Spatial modeling of significant wave height using stochastic partial differential equations

Anders Hildeman, David Bolin, and Igor Rychlik

Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, Sweden

Abstract

The general condition of the ocean surface at a certain location in space and time is described by the sea state. Knowing the distribution of the sea state is, for example, important when estimating the wear and risks associated with a planned journey of a ship. One important characteristic of the sea state is the significant wave height. We propose a spatial model for the logarithm of significant wave height based on a continuously indexed Gaussian random field defined as a solution to a stochastic partial differential equation (SPDE). The SPDE is obtained by combining the SPDE representation of a stationary Gaussian Matérn field with the deformation approach by Sampson and Guttorp [33]. The resulting model can capture both non-stationarity and anisotropy, has beneficial computational properties, and easily interpretable parameters. We also show that the introduction of non-stationarity through the deformation approach allow us to derive theoretical bounds for exceedance probabilities of the field. Such bounds are of importance when modelling extreme loads on ships.

The parameters of the model are estimated on data of the north Atlantic taken from the ERA-Interim data set. 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 computed distributions of fatigue damage and exceedence probability is shown to agree well with data.

1 Introduction

The local distribution of 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.

A common assumption is that the elevation of the sea surface at a fixed spatial location can be modelled as a stationary Gaussian process, typically described through its power spectrum \( S(\omega) \). Several parametric power spectrum models have been proposed, one such is the Pierson-Moskowitz spectrum [12] which is often used to model fully developed seas. An important parameter for such models is the significant wave height, \( H_s \), which traditionally has been defined as the average wave height among the one third highest individual waves [26]. This definition of the significant wave height was intended to mathematically express the average wave-height estimated by a "trained observer". For a stationary Gaussian process, the mathematical definition of \( H_s \) is \( H_s = 4\sigma \), where \( \sigma \) is the standard deviation of the process.

Even though \( H_s \) only provides limited information about the sea state it is useful in many 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 $H_s$ [24, 22]. With a good understanding of the spatial distribution of the sea state and how it interact with the fatigue damage, shipping lines can plan their routes and maintenance intervals of the ship accordingly. This will lead to safer and more cost effective shipping.

The assumption of sea conditions as homogeneous for larger spatial regions or stationary for longer time periods rarely holds. However, by considering the parameters of $S(\omega)$ as varying in time and space it is possible to characterize the distributions of the sea state through stochastic models of the parameters of $S(\omega)$. Probabilistic models of $H_s$ for a fixed point in space and time have been studied extensively. Jasper [20] and Ochi [27] showed that a log-normal distribution fits the bulk of the marginal probability distributions of $H_s$ for data from the North Atlantic. Athanassoulis et al. [4] argued that the tail behavior is often deviating from the log-normal regime but that the log-normal distribution can be used for moderate-value analysis such as fatigue-life analysis and estimation of wave-energy resources.

These models have also been extended to temporal stochastic process models for $H_s$ for fixed points in space. Typically the temporal dependence is modelled using ARMA-processes [25] and the reference within. Temporal models are of importance for example when analyzing accumulated stress/fatigue of static structures in an ocean environment [14]. Spatial modeling of $H_s$ is also of great importance. For instance to interpolate observations or when considering moving ships where the wave state at points visited on the ships route will be highly dependent. Such spatial models for $H_s$ are usually based on Gaussian fields with some chosen parametric stationary covariance function [7]. However, an important property of a spatial model for any larger region is that it allows for spatial non-stationarity [8, 2]. Baxevani et al. [8] modelled non-stationarity by a moving average process using a kernel with location-dependent bandwidth. Ailliot et al. [2] instead considered mutually exclusive subregions of the spatial domain for which they assumed stationarity within.

The aim of this paper is to develop a fully non-stationary model for $H_s$ with greater flexibility than the previously proposed ones. The model is based on a description of the random field through a stochastic partial differential equation (SPDE) which is obtained by combining the SPDE representation of a stationary Gaussian Matérn field [23] with the deformation approach by Sampson and Guttorp [33]. Compared to the covariance based models of [7, 8, 2], the SPDE-based model allows for greater flexibility, has beneficial computational properties, and easily interpretable parameters. We also show that the introduction of non-stationarity in the model through the deformation allow us to derive theoretical bounds for exceedance probabilities of the fields. Such bounds are of importance when modelling extreme loads on ships.

As a motivating example, we use 38 years of data from the North Atlantic during the month of April. Particularly, we want to investigate if the model can be used to estimate damage on ships traveling on the sea. We will both consider damages caused by large waves as well as fatigue damage. For the ship damage tests we focus on the transatlantic shipping route shown in Figure 1 which is an important route in naval logistics.

The structure of the article is as follows. In Section 2 the data for the application is presented. In Section 3 the proposed model is derived. Section 4 presents some important properties of the model, and an upper bound for extreme seas is derived for the model in Section 5. Section 6 introduces a method for estimating the model parameters based on data. The results of the application of the model to the North Atlantic data are presented in Section 7. Finally Section 8 concludes with a discussion of the results and potential further applications and extensions of the model. The manuscript also contains three appendices which present further details on the model and the estimation procedure.
2 Data

We use data from the ERA-Interim global atmospheric reanalysis \cite{15} 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 data set is approximately 80 km and data is available from 1979 to present. The variable of interest in the dataset is significant wave height of wind and ground swells which is available at a temporal resolution of 6 hours. The values produced for a fixed point in time are based on measurements from a temporal window of ±6 hours. Hence there is a 3 hour overlap in between data from consecutive points in time which will cause a smoothing effect. We will not model the temporal evolution and therefore want to approximate data from different points in time as independent. Hence, we choose to thin the data to a temporal resolution of 24 hours.

We restrict our domain of interest to the north Atlantic 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 the pointwise marginal distributions for any chosen month are well approximated by log-Gaussian distributions Jasper \cite{20} and Ochi \cite{27}. We also restrict the analysis to the month of April, but using data from all of the 38 years (1979-2017). The reason for the 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. April was chosen since not only the bulk but also the tails of the logarithmized marginal data was described well by a normal distribution. However, any other month could have been chosen instead since we are mainly concerned with the bulk behavior.

The mean and variance of the logarithmized data can be seen in Figure 2. 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 Figure 5 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 has 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 shapes of anisotropy. However, it is clear that no stationary model (on the sphere or in the plane) can explain the observed behaviour.
3 Model formulation

A Gaussian random field is characterized completely by its mean and covariance functions. The Matérn covariance function is a popular choice due to its flexibility using only three easily interpretable parameters. It is defined as

\[ C(s_1, s_2) = \sigma^2 C_\nu(\kappa h) := \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)} (\kappa h)^\nu K_\nu(\kappa h), \quad (1) \]

where \( h = \| s_1 - s_2 \| \), \( \sigma^2 \) is the marginal variance, \( \nu > 0 \) is a smoothness parameter, and \( \kappa > 0 \) is related to the practical correlation range, \( r \), as \( r = \sqrt{8\nu/\kappa} \) where \( C(r)/\sigma^2 \approx 0.1 \). Whittle [36] showed that if \( D := \mathbb{R}^d \), then a centered Gaussian random field \( X(s) \) with a Matérn covariance corresponds to a solution to the SPDE

\[ (\kappa^2 - \Delta)^{\alpha/2}(\tau X(s)) = W(s), \quad s \in D, \quad (2) \]

where \( W \) is Gaussian white noise, \( \alpha = \nu + d/2 \), and

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

is a constant needed to get the desired marginal variance \( \sigma^2 \).

Lindgren et al. [23] 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 \( D \subset \mathbb{R}^n \). The computational benefits of this approach has made it highly popular.

As pointed out before, the Matérn covariance function will likely not be flexible enough to model \( H_s \) due to its stationarity and isotropy. A non-stationary extension can be obtained by replacing the parameters \( \kappa \) and \( \tau \) in (2) by smoothly spatially varying functions [23], and anisotropy can also be acquired by extending the SPDE as in [17, 16]. However, to acquire a non-stationary and anisotropic model with more easily interpretable parameters, we instead combine the SPDE representation with the geometrical approach of Sampson and Guttorp [33].

The idea of Sampson and Guttorp [33] is to consider the observational domain as a subset of some space, \( \mathcal{G} \), being a deformation of another space, \( \mathcal{D} \). A random field that is stationary and isotropic in \( \mathcal{D} \) can then be mapped to \( \mathcal{G} \) by a function \( F : \mathcal{D} \mapsto \mathcal{G} \) which characterizes the non-stationarity and anisotropy of the random field.

Using this idea, \( H_s \) is modeled as \( \exp(X(s) + \mu(s)) \) for a centered, non-stationary, and anisotropic Gaussian random field \( X \). Furthermore, \( X(s) = X(F(\tilde{s})) := \tilde{X}(\tilde{s}) \), where \( F \) is a bijective and differentiable mapping and \( \tilde{X}(\tilde{s}) \) is defined using the SPDE representation of a Gaussian Matérn field. This procedure will lead to a non-stationary model which, as the stationary Matérn model, has an interpretable parametrization and controllable smoothness. See Figure 3 for a sketch of the procedure of mapping a stationary and isotropic random field to \( \mathcal{G} \).

More specifically, we let \( \tilde{\mathcal{L}} = 1 - \tilde{\nabla} \cdot \tilde{\nabla} \) and define \( \tilde{X} \) through

\[ \tilde{\mathcal{L}}^{\alpha/2}\tilde{X}(\tilde{s}) = W(\tilde{s}), \quad \tilde{s} \in \mathcal{D} := \mathbb{R}^d. \quad (3) \]
Here, $\tilde{\nabla}$ is the gradient operator on $\mathcal{D}$. Thus, $\tilde{X}$ is a Gaussian Matérn field with unit dampening constant, $\kappa$. This makes the final model identifiable since the correlation range can be controlled both by $F$ and by $\kappa$. Using the chain rule, $\tilde{\nabla}$ is equivalent to $J[F](\tilde{s})^{-1}\nabla$, where $\nabla$ is the gradient operator on $\mathcal{G}$ and $J[F]$ denotes the Jacobian of $F$. The right hand side of (3) does not have a pointwise meaning, and the equation therefore has to be interpreted in the weak sense. A weak solution $\tilde{X}$ to (3) satisfies

$$\langle \tilde{L}^{\alpha/2} \tilde{X}, \phi \rangle_{\mathcal{D}} = \langle W, \phi \rangle_{\mathcal{D}},$$

for any $\phi \in V := H^1(\mathcal{D})$, where $\langle \cdot, \cdot \rangle_{\mathcal{D}}$ denotes the $L_2$ inner product on $\mathcal{D}$. To obtain the corresponding SPDE in $\mathcal{G}$-space, we perform a change of variables in the weak formulation. To do so, we first use the results of $[34, 23]$ to get

$$\tilde{L}^{\alpha/2} \tilde{X}(\tilde{s}) = \left(1 - |J[F](\tilde{s})^{-1}(s)| \nabla \cdot \left(\frac{J[F]^{-1}(\tilde{s}) J[F]^T(\tilde{s}) s}{|J[F](\tilde{s})^{-1}(s)|} \right) \right)^{\alpha/2} \tilde{X}(\tilde{s})$$

$$= \left(1 - |J[F^{-1}(s)]| \nabla \cdot \left(\frac{J[F^{-1}-1](s) J[F^{-1}-T](s) s}{|J[F^{-1}-1](s)|} \right) \right)^{\alpha/2} \tilde{X}(s).$$

Here we used that $J[F](\tilde{s}) = J[F^{-1}-1](s)$ $[35]$ and $\tilde{X}(\tilde{s}) := X(s)$. To compress the notation, we denote

$$\kappa(s) := \sqrt{|J[F^{-1}](s)|} \quad \text{and} \quad H(s) := \kappa(s)^2 J[F^{-1}-1](s) J[F^{-1}-T](s).$$

Now, performing the change of variables in the weak formulation, using that $d\tilde{s} = |J[F^{-1}](s)| \, ds$, yields

$$\kappa(s)^2 \left(1 - \kappa(s)^{-2} \nabla \cdot H(s) \nabla \right)^{\alpha/2} X(s) = \kappa(s) W(s).$$

The solution to this SPDE has variance $\Gamma(\alpha - 1)/(\Gamma(\alpha)(4\pi)^{d/2})$. In order to model random fields of arbitrary marginal variances we will include a scaling $\tau(s)$ as for the original Matérn SPDE. Combining the $\kappa(s)$ terms, we can then formulate the final model as

$$\left[\kappa(s)^{\frac{2}{d}-2} (\kappa(s)^2 - \nabla \cdot H(s) \nabla) \right]^{\alpha/2} (\tau(s) X(s)) = W(s). \quad (4)$$

Given that the deformation function $F$ is define such that $H(s)$ is Lipschitz continuous and uniformly positive definite, and $\kappa$ is bounded, existence and uniqueness of the solution to this equation follow from the results of $[11]$. 

---

**Figure 3**: Realization of anisotropic Gaussian random field using deformation method.
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. [23]. Here, the solution is approximated by a basis expansion where the coefficients of the expansion is a multivariate Gaussian random vector with a sparse precision matrix (for integer values of $\alpha$). The details of the procedure for our model are presented in Appendix A.

4 Model properties

We see that the deformation induces anisotropy of the solution to (4) through the matrix $H(s)$. By construction, the variance of the solution is

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

In the stationary case, i.e. when the deformation is not varying spatially, the covariance of the solution to (4) is

$$\mathbb{C}(s_1, s_2) = \sigma^2 C_\nu \left( \kappa \sqrt{(s_1 - s_2)^T H^{-1}(s_1 - s_2)} \right), \quad \sigma^2 = \frac{\Gamma(\alpha - 1)}{\Gamma(\alpha)(4\pi)^{d/2} \tau^2},$$

where $C_\nu$ is the Matérn correlation function in (1) with smoothness parameter $\nu = \alpha - d/2$ [21]. Thus, the parameters $\kappa$, $\nu$ and $\sigma$ has the same interpretations as for the standard Matérn covariance, and $H$ controls the anisotropy where eigenvectors of $H$ with larger eigenvalues corresponds to directions with longer correlation ranges.

The model can be seen as an extension of the deformation model of [23] to general $\alpha$ (they only considered $\alpha = 2$). For the special case $\alpha = 2$, [17] considered the related model

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

which was generalized in [16] to

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

In these models $\tilde{H}(s)$ and $\tilde{\kappa}(s)$ where chosen independently of each other. Because of this, the marginal variance of the solution is affected by the choice of $\tilde{H}(s)$ and $\tilde{\kappa}(s)$. In the stationary case, a direct consequence of comparing the covariance function of our model to the covariance function derived in [16] shows that the two models are equivalent (when using $\alpha = 2$ in our model). However, the important difference between the models (except for the general smoothness in our model) is that our parametrization allows for controlling the variance of the process independently of $\kappa$ and $H$. For the stationary case, this is not crucial, but when changing to non-stationary models it greatly simplifies estimation and parameter interpretations.

5 Exceedance probability and extreme loads on ships

By knowing the distribution of significant wave height it is possible to assess some risks involved in shipping. 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 such extreme loads occurring.

The goal of this section is to find a bound for the probability of a ship encountering a value of significant wave height larger than some level $u$ while traveling on a chosen route over the ocean. The model proposed in Section 3 is strictly spatial, therefore any realization of the
model is assumed to persist unchanged over time. However, we start by formulating results for more general spatio-temporal models that are not necessarily Matérn in $\mathcal{D}$ nor purely spatial. Thus, let $X(s,t) = \log H_s(s,t)$ denote a spatio-temporal Gaussian random field which models the logarithm of the significant wave height, with mean value function $\mu(s,t)$ and standard deviation function $\sigma(s,t)$. Consider a ship traveling a route $\gamma(t) = (s(t), t)^T$, where $s \in \mathcal{G}$ and $t \in [0,T]$, and define $X_s(t) = X(s(t), t)$ as the random field evaluated on the route $\gamma$, with mean $\mu_s(t)$ and standard deviation $\sigma_s(t)$. Using Rice’s method, we can derive an upper bound for the probability of the ship encountering extreme loads along the route. Specifically, the following proposition gives a bound for the probability of $H_s$ exceeding a threshold $e^u$ along the route.

**Proposition 5.1.** Assume that $X(s,t)$ is differentiable in mean square sense and that both $\mu(s,t)$ and $\sigma(s,t)$ are differentiable with continuous partial derivatives. Further assume that the route $\gamma$ is differentiable with continuous partial derivatives, then

$$
\Pr \left( \max_{t \in [0,T]} X_{\gamma}(t) > u \right) \leq \int_0^T \sigma_W(t) \left( a(t) - \frac{u - \mu_s(t)}{\sigma_s(t)} \right) \phi \left( \frac{u - \mu_s(t)}{\sigma_s(t)} \right) \, dt + \frac{1}{2} \int_0^T \sigma_W(t) \left( a(t) - \frac{u - \mu_s(t)}{\sigma_s(t)} - 1 \right) \phi \left( \frac{u - \mu_s(t)}{\sigma_s(t)} \right) \, dt + \Phi \left( \frac{\mu_s(t) - u}{\sigma_s(t)} \right).
$$

Here, $\phi$ and $\Phi$ denotes the pdf and cdf of the standard normal distribution, $a(t) := \frac{u - \mu_s(t)}{\sigma_s(t)} + \frac{1}{\sigma_s(t)} \mu_s(t)$, and $\sigma_W(t)$ is the standard deviation of the mean square derivative of $W(t) := X_{\gamma}(0) - \mu_s(t)$.

A proof of Proposition 5.1 can be found in Appendix D. The upper bound is a good approximation of the actual exceedence probability for large values of $u$, which is our main interest. It is worth noticing that the only spatial dependence in this formula is in the marginal variance of the derivative process $W(t)$, i.e. $\sigma_{W}(t)$. We will now show how this variance can be computed when $X(s,t)$ is defined using a spatial deformation.

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 the deformation approach using a function $F : \mathcal{D}_s \times \mathcal{D}_t \to \mathcal{G}_s \times \mathcal{G}_t$ such that the process on $\mathcal{D}_s \times \mathcal{D}_t$ is stationary, and that for a fixed $t \in \mathcal{D}_t$, $X(s,\tilde{t})$ is stationary and isotropic.

Let $W(\tilde{s}, \tilde{t}) = W(F^{-1}(s,t))$ denote by $W_x'$, $W_y'$, and $W_t'$ its partial derivatives. Then $\sigma_{W}(t)$ can be computed using the following proposition.

**Proposition 5.2.** Let $\sigma^2_{ij} = \mathbb{C}(W_i'(\tilde{s}, \tilde{t}), W_j'(\tilde{s}, \tilde{t}))$, for $i,j \in \{x,y,t\}$. With $\Sigma = \{\sigma^2_{ij}\}_{i,j}$ as the corresponding covariance matrix, we have

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

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

A proof of Proposition 5.2 can be found in Appendix D. For the special case of a model which is constant over time, the elements in the third row and column of $\Sigma$ are zero. For the model in Section 3 the Matérn correlation in $\mathcal{D}$ yields $\sigma^2_{xx} = \sigma^2_{yy} = \frac{1}{2(\nu-1)}$ and hence,

$$
\sigma^2_{W}(t) = \frac{1}{2(\nu-1)} \bar{s}(t)^T J[F^{-1}]^T(s(t)) J[F^{-1}](s(t)) \bar{s}(t)
$$

$$
= \frac{\kappa^2(t)}{2(\nu-1)} \bar{s}(t)^T H^{-1}(s(t)) \bar{s}(t),
$$

It is worth noticing that we need $\nu > 1$ or equivalently $\alpha > 1 + d/2$ for the result to hold since the field is otherwise not mean-square differentiable.

7
6 Parameter estimation

In order to use the proposed model with data, the model parameters first have to be estimated. In this section, we describe how to parametrize the model and then discuss the estimation of these parameters.

6.1 Parametrization

A natural idea would be to parametrize the mapping $F$ directly as in [33]. However, the SPDE only depends on the Jacobian matrix of $F^{-1}$ and we could therefore parametrize this Jacobian matrix instead. We instead choose to parametrize the matrix $H(s)$ directly, which has to be positive definite for any $s \in \mathcal{G}$. Since we have data on $\mathbb{R}^2$, we let $\tilde{H}(s)$ be a $2 \times 2$ matrix with $\tilde{H}_{11}(s) = \exp(h_1(s))$, $\tilde{H}_{22}(s) = \exp(h_2(s))$, and $\tilde{H}_{12}(s) = \tilde{H}_{21}(s) = (2S(h_3(s)) - 1)\exp(0.5(h_1(s) + h_2(s)))$. Here $S$ denotes the sigmoid function and $h_1, h_2, h_3$ are some bounded functions left to be defined. Based on this matrix, we define $\kappa(s) = |\tilde{H}(s)|^{-1/2}$ and set $H(s) = \kappa(s)^2\tilde{H}(s)$.

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

$$h_i(s) = \sum_{p=0}^{k} \sum_{n=0}^{k} \beta_{np}^i \cos(n \frac{\pi s_1}{T}) \cos(p \frac{\pi s_2}{S}), \quad i = 1, 2, 3$$

where $s = (s_1, s_2)$ and $T$ and $S$ denotes the width and height of the bounding box of the locations of observations. The advantage with this parameterization is that we do not have any restrictions on the coefficients $\beta_{np}^i$ in order to obtain a valid model.

Finally, a similar basis expansion can be used to parametrize $\log(\tau(s))$ if a spatially varying variance is needed in the model.

6.2 Likelihood-based estimation

Let $y_{jk}$ denote the observed value of $\log(H_s)$ at location $s_j \in \mathcal{G}$ and time point $k$. We assume the previously proposed Gaussian random field model for the observations but also include a nugget effect to account for small-scale effects that cannot be captured by the model. Thus, the assumed model is $Y_{jk} = \mu(s_j) + X_{jk}(s_j) + \epsilon_{jk}$ where $\epsilon_{jk} \sim N(0, \sigma^2_\epsilon)$ are independent, $\mu(s_j)$ is a possible regression for the mean value of the latent field, and $X_{jk}(s)$ 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

$$Y_k = \mu + AU_k + \epsilon_k, \quad U_k \sim N(0, Q_U^{-1}), \quad \epsilon_k \sim N(0, \sigma^2_\epsilon), \quad k = 1, \ldots, K.$$ 

Here $K$ is the total number of replicates in the data set, $Y_k = (Y_{1k}, \ldots, y_{jk})^T$ is a vector with all observations for time $k$, $\mu$ is a vector with entries $\mu_j = \mu(s_j)$, $A$ is the observation matrix with elements $A_{ji} = \phi_i(s_j)$ where $\phi_i$ is the $i$:th FEM basis function, $U$ is the multivariate Gaussian distribution of the coefficients of basis functions in the FEM approximation. The parameters of the model, which needs to be estimated, are thus the parameters $\{\beta^1_{np}, \beta^2_{np}, \beta^3_{np}\}_{n,p=0}$ and possibly the corresponding parameters for a basis expansion of $\log(\tau(s))$ as well as the regression parameters for the mean and the nugget variance $\sigma_\epsilon$. We parametrize the latter as $\log(\sigma_\epsilon) = \tilde{\sigma}_\epsilon$ and estimate $\tilde{\sigma}_\epsilon$, which has no constraints.

To estimate the proposed model from data, we use a maximum likelihood approach (ML). Since both the nugget effect and the random field is Gaussian, the distribution of $Y_k$ is Gaussian with mean $\mu$ and $\Sigma_Y = AQ_U^{-1}A^T + \sigma^2_\epsilon I$. However, evaluating the log-likelihood of this Gaussian
variable directly would require computing $Q^{-1}$ which is dense. To be able to take advantage of the sparsity of $Q_U$, we use the fact that the density of $Y_k$ can be rewritten as

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

where $u^*$ is an arbitrary value in the sample space of $U$. Let $\Theta$ denote the set of all parameters. Choosing $u^* = 0$ yields the log-likelihood

$$l(\Theta; y) = \frac{K}{2} \left( \log |Q_U| - J \log \sigma^2 + J \log |Q_{U|Y}| + J \log(2\pi) \right) + \frac{1}{2\sigma^2} \sum_{k=1}^{K} (y_k - \mu)^T \left[ \frac{1}{\sigma^2} A Q_{U|Y} A^T - I \right] (y_k - \mu).$$

To evaluate the log-likelihood, computing the log determinants and solving linear systems for 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.

To find the parameter estimates, we perform numerical optimization of the log-likelihood. This can be time consuming since the proposed model can have quite a large number of parameters. To reduce the computation time, and the risk of finding suboptimal 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 to observed locations in space and computing 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.

7 Results

In this section, we present the results of using the proposed model for the data described in Section 2.

7.1 Model fit

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 for each of the spatial locations separately. 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 data set contains 585 replicates.

The proposed model of Section 3 is fitted to the training data using the parametrization of Section 6. 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 has a constant zero mean as well as a unit marginal variance. Hence, except for the coefficients in $h_1(s), h_2(s), h_3(s)$, the only other parameters to be estimated are $\sigma^2$ and $\nu$. The model thus required estimating 77 parameters in total.

A histogram of the local estimates of the smoothness parameter $\nu$ using the method of Appendix B can be seen in Figure 4. All estimated values are between 1 and 3 with the median close to 2, corresponding to $\alpha = 3$. Since the SPDE only is Markov if $\alpha$ is an integer, 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 Figure 5. The figure shows both the empirical correlation function
Figure 4: Histogram of estimated smoothness parameters, $\nu$, from the different local estimates.

Figure 5: 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).

from data and the computed correlation function from the estimated model. Although some differences are visible between data and the estimated model in Figure 5, the overall structures are well captured. Another way of visualizing the non-stationarity of the estimated model is to show the estimated deformation. Figure 6 show how a rectangular grid of the north Atlantic (the $G$-space) has been mapped to $D$ 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 an effect makes sense due to the various effects occurring on the interface between land and ocean, causing spatially (and temporally) more chaotic behavior and hence less large scale dependence. The data is placed on a grid in the longitude/latitude coordinate system. This will make the distance in the longitudinal direction between point 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 $D$ 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 reference frame.

It should be noted that the grid has folded in a few positions close to the coast, suggesting that the estimated $D$ is not well defined. However, the figure still gives a proper visual interpretation of the non-stationarity and anisotropy of the estimated model. See Appendix C for further details on this and on how the $D$-space is estimated from the model parameters.
7.2 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 for the models, we can perform a likelihood ratio test between the two models. 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 \( \Theta_0 \) is the subspace of parameter values which yields a stationary model and \( \Theta \) is the full parameter space of the non-stationary model. 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 \(-2\lambda\) converges to a \( \chi^2 \) distribution \[13, \text{Theorem 10.3.3}\] it is possible to compute \( c \). Here, the degrees of freedom \( df \) equals 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 \( \alpha \) 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 \( \lambda \) 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 \[3\].

7.3 Fatigue damage accumulated by a vessel

Mao et al. \[24\] 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 significant wave height and is based on the narrow-band approximation \[32\]. The approximation was shown to be conservative (i.e. an upper bound) in comparison with actual fatigue damage. According to this formula, the expected instantaneous fatigue damage to a ship on route as a function of
the significant wave height, zero crossing period, ship velocity, and wave velocity is

\[ d(t) \approx \frac{0.47 C^\beta H_s^\beta(s(t))}{\gamma} \left( \frac{1}{T_z(s(t))} - \frac{2\pi V(t) \cos \theta(t)}{gT_z^2(s(t))} \right), \tag{5} \]

where \( g \) is the gravitational constant (\( \approx 9.81 \)), \( T_z \) is the mean zero crossing period of the waves, \( V \) is the speed of the ship, and \( \theta \) is the angle between the heading of the ship and the direction of the traveling waves. Further, \( \gamma \) and \( \beta \) are constants dependent on the material of the ship and \( C \) is a constant depending on the ship design, see [24] for details. Note that the formula concerns a route in space-time parametrized by time, \( t \), just as the route in Section 5.

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 [24, 28] and approximate it as \( T_z \approx 3.75 \sqrt{H_s} \), making the formula in (5) dependent only on the values of \( H_s \) for which we have a spatial model. The total damage up until time \( T \) is computed by numerically integrating \( d \),

\[ D(T) = \int_0^T d(t) dt \approx \sum_i d_i \Delta t, \]

where \( d_i \) denotes a sequence of evaluations of \( d(t) \), \( \Delta t \) apart.

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 Figure 1, which is a scenario similar to the one in Podgórski and Rychlik [28]. 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 hours or equivalently 6.23 days. The heading of the ship, in one of the 100 locations on 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 [28, 24], i.e. \( C = 20, \beta = 3, \) and \( \gamma = 10^{12.73} \). In order to use [5] we also need the propagating waves angle in comparison with the ship. In Podgórski and Rychlik [28] the mean wave propagation in space was estimated for each month of a year. We use their 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 temporal resolution is low (one measurement per day) and we want to compare the empirical distribution from data with equivalent empirical distributions from simulations of the fitted model, which is not temporal. Therefore 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 was computed for the 100 point locations on the route.

We use the validation data to compute the empirical fatigue damage distribution by computing the total fatigue damage on the ship for each of the 585 replicates in the data set. 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 Figure 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, 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 show 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.

### 7.4 Extreme loads

Apart from estimating the fatigue damage, it is often also of interest to estimate the probability of encountering large waves since that corresponds to extreme loads on the ship. We again consider a ship traveling the route of Figure 1.

The exceedance probability of a given threshold for the significant wave height is estimated empirically from data by counting the number of replicates for which the threshold is exceeded on the considered route. We compare this empirical exceedance probability with the same calculations for 200 simulations from the fitted model as well as for the upper bound from Section 5. 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 5.1 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 neighbouring pixels on route.

Figure 8 show the estimated exceedance probabilities for a range of thresholds between 2 metres and 12 metres, for the fitted stationary and non-stationary models. One can see that that the non-stationary model has an envelope that encapsulates the data completely. For the stationary model the exceedance probability is below the envelope in the lower range of 2 to 4 meters, 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
We have developed a non-stationary and 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 [33]. An advantage with this model compared to the similar non-stationary models of [17, 16] is that the marginal variance parameter is independent of the choice of the parameters explaining anisotropy. This 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 [16] 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, such as the upper bound on exceedance probabilities which we derived using Rice’s method.

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 modelling.

Significant wave height data from the ERA-interim reanalysis data set was used to fit and validate the model. Results showed that both the exceedance probability of significant wave height on a ships 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 8 Discussion

above 5 meters in both the stationary and non-stationary model. That the bound performs best for large values is expected as previously explained.
future work. One such extension is to model multivariate random fields similar to [19] [10]. Such models would allow for joint modelling 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 next step is to derive a suitable spatio-temporal extension of the model.

9 Acknowledgements

This work has been supported in part by the Swedish Research Council under grant No. 2016-04187. We would like to thank the European Centre for Medium-range Weather Forecast for the development of the ERA-Interim data set and for making it publicly available. The data used was the ERA-Interim reanalysis dataset, Copernicus Climate Change Service (C3S) (accessed September 2018), available from https://www.ecmwf.int/en/forecasts/datasets/archive-datasets/reanalysis-datasets/era-interim.

A Finite element approximation

To obtain a finite dimensional approximation of the model in Section 3, we first have to redefine the model on a bounded domain. Thus, we assume that (4) is defined for $s \in G$, where $G$ is a bounded, convex, polygonal domain, and we equip the operator with Dirichlet boundary conditions.

The solution of the SPDE is then approximated by a basis expansion $X(s) = \sum_{i=1}^{N} U_i \phi_i$, where $\{\phi_i\}$ are piecewise linear basis functions induced by a mesh of the spatial domain, see Figure 9. The distribution of the coefficients $U = (U_1, \ldots, U_N)^T$ are then computed by replacing the infinite-dimensional test space of the weak formulation with a finite-dimensional subspace spanned by the basis functions $\{\phi_i\}$. For $\alpha = 2$, the distribution of $U$ is computed by solving the following system of linear equations

$$\sum_{j} \langle L \tau \phi_j, \phi_i \rangle U_j, i = 1, \ldots N \right\} \overset{d}{=} \{ \langle W, \phi_i \rangle, i = 1, \ldots N \} \quad (6)$$

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

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

The left-hand side of (6) 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 $C_{ii} = \langle 1, \phi_i \rangle$ [23]. Solving $KU = W$, yields that $U \sim \mathcal{N}(0, Q^{-1})$, where the sparse precision matrix $Q = K^T C^{-1} K$.

For the evaluation of the inner products in Equation (6) the spatially varying parameters are in Section 7 approximated as constant over each triangle. This means that the inner products needed to compute $K$ only corresponds to computing integrals of at most two linear function on each triangle.

Equation (6) provides an approximation of the solution when $\alpha = 2$. For models where $\alpha$ is an even integer Lindgren et al. [23] showed how a solution can be expressed by recursively applying $L$ several times. Also for $\alpha$ being an odd integer, a similar construct could be achieved.
by considering the least squares solution to $L^{1/2}X = W$ and then considering recursive solutions to this. This is possible also for our anisotropic model since $L^{\alpha/2}$ is self-adjoint. For any integer valued $\alpha$, this scheme yields the precision matrices $Q^{(1)} = K$ for $\alpha = 1$, $Q^{(2)} = KC^{-1}K$ for $\alpha = 2$, and $Q^{(\alpha)} = KC^{-1}Q_{\alpha-2}C^{-1}K$ for $\alpha = 3, 4, \ldots$. These matrices become less sparse as $\alpha$ is increased. However, the computational complexity of using the approximation for inference will be $O(N^{3/2})$ but with a larger constant due to less sparsity for larger $\alpha$. This should be compared to the standard cubic increase in computational cost for general covariance based methods in spatial statistics [31].

**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 so that its behaviour 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 $\kappa$. Lindgren et al. [23] 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 $r = \sqrt{8\nu/\kappa}$ 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 width (the largest 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 width be at most a fifth of the local correlation range at the location of the triangle. However, when $\kappa$ 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. [6] 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 $J[F^{-1}](s) = \sqrt{8\nu/\kappa}I$ for all triangles that are exclusively in the exterior of $G$. Here, $r_{\text{min}}$ denotes the minimum local correlation range in the interior. In Figure 10, the blue triangles are part of the spatial domain of interest and the gray triangles are part of the extension.
Figure 10: The mesh of the north Atlantic used in Section 7. Triangles of the extension are shown in gray and triangles from the interior in blue.

Figure 11: Evenly spaced local neighborhoods are chosen for the local estimates. Data from the nine locations in each neighborhood is used for the local estimate.

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 these regions. We use small $3 \times 3$ neighborhoods of pixels for each local estimate, see Figure 11.

For the local estimate, we first estimate the marginal mean and variance for each location using the regular sample mean and sample variance. We then compute the parameters of the constant $H$ matrix of the stationary covariance from Section 4 by numerical optimization of the log-likelihood.

All the local estimates are merged by computing the least squares solution of the values $\{\beta_{np}\}$ based on the local estimates. The merged values are used to acquire good initial values for the numerical maximization of the likelihood function as described in Section 6.

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 $J[F^{-1}]$. During the global optimization $J[F^{-1}]$ will vary, and one 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 $\alpha$ based on the local estimates prior to estimating the other parameters globally.

C Existence and estimation of $D$-space

As mentioned in Section 6, we do not parametrize the model using the deformation function $F^{-1}$ directly. This is not needed since the model only depends on $F^{-1}$ through the $H$ matrix. However, knowing $F^{-1}$ or $D$ can in certain applications be of interest in itself since it might carry with it 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 $D$-space from the estimated $H(s)$, three problems have to be taken care of.

Firstly, $J[F^{-1}]$ is not uniquely defined by $H$. A simple example of this is the stationary case where it is not possible to identify maps, $F^{-1} : G \rightarrow D$, that only differ by one being the mirror of the other with respect to some plane. 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 $H$ since they share eigenvectors and the eigenvalues of $J[F^{-1}]$ are just the inverse square root of those for $H$.

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 components 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 $H$.

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 just consider $F^{-1}$ as a mapping $F^{-1} : G \rightarrow D$, not necessarily injective. The inverse function theorem [35] states that there exists a differentiable injective function in an open set around a point, $s$, if $|J[F^{-1}](s)| \neq 0$. 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 $D$ folds, i.e. several points in $G$ maps to the same point in $D$.

Folding was also a problem using the thin-plate spline parametrization of $F^{-1}$ in Sampson and Guttorp [33]. They reported issues with folding and recommended ameliorating this problem by forcing $F^{-1}$ to be smooth enough in their parametrization. For our model [41], 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. We only have a problem with folding when the interest lies in $D$ itself. It is for us an open question how to find a practically useful parametrization of $H$ or $F^{-1}$ in a way that ensures no folding.

However, even if the estimated model folds, it should not occur in many places since $H$ is parametrized to vary smoothly. It is therefore possible to consider dividing $G$ in to subregions, $\{G_k\}_k$. Each $G_k$ are in turn mapped to a space $D_k$ 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 $\{D_k\}_k$. This was done in Figure 6 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 $G$ and $D$ can be computed as

$$F^{-1}(s) = F^{-1}(s_0) + \begin{bmatrix} \int_0^1 a(r(t))dr_x(t) \\ \int_0^1 c(r(t))dr_x(t) \end{bmatrix} + \begin{bmatrix} \int_0^1 b(r(t))dr_y(t) \\ \int_0^1 c(r(t))dr_y(t) \end{bmatrix}.$$
where

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

and the functions \(a, b,\) and \(c\) are defined through the spatially varying parameters \(h_i\) of the \(H\) matrix. By identifying one node in the FEM mesh on \(G\) as mapping to a distinct point in \(D\), the mapping of all nodes to \(D\) can be acquired by successively mapping each neighboring node in a recursive manner. Points inside a triangle can then easily be approximately mapped to \(D\) 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 [18].

An example of the result is shown in Figure 12, where a realisation of a non-stationary and anisotropic Gaussian random field 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 6. The \(D\) region was then estimated using the algorithm described above, and is shown in the right part of the figure.

D  Proofs

Proof of Proposition 5.1. Let \(N_T(v)\) be the number of times \(X_\gamma(t)\) crosses a threshold, \(u\) on the route. Then by Rice’s method, we have

\[
P \left[ \max_{t \in [0,T]} X_\gamma(t) > u \right] = P [X_\gamma(0) > u] + P [(N_T(u) > 0) \cap (X_\gamma(0) \leq u)]
\leq P (X_\gamma(0) > u) + \frac{1}{2} E [N_T(u)].
\]

By Rice’s formula [5, 29, 30] we have

\[
E [N_T(u)] = \int_0^T E \left[ X_\gamma(t) \big| X_\gamma(t) = u \right] f_{X_\gamma(t)}(u) dt,
\]

where \(\dot{X}_\gamma\) denotes the mean square derivative of \(X_\gamma(t)\). Now, if \(X_\gamma\) would have constant variance, the derivative of the process at \(t\) would not be correlated with the value of the process at \(t\).
section 5.6]. To make use of this property, we introduce \( W(t) = \frac{X_\gamma(t) - \mu_\gamma(t)}{\sigma_\gamma(t)} \) and get
\[
\mathbb{E} \left[ |\dot{X}_\gamma(t)| \bigg| X_\gamma(t) = u \right] = \mathbb{E} \left[ |W(t)\sigma_\gamma(t) + W(t)\dot{\sigma}_\gamma(t) + \dot{\mu}_\gamma(t)| \bigg| X_\gamma(t) = u \right].
\]
Under the condition \( W(t) = \frac{u - \mu(t)}{\sigma(t)} \) we can write the expression inside the absolute value as \( \sigma_\gamma(t)(\dot{W}(t) + a(t)) \), where \( a(t) = \frac{u - \mu_\gamma(t)}{\sigma_\gamma(t)}\dot{\sigma}_\gamma(t) + \frac{\dot{\mu}(t)}{\sigma_\gamma(t)} \). Using that \( \dot{W}(t) \) and \( W(t) \) are uncorrelated for a fixed \( t \) and that \( \dot{W} \) is centered, we have that
\[
\mathbb{E} \left[ |\dot{X}_\gamma(t)| \bigg| X_\gamma(t) = u \right] = 2\sigma(t)\sigma_W(t)\phi \left( \frac{a(t)}{\sigma_W(t)} \right) + a(t) \left[ 1 - 2\Phi \left( \frac{-a(t)}{\sigma_W(t)} \right) \right],
\]
which can be seen by computing the expected value of a folded normal distribution. Now, using Equation (7) and plugging the result in to the first inequality finishes the proof. \( \square \)

**Proof of Proposition 5.2** By the chain rule we have
\[
\dot{W}(t) = \dot{s}(t) \cdot \nabla W(s(t)) = \dot{s}(t)^T J[F^{-1}]^T(s(t)) \nabla \dot{W}(\dot{s}(t)).
\]
Due to the stationarity of \( \dot{W}(\dot{s}(t)) \), the marginal distribution with respect to space-time of the partial derivatives of \( \dot{W}(\dot{s}(t)) \) is a 3-dimensional Gaussian distribution independent of \( s \). Let \( \Sigma \) denote the covariance matrix of this 3 dimensional Gaussian distribution, then
\[
\Sigma := \mathbb{E} \left[ \left( \nabla \dot{W}(\dot{s}) \right) \left( \nabla \dot{W}(\dot{s}) \right)^T \right] = \begin{bmatrix}
\sigma_{xx}^2 & 0 & \sigma_{xt}^2 \\
0 & \sigma_{yy}^2 & \sigma_{yt}^2 \\
\sigma_{xt}^2 & \sigma_{yt}^2 & \sigma_{tt}^2
\end{bmatrix}.
\]
The zeros in \( \Sigma \) is due to the isotropy in the spatial coordinates, making partial derivatives with respect to orthogonal vectors independent \( [1] \). This isotropy also yields that \( \sigma_{xx} = \sigma_{yy} \). The proof is concluded by the fact that
\[
\sigma_W^2(t) = \mathbb{E} \left[ \dot{W}(t)\dot{W}(t)^T \right] = \dot{s}(t)^T J[F^{-1}]^T(t)\Sigma J[F^{-1}](t)\dot{s}(t).
\]
\( \square \)

**References**

[1] R.J. Adler and J.E. Taylor. *Random Fields and Geometry*. Springer, 2007. ISBN 978-0-387-48112-8.

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

[3] H. Akaike. Information theory and an extension of the maximum likelihood principle. *Second International Symposium of Information Theory*, pages 267–281, 1973.

[4] G.A. Athanassoulis, E.K. Skarsoulis, and K.A. Belibassakis. Bivariate distributions with given marginals: with an application to wave climate description. *Applied Ocean Research*, 16:1–17, 1995.

[5] J-M. Azaïs and M. Wschebor. *Level Sets and Extrema of Random Processes and Fields*. Wiley, 1 edition, 2009. ISBN 0470409393.
[6] H. Bakka, J. Vanhatalo, J.B. Illian, D. Simpson, and H. Rue. Non-stationary Gaussian models with physical barriers. *Spatial Statistics*, 29:268–288, 2019.

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

[8] A. Baxevani, S. Caires, and I. Rychlik. Spatio-temporal statistical modelling of significant wave height. *Environmetrics*, 20:14–31, 2009.

[9] D. Bolin and K Kirchner. The rational SPDE approach for Gaussian random fields with general smoothness. *arXiv preprint arXiv:1711.04333v2*, 2018.

[10] D. Bolin and J Wallin. Multivariate Type-G Matérn fields. *arXiv preprint arXiv:1606.08298v2*, 2018.

[11] David Bolin, Kristin Kirchner, and Mihály Kovács. Numerical solution of fractional elliptic stochastic PDEs with spatial white noise. *IMA Journal of Numerical Analysis*, 12 2018. doi: 10.1093/imanum/dry091. URL [https://dx.doi.org/10.1093/imanum/dry091](https://dx.doi.org/10.1093/imanum/dry091)

[12] D.J.T. Carter. Estimation of wave spectra from wave height and period, 1982.

[13] G. Casella and R.L. Berger. *Statistical Inference*, volume Second edition. Thomson Learning, 2002. ISBN 0-534-24312-6.

[14] G. De Masi, R. Bruschi, and M. Drago. Synthetic metocean time series generation for offshore operability and design based on multivariate Markov model. *MTS/IEEE OCEANS 2015 - Genova: Discovering Sustainable Ocean Energy for a New World*, 17 September, 2015.

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

[16] G.A. Fuglstad, F Lindgren, D. Simpson, and H. Rue. Exploring a new class of non-stationary spatial gaussian random fields with varying local anisotropy. *Statistica Sinica*, 25(1):115–133, 2015.

[17] G.A. Fuglstad, D. Simpson, and F Lindgren. Does non-stationary spatial data always require non-stationary random fields? *Spatial Statistics*, 14:505–531, 2015.

[18] M.T Heath. *Scientific Computing*, volume Second edition. McGraw-Hill Professional, 2002. ISBN 0-07-239910-4.

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

[20] N.H. Jasper. Statistical distribution patterns of ocean waves and of wave-induced ship stresses and motions with engineering applications. *Trans. Soc. Naval Architects and Marine Engrs*, 64:375–432, 1956.
[21] U. Khristenko, L. Scarabosio, P. Swierczynski, E. Ullmann, and B. Wohlmuth. Analysis of boundary effects on PDE-based sampling of Whittle-Matérn random fields. *arXiv preprint arXiv:1809.07570v1*, 2018.

[22] E.V. Lewis. *Principles of Naval Architecture*, volume 1. Society of Naval architects and marine engineers, 1988.

[23] F. Lindgren, H. Rue, and J. Lindström. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society*, 73(4):423–498, 2011.

[24] W. Mao, J.W. Ringsberg, I. Rychlik, and G. Storhaug. Development of a Fatigue Model Useful in Ship Routing Design. *Journal of Ship Research*, 54(4):281–293, 2010.

[25] V. Monbet, P. Ailliot, , and M. Prevosto. Survey of stochastic models for wind and sea state time series. *Probabilistic engineering mechanics*, 22:113–126, 2007.

[26] W.H. Munk. Proposed uniform procedure for observing waves and interpreting instrument records. *La Jolla, California: Wave Project at the Scripps Institute of Oceanography*, 22, 1944.

[27] M.K. Ochi. On the long-term statistics for ocean and coastal waves. *Proc. 16th Conf. Coastal Engineering*, pages 59–75, 1978.

[28] K. Podgórski and I. Rychlik. A model of significant wave height for reliability assessment of a ship. *Journal of marine systems*, 130:109–123, 2014.

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

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

[31] H. Rue and L. Held. *Gaussian Markov random fields*, volume 104. Chapman and Hall, 2005. ISBN 0203492021.

[32] I. Rychlik. On the ‘narrow-band’ approximation for expected fatigue damage. *Probabilistic Engineering Mechanics*, 8(1):1–4, 1993.

[33] P.D. Sampson and P. Guttorp. Nonparametric Estimation of Nonstationary Spatial Covariance Structure. *Journal of the American Statistical Association*, 87(417):108–119, 1992.

[34] T. Smith. Change of variables in Laplace’s and other second-order differential equations. *Proc. Phys. Soc.*, 46(3):344–349, 1934.

[35] M. Spivak. *Calculus on Manifolds*. Addison-Wesley, 1965. ISBN 0-8053-9021-9.

[36] P. Whittle. On stationary Processes in the Plane. *Biometrika*, 41(3/4):434–449, 1954.