velocity, and wave velocity is

$$d(t) \approx \frac{0.47 C^\beta H_s^\beta(\boldsymbol{s}(t)), t}{\gamma} \left(\frac{1}{T_z(\boldsymbol{s}(t), t)} - \frac{2\pi V(t) \cos \theta(t)}{g T_z^2(\boldsymbol{s}(t), t)} \right), \tag{1}$$

where g is the gravitational constant (≈ 9.81), T_z is the mean zero crossing period of the waves, V is the speed of the ship, and θ is the angle between the heading of the ship and the direction of the traveling waves. Further, γ and β are constants dependent on the material of the ship and C is a constant depending on the ship design, see Mao et al. (2010) for details. Note that the formula concerns a route in space–time parameterized by time, t, i.e. $H_s(\boldsymbol{s}(t), t)$ is the significant wave height at point $\boldsymbol{s}(t)$ at time t and so on.

From the formula we see that increasing wave height and diminishing wave period will increase damage. In general, the conclusion is that high waves have a greater impact on the fatigue damage but also that steep waves are more damaging than flat ones. Since T_z is not modeled in this work, we follow the procedure of Mao et al. (2010) and Podgórski and Rychlik (2014) and approximate it as $T_z \approx 3.75 \sqrt{H_s}$, making the formula in (1) dependent only on the values of H_s, for which we are developing a spatial model. The total damage up until time T is computed by numerically integrating d,

$$D(T) = \int_0^T d(t) dt \approx \sum_i d_i \Delta t, \tag{2}$$

where d_i denotes a sequence of evaluations of $d(t)$, Δt apart. It should be noted that the formula of Eq. (1) only concerns the conditional expectation of fatigue damage given the sea state parameters. However, the variability conditioned on the sea states are very small in comparison to the variability without conditioning on the sea states.

Given a stochastic model of the significant wave height, from which realizations can be generated quickly, the distribution of accumulated fatigue damage can be approximated by Monte-Carlo simulations.

2.2. Extreme loads

When using a significant wave height model to analyze risks for a ship, it is often of interest to calculate the probabilities of exceeding a given threshold value of H_s when sailing along a chosen ship route. For instance, extreme sea conditions characterized by very large significant wave heights can lead to breakage of the hull instantaneously or over a very short time span. Therefore, it is important to have control of the risks of encountering such extreme sea states and possibly reroute the ship if the risk turns out to be too high.

Mathematically this corresponds to the probability of the random field exceeding a threshold on a given curve in the spatial domain. Consider a curve $\gamma(t) = [\boldsymbol{s}(t), t]^T$, where $\boldsymbol{s}$ a point in the spatial domain of the ocean and t a point in an interval of time. Further, define $X_\gamma(t) = X(\boldsymbol{s}(t), t)$ as the random process acquired by evaluating the spatio-temporal field of significant wave height on the curve γ. The corresponding mean function and marginal variance function of X_γ are denoted as $\mu_\gamma(t)$ and $\sigma_\gamma(t)$. For analyzing the risk of encountering high significant wave heights on a ships journey one needs the exceedance probability function, $\mathbb{P}\left(\max_{t \in [0,T]} X_\gamma(t) > u\right)$.

Just as with accumulated fatigue damage, given a stochastic model of the significant wave height, it would be possible to approximate the exceedance probability using Monte-Carlo simulations. However, when analyzing the risks to a ship, the probability of encountering really high values of significant wave height is typically of interest. High values have a low probability of occurring and hence, the sample size of the Monte-Carlo simulation would need to be large to yield a reliable estimate of the exceedance probability. A better alternative would be to use an analytical formula if available. Although we are not able to derive a sharp formula for the exceedance probability, we can derive an upper limit.

Proposition 2.1. *Assume that $X(\boldsymbol{s}, t)$ is a Gaussian random fields that is differentiable in mean square sense and where both its marginal parameters, $\mu(\boldsymbol{s}, t)$ and $\sigma(\boldsymbol{s}, t)$, are differentiable with continuous partial derivatives. Further assume that the curve γ, is differentiable with continuous partial derivatives. Then,*

$$\mathbb{P}\left(\max_{t \in [0,T]} X_\gamma(t) > u\right) \leq \int_0^T \left(\sigma_{\dot{W}}(t)\phi\left(\frac{a(t)}{\sigma_{\dot{W}}(t)}\right) + a(t)\right) \phi\left(\frac{u - \mu_\gamma(t)}{\sigma_\gamma(t)}\right) dt + \Phi\left(\frac{\mu_\gamma(0) - u}{\sigma_\gamma(0)}\right).$$

Here, ϕ and Φ denote the pdf and cdf of the standard normal distribution, $a(t) := \frac{u - \mu_\gamma(t)}{\sigma_\gamma(t)^2}\dot{\sigma}_\gamma(t) + \frac{1}{\sigma_\gamma(t)}\dot{\mu}_\gamma(t)$, the dot above the symbol of a function denotes the derivative w.r.t. time. The stochastic process $W(t)$ is a marginal standardization of $X_\gamma(t)$, i.e., $W(t) := \frac{X_\gamma(t) - \mu_\gamma(t)}{\sigma_\gamma(t)}$, and $\sigma_{\dot{W}}(t)$ is the standard deviation of the mean square derivative of $W(t)$.

The proof of the proposition can be found in Appendix D. The upper bound is a good approximation of the actual exceedance probability for large values of u, which is our main concern when considering risks to ships.

It is worth noticing that the only spatial dependency in this formula is in the marginal variance of the derivative process $\dot{W}(t)$, i.e., $\sigma_{\dot{W}}(t)$.

3. The deformed SPDE model

Sections 1 and 2 have given the background and motivation to why a non-stationary Gaussian random field model with low computational cost is needed for modeling the logarithm (or for some regions of the worlds oceans possibly another marginal transformation into Gaussianity) significant wave height, X over large regions of ocean. This section introduces our proposed model as well as the main theoretical results of this work. It should be noted that this model, and the results within, are in no way exclusive for modeling significant wave height. The results of this work should be of importance for any one concerned with modeling non-stationary Gaussian random fields using the SPDE-approach (Lindgren et al., 2011).

A Gaussian random field is completely characterized by its mean and covariance functions. For simplicity, and without loss of generality, we will in this section assume a centered Gaussian random field, X, and therefore only be concerned with the model for the covariance structure. The Matérn covariance function is a popular choice due to its flexibility using only three easily interpretable parameters. For $\boldsymbol{s}_1, \boldsymbol{s}_2 \in \mathcal{D} \subseteq \mathbb{R}^d$, it is defined as

$$\mathbb{C}(\boldsymbol{s}_1, \boldsymbol{s}_2) = \sigma^2 C_\nu(\kappa \|\boldsymbol{s}_1 - \boldsymbol{s}_2\|), \quad C_\nu(h) = \frac{1}{2^{\nu-1}\Gamma(\nu)}(h)^\nu K_\nu(h). \tag{3}$$

For a random field with a Matérn covariance function, σ^2 is the marginal variance, $\kappa > 0$ controls the practical correlation range, and $\nu > 0$ determines the smoothness. Whittle (1954) showed that if $\mathcal{D} := \mathbb{R}^d$, a centered Gaussian random field $X(\boldsymbol{s})$ with a Matérn covariance is equal to the solution of the SPDE,

$$\left(\kappa^2 - \Delta\right)^{\alpha/2}(\tau X) = \mathcal{W}, \quad \text{in } \mathcal{D}, \tag{4}$$

where Δ is the Laplacian and $\mathcal{W}$ is Gaussian white noise. The fractional exponent α is related to the smoothness parameter through $\alpha = \nu + d/2$, and

$$\tau = \sqrt{\frac{\Gamma(\nu)}{\Gamma(\nu + d/2)(4\pi)^{d/2}}} \frac{1}{\kappa^\nu \sigma},$$

is a constant needed to get the desired marginal variance, σ^2.

Lindgren et al. (2011) used the SPDE formulation in combination with the finite element method to compute Gaussian Markov random field approximations of Gaussian Matérn fields for bounded domains, $\mathcal{D} \subset \mathbb{R}^n$. The computational benefits of this approach, as well as the fact that it allows for several extensions of stationary Matérn fields on $\mathbb{R}^d$, has made it highly popular.

One important extension is to non-stationary models. Fuglstad et al. (2015b,a) proposed a non-stationary extension based on replacing the differential operator $\left(\kappa^2 - \Delta\right)^{\alpha/2}$ in (4) by $\left(\kappa^2(\boldsymbol{s}) - \nabla \cdot H(\boldsymbol{s})\nabla\right)$. They assumed that H was a continuously differentiable and uniformly positive definite matrix-valued function and that κ was a continuous function. Bolin and Kirchner (2019) proposed a larger class of models with less restrictive assumptions on κ and H, and also allowed for a fractional power of the non-stationary operator as in the original stationary case. One can define random fields with a wide variety of non-stationary behaviors using differential operators on this form. However, the drawback is that κ and H jointly control local anisotropy, correlation range, and marginal variance. That means that the interpretability of the parameters from the original SPDE-approach is lost. Furthermore, in general one cannot derive a formula for the marginal variance of the non-stationary model.

Because of these issues, we will now derive a new SPDE-based model inspired by the deformation method of Sampson and Guttorp (1992). The deformation method considers the observational domain as a subset of some space, $\mathcal{G}$, being a deformation of another space, $\mathcal{D}$. Here, by deformation we mean that there exists a bijection, $F : \mathcal{D} \mapsto \mathcal{G}$. A random field that is stationary and isotropic in $\mathcal{D}$ can then be mapped to $\mathcal{G}$ by F. The function F characterizes the non-stationarity and/or anisotropy of the random field. See Fig. 2 for a sketch of the procedure of mapping a stationary and isotropic random field to $\mathcal{G}$.
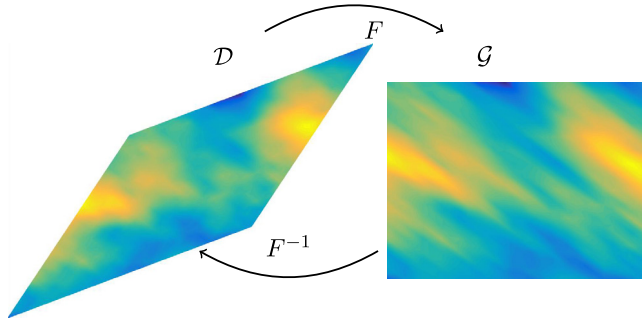
Fig. 2. Realization of anisotropic Gaussian random field using deformation method.

Our idea is that we have the standard Matérn model on $\mathcal{D}$. By the deformation this will yield a non-stationary model with the same beneficial properties as the original stationary SPDE-approach. Moreover, instead of working with a SPDE in some unknown $\mathcal{D}$-space, we want to acquire the corresponding SPDE on $\mathcal{G}$. Then, a finite element approximation can be considered directly on the observational domain of $\mathcal{G}$. To more rigorously define the model, let $\mathcal{D} \subset \mathbb{R}^d$ be an open, bounded, convex polytope and consider the SPDE,

$$(I - \hat{\nabla} \cdot \hat{\nabla})^{\alpha/2} \hat{X} = \hat{\mathcal{W}} \quad \text{in } \mathcal{D}. \tag{5}$$

The notation with a hat, such as in $\hat{X}$, denote that we are considering a formulation in spatial coordinates of $\mathcal{D}$; as opposed to the notation without a hat denoting the $\mathcal{G}$-space counterpart. Here, $\hat{\mathcal{W}}$ is Gaussian white noise on $\mathcal{D}$ and $\hat{\mathcal{L}} = (I - \hat{\nabla} \cdot \hat{\nabla}) : \mathscr{D}(\hat{\mathcal{L}}) \subset L_2(\mathcal{D}) \to L_2(\mathcal{D})$ is equipped with homogeneous Dirichlet boundary conditions on $\partial \mathcal{D}$. By Proposition 3.1 in Bolin and Kirchner (2019), this equation has a unique solution in $L_2(\Omega; L_2(\mathcal{D}))$ whenever $\alpha > d/2$.

We are now interested in defining a non-stationary random field, X, through the deformation approach. The following theorem shows that we can obtain X as the solution to a SPDE directly on $\mathcal{G}$.

Theorem 3.1. *Let* $F : \mathcal{D} \mapsto \mathcal{G} := F(\mathcal{D}) \subset \mathbb{R}^d$ *be a bijective and differentiable function. Then the deformed random field* $X(\boldsymbol{s}) := \hat{X}(F^{-1}(\boldsymbol{s}))$, *where* $\hat{X}$ *is the solution to* (5) *with* $\alpha > d/2$, *satisfies*

$$\left(I - \kappa^{-2} \nabla \cdot H \nabla\right)^{\alpha/2} X = \kappa^{-1} \mathcal{W} \quad \text{in } \mathcal{G},$$

Here $\mathcal{W}$ *is white noise on* $\mathcal{G}$, ∇ *is the gradient operator on* $\mathcal{G}$, $\kappa(\boldsymbol{s}) := \sqrt{\left|J[F^{-1}](\boldsymbol{s})\right|}$, *and* $H(s) := \kappa(\boldsymbol{s})^2 J[F^{-1}]^{-1}(\boldsymbol{s}) J[F^{-1}]^{-T}(\boldsymbol{s})$, *where* $J[F^{-1}]$ *denotes the Jacobian matrix of* F^{-1} *and* $\left|J[F^{-1}]\right|$ *denotes the determinant of this matrix.*

The proof of the theorem is given in Appendix D. The theorem is a generalization of the deformation result stated (but not formally proved) in Lindgren et al. (2011) for $\alpha = 2$. The reason for formulating the result using a unit dampening constant in (5) is that the correlation range can be controlled both by F and this constant. So by fixing the dampening constant to 1, we avoid issues with identifiability when later estimating parameters.

The main advantage with defining a non-stationary model in this way is that X has the same marginal variances as $\hat{X}$, which means that the non-stationarity induced by the deformation does not affect the variance. This greatly simplifies the interpretation of the parameters. In order to model random fields of arbitrary marginal variances we will include a scaling, $\tau(\boldsymbol{s})$, as in the original Matérn SPDE. The final model that we will use is therefore,

$$\left(I - \kappa^{-2} \nabla \cdot H \nabla\right)^{\alpha/2} (\tau X) = \kappa^{-1} \mathcal{W} \quad \text{in } \mathcal{G}. \tag{6}$$

If we would have $\mathcal{D} = \mathbb{R}^d$, the variance of the solution would be,

$$\text{Var}(X(\boldsymbol{s})) = \frac{\Gamma(\alpha - 1)}{\Gamma(\alpha)(4\pi)^{d/2}\tau(\boldsymbol{s})^2}.$$

In this case, if $F(\boldsymbol{s})$ is not varying spatially and τ is a constant, the covariance of the solution to (6) is stationary and given by,

$$\mathbb{C}(\boldsymbol{s}_1, \boldsymbol{s}_2) = \sigma^2 C_\nu \left(\kappa \sqrt{(\boldsymbol{s}_1 - \boldsymbol{s}_2)^T H^{-1}(\boldsymbol{s}_1 - \boldsymbol{s}_2)} \right), \quad \sigma^2 = \frac{\Gamma(\alpha - 1)}{\Gamma(\alpha)(4\pi)^{d/2}} \frac{1}{\tau^2}, \tag{7}$$

where C_ν is the Matérn correlation function in (3) with smoothness parameter $\nu = \alpha - d/2$ (Khristenko et al., 2019). Thus, the parameters κ, ν and σ have the same interpretation as for the standard Matérn covariance, and H controls the anisotropy where eigenvectors of H with larger eigenvalues correspond to directions with longer correlation ranges. The difference when $F(\boldsymbol{s})$ is allowed to vary in space is that $\kappa(\boldsymbol{s})$ and $H(\boldsymbol{s})$ should be interpreted as local correlation range and anisotropy explaining how the space is locally deformed in a neighborhood of point $\boldsymbol{s}$.

For the model to be of practical use, it has to be approximated by a finite dimensional representation. This can be done using the finite element method as in Lindgren et al. (2011). Here, the solution is approximated by a basis expansion where the coefficients of the expansion are given by a multivariate Gaussian random vector with a sparse precision matrix (for integer values of α). The details of the procedure for our model are presented in Appendix A. For the general fractional case, the method of Bolin and Kirchner (2019) can be used.

It should be noted that the model of Eq. (6) approximated by the finite element method is valid as long as H and κ are derived from a function, F^{-1}, that is differentiable almost everywhere and which satisfy the inverse function theorem (Spivak, 1965). This means that the model is valid even though a bijective mapping to a deformed space of stationarity does not actually exist. It is sufficient that each local neighborhood can be deformed to stationarity, see Appendix C.

3.1. Exceedance probability

We presented an analytical formula for an upper limit of the exceedance probability, Proposition 2.1. It was noted that the only quantity that depended on the correlation structure was $\sigma_{\dot{W}}(t)$.

The model proposed in this section is strictly spatial while a ship is sailing on a dynamically evolving spatial domain. It is possible to model a spatio-temporal process based on the spatial model of this work using, for instance, a space–time separable covariance function or even better a space–time separable covariance function in the lagrangian frame of reference as in Baxevani et al. (2009, 2011). However, in this work we will just consider the purely spatial model. Therefore, during a ships journey the realization of the random field is assumed to persist unchanged over time, i.e., we assume that a ship is traveling on a "freezed" realization of the spatial random field. Hence, t will in this work only act as a parameterization of the curve over the spatial domain. For this model we have an explicit formula for $\sigma_{\dot{W}}^2(t)$.

Corollary 3.2. *Assume that the curve, γ, is differentiable with continuous partial derivatives. Then,*

$$\sigma_{\dot{W}}^2(t) = \frac{\kappa^2(\boldsymbol{s}(t))}{2(\nu - 1)} \dot{\boldsymbol{s}}(t)^T H^{-1}(\boldsymbol{s}(t)) \dot{\boldsymbol{s}}(t),$$

where $\dot{\boldsymbol{s}}(t)$ is the vector of partial derivatives of $\boldsymbol{s}(t)$ in space, i.e., the derivative of the curve γ at point t.

The corollary is based on Proposition D.4 and a proof can be found in Appendix D. Proposition D.4 gives an explicit formula for $\sigma_{\dot{W}}^2(t)$ for the more general case of any non-stationary model that corresponds to local deformations to a stationary random field.

4. Inference

In order to use the proposed model for applications the model parameters have to be estimated from data. Even before that, a finite dimensional representation of the spatially varying functions for the non-stationary model of Section 3 has to be chosen.

4.1. Representation of spatially varying parameters

An intuitive approach would be to parameterize the mapping, F, directly as in the original deformation method (Sampson and Guttorp, 1992). However, our SPDE only depends on the Jacobian matrix of F^{-1} and we could therefore parameterize this Jacobian matrix instead. Due to possible identifiability issues, see Appendix C, we do not do that either. Instead we choose to parameterize $H(\boldsymbol{s})$ and $\kappa(\boldsymbol{s})$ through an intermediate step, an auxiliary matrix-valued function,

$$\tilde{H}(\boldsymbol{s}) := J[F^{-1}]^{-1}(\boldsymbol{s})J[F^{-1}]^{-T}(\boldsymbol{s}).$$

Since the data is on $\mathbb{R}^2$, $\tilde{H}(\boldsymbol{s})$ is a 2×2-matrix. This matrix is defined by $\tilde{H}_{11}(\boldsymbol{s}) = \exp(h_1(\boldsymbol{s}))$, $\tilde{H}_{22}(\boldsymbol{s}) = \exp(h_2(\boldsymbol{s}))$, and $\tilde{H}_{12}(\boldsymbol{s}) = \tilde{H}_{21}(\boldsymbol{s}) = (2S(h_3(\boldsymbol{s})) - 1)\exp(0.5(h_1(\boldsymbol{s}) + h_2(\boldsymbol{s})))$. Here, S denotes the sigmoid function and h_1, h_2, h_3 are some bounded functions. Based on this matrix, we can define $\kappa^2(\boldsymbol{s}) = |\tilde{H}(\boldsymbol{s})|^{-1/2}$ and $H(\boldsymbol{s}) = \kappa(\boldsymbol{s})^2\tilde{H}(\boldsymbol{s})$. The reason for this parameterization is that it ensures that $H(\boldsymbol{s})$ symmetric and positive definite, $\kappa^2 > 0$, and the parameterization is identifiable, i.e., there is a one-to-one relationship between κ, H and the functions h_1, h_2, h_3. Moreover, it does not put any constraints on h_1, h_2, h_3 more than that they should be bounded and real-valued.

We define h_1, h_2, h_3 as low-dimensional regressions on cosine functions over the domain of interest:

$$h_i(\boldsymbol{s}) = \sum_{p=0}^{k} \sum_{n=0}^{k} \beta_{np}^i \cos\left(n\frac{\pi s_1}{C_1}\right) \cos\left(p\frac{\pi s_2}{C_2}\right), \quad i = 1, 2, 3, \tag{8}$$

where $\boldsymbol{s} = (s_1, s_2)$ and C_1, C_2 denote the width and height of the bounding box of the observational domain. The advantage with this parameterization is that we do not have any restrictions on the coefficients β_{np}^i in order to obtain a valid model, this is important since it simplifies the optimization problem later on. A truncated cosine series allow for quite flexible modeling of the non-stationarity while ensuring smoothness and a low-dimensional representation of the non-stationarity. The smoothness is important for two reasons; partly it is a prerequisite for the method introduced in Appendix B, and partly it is necessary when approximating the parameters as piecewise constant over each triangle in the FEM mesh, see Appendix A. A similar basis expansion can be used to parameterize $\log(\tau(\boldsymbol{s}))$ and/or the marginal mean, $\mu(\boldsymbol{s})$, if spatially varying marginal parameters are needed in the model.

4.2. Likelihood-based parameter estimation

Let y_{jl} denote an observed value of X at location $\boldsymbol{s}_j \in \mathcal{G}$, where l enumerates replicates. We assume the previously proposed Gaussian random field model for the observations but also include a nugget effect to account for spatially small-scale and independent effects. Thus, the assumed model is $Y_{jl} = \mu(\boldsymbol{s}_j) + X_l(\boldsymbol{s}_j) + \epsilon_{jl}$ where $\epsilon_{jl} \sim \mathbb{N}(0, \sigma_\epsilon^2)$ are independent, $\mu(\boldsymbol{s}_j)$ is the mean value of the latent field at $\boldsymbol{s}_j$, and $\{X_l(\boldsymbol{s})\}_l$ are independent replicates of the Gaussian random field model. Using the FEM discretization of the random field, as explained in Appendix A, we can write the model as

$$\boldsymbol{Y}_l = \boldsymbol{\mu} + A\boldsymbol{U}_l + \epsilon_l, \quad U_l \sim \mathbb{N}\left(0, Q_U^{-1}\right), \quad \epsilon_l \sim \mathbb{N}(0, \sigma_\epsilon^2 I), \quad l = 1, \dots, L.$$

Here, L is the total number of replicates in the data set, $\boldsymbol{Y}_l = (Y_{1l}, \dots, y_{jl})^T$ is a vector with all observations for replicate l, $\boldsymbol{\mu}$ is a vector with entries $\mu_j = \mu(\boldsymbol{s}_j)$, A is the *observation matrix* with elements $A_{ji} = \phi_i(\boldsymbol{s}_j)$ where ϕ_i is the i:th FEM basis function, $\boldsymbol{U}_l$ is the multivariate Gaussian distribution at replicate l of the coefficients of basis functions in the FEM approximation. The parameters of the model, which needs to be estimated, are thus the parameters $\{\beta_{np}^1, \beta_{np}^2, \beta_{np}^3\}_{n,p=0}^k$ and possibly the nugget variance σ_ϵ, as well as the parameters for a representation of $\tau(\boldsymbol{s})$ and/or $\mu(\boldsymbol{s})$. We parameterize $\log(\sigma_\epsilon) = \tilde{\sigma}_\epsilon$ since $\tilde{\sigma}_\epsilon$ has to be positive.

To estimate the proposed model from data we use a maximum likelihood approach. Since both the nugget effect and the random field are Gaussian, the distribution of $\boldsymbol{Y}_l$ is Gaussian with mean $\boldsymbol{\mu}$ and covariance $\Sigma_Y = AQ_U^{-1}A^T + \sigma_\epsilon^2 I$. However, evaluating the log-likelihood of this Gaussian

variable directly would require computing Q_U^{-1}, which is dense. To be able to take advantage of the sparsity of Q_U, we use the fact that the density of Y_l can be rewritten as

$$f(\mathbf{y}) = \left. \frac{f(\mathbf{u})f(\mathbf{y}|\mathbf{u})}{f(\mathbf{u}|\mathbf{y})} \right|_{\mathbf{u}=\mathbf{u}^*},$$

where $\mathbf{u}^*$ is an arbitrary value in the sample space of $\mathbf{U}$. Let Θ denote the set of all parameters, $\mathbf{y}_l$ an observed realization of $\mathbf{Y}_l$ and $\mathbf{y} = \{\mathbf{y}_l\}_{l=1}^L$. Choosing $u^* = 0$ yields the log-likelihood

$$l(\Theta; \mathbf{y}) = \frac{L}{2}\left(\log|Q_U| - J\log|\sigma_\epsilon^2| - \log|Q_{U|Y}| + J\log(2\pi)\right)$$

$$+ \frac{1}{2\sigma_\epsilon^2}\sum_{l=1}^L (\mathbf{y}_l - \boldsymbol{\mu})^T \left[\frac{1}{\sigma_\epsilon^2}AQ_{U|Y}^{-1}A^T - I\right](\mathbf{y}_l - \boldsymbol{\mu}).$$

To evaluate the log-likelihood, computing the log-determinants and solving linear systems w.r.t. the matrix $Q_{U|Y}$ are needed. Since all matrices involved in the likelihood are sparse, symmetric, and positive definite, this can be handled efficiently using sparse Cholesky factorizations.

Since no analytical formula is available for the parameter values maximizing the likelihood, we perform numerical optimization of the log-likelihood. The log-likelihood depend on the parameters, $\{\beta_{np}^i\}_{npi}$, in a rather complex manner and it is not possible to compute the derivatives of the log-likelihood explicitly. Instead, we use a Quasi-Newton method through the *optimization toolbox* in Matlab to acquire the maximum likelihood estimates. This can be time consuming since the proposed model can have a large number of parameters. To reduce the computation time, and the risk of finding sub-optimal local maxima, good starting values are needed for the parameters. We obtain such starting values by first computing local parameter estimates as explained in Appendix B. The method is based on approximating the parameters as constant in small neighborhoods close to observed locations in space and compute local estimates for each such neighborhood. The estimates are then merged to get a starting value through a least-squares procedure. See Appendix B for further details.

5. Results of case study

In this section we are concerned with the performance of the derived non-stationary SPDE-model of Section 3 to the application of modeling significant wave height. We want to confirm that maximizing the likelihood numerically given the paramaterization of Section 4 is possible and to evaluate how the model perform for the two tasks related to shipping presented in Section 2. Furthermore, it is of interest to see if the non-stationary model is needed or an anisotropic Matérn model is sufficient. The data used in the results of this section is the ERA-interim dataset presented in Section 2.

5.1. Model fit

The model presented in Section 3 was fitted to the data described in Section 2. Before fitting the model, the data was logarithmized and then standardized marginally by subtracting the sample mean and dividing with the sample standard deviation for the values of each of the spatial locations separately. This was possible since each location had replicate observations. The data is partitioned into a training set and a validation set by taking every other day into the training set and the remaining days in the test set. By this partition, each dataset contains 585 replicates, no closer than 2 days apart.

The proposed model of Section 3 is fitted to the training data using the parameterization of Section 4. In the cosine series expansions of the h_i functions, $k = 4$ was used yielding $5 \cdot 5 = 25$ parameters to be estimated for each function. Since the data is marginally standardized, we assume that it have a constant zero mean as well as a unit marginal variance. Hence, except for the coefficients in $h_1(\mathbf{s}), h_2(\mathbf{s}), h_3(\mathbf{s})$, the only other parameters to be estimated are σ_ϵ^2 and ν. The model thus required estimating 77 parameters in total.
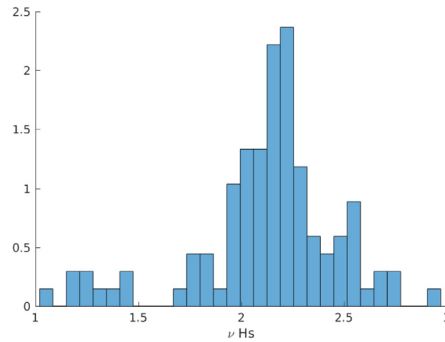
**Fig. 3.** Histogram of estimated smoothness parameters, ν, from the different local estimates.
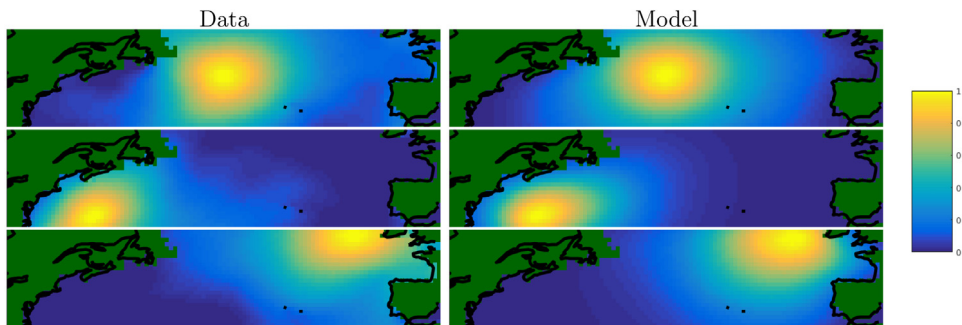
Fig. 4. Empirical (left) and model (right) correlations between reference points and all other points in the spatial domain. The reference point is in the middle of the north Atlantic ocean (top row), close to the west coast of USA (middle row), and to the west coast of central Europe (bottom row).

A histogram of the local estimates of the smoothness parameter, ν, using the method of Appendix B can be seen in Fig. 3. All estimated values are between 1 and 3 with the median close to 2, corresponding to $\alpha = 3$. Since the FEM approximation of Lindgren et al. (2011) only allow for α such that $2\alpha \in \mathbb{N}$, we fix $\alpha = 3$ before estimating the other parameters.

Three examples of correlation functions between one point in the spatial domain with all other points can be seen in Fig. 4. The figure shows both the empirical correlation function from data and the computed correlation function from the estimated model. As can be seen in Fig. 4, the correlation structure is well captured by the model. Another way of visualizing the non-stationarity of the estimated model is to show the estimated deformation. Fig. 5 shows how a rectangular grid of the north Atlantic (the $\mathcal{G}$-space) has been mapped to $\mathcal{D}$ acquired from the parameter estimation of the non-stationary model. In the figure one sees that distances close to the coasts are elongated compared to the middle of the Atlantic ocean. This means that correlation drops of quicker with distance close to the coasts. Such a behavior makes sense due to the various effects occurring on the interface between land and ocean, causing spatially (and temporally) less large scale dependence. The observations are placed on a grid in the longitude/latitude coordinate system. This will make the distance in the longitudinal direction between points further north larger than for points in the south of the domain. If the field was stationary on the sphere, this would correspond to an elongation in the x-direction in $\mathcal{D}$ as the y-value decreases. Such an effect is not clearly visible in the figures, likely because the actual random field is not stationary even in the spherical frame of reference.

It should be noted that the grid has folded in a few positions close to the coast, suggesting that the estimated $\mathcal{D}$ is not well-defined. The folding is hard to avoid, see Appendix C. However, for
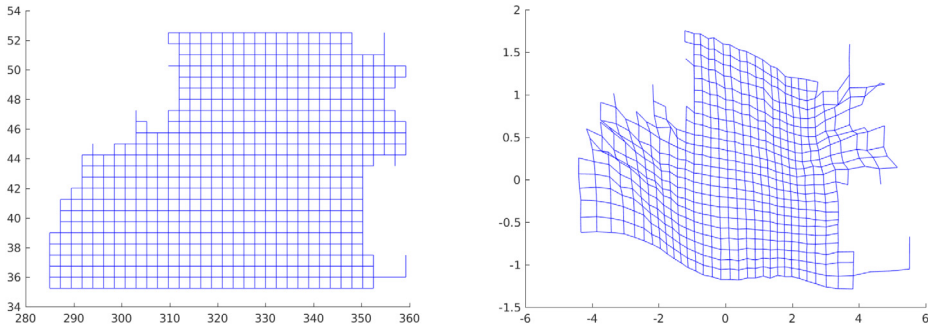
Fig. 5. Mapping of a grid in $\mathcal{G}$ (left) to $\mathcal{D}$ (right). The grid points outside of the ocean have been removed.

our model, folding is not an issue and the model is still valid—this is not the case for the original deformations method (Sampson and Guttorp, 1992). Moreover, even with folding, Fig. 5 still gives an interpretation of the non-stationarity and anisotropy since it is possible to see the local elongations and rotations.

5.2. Comparison stationarity/non-stationarity

The results of the previous subsection suggest that the non-stationary model is needed for the data. However, a relevant question is whether the fitted non-stationary model explains the distribution of the data significantly better than a fitted stationary model. To answer this, we also fitted a stationary anisotropic model to the data. Since the likelihood function is explicitly available and the two models are nested, we can perform a likelihood-ratio test between them. Here, the stationary model is considered to be the baseline model so the non-stationary model represents the more flexible alternative. The statistic we consider is the logarithm of the ratio between the maximum likelihood values of the two models,

$$\lambda = \log \frac{\sup_{\theta \in \Theta_0} L(\theta; y)}{\sup_{\theta \in \Theta} L(\theta; y)} = l_{\text{stationary}} - l_{\text{non-stationary}}.$$

The null hypothesis is that the stationary model explains the data as well as the non-stationary one. The parameter space Θ_0 is the subspace of parameter values which yields a stationary model ($k = 0$ in Eq. (8)) and Θ is the full parameter space of the non-stationary model (in our case $k = 4$ in Eq. (8)). The null hypothesis is rejected if $\lambda < c$ for some critical value c. We cannot compute c for a given significance level explicitly. However, since the number of realizations of the spatial observations are rather big, 585, it is reasonable to consider asymptotic results. Using that the distribution of -2λ converges to a χ^2_{df}-distribution (Casella and Berger, 2002, Theorem 10.3.3) it is possible to compute c. Here, the degrees of freedom, df, equal the difference between the number of parameters in the two models, in our case $df = 72$. This means that $c = -\frac{1}{2} F^{-1}_{\chi^2_{df}}(1-\alpha)$ where α is a chosen significance level and $F^{-1}_{\chi^2_{df}}$ is the quantile function of the chi-square distribution. Choosing for instance $\alpha = 10^{-4}$ yields $c = -62.688$. The observed value of λ is $2.652 \cdot 10^6 - 2.907 \cdot 10^6 = -2.55 \cdot 10^5$, clearly $\lambda \ll -62.688$ and the non-stationary model is significantly better according to the test even with an extremely low significance level. Because of the large likelihood difference, the non-stationary model is also preferable according to the Akaike information criterion (Akaike, 1973).

5.3. Fatigue damage accumulated by a vessel

For model validation we will compare the empirical distribution of accumulated fatigue damage between the data and simulated data from the fitted model for the route shown in Fig. 6, which is a

Fig. 6. Atlantic route.

scenario similar to the one in Podgórski and Rychlik (2014). The continuous route is approximated by line segments between 100 point locations (evenly spaced in geodesic distance).

We set the ships speed to a fixed value of 10 m/s which yields a sailing duration of 149.69 h or equivalently 6.23 days. The heading of the ship, in one of the 100 locations en route, is approximated as the mean between the direction acquired from the two connecting line segments. We set the constants specific to the ship as in Podgórski and Rychlik (2014) and Mao et al. (2010), i.e. $C = 20$, $\beta = 3$, and $\gamma = 10^{12.73}$. In order to use Eq. (1) we also need the propagating waves angle in comparison with the ship. In Podgórski and Rychlik (2014) the mean wave propagation in space was estimated for each month of a year. We use those propagation estimates in order to acquire the instantaneous angle between the ships heading and the expected direction of the traveling waves during the month of April.

Our model is purely spatial and the temporal resolution of the data is low (one measurement per day). We want to compare the empirical distribution of the accumulated damage using data with simulations from our model. Therefore, as mentioned in Section 3, we assume that the sea state is constant in time but spatially varying during the ship's journey. Finally, to acquire the accumulated damage, the numerical integration of Eq. (2) is computed over the 100 point locations on the route.

We use the validation data to compute the empirical fatigue damage distribution by calculating total fatigue damage of the ships journey for each of the 585 replicates in the dataset. We then simulate new H_s data sets of equal size (585 independent replicates) from the model and compute the corresponding fatigue damage distribution. Quantile–quantile (QQ) plots of 200 simulated fatigue damage distributions against the empirical distribution from the data can be seen in Fig. 7. In the figure, three models are compared both traveling from USA to Europe and vice versa. The direction of the route affects the fatigue damage since the waves will have a different angle of attack when the ship is traveling in one direction compared to the opposite. The three models compared are the non-stationary model, the stationary model, and an independent model where each of the 100 point locations on the route are modeled independently using the marginal normal distributions of $\log H_s$. Clearly, a spatial model is needed since the variability is underestimated in the independent model, which does not resemble the distribution of the data at all. The stationary and non-stationary models both seem to capture the bulk of the accumulated fatigue damage probability distribution well. However, in the lower tails the non-stationary model shows a better fit. The difference is most clearly seen on the route from Europe to USA which also seem to yield a greater amount of damage.

5.4. Extreme loads

Section 2 addressed the need for analyzing the risk of encountering extreme significant wave heights during a journey. Also in this applications we will consider a ship planned to travel the route of Fig. 6. We compare the exceedance probability estimated from data with those estimated from the model. Estimated model exceedance probabilities are both acquired by Monte-Carlo simulations and by the explicit upper bound calculated using Proposition 2.1 and Corollary 3.2.
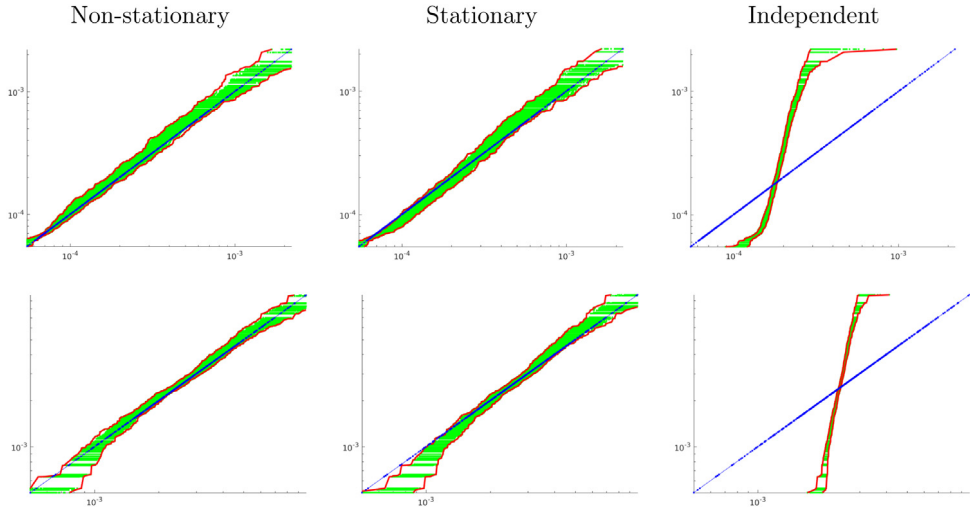
Fig. 7. QQ-plots of fatigue damage for data (blue) and 200 simulations (green). The envelopes of the simulations are drawn in red. Top row shows damage when traveling the route eastwards and bottom row show damage when traveling westwards. The columns from left to right show the results for the non-stationary, stationary, and independent model respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The exceedance probability of a given threshold for the significant wave height is estimated from data by counting the fraction of replicates for which the threshold was exceeded on the considered route. The Monte-Carlo simulations from the model are computed in the same way for 200 simulations from the fitted model—each one with 585 replicates. Since we are considering a purely spatial model there will be no difference in exceedance probability depending on the direction traveled.

To compute the theoretical bound, we approximate the integral of Proposition 2.1 by a sum over the 100 points on the route. The derivatives of the mean and standard deviation fields are approximated numerically using the difference between values of neighboring pixels on route.

Fig. 8 shows the estimated exceedance probabilities for a range of thresholds between 2 to 12 m. The figure compares the empirical exceedance probability from the data with both the stationary and non-stationary model. The non-stationary model has an envelope that encapsulates the data completely, suggesting that the model is sufficient for modeling exceedance probabilities. For the stationary model, the exceedance probability is below the envelope in the lower range of threshold values (up until about 5 m), meaning that the model overestimates the exceedance probability of small values. The upper bound seems to perform well as an approximation of the exceedance probability for values above 5 m for the non-stationary model and values above 6 m for the stationary model, as suggested by the location of the black line in comparison with the bulk of the green lines. For the very high threshold values, which are the exceedance probabilities which are typically of interest, the empirical estimates perform poorly due to lack of sufficient data. This is the range where the upper bound is expected to behave as the real exceedance probability. However, it is hard to validate this since the reference (the blue line) is uncertain in this region.

6. Discussion

We have developed a non-stationary and/or anisotropic Gaussian random field model for significant wave heights. The model is formulated using the SPDE representation of Gaussian Matérn fields in combination with the deformation method of Sampson and Guttorp (1992). Our main result

(a) Non-stationary model   (b) Stationary model
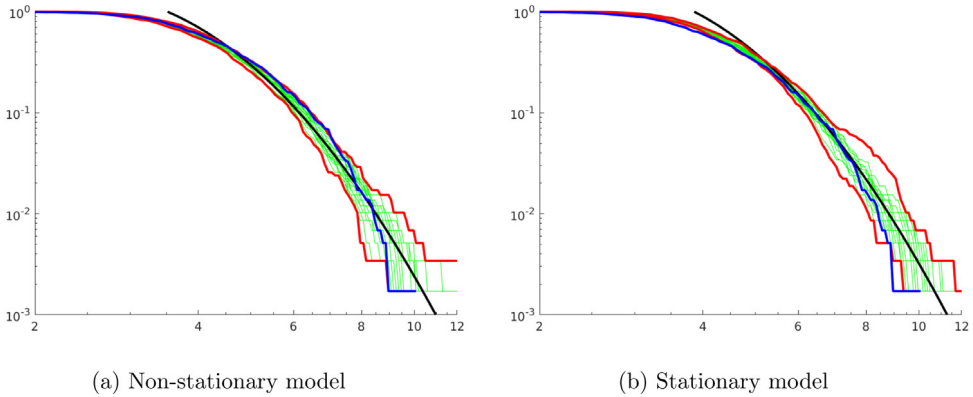
Fig. 8. Comparison of the empirical exceedance probability from data (blue), the corresponding empirical exceedance probabilities from 200 model simulations (green), as well as the upper bound from the model (black). The y-axis shows the exceedance probability for the corresponding threshold on the x-axis. The red lines show the envelope of the 200 simulations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

shows that a deformed Matérn SPDE model can be represented as a solution to another fractional SPDE on the deformed space.

For the special case of $\alpha = 2$, Fuglstad et al. (2015b,a) considered the related model

$$(\tilde{\kappa}^2 - \nabla \cdot \tilde{H}\nabla)u = \mathcal{W},$$

where the functions $\tilde{H}(\boldsymbol{s})$ and $\tilde{\kappa}(\boldsymbol{s})$ could be chosen independently of each other. Because of this, the marginal variance of the solution is affected by the choice of $\tilde{H}(\boldsymbol{s})$ and $\tilde{\kappa}(\boldsymbol{s})$. In the stationary case, a direct consequence of comparing the covariance function of our model to the covariance function derived in Fuglstad et al. (2015a) shows that the two models are equivalent (when using $\alpha = 2$ in our model). However, the important difference between the models is that our parameterization allows for controlling the variance of the process independently of κ and H. For the stationary case, this is not crucial, but when changing to non-stationary models it is important for the interpretability of the parameters and simplifies estimation. In particular, it guarantees that the model parameters are identifiable, which can be a difficult thing to check in more general models. It also solves the problem stated in Fuglstad et al. (2015a) of identifying the marginal variance for a non-stationary SPDE model.

Another important advantage with the proposed model is that the deformation method allows for theoretical derivations of results that are typically difficult to obtain for non-stationary models. One example of such is the upper bound on exceedance probabilities which we derived.

Just as for the more basic SPDE-based models, Gaussian Markov random field approximations of the model can be obtained using the finite element method. This allows for computationally efficient inference and simulation, also for complex spatial domains and spatially irregular measurements. Thus, the proposed model will likely be useful also in many other applications not related to wave modeling.

Significant wave height data from the ERA-interim reanalysis dataset was used to fit and validate the model. Results showed that both the exceedance probability of significant wave height for a specified shipping route, as well as the accumulated fatigue damage distribution, are well explained by the proposed model. The simpler stationary model performed well on the accumulated fatigue damage but underestimated small exceedance probabilities.

There are many possible extensions to the proposed model which will be investigated in future work. One such extension is to model multivariate random fields similar to Hu et al. (2018) and Bolin and Wallin (2020). Such models would allow for joint modeling of significant wave height and mean

zero crossing period, which for example would allow for a more accurate fatigue damage model. Another important aspect of the wave state data is that it is evolving in time. Therefore, another natural continuation is to derive a suitable spatio-temporal extension of the model.

Acknowledgments

Appendix A. Finite element approximation

To obtain a finite dimensional approximation of the model in Section 3 we assume that $\mathcal{G}$ is a bounded and polygonal domain. The solution of the SPDE is then approximated by a basis expansion, $X(\boldsymbol{s}) = \sum_{i=1}^{N} U_i \phi_i$, where $\{\phi_i\}$ are piecewise linear basis functions induced by a triangular mesh on the spatial domain, see Fig. 9. The distribution of the coefficients $\boldsymbol{U} = (U_1, \ldots, U_N)^T$ are computed by replacing the infinite-dimensional test space of the weak formulation of the SPDE with the finite-dimensional subspace spanned by the basis functions, $\{\phi_i\}$.

Furthermore, when computing the inner products in the FEM problem, the spatially varying parameters, $\kappa(\boldsymbol{s})$, $H(\boldsymbol{s})$, are in the implementation of Section 5 approximated as constant over each triangle. The value over a triangle is taken as the value in the mid-point of the triangle. This approximation will make the inner products of the FEM possible to calculate explicitly, it also makes Eq. (6) equivalent with

$$\left(\kappa^{2/\alpha} \mathcal{L} \right)^{\alpha/2} X = \mathcal{W},$$

since κ and $\mathcal{L}$ will commute. Here, $\mathcal{L} := \mathrm{I} - \kappa^{-2} \nabla \cdot H \nabla$. Assuming the parameters to be constant over each triangle should not affect the solution significantly since $\kappa(\boldsymbol{s})$, $H(\boldsymbol{s})$ are varying slowly in comparison with the spatial extent of a triangle in the mesh.

For $\alpha = 2$, the distribution of $\boldsymbol{U}$ is computed by solving the following system of linear equations

$$\left\{ \sum_j \langle \kappa \mathcal{L} \tau \phi_j, \phi_i \rangle U_j, i = 1, \ldots, N \right\} \overset{d}{=} \left\{ \langle \mathcal{W}, \phi_i \rangle, i = 1, \ldots, N \right\}. \tag{9}$$

Using Stokes theorem in combination with the Dirichlet boundary conditions, the inner products on the left hand side can be written as

$$\langle \mathcal{L} \tau \phi_j, \phi \rangle = \langle \kappa \tau \phi_j, \phi \rangle - \langle \kappa^{-1} \nabla \cdot H \nabla (\tau \phi_j), \phi \rangle$$
$$= \langle \kappa \tau \phi_j, \phi \rangle + \langle H \nabla (\tau \phi_j), \nabla (\kappa^{-1} \phi) \rangle.$$

The left-hand side of (9) can therefore be written as $KU = (B + G)U$, where $B_{ij} = \langle \kappa \tau \phi_j, \phi_i \rangle$ and $G_{ij} = \langle H \nabla (\tau \phi_j), \nabla (\kappa^{-1} \phi_i) \rangle$. In order to obtain a Markov random field, the right-hand side is approximated by a centered multivariate Gaussian random variable W with diagonal covariance matrix C, with elements $C_{ii} = \langle 1, \phi_i \rangle$ (Lindgren et al., 2011). Solving $KU = W$, yields $U \sim \mathbb{N}(0, Q^{-1})$, with the sparse precision matrix $Q = K^T C^{-1} K$.

Eq. (9) provides an approximation of the solution when $\alpha = 2$. For models where α is an even integer (Lindgren et al., 2011) showed how a solution can be expressed by recursively applying $\kappa^{2/\alpha} \mathcal{L}$ several times. Also for α being an odd integer, a similar construct could be achieved by considering the least squares solution to $\kappa^{1/2} \mathcal{L}^{1/2} X = \mathcal{W}$ and then considering recursive solutions to this. This is possible also for our anisotropic model if κ is piecewise constant. Hence, for any integer-valued α, this scheme yields the precision matrices $Q^{(1)} := K$ for $\alpha = 1$, $Q^{(2)} := KC^{-1}K$
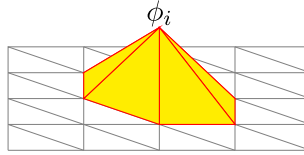
Fig. 9. Example of a test function for node i on a 2-dimensional mesh.

for $\alpha = 2$, and integer precision matrices recursively being defined as $Q^{(\alpha)} := KC^{-1}Q^{(\alpha-2)}$ for $\alpha \in \{3, 4, \ldots\}$. These matrices become less sparse as α is increased. However, the computational complexity of using the approximation for inference will be $\mathcal{O}(N^{3/2})$ but with a larger constant due to less sparsity for larger α. This should be compared to the standard cubic increase in computational cost for general covariance based methods (Rue and Held, 2005).

A.1. Boundary values and meshing

The use of boundary conditions for the bounded domain in the FEM approximation will affect the solution to the SPDE such that behavior close to the boundary deviates from that of the corresponding model on an unbounded domain. However, the boundary conditions will have a negligible effect in spatial regions that are sufficiently far from the boundary due to the positive dampening, κ (Khristenko et al., 2019). Lindgren et al. (2011) used this fact and simply extended the finite element mesh far enough outside of the spatial domain of interest such that the solution in the domain of interest was practically unaffected by the boundary conditions. For a Gaussian random field with a Matérn covariance function the extension distance is often chosen as r or $2r$ where $r = \frac{\sqrt{8\nu}}{\kappa}$ denotes the practical correlation range.

Here there are two effects that have to be balanced: In order to reduce the computational cost, the number of triangles should be as small as possible. At the same time, the mesh diameter (the longest edge in the triangulation) has to be small enough in order to get a good FEM approximation of the true solution. We let the mesh diameter be at most a fifth of the local correlation range at the location of the triangle. However, when κ and H are spatially varying, we do not have a fixed correlation range. It would be a waste of computational resources to extend the mesh based on the largest local correlation range while keeping the mesh fine enough everywhere to resolve the smallest local correlation range. By applying the barrier method of Bakka et al. (2019) we get around this issue enforcing the smallest local correlation range in all of the extended region outside of the observational domain. Hence, the extension distance only has to be two times the smallest local correlation range on the observational domain. This effectively minimizes the number of triangles needed in the extended domain while keeping a good FEM approximation of the true solution. Setting the extended triangles to the smallest correlation range is achieved by setting the Jacobian matrix $J[\mathrm{F}^{-1}](\boldsymbol{s}) = \frac{\sqrt{8\nu}}{r_{\min}}I$ for all triangles that are exclusively in the exterior of $\mathcal{G}$. Here, $r_{\min}$ denotes the minimum local correlation range in the interior. In Fig. 10, the blue triangles are part of the spatial domain of interest and the gray triangles are part of the extension.

Appendix B. Local parameter estimation

To obtain starting values for the parameter estimation, we first estimate the parameter values of the model of Section 3 locally in small spatial regions by approximating the parameters as constant in neighborhoods of 3×3 pixels, see Fig. 11. For each local estimate, we first estimated the marginal mean and variance for each location in the neighborhood using the regular sample mean and sample variance. We then compute the parameters of the constant H matrix by numerical optimization of the log-likelihood when assuming that the covariance function in the local neighborhood is defined as in Eq. (7). It should be noted that we do not estimate local parameters for all points but for a selection of points spread out evenly in the domain.
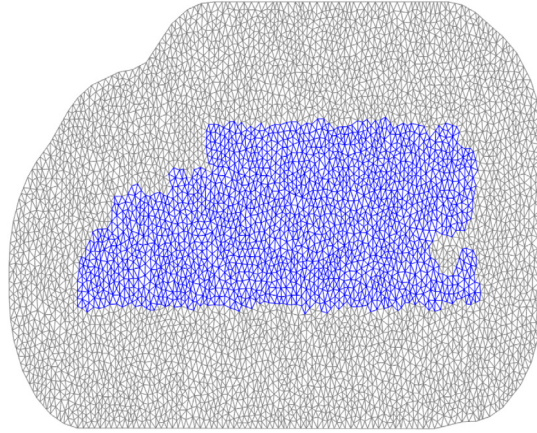
Fig. 10. The mesh of the north Atlantic used in Section 5. Triangles of the extension are shown in gray and triangles from the interior in blue. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
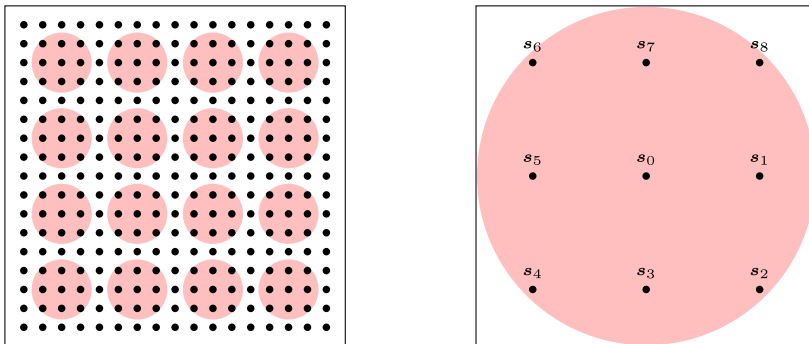
Fig. 11. Evenly spaced local neighborhoods are chosen for the local estimates. Data from the nine locations in each neighborhood is used for the estimates.

The local estimates are merged to get spatially varying parameters according to the series of cosines specified in Section 4. This is accomplished by computing the least squares solution of the values $\{\beta_{np}^i\}$ based on the local estimates. The merged values are then used to acquire good initial values for the numerical maximization of the likelihood function as described in Section 4.

An important feature of first estimating the parameters locally is that it can be used to construct the triangular mesh and choose the smoothness of the random field. It is important that the triangular mesh is fine enough to resolve the Gaussian random field approximation for a given $J[\mathrm{F}^{-1}]$. During the global optimization, $J[\mathrm{F}^{-1}]$ will vary and we should ideally update the mesh in every iteration of the numerical optimization. Unfortunately, meshing is one of the computationally most costly operations and it is therefore not feasible to remesh in each iteration. We instead precompute the mesh prior to the optimization, and use the local estimates to select a suitable maximum mesh resolution. Likewise, the smoothness parameter changes the structure of the FEM matrices drastically, and we therefore fix α based on the local estimates prior to estimating the other parameters globally.

Appendix C. Existence and estimation of $\mathcal{D}$-space

As mentioned in Section 4, we do not parameterize the model using the deformation function, F^{-1}, directly. This is not needed since the model only depends on F^{-1} through the $\tilde{H}$ matrix. However, knowing F^{-1} or $\mathcal{D}$ can in certain applications be of interest in itself since it might convey some physical interpretations, can be used to visualize deformations, and makes it possible to analyze data using techniques that relies on stationarity and/or isotropy. In order to acquire a $\mathcal{D}$-space from the estimated $\tilde{H}(\boldsymbol{s})$, three problems have to be taken care of. Firstly, $J[F^{-1}]$ is not uniquely defined by $\tilde{H}$. A simple example of this is the stationary case where it is not possible to identify maps, $F^{-1} : \mathcal{G} \to \mathcal{D}$, that only differ by one being the mirror of the other with respect to an arbitrary plane through the origin. In this case, we can restrict F^{-1} to the subset of all differentiable maps which have positive definite and symmetric Jacobian matrices. With that restriction we are able to identify $J[F^{-1}]$ from $\tilde{H}$ since they share eigenvectors and the eigenvalues of $J[F^{-1}]$ are just the inverse square root of those for $\tilde{H}$.

Secondly, for non-stationary fields where $J[F^{-1}]$ is spatially varying, the rows of $J[F^{-1}]$ correspond to vector fields. Since these vector fields are potential fields of the elements of F^{-1} they have to be conservative. It is not clear to the authors what further restrictions this puts on the spatially varying, symmetric, and positive definite matrices $\tilde{H}$.

Thirdly, even given a matrix-valued function $J[F^{-1}]$ that is the Jacobian of some function F^{-1}, F^{-1} is not necessarily injective. For the sake of the argument, let us consider F^{-1} as a mapping $F^{-1} : \mathcal{G} \to \mathcal{D}$, not necessarily injective. The inverse function theorem (Spivak, 1965) states that there exists a differentiable injective function in an open set around a point, $\boldsymbol{s}$, if $\left|J[F^{-1}](\boldsymbol{s})\right| \neq 0$. This is a necessary but unfortunately not sufficient condition to guarantee that F^{-1} is injective. When F^{-1} is not injective we say that $\mathcal{D}$ folds, i.e., several points in $\mathcal{G}$ maps to the same point in $\mathcal{D}$. Folding was also a problem using the thin-plate spline parameterization of F^{-1} in the original paper of the deformation method (Sampson and Guttorp, 1992). They reported issues with folding and recommended ameliorating this problem by forcing F^{-1} to be smooth enough in their parameterization. For our model folding is not an issue since F^{-1} only acts in a local sense through the value of the Jacobian. Hence, the inverse function theorem is enough for solutions to the SPDE to behave as intended. For us, folding is only a problem when the interest lies in properties of $\mathcal{D}$ itself. It is for us an open question how to find a practically useful parameterization of F^{-1} in a way that ensures no folding.

However, even if the estimated model folds, it should not occur in many places since $\tilde{H}$ is parameterized to vary smoothly. It is therefore possible to consider dividing $\mathcal{G}$ into subregions, $\{\mathcal{G}_k\}_k$. Each $\mathcal{G}_k$ are in turn mapped to a space $\mathcal{D}_k$ for which the data behaves as isotropic and stationary. Such a division can be used for instance to visualize the behavior of the estimated random field by merging overlapping subsets $\{\mathcal{D}_k\}_k$. This was done in Fig. 5 in order to visualize the deformation.

Assuming that the acquired Jacobian matrix is a Jacobian of some function, F^{-1}, for a two dimensional model, the mapping of points between $\mathcal{G}$ and $\mathcal{D}$ can be computed as

$$F^{-1}(\boldsymbol{s}) = F^{-1}(\boldsymbol{s}_0) + \begin{bmatrix} \int_0^1 a(\boldsymbol{r}(t))dr_x(t) \\ \int_0^1 c(\boldsymbol{r}(t))dr_x(t) \end{bmatrix} + \begin{bmatrix} \int_0^1 c(\boldsymbol{r}(t))dr_y(t) \\ \int_0^1 b(\boldsymbol{r}(t))dr_y(t) \end{bmatrix},$$

where

$$J[F^{-1}](\boldsymbol{s}) = \begin{bmatrix} a(\boldsymbol{s}) & c(\boldsymbol{s}) \\ c(\boldsymbol{s}) & b(\boldsymbol{s}) \end{bmatrix},$$

and the functions a, b, and c are defined through the spatially varying parameters h_i of the $\tilde{H}$ matrix. By identifying one node in the FEM mesh on $\mathcal{G}$ as mapping to a distinct point in $\mathcal{D}$, the mapping of all nodes to $\mathcal{D}$ can be acquired by successively mapping each neighboring node in a recursive manner. Points inside a triangle can be mapped to $\mathcal{D}$ using a linear approximation based on the location of the nodes of the triangle. We compute the integrals numerically by evaluating a, b, c, on the edges in the mesh and estimating the integral using Simpson's rule (Heath, 2002).
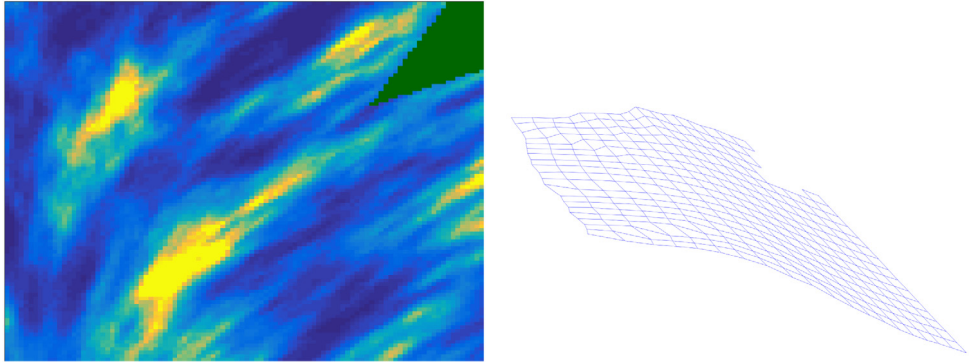
Fig. 12. Realization of a non-stationary and anisotropic Gaussian random field generated through the model of Section 3 (left), and $\mathcal{D}$ acquired from estimating the parameters of the model from data (right). The locations of a regular grid in $\mathcal{G}$ is shown in $\mathcal{D}$.

An example of an acquired $\mathcal{D}$ is shown in Fig. 12, where a realization of a non-stationary and anisotropic Gaussian random field on a non-rectangular domain, defined through the model of Section 3, is shown. Based on 182 realizations of the field, the parameters of the model was estimated using the procedure described in Section 4. The $\mathcal{D}$ region was estimated using the algorithm described above, and is shown in the right part of the figure. When observing such a figure, interpretation of the deformation comes from our knowledge that the connected nodes were on a regular grid in $\mathcal{G}$ before the deformation.

Appendix D. Proofs

D.1. A SPDE equivalent to a deformed Matérn field

To prove Theorem 3.1, we first need some additional notation. As previously mentioned, F : $\mathcal{D} \mapsto \mathcal{G} := F(\mathcal{D}) \subset \mathbb{R}^d$ is a bijective and differentiable function. We consider the slightly more general operator $\hat{\mathcal{L}}_c = (c^2 - \hat{\nabla} \cdot \hat{\nabla})$ on $\mathcal{D}$ for a constant $c > 0$. Similarly, we define an operator $\mathcal{L}_c = c^2 - \kappa^{-2}\nabla \cdot H\nabla$ on $\mathcal{G}$.

Let $\{\lambda_j\}_{j \in \mathbb{N}}$ denote the eigenvalues of $\hat{\mathcal{L}}_c$ in nondecreasing order, and let $\{e_j\}_{j \in \mathbb{N}}$ denote the corresponding eigenfunctions. As in Bolin et al. (2018), we use the spectral definition of the operator $\hat{\mathcal{L}}_c^\beta : \mathscr{D}(\hat{\mathcal{L}}_c^\beta) \to L_2(\mathcal{D})$. That is, for $\beta > 0$ and $\phi \in \mathscr{D}(\hat{\mathcal{L}}_c^\beta)$, the action of the operator is defined by

$$\hat{\mathcal{L}}_c^\beta \phi = \sum_{j \in \mathbb{N}} \lambda_j^\beta \langle \phi, e_j \rangle e_j.$$

We define the space $\dot{H}^{2\beta} := \mathscr{D}(\hat{\mathcal{L}}_c^\beta)$, which is a Hilbert space with inner product $\langle \hat{\mathcal{L}}_c^\beta \phi, \hat{\mathcal{L}}_c^\beta \psi \rangle$. Further, $\dot{H}^{-2\beta}$ is defined as the dual space of $\dot{H}^{2\beta}$ and $\dot{H}^0 := L_2(\mathcal{D})$ (see Bolin et al., 2018, for further details).

Lemma 2.1 in Bolin et al. (2018) shows that for $s \in \mathbb{R}$, $\hat{\mathcal{L}}_c$ can be extended to an isometric isomorphism $\hat{\mathcal{L}}_c : \dot{H}^s \to \dot{H}^{s-2\beta}$. This means that if $\hat{g} \in \dot{H}^s$ $\mathbb{P}$-a.s., for $s \geq -2\beta$, then $\hat{\mathcal{L}}_c^\beta \hat{X} = \hat{g}$ has a unique solution in $L_2(\Omega, L_2(\mathcal{D}))$. In particular, we have by Proposition 2.3 in Bolin et al. (2018) that white noise on $\mathcal{D}$ is in $\dot{H}^{-d/2-\epsilon}$ $\mathbb{P}$-a.s. for any $\epsilon > 0$. For $\hat{g}$, we define the corresponding function $g := \hat{g} \circ F^{-1}$.

The proof of Theorem 3.1 will make use of three lemmas. The following lemma connects the two operators $\hat{\mathcal{L}}_c$ and $\mathcal{L}_c$.

Lemma D.1. *Assume that $\hat{g} \in \dot{H}^s$ $\mathbb{P}$-a.s. for $s \geq -2$ and let $\hat{X}$ be the solution to*

$$\hat{\mathcal{L}}_c \hat{X} = \hat{g} \quad in \ \mathcal{D}.$$

Then, $X(\boldsymbol{s}) := \hat{X}(F^{-1}\boldsymbol{s})$ is the solution to

$$\mathcal{L}_c X = g \quad in \ \mathcal{G}.$$

Proof. The operator $\hat{\mathcal{L}}_c$ induces a symmetric, continuous and coercive bilinear form $B_{\hat{\mathcal{L}}_c}$ on $V_{\mathcal{D}} := H_0^1(\mathcal{D})$,

$$B_{\hat{\mathcal{L}}_c} : V_{\mathcal{D}} \times V_{\mathcal{D}} \to \mathbb{R}, \qquad B_{\hat{\mathcal{L}}_c}(\hat{u}, \hat{v}) := c^2 \langle \hat{u}, \hat{v} \rangle_{L_2(\mathcal{D})} + \langle \hat{\nabla}\hat{u}, \hat{\nabla}\hat{v} \rangle_{L_2(\mathcal{D})}.$$

The solution to $\hat{\mathcal{L}}_c \hat{u} = \hat{g}$ in $\mathcal{D}$ can be viewed as the unique element $\hat{u} \in V_{\mathcal{D}}$ satisfying $B_{\hat{\mathcal{L}}_c}(\hat{u}, \hat{v}) = \hat{g}(\hat{v})$ for all $\hat{v} \in V_{\mathcal{D}}$.

To obtain the desired result we perform a change of variables in this variational formulation. First, we have

$$\langle \hat{u}, \hat{v} \rangle_{L_2(\mathcal{D})} = \int_{\mathcal{D}} \hat{u}(\hat{\boldsymbol{s}})\hat{v}(\hat{\boldsymbol{s}})d\hat{\boldsymbol{s}} = \int_{\mathcal{G}} \frac{\hat{u}(F^{-1}\boldsymbol{s})\hat{v}(F^{-1}\boldsymbol{s})}{|J[F](F^{-1}\boldsymbol{s})|}d\boldsymbol{s}$$

$$= \int_{\mathcal{G}} u(\boldsymbol{s})v(\boldsymbol{s})\kappa^2(\boldsymbol{s})d\boldsymbol{s} = \langle \kappa^2 u, v \rangle_{L_2(\mathcal{G})},$$

where we used that $J[F](F^{-1}(\boldsymbol{s})) = J[F^{-1}]^{-1}(\boldsymbol{s})$ due to the inverse function theorem. We also defined $v(\boldsymbol{s}) = \hat{v}(F^{-1}(\boldsymbol{s}))$, and one should note that every function $v \in V_{\mathcal{G}} := H_0^1(\mathcal{G})$ can be represented in this way. Now, if we let ∇ denote the gradient on $\mathcal{G}$,

$$\langle \hat{\nabla}\hat{u}, \hat{\nabla}\hat{v} \rangle_{L_2(\mathcal{D})} = \int_{\mathcal{D}} \left[\hat{\nabla}\hat{u}(\hat{\boldsymbol{s}}) \right]^T \hat{\nabla}\hat{v}(\hat{\boldsymbol{s}})d\hat{\boldsymbol{s}}$$

$$= \int_{\mathcal{G}} \frac{\left[J[F]^T(F^{-1}\boldsymbol{s})\nabla\hat{u}(F^{-1}\boldsymbol{s}) \right]^T \left[J[F]^T(F^{-1}\boldsymbol{s})\nabla\hat{v}(F^{-1}\boldsymbol{s}) \right]}{\left| J[F](F^{-1}\boldsymbol{s}) \right|}d\boldsymbol{s}$$

$$= \int_{\mathcal{G}} \left[J[F^{-1}]^{-T}(\boldsymbol{s})\nabla u(\boldsymbol{s}) \right]^T \left[J[F^{-1}]^{-T}(\boldsymbol{s})\nabla v(\boldsymbol{s}) \right]\kappa^2(\boldsymbol{s})d\boldsymbol{s}$$

$$= \left\langle \kappa J[F^{-1}]^{-T}\nabla u, \kappa J[F^{-1}]^{-T}\nabla v \right\rangle_{L_2(\mathcal{G})}.$$

Thus, the bilinear form, $B_{\hat{\mathcal{L}}_c}$, corresponds to a bilinear form, $B_{\mathcal{L}} : V_{\mathcal{G}} \times V_{\mathcal{G}} \to \mathbb{R}$

$$B_{\mathcal{L}}(u, v) := c^2 \langle \kappa^2 u, v \rangle_{L_2(\mathcal{G})} + \left\langle \kappa J[F^{-1}]^{-T}\nabla u, \kappa J[F^{-1}]^{-T}\nabla v \right\rangle_{L_2(\mathcal{G})},$$

induced by the operator $\kappa^2 \mathcal{L}_c = \kappa^2 c^2 - \nabla \cdot H \nabla$.

Performing the same change of variables on the right-hand side yields

$$\hat{g}(\hat{v}) = \int_{\mathcal{D}} \hat{v}(\hat{\boldsymbol{s}})\hat{g}(\hat{\boldsymbol{s}})d\hat{\boldsymbol{s}} = \int_{\mathcal{G}} \hat{v}(F^{-1}\boldsymbol{s})\hat{g}(F^{-1}\boldsymbol{s})\kappa^2(\boldsymbol{s})d\boldsymbol{s}$$

$$= \int_{\mathcal{G}} v(\boldsymbol{s})\hat{g}(F^{-1}\boldsymbol{s})\kappa^2(\boldsymbol{s})d\boldsymbol{s}.$$

This implies $\mathcal{L}_c X = g$ in $\mathcal{G}$ since $\kappa(\boldsymbol{s}) > 0, \forall \boldsymbol{s} \in \mathcal{G}$. $\quad\square$

In Eq. (5) the right-hand side is white noise, i.e., $\hat{g} = \hat{\mathcal{W}}$. The following lemma identifies the distribution of g for this case.

Lemma D.2. *If $\hat{g}$ is a white noise on $\mathcal{D}$, then $g \overset{D}{=} \kappa^{-1}\mathcal{W}$. That is, g is a white noise on $\mathcal{G}$ scaled by κ^{-1}.*

Proof. Let $\hat{\mathcal{W}}$ be a white noise on $\mathcal{D}$. Due to the properties of white noise and a change of variables,

$$g(v) := (\hat{g} \circ F^{-1})(v) = \int_{\mathcal{G}} \hat{\mathcal{W}}(F^{-1}\mathbf{s})v(\mathbf{s})d\mathbf{s} = \int_{\mathcal{D}} \hat{\mathcal{W}}(\hat{\mathbf{s}})v(F\hat{\mathbf{s}})\kappa^{-2}(F\hat{\mathbf{s}})d\hat{\mathbf{s}}$$

$$= \hat{\mathcal{W}}\left(\hat{v}(\kappa^{-2} \circ F)\right) \sim \mathbb{N}\left(0, \int_{\mathcal{D}} \left(\hat{v}^2(\hat{\mathbf{s}})\kappa^{-2}(F\hat{\mathbf{s}})\right)^2 d\hat{\mathbf{s}}\right)$$

$$= \mathbb{N}\left(0, \int_{\mathcal{G}} \left(v(\mathbf{s})\kappa^{-2}(\mathbf{s})\right)^2 \kappa^2(\mathbf{s})d\hat{\mathbf{s}}\right) = \mathbb{N}\left(0, \int_{\mathcal{G}} (v(\mathbf{s})\kappa^{-1}(\mathbf{s}))^2 d\hat{\mathbf{s}}\right). \quad \square$$

One can note that Lemma D.1 together with Lemma D.2 can be used to show the result of Theorem 3.1 in the non-fractional case. It is however not clear if $\hat{\mathcal{L}}_c^\beta$ maps to $\mathcal{L}_c^\beta$ for arbitrary β. The following lemma is necessary in order to show this.

Lemma D.3. *Let $\beta := \alpha/2 \in (0, 1)$ and assume that $\hat{g} \in \dot{H}^s$, $\mathbb{P}$-a.s., for $s \geq -2\beta$. Let $\hat{X}$ be the solution to*

$$\hat{\mathcal{L}}_c^\beta \hat{X} = \hat{g} \quad \text{in } \mathcal{D}.$$

Then, $X(\mathbf{s}) := \hat{X}(F^{-1}\mathbf{s})$ is the solution to

$$\mathcal{L}_c^\beta X = g \quad \text{in } \mathcal{G}.$$

Proof. Introducing the composition operator $\mathcal{C}_F : V_{\mathcal{D}} \to V_{\mathcal{G}}$ defined by $\mathcal{C}_F v = v \circ F$, we can state the non-fractional result of Lemma D.1 as

$$X = \mathcal{L}_c^{-1}g = \mathcal{C}_F \hat{\mathcal{L}}_c^{-1}\hat{g}.$$

Now, in the fractional case we have defined the fractional power of the operator using the spectral definition. We can therefore use the following representation of the fractional inverse from the Dunford–Taylor calculus by Balakrishnan (1960). For $\beta \in (0, 1)$,

$$\hat{X} = (c^2 - \hat{\nabla} \cdot \hat{\nabla})^{-\beta}\hat{g} = \frac{2\sin(\pi\beta)}{\pi} \int_0^\infty t^{2\beta-1} \left(I + t^2(c^2 - \hat{\nabla} \cdot \hat{\nabla})\right)^{-1} dt\hat{g}.$$

What we need to show is that the corresponding fractional model on $\mathcal{G}$,

$$X = \left(c^2 - \kappa^{-2}\nabla \cdot H\nabla\right)^{-\beta} g$$

is equal to $\mathcal{C}_F(c^2 - \hat{\nabla} \cdot \hat{\nabla})^{-\beta}\hat{g}$. Using the Balakrishnan representation of the fractional inverse again, we have

$$X = \left(c^2 - \kappa^{-2}\nabla \cdot H\nabla\right)^{-\beta} g$$

$$= \frac{2\sin(\pi\beta)}{\pi} \int_0^\infty t^{2\beta-1} \left(I + t^2(c^2 - \kappa^{-2}\nabla \cdot H\nabla)\right)^{-1} dtg$$

$$= \frac{2\sin(\pi\beta)}{\pi} \int_0^\infty t^{2\beta-3} \left(t^{-2} + c^2 - \kappa^{-2}\nabla \cdot H\nabla\right)^{-1} gdt.$$

At this stage, we can use Lemma D.1 with operator $\mathcal{L}_C$ and $C^2 := t^{-2} + c^2$ to change from $\mathcal{G}$ to $\mathcal{D}$ in the integral. Also using that $\mathcal{C}_F$ is a bounded linear operator, we get

$$X = \frac{2\sin(\pi\beta)}{\pi} \int_0^\infty t^{2\beta-3} \mathcal{C}_F \left(C^2 - \hat{\nabla} \cdot \hat{\nabla}\right)^{-1} \hat{g}dt$$

$$= \mathcal{C}_F \frac{2\sin(\pi\beta)}{\pi} \int_0^\infty t^{2\beta-1}(I + t^2\left(c^2 - \hat{\nabla} \cdot \hat{\nabla}\right))^{-1}\hat{g}dt$$

$$= \mathcal{C}_F(c^2 - \nabla \cdot \nabla)^{-\beta}\hat{g}. \quad \square$$

Given these lemmas we are now ready for the proof of Theorem 3.1.

Proof of Theorem 3.1. Recall that for $s \in \mathbb{R}$, $\hat{\mathcal{L}}_c$ is an isometric isomorphism $\hat{\mathcal{L}}_c : \dot{H}^s \to \dot{H}^{s-2\beta}$. Further recall that white noise on $\mathcal{D}$ is in $\dot{H}^{-d/2-\epsilon}$ $\mathbb{P}$-a.s. for any $\epsilon > 0$. With these facts, Lemma D.1 together with Lemma D.2 directly give the result for $\alpha = 2$. Also, Lemma D.3 give the result for all $\alpha \in (d/2, 1]$.

The spectral definition of fractional operators yields that the operators adhere to the law

$$\mathcal{L}^\alpha X = \mathcal{L}^{\alpha-1}(\mathcal{L}X).$$

Hence, the result for a general $\alpha > d/2$ readily follows by induction and the fact that $\mathcal{L}_c^\beta X = \mathcal{W}$ can be represented as $\mathcal{L}_c^{\tilde{\beta}} u = X_0$, where $\tilde{\beta}$ is the integer part of β and $\mathcal{L}^{\beta-\tilde{\beta}} X_0 = \mathcal{W}$, with $\beta - \tilde{\beta} \in (0, 1)$. $\square$

D.2. Exceedance probability

Proof of Proposition 2.1. Let $N_T(v)$ be the number of upcrossings for $X_\gamma(t)$ at a threshold value, u, on the route. Then by Rice's method of moments (Piterbarg, 1996) we have

$$\mathbb{P}\left[\max_{t \in [0,T]} X_\gamma(t) > u\right] = \mathbb{P}\left[X_\gamma(0) > u\right] + \mathbb{P}\left[(N_T(u) > 0) \cap (X_\gamma(0) \le u)\right]$$

$$\le \mathbb{P}\left(X_\gamma(0) > u\right) + \mathbb{E}\left[N_T(u)\right].$$

Rice's formula (Azaïs and Wschebor, 2009; Rice, 1944, 1945) give an explicit expression,

$$\mathbb{E}\left[N_T(u)\right] = \int_0^T \mathbb{E}\left[|\dot{X}_\gamma(t)|\mathbb{I}\left(\dot{X}_\gamma(t) \ge 0\right)\big| X_\gamma(t) = u\right] f_{X_\gamma(t)}(u) dt, \tag{10}$$

where $\dot{X}_\gamma$ denotes the mean square derivative of $X_\gamma(t)$ and $\mathbb{I}$ is the indicator function. Now, if X_γ would have constant variance, the derivative of the process at t would not be correlated with the value of the process at t (Adler and Taylor, 2007, section 5.6). To make use of this property, we introduce the standardized Gaussian process $W(t) := \frac{X_\gamma(t) - \mu_\gamma(t)}{\sigma_\gamma(t)}$ and get

$$\dot{X}_\gamma(t) = \dot{W}(t)\sigma_\gamma(t) + W(t)\dot{\sigma}_\gamma(t) + \dot{\mu}_\gamma(t).$$

$W(t)$ and $\dot{W}(t)$ for a fixed t are independent. Hence,

$$\mathbb{E}\left[|\dot{X}_\gamma(t)|\mathbb{I}\left(\dot{X}_\gamma(t) \ge 0\right)\big| X_\gamma(t) = u\right] = \sigma_\gamma(t)\mathbb{E}\left[\dot{W}(t)\mathbb{I}\left(\dot{W}(t) \ge -a(t)\right)\right] + \sigma_\gamma(t)a(t),$$

where $a = \frac{\dot{\mu}_\gamma(t)}{\sigma_\gamma(t)} + \dot{\sigma}_\gamma(t)\frac{u-\mu_\gamma(t)}{\sigma_\gamma^2(t)}$. Since $\dot{W}(t)$ is a centered random process,

$$\mathbb{E}\left[\dot{W}(t)\mathbb{I}\left(\dot{W}(t) \ge -a(t)\right)\right] = \sigma_{\dot{W}}(t)\phi\left(\frac{a(t)}{\sigma_{\dot{W}}(t)}\right).$$

Now, using Eq. (10), the fact that $f_{X_\gamma(t)}(u) = \frac{1}{\sigma_\gamma(t)}\phi\left(\frac{u-\mu_\gamma(t)}{\sigma_\gamma(t)}\right)$, and plugging the result into the first inequality finishes the proof. $\square$

The general choice of model for $\mu_\gamma(t)$ and $\sigma_\gamma(t)$, as well as the choice of numerical approximation of the integral in Proposition 2.1, are beyond the scope of this work. However, a quantity present in Proposition 2.1 which is highly relevant to the work of this paper is the variance of the mean square derivative of $W(t)$, i.e., $\sigma_{\dot{W}}^2(t)$. We will now show how this variance can be computed when $X(\boldsymbol{s}, t)$ is a spatio-temporal Gaussian random process defined using a stationary covariance function and a (local) spatial deformation. By local we mean that there exists a function, F^{-1}, satisfying the inverse function theorem (Spivak, 1965) over any Lebesgue measurable set of the spatial domain. This means that the proof will hold for any deformation model (Sampson and Guttorp, 1992), any locally deformed Matérn SPDE model, but also any other local deformation model.

In Section 3 we referred to $\mathcal{D}$ as the space where our spatial model would be stationary, isotropic, and for which the correlation was Matérn. We now make a similar definition for a spatio-temporal process. Assume that X is defined through a local deformation approach using a function

$F^{-1}(\boldsymbol{s}, t) : \mathcal{G}_s \times \mathcal{G}_t \to \mathcal{D}_s \times \mathcal{D}_t(\boldsymbol{s}, t)$ such that the process on $\mathcal{D}_s \times \mathcal{D}_t(\boldsymbol{s}, t)$ is stationary but not necessarily isotropic. Notice how the relaxation of a global deformation requires the product space $\mathcal{D}_s \times \mathcal{D}_t(\boldsymbol{s}, t)$ to be indexed by points in $\mathcal{G}_s \times \mathcal{G}_t$, i.e. a local mapping and a local space for each neighborhood. Let $\tilde{X}'_x, \tilde{X}'_y, \ldots, \tilde{X}'_t$ denote the partial derivatives of $\tilde{X}$. Then, $\sigma_{\dot{W}}(t)$ can be computed using the following proposition.

Proposition D.4. *Let $\sigma_{ij}^2 = \mathbb{C}(\tilde{X}'_i(\tilde{\boldsymbol{s}}, \tilde{t}), \tilde{X}'_j(\tilde{\boldsymbol{s}}, \tilde{t}))$, for $i, j \in \{1, \ldots, d, d+1\}$, i.e., the set of all spatial dimensions and the temporal. With Σ the corresponding matrix $\Sigma_{ij} = \{\sigma_{ij}^2\}$ we have*

$$\sigma_{\dot{W}}^2(t) = \dot{\gamma}(t)^T J[F^{-1}](\gamma(t)) \Sigma J[F^{-1}]^T(\gamma(t)) \dot{\gamma}(t),$$

where $\gamma(t) = [x(t), y(t), \ldots, t]^T$ and $\dot{\gamma}(t) = [\dot{x}(t), \dot{y}(t), \ldots, 1]^T$ is the velocity of the curve at time t.

Proof. By the chain rule we get

$$\dot{W}(t) = \dot{\gamma}(t) \cdot \left[\nabla X(\gamma(t)), \dot{X}(\gamma(t))\right]^T = \dot{\gamma}(t)^T J\left[F^{-1}\right]^T(\gamma(t)) \left[\tilde{\nabla}\tilde{X}(\tilde{\gamma}(t)), \dot{\tilde{X}}(\tilde{\gamma}(t))\right]^T.$$

Due to stationarity of $\tilde{X}(\tilde{\gamma}(t))$, the marginal distribution with respect to space–time of the partial derivatives of $\tilde{X}(\tilde{\gamma}(t))$ is a $d+1$-dimensional Gaussian distribution independent of $\gamma(t)$. Let Σ denote the covariance matrix of this multivariate Gaussian distribution. Then,

$$\Sigma := \mathbb{E}\left[\left[\tilde{\nabla}\tilde{X}(\tilde{\gamma}(t)), \dot{\tilde{X}}(\tilde{\gamma}(t))\right]^T \left[\tilde{\nabla}\tilde{X}(\tilde{\gamma}(t)), \dot{\tilde{X}}(\tilde{\gamma}(t))\right]\right] = \Sigma.$$

The proof is concluded by the fact that

$$\sigma_{\dot{W}}^2(t) = \mathbb{E}\left[\dot{W}(t)\dot{W}(t)^T\right] = \dot{\gamma}(t)^T J[F^{-1}]^T(\gamma(t)) \Sigma J[F^{-1}](\gamma(t)) \dot{\gamma}(t). \quad \square$$

The point of Proposition D.4 is that when a model i specified as having a stationary covariance function given a (local) deformation, that covariance function is usually available in explicit form and well understood. This means that, in most cases, Σ can be computed explicitly and hence also $\sigma_{\dot{W}}(t)$. It should be noted that all pairs of sub-dimensions that are isotropic in $\mathcal{D}$, given the other dimensions being fixed, have partial derivatives of $\tilde{X}$ that are independent of each other (Adler and Taylor, 2007, section 5.7). For a covariance function in the deformed space that is not only stationary but isotropic as well, this yields

$$\sigma_{\dot{W}}^2(t) = \sigma_{ii}^2 \dot{\boldsymbol{s}}(t)^T J[F^{-1}]^T(\boldsymbol{s}(t)) J[F^{-1}](\boldsymbol{s}(t)) \dot{\boldsymbol{s}}(t),$$

for any $i \in \{1, \ldots, d+1\}$ since $\sigma_{xx}^2 = \sigma_{yy}^2 = \cdots = \sigma_{tt}^2$ due to isotropy. Moreover, for the case of a spatio-temporal Matérn covariance in $\mathcal{D}$,

$$\sigma_{ii}^2 = \frac{\sigma^2 \kappa^2}{2(\nu - 1)}, \quad \forall i \in \{1, \ldots, d+1\}.$$

This can be derived from the spectral density of the Matérn covariance function (Stein, 1999) together with the fact that the second spectral moment in the direction of dimension i equals the variance of the partial derivative process in the same direction (Adler and Taylor, 2007, section 5.5). Notice that we need $\nu > 1$ for the result to hold since the field is otherwise not mean-square differentiable.

Proof of Corollary 3.2. For the special case of the model of Section 3 that is constant over time, the elements in the last row and column of Σ are zero since $\dot{\tilde{X}}(\tilde{\boldsymbol{s}}, \tilde{t}) \equiv 0$. In this model, the spatial dimensions have an isotropic Matérn covariance structure with unit dampening and unit variance, hence

$$\sigma_{\dot{W}}^2(t) = \frac{\kappa^2(t)}{2(\nu - 1)} \dot{\boldsymbol{s}}(t)^T H^{-1}(\boldsymbol{s}(t)) \dot{\boldsymbol{s}}(t).$$

It should be clarified that $\kappa(t)$ above refer to the function $\kappa(\boldsymbol{s}(t))$ from Section 3 over the curve, γ. It is not the dampening in the deformed space—which we know is constant and equal to one. $\quad \square$

References

Adler, R.J., Taylor, J.E., 2007. Random Fields and Geometry. Springer, ISBN: 978-0-387-48112-8, p. 464.

Ailliot, P., Baxevani, A., Cuzol, A., Monbet, V., Raillard, N., 2010. Space-time models for moving fields with an application to significant wave height fields. Environmetrics 22 (3), 354–369.

Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. In: Second International Symposium of Information Theory. pp. 267–281.

Azaïs, J.-M., Wschebor, M., 2009. Level Sets and Extrema of Random Processes and Fields, first ed. Wiley, p. 393, ISBN 04704093393.

Bakka, H., Vanhatalo, J., Illian, J.B., Simpson, D., Rue, H., 2019. Non-stationary Gaussian models with physical barriers. Spatial Stat. 29, 268–288.

Balakrishnan, A., 1960. Fractional powers of closed operators and the semigroups generated by them. Pac. J. Math. 10, 419–437.

Baxevani, A., Caires, S., Rychlik, I., 2009. Spatio-temporal statistical modelling of significant wave height. Environmetrics 20, 14–31.

Baxevani, A., Podgorski, K., Rychlik, I., 2011. Dynamically evolving Gaussian spatial fields. Extremes 14 (2), 223–251.

Baxevani, A., Rychlik, I., Wilson, R.J., 2005. A new method for modelling the space variability of significant wave height. Extremes 8 (4), 267–294.

Bolin, D., Kirchner, K., 2019. The rational SPDE approach for Gaussian random fields with general smoothness. J. Comput. Graph. Stat..

Bolin, David, Kirchner, Kristin, Kovács, Mihály, 2018. Numerical solution of fractional elliptic stochastic PDEs with spatial white noise. IMA J. Numer. Anal. http://dx.doi.org/10.1093/imanum/dry091.

Bolin, David, Lindgren, Finn, 2011. Spatial models generated by nested stochastic partial differential equations, with an application to global ozone mapping. Ann. Appl. Statist. 5 (1), 523–550.

Bolin, D., Wallin, J., 2020. Multivariate Type-G Matérn fields. J. R. Stat. Soc. Ser. B 82 (1), 215–239.

Casella, G., Berger, R.L., 2002. Statistical Inference, second ed. Thomson Learning, ISBN: 0-534-24312-6, p. 660.

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Holm, E.V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.J., Park, B.K., Peubey, C., de Rosnay, P., Tavolato, C., Thepaut, J.N., Vitart, F., 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q. J. R. Meterol. Soc. April (A), 553–597.

Fuglstad, G.A., Lindgren, F., Simpson, D., Rue, H., 2015a. Exploring a new class of non-stationary spatial Gaussian random fields with varying local anisotropy. Statist. Sinica 25 (1), 115–133.

Fuglstad, G.A., Simpson, D., Lindgren, F., 2015b. Does non-stationary spatial data always require non-staionary random fields? Spatial Stat. 14, 505–531.

Heath, M.T., 2002. Scientific Computing, second ed. McGraw-Hill Professional, ISBN: 0-07-239910-4.

Higdon, D., 1998. A process-convolution approach to modelling temperatures in the north atlantic ocean. Environ. Ecol. Stat. 5, 173–190.

Hu, X., Simpson, D., Lindgren, F., Rue, H., 2018. Multivariate Gaussian random fields using systems of stochastic partial differential equations. arXiv preprint arXiv:1711.04333v2.

Jasper, N.H., 1956. Statistical distribution patterns of ocean waves and of wave-induced ship stresses and motions with engineering applications. Trans. Soc. Nav. Archit. Mar. Eng. 64, 375–432.

Khristenko, U., Scarabosio, L., Swierczynski, P., Ullmann, E., Wohlmuth, B., 2019. Analysis of boundary effects on PDE-based sampling of Whittle-Matérn random fields. SIAM/ASA J. Uncertain. Quantif. 7 (3), 948–974.

Lewis, E.V., 1988. Principles of Naval Architecture, Vol. 1. Society of Naval architects and marine engineers.

Lindgren, F., Rue, H., Lindström, J., 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. J. R. Stat. Soc. 73 (4), 423–498.

Mao, W., Ringsberg, J.W., Rychlik, I., Storhaug, G., 2010. Development of a fatigue model useful in ship routing design. J. Ship Sesearch 54 (4), 281–293.

Munk, W.H., 1944. Proposed Uniform Procedure for Observing Waves and Interpreting Instrument Records, Vol. 22. Wave Project at the Scripps Institute of Oceanography, La Jolla, California.

Ochi, M.K., 1978. On the long-term statistics for ocean and coastal waves. In: Proc. 16th Conf. Coastal Engineering, pp. 59–75.

Paciorek, C.J., Schervish, M.J., 2006. Spatial modelling using a new class of nonstationary covariance functions. Environmetrics 17 (5), 483–506.

Piterbarg, V.I., 1996. Asymptotic Methods in the Theory of Gaussian Processes and Fields, Vol. 148. American Mathematical Society, ISBN: 0-8218-0423-5, p. 206.

Podgórski, K., Rychlik, I., 2014. A model of significant wave height for reliability assessment of a ship. J. Mar. Syst. 130, 109–123.

Rice, S., 1944. The mathematical analysis of random noise part i. Bell Syst. Tech. J. 23, 282–332.

Rice, S., 1945. The mathematical analysis of random noise part ii. Bell. Syst. Tech. J. 24, 46–156.

Rue, H., Held, L., 2005. Gaussian Markov Random Fields, Vol. 104. Chapman and Hall, ISBN: 0203492021, p. 280.

Rychlik, I., 1993. On the 'narrow-band' approximation for expected fatigue damage. Probab. Eng. Mech. 8 (1), 1–4.

Sampson, P.D., 2010. Constructions for nonstationary spatial processes. In: Gelfand, A.E., Diggle, P.J., Fuentes, M., Guttorp, P. (Eds.), Handbook of Spatial Statistics. Taylor and Francis, ISBN: 9781420072884, p. 575.

Sampson, P.D., Guttorp, P., 1992. Nonparametric estimation of nonstationary spatial covariance structure. J. Amer. Statist. Assoc. 87 (417), 108–119.

Spivak, M., 1965. Calculus on Manifolds. Addison-Wesley, ISBN: 0-8053-9021-9, p. 146.

Stein, M.L., 1999. Interpolation of Data. Springer.

Whittle, P., 1954. On stationary processes in the plane. Biometrika 41 (3/4), 434–449.