Extremes and Trends in Wave Climate

- a regional and global study

Ole Johan Aarnes

Dissertation for the degree of Philosophiae Doctor (PhD)

Geophysical Institute
University of Bergen

October 2014

Dissertation date: 16th of Jan. 2015

Abstract

Wind generated surface waves represent a critical factor for offshore constructions and coastal development, and are highly relevant in scientific questions related to climate and weather. The sea state can be described by several parameters, but significant wave height is particularly important, traditionally defined by the average height of the highest third of individual waves. This study targets two aspects describing wave climate: extremes and trends in significant wave height.

The Norwegian Reanalysis (NORA10) - a combined high-resolution atmospheric downscaling and wave hindcast based on the ERA-40 reanalysis covering the Norwegian Sea, the North Sea, and the Barents Sea is presented and validated. The wind-wave data archive, spanning the period September 1957 to August 2002, shows a significant improvement from ERA-40 over the whole range of data, but particularly in upper percentiles. Given the performance of NORA10, it provides a baseline climatology of significant wave height, which other datasets may be compared against.

An extended version of NORA10, covering 1958 to 2009, is utilized to obtain 100-year return value estimates of significant wave height. The analysis explores three different approaches, where the applied extreme value distribution is dictated by the chosen subset of the initial data. This is done by: (*i*) peaks-over-threshold; (*ii*) annual maxima; and (*iii*) the *r* largest order statistic within blocks of one year. These subsets should conform to the generalized Pareto distribution (*i*) and the generalized extreme value distribution (*ii/iii*), respectively. By assuming stationary conditions, the three models provide results mainly within $\pm 5\%$. In areas where the discrepancy is larger, (*iii*) is found less satisfactory. Method (*i*) yields good conformity when the threshold is set high. Here, the 99.7-percentile is applied. Given the better use of data, (*i*) is preferred over (*ii*). Based on (*i*), the 100-year estimates are peaking around 22 m northwest of Scotland, around 14 m in the North Sea and above 16 m in the Norwegian Sea.

The robustness of a return value estimates often depend on the amount of available data. As most conventional time series of significant wave height do not extend further than 50 years, 100-year return value estimates can only be obtained by extrapolating some fitted theoretical distribution. In this study, a method for estimating return values from aggregated ensemble forecasts is presented. The archived ensemble forecasts originate from the European Centre for Medium-Range Weather Forecasts (ECMWF) and consist of 51 members run twice daily. To ensure that the aggregates are independent representations of the model climate, only 10-day forecasts are retained. By assuming each forecast being representative of a 6-hour interval, collectively 1 year of ensemble forecasts are representative of more than 25 years of data. In two separate papers, aggregates of ensemble forecasts, equivalent to more than 220 years of data, are utilized to obtain 100-year return value estimates of significant wave height and wind speed. The datasets are carefully validated against *in situ* measurements and al-

timeter data, and found representative of the observed climate. Return value estimates are obtained by traditional extreme value models, but also obtained directly from order statistics as the dataset, on average, should contain more than 2 events exceeding the 100-year return value. It is shown that the estimates come close in equalizing corresponding estimates from NORA10, and yield significant improvements compared to ERA-40 and ERA-I.

Reanalyses may be considered homogeneous in the sense that they run with the same model configuration. However, they are highly dependent on assimilation and may suffer from the ever growing observational system. Ideally, assimilation should only be a mean to reduce random errors, but if a model exhibits systematic errors, data assimilation may correct bias. Once a model is run in forecast mode, the effect of assimilation is usually lost, as the model relax towards it biased state. ERA-I is a coupled wind-wave reanalysis produced at ECMWF, spanning the period 1979 onwards. Besides the reanalysis, ERA-I is run as 10-day forecasts. Altimeter observations are the only kind of wave data assimilated, first introduced in August 1991. Here, trends based on different forecast ranges are compared and validated against observations. It is shown that trends in significant wave height from analysis are highly affected by the transition in August 1991, especially in the northeast Atlantic and the eastern tropical Pacific. Finally, the 48-hour forecast range is proposed as a better candidate to obtain realistic trends.

Scientific environment

This PhD study has been carried out at the Norwegian Meteorological Institute (MET-Norway), Division for Oceanography and Maritime Meteorology, and at the University of Bergen (UoB). It has to a large extent been a free standing project financed by MET-Norway, with partial funding from the Norwegian Center for Offshore Wind Energy (NORCOWE) and the Extreme Waves and Climate Change (ExWaCli) project, both funded by the Norwegian Research Council (project numbers: 193821 and 226239). A total of two months were spent at ECMWF in Reading, UK.

Acknowledgements

This PhD would not have been possible without the support from several of my colleagues at the Division for Oceanography and Maritime Meteorology. First and foremost, I would like to express my gratitude to my co-supervisor Øyvind Breivik for pushing me through this challenging period of my professional life. Without your help and energy, this project would probably have come to a standstill. Great thanks to my supervisor Birgitte Furevik, without your initiative, I would not have had this opportunity. Thanks to Magnar Reistad, your level of knowledge and calm demeanour truly creates a pleasant work environment. Further, thanks to Anne Karin Magnusson who introduced me to MET-Norway as a student back in 2003. And, not at least, thanks to all my other colleagues who make me proud of being part of *Vervarslinga på Vestlandet*.

I also like to thank my co-supervisor Tore Furevik. And, Jean Bidlot, Saleh Abdalla and Peter Janssen for making me feel welcome at ECMWF.

These last few years have been a juggle between this thesis, my work at MET-Norway, building a family and creating a home. Through it all, I have found great comfort in being blessed with three beautiful children and the best wife one could ever wish for. Thanks to my in-laws who have been helpful when I have spent long hours at the office.

Mom and dad, thanks for all your support. I'm ever grateful. This thesis is dedicated to my mother who passed away all to early.

Contents

Chapter 1

Outline

The purpose of this thesis is to investigate return value estimates and trends in significant wave height globally and with special emphasis on the northeast Atlantic. Focus is being made on improving corresponding estimates based on well established reanalyses. This is conducted by utilizing conventional and unconventional data, and different statistical methods to highlight uncertainty in the aforementioned estimates.

The manuscript is organized as follows. Chapter 2 gives a brief introduction, accompanied by the main objectives of the thesis. Chapter 3 provides the relevant scientific background. The applied methodology is presented in Chapter 4, while Chapter 5 offers a summary of results and future perspectives. The five papers included in this thesis are listed below and will in the following be referred to by their Roman letters.

Paper I: A high-resolution hindcast of wind and waves for the North Sea, the Norwegian Sea and the Barents Sea
Magnar Reistad, Øyvind Breivik, Hilde Haakenstad, Ole Johan Aarnes, Birgitte R. Furevik and Jean-Raymond Bidlot
J. Geophys. Res., **116**, C05019, 2011

Paper II: Wave Extremes in the Northeast Atlantic
Ole Johan Aarnes, Øyvind Breivik and Magnar Reistad
J. Climate, **25**, 1529–1543, 2012

Paper III: Wave Extremes in the Northeast Atlantic from Ensemble Forecasts
Øyvind Breivik, Ole Johan Aarnes, Jean-Raymond Bidlot, Ana Carrasco and Øyvind Sætra
J. Climate, **26**, 7525–7540, 2013

Paper IV: Wind and wave extremes over the world oceans from very large ensembles
Øyvind Breivik, Ole Johan Aarnes, Saleh Abdalla, Jean-Raymond Bidlot and Peter A.E.M. Janssen
Geophys. Res. Lett., **41**, 5122–5131, 2014

Paper V: Marine wind and wave height trends at different ERA-Interim forecast ranges
Ole Johan Aarnes, Saleh Abdalla, Jean-Raymond Bidlot and Øyvind Breivik
J. Climate (accepted)

Chapter 2

Introduction and objectives

Wind generated surface waves are probably the most striking evidence of the air-sea interaction taking place at the boundary between the atmosphere and ocean. They range in size and shape, from tiny ripples seen on a calm day to violent breaking waves more often associated with stormy conditions. Simplistically, waves are the combined effect of wind and gravity. As wind blows across the ocean surface, momentum is transferred from air to water by wind induced pressure differences, creating a resonance effect, forcing the surface away from its flat initial state. While the restoring force of gravity will strive to maintain equilibrium, inertia will make the surface overshoot/undershoot, creating a oscillating effect. This process and its demise play an important role in weather and climate by exchanging heat and gas between the atmosphere and ocean.

In order to describe the sea state, we rely primarily on a few key parameters that are observable with the human eye. These are: wave height - the vertical displacement between an adjacent crest and trough; wave period - the time elapsed between two successive crossings of the mean surface elevation (zero-crossing); and wave direction - the heading of a wave. However, as the surface elevation to a large extent is a stochastic process, it is not practical to describe the wave condition by the measure of individual waves. More often the sea state is defined by the average conditions taken over some period of time where it is reasonable to assume stationarity, ideally between 15-35 minutes (WMO-group, 1998). From this, the significant wave height is defined as the average height of the highest third of individual waves. Mean period and mean wave direction are obtained in a similar manner. In the following thesis, significant wave height (H_s) is the primary parameter of interest. Maximum individual wave height, important in estimating structural impact loads, is more often obtained by empirical relations and statistical distributions having dependency on H_s (Forristall, 1978). Similarly, freak waves, or rogue waves, are beyond the scope of this work. This study does however emphasise wind, as it is the driving force in generating waves.

In a direct sense, waves are important to humans on different levels. They can affect one's daily recreational activities or have substantial socio-economic impact, e.g in conjunction with coastal settlement in areas exposed to erosion and deposition. Commercial businesses often take a special interest in waves, as wave-induced loads may lead to economic loss, contamination or even fatality. For instance, as the aquaculture is moving further out to the coast, and into harsher wave climate, there is an increasing concern for cultured fish escaping and spreading disease. If we are going to stabilize global greenhouse gas emissions as advised by the United Nations Framework Con-

vention on Climate Change, we need to a see a rapid increase in renewable energy exploitation. With the advent of offshore wind farms and wave energy converters in deep water, a new set of wave related challenges arise. The international shipping community is well aware of the dangers associated with waves, but has historically struggled to come to a common classification standard. While ships have the advantage of being mobile, the offshore petroleum industry is usually confined to fixed locations unable to avoid major weather events. The Petroleum Safety Authority in Norway (PTIL), an independent government regulator, is responsible for maintaining a safe operation. All petroleum related industry on the Norwegian Continental Shelf are imposed by PTIL to meet the recommendation stated by the NORSOK-standards. Typically, structural design and mooring dimensioning must withstand wave events with an annual probability of exceedance equal to 10^{-2} and 10^{-4}, depending on the platform being manned or not.

In order to make return value estimates of H_s with an average recurrence rate once every 100 or 10,000 years, there is a substantial need for long reliable data records. This is not easily achieved, as the first wave observations became available in the late 70s from *in situ* measurements (Gemmrich et al., 2011), later followed by platform-mounted instruments. Satellite-borne altimetery was not operational before the mid 80s. The advent of meteorological reanalyses in the 90s offered a much appreciated contribution. By running numerical weather models over an extended period of time, firmly controlled by historic observations, climate research obtained a relatively coherent dataset, regular both in time and space. Still, reanalyses are only approximate representations of the real climate, and have until recently not covered more than ~ 60 years. They are affected by the ever-improving observational system and have proven inadequate in representing extreme wind and wave events, mainly due to resolution limitations (Cox and Swail, 2001; Caires et al., 2004).

Firstly, the robustness of return value estimates of H_s will primarily rely on three factors: the quality of the data; the size of the dataset; and the choice of statistical method. All three aspects are considered herein. Secondly, reported trends in mean H_s are typically on the order of a few centimetres per year, at most, depending on the area and period in question (Young et al., 2011; Wang et al., 2012; Dodet et al., 2010; Dupuis et al., 2006). If assimilation can impose step changes of some decimetres, this will significantly affect trends.

This PhD thesis is focused around limitations and side effects associated with H_s-data obtained from reanalyses, particularly related to: (*i*) their ability to represent extreme conditions and (*ii*) how inherent trends are influenced by assimilation. The main motivation is to overcome these deficiencies to obtain better estimates of return values and trends. Given the obvious relationship between wind and waves, the majority of the work is supported by corresponding analyses conducted on wind speed.

Objectives

The objective of this thesis has been twofold:

1. Identify regional and global return value estimates of significant wave height by:
    - utilizing a high-quality hindcast developed at the Norwegian Meteorological Institute.
    - investigating the usefulness of unconventional data, in the form of archived ensemble forecasts, to obtain vast datasets.
    - validating different statistical models.
2. Investigate trends in significant wave height from ERA-I and its influence of non-stationary assimilation by:
    - comparing trends obtained at analysis and increased forecast range.

Chapter 3

Scientific background

3.1 Global wave climate

In deep water, wave growth is primarily controlled by three factors: wind speed; fetch - the open stretch of sea surface forced by the wind; and duration - the time the sea surface is affected by wind. With wind speed being the most important factor, the wave climate tends to follow the wind climate. In Figure 3.1 it is shown that the majority of wave energy is located within the mid-latitudes of the Northern and Southern Hemisphere (NH/SH), particularly near the storm tracks (Sterl and Caires, 2005; Stopa et al., 2013). However, as waves are unaffected by the Coriolis effect (wave lengths are rarely above 1 km long), they propagate along great circles until encountering land or shoals, influencing oceans outside their area of generation. The local sea state is often a combination of waves generated locally - *wind sea*, and those generated by distant weather systems - *swell*. As this definition is somewhat imprecise, wave models tend to define swell as wave energy propagating faster, or in another direction than the local wind. While there can be several swell systems present at one time, there can only be one wind sea system. In general, the stronger the wind climate, the more wind sea influenced the wave climate is. Typically, wind speed and significant wave height of the mid-latitudes are more correlated than in the swell-dominated Tropics (Semedo et al., 2011).

As illustrated in Figure 3.1, the annual mean wind and wave conditions are highest in the SH. This is related to a relatively higher density of extratropical cyclones surrounding Antarctica compared to the NH (Bengtsson et al., 2006; Ulbrich et al., 2009).

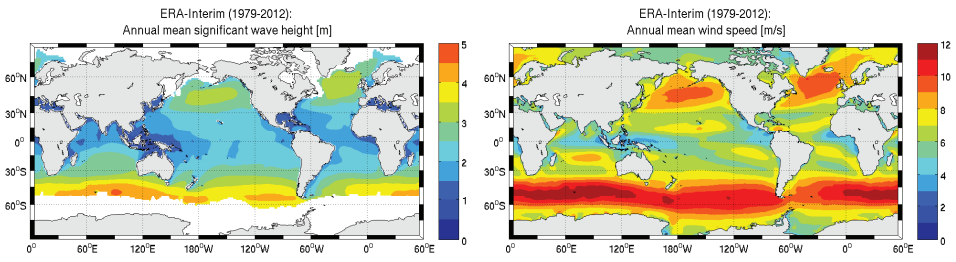

Figure 3.1: Annual mean H_s and U_{10} based on ERA-Interim over the period 1979-2012. Maximum ice extent is illustrated by the H_s data, in white.

Even though the total number of extratropical cyclones per year in the SH are only marginally higher than in the NH (Hodges et al., 2011), the typical life span of a low-pressure system in the SH is longer as these systems move almost unaffected by land (near 60°S), and are therefore able to generate more wave energy.

In terms of annual wind and wave variability, the situation is reversed. The highest extratropical cyclone intensity is found in the North Pacific and the North Atlantic Bengtsson et al. (2006). These low-pressure systems are generated (cyclogenesis) or intensified by the baroclinic instability imposed by the warm Kuroshio and Gulf Stream, east of Japan and off Cape Hatteras/Newfoundland, respectively. This implies that the highest waves and wind speeds globally are found downstream of these areas (Moore et al., 2008; Sampe and Xie, 2007; Alves and Young, 2003; Caires and Sterl, 2005a; Young et al., 2012; Breivik et al., 2014). Here, the nature of these extreme events is investigated, reduced to a matter of estimating the highest wave conditions that may be expected during a certain period of time. In doing so, different types of wave data are assessed.

3.2 Observations

Wave observations are typically obtained in two ways, by *in situ* or remote sensing techniques (Holthuijsen, 2007). The first category include instruments located at the surface (buoys), below the surface (inverted echo-sounders, pressure sensors) or surface-piercing devices (wave poles). All instruments that measure waves from above the surface are by definition remote sensing. These are typically installed onboard ships and platforms, or satellite-mounted. Still, platform-mounted devices are sometimes referred to as *in situ* measurements, due to their vicinity to the surface. A few places this is also the case in this study. However, wave height observations are mainly obtained from buoys and satellite-borne radar altimetry.

Wave buoys are surface following devices which measure acceleration. By integrating the vertical acceleration twice, the vertical motion of the surface elevation (η) is obtained. Some buoys also contain an electronic compass and two additional accelerometers in order to measure the directional components of the wave field. Significant wave height is usually defined as:

$$H_s = 4\mathrm{std}(\eta(t)) \tag{3.1}$$

which is equivalent to the average of the highest third of the waves over time (t).

Radar altimetry measures the vertical displacement directly. However, when performed from space, the footprint of the beam is too large to represent the surface elevation at one single location. Instead, the reflected pulse will represent the roughness of the surface over a larger area, typically a few kilometres. By measuring the shape of the returning signal, an estimate of H_s is obtained. Basically, higher waves lead to a longer return signal, as opposed to smaller waves, which in turn is empirically associated with some H_s (Tucker, 1991; Carter, 1993).

3.3 Wave models

As the evolution of individual waves is too computationally expensive to solve operationally, wave models attempt to predict the statistics of the wave field, i.e. how wave amplitude (surface elevation) is related to individual wave components of different frequencies (f) and directions (θ). The variance density spectrum, $E(f,\theta)$, offers a continuous representation of the surface elevation, formally known as the two-dimensional (2D) wave spectrum. By integrating the 2D wave spectrum over all frequencies and directions, the total variance of the surface elevation is obtained, also known as the zeroth-order moment

$$m_0 = \int_0^{2\pi} \int_0^{\infty} E(f,\theta) df d\theta \tag{3.2}$$

Based on Eq. 3.1, the significant wave height is defined by

$$H_{m_0} = 4\sqrt{m_0} \tag{3.3}$$

Wave modelling dates back to the 50s (Gelci et al., 1957), but has come a long way since then. Today most wave models are of the third-generation type, with no *ad hoc* assumption being made on the shape of the 2D wave spectrum. The most commonly used wave models in forecasting are WAM (WAMDI Group, 1988; Komen et al., 1994), WaveWatch III (WW3) (Tolman, 2009) and SWAN (Holthuijsen, 2007). In deep water, mainly three source terms control the evolution of the 2D spectrum, i.e. wind input (S_{in}), non-linear interaction (S_{nl}) and wave dissipation (S_{ds}). Simplified, the evolution of the 2D wave spectrum can be expressed by:

$$\frac{dE(f,\theta)}{dt} = S_{in} + S_{nl} + S_{ds} \tag{3.4}$$

A comprehensive description of how these source functions are defined and solved, is presented in Cavaleri et al. (2007).

3.4 Reanalyses and hindcasts

In situ wave observations obtained from buoys are a rare commodity. The majority of wave buoys are located in coastal areas or in conjunction with oil rigs (see National Data Buoy Center - NOAA (2014) for the most comprehensive collection of real-time wave buoy data). While their temporal coverage is often quite good, the scarcity of data in the open ocean offers an insufficient spatial representation for wave climate assessments. For this reason, reanalyses and hindcasts are attractive substitutes for observational records. These data archives are built by running coupled numerical weather prediction (NWP) models and wave models to recreate a historic period. This is usually conducted in two ways. In a stand-alone system, winds obtained from the NWP model force the wave model without any feedback to the NWP. In this way, the NWP and the wave model can be run in two separate operations. In a coupled system, the sea surface roughness used by the atmospheric component is dependent on the sea state defined by the wave model, which again affects the winds of the NWP (Janssen, 2004). This set-up requires the models to run simultaneously.

Table 3.1: Commonly applied reanalyses.

Reanalysis	**Period**	**Resolution**	**Reference**
NCEP-NCAR R1	1948-	$\sim$ 210 km	Kalnay et al. (1996)
NCEP-DOE R2	1979-	$\sim$ 210 km	Kanamitsu et al. (2002)
CFSR	1979-	$\sim$ 38 km	Saha et al. (2010)
JRA-25	1979-2004	$\sim$ 125 km	Onogi et al. (2007)
JRA-55	1958-2012	$\sim$ 50 km	Ebita et al. (2011)
MERRA	1979-	$\sim$ 50 km	Rienecker et al. (2011)
ERA-15	1979-1993	$\sim$ 125 km	Gibson et al. (1997)
ERA-40	1957/09-2002/08	$\sim$ 125 km	Uppala et al. (2005)
ERA-Interim	1979-	$\sim$ 79 km	Dee et al. (2011)
NOAA-20CRv2	1871-2011	$\sim$ 210 km	Compo et al. (2011)

In order to recreate the past weather on a day-to-day basis, data assimilation is essential. This is evident from traditional weather forecasting where the skill of the NWP rapidly deteriorates with forecast range (FCR). Operationally, NWP models combine the last available forecast (first guess) with meteorological observations to create an optimal estimate of the atmospheric state at the time of analysis. This may be conducted according to different assimilation methods, e.g. Optimum Interpolation (Lionello et al., 1992), 3 dimensional variational analysis and 4 dimensional variational analysis (Kalnay, 2003). Reanalyses are essentially run in the same way, but have the advantage of continuous data assimilation. In this way, model limitations and inaccuracies imposed in the previous analysis step are unable to evolve forward in time. Since wave models represent a forced system, primarily relying on the quality of the wind, wave data assimilation is secondary and therefore more often run without assimilation, i.e. as hindcasts (Cox and Swail, 2001; Swail and Cox, 2000; Reistad et al., 2011; Chawla et al., 2013).

While the model configurations used in operational NWPs are updated and improved regularly, models used in atmospheric reanalyses are kept unchanged over the period in question with the aim of creating homogeneous datasets. However, as reanalyses are highly dependent on data assimilation, inhomogeneities are difficult to avoid. Since the International Geophysical Year (IGY), 1st of July 1957 to the 31st of December 1958, there has been a tremendous increase of meteorological observations, mainly due to the advent of satellite-borne measurements. This has led to a more effective model-bias correction over the years. However, observations, like models, possess biases that often need to be calibrated before assimilation, a non-trivial task in itself (Zieger et al., 2009; Wan et al., 2010; Vincent et al., 2012).

Because most reanalyses span considerable periods of time, they are computationally expensive to run. Usually developers need to resort to coarser model resolution than normally applied in operational forecast models at the time. There are a number of well established global atmospheric reanalyses that have been extensively used in climate monitoring and research, see Table 3.1. For a more complementary summary, see Dee (2013). None of these provide wave data, but offer forcing data for wave hindcasts (Swail and Cox, 2000; Cox and Swail, 2001; Dodet et al., 2010; Chawla et al., 2013; Caires et al., 2004).

The European Center for Medium-Range Weather Forecasts (ECMWF) is the only center running coupled atmosphere-wave reanalyses. These reanalyses are directly or indirectly incorporated in all 5 papers presented herein. A brief introduction follows below.

3.4.1 European Reanalysis - ERA

The ERA project is a series of reanalyses developed at ECMWF (Gibson et al., 1997; Uppala et al., 2005; Dee et al., 2011). It was motivated by the need for a dataset generated by a modern, consistent and invariant data assimilation system. Up until the mid-nineties, most climate monitoring was based on available observations and archived operational forecasts coming out of models with continuously updated configurations. As a result, the atmospheric reanalysis ERA-15 was completed in 1996 (Gibson et al., 1997). It covered the period December 1978 through February 1994 and had a resolution of $\sim$ 125 km. Even though widely recognised, several deficiencies were detected in ERA-15 (Kallberg et al., 2004). Sterl et al. (1998) used ERA-15-winds to force a stand-alone WAM model, creating the first global wave hindcast. The performance of the wave model was then used to infer the quality of the ERA-15-winds. It was concluded that winds were too weak in areas of high winds due to resolution limitations. In areas of weak winds the study was inconclusive, but it was suggested that WAM itself tended to overestimate wave heights in these areas.

By 2002, a second generation reanalysis was available. The improved and extended ERA-40 was the first reanalysis run with a fully coupled atmosphere-wave model (Uppala et al., 2005; Janssen, 2004), spanning the period September 1957 to August 2002, a total of 45 years. By replacing the Optimal Interpolation (OI) scheme used in ERA-15 with a 3D-Var analysis, the available observations were better utilized, increasing the performance. Caires et al. (2004) carried out a comprehensive intercomparison of different wind-wave hindcasts, including the ERA-40 reanalysis. It was concluded that ERA-40 was best suited to describe global wind and wave height variability and to conduct detailed analyses. In terms of trends, the different datasets offered similar estimates. Still, by collocating ERA-40 with *in situ* buoy measurements and TOPEX altimeter wind and wave height data, Caires and Sterl (2003) confirmed that ERA-40 wave heights are too low at the high percentiles, and also overestimates at the low end. As a consequence, Caires and Sterl (2005b) proposed a new non-parametric method developed to correct these defects, and to compensate for inhomogeneities imposed by assimilated altimeter wave observations. For instance, ERA-40 H_s are of lower quality from December 1991 to May 1993 due to impaired quality in the assimilated ERS-1 wave altimeter data (Uppala et al., 2005). The corrected ERA-40 (C-ERA-40) showed clear improvements over the entire range of wave data, both in scatter and bias.

ERA-Interim (ERA-I) is the third reanalysis produced within the ERA-series and was intended as a preparation for the successor of ERA-40 (Dee et al., 2011). Initially, the reanalysis covered the period 1989 and onwards, but it has since been extended back to 1979. It is based on the Integrated Forecast System (IFS) release Cy31r2, which was operational at ECMWF from 12 December 2006 until 5 June 2007. The most relevant model updates since ERA-40 (Cy23r4) are presented in Dee et al. (2011) (Table II). ERA-I has an increased spatial resolution of $\sim$ 79 km and uses a 4D-Var assimilation scheme. Even though the majority of observations assimilated in ERA-I are the same

as those used in ERA-40 (up until 2002), there are some improvements. The ERA-I project is ongoing and updated in near real-time. An intercomparison study of ERA-I and CFSR marine surface wind speed (U_{10}) and H_s is presented in Stopa and Cheung (2014). Unlike the coupled ERA-I, the CFSR wave data is obtained with a hindcast using WW3. This study showed that ERA-I, like its predecessors, generally underestimates wind and waves at the upper percentiles, while CFSR revealed a small positive bias. By investigating error trends relative to altimeter data, the study concluded ERA-I to be more homogeneous in time and therefore more suitable for analysing long-term variability. This was primarily based on comparison with TOPEX and GFO data, first available in 1992. However, altimeter wave data, the only type of wave data assimilated in ERA-I, was first introduced in August 1991. The homogeneity of ERA-I H_s is mainly affected by this transition, which this study does not capture. This issue is treated in **Paper V**.

3.4.2 NORA10

The Norwegian Reanalysis 10 km (NORA10) is a regional dynamical downscaling of ERA-40's atmospheric component, originally spanning the period September 1957 through August 2002. Beyond this date, the archive utilizes operational analyses from the ECMWF IFS and is regularly extended to the present day. Computed winds are further used to force a stand-alone wave model, forming a coherent wind and wave archive covering the northeast Atlantic, i.e an area west of Ireland and to the north, including the Norwegian Sea, the North Sea and the Barents Sea. NORA10 is developed by the Norwegian Meteorological Institute and funded by the Norwegian Deepwater Program[1] (NDP), a consortium from the Norwegian offshore petroleum industry. The project was primarily motivated by the need for high quality wind and wave data, crucial for surveys of metocean conditions. Besides basic statistics, the archive is typically used to calculate average weather windows, enabling more cost effective operations offshore. The data are frequently used in extreme value analysis in conjunction with mooring analysis and structural design. Getting the high percentiles and extremes right were always a prime motivation.

The atmospheric downscaling is performed with the High Resolution Limited Area Model (HIRLAM) version 6.4.2 (Undén et al., 2002), on a 10-11 km spatial resolution. The model is forced by ERA-40 on the boundaries with temperature, wind, specific humidity and cloud water at all 40 vertical levels, and surface pressure with 6-hourly temporal resolution. HIRLAM is run in 9-hourly sequences, but the last 3 hours are primarily used to obtain better precipitation data. The initial condition is a blend of the preceding 6-h forecast (first guess) and the initialization increment between the two filtered states of ERA-40 and first guess. This is intended to preserve quickly evolving systems and better represent intense extratropical cyclones.

The wave hindcast is generated with a modified WAM Cycle 4 model (Günther et al., 1992; Komen et al., 1994) and run in a nested mode. A coarser 50 km model forced by ERA-40-winds (WAM50), covering the main wave generating area of the North Atlantic, provides boundary conditions to the inner model which is forced by HIRLAM winds and covers the same area (WAM10), see Breivik et al. (2009) for a

[1] www.ndwp.org

description of the handling of the open boundaries. Unlike ERA-40, WAM10 is run with shallow water physics and has twice the directional resolution, i.e. 15°/24 bins.

Paper I presents an in-depth description of NORA10 and validation. **Paper II** uses NORA10 to obtain 100-year return value estimates of H_s, which is further utilized in **Paper III** to validate return values estimates of H_s based on archived ECMWF ensemble data (presented below).

3.5 Ensemble prediction

The forecast skill of NWP models is constantly improving as our ability to measure and model atmospheric and oceanic processes evolves. Increased computational power allows NWP models to run with ever higher resolution, explicitly resolving smaller processes. Still, the chaotic nature of the atmosphere and ocean will always add an element of unpredictability which is not represented by deterministic models. This uncertainty is mainly associated with two factors: limited knowledge about the initial state of the atmosphere/ocean and necessary model simplifications and approximations. By running essentially the same model with perturbed initial conditions and a perturbed model, this uncertainty may be quantified by the spread of the different forecasts. Being computational heavy, ensembles are often run with coarser resolution. Probabilistic forecasts in the form of ensembles have been run operationally at ECMWF (Molteni et al., 1996; Buizza et al., 2007, 2005) and US National Centers for Environmental Prediction (NCEP) (Toth and Kalnay, 1993, 1997) since 1992. As of 2014, ECMWF runs their Ensemble Prediction System (ENS) for medium-range and monthly forecasts, and System 4 (S4) for seasonal forecasts (Molteni et al., 2011). The medium-range forecast is initialized twice a day, at 00 UTC and 12 UTC, and extends 15 days ahead. The atmospheric model is run with a horizontal resolution of $\sim$ 32 km at the surface and with a coarser resolution of $\sim$ 63 km beyond day 10, while the coupled wave component uses the coarser resolution throughout the whole period. The ensemble comprises 51 members, where 50 are run with slightly perturbed conditions of the atmospheric component. The last member is started from the best estimate of the initial state (control), corresponding to the initial state used in the high-resolution (HRES) deterministic model. Any deviation between the control and HRES reflects resolution issues. Twice a week the medium-range runs are extended up to 32 days to create monthly forecasts (Vitart, 2004), while long range seasonal forecasts (1-7 months) are produced once a month.

An example of a medium range ENS forecast of H_s is presented in Figure 3.2. Unlike the atmospheric component, the wave model is not initialized from perturbed conditions, which is reflected by the lack of spread at analysis (ANA). The H_s-forecasts are only affected by different atmospheric forcing. It is further illustrated that the spread of the ensemble is generally increasing with time, representing lower confidence with forecast range (FCR). At day 10, indicated by the black ellipse, the forecast skill of the ENS data is more or less non-existing. The 10-day forecasts constitute the data basis in **Paper III** and **Paper IV**.

As models are integrated forward in time, they are prone to drift, offering slightly different model climate at increased FCR. This implies that prognoses may be interpreted slightly differently depending on FCR. ECMWF has tackled this issue by re-

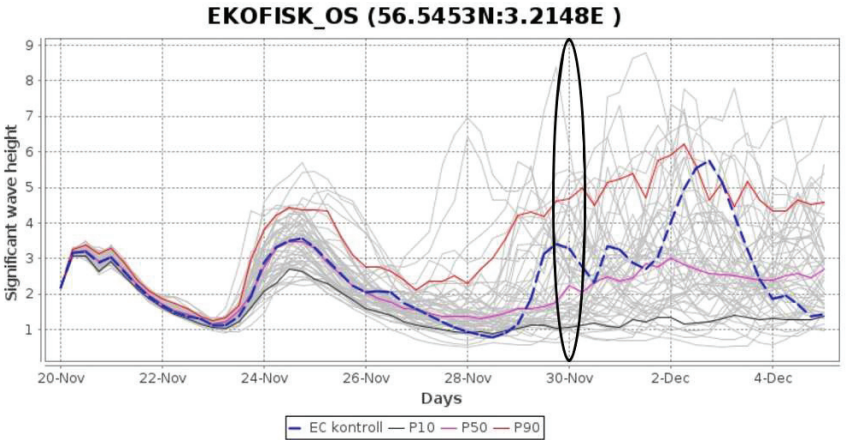

Figure 3.2: Plume diagram of ENS significant wave height at Ekofisk. Each individual forecast is plotted in gray. The control run is presented by the blue dashed line. The 10-, 50- and 90-percentile of the ensemble is illustrated in black, magenta and red, respectively. The 10-day forecast range is indicated by the black ellipse.

running the model configuration in question over a historic period (re-forecasts) (Hagedorn, 2008; Hagedorn et al., 2012). In this way, a representative model climatology at the desired FCR is obtained, in which the real-time ensemble may be compared against. For instance, the extreme forecast index (EFI) is an index representing the shift in distribution between the real-time ensemble and the model climate, ranging from -1 to 1 (Lalaurette, 2003; Prates and Buizza, 2011). In the extreme case where all members of the real-time ensemble are higher than the model climate, the EFI equals one. The model climate is based on five consecutive 32-day re-runs starting Thursdays, centred around the actual prediction time. This is done for the last 18 years, where an ensemble of 5 members are initiated by the ERA-I reanalyses, consisting of four slightly perturbed initial conditions. This offers a model climatology consisting of 5 members x 5 weeks x 18 years, a total of 450 forecasts representative of the FCR and time of year. The obvious advantage of such an index is that model bias, if any, is accounted for without the user needing to be familiar with the performance of the model.

3.6 Return value and trend estimates

100-year return value estimates of H_s

Knowledge about the "worst case scenario" has special relevance in aspects concerning offshore and coastal planning. To define such an event (at a given location), it is often convenient and necessary to relate the estimate to a time period, asking questions such as: what is the highest expected H_s over a period of 100 years? Given the limited availability of wave data, return value estimates of H_s are hard to verify, as the events one seeks to describe usually are larger than those recorded. In this regard, statistical extreme value models are crucial. By fitting theoretical probability distributions to the

data at hand, exceedance levels beyond the span of the data may be investigated.

Historically, different methods have been applied to obtain return value estimates of H_s (Mathiesen et al., 1994; Lopatoukhin et al., 2000; Soukissian and Kalantz, 2006). These can be distinguished by how the raw data is managed. Usually, this is done in three ways, either by using all data, also known as the *initial distribution method* (IDM), by using threshold exceedances, known as *peaks-over-threshold* (POT), or finally, by only retaining the highest value within blocks, where a block refers to some period of time. In this study, we have only applied the latter two, which are presented in Section 4.1 and applied in **Paper II**, **Paper III** and **Paper IV**.

The biggest advantage of the IDM, is that it is not wasteful. It exploits all data. Historically, this has been an attractive approach, since the availability of H_s-data have been limited. However, as most time series of H_s obtained from *in situ* measurements, hindcasts and reanalyses are given at 1-6 hours intervals, the data are serially correlated. This violates the requirement that data should represent independent events or storms. In order to come to terms with this, some decorrelation time scale needs to be applied, which is usually set to 3 hours (Alves and Young, 2003; Cooper and Forristall, 1997; Tucker, 1991; Vinoth and Young, 2011). Even though this is rarely true, 100-year return value estimates obtained with the IDM often correspond fairly well with estimates obtained with POT or by blocking, which are in fact representing independent data. However, the IDM offers usually higher return value estimates when the return period is reduced to 1 or 10 years. This may indicate that also 100-year return value estimates will be biased high as H_s time series grow longer in the future. Also, there is no theoretical justification for a specific distribution function representing the IDM data. Usually, several different functions are tested, where the function providing the best fit to the data are chosen (Vinoth and Young, 2011; Alves and Young, 2003). Further, since the IDM is highly influenced by the bulk of the data, or the lower part of the H_s-distribution, this may have implications when waves are generated by different weather systems. For instance, in areas influenced by both tropical and extratropical cyclones, the H_s-data belong to two different populations. Since extratropical storms are more frequent, the H_s-data generated from these systems will be more weighted when fitting a distribution function. If so, there is a big risk that the 100-year return value estimate will be underestimated as they are typically related to tropical cyclones. Similarly, in areas where water depth becomes a limiting factor on wave growth, the H_s-data may be a combination of deep water waves and shallow water waves. If the majority of H_s-data is unaffected by water depth, the 100-year return value estimates, representing shallow water waves, are likely to be too high.

A recent paper by Vinoth and Young (2011) presents 100-year return value estimates of H_s and U_{10} from altimeter data, which is a follow up of Alves and Young (2003) based on H_s alone. These studies illustrates significant discrepancies between estimates obtained with IDM and POT, where it is concluded that the FT1 (Gumbel) distribution fitted to the IDM data validates best against corresponding estimates based on *in situ* measurements. Here, the POT approach is found unsuitable because altimeter data represent a temporal undersampling of the sea state, i.e. storm peaks are often missed. Given the fact that IDM/FT1 is more affected by the bulk of the data, this approach is less affected by missing peaks. It should be noted that Vinoth and Young (2011) use a relatively low threshold for the POT data, which is set at the 90-percentile. This was probably done intentionally to retain a minimum number of data, but still, this may

have affected the return value estimates significantly. In **Paper II** it is shown that the threshold needs to be set significantly higher to get a satisfactory fit to modelled data in the northeast Atlantic.

Another study based on altimeter data (Topex/Poseidon) was conducted by Chen et al. (2004). Here also the IDM/Gumbel was applied to obtain 100-year return value estimates of H_s and U_{10}. In order to compare the results of these studies, Table 3.2 presents the maximum 100-year estimates obtained globally, i.e. in the northeast Atlantic. These estimates are accompanied by the data period, the binning resolution and the applied method. In those cases where a study presents more than one estimate, the estimate of the preferred method is written in bold. For other regional studies based on altimeter data, see Carter (1993); Wimmer et al. (2006) for 50-year returns in the northeast Atlantic and Panchang et al. (1998) for 50-year returns around North America.

Table 3.2: Maximum 100-year return value estimates of H_s in the northeast Atlantic as obtained from different global studies. From left to right: reference of study; type of data; data period; resolution of data; the applied method; maximum 100-year return value estimate in the northeast Atlantic, where recommended estimates are presented in bold.

Reference	**Data**	**Period**	**Resolution**	**Method**	**100-yr**
Alves and Young (2003)	Altimeter	1986-1995	$2^\circ \times 2^\circ$	POT/3PW	22-24 m
			$4^\circ \times 4^\circ$	POT/3PW	22-24 m
			$2^\circ \times 2^\circ$	IDM/FT1	**24-26** m
Chen et al. (2004)	Altimeter	1993-2000	$1^\circ \times 1^\circ$	IDM/Gumbel	**20-21** m
Caires and Sterl (2005a)	ERA-40 (cal)	1958-2000	$1.5^\circ \times 1.5^\circ$	POT/EXP	**24.5-27.5** m
Vinoth and Young (2011)	Altimeter	1985-2008	$1^\circ \times 1^\circ$	IDM/FT1	**20-22** m
				POT/3PW	18-20 m
				POT/GP	20-22 m
			$2^\circ \times 2^\circ$	IDM/FT1	20-22 m
				POT/3PW	20-22 m
				POT/GP	20-22 m

Caires and Sterl (2005a) conducted a similar global study based on ERA-40 H_s and U_{10} over the period 1958-2000, also presented in Sterl and Caires (2005). Here, the exponential distribution is fitted to POT data above the 93-percentile. As ERA-40 is known to underestimate the high percentiles (Caires and Sterl, 2003), the 100-year return value estimates were calibrated against corresponding estimates based on *in situ* measurements. This study concluded that the exponential distribution in most cases yield realistic estimates, but on the conservative side. In certain regions, like the northeast Atlantic, the estimates were considered too conservative, i.e. overestimates. It was suggested that this could be compensated by applying a higher threshold, e.g. at the 97-percentile. Even though not advised by Caires and Sterl (2005a), there are other studies implying that H_s-data in the northeast Atlantic may belong to the GP-distribution (described in Section 4.1) with a negative shape parameter, i.e. having an upper bound, see e.g. Izaguirre et al. (2011). Estimate of the maximum 100-year H_s from Caires and Sterl (2005a) is incorporated in Table 3.2.

Trend estimates

A simplified way of describing climate change is to represent it by a linear trend. Although trends should be used cautiously to predict the future by extrapolation, they are relevant in describing tendencies. Reanalyses constitute an important resource in doing so. However, trends, especially when weak, put high demands on the accuracy of the underlying data. Typically, reanalyses are affected by the growing number of assimilations over the period it covers, which may lead to deceptive trends.

There are a number of reanalyses, see Table 3.1, and hindcasts whose purpose is to describe the atmosphere-wave climate as precise as possible. However, as long as these datasets yield different representations, there are good reasons to examine what these differences are caused by. For instance, see Simmons et al. (2014), for comparisons of temperature between ERA-I, ERA-40, MERRA and JRA-55.

In Caires et al. (2004) it was concluded that long-term trends taken over the two periods, 1958 to 1967 and 1990 to 1997, showed very similar features, implying trends were indifferent to the dataset being used. It was however stated that ERA-40, showed a slightly different pattern in trends over the period 1990-1997 due to erroneous altimeter data (ERS-1) assimilated over the period January 1992 until May 1993. In a similar study conducted by Stopa and Cheung (2014), H_s and U_{10} from ERA-I and CFSR/CFSR wave hindcast (Chawla et al., 2013) were compared against *in situ* measurements and altimeter observations. By investigating trends in the monthly error between the modelled and observed data, this study concluded that ERA-I was the most homogeneous over time, and therefore more adapted to trend analysis of H_s and U_{10}. This result was mainly based on comparisons with altimeter data (Topex/Poseidon and GEOSAT Follow-on) taken over the period 1992-2008.

While the CFSR wave hindcast is run without any wave data assimilation, both ERA-40 and ERA-I introduced altimeter wave measurements in August 1991. It is a general concern how this may affect the regional wave statistics of the reanalyses. Due to the choice of period, this issue is not captured by Caires et al. (2004) and Stopa and Cheung (2014), but is addressed in **Paper V**. To illustrate the problem, Figure 3.3 presents two time series of the mean H_s for the month of January over the period 1979-2012 at 55°N 40°W. It is shown that trends obtained at ANA and the 48-hour FCR (FC48) portray very different trends. This is directly affected by the step change imposed by altimeter assimilation in August 1991, indicated by the sudden change in discrepancy between the two time series.

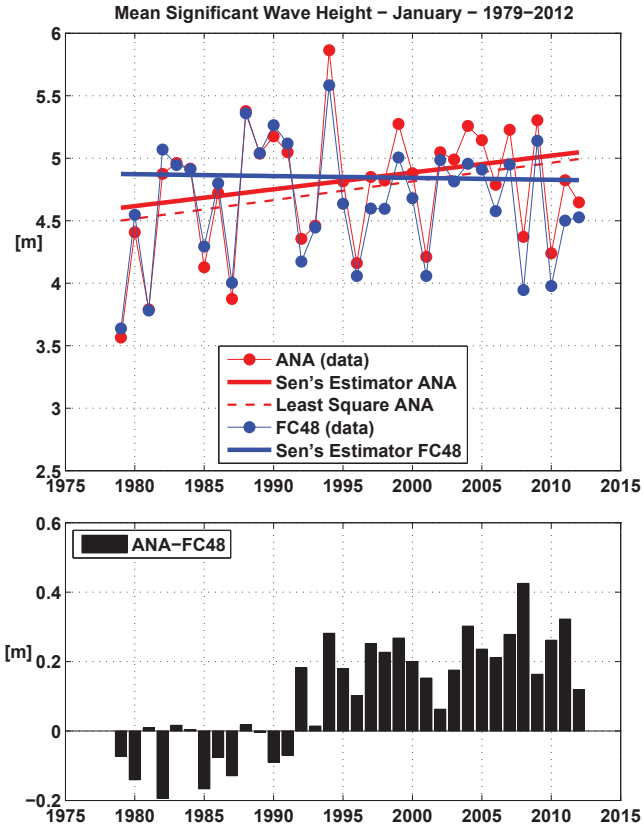

Figure 3.3: *Top:* Mean significant wave height for the month of January over the period 1979-2012 at 55°N 40°W based on ERA-I from analysis (ANA) and 48-hour forecast (FC48). The corresponding Sen estimator for trend is presented by solid lines. For comparison, the trend in ANA based on linear regression is presented by the dashed line. *Bottom:* Discrepancy between ANA and FC48 over the corresponding period.

Chapter 4

Methodology

4.1 Extreme value analysis

Extreme value analysis is a branch within statistics dealing with rare events, typically, more extreme than previously observed. This is accomplished by means of statistical models, i.e. theoretical probability distributions, known to conform to the relevant data at hand. As these models indeed are general approximations, they may be used to investigate probabilities of events beyond the empirical distribution, which only represent a subset of all possible outcomes. This type of extrapolation is necessary, but can be highly sensitive to the applied statistical approach.

Extreme events are scarce by definition and usually shifted away from the main part of the data, represented by the mode. Therefore, it may be inexpedient to model the distribution of all possible outcomes of the random process, but rather to focus on the tail of the distribution. The extreme statistics used in this thesis is mainly based on the work by Coles (2001), and the references given therein. We primarily rely on two methods of extracting the representative data, either by *blocking* or *peaks-over-threshold* (POT). By blocking, we only retain the highest value or a limited number of the highest values within a given time frame, which is here, set to one year. With POT, we retain threshold exceedances above a predetermined threshold. Both methods require the data to be independent events, implying no two H_s entries should originate from the same storm. This is achieved by requiring a minimum of 48 hours between peaks, which is sufficient to separate H_s-data generated by the same extratropical cyclone. Further, data should be identical, i.e. they should originate from the same statistical distribution. This may be violated if there are significant trends in the data, or if the dataset exhibits sudden step changes. Both issues are addressed in the thesis. Trends are presented in **Paper V**, while step changes, due to model updates and non-stationary assimilation, are discussed in **Paper II**, **Paper III** and **Paper V**. Although debatable, we find it reasonable to assume the effect of trends and discontinuities to be secondary compared to the substantial uncertainty associated with the return value estimates presented herein. For all practical purposes, the data are assumed independent and identically distributed (IID).

The *central limit theorem* (CLT) states that the distribution of means based on a large number of IID variables will be approximately normal, regardless of the underlying distribution. Related to the above discussion, this implies that the means based on all independent events within each block would be approximately normally distributed. In extreme value theory there exist an analogy to the CLT, given by the *extremal types*

theorem, which states that the renormalized maxima of all independent events within each block has a limiting distribution defined by the *generalized extreme value* (GEV) distribution:

$$G(z) = \exp\left\{ -\left[1 + \xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi} \right\} \tag{4.1}$$

where z in our case represents H_s. The form of the distribution is controlled by three parameters, i.e. shape (ξ), scale (σ) and location (μ), where $G(z)$ takes the form of the Gumbel ($\xi \to 0$), Frèchet ($\xi > 0$) or Weibull ($\xi < 0$) distribution. By inverting Eq. 4.1 we are left with two terms, depending on ξ, that may be used to obtain **return value estimates** of z:

$$z_p = \begin{cases} \mu - \frac{\sigma}{\xi}[1 - \{-\ln(1-p)\}^{-\xi}], & \text{for } \xi \neq 0 \\ \mu - \sigma \ln\{-\ln(1-p)\}, & \text{for } \xi = 0 \end{cases} \tag{4.2}$$

where $G(z_p) = 1 - p$ and p corresponds to some probability of exceedance. In our case, where G is representing the distribution of annual maxima, $1/p$ will correspond to a return period in years. For instance, for a return period of 100 years, the annual probability of exceedance is $p = 0.01$.

It can further be shown that if the annual maxima have an approximate distribution of G, defined by Eq. 4.1, then threshold excesses, y, above some large enough value u should conform to a distribution within the *generalized Pareto* (GP) family, which is defined by:

$$H(y) = 1 - \left(1 + \frac{\xi y}{\tilde{\sigma}}\right)^{-1/\xi} \tag{4.3}$$

requiring $y > 0$ and $(1 + \xi y/\tilde{\sigma}) > 0$, where $\tilde{\sigma} = \sigma + \xi(u - \mu)$. Similar to GEV, the GP family takes three forms, depending on ξ. In the special case of $\xi \to 0$, Eq. 4.3 reduces to the exponential distribution (EXP). Corresponding N-year return values are defined by

$$z_N = \begin{cases} u + \frac{\sigma}{\xi}[(N n_y \zeta_u)^{\xi} - 1], & \text{for } \xi \neq 0 \\ u + \sigma \ln(N n_y \zeta_u), & \text{for } \xi = 0 \end{cases} \tag{4.4}$$

where n_y represents the number of data per year and ζ_u is the probability of an individual observation exceeding the threshold u.

Model diagnostics

In this study we use the *maximum likelihood* to fit the theoretical distribution to the chosen subset of the initial data (Coles, 2001). Return value plots offer a convenient tool to check the conformity between the two. By plotting $-\ln(-\ln(1-p))$ against z_p, the GEV distribution is linear if $\xi = 0$, concave with an asymptotic limit if $\xi < 0$ and convex with no upper limit if $\xi > 0$. The model diagnostics of the GP distribution can be presented in the same way. Figure 4.1 presents three return value plots where the GP and EXP distribution is fitted to POT data using three different thresholds, i.e. the 97-percentile, the 99.3-percentile and the 99.7-percentile. It is shown that the GP distribution takes three forms depending on threshold, i.e $\xi < 0$, $\xi \sim 0$ and $\xi > 0$, respectively. This has significant impact on the accompanied 100-year return value

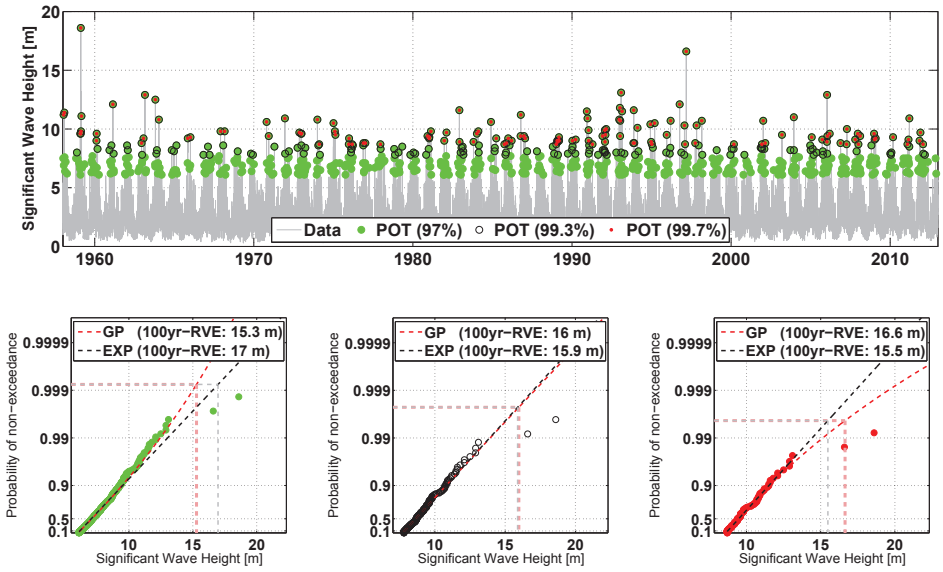

Figure 4.1: *Top:* Time series of H_s from NORA10 over the period 1958-2012 at the position 72°N 5°E marked in gray. Peaks-over-threshold (POT) data separated by a minimum of 48 hours with the threshold set at the 97-percentile (green dots), 99.3-percentile (black circles) and 99.7-percentile (red dots). *Bottom:* Return value plots of the generalized Pareto (GP) (dashed red) and the exponential distribution (EXP) (dashed black) fitted to the different POT-data. Corresponding 100-year return value estimates are given in the legend and indicated by the pink and gray dashed lines, respectively.

estimates, presented in the legends, ranging from 15.3 to 16.6 m. Also, the EXP shows great variation, ranging from 15.5 to 17 m. One of the potential benefits of using a threshold model is the possibility of utilizing more data for the distribution fitting, as opposed to the annual maximum. Even so, there exist no consensus on how to choose an optimal threshold. Often it is necessary to carefully study return value plots to obtain an ideal fit. Since this is not practical when return value estimates are obtained for a large area, a pragmatic way of dealing with this issue, is to apply a predetermined threshold representing some percentile of the initial data (Challenor et al., 2004; Caires and Sterl, 2005a).

4.2 Trend analysis

Mann-Kendall/Seasonal Kendall

A straightforward approach for estimating trend is obtained by linear regression. However, as pointed out by Sen (1968), the least square estimator is vulnerable to gross errors if the parental distribution departs from the normal distribution, which is the case for H_s (Holthuijsen, 2007).

In the following trend analysis presented in **Paper V** we are dealing with monthly

data, mean and maxima, respectively. Let $X_i = (x_{i1}, x_{i2}, \ldots x_{in_i})$ represent the monthly data of H_s and U_{10} in chronological order, where n_i is the total number of entries from month $i = 1, 2, \ldots 12$ (only one entry per year). Based on Sen (1968), a robust estimate of the monotonic trend is given by:

$$b = \text{median}\left(\frac{x_{ij} - x_{il}}{j - l}\right) \quad \forall l < j \tag{4.5}$$

where $1 \leq l < j \leq n_i$. For seasonal estimates of trend, the desired months (i) are chosen accordingly. Notice that the set of slopes are only obtained within each month. In this way, seasonality is accounted for. The annual trend is obtained by the median of all slopes taken over all months. As an example, Figure 3.3 illustrates trends obtained with the Sen estimator and linear regression.

In order to establish whether a trend is statistical significant or not, we we apply the Seasonal Kendall test, a non-parametric test of randomness (H_0) against trend (H_1), an extension of the Mann-Kendall test (Mann, 1945; Kendall, 1948) especially adapted to seasonal data with serial dependence (Dietz and Killeen, 1981; Hirsch and Slack, 1984). The Seasonal Kendall statistics for month i is expressed by

$$S_i = \sum_{k=1}^{n_i - 1} \sum_{j=k+1}^{n_i} \text{sgn}(x_{ij} - x_{ik}) \tag{4.6}$$

In case of missing data at time j or k, $\text{sgn}(x_{ij} - x_{ik})$ is set to zero. From Mann (1945); Kendall (1948); Hirsch et al. (1982) we define $S' = \sum_{i=1}^{12} S_i$, having a mean and variance given by

$$E[S'] = \sum_{i=1}^{12} E[S_i] = 0 \tag{4.7}$$

$$\text{Var}[S'] = \sum_{i=1}^{12} \text{Var}[S_i] + \sum_{i=1}^{12} \sum_{l=1}^{12} \text{cov}(S_i S_l), \quad i \neq l \tag{4.8}$$

where $\text{Var}[S_i] = n_{ig}(n_{ig} - 1)(2n_{ig} + 5)/18$ and n_{ig} represents the number of non-missing data per month ($n_{ig} = n_i$ for complete series). According to Hirsch et al. (1982), $\text{cov}(S_i S_l) = 0$ when S_i and S_l are independent random variables. However, this fails to hold for monthly lag-1 serial correlation as low as 0.2 (Hirsch and Slack, 1984). In the following we use an estimate of the covariance term defined by Dietz and Killeen (1981), which is documented in Hirsch and Slack (1984). A two-sided test for trend is based on the standard normal variate Z defined by

$$Z = \begin{cases} \frac{S'-1}{(\text{Var}[S'])^{1/2}} & \text{if } S' > 0 \\ 0 & \text{if } S' = 0 \\ \frac{S'+1}{(\text{Var}[S'])^{1/2}} & \text{if } S' < 0 \end{cases} \tag{4.9}$$

where H_0 is accepted when $|Z| < 1.96$, using a significance level of $\alpha = 0.05$.

Chapter 5

Summary

This PhD thesis is a survey of regional and global wave climate represented by significant wave height, H_s. The study has had two focus areas: return value estimates and trends. Both aspects are mainly controlled by three factors:

- the quality of the data
- the time span and coverage of the data
- choice of statistical approach

All three issues are addressed by:

- **Paper I:** Validating a high resolution hindcast covering the Northeast Atlantic, the NORA10.
- **Paper II:** Comparing three statistical approaches to obtain 100-year return value estimates based on NORA10.
- **Paper III:** Investigating the potential of utilizing large aggregated ensemble forecasts for return value estimates in the northeast Atlantic.
- **Paper IV:** Using a nonparametric approach based on order statistics to obtain global return value estimates of H_s and U_{10} from aggregated ensembles.
- **Paper V:** Investigating global trends in Hs and U_{10} from ERA-I, and their influence of data assimilation.

A brief summary of the papers are given below:

5.1 Summary of papers

Paper I: High-resolution hindcast of wind and waves for the North Sea, the Norwegian Sea and the Barents Sea

The relative coarse resolution of ERA-40's atmospheric component ($\sim$ 125 km) has been shown to generate H_s data of poor quality in the upper percentiles, where extreme events are quite severely underestimated. For this reason, the raw data have been

considered inappropriate for return value analysis. However, ERA-40 does portray H_s variability very well, a testament to its ability to reproduce synoptic features of the atmosphere. This paper presents the Norwegian Reanalysis 10 km (NORA10), a combined regional downscaling of ERA-40's atmospheric component and a stand-alone wave hindcast covering the northeast Atlantic over the period September 1957 to August 2002. With different model physics and higher spatial resolution, NORA10 winds are on average 0.5 m/s higher than ERA-40 over the open ocean, and in excess of 3 m/s in coastal areas. Validation against synoptic observations and QuikSCAT reveals that NORA10 is performing better on the whole range of data, but are still slightly weak on the highest wind speeds. The improved winds are reflected in the H_s data, which generally show higher correlation and smaller bias when validated against *in situ* measurements and altimeter observations. More importantly, NORA10 shows a clear improvement in the upper percentiles of H_s, establishing NORA10 as a high quality wave hindcast able to reproduce H_s extremes satisfactory in the northeast Atlantic. Given the performance of NORA10 H_s, it provides a baseline climatology, which other datasets may be compared against.

Paper II: Wave Extremes in the Northeast Atlantic

In this paper, an extended version of NORA10 is utilized to obtain 100-year return value estimates of H_s in the northeast Atlantic. The original downscaling of ERA-40 is prolonged in an identical manner based on analysis fields from the Integrated Forecast System (IFS) at the European Centre for Medium-Range Weather Forecasts (ECMWF). The chosen period, 1958 until the end of 2009, spans a total of 52 years. By comparing NORA10 against *in situ* measurements before and after the transition in August 2002, no major discontinuities are detected.

Three commonly applied extreme value models are used to investigate uncertainty in the 100-year return value estimates. The applied statistical distribution is dictated by the way extreme data is extracted from the initial dataset. This is done in two ways: by blocking or as threshold exceedances. Within each block of one year, we retain the annual maximum or a number of the largest entries. This is considered as two different approaches, where both subsets should conform to the generalized extreme value (GEV) distribution. Threshold exceedances are obtained according to the peaks-over-threshold (POT) approach, making sure peaks are separated by a minimum of 48 hours. The threshold selection is based on the Anderson-Darling test and visual inspections of return value plots. Here, the generalized Pareto (GP) distribution is fitted to POT data above the 99.7-percentile to obtain 100-year return value estimates.

For the majority of the model domain, the three approaches offers estimates within $\pm 5\%$. However, in certain areas discrepancies are peaking around $\pm 20\%$. By investigating the conformity between the three data subsets and the corresponding distributions, the POT/GP combination seems to yield better results in these areas.

Based on POT/GP the 100-year estimates are peaking around 22 m northwest of Scotland, around 14 m in the North Sea and above 16 m in the Norwegian Sea. It is further concluded that the return value estimates are too low near the western boundaries of the domain due to influence of the boundary conditions provided by ERA-40 and IFS-ECMWF analysis.

Paper III: Wave Extremes in the Northeast Atlantic from Ensemble Forecasts

Return value estimates of H_s are often based on datasets that are considerable shorter than the return period in question. As most time series do not extend much further than 50 years, 100-year return values are rarely represented in the data. In an effort to overcome this limitation, this study investigates the possibility of utilizing an unconventional dataset that is shown to span far beyond 100 years. Here, represented by archived ensemble forecasts from the ensemble prediction system (ENS) at ECMWF.

The main purpose of the ENS is to provide confidence estimates to the traditional medium range forecasts. This is obtained by running 50 slightly perturbed versions of the deterministic model at lower spatial resolution, plus a control run. In general, the spread of the ensemble is increasing with forecast range, reflecting the uncertainty in the forecast. Beyond day 6, the forecast skill is very low. At day 10, the 51 ensemble members are sufficiently uncorrelated to be considered independent representations of the model climate.

In this study, we aggregate historic ENS data at forecast range +240 h, i.e. day 10, over the period 1999-2009. The ENS is run once a day up until 23 March 2003, and twice a day beyond this data. By assuming each member representing a 6-h interval, the aggregated dataset is equivalent of ~ 226 years of data. This implies that the dataset, on average, should contain more than 2 events exceeding the 100-year return value.

Here, we apply the GEV and GP distribution to obtain 100-year return value estimates of H_s in the northeast Atlantic. The results are compared against corresponding estimates based one NORA10, ERA-40 and ERA-I. It is shown that the ENS data yields estimates comparable to NORA10 return values, and show a clear improvement over ERA-40 and ERA-I, which are known to underestimate the higher percentiles. It is concluded that archived ENS data represents an unused resource that may complement and sometimes yield more precise return values than traditional reanalyses and hindcasts.

Paper IV: Wind and wave extremes over the world oceans from very large ensembles

In this study, we investigate a nonparametric approach to obtain global 100-year return value estimates of H_s and U_{10}. Aggregated ensemble forecasts over the period 26 March 2003 to 25 March 2012 offer approximately 229 years of data. According to definition, these data should, on average, have ~ 2.29 entries that are equal to or exceeding the 100-year return value. So, without resorting to traditional extreme value models, it is possible to obtain direct estimates (DRE) from the weighted interpolation between the second and third highest entries taken from the order statistics.

In the following we compare 100-year return value estimates based on DRE, the GP distribution and the exponential (EXP) distribution, where the latter is a special case within the GP family. Normally, the conformity between the extreme value models and the data are validated by return value plots. Here, this is conducted by comparing the GP/EXP estimates against DRE estimates, which serves the same purpose at the probability level of the 100-year return period. This offers a representation of the be-

haviour of the GP/EXP distributions globally. The discrepancy between GP and EXP is most pronounced in the cyclone prone areas, where GP yields significantly higher return value estimates, caused by a positive shape parameter. In general, GP offers higher return value estimates of H_s and U_{10} and better conformity to the DRE.

Paper V: Marine wind and wave height trends at different ERA-Interim forecast ranges

Reanalyses play an important part in describing the past climate. They provide baseline climatologies that e.g. climate models are compared against. However, as reanalyses are highly affected by assimilation, it begs the question how trends are affected by the ever growing number of observations.

ERA-I is a highly recognized reanalysis produced at ECMWF providing both atmospheric and wave data. Besides the traditional reanalysis, ERA-I is also run as forecasts every 12 hours up until 10 days ahead. In general, models tend to relax towards the model climate at increased forecast range (FCR), being less affected by assimilation. By comparing trends obtained at analysis against increased FCR, it is possible to investigate to what degree trends are affected by assimilation. Here, we put special emphasis on the effect of wave altimeter data, introduced in August 1991, as this is the only type of data directly affecting the wave field. Further, we compare two stand-alone WAM runs forced by ERA-I winds, run with and without wave altimeter assimilation, in order to see effects of different satellite updates. Trends in H_s and U_{10} at different FCR are validated against *in situ* measurements and ENVISAT altimeter winds.

Here, it is concluded that the introduction of wave altimeter assimilation in August 1991 creates a step change in the ERA-I H_s data, especially pronounced in the northeast Atlantic and the eastern tropical Pacific, which impose spurious trends. It is shown that H_s-trends are affected by the different satellite updates, the number of operating satellites and the general availability of wave altimeter data. There are also proof of step changes in the U_{10} data, e.g. seen with the introduction of data from QuikSCAT in the southern hemisphere in 2000, but to a lesser extent.

Trends seem to be damped or slightly shifted towards the negative using data at increased FCR. Even so, the 10-day FCR is capable of representing the main spatial features of the trend found near analysis. Still, there is a trade-off between removing the impact of data assimilation at longer forecast range and getting lower level of uncertainty in the predictions at shorter forecast range. Here, data at the 48-hour FCR is proposed as a better candidate to reproduce realistic H_s-trends. Given the fact that assimilation also introduces step changes in the U_{10} data, trends based on the 24-48 hour FCR will remove some of the spurious effects.

5.2 Conclusions and future perspectives

The robustness of return value estimates are mainly controlled by three factors: the quality of the ground data, the amount of available data and the choice of statistical method. Even so, given the nature of extremes, return value estimates are not easily verified. Increased confidence is first and foremost achieved by coherent results based

on different datasets and statistical approaches. In this regard, the northeast Atlantic makes an interesting proving ground as it represents the area anticipating highest H_s extremes, i.e. the global maximum.

The return value estimates presented in this thesis are mainly based on NORA10 and archived ENS data, but further compared against corresponding estimates from ERA-40, ERA-I and *in situ* measurements. The two primary datasets have different characteristics. NORA10 is a high-resolution dataset with verified high performance on the whole range of data; it is run with the same model configurations throughout the period; it does however have a transition in August 2002, going from ERA-40 to IFS-ECMWF boundary conditions, but if any, the discontinuity is shown to be small; it is only regional; the length of the time series is moderate ($\sim$ 50 years). The archived ENS data has lower spatial resolution; it is based on different model configurations; it is not influenced by assimilation, as it consist of 10-day forecasts; it is global; the dataset is vast ($>$ 220 years).

Given the different spatial resolution, the 100-year return value estimates are not strictly comparable. While NORA10 represents approximately 1-hour average sea states, the ENS data are validated and assumed to represent 6-hour averages. The ENS estimates are therefore biased slightly low compared to NORA10, i.e. when disregarding the westernmost areas of NORA10 which are heavily influenced by the boundary conditions. Still, the ENS data provide significantly higher estimates than ERA-40 and ERA-I, which are known to underestimate the higher H_s percentiles. Based on the ENS data, the global maximum is just below 20 m. In comparison, NORA10 estimates are exceeding 22 m when based on POT/GP. Compared to similar studies, see Table 3.2, the ENS estimates are slightly on the low side, while NORA10 is more centred among the rest.

In this study two families of extreme value distributions have been utilized to obtain 100-year return value estimates, the generalized extreme value (GEV) distribution and the generalized Pareto distribution, where the choice of distribution is dictated by the chosen subset of the initial data. Both distributions may be applied with restrictions set on the shape parameter. If set to zero, the distributions reduce to the Gumbel and the exponential distribution (EXP), respectively. When the GEV and the GP distributions were applied with no restrictions, more than 70% of the model domain conforms to a negative shape parameter based on NORA10 H_s, i.e. providing return value estimates approaching an asymptotic upper limit. This is in accordance with Izaguirre et al. (2011). This implies that the EXP distribution will provide more conservative estimates in the northeast Atlantic, as the distribution is unbounded at the high end. However, we do acknowledge that the EXP and GP distribution tend to converge when applying a sufficiently high threshold in the POT data. Still, the GP distribution may have advantages in situations where the wave data belong to different populations. For instance, in certain areas where bottom friction becomes a limiting factor, data may be a mix of deep water and shallow water waves. If the majority of H_s data are behaving as deep water waves, the fitted EXP is likely to produce overestimates, while the increased adaptability of the GP are more likely to conform better to the highest waves. Similarly, in cyclone-prone areas, where the majority of waves still are generated by extratropical cyclones, the highest H_s-data generated by tropical cyclones are likely to be underestimated by the EXP distribution.

The return value estimates presented in this study is based on extreme value mod-

els assuming stationary conditions, which is in accordance with the studies we have compared against. Still, in future work, non-stationary models should be investigated in order to account for trends and interannual variability (Coles, 2001; Ruggiero et al., 2010; Izaguirre et al., 2011, 2013). Even so, it may be added that trends in monthly maximum H_s over the period 1979-2012, is shown to be weak and non-significant outside the tropics, justifying the assumption of stationary conditions. It should also be added that one of the benefits of using large aggregates of ENS data, is that it does not exhibit long-term trends and low-frequency oscillations since the initial conditions cover a fairly short period of time; in our our case, no more than 11 years.

The use of ensemble aggregates for return value estimates is a new and alternative approach. It represents an unused resource that may complement and sometimes yield more realistic return value estimates than obtained from traditional reanalysis and hindcasts. The aggregates used in this study is shown to represent more than 200 years of data, but may be significantly extended as more historic ENS data becomes available. There are however a few pitfalls. Special attention must made to the homogeneity of the data, as models are constantly updated. And, even though it is shown that the 10-day forecasts represent independent draws from the model climate, this may fail to hold as the forecast skill of the ENS is improving. Still, the method should be applicable to other datasets. By avoiding assimilation altogether, 10-member ensemble integrations spanning the twentieth century (Hersbach et al., 2013) may prove relevant for estimating H_s return value estimates and corresponding trends.

This study has presented trends in H_s and U_{10} based on the ERA-I reanalysis. It is shown that H_s-trends, in particular, should be used with care. The last decades have seen an ever growing number of assimilations and inclusions of new observations which are likely to affect trends. It is shown that the introduction of altimeter wave heights in August 1991 imposes a step change in the H_s data, with most pronounced effect in the northeast Atlantic and the eastern tropical Pacific. These areas are typically associated with wave model bias, which is abruptly corrected by altimeter wave height assimilation. Here, it is shown that data at increased FCR are less affected by spurious effects imposed by assimilation, and therefore may portray H_s and U_{10} trends more realistically.

With the advent of reanalyses spanning the twentieth century (Compo et al., 2011; Poli et al., 2013), special care is taken to tackle the problem of a changing observing system. By omitting all satellite data, assimilating surface observations only, the objective of these reanalyses is to attain homogeneous datasets more adapted for trend analysis. Trends in U_{10}, mean and extremes, have recently been presented in (Boisséson et al., 2014; Donat et al., 2011; Brönnimann et al., 2012), but also H_s-trends based on reconstructions and hindcasts have been conducted Wang et al. (2012); Bertin et al. (2013). As shown in this study, inferences should be made cautiously when compared against H_s-trends obtained with ERA-I at analysis.

Generally, reanalyses like ERA-40 and ERA-I constitute invaluable data archives for climate research. Still, they should not be considered flawless representations of the real climate. They require special attention depending on the topic in question, reflected by their inherent limitations. Often alternative datasets are required to support and verify their validity, which is illustrated by this thesis.

References

Alves, J. H. G. and I. R. Young, 2003: On estimating extreme wave heights using combined GEOSAT, Topex/Poseidon and ERS-1 altimeter data. *Applied Ocean Research*, **25 (4)**, 167 – 186, doi:10.1016/j.apor.2004.01.002.

Bengtsson, L., K. I. Hodges, and E. Roeckner, 2006: Storm tracks and climate change. *Journal of Climate*, **19 (15)**, 3518–3543, doi:10.1175/JCLI3815.1.

Bertin, X., E. Prouteau, and C. Letetrel, 2013: A significant increase in wave height in the North Atlantic Ocean over the 20th century. *Global and Planetary Change*, **106**, 77–83, doi:10.1016/j.gloplacha.2013.03.009.

Boisséson, E., M. Balmaseda, S. Abdalla, E. Källén, and P. Janssen, 2014: How robust is the recent strengthening of the Tropical Pacific trade winds? *Geophysical Research Letters*, doi:10.1002/2014GL060257.

Breivik, Ø., O. J. Aarnes, S. Abdalla, J.-R. Bidlot, and P. A. Janssen, 2014: Wind and Wave Extremes over the World Oceans From Very Large Ensembles. *Geophysical Research Letters*, 10, arXiv:1407.5581, doi:10.1002/2014GL060997.

Breivik, Ø., Y. Gusdal, B. R. Furevik, O. J. Aarnes, and M. Reistad, 2009: Nearshore wave forecasting and hindcasting by dynamical and statistical downscaling. *J Marine Syst*, **78 (2)**, S235–S243, arXiv:1206.3055, doi:10/cbgwqd.

Brönnimann, S., O. Martius, H. Von Waldow, C. Welker, J. Luterbacher, G. Compo, P. Sardeshmukh, and T. Usbeck, 2012: Extreme winds at northern mid-latitudes since 1871. *Meteorologische Zeitschrift*, **21 (1)**, 13–27, doi:10.1127/0941-2948/2012/0337.

Buizza, R., J.-R. Bidlot, N. Wedi, M. Fuentes, M. Hamrud, G. Holt, and F. Vitart, 2007: The new ECMWF VAREPS (variable resolution ensemble prediction system). *Quarterly Journal of the Royal Meteorological Society*, **133 (624)**, 681–695, doi:10.1002/qj.75.

Buizza, R., P. Houtekamer, G. Pellerin, Z. Toth, Y. Zhu, and M. Wei, 2005: A comparison of the ecmwf, msc, and ncep global ensemble prediction systems. *Monthly Weather Review*, **133 (5)**, 1076–1097, doi:10.1175/MWR2905.1.

Caires, S. and A. Sterl, 2003: Validation of ocean wind and wave data using triple collocation. *Journal of Geophysical Research: Oceans (1978–2012)*, **108 (C3)**, doi:10.1029/2002JC001491.

Caires, S. and A. Sterl, 2005a: 100-Year Return Value Estimates for Ocean Wind Speed and Significant Wave Height from the ERA-40 Data. *Journal of Climate*, **18**, 1032–1048, doi:10.1175/JCLI-3312.1.

Caires, S. and A. Sterl, 2005b: A new nonparametric method to correct model data: application to significant wave height from the ERA-40 re-analysis. *Journal of Atmospheric and Oceanic Technology*, **22 (4)**, 443–459, doi:10.1175/JTECH1707.1.

Caires, S., A. Sterl, J. Bidlot, N. Graham, and V. Swail, 2004: Intercomparison of different wind-wave reanalyses. *Journal of Climate*, **17 (10)**, 1893–1913, doi:10.1175/1520-0442(2004)017<1893:IODWR>2.0.CO;2.

Carter, D. J. T., 1993: Estimating extreme wave heights in the NE Atlantic from GEOSAT data. Tech. rep., Offshore Tech. Rep. OTH 93 396, Health and Safety Executive, London, United Kingdom. URL `http://mail.hsebooks.com/research/othpdf/200-399/oth396.pdf`.

Cavaleri, L., et al., 2007: Wave modelling–the state of the art. *Progress in oceanography*, **75 (4)**, 603–674, doi:10.1016/j.pocean.2007.05.005.

Challenor, P., W. Wimmer, and I. Ashton, 2004: Climate change and extreme wave heights in the north atlantic. *Proceedings of the 2004 Envisat & ERS Symposium*, 6–10.

Chawla, A., D. M. Spindler, and H. L. Tolman, 2013: Validation of a thirty year wave hindcast using the climate forecast system reanalysis winds. *Ocean Modelling*, **70**, 189–206, doi:10.1016/j.ocemod.2012.07.005.

Chen, G., S.-W. Bi, and R. Ezraty, 2004: Global structure of extreme wind and wave climate derived from topex altimeter data. *International Journal of Remote Sensing*, **25 (5)**, 1005–1018, doi:10.1080/01431160310001598980.

Coles, S., 2001: *An introduction to Statistical Modelling of Extreme Values*. Springer-Verlag, 208 pp., doi:10.1007/978-1-4471-3675-0.

Compo, G. P., et al., 2011: The twentieth century reanalysis project. *Quarterly Journal of the Royal Meteorological Society*, **137 (654)**, 1–28, doi:10.1002/qj.776.

Cooper, C. K. and G. Z. Forristall, 1997: The Use of Satellite Altimeter Data to Estimate the Extreme Wave Climate. *J. Atmos. Oceanic Technol.*, **14 (2)**, 254–266, doi:10.1175/1520-0426(1997)014<0254:TUOSAD>2.0.CO;2, URL `http://dx.doi.org/10.1175/1520-0426(1997)014<0254:TUOSAD>2.0.CO;2`.

Cox, A. T. and V. R. Swail, 2001: A global wave hindcast over the period 1958–1997: validation and climate assessment. *Journal of Geophysical Research: Oceans (1978–2012)*, **106 (C2)**, 2313–2329, doi:10.1029/2001JC000301.

Dee, D., 2013: The Climate Data Guide: ERA-Interim. Last modified 13 Nov 2013., https://climatedataguide.ucar.edu/climate-data/era-interim., last modified 13 Nov 2013.

Dee, D., et al., 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, **137 (656)**, 553–597, doi:10.1002/qj.828.

Dietz, E. J. and T. J. Killeen, 1981: A nonparametric multivariate test for monotone trend with pharmaceutical applications. *Journal of the American Statistical Association*, **76 (373)**, 169–174, doi:10.1080/01621459.1981.10477624.

Dodet, G., X. Bertin, and R. Taborda, 2010: Wave climate variability in the North-East Atlantic Ocean over the last six decades. *Ocean Modelling*, **31 (3)**, 120–131, doi:10.1016/j.ocemod.2009.10.010.

Donat, M., D. Renggli, S. Wild, L. Alexander, G. Leckebusch, and U. Ulbrich, 2011: Reanalysis suggests long-term upward trends in european storminess since 1871. *Geophysical Research Letters*, **38 (14)**, doi:10.1029/2011GL047995.

Dupuis, H., D. Michel, and A. Sottolichio, 2006: Wave climate evolution in the bay of biscay over two decades. *Journal of Marine Systems*, **63 (3)**, 105–114.

Ebita, A., et al., 2011: The Japanese 55-year Reanalysis" JRA-55": an interim report. *Sola*, **7**, 149–152, doi:10.2151/sola.2011-038.

Forristall, G. Z., 1978: On the Statistical Distribution of Wave Heights in a Storm. *JOURNAL OF GEOPHYSICAL RESEARCH*, **83**, 2353–2358, doi:10.1029/JC083iC05p02353.

Gelci, R., H. Cazalé, and J. Vassal, 1957: Prévision de la houle. La méthode des densités spectroangulaires. *Bulletin d\'information du Comité d\'Océanographie et d\'Etude des Côtes*, **9**, 416–435.

Gemmrich, J., B. Thomas, and R. Bouchard, 2011: Observational changes and trends in northeast Pacific wave records. *Geophysical Research Letters*, **38 (22)**, doi:10.1029/2011GL049518.

Gibson, J., P. Kallberg, and S. Uppala, 1997: ERA description. Tech. rep.

Günther, H., S. Hasselmann, and P. A. Janssen, 1992: The WAM model cycle 4. Tech. rep., Deutsches Klimarechenzentrum (DKRZ), Hamburg (Germany).

Hagedorn, R., 2008: Using the ECMWF reforecast dataset to calibrate EPS forecasts. Ecmwf newsletter, European Centre for Medium-Range Weather Forecasts, 8–13 pp.

Hagedorn, R., R. Buizza, T. Hamill, M. Leutbecher, and T. Palmer, 2012: Comparing TIGGE multimodel forecasts with reforecast-calibrated ECMWF ensemble forecasts. *Q J R Meteorol Soc*, **138**, 14, doi:10.1002/qj.1895.

Hersbach, H., C. Peubey, A. Simmons, P. Poli, D. Dee, and P. Berrisford, 2013: Era-20cm: A twentieth century atmospheric model ensemble. *ERA Report Series*, **16**.

Hirsch, R. M. and J. R. Slack, 1984: A nonparametric trend test for seasonal data with serial dependence. *Water Resources Research*, **20 (6)**, 727–732, doi:10.1029/WR020i006p00727.

Hirsch, R. M., J. R. Slack, and R. A. Smith, 1982: Techniques of trend analysis for monthly water quality data. *Water resources research*, **18 (1)**, 107–121, doi:10.1029/WR018i001p00107.

Hodges, K., R. Lee, and L. Bengtsson, 2011: A comparison of extratropical cyclones in recent reanalyses ERA-Interim, NASA MERRA, NCEP CFSR, and JRA-25. *Journal of Climate*, **24 (18)**, 4888–4906, doi:10.1175/2011JCLI4097.1.

Holthuijsen, L. H., 2007: *Waves in oceanic and coastal waters*. Cambridge University Press.

Izaguirre, C., F. J. Méndez, A. Espejo, I. J. Losada, and B. G. Reguero, 2013: Extreme wave climate changes in Central-South America. *Climatic change*, **119 (2)**, 277–290, doi:10.1007/s10584-013-0712-9.

Izaguirre, C., F. J. Méndez, M. Menéndez, and I. J. Losada, 2011: Global extreme wave height variability based on satellite data. *Geophysical Research Letters*, **38 (10)**, doi:10.1029/2011GL047302.

Janssen, P., 2004: *The interaction of ocean waves and wind*. Cambridge University Press.

Kallberg, P., S. Simmons, S. Uppala, and M. Fuentes, 2004: The ERA-40 Archive, ERA-40 Project Report Series No. 17. Tech. rep. URL `http://old.ecmwf.int/publications/library/do/references/list/192`.

Kalnay, E., 2003: *Atmospheric modeling, data assimilation, and predictability*. Cambridge university press.

Kalnay, E., et al., 1996: The NCEP/NCAR reanalysis 40-year project. *Bull. Am. Meteorol. Soc*, **77 (3)**, 437–471, doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. Hnilo, M. Fiorino, and G. Potter, 2002: Ncep-doe amip-ii reanalysis (r-2). *Bulletin of the American Meteorological Society*, **83 (11)**, 1631–1643, doi:10.1175/BAMS-83-11-1631.

Kendall, M. G., 1948: *Rank correlation methods.* Griffin.

Komen, G. J., M. Cavaleri, M. Donelan, K. Hasselmann, S. Hasselmann, and P. A. E. M. Janssen, 1994: *Dynamics and Modelling of Ocean Waves*. Cambridge University Press.

Lalaurette, F., 2003: Early detection of abnormal weather conditions using a probabilistic extreme forecast index. *Quarterly Journal of the Royal Meteorological Society*, **129 (594)**, 3037–3057, doi:10.1256/qj.02.152.

Lionello, P., H. Günther, and P. A. Janssen, 1992: Assimilation of altimeter data in a global third-generation wave model. *Journal of Geophysical Research: Oceans (1978–2012)*, **97 (C9)**, 14 453–14 474, doi:10.1029/92JC01055.

Lopatoukhin, L. J., V. A. Rozhkov, V. E. Ryabinin, V. R. Swail, A. V. Boukhanovsky, and A. B. Degtyarev, 2000: Estimation of extreme wind wave heights. Jcomm tech. rep. 9, Joint WMO/IOC Technical Commission for Oceanography and Marine Meteorology. URL `\url{http://jdoss.jodc.go.jp/info/ioc_doc/JCOMM_Tech/TR09.pdf}`.

Mann, H. B., 1945: Nonparametric test against trend. *Econometrika*, **13**, 245–259, URL `http://www.jstor.org/stable/1907187`.

Mathiesen, M., Y. Gdoda, P. Hawkes, E. Mansard, M. J. Martin, and E. P. E. T. adn G. van Vledder, 1994: Recommended practice for extreme wave analysis. *Journal of Hydraulic Research*, **32(6)**, 813–814, doi:10.1080/00221689409498691.

Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliagis, 1996: The ecmwf ensemble prediction system: Methodology and validation. *Quarterly Journal of the Royal Meteorological Society*, **122 (529)**, 73–119, doi:10.1002/qj.49712252905.

Molteni, F., et al., 2011: The new ecmwf seasonal forecast system (system 4). Tech. Rep. 656. URL `http://old.ecmwf.int/publications/library/do/references/list/14`.

Moore, G., R. S. Pickart, and I. Renfrew, 2008: Buoy observations from the windiest location in the world ocean, cape farewell, greenland. *Geophysical Research Letters*, **35 (18)**, doi:10.1029/2008GL034845.

National Data Buoy Center - NOAA, 2014: http://www.ndbc.noaa.gov/.

Onogi, K., et al., 2007: The JRA-25 reanalysis. *Journal of the Meteorological Society of Japan*, **85 (3)**, 369–432, doi:10.2151/jmsj.85.369.

Panchang, V., L. Zhao, and Z. Demirbilek, 1998: Estimation of extreme wave heights using GEOSAT measurements. *Ocean Engineering*, **26 (3)**, 205–225, URL `http://www.sciencedirect.com/science/article/pii/S0029801897100269`.

Poli, P., et al., 2013: The data assimilation system and initial performance evaluation of the ecmwf pilot reanalysis of the 20th-century assimilating surface observations only (era-20c). *ERA Report*, **14**.

Prates, F. and R. Buizza, 2011: Pret, the probability of return: a new probabilistic product based on generalized extreme-value theory. *Quarterly Journal of the Royal Meteorological Society*, **137 (655)**, 521–537, doi:10.1002/qj.759.

Reistad, M., O. Breivik, H. Haakenstad, O. J. Aarnes, and B. R. Furevik, 2011: A high-resolution hindcast of wind and waves for The North Sea, The Norwegian Sea and The Barents Sea. *Journal of Geophysical Research, Oceans.*, **116**, doi:10.1029/2010JC006402, c05019.

Rienecker, M. M., et al., 2011: MERRA: NASA's modern-era retrospective analysis for research and applications. *Journal of Climate*, **24 (14)**, 3624–3648, doi:10.1175/JCLI-D-11-00015.1.

Ruggiero, P., P. D. Komar, and J. C. Allan, 2010: Increasing wave heights and extreme value projections: The wave climate of the US Pacific Northwest. *Coastal Engineering*, **57 (5)**, 539–552, doi:10.1016/j.coastaleng.2009.12.005.

Saha, S., et al., 2010: The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, **91 (8)**, 1015–1057, doi:10.1175/2010BAMS3001.1.

Sampe, T. and S.-P. Xie, 2007: Mapping high sea winds from space: A global climatology. *Bulletin of the American Meteorological Society*, **88 (12)**, 1965–1978, doi:10.1175/BAMS-88-12-1965.

Semedo, A., K. Sušelj, A. Rutgersson, and A. Sterl, 2011: A global view on the wind sea and swell climate and variability from ERA-40. *Journal of Climate*, **24**, 1461–1479, doi:10.1175/2010JCLI3718.1.

Sen, P. K., 1968: Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association*, **63 (324)**, 1379–1389, doi:10.1080/01621459.1968.10480934.

Simmons, A., P. Poli, D. Dee, P. Berrisford, H. Hersbach, S. Kobayashi, and C. Peubey, 2014: Estimating low-frequency variability and trends in atmospheric temperature using era-interim. *Quarterly Journal of the Royal Meteorological Society*, **140 (679)**, 329–353, doi:10.1002/qj.2317.

Soukissian, T. H. and G. D. Kalantz, 2006: Extreme value analysis methods used for extreme wave prediction. *Proceedings of the sixteenth international offshore and polar engineering conference. San Francisco, California, USA, May 28-June 2, 2006.*, URL `\url{http://207.56.194.103/publications/proceedings/ISOPE/ISOPE%202006/papers/2006_TS_01.pdf}`.

Sterl, A. and S. Caires, 2005: Climatology, variability and extrema of ocean waves: The Web-based KNMI/ERA-40 wave atlas. *International Journal of Climatology*, **25 (7)**, 963–977, doi:10.1002/joc.1175.

Sterl, A., G. Komen, and P. Cotton, 1998: Fifteen years of global wave hindcasts using winds from the european centre for medium-range weather forecasts reanalysis: Validating the reanalyzed winds and assessing the wave climate. *Journal of Geophysical Research: Oceans (1978–2012)*, **103 (C3)**, 5477–5492, doi:10.1029/97JC03431.

Stopa, J. E. and K. F. Cheung, 2014: Intercomparison of wind and wave data from the ECMWF Reanalysis Interim and the NCEP Climate Forecast System Reanalysis. *Ocean Modelling*, **75**, 65–83, doi:10.1016/j.ocemod.2013.12.006.

Stopa, J. E., K. F. Cheung, H. L. Tolman, and A. Chawla, 2013: Patterns and cycles in the climate forecast system reanalysis wind and wave data. *Ocean Modelling*, **70**, 207–220, doi:10.1016/j.ocemod.2012.10.005.

Swail, V. R. and A. T. Cox, 2000: On the use of ncep-ncar reanalysis surface marine wind fields for a long-term north atlantic wave hindcast. *Journal of Atmospheric*

and Oceanic Technology, **17 (4)**, 532–545, doi:10.1175/1520-0426(2000)017<0532:OTUONN>2.0.CO;2.

Tolman, H. L., 2009: User manual and system documentation of wavewatch iii tm version 3.14. *Technical note, MMAB Contribution*, **(276)**.

Toth, Z. and E. Kalnay, 1993: Ensemble forecasting at nmc: The generation of perturbations. *Bulletin of the American Meteorological Society*, **74 (12)**, 2317–2330, doi:10.1175/1520-0477(1993)074<2317:EFANTG>2.0.CO;2.

Toth, Z. and E. Kalnay, 1997: Ensemble forecasting at ncep and the breeding method. *Monthly Weather Review*, **125 (12)**, 3297–3319, doi:10.1175/1520-0493(1997)125<3297:EFANAT>2.0.CO;2.

Tucker, M., 1991: Waves in ocean engineering: Measurement. *Analysis, Interpretation*.

Ulbrich, U., G. Leckebusch, and J. Pinto, 2009: Extra-tropical cyclones in the present and future climate: a review. *Theoretical and Applied Climatology*, **96 (1-2)**, 117–131, doi:10.1007/s00704-008-0083-8.

Undén, P., et al., 2002: HIRLAM-5 Scientific Documentation, HIRLAM-5 Project. Tech. rep., Available from SMHI, S-601767 Norrköping, Sweden.

Uppala, S. M., et al., 2005: The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, **131 (612)**, 2961–3012, doi:10.1256/qj.04.176.

Vincent, L. A., X. L. Wang, E. J. Milewska, H. Wan, F. Yang, and V. Swail, 2012: A second generation of homogenized Canadian monthly surface air temperature for climate trend analysis. *Journal of Geophysical Research: Atmospheres (1984–2012)*, **117 (D18)**, doi:10.1029/2012JD017859.

Vinoth, J. and I. Young, 2011: Global estimates of extreme wind speed and wave height. *Journal of Climate*, **24 (6)**, 1647–1665, doi:10.1175/2010JCLI3680.1.

Vitart, F., 2004: Monthly forecasting at ecmwf. *Monthly Weather Review*, **132 (12)**, 2761–2779.

WAMDI Group, 1988: The WAM model-a third generation ocean wave prediction model. *Journal of Physical Oceanography*, **18 (12)**, 1775–1810, doi:10.1175/1520-0485(1988)018<1775:TWMTGO>2.0.CO;2.

Wan, H., X. L. Wang, and V. R. Swail, 2010: Homogenization and Trend Analysis of Canadian Near-Surface Wind Speeds. *Journal of Climate*, **23 (5)**, doi:10.1175/2009JCLI3200.1.

Wang, X. L., Y. Feng, and V. Swail, 2012: North Atlantic wave height trends as reconstructed from the 20th century reanalysis. *Geophysical Research Letters*, **39 (18)**, doi:10.1029/2012GL053381.

Wimmer, W., P. Challenor, and C. Retzler, 2006: Extreme wave heights in the North Atlantic from altimeter data. *Renewable Energy*, **31 (2)**, 241–248, doi:10.1016/j.renene.2005.08.019.

WMO-group, 1998: *Guide to wave analysis and forecasting*. second edition ed., WMO-No.702, World Meteorological Organization.

Young, I., J. Vinoth, S. Zieger, and A. Babanin, 2012: Investigation of trends in extreme value wave height and wind speed. *Journal of Geophysical Research: Oceans (1978–2012)*, **117 (C3)**, doi:10.1029/2011JC007753.

Young, I., S. Zieger, and A. Babanin, 2011: Global trends in wind speed and wave height. *Science*, **332 (6028)**, 451–455, doi:10.1126/science.1197219.

Zieger, S., J. Vinoth, and I. Young, 2009: Joint calibration of multiplatform altimeter measurements of wind speed and wave height over the past 20 years. *Journal of atmospheric and oceanic technology*, **26 (12)**, 2549–2564, doi:10.1175/2009JTECHA1303.1.

Errata

Equation (3) on page 85 and equation (1) on page 110 are both written:

$$H(y) = 1 - \left(1 + \frac{\xi y}{\tilde{\sigma}}^{-1/\xi}\right)$$

, but should read:

$$H(y) = 1 - \left(1 + \frac{\xi y}{\tilde{\sigma}}\right)^{-1/\xi}$$

Corrections

1. Paper V (page 3 and 123) is no longer *in revision*, but has been *accepted*
2. Wrong reference on page 128, under "RHtestsV4 - Homogenization", second paragraph: "Wang et al. (2010)" not "Wan et al. (2010)"
3. Table 3.1 on page 12: should be "ERA-40 / 1957/09-2002/08" not "ERA-40 / 1979-1993" and "ERA-Interim / 1979-" not "ERA-Interim / 1979-1993"