



Bayesian hierarchical modelling of North Atlantic windiness

E. Vanem and O. N. Breivik

Department of Mathematics, University of Oslo, P.O. Box 1053 Blindern, 0316 Oslo, Norway

Correspondence to: E. Vanem (erikvan@math.uio.no)

Received: 21 November 2012 – Published in Nat. Hazards Earth Syst. Sci. Discuss.: –

Revised: 21 January 2013 – Accepted: 23 January 2013 – Published: 4 March 2013

Abstract. Extreme weather conditions represent serious natural hazards to ship operations and may be the direct cause or contributing factor to maritime accidents. Such severe environmental conditions can be taken into account in ship design and operational windows can be defined that limits hazardous operations to less extreme conditions. Nevertheless, possible changes in the statistics of extreme weather conditions, possibly due to anthropogenic climate change, represent an additional hazard to ship operations that is less straightforward to account for in a consistent way. Obviously, there are large uncertainties as to how future climate change will affect the extreme weather conditions at sea and there is a need for stochastic models that can describe the variability in both space and time at various scales of the environmental conditions. Previously, Bayesian hierarchical space-time models have been developed to describe the variability and complex dependence structures of significant wave height in space and time. These models were found to perform reasonably well and provided some interesting results, in particular, pertaining to long-term trends in the wave climate. In this paper, a similar framework is applied to oceanic windiness and the spatial and temporal variability of the 10-m wind speed over an area in the North Atlantic ocean is investigated. When the results from the model for North Atlantic windiness is compared to the results for significant wave height over the same area, it is interesting to observe that whereas an increasing trend in significant wave height was identified, no statistically significant long-term trend was estimated in windiness. This may indicate that the increase in significant wave height is not due to an increase in locally generated wind waves, but rather to increased swell. This observation is also consistent with studies that have suggested a poleward shift of the main storm tracks.

1 Introduction

Ship operations are vulnerable to extreme weather conditions and extreme seas may be a direct cause or contributing factor in ship accidents. Hence, such conditions need to be accounted for in the design and operation of ships. A correct understanding of the statistics of such extreme conditions is, therefore, important to ship safety and there is a need for adequate statistical descriptions of different environmental parameters such as significant wave height, wave period and wind speed. However, it has been increasingly evident in recent years that the climate may change due to increased anthropogenic forcing of the globe, i.e., that the globe is warming (IPCC, 2007, 2012). As a result of this, the dynamics of the ocean-atmosphere system may be changing and projections of future climate change indicate that frequencies and intensities of some extreme weather events are likely to increase.

In order to investigate the impact of climate change on the significant wave height, one of the most important sea state parameters for ship operations, the framework of Bayesian hierarchical space-time models (Wikle et al., 1998, 2001) was applied to significant wave height data in Vanem et al. (2012b). Extensions of this model including a logarithmic transformation of the data (Vanem et al., 2012d) and a regression component on atmospheric levels of CO₂ (Vanem et al., 2012a) have also been investigated. Overall, these models seemed to perform well in describing the spatial and temporal variability in the significant wave height data and for the area in the North Atlantic ocean an increasing trend was extracted. The models were also applied to the monthly maximum significant wave height and identified an even stronger trend in the monthly extremes (Vanem et al., 2012c,a). Furthermore, in Vanem and Bitner-Gregersen (2012) it was demonstrated how projections of future wave

climate could be included in load calculations of ship structures and it was indicated that the effect may not be negligible. Thus, possible long-term trends in extreme seas due to climate change represent additional hazards to ship operations.

Studies of trends in the ocean wave climate and future projections under climate change scenarios have also been reported in e.g., Caires et al. (2006); Wang and Swail (2001, 2002); Wang et al. (2004); Wang and Swail (2006); Kushnir et al. (1997); Grabemann and Weisse (2008); Debernard and Røed (2008). See also a more comprehensive review in Vanem (2011). The variability of estimated trends and future projections is obviously great, but overall, the results from the Bayesian hierarchical space-time models were found to agree reasonably well with other studies. Furthermore, a number of time-series trend analysis techniques on the significant wave height data were found to yield similar long-term trends (Vanem and Walker, 2013).

Although the physical mechanisms for wave generation are complex and not completely understood, it is generally accepted that the main physical mechanism for wave generation is transfer of energy from the atmosphere to the ocean due to wind friction on the sea surface (see e.g., Talley et al., 2011). Waves that are generated by local wind fields are referred to as wind-sea, whereas swell is used to refer to waves that remain after the wind has died out and that can travel considerable distances. Hence, the origin of swell may be very remote from where the waves are measured. In the open ocean, waves of many different directions and frequencies are present and wave spectra often display two distinctive frequency modes associated with wind-sea and swell (bimodal waves).

In the significant wave height data used in the Bayesian hierarchical space-time models, it was not distinguished between the wind-sea and the swell contributions and the models were not able to distinguish the origin of the waves contributing to the significant wave height. Therefore, it could not determine whether the estimated increase in significant wave height was due to increased wind-sea or increased swell or both. In order to investigate this, this paper applies a similar stochastic model on the wind speed over the same area in the North Atlantic ocean. It will then be argued that if a similar increasing trend is estimated for wind speed as for significant wave height, the increase in significant wave height can be attributed, at least in part, to an increase in wind-sea. Conversely, if there are no detectable trends or even decreasing trends in the windiness over the same area, the estimated increase should be ascribed to increasing swell. Furthermore, it is believed that investigating the spatial and temporal variability of North Atlantic windiness is of interest in itself and that it will be interesting to explore how the Bayesian hierarchical modelling framework performs for wind speed.

The model presented in this paper is a stochastic model, a probabilistic counterpart to physical models that are more deterministic, with treatment of uncertainties as an integral

part of the model. In a historic perspective, it is noted that for a long time following the scientific revolution in the 16th century, the predominant world-view was deterministic. It was believed that if exact knowledge of initial conditions and causal laws governing a system were available, the exact state of the system could be determined at any later point in time. In such a mechanistic world, randomness would not exist and failure to precisely predict future events would be entirely due to incomplete knowledge of initial states and universal laws. However, in the late 19th and early 20th century, new scientific discoveries cast serious doubts on a strictly deterministic world-view. Chaos theory explained how even an infinitesimally small perturbation of initial conditions of a purely deterministic nonlinear system can lead to large changes in the development of the system (the butterfly effect). Furthermore, the development of quantum mechanics and the formulation of the Heisenberg uncertainty principle demonstrated that reality, at least at atomic scales, does not seem to be absolutely deterministic, suggesting a more probabilistic understanding of the world.

Regardless of whether the world is fundamentally probabilistic or if it is deterministic, but with uncertain knowledge of the underlying physical laws, physical environmental processes inevitably display some seemingly causal relationships along with a considerable degree of randomness. Hence, it is argued that it would make sense to describe such phenomena probabilistically, i.e., using probability theory and statistics to model physical processes as stochastic processes where there are several possible ways for a system to evolve. The model presented in this paper is such a stochastic model for describing the statistics of the North Atlantic windiness.

The remaining part of this paper is structured as follows: first, the wind data and the area selected for investigation will be described and a crude initial data analysis will be performed. Then, in Sect. 3 the Bayesian hierarchical space-time model will be outlined, including a description of the main model and some model alternatives. In Sect. 4 the implementation of the model by way of Markov chain Monte Carlo simulations is presented and Sect. 5 presents the results for the various model components. Finally, Sects. 6 and 7 provide some discussion and a conclusion with some final remarks. The posterior distributions of the different model parameters are presented in a table in Appendix A.

2 Data description and area of study

2.1 Area description

For the purpose of this study, an area in the mid-latitudes of the North Atlantic ocean was selected for investigation. One important reason for selecting an area in the North Atlantic ocean is that North Atlantic conditions are normally used as a basis for ship design, so this area is particularly interesting

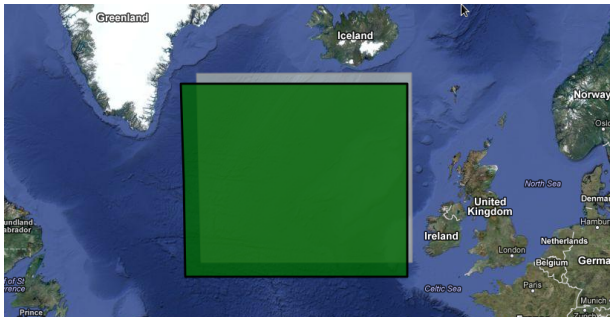


Fig. 1. The areas in the North Atlantic ocean for analysing windiness (dark green) and significant wave height (light grey).

with regards to ship operations. Ideally, the selected area should be identical to the area used for analysing significant wave height. However, for the significant wave height, a corrected dataset with resolution $1.5^\circ \times 1.5^\circ$ was used whereas the dataset for wind speed that is downloadable from the European Centre for Medium-Range Weather Forecasts (ECMWF) website has a spatial resolution of $2.5^\circ \times 2.5^\circ$ degrees. Thus, the areas are not entirely identical, but they are largely overlapping and this is assumed to be satisfactory. The area included in this study ranges from 50° to 62.5° N and 322.5° to 347.5° E corresponding to a grid of $6 \times 11 = 66$ data points. Due to the reduced spatial resolution, the distance between the grid points are larger than for the dataset used in the significant wave height modelling even though the areas are essentially overlapping. It is assumed that the different spatial resolution of the wind data compared to the wave data does not influence the overall results from the models and particularly that any long-term trends would not be very sensitive to the difference in spatial resolution. The selected area is illustrated in Fig. 1 where also the area for the significant wave height analysis is indicated.

2.2 Wind data

This study exploits the ERA-40 reanalysis data for wind speed (Uppala et al., 2005), which covers most of the entire globe and a period of 45 yr at six-hourly resolution (from September 1957 until August 2002). It is important to note that the ERA-40 wind data is from the same re-analysis study as the ERA-40 significant wave height data that was used in the previous models for significant wave height (Vanem et al., 2012b,d). These significant wave height data were selected based on its spatial and temporal resolutions. Hence, the datasets should be consistent and allow for comparison of results for waves and winds. The data are freely available for research purposes and may be downloaded from the ECMWF website¹.

¹<http://data-portal.ecmwf.int/>

The ERA-40 data contain the two parameters U and V for the 10 m wind speed, i.e., the wind speed at height 10 m above the sea surface. U refers to the wind velocity component in the east-west direction, whereas V refers to the north-south component of the wind velocity, both in terms of metres per second (m s^{-1}). Jointly, U and V , therefore, describe both the magnitude of the wind speed and the wind direction. This study, however, is restricted to analysis of the wind speed, and the wind direction will not be considered. Hence, the parameter of interest in this study, for which the Bayesian hierarchical space-time model has been applied, is the absolute value of the wind speed, W :

$$W = \sqrt{U^2 + V^2}. \quad (1)$$

The temporal resolution of the data is six hours. However, for the purpose of this study, the monthly maximum wind speed in each spatial location is analysed, in line with the analyses reported in Vanem et al. (2012c,a). This corresponds to time series of 540 monthly maxima, corresponding to the 45 yr of data, at each spatial location. It is noted that in the raw data there are scale factors to be multiplied and offsets to be added that are different for different months in the ERA-40 data for wind speed, and care should be taken to always include the correct factors and offsets. The actual values of these correction factors and offsets for each month are included in the actual data.

The ERA-40 10 m wind speed data have been compared to satellite and buoy measurements in Wallcraft et al. (2009). Comparison of monthly means indicates good agreement between the ERA-40 wind speeds and the satellite measurements, with high correlations and low root-mean-square values for wind speed difference, although ERA-40 tends to underestimate the monthly mean wind speed. Comparison on shorter time scales, i.e., comparing with daily satellite winds, reveals larger differences, but it is assumed that the ERA-40 wind speed data for monthly maximum values are reasonable. Also the KNMI/ERA-40 wave atlas² states that the 10-m wind speeds compare quite well with observations. It is noted that the results from the modelling, presented in this paper, are conditional on the data and no attempt has been made to correct possible biases. Therefore, it is reassuring that the monthly mean 10 m wind speeds from ERA-40 agree well with satellite data.

Before applying the Bayesian hierarchical space-time model on the data, a crude data inspection will be carried out and some main features of the raw data pertaining to the selected area will be presented. The highest wind speed recorded in the monthly maximum data overall is just above 32 m s^{-1} and the minimum value in the dataset is about 8.6 m s^{-1} . The estimation of possible long-term trends is one of the main motivations for applying the stochastic model, and a very crude approach to see if there are any likely trends is to fit a straight line to the raw data. In

²<http://www.knmi.nl/waveatlas>

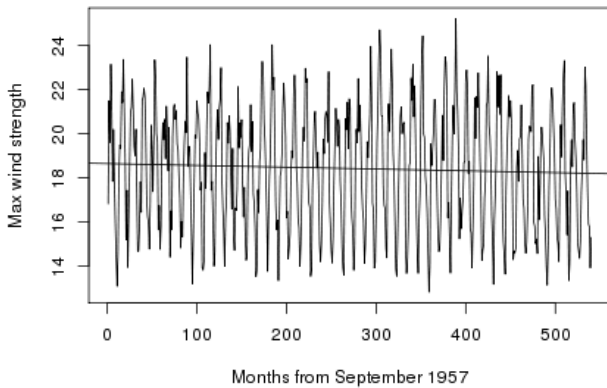


Fig. 2. Fitting a straight line to the raw data indicates a negative trend.

Fig. 2, the spatially averaged data are shown with a straight line fitted to it. It is observed that this line has a negative slope of $-0.0008171 \text{ m s}^{-1}$ per month with an intercept of 18.64 m s^{-1} , corresponding to slightly decreasing monthly maximum wind speeds of about 0.44 m s^{-1} overall throughout the period. It is also interesting to observe the seasonality in the raw wind data as illustrated by Fig. 3, where the spatially averaged data for the first ten years are shown.

3 The stochastic model

The Bayesian hierarchical space-time model resembles the model for significant wave height (Vanem et al., 2012b) and the spatiotemporal data is indexed in a similar way by two indices; an index x to denote the spatial location, with $x = 1, 2, \dots, X = 66$ and an index t to denote time (i.e., month), where $t = 1, 2, \dots, T = 540$, and $t = 1$ corresponds to September 1957 and $t = T = 540$ corresponds to August 2002. The maximum windiness at location x in month t is expressed by $W(x, t)$. The structure of the main model will be outlined below, and a few alternative models that were also explored will be described. A brief discussion of the prior distributions applied to the model parameters will also be given, and this fully specifies the model. It is noted that the derivation of the full conditionals used in the Gibbs sampler is completely analogous to the significant wave height model and reference is made to Vanem et al. (2012b) for details; see also Natvig and Tvette (2007).

It is emphasised that this model is a stochastic model rather than a physics-based model and the physical mechanisms in the atmosphere responsible for generating wind are not explicitly included in the model. It is rather the complex stochastic dependence structures, in space and time at various scales, in the wind data itself that have been modelled. However, it is argued that all the physics are undeniably inherent in the data, so all the relevant physical mechanisms are implicitly taken into account by the model, by way of the

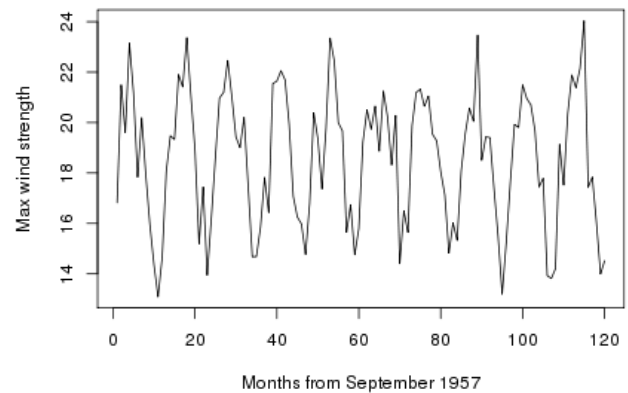


Fig. 3. The seasonality of the raw data.

data. Hence, this probabilistic model is proposed as a complement and an alternative to more deterministic, geophysics based models.

3.1 Main model specification

At the first level, the observation or data equation (Eq. 2) models the maximum windiness at location x and month t as a latent (or hidden) process H , corresponding to an underlying wind speed process that may normally be construed as the true process, and some random noise, ε_W , which may be construed as measurement error or data uncertainty.

$$W(x, t) = H(x, t) + \varepsilon_W(x, t) \quad \forall x, t. \quad (2)$$

It is noted that in Eq. (2) and in the following (Eqs. 4–8), all the noise terms in the model are assumed mutually independent and independent in space and time, having a zero-mean Gaussian distribution with some random but identical variance. With generic notation, $\varepsilon_N \stackrel{i.i.d.}{\sim} N(0, \sigma_N^2)$. Furthermore, all system equations are assumed to apply for all locations and months (i.e., $\forall x, t$) as applicable.

The underlying process for monthly maximum windiness is modelled by the state or system model which is split into a time-independent part $\mu(x)$, a short-term temporally and spatially dependent part $\theta(x, t)$ and two spatially independent parts $M(t)$ and $T(t)$ for seasonality and long-term trends, respectively (Eq. 3).

$$H(x, t) = \mu(x) + \theta(x, t) + M(t) + T(t) \quad (3)$$

The time-independent part is modelled as a first-order Markov Random Field (MRF), conditional on its nearest neighbours in all cardinal directions, and with different dependence parameters in lateral and longitudinal direction, as shown in Eq. (4). For the remainder of this paper, the following notation is used for neighbouring locations of x in space: x^D = the location of the nearest grid point in direction D from x , where $D \in \{N, S, W, E\}$ and $N = \text{North}$, $S = \text{South}$, $W = \text{West}$ and $E = \text{East}$. If x is at the border of the area, the

value at the corresponding neighbouring grid point outside the data area is taken to be zero. Hence, no particular adjustment is made to account for edge effects.

$$\begin{aligned} \mu(x) = & \mu_0(x) + a_\phi \left\{ \mu(x^N) - \mu_0(x^N) + \mu(x^S) - \mu_0(x^S) \right\} \\ & + a_\lambda \left\{ \mu(x^E) - \mu_0(x^E) + \mu(x^W) - \mu_0(x^W) \right\} \\ & + \varepsilon_\mu(x). \end{aligned} \tag{4}$$

In the equation above, $\mu_0(x)$ is the Markov Random Field mean at grid point x and a_ϕ and a_λ are spatial dependence parameters in lateral (i.e., North–South) and longitudinal (i.e., East–West) direction, respectively. σ_μ^2 is the homogeneous Markov Random Field noise variance. The spatially specific mean, $\mu_0(x)$, is modelled as having a quadratic form with an interaction term in latitude and longitude, as shown in Eq. (5). Letting $m(x)$ and $n(x)$ denote the longitude and latitude of location x , respectively, it is assumed that

$$\begin{aligned} \mu_0(x) = & \mu_{0,1} + \mu_{0,2}m(x) + \mu_{0,3}n(x) + \mu_{0,4}m(x)^2 \\ & + \mu_{0,5}n(x)^2 + \mu_{0,6}m(x)n(x) \end{aligned} \tag{5}$$

The spatiotemporal dynamic term $\theta(x, t)$ is modelled as a vector autoregressive model of order one, conditionally specified on its nearest neighbours in all cardinal directions, as shown in Eq. (6).

$$\begin{aligned} \theta(x, t) = & b_0\theta(x, t - 1) + b_N\theta(x^N, t - 1) + b_E\theta(x^E, t - 1) \\ & + b_S\theta(x^S, t - 1) + b_W\theta(x^W, t - 1) + \varepsilon_\theta(x, t) \end{aligned} \tag{6}$$

The seasonal component is modelled as a combination of an annual and a semi-annual seasonal contribution, where the seasonal contribution is assumed independent of space, see Eq. (7). The period of the seasonal cycle is one year, so $\omega = \frac{2\pi}{12}$ for monthly maximum data, and both the first and second harmonics are included.

$$\begin{aligned} M(t) = & c \cos(\omega t) + d \sin(\omega t) \\ & + f \cos(2\omega t) + g \sin(2\omega t) + \varepsilon_m(t) \end{aligned} \tag{7}$$

The long-term trend is modelled as a simple Gaussian process with a linear trend, as shown in Eq. (8), and is perhaps the component of most interest in this study.

$$T(t) = \gamma t + \varepsilon_T(t). \tag{8}$$

3.2 Alternative models

Two model alternatives have also been explored in this study. First, a model where the trend component has been removed is applied, and this model is simply identical to the model in Eqs. (2) and (3) with $T(t) = 0$ and all other components unchanged, as presented in Eq. (9)

$$\begin{aligned} W(x, t) = & H(x, t) + \varepsilon_W(x, t) \\ = & \mu(x) + \theta(x, t) + M(t) + \varepsilon_W(x, t). \end{aligned} \tag{9}$$

This model would then explain the data under the assumption of no trend.

Another model alternative that is introduced in order to account for possible heteroscedastic features of the data and possible stronger trends in the extremes compared to the means is to take the logarithmic transformation of the data. By taking this log-transform, the model effectively changes into the model in Eq. (10), where $H(x, t)$ is still modelled as in Eqs. (3)–(8), see also Vanem et al. (2012d).

$$Y(x, t) = \ln W(x, t) = H(x, t) + \varepsilon_Y(x, t). \tag{10}$$

An equivalent representation of this model alternative on the original scale is

$$W(x, t) = e^{\mu(x)} e^{\theta(x, t)} e^{M(t)} e^{T(t)} e^{\varepsilon_Y(x, t)} \tag{11}$$

where now the contributions from the individual components to the wind speed have become multiplicative factors rather than additive contributions. This gives a different interpretation of the model components.

In principle, there are some identifiability problems with respect to the two temporal noise terms ε_m and ε_T and there is no way for the model to distribute the temporal noise between these two terms. It turns out that when both terms are included in the model, the temporal noise is evenly distributed between the two terms, and it should be acknowledged that this is somewhat arbitrary. One way around this is to merge the terms $M(t)$ and $T(t)$ into a joint temporal term with only one temporal noise term. However, it turns out that doing this does not really influence the overall results and the identifiability problems seem to be isolated to the two noise terms. The other model parameters and, hence, the model results do not seem to be sensitive to these differences. Hence, for the purpose of this study only results from the model with two noise terms are reported. Similar conclusions were also arrived at for the significant wave height model, and in Vanem et al. (2012b,d) results for both alternatives were presented.

3.3 Prior distributions

In Bayesian hierarchical models, uncertainties of the model parameters are included in the model by way of prior distributions. Thus, specifying prior distributions for all model parameters completes the specification of the model. The priors adopted in this study are summarised in Table 1, and the rationale for assigning these priors are similar as for the significant wave height model, as discussed in Vanem et al. (2012b). $IG(\alpha, \beta)$ denotes the inverse gamma distribution with parameters α and β and $N(\mu, \sigma^2)$ denotes the Gaussian distribution with mean μ and variance σ^2 .

It is noted that the final results are not very strongly dependent on the particular hyper-parameters in the prior distributions, and this is reasonable in light of the large amount of data; priors tend to become asymptotically irrelevant as the amount of data increases. A few simulations have also been run with non-informative priors on the noise variances, but the overall results did not seem to be notably affected by this.

Table 1. Prior distributions for the model parameters.

Parameters	Prior distributions
$\sigma_W^2, \sigma_\mu^2, \sigma_\theta^2, \sigma_m^2$ and σ_T^2	IG (3, 2)
a_ϕ, a_λ	$N(0.2, 0.25)$
b_0, b_N, b_S, b_E and b_W	$N(0.2, 0.25)$
$\mu_{0,i}$ for $i = 2, \dots, 6$	$N(0, 9)$
$\mu_{0,1}$	$N(20, 9)$
c, d, f, g	$N(0, 4)$
γ	$N(0, 0.1)$
$\theta(x, 0), \forall x$	$N(0, 15)$

4 MCMC simulations

Letting Θ denote the model parameters, the model as specified above is essentially on the form of a product of an observation model, $f(W|H, \Theta)$, a state model $f(H|\Theta)$ and a parameter model $f(\Theta)$:

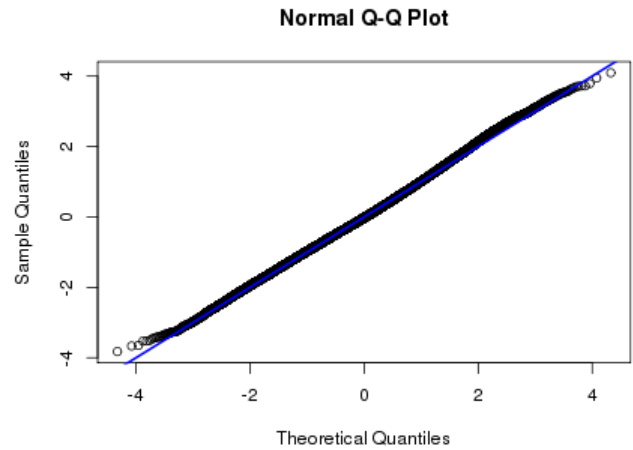
$$f(W, H, \Theta) = f(W|H, \Theta)f(H|\Theta)f(\Theta). \quad (12)$$

Now, the conditional model of the latent process and the model parameters given the data, referred to as the posterior distribution, can be found from Bayes' theorem,

$$f(H, \Theta|W) = \frac{f(W|H, \Theta)f(H|\Theta)f(\Theta)}{f(W)} \quad (13)$$

and it is from this posterior distribution all inference and predictions on the process H and the parameters Θ are made. The high dimension of Θ makes this posterior distribution difficult to compute analytically, but using Markov chain Monte Carlo (MCMC) methods it is possible to simulate samples from it. In order to simulate from the model in this study, Markov chains are constructed using the Gibbs sampler with additional Metropolis-Hastings steps for the a_ϕ and a_λ parameters. The Gibbs sampler requires the full conditionals for all the parameters involved, and since the model has been specified conditionally these are mostly quite straightforward to derive (Vanem et al., 2012b). The pseudo-distribution of the (a_ϕ, a_λ) -pair was used as proposal for the Metropolis-Hastings step, as explained in Vanem et al. (2012b).

The MCMC simulations performed in this study used a burn-in period of 100 000 samples and a batch size of 20 (i.e., keeping every 20th sample after the burn-in period) corresponding to a total of 120 000 simulations to obtain a collection of 1000 samples of the multi-dimensional parameter vector. The Metropolis-Hastings steps were repeated six times to obtain an overall acceptance rate of about 70%. The burn-in period is much longer than what was used for the significant wave height model, and it ensures that the Markov chain converges. The assumption that the chain converges within the burn-in period is confirmed by trace plots and a few control runs. A simulation according to these

**Fig. 4.** Normal probability plot of the residuals.

specifications completes in about 5 h on a computer with an Intel Core i5-2500 CPU @ 3.30 GHz processor.

A normal probability plot of the residuals is investigated in order to check the main model assumption of normality at the observation level, and this is shown in Fig. 4. This indicates that the assumption is reasonable. The figure shows the residuals for the main model with a linear trend, but normal probability plots of the residuals for the model without trend and also for the model applied to log-transformed data look equally reasonable.

5 Results

In this section, the results for the different model alternatives and pertaining to the different model components will be reported. The estimated marginal posterior distributions (mean and standard deviation) are presented in the table in Appendix A (Table A1). It is noted that the parameter estimates for the log-transformed data are not directly comparable to the estimates pertaining to the original model.

5.1 Results from the main model

First, the results for the various components for the main model, including a linear long-term trend component will be presented. Overall, this model seems to perform well and the results appear stable; running a few control simulations provides very similar results.

5.1.1 Spatial component

The parameters $\mu_{0,i}$, together with a_ϕ and a_λ determine the contribution from the time-independent spatial field $\mu(x)$. Over the area, the mean of this contribution varies between 17.0 and 19.2 m s⁻¹ and this component, thus, explains some of the spatial variation in the wind speed data. The mean estimated posterior field $\mu(x)$ is illustrated in Fig. 5 and it is

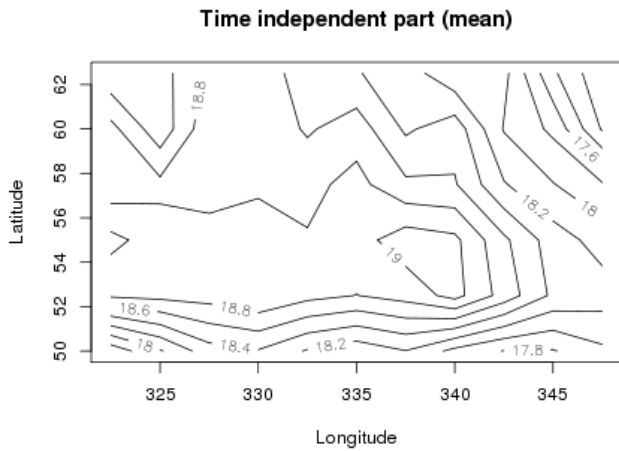


Fig. 5. The spatial field $\mu(x)$.

observed that there is variability in both north-south and in east-west directions, with generally higher wind speeds to the east and the north of the area.

5.1.2 Short-term dynamic component

The space-time dynamic component $\theta(x, t)$ is described by the b . parameters, and the mean of this component is found to vary between -1.63 to 1.88 m s^{-1} over all times (except $t = 0$) and locations. Hence, a notable part of the variability of wind speeds are captured by this component. The mean contribution of this component, averaged over all times, are zero, meaning that this component is stationary over long time scales as it should, not contributing to the long-term trend part of the model.

5.1.3 Seasonal component

The estimated seasonal contribution according to the model is illustrated in Fig. 6, and it is seen that this captures the seasonal characteristics of the raw data quite well. It is noted that the figure only shows the seasonal contribution for the first ten years, but the contribution is valid for the complete time span of the data, and all data have been used in estimating the seasonal contribution. The seasonal component was modelled as a combination of an annual and a semi-annual part and is described by the parameters c, d, f and g . The estimated parameters correspond to a first harmonic with mean amplitude $A_1 = \sqrt{c^2 + d^2} \approx 3.5 \text{ m s}^{-1}$ and a second harmonic with mean amplitude $A_2 = \sqrt{f^2 + g^2} \approx 0.48 \text{ m s}^{-1}$ corresponding to a seasonal contribution varying between 3.0 m s^{-1} in February and about -3.9 m s^{-1} in August. It is observed that the annual contribution is dominating, but the contribution from the semi-annual component is not negligible. The mean sampled seasonal contribution $M(t)$ has an overall minimum of -4.66 and a maximum of 4.75 m s^{-1} .

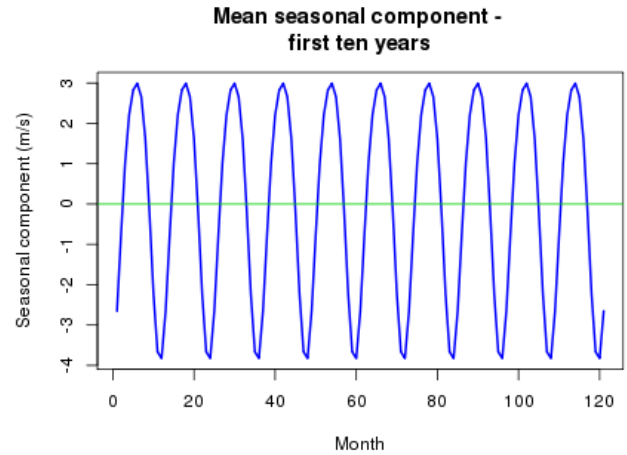


Fig. 6. The posterior mean seasonal contribution.

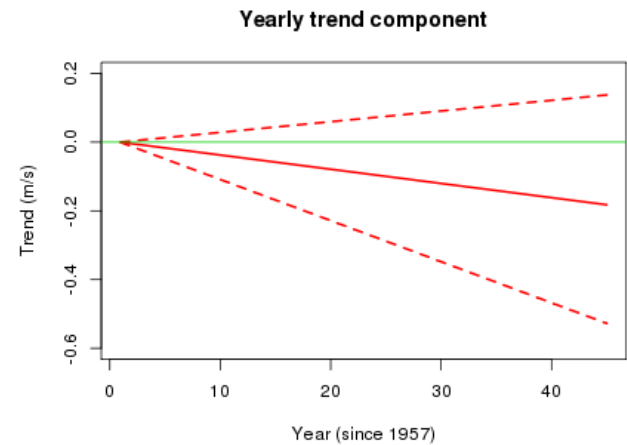


Fig. 7. The posterior mean long-term trend component with 90 % credible interval of γt ; green line corresponds to no trend.

5.1.4 Long-term trend component

The long-term trend contribution is perhaps the one of most interest, and a linear trend was assumed in the main model. The linear trend in the model is determined by the parameter γ and the mean posterior γ is estimated to $-0.000345 \text{ m s}^{-1}$ per month. This corresponds to an overall decrease of about 0.19 m s^{-1} over the whole period which is slightly less than the straight line fitted to the raw data in Fig. 2. The estimated mean long-term trend together with a 90 % credible interval of the mean are illustrated in Fig. 7. The estimated trend corresponds to a mean decrease of about 19 cm s^{-1} with a 90 % credible interval ranging from negative to positive trends in the monthly maximum wind speed over the whole period. Hence, even though the data indicate that there might be a slight decreasing trend in the wind data no statistically significant trend in the wind speed is estimated by the model.

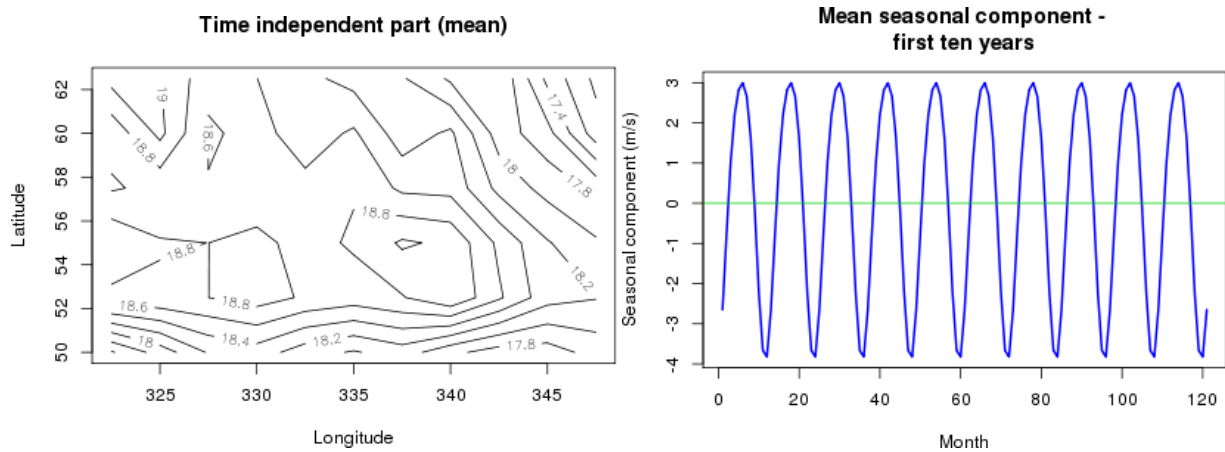


Fig. 8. Results from the model without trend; the mean spatial field and seasonal contribution.

5.2 Results from alternative models

Two alternative models were also tried out, one without a long-term trend and one on the log-transformed data. The estimated posterior distributions of these simulations are also included in Table A1, and the results will be briefly presented in the following.

5.2.1 Without trend

The model was also tried without any long-term trends (i.e., by letting $T(t) = 0$), and apart from the absence of any trends, the results were very similar to the results from the main model. The values of the estimated spatial field are comparable, although perhaps slightly lower, without any trend, with values ranging from 17.0 to 19.0 m s^{-1} , which seems reasonable. The contributions from the short-term dynamic (mean $\approx 0.00 \text{ m s}^{-1}$ and ranging from -1.67 to 1.92 m s^{-1}) and seasonal (ranging from -3.9 to 3.0 m s^{-1}) components are very similar, as can also be seen from comparing the estimates in Table A1. The estimated mean spatial field and seasonal contribution are illustrated in Fig. 8. The fact that the models with a linear trend and without any trend are similar is not surprising, especially since the model with a trend component failed to identify a statistically significant trend. Hence, for all practical purposes the models can be regarded as identical.

5.2.2 With a logarithmic data transformation

When the model was run with a logarithmic transformation of the data, the results are not directly comparable, but similar main features are identified. The estimated mean spatial field now varies between 16.9 and 19.1 m s^{-1} , with the contributions from the short-term dynamic, seasonal and long-term trend parts now being multiplicative factors. The estimated spatial field and the seasonal contribution are illustrated in Fig. 9. The estimated seasonal contribution varies

between factors of 0.80 for calm seasons and 1.2 for windy seasons. The mean factor reflecting the contribution from the short-term dynamic component, $E[e^{\theta(x,t)}]$ varies between 0.79 and 1.2 for all locations and times except $t = 0$. It is noted that the figures display the results on the re-transformed original scale, and bias corrections have been applied when necessary, see e.g., Vanem et al. (2012d); Ferguson (1986); Beauchamp and Olson (1973); Stow et al. (2006) for more details.

The estimated long-term trend factor is illustrated in Fig. 10. The mean long-term trend is found to be decreasing, but the 90% credible interval ranges from decreasing to increasing trends. The mean estimated trend for the whole period corresponds to a factor of 0.98 . The 90% credible interval for the expected trend factor ranges from 0.94 to 1.02 , hence, the decreasing trend is not significant at the 90% level. For a typical monthly maximum wind speed of about 18 m s^{-1} the estimated mean trend corresponds to a decrease of about 36 cm s^{-1} over the whole period, but for more extreme wind speeds, say 30 m s^{-1} , the trend factor corresponds to a decrease of 60 cm s^{-1} . Hence, the model on the log-transformed data yields larger trends for extremes compared to averages.

Comparing the results from the model for log-transformed data to the results pertaining to the original data, it is observed that a slightly more decreasing mean trend is estimated, but both models fail to identify any significant trends. Hence, the results generally agree. It is questionable whether the log-transform represents an improvement, and the results indicate that the original model might perform better. For example, short-term prediction losses are smaller for the original model.

5.3 Results pertaining to another ocean area

It seems obvious that the identified increase in significant wave height cannot be explained by the absence of any

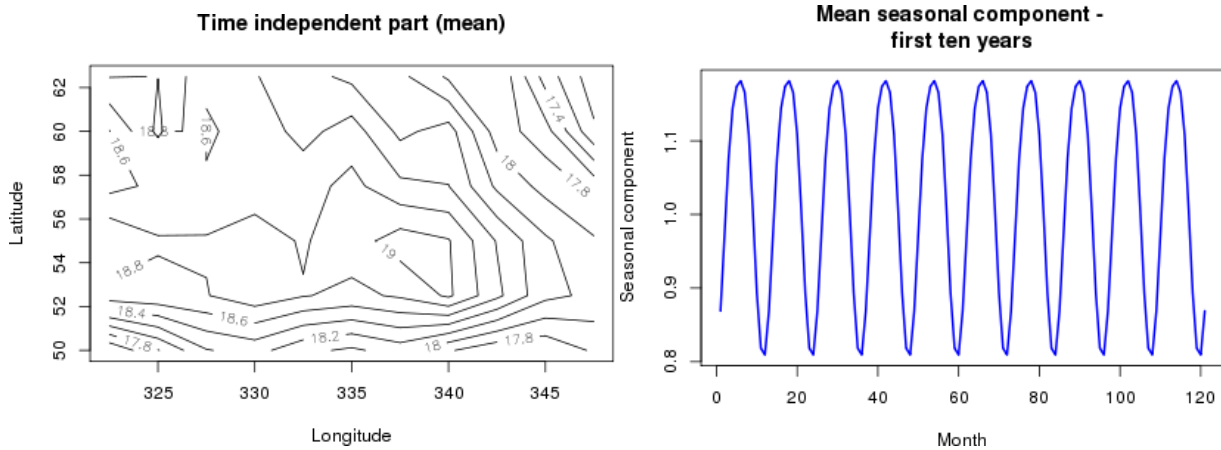


Fig. 9. Results from the model with log-transform of the data: spatial field and seasonal factor

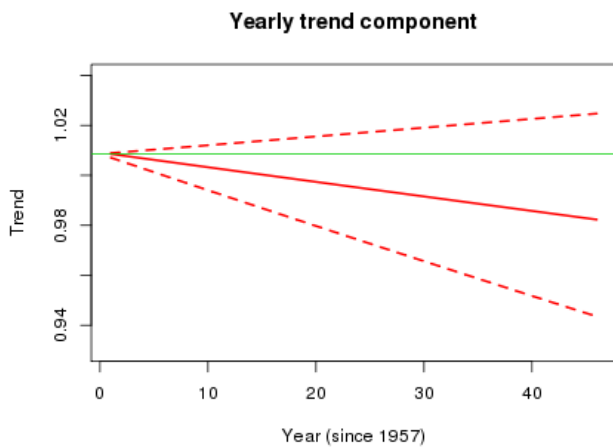


Fig. 10. The estimated trend factor for the log-transformed data, with 90 % credible interval.

increase or even a slight *decrease* (although not statistically significant) in wind speed over the area. Hence, the increasing significant wave height might be a result of increased swell from increased windiness in other areas. In order to investigate this, another area to the north of the initially investigated area is analysed. The spatial resolution and temporal span are the same as for the initial area. The coordinates of this other area are from 67.5° to 77.5° N and from 345° to 357.5° W, corresponding to a grid with $5 \times 6 = 30$ grid points north of Iceland. The lowest and highest values for monthly maximum wind speed in this area are 6.8 and 30.5 m s^{-1} , respectively, somewhat less than for the area initially studied.

The results for the new area look reasonable for the spatial field, the short-term spatiotemporal part and the seasonal component. The mean spatial field varies between 13.9 and 16.6 m s^{-1} over the area with a mean contribution of 15.7 m s^{-1} ; the short-term dynamic part $\theta(x, t)$ varies between -3.69 m s^{-1} and 3.53 m s^{-1} with a mean of about

zero m s^{-1} over the area and entire period; the seasonal contribution varies between -4.47 and 2.93 m s^{-1} . The results for the spatial field and the seasonal contribution are illustrated in Fig. 11.

It is interesting to observe that the model picks up a significantly positive long-term trend when applied to this area north of Iceland. The estimated trend is illustrated in Fig. 12, and the mean estimated trend corresponds to an increase in monthly maximum wind speed of about 0.75 m s^{-1} over the period. The 90 % credible interval of the expected trend ranges from 0.37 to 1.1 m s^{-1} and is, hence, entirely positive. Therefore, even though there seems to have been an insignificant decrease in wind speeds in the area first investigated, this area more north seems to have experienced an overall increase in wind speeds. It is noted, however, that even though the increase was found to be statistically significant, an increase of 0.75 m s^{-1} is not necessarily practically significant with respect to wave generation and the effects on the significant wave height.

6 Discussion

The Bayesian hierarchical space-time model applied to monthly maximum wind speeds over an area in the mid-latitude North Atlantic has identified a possible slightly negative trend for the initial area in the North Atlantic, albeit not statistically significant. When significant wave height data were modelled with a similar model for the same area, an increase in monthly maximum significant wave height was discovered, and the results from the analysis of the wind data indicate that this increase is not due to increased wind sea. If the increases in significant wave height are to be explained by increased wind sea, it would necessarily need to be accompanied by an increase in local wind speeds.

The increase in significant wave height could then be explained by increased swell, i.e., remains of wind sea

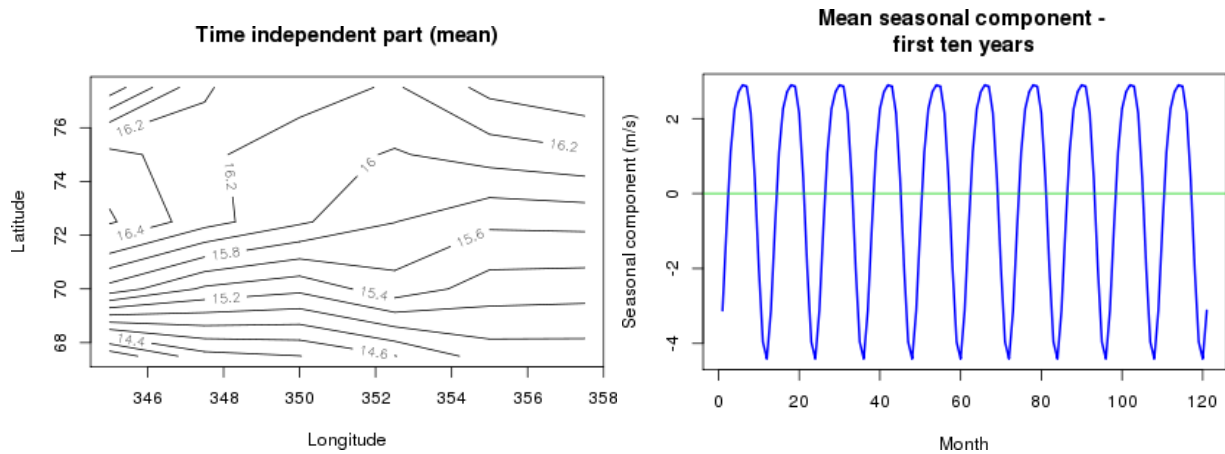


Fig. 11. Results from the model applied to an alternative area: spatial field and seasonal factor.

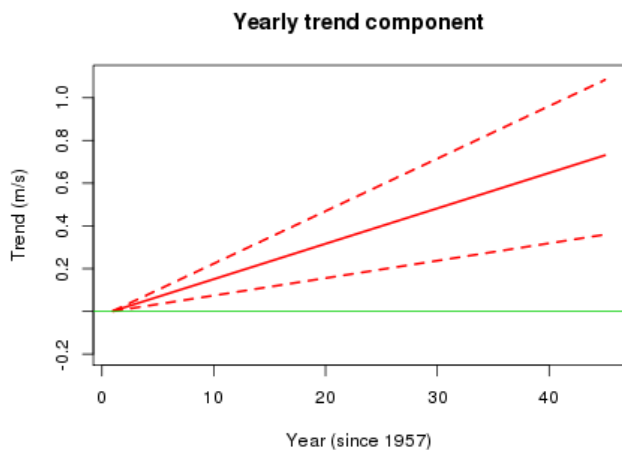


Fig. 12. The estimated mean trend for the area north of Island with 90 % credible intervals.

generated by wind forces outside the area and propagated into the area that has been analysed. This assumption is substantiated by the results obtained when the wind speed model was applied to an area further north. For this area, a significant increase of monthly maximum wind speeds was detected which would possibly lead to increased wind sea in this alternative area. This increased wind sea could then possibly propagate as increased swell to the initial area and explain, at least partly, the observed increase in significant wave height here. It should be noted, however, that only wind speeds have been analysed and the direction of swell propagation would obviously be highly dependent on wind direction.

The absence of a long-term trend (or possibly a slight decreasing trend) in windiness in one area together with an increase in windiness in an area further north indicate that there might have been a change in the main storm tracks, with storm tracks generally moving more to the north. This would

be in agreement with other studies that have described such an effect, i.e., that the storm tracks have experienced a poleward shift as a result of climate change, see e.g., Yin (2005); Gastineau and Soden (2009); Bender et al. (2012). Such a poleward shift of storm tracks is associated with a poleward shift in surface winds, which would be in agreement with the results presented herein. Bader et al. (2011) presents a review of recent studies and states that an observed poleward shift of mid-latitude storms is the most agreed on result. Hence, it is reassuring that the Bayesian hierarchical space-time model arrives at similar results and is able to pick up this signal in the data.

It is noted that this model only analyses the trend in the monthly maximum wind speed, and this alone might not be a sufficient measure of the wave generating forces. If, for example, the duration of extreme wind speeds increases this would lead to larger waves even without an increase in the wind speed itself. Such effects would not be picked up by the model when applied to monthly maximum wind speeds. Another effect could be the frequencies of strong winds. If there are not sufficiently long calm periods between storms for the sea to quiet down, waves could aggregate to larger heights even without higher wind velocities. Possibly, increasing frequencies and prolonged duration of storms would not necessarily be reflected in the monthly maximum wind speed data and further analyses would be needed to look into this. This is left for further study. Furthermore, the wind direction is important for wave generation, and wind directions have not been analysed in the present study. A possible extension of the model could be to include wind direction as well as magnitude. It is also assumed that possible effects of changes in fetch due to Arctic ice reduction is negligible.

It has already been emphasised that even though the governing physics are not explicitly included in the model presented in this study it is undeniably incorporated in the model by way of the data. The Bayesian hierarchical model is a purely probabilistic model and as such it is different from

Table A1. Posterior marginal distributions (mean; standard deviation).

	Main model	No trend	Log-transform	Alternative area
γ	-0.00034; 0.084	–	$-4.9 \times 10^{-5}; 4.8 \times 10^{-5}$	0.0014; 0.00042
$\mu_{0,1}$	20; 3.0	20; 3.0	3.1; 2.0	20; 5.2
$\mu_{0,2}$	-0.41; 0.13	-0.42; 0.13	-0.026; 0.076	-0.56; 0.37
$\mu_{0,3}$	2.6; 0.77	2.7; 0.79	0.16; 0.45	2.9; 2.3
$\mu_{0,4}$	0.00093; 0.00033	0.00094; 0.00034	$5.1 \times 10^{-5}; 0.00019$	0.0012; 0.0010
$\mu_{0,5}$	-0.010; 0.0047	-0.011; 0.0047	-0.00089; 0.0024	-0.012; 0.0076
$\mu_{0,6}$	-0.0043; 0.0019	-0.0044; 0.0020	-0.00018; 0.0010	-0.0038; 0.0062
a_ϕ	0.17; 0.080	0.17; 0.082	0.12; 0.079	0.14; 0.093
a_λ	0.29; 0.084	0.28; 0.089	0.13; 0.081	0.19; 0.11
c	-2.7; 0.078	-2.7; 0.077	-0.15; 0.011	-3.0; 0.094
d	2.2; 0.078	2.2; 0.076	0.12; 0.012	2.1; 0.095
f	-0.0037; 0.078	-0.0078; 0.073	-0.0023; 0.011	-0.13; 0.093
g	0.48; 0.078	0.48; 0.073	0.035; 0.010	0.80; 0.093
b_0	0.30; 0.016	0.30; 0.016	0.018; 0.0048	0.20; 0.018
b_N	0.056; 0.020	0.051; 0.021	0.010; 0.0055	0.84; 0.040
b_E	0.24; 0.018	0.24; 0.017	0.012; 0.0048	0.20; 0.019
b_S	-0.14; 0.027	-0.15; 0.030	-0.0011; 0.0052	-0.041; 0.010
b_W	0.26; 0.019	0.26; 0.017	0.018; 0.0048	0.16; 0.016
σ_W^2	3.2; 0.030	3.2; 0.031	0.0057; 0.00015	2.7; 0.032
σ_μ^2	0.11; 0.022	0.11; 0.022	0.062; 0.011	0.25; 0.063
σ_θ^2	0.25; 0.018	0.26; 0.017	0.0057; 0.00015	0.055; 0.0073
σ_m^2	0.77; 0.25	1.5; 0.090	0.017; 0.0013	1.2; 0.47
σ_T^2	0.72; 0.24	–	0.017; 0.0013	1.0; 0.46

many meteorological and geophysical models based on deterministic relationships such as the Navier-Stokes equations for describing the climate and the atmospheric circulation and for projecting climate change. Physical models remain the primary approach for investigating the impacts of climate change and ensemble studies are carried out in order to quantify uncertainties, where different climate models and small perturbations of the initial conditions give different results. However, it has been acknowledged that there are notable statistical challenges related climate change projections based on such ensemble studies, see e.g., Tebaldi and Knutti (2007); Stephenson et al. (2012); Collins et al. (2012). The models presented in this paper offer an alternative approach to modelling the impacts of climate change on the wind climate with a more direct approach to modelling of uncertainties, and it should rather be regarded as a complement to the efforts made in developing physical models than a competitor.

7 Conclusions

This paper has presented a Bayesian hierarchical spatiotemporal model for 10 m wind speeds over an area in the North Atlantic ocean. Overall, the model seems to perform well in capturing the dominating spatial and temporal dependence structures in the wind speed data. Hence, this paper suggests this modelling framework as an alternative to physical

models for analysing environmental processes such as wind speed in space and time.

A similar model has previously been applied to extreme wave climate over the same area and identified inter alia, increasing trends in the monthly maximum significant wave height. Previous studies have also demonstrated that such increasing trends may have an impact on ship structural loads and that it, therefore, represents an additional hazard to ship operations. The results from the wind speed models do not suggest any corresponding increases in the monthly maximum wind speeds. On the contrary, a slight decreasing trend was estimated although this was not statistically significant. Hence, the results indicate that the roughening of the wave climate could not be explained by increases in locally generated wind sea. Possibly, the increased significant wave height can be explained by increased swell in the area. The monthly maximum wind speed was also analysed for another ocean area further north and in this area a significant positive trend in the 10 m wind speed was identified. These results agree with various previous studies that suggest that the North Atlantic storm tracks shift polewards due to a warming climate and could also explain, at least partly, an increase in swell in the original area. Thus, the results presented in this paper suggest that the observed increased significant wave height might be mostly due to increased swell.

Appendix A

Posterior distributions

The estimated marginal posterior distributions for all the model parameters for the different simulations are presented in Table A1, in terms of estimated mean and standard deviations.

Acknowledgements. Thanks to Bent Natvig and Ida Scheel at the statistics group, Department of Mathematics, University of Oslo for valuable discussions and comments and for general support and encouragement. Thanks also to the Statistics for Innovation (SFI²) Centre for Research-based Innovation for financial support to cover the service charges.

Edited by: A. Toffoli

Reviewed by: three anonymous referees

References

- Bader, J., Mesquita, M. D. S., Hodges, K. I., Keenlyside, N., Østerhus, S., and Miles, M.: A review on Northern Hemisphere sea-ice, storminess and the North Atlantic Oscillation: Observations and projected changes, *Atmos. Res.*, 101, 809–834, 2011.
- Beauchamp, J. J. and Olson, J. S.: Corrections for Bias in Regression Estimates After Logarithmic Transformation, *Ecology*, 54, 1403–1407, 1973.
- Bender, F. A.-M., Ramanathan, V., and Tselioudis, G.: Changes in extratropical storm track cloudiness 1983–2008: observational support for a poleward shift, *Clim. Dynam.*, 38, 2037–2053, 2012.
- Caires, S., Swail, V. R., and Wang, X. L.: Projection and Analysis of Extreme Wave Climate, *J. Climate*, 19, 5581–5605, 2006.
- Collins, M., Chandler, R. E., Cox, P. M., Huthnance, J. M., and Rougier, J.: Quantifying future climate change, *Nat. Climate Change*, 2, 403–409, 2012.
- Debernard, J. B. and Røed, L. P.: Future wind, wave and storm surge climate in the Northern Seas: a revisit, *Tellus*, 60A, 427–438, 2008.
- Ferguson, R.: River Loads Underestimated by Rating Curves, *Water Resour. Res.*, 22, 74–76, 1986.
- Gastineau, G. and Soden, B. J.: Model projected changes of extreme wind events in response to global warming, *Geophys. Res. Lett.*, 36, L10810, doi:10.1029/2009GL037500, 2009.
- Grabemann, I. and Weisse, R.: Climate change impact on extreme wave conditions in the North Sea: an ensemble study, *Ocean Dynam.*, 58, 199–212, 2008.
- IPCC: Climate Change 2007: The Physical Sciences Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 2007.
- IPCC: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation, Cambridge University Press, 2012.
- Kushnir, Y., Cardone, V., Greenwood, J., and Cane, M.: The Recent Increase in North Atlantic Wave Heights, *J. Climate*, 10, 2107–2113, 1997.
- Natvig, B. and Tvette, I. F.: Bayesian Hierarchical Space-time Modeling of Earthquake Data, *Methodol. Comput. Appl. Probability*, 9, 89–114, 2007.
- Stephenson, D. B., Collins, M., Rougier, J. C., and Chandler, R. E.: Statistical problems in the probabilistic prediction of climate change, *Environmetrics*, 23, 364–372, 2012.
- Stow, C. A., Reckhow, K. H., and Qian, S. S.: A Bayesian Approach to Retransformation Bias in Transformed Regression, *Ecology*, 87, 1472–1477, 2006.
- Talley, L. D., Pickard, G. L., Emery, W. J., and Swift, J. H.: Descriptive Physical Oceanography an introduction, Elsevier, 6th Edn., 2011.
- Tebaldi, C. and Knutti, R.: The use of the multi-model ensemble in probabilistic climate projections, *Phil. Trans. Roy. Soc. A*, 365, 2053–2075, 2007.
- Uppala, S. M., Kållberg, P. W., Simmons, A. J., Andrae, U., Da Costa Bechtold, V., Fiorino, M., Gibson, J. K., Haseler, J., Hernandez, A., Kelly, G. A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R. P., Andersson, E., Arpe, K., Balmaseda, M. A., Beljaars, A. C. M., Van de Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B. J., Isaksen, I., Janssen, P. A. E. M., Jenne, R., McNally, A. P., Mahfouf, J.-F., Morcrette, J.-J., Rayner, N. A., Saunders, R. W., Simon, P., Sterl, A., Trenberth, K. E., Untch, A., Vasiljevic, D., Vitebro, P., and Woolen, J.: The ERA-40 re-analysis, *Q. J. Roy. Meteorol. Soc.*, 131, 2961–3012, 2005.
- Vanem, E.: Long-term time-dependent stochastic modelling of extreme waves, *Stochastic Environ. Res. Risk Assess.*, 25, 185–209, 2011.
- Vanem, E. and Bitner-Gregersen, E.: Stochastic modelling of long-term trends in the wave climate and its potential impact on ship structural loads, *Appl. Ocean Res.*, 37, 235–248, 2012.
- Vanem, E. and Walker, S.-E.: Identifying trends in the ocean wave climate by time series analyses of significant wave height data, *Ocean Eng.*, 60, 148–160, 2013.
- Vanem, E., Huseby, A. B., and Natvig, B.: Bayesian Hierarchical Spatio-Temporal Modelling of Trends and Future Projections in the Ocean Wave Climate with a CO₂ Regression Component, submitted, 2012a.
- Vanem, E., Huseby, A. B., and Natvig, B.: A Bayesian Hierarchical Spatio-Temporal Model for Significant Wave Height in the North Atlantic, *Stochastic Environ. Res. Risk Assess.*, 26, 609–632, 2012b.
- Vanem, E., Huseby, A. B., and Natvig, B.: A Stochastic Model in Space and Time for Monthly Maximum Significant Wave Height, in: Proc. Ninth International Geostatistics Congress (Geostats 2012), 505–517, 2012c.
- Vanem, E., Huseby, A. B., and Natvig, B.: Modeling Ocean Wave Climate with a Bayesian Hierarchical Space-Time Model and a Log-Transform of the Data, *Ocean Dynam.*, 62, 355–375, 2012d.
- Wallcraft, A. J., Kara, A. B., Barron, C. N., Metzger, E. J., Pauley, R. L., and Bourassa, M. A.: Comparison of monthly mean 10 m wind speeds from satellites and NWP products over the global ocean, *J. Geophys. Res.*, 114, D16109, doi:10.1029/2008JD011696, 2009.
- Wang, X. J., Zwiers, F. W., and Swail, V. R.: North Atlantic Ocean Wave Climate Change Scenarios for the Twenty-First Century, *J. Climate*, 17, 2368–2383, 2004.

- Wang, X. L. and Swail, V. R.: Changes of Extreme Wave Heights in Northern Hemisphere Oceans and Related Atmospheric Circulation Regimes, *J. Climate*, 14, 2204–2221, 2001.
- Wang, X. L. and Swail, V. R.: Trends of Atlantic Wave Extremes as Simulated in a 40-Yr Wave hindcast Using Kinematically Reanalyzed Wind Fields, *J. Climate*, 15, 1020–1035, 2002.
- Wang, X. L. and Swail, V. R.: Climate change signal and uncertainty in projections of ocean wave heights, *Climate Dynamics*, 26, 109–126, 2006.
- Wikle, C. K., Berliner, L. M., and Cressie, N.: Hierarchical Bayesian space-time models, *Environ. Ecol. Stat.*, 5, 117–154, 1998.
- Wikle, C. K., Milliff, R. F., Nychka, D., and Berliner, L. M.: Spatiotemporal Hierarchical Bayesian Modeling: Tropical Ocean Surface Winds, *J. Am. Stat. Assoc.*, 96, 382–397, 2001.
- Yin, J. H.: A consistent poleward shift of the storm tracks in simulations of 21st century climate, *Geophys. Res. Lett.*, 32, L1801, doi:10.1029/2005GL023684, 2005.