

Long-term time-dependent stochastic modelling of extreme waves

Erik Vanem

Published online: 5 August 2010

© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract This paper presents a literature survey on time-dependent statistical modelling of extreme waves and sea states. The focus is twofold: on statistical modelling of extreme waves and space- and time-dependent statistical modelling. The first part will consist of a literature review of statistical modelling of extreme waves and wave parameters, most notably on the modelling of extreme significant wave height. The second part will focus on statistical modelling of time- and space-dependent variables in a more general sense, and will focus on the methodology and models used also in other relevant application areas. It was found that limited effort has been put on developing statistical models for waves incorporating spatial and long-term temporal variability and it is suggested that model improvements could be achieved by adopting approaches from other application areas. In particular, Bayesian hierarchical space–time models were identified as promising tools for spatio-temporal modelling of extreme waves. Finally, a review of projections of future extreme wave climate is presented.

Keywords Extreme waves · Stochastic modelling · Spatiotemporal modelling · Climate change · Risk assessment

1 Introduction

According to casualty statistics, one of the major causes of ship losses is bad weather (Guedes Soares et al. 2001),

which stresses the importance of taking extreme sea state conditions adequately into account in ship design. Therefore, a correct and thorough understanding of meteorological and oceanographic conditions, most notably the extreme values of relevant wave and wind parameters, is of paramount importance to maritime safety. Thus, there is a need for appropriate statistical models to describe these phenomena.

When designing ships and other marine and offshore structures, relevant safety regulations and design standards should be based on the best available knowledge. Meteorological data for the last 50+ years are available and this is often assumed to be representative also for the current situation. However, ships and other marine structures are designed for lifetimes of several decades and design codes and standards should be based on knowledge about the operating environment throughout the expected lifetime of the structure—several decades into the future. Such knowledge will also be crucial for any risk assessment of maritime transportation or offshore operations.

According to the IPCC Fourth Assessment Report (IPCC 2007), the globe is currently experiencing climate change and the Earth is warming. It is also very likely that human activities and emission of greenhouse gasses are mainly responsible for the recent rise of global temperatures. Projections of future climate indicate that it is very likely that frequencies and intensities of extreme weather events will increase (IPCC 2007). Model projections also show a poleward shift of the storm tracks with more extreme wave heights in those regions.

Thus, it is increasingly evident that climate change is a reality. An overwhelming majority of researchers and scientists agree on this and it is reasonable to assume that the averages and extremes of sea states are changing and cannot be considered stationary. Hence, it is no longer

E. Vanem (✉)
Statistics Division, Department of Mathematics, University
of Oslo, P.O. Box 1053, Blindern, 0316 Oslo, Norway
e-mail: erikvan@math.uio.no

sufficient to base design codes on stationary wave parameters without any consideration of how these are expected to change in the future. There is a need for time-dependent statistical models that can take the time-dependency of the integrated wave parameters into account, and also adequately model the uncertainties involved, in order to predict realistic operating environments throughout the lifetime of ships and marine structures.

This paper aims at providing a comprehensive, up-to-date review of statistical models proposed for modelling long-term variability in extreme waves and sea states as well as a review of alternative approaches from other areas of application. The paper is organized as follows: Section 2 outlines alternative sources of wave data, Sect. 3 comprises a review of statistical models for extreme waves, Sect. 4 presents a review of relevant spatio-temporal statistical models from other areas of application, Sect. 5 reviews projections of future wave climate and Sect. 6 concludes with some recommendations for further research. An abbreviated version of this work was presented at the OMAE conference this year (Vanem 2010).

Efforts have been made to include all relevant and important work to make this literature survey as complete as possible, and this has resulted in a rather voluminous list of references at the end of the paper. Notwithstanding, due to the enormous amount of literature in this field some important works might inevitably have been omitted. This is unintended and it should be noted that important contributions to the discussion herein might exist of which I have not been aware. Nevertheless, it is believed that this literature study contains a fair review of relevant literature and as such that it gives a good indication of state-of-the-art within the field and may serve as a basis for further research on stochastic modelling of extreme waves and sea states.

1.1 Integrated sea state parameters

The state of the sea changes constantly, and it is therefore neither very practical nor very useful to describe the sea for an instantaneous point in time. Therefore, sea states are normally described by different averages and extreme values for a certain period of time, often referred to as integrated sea state parameters. Typically, such integrated parameters include the significant wave height,¹ mean wave period, mean main wave direction, spread of the wave direction and mean swell. Such integrated wave parameters represent averages over a defined period of time, typically in the order of 20–30 min.

¹ Significant waveheight, denoted SWH, H_S or $H_{1/3}$ is often defined as the average wave height, from trough to crest, of the one-third largest waves that is observed during the period.

Integrated wave parameters, which are averages over different periods of time, will have its own averages and extremes. Of particular interest may be the m -year return value of the significant wave height, SWH_m , which is defined as the value of H_S that is exceeded on average once every m years. In ship design, the SWH_{20} has traditionally been of particular interest since ships are normally designed for a lifetime of 20 years. The modelling of such extreme values, for example for the significant wave height, is therefore of interest.

It is also of interest to investigate how such average wave parameters vary over time. In particular, long term variations (i.e. how these parameters will vary in the next 50–100 years) will be an important basis for design of marine and offshore structures with expected lifetimes in the range of several decades and also for maritime risk analyses. This is of particular importance at times where climate change indicates that the future is not well represented by today's situation (i.e. where an increase in extreme weather and sea state is expected).

1.2 Waves as stochastic processes

Although the dynamics of the sea and the mechanisms underlying the generation of waves on the sea surface inevitably follows the laws of physics and therefore, in principle, the sea state could be described deterministically, in reality this is not possible due to the complexity of the system. Hence, the description of waves and the sea must be done probabilistically. The sea is a dynamic system that is influenced by innumerable factors and an infinite number of interrelated parameters would be needed in order to provide an exact description of the sea in any given point in time. It is simply not possible to know all and every one of these parameters. The unknown parameters introduce uncertainties to any description of the system and an exact description of the sea is therefore not feasible. Thus, the problem of describing the sea turns into a statistical problem, and probabilistic models are needed in order to represent waves on the sea surface and to provide a better understanding of the maritime environment in which ships operate. In this regard stochastic models would seem to be the most appropriate approach to describe extreme waves. Also, the fact that the sea state is normally described through different average and extreme properties, as discussed briefly above, indicates that statistical tools are appropriate to model waves and sea states. A comprehensive overview of statistical techniques, methodologies, theories and tools used in climatic analyses is presented in von Storch and Zwiers (1999).

Stochastic modelling of ocean waves can be performed on two very different time scales. In the short-term models, the parameters of most concern are those for individual

waves such as individual wave height, wave length and period, etc. The times involved in such models are normally in the order from a few seconds to a couple of hours. The long-term models mainly refer to the description of spectral parameters, and the times that are involved normally span over many years. It is the latter time scales that are of main interest in the present work, considering modelling of possible long-term trends due to climate change.

1.3 Predicting the impact of climate change on extreme sea states

The state of the oceans and the characteristics of the waves are influenced by innumerable external factors, and the most influential boundary conditions are related to the atmosphere and the global and local climate in general. Atmospheric pressure, wind, temperature, precipitation, solar radiation and heat, tidal movements, the rotation of the earth and movements of the seabed (e.g. from earthquakes or volcanic activities) are examples of external factors that jointly influence the generation of waves on the sea surface. In one sense, some of the average and extreme properties of the sea state can be regarded as stationary if the overall average boundary conditions does not change. That is, in spite of the continuous variations of sea states over time, the averages such as seasonal average wave heights and return periods for extreme waves can be considered as stationary if the average boundary conditions (e.g. average atmospheric pressure, average wind, average temperatures, etc.) remain stationary.

However, in recent years it has become increasingly apparent that the climate system overall is not stationary and that the climate will change in the near future—in fact it has been observed that the climate is already undergoing a change with a global long-term trend towards higher temperatures and more frequent and intense severe weather events, although local and regional trends may differ from this global trend. These climate changes—man-made or not—will thus change the overall boundary conditions for the sea, and the assumption that the average sea states can be regarded as stationary ceases to be valid.

In order to predict future trends in sea state parameters in the non-stationary case, one may therefore start with predicting the trends in the boundary conditions such as temperature, atmospheric pressure and wind. Assuming that a significant part of the climate change is man-made and can be ascribed to the increasing emission of greenhouse gases, most notably CO₂, and aerosols, predictions of climate change can be made based on various emission scenarios or forcing scenarios (Nakićenović et al. 2000). These forcing scenarios can then be fed into climate models to predict global trends in meteorological variables,

which can again be used to predict trends in average and extreme properties of sea waves. However, most wave models are deterministic and not able to handle the inherent uncertainties involved in a rigorous manner.

Estimates of future H_S return values are difficult since there are no projections of future H_S fields. However, projections of sea level pressure provided by climate models are reasonable reliable and it is known that the H_S fields are highly correlated with sea level pressure fields. Therefore, one approach could be to model H_S fields by regressing on projected sea level pressure fields, as was done in Wang et al. (2004). Other covariates may also be used to predict changes in extreme wave climate from projected changes in the overall climate, and the utilization of such dependencies may prove important in modelling long-term trends in extreme waves.

2 Wave data and data sources

As in all statistical modelling, a crucial prerequisite for any sensible modelling and reliable analysis is the availability of statistical data. For example, if models describing the spatio-temporal variability of extreme waves are to be developed, wave data with sufficient spatio-temporal resolution is needed. Furthermore, the lack of adequate coverage in the data will restrict the scope of the statistical models that can be used.

Wave data can be obtained from buoys, laser measurements, satellite images, shipborne wave recorders or be generated by numerical wave models. Of these, buoy measurements are most reliable, but the spatial coverage is limited. For regions where buoy data are not available, satellite data may be an alternative for estimation of wave heights (Krogstad and Barstow 1999; Panchang et al. 1999), and there are different satellites that collect such data. Examples of satellite missions are the European Remote Sensing Satellites (ERS-1 and ERS-2), the Topex/Poseidon mission and Jason-1 and -2 missions.

Wave parameters derived from satellite altimeter data were demonstrated to be in reasonable agreement with buoy measurements by the end of last century (Hwang et al. 1998). More recently, further validation of wave heights measured from altimeters have been performed, and the agreement with buoy data is generally good (Queffelec 2004; Durrant et al. 2009). However, corrections due to biases may be required, and both negative and positive biases for the significant wave height have been reported, indicating that corrections are region-dependent (Meath et al. 2008). Sea state parameters such as significant wave height derived from synthetic aperture radar images taken from satellites were addressed in Lehner et al. (2007).

Ship observations are another source of wave data which covers areas where buoy wave measurements are not available. The Voluntary Observing Ship (VOS) scheme has been in operation for almost 150 years and has a large set of voluntary collected data. However, due to the fact that ships tend to avoid extreme weather whenever possible, extreme wave events are likely to be under-represented in ship observations and hence such data are not ideally suited to model extreme wave events (DelBalzo et al. 2003; Olsen et al. 2006).

Recently, a novel wave acquisition stereo system (WASS) based on a variational image sensor and video observational technology in order to reconstruct the 4D dynamics of ocean waves was developed (Fedele et al. 2009). The spatial and temporal data provided by this system would be rich in statistical content compared to buoy data, but the availability of such data are still limited.

In general, measurements of wave parameters are more scarce than meteorological data such as wind and pressure fields which are collected more systematically and covering a wider area. An alternative is therefore to use output from wave models that uses meteorological data as input rather than to use wave data that are measured directly.

Wave models are normally used for forecast or hindcast of sea states (Guedes Soares et al. 2002). Forecasts typically predicts sea states up to 3–5 days ahead. Hindcast modelling can be used to calibrate the models after precise meteorological measurements have been collected. It can also be used as a basis for design but it is stressed that quality control is necessary and possible errors and biases should be identified and corrected (Bitner-Gregersen and de Valk 2008).

Currently, data are available from various reanalysis projects (Caires et al. 2004). For example, 40 year of meteorological data are available from the NCEP/NCAR reanalysis project (Kalnay et al. 1996) that could be used to run wave models (Swail and Cox 2000; Cox and Swail 2001). A more recent reanalysis project, ERA-40 (Uppala et al. 2005), was carried out by the European Centre for Medium-Range Weather Forecasts (ECMWF) and covers a 45-year period from 1957 to 2002. The data contain six-hourly fields of global wave parameters such as significant wave height, mean wave direction and mean wave period as well as mean sea level pressure and wind fields and other meteorological parameters. A large part of this reanalysis data are freely available for download from their website for research purposes.²

It has been reported that the ERA-40 dataset contains some inhomogeneities in time and that it underestimates high wave heights (Sterl and Caires 2005), but corrected datasets for the significant wave height have been produced

(Caires and Sterl 2005). Hence, a new 45-year global six-hourly dataset of significant wave height has been created, and the corrected data shows clear improvements compared to the original data. In Caires and Swail (2004) it is stated that this dataset can be obtained freely from the authors for scientific purposes.

3 Review of statistical models for extreme waves

In order to model long-term trends in the intensity and frequency of occurrence of extreme wave events or extreme sea states due to climate change, appropriate models must be used. There are numerous stochastic wave models proposed in the literature, but most of these are developed for other purposes than predicting such long-term trends. Models used for wave forecasting, for example in operational simulation of safety of ships and offshore structures typically have a short-term perspective, and cannot be used to investigate long-term trends. Also, many wave models assume stationary or cyclic time series, which would not be the case if climate change is a reality.

There are different approaches to estimating the extreme wave heights at a certain location based on available wave data, and some of the most widely used are the initial distribution method, the annual maxima method, the peak-over-threshold method and the MEan Number of Up-crossings (MENU) method. The initial distribution method uses data (measured or calculated) of all wave heights and the extreme wave height of a certain return period is estimated as the quantile h_p of the wave height distribution $F(h)$ with probability p . The annual maxima approach uses only the annual (or block) maxima and the extreme wave height will have one of the three limit distributions referred to as the family of the generalized extreme value distribution. The peak-over-threshold approach uses data with wave heights greater than a certain threshold, and thus allows for increased number of samples compared to the annual maxima approach. Waves exceeding this threshold would then be modelled according to the Generalized Pareto distribution. However, the peaks-over-threshold method has demonstrated a clear dependence on the threshold and is therefore not very reliable. The MENU method determines the return period of an extreme wave of a certain wave height by requiring that the expected or mean number of up-crossings of this wave height will be one for that time interval.

Another approach useful in extreme event modelling is the use of quantile functions, an alternative way of defining a probability distribution (Gilchrist 2000). The quantile function, Q , is a function of the cumulative probability of a distribution and is simply the inverse of the cumulative density function: $Q(p) = F^{-1}(p)$ and $F(x) = Q^{-1}(x)$. This function can then be used in frequency analysis to find

² Data available from url: http://data-portal.ecmwf.int/data/d/era40_daily/.

useful estimates of the quantiles of relevant return periods T of extreme events in the upper tail of the frequency distribution, $Q_T = Q(1 - 1/T)$.

Yet another approach for estimating the maxima of a stationary process is to model the number of extreme events, defined as the number of times the process crosses a fixed level u in the upward direction, as a Poisson process (a counting process $\{N(t), t \geq 0\}$ with $N(0) = 0$, independent increments and with number of events in a time interval of length t Poisson distributed with mean λt is said to be a Poisson process with rate λ) and apply the Rice formula to compute the intensity of the extreme events (see e.g. Rychlik 2000).

In the following, a brief review of some wave models proposed in the literature will be given. This includes a brief description of some short-term and stationary wave models as well as a more comprehensive review of proposed approaches to modelling long term trends due to global climatic changes. An introduction to stochastic analysis of ocean waves can be found in Ochi (1998) and Trulsen (2006), albeit the latter with a particular emphasis on freak or rogue waves.

3.1 Short-term stochastic wave models

Waves are generated from wind actions and wave predictions are often based on knowledge of the generating wind and wind-wave relationships. Most wave models for operational wave forecasting is based on the energy balance equation; there is a general consensus that this describes the fundamental principle for wave predictions, and significant progress have been made in recent decades (Janssen 2008). Currently, the third-generation wave model WAM is one of the most widely used models for wave forecasting (The WAMDI Group 1988; Komen et al. 1994) computing the wave spectrum from physical first principles. Other widely used wave models are Wave Watch and SWAN, and there exist a number of other models as well (The Wise Group et al. 2007). However, wave generation is basically an uncertain and random process which makes it difficult to model deterministically, and in Deo et al. (2001), Bazargan et al. (2007) approaches using neural networks were proposed as an alternative to deterministic wave forecasting models.

There are a number of short-term, statistical wave models for modelling of individual waves and for predicting and forecasting sea states in the not too distant future. Most of the models for individual waves are based on Gaussian approaches, but other types of stochastic wave models have also been proposed to account for observed asymmetries (e.g. adding random correction terms to a Gaussian model (Machado and Rychlik 2003) or based on Lagrangian models (Lindgren 2006; Aberg and Lindgren 2008)). Asymptotic models for the distribution of maxima

for Gaussian processes for a certain period of time exist, and under certain assumptions, the maximum values are asymptotically distributed according to the Gumbel distribution. However, as noted in Rydén (2006), care should be taken when using this approximation for the modelling of maxima of wave crests. A similar concern was expressed in Coles et al. (2003), albeit not related to waves.

Given the short-term perspective of these types of models, they cannot be used to describe long-term trends due to climate change, nor to formulate design criteria for ships and offshore structures, even though they are important for maritime safety during operation. Improved weather and wave forecasts will of course improve safety at sea, but the main interest in the present study is on long-term trends in ocean wave climate, and the effect this will have on maritime safety and on the design of marine structures. Therefore, short-term wave models will not be considered further herein.

3.1.1 Significant wave height as a function of wind speed

The significant wave height for a fully developed sea, sometimes referred to as the equilibrium sea approximation, given a fixed wind speed have been modelled as a function of the wind speed in different ways, for example as $H_S \propto U^{5/2}$ or $H_S \propto U^2$ (Kinsman 1965). This makes it possible to make short-term predictions of the significant wave height under the assumptions of a constant wind speed and assuming unlimited fetch and duration. For developing sea conditions, with limited fetch or limited wind duration, the significant wave height as a function of wind speed, U (m/s) and respectively fetch X (km) and duration D (h) has been modelled in different ways, for example as $H_S \sim X^{1/2}U$ and $H_S \propto D^{5/7}U^{9/7}$ (Özger and Şen 2007).

However, it is observed that the equilibrium wind sea approximation is seldom valid, and an alternative model for predicting the significant wave height for wind waves, H_S from the wind speed U_{10} at a reference height of 10 m were proposed in Andreas and Wang (2007), using a different, yet simple parametrization. 18 years of hourly data of significant wave height and winds speed for 12 different buoys were used in order to estimate the model which can be written on the following form:

$$H_S = C(D)I(U_{10} \leq 4m/s) + [a(D)U_{10}^2 + b(D)]I(U_{10} > 4m/s) \quad (1)$$

D denotes the water depth and C , a and b are depth-dependent parameters. Based on comparison with measurements it was concluded that this model is reliable for wind speeds up to at least $U_{10} = 25$ m/s.

It is out of scope of the present literature survey to review all models for predicting wave heights from wind speed or other meteorological data. Such models are an integral part of the various wave models available for wave forecasting, but cannot be used directly to model long-term variations in wave height. However, given adequate long-term wind forecasts, such relationships between wind speed and wave height may be exploited in simulating long-term wave data for long-term predictions of wave climate.

3.2 Stationary models

A thorough survey of stochastic models for wind and sea state time series is presented in Monbet et al. (2007). Only time series at the scale of the sea state have been considered without modelling events at the scale of individual waves, and only at given geographical points. One section of Monbet et al. (2007) is discussing how to model non-stationarity such as trends in time series and seasonal components, but for the main part of the paper it is assumed that the studied processes are stationary. The models have been classified in three groups: Models based on Gaussian approximations, other non-parametric models and other parametric models. In the following, the main characteristics for these different types of wave models are highlighted.

Even though ocean wave time series cannot normally be assumed to be Gaussian, it may be possible to transform these time series into time series with Gaussian marginal distributions when they have a continuous state space (Monbet et al. 2007). The transformed time series can then be simulated by using existing techniques to simulate Gaussian processes. If $\{Y_t\}$ is a stationary process in \mathbf{R}^d , assume that there exists a transformation $f: \mathbf{R}^d \rightarrow \mathbf{R}^d$ and a stationary Gaussian process $\{X_t\}$ so that $Y_t = f(X_t)$. Such a procedure consists of determining the transformation function f , generation of realizations of the process $\{X_t\}$ and then transforming the generated samples of $\{X_t\}$ into samples of $\{Y_t\}$ using f . A number of such models for the significant wave height have been proposed in the literature (e.g. Cunha and Guedes Soares (1999), Walton and Borgman (1990) for the univariate time series for significant wave height, H_s , Guedes Soares and Cunha (2000), Monbet and Prevosto (2001) for the bivariate time series for significant wave height and mean wave period, (H_s, T) and DelBalzo et al. (2003) for the multivariate time series for significant wave height, mean wave period and mean wave direction, (H_s, T, Θ_m)). However, it is noted that the duration statistics of transformed Gaussian processes has been demonstrated not to fit too well with data, even though the occurrence probability is correctly modelled (Jenkins 2002).

Multimodal wave models for combined seas (e.g. with wind-sea and swell components) have also been discussed in the literature (see e.g. Torsethaugen 1993; Torsethaugen

and Haver 2004; Ewans et al. 2006), but these are generally not required to describe severe sea states where extremes occur (Bitner-Gregersen and Toffoli 2009).

A few non-parametric methods for simulating wave parameters have been proposed, as reported in Monbet et al. (2007). One may for example assume that the observed time series are Markov chains and use non-parametric methods such as nearest-neighbor resampling to estimate transition kernels. In Caires and Sterl (2005), a non-parametric regression method was proposed to correct outputs of meteorological models. A continuous space, discrete time Markov model for the trivariate time-series of wind speed, significant wave height and spectral peak period was presented in Monbet and Marteau (2001). However, one major drawback of non-parametric methods is the lack of descriptive power.

An approach based on copulas for multivariate modelling of oceanographic variables, accounting for dependencies between the variables, were proposed in de Waal and van Gelder (2005) and applied to the joint bivariate description of extreme wave heights and wave periods.

Parametric models for wave time series include various linear autoregressive models, nonlinear retrogressive models, finite state space Markov chain models and circular time series models. A modified Weibull model was proposed in Muraleedharan et al. (2007) for modelling of significant and maximum wave height. For short-term modelling of wave parameters, different approaches of artificial neural networks (see e.g. Deo et al. 2001; Mandal and Prabakaran 2006; Arena and Puca 2004; Makarynsky et al. 2005) and data mining techniques (Mahjoobi and Etemad-Shahidi 2008; Mahjoobi and Mosabbebeh 2009) have successfully been applied. A non-linear threshold autoregressive model for the significant waveheight was proposed in Scotto and Guedes Soares (2000).

3.3 Non-stationary models

Many statistical models for extreme waves assume the stationarity of extreme values, but there are some non-stationary models proposed in the literature. In the following, some non-stationary models for extreme waves that are known and previously presented in the literature will be reviewed. A review of classical methods for asymptotic extreme value analysis used in extreme wave predictions are presented in Soukissian and Kalantzi (2006).

3.3.1 Microscopic models

A number of statistical models have been presented in the literature where the focus has been to use sophisticated statistical methods to estimate extreme values at certain specific geographical points (e.g. based on data measurements at that

location). This approach is natural, given the limited spatial resolution of available wave data, and aims at exploiting available data measurements at certain locations to the maximum, i.e. to obtain as good predictions as possible for locations where wave data are available. In the following, some of these will be briefly reviewed, even though it is noted that the aim of this study is to extend the scope and broaden the perspective of the statistical models to also include the spatial dimension.

A method for calculating return periods of various levels from long-term nonstationary time series data of significant wave height based on a new definition of the return period is presented in Stefanakos and Monbet (2006) and Stefanakos and Athanassoulis (2006). This definition is based on the mean number of upcrossings of a particular level and was first introduced in the context of prediction of sea-level extremes in Middleton and Thompson (1986). In Soukissian and Kalantzi (2007) and Guedes Soares and Scotto (2004), new de-clustering methods and filtering techniques are proposed in order to apply the r -largest-order statistics for long-term predictions of significant wave height. A new de-clustering method was also suggested in Soukissian et al. (2006) for applying the peaks-over-threshold method for H_s time series. An approach using stochastic differential equations for clarifying the relationship between long-term time-series data and its probability density functions in order to extrapolate long-term predictions from shorter historical data is proposed in Minoura and Naito (2006). Two approaches for estimating long term extreme waves are discussed in Hagen (2009) (i.e. an initial distribution approach and a Peak Over Threshold (POT) approach for storm events) and issues related to sampling variability, model fitting and threshold selection (for the POT analysis) are addressed.

Duration statistics of long time series of significant wave height H_s (i.e. the duration of sea states with different intensities) were analyzed in Soukissian and Samalekos (2006) using a bottom-up segmentation algorithm. This analysis makes use of the increasing or decreasing intensity of successive sea state conditions, and subdivide long-term H_s time series into subsequent series of monotonically increasing or decreasing intensities. This would correspond to developing and decaying sea states, and the segmentation algorithm should ensure that a meaningful subdivision of the long-term time series are obtained. A sensitivity analysis of this approach, investigating the effect of the maximum allowed error on the segmentation of the H_s time series is reported in Soukissian and Photiadou (2006).

Return periods of storms with an extreme wave above a certain threshold are found based on an equivalent triangular storm model in Arena and Pavone (2006). This approach is extended to find return periods analytically for storms with two or more waves exceeding the threshold in Arena et al. (2009), Arena and Pavone (2009). The basic

idea behind the equivalent triangular storm model is that it, for a fixed location, associates a triangle to each actual storm and represents a significant wave height time series by means of a sequence of triangular storms. The triangle height is the maximum significant wave height during the actual storm and the triangle base is such that the maximum expected wave height in the actual storm equals the maximum expected wave height in the triangular storm model (Bocotti 2000). The equivalent power storm model was presented in Fedele and Arena (2009) as a generalization of the equivalent triangular storm model to predict return periods for waves above a certain threshold. It is noted that the equivalent triangular storm is firmly based on what has become known as the Borgman Integral (Borgman 1973), which gives the distribution function for the largest wave, $F_m(h) = P(H_m \leq h)$ as follows, with H_m denoting the largest wave height, $a^2(t)$ time varying Rayleigh parameter and $T(t)$ typical wave period at time t :

$$F_m(h) = e^{\int \log[1 - e^{-h^2/a^2(t)}] \frac{dt}{T(t)}} \quad (2)$$

A nonstationary stochastic model for long-term time series of significant wave height was presented in Athanassoulis and Stefanakos (1995) which were modelled by decomposing detrended time series to a periodic mean value and a residual time series multiplied with a periodic standard deviation: $X(\tau) = \bar{X}_{\text{trend}}(\tau) + \mu(\tau) + \sigma(\tau)W(\tau)$. It was then showed that $W(\tau)$ could be considered stationary. Short-term and long-term wave characteristics of ocean waves were combined in order to develop nested, stochastic models for the distribution of maximum wave heights in Prevesto et al. (2000). Different time scales were introduced, i.e. fast time and slow time, and a stochastic process was modelled in the fast time where the state variables were modelled as a stochastic process in the slow time.

The seasonal effect on return values of significant wave height were investigated in Menéndez et al. (2009), where a time-dependent generalized extreme value model was used for monthly maxima of significant wave height. Non-stationarity representing annual and semiannual cycles is introduced in the model via the location, scale and shape parameters and the inclusion of seasonal variabilities is found to reduce the residuals of the fitted model substantially. Hence, the model provides a way of quantitatively examining the long-term seasonal distribution of significant wave height.

Various other models for the long-term distribution of significant wave height have been suggested (e.g. using the Beta and Gamma models (Ferreira and Guedes Soares 1999), using the Annual Maxima and Peak Over Threshold methods (Guedes Soares and Scotto 2001) using non-linear threshold models (Scotto and Guedes Soares 2000), using time-dependent Peak Over Threshold models for the intensity

combined with a Poisson model for frequency (Méndez et al. 2006, 2008), employing different autoregressive models (Guedes Soares and Ferreira 1996; Guedes Soares et al. 1996), and using a transformation of the data and a Gaussian model for the transformed data (Ferreira and Guedes Soares 2000)). Short- and long-term statistics were combined in Krogstad (1985) in order to establish distributions of maximum wave heights and corresponding periods. Some considerations of bias and uncertainty in methods of extreme value analysis were discussed in Gibson et al. (2009), leading to some recommended approaches for such analyses and applied on a set of wave data.

More recently, an interesting approach to long-term predictions of significant wave height, combining Bayesian inference methodology, extreme value techniques and Markov Chain Monte Carlo (MCMC) procedures is presented in Scotto and Guedes Soares (2007). The benefits of using a Bayesian approach compared to a traditional likelihood-based approach is that prior knowledge about parameter values θ can be used together with observed data x to update a posterior distribution $\pi(\theta|x)$. Simulations of this posterior distribution can be obtained by constructing a Markov Chain whose invariant distribution, or target distribution, is proportional to the posterior distribution by employing the Metropolis-Hastings algorithm (see Robert and Casella 2004). This Bayesian approach was used to analyze a dataset of significant wave height collected in the northern North Sea.

Another Bayesian approach to estimating posterior distributions of return periods for extreme waves is proposed in Egozcue et al. (2005). Here, the occurrence of extreme events is modelled as a Poisson-process with extreme wave heights distributed according to a generalized Pareto distribution.

3.3.2 Combining long and short term wave height statistics

The Borgman Integral (Eq. 2) is a fundamental tool for combining the long term distribution of significant wave height with short term distribution for the individual wave heights (Borgman 1973). This is often desired for estimating the maximum wave or crest height occurring in a long return interval. A similar method was proposed by Battjes (1972). It is noted that the particular expression of the Borgman Integral as presented in Eq. 2 is based on the assumption of a Rayleigh distribution for the individual wave height. A more general form would be, letting $P(h|H_s)$ denote the short-term distribution of the individual wave height conditioned on the sea state,

$$F_m(h) = e^{-\int \log[P(h|H_s)] \frac{dH_s}{H_s}} \quad (3)$$

Long time series of individual wave heights are typically not available and calculations must therefore be based on

time series of sea state parameters such as the significant wave height. Hence, the problem of modelling the maximum wave height in a long time interval comprises three aspects: modelling of long term sea state parameters (e.g. significant wave height), short term modelling of individual wave heights conditioned on the sea state and combining the two distributions. This can be done by first fitting a short-term distribution and then apply the Borgman Integral to this distribution. Integration of short-term second order models over time series of measured sea states was performed by Krogstad and Barstow (2004). A recent study concerned with finding the most accurate method for combining long and short term wave statistics was reported in Forristall (2008).

3.3.3 Spatio-temporal models for extreme waves

The spatial-temporal variability of ocean wave fields are complex, and the fields will generally be inhomogeneous in space and non-stationary in time, with strong temporal and spatial variation (Jönsson et al. 2002). Different models have been proposed in the literature for modelling these variabilities and for analyzing and synthesizing spatio-temporal wave data.

There has been significantly more focus on the temporal variability compared to the spatial variability of wave fields, but the spatial behavior (i.e. the spatial interdependence and radius of influence of a set of spatially distributed stations) of significant wave height is investigated in Altunkaynak (2005). The methodology is based on the concept of trigonometric point cumulative semivariograms, consisting of cumulative broken lines where the angle between two successive lines connecting two station records is a measure of the regional dependence, ranging from 0 (complete independence) to 1 (complete dependence). Another approach for predicting the maximum wave height over a spatial area were proposed in Fedele et al. (2009), based on 4D video data of sea states acquired through a wave acquisition stereo system (WASS) and using Euler Characteristics' theory. A regional frequency analysis of extreme wave heights, analyzing peaks-over-threshold wave data from nine locations along the Dutch North Sea coast was reported in Van Gelder et al. (2001). The different locations could be considered as a homogeneous region and it was shown that the Generalized Pareto Distribution is an optimal regional probability distribution for the extreme wave heights for the region. Notable differences were found for the regional quantile estimates compared to the at-site quantile estimates, indicating that it would be better to rely on the regional estimates in decision making.

Models for stochastic simulation of the annual (Boukhanovsky et al. 2003a) and synoptic (Boukhanovsky et al.

2003b) variability of inhomogeneous metocean fields were proposed as expansions of the field $\zeta(r, t)$ in terms of periodical empirical orthogonal functions in Boukhanovsky et al. (2003a, b):

$$\zeta(\mathbf{r}, t) = m(\mathbf{r}, t) + \sum_k a_k(t) \phi_{kt}(\mathbf{r}, t) + \epsilon(\mathbf{r}, t) \tag{4}$$

where $m(\mathbf{r}, t)$ are the mathematical expectations, $\phi_{kt}(\mathbf{r}, t)$ are the spatio-temporal basis functions, $\epsilon(k, t)$ are inhomogeneous white noise and $a_k(t)$ are the coefficients. \mathbf{r} denotes the geographical coordinates and t time. The results of simulating these models is a set of simulated metocean fields $\zeta(r, t)$ in a discrete set of grid points and at discrete times. They could then be used to investigate the field extremes and rare events in terms of both spatial and temporal extremes, and wave data from the Barents Sea region have been used to test the models with reasonable agreement. The stochastic models for annual variability can be regarded as field generalizations of periodically correlated stochastic processes. The model for synoptic variability uses a Lagrangian approach and the temporal sequence of storm centres are modelled as a finite-state Markov Chain with the storm extensions and field properties as spatio-temporal impulses.

Recently, spatio-temporal statistical models for the significant wave height have been reported that describes the variability of significant wave height over large areas by stochastic fields (Baxevani et al. 2005, 2009). This is based on constructing a homogeneous model valid for a small region and then extending this to a non-homogeneous model valid for large areas. Global wave measurements from satellites have been used for model fitting, providing wave data of spatial variability, but limited physical knowledge about the wave phenomena have been incorporated into the models. The resulting models can then be used to estimate the probability of a maximum significant wave height to exceed a certain level or to estimate the distribution of the (spatial) length of a storm (Baxevani et al. 2007). However, the temporal validity of this model is limited to the order of hours (Baxevani et al. 2009), and therefore it does not seem suited for studying long-term trends and the effects of climate change.

Caires et al. (2006b) used two approaches to model the extremes of non-stationary time series, i.e. the non-homogeneous Poisson process and a non-stationary generalized extreme value model. The non-homogeneous Poisson process was used to model extreme values of the significant wave height, obtained from the 40-year ECMWF re-analysis (ERA-40) (Uppala et al. 2005) and compared to estimates obtained using a nonstationary generalized extreme value model (NS-GEV). The parameter of the Poisson distribution in this model was on the form $\lambda = \iint \lambda(t, x) dt dx$, where:

$$\lambda(t, x) = \frac{1}{\sigma(t)} \left[1 + \xi(t) \frac{x - \mu(t)}{\sigma(t)} \right]_+^{-(\frac{1}{\xi(t)})-1} \tag{5}$$

From projections of the sea level pressure under three different forcing scenarios (Nakićenović et al. 2000; Boer et al. 2000), projections of the parameters in the non-homogeneous Poisson process are made up to the end of the twenty-first century. Trends in these parameters are then determined, projections of return value estimates of H_5 are projected and their uncertainties are assessed.

4 Relevant statistical models from other areas of application

Extreme value analysis has a wide area of applications aside from ocean waves, in particular in various environmental sciences where events are also associated with spatio-temporal variations, and it is believed that some lessons can be learned by examining different statistical models for other types of extreme events.

An interesting discussion on the use of asymptotic models for the description of the variation of extremes is available in Coles et al. (2003), within the context of extreme rainfall modelling. It is concerned with the lack of ability of such models to predict extreme, catastrophic events leading to inadequate designs and lack of preparedness for such rare events. One of the reasons for this, according to Coles et al. (2003) is models that do not take the uncertainties in both model and predictions adequately into account. For example, it is argued that even in cases where data support the reduction of the generalized extreme value model to a Gumbel model, this should not be done without an appraisal of the uncertainty this decision introduces and as a general advice it is suggested to use the generalized extreme value model rather than Gumbel reduction. Furthermore, the preference for Bayesian analysis over the classical likelihood analysis is emphasized, even if using diffuse priors.

In this section, a review of relevant time- and space-dependent statistical models from other areas of application is presented. Further work will then focus on how these approaches can be used for statistical modelling of extreme waves and sea states.

4.1 Bayesian hierarchical space–time models

Modelling of wave data in space and time is an alternative to the common approach of extreme value analysis based on a point process representation, provided that adequate space–time wave data can be obtained. Wikle et al. (1998) propose a hierarchical Bayesian space–time model as an alternative to traditional space–time statistical models and

apply it on an atmospheric data set of monthly maximum temperatures. Such models generally consist of three stages often referred to as the data stage, the process stage and the parameter stage.

Similar models has also been used for modeling tropical ocean surface winds (Wikle et al. 2001), North Atlantic sea surface temperatures (Lemos and Sansó 2009), concentrations of PM₁₀ pollution³ (Cocchi et al. 2007), ozone levels (Sahu et al. 2007) and earthquake data (Natvig and Tvette 2007) A brief overview of hierarchical approaches applied to environmental processes is presented in Wikle (2003). More recently, various hierarchical Bayesian space–time models for extreme precipitation events were proposed in Sang and Gelfand (2009). As a proxy for these Bayesian hierarchical space–time models, the model for earthquake data will be briefly reviewed in the following.

The modelling of earthquake data for earthquake prediction, using a Bayesian hierarchical space–time approach in Natvig and Tvette (2007) considered a spatial resolution of 0.5 × 0.5° (about 50 × 50 km) and a temporal resolution of 4 months. The observations are, for each time period, the magnitude of the largest earthquakes (by the Richter scale) observed within each grid. The model is implemented within a Markov Chain Monte Carlo framework using Gibbs sampling and additional Metropolis-Hastings steps. Four different model alternatives were suggested in a hierarchical structure, one main model and three levels of simplified models, nested within the model at the higher level. The main features of the main model are briefly outlined in the following.

Denoting the discrete spatial and temporal locations $x = 1, \dots, X$ and $t = 1, \dots, T$, respectively, the observed maximal magnitude earthquake at location x and time t is $M(x, t)$. The latent variable $Y(x, t)$ is defined in order to describe the maximal earthquake without a cut point at 0:

$$M(x, t) = \begin{cases} Y(x, t) & \text{if } Y(x, t) > 0 \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

Corresponding to the latent variable $Y(x, t)$, for every point in space–time (x, t) there is an underlying potential for a maximal magnitude earthquake, modelled by the hidden system state variables $\theta(x, t)$. The latent variable is then modelled as this potential and a random noise term as follows:

$$Y(x, t) = \theta(x, t) + \epsilon_Y(x, t), \quad \begin{matrix} x = 1, \dots, X, \\ t = 1, \dots, T \end{matrix} \quad (7)$$

where the random noise are assumed independent and normally distributed $\epsilon_Y(x, t) \sim N(0, \sigma_Y^2)$, σ_Y^2 being a random quantity. Furthermore, the earthquake potential is

assumed to be decomposed into a time-independent contribution $\mu(x)$ and a time-dependent distribution with a spatial description $\theta_S(x, t)$:

$$\theta(x, t) = \mu(x) + \theta_S(x, t), \quad \begin{matrix} x = 1, \dots, X, \\ t = 1, \dots, T \end{matrix} \quad (8)$$

Now, the time-independent term is modelled as a Gaussian Markov Random Field, with spatial dependence only on its nearest neighbours, with the notation N = North, E = East, S = South, W = West, and e.g. $\mu(x)^N$ the μ value in the grid cell north of x :

$$\begin{aligned} \mu(x) = & \mu_0(x) + c_{NS} \{ \mu(x)^N - \mu_0(x)^N + \mu(x)^S - \mu_0(x)^S \} \\ & + c_{EW} \{ \mu(x)^E - \mu_0(x)^E + \mu(x)^W - \mu_0(x)^W \} \\ & + \epsilon_\mu(x), \quad x = 1, \dots, X \end{aligned} \quad (9)$$

c_{NS} , c_{EW} are spatial dependence parameters in the north–south and east–west directions respectively. The noise terms are again normally distributed with variance σ_μ^2 .

In the expression above, $\mu_0(x)$ denote the Markov Random Field mean in grid point x and this is modelled as having a quadratic form, letting $m(x)$ and $n(x)$ denote longitude and latitude corresponding to the grid point x , in the following way (for $x = 1, \dots, X$):

$$\begin{aligned} \mu_0(x) = & \mu_0[1] + \mu_0[2]m(x) + \mu_0[3]n(x) \\ & + \mu_0[4](m(x))^2 + \mu_0[5](n(x))^2 + \mu_0[6]m(x)n(x) \end{aligned} \quad (10)$$

The space–time dynamic term are modelled with a vector autoregressive model of order 1, with spatial dependence only on its nearest neighbours (with same notation for North, South, etc. as above, and for $x = 1, \dots, X, t = 1, \dots, T$):

$$\begin{aligned} \theta_S(x, t) = & (a(x)\theta_S(x, t - 1) + \beta(x)) \\ & \times e^{-\kappa(M(x,t-1)-3)^2 I(M(x,t-1) > 3)} \\ & + a_N\theta_S(x, t - 1)^N + a_E\theta_S(x, t - 1)^E \\ & + a_S\theta_S(x, t - 1)^S + a_W\theta_S(x, t - 1)^W + \epsilon_{\theta_S}(x, t) \end{aligned} \quad (11)$$

The exponential term in the equation above is included to account for strain which is built-up in periods with only small earthquakes and reduced when tension is released as a major earthquake occurs. κ is a positive parameter that regulates the reduction in strain according to the largest earthquake at the previous time period for a grid cell. $a(x)$ and $\beta(x)$ are modelled in completely the same way as $\mu(x)$ outlined above, introducing parameters a_{NS} , a_{EW} , b_{NS} and b_{EW} accordingly. The noise term $\epsilon_{\theta_S}(x, t)$ is Gaussian $N(0, \sigma_{\theta_S}^2)$ and independent in space and time. The various simplified, nested model alternatives are obtained by

- i. setting $\beta(x) = 0$
- ii. setting $e^{-\kappa(M(x,t-1)-3)^2 I(M(x,t-1) > 3)} = 1$ to remove the strain term

³ PM₁₀ is the fraction of aerosol particles with aerodynamic diameter less than 10 μm .

iii. setting $\theta_S(x, t) = 0$ to obtain a time-independent model

The model outlined above contains a large number of parameters, and all prior parameter distributions are considered independent. A Markov Chain Monte Carlo approach using the Gibbs sampler and also an additional Metropolis-Hastings step for some of the parameters, was adopted for generating independent samples from the posterior distributions in order to arrive at posterior estimates and predictions. For an introduction to Markov Chain Monte Carlo methods, including the Gibbs sampler and the Metropolis-Hastings algorithm, reference is made to Robert and Casella (2004) or similar textbooks.

4.2 Continuous space models

Even though wave data are generally only available at certain specific locations, extreme waves should in principle be considered as a continuous process in space and time rather than a discrete process. Considering the continuous space modelling of a process' extremes, this would require the specification of a continuous space model for the marginal behaviour of the extremes of the process and a continuous space specification of the dependence structure of the extremes. Hence, a generalization of the dependence structure of multivariate extremes to the infinite dimensional case is needed, and one such generalization is provided by the theory of max-stable processes (de Haan 1984). By definition, a stochastic process $\{Y_t\}$ is called a max-stable process if the following property holds:

If $\{Y_t^{(i)}\}_{t \in T}$, $i = 1, \dots, r$, are independent copies of the process then the process $\{\max_{i \leq r} Y_t^{(i)}\}_{t \in T}$ has the same distribution as $\{rY_t^{(1)}\}_{t \in T}$.

In the following, a procedure for using the theory of max-stable processes for modelling data which are collected on a grid of points in space are reviewed. This approach is considered as an infinite dimensional extension of multivariate extreme value theory and has the advantage that it can be used to aggregate the process over the whole region and for interpolation to anywhere within the whole region. Models based on the resulting family of multivariate extreme value distributions are suitable for a large number of grid points.

In Coles (1993), Coles and Tawn (1996) a class of max-stable process models for regional modelling of extreme storms were specified which can be estimated using all relevant extreme data and which are consistent with the multivariate extreme structure of the data. The essence of this approach is to describe the process of storms by the following components:

- i. A phase space S of storm types so that the storm type is independent of their size
- ii. An index space T for the region, conveniently referred to as the region itself
- iii. A measure $\nu(ds)$ on S describing the relative frequency of storm types
- iv. A function $f(s, t)$ interpreted as the proportion of a storm of type s observed at t

With x_j interpreted as the size of the j th storm, s_j the type of the j th storm, if $\{(x_i, s_i); i = 1, \dots\}$ are taken to be the points of a Poisson process on $(0, \infty) \times S$ with intensity $\mu(dx, ds) = x^{-2} dx \nu(ds)$ and letting $f(x, s)$ be a positive function on $S \times T$, then the process

$$Z_t = \max_i \{X_i f(S_i, t)\} \tag{12}$$

is a max-stable process for $t \in T$.

For statistical modelling of extreme storms as such a max-stable process, it was assumed that the spatial variability of storms could be described adequately by variability within a subset of data sites T_1 (Coles 1993). Then, a multivariate extreme value model was fitted to the data for this subset and the model is extended smoothly as a max-stable process through suitable functions $f(\cdot, \cdot)$ on the basis of information from the remaining data sites. Such a model was fitted for rainfall data collected from 11 sites, and in spite of some interesting qualitatively observations, the quality of fit of the model was rather poor.

In Buishand et al. (2008) a somewhat different approach of using max-stable processes for the modelling of spatial extreme rainfall is proposed based on random fields. Whereas Coles (1993), Coles and Tawn (1996) indicate how to analytically calculate quantiles of areal rainfall, in Buishand et al. (2008), the 100-year quantile of the total rainfall over an area in Holland is found by simulating synthetic daily rainfall fields using their estimated model. An extended Gaussian max-stable model for spatial extreme rainfall was also presented in Smith and Stephenson (2009), where Bayesian techniques are used in order to incorporate information other than data into the model, i.e. by using informative priors for the marginal site parameters and non-informative priors for parameters relating to the dependence structure of the process. The extended model is estimated using a pairwise likelihood within the Bayesian analysis and Markov chain Monte Carlo techniques were used to simulate from the posterior distributions, using a Gibbs sampler with a Metropolis step. Max-stable processes have also been applied to e.g. modelling of extreme wind speeds (Coles and Walshaw 1994).

4.3 Process convolution models

Several models for spatio-temporal processes based on process convolution have been proposed in the literature

(e.g. Higdon 1998; Calder 2007, 2008; Sansó et al. 2008). The main idea is to convolve independent processes to construct a dependent process by some convolution kernel. This kernel may evolve over space and time thus specifying models with non-stationary dependence structure.

The model proposed in Higdon (1998) is motivated by estimation of the mean temperature field in the North Atlantic Ocean based on 80 year of temperature data for a region. First, the temperature field $y(s, t)$ is modelled as a process over space s and time t as the sum of two processes

$$y(s, t) = z(s, t) + \epsilon(s, t) \quad (13)$$

where $z(s, t)$ is a smooth Gaussian process and $\epsilon(s, t)$ is an independent error process. The smooth process $z(s, t)$ is constructed to model the data by taking the convolution of a 3-dimensional lattice process. Given a grid process $x = (x_1, \dots, x_m)$ with space–time coordinates $(\omega_1, \tau_1), \dots, (\omega_m, \tau_m)$, the smooth field is expressed as

$$z(s, t) = \sum_{j=1}^m K_s(s - \omega_j, t - \tau_j) \cdot x_j \quad (14)$$

where the properties of the convolution kernel determine the smoothness of z . A separable kernel were used (i.e. a product of a kernel that smooths over space and one that smooths over time): $K_s(\Delta s, \Delta t) = C_s(\Delta s) \cdot R(\Delta t)$. Inference on the resulting model was made using a Bayesian approach and simulating the posterior distribution of the mean temperature field over space and time using Markov chain Monte Carlo methods.

Following a similar approach, but using nonseparable, discrete convolution kernels, regional temperature measurements were modelled in space and time in Sansó et al. (2008). Two alternative set of models were suggested. The first was to convolute spatial Gaussian processes with a kernel providing temporal dependencies and the second was to convolute autoregressive models with a kernel providing spatial interactions. In other words, the data could either be considered as a number of time series at each location (temporal convolution model) or as a number of realizations of spatial processes observed at some locations (spatial convolution model).

A dynamic process convolution model extends the discrete process convolution approach by defining the underlying process x to be a time-dependent process that is spatially smoothed by a smoothing kernel at each time-step (Calder 2007, 2008). Such models have been used in air quality assessment (e.g. in bivariate modelling of levels of particulate matter $PM_{2.5}$ and PM_{10} in Calder (2008) and for multivariate modelling of the concentration of five pollutants in Calder (2007)). A continuous version in space and time is considered in Brown et al. (2000), where a model is formulated in discrete time and continuous space and a

limit argument is applied to obtain continuous time as well. A general approach using cross-convolution of covariance functions for modelling of multivariate geostatistical data were proposed in Majumdar and Gelfand (2007). All of the convolution models discussed above used bayesian approaches and Markov chain Monte Carlo methods for model specification.

Finally, it is noted that some limitations to the convolution model approach is reported in Higdon (1998) and Calder (2008). One is the impact of prior assumptions on the posterior distributions. Furthermore, it is stated that it would be preferable to allow the data to determine the kernels, which could depend on space and time, rather than specifying it a priori. In addition, the model for particulate matter are not able to handle extreme observations very well and permits nonsensible predictions.

4.4 Nonstationary covariance models

Many spatiotemporal models assumes separability in space and time so that the space–time covariance function can be represented as the product of two models: one as a function of space and the other as a function of time. However, the rationale for using a separable model is often convenience rather than the ability of such models to describe the data well, and the assumption is often unrealistic. Other simplifying assumptions often employed are stationarity (e.g. second order stationarity which means that the mean function is assumed constant and the space–time covariance function is assumed to depend only on the directional distance between measurement sites) and isotropy (i.e. that the covariance function is dependent only on the length of the separation and not on its direction). An example of a spatio-temporal covariance model where the assumptions of stationarity and separability is relaxed is presented in Bruno et al. (2009), applied to tropospheric ozone data.

Due to the increased availability of satellite measurements of many geophysical processes, global data are increasingly available. Such data often show strong non-stationarity in the covariance structure. For example, processes may be approximately stationary with respect to longitude but with highly dependent covariance structures with respect to latitude. In order to capture the nonstationarity in such global data, with a spherical spatial domain, a class of parametric covariance models are proposed in Jun and Stein (2008). These assume that processes are axially symmetric, i.e. that they are invariant to rotations about the earth's axis and hence stationary with respect to longitude.

Assuming a homogeneous, zero-mean process Z_0 , a zero-mean nonstationary process Z may be defined by applying differential operators with respect to latitude and

longitude, letting L and l denote latitude and longitude respectively (Jun and Stein 2007),

$$Z(L, l) = \left\{ A(L) \frac{\partial}{\partial L} + B(L) \frac{\partial}{\partial l} \right\} Z_0(L, l) + CZ_0(L, l) \quad (15)$$

Now, A and B denote nonrandom functions depending on latitude (and may also in principle depend on longitude, but this would break the axial symmetry). A non-negative constant C corresponds to including homogeneous models for the case $A(L) = B(L) = 0$. In order to apply this model to real applications, the A and B functions need to be estimated, and it is suggested to use linear combinations of Legendre polynomials (Jun and Stein 2008).

The covariance model is applied to global column ozone level data and it is shown that the strong nonstationarity with respect to latitudes as well as the local variation of the process can be well captured with only a modest number of parameters. Thus, it may be a promising candidate for modelling spatially dependent data on a sphere. Furthermore, an extension to spatio-temporal processes would be obtained by introducing a similar differential operator with respect to time in addition to the ones with respect to latitude and longitude. Then, such models should be able to capture spatial-temporal nonstationary behaviour and to create flexible space–time interactions such as space–time asymmetry. Review of various methods and recent developments for the construction of spatio-temporal covariance models are presented in Kolovos et al. (2004) and Ma (2008).

4.5 Coregionalization models

A multivariate spatial process is a natural modelling choice for multivariate, spatially collected data. When the interest is in modelling and predicting such joint processes it will be important to account for the spatial correlation as well as the correlation among the different variables. If this is modelled using a Gaussian process, the main challenge is the specification of an adequate cross-covariance function (Schmidt and Gelfand 2003), which can be developed through linear models of coregionalization (LMC). The linear model of coregionalization is reviewed in Gelfand et al. (2004) where the notion of spatially varying LMC are proposed in order to enhance the usefulness by providing a class of multivariate nonstationary processes.

Traditionally, linear models of coregionalization have been used to reduce dimensions, approximating a multivariate process through a lower dimensional representation. However, it may also be used in multivariate process construction, i.e. obtaining dependent multivariate processes by linear transformation of independent processes. A general multivariate spatial model could be on the form

$$\mathbf{Y}(s) = \boldsymbol{\mu}(s) + \mathbf{v}(s) + \epsilon(s) \quad (16)$$

where $\epsilon(s)$ is a white noise vector (i.e. $\epsilon(s) \sim N(0, \mathbf{D})$ where \mathbf{D} is a diagonal matrix with $(D_{jj}) = \tau_j^2$), $\mathbf{v}(s)$ arises from a linear model of coregionalization from independent spatial processes $\mathbf{w}(s) = (w_1(s), \dots, w_p(s))$: $\mathbf{v}(s) = \mathbf{A} \mathbf{w}(s)$ and where $\boldsymbol{\mu}(s)$ may be assumed to arise linearly in the covariates, i.e. $\mu_j(s) = \mathbf{X}_j^T(s) \beta_j$ where each component may have its own set of covariates \mathbf{X}_j and its own coefficient vector β_j . If ignoring the term $\boldsymbol{\mu}(s)$ and the $w_j(s)$ processes are assumed to have mean 0, variance 1 and a stationary correlation function $\rho_j(h)$, then $E(\mathbf{Y}(s)) = 0$ and the cross-covariance matrix associated with $\mathbf{Y}(s)$ becomes

$$\Sigma_{\mathbf{Y}(s), \mathbf{Y}(s')} \equiv C(s - s') = \sum_{j=1}^p \rho_j(s - s') \mathbf{T}_j, \quad \mathbf{T}_j = \mathbf{a}_j \mathbf{a}_j^T \quad (17)$$

with \mathbf{a}_j the j th column of \mathbf{A} . Priors on the model parameters $\boldsymbol{\theta}$ consisting of $\{\beta_j\}$, $\{\tau_j^2\}$, \mathbf{T} and ρ_j , $j = 1, \dots, p$ would then complete the model specification in a Bayesian setting, obtaining the posterior distribution of the model parameters

$$\pi(\boldsymbol{\theta} | \mathbf{Y}) \propto f(\mathbf{Y} | \{\beta_j\}, \mathbf{D}, \{\rho_j\}, \mathbf{T}) \pi(\boldsymbol{\theta}) \quad (18)$$

The extension to a spatially varying linear model of coregionalization is obtained by letting \mathbf{A} be spatially dependent, i.e. replacing \mathbf{A} with $\mathbf{A}(s)$ in $\mathbf{v}(s) = \mathbf{A}(s) \mathbf{w}(s)$ (Gelfand et al. 2004). $\mathbf{v}(s)$ will then no longer be a stationary process. Further extensions to spatio-temporal versions of the model, modelling $\mathbf{v}(s, t) = \mathbf{A}(s, t) \mathbf{w}(s, t)$, where the components of $\mathbf{w}(s, t)$ are independent spatio-temporal processes may also be feasible, but this was not further investigated.

A stationary Bayesian linear coregionalization model for multivariate air pollutant data was presented in Schmidt and Gelfand (2003) and Gelfand et al. (2004) presents a commercial real estate example of a spatially varying model. Rather than taking the Bayesian approach, an Expectation-Maximization (EM) algorithm for the maximum-likelihood estimation of the parameters in a linear coregionalization model is developed in Zhang (2007), and applied on a spatial model of soil properties.

4.6 Generalized extreme value models

The generalized extreme value distribution is a cornerstone of extreme value modelling, and in Huerta and Sansó (2007) non-stationary, location-dependent processes are studied using the GEV distribution where the parameters are allowed to vary in space and time. The modeling is based on a hierarchical structure assuming an underlying spatial model. Parameter changes over time (i.e. for the

location, scale and shape parameters) are modelled by use of Dynamic Linear Models (West and Harrison 1997) which is a very general class of time series models. Now, the trends are not constrained to have a specific parametric form and the significance of short term changes can be assessed together with the long term changes. It is also possible to estimate how the effects of covariates change over time. An extension of this model to include changes in space as well as in time is made using a process convolution approach in defining a Dynamic Linear Model on the parameters.

Several approaches for estimation of parameters and quantiles of the GEV distribution have been applied, such as maximum likelihood estimation, L-moments estimation, Probability Weighted Moments estimation and the method of moments. Recently, an alternative to these, employing a full Bayesian GEV estimation method which contains a semi-Bayesian framework of generalized maximum likelihood estimators and considers the shape, location and scale parameters as random variables were developed (Yoon et al. 2010). However, these approaches do not consider non-stationarity.

A generalized Probability Weighted Moments (PWM) method was suggested in Ribereau et al. (2008) to model temporal covariates and provide accurate estimation of return levels from maxima of non-stationary random sequences modelled by a GEV distribution. This is a generalization of the PWM method that has proved to be efficient in estimating the parameters of the GEV distribution for IID processes and is an alternative to Maximum Likelihood Estimation (MLE) for cases when the IID assumption is violated (e.g. in non-stationary cases). The approach is illustrated by applying it on time series of annual maxima of CO₂ concentrations and seasonal maxima of cumulated daily precipitations.

An alternative to GEV models could be to use threshold models (Behrens et al. 2004). For example, various statistical methods for exploring the properties of extreme events in large grid point datasets were presented in Coelho et al. (2008), and a flexible generalized Pareto model that are able to account for spatial and temporal variation in the distribution of excesses were outlined. The generalized Pareto distribution parameters may incorporate the dependence of the extreme values and different explanatory variables related to spatial and temporal changes such as climate change. The methods were illustrated using mean surface temperatures of the Northern Hemisphere.

A generalized PWM method was introduced in Diebolt et al. (2007) in order to estimate the parameters of the generalized Pareto distribution (GPD) from finite length time series. A Bayesian framework for analysis of extremes in a non-stationary context was proposed in Renard et al. (2006) with a case study on peak-over-threshold data.

Several probabilistic models, including stationary, step-wise changing and linear trend models, and different extreme value distributions were considered allowing modelling uncertainty to be taken into account.

An alternative to the standard approach of modelling non-stationarity in threshold models (i.e. retaining a constant threshold and letting the parameters of the GPD be functions of some covariates) is proposed in Eastoe and Tawn (2009). This involves preprocessing; attempting to model the non-stationarity in the entire data set and then removing this non-stationarity from the data. If this preprocessing is successful, the extremes of the preprocessed data will have most, if not all, of its non-stationarity removed and a simple extreme value analysis of the preprocessed data can be employed. It is argued that this approach provides improved description of the non-stationarity of the extremes, clearer interpretation, easier threshold selection and reduced threshold sensitivity. The approach was also found to be superior to approaches with continuous varying thresholds.

4.7 Optimality models

One type of statistical models that has recently been applied in evolutionary sciences are optimality models. These assume the evolution of some biological trait towards an optimal state dictated by the environmental conditions. Due to a randomly changing environment, the optimal state is assumed to change over time, and the species are assumed to be adapting to this changing optimality with a certain phylogenetic inertia. One choice of process models for analysing such an adaptation-inertia problem is the Ornstein-Uhlenbeck process, as suggested in Hansen et al. (2008), represented by the stochastic differential equation

$$dy = -\alpha(y - \theta)dt + \sigma_y dW_y \quad (19)$$

Here, dy is the change in some random variable y over a time step dt , α is a parameter measuring the rate of adaptation toward the optimum θ , dW_y is a random noise process and σ_y is the standard deviation of the random changes. Thus, evolution according to this model has two components: one is a deterministic pull toward the primary optimum and the other is a stochastic change without direction.

A layered process is introduced for modelling adaptation to a randomly changing optimum, assuming that the optimum at any point on the phylogeny (that is, the history of organismal lineages as they change through time) is a function of a randomly changing predictor variable x . Thus, the model is extended to the coupled stochastic differential equations below where the predictor indirectly influences the trait through its influence of the optimum.

$$dy = -\alpha(y - \theta(x))dt + \sigma_y dW_y \tag{20}$$

$$dx = \sigma_x dW_x \tag{21}$$

Additional layers of hidden processes may also be modelled in this way, where each layer is responding to changes in the layer beneath. The model may also be extended in that the predictor variable itself may be modelled as an Ornstein-Uhlenbeck process, tracing some optimum. The Ornstein-Uhlenbeck process has also been proposed for modelling of drought and flood risks (Unami et al. 2010) and survival data (Aalen and Gjessing 2004) and has been widely used in financial modelling (Stein and Stein 1991; Barndorff-Nielsen and Shephard 2001; Benth et al. 2007).

It could be worthwhile to investigate whether an analogy to this approach would be appropriate for the development of extreme waves, i.e. whether the distribution of extreme sea states are trying to adapt to a changing mean state due to the changing environment. For example, will there be a certain average wave climate given the changing environmental conditions such as the level of CO₂ concentration in the atmosphere, global temperatures, greenhouse gas emissions etc.? In other words, it could be investigated whether the distribution of extreme waves in a changing environment could be adequately modelled using layered Ornstein-Uhlenbeck processes in some way.

4.8 Bayesian maximum entropy models

Bayesian maximum entropy (BME) models have been used to model spatiotemporal random fields. For example, in Choi et al. (2009), this approach was used for developing a systematic epidemic forecasting methodology used to study the space–time risk patterns of influenza mortality in California during wintertime. Influenza mortality rates were represented as spatiotemporal random fields and the Bayesian maximum entropy method was used to map the rates in space and time and thus generate predictions. Bayesian maximum entropy models have also been used for space–time mapping of soil salinity (Douanik et al. 2004), urban climate (Lee et al. 2008) and the contamination pattern from the Chernobyl fallout (Savelieva et al. 2005) and for modelling geographic distributions of species (Phillips et al. 2006).

In short, the principle of maximum entropy states that the probability distribution best representing the current state of knowledge, which may be incomplete, is the one with the largest entropy. If some testable information about a probability distribution function is given, then, considering all trial probability distributions that encode this information, the probability distribution that maximizes the information entropy is the true probability distribution with respect to the testable information. This principle is applicable to problems of inference with a well-defined hypothesis space and incomplete data without noise and the Bayesian maximum

entropy method can be used to predict the value of a spatiotemporal random field at an unsampled point in space–time based on precise (hard) and imprecise (soft) data.

The BME method applied to influenza mortality risk (Choi et al. 2009) consists of three stages with different knowledge bases at each stage: the general knowledge base (core knowledge), the specificatory knowledge base (case-specific knowledge) and the integration knowledge base (union of the general and specificatory knowledge bases). The influenza risk is represented as a spatiotemporal random field $X(\mathbf{p})$ defined at each space–time point $\mathbf{p} = (s, t)$. The influenza modelling approach then follows the three BME stages:

- a. A probability density function, $f_g(\mathbf{x}_{\text{map}})$ is constructed on basis of the general knowledge base, where the vector \mathbf{x}_{map} denotes a possible realization of the random field associated with the point vector \mathbf{p}_{map} . The \mathbf{x}_{map} generally includes hard data $\mathbf{x}_{\text{hard}} = (x_1, \dots, x_h)$ at points $\mathbf{p}_{\text{hard}} = (\mathbf{p}_1, \dots, \mathbf{p}_{m_h})$, soft data $\mathbf{x}_{\text{soft}} = (x_{m_h+1}, \dots, x_m)$ at points $\mathbf{p}_{\text{soft}} = (\mathbf{p}_{m_h+1}, \dots, \mathbf{p}_m)$ and the unknown estimates x_k at points \mathbf{p}_k .
- b. At the specificatory stage, the specificatory knowledge base considers hard data and soft data
- c. At the integration stage, the general and specificatory knowledge bases are combined in a total knowledge base to give the integration pdf $f_k(x_k)$ at each mapping point \mathbf{p}_k using the operational Bayesian formula

$$f_k(x_k) = A^{-1} \int_D f_g(\mathbf{x}_{\text{map}}) d\mathcal{E}_S(x_{\text{soft}}) \tag{22}$$

where A is a normalizing constant and \mathcal{E}_S and D denote an integration operator and the range determined by the specificatory knowledge base respectively.

4.9 Stochastic diffusion models

A continuous time parameter stochastic process is referred to as a diffusion process if it possesses the Markov property and its sample paths $X(t)$ are continuous functions of time t . Many physical and other phenomena can be reasonably modelled by diffusion processes. Diffusion processes may be characterized by two infinitesimal parameters describing the mean and the variance of the infinitesimal displacements, defined as the following limits: Let the increment of the process accrued over a time interval h be $\Delta_h X(t) = X(t+h) - X(t)$, then the infinitesimal parameters of the process are:

$$\mu(x, t) = \lim_{h \rightarrow 0} E[\Delta_h X(t) | X(t) = x] \tag{23}$$

$$\sigma^2(x, t) = \lim_{h \rightarrow 0} E[\{\Delta_h X(t)\}^2 | X(t) = x] \tag{24}$$

$\mu(x, t)$ are sometimes referred to as the drift parameter, infinitesimal mean or the expected infinitesimal

displacement and $\sigma^2(x, t)$ is called the diffusion parameter or the infinitesimal variance and these are generally continuous functions in x and t . Alternative characterizations of diffusion processes exist, e.g. based on stochastic differential equations.

A methodology for analysing secular trends in the time evolution of certain variables, modelling the variables by nonhomogeneous stochastic diffusion processes with time-continuous trend functions is proposed in Gutiérrez et al. (2008). The methodology was applied to the evolution of CO₂ emissions in Spain with the Spanish GDP as an exogenous factor affecting the trend component and hence introducing nonhomogeneity. The trend can be analysed by means of statistical fit of the trend functions of the stochastic diffusion model to the observed data, and the models were also found appropriate for medium-term forecasts.

Stochastic diffusion models have been applied to other temporal or spatial problems as well, such as modelling of tumor growth (Albano and Giorno 2006), ion channel gating (Vaccaro 2008), financial volatility (Todorov 2009) and scaling behaviour of precipitation statistics (Kundu and Bell 2006).

4.10 Regional frequency analysis

A method commonly used in hydrology, referred to as regional frequency analysis, utilizes data from several similar sites in order to estimate event frequencies, typically extreme events, at a particular site. The main idea is that data from neighboring or other sites where the frequency of the event to be investigated are similar provide additional information and hence yield more accurate predictions than data from the particular site alone. This approach can also be used to interpolate to ungauged sites where there are no data, based on data from similar sites.

The main idea is that, given data from N similar sites so that one may assume the sites to form a homogeneous region, i.e. that the frequency distributions of the different sites are identical apart from a site-specific scaling factor, the quantile function of the frequency distribution at a site i can be modelled by this scaling factor and a regional quantile function common to every site, referred to as the regional growth curve:

$$Q_i(F) = \mu_i q(F), \quad i = 1, \dots, N \quad (25)$$

In the equation above, $Q_i(F)$ denotes the site-specific quantile function at site i , μ_i denotes the site-specific scaling factor, often referred to as the index flood, and $q(F)$ is the regional growth curve. F is the cumulative distribution function of the frequency distribution of the quantity of interest (e.g. significant wave height).

A typical regional frequency analysis will consist of the following four steps:

- i. Screening of the data: Eliminating gross errors and inconsistencies and checking whether the data are homogeneous over time
- ii. Identification of homogeneous regions: Assign the sites to regions whose frequency distribution are similar
- iii. Choice of regional frequency distribution
- iv. Estimating the frequency distribution: Estimating the distribution at each site to give a regional average

A thorough description of the regional frequency analysis approach is given in Hoskins and Wallis (1997), together with an outline of regional model estimators based on L-moments, a widely used approach for estimation in regional frequency analysis. Estimation based on Bayesian Markov chain Monte Carlo methods in regional frequency analysis were proposed in Gaume et al. (2010). Regional frequency analysis are widely used in hydrology and there are abundant literature on applications to extreme rainfall (Yang et al. 2010; Nguyen et al. 2002; Fowler and Kilsby 2003) and flooding (Leclerc and Ourada 2007; Saf 2010). Regional frequency analysis has also been applied in ocean engineering problems such as modelling of significant wave heights (Van Gelder et al. 2001; Ma et al. 2006) and the height of wave crests (Izadparast and Niedzwecki 2009).

4.11 Selecting a modelling approach

In the preceding sections of the paper, a number of different modelling approaches have been reviewed, which may all be appropriate for modelling long-term trends in extreme wave climate. However, the approach based on Bayesian hierarchical space–time models is believed to be superior and offer several benefits compared to the other approaches that has been reviewed. The hierarchical Bayesian approach to modelling data and processes with different scales of spatial and temporal variability consist of different stages (e.g. the data stage, the process stage and the parameter stage) and such models are generally very flexible and intuitive to work with, as outlined in Wikle et al. (1998).

Some of the advantages of an hierarchical approach are the flexibility such models allow. One may build up the models in a modular, hierarchical manner where different aspects of the model can be treated separately. Extensions to the model may easily be incorporated, if necessary, and the different modules may be updated individually as knowledge and insight increase. Knowledge about physical aspects may be incorporated in the models, as illustrated by the earthquake model in Natvig and Tsvete (2007) and such models are very flexible with regard to how they are built up. Furthermore, hierarchical models, incorporating knowledge about the physical phenomena they represent, performs better with regards to interpretation of results.

One crucial assumption applied in the model of earthquakes (Natvig and Tvette 2007) was the Markovian assumption (i.e. that the spatial process or field in one location is only dependent on its nearest neighbors). Although this assumption needs to be challenged on a case by case basis, it is presumably a reasonable assumption also for ocean waves. Hence, it may be reasonable to model ocean waves as a random Markov field along the lines of Natvig and Tvette (2007).

There are also benefits from utilizing a Bayesian approach, related to the fact that knowledge about the physical process and its characteristics may be exploited by way of the prior distributions. This is clearly an advantage in modelling of physical phenomena where such knowledge are available, as is the case of ocean waves. Furthermore, by adopting a Bayesian approach the uncertainty in model parameters is taken into account. Hence, of all the modelling approaches reviewed herein, Bayesian hierarchical space–time models are believed to be the most promising candidate for further developments in long-term time-dependent stochastic modelling of extreme waves and it is suggested that further research and model development are focused in this direction.

5 Wave climate projections

5.1 Climate change

The IPCC's fourth assessment report states that "Warming of the climate system is unequivocal, as is now evident from observations of increases in global average air and ocean temperatures, widespread melting of snow and ice and rising global average sea level" (IPCC 2007). It predicts further global warming and that many changes in the global climate are very likely to be larger during the twenty-first century than what has been observed during the twentieth century. Furthermore, the frequency and intensity of extreme events are expected to change as the global climate changes, some of which has already been observed. A more recent up-to-date overview on climate change research (Richardson et al. 2009) has as one of its key messages that recent observations indicate that the climate change may be more severe and occur earlier than the fourth assessment report predicts.

However, in spite of climate change being a global phenomenon, regional variability is large, and it has been observed that for example the Arctic has warmed at double rate compared to the rest of the world in recent decades (ACIA 2005). Regional differences are also presumed to be predominant in future changes of the climate, but overall, the globe is expected to warm and the intensity and frequency of extreme climatic events are likely to increase.

Hence, one important question for the stakeholders involved in maritime transport is to what extent the observed and projected global warming will influence the wave climate on short and long term and what impact this will have on the safety of maritime transportation. In the following, a review of some projections of wave climate within the context of this global warming will be presented as well as analyses of previous and current trends.

5.2 Current trends in the wave climate

Evidence for a statistical significant increasing trend in mean wave height in the North Atlantic was observed more than 30 years ago (Carter and Draper 1988; Bacon and Carter 1991). Since then, there are a number of studies reported in the literature which try to identify and assess previous and current trends in extreme wave climate, most with a focus on the North Atlantic, by different hindcast and reanalysis techniques combined with statistical analyses (see e.g. Kalnay et al. 1996; Uppala et al. 2005; Athanassoulis and Stefanakos 1995). Some of these will be briefly outlined in the following.

Seasonal trends in extremes of significant wave height were assessed in Wang and Swail (2001), Wang and Swail (2002) for the North Atlantic and the North Pacific by simulating a 40-year global wave hindcast. For both oceans, no statistically significant changes were observed for the last century, but significant changes were found in some regions and for some seasons for the last four decades. Most notably an increase was found for the winter season in the North Atlantic, matched by a decrease in the subtropical North Atlantic and a significant increase in the North Pacific for the winter and spring seasons. Extensive validation of a 40 year global wave hindcast against available wave observations (from buoys, platforms, ships and satellites) has shown generally good agreement over the entire frequency distribution for such reanalyzed data (Cox and Swail 2001).

A previous study, somewhat limited in scope with regards to the period and area covered compared to the assessments in Wang and Swail (2001, 2002), reported a similar increasing trend in significant wave height at several northeast Atlantic locations since 1960, as well as a decrease south of 40°N (Kushnir et al. 1997). Similar patterns were also suggested in Sterl et al. (1998). An increase in frequency of extreme events in the last four decades were reported for the North Sea in Weisse and Stawarz (2004), although no significant changes were found with regards to intensity and duration. Also, an analysis of wave hindcast for 1955–1994 reported in The WASA Group (1998) suggests an increasing trend in both the North Sea and the Norwegian Sea, but with decreasing trends in other regions. A global wave climate trend

analysis was reported in Caires and Swail (2004) where a ERA-40 dataset, corrected for inhomogeneities, was used and significant increasing trends for mean, 90- and 99-percentiles were found for a large part of the global oceans. An intercomparison of significant wave height data derived from different reanalyses was presented in Caires et al. (2004), and in spite of differences in the data quality and scope, it was reported that most of the long-term characteristics such as trends and variability, were equally present in all datasets.

Hence, most studies from the turn of the century generally agree that the wave climate of the North Atlantic became rougher in the last decades of the twentieth century. These general conclusions have been supported by analysis of microseismological data (Grevemeyer et al. 2000), by significant wave height data from ship observations (Gulev and Hasse 1999) and by satellite altimetry data (Carter 1999; Woolf et al. 2002). Buoy measurements have also suggested an increase in wave height for the western Atlantic ocean, but only for the summer hurricane season (Komar and Allan 2008). More recent studies observe the continuation of this increasing trend into the twenty-first century (Caires et al. 2006a; Dodet et al. 2010), although there are still uncertainties as to whether, or to what extent, the trend can be ascribed to global warming (Wolf and Woolf 2006; Wang et al. 2009). Increasing trends have also been found in other oceans than the Atlantic (Sasaki et al. 2005; Gower 2002; Caires and Swail 2004; Ruggiero et al. 2010).

However, it is noted that opposite trends have been reported for different regions, some studies reporting decreasing trends for particular regions (Martucci et al. 2009; Dupuis et al. 2006) and for different seasons (e.g. decreasing trends for the months February was reported in contrast to increasing trends in January for the Baltic Sea in Rózyński (2010)) so care should always be exercised when extrapolating conclusions arrived at from one location to another or from one season to another. Notwithstanding, there are evidence for a general overall trend of rougher wave climate in the North Atlantic as well as in various other ocean areas.

5.3 Prediction of future trends in the wave climate

In light of the observed increasing trends in recent extreme wave climate in many areas of the world, a much relevant question is whether, or to what extent, this trend will prevail in the coming decades and how the future wave climate will develop. In the following, a review of some attempts to make projections of future trends in the wave climate will be presented, with an emphasis on the trends for extreme waves.

The modelling approach outlined in Caires et al. (2006b), modelling extreme waves as non-homogeneous

Poisson processes (NPP), utilizing the statistical relationship between wave height and sea level pressure, and compared to a non-stationary generalized extreme value model (NS-GEV), is already discussed briefly in previous sections. Previous work focusing on the relationship between wave height and sea level pressure includes Wang et al. (2004) and Wang and Swail (2006a, b). One interesting finding is that the regression model best describing the 20-year significant wave height time series towards 2099 is quadratic in time, in contrast to the present climate where the trends are linear (Caires et al. 2002).

The changes of the projected extreme wave climate towards 2099 arrived at from the NPP model were found to be dependent on season and location and the spatial patterns were very similar for the different scenarios that were investigated. However, the magnitude of the estimated changes as well as the time evolution of the projections (i.e. how fast the changes will occur) were scenario dependent. Comparing the projected times series of significant wave height return values obtained from the NPP and the NS-GEV models, it was found that they are highly correlated and hence compatible in some sense, but significant differences in means and variances were found, mainly in tropical areas. The seasonal projections of the 20-year significant wave height towards 2099, under different forcing scenarios and using a NPP model with parameters estimated from present climate data (Caires et al. 2006b), predict significant changes in SWH_{20} in different regions of the world. The rate of future changes depends on the scenario, but under all scenarios considered, significant positive trends are predicted in the North Pacific. This is in agreement with the projections made in Sasaki et al. (2006), which predicts increases in significant wave height by up to 0.4 m over a wide area of the western North Pacific.

Projections of future wind, wave and storm surge climate in the North Atlantic based on regional wave models are reported in Debernard et al. (2002) and Debernard and Røed (2008), where a future climate for the period 2030–2050 was compared to a control climate for the period 1980–2000. The initial study did not identify statistically significant changes in wave height (Debernard et al. 2002), but when the study was revisited with more recent IPCC scenarios, the following statistically significant changes in extreme significant wave height were found: Significant increases west of the British Isles and in the eastern North Sea and in the Skagerrak and significant decreases west of 30°W (Debernard and Røed 2008).

Projections of extreme wave heights for the Northwest Pacific ocean towards the year 2032 were presented in Ruggiero et al. (2010), based on various approaches, i.a. a time-dependent Generalized Pareto-Poisson model with time-dependent event rate, generalized extreme value-

model with time-dependent trends in the location and scale parameters of the GEV distribution and based on non-stationary *r*-largest extreme value analysis. The projected 100-year return level from the different modelling approaches showed a robust trend but with significant spread for the year 2032. Projections further into the future would yield still greater spread in the model projections. It is therefore cautioned against using projected extreme values in actual engineering problems and such projections should be considered as uncertain until the underlying causes of the long-term trends are better understood.

Other wave climate projections based on dynamical downscaling of projections from global atmosphere-ocean climate models are reported for different regions of the world in Andrade et al. (2007), Caires et al. (2008), Hemer et al. (2009b), Lionello et al. (2008) and Perrie et al. (2004), the details of which will not be covered herein. In order to extend the confidence and coverage of future wave climate projections a proposal for a coordinated effort towards global wave climate projections was made in Hemer et al. (2009a), suggesting a shift in focus from regional projections.

Various downscaling methods for estimation of statistics for significant wave height were investigated in Wang et al. (2010), evaluated against the ERA-40 wave data. Statistical downscaling approaches, typically based on the observed statistical relationship between atmospheric variables and wave height were deemed better than dynamical methods, which typically involves using atmospheric variables to drive ocean wave models. Furthermore, different atmospheric covariates were analysed in nested regression models (i.e. sea level pressure anomalies, sea level pressure gradients and anomalies of seasonal mean squared surface wind speeds) and analysis of the various models suggests that it is sufficient to use the wind-based predictor alone since this model performs very similar to the full model. Projections made from the different approaches show similar patterns for both seasonal means and extremes. In wintertime, increases in the eastern and western subtropical North Atlantic and decreases most other places were the predominant pattern whereas autumn projections were characterized by increases in the mid-latitudes and eastern subtropical North Atlantic and decreases in some other areas.

Dynamic and statistical downscaling techniques were also investigated in Gaslikova and Weisse (2006), and a combination of dynamical and statistical approaches was proposed as a faster, less computational-intensive alternative to purely dynamical methods for downscaling of medium-scale wave data, a conclusion supported by Breivik et al. (2009). The method demonstrated reasonable agreement with observed wave conditions for simulations of a near-shore area around Helgoland.

The uncertainty of the impact of climate change on future extreme wave conditions in the North Sea was investigated in Grabemann and Weisse (2008) by running the WAM wave model (The WAMDI Group 1988) over an ensemble of four different climate change realizations for the 30-year period 2071–2100. Wind field data sets were obtained by simulation outputs from two global circulation models for two emission scenarios, and compared to two control scenarios. The study revealed that there are large uncertainties in the magnitude and the spatial patterns of the climate change signals, and results indicate that the uncertainties due to different climate models are larger, by a factor exceeding five in some regions, than the uncertainties related to the different scenarios. Notwithstanding these uncertainties, it was general agreement between the simulations in that extreme wave heights will increase in large parts of the North Sea and that the future frequency of severe sea states will increase due to global warming.

6 Summary and conclusions

This paper has presented a comprehensive review of the literature concerning probabilistic modelling of ocean waves and sea states. It has addressed the importance of available wave data in order to develop sensible probabilistic models, and although buoy measurements are generally regarded as most reliable, the spatial coverage of such data may be inadequate for spatial models. Alternatives exist in satellite data and in reanalysis data obtained from wave models forced by various meteorological parameters. In particular, data from the ERA-40 reanalysis project that are freely available for scientific purposes have been identified.

Numerous statistical models for extreme waves have been reported in the literature, and some of these have been presented herein. Most of these either have short-term scope or are microscopic in the sense that they focus on a particular location where wave data have been available. That is, the spatial variability is not covered by many of these models. The long-term trends and time-dependencies due to observed and projected climate change are also poorly incorporated in some of these models. There have been some attempts to develop spatio-temporal models for extreme waves, and these have been discussed herein. However, due to the modest number of attempts to establish spatio-temporal models for extreme waves, a glance at models proposed in other areas of application has also been reported. Hence, a review of some models used in earthquake modelling, storm modelling, temperature modelling and air pollution modelling has been presented. It is suggested that similar approaches might be appropriate for spatio-temporal modelling of extreme waves and further work should focus on developing such models.

In particular, the framework of Bayesian hierarchical space–time models have been identified as a promising candidate for further development of long-term stochastic models of extreme wave climate. It is believed that such a framework offer significant improvements in the statistical understanding and modelling of extreme waves and may be used in modelling and projecting long-term trends due to climate change.

Following the review of different stochastic models, a review of projections of future wave climate has been presented. Most of these predict changes in the global wave climate towards the end of the century, but the changes are very region-dependent. However, the overall message is that, at least for the Northern Atlantic, the wave climate will tend to become rougher during the present century. This means that historic wave data may no longer be adequate as basis for design of ships and offshore structures or for use in risk assessment and that new knowledge about the time-dependence and long-term trends of extreme wave climate is of crucial importance. This should motivate further research into the statistical relationships and development of improved spatio-temporal models for extreme waves.

Acknowledgments Thanks to Dr. Elzbieta Bitner-Gregersen (DNV) for contributing literature for this review and to Professor Bent Natvig (UiO) for valuable discussions and comments. Thanks also to Associate Professor Arne Bang Huseby (UiO) for general support and encouragement.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Aalen OO, Gjessing HK (2004) Survival models based on the Ornstein-Uhlenbeck process. *Lifetime Data Anal* 10:407–423
- Aberg S, Lindgren G (2008) Height distribution of stochastic Lagrange ocean waves. *Probab Eng Mech* 23:359–363
- ACIA (2005) Arctic Climate Impact Assessment. Cambridge University Press, Cambridge
- Albano G, Giorno V (2006) A stochastic model in tumor growth. *J Theor Biol* 242:329–336
- Altunkaynak A (2005) Significant wave height prediction by using a spatial model. *Ocean Eng* 32:924–936
- Andrade C, Pires H, Taborda R, Freitas M (2007) Projecting future changes in wave climate and coastal response in Portugal by the end of the 21st century. *J Coast Res SI* 50:253–257
- Andreas EL, Wang S (2007) Predicting significant wave height off the northeast coast of the United States. *Ocean Eng* 34:1328–1335
- Arena F, Pavone D (2006) Return period of nonlinear high wavecrests. *J Geophys Res Ocean* 111:C08004
- Arena F, Pavone D (2009) A generalized approach for long-term modelling of extreme crest-to-trough wave heights. *Ocean Model* 26:217–225
- Arena F, Puca S (2004) The reconstruction of significant wave height time series by using a neural network approach. *J Offshore Mech Arct Eng* 126:213–219
- Arena F, Barbaro G, Romolo A (2009) Return periods of a sea storm with at least two waves higher than a fixed threshold. In: Proceedings of the 28th international conference on offshore mechanics and Arctic engineering (OMAE 2009), American Society of Mechanical Engineers (ASME)
- Athanassoulis G, Stefanakos CN (1995) A nonstationary stochastic model for long-term time series of significant wave height. *J Geophys Res* 100(C8):16149–16162
- Bacon S, Carter D (1991) Wave climate changes in the North Atlantic and North Sea. *Int J Climatol* 11:545–558
- Barndorff-Nielsen OE, Shephard N (2001) Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *J R Stat Soc B* 63:167–241
- Battjes JA (1972) Long-term wave height distribution at seven stations around the British Isles. *Ocean Dyn* 25:179–189
- Baxevani A, Rychlik I, Wilson RJ (2005) A new method for modelling the space variability of significant wave height. *Extremes* 8:267–294
- Baxevani A, Borgel C, Rychlik I (2007) Spatial models for the variability of the significant wave height on the world oceans. In: Proceedings of the 17th international offshore and polar engineering conference (ISOPE 2007), The International Society of Offshore and Polar Engineering (ISOPE)
- Baxevani A, Caires S, Rychlik I (2009) Spatio-temporal statistical modelling of significant wave height. *Environmetrics* 20:14–31
- Bazargan H, Bahai H, Aryana F (2007) Simulation of the mean zero-up-crossing wave period using artificial neural networks trained with a simulated annealing algorithm. *J Mar Sci Technol* 12(1): 22–33
- Behrens CN, Lopes HF, Gaman D (2004) Bayesian analysis of extreme events with threshold estimation. *Stat Modelling* 4:227–244
- Benth FE, Kallsen J, Meyer-Brandis T (2007) A non-gaussian Ornstein-Uhlenbeck process for electricity spot price modeling and derivatives pricing. *Appl Math Finance* 14:153–169
- Bitner-Gregersen EM, de Valk C (2008) Quality control issues in estimating wave climate from hindcast and satellite data. In: Proceedings of the 27th international conference on offshore mechanics and Arctic engineering (OMAE 2008), American Society of Mechanical Engineers (ASME)
- Bitner-Gregersen EM, Toffoli A (2009) Uncertainties of wind sea and swell prediction from the Torsethaugen spectrum. In: Proceedings of the 28th international conference on offshore mechanics and Arctic engineering (OMAE 2009), American Society of Mechanical Engineers (ASME)
- Bocotti P (2000) Wave mechanics for ocean engineering. Elsevier oceanography series. Elsevier Science B.V
- Boer GJ, Flato G, Reader MC, Ramsden D (2000) A transient climate change simulation with greenhouse gas and aerosol forcing: experimental design and comparison with the instrumental record for the twentieth century. *Clim Dyn* 16:405–425
- Borgman LE (1973) Probabilities for highest wave in hurricane. *J Waterw Harb Coast Eng Div ASCE* 99:185–207
- Boukhanovsky AV, Krogstad HE, Lopatoukhin LJ, Rozhkov VA (2003a) Stochastic simulation of inhomogeneous metocean fields. Part i: Annual variability. In: Sloot P, Abramson D, Bogdanov A, Dongarra J, Zomaya A, Gorbachev Y (eds) Computational science—ICCS 2003, lecture notes in computer science, vol 2658. Springer, Berlin, pp 213–222
- Boukhanovsky AV, Krogstad HE, Lopatoukhin LJ, Rozhkov VA, Athanassoulis GA, Stephanakos CN (2003b) Stochastic simulation of inhomogeneous metocean fields. partii: synoptic variability and rare events. In: Sloot P, Abramson D, Bogdanov A, Dongarra J, Zomaya A, Gorbachev Y (eds) Computational science—ICCS 2003, lecture notes in computer science, vol 2658. Springer, Berlin, pp 223–233

- Breivik V, Gusdal Y, Furevik BR, Aarnes OJ, Reistad M (2009) Nearshore wave forecasting and hindcasting by dynamical and statistical downscaling. *J Mar Syst* 78:235–243
- Brown PE, Ksresen KF, Roberts GO, Tonellato S (2000) Blur-generated non-separable space-time models. *J R Stat Soc B* 62:847–860
- Bruno F, Guttorp P, Sampson PD, Cocchi D (2009) A simple non-separable, non-stationary spatiotemporal model for ozone. *Environ Ecol Stat* 16:515–529
- Buishand T, de Haan L, Zhou C (2008) On spatial extremes: with application to a rainfall problem. *Ann Appl Stat* 2:624–642
- Caires S, Sterl A (2005) A new nonparametric method to correct model data: application to significant wave height from ERA-40 re-analysis. *J Atmos Ocean Technol* 22:443–459
- Caires S, Swail V (2004) Global wave climate trend and variability analysis. In: Preprints of 8th international workshop on wave hindcasting and forecasting
- Caires S, Sterl A, Bidlot JR, Graham N, Swail V (2002) Climatological assessment of reanalysis ocean data. In: Preprints of 7th international workshop on wave hindcasting and forecasting
- Caires S, Sterl A, Bidlot JR, Graham N, Swail V (2004) Intercomparison of different wind-wave reanalyses. *J Clim* 17:1893–1913
- Caires S, Groeneweg J, Sterl A (2006a) Changes in the North Sea extreme waves. In: Preprints of 9th international workshop on wave hindcasting and forecasting
- Caires S, Swail VR, Wang XL (2006b) Projection and analysis of extreme wave climate. *J Clim* 19:5581–5605
- Caires S, Groeneweg J, Sterl A (2008) Past and future changes in the North Sea extreme waves. In: Proceedings of the 31st international conference on coastal engineering (ICCE 2008)
- Calder CA (2007) Dynamic factor process convolution models for multivariate space-time data with application to air quality assessment. *Environ Ecol Stat* 14:229–247
- Calder CA (2008) A dynamic process convolution approach to modeling ambient particulate matter concentrations. *Environmetrics* 19:39–48
- Carter D (1999) Variability and trends in the wave climate of the North Atlantic: a review. In: Proceedings of the 9th international offshore and polar engineering conference (ISOPE 1999), The International Society of Offshore and Polar Engineering (ISOPE)
- Carter D, Draper L (1988) Has the north-east Atlantic become rougher? *Nature* 332:494
- Choi KM, Yu HL, Wilson ML (2009) Spatiotemporal statistical analysis of influenza mortality risk in the State of California during the period 1997–2001. *Stoch Environ Res Risk Assess* 22:S15–S25
- Cocchi D, Greco F, Trivisano C (2007) Hierarchical space-time modelling of PM₁₀ pollution. *Atmos Environ* 41:532–542
- Coelho C, Ferro C, Stephenson D, Steinskog D (2008) Methods for exploring spatial and temporal variability of extreme events in climate data. *J Clim* 21:2072–2092
- Coles SG (1993) Regional modelling of extreme storms via max-stable processes. *J R Stat Soc B* 55:797–816
- Coles SG, Tawn JA (1996) Modelling extremes of the areal rainfall process. *J R Stat Soc B* 58:329–347
- Coles S, Walshaw D (1994) Directional modelling of extreme wind speeds. *J R Stat Soc* 43:139–157
- Coles S, Pericchi LR, Sisson S (2003) A fully probabilistic approach to extreme rainfall modeling. *J Hydrol* 273:35–50
- Cox AT, Swail VR (2001) A global wave hindcast over the period 1958–1997: validation and climate assessment. *J Geophys Res* 106:2313–2329
- Cunha C, Guedes Soares C (1999) On the choice of data transformation for modelling time series of significant wave height. *Ocean Eng* 26:489–506
- Debernard JB, Røed LP (2008) Future wind, wave and storm surge climate in the Northern Seas: a revisit. *Tellus* 60A:427–438
- Debernard J, Sætra Ø, Røed LP (2002) Future wind, wave and storm surge climate in the northern North Atlantic. *Clim Res* 23:39–49
- de Haan L (1984) A spectral representation for max-stable processes. *Ann Probab* 12:1194–1204
- de Waal D, van Gelder P (2005) Modelling of extreme wave heights and periods through copulas. *Extremes* 8:345–356
- DelBalzo DR, Schultz JR, Earle MD (2003) Stochastic time-series simulation of wave parameters using ship observations. *Ocean Eng* 30:1417–1432
- Deo M, Jha A, Chaphekar A, Ravikant K (2001) Neural networks for wave forecasting. *Ocean Eng* 28:889–898
- Diebolt J, Guillou A, Rached I (2007) Approximation of the distribution of excesses through a generalized probability-weighted moments method. *J Stat Plan Inference* 137:841–857
- Dodet G, Bertin X, Tabora R (2010) Wave climate variability in the North-East Atlantic Ocean over the last six decades. *Ocean Model* 31:120–131
- Douanik A, van Meirvenne M, Tóth T, Serre M (2004) Space-time mapping of soil salinity using probabilistic bayesian maximum entropy. *Stoch Environ Res Risk Assess* 18:219–227
- Dupuis H, Michel D, Sottolichio A (2006) Wave climate evolution in the Bay of Biscay over two decades. *J Mar Syst* 63:105–114
- Durrant TH, Greenslade DJ, Simmonds I (2009) Validation of Jason-1 and Envisat remotely sensed wave heights. *J Atmos Ocean Technol* 26:123–134
- Eastoe EF, Tawn JA (2009) Modelling non-stationary extremes with application to surface level ozone. *J R Stat Soc C* 58:25–45
- Egozcue J, Pawlowsky-Glahn V, Ortego M (2005) Wave-height hazard analysis in eastern coast of Spain—bayesian approach using generalized Pareto distribution. *Adv Geosci* 2:25–30
- Ewans K, Bitner-Gregersen E, Guedes Soares C (2006) Estimation of wind-sea and swell components in a bimodal sea state. *J Offshore Mech Arct Eng* 128(4):265–270
- Fedele F, Arena F (2009) The equivalent power storm model for long-term predictions of extreme wave events. In: Proceedings of the 28th international conference on offshore mechanics and Arctic engineering (OMAE 2009), American Society of Mechanical Engineers (ASME)
- Fedele F, Sampath P, Benetazzo A, Forristall G, Gallego G, Yezzi A, Tayfun M, Cavaleri L, Sclavo M, Bastianini M (2009) Beyond waves & spectra: Euler characteristics of ocean sea states. In: Proceedings of the 28th international conference on offshore mechanics and Arctic engineering (OMAE 2009), American Society of Mechanical Engineers (ASME)
- Ferreira J, Guedes Soares C (1999) Modelling the long-term distribution of significant wave height with the Beta and Gamma models. *Ocean Eng* 26:713–725
- Ferreira J, Guedes Soares C (2000) Modelling distributions of significant wave height. *Coast Eng* 40:361–374
- Forristall GZ (2008) How should we combine long and short term wave height distributions? In: Proceedings of the 27th international conference on offshore mechanics and Arctic engineering (OMAE 2008), American Society of Mechanical Engineers (ASME)
- Fowler H, Kilsby C (2003) A regional frequency analysis of United Kingdom extreme rainfall from 1960 to 2000. *Int J Climatol* 23:1313–1334
- Gaslikova L, Weisse R (2006) Estimating near-shore wave statistics from regional hindcasts using downscaling techniques. *Ocean Dyn* 56:26–35
- Gaume E, Gaál L, Viglione A, Szolgay J, Kohnová S, Blöschl G (2010) Bayesian MCMC approach to regional flood frequency analysis involving extraordinary flood events on ungauged sites. *J Hydrol*. doi:10.1016/j.jhydrol.2010.01.008
- Gelfand AE, Schmidt AM, Benerjee S, Sirmans C (2004) Non-stationary multivariate process modeling through spatially varying coregionalization. *Test* 13:263–312

- Gibson R, Forristall GZ, Owrld P, Grant C, Smyth R, Hagen O, Leggett I (2009) Bias and uncertainty in the estimation of extreme wave heights and crests. In: Proceedings of the 28th international conference on offshore mechanics and Arctic engineering (OMAE 2009), American Society of Mechanical Engineers (ASME)
- Gilchrist WG (2000) Statistical modelling with quantile functions. Chapman & Hall, London
- Gower J (2002) Temperature, wind and wave climatologies, and trends from marine meteorological buoys in the northeast Pacific. *J Clim* 15:3709–3718
- Grabemann I, Weisse R (2008) Climate change impact on extreme wave conditions in the North Sea: an ensemble study. *Ocean Dyn* 58:199–212
- Grevemeyer I, Herber R, Essen HH (2000) Microseismological evidence for a changing wave climate in the northeast Atlantic Ocean. *Nature* 408:349–352
- Guedes Soares C, Cunha C (2000) Bivariate autoregressive models for the time series of significant wave height and mean period. *Coast Eng* 40:297–311
- Guedes Soares C, Ferreira A (1996) Representation of non-stationary time series of significant wave height with autoregressive models. *Probab Eng Mech* 11:139–148
- Guedes Soares C, Scotto M (2001) Modelling uncertainty in long-term predictions of significant wave height. *Ocean Eng* 28:329–342
- Guedes Soares C, Scotto M (2004) Application of the r largest-order statistics for long-term predictions of significant wave height. *Coast Eng* 51:387–394
- Guedes Soares C, Ferreira A, Cunha C (1996) Linear models of the time series of significant wave height on the Southwest Coast of Portugal. *Coast Eng* 29:149–167
- Guedes Soares C, Bitner-Gregersen E, Antão P (2001) Analysis of the frequency of ship accidents under severe north Atlantic weather conditions. In: Proceedings of the design & operations in abnormal conditions II. Royal Institution of Naval Architects (RINA), London
- Guedes Soares C, Weisse R, Carretero JC, Alvarez E (2002) A 40 years hindcast of wind, sea level and waves in European waters. In: Proceedings of the 21st international conference on offshore mechanics and arctic engineering (OMAE 2002), American Society of Mechanical Engineers (ASME)
- Gulev SK, Hasse L (1999) Changes of wind waves in the North Atlantic over the last 30 years. *Int J Climatol* 19:1091–1117
- Gutiérrez R, Gutiérrez-Sánchez R, Nafidi A (2008) Trend analysis using nonhomogeneous stochastic diffusion processes. Emission of CO₂; Kyoto protocol in Spain. *Stoch Environ Res Risk Assess* 22:57–66
- Hagen V (2009) Estimation of long term extreme waves from storm statistics and initial distribution approach. In: Proceedings of the 28th international conference on offshore mechanics and Arctic engineering (OMAE 2009), American Society of Mechanical Engineers (ASME)
- Hemer M, Church J, Swail V, Wang X (2009a) Coordinated global wave climate projections. In: Preprints of 11th international workshop on wave hindcasting and forecasting and coastal hazards symposium
- Hemer MA, McInnes K, Ranasinghe R (2009b) Future projections of the East Australian wave climate. In: Preprints of 11th international workshop on wave hindcasting and forecasting and coastal hazards symposium
- Hansen TF, Pienaar J, Orzack SH (2008) A comparative method for studying adaptation to a randomly evolving environment. *Evolution* 62:1965–1977
- Higdon D (1998) A process-convolution approach to modelling temperatures in the North Atlantic Ocean. *Environ Ecol Stat* 5:173–190
- Hoskins JRM, Wallis JR (1997) Regional frequency analysis. Cambridge University Press, Cambridge
- Huerta G, Sansó B (2007) Time-varying models for extreme values. *Environ Ecol Stat* 14:285–299
- Hwang PA, Teague WJ, Jacobs GA (1998) A statistical comparison of wind speed, wave height, and wave period derived from satellite altimeters and ocean buoys in the Gulf of Mexico region. *J Geophys Res* 103:10451–10468
- IPCC (2007) Climate change 2007: synthesis report. Technical report. Intergovernmental Panel on Climate Change
- Izadparast AH, Niedzwecki JM (2009) Estimating wave crest distributions using the method of L-moments. *Appl Ocean Res* 31:37–43
- Janssen PA (2008) Progress in ocean wave forecasting. *J Comput Phys* 227:3572–3594
- Jenkins AD (2002) Wave duration/persistence statistics, recording interval and fractal dimension. *Int Offshore Polar Eng* 12:109–113
- Jönsson A, Broman B, Rahm L (2002) Variations in the Baltic Sea wave fields. *Ocean Eng* 30:107–126
- Jun M, Stein ML (2007) An approach to producing space-time covariance functions on spheres. *Technometrics* 49:468–479
- Jun M, Stein ML (2008) Nonstationary covariance models for global data. *Ann Appl Stat* 2:1271–1289
- Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo K, Ropelewski C, Wang J, Leetmaa A, Reynolds R, Jenne R, Joseph D (1996) The NCEP/NCAR 40-year reanalysis project. *Bull Am Meteorol Soc* 77:437–471
- Kinsman B (1965) Wind waves. Their generation and propagation on the ocean surface. Prentice-Hall, Englewood Cliffs
- Kolovos A, Christakos G, Hristopoulos D, Serre M (2004) Methods for generating non-separable spatiotemporal covariance models with potential environmental applications. *Adv Water Resour* 27:815–830
- Komar PD, Allan JC (2008) Increasing hurricane-generated wave heights along the U.S. east coast and their climate controls. *J Coast Res* 24:479–488
- Komen G, Cavaleri L, Donelan M, Hasselmann K, Hasselmann S, Janssen P (1994) Dynamics and modelling of ocean waves. Cambridge University Press, Cambridge
- Krogstad HE (1985) Height and period distributions of extreme waves. *Appl Ocean Res* 7:158–165
- Krogstad HE, Barstow SF (1999) Satellite wave measurements for coastal engineering applications. *Coast Eng* 37:283–307
- Krogstad HE, Barstow SF (2004) Analysis and applications of second-order models for maximum crest height. *J Offshore Mech Arct Eng* 126:66–71
- Kundu PK, Bell TL (2006) Space-time scaling behaviour of rain statistics in a stochastic fractional diffusion model. *J Hydrol* 322:49–58
- Kushnir Y, Cardone V, Greenwood J, Cane M (1997) The recent increase in north Atlantic wave heights. *J Clim* 10:2107–2113
- Leclerc M, Ourada TB (2007) Non-stationary regional flood frequency analysis at ungauged sites. *J Hydrol* 343:254–265
- Lee SJ, Balling R, Gober P (2008) Bayesian maximum entropy mapping and the soft data problem in urban climate research. *Ann Assoc Am Geogr* 98:309–322
- Lehner S, Koenig T, Schulz-Stellenfleth J (2007) Global statistics of extreme windspeed and sea state from SAR. In: Proceedings of the Envisat symposium 2007, European Space Agency
- Lemos RT, Sansó B (2009) A spatio-temporal model for mean, anomaly and trend fields of north Atlantic sea surface temperature. *J Am Stat Assoc* 104:5–18
- Lindgren G (2006) Slepian models for the stochastic shape of individual Lagrange sea waves. *Adv Appl Probab* 38:430–450

- Lionello P, Cogo S, Galati M, Sanna A (2008) The Mediterranean surface wave climate inferred from future scenario simulations. *Glob Planet Change* 63:152–162
- Ma C (2008) Recent developments on the construction of spatio-temporal covariance models. *Stoch Environ Res Risk Assess* 22:S39–S47
- Ma QS, Li YB, Li J (2006) Regional frequency analysis of significant wave heights based on L-moments. *China Ocean Eng* 20:85–98
- Machado U, Rychlik I (2003) Wave statistics in non-linear random sea. *Extremes* 6:125–146
- Mahjoobi J, Etemad-Shahidi A (2008) An alternative approach for the prediction of significant wave heights based on classification and regression trees. *Appl Ocean Res* 30:172–177
- Mahjoobi J, Mosabbeh EA (2009) Prediction of significant wave height using regressive support vector machines. *Ocean Eng* 36:339–347
- Majumdar A, Gelfand AE (2007) Multivariate spatial modeling for geostatistical data using convolved covariance functions. *Math Geol* 39:225–245
- Makarynsky O, Pires-Silva A, Makarynska D, Ventura-Soares C (2005) Artificial neural networks in wave predictions at the west coast of Portugal. *Comput Geosci* 31:415–424
- Mandal S, Prabakaran N (2006) Ocean wave forecasting using recurrent neural networks. *Ocean Eng* 33:1401–1410
- Martucci G, Carniel S, Chiggiato J, Sclavo M, Lionello P, Galati M (2009) Statistical trend analysis and extreme distribution of significant wave height from 1958 to 1999—an application to the Italian seas. *Ocean Sci Discuss* 6:2005–2036
- Meath SE, Aye L, Haritos N (2008) Accuracy of satellite-measured wave heights in the Australian region for wave power applications. *Bull Sci Technol Soc* 28:244–255
- Méndez FJ, Menéndez M, Luceño A, Losada IJ (2006) Estimation of the long-term variability of extreme significant wave height using time-dependent Peak Over Threshold (POT) model. *J Geophys Res* 111:C07024
- Méndez FJ, Menéndez M, Luceño A, Medina R, Graham NE (2008) Seasonality and duration in extreme value distributions of significant wave height. *Ocean Eng* 35:131–138
- Menéndez M, Méndez FJ, Izaguirre C, Luceño A, Losada IJ (2009) The influence of seasonality on estimating return values of significant wave height. *Coast Eng* 56:211–219
- Middleton JF, Thompson KR (1986) Return periods of extreme sea levels from short records. *J Geophys Res Ocean* 91:11707–11716
- Minoura M, Naito S (2006) Stochastic sea climate simulation based on hindcast data. In: Proceedings of the 16th international offshore and polar engineering conference (ISOPE 2006), The International Society of Offshore and Polar Engineering (ISOPE)
- Monbet V, Marteau PF (2001) Continuous space discrete time Markov models for multivariate sea state parameter processes. In: Proceedings of the 11th international offshore and polar engineering conference (ISOPE 2001), The International Society of Offshore and Polar Engineering (ISOPE)
- Monbet V, Prevosto M (2001) Bivariate simulation of non stationary and non Gaussian observed processes—application to sea state parameters. *Appl Ocean Res* 23:139–145
- Monbet V, Ailliot P, Prevosto M (2007) Survey of stochastic models for wind and sea state time series. *Probab Eng Mech* 22:113–126
- Muraleedharan G, Rao A, Kurup P, Unnikrishnan Nair N, Sinha M (2007) Modified Weibull distribution for maximum and significant wave height simulation and prediction. *Coast Eng* 54:630–638
- Nakićenović N, Alcamo J, Davis G, de Vries B, Fenhann J, Gaffin S, Gregory K, Grügler A, Jung TY, Kram T, La Rovere EL, Michaelis L, Mori S, Morita T, Pepper W, Pitcher H, Price L, Riahi K, Roehrl A, Rogner HH, Sankovski A, Schlesinger M, Shukla P, Smith S, Swart R, van Rooijen S, Victor N, Dadi Z (2000) Emissions scenarios. Cambridge University Press, Cambridge
- Natvig B, Tvete IF (2007) Bayesian hierarchical space-time modeling of earthquake data. *Methodol Comput Appl Probab* 9:89–114
- Nguyen V, Nguyen T, Ashkar F (2002) Regional frequency analysis of extreme rainfalls. *Water Sci Technol* 45:75–81
- Ochi MK (1998) Ocean waves the stochastic approach. Cambridge ocean technology series 6. Cambridge University Press, Cambridge
- Olsen A, Schrøter C, Jensen J (2006) Wave height distribution observed by ships in the North Atlantic. *Ships Offshore Struct* 1:1–12
- Özger M, Şen Z (2007) Triple diagram method for the prediction of wave height and period. *Ocean Eng* 34:1060–1068
- Panchang V, Zhao L, Demirbilek Z (1999) Estimation of extreme wave heights using GEOSAT measurements. *Ocean Eng* 26:205–225
- Perrie W, Jiang J, Long Z, Toulany B, Zhang W (2004) NW Atlantic wave estimates and climate change. In: Preprints of 8th international workshop on wave hindcasting and forecasting
- Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecol Modell* 190:231–259
- Prevosto M, Krogstad HE, Robin A (2000) Probability distributions for maximum wave and crest heights. *Coast Eng* 40:329–360
- Queffeuilou P (2004) Long-term validation of wave height measurements from altimeters. *Mar Geodesy* 27:495–510
- Renard B, Lang M, Bois P (2006) Statistical analysis of extreme events in a non-stationary context via a Bayesian framework: case study with peak-over-threshold data. *Stoch Environ Res Risk Assess* 21:97–112
- Ribereau P, Guillou A, Naveau P (2008) Estimating return levels from maxima of non-stationary random sequences using the Generalized PWM method. *Nonlinear Process Geophys* 15:1033–1039
- Richardson K, Steffen W, Schellnhuber HJ, Alcamo J, Barker T, Kammen DM, Leemans R, Liverman D, Munasinghe M, Osman-Elasha B, Stern N, Wæver O (2009) International scientific congress climate change: global risks, challenges & decisions—synthesis report. Technical report. International Alliance of Research Universities
- Robert CP, Casella G (2004) Monte Carlo statistical methods, 2nd edn. Springer, New York
- Rózyński G (2010) Long-term evolution of Baltic Sea wave climate near a coastal segment in Poland; its drivers and impacts. *Ocean Eng* 37:186–199
- Ruggiero P, Komar PD, Allan JC (2010) Increasing wave heights and extreme value projections: the wave climate of the U.S. Pacific Northwest. *Coast Eng* 57:539–552
- Rychlik I (2000) On some reliability applications of rice's formula for the intensity of level crossings. *Extremes* 3:331–348
- Rydén J (2006) A note on asymptotic approximations of distributions for maxima of wave crests. *Stoch Environ Res Risk Assess* 20:238–242
- Saf B (2010) Assessment of the effects of discordant sites on regional flood frequency analysis. *J Hydrol* 380:362–375
- Sahu SK, Gelfand AE, Holland DM (2007) High resolution space-time ozone modeling for assessing trends. *J Am Stat Assoc* 102:1212–1220
- Sang H, Gelfand AE (2009) Hierarchical modeling for extreme values observed over space and time. *Environ Ecol Stat* 16:407–426
- Sansó B, Schmidt AM, Nobre AA (2008) Bayesian spatio-temporal models based on discrete convolutions. *Can J Stat* 36:239–258
- Sasaki W, Iwasaki S, Matsuura T, Iizuka S (2005) Recent increase in summertime extreme wave heights in the western North Pacific. *Geophys Res Lett* 32:L15607

- Sasaki W, Hibiya T, Kayahara T (2006) Interannual variability and future projections of summertime ocean wave heights in the western North Pacific. *Ocean Sci Discuss* 3:1637–1651
- Savelieva E, Demyanov V, Kanevski M, Serre M, Christakos G (2005) BME-based uncertainty assessment of the Chernobyl fallout. *Geoderma* 128:312–324
- Schmidt AM, Gelfand AE (2003) A Bayesian coregionalization approach for multivariate pollutant data. *J Geophys Res* 108:STS 10.1–STS 10.8
- Scotto M, Guedes Soares C (2000) Modelling the long-term time series of significant wave height with non-linear threshold models. *Coast Eng* 40:313–327
- Scotto M, Guedes Soares C (2007) Bayesian inference for long-term prediction of significant wave height. *Coast Eng* 54:393–400
- Smith EL, Stephenson AG (2009) An extended gaussian max-stable process model for spatial extremes. *Journal of Statistical Planning and Inference* 139:1266–1275
- Soukissian TH, Kalantzi GD (2006) Extreme value analysis methods used for extreme wave prediction. In: Proceedings of the 16th international offshore and polar engineering conference (ISOPE 2006), The International Society of Offshore and Polar Engineering (ISOPE)
- Soukissian TH, Kalantzi GD (2007) A new method for applying the r-largest maxima model for design sea-state prediction. In: Proceedings of the 17th international offshore and polar engineering conference (ISOPE 2007), The International Society of Offshore and Polar Engineering (ISOPE)
- Soukissian TH, Photiadou CS (2006) A sensitivity analysis of the bottom-up algorithm for the segmentation of H_s -time series. In: Proceedings of the 16th international offshore and polar engineering conference (ISOPE 2006), The International Society of Offshore and Polar Engineering (ISOPE)
- Soukissian TH, Samalekos PE (2006) Analysis of the duration and intensity of sea states using segmentation of significant wave height time series. In: Proceedings of the 16th international offshore and polar engineering conference (ISOPE 2006), The International Society of Offshore and Polar Engineering (ISOPE)
- Soukissian TH, Kalantzi G, Karagali I (2006) De-clustering of H_s -time series for applying the peaks-over-threshold method. In: Proceedings of the 16th international offshore and polar engineering conference (ISOPE 2006), The International Society of Offshore and Polar Engineering (ISOPE)
- Stefanacos CN, Athanassoulis GA (2006) Extreme value predictions based on nonstationary time series of wave data. *Environmetrics* 17:25–46
- Stefanacos CN, Monbet V (2006) Estimation of wave height return periods using a nonstationary time series modelling. In: Proceedings of the 25th international conference on offshore mechanics and Arctic engineering (OMAE 2006), American Society of Mechanical Engineers (ASME)
- Stein EM, Stein JC (1991) Stock price distributions with stochastic volatility: an analytic approach. *Rev Financ Stud* 4:727–752
- Sterl A, Caires S (2005) Climatology, variability and extrema of ocean waves: the web-based KNMI/ERA-40 wave atlas. *Int J Climatol* 25:963–977
- Sterl A, Komen G, Cotton P (1998) Fifteen years of global wave hindcasts using winds from the European Centre for Medium-Range Weather Forecast reanalysis: validating the reanalyzed winds and assessing the wave climate. *J Geophys Res* 103:5477–5492
- Swail VR, Cox AT (2000) On the use of NCEP-NCAR reanalysis surface marine wind fields for a long-term North Atlantic wave hindcast. *J Atmos Ocean Technol* 17:532–545
- The WAMDI Group (1988) The WAM model—a third generation oceanwave prediction model. *J Phys Oceanogr* 18:1775–1810
- The WASA Group (1998) Changing waves and storms in the Northeast Atlantic. *Bull Am Meteorol Soc* 79:741–760
- The Wise Group, Cavaleri L, Alves JH, Ardhuin F, Babanin A, Banner M, Belibassakis K, Benoit M, Donelan M, Groeneweg J, Herbers T, Hwang P, Janssen P, Janssen T, Lavrenov I, Magne R, Monbaliu J, Onorato M, Polnikov V, Resio D, Rogers W, Sheremet A, McKee Smith J, Tolman H, van Vledder G, Wolf J, Young I (2007) Wave modelling—the state of the art. *Prog Oceanogr* 75:603–674
- Todorov V (2009) Estimation of continuous-time stochastic volatility models with jumps using high-frequency data. *J Econom* 148:131–148
- Torsethaugen K (1993) A two-peak wave spectral model. In: Proceedings of the 12th international conference on offshore mechanics and Arctic engineering (OMAE 1993), American Society of Mechanical Engineers (ASME)
- Torsethaugen K, Haver S (2004) Simplified double peak spectral model for ocean waves. In: Proceedings of the 14th international offshore and polar engineering conference (ISOPE 2004), The International Society of Offshore and Polar Engineering (ISOPE)
- Trulsen K (2006) Weakly nonlinear and stochastic properties of ocean wave fields. Application to an extreme wave event. In: Grue J, Trulsen K (eds) *Waves in geophysical fluids—tsunamis, rogue waves, internal waves and internal tides*, Chap 2. CISM courses and lectures no. 489. Springer, New York, pp 49–106
- Unami K, Abagale FK, Yangyuru M, Alam AHMB, Kranjac-Berisavljevic G (2010) A stochastic differential equation model for assessing drought and flood risks. *Stoch Environ Res Risk Assess* 24:725–733
- Uppala SM, Kållberg PW, Simmons AJ, Andrae U, Da Costa Bechtold V, Fiorino M, Gibson JK, Haseler J, Hernandez A, Kelly GA, Li X, Onogi K, Saarinen S, Sokka N, Allan RP, Andersson E, Arpe K, Balmaseda MA, Beljaars ACM, 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 BJ, Isaksen L, Janssen PAEM, Jenne R, McNally AP, Mahfouf JF, Morcrette JJ, Rayner NA, Saunders RW, Simon P, Sterl A, Trenberth KE, Untch A, Vasiljevic D, Vitebro P, Woolen J (2005) The ERA-40 re-analysis. *Q J R Meteorol Soc* 131:2961–3012
- Vaccaro S (2008) Position-dependent stochastic diffusion model of ion channel gating. *Phys Rev E* 78:061915
- Van Gelder P, De Ronde J, Neykov NM, Neytchev P (2001) Regional frequency analysis of extreme wave heights: trading space for time. In: Edge BL (ed) *Coastal engineering (2000)*, proceedings of the 27th international conference on coastal engineering held in Sydney, Australia, 16–21 July 2000, American Society of Civil Engineers, pp 1099–1112
- Vanem E (2010) Stochastic models for long-term prediction of extreme waves: a literature survey. In: Proceedings of the 29th international conference on offshore mechanics and Arctic engineering (OMAE 2010), American Society of Mechanical Engineers (ASME)
- von Storch H, Zwiers FW (1999) *Statistical analysis in climate research*. Cambridge University Press, Cambridge
- Walton TL, Borgman LE (1990) Simulation of nonstationary, non-Gaussian water levels on Great Lakes. *J Water Port Coast Ocean Eng* 116:664–685
- Wang XL, Swail VR (2001) Changes of extreme wave heights in northern hemisphere oceans and related atmospheric circulation regimes. *J Clim* 14:2204–2221
- Wang XL, Swail VR (2002) Trends of Atlantic wave extremes as simulated in a 40-yr wave hindcast using kinematically reanalyzed wind fields. *J Clim* 15:1020–1035
- Wang XL, Swail VR (2006a) Climate change signal and uncertainty in projections of ocean wave heights. *Clim Dyn* 26:109–126

- Wang XL, Swail VR (2006b) Historical and possible future changes of wave heights in northern hemisphere ocean. In: Perrie W (ed) Atmosphere-ocean interactions, vol 2, Chap 8. Advances in fluid mechanics, vol 39. WIT Press, pp 185–218
- Wang XJ, Zwiers FW, Swail VR (2004) North Atlantic ocean wave climate change scenarios for the twenty-first century. *J Clim* 17:2368–2383
- Wang XL, Swail VR, Zwiers FW, Zhang X, Feng Y (2009) Detection of external influences on trends of atmospheric storminess and northern oceans wave heights. *Clim Dyn* 32:189–203
- Wang XL, Swail VR, Cox A (2010) Dynamical versus statistical downscaling methods for ocean wave heights. *Int J Climatol* 30:317–332
- Weisse R, Stawarz M (2004) Long-term changes and potential for future developments of the North Sea wave climate. In: Preprints of 8th international workshop on wave hindcasting and forecasting
- West M, Harrison J (1997) Bayesian forecasting and dynamic models, 2nd edn. Springer-Verlag, New York
- Wikle CK (2003) Hierarchical models in environmental science. *Int Stat Rev* 71:181–199
- Wikle CK, Berliner LM, Cressie N (1998) Hierarchical bayesian space-time models. *Environ Ecol Stat* 5:117–154
- Wikle CK, Milliff RF, Nychka D, Berliner LM (2001) Spatiotemporal hierarchical bayesian modeling: tropical ocean surface winds. *J Am Stat Assoc* 96:382–397
- Wolf J, Woolf DK (2006) Waves and climate change in the north-east Atlantic. *Geophys Res Lett* 33:L06604
- Woolf D, Challenor P, Cotton P (2002) Variability and predictability of the North Atlantic wave climate. *J Geophys Res* 107:9-1–9-14
- Yang T, Shao Q, Hao ZC, Chen X, Zhang Z, Xu CY, Sun L (2010) Regional frequency analysis and spatio-temporal pattern characterization of rainfall extremes in the Pearl River Basin, China. *J Hydrol* 380:386–405
- Yoon S, Cho W, Heo JH, Kim CE (2010) A full Bayesian approach to generalized maximum likelihood estimation of generalized extreme value distribution. *Stoch Environ Res Risk Assess* 24:761–770
- Zhang H (2007) Maximum-likelihood estimation for multivariate spatial linear coregionalization models. *Environmetrics* 18: 125–139