THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

A long-term analysis of the wave climate in the North East Atlantic and North Sea

Atul Agarwal

Doctor of Philosophy

THE UNIVERSITY OF EDINBURGH

2015

To my parents

Lay Summary

Over the last thirty five years, studies have reported that wave heights in the North East Atlantic and North Sea have been steadily increasing. Moreover, with the potential impacts of climate change not completely known, this observation is a matter of concern for users of these waters which include wave energy converters in the foreseeable future.

In light of this, the objectives of the doctoral research are threefold - (a) to identify whether the changes reported in the past are real and continue to persist; (b) to identify whether these changes, if any, are linked to climate change and atmospheric greenhouse gas levels; and (c) to gain insight into the potential future wave climate in the region.

The research uses computer simulations to produce a dataset of wave conditions for the region spanning 140 years between 1871 and 2010 which are analysed for trends and changes. Similarly, a dataset of possible future wave climates corresponding to the high, medium and low greenhouse gas emissions scenarios is also produced for the period between 2001 and 2100.

The study observed increasing trends in large parts of the region for the different parameters of historical wave climate; conversely, decreasing trends were observed in these across all three future emissions scenarios. An inter-comparison of the wave data for different concentrations of greenhouse gases in the atmosphere indicates that it is unlikely that the observed trends in the historical wave climate are related to anthropogenic climate change. Finally, statistical methods are used to estimate extreme conditions in the 21st century.

The results of the analyses are presented graphically in the form of an atlas to enable ease of reference by stakeholders.

Abstract

Changes in the marine environment have been reported for over three decades in terms of mean annual wave heights, exceedance probabilities and extreme conditions. More recently, the existence of a link between these changes and anthropogenic climate change has been postulated. This is not unreasonable, as climatic changes in regional warming and cooling are likely to alter wind patterns, and therefore the wave climate as well.

In an attempt to mitigate climate change and increase energy security, the offshore environment is being looked at to provide sustainable energy from wind, waves and tides. As a result the number of marine structures is only likely to increase. While survivability in this environment is essential for all such installations, some devices such as wave energy converters also rely on the environment for energy production. In designing these offshore structures to survive the harshest conditions as well as to ensure optimum operation, knowledge of the evolution of the wave climate is essential.

This study aims to identify and evaluate any historical trends that may be exhibited by the wave climate in the North East Atlantic and North Sea region. The study also aims to investigate the link between any observed changes and atmospheric greenhouse gas levels and projected wave conditions for the 21st century.

This is achieved by producing a long-term, high resolution hindcast of wave conditions for 1871-2010 using the third-generation spectral wave model WAVEWATCH III. A dataset of wave climate projections for the high, medium and low emissions scenarios is also prepared by forcing the model with GCM winds for 2001-2100.

In addition to dynamically projecting the wave climate in the 21st century for different IPCC climate change scenarios, statistical methods were applied to historic data to estimate extreme events in terms of 100-year return values of significant wave height. These, together, provide some idea of the plausible wave climate up to 2100.

The results of the work show the existence of long-term trends in the historical wave climate in the region from 1921 onwards. However, based on the findings of the study, it is unlikely that these are a result of changes in atmospheric greenhouse gas concentrations and are more likely due to internal variability in the system.

Acknowledgements

I am indebted to my supervisors, Dr. Vengatesan Venugopal and Prof. Gareth Harrison, for their encouragement and guidance during the course of the project. I am grateful to them for the degree of autonomy that they allowed me as well as for always making time to advise me when I struggled.

My gratitude to my family is too great for words. They backed all my decisions, even when they didn't agree with them, and were always there for support when I needed it. I am, in no small part, what I am because of them.

I wish to make special mention of Ajay Kamath, Amar Simha, Satyavrat Bondre and Vaibhav Devanathan for the inside jokes developed over more than a decade of friendship, and more recently for the times they talked sense into me and on occasion kept me from quitting. I am grateful to my office-mates Aby Iyer, Andy MacGillivray, Angus Creech, Dougal Burnett, Lucy Cradden, N.T. Reddy, Samantha Quinn and Zhi Yung Tay for allowing me to pick their brains without prior appointment, and for putting up with my quirks without (too much) complaint. I am also grateful to Mathew Topper for help with Python and for developing the LaTeX template which took the pain out of producing this thesis. I want to thank George Lavidas for being a patient sounding board (and not a patient sounding bored) and 'auntie' Pauline for all the banter and cake. A concluding thank you to all the lovely people at IES for making the last three years a pleasure.

The studentship from the Engineering and Physical Research Council (EPSRC) via the Adaptation and Resilience in Energy Systems (ARIES) consortium is gratefully acknowledged.

I would like to acknowledge Dr. Hendrik Tolman, Dr. Arun Chawla and Dr. Henriques Alves of NCEP for their advice on wave modelling as well as for letting me have WAVEWATCH III version 4.18 (beta) well before its official release. Hindcast wind data used in the study were obtained from the European Centre for Medium-Range Weather Forecasts, National Oceanic and Atmospheric Administration (via the Research Data Archive at NCAR) and the National Centers for Environmental Prediction. Bathymetry was obtained from the National Geophysical Data Center. GCM projection data for climate change scenarios were sourced from the Coupled Model Intercomparison Project phase 3.

The numerical modelling involved in this work was performed using the UK National Supercomputing Resources HECToR and Archer.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

Atul Agarwal

Contents

Figures and Tables

Figures

Tables

Nomenclature

β_{max}	Linear wind growth parameter
Δt_s	Source term time step
Δt_k	Intraspectral propagation time step
Δt_{xy}	Spatial propagation time step
Δtg	Global model time step
Γ	JONSWAP Bottom friction coefficient
$\hat{\mu}$	Normalised mean
$\hat{\sigma}$	Normalised standard deviation
λ	Longitude
D_m	Mean wave direction
H_max	Maximum significant wave height
H_s	Significant wave height
H_{95}	95th percentile significant wave height
N	Frequency-direction wave spectrum
T_e	Mean energy period
T_p	Average peak wave period
T_p	Peak wave period
T_{95}	95th percentile energy period
T_{max}	Maximum energy period
μ	Mean
ν,var	Variance
$\overline{H}_s$	Mean significant wave height
ϕ	Latitude
σ	Standard deviation
θ	Direction
c_g	Group velocity
d	Depth
f	Frequency
g	Acceleration due to gravity (=9.81m/s)
k	Wave number
P	Wave power density
S	Net effect of sources and sinks
s_1	Swell Feedback parameter
SI	Scatter Index

Abbreviations

20CR	20th Century Reanalysis Project
AMIP	Atmospheric Model Intercomparison Project
ANOVA	Analysis of Variance
AOGCM	Atmosphere-Ocean General Circulation Model
AR4	IPCC Fourth Assessment Report
AR5	IPCC Fifth Assessment Report
CEFAS	Centre for Environment, Fisheries and Aquaculture Science
CFL	Courant-Friedrichs-Lewy Number
CFS	NCEP Climate Forecast System
CFSR	Climate Forecast System Reanalysis
CMIP3	Coupled Model Intercomparison Project phase 3
CMIP5	Coupled Model Intercomparison Project phase 5
CO2	Carbon dioxide
DoE	Design of Experiments
ECMWF	European Centre for Medium-Range Weather Forecasts
ERA-40	ECMWF 40-year reanalysis project
ERA-I	ECMWF Interim reanalysis project
EU	European Union
FAR	IPCC First Assessment Report
GCM	General Circulation Model
GEV	Generalised extreme value
GHG	Greenhouse Gas
GPD	Generalised Pareto Distribution
GSE	Garden Sprinkler Effect
IEA	International Energy Agency
IFS	ECMWF Integrated Forecast System
IPCC	Intergovernmental Panel on Climate Change
JODC	Japan Oceanographic Data Center
JONSWAP	Joint North Sea Wave Project
LHS	Latin Hypercube Sampling scheme
MCS	Monte-Carlo Simulation
MPI	Max Planck Institute, Germany
MSLP	Mean Sea Level Pressure
NASA	National Aeronautics and Space Administration, USA

NCAR	National Center for Atmospheric Research, USA
NCEP	National Centers for Environmental Prediction, USA
NDBC	National Buoy Data Center, USA
NGDC	National Geophysical Data Center, USA
NOAA	National Oceanic and Atmospheric Administration, USA
NWM	Numerical Weather Model
PDF	Probability Density Function
POT	Peaks-over-threshold
RCM	Regional Climate Model
RSM	Response Surface Modelling
SAR	IPCC Second Assessment Report
SRES	IPCC Special Report on Emissions Scenarios
SWAN	Simulating WAves Nearshore
TAR	IPCC Third Assessment Report
UKCIP02	UK Climate Impacts Programme
UKCP09	UK Climate Projections 2009
UKMO	UK Meteorological Office
UNEP	United Nations Environment Programme
UNFCCC	United Nations Framework Convention on Climate Change
WEC	Wave energy converter
WMO	World Meteorological Organization
WW3	WAVEWATCH III

Chapter 1

Introduction

Climate change is potentially the greatest challenge faced by mankind today. Initially used to describe an unprecedented increase in global mean surface temperatures since the mid 1970s the moniker today, encompasses changes in climatological variables including temperature, precipitation, pressure and circulation patterns, to name a few. A substantial body of evidence suggests that these changes are directly linked to atmospheric levels of greenhouse gases (GHGs), such as carbon dioxide, altered by human activity.

Climate change is closely followed by the ever-increasing demand for energy of a burgeoning global population. Energy consumption worldwide has grown from a little over 4000 million tonnes of oil equivalent (Mtoe) in 1971 to almost 9000Mtoe in 2012. In the same period, global electricity consumption has nearly quadrupled from 440Mtoe to 1626Mtoe (IEA, 2014). Considering that large parts of the developing world are yet to be provided access to an uninterrupted power supply and taking into account the rate of population growth, these figures are only likely to grow larger in the future.

In a bid to assuage climate change and secure energy supply, there has been significant effort at increasing electricity generation from low-carbon sources. Although a global commitment towards this end is yet to be reached, regional efforts towards achieving targets of GHG emissions and increasing the contribution to the energy mix from renewables can already be seen.

A notable example is the EU 20-20-20 strategy to reduce GHG emissions by 20%, increase energy efficiency by 20% and to source 20% of the energy consumed from renewables by the year 2020 (Council of the European Union, 2008). As part these targets, the United Kingdom is committed to meeting 15% of its energy demand from renewable sources (Department of Energy and Climate Change, 2011).

The roadmap to 2020 includes various mechanisms such as the Renewables Obligations scheme and feed-in-tariffs to encourage the integration of renewable energy technologies into the generation mix. An illustrative scenario of technologies deployed for achieving the UK's 2020 renewable energy targets is shown in Figure 1.1.

From the figure, it can be seen that mature technologies such as onshore wind and hydro are likely to contribute to the targets in a large way. However, an almost equal contribution

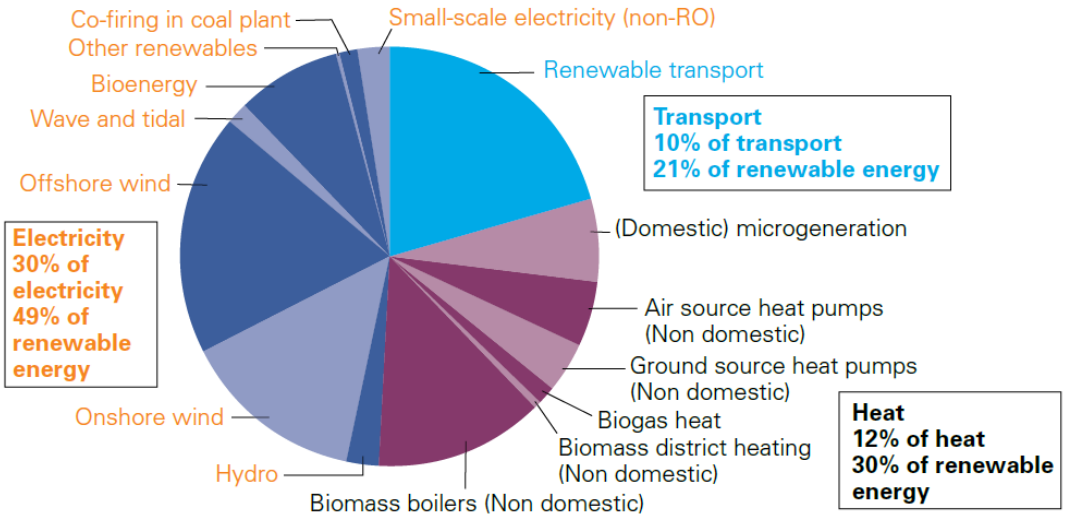

Figure 1.1: Lead scenario of technology mix for achieving the UK's 2020 target of 15% energy from renewables. Sourced from HM Government (2009)

can be expected from offshore sources with the deployment of offshore wind farms and the development of wave and tidal energy converters.

The development of commercial scale marine energy converters is particularly evident in Scotland, which boasts of 25% of Europe's tidal resource and 10% of Europe's wave resource. Potential wave energy capacity in the region, by itself, is in the region of 14GW (Renewable Policy Team, 2008). The recent leasing of marine energy sites in UK waters is a sign of the impending large-scale deployment of these devices (The Crown Estate, 2014).

It is known that the marine environment has the potential to be extremely destructive. As a result, all offshore structures, whether they be oil platforms or renewable energy projects, need to be designed for survivability under the most extreme loading conditions such as storms and hurricanes. In addition, structures such as a wave energy converters (WECs) rely on the marine environment for the generation of electricity. Changes in the wave climate could adversely affect the output of these devices resulting in the projects being financially non-viable over their lifetimes.

As a result of the uncertainty in extreme events and changes in the marine environment, the risk associated with offshore energy projects is not small. Before large-scale offshore energy projects can be realised, improved knowledge of the environment and nature of the changes is required so that the risk associated with them can be quantified and mitigated.

This research project attempts to bridge the gap in the understanding of long term trends and climatic changes in the wave climate by analysing historical and projected wave conditions in the North East Atlantic and North Sea. The objectives of the work are elaborated on in the following sections.

1.1 Research objectives

This doctoral research attempts to test the hypothesis that long-term climatic trends exist in the wave climate of the North East Atlantic and North Sea which are a consequence of greenhouse gas induced climate change.

Towards this end, the research seeks to answer the following questions:

- Has the wave climate in the North East Atlantic and North Sea changed over the last century?
- Are these changes, if any, significant trends or merely a result of short-term variability?
- Are these changes, if any, a consequence of greenhouse-gas (GHG) induced climate change? If so, what are the possible states of the wave climate in the 21st century for different climate change scenarios?
- What are the design wave conditions, estimated as 100-year extreme wave heights, for the 21st century?

It is expected that the production of a homogenous dataset of hindcast wave conditions spanning over centennial scales would enable analyses of wave conditions for the detection of climatic changes in the region. In addition, a dataset of projected wave conditions in the 21st century for different climate change experiments would allow the exploration of a causal link with atmospheric greenhouse gas (GHG) concentrations. Possible scenarios of the wave climate in the 21st century can also be developed based on these.

The project also aims to develop a estimates of extreme wave heights in the region based on statistical methods to aid in the design and deployment of marine structures including wave energy converters (WECs), tidal turbines and offshore wind farms.

In order to achieve the aims outlined above, the project is divided into the following tasks:

- Set up a suitable numerical wave model to accurately simulate wave conditions in the region of interest. This includes assessment of uncertainty in model output, sensitivity to model tuning parameters and development of a robust approach for model calibration;
- Develop a high-resolution gridded time-series dataset of historical significant wave height, mean energy wave period and mean wave direction spanning multiple decades;
- Develop a high-resolution gridded time-series dataset of projected significant wave height, mean energy wave period and mean wave direction for the 21st century for different climate change scenarios;
- Identify and apply suitable methods of time-series analyses for the detection of climatic trends in the wave climate;
- Investigate the existence of a causal link between atmospheric GHG levels and the wave climate using suitable statistical methods; and
- Use extreme value analysis methods to produce a gridded dataset of 100-year return values of significant wave height to estimate extreme conditions in the region for the

21st century.

1.2 Contribution to knowledge

The need for this work arises out of the lack of consensus about historically observed trends and future projections of the wave climate in the North East Atlantic and North Sea.

Studies, in the past, have established an upward trend in wave heights in the region, however, little work appears to have been done to attribute these to any physical phenomena. To a large extent, this has been due to the lack of a coherent, homogenous dataset of wave conditions that extends over centennial scales which can be used to differentiate climatic trends from shorter term variations.

Several datasets of wave conditions are generated which are an essential part of this study. These include a long-term, hourly hindcast dataset of wave conditions in the region for the period 1871-2010 as well an hourly dataset of projected wave climate for the high, medium and low emissions scenarios, prepared by forcing the wave model with general circulation model (GCM) winds. In addition to being invaluable for study, it is anticipated that these datasets could prove useful to a host of future analyses and research projects.

Currently, several oceanic and coastal models exist which are capable of simulating wave conditions to a high degree of accuracy. To do so, however, they first need to be 'tuned' by adjusting the values of parameters that determine the effect of various physical processes which usually differ from one region to another. This study proposes and demonstrates a structured approach for identifying the optimum parameter values for accurately simulating wave conditions in the region. In addition, the uncertainty in WAVEWATCH III version 4.18 due to selected tuning parameters is assessed and the model's sensitivity to these is quantified.

A methodology for estimating long-period return values by applying the peaks-over-threshold to numerical model data is demonstrated. This includes selection of an appropriate extreme-value distribution, threshold selection and size of dataset. A method of calibrating the raw sample is also proposed.

Regional maps of historical and projected trends in the wave climate are prepared for several wave parameters. Maps of 100-year return significant wave heights are also prepared along with the 90% confidence interval to assist in the design of marine structures and site selection for marine energy projects.

1.3 Thesis outline

This thesis comprises of 9 Chapters and 3 Appendices. The current chapter introduces the context of the work and outlines the scope and objectives of the study.

Chapter 2 presents an introduction to climate change followed by a survey of notable work done so far in the area of variability and climate change impact on ocean waves. Global and regional efforts in the identification and attribution of historical changes are briefly described along with an introduction to scenarios used for the assessment of future impacts of these changes thus setting the background of the doctoral work undertaken.

Chapter 3 introduces the 3rd generation spectral wave model, WAVEWATCH III version 4.18, used in this study for generating data describing the wave climate in the region of interest. The governing equations as solved by the model are briefly introduced followed by a detailed description of the numerical approaches, propagation schemes and parameterisation of physics for the generation, propagation and dissipation of wind seas and swells selected for the study from the different available options.

Chapter 4 describes the process of setting up WAVEWATCH III for accurately modelling the North East Atlantic and North Sea. The available datasets for forcing the wave model are introduced and a score-based comparison is described, based on which the 20th Century Reanalysis (20CR) dataset is selected for generating a long-term hindcast of wave conditions. A Monte-Carlo type analysis is presented for quantifying the uncertainty in the model's output due to epistemic uncertainty in the values of parameters governing wind growth, swell feedback and bottom friction. The sensitivity of the model to these parameters is also tested. A simultaneous optimisation approach for selecting a suitable combination of input parameter values using the Response Surface Methodology (RSM) to minimise bias in significant wave height and peak wave period is demonstrated and the output of the model calibrated in this manner is validated with buoy measurements.

Chapter 5 introduces and compares trend analysis methods used with other climatological parameters such as temperature. The hinge-fit method for trend analysis is selected for use with wave parameters and trends in the historical wave climate between 1871 and 2010 are analysed at site and regional levels. Observed trends in mean significant wave height, energy period, mean direction as well as storm events represented by the annual 95th percentile and annual maximum significant wave heights and the wave energy resource are discussed. Statistical significance is tested using the Student's t-test.

Chapter 6 presents an analysis of wave climate projections for the 21st century. Wave data is generated by forcing the wave model with GCM winds for high, medium and low emissions experiments and analysed at site and regional levels. The Analysis of Variance (ANOVA) technique is applied to wave data from the scenarios for the detection of a causal link between atmospheric GHG levels and the wave climate. Finally, a linear least-squares function is fit to

the projected wave scenarios and trends are analysed at site and regional levels.

Chapter 7 presents estimates of extreme wave heights in the 21st century based on statistical analyses of historical data. A brief review of methods of extreme value analysis is presented followed by a detailed description of the 'Peaks-over-threshold' method used in the study. The performance of the Generalised Extreme Value (GEV) family of distributions and the Generalised Pareto Distribution (GPD) is compared and the latter is chosen as being more appropriate. A calibration method is proposed and estimates of 100-year return values for significant wave height are presented for the North East Atlantic and North Sea region along with 5% and 95% confidence bounds. The preliminary research presented in this chapter draws significantly from early work during the doctoral research using ERA-Interim data which has since been published as a peer-reviewed paper (Agarwal *et al.*, 2013).

Chapter 8 discusses aspects of the study that affect the body of work as a whole. These include constraints and limitations in computing and data storage, inaccuracies in the output dataset due to the Garden Sprinkler Effect, uncertainty in historical wind, dependence of projected wave scenarios on GCM skill and advances in modelling and climate scenarios since the IPCC Fourth Assessment Report.

Finally, the conclusions of the research project are presented in Chapter 9.

Chapter 2

Background and literature review

2.1 Climate change

Climate change is one of the greatest scientific, political and socio-economic challenges facing the world today. The effects of this phenomenon encompass increases in surface temperatures, changes in precipitation patterns, rise in sea levels and an increase in extreme events, which pose great risks to communities across the globe. Considering the seriousness of the impacts, significant scientific effort has been dedicated to the study climate change.

In this section, an overview of the subject is presented along with definition of key concepts and terms and methods of attribution and detection.

2.1.1 Intergovernmental Panel on Climate Change (IPCC)

The Intergovernmental Panel on Climate Change (IPCC) is a scientific body established in 1988 under the United Nations Environment Programme (UNEP) and the World Meteorological Organisation (WMO) with the objective of assessing and disseminating the latest scientific, technical and socio-economic information on climate change.

While the IPCC does not conduct any scientific research in the area, it publishes periodic reports, once every four to six years, providing what its panel of experts consider the most relevant scientific information on climate change and its potential impacts, as well as strategies for mitigation and adaptation to guide policy makers. These reports are comprehensive treatises which present the scientific basis of climate change, the methodologies followed for the analyses and a detailed disussion of risks and impacts at a global scale.

The most recent report, published in late 2013, is the Fifth Assessment Report (AR5). For this research project, however, the earlier Fourth Assessment Report (AR4) published in 2007 is used on the basis of being, for large duration of the project, the most recent and widely accepted source of information and data on the subject.

2.1.2 UK Climate Projections 2009 (UKCP09)

While the IPCC presents climate change information on a global scale, a regional analysis of projected climate for the UK is presented in the UK Climate Projections (UKCP09) report prepared by the UK Meteorological Office (UKMO). This is the successor of the earlier UK Climate Impacts Programme (UKCIP02) .

Climate projections were developed for high, medium and low GHG emissions scenarios, which are equivalent to the A1F1, A1B and B1 SRES scenarios. However, for brevity, only projections and impacts for the medium emissions scenario (SRES A1B) are presented in the report.

The report uses data from the third ocean-atmosphere coupled GCM at the Hadley Centre, HadCM3, run at a resolution of 2.5° latitude × 3.75° longitude with 19 vertical levels over land. The ocean component of the model is run at 1.25° latitude × 1.25° longitude with 20 vertical levels, which can be visualised in Figure 2.1.

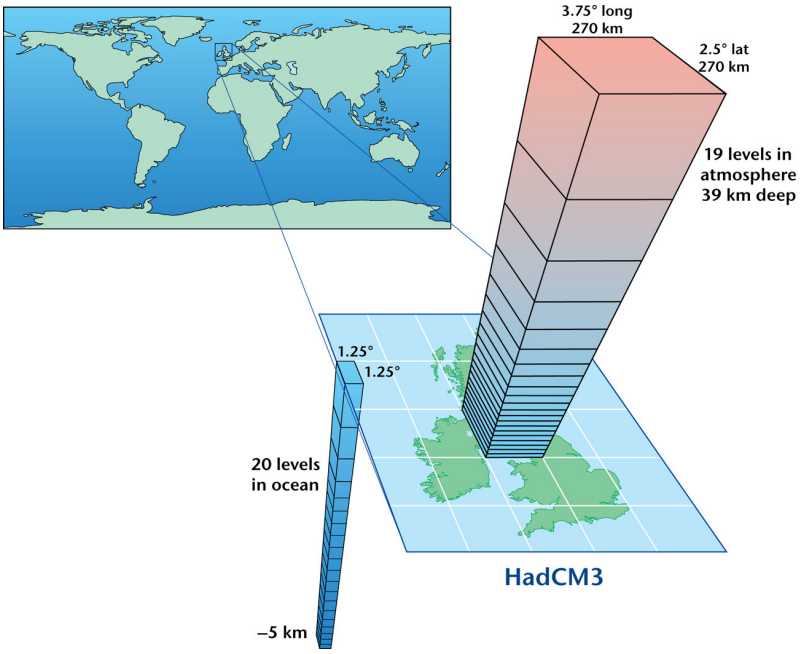

Figure 2.1: Schematic representation of the HadCM3 AOGCM. Sourced from Lowe *et al.* (2009)

The UK Climate Projections report presents probabilistic projections of future climate for the three emissions scenarios simulated. This is done by using a perturbed physics ensemble (PPE), which is generated by running the climate model for multiple realisations and varying the values of parameters whose values are uncertain. The atmospheric PPE is generated by running 280 variants of HadSM3 (the slab-ocean version of HadCM3) by perturbing physical parameters within plausible ranges. A statistical emulator is then trained using data from the PPE to explore a larger range of variability within the model. In this manner, the model

uncertainty and the epistemic uncertainty due to physical parameters can be quantified for the simulated projections of the climate.

2.1.3 Definitions of key terms

Climate: The term 'climate' is a statistical description of long-term weather conditions usually in terms of the mean and variability over a suitable period of time. This period may range from months to thousands of years, depending on the parameter being studied. However, an averaging period of 30 years is commonly used to determine the mean 'climate', on the recommendations of the World Meteorological Organization (1989).

Climate change: This refers to a change in the state of the climate that persists on decadal or longer time scales. These can be identified by changes in the mean or variability of the climatological parameter by various methods such as statistical tests. The United Nations Framework Convention on Climate Change (UNFCCC) defines the term 'anthropogenic climate change' to describe climatic changes that are directly or indirectly a consequence of alterations in the composition of the global atmosphere caused by human activity.

Detection: The IPCC defines detection of climate change as "*the process of demonstrating that the climate has changed in a defined statistical sense without providing a reason for the change*".

Attribution: Attribution is defined by the IPCC as the process of establishing the most likely causes for the detected changes in the climatological variables with some level of confidence.

Internal variability: Sometimes also referred to as natural climate variability, this term refers to variations in the statistical parameters describing the state of the climate due to natural, internal processes.

2.1.4 Detection and attribution of climate change

In order to establish the occurrence of the phenomenon of climate change, it is important to show that the observed changes in climatological parameters are not a result of natural variability in the climate system. Examination of historical records of climate variables could provide insight into the natural variability of the system. In the absence of sufficiently accurate records extending over 150 years into the past, proxies such as tree rings have been used to reconstruct the state of the climate. While these are informative, it must be observed that these are not infallible.

It can be observed in historical records of temperature that the mean value of the parameter has increased over the last one-and-half centuries (see Figure 2.2). A sharp upwards trend in global mean temperatures is also seen from the mid 1970s.

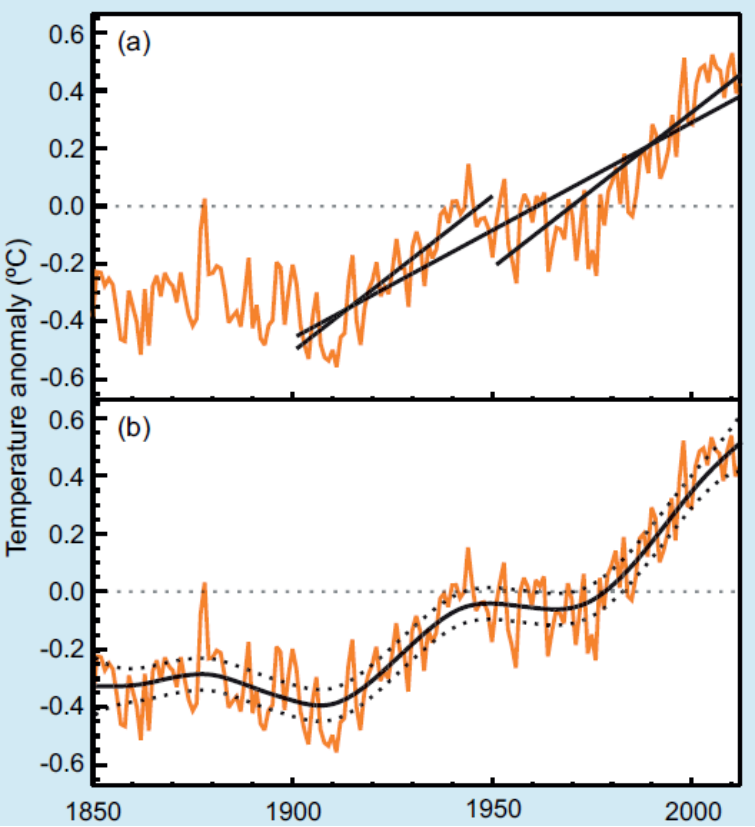

Figure 2.2: Observed annual global mean surface temperature anomaly from 1850 to 2012. Top: Simple trends fitted to the data; bottom: moving average and 90% confidence interval fitted to data. Figure sourced from Hartmann *et al.* (2013)

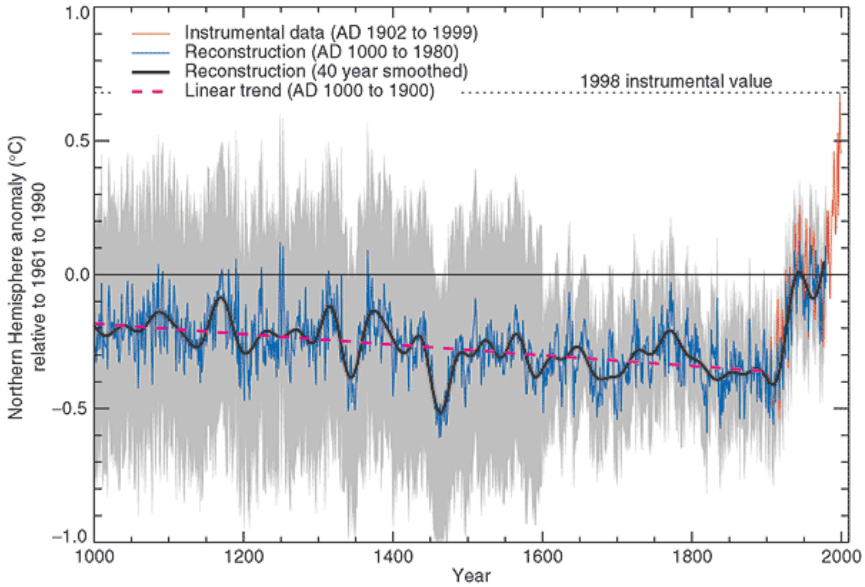

Figure 2.3: Reconstructed timeseries of temperature from AD 1000 to 1999. Sourced from Folland *et al.* (2001)

A reconstruction of millenial timeseries of temperature from AD 1000 to 1999 from proxies, plotted in Figure 2.3 shows that the changes in the parameter over the last century are, indeed, unprecedented. Consequently, the IPCC concluded that it is 'very likely' (confidence >90%) that the climate change trend observed since the late 20th century is not a result of internal variability (IPCC, 2007).

The process of detection described above is not yet possible for all climatological variables for various reasons. These may be weak response to external drivers and forcings, less reliable observations and proxies, and inadequate ability of the model in simulating these variables.

Attribution of these climatic changes to anthropogenic causes is more complex than detection even for key indicators such as temperature. Due to the nature of the climate system, it is not possible to conduct controlled experiments by varying the strength of suspected human drivers of the climate system and observing the response in the variable.

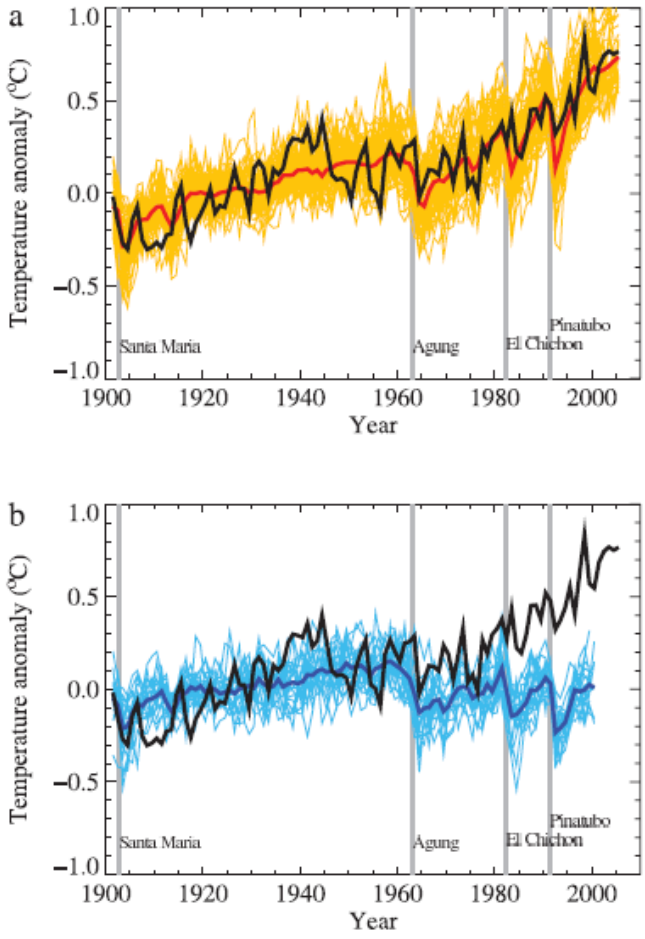

Figure 2.4: Comparison between observed (black line) and simulated global mean temperature anomalies from AOGCM runs. (a) Natural and anthropogenic forcings (b) Natural forcings only. Figure sourced from Hegerl *et al.* (2007)

With the availability of sophisticated climate models, GCMs and AOGCMs, it is possible to simulate experiments with the climate system and compare them with observations. For attribution of the observed climate change to anthropogenic causes, these climate models were run with (i) only natural forcings, (ii) only anthropogenic forcings and (iii) both types of forcings. Anthropogenic forcings include levels of atmospheric greenhouse gases and aerosols, changes in land use, etc.

The large discrepancy between simulated and observed data from experiments with only natural forcings coupled with the good agreement between simulated and observed data from experiments with both types of forcings indicates that the changes in climate observed are, in most cases, 'likely' as a result of human activity. The confidence level of >66% is associated with the term 'likely'.

2.1.5 IPCC scenarios for the 21st century

Long-term projections of future climate are fraught with uncertainty, which may arise out of uncertainties in future atmospheric concentrations of greenhouse gases and aerosols, uncertainties due to the natural variability of the climate system, and the uncertainty in global climate models (GCMs) (Fronzek *et al.*, 2012).

To be able to achieve some degree of quantification of this uncertainty, several 'what-if' scenarios can be designed and analysed. It is essential that these scenarios satisfy a few basic requirements (Carter and Task Group on Data and Scenario Support for Impact and Climate Assessment, 2007):

1. They should be representative of the climate, as assessed for a suitable control period,
2. They should be physically plausible, and
3. They should be representative of the potential range of future climates

Assuming that these scenarios are comprehensive, it can be said that the real future climate at any time is likely to lie within the envelope bounded by the extreme scenarios (Hayes, 2011).

In 2000, the Intergovernmental Panel on Climate Change (IPCC) published the Special Report on Emission Scenarios (SRES) (Nakicenovic *et al.*, 2000) describing the different emissions scenarios for climate studies in exploring the future state of the global environment. These scenarios differ from each other not only in the emissions and stabilization levels of greenhouse gases and aerosols, but also in the demographic, social, economic, technological and environmental developments. The salient features of the SRES marker scenarios, representative of high, medium and low emissions trajectories, used in this research are as under (Carter and Task Group on Data and Scenario Support for Impact and Climate Assessment, 2007):

- *SRES A2 scenario:* This is a high emissions scenario in which the levels of greenhouse gases and aerosols do not stabilise. The associated socio-economic state is that of a

continuously growing population with fragmented, regional economic growth that is slower than in other scenarios.

- *SRES A1B scenario:* This is the representative scenario from the A1 family, describing a case with intermediate emissions which stabilise at the 750 ppm (parts per million of CO_2). The family of scenarios describes a world of rapid economic growth with rapid introduction of new and efficient technology. Global population peaks in mid-century and then declines.
- *SRES B1 scenario:*This scenario is representative of the low end of emissions in the SRES range, stabilisation achieved at 550 ppm. It describes a future world with a continuously growing population (lower than the A2 scenario), with local solutions to economic, social and environmental sustainability.

Climate models (GCMs and AOGCMs) are then run for experiments with these scenarios as input and corresponding projections of the future climate are obtained. By developing projections for multiple scenarios, the uncertainty due to future emissions can, to some extent, be quantified.

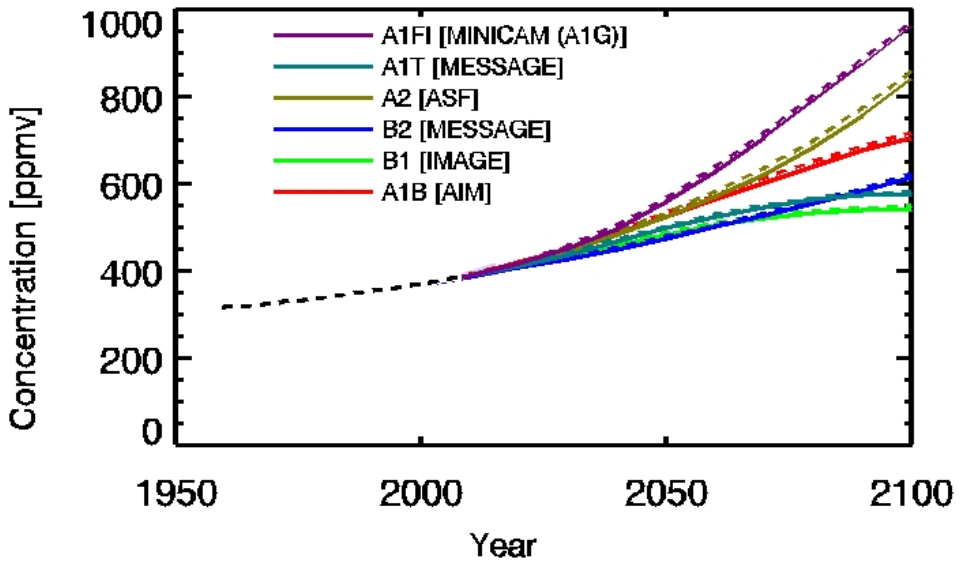

Figure 2.5: Atmospheric carbon dioxide levels observed at Mauna Loa from 1958-2008 (dashed black line) and projections for 6 SRES marker scenarios. Sources from IPCC (2014)

Figure 2.5 shows the trajectories of atmospheric CO2 concentrations for different SRES marker scenarios. It can be seen that the B1 and A1F1 represent the lower and upper bounds of the envelope respectively. However, as sub-daily GCM data is not currently available for the A1F1 experiment, the A2 experiment is taken as representative of the upper bound for this study.

2.2 A changing wave climate

As an integral element of the climate system the possibility of ocean waves being affected by climate change has not escaped the scientific community. As ocean waves are directly related to the wind patterns, any changes in these are likely to cause changes in the prevailing wave climate. With studies (e.g. Briffa *et al.* (1990); Watson *et al.* (2001)) indicating the existence of changes in wind patterns and speeds observed in the recent past, it is not unreasonable to consider that, as a consequence, wave conditions have changed as well.

Although little attention has been given to the ocean wave climate by the Intergovernmental Panel on Climate Change (IPCC) in their assessments, historical changes and possible future scenarios of wave conditions have been studied and documented by others.

In Section 2.2.1, noteworthy studies of the historical wave climate are presented providing background knowledge in the area relevant to this doctoral research. These studies explore short-term and long-term variability along with trends and changes in the wave climate. More recent studies from the era of climate-change research explore potential links between wave conditions and anthropogenic climate change. The section discusses at suitable length methods, data, conclusions as well as limitations of these studies.

Section 2.2.2 briefly discusses studies that have attempted, by various methods, to project the wave climate for the 21st century. Some of these rely on stochastic relationships between the wave climate and other climatological variables, whereas others use numerical modelling approaches.

2.2.1 Analysis of historical wave climate

Analysis of visual observations

As early as three decades ago, Neu (1984), while studying variations in the wave climate of the North Atlantic, observed a continuous increase in wave heights over the period 1970-1982. The study used synoptic charts prepared from visual observations from approximately 200 voluntary observing ships (VOS) reported every 6 hours and instrument data from buoys and oil exploration platforms.

Neu adopted the method of fitting the log-normal distribution to annual significant wave height data and preparing exceedance diagrams, as shown in Figure 2.6 (left panel). From the exceedance distributions obtained, Neu proceeded to plot the variation in significant wave height for different exceedance levels for the 13-year period, as shown in Figure 2.6 (right panel).

Neu observed a decreasing trend in the slope of the distributions fitted in Figure 2.6 (left panel) indicating that the wave height associated with an exceedance level was increasing from one year to the next. This was confirmed by upward slopes in Figure 2.6 (right panel) showing an increasing trend over the period studied.

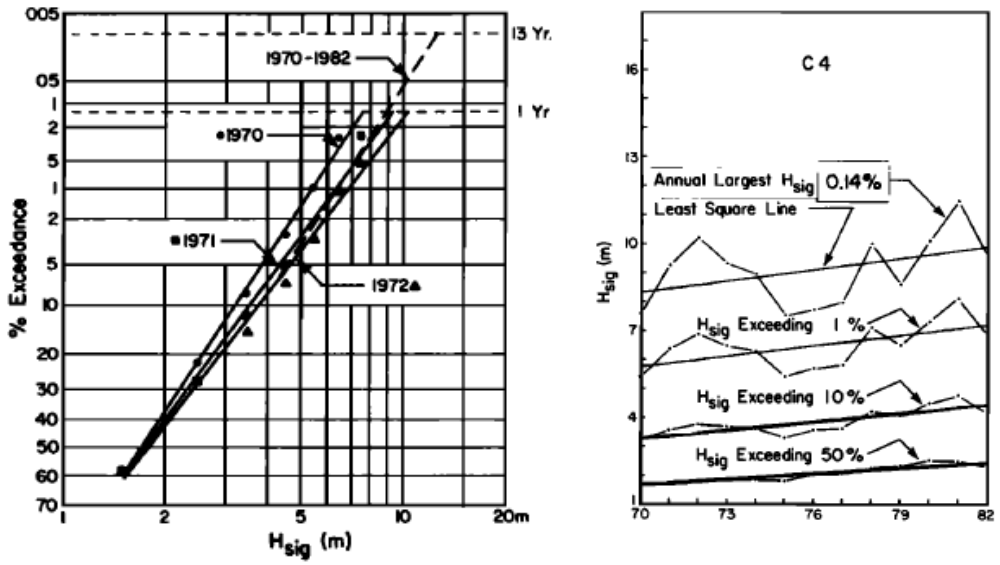

Figure 2.6: Left panel: Wave height exceedance distributions for annual H_s data from 1970 to 1972; Right panel: Variation in H_s for an area in the North Atlantic for 0.14%, 1%, 10% and 50% exceedance levels for the period 1970-1982. Figure sourced from Neu (1984)

One of the major drawbacks of this study is that data is obtained from interpolated charts prepared from VOS reports. Problems associated with VOS reports have been extensively studied and documented (e.g. Gulev (2003); Gulev *et al.* (2003); Kent and Berry (2005)). These include concentration of reports along shipping lanes, high degree of reliance on observer experience and inaccuracies such as differences in day and night observations. Moreover, a 13-year period is not adequate for estimating trends that may occur on decadal and longer timescales.

In the absence of conclusive evidence of long-term trends in wave conditions on a global scale, Gulev and Grigorieva (2004) attempted a centennial scale analysis using Voluntary Observing Ships (VOS) data from the International Comprehensive Ocean-Atmosphere Data Set (ICOADS) which contains observations going as far back as 1784.

The number of observations per region per month varies significantly in this dataset, from 15-20 in the late 1800s to several thousands from 1960 onwards, which is likely to introduce time dependent sampling errors. In order to homogenise these data, monthly means were calculated using randomly sampled subsets of the available observations for each month. The authors observed a high correlation between means calculated using different sample sizes ($r > 0.9$) and results based on 15-observation samples were reported.

The authors report weak decreasing trends on centennial time scales of -5.2cm/decade along the North Atlantic storm track in the western part of the ocean and weak positive trends of $+4.6$cm/decade off the coast of Gibraltar and their findings are shown in Figure 2.7. Statisti-

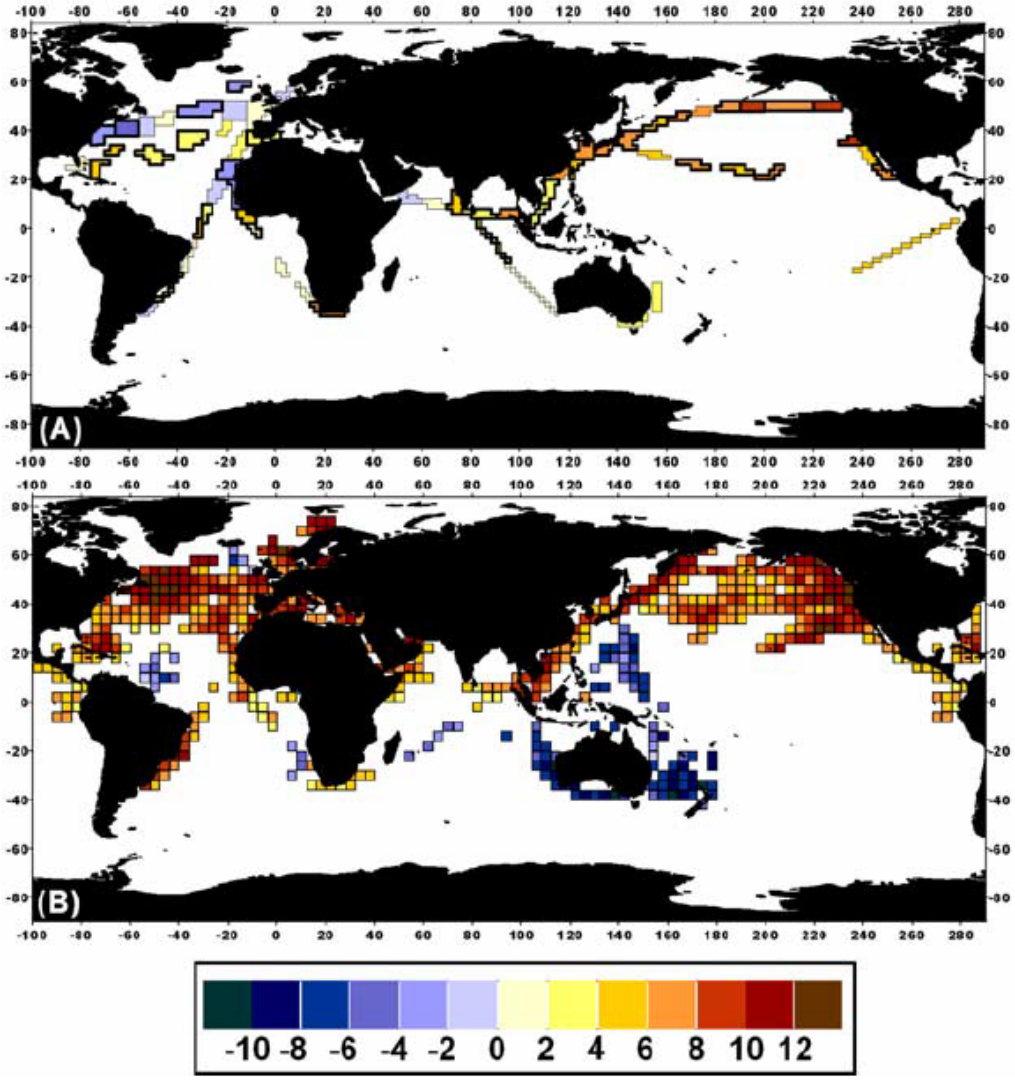

Figure 2.7: Linear trends in Hs based on VOS data (cm/decade); (A) for the period 1900-2002, (B) for the period 1950-2002.

cally significant positive trends were also observed for large parts of the North Atlantic between 1958 and 2002 which are visible in Figure 2.7(B). For the period 1976-1988, a declining trend of ≈6% was observed at the OWS station Lima, contradicting the findings of Bacon and Carter (1991) for the same period.

Gulev and Grigorieva observe that while the data is homogenised for random sampling errors, it may still be susceptible to fair weather biases as ships tend to avoid stormy conditions, and other drawbacks of visual estimates.

Analyses of measurements at individual locations

Evidence of changes in the wave climate of the North Atlantic based on instrument measurements was first provided by Draper (1986), using wave height data from 1970-1975 from shipborne wave recorders fitted on ships occupying Ocean Weather Stations (OWSs) India (located at 59°N, 19°W) and Juliett (located at 52.5°N, 16°W) for the analysis.

On comparing wave heights from 1970-1975 with similar, previously published results for data from 1955-1965, Draper observed that annual median wave heights at India and Juliett had increased by 27% and 11% respectively. Over the same period, the annual maxima were also found to increase by 35% and 30% at India and Juliett respectively. Seasonal analyses of median wave heights revealed increases across the year, ranging from 16% in Autumn to 36% in Winter at India and 0% in Spring to 21% in Summer at Juliett. Draper also observed a decrease in the mode of the distribution of wave periods of approximately 6% at both locations.

After a detailed investigation into possible errors due to instruments, prevailing climate and analysis methods including using wind as a proxy and recalculation of statistics from the earlier period, Draper concluded that these factors could not have induced the observed differences between wave heights in the sixties and the seventies. In the case of wave periods, however, Draper does not accept the decrease as a real change.

Draper concluded that despite the absence of a meteorological explanation, it is likely that the severity of wave conditions at the two sites had changed from the early 1960s to the early 1970s. He notes that these could be a result of variability from the short samples of instrument measurements. However, Draper adds that while the evidence is too little to conclude that an upward trend exists, equally, the conclusions reached by Neu (1984) cannot be disproved, and that more data is required before any such inferences can be made.

Carter and Draper (1988) also made similar observations about the roughening of the wave climate in the North East Atlantic. They too used instrument data, analysing wave heights measured by the shipborne wave recorder fitted on the Seven Stones Light Vessel, providing further evidence of an increasing trend, albeit at a single location.

Carter and Draper calculated annual mean significant wave height from 1962 to 1985 and fitted a linear regression function to the data to identify any trends in the wave climate, shown in Figure 2.8. From the slope of this line, they found an increase of 0.034m yr^{-1} and established the significance of the trend at the 99% confidence level by the Wilcoxon rank-sum test.

They observed that they were limited in their ability to determine long-term trends in H_s on account of the large interannual variability. They also noted that their analysis only describes conditions at one site and cannot be used to generalise conditions on a wider scale.

In an attempt to explain the observed trends in the wave climate at various locations on the North Atlantic side of the UK, Bacon and Carter (1993) attempted to correlate wave height

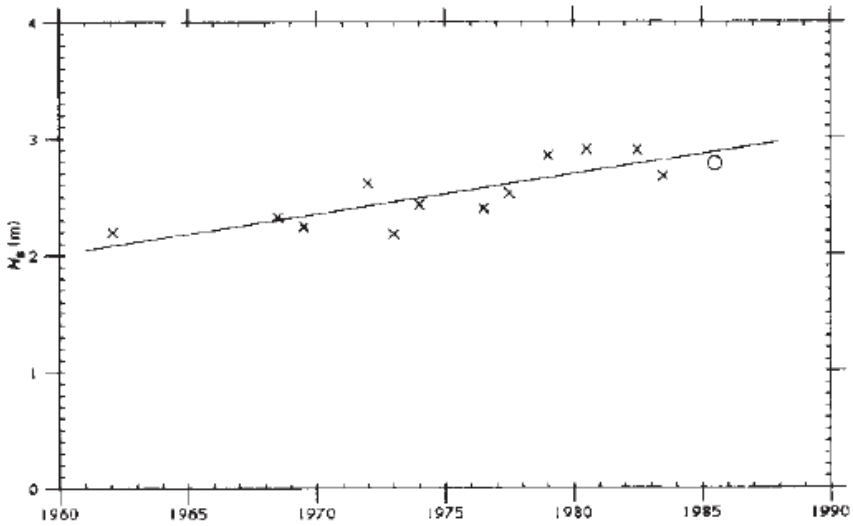

Figure 2.8: Annual mean significant wave height at the Seven Stones Light Vessel. Figure sourced from Carter and Draper (1988)

with atmospheric conditions in the North Atlantic.

A dataset of monthly pressure gradients over the North Atlantic was derived by calculating the difference in surface pressures between the central minimum of the Iceland Low (within 50° − 70°N and 60°W − 10°E) and the central maximum of the Azores High (within 20° − 40°N and 20° − 50°W). For any month m, the trends in the monthly mean significant wave height, $\overline{H}_s$ and monthly pressure gradient, P, were calculated as

$$\overline{H}_s = A_1 + B_1 \cos\left(\frac{\pi(m-0.5)}{6} + \phi\right) + \varepsilon \tag{2.1}$$

$$P = A_2 + B_2 \cos\left(\frac{\pi(m-0.5)}{6} + \phi\right) + \eta \tag{2.2}$$

where ϕ is the phase lag between the two climate parameters and ε and η are residual errors. Bacon and Carter took the two cycles to be in phase, which reduced ϕ to 0. Finally, a linear relationship between the two parameters was fit as

$$\overline{H}_s = \kappa P + (A_1 - A_2\kappa + \varepsilon - \eta\kappa) \tag{2.3}$$

where $\kappa = B_1/B_2$. The slope of this function was found to be 0.94 with significance $> 99.9\%$.

Bacon and Carter thus demonstrated that a statistically significant relationship does exist between the atmospheric pressure gradient and the mean significant wave height at the tested locations on a monthly scale. They also observed that these increases in wave height occurred despite little change in the mean wind speeds.

This study formed the basis of several later works exploring the link between variability in

wave conditions and the dominant teleconnection in the region, the North Atlantic Oscillation, e.g. Woolf *et al.* (2002); Swail *et al.* (2003); Mackay *et al.* (2010).

Analysis of satellite measurements

Given the direct dependence of ocean waves on wind patterns, Young *et al.* (2011) investigated whether systematic changes in oceanic wind speeds and wave heights had occurred. This analysis was performed on a global scale using data from satellite missions covering the period 1991-2008 for wind speeds and 1985-2010 for wave heights. As the study used data from multiple satellite missions, consistent calibration and validation of the dataset were essential considerations.

The global domain was expressed as a $2^\circ \times 2^\circ$ grid and the mean, 90th percentile and 99th percentile values were analysed. Significance of any obtained trends was tested at the 90% and 95% confidence levels.

Young *et al.* observed that there was an overall increase in global wind speeds for all three statistics, with similar trends observed for the mean and 90th percentile values. However, the magnitude of increase in the 99th percentile value was found to be larger, leading them to infer that the intensity of extreme events was increasing at a greater rate than mean conditions.

In the analysis of wave heights, a weak negative trend of $\approx$ -0.25% per year was observed for large parts of the North Pacific and North Atlantic Oceans in the mean value. The 90th and 99th percentile values exhibited statistically significant positive trends of $\approx$ 0.25% and 0.50% respectively in the higher latitudes. The results of the study for the mean values are shown in Figure 2.9.

Although the analysis detected strong trends in the wind speeds and complex spatial trends in wave heights, a causal relationship between the two could not be established. It is likely that this is because of complex factors affecting wave conditions such as the effect of local wind seas, remotely generated swells and local effects such as bathymetry.

2.2.2 Projections of the wave climate

Stochastic and statistical methods

Variations in the significant wave height in the North Atlantic have been found to be closely associated with seasonal mean variations in the sea level pressure (SLP) in the region. Wang *et al.* (2004) used this relationship between the two variables to predict and project significant wave heights. As the strength of this relationship is stronger during winter and fall, the analysis was restricted to these seasons.

Wang *et al.* produced a hindcast of wave conditions for 1958-1997 by forcing the third generation wave model OWI 3G with reanalysis winds and seasonal means for the period taken from

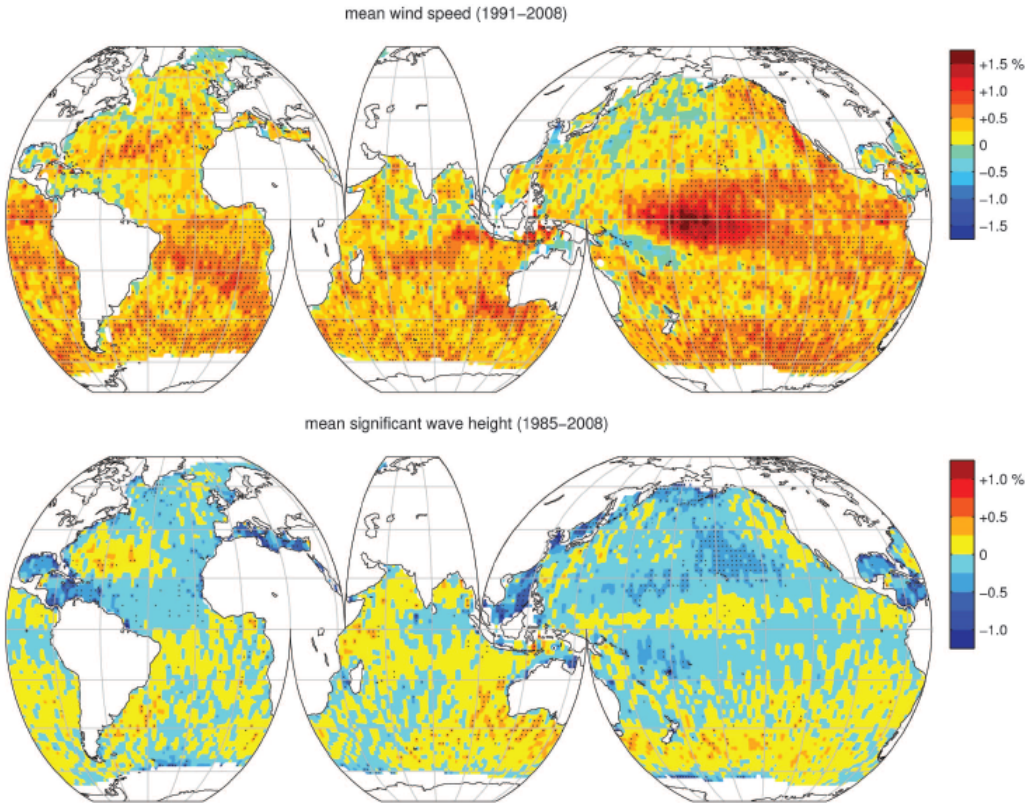

Figure 2.9: Global trends in mean wind speed (top) and mean significant wave height (bottom) in % per year (Young *et al.*, 2011)

the NCEP/NCAR reanalysis. The redundancy analysis approach was used to fit an empirical model relating seasonal mean H_s to the seasonal mean SLP and non-stationary generalised extreme value (GEV) analyses were used to relate extreme wave heights to SLP.

SLP data were obtained from the Canadian climate model CGCM2 forced for 1% increase per yer in GHGs (IS92a), SRES A2 and SRES B2 emissions scenarios. The empirical models trained using hindcast data was then used with the projected SLP data to predict seasonal statistics of significant wave height.

Strong temporal trends were found for the IS92a and A2 scenarios and a weaker trend was found for the B2 scenario. These trends were found to be quadratic in the Norwegian Sea and North Sea. In some cases, a decreasing trend was observed initially followed by increases in the latter part of the 21st century. Between 1990 and 2099, increases of up to 32cm or 11% of the baseline climate were observed in mean H_s for the A2 scenario, whereas no significant changes were predicted for the B2 scenario. Similar trends were observed for the 90th percentile winter significant wave heights with small differences in lower latitudes.

Increases in the return wave height (or reduction in the waiting time between large wave height events) were observed from the projected time-series of extreme conditions. 20-year return

values in the Norwegian and North Seas were estimated to increase between 15cm and 45cm for the A2 scenario. The largest increase in the Northeast Atlantic was found to occur during fall.

The authors note that although the empirical relationships derived by them are robust, their analysis is dependent on the assumption that these do not change for different forcing scenarios and future climates.

Wave energy converters (WECs) rely on the wave climate, specifically the prevailing significant wave height and wave energy period, for the generation of electricity. Changes in wind patterns and speeds, especially in oceanic waters, as a result of anthropogenic climate change could have a direct effect on prevailing wave conditions and therefore on the performance of WECs. Realising this, Harrison and Wallace (2005) were among the first to assess the sensitivity of the production and economics of wave energy to these changes.

They proposed a methodology linking the Pierson-Moskowitz spectrum for wind-waves with the Raleigh distribution for local wind speeds to obtain H_s and T_e data for a range of incremental wind speeds. Using these, the annual energy production was calculated as a function of the available resource and device efficiency from which cost of electricity could be obtained.

The authors demonstrated the approach for a wave energy site off the west coast of Scotland using the Pelamis wave energy converter and varying the mean annual wind speed between $\pm$ 20% in steps of 10% and evaluating the energy production and other metrics of economic performance. Their findings are summarised in Table 2.1.

Measure	Annual mean wind speed change			
	-20%	-10%	10%	20%
Mean H_s (m)	1.73	2.19	3.27	3.88
Mean T_e (s)	5.00	5.63	6.88	7.50
Mean wave power (kW/m)	27.5	49.5	134.4	205.6
Mean output (kW)	134.4	183.8	279.4	322.5
Production (GWh/yr)	1.18	1.61	2.45	2.83
Load Factor (%)	17.9	24.5	37.3	43.0
Time at idle (%)	48.7	41.0	29.7	25.7
At capacity (%)	0.9	2.4	8.2	12.2
IRR (%)	2.45	6.18	12.16	14.63
Unit cost (p/kWh), 8%	9.43	6.89	4.54	3.93
Unit cost (p/kWh), 15%	14.68	10.74	7.06	6.12

Table 2.1: Summary of changes with variation in wind speeds (Harrison and Wallace, 2005)

The authors acknowledge that the methodology proposed by the study is simplistic in directly linking changes in local mean wind speeds with prevailing wave conditions by using simple statistical modelling techniques. It has been demonstrated by previous studies that the local wave climate is a combination of locally generated wind-seas as well as remotely generated

swells. Later studies, e.g. Young *et al.* (2011), have also demonstrated that a trend in wind speeds is not necessarily accompanied by an equivalent trend in wave conditions indicating that the statistical model used by Harrison and Wallace may not be accurate. Despite its limitations, the study is useful in that for the first time, a methodology was proposed for assessing the impact of changes in the wave climate on the feasibility of wave energy generation.

Numerical modelling approaches

A notable project for research into changes in the wave climate, the European Waves and Storms in the North Atlantic (WASA) Project, was set up in the 1990s to investigate the possible roughening of the wave climate and intensification of storms in the region. The establishment of this research group followed reports of increased storm related damage and higher-than-before extreme waves (WASA Group, 1998).

One of the biggest challenges faced by the WASA project was the non-availability of long-term data which was free from inhomogeneities introduced by changes in observation techniques and instruments. As a solution, they generated a 40-year dataset of wave conditions for the period 1955 through 1994 by forcing the wave model WAM (The WAMDI Group, 1988) with hindcast wind fields from Det Norske Meteorologiske Institutt (DNMI), Norway and supplemented by wind data from the Fleet Numerical Meteorology and Oceanography Center (FNMOC), USA (Gunther *et al.*, 1997).

In addition to dynamic methods, the study also used a statistical model based on the work of von Storch and Reichardt (1997) for reconstructing historic percentile time-series data for the analysis of storm events. Time-series' of historical 50-, 80- and 90-percentile intramonthly wave heights were produced and analysed for two locations, Brent and Ekofisk for the period 1899 through 1994.

For the assessment of the impact of changing atmospheric carbon dioxide (CO_2) concentrations, a general circulation model (GCM) was run to simulate a climate scenario with doubled CO_2 levels from the control ($2\times CO_2$) and the output was used to drive the wave model. 6-year time slices, circa 2050, were generated for the control and $2\times CO_2$ scenarios for this analysis.

The WASA group concluded that there were no significant systematic changes in either the storm or wave climates in the region and that the observed changes were due to the natural variability in the system. Furthermore, based on an empirical analysis of the 40-year wave hindcast, they concluded that the observed increase in mean significant wave heights was related, in part, to the change in the strength of the North Atlantic Oscillation observed during the same period. Based on evidence from the analysis of the projected wave climate scenarios, the WASA group cautiously suggested moderate increases in the wave and surge climates in the region.

The study by the WASA group is significant in that a hitherto unused methodology, involving

numerical prediction of historical and projected wave conditions, was implemented for the analysis of trends and climatic changes in the wave climate of the North Atlantic. However, its findings must be treated with care for several reasons.

It is unlikely that 40-years of historical data is adequate for differentiating long-term climatic changes from interdecadal variability in the system. Moreover, in the detection of trends on centennial scales no generalisation can be made for the entire region based on merely two locations. This is especially true as the North Sea, where the sites are located, experiences a markedly different wave climate from the North Atlantic (Hasselmann *et al.*, 1973). Another drawback of the study is that the statistical regression model used for reconstructing data for Brent and Ekofisk assumes that the relationship between the mean air pressure (predictor) and the percentile significant wave height (predictand) is stationary in time, which may not be true.

As a part of the UK Climate Change Projections 2009 (UKCP09) described earlier, the impacts of climate change on the marine and coastal environment in the North Atlantic and North West European continental shelf were assessed and reported by Lowe *et al.* (2009).

The report presents projections of sea level rise, changes in extreme water levels including from changes in storm surge, the offshore wave climate in terms of significant wave height, period and direction, and other ocean parameters such as temperature, salinity and currents.

Of most interest to this doctoral research project are the projections of the wave climate, reported for the medium emissions scenario. These were developed by forcing a modified version of WAM cycle-4 with GCM winds from the perturbed physics ensemble (PPE).

Lowe *et al.* found changes in the winter mean wave height ranging from -35cm to +5cm and changes in annual maxima between -1.5m and +1m over the period studied. An example of the findings are shown in Figure 2.10, where the projected changes in the seasonal significant wave height (in metres) from 1960-1990 to 2070-2100 are plotted.

Based on the output of the Hadley Centre PPE of climate models, there is evidence of weakening of storms accompanied by a southward movement of storm tracks. Consequently, an increase in seasonal mean and extreme waves to the southwest of the UK and a decrease to the north of the UK are reported. The North Sea is expected to experience little change in these parameters.

Importantly, Lowe *et al.* observe that on account of the large interannual variability in wave parameters, differences between 30-year time-slices cannot be used effectively in detecting trends. It may be inferred from this that more value may be obtained if the entire time-series was considered for this purpose.

Developing further the work of Harrison and Wallace (2005), Reeve *et al.* (2011) investigated the impacts of climate change on wave energy generation by using a numerical modelling approach, rather than a statistical one. In order to assess the sensitivity of the wave conditions

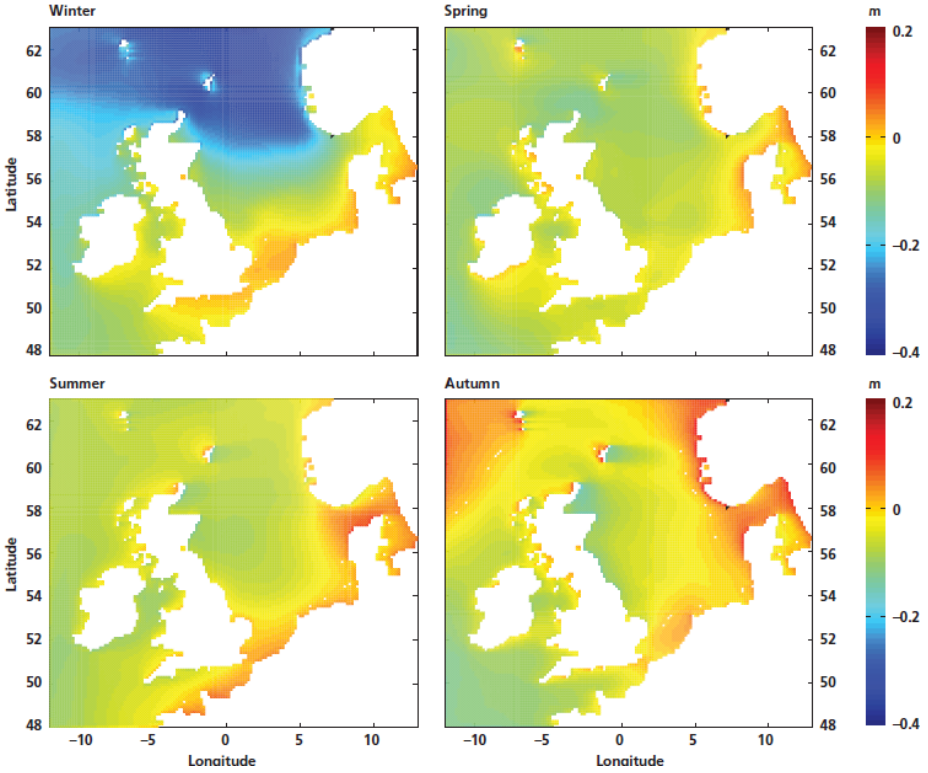

Figure 2.10: UKCP09 projected changes in seasonal Hs (in metres) from 1960-1990 to 2070-2100. Sourced from Lowe *et al.* (2009)

at the Wave Hub wave energy test site, the numerical wave model WAVEWATCH III (version 2.22) was forced using data for future wind conditions from the A1B (medium) and B1 (low) emissions scenarios from a global climate model.

Data for a control period, between 1961-2000, was generated to represent the current climate and projected climate data was simulated for the two scenarios for 2061-2100. The energy produced by the wave energy converter with wave conditions from these data was calculated and compared.

The authors found that the available wave power would increase by 2-3% for the A1B scenario from current levels, and be 4-6% less for the B1 in comparison with the A1B scenario. Contrarily, it was found that for both scenarios the wave energy yield would decrease by 2-3% due to the inability of the selected device to exploit steeper waves.

Reeve et al. noted that the changes in the climate change projections were small in comparison to the natural variability, and the Wilcoxon rank-sum test was used to establish statistical significance at the 5% level. However, these results may be misleading.

The outcomes of this study are questionable on account of inhomogeneities in the control and projected wave datasets. While the former were produced from historic wind data, the projected datasets are generated from climate model runs initialised in the year 2000 from a GCM run

simulating 20th century forcing (20C3M). As a result, a large number of climate parameters (e.g. wind patterns and sea level pressures) from the 202C3M-initialised run are found not to agree with historical observations, even when aggregated over suitable periods. This mismatch between the modelled and observed 'current climate' may be true even for wave conditions.

This type of homogeneity can be eliminated by using data from the 20C3M experiment to obtain the state of the current wave climate and use these wave data for the comparison with projected data. Results of significance tests applied to compare these can be used with greater confidence.

de Winter *et al.* (2012) studied the effect of climate change on wave conditions in the North Sea and along the Dutch coast with a particular interest in the impact on extreme conditions. To assess any changes in the wave climate, they forced a numerical wave model with an ensemble of 17 GCM runs simulating greenhouse gas emissions following the A1B climate change scenario. The period from 1950 to 2100 was modelled and the wave climate in 2071-2100 was compared with the climate in 1961-1990 to identify any significant trends or changes.

They observed near-zero changes in the mean wave conditions (H_s and T_m) between the two periods. Examples of their results are shown in Figures 2.11 and 2.12. de Winter *et al.* used the Mann-Whitney U-test (also called the Wilcoxon rank sum test) to establish statistical significance of their findings.

In their analysis of extreme conditions, they projected a decrease in annual maximum H_s of 0.2m to 0.5m depending upon the location, and near-zero changes in the parameter along the Dutch coast. The authors also found a decrease of 0.3s to 0.6s in the mean wave period, T_m, at all the locations modelled. These decreases were found to be significant at the 95% confidence level.

Changes were also seen in the direction from which these annual maxima originate, with a projected decrease of 40% in H_s maxima arriving from the north and north-west and about 60% decrease in the frequency of T_m maxima arriving from the same directions, with an increase of 30% in the maxima arriving from the south-west.

Using the Generalised Extreme Value (GEV) family of distributions, de Winter *et al.* also compared high period return values estimated using data from 1961-1990 and 2071-2100 for three locations. It was observed that the distribution remained approximately unaltered for one location and a decrease was found in the 10,000-year return H_s at the other two. These, however, were not found to be statistically significant.

This is a noteworthy study in that an ensemble of multiple GCM runs for the same emissions scenario was used for the assessment of impacts on extreme conditions. It should be mentioned that establishing the statistical significance of return values estimated from theoretical distributions is a difficult task even with the use of multi-member ensembles, and it is not surprising that the authors were unable to reject the null hypothesis at the 95% confidence level.

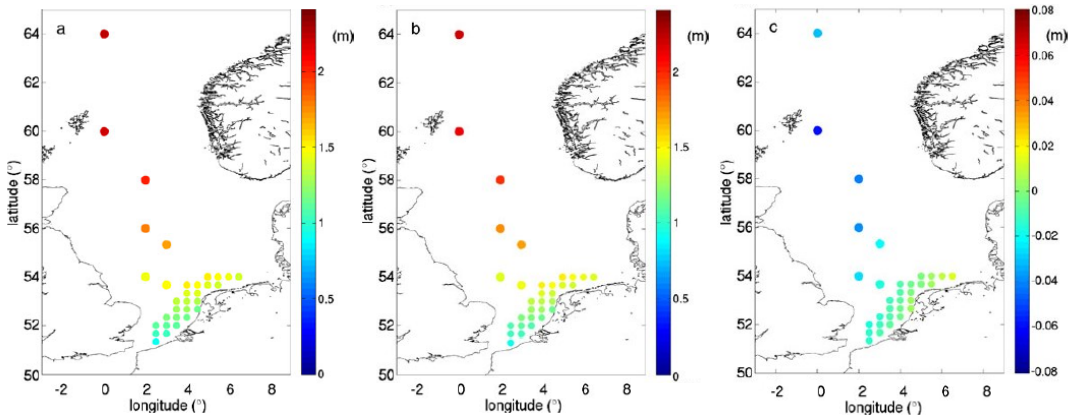

Figure 2.11: (a) Mean annual H_S in 1961-1990 compared with (b) $H_S$in 2071-2100 and (c) the difference between the periods 1961-1990 and 2071-2100 (de Winter *et al.*, 2012)

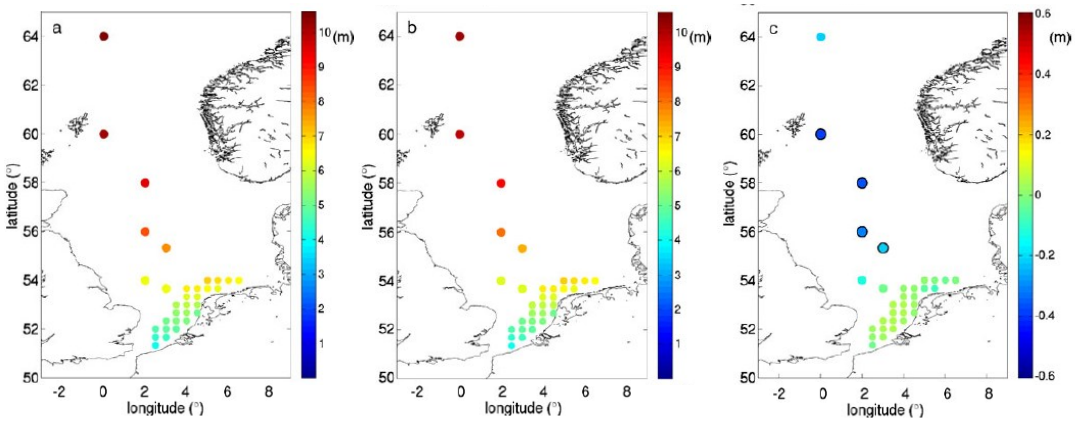

Figure 2.12: (a) Average annual maximum H_S in 1961-1990 compared with (b) average annual maximum $H_S$in 2071-2100 and (c) the difference between the periods 1961-1990 and 2071-2100 (de Winter *et al.*, 2012)

However, as with Lowe *et al.* (2009), only one climate change scenario was simulated. As real GHG emissions are unlikely to follow the emissions according to the scenarios, a more robust approach would be to use multiple scenarios in order to obtain an envelope of potential impacts by which a better quantification of risk can be derived taking scenario uncertainty into account.

In several studies of wave climate variability, the data cannot be considered as being free from systematic errors and inhomogeneities. This is especially true of the studies conducted using charts (Neu, 1984) and VOS observations (Gulev *et al.*, 2003). The trends obtained from analysing instrument measurements (e.g. Carter and Draper (1988); Bacon and Carter (1991)) are limited in that they are for single locations only and cannot be extrapolated for the region. Moreover, these data do not extend over sufficiently long periods to be able to differentiate

between climatic trends and internal variability.

While satellite observations overcome the spatial limitations of in situ measurements, the interval between consecutive measurements is quite large and is dependent on the period of orbit of the satellite. As a result, it is possible that events such as storms and hurricanes are not captured in the dataset and are omitted from the analysis.

Numerical modelling has been shown to be successful for generating data describing the wave climate suitable for spatial analyses. Most research using these methods suffered from inadequate data (e.g. 40 years of data used by the WASA Group) because of which climatic trends cannot be separated from natural variability.

Although the future wave climate has been simulated using GCM output, these studies only attempt to evaluate changes from the current climate. In addition, apart from the work of Lowe *et al.* (2009) only a single greenhouse gas emissions scenario was simulated for this purpose (e.g WASA Group (1998); de Winter *et al.* (2012)). No documented attempt could be found investigating a possible link between anthropogenic climate change and the variability in ocean wave conditions.

It is thus evident from the succinct review of literature presented in this section that a robust analysis of variability in the wave climate is necessary, especially if a causal link with greenhouse gas levels is to be investigated.

For the sake of brevity only key studies are presented in this section to provide a background to the thesis. Literature specific to the analyses are introduced and discussed in the relevant chapters.

2.3 Chapter summary

- Climate change is introduced along with a description of future emissions scenarios, methods of detection and attribution of causes;
- Notable studies of variability in the historical and projected wave climate are discussed along with their results and limitations;
- The studies are also discussed in terms of data (e.g. VOS observations, shipborne wave recorders, satellite observations and numerical model data) and in terms of methods used (e.g statistical methods and dynamical modelling);
- On the basis of the shortcomings of the research discussed, the need for this work is established.

Chapter 3

WAVEWATCH III

WAVEWATCH III (Tolman *et al.*, 2014a) is a third generation numerical wave model developed and maintained by the National Oceanic and Atmospheric Administration at the National Centers for Environmental Prediction (NOAA/NCEP). Although the wave model is based on WAVEWATCH I and WAVEWATCH II, it differs greatly from its predecessors in that the governing equations and numerical and physical approaches are inspired by the wave model WAM (Tolman, 1991).

WAVEWATCH III is different from other third-generation models in the sense that it is designed to be a modular modelling framework in addition to a numerical model. By itself, it is a capable deep water model with some nearshore modelling ability. In addition, it is designed to allow the easy integration of user-developed modules into the main body of the model. It is a freely available wave model and the source code can be obtained from (National Weather Service - Environmental Modelling Center (NOAA), 2012). The open-source nature of the model encourages and enables constant improvement and development of the model and physics driven by the larger wave modelling community, e.g. SHOWEX parameterisation for sandy bottoms (Ardhuin *et al.*, 2003), ST4 source term package Ardhuin *et al.* (2010), and nearshore processes like triad interaction Van der Westhuysen (2012).

The most recent release of the model, WAVEWATCH III version 4.18, was selected for use in this research project. The model was obtained from the developers for beta testing much earlier than the official release to the public enabling its use for the work undertaken.

In the following sections, the equations governing wave propagation, sources and sinks, and numerical approaches available are explained along with a brief description of the model structure and program elements.

3.1 Governing equations for wave propagation

It is common practice to describe the sea state in terms of a variance density spectrum $F(k;\theta;x;t)$ which is a function of wavenumber (k) and direction (θ) varying in space (x) and time (t). On account of its invariance with respect to physics for temporally varying depths (Whitham, 2006), this spectrum is used in WAVEWATCH III (Tolman *et al.*, 2014a).

Integrating such spectra over all directions gives one-dimensional spectra in the frequency domain and integrating over the entire spectrum yields the total variance (E), sometimes referred to as the wave energy in wave modelling. With the effect of currents on waves, the energy or variance of the spectrum is not conserved as is the case in the absence of currents. However, in general, the wave action is conserved and the action density spectrum $N(k;\theta)$ becomes equivalent to $F(k;\theta)/\sigma$ (Bretherton and Garrett, 1969). According to the linear wave theory, for any depth d, σ can be obtained from the dispersion equation:

$$\sigma^2 = gk\tanh kd \tag{3.1}$$

Based on this, in WAVEWATCH III wave propagation is described by:

$$\frac{DN}{Dt} = \frac{S}{\sigma} \tag{3.2}$$

where D/Dt represents the total derivative, N is the action density spectrum varying in space and time ($N(k,\theta;x,t)$), S is the net effect of all source and sinks terms. S is expanded into its constituent source and sink terms in Section 3.2.

For implementation in a numerical model, Equation 3.2 needs to be in an Eulerian form. WAVEWATCH III formulates this as a conservation equation, conserving the total wave energy, which for large-scale applications becomes Tolman *et al.* (2014a):

$$\frac{\partial N}{\partial t} + \frac{1}{\cos\phi}\frac{\partial}{\partial\phi}\dot{\phi}N\cos\theta + \frac{\partial}{\partial\lambda}\dot{\lambda}N + \frac{\partial}{\partial k}\dot{k}N + \frac{\partial}{\partial\theta}\dot{\theta}_g N = \frac{S}{\sigma} \tag{3.3}$$

$$\dot{k} = -\frac{\partial\sigma}{\partial d}\frac{\partial d}{\partial s} - k\cdot\frac{\partial U}{\partial s} \tag{3.4}$$

$$\dot{\theta} = -\frac{1}{k}\left[\frac{\partial\sigma}{\partial d}\frac{\partial d}{\partial m} - k\cdot\frac{\partial U}{\partial m}\right] \tag{3.5}$$

$$\dot{\phi} = \frac{c_g\cos\theta + U_\phi}{R} \tag{3.6}$$

$$\dot{\lambda} = \frac{c_g\sin\theta + U_\lambda}{R\cos\phi} \tag{3.7}$$

$$\dot{\theta}_g = \dot{\theta} - \frac{c_g\tan\phi\cos\theta}{R} \tag{3.8}$$

$$c_g = \frac{\partial \sigma}{\partial k} = \left[\frac{1}{2} + \frac{kd}{\sinh 2kd}\right]\frac{\sigma}{k} \tag{3.9}$$

where λ and ϕ are the longitude and latitude, s is a coordinate in direction θ and m is a coordinate perpendicular to s, R is the radius of the earth, U is the depth and time averaged current velocity, U_λ and U_ϕ are current components and c_g is the group velocity.

3.2 Source and sink terms

The net source term S in Equations 3.2 and 3.3 represents the combined effect of the energy input and dissipation due to various physical processes. In WAVEWATCH III, the net source term is expanded as (Tolman and Chalikov, 1996; Tolman *et al.*, 2014a):

$$S = S_{ln} + S_{in} + S_{nl} + S_{ds} + S_{bot} + S_{db} + S_{tr} + S_{sc} + S_{xx} \tag{3.10}$$

where S is the net effect of initial linear input (S_{ln}), exponential wind growth (S_{in}), nonlinear wave-wave interactions (S_{nl}), white-capping (S_{ds}), interaction of waves with the bottom (S_{bot}), depth induced wave breaking (S_{db}), triad wave-wave interactions (S_{tr}), and scattering of waves by the bottom (S_{sc}). WAVEWATCH III provides the ability to introduce additional, user-defined sources or sinks using the S_{xx} term.

In deep water, the net source term, neglecting initial conditions, reduces to:

$$S = S_{in} + S_{nl} + S_{ds} + S_{xx} \tag{3.11}$$

as wind growth, nonlinear wave interactions (e.g. quadruplets) and white-capping are the dominant physical processes.

For this research the input and dissipation package developed by (Ardhuin *et al.*, 2010) (hereafter referred to as ST4) was used for the deep-water sources and sinks described in Equation 3.11. This is the most recently developed package and has been reported to estimate wave conditions more accurately than older packages in WAVEWATCH III (Ardhuin *et al.*, 2010).

In addition to the ST4 package, the JONSWAP formulation for bottom friction (Hasselmann *et al.*, 1973) and the depth-induced wave breaking algorithms from Battjes and Janssen (1978) were included to model processes in shelf and shallow waters. The triad interaction term was disabled because of the computationally intensive nature of the approximations. The bottom scattering term was disabled because bottom spectrum data, usually obtained from multi-beam bathymetric surveys, was not available for the region.

The source and sink terms from the parameterisation packages used are discussed below.

3.2.1 Wind-wave interactions

The source term for the interaction of the atmosphere with the ocean surface (S_{in}) can be resolved into a positive part indicating input of energy to the waves and a negative part, indicating dissipation of energy from the waves. If S_{atm-in} represents the input of energy to waves and $S_{atm-out}$ represents the dissipation, S_{in} is written as

$$S_{in} = S_{atm-in} + S_{atm-out} \tag{3.12}$$

In the ST4 package, the positive part of the source term (S_{atm-in}) comprises of the wind input which is based on the parameterisation proposed by Janssen (1991) as adapted for ECWAM, the wave model operated by ECMWF (Bidlot *et al.*, 2007). It is formulated as (Ardhuin *et al.*, 2010; Tolman *et al.*, 2014a)

$$S_{atm-in}(k,\theta) = \frac{\rho_a}{\rho_w}\frac{\beta_{\max}}{\kappa^2}\mathrm{e}^Z Z^4 \left(\frac{u'_*}{C}\right)^2 \times \max[\cos(\theta-\theta_u),0]^p \sigma F(k,\theta) \tag{3.13}$$

where $\beta_{\max}$ is a constant non-dimensional growth parameter, κ is the von Karman constant, ρ_a and ρ_w are the densities of air and water respectively, $C = 2.5\times10^7$, $p = 2$ and Z is the wave age calculated as (Ardhuin *et al.*, 2010)

$$Z = \log(kz_1) + \frac{\kappa}{[\cos(\theta-\theta_u)(u_*/C+z_\alpha)]} \tag{3.14}$$

If u_* is the wind friction velocity, τ_w is the wave supported stress, $z_\alpha = 0.006$ is the wave age tuning parameter, and $\tau = u_*^2$, z_1 is the roughness length due to τ_w calculated as

$$U_{10} = \frac{u_*}{\kappa}\log\left(\frac{z_u}{z_1}\right) \tag{3.15}$$

$$z_0 = \min\left(\alpha_0\frac{\tau}{g}, z_{0,\max}\right) \tag{3.16}$$

$$z_1 = \frac{z_0}{\sqrt{(1-\tau_w)/\tau)}} \tag{3.17}$$

where z_u is the height of specified wind (10m in this case). z_0 is used to prevent unrealistic wind stresses at high wind speeds limited by the value of $z_{0,max} = 1.002$.

The wind friction velocity (u_*) in the original formulation of S_{atm-in} Bidlot *et al.* (2007) is replaced by wave-number dependent friction velocity (u'_*) in Equation 3.13 to allow a balance with saturation based dissipation. This is obtained from (Ardhuin *et al.*, 2010; Tolman *et al.*, 2014a)

$$(u'_*)^2 = \left| u_*^2(\cos\theta_u,\sin\theta_u) - |s_u|\int_0^k\int_0^{2\pi}\frac{S_{in}(k',\theta)}{C}(\cos\theta,\sin\theta)dk'd\theta \right| \tag{3.18}$$

where $|s_u| \approx 1$ is the sheltering coefficient.

The negative part of the source term, $S_{atm-out}$ represents the transfer of momentum from the ocean waves to the wind at the surface which causes the dissipation of swells. This dissipation has been found to be non-linear, and can be represented as the sum of viscous dissipation and turbulent dissipation as

$$S_{atm-out} = r_{visc} S_{visc} + r_{turb} S_{turb} \tag{3.19}$$

The dissipation is said to be viscous if the Reynold's number does not exceed the critical value, i.e. $Re \leq Re_c$. The Reynold's number is calculated as

$$Re = \frac{4 u_{orb} a_{orb}}{\nu_a} \tag{3.20}$$

where u_{orb} is the surface orbital velocity, a_{orb} is the displacement amplitude and ν_a is the viscosity of air at the surface. The critical Reynold's number is estimated by (Ardhuin *et al.*, 2010)

$$Re_c = \frac{2 \times 10^5}{\mathrm{H_s}} \tag{3.21}$$

For $Re \leq Re_c$, the viscous dissipation term is formulated as (Tolman *et al.*, 2014a)

$$S_{visc}(k,\theta) = -s_5 \frac{\rho_a}{\rho_w} \left\{ 2k\sqrt{2\nu\sigma} \right\} N(k,\theta) \tag{3.22}$$

and the turbulent dissipation term for $Re > Re_c$ is formulated as

$$S_{turb}(k,\theta) = -\frac{\rho_a}{\rho_w} \left\{ \frac{16 f_e \sigma^2 u_{orb}}{g} \right\} N(k,\theta) \tag{3.23}$$

where $s_5 = 1.2$ is a tuning constant, σ is obtained from the dispersion relation in Equation 3.1. The orbital velocity, u_{orb} is defined by

$$u_{orb} = 2 \left[\int \int \sigma^3 N(k,\theta) dk d\theta \right]^{\frac{1}{2}} \tag{3.24}$$

and the weights r_{visc} and r_{turb} used in Equation 3.19 are found by

$$r_{visc} = 0.5(1 - \tanh(Re - Re_c)/s_7) \tag{3.25}$$

$$r_{turb} = 0.5(1 + \tanh(Re - Re_c)/s_7) \tag{3.26}$$

It should be noted that the tuning constants s_1 through s_7 are the same as those used in WAM. However, not all the constants are used in the ST4 parameterisation (e.g. s_6).

3.2.2 Nonlinear wave-wave interactions

In this study, nonlinear quadruplet interactions are modelled using discrete interaction approximation (DIA) (Hasselmann *et al.*, 1985). Although the DIA approach is a part of the ST4 parameterisation, it is not unique to this package and has been used in wave models since WAM Cycle 3 (The WAMDI Group, 1988).

In this approach, for resonant nonlinear interactions to occur between four wave components with wave numbers k_1, k_2, k_3 and k_4 such that $k_1 = k_2 = k$ the following conditions need to be satisfied (Hasselmann *et al.*, 1985; Tolman *et al.*, 2014a):

$$k_1 + k_2 = k_3 + k_4 \tag{3.27}$$

$$\sigma_2 = \sigma_1 \tag{3.28}$$

$$\sigma_3 = (1+\lambda_{nl})\sigma_1 \tag{3.29}$$

$$\sigma_4 = (1-\lambda_{nl})\sigma_1 \tag{3.30}$$

where λ_{nl} is a constant. For each frequency-direction combination, (f_r, θ), the contribution of these quadruplets to the interaction (δS_{nl}) can be calculated as

$$\begin{pmatrix} \delta S_{nl1} \\ \delta S_{nl3} \\ \delta S_{nl4} \end{pmatrix} = D \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} C g^{-4} f_{r_1}^{11} \times \left[F_1^2 \left(\frac{F_3}{(1+\lambda_{nl})^4} + \frac{F_4}{(1-\lambda_{nl})^4} \right) - \frac{2F_1F_3F_4}{(1-\lambda_{nl}^2)^4} \right] \tag{3.31}$$

where the variance spectra are $F_1 = F(f_{r_1}, \theta_1)$, $F_3 = F(f_{r_3}, \theta_3)$, $F_4 = F(f_{r_4}, \theta_4)$ and C is a proportionality constant. D is a scaling factor (from WAM) for shallow water internally calculated as

$$D = 1 + \frac{c_1}{\bar{k}d}\left[1 - c_2\bar{k}d\right]\mathrm{e}^{-c_3\bar{k}d} \tag{3.32}$$

For depth d, $\bar{k}d$ is the mean relative depth, calculated as

$$\bar{k}d = 0.75\hat{k}d \tag{3.33}$$

$$\hat{k} = \left(\frac{1}{\sqrt{k}}\right)^{-2} \tag{3.34}$$

This correction is only valid for intermediate depths and the value of $\bar{k}d$ is not allowed to become smaller than 0.5 (The WAMDI Group, 1988).

The default values of the constants used for the approximation of nonlinear wave interactions using the ST4 source term package are tabulated in Table 3.1. These values may be adjusted by the modeller during calibration.

Alternatively, the full Boltzmann integral based on the Webb-Resio-Tracy (WRT) method (van Vledder, 2006) may be used to exactly calculate nonlinear wave interactions. However, this is

Parameter	Value
λ_{nl}	0.25
C	2.5×10^7
c_1	5.5
c_2	5/6
c_3	1.25

Table 3.1: Model default values for constants used for Discreet Interaction Approximation of nonlinear wave interactions (Hasselmann *et al.*, 1985; Ardhuin *et al.*, 2010)

computationally expensive, reportedly requiring 10^3 to 10^4 times the computational resources of the DIA approach (Tolman *et al.*, 2014a) and was deemed impractical for the research.

3.2.3 Deep water wave breaking (white-capping)

The ST4 package models dissipation in deep-water waves due to breaking and white-capping using the saturation based approach (van der Westhuysen *et al.*, 2007). In this method, a saturation spectrum B' is defined as (Ardhuin *et al.*, 2010):

$$B'(k,\theta) = \int_{\theta-\Delta\theta}^{\theta+\Delta\theta} \sigma k^3 \cos^{S_B}(\theta-\theta')N(k,\theta')d\theta' \tag{3.35}$$

where $\Delta\theta$ is the sector half-width. Default values of the constant terms are listed in Table 3.2. The dissipation term, S_{ds} is then formulated as (Ardhuin *et al.*, 2010):

$$\begin{aligned} S_{ds}(k,\theta) =& \sigma \frac{C_{ds}^{\text{sat}}}{B_r^2}\big[\delta_d \max\{(B(k)-B_r),0\}^2 \\ &+ (1-\delta_d)\max\{(B'(k,\theta)-B_r),0\}^2\big]N(k,\theta) \\ &+ S_{bk,cu}(k,\theta) + S_{turb}(k,\theta) \end{aligned} \tag{3.36}$$

In Equation 3.36, the isotropic part (the δ_d term) and the direction-dependent part (the $(1-\delta_d)$ term) are used to control the directional spread of the resulting spectrum. B_r is a saturation threshold, $S_{bk,cu}$ is the cumulative breaking term which includes dissipation due to spontaneous and induced breaking of waves and S_{turb} represents the interaction between surface waves and upper ocean turbulence. $B(k)$ is the direction-integrated spectral saturation given by:

$$B(k) = \max\{B'(k,\theta)\}, \theta \in [0, 2\pi] \tag{3.37}$$

The spectral density of breaking crest length, $\Lambda(k)$ is calculated as a function of the spectral density of wavelength per unit surface, $l(k)$, and the probability of breaking, $P(k0)$ as $\Lambda(k) = l(k)P(k)$, where

$$l(k) = \frac{1}{2\pi^2 k} \tag{3.38}$$

$$P(k) = 28.4\left(\max\left\{\left(\sqrt{B'(k,\theta)} - \sqrt{B'_r}\right), 0\right\}\right)^2 \tag{3.39}$$

If big breakers of celerity C' induce breaking in smaller waves of phase speed C, the relative velocity of crests can be calculated as $\Delta C = |C - C'|$. The cumulative breaking term, $S_{br,cu}$, can then be formulated as (Ardhuin *et al.*, 2010):

$$S_{br,cu}(k,\theta) = -C_{cu} N(k,\theta) \int_{f'<(r_{cu}\times f)} \Delta C \Lambda(k') dk' \tag{3.40}$$

where r_{cu} defines the maximum ratio of frequencies for which long waves wipe out short waves. Equation 3.40 can be rewritten as:

$$\begin{aligned} S_{bk,cu}(k,\theta) = & \frac{-14.2 C_{cu}}{\pi^2} N(k,\theta) \times \\ & \int_0^{r_{cu}^2 k} \int_0^{2\pi} \max\left\{\left(\sqrt{B(f',\theta')} - \sqrt{B_r}\right), 0\right\}^2 d\theta' dk' \end{aligned} \tag{3.41}$$

The wave-turbulence interaction term, S_{turb} is formulated as (Teixeira and Belcher, 2002; Ardhuin and Jenkins, 2006):

$$S_{turb}(k,\theta) = -2 C_{turb} \sigma \cos(\theta_u - \theta) k \frac{\rho_a u_\star^2}{g \rho_w} N(k,\theta) \tag{3.42}$$

where C_{turb} is a constant of order 1 which adjusts the magnitude of the interaction for ocean stratification and wave groupiness. This term can be activated by setting $C_{turb} > 0$.

Parameter	WW3 default
C_{ds}^{sat}	-2.4 $\times$ 10^{-5}
$\Delta\theta$	80°
B_r	0.00009
C_{cu}	-0.40344
S_B	2.0
C_{turb}	0.0
δ_d	0.3

Table 3.2: Default values in WAVEWATCH III for parameters affecting the wave breaking source term S_{ds}

3.2.4 JONSWAP bottom friction

The Joint North Sea Wave Project (JONSWAP) (Hasselmann *et al.*, 1973) proposed a simple, empirical linear parameterisation for bottom friction which is adapted for WAVEWATCH III as (Tolman, 1991):

$$S_{bot}(k,\theta) = 2\Gamma\frac{n-0.5}{gd}N(k,\theta) \tag{3.43}$$

Here Γ is the empirical constant for bottom roughness estimated at $-0.038m^2s^{-3}$ for swells and $-0.067m^2s^{-3}$ for wind seas (Hasselmann *et al.*, 1973). However, following the findings of van Vledder *et al.* (2010), Γ is set to $-0.038m^2s^{-3}$. The negative sign indicates dissipation of energy. g is acceleration due to gravity, d is the depth and n is the ratio of phase velocity to group velocity, calculated as:

$$n = \frac{1}{2} + \frac{kd}{\sinh 2kd} \tag{3.44}$$

The value of Γ can be adjusted by the user.

3.2.5 Depth induced wave breaking

The depth-induced wave breaking algorithms in WAVEWATCH III are based on the proposals of Battjes and Janssen (1978) for energy dissipation in shallow waters. This method defines a threshold wave height as a function of the bottom topography, and assumes that all waves in a random wave field that exceed this threshold height will break.

The bulk rate of spectral energy density dissipation for the fraction of breaking waves (δ) is estimated as:

$$\delta = 0.25Q_b f_m H_{\max}^2 \tag{3.45}$$

where Q_b is the fraction of breaking waves in the wave field, f_m is the mean frequency and $H_{\max}$ is the maximum wave height that individual wave components can reach without breaking. $H_{\max}$ can be estimated by:

$$H_{\max} = \gamma d \tag{3.46}$$

and γ is an empirically derived constant and d is the local depth.

The fraction of breaking waves Q_b is estimated using a Rayleigh distribution truncated at $H_{\max}$ (as waves for which $H > H_{\max}$ break). From the Rayleigh distribution, the following relation is obtained:

$$\frac{1-Q_b}{-\ln Q_b} = \frac{H_{\text{rms}}}{H_{\max}} \tag{3.47}$$

Equation 3.47 is solved iteratively for Q_b.

The Battjes-Janssen approach assumes that δ is distributed evenly over the entire spectrum and that the shape of the spectrum does not change. Based on this, the following formulation for

the source term S_{db} is obtained:

$$S_{db}(k,\theta) = -\alpha \frac{\delta}{E} F(k,\theta) \tag{3.48}$$

where α is a linear tunable parameter and E is the total spectral energy. Substituting from Equation 3.45 in Equation 3.48, we have

$$S_{db}(k,\theta) = -0.25\alpha Q_b f_m \frac{H_{\max}^2}{E} F(k,\theta) \tag{3.49}$$

The model defaults for the constant parameters are $\alpha = 1.0$ and $\gamma = 0.73$, which can be adjusted by the modeller.

3.3 Numerical approaches

The numerical approaches used in WAVEWATCH III for solving the wave action spectrum defined in Equation 3.3 in the wavenumber-direction grid are briefly described in this section. These are the fractional timestep approach, the Ultimate Quickest third order scheme for spatial and spectral propagation and the spatial averaging method for the alleviation of the Garden Sprinkler Effect (GSE alleviation).

3.3.1 The fractional timestep approach

Similar to other wave models, e.g. WAM, WAVEWATCH III uses the fractional timestep method to solve the equations in the wavenumber-direction domain for the generation, propagation and dissipation of waves. These fractional steps are used to split the effect of temporal variations in depth, ensuring the invariance of the (k,θ) grid as well as for spatial propagation, intra-spectral propagation and for the integration of source terms.

Depth variations in time

Depth variations in time become an important consideration for sites which experience large tidal ranges. In such cases, the model requires additional input data describing the variation of the sea level with time.

As the wavenumber spectrum is invariant with changes in depth over time, the spectrum for any timestep can be obtained by simple interpolation of the spectrum from the previous time step updating term using the dispersion equation (Equation 3.1) for the new depth. The fractional timestep used for updating the spectrum depends upon the timescale of the variations in local depth (Tolman *et al.*, 2014a).

In this study, temporal variations in depth are not considered and it is assumed that the sea level (datum) remains constant over the entire region and period simulated.

Solution of source terms

To solve for the effect of the source terms, the semi-implicit dynamic timestepping scheme is applied in WAVEWATCH III. In this scheme, integration is performed in several dynamically adjusted timesteps Δt_d until the global timestep Δt_g is bridged. The dynamic timestep depends upon the net source term S, the maximum change in the action density spectrum ΔN_m and the remaining interval in Δt_g. The n^{th} dynamic timestep Δt_g^n is calculated by the following steps (Tolman *et al.*, 2014a):

$$\Delta t_d^n = \min\left[\frac{\Delta N_m}{|S|}, \forall f\left(1+\varepsilon D\frac{\Delta N_m}{|S|}\right)^{-1}\right] \tag{3.50}$$

$$\Delta t_d^n = \max\left[\Delta t_d^n, \Delta t_{\min}\right] \tag{3.51}$$

$$\Delta t_d^n = \min\left[\Delta t_d^n, \Delta t_g - \sum_{i=1}^{n-1}\Delta t_d^i\right] \tag{3.52}$$

where $\Delta t_{\min}$ is the minimum allowable timestep which is set by the user.

The maximum change in the action density spectrum, N_m is calculated as

$$\Delta N_m(k,\theta) = \min\left[\Delta N_p(k,\theta), \Delta N_r(k,\theta)\right] \tag{3.53}$$

$$\Delta N_p(k,\theta) = X_p\frac{\alpha}{\pi}\frac{(2\pi)^4}{g^2}\frac{1}{\sigma k^3} \tag{3.54}$$

$$\Delta N_r(k,\theta) = X_r \times \max\left[N(k,\theta), N_f\right] \tag{3.55}$$

$$N_f = \max\left[\Delta N_p(k_{\max},\theta), X_f \times \max_{\forall k,\theta}\left\{N(k,\theta)\right\}\right] \tag{3.56}$$

where ΔN_p and ΔN_r are the parametric and filtered relative changes in the action density spectrum respectively. X_p, X_r and X_f are user defined constants with default values 0.15, 0.10 and 0.05 respectively. α is a Pierson-Moskowitz energy level set at 0.62×10^{-4} and $k_{\max}$ is the largest discrete wave number in the spectral grid.

The net effect of source terms are recalculated for each timestep. These dynamic timesteps are calculated separately for each wet cell in the grid, i.e. not land. As a result, when the timestep is reduced for events during which the spectrum is subject to rapid changes, such as storms, additional computational effort is only allocated to wet cells as necessary and not for the entire grid, ensuring the most economical integration of source terms without compromising on accuracy.

3.3.2 Third order accurate propagation scheme

The default scheme for the spatial and spectral propagation of wave fields in WAVEWATCH III is the third order accurate scheme with the 'ULTIMATE QUICKEST' (Leonard, 1991) total variance diminishing limiter. This scheme is applied separately in longitudinal and latitudinal directions and the direction to be treated first is alternated between the two (Tolman, 1995; Tolman *et al.*, 2014a).

In this scheme, for spatial propagation the flux in the ϕ-space (along latitude) between grid points i and $i-1$ is calculated as

$$\mathbf{F_{i-}} = \left[\dot{\phi}_b N_b\right]^n_{j,l,m} \tag{3.57}$$

$$\dot{\phi}_b = 0.5\left(\dot{\phi}_{i-1} + \dot{\phi}_i\right) \tag{3.58}$$

$$N_b = \frac{1}{2}\left[(1+C)N_{i-1} + (1-C)N_i\right] - \left(\frac{1-C^2}{6}\right)\mathbf{C_U}\Delta\phi^2 \tag{3.59}$$

$$\mathbf{C_U} = \begin{cases} (N_{i-2} - 2N_{i-1} + N_i)\Delta\phi^{-2} \text{ for } \dot{\phi}_b \geq 0 \\ (N_{i-1} - 2N_i + N_{i+1})\Delta\phi^{-2} \text{ for } \dot{\phi}_b < 0 \end{cases} \tag{3.60}$$

where $\mathbf{C_U}$ is the upstream curvature of the action density distribution and C is a CFL number with the sign indicating the direction of propagation. C can be calculated as

$$C = \frac{\dot{\phi}_b \Delta t}{\Delta\phi} \tag{3.61}$$

and for a stable solution it is essential that $|C| \leq 1$.

The ULTIMATE limiter is used with this scheme to ensure that no aphysical extreme are generated. This limiter uses the central, upstream and downstream action densities (denoted by the suffixes C, U, and D respectively) defined as

$$N_C = N_{i-1}, N_U = N_{i-2}, N_D = N_i \qquad \text{for } \dot{\phi}_b \geq 0 \tag{3.62}$$

$$N_C = N_i, N_U = N_{i+1}, N_D = N_{i-1} \qquad \text{for } \dot{\phi}_b < 0 \tag{3.63}$$

The normalised action $\tilde{N}$ is calculated as

$$\tilde{N} = \frac{N - N_U}{N_D - N_U} \tag{3.64}$$

If $0 \leq \tilde{N} \leq 1$ (monotonic state), then

$$\tilde{N_C} \leq \tilde{N_b} \leq 1, \tilde{N_b} \leq \tilde{N_C} C^{-1} \tag{3.65}$$

else,

$$\tilde{N}_b = \tilde{N}_c \tag{3.66}$$

The limiter is represented by Equations 3.65 and 3.66 for monotonic and non-monotonic initial states. If the grid point is adjacent to land points, $\dot{\phi}_b$ is obtained by averaging over adjacent sea points. The final propagation relation is formulated as

$$N^{n+1}_{i,j,l,m} = N^{n}_{i,j,l,m} + \frac{\Delta t}{\Delta \phi}\left[\mathbf{F}_{\mathbf{i}-} - \mathbf{F}_{\mathbf{i}+}\right] \tag{3.67}$$

The scheme for propagation in the λ-space (along longitude) is obtained by replacing $\dot{\phi}$ by $\dot{\lambda}$ and rotating the indices in the Equations 3.57 to 3.67 to represent the new direction.

Intraspectral propagation (wavenumber shift) of the wave field requires the solution of (Tolman *et al.*, 2014a)

$$\frac{\partial N}{\partial t} + \frac{\partial}{\partial k}\dot{k}_g N + \frac{\partial}{\partial \theta}\dot{\theta}_g N = 0 \tag{3.68}$$

$$\dot{k}_g = \frac{\partial \sigma}{\partial d}\frac{U \cdot \nabla_x d}{c_g} - k \cdot \frac{\partial U}{\partial s} \tag{3.69}$$

where $\dot{k}_g$ is the wavenumber velocity relative to the grid. The solution of these relations requires boundary conditions in the k-space but not in the θ-space.

The implementation of the ULTIMATE QUICKEST propagation scheme in the θ-space is similar to the spatial propagation scheme, however, some modifications are required for its implementation in the k-space due to variable grid spacing on account of the logarithmic increment (Leonard, 1979). The relations are modified as

$$\mathbf{F}_{\mathbf{m}-} = \left[\dot{k}_{g,b} N_b\right]^n_{i,j,l} \tag{3.70}$$

$$\dot{k}_{g,b} = 0.5\left(\dot{k}_{g,m-1} + \dot{k}_{g_m}\right) \tag{3.71}$$

$$N_b = \frac{1}{2}\left[(1+C)N_{i-1} + (1-C)N_i\right] - \frac{1-C^2}{6}\mathbf{C}_{\mathbf{U}}\Delta k^2_{m-1/2} \tag{3.72}$$

$$\mathbf{C}_{\mathbf{U}} = \begin{cases} \frac{1}{\Delta k_{m-1}}\left[\frac{N_m - N_{m-1}}{\Delta k_{m-1/2}} - \frac{N_{m-1} - N_{m-2}}{\Delta k_{m-3/2}}\right] & \text{for } \dot{k}_b \geq 0 \\ \frac{1}{\Delta k_m}\left[\frac{N_{m+1} - N_m}{\Delta k_{m+1/2}} - \frac{N_m - N_{m-1}}{\Delta k_{m-1/2}}\right] & \text{for } \dot{k}_b < 0 \end{cases} \tag{3.73}$$

$$C = \frac{\dot{k}_{g,b}\Delta t}{\Delta k_{m-1/2}} \tag{3.74}$$

3.3.3 Garden Sprinkler Effect alleviation

The Garden Sprinkler Effect (GSE) (Booij and Holthuijsen, 1987) is a phenomenon observed in wave model output when a swell field disintegrates into discrete fields due to the inadequate resolution of the spectral grid. The occurrence of this effect is higher in lower order propagation schemes (i.e. first and second order accurate schemes) and relatively lower for the third order ULTIMATE QUICKEST scheme, as used in this study. However, even this propagation scheme is not completely free from GSE and there are several options in WAVEWATCH III available for its alleviation.

For this research project the spatial averaging method provided by Tolman (2002a) is used. This method has been found to be effective in alleviating GSE and is significantly more economical in comparison with other methods (e.g. diffusion tensor method (Booij and Holthuijsen, 1987), divergent advection (Tolman, 2001)). This method only affects spatial propagation of swells and does not influence the timestep as it separate from actual swell and wave propagation.

In this approach the field of energy densities is directly averaged for a given spectral component. The area over which the field is averaged is resolved along the direction of propagation (**s**) and perpendicular to the direction (**n**). This is graphically depicted in Figure 3.1 and the extent of the area averaged along **s** and **n** is given as

$$\pm \gamma_{a,s} \Delta c_g \Delta t \mathbf{s}, \pm \gamma_{a,n} c_g \Delta\theta \Delta t \mathbf{n} \tag{3.75}$$

$\gamma_{a,s}$ and $\gamma_{a,n}$ are constants whose value (set at 1.5 by default) can be tuned by the user.

The vectors $\vec{s}$ and $\vec{n}$ in the (x,y) space can be written as

$$\vec{s} = \begin{pmatrix} 0.5\alpha_s(\gamma - \gamma_{-1}) c_g \Delta t \cos(\theta)(\Delta x)^{-1} \\ 0.5\alpha_s(\gamma - \gamma_{-1}) c_g \Delta t \sin(\theta)(\Delta y)^{-1} \end{pmatrix} \tag{3.76}$$

$$\vec{n} = \begin{pmatrix} -\alpha_n \Delta\theta c_g \Delta t \sin(\theta)(\Delta x)^{-1} \\ \alpha_n \Delta\theta c_g \Delta t \cos(\theta)(\Delta y)^{-1} \end{pmatrix} \tag{3.77}$$

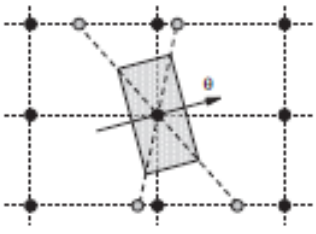

Figure 3.1: The spatial averaging process for GSE alleviation. Solid circles and dashed lines represent the spatial grid; hatched region represents the averaging area. Corner values are obtained from interpolating from adjacent computation grid points. Figure sourced from Tolman *et al.* (2014a)

The vectors $\vec{r}_1$ to $\vec{r}_4$ representing the 4 corners of the averaging area become

$$\vec{r}_1 = \vec{s} + \vec{n} \tag{3.78}$$

$$\vec{r}_2 = -\vec{s} + \vec{n} \tag{3.79}$$

$$\vec{r}_3 = -\vec{s} - \vec{n} \tag{3.80}$$

$$\vec{r}_4 = \vec{s} - \vec{n} \tag{3.81}$$

and the origin of $\vec{r}$ is the grid point at the centre of the averaging area.

The average value of the variance density spectrum is then calculated as

$$F_{avg} = \frac{1}{6}\sum_{n=1}^{4} F(\vec{r}_n) + \frac{1}{3}F(\vec{r}=0) \tag{3.82}$$

Converting to the physical (x,y) space, this can be rewritten as

$$F_{avg}(f_i,\theta_j;x_i,y_j) = \sum_{L=-1}^{1}\sum_{M=-1}^{1} a_{L,M}(x_i,y_j) \times F(f_i,\theta_j;x_{l+L},y_{m+M}) \tag{3.83}$$

To ensure the conservation of energy, it is necessary that

$$\sum_{L=-1}^{1}\sum_{M=-1}^{1} a_{L,M}(x_{l-L},y_{m-M}) = 1 \tag{3.84}$$

This process is introduced as an additional step in the fractional solution. On account of its simplicity in calculation and implementation, it is significantly more economical in terms of the computational resource required in comparison with alternative methods especially for large grids.

3.4 Model Structure

WAVEWATCH III has been designed and implemented as a subroutine which can either be called on its own or by other programs which require dynamically updated wave data. The basic elements of the wave model along with the data flow are shown in Figure 3.2.

In this section, the different modules in WAVEWATCH III are introduced along with a brief description of their functions, input required and output generated.

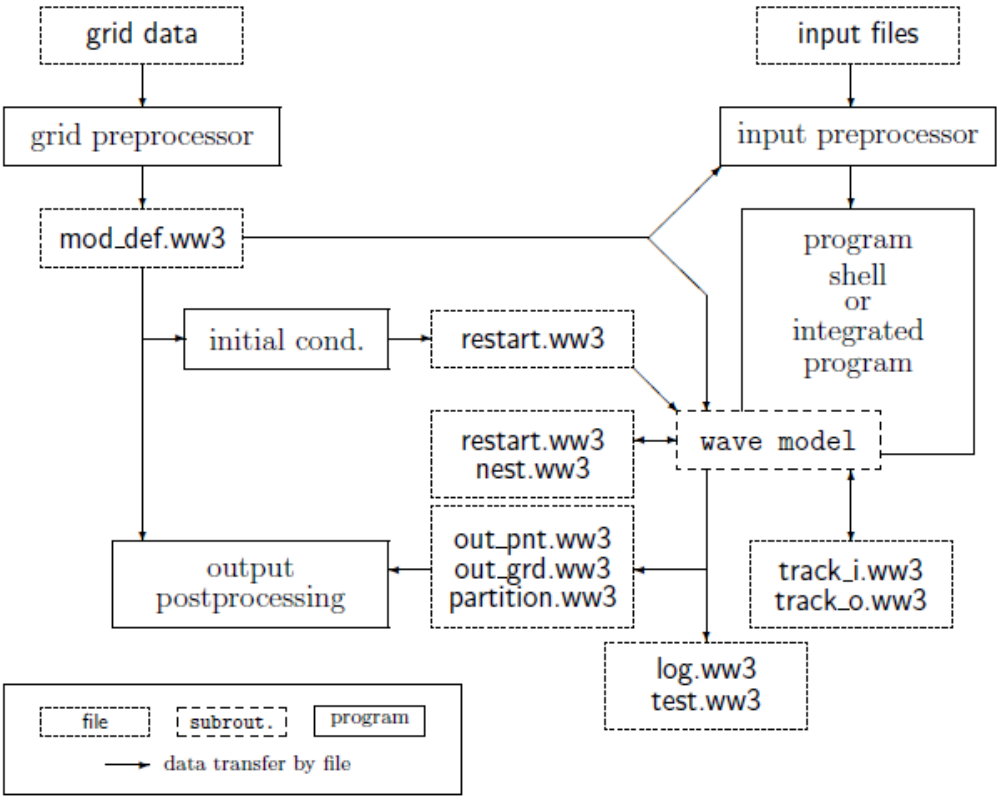

Figure 3.2: Program elements and data flow in WAVEWATCH III v4.18. Sourced from Tolman *et al.* (2014a).

3.4.1 The grid preprocessor module

The grid preprocessor module **ww3_grid** is the first auxillary program in the WAVEWATCH III suite to be executed.

The model definition file output by this program (mod_def.ww3) is essential for running the main wave model subroutine as it contains all the information regarding the spatial and spectral grids, model timesteps and user-defined values for tuning parameters for each computational grid. The mod_def files are also required by other preprocessing and postprocessing programs as can be seen from Figure 3.2. It requires as input:

- User defined model timesteps:Δt_g, Δt_{xy}, Δt_k and Δt_s;
- Spectral grid information: first frequency in the spectral grid, logarithmic increment between frequency bins and number of frequency bins and number of directional bins;
- Spatial grid information: Position of origin of the grid, grid spacing in x- and y-directions, extent of the grid (number of nodes) in x- and y- directions;
- Bathymetry data (usually read from file);

- Land-sea mask (usually read from file);
- x- and y-obstruction data (optional);
- List of nodes at which boundary data are available for input;
- List of nodes at which boundary data should be output for nesting; and
- User defined values of tuning parameters.

Individual mod_def files are created for each computational grid if a nesting scheme with multiple grids is used allowing the user to calibrate each grid with different tuning values if needed. In addition, it is sometimes used to create a dummy grid for input data (e.g. wind, ice, currents, etc.) from which only the spatial grid information is used.

3.4.2 The initial conditions module

The main wave model subroutine requires an initial description of the sea state for each computational grid for the zeroth timestep ($T = 0$). In operational models this is usually produced by previous runs of the wave model in which the last simulated timestep corresponds with the first timestep of the current run.

In the absence of a restart file containing this spectral data (e.g. for the first run of the wave model) the initial conditions program, **ww3_strt** can be used to prepare a restart file. The different types of initial conditions that can be generated are:

- Gaussian type spectrum based on user defined peak frequency and spread, cosine type direction, spatial grid information and maximum wave height;
- JONSWAP (Hasselmann *et al.*, 1980) spectrum described by user defined α;
- Fetch limited JONSWAP spectrum estimated from local wind speed and direction;
- User defined (f, θ) spectrum applied to all wet cells; and
- Start from calm conditions ($H_s = 0$).

This module requires the mod_def file for the grid and a choice of spectrum from the user based on which a restart file (restart.ww3) with initial spectral data is produced for input to the main wave model routine.

3.4.3 The input field preprocessor

The main wave model routine can be run without any input fields for wave growth and dissipation using only initial conditions or spectral data at the boundaries. However for realistic simulation of the sea states, appropriate input data is required. Input files with these fields are prepared by **ww3_prep**, the input preprocessor module.

This module preprocesses input data which may be water levels, wind fields, currents, sea ice concentrations and location specific spectral data for assimilation into the wave model. It is run separately for each input field, and produces corresponding output files (e.g. wind.ww3, current.ww3 and level.ww3). In the event that the grid to which the input fields are mapped

is not identical to the computational grid, **ww3_prep** interpolates using conservative bilinear weighting to remap the input grid.

WAVEWATCH III v.4.18 has the ability to prepare the input field files by reading data directly from self defining NetCDF files without any input from the user. However, the format of NetCDF files required by the program does not follow standard meteorological conventions, and as a result, input fields were input as ASCII files in this project.

3.4.4 The wave model subroutine

The main wave model is designed to operate as a standalone program called in a shell, or as a subroutine which can be integrated into other programs. It is available as two subroutines - **ww3_shel** for modelling one grid at a time (and for 1-way nesting) and **ww3_multi** for 2-way multi grid nesting. These subroutines are the only programs in the WAVEWATCH III modelling suite that are designed for parallelised execution.

The wave model requires the model definition files (mod_def.ww3) files for each computational grid in the nesting scheme, restart files, and input field files. If the 1-way nesting approach is followed using **ww3_shel**, then runs for the nested grids require boundary spectra files (nest.ww3) generated from coarse (outer) grid runs of the model. If the 2-way nesting scheme is used with **ww3_multi**, nest.ww3 files are not required as this is handled internally by the subroutine.

The program produces gridded output (out_grd.ww3) of wave fields for each computational grid as well as spectral point output (out_pnt.ww3) for specific locations. In addition, boundary spectra (nest.ww3), along-track output (track_o.ww3) for comparison with satellite observations and restart files (restart.ww3) are also output if activated by the user. The model also produces a log file, updated in real time, which contains a record of the timestep, the pass and any input or output operations that are performed at this step.

The gridded and point output files produced by this routine need to be processed from the native WAVEWATCH III format before they can be read by external programs.

3.4.5 The gridded output postprocessor module

This program is used to postprocess gridded output data from the native WAVEWATCH III format to a format that can be read by other programs. There are two versions of this module in the WAVEWATCH III suite - **ww3_outf** to produce ASCII files of gridded output and **ww3_ounf** to produce a selfdefining NetCDF file with gridded output data.

This module takes the raw gridded output file (out_grd.ww3) and the model definition file (mod_def.ww3) for each computational grid and produces gridded output files as desired by the user. One ASCII file is produced for each timestep with the gridded data arranged as rows and columns delimited by a single whitespace.

In this study, **ww3_ounf** was used to generate NetCDF type gridded output files. Despite not conforming to the standard meteorological conventions, these files offer several benefits such as:

- Compactness: As data are stored as multidimensional integer arrays the combined dataset requires less memory for storage than ASCII files in which data is stored as characters; and
- Ease of file handling: A single NetCDF file file contains gridded data for all the output wave parameters at all timesteps, whereas individual ASCII files are created per parameter per timestep.

Coupled with the Grid Analysis and Data Sytem (GrADS) libraries, this postprocessing module can also visualize gridded data for selected timesteps.

3.4.6 The point output postprocessor module

This module is run to generate ASCII files of point output that can be read by other programs. It takes, as input, the raw point output file containing spectral data (out_pnt.ww3) and a list of output locations and timesteps for which ASCII output is required. For each output location, the program produces an output file containing data specified by the user. These may be

- Print plots of 2-D wave spectra for each timestep;
- Tables of 1-D or 2-D spectra for each timestep;
- Tables of mean wave parameters; or
- Tables of 1-D and 2-D source spectra $S(k,\theta)$

In these output files, direction is tabulated in radians and follows the standard meteorological convention. According to this, 0° indicates waves are coming from the North and 90° indicates waves are coming from the East.

3.5 Parallelisation of WAVEWATCH III

Third generation ocean wave models such as WAVEWATCH III are computationally intensive owing to the calculation of complex physical processes (e.g. nonlinear wave interactions), the relatively small steps because of the fractional time-step approach used, and the size and resolution of the geographic and spectral domains over which wave conditions are estimated. Today, computer systems range from desktop workstations with as many as 16 processors to massively parallelised supercomputers with 10^3 to 10^4 cores. As a consequence, the unavailability of computing resource is no longer a major problem for running such models to predict the sea states in a timely manner.

For this doctoral project, access was provided to the UK's national supercomputers HECToR and its successor Archer. As the resource allocated is not unlimited, it is important to optimise

its usage, which requires an understanding of the parallel concepts implemented in WAVEWATCH III.

WAVEWATCH III uses OpenMP (OMP) and Message Passing Interface (MPI) protocols for parallelisation of the model in shared and distributed memory systems. This section briefly explains the decomposition of the domains and implementation of methods for parallel computing in WAVEWATCH III. This section draws significantly from the information presented by Tolman (2002b) and Tolman *et al.* (2014a).

3.5.1 Numerics in WAVEWATCH III

In WAVEWATCH III, for the solution of the wave action density spectrum, the physical space (x) and spectral space (k, θ) are discretised and Equation 3.2 is solved forwards in time (t) using fractional time-steps. This approach addresses spatial propagation, spectral propagation and source terms separately as Equations 3.85 to 3.87 which are solved consecutively Tolman (2002b).

$$\frac{\partial N}{\partial t} + \nabla_x \cdot c_x N = 0 \tag{3.85}$$

$$\frac{\partial N}{\partial t} + \nabla_{k,\theta} \cdot c_{k,\theta} N = 0 \tag{3.86}$$

$$\frac{\partial N}{\partial t} = \frac{S}{\sigma} \tag{3.87}$$

N is the wave action density spectrum as a function of wave number and direction, varying in space and time, ∇_x and $\nabla_{k,\theta}$ represent differential operations in the physical and spectral spaces, and c_x and $c_{k,\theta}$ are the propagation velocity vectors in physical and spectral spaces.

3.5.2 Parallel concepts implemented in WAVEWATCH III

In this section, the principal parallel concepts and paradigms are introduced and their implementation in WAVEWATCH III is explained.

- **Vectorisation:** Vectorisation is a concept used since the first supercomputers wherein computational performance gains are achieved by simultaneously performing identical calculations on large sets of data on multiple processors. Vectorisation is implemented by organising the computer program into long loop structures which can be iterated simultaneously over multiple data items.
- **Parallelisation:** Parallelisation is a form of high performance computing in which the job is spread over multiple processors and different calculations are performed simultaneously. Again, the program structure is essential in the implementation of this paradigm which requires the algorithm to be splittable into independent parts which can be solved simultaneously by multiple processors. Efficient sharing of data between processors is

critical for performance gains with this paradigm, and is achieved using either shared memory protocols (OMP) or distributed memory protocols (MPI).

- **Domain decomposition:** Data decomposition is a method of splitting the data for distribution to the available processors in which the domain modelled is divided into contiguous blocks of data points. The blocks of data are assigned to individual processors and continuity of the computations in the decomposed domain is maintained by communication of boundary data between processors.
- **Transposition:** In this parallel paradigm, all the necessary data for a computation is gathered at a single processor before the calculation is performed. This requires continuous reorganisation of data distributed over the available processors.

In third generation wave models spatial propagation given by Equation 3.85 is independent for any spectral component (k_i, θ_j) and can be solved for all spatial nodes simultaneously. As a result, this step can be easily vectorised and parallelised simultaneously (Tolman, 2002b). In the time-step approach, a maximum propagation time-step Δt_p is defined internally and if this is less than the global time-step Δt_g, propagation is performed in steps until Δt_g is reached.

In Equations 3.86 and 3.87, spectral propagation and the effects of source terms are solved simultaneously for all discrete (k_i, θ_j) spectral nodes. These computations are identical for all spatial nodes, and as a result, it is possible to parallelise the spatial domain and vectorise the spectral domain.

Domain decomposition is a parallel approach popular in numerical wave modelling, and has been implemented in WAM even before the advent of parallel computing to allow large models to be run on computers with limited memory (Gunther, 2002). However, this approach can be applied only to the spatial domain because nonlinear interactions between waves requires data from all spectral bins to be available.

Figure 3.3 presents an example of domain decomposition approaches for wave models. In the figure, nodes in the spatial domain are divided among 4 processors numbered 0-3, grey areas denote land and the hatched areas at the boundaries of subdomains identify 'halo' regions where communication of boundary data between processors is necessary.

Figure 3.3(a) shows the traditional approach, as used in WAM, where decomposition is achieved by dividing the domain of interest into rectangular blocks containing an approximately equal number of nodes and assigning one block per available processor.

Although such an approach is simple to implement and allows the solution of Equations 3.85 to 3.87 at the same node, it is susceptible to load imbalances. For example, Equation 3.87 is solved using a dynamically adjust time-step for the source terms. These time-steps can be quite small during events associated with rapid spectral changes such as storms significantly increasing the computational effort required. As these are usually locally occurring events covering a small number of spatially contiguous nodes, it is likely that a very small number of processors will

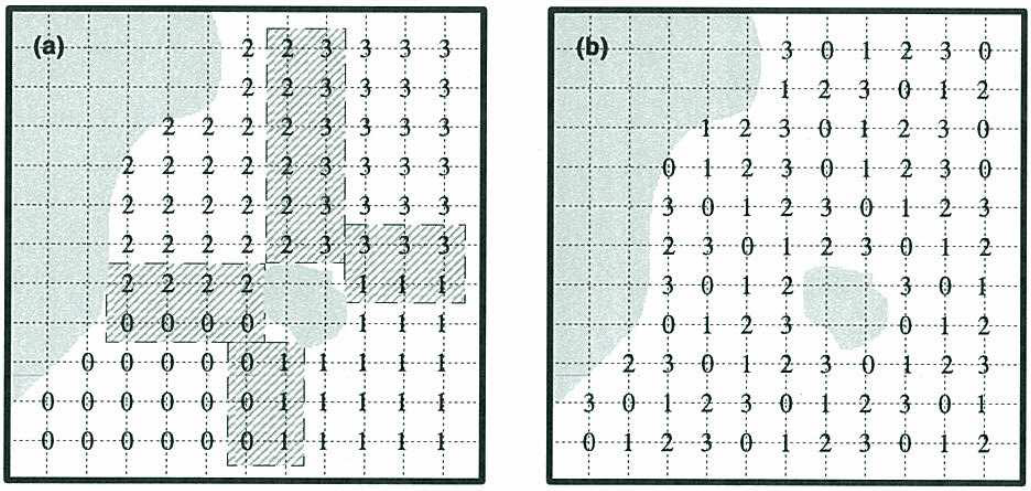

Figure 3.3: Domain decomposition approaches for wave models demonstrated for a 4-processor system. (a) Traditional approach and (b) Spatially non-coherent approach. Figure sourced from Tolman (2002b)

be affected by this, leading to an imbalance of the load.

To prevent such load imbalances, WAVEWATCH III implements an alternative approach for decomposition by which spatial nodes are assigned to available processors in a non-coherent way. This is demonstrated in Figure 3.3(b), where the grid counter m starts in the lower left corner of the grid and works upwards along sea points line-by-line. The allocation of grid-points to processors is done by

$$n = (m-1)\%N \tag{3.88}$$

where grid-point identified by the counter m is allocated to processor n for a total number of available processors N. The '%' operator is used to represent the remainder after division. This form of decomposition ensures that the number of grid points as well as the computational load is approximately equally spread over all the available processors.

However, such a spatially non-coherent method of domain decomposition makes it impossible to solve Equation 3.85 for spatial propagation at each processor. The solution of Equation 3.85 requires that the data for all spatial nodes for a spectral bin (k_i, θ_j) be gathered at a single processor before calculation. Similarly, the solution of Equations 3.86 and 3.87 requires all the spectral data for any spatial node to be gathered at a single processor.

Consequently, data at each processor is transposed twice during each computational cycle - once to solve for spatial propagation and once to solve for spectral propagation and source term effects. The developers of WAVEWATCH III observed that despite the large increase in data communicated between processors, the data transpose method yielded greater overall performance gains than the traditional domain decomposition method, and the former approach

was implemented in the wave model.

Because of the parallelisation schemes implemented in WAVEWATCH III, the limiting number of processors over which the model can be executed is the smaller of the number of spectral bins or the number of wet cells in the spatial computational grid.

3.5.3 Execution of WAVEWATCH III in parallel for this study

Increasing the number of processors over which the job is spread results in performance gains up to a point. Adding processors beyond this yields negligible gains and sometimes can even cause in an increase in the overall runtime.

According to Amadahl's law a decrease in runtime is achieved up to a theoretical limit which is determined by the parallel fraction of the code. At this limit, the serial part of the code, preprocessing required to split the job for each processor and the speed of and bandwidth for communication between processors and memory begin to dominate the overall runtime and no further performance gains can be achieved.

Parallel efficiency (ζ) is a commonly used metric for assessing the average performance gain for a given number of parallel tasks, and can be defined as (Tolman, 2002b)

$$\zeta = \frac{t_1}{t_N N} \tag{3.89}$$

where t_1 is the runtime using a single processor and t_N is the runtime using N processors. $\zeta = 1$ would indicate a perfectly parallel behaviour of the code.

The results of benchmarking tests performed by the developers of WAVEWATCH III in terms of runtimes and efficiency of parallel performance as a function of the logarithm of the number of processors are plotted in Figures 3.4 and 3.5 (Tolman, 2002b) respectively.

It can be observed form Figure 3.4 that the run-time decreases approximately linearly according to the natural logarithm of the number of processors up to $N \approx 80$. These diminishing returns are in keeping with Amadahl's law. From Figure 3.5 it can be noted that the parallel efficiency is greater than 0.9 for $N \leq 3$ and between 0.8 and 0.85 up to N between 3 and ≈ 30 beyond which there is a systematic decrease in ζ.

From these, it may be inferred that the optimum number of processors over which the wave model subroutine is spread, N, is less than 80. It would appear that keeping $N \leq 30$ would optimise for parallel efficiency provided that the runtime ($t_{N \leq 30}$) is acceptable.

In addition to these, restrictions imposed by the queuing systems of high performance computers need to be considered when allocating processors to the model. These include:

- A hard runtime limit of 12 hours, after which any running job will be forcibly terminated; and

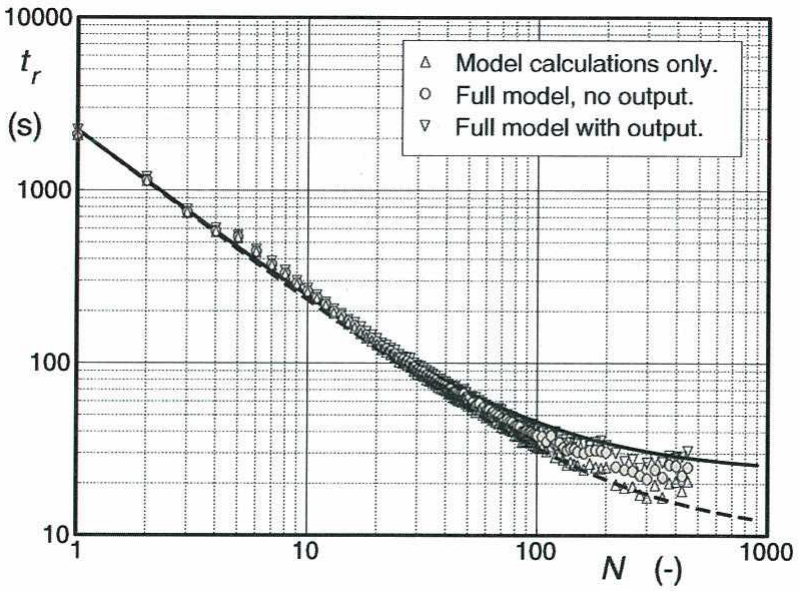

Figure 3.4: Run times t_r as a function of the number of processors N. Figure sourced from Tolman (2002b)

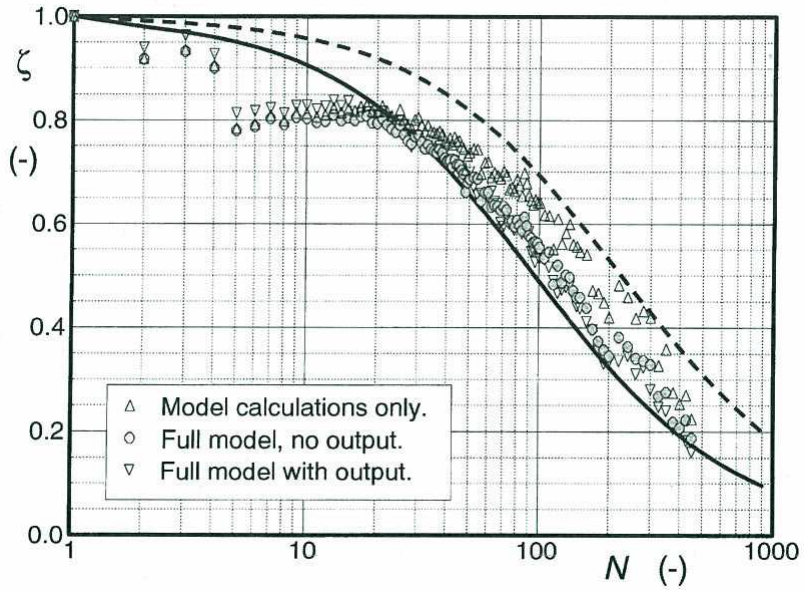

Figure 3.5: Parallel efficiency ζ according to Equation 3.89. Figure sourced from Tolman (2002b)

- Accounting of resources used is based on computing nodes utilised.

These restrictions mean that the total runtime of the job cannot exceed 12 hours, and the number of allocated processors should be an integer multiple of the number of cores per processing node (32 for HECToR and 24 for Archer).

After benchmarking trials with the model set up for this study, it was found that the running the model over a single node (utilising 32 processors on HECToR and 24 processors on Archer) yielded the lowest cost of simulation. Approximate runtimes for simulating the wave climate for one year were $\approx$ 8 hours on HECToR and $\approx$ 5 hours on Archer.

Running the model on a single computing node involves faster communication between processors and memory as shared memory protocols (OMP) are used, wherein the hardware as well as the controlling software enables higher speed data transfer than for distributed memory protocols (MPI).

3.6 Chapter summary

- The third generation spectral wave model, WAVEWATCH III, chosen for the doctoral research is introduced.
- The partial differential equations governing the generation and propagation of ocean waves in terms of frequency and direction varying in space and time, as solved by WAVEWATCH III are described.
- The source and sink terms governing the physical processes relevant to the study are described in detail. These include wind-wave interactions, nonlinear wave interactions, deep-water wave breaking, bottom friction and depth-induced wave breaking.
- Numerical approaches such as fractional time-step approach and propagation scheme employed for the solution of the wave spectrum are introduced. The garden sprinkler alleviation method selected for this research is also described.
- The program structure of WAVEWATCH III is introduced with an explanation of the functions and interdependence of each module within the model.
- Parallel concepts implemented in the model are explained and the optimum level of parallelisation for this research is obtained.

Chapter 4

Numerical modelling

In the absence of a 'perfect' dataset of instrument measurements for the long-term analysis of the wave climate, numerical simulation of wave conditions using a suitable spectral wave model is commonly used to generate data. In this study, the 3rd generation spectral wave model WAVEWATCH III is used for simulating historical and projected wave data for analysis.

In this chapter, the different types and sources of input data used in the research are introduced, followed by a detailed explanation on the model setup. Critical analyses such as selection of forcing wind, estimating uncertainty due to tuning parameters, sensitivity analysis, parameter optimisation and validation of the calibrated model are presented in the following sections.

To the best of the author's knowledge, this is the first documented instance of a structured approach with designed experiments being used for the analysis of uncertainty in and sensitivity of a wave model. In addition to also being the first instance of response surface modelling (RSM) being used for parameter optimisation, the approach is novel in that two response variables are simultaneously optimised using non-linear programming.

4.1 Description of Data Sources

To set up a numerical wave model for the desired region and to simulate wave conditions, the model requires several types of input. Of these, bathymetry and wind forcing are considered essential. This section provides a brief description of the input datasets used to set up, force and validate WAVEWATCH III for simulating the wave climate in the North Atlantic Ocean and North Sea.

4.1.1 Bathymetry data from ETOPO-1

ETOPO1 is a global relief model developed by the US National Geophysical Data Center (NGDC) which integrates land topography and ocean bathymetry into a single dataset. It extends from -90° to $+90^\circ$ latitude and -180° to $+180^\circ$ longitude at a horizontal resolution of 1 arc-minute, which is approximately 1.85km along longitude $\times$ 1km along latitude ($\pm$ 0.25km

from south to north) in the region of interest. It is available for public access online (National Geophysical Data Center, 2008).

ETOPO1 was developed as a successor to the ETOPO2v2 relief model, improving on its accuracy and resolution in order to support tsunami modelling and forecasting as well as ocean circulation modelling (Amante and Eakins, 2009). It incorporates bathymetric data from several sources including hydrographic survey soundings from the Japan Oceanographic Data Center (JODC), multibeam swath sonar bathymetric surveys from NGDC and the Mediterranean Science Comission (CIESM) and digitised depth contours of the Caspian Sea. Topography data was obtained fron NGDC, NASA and the National Snow and Ice Data Center.

The ETOPO1 dataset was used to prepare the bathymetry input file for WAVEWATCH III by interpolating the 1-arc-minute depth data to the resolution of the computational grids.

4.1.2 The 20th Century Reanalysis v2 (20CR)

The 20th Century Reanalysis (20CR) (Compo *et al.*, 2011) is a global atmospheric circulation dataset provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA. The dataset spans from January 1871 to December 2010, and contains several data products gridded as a T62 gaussian grid and sampled at 3-hourly intervals.

It is produced using the NCEP's Climate Forecast System (CFS), assimilating only surface pressure reports, observations of sea-surface temperature and sea-ice distributions as boundary conditions and using an Ensemble Kalman Filter. The feasibility of such an approach has been discussed in detail by Whitaker *et al.* (2004) and Compo *et al.* (2006) previously. The dataset has been validated against radiosonde data, and intercomparisons of the latter half of the 20th century with other reanalysis datasets show good agreement especially in the Northern Hemisphere (Compo *et al.*, 2011).

From this dataset, 'u' and 'v' components of surface winds at 10m height (in ms^{-1}) are of interest to force the numerical wave model in order to produce a hindcast of the wave climate. The data is remapped to a regular grid with a spatial resolution of $1.875^\circ \times 1.875^\circ$. This resolution is nearest to the original T62 form is likely to introduce the least bias during remapping.

4.1.3 The NCEP Climate Forecast Reanalysis (CFSR)

The National Centers for Environment Prediction (NCEP) Climate Forecast Reanalysis (CFSR) (Saha *et al.*, 2010) is a global, high-resolution dataset of gridded atmospheric, oceanic and land surface data products extending forwards in near-real-time from January 1979.

The atmospheric component of the model is run over a T382 Gaussian grid which approximates to a resolution of 0.3125°x0.3125° (or $\approx$ 38km) and 64 vertical levels. Apart from the

resolution and timestep, improvements in this reanalysis include a sophisticated Global Data Assimilation System (GDAS) and the use of atmospheric carbon dioxide (CO_2) concentrations as a function of time.

The inclusion of atmospheric greenhouse gas concentrations in the reanalysis makes it especially useful for climate and climate change research. It is anticipated that local wind patterns will be better represented by this dataset which will lead to higher accuracy simulations when modelling wave energy sites. This reanalysis dataset is recommended by the developers of WAVEWATCH III and is used in the operational wave model at NOAA/NCEP (Tolman *et al.*, 2014b).

4.1.4 The ERA-Interim Reanalysis (ERA-I)

The ERA-Interim (ERA-I) reanalysis (Dee *et al.*, 2011) is the third and most recent global reanalysis dataset produced by the European Centre for Medium-Range Weather Forecasts (ECMWF). It is produced using the ECMWF Integrated Forecast System (IFS) and gridded data products are available from January 1979 extending forwards in near-real-time. The gridded datasets, sampled at 3-hourly or 6-hourly intervals, include parameters describing the troposphere, stratosphere, atmospheric fluxes and other derived fields. It can be accessed online at full spatial and temporal resolution (European Centre for Medium-Range Weather Forecasts, 2012) and a comprehensive description of the data can be obtained from Berrisford *et al.* (2009).

The wave model component of the IFS is based on the Wave Action Model (WAM) (Komen *et al.*, 1994), with several enhancements including a reformulation of the dissipation source term (Bidlot *et al.*, 2007). The wave model is run at a horizontal resolution of 110km and the wave spectra are discretised over 24 directions and 30 frequencies.

The dataset has been shown to be capable of reproducing weather conditions to a high degree of accuracy in comparison to other similar reanalyses (Bidlot *et al.*, 2002). The high accuracy of the reanalysis can be attributed, to a great extent, to the superior data assimilation system within the IFS. The atmospheric component incorporates scatterometer ocean surface wind data from the satellite missions ERS-I, ERS-II and QuikSCAT, and the wave model incorporates reprocessed altimeter wave height data from ERS-I and ERS-II, as well as near-real-time data from the ENVISAT, JASON-1 and JASON-2 missions (Dee *et al.*, 2011).

For this study, the datasets of interest are gridded surface winds at 10m height (in terms of 'u' and 'v' components) for forcing WAVEWATCH III, and the ocean wave parameters significant wave height (H_s), Mean wave period (T_{m02}) and Mean wave direction (D_m) for a comparative evaluation of the model output.

4.1.5 Buoy data

Data buoys, considered to an excellent source of wave observations, provide hourly measurements of significant wave height (H_s) and mean zero crossing perod (T_z or T_{m02}). Some datasets contain the peak period (T_p) instead and more recent buoys, such as Waveriders, also provide directional data in the form of mean and peak wave directions. In some cases, the frequency-direction spectra of waves measured by the buoys can also be obtained (e.g. from Fisheries and Oceans Canada (2013)).

Buoy data used in this research have been obtained from several sources in order to include deep waters as well as coastal locations in the North Sea as well as the North Atlantic Ocean. Data from the North American buoys have been obtained from the US National Data Buoy Center (NDBC) (National Data Buoy Center, 2013) and Fisheries and Oceans, Canada (Fisheries and Oceans Canada, 2013). Wave observations in European waters have been obtained from buoys maintained by the UK's Wavenet project (Department for Environment Food and Rural Affairs, 2013) and the Marine Institute, Ireland (Marine Institute, 2013).

In this study, buoy data has been used mainly for calibrating and validating the wave model set up for the region of interest.

4.1.6 Projection Data

During the preparation of the IPCC's Fourth Assessment Report (AR4) (IPCC, 2007), an ensemble of climate data was produced by running global climate models (GCMs) developed and operated by various institutions across the world in a coordinated manner for defined experiments for the assessment of climate change and its impacts. These datasets have since been collated and compared with one another under the third Coupled Model Intercomparison Project (CMIP3) (Meehl *et al.*, 2007). One of the objectives of this project was to provide access to these archived datasets to climate change researchers.

Synoptic data in the form of daily, monthly and annual means of various climatological parameters are freely available from this project for the different experiments from all the participating institutions. In addition, for a few GCMs, e.g. ECHAM5 (Roeckner *et al.*, 2003) and CNRM-CM3 (Salas-Mélia *et al.*, 2005), data sampled at sub-daily rates are also available for certain experiments.

In this study 10m surface winds sampled at 6-hourly intervals in the form of 'u' and 'v' components generated from simulations of ECHAM5 for the high (A2), medium (A1B) and low (B1) emissions scenarios were identified for modelling wave conditions in the 21st century with WAVEWATCH III. This model was chosen for its comparatively finer spatial resolution (T62 Gaussian grid) and the reported superior ability in simulating the climate (Reichler and Kim, 2008).

During the period of this study, the IPCC's Fifth Assessment Report (IPCC, 2013) on climate change was being prepared. Although data from the climate models involved in the process were being collated, evaluated and compared under the CMIP5 project only synoptic data, e.g. monthly means, were available in the public domain. As sub-daily surface wind data from these model runs were not available for use, data from AR4 were used in this study.

4.2 Wave model setup

4.2.1 Computational grid set up

The region of interest for this doctoral study is the North East Atlantic Ocean and the North Sea. It is known that waves have the ability to travel great distances and, as a result, the wave climate in the western part of the domain is influenced to a great extent by wind seas and swells generated elsewhere in the Atlantic Ocean. For this reason, it is necessary to nest the region of interest in a larger grid covering the North Atlantic Ocean. The wave fields entering the inner grid from the larger grid are introduced in the form of spectral boundary data.

The computational grid in WAVEWATCH III v4.18 may be of the regular (or structured), unstructured or spherical types. For computational efficiency, ease of grid generation and simplicity of nesting, the computational grid was set up in the form of regular structured meshes.

The grid preprocessor module, **ww3_grid**, requires the following input for defining the computational grid for WAVEWATCH III:

1. X and y limits (in metres or degrees);
2. Grid spacing in x and y directions (in metres or degrees);
3. Bathymetry data at each node in the computational mesh;
4. Land-sea mask for the domain;
5. Optional x and y obstruction data for unresolved islands in the grid; and
6. Co-ordinates of nodes at boundaries for data exchange between grids.

ASCII files containing bathymetric depth data, the land-sea mask and x and y obstruction data can be easily prepared using 'gridgen', a software toolbox developed at NOAA. The output files from this toolbox are in a format that can be read by the grid preprocessor, **ww3_grid**, without any modification.

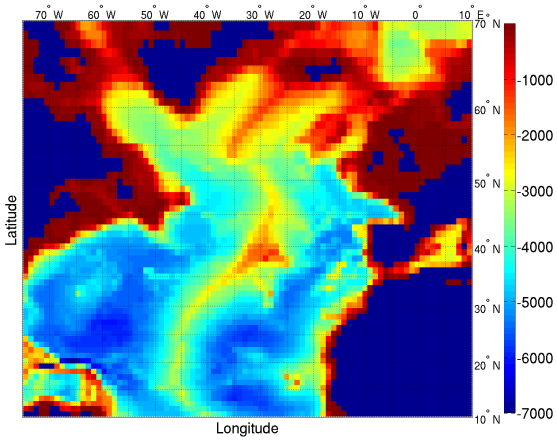

Figure 4.1: Bathymetry (m) for the North Atlantic grid (grd1). Negative values indicate depth below the datum or land

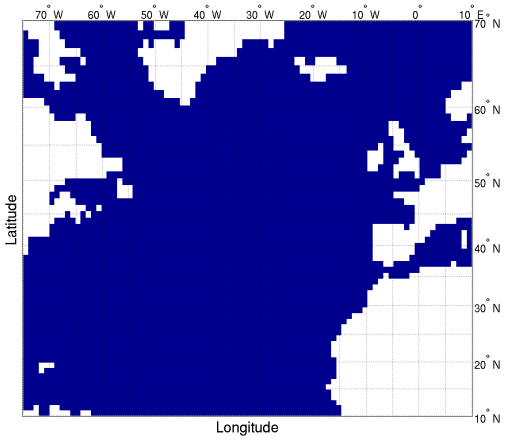

Figure 4.2: Land-sea mask for the North Atlantic grid (grd1)

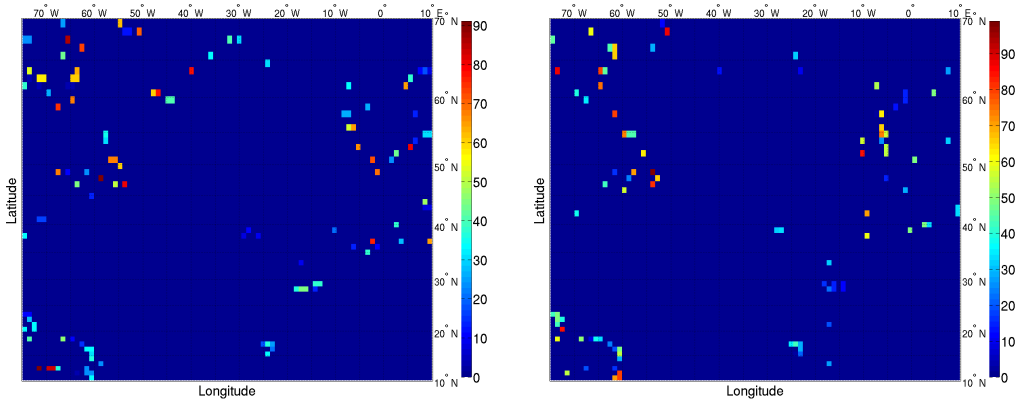

Figure 4.3: X (left) and y (right) obstruction data (%) for the North Atlantic grid (grd1)

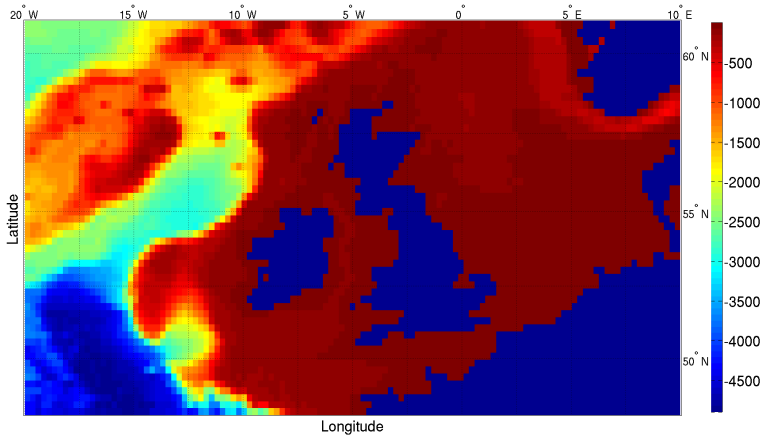

Figure 4.4: Bathymetry (m) for the North East Atlantic and North Sea grid (grd2). Negative values indicate depth below the datum or land

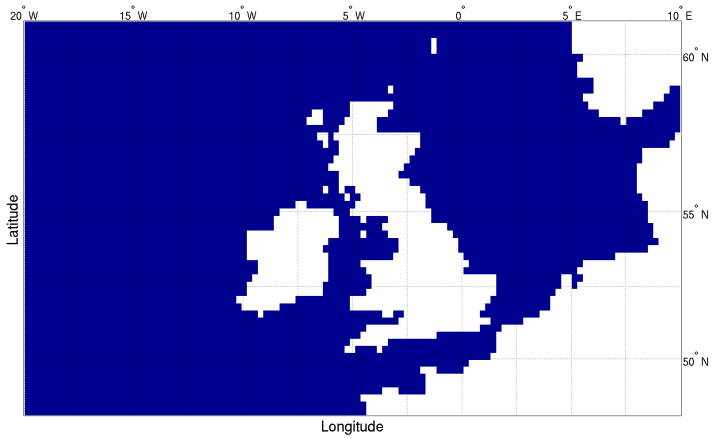

Figure 4.5: Land-sea mask for the North East Atlantic and North Sea grid (grd2)

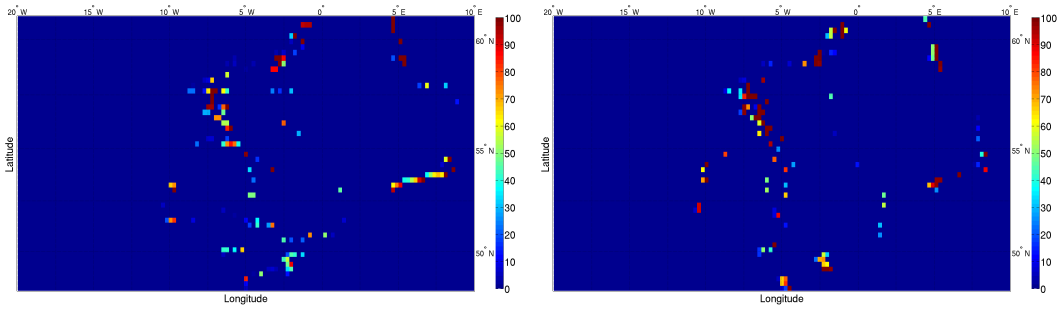

Figure 4.6: X (left) and y (right) obstruction data (%) for the North East Atlantic and North Sea grid (grd2)

The North Atlantic

The North Atlantic grid (hereafter referred to as **grd1**) is the outer grid for the region bounded by the latitudes 10°N and 70°N and the longitudes 75°W and 10°E. As the purpose of this grid is to generate boundary data for the nested grid, it is not necessary that it be of very high resolution and a spatial resolution of 1° x 1° was deemed adequate.

The depth data, land-sea mask and the x and y obstruction are presented visually in Figures 4.1 to 4.3.

The NE Atlantic and North Sea

The North East Atlantic and the North Sea grid (hereafter referred to as **grd2**) covers the area of interest for this research. The domain is bounded by the latitudes 48°N and 61°N and the longitudes 20°W and 10°E. Gridded output from this grid is intended for various analyes as part of the research. This mesh was prepared at 0.25° x 0.25° to produce as high resolution data as possible while keeping the computational resource required within reasonable limits.

Nodes at the southern, western and northern boundaries of this grid at 1° intervals were flagged as active boundary points for the exchange of data between grids. To prevent 'hotspots' at the boundaries, the wave model interpolates the spectrum between active boundary nodes.

The depth data, land-sea mask and the x and y obstruction are presented visually in Figures 4.4 to 4.6.

4.2.2 Spectrum setup

For a spectral wave model to accurately simulate the sea state, the computational frequency-direction spectrum needs to be adequately resolved. A high resolution spectrum will result in parameter estimates of greater accuracy as a smooth spectrum can be calculated. In addition, problems such as the Garden Sprinkler Effect can be avoided if the resolution is high enough. This, however, comes at an increased computational cost on account of the larger number of frequency-direction combinations at which the partial differential equations need to be solved.

The resolution of the spectrum depends upon the space that needs to be simulated. In the case of direction, θ, it is known that this space is bounded 0° and 360°. In the case of wind-generated seas, the wave period is usually between 2s and 22s, thus defining the limits of the frequencies integrated over.

It is the modern practice in operational models to resolve the frequency-direction space into 36 frequencies and 36 directions (Tolman *et al.*, 2014b) such that $\Delta f = 1.1f$ and $\Delta\theta = 10°$ where Δf is the increment between frequency bins and $\Delta\theta$ is the increment between directional bins. An example of such a spectrum as solved by a spectral wave model is shown in Figure 4.7.

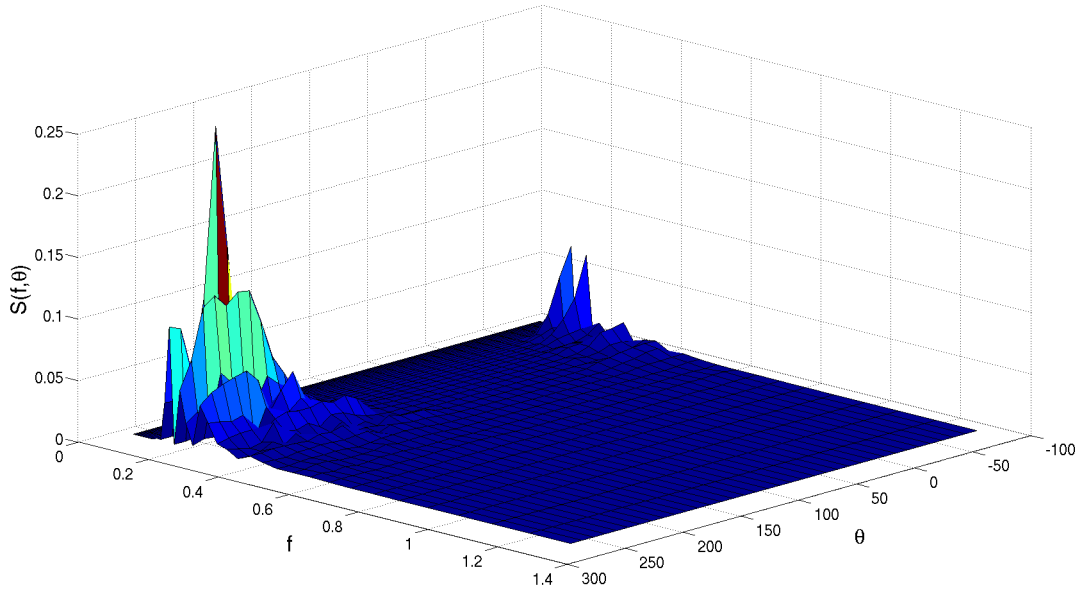

Figure 4.7: A sample wave spectrum resolved over 36 frequencies and 36 directions as solved by a numerical wave model. Data sourced from Lavidas (2014).

Alternatively, the frequency-direction space can be resolved into 25 frequencies and 24 directions, keeping $\Delta f = 1.1f$ and setting $\Delta\theta = 15^{\circ}$. The first frequency in this scheme is 0.0418Hz. By doing so the range of frequency bins still covers the wave periods of wind generated seas and swells. In comparison, the computational requirements with this scheme are approximately 55% lower.

Taking into account the temporal span of the data to be generated in this study, the latter spectral resolution was selected at the cost of slightly lower accuracy of the computed spectrum.

4.2.3 Selection of model time steps

WAVEWATCH III requires the user to define 4 time steps during the setup of the computational grid using the grid preprocesser module, **ww3_grid**. These are the global time step (Δt_g), the spatial propagation time step (Δt_{xy}), the intraspectral propagation time step (Δt_k) and the source term time step (Δt_s) in seconds. The selection of appropriate values of these time steps is critical as it greatly affects the accuracy of the model's output and the computational economy. The recommended procedure for selecting these time steps is described below:

The spatial propagation time step (Δt_{xy})

This time step, as the name suggests, defines the maximum time step for spatial propagation and is the first to be determined by the user. It is essential for accurate simulations that Δt_{xy} satisfies the Courant-Friedrichs-Levy (CFL) criterion, i.e. the speed of the fastest modelled waves must be less than or equal to the ratio of grid spacing and time step. According to this definition, Δt_{xy} is dependent on the grid spacing and must be calculated for each grid individually. This can be done using (Spindler and Tolman, 2008):

$$\Delta t_{xy} = \frac{\Delta x}{C_g} \tag{4.1}$$

$$= \frac{40 \times 10^6 \times \Delta x \times \cos(Lat_{max}) \times 4\pi \times f}{360 \times 1.15 \times g} \tag{4.2}$$

$$\approx 123766 \times \Delta x \times \cos(Lat_{max}) \times f \tag{4.3}$$

where Δx is the longitudinal grid spacing in degrees, Lat_{max} is the maximum latitude in degrees, f is the first frequency (0.0418 Hz) and g is the acceleration due to gravity. It is usual practice to select an integer close to and preferably less than the solution of Equation 4.3.

The global time step (Δt_g)

This is the interval at which the entire solution is propagated and is the first time step to be input into the model. Input data for wind and currents are interpolated and assimilated into the solution at this sampling rate. In addition, for a nested scheme of grids, this is the interval at which the grids exchange information at the boundaries.

The selected value of the time step should take into consideration the time intervals at which forcing data is sampled as well as the global time steps of nested grids. The value of this time step may be set at 2 to 3 times the value of Δt_{xy} and it is advisable to select these for the individual grids such that they are integer multiples to ensure regular intersection of the time steps for the grids.

The intraspectral propagation time step (Δt_k)

This is the third time step given as input to the model and is sometimes referred to as the maximum refraction time step or the wavenumber shift time step.

For deep-water grids, where the effect of bottom topography is minimal, this time step is usually chosen to be equal to the global time step to maximise the computational economy of the model. However, for shallow waters and regions with strong refraction, selecting $\Delta t_k = \Delta t_g$ can result in undulations of wave parameters. To avoid this, Δt_k is usually selected as being an even integer fraction of Δt_g for coastal and shallow water grids (Tolman *et al.*, 2014a).

The source term time step (Δt_s)

This is the final time step that needs to be selected by the user. This time step defines the minimum value of the internal time step during dynamical source term integration. This time step typically takes a value between 5 and 15s depending on the resolution of the computational grid.

It should be noted that selecting a large value of Δt_s will not necessarily improve model economy. This is because a larger minimum time step can cause increased spectral noise during source term integration resulting in a larger value of the average time step, thus increasing the overall computational cost.

Following the procedure described above, appropriate time steps were calculated for the North Atlantic grid (outer grid) and the grid for the North East Atlantic and North Sea (inner grid). These are tabulated in Table 4.1 in the order of input to the model.

Time step	North Atlantic grid	NE Atlantic and North Sea grid
Δt_g	3600	1800
Δt_{xy}	1200	600
Δt_k	3600	900
Δt_s	15	10

Table 4.1: Model time steps (s) calculated for the outer and inner computational grids for WAVEWATCH III

4.2.4 Selection of output parameters

Gridded output

For a computational grid with a large number of nodes, e.g. 121x53 for **grd2** in this study, it is not feasible to output and store the frequency-direction spectrum for each point at each timestep, especially if the model is run to simulate long periods. As a compromise, at the end of each computational step, WAVEWATCH III processes the solved wave spectrum and calculates wave parameters selected by the modeller which are then written to file.

The following wave parameters were identified for output and storage from the wave model:

- Significant wave height, H_s or H_{m0};
- Mean ave length, L;
- Wave period from the first spectral moment, T_{m01};
- Wave period from the second spectral moment, T_{m02};
- Energy averaged wave period, T_e or $T_{m-1,0}$;
- Peak frequency, f_p;
- Mean wave direction, D_m;

- Mean directional spread, σ_θ; and
- Peak wave direction, D_p.

These wave parameters were output and stored for the North Atlantic region (grd1) as well as the North East Atlantic and North Sea (grd2). However, only data for the latter grid were processed from the raw format for the analyses presented in subsequent chapters.

Although only the wave parameters H_s, T_e and, to some extent, D_m are studied in detail in this body of research, other wave parameters were also produced and stored as they may be useful for a variety of analyses and other research projects in the future.

Point output

Despite the large storage requirements of writing the f-θ spectrum to file, it is desirable to save this data for a few select locations in the grid. These may correspond to sites with measuring instruments (e.g. buoys and ADCPs), boundaries for nesting grids and other wave models (e.g. SWAN) or simply a location of interest (e.g. a prospective wave energy site). These spectral output files can be processed with **ww3_outp**, the point output post-processing module of WAVEWATCH III, to produce tabular data of standard wave parameters for the desired locations.

For this research, f-θ spectral data was output and stored for the sites listed in Table 4.2. These locations correspond to buoys operated by various oceanographic agencies and point output is stored for comparison, calibration, validation and other analyses. Saving spectral data to disk for providing boundary conditions from the outer grid to the inner grid is not necessary with the multi-grid version of the wave model, **ww3_multi** as this is taken care of internally.

4.3 Selection of forcing wind

As described in section 4.1, there are several available reanalysis datasets which can be used to drive the wave model. These datasets differ not only in their spatial resolution, sampling frequency and the duration over which they extend, but also in the method of preparation. On account of these differences, the quality of output from the wave model is also likely to differ when different datasets are used.

Previous studies (e.g. Zhao *et al.* (2003); Ardhuin *et al.* (2007)) have demonstrated that the spatial and temporal resolution of the driving datasets have a direct influence on the accuracy of output wave parameters from the model. Increasing the resolution has the effect of decreasing the error between modelled and measured parameters (Cavaleri and Bertotti, 2006).

Intuitively, it would seem that the model forced with CFSR data is likely to be the most accurate. However, this is not a certainty, and despite the relatively poorer resolutions of

Location/Buoy	Depth (m)	Latitude (°N)	Longitude (°E)	Description
C44138	1500	44.25	-53.62	Canadian buoy
C44140	1500	42.87	-51.47	Canadian buoy
C44141	4500	43.00	-58.00	Canadian buoy
C44150	28	42.51	-64.02	Canadian buoy
M3	155	51.217	-10.551	Irish buoy
M4	72	54.998	-9.992	Irish buoy
M6	3000	53.075	-15.881	Irish buoy
41041	3485	14.329	-46.082	NDBC buoy
41044	4536	21.562	-58.644	NDBC buoy
41048	5261	31.95	-69.497	NDBC buoy
41049	5261	27.537	-62.945	NDBC buoy
Blackstones	97	56.062	-7.057	CEFAS buoy
Firth of Forth	65	56.188	-2.504	CEFAS buoy
Hastings	43	50.747	0.755	CEFAS buoy
Moray Firth	54	57.966	-3.333	CEFAS buoy
Poole Bay	28	50.634	-1.719	CEFAS buoy
Scarweather	35	51.433	-3.933	CEFAS buoy
West Gabbard	33	51.983	2.082	CEFAS buoy
West of Hebrides	99	57.292	-7.915	CEFAS buoy
Dowsing	22	53.531	1.054	CEFAS buoy
Tyne/Tees	63	54.919	-0.749	CEFAS buoy

Table 4.2: List of point output locations

the other datasets, they may yield more accurate simulations. Moreover, for climate studies, datasets which extend over long periods, such as the 20CR dataset which extends from 1871 to 2010, are invaluable as centennial scale analyses may be performed.

Thus, the selection process must take into account the quality of output from the wave model as well as the span of data that can be produced, striking a compromise between the two. To do this in a systematic manner the methodology presented as a flowchart in Figure 4.8 was followed. The wave model was forced with wind data from the 20CR, ERA-I and CFSR reanalyses for March 2009. The output wave parameters were compared against buoy measurements, and the reanalysis datasets were evaluated.

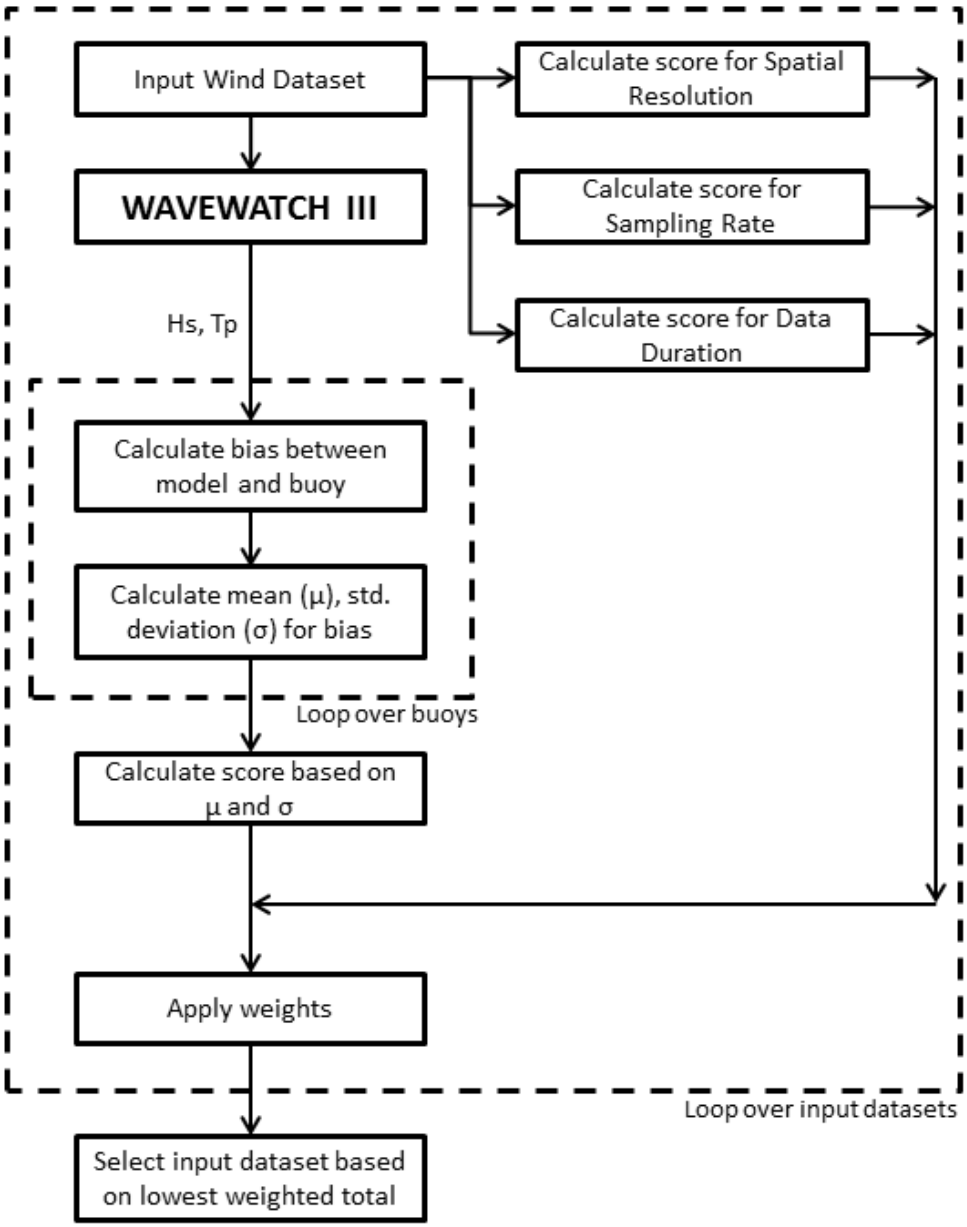

Figure 4.8: Flowchart describing the input dataset selection process

4.3.1 Identification of decision criteria

It is necessary to identify criteria for comparing the performance of the different forcing datasets such that they take into consideration all the aspects relevant to the study. The quality of the wave output is the single most important consideration, followed by other factors like the time span. Taking this into account, the following criteria were identified:

- Input grid resolution
- Input data sampling rate
- Temporal length of data
- Mean bias for H_s , T_p
- Standard deviation for H_s , T_p

The mean bias and standard deviation of bias calculated for H_s and T_p are expected to yield quantitative measures for comparing the quality of wave output from the different forcing datasets. The mean bias is used as an indicator of the accuracy of the output and the standard deviation is used as an indicator of its precision.

In addition to these measures, the spatial and temporal resolutions are also identified for assessing the quality of the forcing datasets. These, however, are minor factors as their effect is likely to be manifested in the quality of wave parameters simulated by the model. The common feature of these factors is that a smaller absolute or numerical value is indicative of superior performance.

For long-term analyses of the climate, datasets of longer spans are more desirable. For this reason, the temporal extent of the datasets was also included in the comparison.

Once these criteria are identified, they need to be weighted in a manner that reflects the relative importance to the study. The mean bias and standard deviation of the wave output was given weights of 30% each, thus emphasizing the importance of the quality of output. The span of data was weighted at 25% and the remaining 15% was equally divided between spatial and temporal resolutions.

4.3.2 Comparison and evaluation of forcing datasets

Following the procedure described in Figure 4.8, the wave model was run with the different forcing datasets and the mean (μ) and standard deviation (σ) of bias were calculated for significant wave height and peak wave period for different buoy locations. These statistics were normalised across the buoy locations to obtain the dimensionless normalised mean ($\widehat{\mu}$) and normalised standard deviation ($\widehat{\sigma}$). As these parameters are dimensionless and have the same scale (from 0 to 1) across all three experiments, they can be used for quantitatively comparing the forcing wind datasets. These statistics are tabulated for H_s in Tables 4.3 and 4.4 and for T_p in Tables 4.5 and 4.6.

	20CR		CFSR		ERA-I	
Buoys	μ	$\hat{\mu}$	μ	$\hat{\mu}$	μ	$\hat{\mu}$
41041	0.1355	0.6924	0.1957	1.0000	-0.0750	0.3832
41048	0.2177	1.0000	0.1022	0.4695	-0.1979	0.9090
W. of Heb	-0.1422	0.2048	-0.1797	0.2588	-0.6943	1.0000
F. of Forth	0.2705	1.0000	-0.1819	0.6725	-0.1607	0.5941
Moray Firth	-0.0552	0.1729	-0.3193	1.0000	-0.1352	0.4234
M3	0.2639	0.9234	0.2858	1.0000	-0.1418	0.4962
M4	0.5006	0.8188	0.6114	1.0000	0.1152	0.1884
M6	1.9088	1.0000	1.6814	0.8809	1.1224	0.5880

Table 4.3: Mean (μ) and normalised mean ($\hat{\mu}$) for bias in Significant Wave Height (H_s)

	20CR		CFSR		ERA-I	
Buoys	σ	$\hat{\sigma}$	σ	$\hat{\sigma}$	σ	$\hat{\sigma}$
41041	0.2939	1.0000	0.2451	0.8430	0.1778	0.6050
41048	0.5045	1.0000	0.3382	0.6704	0.3618	0.7171
W. of Heb	0.7615	1.0000	0.5014	0.6584	0.5663	0.7437
F. of Forth	0.4940	1.0000	0.2610	0.5283	0.3093	0.6261
Moray Firth	0.4028	1.0000	0.1998	0.4960	0.2931	0.7277
M3	0.7328	1.0000	0.4827	0.6587	0.4728	0.6452
M4	0.7166	1.0000	0.6045	0.8436	0.4660	0.6503
M6	0.9900	1.0000	0.8794	0.8883	0.5020	0.5071

Table 4.4: Standard deviation (σ) and normalised standard deviation ($\hat{\sigma}$) for bias in Significant Wave Height (H_s)

	20CR		CFSR		ERA-I	
Buoys	μ	$\hat{\mu}$	μ	$\hat{\mu}$	μ	$\hat{\mu}$
41041	0.4649	0.4142	1.1224	1.0000	0.3288	0.2929
41048	-0.3991	0.3699	-0.3987	0.3695	-1.0790	1.0000
W. of Heb	0.1432	0.2464	0.0070	0.0120	-0.5812	1.0000
F. of Forth	-1.3619	0.8920	-1.1143	0.7298	-1.5268	1.0000
Moray Firth	-1.9599	0.9140	-1.7479	0.8151	-2.1444	1.0000
M3*	0.1436	0.2528	-0.1160	0.2086	-0.5561	1.0000
M4*	0.0423	0.0616	-0.2265	0.3299	-0.6866	1.0000
M6*	-0.2876	0.3133	-0.4021	0.4380	-0.9180	1.0000

Table 4.5: Mean (μ) and normalised mean ($\hat{\mu}$) for bias in Wave Period (T_p). ${}^*T_e$ used for calculations.

	20CR		CFSR		ERA-I	
Buoys	σ	$\hat{\sigma}$	σ	$\hat{\sigma}$	σ	$\hat{\sigma}$
41041	1.9372	0.8320	2.2328	1.0000	1.8918	0.8125
41048	1.7263	1.0000	1.3582	0.7868	1.4289	0.8277
W. of Heb	1.5842	1.0000	1.1160	0.7045	1.1566	0.7301
F. of Forth	2.9078	1.0000	2.6722	0.9190	2.8124	0.9672
Moray Firth	3.8776	1.0000	3.5251	0.9091	3.6541	0.9424
M3*	1.2208	1.0000	0.9714	0.7957	0.9535	0.7810
M4*	1.2024	1.0000	0.9327	0.7757	0.9804	0.8154
M6*	1.0491	1.0000	0.8722	0.8314	0.8900	0.8483

Table 4.6: Standard deviation (σ) and normalised standard deviation ($\hat{\sigma}$) for bias in Wave Period (T_p). ${}^*T_e$ used for calculations

For the forcing datasets, the absolute value (in degrees) of the grid interval and the sampling rate (in hours) were taken for scoring based on spatial and temporal resolution; for the indicators of quality, the normalised mean bias was averaged across all buoys for H_s and T_p; and the inverse of the span of data (in years) was used for scoring based on the extent of the dataset. As observed earlier, a smaller value of each of these scores is indicative of greater desirability.

An example of calculating the dimensionless raw score is shown for the spatial resolution category. Under this criterion, the poorest raw score corresponds to the 20CR dataset which is resolved at 1.875° and is therefore assigned a raw score of 10.00. The dimensionless raw score for the ERA-I dataset, resolved at 1.0° is then calculated as:

$$\begin{aligned} \text{score(ERA-I)} &= \frac{1.0}{1.875} \times 10 \\ &= 5.33 \end{aligned}$$

In this manner, dimensionless raw scores are calculated for all the datasets under all the identified criteria. Final scores are obtained by weighting the raw scores. The weights, raw scores and weighted scores for each forcing dataset are tabulated in Table 4.7.

		Raw Score			Weighted Score		
Factor	Weight	20CR	CFSR	ERA-I	20CR	CFSR	ERA-I
Spatial Resolution	7.5%	10.00	1.67	5.33	0.0750	0.0125	0.0400
Sampling Rate	7.5%	5.00	1.67	10.00	0.0375	0.0125	0.0750
Average $\hat{\mu}$	30%	7.82	8.58	10.00	0.2345	0.2573	0.3000
Average $\hat{\sigma}$	30%	10.00	7.77	7.55	0.3000	0.2331	0.2264
Duration of data	25%	2.5	10.00	10.00	0.0625	0.2500	0.2500
Score	100%				**0.7095**	**0.7654**	**0.8914**

Table 4.7: Scores and weighted scores of the performance of different datasets under consideration

From the information presented in Table 4.7, it is observed that the 20CR dataset is the most suitable for this study and the ERA-I is the least suitable. The effect of the span of data has affected this outcome in a large way, with the 20CR dataset being four times as desirable as its alternatives. However, if only the quality of wave output were considered, the 20CR dataset is found to be the most desirable in terms of accuracy (indicated by the score for mean bias) and least desirable in terms of precision (indicated by the score for standard deviation of bias). The implication of this is that the 20CR datasets may be unsuited for accurate estimation of individual events, but is likely to yield reliable statistics calculated over long periods.

Based on these outcomes and inferences, the 20CR dataset was selected for forcing WAVEWATCH III to produce a long term hindcast of wave parameters in the region for long-term climate analyses.

4.4 Uncertainty analysis for model tuning parameters

In any analysis where the results are not deterministic, it is important to understand and, wherever possible, quantify the uncertainty at every stage of the analysis. In a study such as this, the overall uncertainty may be composed of the uncertainties in the model, input datasets, systematic errors, measurement errors, sampling uncertainty, and scenario uncertainty (Hayes, 2011). This analysis seeks to quantify one such source of uncertainty.

In a numerical model like WAVEWATCH III, there are several tunable parameters (e.g. bottom friction coefficient, swell feedback factor and white-capping coefficient) whose values can be adjusted by the modeller for the purpose of calibration. These parameters affect the magnitude of various physical processes which affect the generation and propagation of waves and they can take a wide range of values. Moreover, the value of these parameters can vary from one region to another, depending on local conditions. As a consequence, epistemic uncertainty due to the not knowing the optimum value of these tuning parameters is a large constituent of the overall uncertainty in the model's output.

In numerical modelling, it is the predictive nature of the model that is of interest to the users. In such situations, it is usual practice to represent the uncertainty in the form of probabilistic confidence intervals. Several methods have been used in the past for similar assessments of uncertainty, such as the 'Generalized Likelihood Uncertainty Estimation (GLUE)' method (Beven and Binley, 1992), the Monte-Carlo Simulation (MCS) methodology (Ang and Tang, 2007) and the Point Estimation Method (PEM) (Harr, 1989). This study follows the MCS methodology for uncertainty estimation adapting it for application to wave modelling with WAVEWATCH III.

In this section, the traditional MCS method is introduced along with a description of the methodology adapted specifically for the estimation of epistemic uncertainty due to tuning parameters in WAVEWATCH III for the doctoral research. Finally, the results of the analysis are presented.

4.4.1 The Monte-Carlo Simulation (MCS) methodology

The Monte-Carlo Simulation approach is a computer-based method that uses data from a large number of trials obtained from a random or pseudo-random sampling scheme to obtain a probabilistic estimation of the desired quantity which, in this case, is the uncertainty (Hayes, 2011).

The general procedure for uncertainty estimation using the MCS approach is:

1. Model input parameters whose values are not precisely known are identified;
2. The range within which the values of these parameters may lie is defined;
3. A large number of unique combinations of parameter values are generated using a suitable random or pseudo-random sampling scheme;

4. The model is run for all the sample sets and the bias between the simulated output and measured data is calculated
5. A probability density function is fitted to the bias and the uncertainty defined by a suitable confidence interval is estimated.

In this study, a modified form of the MCS approach is used for estimating the uncertainty due to tuning parameters in WAVEWATCH III when forced with 20CR winds. The following section explains the methodology followed in greater detail.

4.4.2 Procedure for uncertainty analysis

Identification of critical parameters

There are a large number of physical processes that affect the generation and propagation of wind generated waves which vary from deep water to shallow water. The magnitude of several of these processes can be 'tuned' by the modeller during the process of model calibration so that the wave model's output resembles measured data as closely as possible. These effects are listed in Table 4.8 (Holthuijsen, 2007).

	Oceanic Waters	Coastal Waters		
Process		Shelf Seas	Nearshore	Harbour
Wind generation	●●●	●●●	●	○
Quadruplet interactions	●●●	●●●	●	○
White-capping	●●●	●●●	●	○
Bottom friction	○	●●	●●	○
Shoaling	○	●●	●●●	●●
Depth induced breaking	○	●	●●●	○
Triad interactions	○	○	●●	●
Reflection	○	○	●/●●	●●●
Diffraction	○	○	●	●●●

Key: ●●● - dominant; ●● - significant but not dominant; ● - of minor importance; ○ - negligible.

Table 4.8: Relative importance of ocean wave processes. Sourced from Holthuijsen (2007)

Although uncertainty estimation methods involving computer simulated experiments are comparatively less expensive in comparison with physical experiments, computational economy must be considered in the case of climatological models. A large number of tuning parameters would require an experiment with the number of trials in the tens or hundreds of thousands requiring extremely large computational resources. Considering this, it is important to identify a small number of parameters governing the most critical processes so that the size of the experiment is kept within practical bounds.

Based on the information in Table 4.8 and on the advice of the developers of WAVEWATCH III (Tolman *et al.*, 2014b), the following tuning parameters were identified for the uncertainty analysis:

1. **Linear wind growth parameter ($\beta_{\max}$):** The parameterisation in the 'ST4' (Ardhuin *et al.*, 2010) source term package selected for the doctoral research is adapted from the Wave Action Model (WAM) (The WAMDI Group, 1988) using the positive part of the wind input. $\beta_{\max}$ is the non-dimensional growth parameter and its value has been found to vary from $\beta_{\max} = 1.2$ (WAM default) to $\beta_{\max} = 1.55$ (Ardhuin *et al.*, 2010) in literature. The default value in WAVEWATCH III for a global grid is $\beta_{\max} = 1.52$. A range for $\beta_{\max}$ between 1.25 and 1.75 was arbitrarily selected such that the WW3 default is approximately in the middle.
2. **Swell feedback parameter (s_1):** This parameter, also a part of the 'ST4' source term package, is used to adjust the friction factor as a function of wind speed and direction. This affects the magnitude of the momentum feedback from the ocean surface to the atmosphere and consequently controls the dissipation of swells. The WW3 default value for this parameter is $s_1 = 0.8$. Keeping this in the middle, a range of 0.65 and 0.95 was chosen for the uncertainty analysis.
3. **JONSWAP bottom friction coefficient (Γ):** The bottom friction coefficient determines the magnitude of energy dissipated due to the interaction of the wave with the ocean floor. In this study, the JONSWAP formulation for bottom friction (Hasselmann *et al.*, 1973) was chosen, and the WW3 default value for this parameter is $\Gamma = 0.038\mathrm{m}^2\mathrm{s}^{-3}$. Previous studies have used values as high as $\Gamma = 0.067\mathrm{m}^2\mathrm{s}^{-3}$ (Cialone and Smith, 2007), however, this is not recommended (van Vledder *et al.*, 2010). Values as low as $\Gamma = 0.019\mathrm{m}^2\mathrm{s}^{-3}$ (Tolman *et al.*, 2005) have also been used for modelling waves generated during hurricanes. Keeping the model default approximately in the middle, a range for Γ between 0.017 and $0.059\mathrm{m}^2\mathrm{s}^{-3}$ was selected for the investigation.

Experiment design

For an uncertainty analysis using the Monte-Carlo Simulation method, an experiment with a large number of unique parameter sets (trials) needs to be designed. Traditionally, a random sampling scheme such as the Monte-Carlo sampling method is used to generate samples of parameter combinations, usually numbering in the tens of thousands.

As this scheme follows a random sampling approach, there is no certainty that the experiment design will adequately represent the space to be investigated, or that the density of samples will be uniform across the space. To avoid this, it is necessary to generate a very large number of samples, which is a major drawback when economy is a significant factor in the design of experiments.

To overcome this, a Latin Hypercube sampling (LHS) scheme (Mckay *et al.*, 2000) was used instead of the Monte-Carlo sampling scheme for generating the unique samples (trials) for the uncertainty estimation experiment. The LHS scheme is a popular space-filling design which provides an economical alternative to the Monte-Carlo scheme. It is a stratified, iterative, pseudo-

random sampling method in which samples are generated from an n-dimensional square matrix where each dimension corresponds to one parameter. Samples are selected such that every level of each dimension is represented once, and that the density of samples is uniform across the volume under investigation (Deutsch and Deutsch, 2012).

By using the LHS approach, a much smaller number of trials are needed in the experiment to provide results with the same confidence as a Monte-Carlo type experiment with a larger number of samples (Janssen, 2013). In this experiment, a Latin Hypercube scheme was used to generate 500 unique samples of parameter combinations for the experiment. This was achieved by using the 'lhs' software package for the statistical language R (Carnell, 2012).

Filtering of behavioural trials

In most sensitivity and uncertainty analysis methods, including the Monte-Carlo method, no distinction is made between behavioural and non-behavioural trials. In this study, however, such a distinction is necessary to prevent the analyses from being distorted by unrealistic outcomes of the simulated experiment. This differentiation between plausible and implausible trials has been adapted from the Generalized Likelihood Uncertainty Estimation (GLUE) approach (Beven and Binley, 1992).

A behavioural trial may be defined as a trial whose outcome is plausible and realistic, and may be a representative of the true state. The difference between a behavioural and non-behavioural trial is illustrated in Figures 4.9 and 4.10. In the figures H_s and T_p simulated for Dowsing are plotted for one behavioural and one non-behavioural trial along with buoy observations.

A close examination of Figure 4.9 reveals that the simulated time-series of H_s from the non-behavioural trial exhibits some 'jitter', whereas the output from the behavioural trial is a smooth curve. Despite the bias (which can be corrected by calibration) being higher for the non-behavioural trial than for the behavioural trial, the correlation of the H_s data from both with buoy observations is high. The difference between the two trials is more obvious in the plot of the peak wave period. It can be observed in Figure 4.10 that for the non-behavioural trial, a large proportion of the simulated data corresponds to the first frequency bin, i.e. 0.0418Hz or $\approx$ 22s. It is possible that under certain combinations of tuning parameters and wind input, the spectrum becomes unstable, resulting in non-behavioural trials (Tolman *et al.*, 2014b). Data from the behavioural trial, on the other hand, is more realistic and exhibits low bias with buoy measurements.

The filtering out of non-behavioural trials is a debatable action. It can be argued that all the trials must be included in the analysis to provide a more complete representation of the uncertainty, as even non-behavioural output has a high probability of occurrence. However, when such output is encountered, it is obvious to the modeller, without any analysis more complex than visual inspection, that the model needs to be recalibrated and rectified.

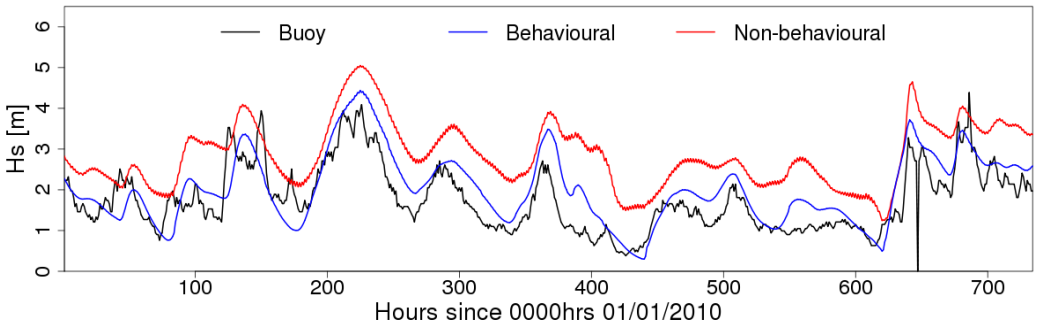

Figure 4.9: Significant wave height output for Dowsing from a behavioural and non-behavioural trial along with buoy observations

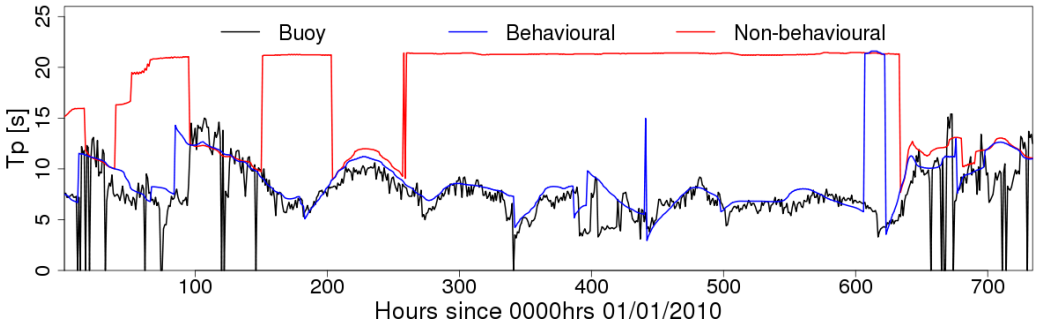

Figure 4.10: Peak wave period output for Dowsing from a behavioural and non-behavioural trial along with buoy observations

The counter-argument is that uncertainty is created when model output is plausible but of unknown accuracy. It is in such situations, that uncertainty analyses are useful because they capture the range of plausible outputs from the model for variations in the tuning parameters. This view is taken in this study.

It is evident from Figures 4.9 and 4.10 that non-behavioural trials cannot be easily identified from H_s data simulated by the wave model. It was found that the peak wave period for non-behavioural timesteps corresponds with the period associated with the first frequency bin (i.e. 0.0418Hz). The following algorithm, applied to T_p data, was used to filter out non-behavioural trials:

1. Count the number of time-steps, n, for which $T_p = 0.0418^{-1}$s
2. If $n < 5\%$ of the total number of timesteps, accept the trial.
3. Else, reject the trial.

All trials were found to be behavioural for the sites located in the North East Atlantic. After applying the filtering algorithm to locations in the North Sea, 173 trials were found to be behavioural across all sites. Two-dimensional projections of the search space are presented in Figure 4.11 with behavioural and non-behavioural parameter sets marked as green dots and red crosses respectively.

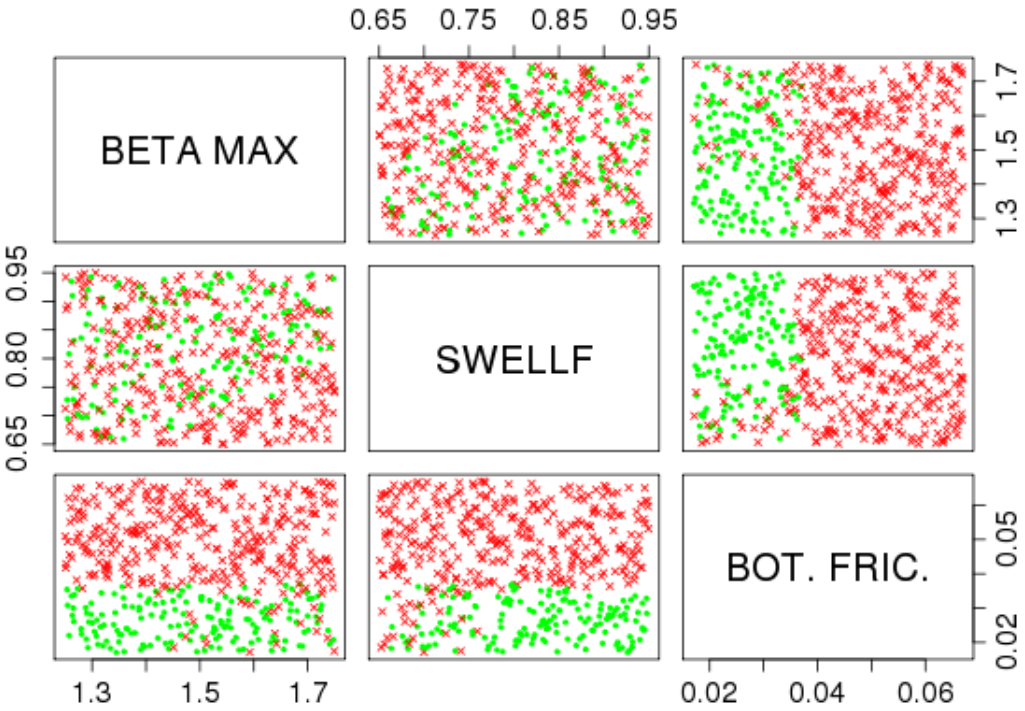

Figure 4.11: Scatter plots showing 2-dimensional projections of parameter sets plotted as pairs. Green dots - parameter sets for behavioural realisations; red crosses - parameter sets for non-behavioural realisations.

It is interesting to note in Figure 4.11 that a distinct region exists where all the trials are behavioural. It is observed that all trials for JONSWAP bottom friction $\Gamma > 0.038$ are non-behavioural and trials for which $\Gamma \leq 0.038$ and $s_1 \geq 0.75$ were all behavioural. The observation of the threshold at $\Gamma = 0.038$ is consistent with the findings of van Vledder *et al.* (2010).

As mentioned earlier, non-behavioural trials were found to exist only in the North Sea. This would imply that in the North Sea, the bottom friction coefficient plays an important part in ensuring the stability of the computed wave spectrum. It is, however, unlikely that bottom interaction is a dominant process except for very long waves and in shallow waters.

Data from the behavioural trials were used in further analysis.

4.4.3 Quantification and communication of uncertainty

A common method of quantifying and communicating uncertainty in the output of a numerical model is in the form of a confidence interval. In such an approach, a suitable probability interval is selected, usually 90%, 95% or 99% depending on application and acceptable level of risk (Loucks *et al.*, 2005). The uncertainty for the selected confidence interval is quantified as the difference between the output at the upper and lower bounds. In this study, the 90% interval calculated between the 5% and 95% probability levels is selected, as is common practice in engineering applications.

Such a method is effective in quantifying the level of uncertainty when the output is a discrete quantity or event. However, this approach is not easily applied to an ensemble of time-series data. In such situations, the following methods may be used for communicating uncertainty (Loucks *et al.*, 2005):

1. Histograms of important discrete events (graphical);
2. Continuous time-series plots of the confidence interval calculated at each time-step (graphical); and
3. Width of the interval averaged over the entire time-series (quantitative).

Probability density functions were fitted to the sample of behavioural trials for each timestep, assuming normality of the sample. This is a valid assumption according to the Central Limit Theorem which states that the distribution of the sample will more closely approximate the Normal Distribution as the size of the sample increases (Rice, 1994).

From the PDF, for each time-step, the 5% and 95% quantiles were calculated for H_s and T_p. Histograms of model output at 0000hrs January 6, 2010 at the locations West of Hebrides and Tyne/Tees are shown in Figure 4.12. Similar histograms for the other test sites are presented in Figures A.1 to A.3 in Appendix A.

It is not feasible to present histograms such as those in Figure 4.12 for each timestep in the time-series. As an alternative, the uncertainty is presented in the form of a time-series plot in Figures 4.13 and 4.14 where the lower (5%) and upper (95%) bounds for each timestep are plotted along with buoy measurements for H_s and T_p at the locations West of Hebrides and Tyne/Tees. The length of the plotted time-series is less than 744 (=31 days x 24 hours) because records for which no buoy data was available have been removed and not included in the analysis. Similar time-series plots for the other test sites are presented in Figures A.4 to A.14 in Appendix A.

It may be observed in the time-series plots that a large number of buoy observations are not within the confidence bounds, estimated in the study. This indicates that the 90% interval selected is too narrow and that a wider interval, such as the 95% or 99% intervals would yield confidence bounds that would envelop buoy measurements. Moreover, some bias between simulated and observed wave conditions is to be expected on account of inaccuracies in the

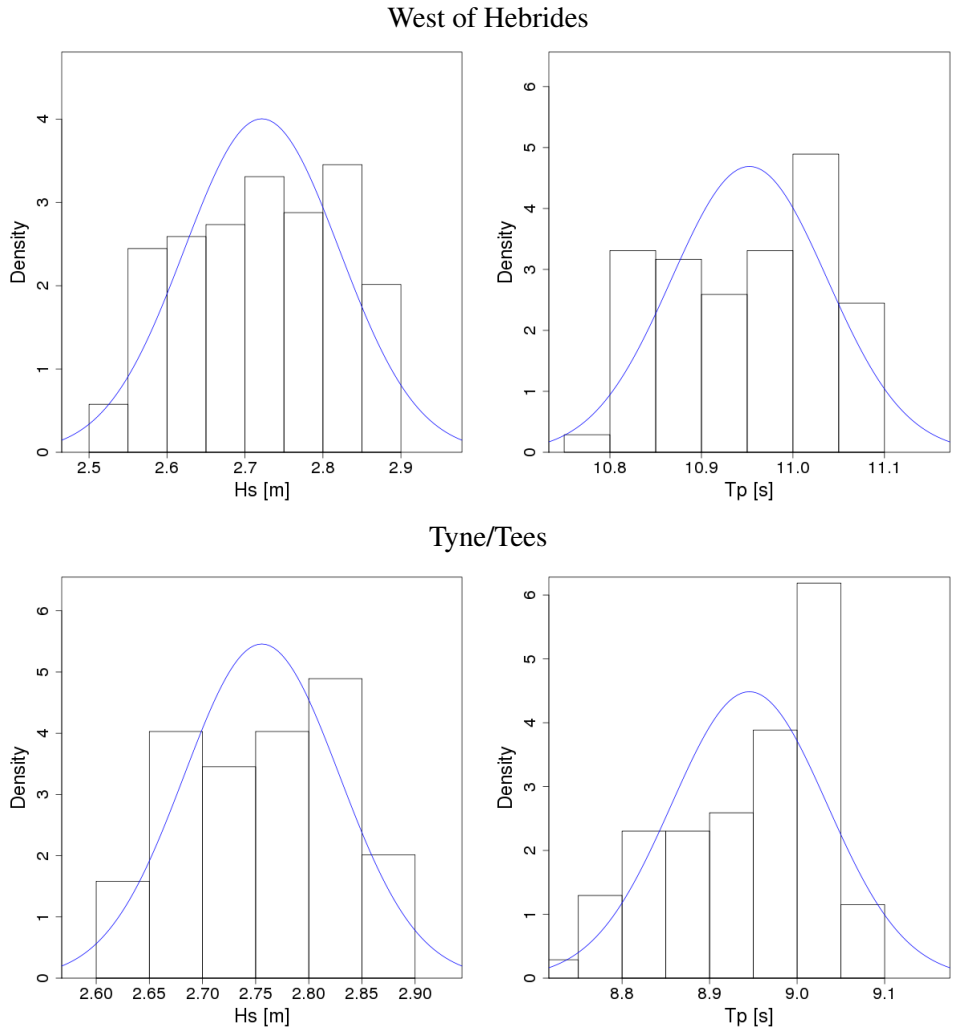

Figure 4.12: Histograms of output with the theoretical probability function from behavioural trials for 0000hrs January 6, 2010.

forcing dataset and errors in buoy measurements.

Selecting a wider interval will change the visual representation of uncertainty, and possibly include observed data from the buoys. Howevever, quantitatively, it has no effect on the uncertainty. For this reason, the 90% interval is persisted with as a standardised confidence interval.

The width of the uncertainty interval, calculated as the difference between the upper (95%) and lower (5%) probability bounds, averaged over the time-series are presented in Table 4.9. These values offer a quantitative estimate of the uncertainty in the output of the wave model due to the chosen tuning parameters.

It can be seen from the data in Table 4.9 that the uncertainty varies from ≈0.25m to ≈0.66m in H_s and from ≈0.63s to ≈3.04s in T_p. While this interval may appear narrow in absolute terms, in relative terms it is between 20% to 30% of the mean annual H_s at the locations tested.

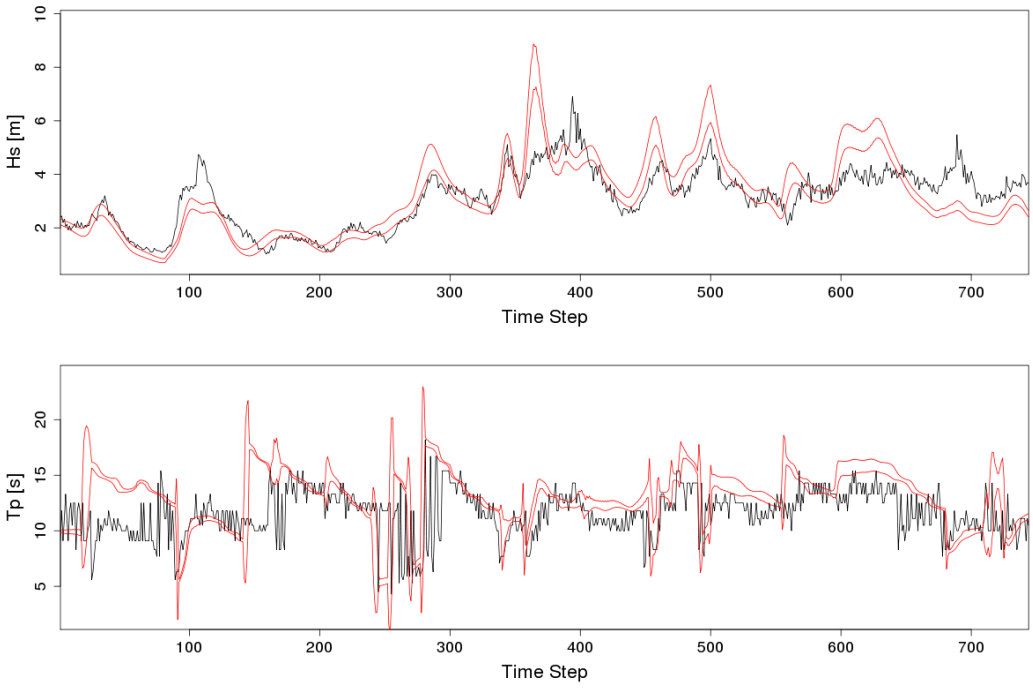

Figure 4.13: Time-series plots for January 2010 along with buoy measurements for the location West of Hebrides

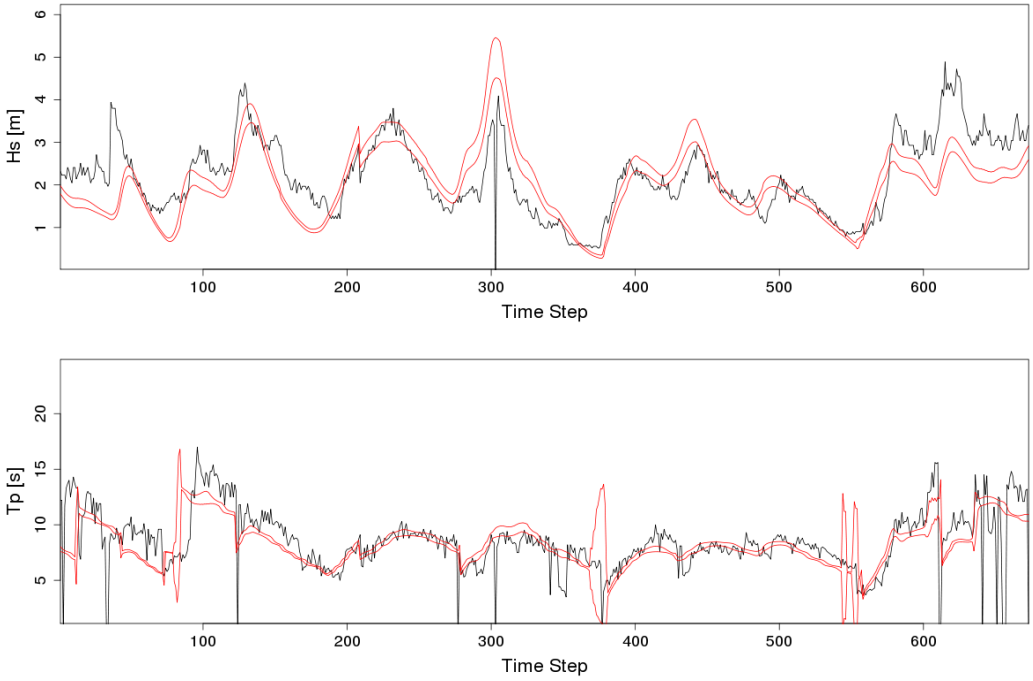

Figure 4.14: Time-series plots for January 2010 along with buoy measurements for the location Tyne/Tees

Site	H_s (m)	T_p (s)
41041	0.316	1.410
41044	0.325	1.147
41049	0.361	1.183
C44138	0.648	0.970
C44140	0.657	0.785
C44141	0.581	1.056
C44150	0.416	1.460
Blackstones	0.380	1.321
Dowsing	0.344	3.042
Forth	0.255	0.712
Moray Firth	0.243	0.633
Tyne/Tees	0.266	0.786
W. of Heb	0.473	1.232

Table 4.9: Width of the uncertainty interval for different test sites averaged over the timeseries for January 2010

Similarly, for T_p, the relative width of the interval would be between 10% and 25% of the annual mean of observed peak period.

In view of the large relative width of the confidence interval, it would appear prudent to calibrate the wave model to minimise the bias between simulated and observed wave conditions.

4.5 Sensitivity of the model to tuning parameters

Sensitivity analysis often accompanies uncertainty analysis as the same model runs can be used to perform both. While the latter analysis seeks to describe the range or interval of possible outcomes, the objective of a sensitivity analysis is to determine the effect of changes in the input parameters (or tuning parameters in this study) on the model's output.

A drawback of the MCS method is that the combined uncertainty due to the tuning parameters is estimated. From this type of uncertainty estimate, it is difficult to ascertain uncertainty in the model's output due to individual parameters. In other words, it is difficult to assess the sensitivity of the model to each parameter from the data collected from an MCS type experiment. The Least Squares Linearization (LSL) method (Lei and Schilling, 1996) is an approach often used with MCS type uncertainty analyses which splits the combined uncertainty from an MCS analysis into components corresponding with individual parameters.

The Least Squares Linearization method offers a computationally inexpensive, easy to implement method of relating the overall uncertainty with the uncertainty due to the individual parameters tested (Saltelli *et al.*, 2000; Parysow *et al.*, 2000; Verbeeck *et al.*, 2006). In this method, sensitivity coefficients are calculated for each input parameter which are indicative of the contribution of the parameter to the overall uncertainty.

The LSL method relates the variance of the model's output with the variance in the input parameters for each realisation in the Monte-Carlo experiment. The LSL method, adapted for this study is derived in the following section.

4.5.1 Derivation of the sensitivity coefficients

Consider H_s, an output parameter of interest. For a behavioural trial, i of the MCS experiment, we have:

$$H_{si} = \overline{H_s} + \Delta H_{si} \tag{4.4}$$

where H_{si} is the significant wave height output from the trial, $\overline{H_s}$ is the mean H_s calculated over all behavioural trials, and ΔH_{si} is the change in the output from the mean value.

Similarly, the small changes in the input parameters (identified in Section 4.4.2) can be written as:

$$\Delta\beta_{max_i} = \beta_{max_i} - \overline{\beta_{max}} \tag{4.5}$$

$$\Delta s_{1i} = s_{1i} - \overline{s_i} \tag{4.6}$$

$$\Delta\Gamma_i = \Gamma_i - \overline{\Gamma} \tag{4.7}$$

Assuming a linear relationship between the changes in output and the changes in tuning parameters, we get:

$$\Delta H_s = \frac{\partial H_s}{\partial \beta_{\max}}\Delta\beta_{\max} + \frac{\partial H_s}{\partial s_1}\Delta s_1 + \frac{\partial H_s}{\partial \Gamma}\Delta\Gamma \tag{4.8}$$

where the partial differential terms are the sensitivity coefficients (constant) calculated in deterministic approaches.

Substituting these into Equation 4.4:

$$H_{si} = \overline{H_s} + \frac{\partial H_s}{\partial \beta_{max}}(\beta_{max_i} - \overline{\beta_{max}}) + \frac{\partial H_s}{\partial s_1}(s_{1i} - \overline{s_1}) + \frac{\partial H_s}{\partial \Gamma}(\Gamma_i - \overline{\Gamma}) \tag{4.9}$$

Rearranging the terms in Equation 4.9:

$$H_{si} = \frac{\partial H_s}{\partial \beta_{max}}\beta_{max_i} + \frac{\partial H_s}{\partial s_1}s_{1i} + \frac{\partial H_s}{\partial \Gamma}\Gamma_i + \left(\overline{H_s} - \frac{\partial H_s}{\partial \beta_{max}}\overline{\beta_{max}} - \frac{\partial H_s}{\partial s_1}\overline{s_1} - \frac{\partial H_s}{\partial \Gamma}\overline{\Gamma}\right) \tag{4.10}$$

$$H_{si} = \frac{\partial H_s}{\partial \beta_{max}}\beta_{max_i} + \frac{\partial H_s}{\partial s_1}s_{1i} + \frac{\partial H_s}{\partial \Gamma}\Gamma_i + b \tag{4.11}$$

$$H_{si} = w_1\beta_{max_i} + w_2 s_{1i} + w_3\Gamma_i + b \tag{4.12}$$

where w_1, w_2 and w_3 are coefficients for the tuning parameters and b is the constant intercept. These can be estimated by linear regression methods e.g. Least Squares Estimation. b

represents the group of constant terms:

$$b = \overline{H_s} - \frac{\partial H_s}{\partial \beta_{max}}\overline{\beta_{max}} - \frac{\partial H_s}{\partial s_1}\overline{s_1} - \frac{\partial H_s}{\partial \Gamma}\overline{\Gamma} \tag{4.13}$$

The uncertainty from the MCS approach in Section 4.4, presented in the form of the interval between selected quantiles, is estimated from a PDF fitted to data from behavioural trials. This metric is closely related to the variance of the sample. In other words, the variance is another metric that can be used to quantify the uncertainty in the output of the wave model.

It is known for independent variables that $\mathrm{var}(aX + bY) = a^2\mathrm{var}(X) + b^2\mathrm{var}(Y)$ (Rohatgi and Saleh, 2011). Applying this property of variance to Equation 4.12, we get:

$$\mathrm{var}(H_s) = {w_1}^2\mathrm{var}(\beta_{max}) + {w_2}^2\mathrm{var}(s_1) + {w_3}^2\mathrm{var}(\Gamma) \tag{4.14}$$

Based on Equation 4.14, the sensitivity of H_s (S) to individual parameters can be calculated as:

$$S(\beta_{max}) = \frac{{w_1}^2\mathrm{var}(\beta_{max})}{\mathrm{var}(H_s)} \times 100 \tag{4.15}$$

$$S(s_1) = \frac{{w_2}^2\mathrm{var}(s_1)}{\mathrm{var}(H_s)} \times 100 \tag{4.16}$$

$$S(\Gamma) = \frac{{w_3}^2\mathrm{var}(\Gamma)}{\mathrm{var}(H_s)} \times 100 \tag{4.17}$$

Similarly, for T_p:

$$S(\beta_{max}) = \frac{{w_1}^2\mathrm{var}(\beta_{max})}{\mathrm{var}(T_p)} \times 100 \tag{4.18}$$

$$S(s_1) = \frac{{w_2}^2\mathrm{var}(s_1)}{\mathrm{var}(T_p)} \times 100 \tag{4.19}$$

$$S(\Gamma) = \frac{{w_3}^2\mathrm{var}(\Gamma)}{\mathrm{var}(T_p)} \times 100 \tag{4.20}$$

4.5.2 Sensitivity of model output

The relations for the sensitivity of H_s and T_p derived in Section 4.5.1 were applied to the model output from behavioural trials at the test locations. The calculated values of sensitivity the output to each tuning parameter are tabulated in Table 4.10.

The wave model was found to be most sensitive to the wind growth term $\beta_{\max}$ for the majority of sites, especially in oceanic waters. This was observed for both, H_s and T_p. An examination of the sensitivity of H_s to $\beta_{\max}$ reveals that it is lower for the southern sites (buoy locations 41041, 41044 and 41049) than for the remaining sites. It is likely that this is because of less energetic winds in the lower latitudes compared to the winds in the higher latitudes. For these

	H_s			T_p		
Location	β_{max}	s_1	Γ	β_{max}	s_1	Γ
41041	69.65	30.35	0.00	83.48	16.27	0.25
41044	68.84	31.16	0.00	88.36	9.40	2.24
41049	77.78	22.20	0.02	84.52	13.41	2.07
C44138	84.05	14.97	0.98	91.53	4.91	3.56
C44140	86.19	13.26	0.55	95.71	2.78	1.51
C44141	87.91	11.83	0.26	89.13	5.54	5.33
C44150	88.69	11.15	0.16	69.76	20.52	9.72
Dowsing	55.57	2.05	42.38	12.31	0.85	86.84
Firth of Forth	97.49	1.65	0.86	90.12	0.98	8.90
Blackstones	84.97	14.68	0.35	76.55	22.42	1.03
Moray Firth	98.80	1.11	0.09	91.00	0.75	8.25
Tyne/Tees	95.91	2.88	1.21	82.95	0.91	16.14
West of Hebrides	85.93	13.84	0.23	87.97	11.70	0.33

Table 4.10: Sensitivity H_s and T_p (in %) to tuning parameters

three locations, the sensitivity to swell dissipation parameter s_1 was found to be quite large ($\approx$ 20% to 30%). This, again is explained by the latitudinal location of the sites. This is likely because the dissipation of swells originating further north in the Atlantic Ocean is a significant physical process in this part of the domain.

It can be seen that the uncertainty for sites in the North Sea is dominated, again, by the wind growth parameter. The effect of the swell feedback parameter in this area is smaller than for sites in the North Atlantic. This may be because the North Sea is sheltered from the swells generated in the North Atlantic. It is known that the wave climate in the North Sea comprises of partially developed wind seas (Hasselmann *et al.*, 1973). Because of this, dissipation of swells may be secondary to wave growth, also explaining the difference in relative sensitivity of H_s and T_p to these parameters in the North Sea from the North Atlantic.

Considering the spatial resolution of the wave model, very few nodes in the grid are shallow enough to be affected greatly by bottom friction. Only the longest waves will interact with the ocean floor in shelf waters in the North East Atlantic and in the North Sea, and no interaction will occur in deeper oceanic waters. Consequently the model is not very sensitive to this tuning parameter which is reflected in the tabulated sensitivity values. It would appear, however, that bottom friction has a greater relative effect on the wave period predicted by the model than in the predicted wave height, especially in areas like Moray Firth and the Firth of Forth.

The site at Dowsing is anomalous in comparison with the other sites in terms of the results of the sensitivity analysis. From Table 4.10, it can be noticed that the sensitivity due to bottom friction is significantly higher for this location (42% for H_s and 87% for T_p). This is likely on account of the water depth at the Dowsing site being merely 22m (see Table 4.2). At this depth, the interaction of waves with the sea floor is likely to be a dominant process, as reflected by the

results of the analysis.

4.6 Parameter Optimisation

From the uncertainty and sensitivity analyses in Sections 4.4 and 4.5 it is evident that running WAVEWATCH III with arbitrary values of the tuning parameters is likely to yield sub-optimal output. This is especially true for the simulated wave periods due to the risk of non-behavioural model output.

The default values of these parameters provided in WAVEWATCH III have been identified by the developers from trials in which specific forcing wind data (e.g. CFSR, ERA-I) was used and wave conditions were simulated for the global domain (Ardhuin *et al.*, 2010). Consequently, when run with these values, WAVEWATCH III will produce the lowest bias averaged over then entire global domain as long as the same forcing dataset is used. Considering the relatively unique characteristics of the wave climate in the North East Atlantic and North Sea, it is possible that a different set of parameter values will produce the lowest bias when averaged over this region. Moreover, the 20CR dataset was identified for forcing the model in this research instead of the CFSR and ERA-I datasets.

At this time, there is no literature on the optimal parameter values specific to the North East Atlantic and North Sea, especially when forced with 20CR wind. This necessitates the calibration of the model, specific to the region and forcing wind, before it can be run to simulate wave conditions. To achieve this in a structured and scientific manner, the use of the Response Surface Methodology (RSM) is explored.

4.6.1 Theory of Response Surface Methodology (RSM)

Response Surface Methodology (RSM) is an application of experiment design (DoE) which integrates DoE methods with regression modelling and optimisation techniques such as linear and nonlinear programming (Beck and Arnold, 1977).

In RSM, as with other DoE methods, there exists a variable of interest, called the response variable or the predictand, whose value depends upon the values of a set of input variables (sometimes referred to as predictors). In most cases where RSM is employed, the relationship between the predictors and the predictand is either not known, or not easily modelled. In the response surface methodology, a set of experiments is designed to explore the space within which the predictors can vary, and the resulting response is measured. Statistical procedures like regression modelling are then applied to these data to approximate an empirical relationship between the input and response variables. This relationship is called the Response Surface. This can be represented mathematically as the generic function

$$y = f(x_1, x_2, ..., x_k) + \varepsilon \tag{4.21}$$

where y is the response, $x_1, x_2, ..., x_k$ are the predictors or input variables and ε is error.

In most real cases, the relationship between the response and input variables is seldom linear. Modelling a second order response function is considered a safe option for modelling such surfaces, with the increased accuracy coming at a higher computational cost. Equation 4.21 can be expanded to represent a second order response function as:

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum \sum_{i<j=2}^{k} \beta_{ij} x_i x_j \tag{4.22}$$

where y is the response, $x_1, x_2, ..., x_k$ are the predictors or input variables, and β is the coefficient for each term.

Once experiments from a suitable design are executed, the values of input variables and the resulting output or response is known. For 'M' variables and 'N' trials, the response and input data may be represented in matrix form (Triefenbach, 2008) as:

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} \tag{4.23}$$

$$X = \begin{bmatrix} x_{1,1} & x_{2,1} & \cdots & x_{M,1} \\ x_{1,2} & x_{2,2} & \cdots & x_{M,2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1,N} & x_{2,N} & \cdots & x_{M,N} \end{bmatrix} \tag{4.24}$$

From these matrices, the matrix of coefficients β can be found by solving (Myers and Montgomery, 2002; Triefenbach, 2008):

$$\beta = [X^T X]^{-1} \cdot X^T Y \tag{4.25}$$

The response function is then obtained by substituting the coefficients calculated from solving Equation 4.25 into Equation 4.22. This process may be simplified by using available toolboxes, e.g. 'rsm' for the statistical language 'R' (Lenth, 2009).

Optimisation algorithms such as linear and non-linear programming may be applied to the response surface to identify the values of input parameters for which the response is at the desired level depending on the nature of the response, i.e. maximum for response parameters like yield and minimum for bias.

The most important characteristics of a good response surface are:

- The model fits the data well;

- There is sufficient information to perform goodness-of-fit tests;
- The design is robust to outliers and random errors;
- The design is cost effective.

A good design of experiments for RSM strikes a balance between the cost of resource utilised and an acceptable level of accuracy of the response surface. To achieve this, there are several designs that may be used such as the Central Composite Design (CCD), Box-Behnken design and the D-optimal design. In this study, this is achieved by using data gathered from the experiment for the uncertainty and sensitivity analyses.

4.6.2 RSM for parameter optimisation in WAVEWATCH III

In numerical models like WAVEWATCH III, WAM and SWAN (Simulating WAves Nearshore), the exact relationship between parameters governing the physics of wave generation, propagation and dissipation and the output parameters like significant wave height, mean wave period, mean wave direction, etc. may be known. However, the process of calibrating the model involves minimising the biases in the output parameters. Deriving a deterministic relationship between the tuning parameters and the response (e.g. bias) is not feasible.

To overcome this, the model is treated as a black-box and an empirical function relating the tuning parameters and the response is identified using RSM, which can be used for the purpose of calibrating the wave model. The key steps in the application of RSM for parameter optimisation in numerical wave modelling with WAVEWATCH III are shown in Figure 4.15 and explained below.

Identification of critical parameters

Prior to response function modelling, it is common practice to conduct preliminary experiments called 'screening experiments', especially if there are several parameters which may or may not have an effect on the response. As the name suggests, these experiments are designed to screen all the parameters and reject those that do not significantly affect the response.

In this case, screening experiments are not performed as the parameters to be optimised are already known from the uncertainty analysis described in Section 4.4 - the JONSWAP bottom friction coefficient (Γ), wind growth parameter (β_{max}) and swell feedback parameter (s_1).

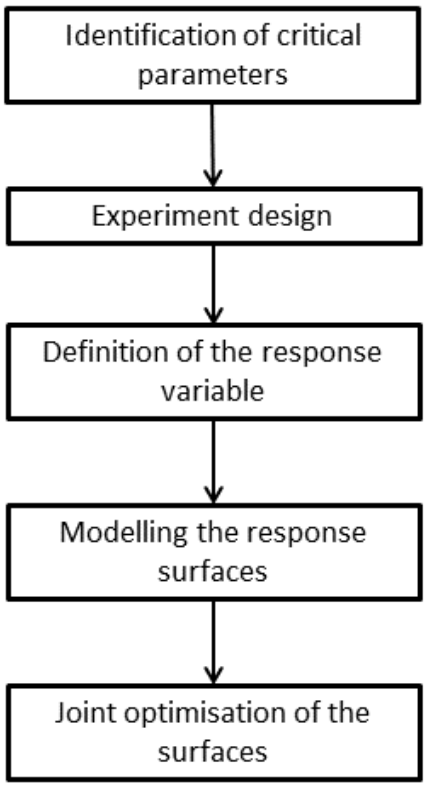

Figure 4.15: Key steps in Response Surface Modelling (RSM) for parameter optimisation in numerical wave modelling

Experiment Design

Expanding Equation 4.22 for 3 variables we obtain 10 terms (3 quadratic terms, 3 interaction terms, 3 linear terms and 1 constant term) for which the coefficients are unknown.

$$\begin{aligned} y = \beta_{11}x_1^2 + \beta_{22}x_2^2 + \beta_{33}x_3^2 + \beta_{12}x_1x_2 + \beta_{13}x_1x_3 + \\ \beta_{23}x_2x_3 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_0 \end{aligned} \tag{4.26}$$

To solve for 10 unknown coefficients, the minimum number of trials required in the experiment is also 10 if a design like the D-optimal design is used. In this study, data from the uncertainty estimation experiment is used for fitting the response surface. The benefits of this are:

- There is no additional cost for performing the simulation experiments
- A response surface of high accuracy can be fit using data from the large number of behavioural trials in the LHS design

Definition of the response variable

From the perspective of ocean waves as an energy resource, the most important output parameters from the wave model are the significant wave height (H_s) and the energy period (T_e). As a consequence of having more than one output parameter of importance, the calibration procedure is no longer straightforward as it needs to optimise the model taking both the output parameters into account. This is because it is not necessary that a parameter set yielding low bias in H_s would also yield low bias in T_e. To overcome this, two response variables were defined - one for wave height (r_H) and another for wave period (r_T).

The response is calculated in two steps. First, the normalised root-mean-square-errors (denoted by H_{rmse} for significant wave height and T_{rmse} for wave period) are obtained for each trial (i) and buoy location (j) from Equations 4.27 and 4.28. As energy period (T_e) is not readily available from buoy records, peak wave period (T_p) is used for calibrating the model.

$$H_{rmse_{i,j}} = \sqrt{\frac{\left[\frac{1}{n}\sum_{t=1}^{n}(H_t - \widehat{H}_t)^2\right]}{\max\left[(H_t - \widehat{H}_t)^2\right]}} \tag{4.27}$$

$$T_{rmse_{i,j}} = \sqrt{\frac{\left[\frac{1}{n}\sum_{t=1}^{n}(T_t - \widehat{T}_t)^2\right]}{\max\left[(T_t - \widehat{T}_t)^2\right]}} \tag{4.28}$$

In the above equations, i represents the trial, j represents the buoy location, t is the timestep, H_t and T_t are parameters recorded from the wave model at time t, and $\widehat{H}_t$ and $\widehat{T}_t$ are parameters recorded by the buoys at time t. To illustrate this, $H_{rmse_{200,3}}$ is the normalised root-mean-square-error in H_s for the 200^{th} trial at the 3^{rd} buoy location. Normalisation of the error is necessary to ensure that the response at every buoy location is equally reflected in the final response.

The overall response of the model for each trial is then calculated by averaging the calculated normalised root-mean-square-error over all the buoy locations.

$$r_{H_i} = \frac{1}{S}\left[\sum_{j=1}^{S} H_{rmse_{i,j}}\right] \tag{4.29}$$

$$r_{T_i} = \frac{1}{S}\left[\sum_{j=1}^{S} T_{rmse_{i,j}}\right] \tag{4.30}$$

The response, as calculate from Equations 4.27 to 4.30 is a derived form of bias whose value lies between 0 and 1. This enables the use of simple minimisation algorithms to find the optimum combination of parameter values.

Response surface modelling and optimisation

In order to find a combination of parameter values that produces the lowest bias in both, H_s and T_p, the procedure treats each output parameter separately, fitting an individual response surface for each. These response functions were calculated using the 'rsm' toolbox for the statistical language 'R' (Lenth, 2009).

Slices of the response surfaces for H_s and T_p are shown in Figures 4.16 and 4.17 respectively. From these, it can be seen that the two response surfaces have extremely different characteristics from one another. The difference in variation of the response to input variables can be inferred from the dissimilar patterns of the response contours in the slices of the response surface presented for the two output wave parameters.

The difference in these characteristics confirms that optimising based on the response of one parameter or the average response of the two parameters is likely to yield a sub-optimal set of parameter values. To avoid this, it is necessary to simultaneously optimise both response surfaces. In such a situation, it is unlikely that the solution set of parameter values will yield the lowest bias in either output parameter. The solution set is likely to lie in an area of compromise

Once the individual response functions are estimated, they are simultaneously optimised using the Joint Optimization Plot (JOP) toolbox for the statistical language 'R' (Kuhnt and Rudak, 2013) which applies non-linear programming (NLP) algorithms to find the 'best' set of input parameter values.

4.6.3 Optimum Parameter Values

The 'optimal' set of parameter values obtained by applying the RSM approach are tabulated in Table 4.11 along with the values obtained by prior studies for other forcing wind datasets for the global domain.

Parameter	20CR*	ERA-I$^{\#}$	CFSR$^{\#}$
Wind Growth (β_{max})	1.32	1.52	1.33
Swell Feedback (s_1)	0.83	0.8	0.8
Bottom Friction (Γ)	0.032	0.037†	0.037†

Table 4.11: Optimised parameter values for 20CR wind from this study tabulated alongside values for ERA-I$^{\#}$ and CFSR$^{\#}$ wind from other studies. *This study; †Model defaults; $^{\#}$Data from Tolman *et al.* (2014a)

It may be inferred from the values tabulated in Table 4.11 that calibration of the wave model is necessary for identifying optimal parameter values for different forcing winds and geographical regions. An interesting observation is that the results of the optimisation procedure are not very different from the recommended values for a global model forced with CFSR wind. While this may be a mere coincidence, it may also be because both reanalyses datasets are prepared using

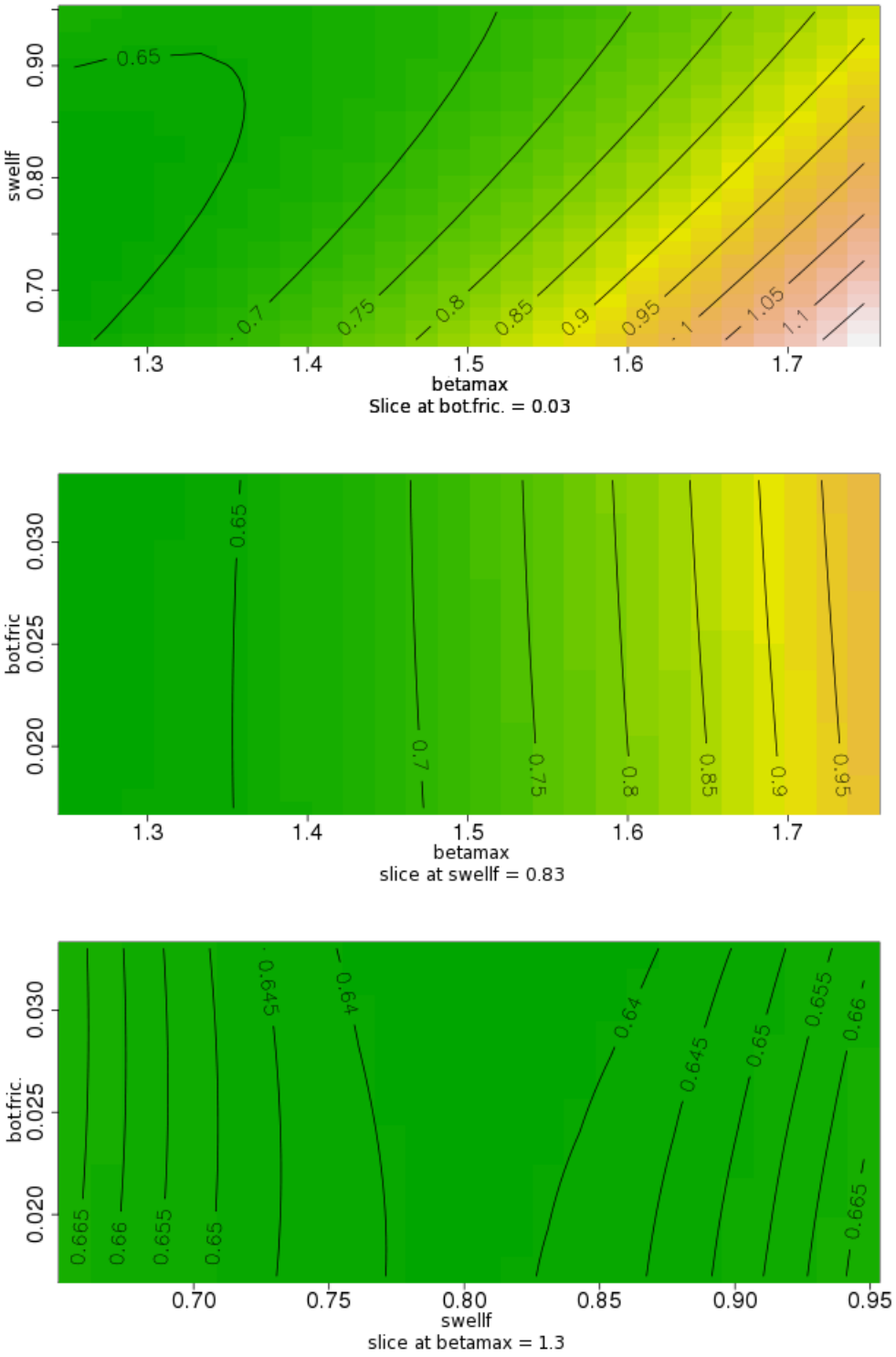

Figure 4.16: Slices of the response surface for significant wave height

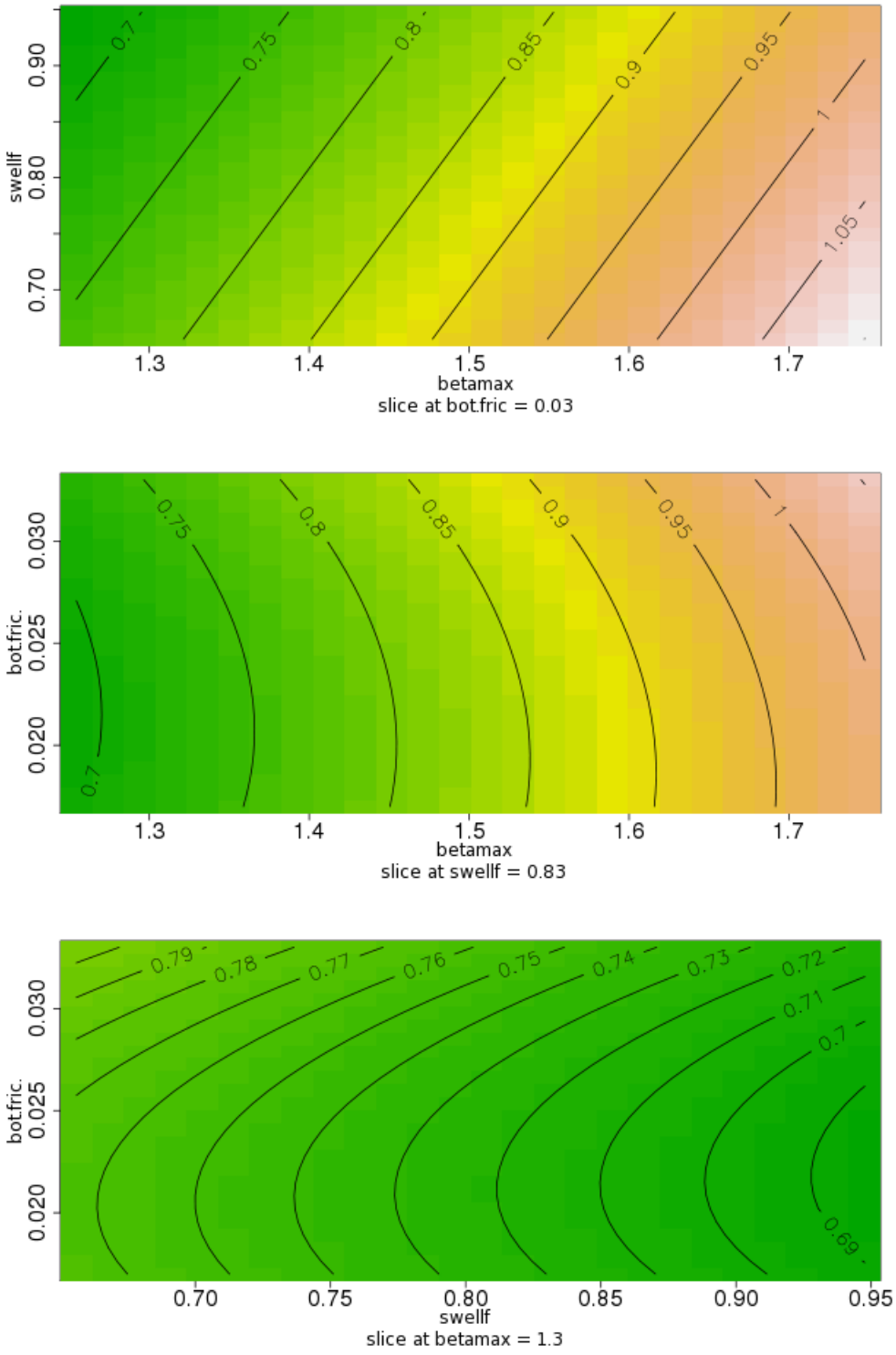

Figure 4.17: Slices of the response surface for wave period

the NCEP Climate Forecast System (CFS). Another possible explanation is that the sampling rates of the 20CR dataset (3-hourly) and the CFSR (hourly) are better able to capture sub-daily variations in wind than the 6-hourly ERA-I dataset.

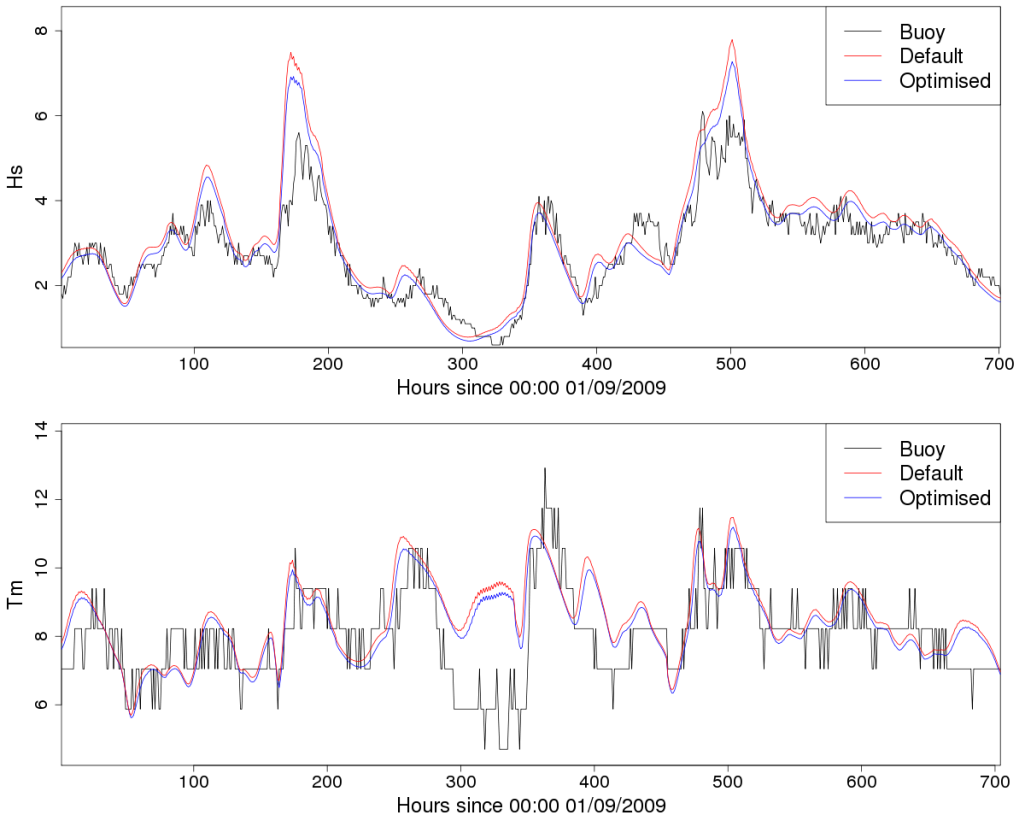

Figure 4.18: Timeseries of default and optimised model simulations along with buoy observations for September 2009 at the M4 buoy location. Top - significant wave height, bottom - mean wave period.

A comparison of the output from the wave model for September, 2009 with default and optimised parameters is presented in Figure 4.18. As it is difficult to identify any improvement in the quality of output from visual inspection of the plots, statistics describing the quality of the two simulations (mean bias, standard deviation and rmse) were computed for various buoy locations and are tabulated in Tables 4.12 and 4.13.

An examination of Tables 4.12 and 4.13 reveals that a reduction in the value of the mean bias (μ_e) and the rmse is observed for all the locations tested. In addition, the standard deviation of the bias (σ_e) is also seen to decrease, indicating that the spread of error between the model and the buoy has also reduced. From these data, the improvement in the model's output as a result of optimisation can be seen in quantitative terms.

	Default			Optimised		
Buoy Location	μ_e (m)	σ_e (m)	rmse (m)	μ_e (m)	σ_e (m)	rmse (m)
41041	0.0487	0.4111	0.4137	-0.0513	0.3622	0.3656
41048	-0.0065	0.4098	0.4095	-0.1025	0.3913	0.4042
W. of Hebrides	0.1444	0.6023	0.6191	0.0287	0.5571	0.5575
Moray Firth	-0.0757	0.3746	0.3819	-0.0698	0.3242	0.3314
M3	0.2734	0.4749	0.5476	0.1164	0.4288	0.4439
M4	0.3169	0.5982	0.6765	0.1294	0.5231	0.5385
M6	0.7019	0.6719	0.9713	0.4947	0.5778	0.7603

Table 4.12: Mean bias (μ_e), standard deviation (σ_e) and root-mean-square-error (rmse) statistics for H_s from default and optimised model runs

	Default			Optimised		
Buoy Location	μ_e (s)	σ_e (s)	rmse (s)	μ_e (s)	σ_e (s)	rmse (s)
41041	0.9171	1.7439	1.9691	0.6629	1.6777	1.8028
41048	0.4921	2.2584	2.3098	0.3056	2.1922	2.2119
W. of Hebrides	0.8546	1.4591	1.6901	0.6716	1.4324	1.5811
Moray Firth	-1.8238	3.4187	3.8726	1.2120	2.2993	2.5978
M3	0.9496	1.4861	1.7627	0.7174	1.406	1.5776
M4	0.7578	1.3046	1.5078	0.5513	1.2424	1.3584
M6	0.6588	1.1639	1.3367	0.4189	1.0895	1.1665

Table 4.13: Mean bias (μ_e), standard deviation (σ_e) and root-mean-square-error (rmse) statistics for T_p from default and optimised model runs

4.7 Validation

To assess the ability of WAVEWATCH III to simulate the sea state when calibrated with the identified set of 'optimal' parameter values, the model was run to hindcast wave conditions for the months April through June (summer months) and October through December (winter months) of 2009. For the validation of the model, simulations were run for two periods to ensure that the model performed satisfactorily in low and high energy conditions. H_s and T_p data from these simulations were compared with buoy observations from 5 locations.

Figures 4.19 and 4.20 show the H_s and T_p timeseries plots the NDBC buoy 41048 and the CEFAS buoy located West of Hebrides for the months April through June of 2009. Scatter density plots of model output plotted against buoy observations for the two buoys are presented in Figure 4.21. In these plots the density of points is indicated by the colour, red representing the highest and white representing the lowest. The plots presented here are for one deep water and one coastal location.

From the Figures 4.19 and 4.20 good agreement between model output and buoy measurements can be observed. This is more clearly visible in the scatter plots shown in Figure 4.21, indicated by the high density of points (indicated by the red shading) clustered about the unity diagonal.

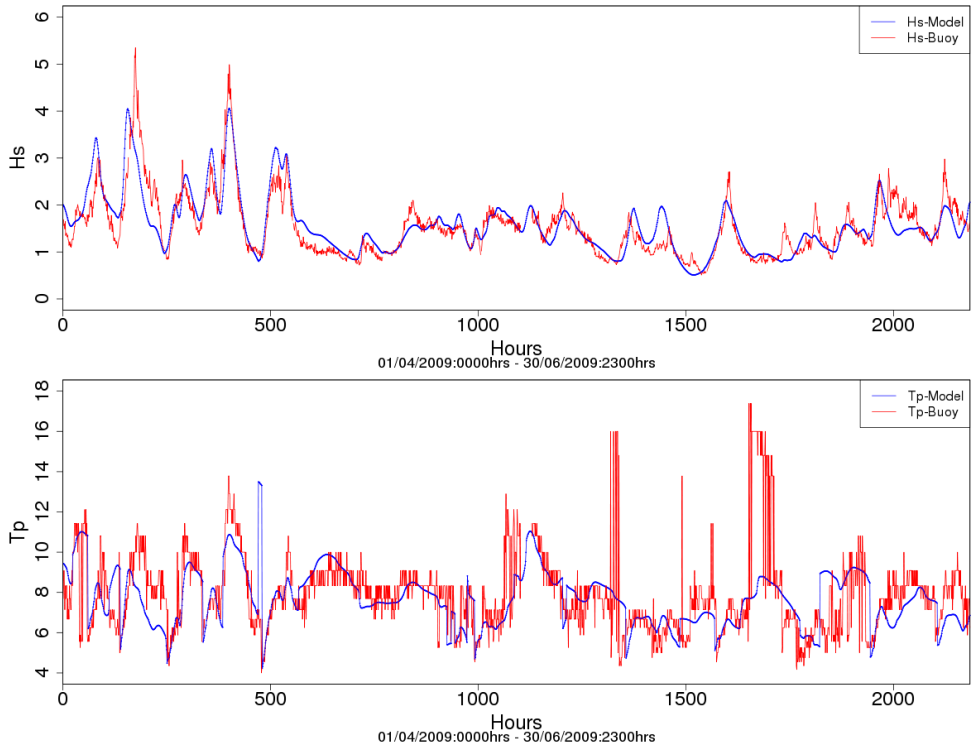

Figure 4.19: Validation plots for NDBC Buoy 41048 for the months April through June of 2009. Top - H_s (m); bottom - T_p (s)

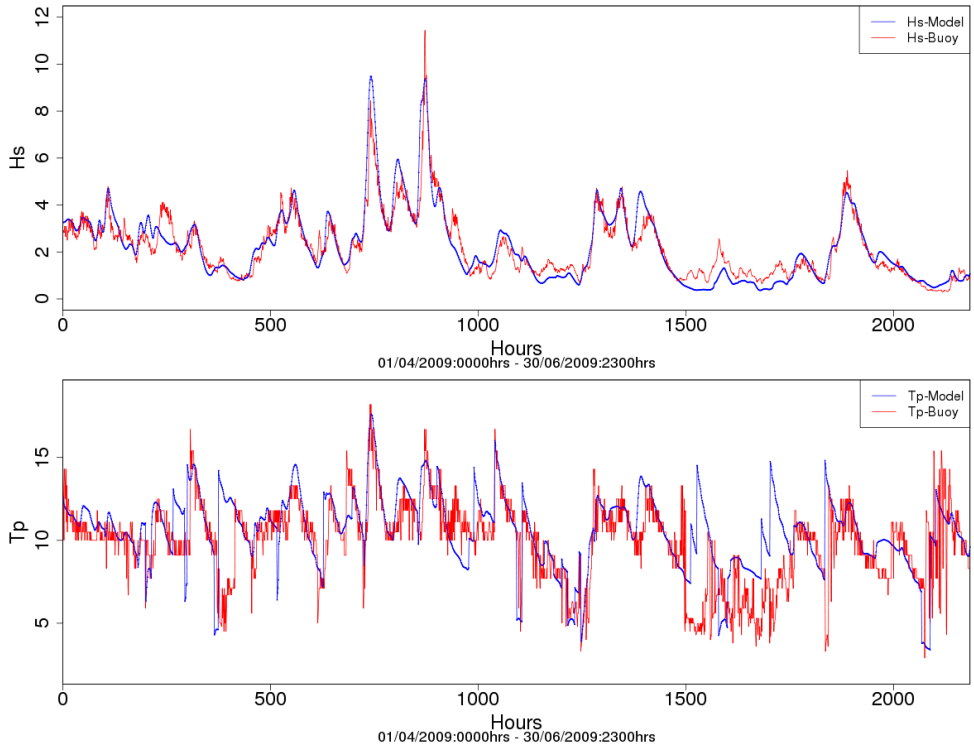

Figure 4.20: Validation plots for the CEFAS buoy located West of Hebrides for the months April through June of 2009. Top - H_s (m); bottom - T_p (s)

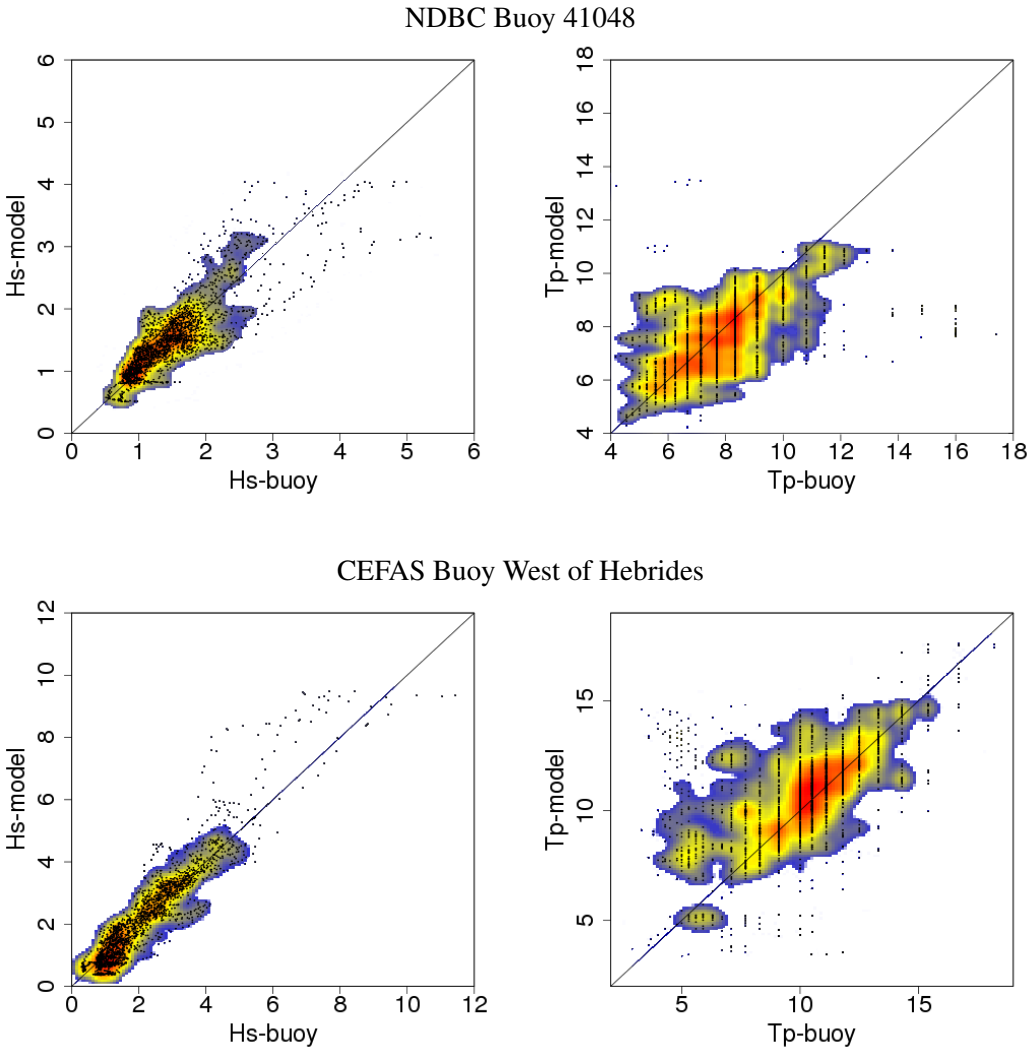

Figure 4.21: Scatter density plots of model output plotted against buoy observations for the NDBC Buoy 41048 (top) and CEFAS buoy located West of Hebrides (bottom); Left column - H_s (m); right column - T_p (s)

Not relying purely on visual inspection for validation, the model's performance was quantified using various indices such as mean bias, root-mean-square error, and scatter index (SI) for both, H_s and T_p. The scatter index (SI) is defined as the rms error normalised by the average observed value ($SI = rmse/\overline{X}$). These are tabulated in Tables 4.14 and 4.15.

In addition to the validation buoys, performance indices were also computed for the calibration buoys. These data can be treated as being independent as the validation period is different from the period from which the data were used for calibration. It can, however, be argued that these data are biased as the model was calibrated based on these locations. This notwithstanding, the performance indicators were calculated for these buoys to provide additional evidence of the model's performance. Some values are absent in the tables because data for some buoys were

	April - June			October - December		
Validation Buoy	Bias (%)	RMSE (m)	SI	Bias (%)	RMSE (m)	SI
41048	1.72	0.356	0.2312	5.20	0.493	0.2269
c44138	-	-	-	1.78	0.769	0.2292
Moray Firth	-0.39	0.308	0.3669	-7.43	0.502	0.3349
W. of Hebrides	1.20	0.570	0.2587	-4.20	0.802	0.2447
M4	8.29	0.607	0.2580	1.30	0.751	0.2125
Calibration Buoy	Bias	RMSE	SI	Bias	RMSE	SI
41041	6.98	0.236	0.1063	-8.09	0.346	0.1626
41044	4.18	0.206	0.1349	-5.49	0.336	0.1662
41049	-4.92	0.282	0.1945	-2.57	0.443	0.2082
c44140	-	-	-	7.01	0.711	0.2212
c44141	5.56	0.510	0.2601	4.10	0.693	0.2255
Blackstones	-1.54	0.479	0.2384	-3.24	0.629	0.2243
F. of Forth	6.51	0.407	0.4933	-2.61	0.528	0.3877
Tyne/Tees	-4.92	0.402	0.4193	-6.24	0.564	0.3341
M3	5.31	0.589	0.2398	-5.28	0.701	0.1907
M6	1.78	0.432	0.2148	1.58	0.841	0.2016

Table 4.14: Performance indicators for H_s calculated for validation buoys (top section) and calibration buoys (bottom section)

	April - June			October - December		
Validation Buoy	Bias (%)	RMSE (s)	SI	Bias (%)	RMSE (s)	SI
41048	-3.52	1.845	0.2295	2.60	2.394	0.2585
c44138	-	-	-	4.80	2.148	0.2046
Moray Firth	-6.10	1.335	0.3078	1.80	1.120	0.2040
W. of Hebrides	7.20	2.236	0.2281	5.40	2.375	0.2071
M4*	2.91	1.384	0.1627	-3.60	1.675	0.1641
Calibration Buoy	Bias	RMSE	SI	Bias	RMSE	SI
41041	2.85	1.733	0.2175	13.27	2.904	0.2869
41044	-1.85	1.359	0.1734	4.46	2.145	0.2083
41049	-1.84	1.679	0.2129	4.29	2.099	0.2129
c44140	-	-	-	-1.24	1.771	0.1604
c44141	-3.02	2.400	0.2698	4.54	2.311	0.2418
Blackstones	10.71	2.622	0.2535	4.49	2.087	0.2025
F. of Forth	-4.07	1.806	0.301	-12.18	2.077	0.2691
Tyne/Tees	-4.25	1.929	0.305	-5.15	2.587	0.3375
M3*	6.32	1.002	0.1581	6.75	1.343	0.1821
M6*	-0.81	0.815	0.1246	-0.21	0.995	0.1291

Table 4.15: Performance indicators for T_p calculated for validation buoys (top section) and calibration buoys (bottom section). *T_e used for calculations

not available for April-June, 2009.

It is important to consider all the performance indices when assessing the model's performance. For example, the scatter index (SI) for H_s at the Moray Firth site during April-June is 0.3669, which would imply that the model is inadequate in modelling the sea state at the location. However, the normalised bias for the same location and period is -0.39% and the root-mean-square-error (rmse) is 0.308m. From these statistics, it might be inferred that the error in the output is within acceptable limits. Similar observations can be made for the statistics computed at the Firth of Forth and Tyne/Tees.

It may also be that the resolution of the computational grid is too low to effectively model the shallow water processes in these areas. Moreover, some shallow water interactions, such as triads, were disabled because of their high demands on computing time. It is expected that increasing the resolution of the model, perhaps by using an unstructured computational grid, using improved local winds from a regional climate model (RCM), and enabling expensive shallow water physics would greatly improve the estimation of nearshore wave conditions.

It expected that the model performance with respect to wave period may be better than indicated in Table 4.15 due to the discrete nature of T_p data from buoys. The calculated bias and SI might improve if wave periods obtained from wave spectra were used for comparison, as these are likely to be continuous timeseries data. However, these were not used because they would have to be derived from the spectral output by the model, whereas T_p could be directly compared.

An intercomparison of data from WAVEWATCH III and the ERA-Interim reanalysis dataset is presented in Figures 4.22 for wave parameters averaged over the year 2000. From these panels, it can be seen that there is a high degree of agreement in significant wave height (H_s) and mean wave direction (D_m) between the two datasets. It would appear that in the North East Atlantic, the mean H_s from ERA-I is approximately 5% to 10% lower than that from WW3. This is most likely a consequence of the difference in sampling rates of the two datasets. The hourly WW3 data is more likely to capture peaks and other short term variations in the wave height than the 6-hourly ERA-I data.

Although the energy period (T_e) follows similar trends in the North East Atlantic, a noticeable difference can be observed in the contours in the North Sea. This may be due to the difference in temporal and spatial resolutions of the two datasets. In the absence of measurements of the wave period in the middle of the North Sea, it is difficult to ascertain which of the models is more accurate.

The normalised bias, rmse and SI values for the tested locations appear to be within acceptable limits for both simulated time periods. From these and the plots presented in Figures 4.19 to 4.21 it can be concluded WAVEWATCH III, when tuned as described in Section 4.6, is adequately able to simulate wave conditions described by significant wave height (H_s) and peak wave period (T_p). The overall agreement between the data from the wave model and from

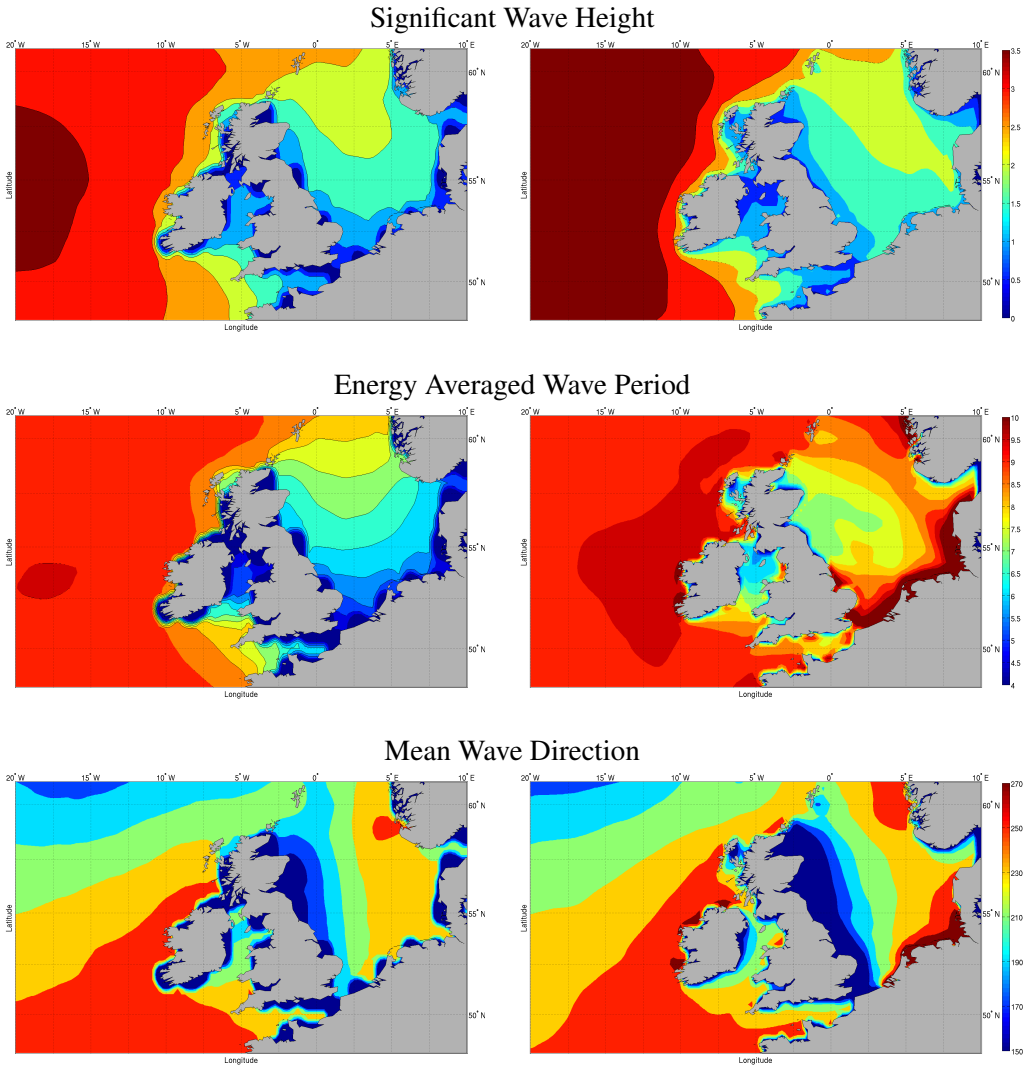

Figure 4.22: Intercomparison of H_s (top), T_e (centre), and D_m (bottom) from ERA-Interim (left) and WAVEWATCH III (right) averaged for the year 2000.

the ERA-I dataset corroborates this.

From the comparison of WW3 output with buoy observations and ERA-Interim reanalysis data, the performance of the model is deemed adequate for simulating long-term data for climate analyses.

4.8 Chapter summary

- The datasets to be used in the study are introduced. These include bathymetry, available reanalysis winds, buoy observations and projected winds for different climate-change experiments.
- The set-up of the model is described. This includes a description of the nesting scheme and computational grids, frequency-direction spectrum resolution and selection of appropriate model time steps according to the selected computational grids. Wave parameters selected for output at hourly intervals are also mentioned.
- A score-based comparison method is presented for the selection of an appropriate reanalysis dataset taking into account factors such as accuracy and precision of output wave parameters and temporal span of the data with appropriate weighting. Based on these, the 20th Century Reanalysis (20CR) dataset was selected for the preparation of a long-term hindcast of wave conditions for the research project.
- A modified Monte-Carlo type experiment is designed using the Latin Hypercube sampling scheme with 500 unique permutations of the tuning variables governing wind-growth, swell feedback and bottom friction. WAVEWATCH III is run to simulate wave conditions for January 2010.
- Data from the experiment is used to perform an uncertainty analysis to quantify model uncertainty due to the selected tuning parameters. Further, the sensitivity of the model to these parameters is also assessed. Wind growth was found to be a dominant parameter followed by swell feedback. Bottom friction was found to have negligible effect on the wave modelŠs output.
- Response surfaces relating normalised bias and the selected tuning parameters were modelled for Hs and Tp using data from the experiment. These were simultaneously optimised using nonlinear programming methods for minimisation and an ‘optimum’ set of values was obtained. The optmimum parameter set for the selected region and forcing wind is tabulated along with model default values.
- The tuned model is validated for two 3-month periods against buoy measurements. Performance indicators such as bias, root mean square error (RMSE) and scatter index (SI) are computed.

Chapter 5

Analysis of the wave climate - I

5.1 Introduction

The Intergovernmental Panel on Climate Change (IPCC) defines detection of climate change as "*the process of demonstrating that the climate has changed in some defined statistical sense without providing a reason for that change*" (Hegerl *et al.*, 2009). Any change that is detected is considered to be statistically significance if the probability of occurrence of the trend by chance or due to internal variability is less than a pre-defined level, usually 10%.

Several studies have indicated changes in surface wind patterns with positive and negative trend bands observed in the North Atlantic Ocean in data from 1988-2010 (Hartmann *et al.*, 2013). An assessment of offshore winds revealed an increase of 15 to 20% in winter wind speeds in the UK (Watson *et al.*, 2001). The confidence associated with such trends, however, is low on account of sparseness of long-term measurements, inhomogeneity in datasets and large differences between global datasets (Field *et al.*, 2012). Due to the nature of wind-generated waves, changes in surface winds are likely to produce changes in the regional wave regimes.

The wave climate in the North Atlantic has been studied extensively for long-term changes since the 1980s (Draper, 1986; Carter and Draper, 1988) with varying degrees of success. Common problems faced by previous studies included inhomogeneity in the data and an inadequate span of available datasets for performing robust analyses of long-term trends (WASA Group, 1998). As a result it could not be concluded with any degree of confidence whether the trends were decadal or climatic.

In this chapter, an analysis of the historical wave climate is presented. A long-term dataset extending between 1871 and 2010 was prepared by forcing the wave model set up in Chapter 4 with winds from the 20th Century Reanalysis (20CR) project (Compo *et al.*, 2011). This dataset is analysed for long-term trends and changes in significant wave height, mean energy period, mean wave direction and wave power density.

The following sections introduce methods of detection of climate change followed by the methodology applied to selected sites in the region. Finally, past and current states of the wave climate are compared at a regional level, and the results are presented in the form of maps of wave parameters.

5.2 Methods of trend detection and analysis

In the past, different methods have been used the analysis of climate and detection of trends in variables like temperature and precipitation including empirical methods and numerical modelling approaches using General Circulation Models (GCMs). In this section, popular empirical methods are described which may be applied to the data from WAVEWATCH III for the detection of trends and changes in the wave climate of the North East Atlantic and North Sea.

5.2.1 Climate normals

A climate normal, according to the recommendations of the World Meteorological Organization (WMO), can be simply defined as a thirty year moving average of a climatological variable, updated every decade (World Meteorological Organization, 1989). This metric has been used not only for predicting the future climate, but also to establish a background reference state which can be used for comparing the present and past climate.

There are five important characteristics of such a normal (Arguez and Vose, 2011):

- It is a temporal mean;
- It is unweighted;
- It is calculated over 30-year periods;
- It is a causal filter using past and present observations only; and
- It is updated decadally (at the discretion of member states).

The 30-year averaging period is based on a 'rule-of-thumb' with the expectation that statistics so calculated would capture the long-term characteristics of the climate without being biased by inter-annual and intra-annual variations. The implicit assumption here is that the climate is stationary over the 30-year period with random variation about the mean value. If this assumption were, indeed, invalid, the predictive properties of such a normal would come under question, however, its ability to represent the current (or past) state of the climate would remain intact (Arguez and Vose, 2011).

Several studies have questioned the use of a 30-year averaging period (e.g. Lamb and Changnon (1981); Dixon and Shulman (1984); Angel *et al.* (1993)) and have suggested alternate averaging periods, such as 10 years for temperature and 15 years for precipitation (Huang *et al.*, 1996), on the grounds of being able predict future climate better. These variations of the metric are sometimes referred to as Optimized Climate Normals (OCN) (Huang *et al.*, 1996). These normals, however, are susceptible to decadal variability and may not provide a robust estimate of the state of the parameter for climate change analyses.

There have been, as yet, no studies regarding the optimum averaging period for parameters describing the wave climate. The value of the climate normal, to this study, is in its ability to

describe the state of the climate and not in its predictive ability. Considering this, the standard 30-year period was used for calculating standard climate normals (referred to as the WMO normal hereafter) for analysis.

5.2.2 Least Squares Linear Fit

The least squares linear fit method is a simple method of trend analysis for auto-correlated time-series data of climatological variables. If the time-series can be represented as $y(t)$, the least square moments technique can be used to estimate a regression line for the variable as a function of time. This may be represented as (Polyak, 1996):

$$Y(t) = a + bt \tag{5.1}$$

where $Y(t)$ is the expected value of the variable at time t, a is the intercept and b is the slope. This type of function may be considered as a time-dependent climate normal.

Any trend that may exist in the time-series can be detected from the value of the slope:

- $b \geq$ indicates an increasing trend;
- $b \leq 0$ indicates a decreasing trend; and
- $b = 0$ indicates that the variable is invariant in time.

In the analysis of trends using a simplistic approach as this, it is important to ensure that the trend is significant. A quick way of doing this is by computing the p-value during the regression analysis and ensuring that the probability of random chance is less than 10%.

Despite its simplicity the linear fit analysis method was found to perform better than the 30-year climate normal as well as optimized climate normals (OCNs) in predicting short-term future climate (Livezey *et al.*, 2007). Owing to its ease of application and analysis, it has been used in several analyses of variation in significant wave height to time-series data of varying lengths (e.g. Carter and Draper (1988); WASA Group (1998); Lowe *et al.* (2009); Bertin *et al.* (2013)).

5.2.3 Hinge Fit

The hinge-fit analysis is a modification of the linear trend analysis in which the time-series is divided into segments and a piece-wise continuous linear regression function is fitted to the data. The use of such a non-uniform function originates to the late 1990s when the U.S. National Weather Service's Climate Prediction Center (CPC) devised it for the prediction of cold-season temperatures (Livezey *et al.*, 2007).

This type of analysis is particularly useful when the trend in the climatological parameter undergoes an abrupt change. For example, various climate-change related analyses of long-term data of surface temperatures have indicated that there is little or no trend visible between 1940 and the mid 1970s. It is circa 1975 that an upward trend in mean temperatures can be observed

(Blunden and Arndt, 2013). It is in such situations that the hinge-fit method outperforms other approaches for trend detection and analysis (Livezey *et al.*, 2007).

For the time-series of the climate variable represented by $y(t)$, the hinge fit can be modelled by:

$$Y(t) = \begin{cases} a_1 + b_1 t & \text{for } t \leq T \\ a_2 + b_2 t & \text{for } t > T \end{cases} \tag{5.2}$$

where $Y(t)$ is the expected value of the variable, a_1 and a_2 are the intercepts, b_1 and b_2 are the slopes for the cases and T is the temporal position of the hinge.

Given the current context of climate change, it has been argued that the hinge-fit approach is a better time-varying representation of the base climate. In addition, by the use of hinges this method is able to capture changes in the state of the climate as well as the period during which the change occurred.

5.3 Site Selection for analysis

There are, as described in the previous section, different methods of fitting and analysing trends in climatological parameters. Given the size of the regional dataset of historical wave conditions it is not feasible to apply all the methods for the entire domain. The impracticality of this is magnified when the trend detection methods need to be applied to annual and seasonal data for multiple wave parameters. It is therefore necessary to identify, from among the available trend analysis methods, the most appropriate approach. This can be achieved by testing and evaluating the methods for selected locations.

Eight sites of importance to the marine energy sector were identified for analysis which include test sites for wave energy converters as well as proposed locations for utility-scale wave energy projects. These are listed in Table 5.1. As the wave power density is greatest towards the north and west of the UK and Ireland, it is not surprising that most of the sites selected lie in that region.

Site	Latitude	Longitude	Description
Farr Point	58.50°N	4.25°W	50MW wave energy project
EMEC	59.00°N	3.50°W	Wave energy test site
West of Hebrides	57.25°N	8.00°W	10MW wave energy project
Galway Bay	53.25°N	9.25°W	$^1/_4$-scale wave energy test site
Belmullet	54.25°N	10.25°W	Wave energy test site
Beatrice	58.00°N	1.50°W	North Sea wind farm
Wave Hub	50.25°N	5.50°W	Wave energy test site
Burghead	58.00°N	3.25°W	Wave energy project (in planning)

Table 5.1: Locations selected for site-level analyses of wave climate

The wave climate in the region under study has been known to exhibit a significant spatial variation in the wave climate. For example, wave conditions in the North East Atlantic can be modelled using a generalised Pierson-Moskowitz (PM) (Cahill and Lewis, 2013) while the wave climate in the North Sea is best described by the JONSWAP spectrum (Hasselmann *et al.*, 1973). Moreover, it is known that the lower latitudes experience less energetic waves than higher latitudes. Taking these into account, sites were selected to allow the analysis of a diverse range of locations.

5.4 Site level analysis of wave model output

5.4.1 Data used

The calibrated wave model was forced with winds from the 20CR dataset and wave conditions in the region were simulated for the period 1871-2010. Location specific data was extracted from the 0.25°x0.25° gridded, hourly hindcast dataset thus obtained for use in this analysis. The time-series for significant wave height (H_s), energy period (T_e) and mean wave direction (D_m) spanning from January 1, 1871 to December 31 2010 were processed and the following quantities were derived for the analysis of long-term trends:

- Annual means of wave parameters (H_s, T_e, and D_m);
- 95th percentiles of annual significant wave height (H_{95});
- Annual maxima (H_{max}); and
- Annual mean wave power density ($P = 0.49 H_s^2 T_e$).

5.4.2 Trend fitting

WMO normal

The climate normal defined by the WMO has been calculated surface parameters such as temperature and precipitation in the past establishing its utility in describing the state of the climate. The standard WMO normals for describing the past and current climate have been calculated for the periods 1901-1930, 1931-1960, and 1961-1990. The next standard normal will be calculated for 1991-2020.

Given the extent of the available data in this study and the proximity of this study to 2020, the WMO normals used for describing the state of the climate are calculated for 30-year periods beginning in 1871, updated every 10 years. By this method, the climate normal is calculated for several non-mandatory periods as well and the state of the current climate can be assumed to be represented by the normal calculated for 1981-2010.

Linear Fit

A simple linear trend function, of the type described by Equation 5.1, was fitted to the time-series of wave climate data (described in Section 5.3) by estimating the slope and intercept using the least-squares method. The time-series of the expected value of the parameter was estimated using the fitted function for all 8 locations.

Hinge Fit

Empirical studies of climate change have suggested that the most recent period of global warming began in the mid 1970s (Hartmann *et al.*, 2013; Blunden and Arndt, 2013). Based on this information, long-term analyses of local and global near-surface temperatures which have used the hinge fit method divided the time-series at 1975 (Livezey *et al.*, 2007).

For the analysis of wave parameters, the year 1975 cannot be used without verification as no causal link has been established, as yet, between global temperature trends and the wave climate. Before fitting a hinge type function to the data, the optimal location of the hinge needs to be identified.

An optimally located hinge ensures that the trends on either side provide the best representation of the state of the climate at any time. The following procedure was applied to the time-series data of annual means of significant wave height (H_s), energy period (T_e) and mean wave direction (D_m) for the identification of the optimal position of the hinge:

1. Based on the accepted notion that statistics computed over 30 years are unaffected by short-term interannual variations, the limiting positions of the hinge were set at 1900 and 1980. These limits ensure that at any position of the hinge, at least 30 years of data is available for fitting the trend functions.
2. Hinge type functions, as described by Equation 5.2, are fitted to the data divided at the break point using the least-squares method for estimation of the slope and intercept. The time-series of expected values according to the fitted functions $Y(t)$ is estimated for the period 1871-2010.
3. The bias between the expected and observed data is calculated for the entire 140-year period.
4. The standard deviation (σ) of the bias is taken as a measure of how well the hinge fit function represents the state of the climate. A lower value of σ would indicate a better representation of the climate than a higher value of σ.
5. The above steps are repeated for hinges positioned in steps of 5 years between the identified limits (1900 to 1980).
6. For any site, the values of σ calculated for all possible hinge positions are normalised by dividing it by the maximum calculated value for the site

$$\widehat{\sigma}_t = \frac{\sigma_t}{\sigma_{\max}} \tag{5.3}$$

where $\widehat{\sigma}_t$ is the normalised standard deviation for the function fitted with the hinge at year t, σ_t is the standard deviation for the function and σ_{max} is the maximum standard deviation calculated for the site.

The procedure described above was repeated for 8 test locations and 3 wave parameters. The raw and normalised standard deviations are tabulated in Tables 5.2 to 5.7. At every tested hinge position, the normalised standard deviations were summed across sites for H_s, T_e and D_m to allow easy comparison of data for the selection of the optimal hinge position. These data are tabulated in Table 5.8 and presented visually in Figure 5.1.

Upon examining the variation of the cumulative normalised standard deviations with the position of the hinge in Figure 5.1, distinct minima can be observed at 1920 for the energy period and mean wave direction. Although the plot for significant wave height does not show much variation for the hinge located between 1905 and 1935, it can be observed that the minima lies between 1915 and 1930. This is confirmed from the data in Table 5.8 where it can be seen that $\Sigma\sigma_{Hs}$ is minimum for the hinge located at 1925.

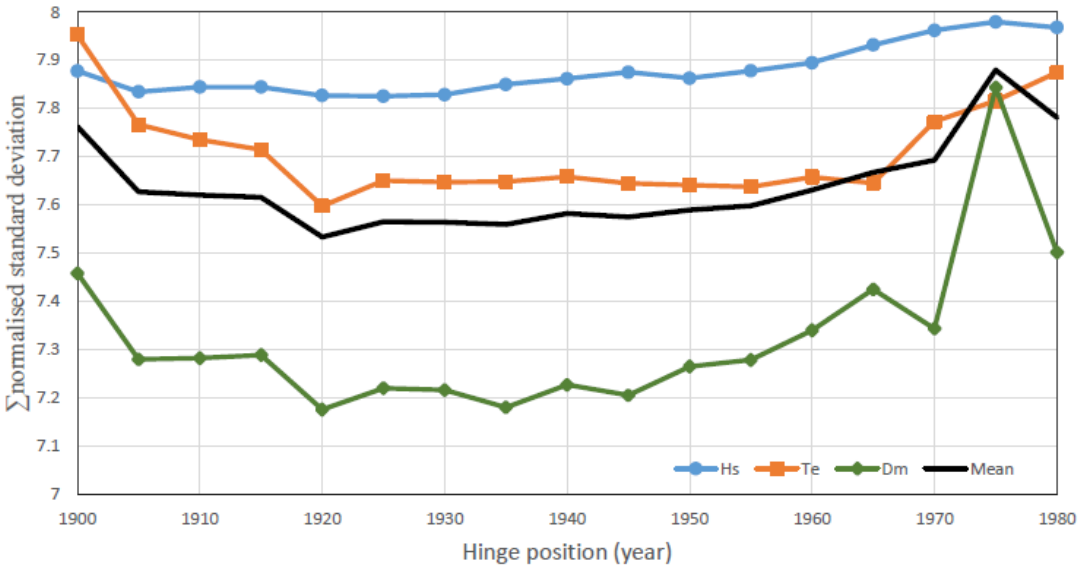

Figure 5.1: Average and individual cumulative normalised standard deviation ($\Sigma\widehat{\sigma}$) for wave parameters parameters plotted for different positions of the hinge.

Based on these and after examination of the plot of the average score across all three parameters the 1920 was chosen as the optimal position of the hinge for further analysis. A hinge fit function was fit to the data as:

$$Y(t) = \begin{cases} a_1 + b_1 t & \text{for } t \leq 1920 \\ a_2 + b_2 t & \text{for } t > 1920 \end{cases} \tag{5.4}$$

The parameters a and b for the cases in Equation 5.4 were estimated using the least-squares

Hinge	Farr	EMEC	W. Hebrides	Galway Bay	Belmullet	Beatrice	Wave Hub	Burghead
1900	0.1696	0.2138	0.2372	0.0992	0.2200	0.0863	0.1859	0.0605
1905	0.1696	0.2136	0.2363	0.0983	0.2176	0.0861	0.1833	0.0603
1910	0.1699	0.2140	0.2365	0.0981	0.2175	0.0865	0.1824	0.0608
1915	0.1697	0.2141	0.2361	0.0982	0.2169	0.0865	0.1838	0.0606
1920	0.1691	0.2135	0.2349	0.0978	0.2155	0.0867	0.1832	0.0609
1925	0.1693	0.2137	0.2352	0.0978	0.2158	0.0865	0.1831	0.0607
1930	0.1690	0.2139	0.2358	0.0977	0.2162	0.0869	0.1808	0.0612
1935	0.1685	0.2135	0.2360	0.0982	0.2170	0.0872	0.1821	0.0616
1940	0.1666	0.2121	0.2358	0.0990	0.2183	0.0873	0.1851	0.0616
1945	0.1670	0.2123	0.2353	0.0990	0.2180	0.0873	0.1877	0.0616
1950	0.1666	0.2120	0.2349	0.0986	0.2173	0.0875	0.1870	0.0617
1955	0.1665	0.2119	0.2353	0.0991	0.2180	0.0875	0.1882	0.0617
1960	0.1674	0.2129	0.2366	0.0995	0.2194	0.0871	0.1877	0.0616
1965	0.1680	0.2135	0.2379	0.0998	0.2208	0.0880	0.1876	0.0620
1970	0.1681	0.2127	0.2387	0.1007	0.2231	0.0876	0.1908	0.0619
1975	0.1681	0.2128	0.2385	0.1006	0.2233	0.0888	0.1914	0.0620
1980	0.1677	0.2120	0.2380	0.1003	0.2233	0.0887	0.1912	0.0621

Table 5.2: Standard deviation of bias between observed and expected mean annual significant wave height (σ_{Hs}) for different hinge points

Hinge	Farr	EMEC	W. Hebrides	Galway Bay	Belmullet	Beatrice	Wave Hub	Burghead
1900	0.9982	0.9986	0.9937	0.9851	0.9852	0.9718	0.9713	0.9742
1905	0.9982	0.9977	0.9899	0.9762	0.9745	0.9696	0.9577	0.9710
1910	1.0000	0.9995	0.9908	0.9742	0.9740	0.9741	0.9530	0.9791
1915	0.9988	1.0000	0.9891	0.9752	0.9713	0.9741	0.9603	0.9758
1920	0.9953	0.9972	0.9841	0.9712	0.9651	0.9764	0.9572	0.9807
1925	0.9965	0.9981	0.9853	0.9712	0.9664	0.9741	0.9566	0.9775
1930	0.9947	0.9991	0.9879	0.9702	0.9682	0.9786	0.9446	0.9855
1935	0.9918	0.9972	0.9887	0.9752	0.9718	0.9820	0.9514	0.9919
1940	0.9806	0.9907	0.9879	0.9831	0.9776	0.9831	0.9671	0.9919
1945	0.9829	0.9916	0.9858	0.9831	0.9763	0.9831	0.9807	0.9919
1950	0.9806	0.9902	0.9841	0.9791	0.9731	0.9854	0.9770	0.9936
1955	0.9800	0.9897	0.9858	0.9841	0.9763	0.9854	0.9833	0.9936
1960	0.9853	0.9944	0.9912	0.9881	0.9825	0.9809	0.9807	0.9919
1965	0.9888	0.9972	0.9966	0.9911	0.9888	0.9910	0.9801	0.9984
1970	0.9894	0.9935	1.0000	1.0000	0.9991	0.9865	0.9969	0.9968
1975	0.9894	0.9939	0.9992	0.9990	1.0000	1.0000	1.0000	0.9984
1980	0.9871	0.9902	0.9971	0.9960	1.0000	0.9989	0.9990	1.0000

Table 5.3: Normalised standard deviation of bias between observed and expected mean annual significant wave height ($\widehat{\sigma}_{\mathrm{Hs}}$) for different hinge points

Hinge	Farr	EMEC	W. Hebrides	Galway Bay	Belmullet	Beatrice	Wave Hub	Burghead
1900	0.2885	0.2691	0.2395	0.3133	0.2570	0.3509	0.3046	0.3618
1905	0.2753	0.2694	0.2317	0.3053	0.2484	0.3448	0.2989	0.3568
1910	0.2767	0.2711	0.2312	0.3047	0.2469	0.3415	0.2963	0.3513
1915	0.2762	0.2711	0.2321	0.3066	0.2470	0.3352	0.2977	0.3454
1920	0.2697	0.2712	0.2274	0.2992	0.2389	0.3361	0.2923	0.3436
1925	0.2718	0.2712	0.2289	0.3019	0.2414	0.3385	0.2941	0.3466
1930	0.2700	0.2712	0.2290	0.3021	0.2415	0.3394	0.2933	0.3473
1935	0.2676	0.2707	0.2290	0.3022	0.2411	0.3407	0.2953	0.3480
1940	0.2656	0.2683	0.2298	0.3031	0.2418	0.3428	0.2978	0.3491
1945	0.2649	0.2688	0.2288	0.3030	0.2407	0.3426	0.2969	0.3488
1950	0.2645	0.2694	0.2285	0.3024	0.2401	0.3427	0.2966	0.3494
1955	0.2646	0.2698	0.2284	0.3017	0.2399	0.3421	0.2968	0.3490
1960	0.2653	0.2719	0.2285	0.3005	0.2403	0.3449	0.2954	0.3524
1965	0.2670	0.2717	0.2295	0.3021	0.2416	0.3366	0.2963	0.3481
1970	0.2698	0.2687	0.2326	0.3066	0.2470	0.3472	0.3051	0.3561
1975	0.2696	0.2706	0.2328	0.3071	0.2478	0.3528	0.3066	0.3601
1980	0.2718	0.2719	0.2338	0.3082	0.2500	0.3574	0.3068	0.3661

Table 5.4: Standard deviation of bias between observed and expected mean annual energy period (σ_{Te}) for different hinge points

Hinge	Farr	EMEC	W. Hebrides	Galway Bay	Belmullet	Beatrice	Wave Hub	Burghead
1900	1.0000	0.9897	1.0000	1.0000	1.0000	0.9818	0.9928	0.9883
1905	0.9542	0.9908	0.9674	0.9745	0.9665	0.9647	0.9743	0.9746
1910	0.9591	0.9971	0.9653	0.9726	0.9607	0.9555	0.9658	0.9596
1915	0.9574	0.9971	0.9691	0.9786	0.9611	0.9379	0.9703	0.9435
1920	0.9348	0.9974	0.9495	0.9550	0.9296	0.9404	0.9527	0.9385
1925	0.9421	0.9974	0.9557	0.9636	0.9393	0.9471	0.9586	0.9467
1930	0.9359	0.9974	0.9562	0.9643	0.9397	0.9496	0.9560	0.9486
1935	0.9276	0.9956	0.9562	0.9646	0.9381	0.9533	0.9625	0.9506
1940	0.9206	0.9868	0.9595	0.9674	0.9409	0.9591	0.9707	0.9536
1945	0.9182	0.9886	0.9553	0.9671	0.9366	0.9586	0.9677	0.9527
1950	0.9168	0.9908	0.9541	0.9652	0.9342	0.9589	0.9668	0.9544
1955	0.9172	0.9923	0.9537	0.9630	0.9335	0.9572	0.9674	0.9533
1960	0.9196	1.0000	0.9541	0.9591	0.9350	0.9650	0.9628	0.9626
1965	0.9255	0.9993	0.9582	0.9643	0.9401	0.9418	0.9658	0.9508
1970	0.9352	0.9882	0.9712	0.9786	0.9611	0.9715	0.9945	0.9727
1975	0.9345	0.9952	0.9720	0.9802	0.9642	0.9871	0.9993	0.9836
1980	0.9421	1.0000	0.9762	0.9837	0.9728	1.0000	1.0000	1.0000

Table 5.5: Normalised standard deviation of bias between observed and expected mean annual energy period ($\widehat{\sigma}_{\mathrm{Te}}$) for different hinge points

Hinge	Farr	EMEC	W. Hebrides	Galway Bay	Belmullet	Beatrice	Wave Hub	Burghead
1900	9.5320	8.4783	4.8617	4.4671	5.9535	10.4090	6.6045	9.7104
1905	9.2150	8.0942	4.6500	4.3618	5.7015	10.3997	6.5354	9.7203
1910	9.1975	7.9664	4.5744	4.4056	5.7079	10.3599	6.6813	9.7404
1915	9.9582	7.7569	4.5688	4.4028	5.6304	10.1559	6.6495	9.6730
1920	9.0167	7.5978	4.5834	4.4001	5.6180	10.1495	6.6746	9.6568
1925	9.1331	7.7734	4.6370	4.3748	5.6585	10.2428	6.6074	9.6969
1930	9.1135	7.7359	4.6375	4.4029	5.6664	10.1504	6.6475	9.6833
1935	9.1098	7.2418	4.6398	4.4225	5.6759	10.1288	6.7867	9.6744
1940	9.1204	7.5405	4.6399	4.4556	5.6797	10.0640	6.8723	9.6423
1945	9.1073	7.7581	4.6199	4.3465	5.6228	10.0598	6.7978	9.6448
1950	9.1160	7.8155	4.6560	4.4905	5.7193	10.0240	6.8313	9.6369
1955	9.1777	7.8481	4.6398	4.5095	5.7206	10.0302	6.8452	9.6278
1960	9.2756	7.9869	4.6832	4.5376	5.7297	10.0687	6.9506	9.6613
1965	9.3297	8.0358	4.7509	4.6163	5.8677	10.0864	7.1165	9.6630
1970	9.3894	7.1789	4.7736	4.6264	5.9015	10.0851	7.1587	9.6366
1975	9.3423	8.1900	7.7575	4.6148	5.9078	10.0683	7.1784	9.6448
1980	9.4047	8.3237	4.8112	4.6370	5.9870	10.1121	7.1338	9.6464

Table 5.6: Standard deviation of bias between observed and expected annual mean wave direction (σ_{Dm}) for different hinge points

Hinge	Farr	EMEC	W. Hebrides	Galway Bay	Belmullet	Beatrice	Wave Hub	Burghead
1900	0.9572	1.0000	0.6267	0.9634	0.9944	1.0000	0.9201	0.9969
1905	0.9254	0.9547	0.5994	0.9407	0.9523	0.9991	0.9104	0.9979
1910	0.9236	0.9396	0.5897	0.9501	0.9534	0.9953	0.9308	1.0000
1915	1.0000	0.9149	0.5890	0.9495	0.9404	0.9757	0.9263	0.9931
1920	0.9055	0.8961	0.5908	0.9489	0.9384	0.9751	0.9298	0.9914
1925	0.9171	0.9169	0.5977	0.9435	0.9451	0.9840	0.9205	0.9955
1930	0.9152	0.9124	0.5978	0.9495	0.9465	0.9752	0.9260	0.9941
1935	0.9148	0.8542	0.5981	0.9537	0.9480	0.9731	0.9454	0.9932
1940	0.9159	0.8894	0.5981	0.9609	0.9487	0.9669	0.9574	0.9899
1945	0.9146	0.9151	0.5955	0.9374	0.9392	0.9665	0.9470	0.9902
1950	0.9154	0.9218	0.6002	0.9684	0.9553	0.9630	0.9516	0.9894
1955	0.9216	0.9257	0.5981	0.9725	0.9555	0.9636	0.9536	0.9884
1960	0.9315	0.9420	0.6037	0.9786	0.9570	0.9673	0.9683	0.9919
1965	0.9369	0.9478	0.6124	0.9955	0.9801	0.9690	0.9914	0.9921
1970	0.9429	0.8467	0.6154	0.9977	0.9857	0.9689	0.9973	0.9893
1975	0.9382	0.9660	1.0000	0.9952	0.9868	0.9673	1.0000	0.9902
1980	0.9444	0.9818	0.6202	1.0000	1.0000	0.9715	0.9938	0.9903

Table 5.7: Normalised standard deviation of bias between observed and expected annual mean wave direction ($\widehat{\sigma}_{\mathrm{Dm}}$) for different hinge points

Hinge	$\Sigma\widehat{\sigma}_{\mathrm{Hs}}$	$\Sigma\widehat{\sigma}_{\mathrm{Te}}$	$\Sigma\widehat{\sigma}_{\mathrm{Dm}}$
1900	7.8781	7.9526	7.4587
1905	7.8348	7.7670	7.2799
1910	7.8447	7.7357	7.2825
1915	7.8446	7.7150	7.2889
1920	7.8272	7.5979	7.1760
1925	7.8257	7.6505	7.2203
1930	7.8288	7.6477	7.2167
1935	7.8500	7.6485	7.1805
1940	7.8620	7.6586	7.2272
1945	7.8754	7.6448	7.2055
1950	7.8631	7.6412	7.2651
1955	7.8782	7.6376	7.2790
1960	7.8950	7.6582	7.3403
1965	7.9320	7.6458	7.4252
1970	7.9622	7.7730	7.3439
1975	7.9799	7.8161	7.8437
1980	7.9683	7.8748	7.5020

Table 5.8: Cumulative normalised standard deviation $\Sigma\widehat{\sigma}$ for wave parameters at different positions of the hinge

method, fitting regression lines to each segement of the time-series. This process was repeated for all eight locations.

5.5 Evaluation of trend analysis methods

In order to compare the performance of the different trend analysis methods fitted in Section 5.4.2, the bias between the expected and observed climatic parameters was analysed. For the WMO normal approach, the bias was calculated for each 30-year block as the difference between the observed value and the climate normal. For the linear fit function (as represented by Equation 5.1) and the hinge fit function (as represented by Equation 5.4), the bias was calculated using:

$$e(t) = Y(t) - y(t) \tag{5.5}$$

where $e(t)$ is the bias, $Y(t)$ is the expected value and $y(t)$ is the observed value at time step t.

The sample of bias was calculated for the annual means of H_s, T_e and D_m for the 8 selected locations and the standard deviation (σ) was calculated for each sample. These are represented visually in Figures 5.2 to 5.4, and the data behind the plots is presented in Tables 5.9 to 5.11.

It is evident from an examination of the standard deviations plotted for the three wave parameters that the linear trend, fitted to the entire time-series, is the worst performing method, indicated by the standard deviation being the highest. This is not surprising, considering that

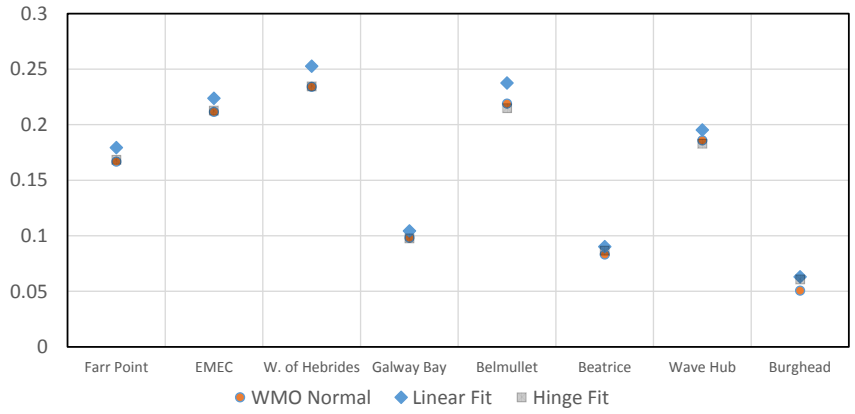

Figure 5.2: Standard deviation of bias (m) calculated for significant wave height for the different trend analysis methods

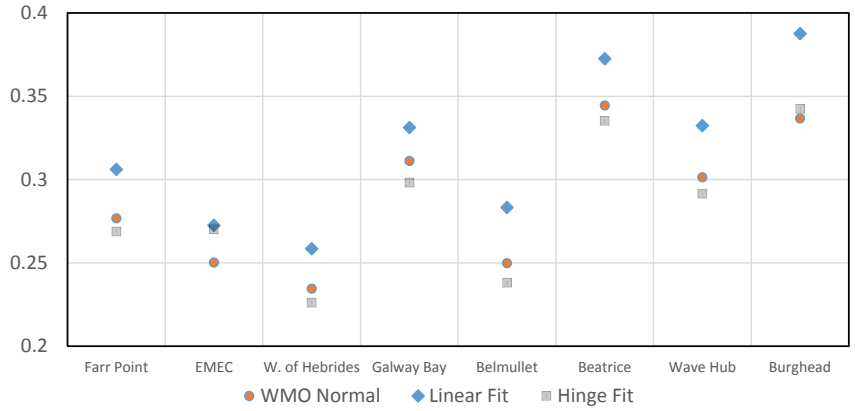

Figure 5.3: Standard deviation of bias (s) calculated mean energy period for the different trend analysis methods

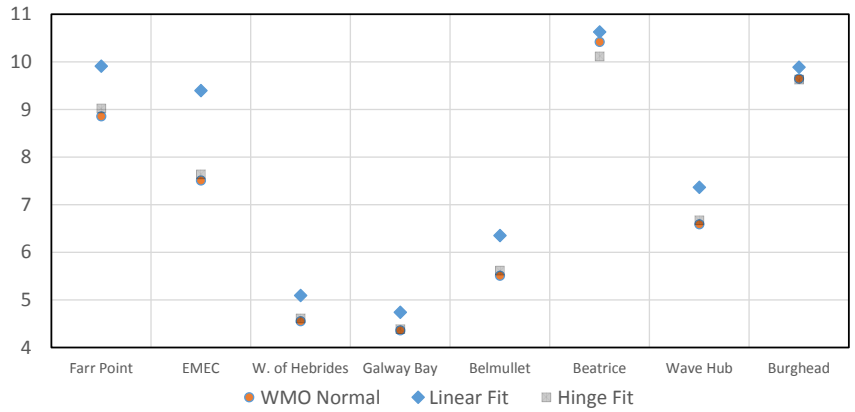

Figure 5.4: Standard deviation of bias (deg.) calculated for mean wave direction for the different trend analysis methods

Site	WMO Normal	Linear Fit	Hinge Fit
Farr Point	0.1668	0.1794	0.1684
EMEC	0.2115	0.2239	0.2127
W. of Hebrides	0.2341	0.2528	0.2345
Galway Bay	0.0983	0.1044	0.0977
Belmullet	0.2190	0.2375	0.2151
Beatrice	0.0831	0.0902	0.0867
Wave Hub	0.1859	0.1954	0.1829
Burghead	0.0506	0.0631	0.0609

Table 5.9: Standard deviation of bias (m) between observed and expected mean annual H_s for the different trend fitting methods

Site	WMO Normal	Linear Fit	Hinge Fit
Farr Point	0.2768	0.3061	0.2689
EMEC	0.2502	0.2726	0.2703
W. of Hebrides	0.2345	0.2585	0.2261
Galway Bay	0.3112	0.3312	0.2982
Belmullet	0.2499	0.2833	0.2381
Beatrice	0.3445	0.3726	0.3352
Wave Hub	0.3014	0.3324	0.2915
Burghead	0.3367	0.3876	0.3425

Table 5.10: Standard deviation of bias (s) between observed and expected mean annual T_e for the different trend fitting methods

Site	WMO Normal	Linear Fit	Hinge Fit
Farr Point	8.8543	9.9117	9.0222
EMEC	7.5094	9.3974	7.6398
W. of Hebrides	4.5548	5.0912	4.6103
Galway Bay	4.3588	4.7387	4.3874
Belmullet	5.5086	6.3523	5.6151
Beatrice	10.4201	10.6293	10.1159
Wave Hub	6.5898	7.3654	6.6755
Burghead	9.6482	9.8885	9.6286

Table 5.11: Standard deviation of bias (deg.) between observed and expected mean annual D_m for the different trend fitting methods

the samples of wave parameters studied spans 140 years, and that a change in the trend occurs circa 1920 as established earlier. In such a situation, a simple linear fit is bound to be inadequate at representing the varying state of the climate.

From the plots of standard deviations for H_s and D_m presented in Figures 5.2 and 5.4, it can be seen that the markers for the WMO normal and the hinge fit are extremely close to one another, even overlapping in several cases. It is only from an examination of the data tabulated in Tables 5.9 and 5.11 that the better method can be identified. For these two wave parameters, neither trend analysis method is consistently better than the other, with the performance varying from site to site and parameter to parameter. In the case of T_e the hinge fit approach was found to be superior to the other methods for majority of the test locations, as evidenced by Figure 5.4.

Based on this comparison of the three trend analysis methods facilitated by the standard deviation of bias, the hinge fit approach was chosen as being the most appropriate for detecting and quantifying trends in the variation of the wave climate.

5.6 Statistical significance of trends

The hinge-fit trend analysis method identified in the previous section is a linear regression function fitting the observed parameter values to time. It is highly improbable that the parameter is, in fact, a direct function of time and as a result, it is highly unlikely that the p-values obtained from this type of regression analysis would indicate that the observed trends are statistically significant. A statistically more robust approach is therefore required to establish that the evolution of the wave climate over time is statistically significant and not a result of the internal and short-term variability in the system.

In an attempt to establish the statistical significance of the trends, the gridded data used in the previous section were subjected to the Student's t-test. This test is used to determine whether two samples are statistically different from each other. The null hypothesis, therefore, is that both the samples have the same mean (or that they are both drawn from the same population). This test assumes that both samples being compared are normally distributed. This is a reasonable assumption in the case of annually averaged wave parameters.

For two samples, X_1 and X_2 with n elements each, The test statistic, t, can be calculated as (Cramér, 1946)

$$t = \frac{\left|\overline{X}_1 - \overline{X}_2\right|}{\sigma_{X_1X_2} \cdot \sqrt{\frac{2}{n}}} \tag{5.6}$$

$$\sigma_{X_1X_2} = \sqrt{\frac{1}{2}(\sigma_{X_1}^2 + \sigma_{X_2}^2)} \tag{5.7}$$

where $\overline{X}_1$ and $\overline{X}_2$ are the means of the two samples and σ_{X_1} and σ_{X_2} are the standard deviations of the samples.

The test statistic follows Student's t-distribution, and its critical value for $2n-2$ degrees of freedom at a significance level α can be obtained from tables or computed from the distribution. If the calculated value of t from Equations 5.6 and 5.7 is greater than the critical value, the null hypothesis can be rejected with the conclusion that the compared samples are statistically different.

Alternatively, existing software such as 't.test()' for the statistical language 'R' automatically yield the p-value from the test allowing for direct comparison with the selected significance level α. The null hypothesis can be rejected if the p-value is less than α.

To establish statistical significance of the time evolution of the wave climate, two samples of the annually averaged wave parameters were drawn for comparison. The first sample comprised of data from 1921-1950 and the second comprised of data from 1981-2010. The selection of blocks in the time-series was done so that both blocks lie on the same side of the hinge, i.e. forwards in time from 1921. Also, they are sufficiently apart from each other for determining whether or not the wave climate is statistically different during the two periods considered.

A 30-year period was chosen for each sample following the recommendations of the World Meteorological Organisation (WMO) (World Meteorological Organization, 1989) according to which it is expected that statistics obtained from this period (means and standard deviations) are likely to be an adequate representation of the base state of the climate free from the effects of short-term variability.

A significance level of $\alpha = 0.10$, corresponding with 90% confidence in the outcome was chosen for rejecting the null hypothesis. If the t-test indicates that the samples are statistically different from one another, it can be inferred that the wave climate has evolved over time.

5.7 Trends in the historical wave climate

For the parameters identified in Section 5.4.1 at the 8 wave energy sites, the prepared time-series data were divided into two segments at the year 1920. A hinge fit type function was fitted to the data for the detection of any trends in the long-term wave climate. This process was applied to data for the selected test locations and also applied iteratively to the gridded data for the North East Atlantic and North Sea. The results of the analysis are presented and discussed here.

In the interests of brevity, only plots for the sites located West of Hebrides (North East Atlantic) and at Burghead (Moray Firth) are presented in this section. Plots for the remaining sites are presented in Appendix B.

5.7.1 Climatic changes in means of wave parameters

In this study, the wave parameters significant wave height (H_s), energy averaged wave period (T_e) and the mean wave direction (D_m) have been used to describe wave conditions. The average prevailing wave climate of the region can be represented using seasonal and annual means of these wave parameters.

The description of the 'normal' or 'mean' state of the climate is estimated from the hinge fit functions fitted to the time-series of seasonal and annual means of these wave parameters. The 'normal' state of the annual mean wave climate in 1921 and 2010 for the North East Atlantic and North Sea obtained from the hinge-fit method are presented as contour maps in Figures 5.5 to 5.7.

As expected, the p-values calculated directly from the trend analysis methods were found to be larger than the significance level (set at $\alpha = 0.10$ or 10%) for this study. Based on these, the null hypothesis in the analysis, which is that no temporal trend exists in the data cannot be rejected based on the p-values obtained. More conclusive evidence was obtained by subjecting the data to the Student's t-test, as described in Section 5.6 the results of which are mapped in Figures 5.11 to 5.13

It should be noted that it can be said with confidence that the wave climate has changed between 1921 and 2010 in a statistically significant manner in the blue regions in the significance plots in Figures 5.11 to 5.13. In the rest of the region, indicated by the red shading, this conclusion cannot be reached. It may be that the wave parameters in this region have evolved over the 90-year period, as indicated by the trends. However, it is more likely that the computed trends are a result of large interannual variations in the parameters.

The contour plots in Figures 5.5 and 5.6 reveal that over the 90 year period studied, the annual means of significant wave height and energy period have changed noticeably in the North East Atlantic. In the North Sea, however, these parameters exhibit little change. The 'normal' mean wave direction in 2010 appears to be almost identical to the 'normal' state in 1921 in Figure 5.7, showing only small changes in the eastern part of the North Sea.

It is difficult to assess the magnitude of the change, if any, from these plots. For this reason, the absolute and relative year-on-year changes in the annual means are presented for the region in Figures 5.8 to 5.10. In addition, the seasonal means along with the hinge-fit trend is presented for individual locations in Figures 5.14 to 5.16 and in Appendix B.

From the regional plots of mean annual significant wave height in Figure 5.8, it can be seen the magnitude of the change is greatest in the northwestern and western parts of the region where the absolute change ranges from 0.25 to 0.45cm-year^{-1}. Relative to the mean wave height in 1921, this translates to a change of $\approx$ 0.08 to 0.14% per year in the same region. The slope of the trend in the North East Atlantic is found to decrease southwards in both, absolute and relative terms.

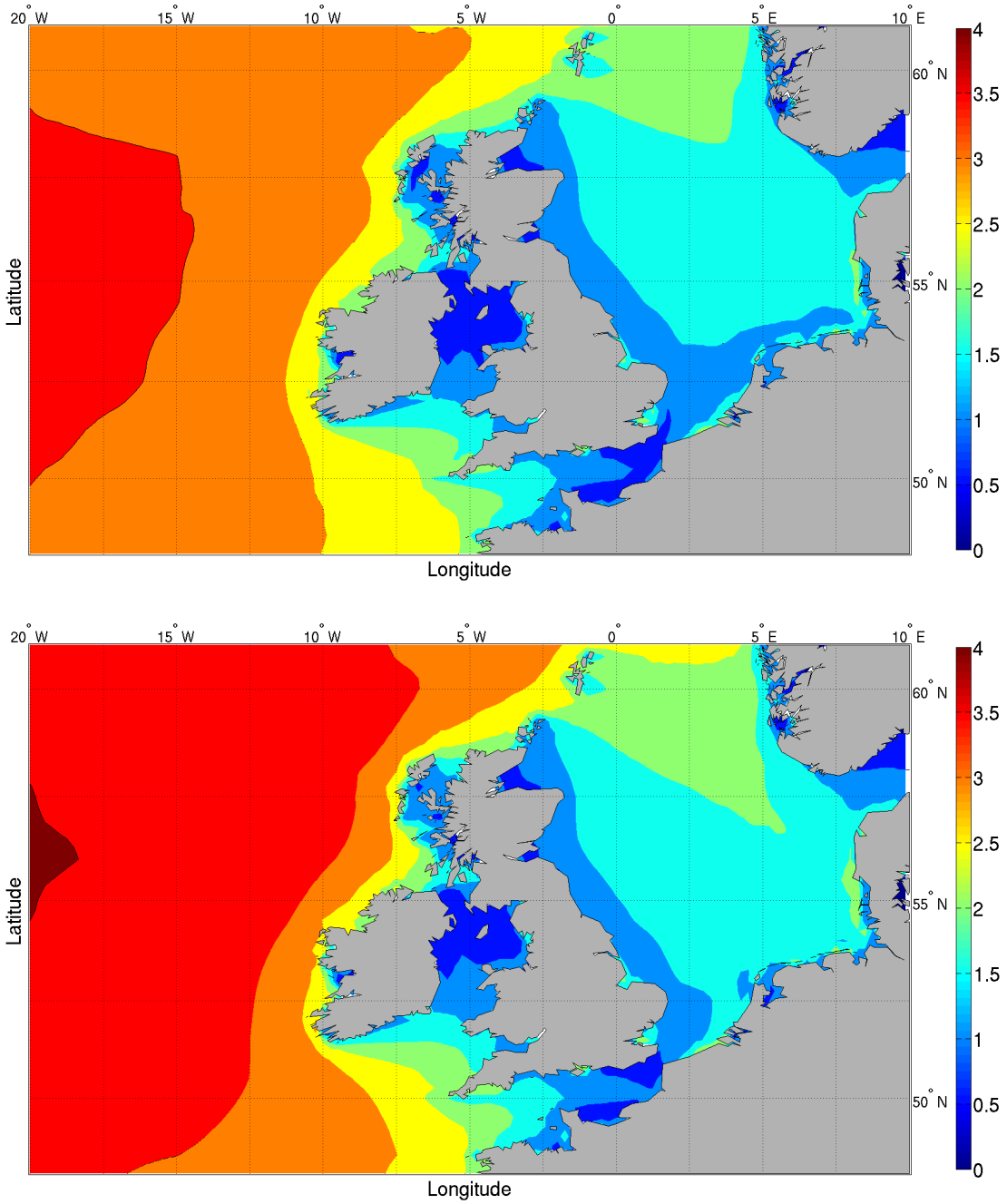

Figure 5.5: 'Normal' annual mean H_s (m) in 1921 (top) and 2010 (bottom) as estimated by the hinge fit function

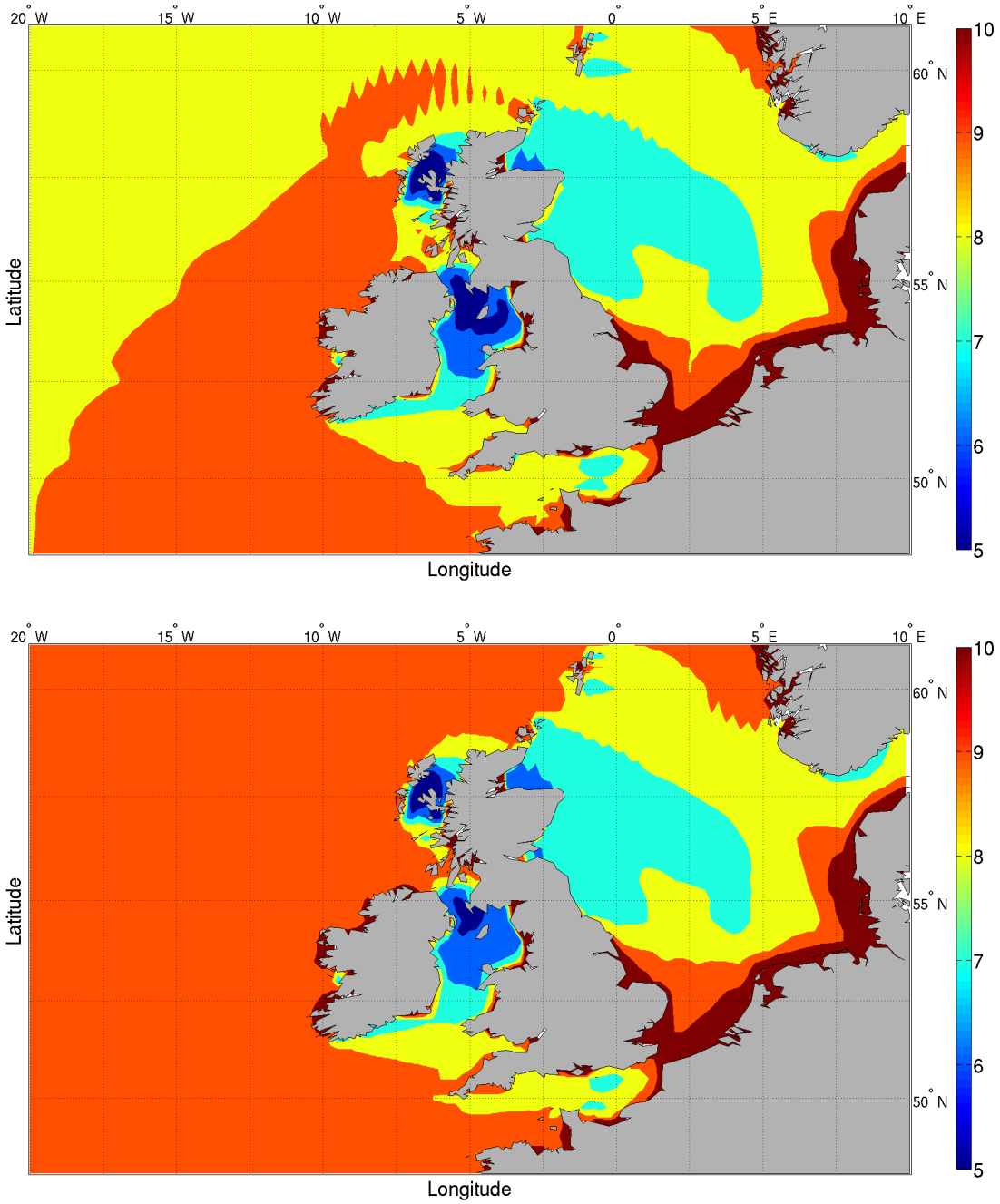

Figure 5.6: 'Normal' annual mean T_e (s) in 1921 (top) and 2010 (bottom) as estimated by the hinge fit function

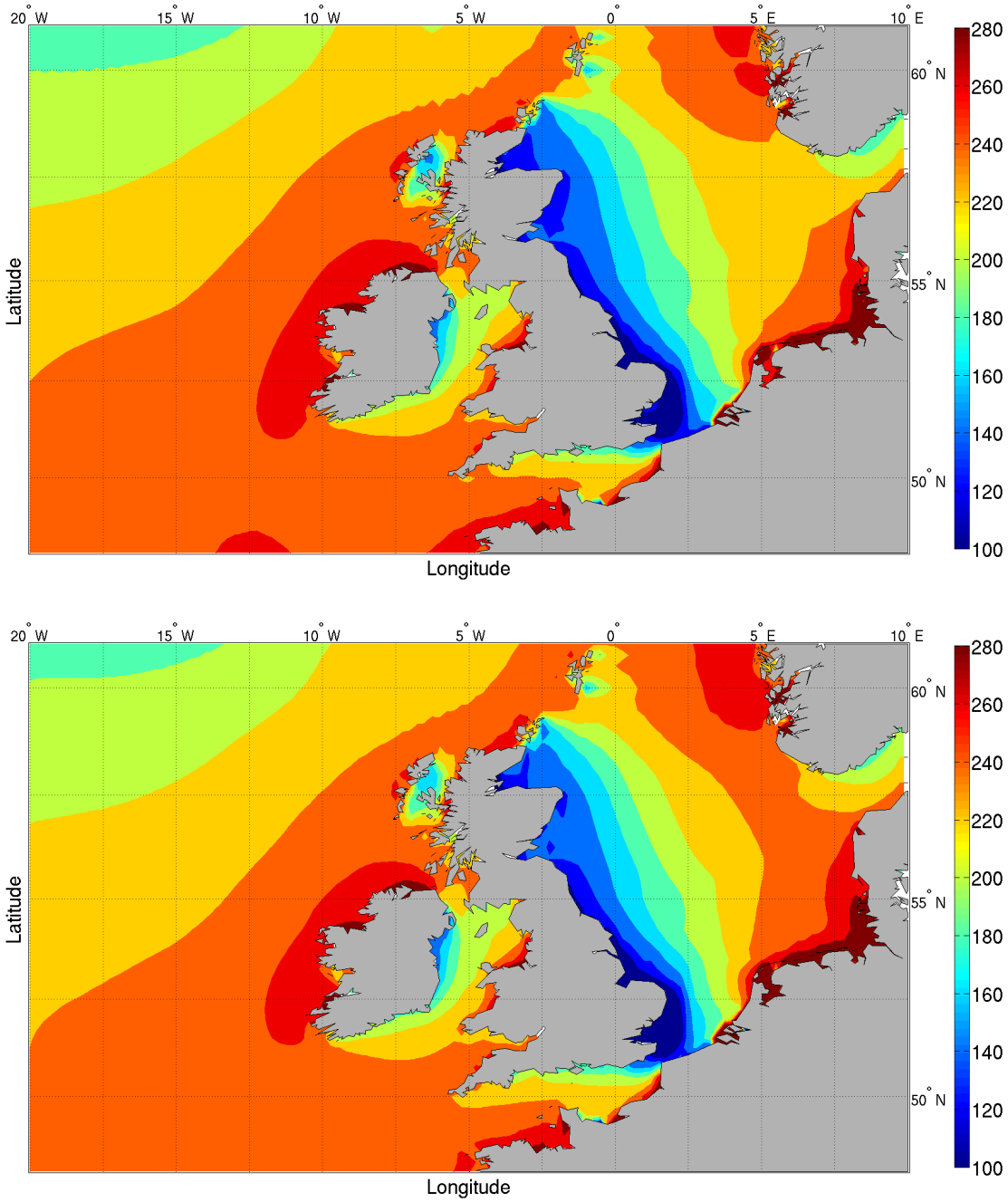

Figure 5.7: 'Normal' annual mean D_m (deg.) in 1921 (top) and 2010 (bottom) as estimated by the hinge fit function

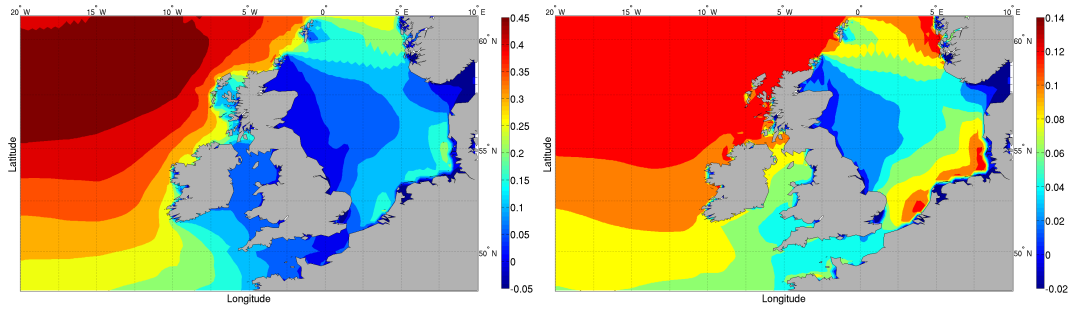

Figure 5.8: Changes in the annual mean significant wave height computed between 1921 and 2010. Left - absolute change (cm per year); right - relative change (% per year)

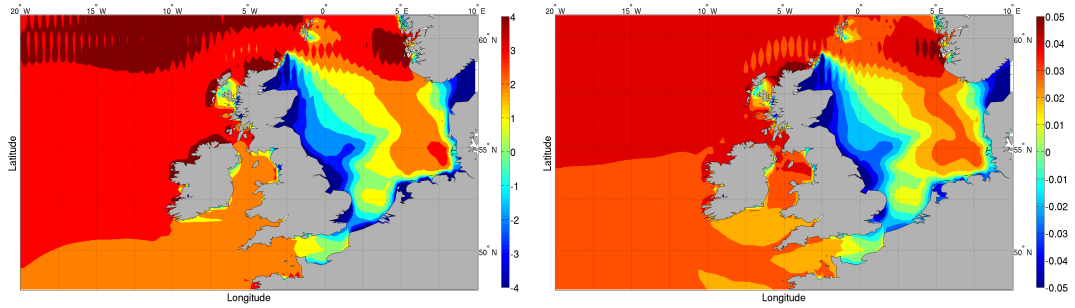

Figure 5.9: Changes in the annual mean energy period computed between 1921 and 2010. Left - absolute change (x10^{-3}s per year); right - relative change (% per year)

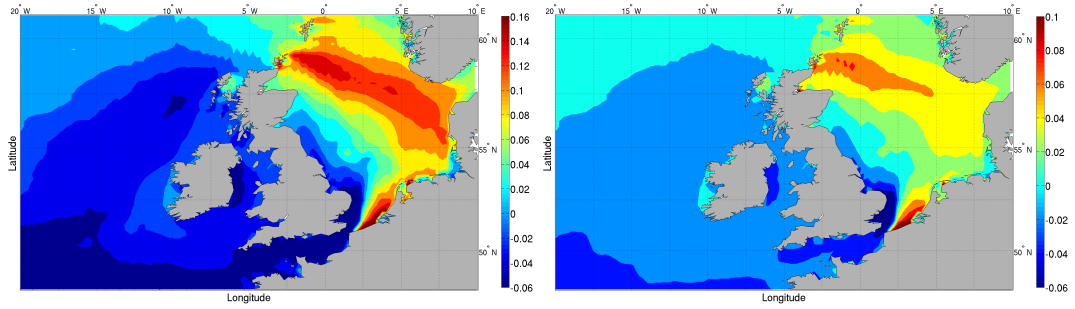

Figure 5.10: Changes in the annual mean wave direction computed between 1921 and 2010. Left - absolute change (deg. per year); right - relative change (% per year)

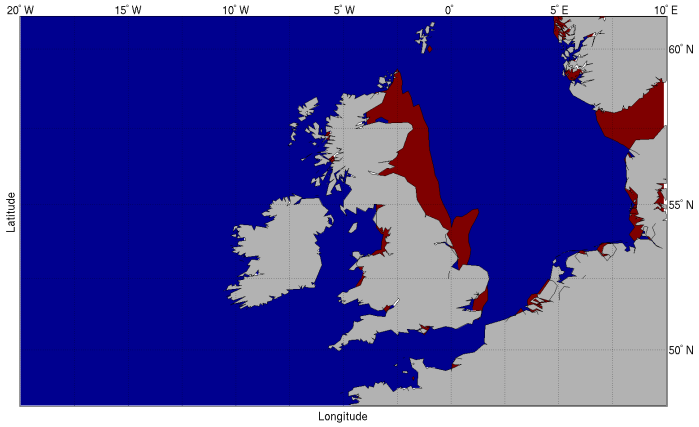

Figure 5.11: Statistical significance of trends observed in annual mean significant wave height. Blue areas are statistically significant at $\alpha = 0.1$.

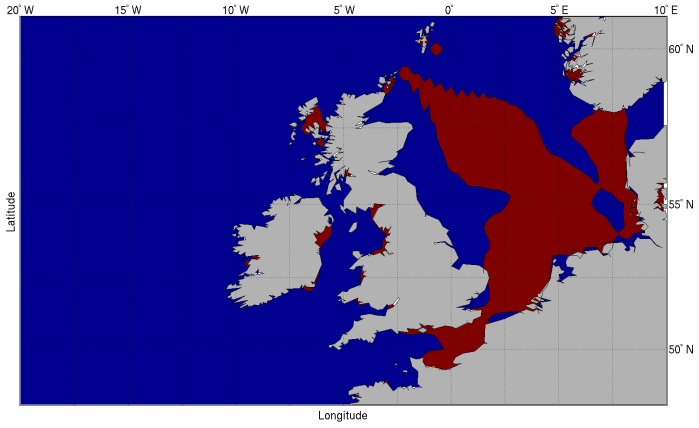

Figure 5.12: Statistical significance of trends observed in annual mean energy period. Blue areas are statistically significant at $\alpha = 0.1$.

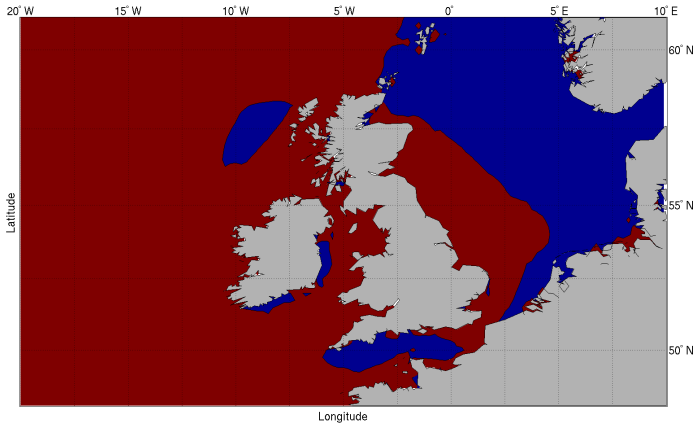

Figure 5.13: Statistical significance of trends observed in annual mean wave direction. Blue areas are statistically significant at $\alpha = 0.1$.

The magnitude of change in H_s is noticeably smaller in the North Sea. The absolute change ranges from 0 to 0.1cm-year^{-1} in the region, even exhibiting a negative trend in the sea between Norway and Denmark. The coast along the UK in the North Sea shows a very small upward trend of 0.05cm-year^{-1}. The magnitude of the positive trend in the North Sea increases southwards to $\approx$ 0.15cm-year^{-1}

These observations are more clearly visible in the timeseries plots of annual mean significant wave height plotted from 1871 to 2010 in Figure 5.14. The slope of the trend from 1920 onwards is greater for the North East Atlantic site (West of Hebrides) than for the North Sea site (Burghead). Similar observations can be made from the plots from the remaining sites in Appendix B.

From Figure 5.11 it is observed that the observed trends in H_s are statistically significant in most of the region, except for coastal UK waters in the North Sea and the region between Norway and Denmark. The implication of this is that the trends observed in H_s at the North Sea locations are not statistically significant and likely due to short-term variability

As indicated by the contour maps of energy period in Figures 5.6 the greatest increase in the parameter is observed in the northwest part of the region, and this is corroborated by the absolute and relative change plotted in Figure 5.9. The trends in the North East Atlantic range from 4x10$^-$3s-year^{-1} or 0.04%-year^{-1} in the north to 2x10$^-$3s-year^{-1} or 0.02%-year^{-1} in the south.

Although Figure 5.6 would indicate the absence of trends in the North Sea, the year-on-year changes shown in Figure 5.9 indicate the existence of weak positive trends in the central and eastern parts of the region and negative trends along the UK coast in the North Sea. These range from -4x10^{-3}s-year^{-1} or -0.04 %-year^{-1} along the UK coast to 2x10$^-$3s-year^{-1} or 0.02 %-year^{-1} in the eastern part of the sea. However, from Figure 5.12, it can be seen that the trends in large parts of the North Sea are not statistically significant, and must be treated with caution.

These trends are, once again, more clearly visible in the timeseries plots of annually averaged mean energy period in Figure 5.14. From the hinge-fit functions, a sharp difference may be observed between the pre-1920 and post-1920 trends. This must not be interpreted as a sudden change in the climate in 1920, and must instead be treated as being representative of the change in the base climate averaged over the two periods.

Of more interest, however, is the noticeable reduction in the inter-annual variability in T_e after 1920. For example, at Burghead, the range of variability reduces from $\approx$ 2s before 1920 to $\approx$ 1s after 1920. This can be observed for several other locations as well (Appendix B). While it is possible that the variability in the climate decreased, it is more likely that this is caused by the improvement in the quality of the input dataset over time.

Although the magnitude of the year-on-year changes in the wave period appear small (of the

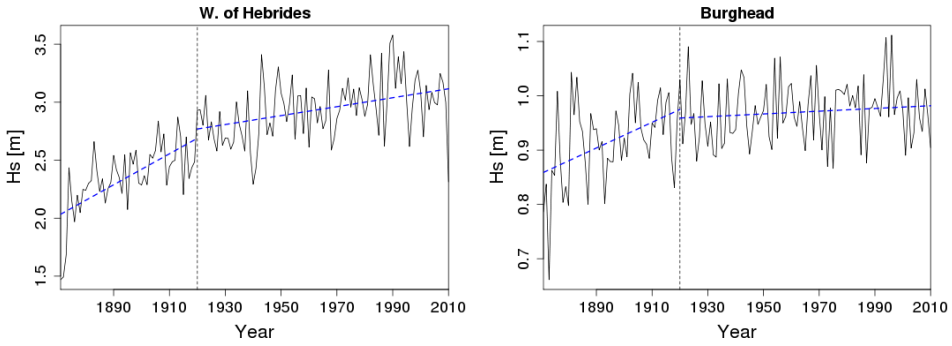

Figure 5.14: Time series data of annual mean significant wave height computed between 1871 and 2010 along with the hinge fit funtion

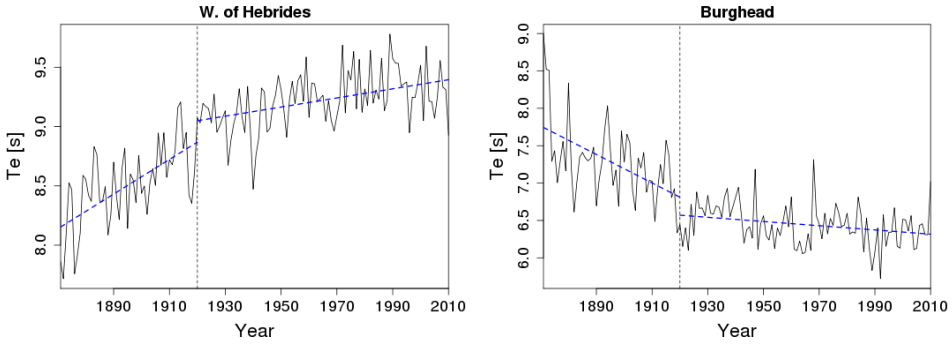

Figure 5.15: Time series data of annual mean energy period computed between 1871 and 2010 along with the hinge fit funtion

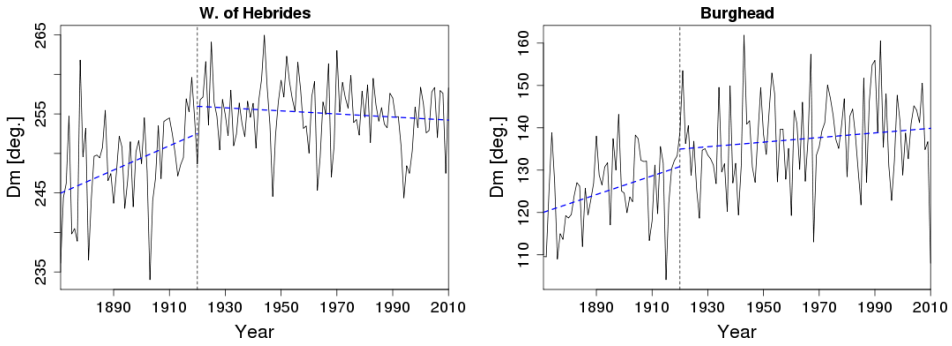

Figure 5.16: Time series data of annual mean wave direction computed between 1871 and 2010 along with the hinge fit funtion

order of 10^{-3}s) in both, the North East Atlantic as well as the North Sea, they are representative of a significant change (-0.2 to 0.5s) when compounded over a century.

The significance plots for annual means of wave direction in Figure 5.13 indicate that the trends observed in the North Atlantic and large parts of UK waters in the North Sea in Figure 5.10 are not significant at the 10% level. It is likely that the computed trends are a result of the large interannual variability, visible in the time-series plots of test locations in Figure 5.16 and in Figure B.3 in Appendix B

The eastern parts of the region, especially central and eastern North Sea where the trend in D_m are statistically significant also exhibit the largest upwards trends in the region, varying from 0.06°-year^{-1} to almost 0.16°-year^{-1}. Due to the nature of the quantity and the scale of the metric for mean wave direction, the relative change is not as informative as the absolute change in the parameter over the 90 year period studies and is therefore omitted from the discussion.

Long term trends in seasonal characteristics of the wave climate are presented as regional maps in Figures B.4 to B.6 and for individual locations in in Figures B.7 to B.18 in Appendix B.

5.7.2 Changes in extreme conditions

Currently, extreme sea conditions due to storms, swells and other high-energy events are described only using wave height and other parameters, such as wave period, are estimated by other means from the wave height (Det Norske Veritas, 2010; Goda, 2010). This precedent of keeping H_s as the focus is followed in this study as well, and for detecting and estimating changes in extreme occurrences the time-series of significant wave height is analysed.

The 95th percentile value of the time-series of H_s is used as a metric for analysing the average intensity of storm events. In an hourly time-series of measurements for a 365-day year, 438 observations exceed the value at this level. On average, 9.1 storm and hurricane events are observed in the North Atlantic Ocean per year (Mcadie *et al.*, 2009) and the duration of these events varies from 20 to 140 hours (Rangel-Buitrago and Anfuso, 2013). Based on this, one can expect about 700 hours of storm activity that will influence the wave climate in the region. Considering this, the 95th percentile value, which is likely to capture the more intense half of the high-energy events, is considered an adequate representation of the average intensity of storms.

The annual maxima of the time-series of H_s are taken as being representative of the intensity of the most-severe storm and are analysed for any long-term trends that may exist during the period of the study.

The plots of annual 95th percentile H_s for individual locations (Figure 5.18) as well as for the entire domain show an increasing trend overall in the region. It can be seen from Figure 5.17

that the magnitude of this trend is much larger in the North East Atlantic, varying from 0.4cm per year to 1.2cm per year. An upward trend can be observed in the south and east of the North Sea with the estimated annual increase varying from 0.2cm per year to 0.6cm per year.

The 95th percentiles of H_s appear to show little change in the western part of the North Sea along the Scottish and English coasts. This can be interpreted as being indicative of the increase in the average intensity of storm events in a large part of the modelled domain while remaining invariant in time in the central and eastern parts of the North Sea. These observations are consistent with the findings of the UKMO as reported in 'The recent storms and floods in the UK' (Slingo *et al.*, 2014). It should also be noted that the trend was not found to be statistically significant, as can be seen in the contour map in Figure 5.21.

In the North East Atlantic, trends observed in the annual maxima of H_s in the area between 50°N and 60°N approximately were found to be statistically significant as can be seen in Figure 5.22. It can also be observed from the figure that statistical significance cannot be established for the North Sea except for a small region in the south.

In oceanic waters in the western parts of the region, the greatest upward trend is observed in the north west corner with a magnitude of as much as 3.5cm-year^{-1}. However, based on the p-values from the t-test, it cannot be concluded that these trends are real and not a result of interannual variability. In the region where statistical significance of the trends is established at the level $\alpha = 0.1$, the increase in H_{max} was found to be between 1.5cm and 2.5cm per year.

In the North East part of the region, off the Norwegian coast, the annual maximum H_s increased by approximately 2.5cm or 0.25% in year-on-year terms. In the North Sea, towards the south and south east, the magnitude of H_{max} increased by ≈0.5cm to nearly 1.0cm per year between 1921 and 2010.

	Absolute change (cm per year)		Relative change (%)	
Location	H_{95}	H_{max}	H_{95}	H_{max}
Farr Point	0.79	1.02	0.19	0.12
EMEC	1.02	1.55	0.20	0.15
W. of Hebrides	1.12	2.07	0.19	0.18
Galway Bay	0.46	0.54	0.23	0.10
Belmullet	1.05	2.20	0.18	0.20
Beatrice	0.07	-0.92	0.02	-0.11
Wave Hub	0.57	0.17	0.11	0.02
Burghead	0.05	-0.74	0.02	-0.13

Table 5.12: Year-on-year changes between 'normal' H_{95} and H_{max} as estimated by the hinge fit trend between 1921-2010

From the year-on-year variation computed for the test locations and tabulated in Table 5.12 and from the contour plots of the regions presented in Figures 5.17 and 5.19, it can be seen

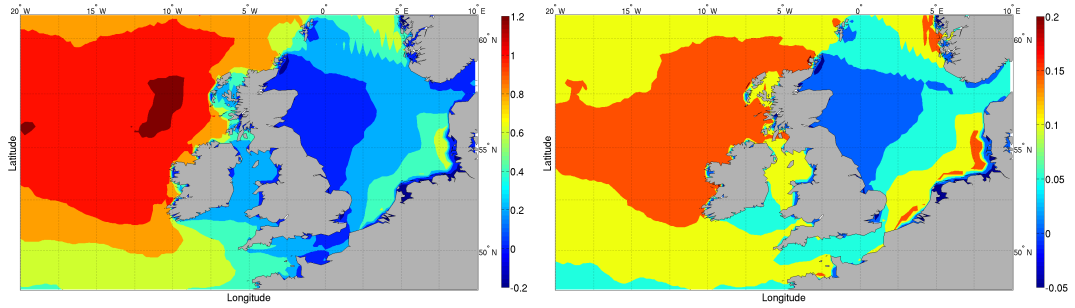

Figure 5.17: Changes in the annual 95th percentile significant wave height computed between 1921 and 2010. Left - absolute change (cm per year); right - relative change (% per year).

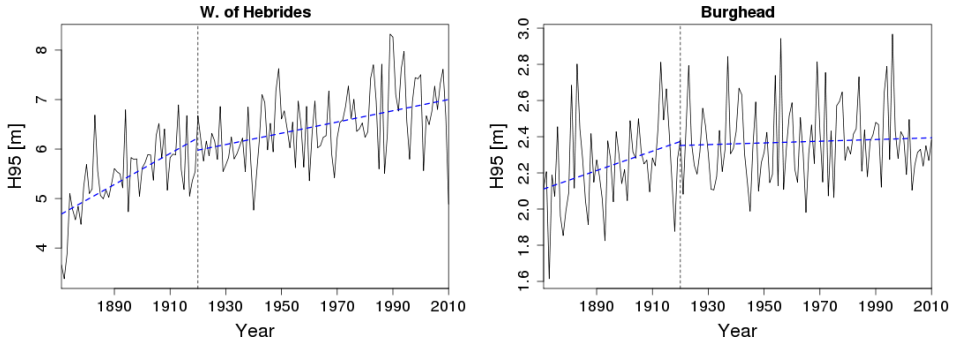

Figure 5.18: Time series data of annual 95th percentile significant wave height computed between 1871 and 2010 along with the hinge fit function for two locations

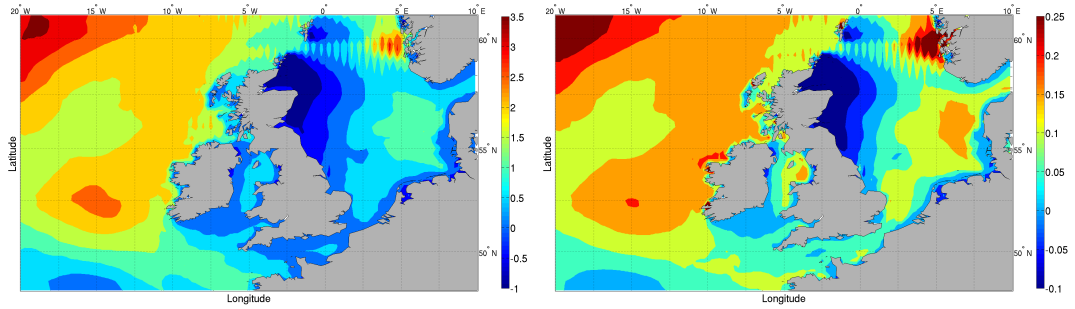

Figure 5.19: Changes in the annual maximum significant wave height computed between 1921 and 2010. Left - absolute change (cm per year); right - relative change (% per year).

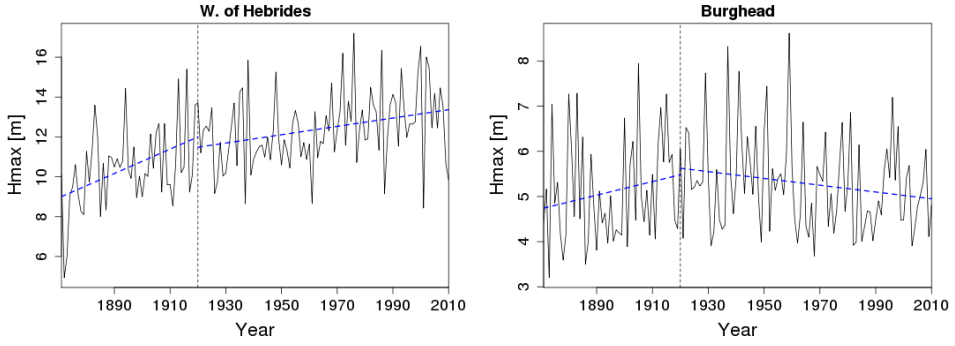

Figure 5.20: Time series data of annual maximum significant wave height computed between 1871 and 2010 along with the hinge fit function for two locations.

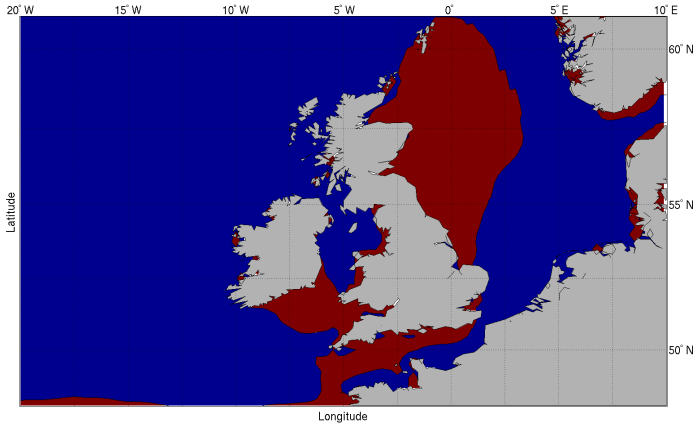

Figure 5.21: Statistical significance of trends observed in annual 95th percentile significant wave height. Blue areas are statistically significant at $\alpha = 0.1$.

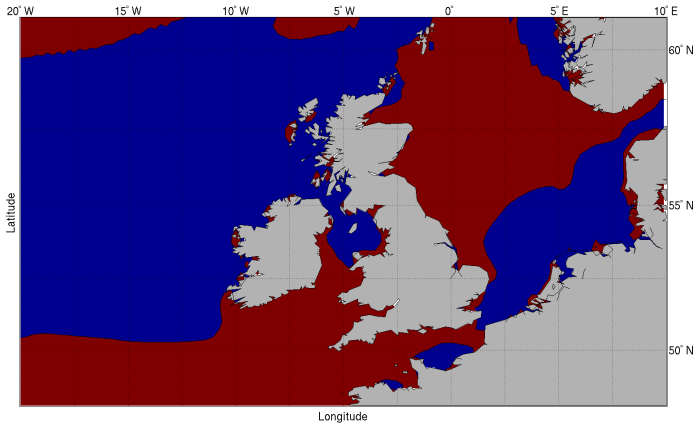

Figure 5.22: Statistical significance of trends observed in annual maximum significant wave height. Blue areas are statistically significant at $\alpha = 0.1$.

that trends in annual H_{95} and H_{max} exhibit very different characteristics. Although all the test locations in the north and west of the region exhibit increasing trends for both H_{95} and H_{max}, the rate of increase in absolute terms is greater for H_{max}, whereas the relative rate of change is greater for H_{95}.

The shapes and positions of the contours are also different for H_{95} and H_{max} in the North East Atlantic. While the magnitude of the upward trend for H_{95} appears to be greatest around 57°N with the year-on-year increase reducing to the north and south, the magnitude of the upward trend for H_{max} is greatest in the north west corner of the domain decreasing southwards.

The differences between the characteristics of the two parameters are more distinct in the North Sea. Only a very small region immediately east of the Orkney Islands shows a year-on-year decrease in H_{95}, however, for H_{max} large areas off the coast of central and northern UK exhibit

decreasing trends.

The absence of a simplistic relationship between the trends in H_{95} and H_{max} is most noticeable at the two North Sea locations, Beatrice and Burghead. It can be seen from Table 5.12 that these sites exhibit a weak increase in the average intensity of storms (indicated by H_{95}) accompanied by a decrease in the intensity of the most intense storm event (indicated by H_{max}).

These observations, particularly for the North Sea, might be explained by the fact that storm tracks in the North Atlantic Ocean have been observed to be shifting northwards over the latter half of the 20th century (Paciorek *et al.*, 2002; Wang *et al.*, 2006). Consequently, storm swells arriving in the region would also originate in a more northerly direction. This is likely to cause changes in shadow regions affecting the wave climate. It must be kept in mind, however, that these trends could not be confirmed as being significant.

The analysis of long-term trends in seasonal storm activity is presented in the form of regional maps of H_{95} and H_{max} in Figures B.21 and B.22 and as time-series plots for test locations in Figures B.23 to B.30 in Appendix B.

5.7.3 Changes in the wave energy resource

The theoretical wave power resource is often estimated in terms of power density, which is the mean wave power transported per meter of wave crest. It can be calculated as a function of the significant wave height and the energy period of the wave as

$$P = \frac{\rho_w g^2}{64\pi} H_s^2 T_e \text{W/m} \tag{5.8}$$

$$\approx 0.49 H_s^2 T_e \text{ kW/m} \tag{5.9}$$

The timeseries so obtained was subjected to similar analyses as for mean wave parameters (Section 5.7.1) and extreme conditions (Section 5.7.2). The computed trends in this parameter along with statistical significance are presented as regional contour maps in Figures 5.23 and 5.24 respectively, and as timeseries plots for sites located West of Hebrides and at Burghead in Figure 5.25.

The regional distribution of statistical significance in Figure 5.24 is very similar to the distribution for annual mean significant wave height in Figure 5.11. This is likely on account of the power density varying according to the square of H_s. The trends obtained from the analysis were found to be significant in most of the region except to the east of the Shetland Islands, and along the UK's coast and between Norway and Denmark in the North Sea. Once again, at the North Sea locations, Beatrice and Burghead, it cannot be concluded that the trends are not a result of short-term variability.

From an examination of the contour maps of trends in Figure 5.23, it can be observed that in the North East Atlantic, the largest absolute year-on-year increase is between $\approx 54^\circ$N and

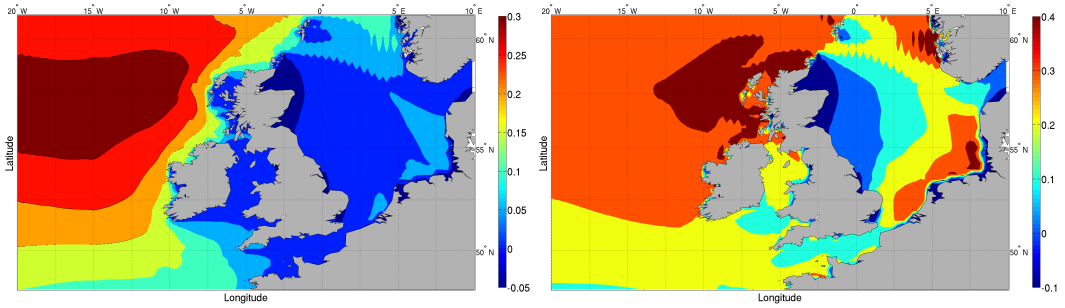

Figure 5.23: Changes in the annual mean wave power density computed between 1921 and 2010. Left - absolute change (cm per year); right - relative change (% per year).

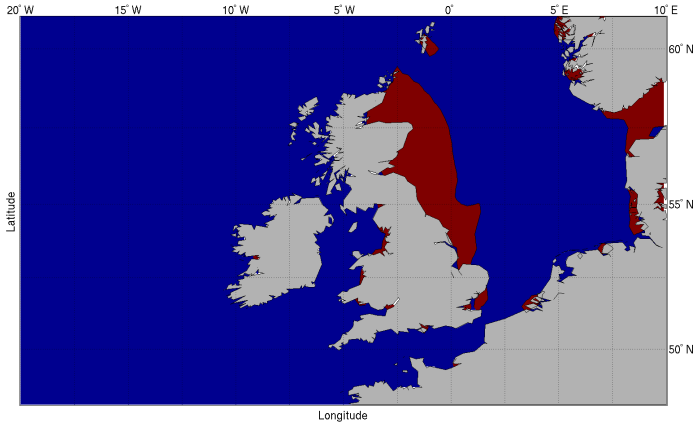

Figure 5.24: Statistical significance of trends observed in annual mean wave power density. Blue areas are statistically significant at $\alpha = 0.1$.

60°N amounting to 0.3kW/m per year. The largest relative increase, of approximately 0.4% per year, is observed to the West and North of Scotland, encompassing sites of interest to the wave energy sector at EMEC, West of Hebrides, Farr Point, to name a few.

Relative trends of similar magnitudes are seen off the coast of Norway and Germany. However, at the latter location, the trend is not significant at the $\alpha = 0.1$ level. A decreasing trend in the power density was only observed in the Moray Firth and Firth of Forth, where statistical significance cannot be established. This can be visually corroborated from an examination of the time-series plot of annual mean wave power density at Burghead which exhibits high interannual variability. In the rest of the North Sea, small increases are observed, ranging from 0kW/m per year to 0.1kW/m per year.

Seasonal trend analyses for the region are presented as contour maps of absolute and relative year-on-year changes in Figure B.32 and as timeseries plots for the selected test locations in Figures B.33 to B.36.

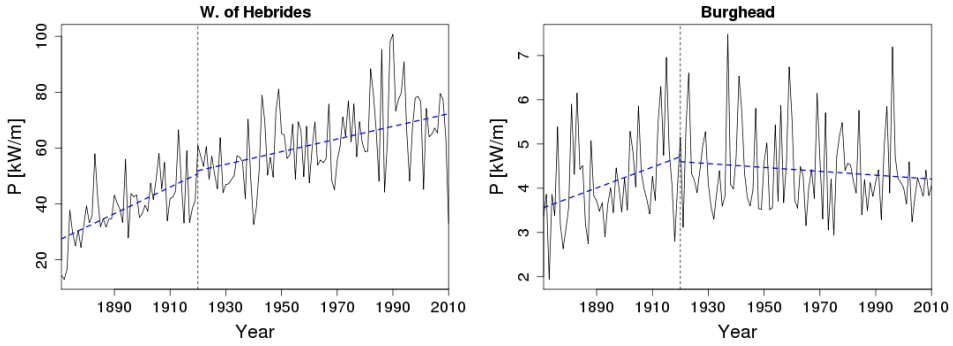

Figure 5.25: Time series data of annual mean wave power density computed between 1871 and 2010 along with the hinge fit funtion.

5.8 Chapter summary

- Three methods for the detection of long-term climatic changes were tested - 30-year WMO-standard climate normals, least squares linear fit and the hinge fit;
- Of these, the hinge-fit was found to be most appropriate for detecting the existence and estimating the magnitude of long-term trends in wave climate parameters;
- Long-term variations in descriptors of the wave climate between 1921 and 2010 were assessed for 8 sites and for the entire gridded dataset. The results were presented as time-series plots and contour maps;
- Statistical significance of detected trends was tested using Student's t-test between samples from 1921-1950 and 1981-2010. Large parts of the region exhibit statistically significant trends, whereas trends observed in other regions are likely a result of short-term variability; and
- It can be concluded that long-term climatic changes have been observed in several wave parameters. The magnitude and direction of these trends vary from parameter to parameter and region to region.

Chapter 6

Analysis of the wave climate - II

6.1 Introduction

In Chapter 5, the historical wave climate in the North East Atlantic Ocean and North Sea was analysed for the detection of climatic changes in the sea state. The analysis was restricted to the detection of such changes without attempting to attribute them to any factors such as atmospheric concentrations of greenhouse gases (GHGs).

This chapter presents the analysis of wave climate data obtained by forcing WAVEWATCH III with atmosphere-ocean global circulation model (AOGCM) data for the B1 (low emissions), A1B (medium emissions) and A2 (high emissions) climate change scenarios.

The analyses presented have the following objectives:

1. Estimation of the changes in the wave climate in the 21st century according to these scenarios; and
2. Attribution of the observed trend in the wave climate to atmospheric GHG levels.

6.2 Description of the dataset

For the analyses proposed, a suitable dataset of wave parameters is required, preferably in the form of a gridded timeseries. Various studies in the past have used the 'time-slice' method, e.g. Gunther *et al.* (1997); Kaas *et al.* (2001), in which numerically modelled datasets are generated for a control period, usually the current climate, and for an equal period in the future according to a predefined emissions scenario such as a doubling of CO_2 levels. These are compared and the differences analysed.

A major shortcoming of this method is that as the two 'slices' of data are not contiguous, it cannot be concluded that the variation from the control is indeed a trend and not the result of shorter term variations. Moreover, it also cannot be concluded from a single scenario that the changes are a result of greenhouse gas forced climate change.

To produce a suitable dataset for analysing the causal link between greenhouse gas levels and the wave climate as well as identifying any trends that may exist the numerical wave

model WAVEWATCH III forced with winds from the ECHAM5 Global Circulation Model (GCM) corresponding to different climate change experiment runs are used to simulate wave conditions.

Wind data from the GCM for the marker scenarios corresponding with high (A2), medium (A1B) and low (B1) emissions scenarios from 2001 to 2100 are used to generated corresponding wave conditions. As the GCM runs for these scenarios are intialised using data from the end of the year 2000 from the 20th century control run (20C3M), the wave model is also forced with winds from the control experiment to homogenously extend the dataset backwards in time to January 1981 to allow the evaluation of the state of the current climate.

It is assumed that the parameters governing the physical processes in the generation, propagation and dissipation of wind waves in the region do not vary from the values identified in the parameter optimisation process in Section 4.6. Specifically, it is assumed that the wind-growth term, $\beta_{\max}$ obtained for the 20^{th} Century Reanalysis (20CR) dataset is valid for the projected winds from the MPI ECHAM5 AOGCM for the climate change scenarios described previously.

6.3 Comparison of current climate

6.3.1 Need for such intercomparison

The ability of a numerical model in simulating the state of the climate needs to be evaluated before the model can be used with confidence to predict future conditions. A common method of such an evaluation, especially for climate models such as AOGCMs, is by testing the ability of the model to simulate contemporary climatic conditions.

In other words, the state of the current climate, as predicted by the climate model is compared with the observed state of the current climate for evaluating the model's skill (Randall *et al.*, 2007). Several such evaluations have been demonstrated for mean and extreme conditions for various climatological parameters such as temperature and precipitation in the recent Fifth Assessment Report by the IPCC (Flato *et al.*, 2013).

The IPCC recommends that, during such an evaluation, differences between simulated and observed climates can be ignored if they are within the range of internal variability and other uncertainties (e.g. difference in forcing). However, a poor agreement between spatial and temporal patterns between the two could indicate that one or more key processes within the climate system have been incorrectly or inadequately parameterised within the model.

This evaluation is especially necessary in a study such as this as the climatic parameters of interest are not directly output by the climate models. Instead, the output from climate models is used to force a secondary model (WAVEWATCH III) to simulate time-series of parameters of interest.

6.3.2 Methodology

It is known that although the 20th Century experiment (20C3M) was run using forcing data representative of observations during the period, the simulated and observed climate cannot be directly compared on an event-by-event or pairwise basis. Therefore, to evaluate the ability of the ECHAM5-WAVEWATCH III coupling in simulating wave conditions, the following methodology is proposed.

- **Period of analysis**: As the timeseries of simulated and observed historical climate cannot be compared directly, relevant statistics need to be computed over a suitable timeframe and compared. For this purpose data for the 30-year period between 1981 to 2010 is chosen as being representative of the current climate. The rationale for this block size has been explained in Chapter 5.
- **Preparation of Datasets:** The timeseries of H_s, T_e and D_m simulated by WAVEWATCH III when forced with the 20th Century Reanalysis (20CR) dataset is taken to represent the observed climate. To represent the simulated contemporary climate, data from 2001 to 2010 from the wave model forced with winds from the A1B experiment is merged with data from 1981 to 2000 from the wave model forced with the 20th Century (20C3M) control experiment. These data are ordered to allow a ranked comparison in addition to the comparison of statistical metrics.
- **Site level comparison:** The mean, 95th percentile and maximum values from the two datasets will be computed and compared for H_s and T_e data for 8 sites (same as those used in previous analyses). Due to the nature of the scale of wave-direction measures such as 95th percentile and 30-year maximum value have little significance. For wave direction, the 30-year mean and the modal histogram bin are used for comparison. In addition, for the observed and simulated H_s, T_e and D_m data histograms with fitted probability density functions and ordered quantile plots are also prepared for visual comparison.
- **Regional comparison:** For the evaluation of spatial patterns, the 30-year mean value is computed for the gridded datasets of significant wave height, mean energy period and mean wave direction and plotted as contour maps of the North East Atlantic and North Sea region.

6.3.3 Results of the comparison

The statistical quantities identified in Section 6.3.2 as being representative of the base state of the climate for comparison were computed for the 8 test locations and are presented in Tables 6.1 to 6.3.

From these statistical measures presented in the tables it can be seen that there is some underestimation in the simulated significant wave heights and wave periods when compared with reanalysis data. Good agreement is observed in the 30-year average of the mean wave direction as well as the dominant wave direction, indicated by the modal directional bin, which is seen

Site	$\overline{H_s}$		H_{95}		H_{max}	
	20CR	A1B	20CR	A1B	20CR	A1B
Farr Point	1.962	1.654	4.768	3.735	11.450	10.136
EMEC	2.502	2.220	6.012	5.103	13.984	14.555
West of Hebrides	3.058	2.906	7.010	6.299	16.551	16.399
Galway Bay	1.086	0.815	2.924	1.961	6.067	4.343
Belmullet	2.998	3.092	6.761	6.551	17.145	14.101
Beatrice	1.528	1.537	3.677	3.592	10.038	10.060
Wave Hub	2.460	2.052	5.975	4.342	9.308	8.073
Burghead	0.975	0.931	2.391	2.289	7.197	6.480

Table 6.1: A comparison of the current climate as estimated by the wave model forced with reanalysis (20CR) and GCM (20C3M+A1B) data in terms of mean, 95th percentile and maximum H_s (in m) between 1981 and 2010

Site	$\overline{T_e}$		T_{95}		T_{max}	
	20CR	A1B	20CR	A1B	20CR	A1B
Farr Point	8.706	7.554	13.370	11.412	18.943	17.891
EMEC	9.625	8.594	13.705	12.391	18.708	17.977
West of Hebrides	9.314	9.136	13.070	12.642	17.978	17.647
Galway Bay	9.171	6.211	13.856	10.561	19.521	17.407
Belmullet	9.534	9.192	13.357	12.677	18.101	17.298
Beatrice	7.556	5.762	14.145	8.309	21.262	16.131
Wave Hub	9.420	8.437	13.966	11.950	19.736	17.759
Burghead	6.349	4.694	11.848	7.133	21.549	12.853

Table 6.2: A comparison of the current climate as estimated by the wave model forced with reanalysis (20CR) and GCM (20C3M+A1B) data in terms of mean, 95th percentile and maximum T_e (in s) between 1981 and 2010

Site	Mean D_m		Modal D_m	
	20CR	A1B	20CR	A1B
Farr Point	270.3	256.2	295.0	295.0
EMEC	261.0	268.2	275.0	275.0
West of Hebrides	254.2	254.5	255.0	255.0
Galway Bay	244.0	236.8	255.0	255.0
Belmullet	265.4	263.6	275.0	265.0
Beatrice	156.7	167.4	125.0	15.0
Wave Hub	254.1	255.5	275.0	275.0
Burghead	139.0	145.7	95.0	95.0

Table 6.3: A comparison of the current climate as estimated by the wave model forced with reanalysis (20CR) and GCM (20C3M+A1B) data in terms of mean and modal D_m (in degrees) between 1981 and 2010

to be the same for most locations. However, further investigation was deemed necessary on account of the discrepancies between the datasets for H_s and T_e.

To analyse the differences between the two datasets of contemporary climate, significant wave height, mean energy period and mean wave direction data were sorted in bins of 1m, 1s and 10° respectively, and density histograms were plotted. Suitable probability distributions from among the Gumbel, Frechet and Weibull distributions were fitted to the observed as well as simulated climate data using maximum likelihood. This was not done for wave direction. These plots are presented for one North Atlantic site (West of Hebrides) and one North Sea site (Burghead) in Figure 6.1. Similar plots for the other 6 locations are presented in Figures C.16 to C.18 in Appendix C. The values of the parameters describing the PDFs are tabulated in Tables C.5 and C.6 in Appendix C.

A large overlap can be observed between the histograms of the observed and simulated H_s and D_m indicating good agreement between the datasets for all the test sites. Although differences between the datasets for H_s are more noticeable in the fitted PDFs, the bulk of samples can be seen to coincide. These observations are seen for all the sites, irrespective of geographic location.

Although it is not immediately obvious from an examination of the histograms and PDFs for mean energy period, a closer examination reveals that in addition to yielding lower estimates of the parameter, the spread of values is also narrower in the ECHAM5+WWIII dataset. As a result the calculated means are significantly shifted to the left, especially in low-energy areas such as the North Sea, and this is manifested in the mismatch between contours in the maps of spatial patterns in T_e presented in Figure 6.3.

In the quantile plots of ordered data presented in Figure 6.2, the high degree of linearity and proximity to the unity diagonal for H_s and D_m indicate that despite underestimations in some cases, the wave model forced with GCM winds is able to replicate these wave parameters. In the case of T_e similar trends can be seen for sites with an energetic wave climate, e.g West of Hebrides, EMEC, and Belmullet. However, for low energy sites, such as the North Sea sites Beatrice and Burghead and the Irish site in Galway Bay, no such trend between the two datasets can be observed.

The 30-year mean calculated from the gridded regional dataset of wave parameters are presented as contour maps in Figure 6.3. In the maps presented for wave direction, some coastal cells, such as along the south east coast of England, have been masked out (coloured white). This is because the quality control process detected infinity as the wave direction for some timesteps. Such a value is not realistic, and hence the cells have been masked out.

From the maps of significant wave height, it is observed that the spatial patterns are similar for the observed and simulated datasets, as indicated by the contours. A high degree of agreement can be seen in the North East Atlantic, and to a large extent in the North Sea. Although the

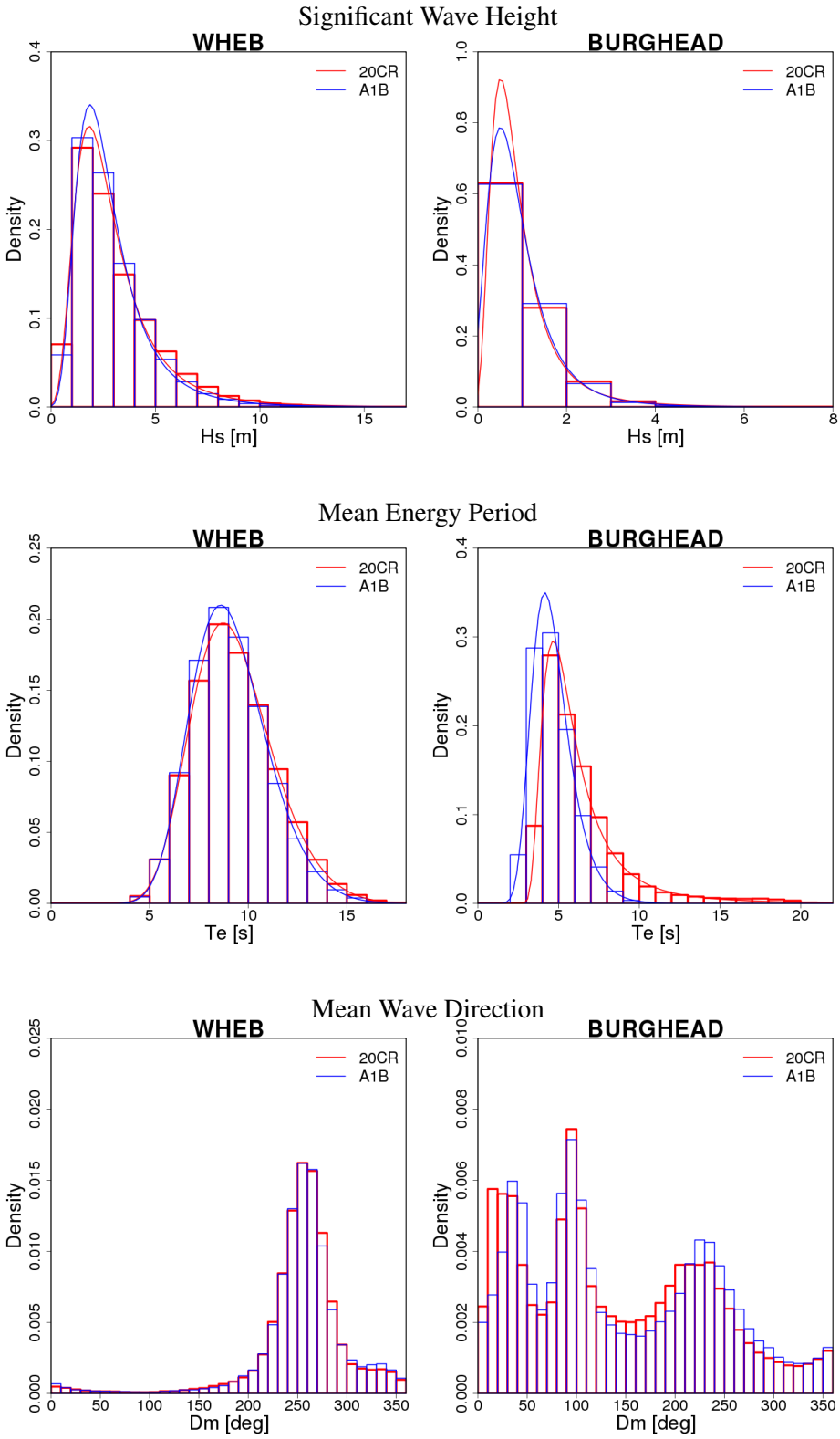

Figure 6.1: Histograms of wave data from model runs forced with reanalysis (20CR) and GCM (20C3M+A1B) data describing the climate between 1981 and 2010 for the sites West of Hebrides (WHEB) and Burghead

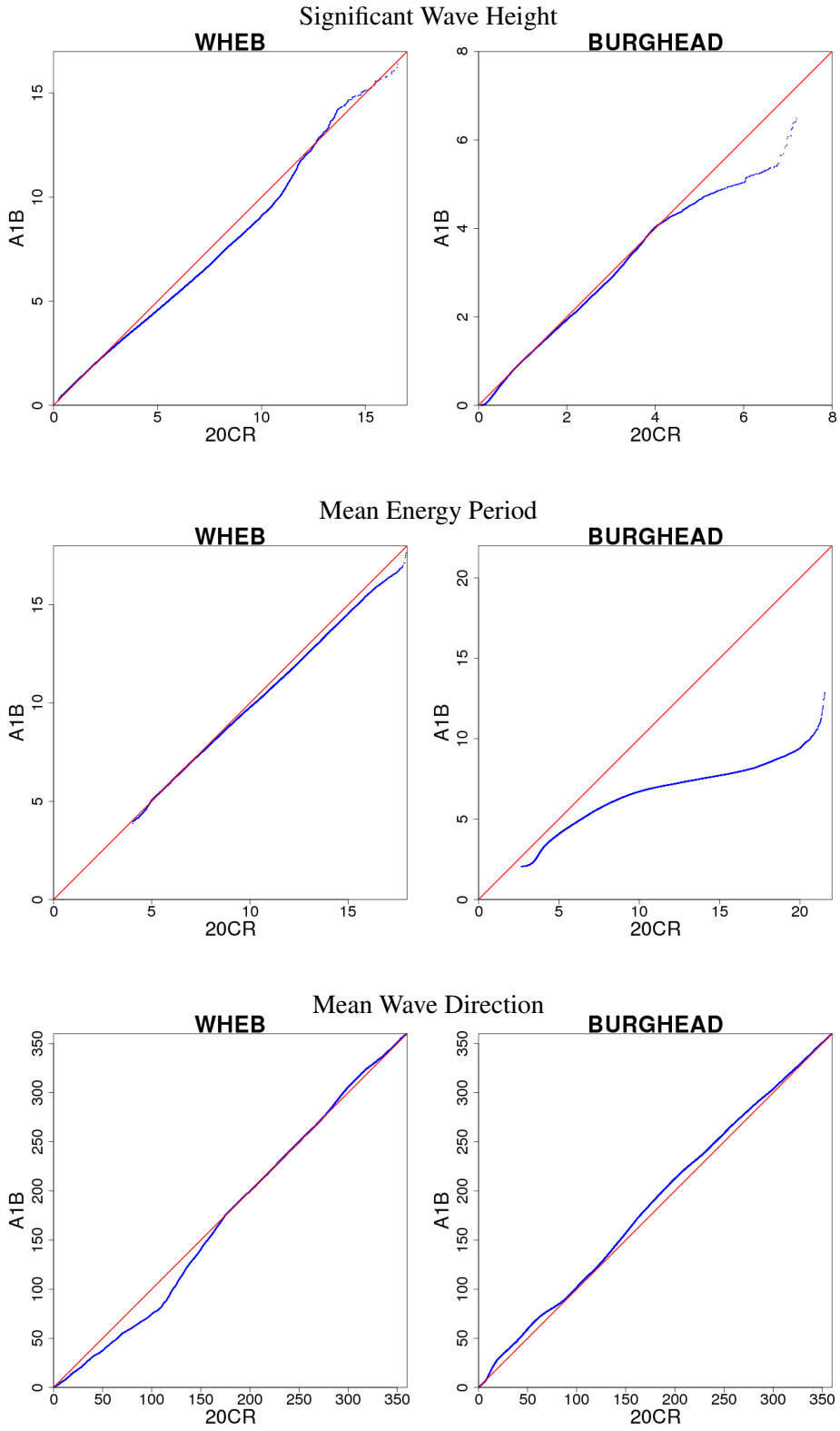

Figure 6.2: Quantile plots of ordered data from model runs forced with reanalysis (20CR) and GCM (20C3M+A1B) data describing the climate between 1981 and 2010 for the sites West of Hebrides (WHEB) and Burghead

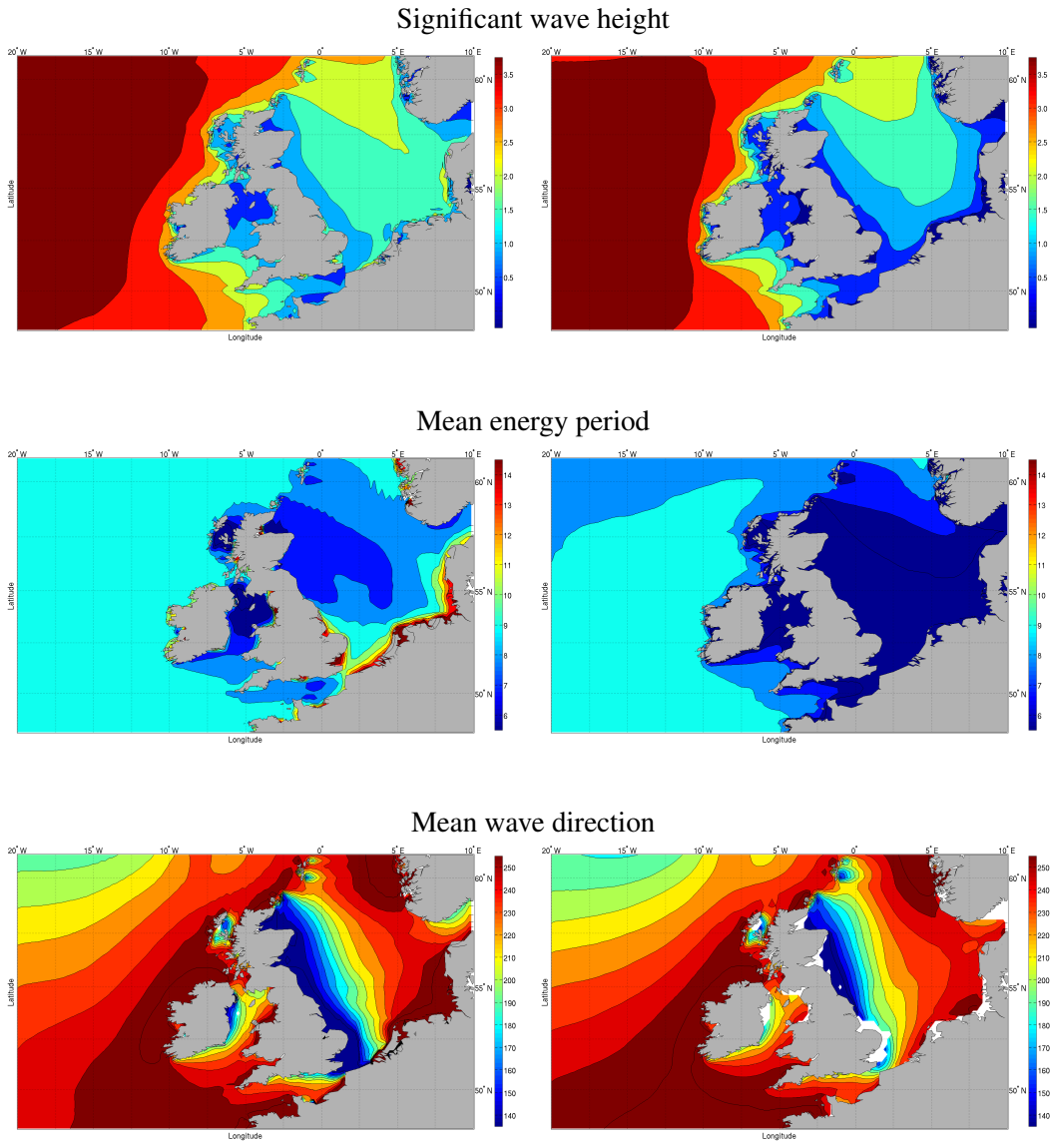

Figure 6.3: A comparison of the wave climate averaged between 1981 and 2010 from hindcasts and climate change experiments. Left - Historical, right - from the 20C3M (1981-2000) and A1B (2001-2010) datasets

contours are similar, it would appear that the ECHAM5-WWIII coupled dataset yields lower estimates of H_s in the North Sea.

The contour maps of 30-year average of mean energy period appear to indicate that in addition to underestimations, the spatial variation of the parameter are dissimilar between the observed and simulated datasets. It is difficult to infer conclusively for the North East Atlantic, however, in the North Sea, these dissimilarities are more obvious corroborating the inferences made from the histograms and quantile plots for less energetic locations.

Especially good agreement can be observed between the magnitudes and positions of contours in the maps of the mean wave direction. This agreement is seen in magnitude as well as spatial patterns in the North East Atlantic as well as the North Sea, which is greatly encouraging.

6.3.4 Implications of the results

It can be concluded from the site-level and regional comparisons of the contemporary wave climate, that despite underestimating significant wave height in comparison with reanalysis data, the ECHAM5-WWIII coupling is able to represent spatial trends in the parameter adequately. However, the severe underestimations in wave energy period, especially in the North Sea, along with a mismatch between observed and simulated spatial patterns are a matter of concern. The agreement between the observed and simulated contemporary climate in terms of mean wave direction at site and regional levels indicates a good ability on the part of the GCM in simulating wind patterns.

On account of the underestimation, a direct projection of future climate cannot be performed from observed historical climate. Instead, these projections must be computed in terms of the difference, absolute and relative, between the ECHAM5+WWIII datasets of projected and contemporary climate. Caution is advised in interpretation of trends relating to wave period and derived quantities, e.g. power density, because of the differences observed.

6.4 Wave climate projections

Having established that the contemporary wave climate obtained from the ECHAM5-WW3 coupling shows acceptable agreement with the wave climate obtained using reanalysis winds, plausible future trends and variability can be analysed. This can be achieved by fitting trends discussed in the previous chapter to future wave data for different scenarios.

However, based on the results of Section 6.3, while some degree of confidence can be associated with trends in H_s and D_m, caution is advised in the treatment of any trends that may be detected in the energy period (T_e).

6.4.1 Site level analysis

For the 8 test sites, the following measures were computed from the annual timeseries data between 2001 and 2100 for the A2, A1B and B1 climate change experiments:

- Annual mean H_s, T_e and D_m;
- Annual 95th percentile H_s; and
- Annual maximum H_s.

In addition, these quantities were computed from 1981-2000 from the 20th Century control (20C3M) experiment, and the data were plotted against time for the test sites.

A rudimentary inspection of the data was performed and no apparent change in trends, indicating the presence of a 'hinge', could be identified between 2001 and 2100. Following the example of Lowe *et al.* (2009), a linear regression function was fit using the least-squares method to the data from the 3 climate change scenarios. These were constrained so that the trend lines from the climate change experiments diverge from the 'historical' trend fit to the 20C3M data from the year 2001. The annual quantities, as identified earlier, along with the least-squares linear functions were plotted. These are presented in Figures 6.4 to 6.8 for two sites, West of Hebrides and Burghead. Plots for the remaining locations are presented in Figures C.22 to C.36 in Appendix C.

In addition, the magnitudes of the trends in absolute as well as relative terms were computed for the 8 locations and tabulated in Tables 6.4 to 6.8. The significance of these trends was tested using the Student's t-test and the p-values are presented in Table 6.9.

Weak negative trends in the mean annual significant wave height were observed at most of the sites tested for the A2 and A1B scenarios. The magnitude of the decreasing trends were found to be slightly larger for the B1 scenario at these locations. These trends, however, were not found to be statistically significant, except under the A2 scenario at the two North Sea sites. It is highly likely, as indicated by the large p-values in Table 6.9(a) that these trends are a result of the variability of the system.

Similarly, weak negative trends were observed in the mean annual T_e at all the sites. Despite the small magnitude of these trends, they were found to be statistically significant for several locations and scenarios at the 10% significance level.

Upward trends were detected in the mean annual D_m with moderate strength in some locations. The relative change, for this parameter is unfortunately not very informative. Locations such as EMEC and West of Hebrides may experience a northwards shift in the parameter by 1° to 4° over a 100-year period, depending on the emissions pathway. Over the same period, other sites, such as Beatrice and Burghead in the North Sea experience a shift of 6° to 17°. The trends in mean annual D_m were found to be statistically significant for several scenarios at various locations (e.g. all scenarios at the two North Sea sites).

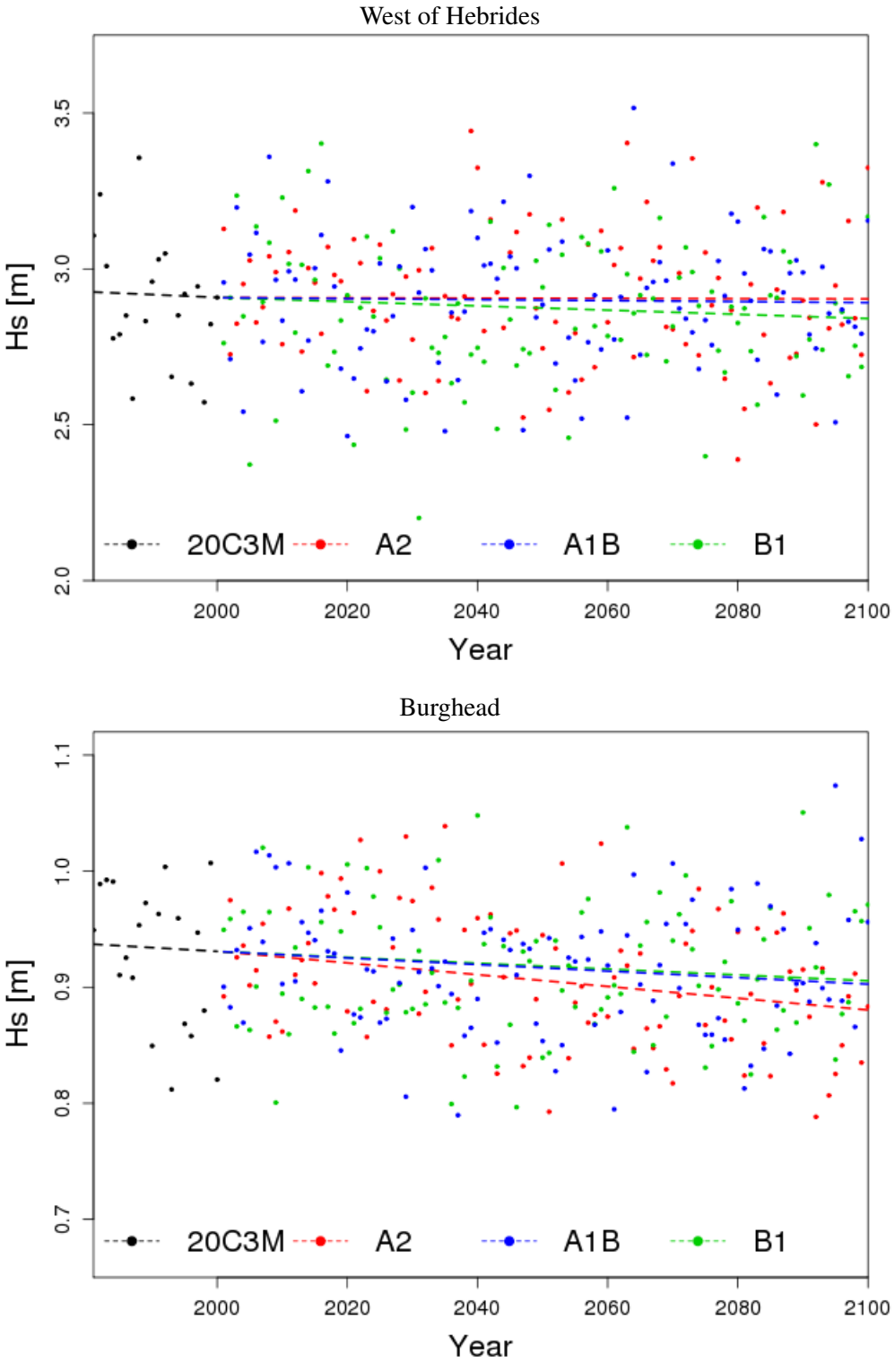

Figure 6.4: Projections of annual mean H_s from the high (A2), medium (A1B) and low (B1) emissions scenarios

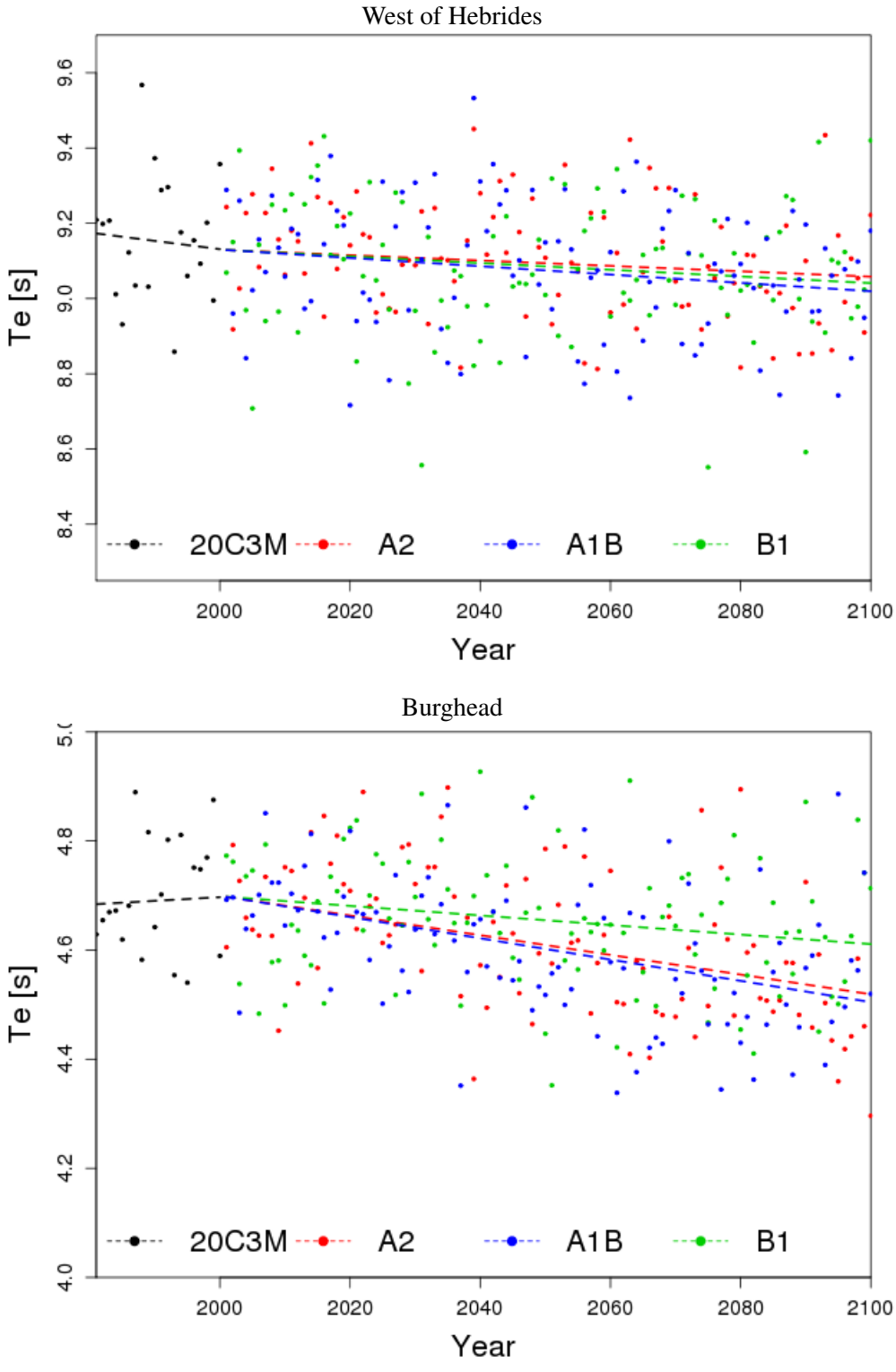

Figure 6.5: Projections of annual mean T_e from the high (A2), medium (A1B) and low (B1) emissions scenarios

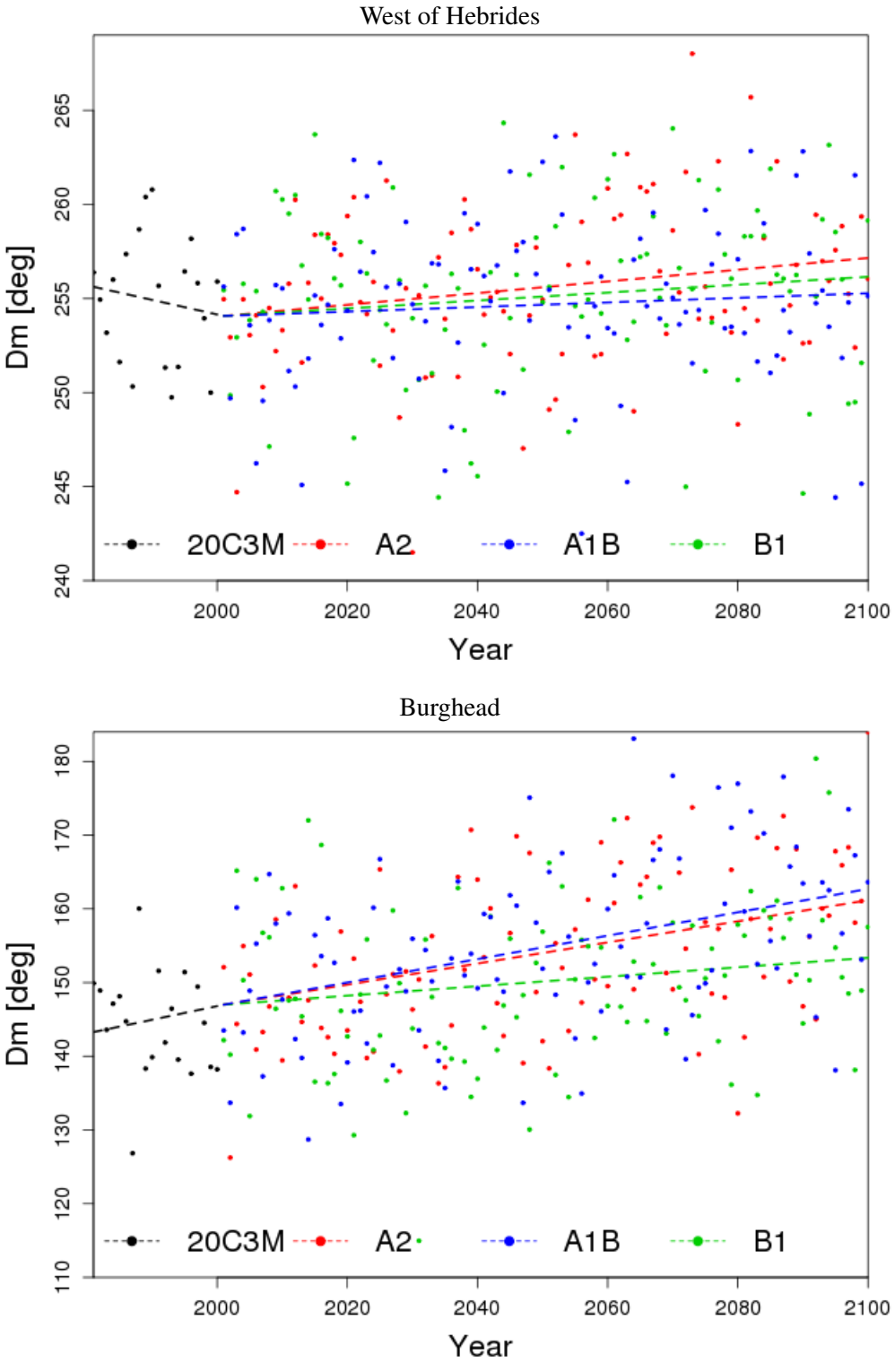

Figure 6.6: Projections of annual mean D_m from the high (A2), medium (A1B) and low (B1) emissions scenarios

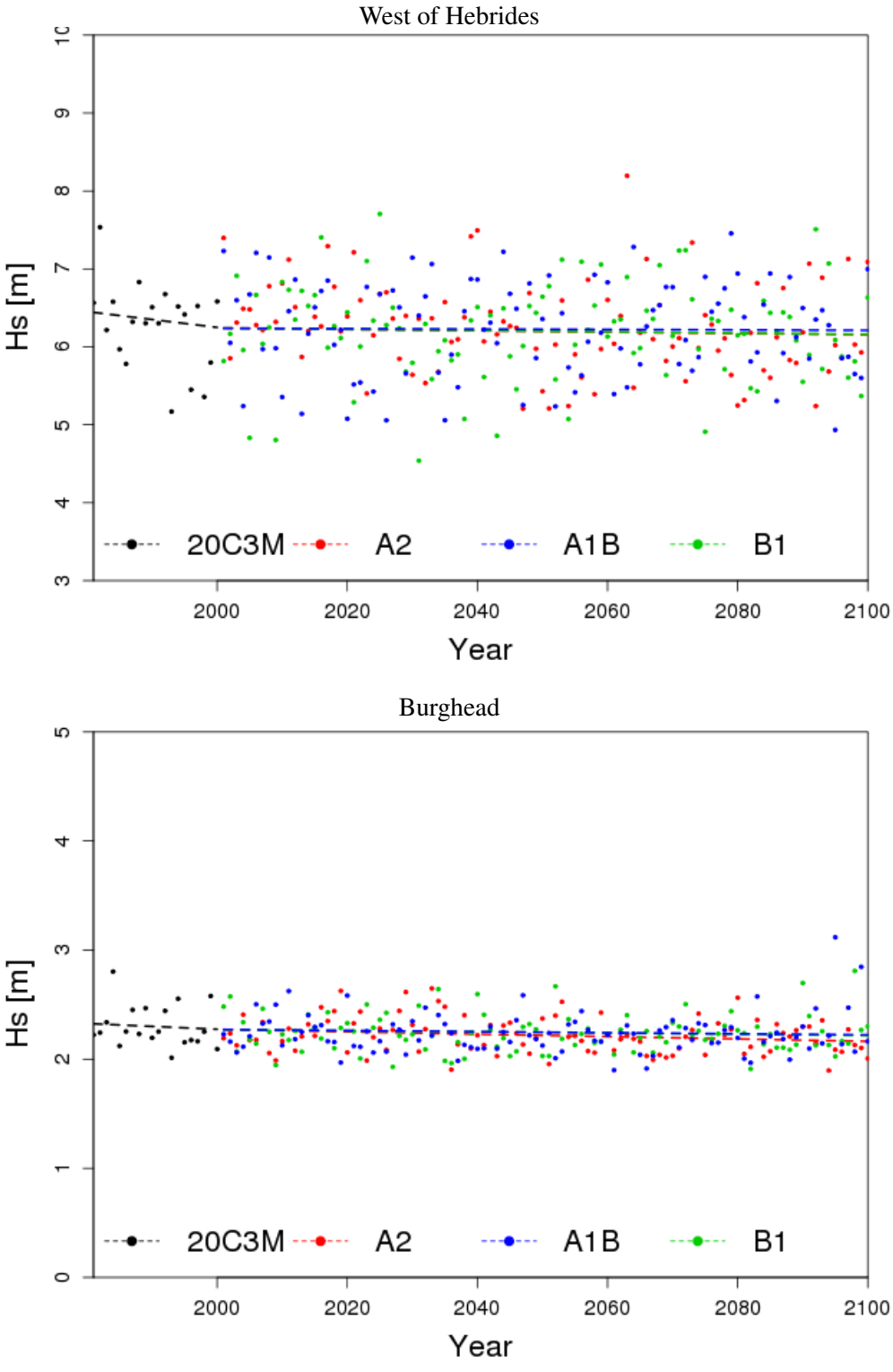

Figure 6.7: Projections of annual 95th percentile H_s from the high (A2), medium (A1B) and low (B1) emissions scenarios

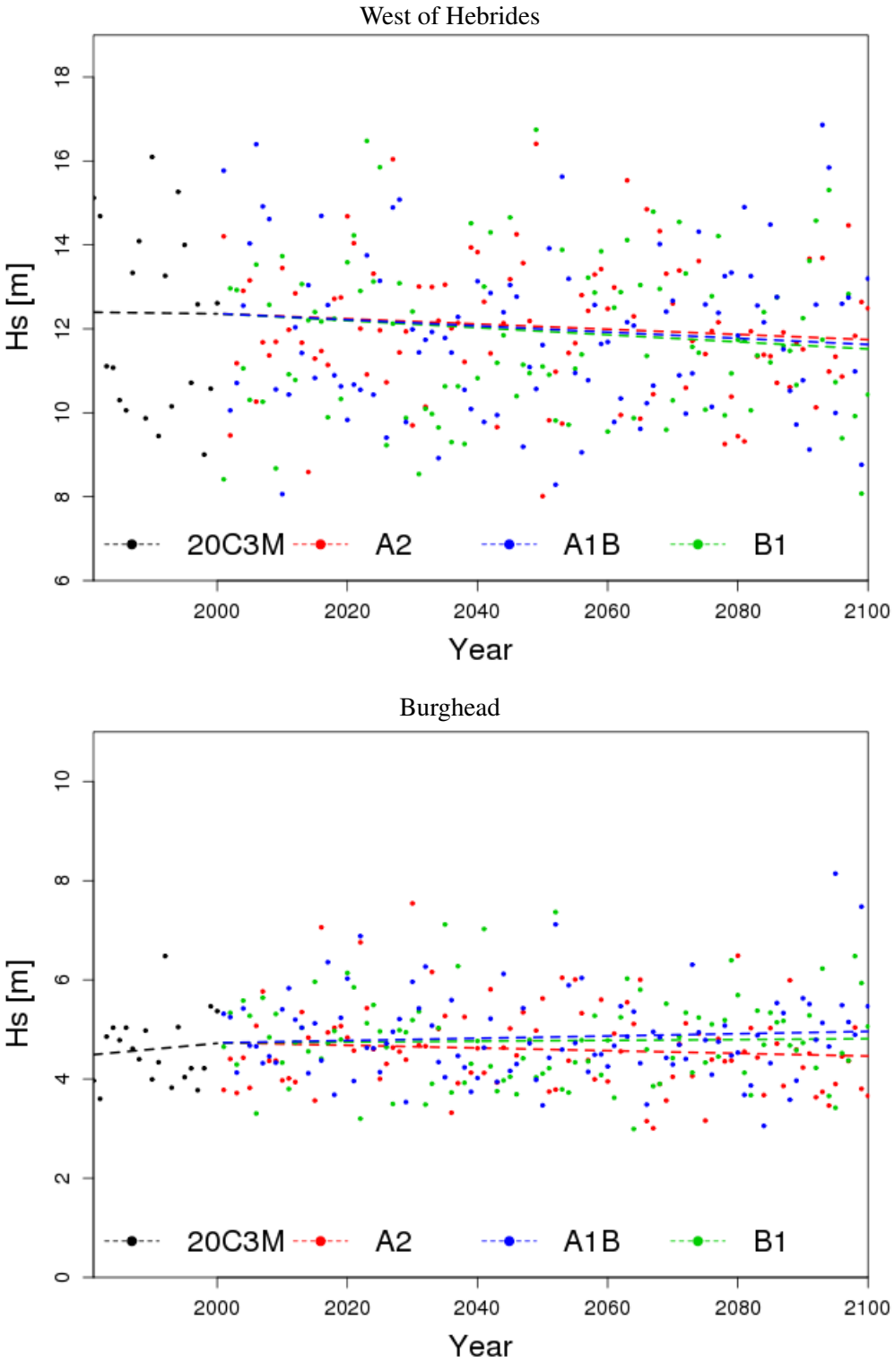

Figure 6.8: Projections of annual maximum H_s from the high (A2), medium (A1B) and low (B1) emissions scenarios

	Absolute change (cm/yr)			Relative change (%/yr)		
Site	A2	B1	A1B	A2	B1	A1B
Farr Point	-0.02	-0.04	-0.04	-0.013	-0.027	-0.023
EMEC	0.02	-0.03	0.00	0.011	-0.016	-0.002
West of Hebrides	0.00	-0.07	-0.02	-0.001	-0.023	-0.005
Galway Bay	0.00	-0.02	0.01	0.000	-0.021	0.007
Belmullet	-0.04	-0.08	-0.05	-0.014	-0.025	-0.015
Beatrice	-0.06	-0.04	-0.04	-0.040	-0.026	-0.026
Wave Hub	-0.05	-0.05	-0.03	-0.023	-0.025	-0.016
Burghead	-0.05	-0.03	-0.03	-0.054	-0.027	-0.030

Table 6.4: Relative and absolute trends observed in annual mean H_s for the high (A2), medium (A1B) and low (B1) emissions scenarios

	Absolute change ($\times 10^{-3}$s/yr)			Relative change (%/yr)		
Site	A2	B1	A1B	A2	B1	A1B
Farr Point	-0.71	-0.81	-1.59	-0.009	-0.011	-0.021
EMEC	-0.59	-0.73	-1.26	-0.007	-0.009	-0.015
West of Hebrides	-0.71	-0.88	-1.09	-0.008	-0.010	-0.012
Galway Bay	-0.54	-0.67	-0.76	-0.009	-0.011	-0.012
Belmullet	-0.98	-1.18	-1.15	-0.011	-0.013	-0.013
Beatrice	-1.58	-1.13	-1.86	-0.027	-0.020	-0.032
Wave Hub	-1.13	-1.17	-1.03	-0.013	-0.014	-0.012
Burghead	-1.79	-0.87	-1.93	-0.038	-0.018	-0.041

Table 6.5: Relative and absolute trends observed in annual mean T_e for the high (A2), medium (A1B) and low (B1) emissions scenarios

	Absolute change ($^\circ$/yr)			Relative change (%/yr)		
Site	A2	B1	A1B	A2	B1	A1B
Farr Point	0.14	0.08	0.09	0.055	0.030	0.035
EMEC	0.04	0.01	0.01	0.015	0.005	0.004
West of Hebrides	0.03	0.02	0.01	0.012	0.008	0.005
Galway Bay	0.05	0.03	0.04	0.022	0.013	0.016
Belmullet	0.04	0.03	0.02	0.016	0.011	0.008
Beatrice	0.15	0.08	0.17	0.089	0.050	0.098
Wave Hub	0.05	0.04	0.04	0.020	0.016	0.015
Burghead	0.14	0.06	0.16	0.096	0.043	0.107

Table 6.6: Relative and absolute trends observed in annual mean D_m for the high (A2), medium (A1B) and low (B1) emissions scenarios

	Absolute change (cm/yr)			Relative change (%/yr)		
Site	A2	B1	A1B	A2	B1	A1B
Farr Point	-0.16	-0.18	-0.15	-0.042	-0.047	-0.041
EMEC	-0.14	-0.13	-0.15	-0.027	-0.026	-0.030
West of Hebrides	-0.08	-0.03	-0.08	-0.013	-0.004	-0.012
Galway Bay	0.03	0.07	0.01	0.017	0.037	0.004
Belmullet	-0.12	-0.02	-0.13	-0.019	-0.004	-0.020
Beatrice	-0.13	-0.07	-0.09	-0.036	-0.020	-0.026
Wave Hub	0.05	0.12	-0.01	-0.011	0.027	-0.003
Burghead	-0.11	-0.05	-0.05	-0.048	-0.022	-0.024

Table 6.7: Relative and absolute trends observed in annual 95th percentile H_s for the high (A2), medium (A1B) and low (B1) emissions scenarios

	Absolute change (cm/yr)			Relative change (%/yr)		
Site	A2	B1	A1B	A2	B1	A1B
Farr Point	-0.54	-0.55	-0.45	-0.069	-0.071	-0.057
EMEC	-0.42	-0.61	-0.36	-0.040	-0.059	-0.035
West of Hebrides	-0.61	-0.84	-0.73	-0.050	-0.068	-0.059
Galway Bay	0.02	0.04	0.13	0.005	0.012	0.035
Belmullet	-0.23	0.02	0.11	-0.020	0.002	0.010
Beatrice	-0.67	-0.06	-0.08	-0.093	-0.009	-0.011
Wave Hub	-0.07	0.00	0.09	-0.009	-0.001	0.012
Burghead	-0.27	0.08	0.23	-0.057	0.017	0.048

Table 6.8: Relative and absolute trends observed in annual maximum H_s for the high (A2), medium (A1B) and low (B1) emissions scenarios

Site	A2	B1	A1B
Farr Point	0.458	0.964	0.827
EMEC	0.854	0.927	0.651
West of Hebrides	0.454	0.704	0.972
Galway Bay	0.759	0.808	0.687
Belmullet	0.160	0.404	0.590
Beatrice	0.004	0.836	0.329
Wave Hub	0.223	0.298	0.215
Burghead	0.001	0.934	0.432

(a) Mean annual H_s

Site	A2	B1	A1B
Farr Point	0.069	0.201	0.101
EMEC	0.112	0.404	0.106
West of Hebrides	0.005	0.272	0.044
Galway Bay	0.312	0.109	0.095
Belmullet	0.002	0.218	0.033
Beatrice	0.000	0.038	0.000
Wave Hub	0.055	0.244	0.019
Burghead	0.000	0.136	0.000

(b) Mean annual T_e

Site	A2	B1	A1B
Farr Point	0.000	0.144	0.006
EMEC	0.035	0.862	0.823
West of Hebrides	0.053	0.854	0.787
Galway Bay	0.013	0.116	0.267
Belmullet	0.016	0.305	0.894
Beatrice	0.000	0.017	0.000
Wave Hub	0.022	0.153	0.065
Burghead	0.000	0.053	0.000

(c) Mean annual D_m

Site	A2	B1	A1B
Farr Point	0.042	0.951	0.648
EMEC	0.123	0.799	0.921
West of Hebrides	0.030	0.524	0.970
Galway Bay	0.673	0.380	0.376
Belmullet	0.058	0.205	0.858
Beatrice	0.377	0.687	0.742
Wave Hub	0.623	0.296	0.725
Burghead	0.106	0.913	0.731

(d) Annual 95th percentile H_s

Site	A2	B1	A1B
Farr Point	0.765	0.346	0.473
EMEC	0.370	0.121	0.728
West of Hebrides	0.259	0.535	0.937
Galway Bay	0.823	0.425	0.418
Belmullet	0.193	0.880	0.823
Beatrice	0.069	0.987	0.888
Wave Hub	0.912	0.794	0.334
Burghead	0.131	0.752	0.882

(e) Annual maximum H_s

Table 6.9: P-values indicating statistical significance of trends in projections of mean annual (a)H_s, (b) T_e, (c) D_m according to the different emissions scenarios

In the analysis of high-energy events represented in terms of the annual 95th percentile H_s and annual maximum H_s, a decreasing trend was observed in general. The magnitude of these was found to be quite large in some cases, e.g. a decrease of 0.84m in maximum H_s between 2001 and 2100 at West of Hebrides if the B1 trajectory is followed. A small upwards trend was observed for all the scenarios at the Galway Bay site for both parameters. At other locations, a negative trend in one parameter is not necessarily accompanied by a negative trend in the other parameter, or vice versa. The statistical significance of these trends could not be established.

If the hypothesis is true that the trends in the wave climate were a result of atmospheric GHG levels, it would be reasonable to expect the envelope of wave climate projections to be bounded by projections for the A2 and B1 scenarios. While this is seen in some cases, e.g. mean annual H_s at West of Hebrides and Burghead, this observation is not consistent. Moreover, on account of the failure to establish statistical significance of trends for several measures of the wave climate, this manner of reasoning cannot be used with any confidence.

6.4.2 Regional analysis

The analyses demonstrated at site level in Section 6.4.1 are performed cell-by-cell on the gridded data of projected wave parameters for the three emissions scenarios to obtain the absolute and relative changes over the period 2001-2100. The trends and and changes obtained are presented as contour maps of the region in Figures 6.9 to 6.11 for annually averaged H_s, T_e and D_m, and in Figures 6.12and 6.13 for extreme conditions represented by the annual 95th percentile and annual maximum H_s.

Mean annual significant wave height

In Figure 6.9, from the plots of changes in mean annual significant wave height, similarities in spatial patterns can be observed across all three scenarios. In general, a decreasing trend in the parameter can be observed in the North East Atlantic with the magnitude of the trend varying from -0.2cm/yr in oceanic waters to 0cm/yr along the coast of Ireland and western Scotland. For the high emissions (A2) scenario, the trend becomes positive west of the Hebrides indicating an increase of as much as 0.04cm/yr.

A small zone with an upwards trend of approximately 0.06 to 0.08 cm/yr is observed west of the Shetland Islands for the A1B and A2 scenarios. A similar zone is also seen east of Orkney for the A2 scenario.

In the North Sea, a decrease in the significant wave height is observed along the UK coast with similar spatial patterns for all three climate change scenarios. The trend gradually increases eastwards to 0.06 to 0.08 cm/yr along the coasts of Denmark and Norway. It is interesting that the contours in these figures, especially indicating the relative change, appear to follow the meridians of longitude in the North Sea.

While the magnitudes of the spatial patterns and trends, positive and negative, appear larger for the A2 scenario, little difference is observed between the A1B and B1 scenarios, making it difficult to attribute these changes to atmospheric greenhouse gas levels.

Mean annual energy period

From the trends in mean annual T_e presented in Figure 6.10, similarities in spatial patterns can be observed between the A2 and A1B data. In general, a decreasing trend can be said to prevail, especially in the North East Atlantic and in the North Sea along the UK coast. A small upwards trend of approximately 0.5×10^{-3}s/yr is observed off the Danish coast for the A2 and A1B scenarios. A negative trend is seen across the region for the B1 scenario varying from 0s/yr off the coast of Northern Ireland to -2.5×10^{-3}s/yr between Wales and Ireland.

It should be mentioned that the calculated trends were found to be very weak and statistical significance could not be established. This is in line with the findings for annual mean T_e for individual sites in Section 6.4.1. While it may be that the trends calculated are a result of short term and internal variability, the existence of distinct spatial patterns would suggest otherwise. It appears that more detailed investigation into these is required, and until then it is advised that the contour plots and any information presented for T_e be treated with caution especially as the coupled model fared poorly in the evaluation of contemporary climate in Section 6.3.

Mean annual wave direction

It can be seen in the plots of absolute changes in mean wave direction in Figure 6.10, that an upwards trend exists in most parts of the region, off the west coast of Norway being an exception. The strength of the trend in oceanic waters is seen to increase northwards as the GHG trajectory varies from the B1 scenario to the A2 scenario via the A1B scenario.

A similar variation to scenario can be observed off the Scottish coast in the North Sea, especially in the Moray Firth, with the magnitude of the trend increasing from 0.1° per year following the B1 scenario to 0.175° following the A2 scenario. Although most noticeable off the coast of Scotland, this pattern can be seen to exist even southwards, off the coast of England.

Annual 95th percentile and maximum significant wave height

An examination of the contour maps for the 95th percentile H_s presented in Figure 6.12 reveals that some similarity exists between the spatial variation of the trends in the A2 and A1B scenarios in oceanic waters. In oceanic waters a decreasing trend can be said to prevail for these two scenarios. However, in the same regions, a weak upwards trend is observed for the B1 scenario. In the North Sea, however, the spatial patterns appear to differ for all three scenarios, as evidenced by the plots of the relative changes.

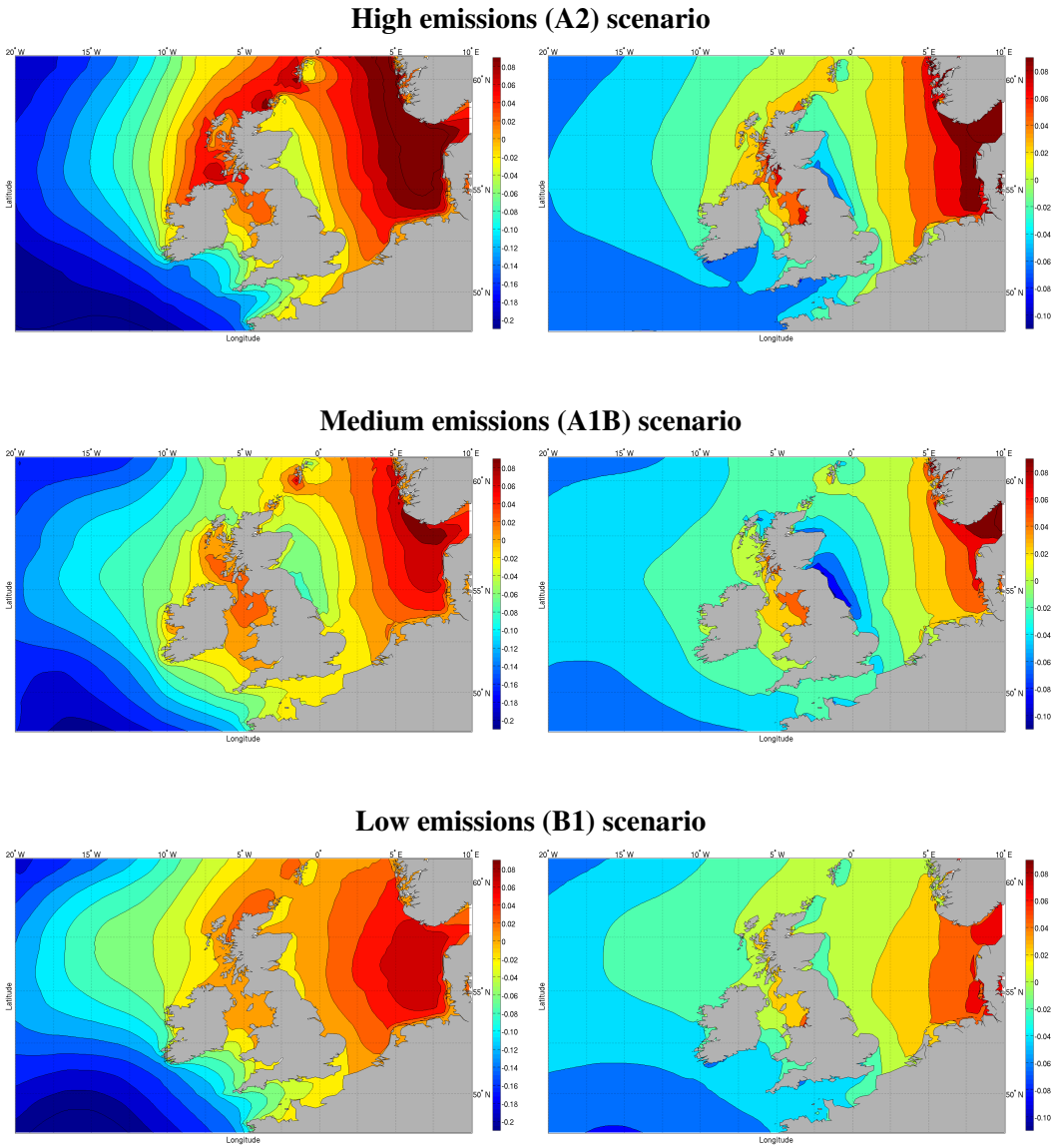

Figure 6.9: Absolute (left) and relative (right) changes per year in cm/yr and %/yr respectively in annual mean significant wave height

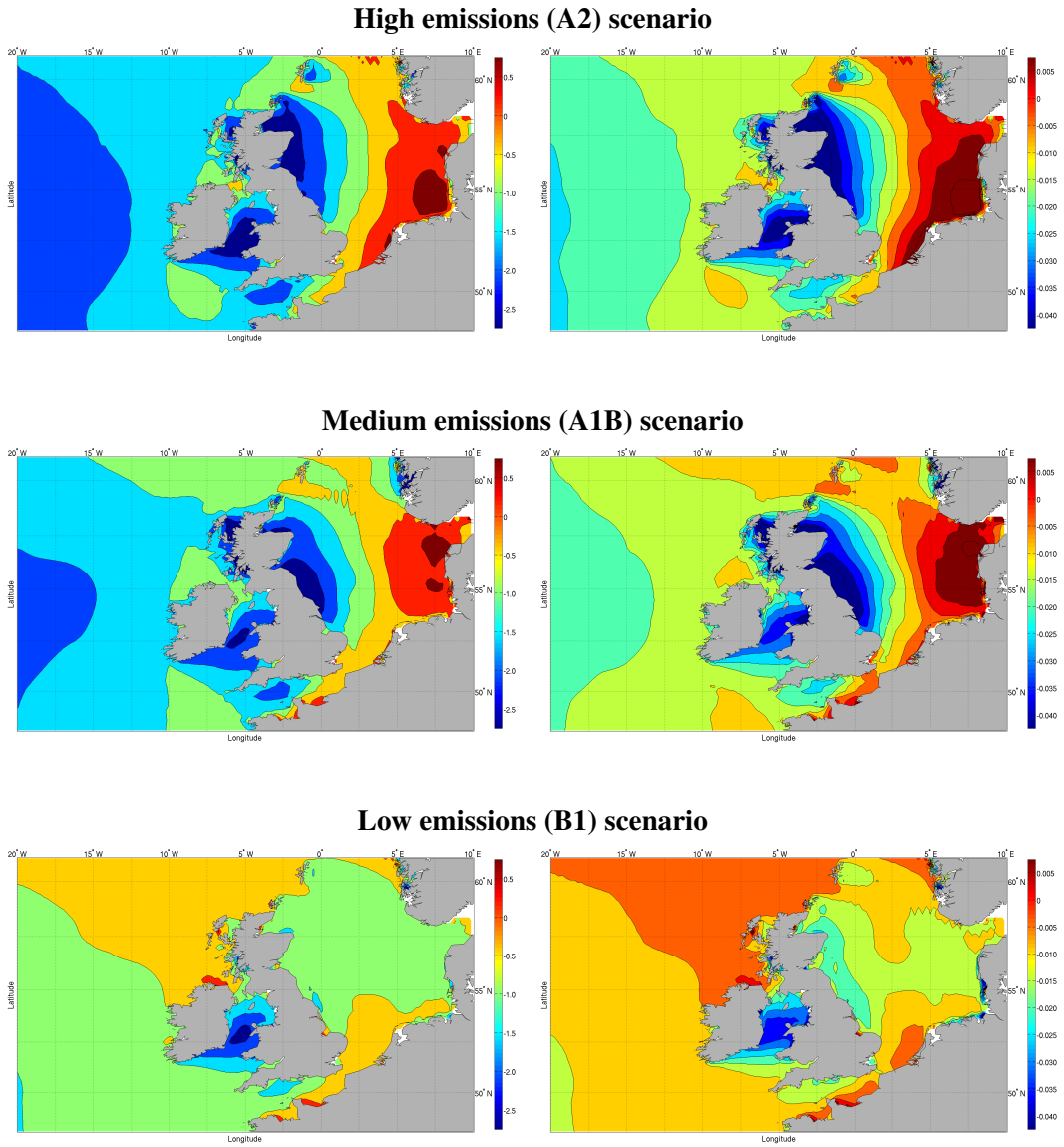

Figure 6.10: Absolute (left) and relative (right) changes per year in $\times 10^{-3}$s/yr and %/yr respectively in annual mean energy period

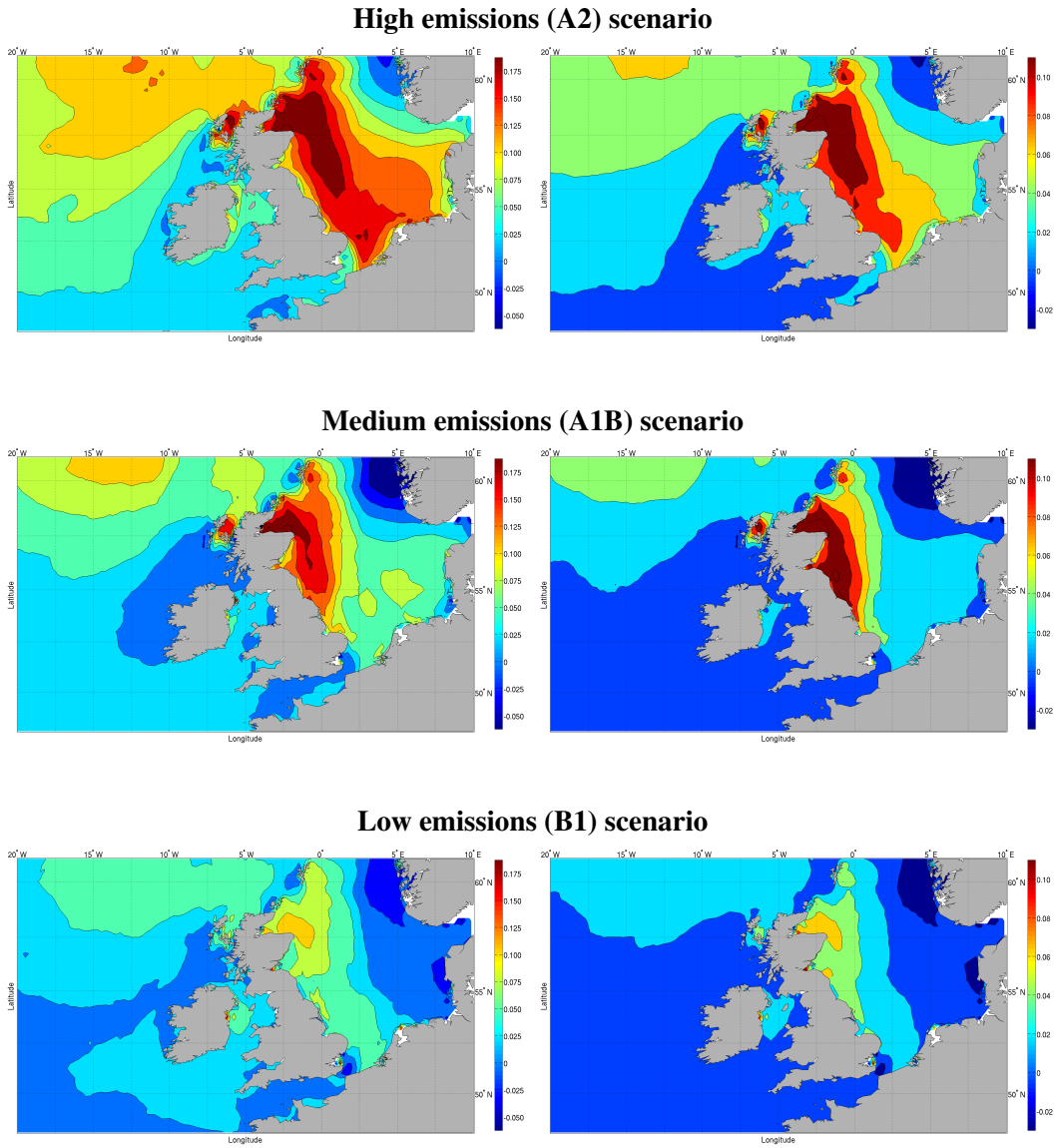

Figure 6.11: Absolute (left) and relative (right) changes per year in °/yr and %/yr respectively in annual mean wave direction

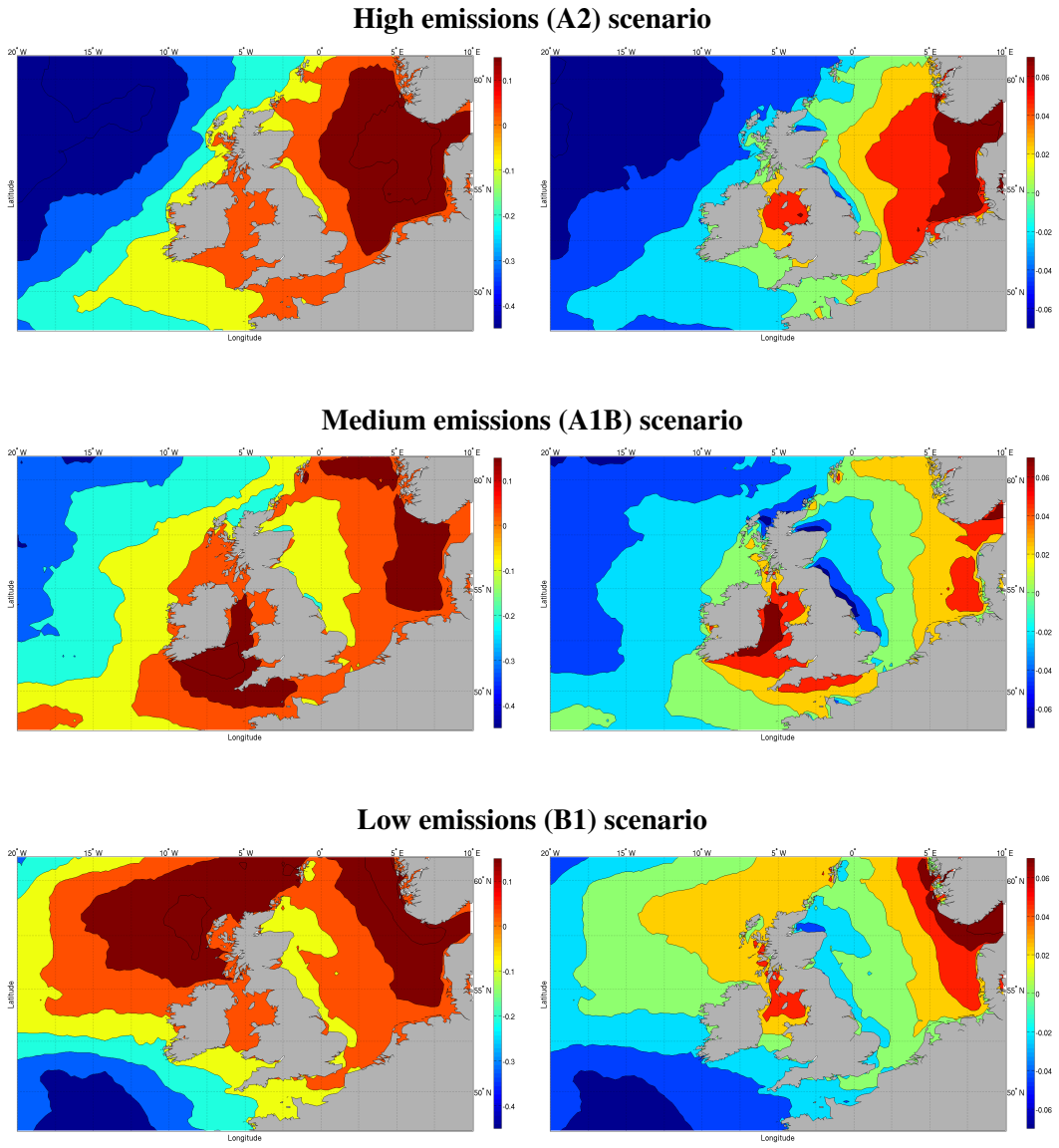

Figure 6.12: Absolute (left) and relative (right) changes per year in cm/yr and %/yr respectively in annual 95th percentile significant wave height

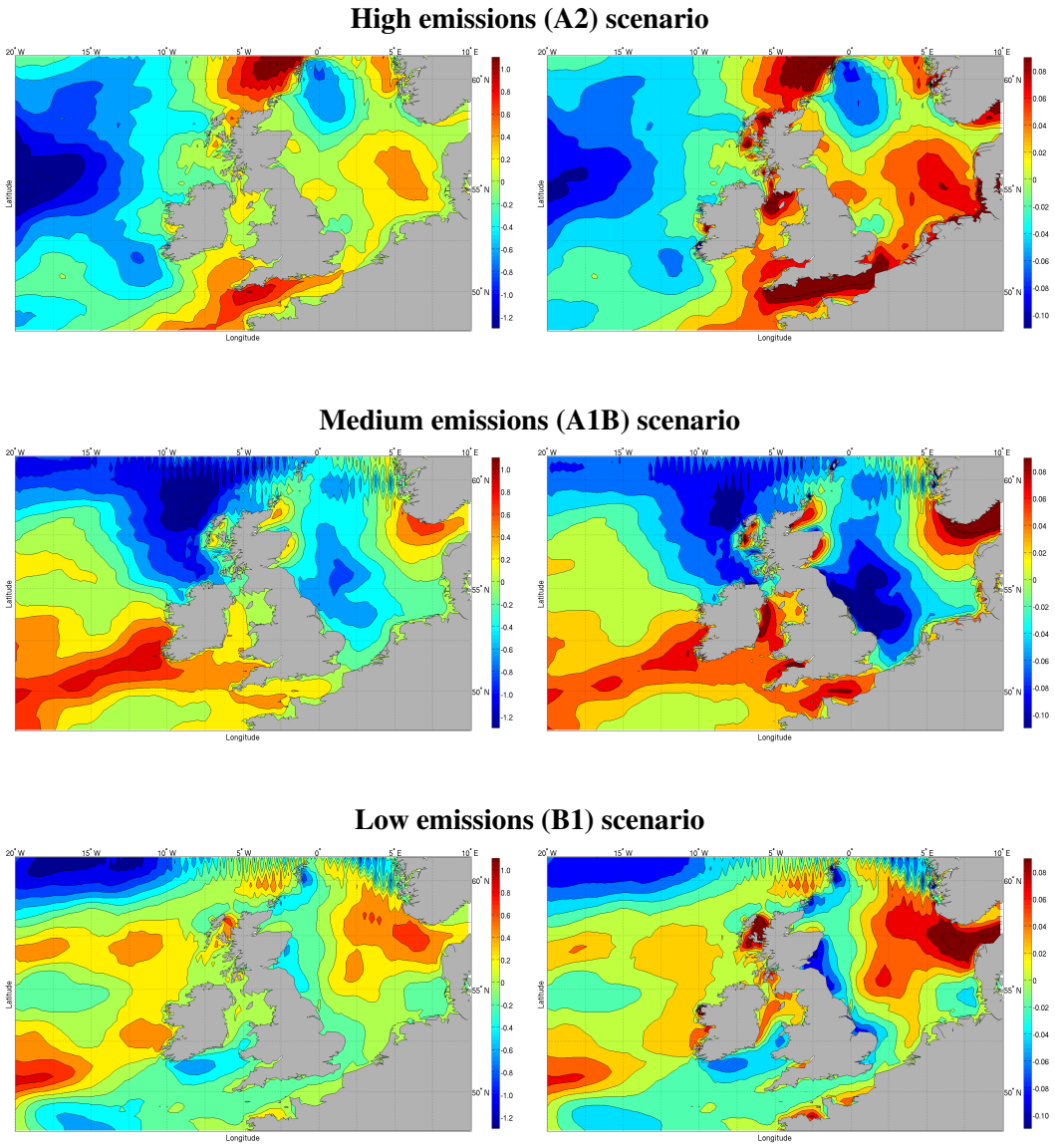

Figure 6.13: Absolute (left) and relative (right) changes per year in cm/yr and %/yr respectively in annual maximum significant wave height

The similarities between the data for the medium and high emissions scenarios may be explained by the fact that the atmospheric GHG levels are not too different, at 600-700ppmv and 600-850ppmv respectively between 2071 and 2100. The GHG levels for the B1 scenario appear to remain stable between 525 and 550ppmv during the same period. If these differences in level were to explain the trends observed, it may be inferred that in oceanic waters, increasing the GHG levels might increase the average intensity of storms to a point. Increasing the GHG levels further would sharply decrease the average intensity of storms, assuming that the hypothesis of a causal link between the two is true.

Such an inference cannot be made for the annual maximum wave height. There appear to be no similarities between the spatial patterns and contours plotted for H_{max} (see Figure6.13) for any of the emissions scenarios. From the site level analysis of this parameter and from the contour maps, it may be inferred that the weak trends observed are likely a result of internal variability of the system.

From the spatial patterns and the magnitudes of variation in the regional plots, it would appear that a systematic trend might exist between the projection datasets on account of the differences in the scenarios, e.g. in mean wave direction.

On account of the low confidence in the results for T_e, analysis of derived parameters such as mean wave power density is not presented.

6.5 Attribution of wave climate changes

6.5.1 Analysis of Variance (ANOVA)

The Analysis of Variance (ANOVA) (Fisher, 1925) is a methodology developed for the study of agricultural experiments in the 1920s, but has, since, found wide application in the fields of biology (Underwood, 1981), psychology (Kiecolt-Glaser *et al.*, 1984), pharmacology (Cleophas *et al.*, 2006) and engineering (Golinkin *et al.*, 1997), to name a few.

In ANOVA, the means of three or more samples (or groups) are compared to determine whether they belong to the same population. It is an extension of the t-test to include more than two groups of data. ANOVA can also be used when the effect of one or more factors is to be detected provided that each factor has at least three levels.

The simple null hypothesis in this test is that there is no statistically significant difference in the means of the groups of observations. When used to determine the effect of a factor (or treatment) on the outcome, the null hypothesis that the factors or treatments have no effect on the outcomes is tested.

In order to assess whether the means of the groups are statistically different from one another, ANOVA uses the variation between groups and the variation within the groups to obtain the F-

statistic. For a dataset with k groups having n elements in each group, the F-value is calculated by (Sahai and Ageel, 2000)

$$\mathrm{MS_B} = \frac{1}{k-1}\left[\sum_{j=1}^{k}\left(\frac{\left(\sum_{i=1}^{n} y_{ij}\right)^2}{n}\right) - \frac{\left(\sum_{j=1}^{k}\sum_{i=1}^{n} y_{ij}\right)^2}{nk}\right] \tag{6.1}$$

$$\mathrm{MS_W} = \frac{1}{n-k}\left[\sum_{j=1}^{k}\sum_{i=1}^{n} y_{ij}^2 - \sum_{j=1}^{k}\frac{\left(\sum_{i=1}^{n} y_{ij}\right)^2}{n}\right] \tag{6.2}$$

$$F = \frac{\mathrm{MS_B}}{\mathrm{MS_W}} \tag{6.3}$$

In Equation 6.3, the numerator ($\mathrm{MS_B}$) is the between-group variance which describes how much the different groups vary from one another. The denominator ($\mathrm{MS_W}$), the within-group variance is a measure of the average variance within each sample. The numerator is also considered as a measure of sampling error as well as differences between groups due to applied treatments and the denominator is considered purely as a measure of sampling error (Lewin, 1999).

In the event of the null hypothesis being true, the difference between groups due to the treatments would be 0, and both, the numerator and denominator would measure the sampling error. As a consequence, F would approach unity.

For a selected significance level (e.g. $\alpha = 0.05$), knowing the number of groups (k) and number of elements in each group (n), the critical F-value can be obtained either from the F-distribution or from equivalent tables. If the calculated F-value exceeds the critical F-value, the null hypothesis can be rejected.

In this section, one-way ANOVA is used to determine whether the wave climate is sensitive to greenhouse gas (GHG) forcings by analysing the projected wave climate corresponding to the high (A2), medium (A1B) and low (B1) emissions scenarios for the 21^{st} century.

6.5.2 Sample selection

The attribution analysis reqiures appropriate samples to be selected from the timeseries data obtained for the different emissions scenarios. It can be seen from Figure 2.5 in Chapter 2 that the atmospheric concentrations of greenhouse gases (GHGs) for the three scenarios are vary distinctly from one another between 2071 and 2100.

From the GHG concentration trajectories in the figure, the low (B1) emissions scenario, atmospheric CO_2 levels show little variation between 2071 and 2100, stabilising at $\approx 525 - 550$ parts per million by volume (ppmv). Following the medium (A1B) and high (A2) emissions marker scenarios, however, the GHG concentrations show greater variations, increasing from $\approx$ 600ppmv to 700ppmv and $\approx$ 600ppmv to 850ppmv respectively.

Based on this difference in CO_2 levels among the scenarios, data from 2071-2100 was selected for comparison. As this time-slice is 70 years apart from the common initialisation point, it is expected that the climate variables will be sufficiently different from each other based on the effect of inputs as defined by the scenarios. The descriptors of the wave climate identified earlier are derived from the hourly timeseries of wave data from this time-slice for analysis.

6.5.3 Results of the ANOVA

The Analysis of Variance was used to compare the samples of wave parameter data for the time-slice extending from 2071 to 2100. This test indirectly evaluates the effect of the scenario on the wave climate during the period. If the samples are significantly different from one another, indicated by the F-value and p-value yielded by the test, then it can be concluded that the difference in the description of the scenarios affects the wave climate.

	H_s		T_e		D_m	
	F-value	P-value	F-value	P-value	F-value	P-value
Farr Point	0.0013	0.9270	0.5150	0.5990	2.0660	0.1330
EMEC	0.1240	0.8830	0.5410	0.5840	1.5780	0.2120
West of Hebrides	0.0590	0.9430	0.6630	0.5180	1.0510	0.3540
Galway Bay	0.1960	0.8220	0.4290	0.6520	0.9430	0.3930
Belmullet	0.0150	0.9850	0.6360	0.5320	2.0060	0.1410
Beatrice	0.9770	0.3810	1.7990	0.1720	2.8780	0.0618
Wave Hub	0.0760	0.9270	0.6170	0.5420	0.3060	0.7370
Burghead	1.8060	0.1700	4.5830	0.0128	3.3480	0.0397

Table 6.10: Computed F-values and p-vales from the ANOVA method applied to samples of mean wave parameter data from A2, A1B and B1 scenarios

The samples of mean wave parameters (H_s,T_e and D_m) for the climate change scenarios subjected to the ANOVA method are shown as box plots for the sites West of Hebrides and Burghead in Figure 6.14 and for the remaining sites in Figures C.1 to C.15 in Appendix C. The F-values and p-values obtained from the test are also tabulated in Table 6.10.

The plots show the sample mean and the 50 and 98 percentile ranges as the box and whiskers respectively along with outliers. It can be seen in these plots that the means of the samples lie very close to one another, but more importantly, the interquartile range (indicated by the box) of each sample would include the mean value of any other sample for a given wave parameter. This is a visual indication of the absence of statistical significance between data from the different scenarios.

The critical F-value for an experiment with 3 samples with 30 elements each is 2.489 at $\alpha = 0.10$ and 3.316 at $\alpha = 0.05$. An examination of the obtained values for the 8 sites in Table 6.10 reveals that for most sites, the samples are not different in a statistically significant sense for any wave parameter. In the North Sea sites, however, statistical significance is seen for some

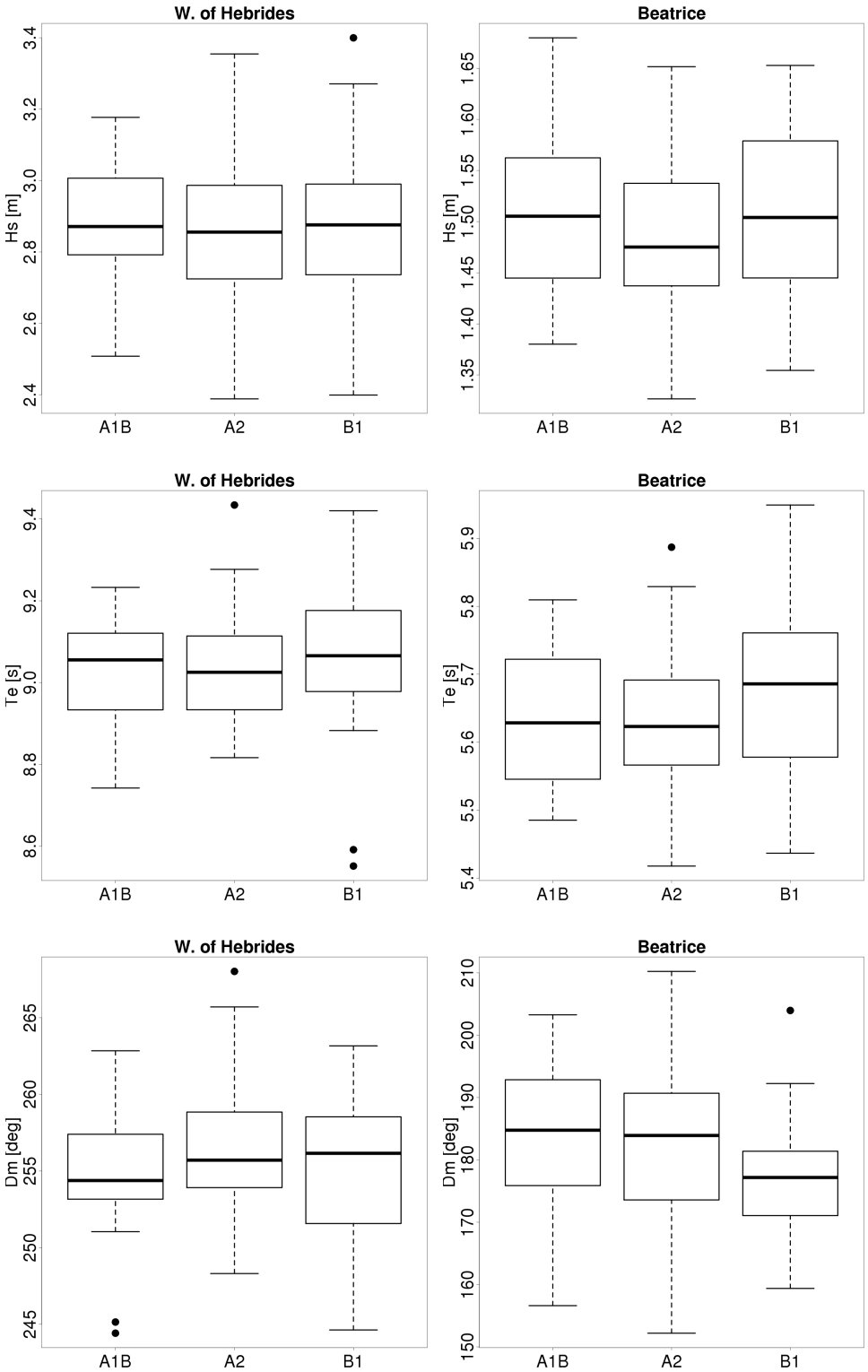

Figure 6.14: Boxplots of samples of mean wave parameters used in the ANOVA test shown for 2 sites.

parameters. At Burghead, the F-values for T_e and D_m exceed the critical value and the same can be noted for D_m at Beatrice.

Further investigation of these reveals that for all three occasions noted above where the null hypothesis can be rejected, the sample B1 differs significantly from the other two samples. However, in all three cases, the samples A2 and A1B cannot be said to differ significantly.

The ANOVA test was also applied recursively to each cell in the gridded dataset to assess statistically significant differences that may exist at points apart from the selected locations as well as to detect any spatial patterns that may exist. Some areas of statistical significance were identified in the analysis of D_m and here too, the B1 scenario was found to be statistically different from the other scenarios. These data are plotted as contour maps in Figure 6.15, where the blue areas indicate regions with statistically different samples.

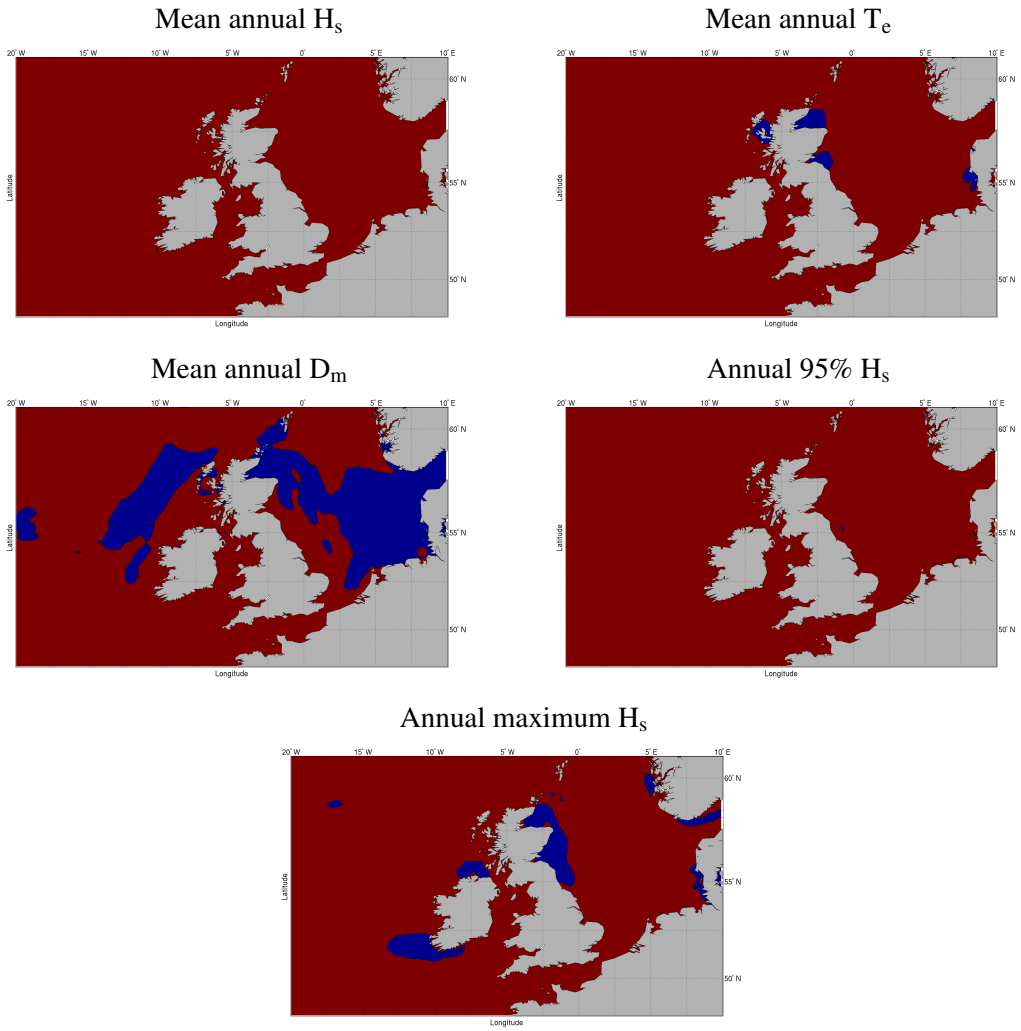

Figure 6.15: Results of ANOVA applied to regional datasets from the A2, A1B and B1 samples for 2071-2100. Blue - statistically significant; Red - not statistically significant

From the small F-values and the consequently large p-values obtained for the 8 test locations as well as the evidence in Figure 6.15, it can be inferred that it is highly unlikely that atmospheric GHG concentrations have an effect on the wave climate in the region.

6.6 Chapter summary

- A high resolution dataset of projected wave climate following the high (A2), medium (A1B) and low (B1) emissions scenarios was generated for 2001-2100 by forcing WAVEWATCH III with sub-daily winds from a ECHAM5.
- The contemporary climate as simulated from hindcast winds (using the 20CR dataset) was compared with GCM winds for the A1B scenario to evaluate the performance of the ECHAM5-WWIII coupled model. Despite some underestimation, good agreement was observed in spatial patterns for H_s and D_m. The model fared poorly in the estimation of contemporary T_e in both magnitude as well as spatial patterns.
- Trends in the wave climate were computed by fitting a regression function using the least-squares method. This was done for individual sites as well as on a regional level.
- The effect of atmospheric greenhouse gas (GHG) levels on the wave climate was tested at site and regional levels by subjecting samples from the A2, A1B and B1 futures to the Analyis of Variation (ANOVA).
- For most locations, no statistically significant difference was found from this analysis. It may be inferred from this that it is unlikely that a causal link exists between GHG levels and the wave climate. This was found to hold in general on a regional level as well.
- At the regional level, it was observed that the B1 future was statistically different from the other two scenarios in some areas, especially in annual mean D_m.

Chapter 7

Extreme wave estimates for the 21st century

7.1 Introduction to extreme wave analysis

For the safe and successful operation of wave, tidal and offshore wind energy devices, a better knowledge of the environment, particularly the extreme conditions, is required. This is especially true as the financial risk associated with failure of offshore structures is substantial. This need was recognised several decades ago, then for its importance to offshore oil platforms, and consequently there have been many studies on the subject. A brief overview of methods and data previously employed is presented in Section 7.2.

Extreme waves, often referred to in terms of m-year return values, can be interpreted simply as wave heights that are likely to be exceeded, on average, once in every 'm' years. While the design life of a marine energy converter is usually about 20 to 30 years, estimates of return values over longer periods, typically 50 or 100 years, are considered in the design process.

Accurate estimation of these design extremes is essential to ensure survivability as well as economic viability of the marine energy converters. An underestimation of the maximum wave height for design could lead to catastrophic failure of the device, while a conservative overestimation would result in needless overdesign of the device, making the capital costs prohibitively high.

In this chapter, the previously produced long term hindcast of wave climate (see Chapter 4) is analysed to estimate 100-year return periods. The results of this work are compared with other published estimates and the efficacy of the wave model in extreme analysis is studied. The analysis considers only extrema in wind generated seas and swells and not storm surges and rogue waves generated by resonant nonlinear wave interactions.

The following sections present a brief historical overview of extreme wave estimation and a description of the peaks-over-threshold (POT) method. The identification of suitable distributions, thresholds and datasets and calibration process are described in Section 7.5 which are applied to the gridded dataset of significant wave height for the North Atlantic Ocean and North Sea

to produce a regional map of 100-year return value estimates for H_s. Finally, some aspects of the study pertaining to the methodology are discussed in Section 7.7. The implications of the results of this study on wave energy converters (WECs) and other marine structures in the 21^{st} century are discussed in Chapter 8.

7.2 Extreme wave analysis - A review

7.2.1 Methods

Initial Distribution Method (IDM): An early method of estimating extreme waves was the initial distribution method (IDM) in which a suitable parametric model was fit to all the available H_s data. In order to estimate the extreme value, which likely lies outside the range of the sample, the model is extrapolated to the desired probability level and the corresponding value of H_s is taken as the estimate. The Gumbel, Weibull, and the log-normal parametric models were popular for describing the sample of H_s data (Vinoth and Young, 2011).

A significant drawback of the initial distribution method was the lack of scientific justification for using one parametric model over another. While goodness-of-fit tests like the χ^2 and Kolmogorov-Smirnov tests provide some measure of how well a distribution fits the bulk of the sample, they are not indicative of how well it represents the tail. Thus, extrapolating the distribution for extreme events may yield incorrect estimates (Ferreira and Guedes Soares, 1998; Vinoth and Young, 2011). In addition, the sample consists of all H_s measurements, usually from a more or less contiguous time-series. As wave heights exhibit the Poisson property, the sample violates the assumption of independence of data. However, despite the drawbacks of IDM, it was an often-used approach for extreme wave estimation (e.g. Goda (1988); Tucker (1991); Ochi (1992)).

Block Maxima Method: The block maxima method overcomes the limitations of the initial distribution method. In this method, the data is divided into blocks and only the maximum value from each block is used in the analysis. According to the extreme value theory, the cumulative distribution of the block maxima will follow the generalized extreme value (GEV) distribution (Castillo and Sarabia, 1992).

The blocks are usually of fixed length and the time-series is divided into monthly (e.g. Carter and Challenor (1978)) or annual blocks. In the latter case, the block-maxima method is also referred to as the annual maximum method (AMM). If monthly maxima are used, the seasonal variations in wave height would need to be taken into account and corrected to get a good fit with the parametric model.

In the block maxima method, independence of the sample is ensured by the sampling method and the distribution fitting is based on statistically sound principles. The drawback of this

method, however, is that the sample size is often too small to fit a distribution with reasonable confidence.

Peaks-over-threshold (POT) method: This method has come to be accepted as the best compromise for extreme wave height estimation. In this method, a suitably high threshold value is selected, and all the peaks in the measured data exceeding the threshold are analysed (Ferreira and Guedes Soares, 1998). According to the extreme value theory, the distribution of maxima exceeding the threshold follows the generalised Pareto distribution (GPD) (Castillo and Sarabia, 1992; Coles, 2001).

Apart from the GPD, Gumbel and Weibull distributions have also often used to fit the exceedances (e.g. Haver and Nyhus (1986); Mathiesen *et al.* (1994); Neelamani *et al.* (2006)). This is because the distribution of maxima reduces to a Gumbel distribution when the parent distribution describing the complete sample is either Weibull or log-normal (Coles, 2001).

A detailed description of the POT methodology, as applied in this study, is presented in Section 7.3.

7.2.2 Data Sources

Limited by the data available, early analyses of extreme waves were conducted on a site-by-site basis, for example, at the Dowsing lightship (Draper, 1976) and the Seven Stones lightship and the rescue ship Famita (Carter and Challenor, 1978). These studies used measurements of wave heights from shipborne wave recorders with the length of data ranging from a single year (Draper, 1976) to approximately a decade (Carter and Challenor, 1978). In these studies, the significant wave height is calculated as the mean of the highest third of observed wave heights.

In the earliest instance of regional estimation of design wave heights for UK waters, wave forecast charts supplemented by instrument data were used and return values interpolated between sites (Draper, 1973). In the last three decades, such regional analyses have been performed using satellite measurements of wave heights and hindcast data from numerical models.

Long term datasets of wave height from altimeters from several satellite missions like GEOSAT, ERS-1, ERS-2 and TOPEX, have been available since March 1985. These datasets provided along-track measurements of significant wave height on a global scale which could be used to study global climatology and variability of ocean waves. The use of satellite data for estimating extreme wave heights has been demonstrated by several studies (e.g Carter (1993); Panchang *et al.* (1998); Anderson *et al.* (2001)). There are, however, some difficulties associated with this data, principal among them the fact that the data for any location or gridpoint is not contiguous and contains gaps equal to the ground track repeat period. For example, GEOSAT had a ground track repeat period of 17 days, which is very large and it is possible that storms may be missed in the altimeter records.

Hindcast wave data from numerical models have been used as early as 1978, when a second generation numerical wave model was used to simulate data for 42 storms in the period 1966 to 1976 to estimate 50-year wave heights in the North Sea (Ewing *et al.*, 1979). More recently, with advances in numerical modelling, other studies have used data from numerical models with similar results to each other (e.g. ERA-40 (Caires and Sterl, 2005), NEXTRA (Williams, 2008) and ERA-I (Agarwal *et al.*, 2013)).

7.2.3 Theoretical Distributions

In recent years, it has become common practice to use the peaks-over-threshold method along with a suitable distribution to model the sample of maxima or exceedances. A brief theoretical background on two such distributions - the Generalised Extreme Value (GEV) distribution and the Generalised Pareto Distribution (GPD) - is presented here.

The Generalised Extreme Value (GEV) distribution is a family of distributions comprising of the Gumbel, Frechet and Weibull probability density functions. It is sometimes referred to as the Fisher-Tippett (FT) distribution and the three constituents are the FT-I, FT-II, and FT-III distributions. It is used to model the tails (maxima or minima) of samples of random variables, and hence finds application in the estimation of extreme waves. The cumulative distribution function (cdf) of the GEV distribution is given by:

$$G(y) = \exp\left\{ -\left[1+\xi\left(\frac{y-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}} \right\} \tag{7.1}$$

where y is an element of the ordered set of maxima (such that $y_1 \leq y_2 \leq ... \leq y_k$), μ is the location parameter, σ is the scale parameter and ξ is the shape parameter. The cdf reduces to the Gumbel, Frechet and Weibull forms when $\xi = 0$, $\xi > 0$ and $\xi < 0$ respectively.

The Generalised Pareto Distribution (GPD) is another family of continuous probability distributions that can be used to model exceedances above a sufficiently high threshold value for a sample of random variables (Coles, 2001). This distribution has gained popularity in extreme wave estimation when applied in conjunction with the POT approach to derive m-year return values (e.g. (Anderson *et al.*, 2001; Caires and Sterl, 2005)).

For fitting a GPD, given a threshold μ, the set of exceedances can be found by:

$$z_i = y_i - \mu \tag{7.2}$$

where y_i is an element from the ordered set of measurements exceeding the threshold (such that $y_1 < y_2 < ... < y_k$). The ordered set of exceedances can then be described by the Generalised Pareto Distribution (GPD) whose cumulative distribution function (cdf) is given by:

$$G(z) = 1 - \left(1+\frac{\xi z}{\sigma}\right)^{-\frac{1}{\xi}} \tag{7.3}$$

where σ is the scale parameter, ξ is the shape parameter and $\xi \neq 0$.

7.3 Peaks-over-Threshold (POT) methodology

In this study, the peaks-over-threshold method is applied for the estimation of 100-year return values of significant wave height. This section briefly explains the methodology of the POT approach, which is also presented as a flowchart in Figure 7.1.

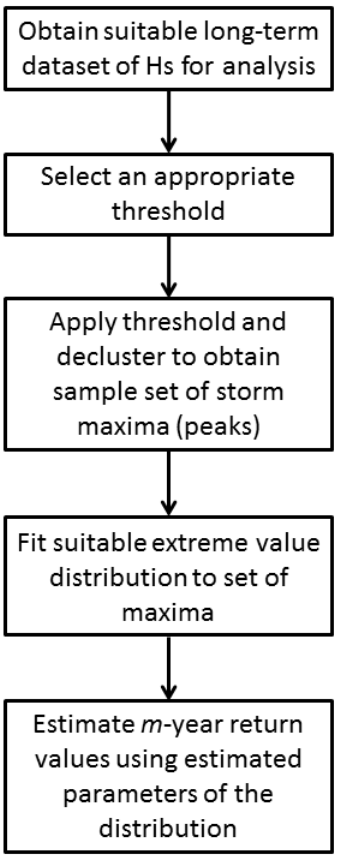

Figure 7.1: Flowchart describing the Peaks-Over-Threshold (POT) method.

7.3.1 Dataset selection and sample size

For POT analysis, time-series data obtained from buoys, satellites or numerical models may be used for extracting the sample of exceedances for analysis. It is desirable that this dataset is continuous, extends over a sufficiently long period and is of high temporal resolution. EquiMar (2011) recommends that the dataset for analysis spans no less than 20% of the return period, implying that the minimum period of data for 50-year and 100-year return value estimation be no less than 10 years and 20 years respectively, and according to Goda (2010) the period of data should be no less than 10 years for a 100-year return period. It was also found when

estimated for wave energy sites (e.g Wave Hub, EMEC and Outer Hebrides) that the 100-year return value exhibited little change when the period of data was increased beyond 15 years (Agarwal *et al.*, 2013).

7.3.2 Selection of a suitable threshold

The selection of a suitable threshold is key in the POT approach. A single value (e.g. 5m for the North Sea) cannot be selected as the wave regimes vary significantly from one location to another due to several factors such as local bathymetry, fetch and latitude. The threshold must be selected for each location (or grid point) based on its local wave climate, making sure that the threshold

- is sufficiently large to ensure that the extreme value distribution fits the data well; and
- is low enough so that the size of the sample analysed is adequate.

Several methods have been proposed in the past for selecting a suitable threshold such as visual inspection of the time-series, a robustness estimator based approach (Dupuis, 1998) and Bayesian methods (Tancredi *et al.*, 2006). In the context of this study, these methods are computationally too expensive and complex for recursively applying cell-by-cell to the gridded dataset. Simpler methods that have been successfully used in the past include selecting the threshold for a predetermined size of the sample of exceedances (Wimmer *et al.*, 2006), or based on a high quantile of the total sample (e.g. 93% quantile (Caires and Sterl, 2005)).

7.3.3 Declustering

In the POT method, it is recommended that the set of exceedances be declustered in order to prevent individual storms from dominating the analysis. The process of declustering ensures that only the peak value recorded during high-energy events is used, thereby ensuring that each storm occurrence is equally represented by the parametric model.

Declustering of the timeseries is achieved by first extracting elements that are not less than the selected threshold to obtain the sample of exceedances. A minimum period between peaks is specified, usually 3 to 6 days depending on the mean storm duration, and maxima are selected from the sample of exceedances, such that they are apart from each other by at least this period. In the event that there are multiple maxima within the specified separation period, the largest value is taken. This process yields the declustered sample of storm maxima for analysis.

7.3.4 Distribution fitting and return value estimation

Once the sample of peaks has been isolated from the timeseries, a suitable extreme value distribution, usually from the GEV family or the generalised Pareto distribution, can be fitted. The shape, scale and location parameters of the GEV and GP distributions can be estimated using the Maximum Likelihood Estimation (MLE) method, made easy by the use of available software (Pfaff and McNeil, 2012). Prior to the grouping of distributions in the GEV family, Gumbel and Weibull distributions were popularly used to describe the sample (Mathiesen *et al.*, 1994) (Neelamani *et al.*, 2006). Now, such a choice is unnecessary as estimating the parameters of the GEV family will automatically yield the best fitting constituent distribution which may be identified by the shape parameter (ξ).

Once the parameters of the extreme value distribution have been estimated, the m-year return value can be calculated by using the equations 7.4 and 7.5 for GEV and GPD respectively (Coles, 2001).

$$H_m = \begin{cases} \mu - \sigma\left[-\left(\ln\frac{1}{m\lambda}\right)\right] & \text{if } \xi = 0 \\ \mu - \frac{\sigma}{\xi}\left[1 - \left\{-\ln\left(1-\frac{1}{m\lambda}\right)^{-\xi}\right\}\right] & \text{if } \xi \neq 0 \end{cases} \tag{7.4}$$

$$H_m = \mu + \frac{\sigma}{\xi}\left[(m\lambda)^{\xi} - 1\right] \tag{7.5}$$

where ξ, σ and μ are the shape, scale and location parameters respectively and λ is the average number of exceedances per year. μ is taken as the threshold value for a GPD. If k is the total number of exceedances in the sample, n is the number of years over which the sample spans, then $\lambda = k/n$.

7.4 Evaluation of fits

Once the parameters of a theoretical distribution have been estimated for the sample of data, there are several ways of testing the goodness of fit. These include visual inspection of the probability curves and quantile (QQ) plots as well as quantitative goodness-of-fit tests such as the chi-squared (χ^2) test and Kolomogorov-Smirnov (K-S) test. Apart from providing a measure of how well the selected distribution fits the data, these methods also enable the comparison of different distributions used to fit a sample.

The Pearson's Chi-squared (χ^2) test, when used to test goodness-of-fit, is used to compare counts of observed data with an expected theoretical distribution. The null hypothesis tested is that the frequency distribution of the sample of empirical data is consistent with a particular distribution.

The two-sample Kolmogorov-Smirnov (K-S) test is a non-parametric goodness-of-fit test used to compare two datasets, one of observed data and the other of expected data. The null hypothesis in the two-sample test is that both samples are drawn from populations having identical distributions.

Relevant to this study, the most useful statistic calculated in both tests is the p-value, which indicates whether or not the null hypothesis can be rejected. A suitable significance level, often 0.01 or 0.05, is selected, and if the p-value calculated by the test is less than the significance level, the null hypothesis is rejected in favour of the alternate hypothesis. If the p-value is greater than the significance level, neither the null hypothesis nor the alternate hypothesis can be rejected. However, if the p-value is much larger than the significance level, it can be inferred that the null hypothesis is likely to be true.

7.5 Preliminary Analyses

It is evident that a large number of options are available for the selection of threshold and parametric models for fitting the data in extreme analysis. From these, it is necessary to identify the threshold and distribution that gives the best estimates when applied to data obtained from the numerical model.

Moreover, as the hindcast dataset spans over 140 years, the sample of H_s values can be as large as needed, even extending as much as the return period. While it may be that using 50 years, or 100 years, or even all 140 years of data would yield a good estimate of the return value, a sample of that size is likely to be unwieldy, and the computational requirements for analysing it may be needlessly high. Efficiency would require a balance between the size of the sample and the accuracy of the estimate.

In this section, different combinations of sample size, distribution and threshold are tested, and the best combination is identified for producing a regional map of 100-year return values. These preliminary analyses are carried out at locations where buoy data is available for comparison as necessary.

7.5.1 Selection of an appropriate parametric model

To compare the performance of the Generalised Extreme Value (GEV) distribution and the Generalised Pareto distribution, data from the period 00:00 hrs on January 1, 1981 to 23:00 hrs on December 31, 2000 was chosen for six buoy locations. The 95th percentile value was arbitrarily chosen as the threshold and a separation period of 3 days was used to obtain the declustered dataset of exceedances. On average, the sample contained 340 elements which are adequate for fitting distributions with a high degree of confidence. Maximum Likelihood

Estimation (MLE) (Pfaff and McNeil, 2012) was used to estimate the parameters of the mathematical models (i.e. GEV and GPD) which are used in the analysis. The histogram of peaks with the theoretical distributions superimposed is shown in Figure 7.2

If y_i is the i^{th} element in the ordered dataset of exceedances with k elements, such that $y_1 \leq y_2 \leq ... \leq y_k$, then the empirical probability of non-exceedance of y_i can be found from:

$$P(y_i) = \left(\frac{i}{k+1}\right) \tag{7.6}$$

The corresponding probability of non-exceedance for the GEV distribution can be calculated from equation 7.1, which may be rewritten as

$$G(y_i) = \begin{cases} \exp\left[-\exp\left\{-\left(\frac{y_i-\mu}{\sigma}\right)\right\}\right] & \text{if } \xi = 0 \\ \exp\left[-\left\{1+\xi\left(\frac{y_i-\mu}{\sigma}\right)\right\}^{-\frac{1}{\xi}}\right] & \text{if } \xi \neq 0 \end{cases} \tag{7.7}$$

Similarly, for the GPD, the probability of non-exceedance may be calculated from equation 7.8.

$$G(y_i) = 1 - \left(1 + \frac{\xi(y_i-\mu)}{\sigma}\right)^{-\frac{1}{\xi}} \tag{7.8}$$

If the model fits the sample well then $P(y_i) \approx G(y_i)$. A probability plot consisting of the points $\{P(z_i), G(z_i), \text{ for } i = 1,2,...,k\}$ should be linear, lying along the unit diagonal. Deviation from linearity and the unit diagonal would indicate the inability of the model to describe the data. Probability plots for the GEV distribution and the GPD are presented for three locations in Figure 7.3.

The theoretical quantiles can also be plotted against empirical data to produce the QQ-plot which offers another visual means of evaluating the goodness-of-fit of the parametric model. The theoretical quantiles ($\widehat{y_i}$) corresponding to the empirical probability $P(y_i)$ can be calculated using Equation 7.9 for the GEV distribution

$$\widehat{y_i} = \begin{cases} \mu - \sigma \ln\left[-\ln\left(\frac{i}{k+1}\right)\right] & \text{if } \xi = 0 \\ \mu - \frac{\sigma}{\xi}\left[1 - \left\{-\ln\left(\frac{i}{k+1}\right)\right\}^{-\xi}\right] & \text{if } \xi \neq 0 \end{cases} \tag{7.9}$$

and using Equation 7.10 for the fitted GPD.

$$\widehat{y_i} = \mu + \frac{\sigma}{\xi}\left[\left(\frac{i}{k+1}\right)^{-\xi} - 1\right] \tag{7.10}$$

If the fitted model describes the sample well then $y_i \approx \widehat{y_i}$. As with the probability plots, the QQ plots obtained by plotting $\{y_i, \widehat{y_i}, \text{ for } i = 1,2,...,k\}$ should also be linear lying along the unit diagonal. Significant departure from linearity or from the unit normal would indicate the inability of the fitted model to describe the data. QQ-plots for the GEV distribution and the

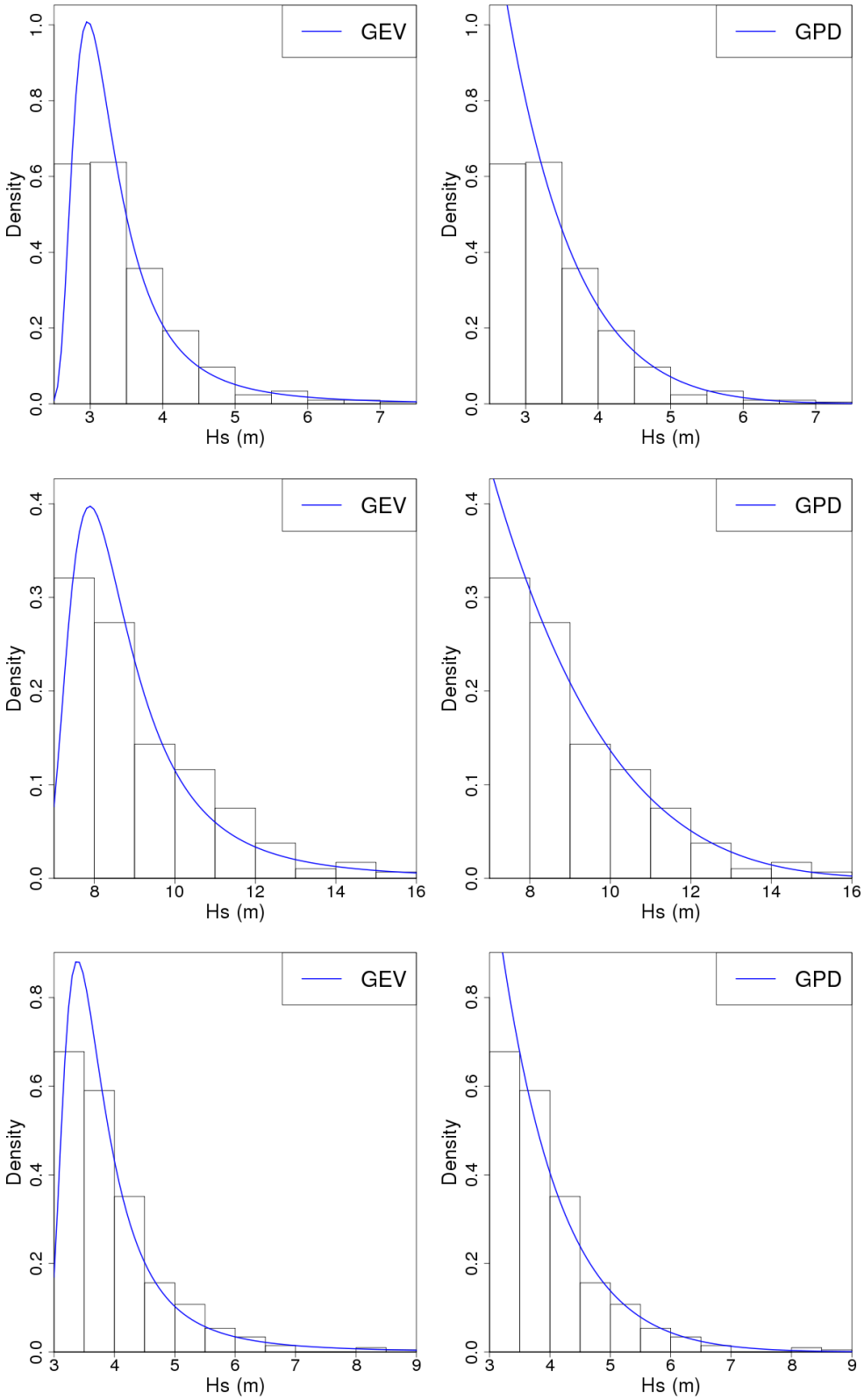

Figure 7.2: Histogram of peaks with GEV distribution (left) and GPD (right) superimposed. Top - Firth of Forth; middle - M3; bottom - Tyne/Tees

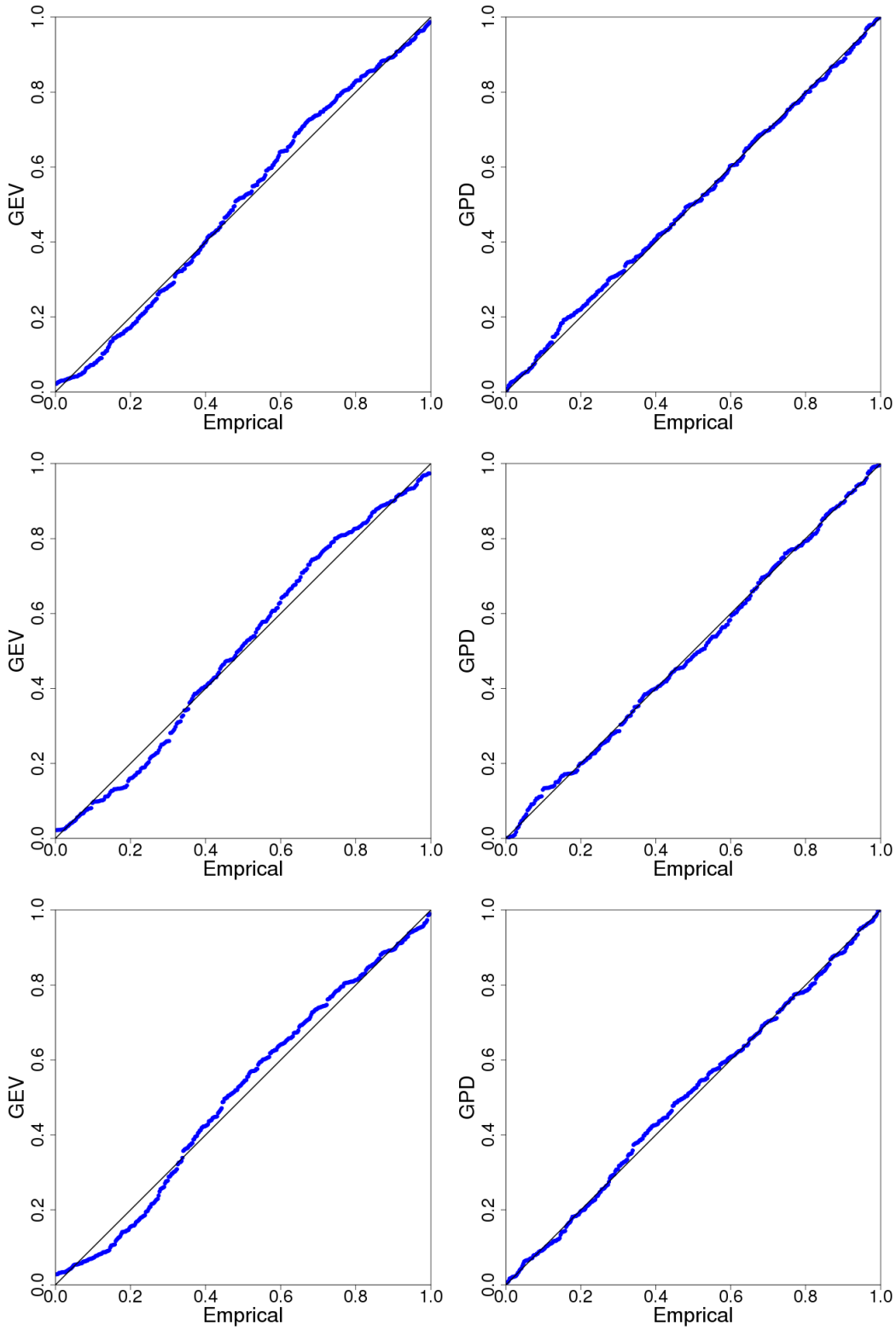

Figure 7.3: Probability plots for GEV distribution (left column) and GPD (right column). Top row - Firth of Forth, middle row - M3, bottom row - Tyne/Tees

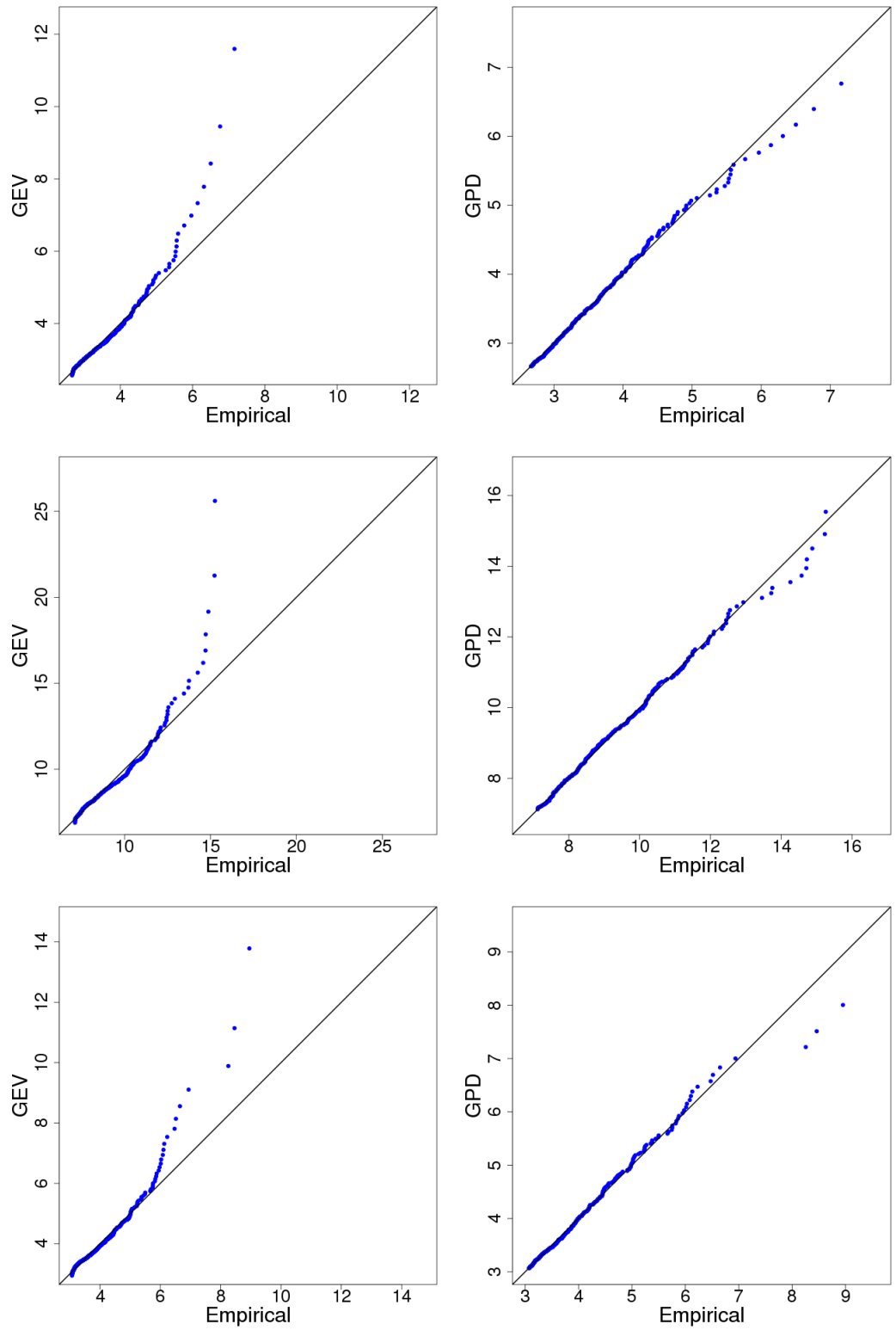

Figure 7.4: Quantile (QQ) plots for GEV distribution (left column) and GPD (right column). Top row - Firth of Forth, middle row - M3, bottom row - Tyne/Tees

GPD are shown in Figure 7.4 for three locations.

A visual comparison of the GEV model and the GPD from Figures 7.3 and 7.4 shows that the GPD is the superior distribution for modelling the sample of exceedances. The probability plots for the GEV distribution shows small deviations from the unit diagonal, however, in comparison, the probability plots for the GPD show greater linearity and smaller departure from the unit diagonal.

The difference in the capability of the two models is seen more clearly in the QQ-plots in Figure 7.4. The QQ-plots for the GEV model show that the model describes low exceedances well, as seen by the linearity and conformity with the unity diagonal. However, the model appears to severely overestimate low-probability events. This has particular significance in extreme-wave estimation as the return values, by definition, have an extremely low probability of occurance associated with them. This implies that the return values estimated using the GEV model are likely to be unrealistically high.

The QQ-plots for the fitted GPD show good agreement between the model estimates and empirical values, with the exception of a few of the highest exceedances where some departure from the unity line is exhibited. However, this deviation is not as severe as for the GEV. Overall, the fitted GPD appears to model exceedances in wave height more accurately than the fitted GEV distribution, especially the low-probability events.

The models were subjected to the Kolmogorov-Smirnov (K-S) and Chi-squared (χ^2) to provide quantitative measures of goodness-of-fit and the obtained p-values are tabulated in Table 7.1.

	Kolmogorov-Smirnov test		χ^2 test	
Buoy site	GEV	GPD	GEV	GPD
M3	0.6379	0.9956	0.0012	0.2957
M4	0.6669	0.8571	0.0169	0.5752
M6	0.5336	0.8943	0.0005	0.3727
F. of Forth	0.6027	0.9166	0.0054	0.6505
Moray Firth	0.1847	0.6268	0.0202	0.1824
Tyne/Tees	0.5965	0.9466	0.0093	0.6642

Table 7.1: p-values from goodness-of-fit-tests for the Generalised Extreme Value distribution and the Generalised Pareto distribution

From the p-values tabulated in Table 7.1, it can be seen that the GPD outperforms the GEV distribution in modelling peaks. Although the p-values from the K-S test for the GEV model are inconclusive in that the null hypothesis cannot be rejected, the p-values for the GPD are comparatively much larger for every location tested. This means that although the K-S test does not reject either distribution, it indicates that the GPD is the superior of the two distributions for modelling exceedances in the POT method.

The p-values calculated from the χ^2 test are particularly significant as they are all smaller than 0.05 for the GEV model. This implies that null hypothesis, i.e. the cdf of the sample of

exceedances is consistent with the GEV distribution, can be rejected at a significance level of 0.05. In 4 out of the 6 locations tested, the null hypothesis can be rejected even at a significance level of 0.01 for the GEV distribution. In comparison, all the p-values from this test for the fitted GPD are much larger than the significance levels, indicating that the null hypothesis is likely to be true.

From the visual evidence presented in Figures 7.3 and 7.4 and the results of goodness-of-fit tests in Table 7.1, the GPD was selected as being suitable for extreme value estimation in the POT approach, and the GEV model was rejected.

7.5.2 Dataset size and threshold selection

One of the objectives of the study is to produce a map of 100-year return values. As this involves calculating the extremes cell-by-cell over the gridded dataset for the region, it is essential that the chosen threshold be simple, easy to automate, and computationally inexpensive.

It was demonstrated by Caires and Sterl (2005) that the 93rd percentile value, when chosen as the threshold, resulted in a good fit in most gridpoints. A higher threshold value can only improve the fit further, provided that the resulting sample is adequately sized (Agarwal *et al.*, 2013). From this evidence, it was decided to test the 95th percentile value and the 98th percentile value as possible thresholds when used with hourly datasets spanning 15 and 20 years.

The reasoning behind testing the threshold and period of data together is that the two are not independent of each other in determining the size of the sample of peaks. A lower threshold, say the 95^{th} percentile value, would yield a larger sample of peaks than the higher threshold, say the 98^{th} percentile value. Similarly, the sample obtained from a longer timeseries would contain more storms, and therefore more peaks, than the sample from a shorter timeseries.

Having the luxury of a hindcast dataset that is longer than the return period, it is possible to estimate the 100-year extreme wave and compare it with simulated annual maxima over the return period. In this way, the effectiveness of the threshold and the length of data can be assessed.

Having identified the GPD as a suitable extreme value distribution for modelling the sample of peaks the POT methodology was applied to timeseries data for 6 locations using this distribtution. Data from 1886 to 1900 (15-year sample) and 1881 to 1900 (20-year sample) were used with thresholds at 95^{th} and 98^{th} percentile levels, yielding 4 unique samples for each location for further analysis. A GPD was fit to each sample of peaks and 100-year return values were estimated for the period 1901-2000 for the selected locations.

The H_{100} values calculated for each sample are tabulated in Table 7.2 along with the maximum H_s value simulated by WWIII between 1901 and 2000. The number of storm peaks that exceed the estimated H_{100} value for each sample are also presented in Table 7.3. From these two tables

it is evident that the estimates obtained by using 20 years of data are superior to those obtained using 15 years. For the 20 year sample, the estimates using the 98^{th} percentile threshold are noticeably better than those obtained using the 95^{th} percentile threshold. However, for the sites to the west of the UK and Ireland, it is difficult to decide between the two.

	15 year sample		20 year sample		
Site	H_{100} Q = 0.95	H_{100} Q = 0.98	H_{100} Q = 0.95	H_{100} Q = 0.98	20^{th} century H_{max}
M3	16.079	15.438	16.895	16.498	16.569
M4	18.067	16.262	19.106	16.774	17.884
M6	16.428	17.413	17.683	18.417	20.308
Firth of Forth	8.680	11.076	9.499	9.658	9.060
Moray Firth	6.941	7.087	7.746	8.461	8.414
Tyne/Tees	9.569	9.237	9.423	9.112	9.779

Table 7.2: 100-year return value estimates (m) from the POT analysis applied to samples with different lengths and threshold values

	15 year sample		20 year sample	
Site	n($H_s > H_{100}$) Q = 0.95	n($H_s > H_{100}$) Q = 0.98	n($H_s > H_{100}$) Q = 0.95	n($H_s > H_{100}$) Q = 0.98
M3	1	2	0	1
M4	0	4	0	2
M6	11	8	6	3
Firth of Forth	1	0	0	0
Moray Firth	7	7	2	0
Tyne/Tees	1	1	1	1

Table 7.3: Number of peaks exceeding the estimated extreme wave height for the 100-year return period

In general, it was found that using the 98^{th} percentile value as threshold with a 20 year time-series, the smallest sample of declustered peaks contained 179 elements, satisfying the requirements outlined in Section 7.3.2. Based on these, it was decided that the 98^{th} percentile value be applied as the threshold to the 20-year timeseries for the estimation of 100-year return values for the 20^{th} century.

7.5.3 Calibration of return value estimates

Numerical models are tuned so that the overall bias, averaged across seasons and regions, is minimised. As a consequence, it may be that the wave model simulates 'average' conditions well, but may be unable to simulate the rarer high and low energy events as pointed out by Cavaleri (2009). In other words, while it may that the body of the distribution of simulated wave heights aggregated over a long period may be similar or identical to the body of the distribution of observations, it is possible that the tails of the two distributions might differ

significantly. The implication of this is that model output, in its raw form, may be unsuitable for the estimation of extreme wave heights.

The simplest way to correct this would be to calibrate the model's output using observations from buoys or satellites. Alternatively, the approach followed by Williams (2008) may be used where H_{100} estimates from model data were corrected using H_{100} estimates from other sources. This approach, however, has numerous arguments against it, most notable among them the fact that any biases in the validation dataset would be artificially introduced into the analysis.

This study adopts the following approach for calibration:

- Assess the need for calibration;
- Calculate a calibration function for the sample of peaks using buoy measurements;
- Calculate the return value for the raw and corrected data for the selected sites;
- Obtain a generalised correction function, if possible, using the raw and corrected H_{100} estimates; and
- Obtain a confidence interval for these estimates using the correction function.

Observations from several buoys were available for study, however, only 10 buoys had overlapping periods of measurements with each other and with the hindcast of two or more years. As a result, these buoys were chosen for calibrating the sample of peaks obtained from the model. The period for this analysis was January 2009 to December 2010. Although fractional parts of 2008 were also available, they were not considered in order to ensure that the seasons were equally represented in the sample. The buoys used in the analsis are:

- CEFAS buoys (UK): Dowsing, Firth of Forth, Hastings, Moray, Scarweather, Tyne/Tees and West Gabbard; and
- Marine Institute buoys (Ireland): M3, M4 and M6.

The sample of peaks was extracted from buoy records and model simulations and compared. A calibration function was found for each location and used to correct the sample of model peaks. The raw and corrected samples of modelled peaks is plotted against corresponding buoy measurements for six buoy locations in Figure 7.5.

The calibration functions are tabulated in Table 7.4 along with the raw and corrected H_{100} estimates. These extreme wave heights were obtained by fitting a GPD to peaks above the 98th percentile from modelled H_s between January 1981 to December 2000 and using Equation 7.5.

Using the data from Table 7.4, a general correction function was found:

$$\widehat{H}_{100} = 0.8466 H_{100} + 1.5174 \tag{7.11}$$

where H_{100} and $\widehat{H}_{100}$ are the raw and corrected 100-year return values.

As this function was obtained by linear regression using only 10 points, it is important to obtain functions for the lower (5%) and upper (95%) confidence bounds to quantify the un-

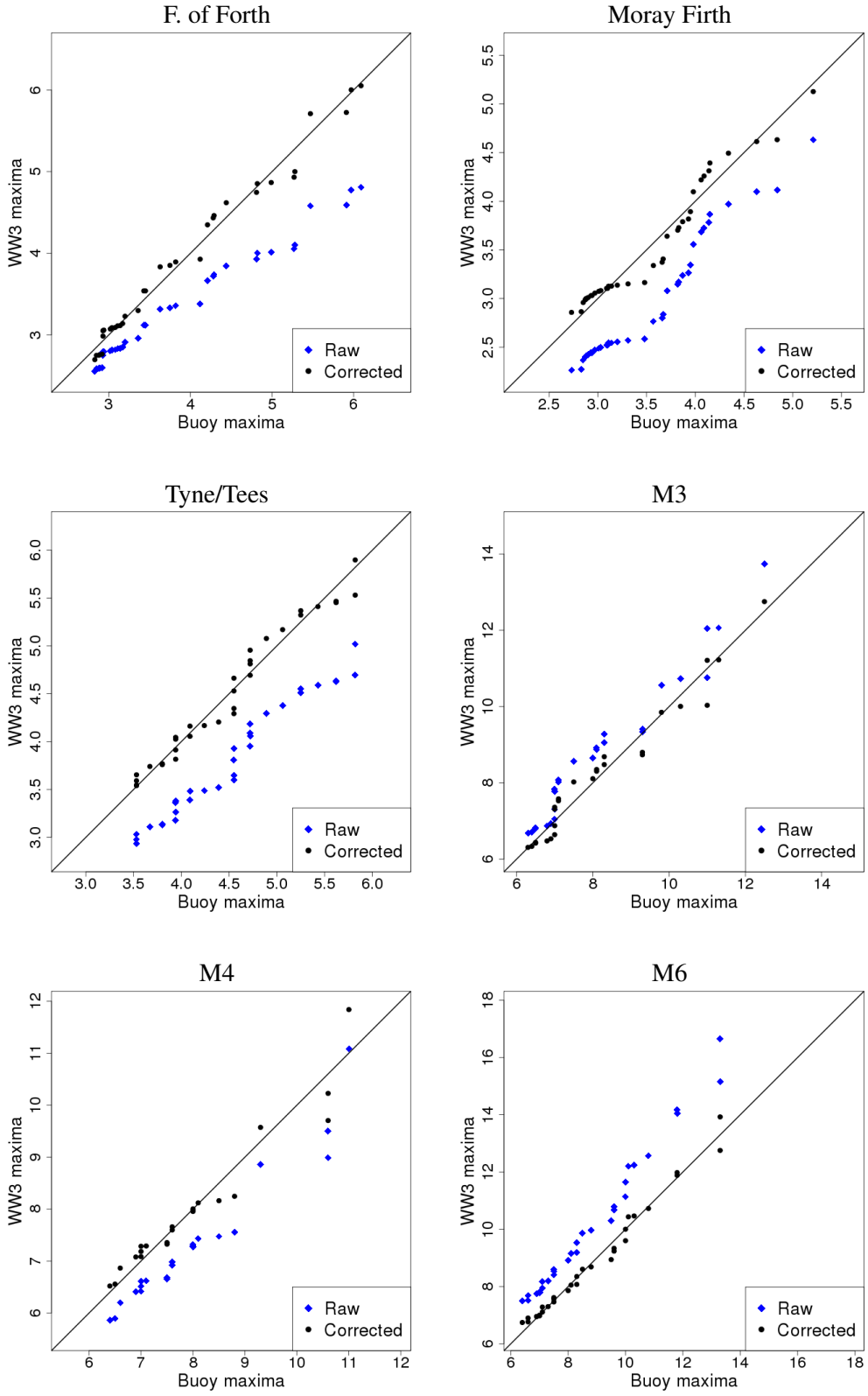

Figure 7.5: Scatter plot of raw and corrected model peaks (m) plotted against corresponding buoy peaks

Buoy	$f_1(H_{raw}):\ H_{corr1} = aH_{raw} + b$	H_{100} (raw)	H_{100} (corr.)
F. of Forth	$H_{corr1} = 1.487H_{raw} - 1.099$	7.9021	10.6507
Moray Firth	$H_{corr1} = 0.958H_{raw} + 0.689$	8.7793	9.1021
Tyne/Tees	$H_{corr1} = 1.130H_{raw} + 0.225$	9.7939	11.2945
Dowsing	$H_{corr1} = 0.964H_{raw} + 0.007$	6.1746	5.9593
Hastings	$H_{corr1} = 0.790H_{raw} + 1.062$	9.2228	8.3479
Scarweather	$H_{corr1} = 0.642H_{raw} + 0.559$	8.5771	6.0653
W. Gabbard	$H_{corr1} = 1.081H_{raw} - 0.129$	7.8881	8.3978
M3	$H_{corr1} = 0.913H_{raw} + 0.204$	16.8162	15.5575
M4	$H_{corr1} = 1.019H_{raw} + 0.547$	19.4605	20.3817
M6	$H_{corr1} = 0.784H_{raw} + 0.873$	21.4933	17.7259

Table 7.4: Calibration functions for model peaks obtained from buoy observations, tabulated along with raw and corrected H_{100} estimates for selected buoy locations

certainty associated with the estimate. Functions for the lower and upper bounds are given in Equations 7.12 and 7.13 respectively.

$$\widehat{H}_{100} = 0.9148H_{100} - 0.8719 \tag{7.12}$$

$$\widehat{H}_{100} = 1.0852H_{100} + 0.8719 \tag{7.13}$$

It must be noted that the calibration data from the buoys only spans 2 years, with samples containing approximately 15 to 25 peaks for each buoy location. Moreover, the calibration was carried out using data from 2009-2010, while the return values were estimated using data from 1981-2000. It is entirely possible that the sample is too small to model the true correction function. It is also a possibility that the implicit assumption that the relationship between model output and buoy observations does not change (a form of the assumption of stationarity of climate) is false.

7.6 Estimation of 100-year extremes

For the estimation of extreme waves at a regional level, significant wave height data from the regional grid with spatial resolution of 0.25° x 0.25° and sampled at hourly intervals would be subjected to the POT analysis. The analysis is carried out cell-by-cell to obtain return values in a gridded form for the region. Based on the findings of the preliminary analyses described in Section 7.5,

- the timeseries data used for the analysis spans over 20 years (1981 to 2000);
- the threshold value is taken as the 98^{th} percentile value of the local time series; and
- the generalised pareto distribution is fitted to the sample of declustered peaks for estimating the return value.

From the parameters estimated, the raw 100-year return value was obtained from Equation 7.5. These values were then corrected using the relation in Equation 7.11. The 5% and 95% confidence bounds were also calculated from Equations 7.12 and 7.13 respectively. The contour maps of the 100-year return values thus calculated are presented in Figures 7.6 to 7.9.

Previous studies (e.g. Williams (2008); Agarwal *et al.* (2013)) have computed 100-year return values for the 21^{st} century for several buoy and offshore platform locations in the North Sea and to the west of the UK. A comparison of these estimates with the return-values computed here are presented in Table 7.5.

The return values estimated by Williams (2008) are based on NEXTRA model data and corrected with established criteria, based on measurements from buoys and offshore platforms. Return values based on 33-years of H_s data from the ERA-I reanalysis dataset were computed by Agarwal *et al.* (2013). Locational estimates for comparison from this study have been obtained from the nearest gridpoint.

Buoy/Platform	H_{100} (raw)	H_{100} (corr.)	H_{100} (lower bound)	H_{100} (upper bound)	H_{100} *	H_{100} #
K2	20.78	19.11	18.14	23.42	18.7	15.21
K4	20.85	19.17	18.20	23.49	19.2	15.11
Thistle	17.81	16.60	15.42	20.20	15.9	12.26
Rhum	14.90	14.13	12.76	17.04	14.2	11.22
Miller	14.25	13.58	12.16	16.34	14.2	11.19
Andrew	14.28	13.61	12.19	16.37	13.5	10.92
Goldeneye	12.52	12.20	10.58	14.46	13.3	9.36
Buchan	13.85	13.25	11.80	15.90	13.2	10.25
Fulmar	13.22	12.71	11.22	15.22	13.3	11.08
Curlew	13.07	12.58	11.08	15.06	13.3	10.55
Auk	13.49	12.94	11.47	15.51	13.3	10.56
Tyne	10.69	10.57	8.91	12.47	9.3	9.16
Cleeton	8.74	8.92	7.12	10.36	10.3	8.19
Carrack	11.23	11.02	9.40	13.06	9.2	7.57
Leman	10.25	10.19	8.50	11.99	7.9	7.27
Clair	16.61	15.57	14.32	18.90	16.6	12.97
Foinaven	17.29	16.15	14.94	19.64	18.0	14.10
Morecambe N.	8.50	8.71	6.90	10.10	8.7	7.66

Table 7.5: A comparison 100-year extreme waves estimated in this study with H_{100} estimates for the 21^{st} century obtained from other studies. All estimates in metres. * Data from Williams (2008), # Data from Agarwal *et al.* (2013).

It is interesting to note from Table 7.5 that the estimates obtained from the ERA-I dataset (Agarwal *et al.*, 2013) are significantly lower than the corrected estimates from this study for all the compared locations, so much so that they are lower than even the estimates computed at the 5% confidence level. One possible reason for this might be the difference in sampling rates

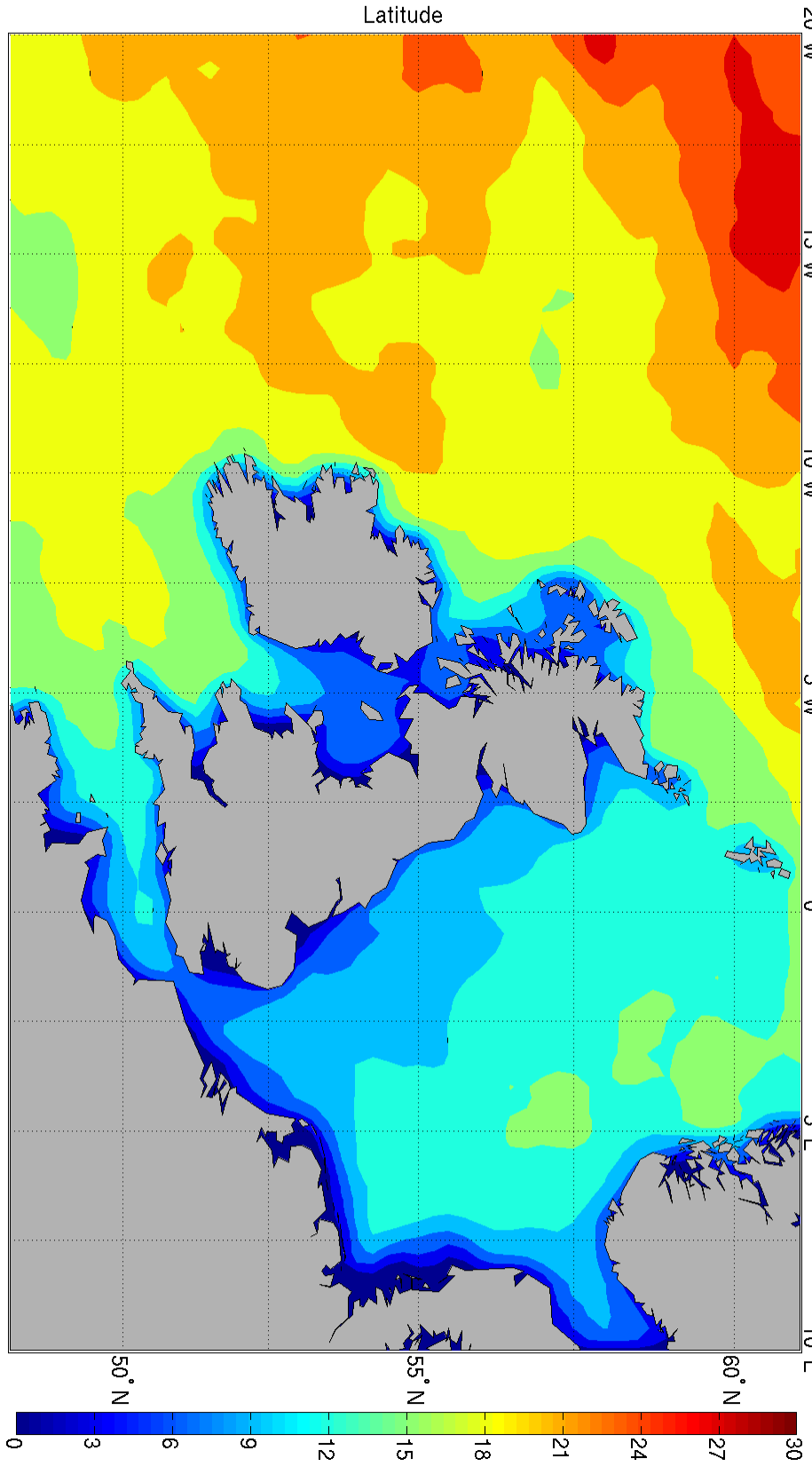

Figure 7.6: Raw 100-year return values for the region (m)

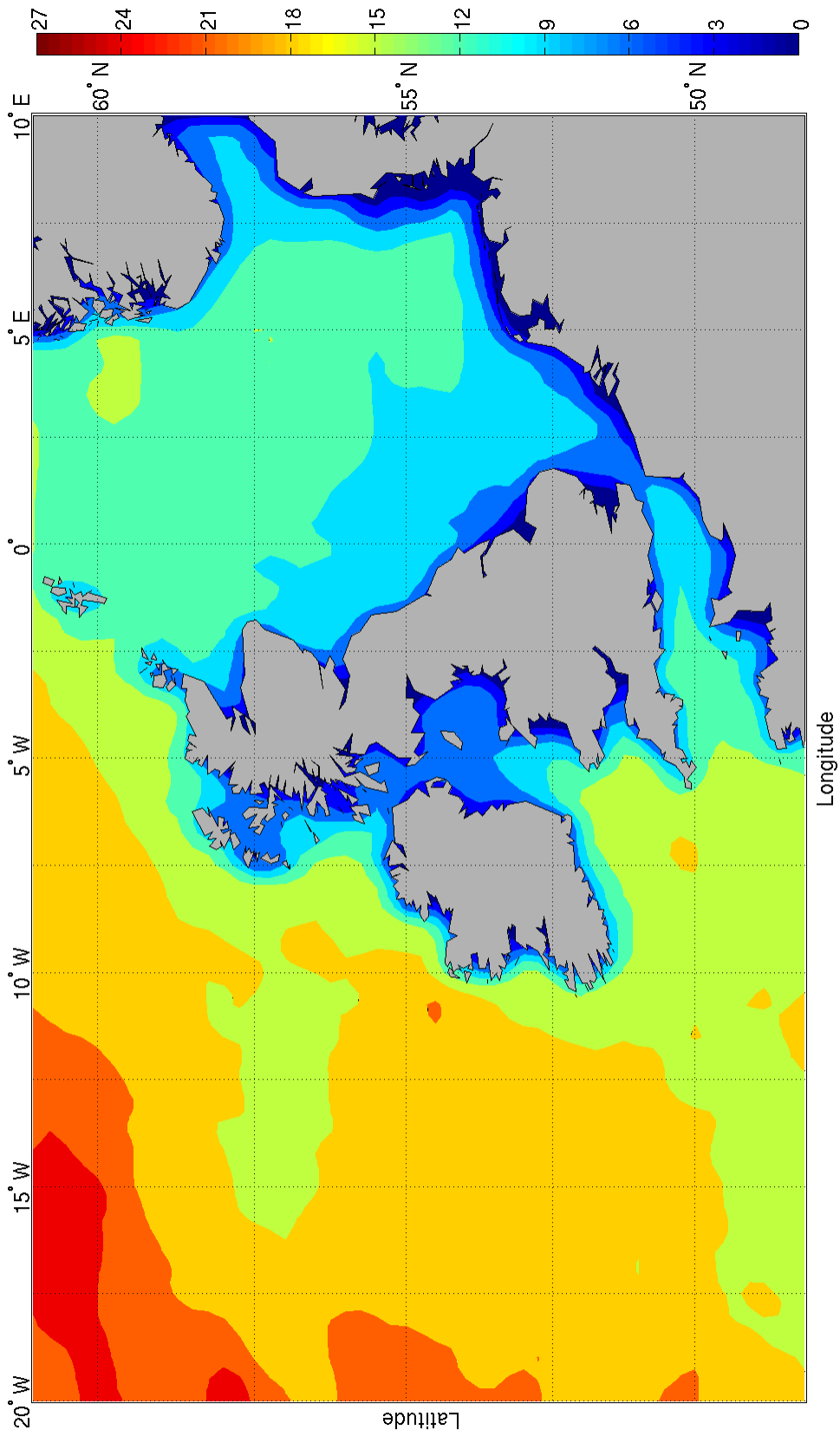

Figure 7.7: Corrected 100-year return values for the region (m)

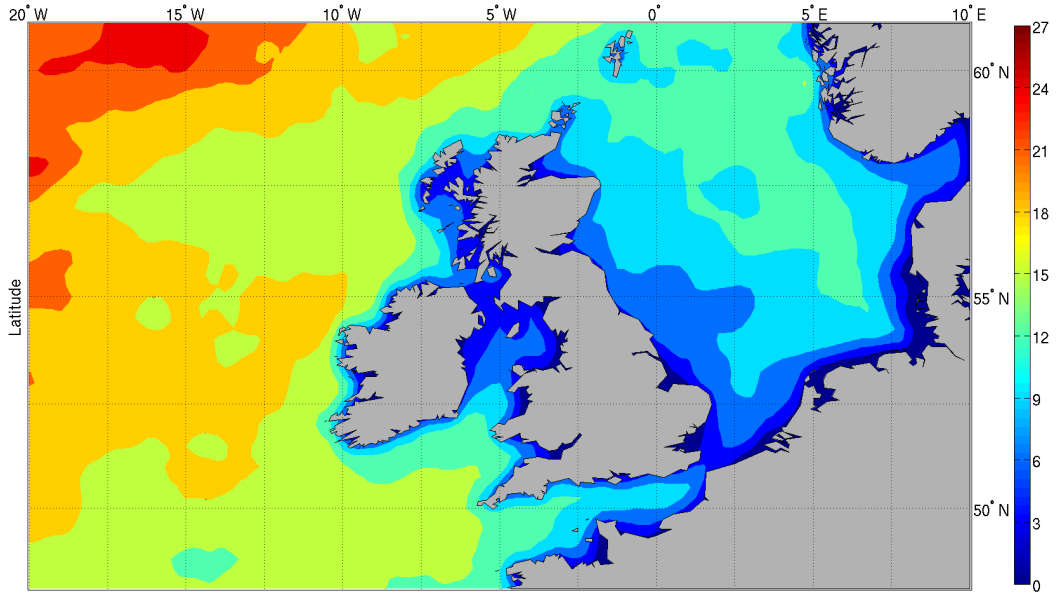

Figure 7.8: Lower (5%) confidence bound for corrected H_{100} estimates (m)

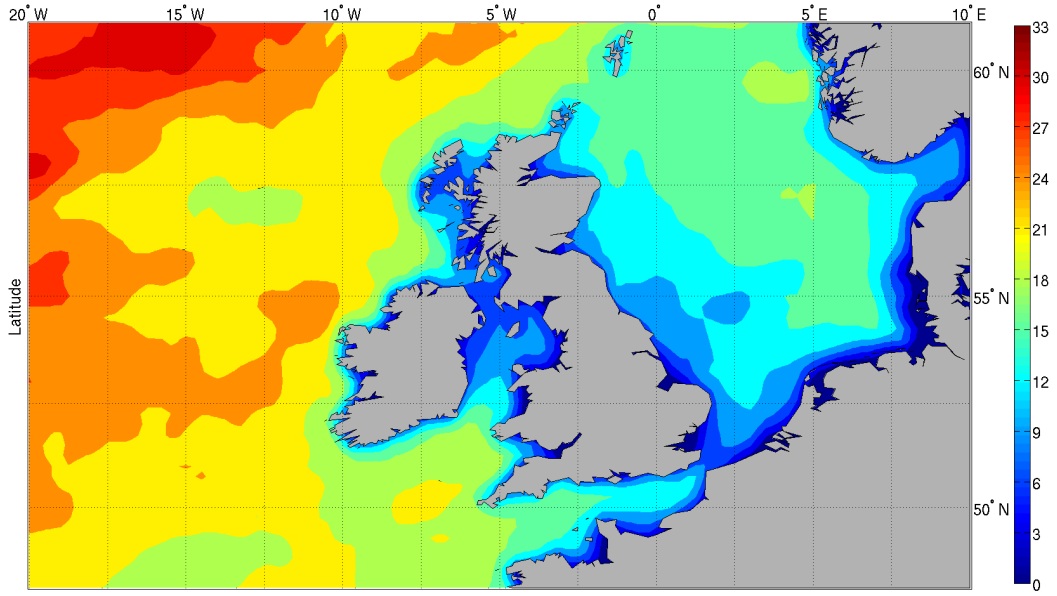

Figure 7.9: Upper (95%) confidence bound for corrected H_{100} estimates (m)

of the underlying datasets. The ERA-I reanalysis dataset, sampled at 6-hourly intervals, may not be able to capture the maximum wave heights that occur during storms, which are likely better captured by the dataset used in this study due to its hourly sampling rate.

On the other hand, the results from this study appear to agree with the estimates obtained from Williams (2008). For most of the locations the corrected H_{100} values are within 10% of the H_{100} values from Williams (2008). Moreover, all the extreme value estimates from Williams (2008) lie within the 90% confidence interval calculated in this study. This is interesting on account of the different datasets, periods and methods used by the two studies, especially for correcting the raw estimates.

7.7 Discussion

During the course of the study, there were several interesting outcomes pertaining to the methodology and approach which are discussed in this section.

A comparsion of the generalised Pareto distribution (GPD) and the generalised extreme value (GEV) distributions showed that the former clearly outperformed the latter, especially for modelling very low probability events. This, however, may be the case only in the region studied, i.e. the Northeast Atlantic Ocean and North Sea, and it may be that the GEV is a superior distribution in other areas.

The selection of threshold based on the 98^{th} percentile, although an extremely simplistic approach, appears to yield samples that are adequately sized for fitting the GPD with confidence. It is easy to apply recursively over the grid and computationally inexpensive. One drawback of this approach, however, is that even when long-period datasets are used, individual years and storms may dominate the sample of exceedances, especially if the sampling rate is high. It may be necessary to dynamically adjust the quantile of the threshold, say between the 93^{rd} and the 98^{th} percentile to ensure that storms over the entire period are represented in the sample of peaks.

It has been pointed out that the use of model data is unsuitable for extreme value estimation (Cavaleri, 2009). However, this can be remedied by a process of calibration against instrument observations. A calibration function for the sample of excesses rather than the entire sample, as done in this study, is likelier to yield better results. The latter calibration approach would shift the mean and variance of the entire sample without necessarily correcting the tail of the sample, which is the important data for extreme value analysis.

In the derivation of a calibration function for the sample of peaks in this study, several assumptions were made which may not hold in reality. The first is that the climate is stationary to the extent that the function derived by comparing data from 2009-2010 would be applicable to the

data from 1981-2010. The other assumption is that the available data, from 10 locations, is adequate to derive a representative function for the entire region.

If satellite data, which spans over more than three decades, is used for calibration, it is possible that a function with a higher degree of confidence may be obtained. Moreoever, it may be possible to obtain calibration functions for each individual datapoint in the regional grid avoiding the need to generalise based on a small number of locations. It should be kept in mind, however, that owing to the gaps in the data due to the track repeat period, some storms may be missed. However, if the dataset extends over a long enough period, there may not be a significant difference between the obtained and true calibration functions.

7.8 Chapter summary

- The different methods of extreme wave analysis are introduced. Of these, the peaks-over-threshold (POT) approach is chosen as being most appropriate;
- The Generalised Extreme Value (GEV) family of distributions is compared with the Generalised Pareto Distribution (GPD) and the latter was found to describe the sample of peaks exceeding the threshold better;
- From different permutations of threshold and sample size, a 20-year sample with the annual 98th percentile significant wave height as the threshold was deemed suitable for the estimation of 100-year return values;
- A method of calibration of the sample of exceedances based on buoy measurements was demonstrated;
- 100-year return values were calculated for the region along with lower (5%) and upper (95%) confidence intervals and presented as contour maps of the region.

Chapter 8

Discussion

In any body of research such as this, there are limitations and difficulties which are encountered and, in many cases, overcome. Some of these may be overcome by making reasonable assumptions, whereas others may require greater advances in knowledge of the climate system. It is important to acknowledge and understand these, not only to define the boundaries of the outcomes of the study, but also to give direction to future work.

In addition, there are certain aspects that come to light when the entire body of work is considered together. For example, the historical and projected wave climate, when treated together may yield interesting insights.

In this chapter the limitations and difficulties during and interesting findings from the study, treated as a single entity, are discussed.

8.1 Computational and storage requirements

The ocean wave model WAVEWATCH III, like other models of complex climate phenomena, is computationally intensive. The ability to run the model in parallel alleviates this to a great extent allowing for simulations of the wave climate to be performed at reasonable time scales. However, without access to large, high-performance computing systems like UK Research Council's HECToR and Archer, it would not be possible to generate the datasets from this study within the time frame of a doctoral research project.

Even with access to these, some compromises had to be made, such as

- Limiting the spatial resolution of the inner grid to 0.25° and abandoning the use of a third, nested grid of resolution 0.1° as originally intended.
- Resolving the spectral grid over 25 frequencies and 24 directions, rather than 36 frequencies and 36 directions as used by operational models. This reduces the computational requirement to approximately half.
- Designing the Monte-Carlo style experiment with only 500 trials.

The Monte-Carlo type experiment (equivalent of approximately 42 years), simulation of hindcast wave conditions for 1871-2010, simulation of contemporary climate using GCM winds

from the 20th Century (20C3M) experiment for 1981-2000, and simulation of projected wave climate with GCM winds from 3 climate change experiments for 2001-2100 required computational resources of approximately 4 million processor hours.

In total, nearly 6TB of storage was required to store the input and output datasets from the study. Major datasets include:

- Hindcast and GCM wind data (netCDF format): ≈120GB;
- Hindcast and GCM wind data (ASCII format for WAVEWATCH III): ≈300GB;
- Raw WAVEWATCH III data (.ww3 format): 1.9TB;
- Processed WAVEWATCH III data for North East Atlantic and North Sea grid only (netCDF format): ≈1.4TB; and
- Derived dataset of wave power density: ≈360GB;
- Wave data (raw and processed) from the Monte-Carlo type experiment: ≈280GB.

Preprocessing and postprocessing of data between formats recognised by WAVEWATCH III and standard formats such as netCDF, as well as analysis of output wave data were performed on a local workstation. Although capable of parallel execution of jobs with up to 16 simultaneous processes, most of these tasks were performed serially. As the data were stored externally, it was found that the speed of the operations was limited by the transfer rate between the storage devices and the processing unit. In such a case no speed benefits can be gained by parallelising the operations.

This was found to be a major constraint on the type and number of analyses that could be performed. For example, calculating the historical trends between 1921 and 2010 for a single wave parameter, say H_s, cell by cell to produce a gridded regional contour map (e.g. Figure 5.8) took approximately 11 hours, with data retrieval accounting for the bulk of total time. Derived parameters such as maximum H_s, or wave power density took significantly longer.

8.2 The Garden Sprinkler Effect (GSE)

In several panels of significant wave height presented in the study (e.g. Figure 5.6), a series of 'hotspots' can be seen off the northern coast of Norway. These hotspots are most likely caused by the Garden Sprinkler Effect (GSE), which occurs when a continuous swell field gets disintegrated into discrete swell fields due to the inadequate resolution of the frequency-direction spectrum. GSE is most noticeable when swell is the major source of local energy (Tolman, 2002a).

The obvious solution to this would be to increase the spectral resolution of the model. However, to eliminate GSE for a model of the North Atlantic, the resolution of the spectral grid would have to be such that $\Delta f = 0.03f$ and $\Delta\theta \approx 1.5^\circ$ (Booij and Holthuijsen, 1987). This is not a practical solution, given the available computational resource. As a more feasible solution for

operational models, Booij and Holthuijsen (1987) proposed the use of a diffusion tensor to explicitly model sub-grid diffusion of swells. This technique has proven to be an effective, albeit computationally intensive way to alleviate the Garden Sprinkler Effect.

WAVEWATCH III, like other spectral wave models, is not immune to GSE. In WW3, while the overall accuracy of simulations increases as higher order propagation schemes are used, the Garden Sprinkler Effect also becomes more apparent (Tolman, 2001). For the alleviation of GSE, options including the diffusion tensor, averaging methods and divergence methods are included in WAVEWATCH III (Tolman, 2002a).

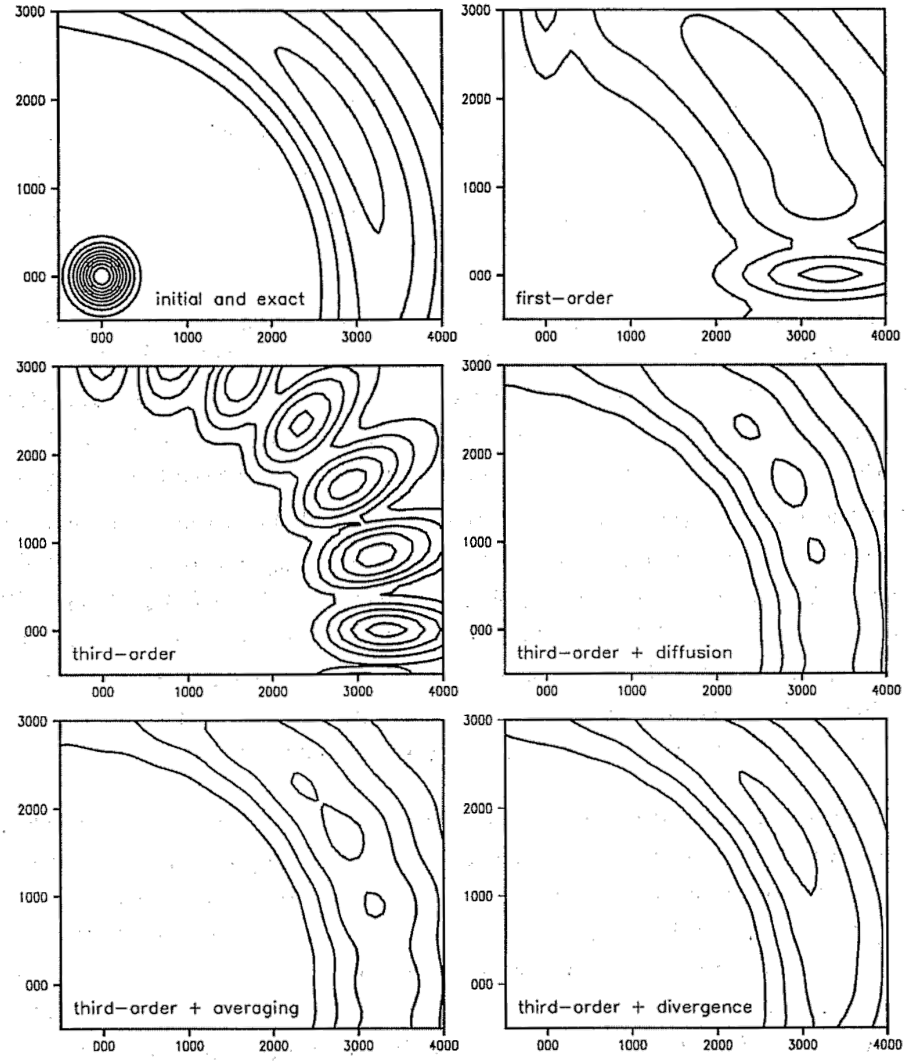

Figure 8.1: Results of swell propagation tests and garden sprinkler effect after 5 days. Topleft: Inital swell field and near exact solution; top-right: first order scheme; middle-left: third order scheme; middle-right: third order scheme with Booij and Holthuijsen (1987) diffusive correction; bottom-left: third order scheme with averaging technique; bottom-right: third order scheme with divergent advection field. Figure sourced from Tolman (2001)

Figure 8.1 shows a results from a test performed by Tolman (2001) to demonstrate the Garden Sprinkler Effect and the different alleviation methods available in WAVEWATCH III. The Garden Sprinkler Effect can be seen from the subplot of the test with no alleviation method applied. The averaging technique yields results comparable with the diffusion tensor technique proposed by Booij and Holthuijsen (1987), but is computationally more economical. The divergence method, while the most effective method for alleviating GSE, is computationally comparable to the diffusion tensor approach. It can be observed from the figure that while the averaging method diffuses the swell field to a great extent, it does not eliminate GSE completely as evidenced by the hotspots (bottom left panel). Tolman (2002a) provide a detailed description of all the available modes of GSE alleviation in WAVEWATCH III along with their comparative efficiencies and effectiveness.

In this study, the third order propagation scheme with the averaging method for GSE alleviation was used for its overall accuracy and computational economy. The appearance of hotspots was not considered to be a problem as it did not affect wave conditions in UK waters, which is the region of interest.

8.3 Uncertainty in historical winds

The accuracy of wave fields simulated by WAVEWATCH III is directly dependent on the accuracy of the forcing dataset, as is the uncertainty associated with it. In the reanalysis of climate data, such as the 20^{th} Century Reanalysis, the accuracy of the product is directly linked to the number and diversity of assimilated climate data. An unavoidable hurdle in the preparation of long-term reanalysis products is that the number and quality of available observations decreases the further back in time one goes.

Specific to the 20CR dataset, which extends back in time to January 1871, this is visualised in Figure 8.2 which shows the number of assimilated records along with the expected (red) and first-guess (blue) root-mean-square (r.m.s.) difference from pressure observations (Compo *et al.*, 2011). It can be seen that the number of assimilated records was of the order of tens in the 19^{th} century and hundreds in the early part of the 20^{th} century. A noticeable decrease in the r.m.s. difference can be observed as the assimilated data increases.

As radiosonde measurements were not assimilated into the CFS, they provide an independent source for validation. An example of such validation performed by Compo *et al.* (2011) is shown in Figure 8.3 where measured 500hPa geopotential height anomalies (m) and 850hPa air temperature anomalies (K) from radiosoundings at the Lindenberg Observatory (Germany) are plotted against reanalysed anomalies. In the figure, blue dots represent data from 1950 to 2011 and red dots represent data spanning 1905 to 1938. Good agreement between reanalysis and observed data can be observed for both periods and parameters, indicated by the clustering

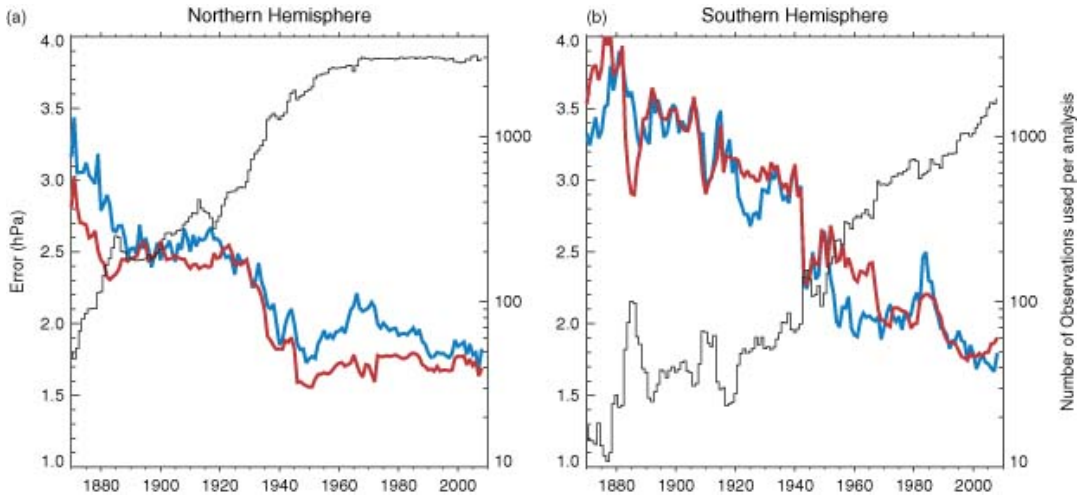

Figure 8.2: Number of assimilated pressure datasets (black) along with expected (red) and first-guess (blue) r.m.s. difference from pressure observations. Sourced from Compo *et al.* (2011)

the accuracy of the reanalysis product in the 20^{th} century. It must be kept in mind that the data presented is for a single location, and it is not necessary that the dataset is equally accurate in other regions.

Other parameters such as mean Sea Level Pressure (SLP) and zonal wind (U10) from the 20CR dataset have been compared with other reanalysis products and with observations from Voluntary Observing Ships (VOS) (Bertin *et al.*, 2013). The results of the intercomparison of the datasets performed by Bertin *et al.* (2013) for the North Atlantic are presented in Table 8.1 and visually in Figure 8.4.

Period	Dataset	SLP		U_{10}		H_s	
		RMSD (hPa)	Bias (hPa)	RMSD (ms^{-1})	Bias (ms^{-1})	RMSD (m)	Bias (m)
1958-2001	NCEP	1.9	-0.06	3.71	-1.29	0.13	0.00
	ERA40	1.82	-0.06	3.42	-2.02	0.51	-0.50
	20CR	2.21	0.22	3.58	-0.94	0.18	0.11
1900-1950	20CR	2.43	0.11	3.91	-0.79	-	-

Table 8.1: Error statistics for different reanalysis datasets calculated from VOS observations. Data sourced from Bertin *et al.* (2013).

From Table 8.1 and Figure 8.4 it can be inferred that the performance of the 20CR dataset is at par with other reanalyis datasets in the latter half of the 20^{th} century. The accuracy of the 20CR data between 1900 and 1950 appears to detiorate slightly, indicated by an increase of approximately 10% in the r.m.s. deviation calculated for the period, when compared to the r.m.s. deviation for 1958-2001. Although Bertin *et al.* (2013) have computed these statistics for the last two decades of the 19^{th} century, they qualify that these are not properly quantified due

to the availability of too few observations.

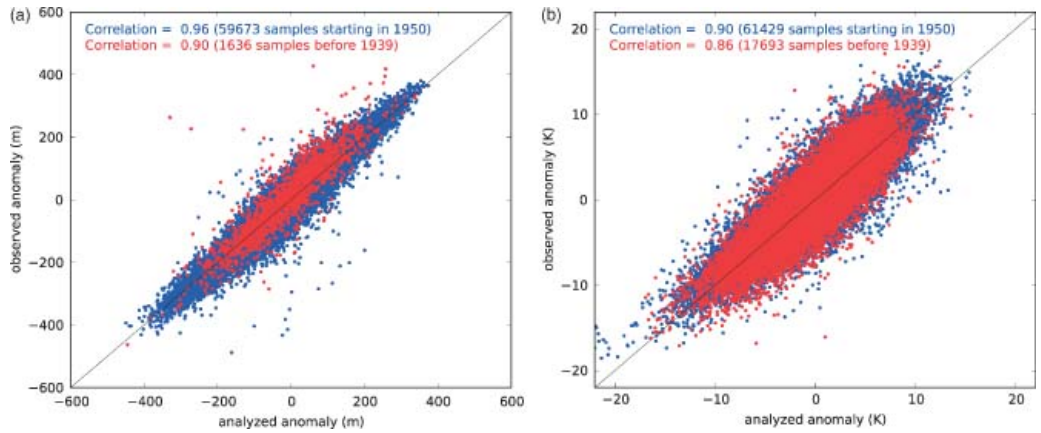

Figure 8.3: Comparison of (a) 500hPa geopotential height anomalies (m) and (b) 850hPa air temperature anomalies (K) at Lindenberg, Germany. Blue dots - anomalies from 1950-2011; red dots = 1905-1938. Anomalies are calculated with respect to the mean annual cycle. Sourced from Compo *et al.* (2011)

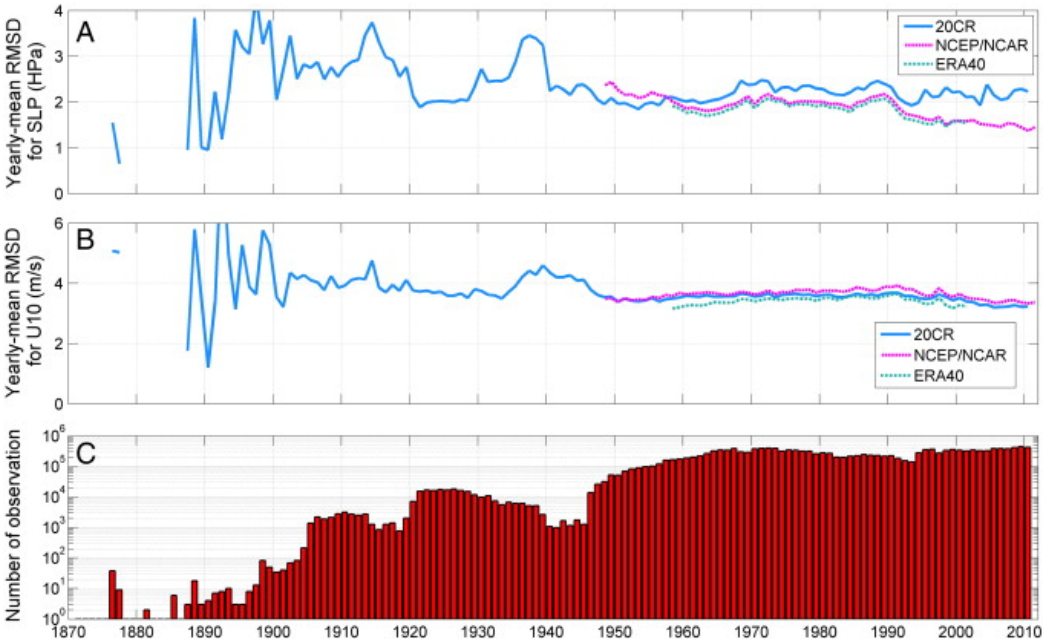

Figure 8.4: Intercomparison of bias statistics for the 20CR dataset with other reanalyses. Root-mean-square difference in A - Sea Level Pressure (SLP) and B - Zonal wind (U10) between reanalysis data and VOS observations. C - Number of VOS observations used in calculations. Sourced from Bertin *et al.* (2013)

It can be concluded that the 20CR dataset is suitable for using with WAVEWATCH III for the 20^{th} century, but care must be exercised using wave fields simulated for the end of the 19^{th} century. It is possible that the accuracy of the dataset between 1871-1900 does not suffer to the extent that it has no value. However, due to the lack of adequate independent data to compare it with, its usefulness is limited.

With more centennial scale reanalysis projects being undertaken, such as the ECMWF's ERA-20C (European Centre for Medium-Range Weather Forecasts, 2014), a greater range of input datasets will be available produced using different climate models and forecast systems. With the paucity of data available for assimilation, it is likely that short-term and small-scale variations in the climate may not be captured by the climate models. However, with an ensemble of reanalysis datasets, uncertainty and error can be quantified for long-term statistics, e.g. 30-year normals of meteorological parameters, in the early periods of the reanalyses.

8.4 Trends in historical and projected data

The comparison of 'observed' and simulated contemporary climate as predicted by WAVEWATCH III when forced with the 20th Century Reanalysis (20CR) dataset and the 20th Century control (20C3M) and medium emissions (A1B) scenario experiments showed good agreement between the spatial patterns in annually averaged significant wave height (H_s) and mean wave direction (D_m), despite some differences in the numerical values of Hs. This was not observed in mean annual energy period (T_e).

If, based on this agreement, it is assumed that the directions of trends computed based on data from the reanalysis and GCM winds are also correct allowing for differences in the magnitudes, a comparison can be made between the historical trends computed in Chapter 5 and the projected trends computed in Chapter 6. This may be done with greater confidence for H_s and D_m than for T_e.

Examining the plots of historical mean annual H_s in Figure 5.14 and projected mean annual H_s in Figure 6.4, we can see a strong increasing trend at the locations West of Hebrides and a weak upwards trend at Burghead between 1921 and 2010. The trend fit to the 20C3M data at both locations appears to be negative, and continues to be thus when fit to data from the climate change scenarios.

At first glance these appear to be contrary to one another, suggesting an inability on the part of the GCM in simulating wind speeds and patterns in the area. However, it must be remembered that the historical trend is fit to all the data between 1921 and 2010, and the location of the hinge was only tested at years that ensured a minimum of 30 years on either side of the hinge. This means that any change in the trend that may have occurred after 1981 is not considered in the search for the hinge position.

Upon re-examining the historical trends in mean annual H_s with this in mind, it is seen that a change in the trend may have occurred circa 1990 (e.g. in Figure 5.14). This is seen more clearly in the plot for the site West of Hebrides, but also observed at other sites such as EMEC, Farr Point and Belmullet. At this stage, however, we do not have enough historical data to confirm this to be a change in trend and not merely a result of interannual variability.

A similar analysis of mean annual D_m in Figures 5.16 and 6.6 reveals more consistency in the directions of historical and projected trends. At the site West of Hebrides, a negative trend is seen between 1921 and 2010. In the GCM data from the 20C3M and A1B experiments too, a negative trend is seen between 1981 and 2010, after which it changes to become weakly positive. This can be seen in the plots of mean annual D_m at EMEC, Farr Point, Belmullet, Galway Bay and Wave Hub as well.

At the Burghead site, the upwards historical trend between 1921 and 2010 appears to exist even in the GCM dataset, continuing upwards up to 2100. This is also observed at the other North Sea location, Beatrice.

From the comparison of the observed historical trends and the projected trends from GCM data, it is highly likely that the mean annual significant wave height will decrease during the 21st century. This is in accordance with the findings of Lowe *et al.* (2009). This is possibly on account of long-term internal variability in the system. Although there is insufficient supporting evidence at this stage, it is possible that Hs peaked around 1990. The frequency of this cycle, if genuine, is yet to be determined.

8.5 Dependence on GCM skill

This research project, like others before it (e.g.WASA) is greatly dependent on the ability of the atmosphere-ocean coupled general circulation models (AOGCMs) in simulating the climate. These models have been shown to be effective in modelling the climate on large temporal and spatial scales espeically when data is aggregated over annual or longer periods. However, some errors continue to exist because small scale processes, which have an important bearing on the climate, are not modelled explicitly and are, at best, incorporated as approximations (Randall *et al.*, 2007). While monthly and annual statistics of wind data from GCMs has been shown to be reasonably accurate Cradden (2009), there is no certainty that the sub-daily patterns, crucial in defining the wave climate, will be of equal accuracy.

Efforts have been made to identify any systematic errors in the GCM that may affect the wave climate in the region. These include site-level and regional comparisons of the 'observed' contemporary wave climate (1981-2010) simulated by forcing WAVEWATCH III with the 20th Century Reanalysis (20CR) dataset and simulated contemporary climate simulated by forcing the wave model with GCM data for the 20th Century (20C3M) control and medium emissions (A1B) experiment. It is observed that, as with other climatological parameters, reasonable skill is demonstrated by the ECHAM5-WWIII coupling in simulating wave conditions, especially significant wave height and wave direction over large temporal and spatial scales.

Although the projected wave climate scenarios developed from the ECHAM5-WWIII coupling seem plausible and consistent with previously reported results (e.g. Lowe *et al.* (2009)), they

are heavily reliant on the GCM's skill and are prone to errors propagating through to the wave model. Moreover, as the errors are a complex function of various factors, it is not feasible to quantify the range of possible errors in the results.

Given the apparent insensitivity to atmospheric greenhouse gas levels observed in the wave climate from the three climate change experiments, it appears that treating the GCM experiments as members of an ensemble might be the way to obtain probabilistic projections of the future wave climate. Such an ensemble will require more than merely the three members used in this research project and must include data from other GCMs with different parameterisations and climate change experiments. This, however, requires significantly more time and computational resources than were available for the doctoral research project.

8.6 Advances since the IPCC Fourth Assessment Report

This research project used 6-hourly surface wind data from the GCM experiments conducted for the preparation of the IPCC Fourth Assessment Report (AR4) (IPCC, 2007). While the report was released in 2007, the development of the general circulation models and simulation of the climate for different experiments occurred prior to that.

More recently, the IPCC Fifth Assessment Report (AR5) IPCC (2013) was released. The data behind this report was generated with more recent climate models with at higher resolutions and parameterisations of climatological processes enabled by advances in knowledge of the climate system as well as available computing infrastructure.

The atmospheric component of the Max Planck Institute's Earth System Model (MPI-ESM), ECHAM6 too, has several improvements over ECHAM5. These include a better representation of transfer of shortwave radiations, new descriptions of aerosols and surface albedo and a significantly improved representation of the middle atmosphere. Minor changes in the representation of convective processes have been affected through changes in model parameters and a different vertical descretisation scheme within the troposphere Moss *et al.* (2010). It is possible that these changes might result in the simulation of more accurate wind fields in the lower atmosphere.

In addition, the definitions of the different scenarios serving as markers of climate change have also change significantly since the last report. The socio-economic scenarios linking potential demographic, political, societal, economical and technological changes with atmospheric greenhouse gas and aerosol content (Nakicenovic *et al.*, 2000) have been replaced with Representative Concentration Pathways (RCPs).

Under the new method of scenario definition, four radiative forcing pathways (RCP8.5, RCP6.0, RCP4.5 and RCP2.6) are defined which can be achieved by multiple socio-economic and technological scenarios. Anticipated benefits of this approach include reductions in the timelag

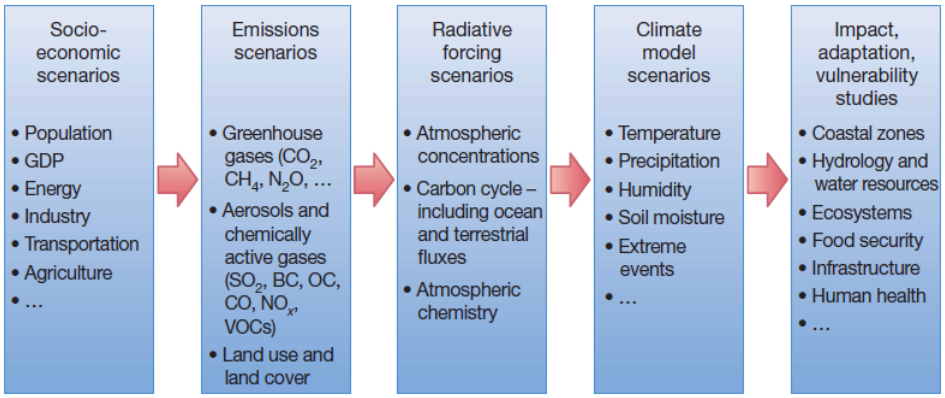

Figure 8.5: The sequential process of scenario development used before AR5. Sourced from Moss *et al.* (2010)

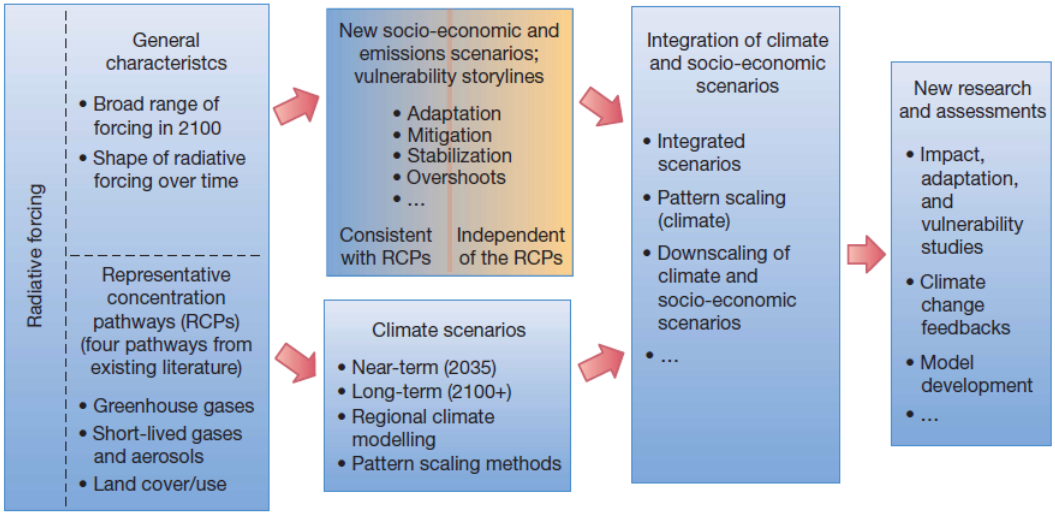

Figure 8.6: Parallel process of scenario development used in AR5. Sourced from Moss *et al.* (2010)

between scenario creation and climate modelling, and increased collaboration in research on vulnerability, impacts and adaptation (Moss *et al.*, 2010). The RCPs are defined so that plausible extreme radiative forcing pathways are included (i.e. RCP8.5 and RCP2.6) and the resulting GCM output is likely to provide a more comprehensive envelope of potential future climatic conditions.

At the time of this research, sub-daily GCM data from the more recent models (e.g. ECHAM6) and RCPs were not available for forcing the wave model. It is possible that the results obtained might be different if these data were used on account of advances in the GCMs. However, from the results of the ANOVA method, it is unlikely that there would be significant differences purely because of the new definitions of emissions scenarios.

8.7 Chapter summary

- The computational and data storage requirements of the research project are briefly discussed along with the consequent constraints and bottlenecks in the analysis of wave climate data.
- The occurrence of 'hotspots' and the Garden Sprinkler Effect as a possible cause are discussed. As these hotspots do not occur within UK territorial waters, they are not considered to be critical. Future work on the prepared dataset can focus on eliminating these by dynamical or statistical filtering approaches.
- The uncertainty in the 20th Century Reanalysis (20CR) winds is discussed, especially in the latter part of the 19th century and early 20th century.
- Historical and projected trends observed at individual locations are discussed. Although some difference exists in the magnitudes, good agreement is seen in the directions. A possible change in trend observed circa 1990 is discussed.
- Dependence of wave climate projections on the GCM skill is discussed.
- Advances since IPCC AR4 are discussed along with their potential impacts on the results obtained from the study. These advances include changes in the climate models as well as a vastly different approach in defining future climate scenarios.

Chapter 9

Conclusions

The research project described in this thesis investigates the existence of long-term trends and changes in the historical wave climate in the North East Atlantic and North Sea region by applying statistical methods to wave data from a spectral wave model. Further, the link between atmospheric greenhouse gas levels and wave conditions in the region is investigated and projections of the wave climate for different climate change scenarios is prepared. Finally, statistical methods for extreme value analysis are used to prepare estimates of 100-year return values of significant wave height in the region. The conclusions from the study are presented below.

9.1 Numerical Modelling

- The 3rd generation phase-averaging spectral wave model WAVEWATCH III version 4.18 was used to simulate hindcast and projected future wave climate for the region. A Monte-Carlo type experiment with 500 trials was designed for testing the uncertainty and sensitivity of the model to wind-growth, swell feedback and bottom friction. For the resolution at which the region was modelled ($0.25^\circ \times 0.25^\circ$), the output of the wave model was found to be most sensitive to wind-growth, followed by swell feedback. Bottom friction appears to have little effect on the uncertainty in the model's output. However, it appears to be a critical parameter in determining the stability of the computed spectrum.
- An application of the Response Surface Modelling approach was demonstrated for calibrating the wave model. Data from the Monte-Carlo type experiment was used to model individual response surfaces for significant wave height and peak wave period, and a joint optimisation procedure was used to simultaneously minimise the bias in both parameters. The wave model was calibrated using the 'optimum' parameter set obtained from this process and validated against buoy data.

9.2 Analysis of the wave climate - I

- Three methods of trend analysis (30-year climate normals, least-squares linear fit and hinge-fit) were compared. Of these, the hinge-fit method with the hinge located at 1920 was found to represent the base state of the wave climate best.
- The hinge-fit type function was fit to point data at 8 locations as well as to the gridded regional dataset. An upwards trend was observed in mean annual significant wave height and mean annual energy period. The strength of the trend varied across the region, being greater in oceanic waters and smaller in the North Sea. Statistical significance of these trends was established using Student's t-test.
- Increasing trends were observed in oceanic waters in annual 95th percentile H_s and annual maximum H_s. However, no link could be established between the two parameters indicating that the changes in the average intensity of storms are not accompanied by equivalent changes in the peak intensity of storms. The trends were found to be statistically significant in large areas of the North East Atlantic and parts of North sea.
- The relative and absolute year-on-year trends calculated for mean annual H_s, T_e and D_m, along with mean wave power density and annual 95th percentile and annual maximum $H_s$were presented as contour maps of the region. Regional plots indicating areas of statistical significance of the trends were also prepared.

9.3 Analysis of the wave climate - II

- From the application of the ANOVA method, it appears unlikely that a causal link exists between atmospheric greenhouse gas levels and the wave climate described by significant wave height and wave period.
- The wave model WAVEWATCH III, when coupled with the AOGCM ECHAM5, shows good ability in modelling significant wave height and wave direction as evaluated based on the comparison of hindcast and simulated current climate. However, it is unable to simulate wave period to an adequate level of accuracy.
- A declining trend is observed in oceanic waters as well as along the UK coast in the North Sea for significant wave height across all projection experiments. Regional contour maps of the absolute and relative trends were plotted in addition to plots for individual locations.
- Other parameters were also analysed and results presented as contour maps and plots. Trends in wave period were not found to be weak and not statistically significant. Caution is advised in the treatment of these and derived quantities such as wave power density.
- Based on the insensitivity of the wave climate to GHG levels, is recommended that data from the different climate change scenarios be treated as members of an ensemble to obtain probabilistic future estimates of wave conditions.

9.4 Extreme waves in the 21st century

- Protocols for estimating extreme wave heights using numerical model wave data were developed in Chapter 7. This work develops on the work presented in Agarwal *et al.* (2013).
- The Peaks-over-threshold (POT) method using the Generalised Pareto Distribution was found to be effective for extreme wave estimation with numerical model data. A suitably high threshold value, such as the 98th percentile value applied to the hourly timeseries data extending over 20 years yielded plausible estimates of 100-year return wave heights.
- A method for calibrating the dataset before applying the extreme value methods, rather than after, was proposed and demonstrated. Regional maps of raw and corrected 100-year wave significant wave height estimates were prepared along with upper (95%) and lower (5%) confidence bounds.

9.5 Suggestions for further work

- The large number of non-behavioural outcomes from the Latin Hypercube experiment in Section 4.4 and the thresholds observed in Figure 4.11 indicate that bottom friction may significantly influence the stability of the spectrum. Further investigation into this is warranted.
- The response surface method for calibration estimates a single value for the tuning parameters which is applied across the region to yield the best performance of the model, averaged across the region. However, some of these parameters, e.g. JONSWAP bottom friction, will vary from one location to another depending upon the type and nature of sea floor. A means of estimating the values of these parameters for regions in a gridded form and incorporating them into the wave model as a gridded data may greatly increase the accuracy and precision of wave predictions.
- It would be interesting to see how the findings of this research would compare with similar datasets of wave conditions simulated by using different numerical models (e.g. WAM coupled with SWAN) and when forced with different datasets (e.g. the 100-year ERA reanalysis). This would, to some extent, reveal the sensitivity of the simulated wave climate to the choice of models as well as to the forcing datasets.
- Running the model at a higher spatial resolution, say $0.1^\circ \times 0.1^\circ$ for shelf and coastal waters 36 frequencies, and with a finer spectral grid, say 36 frequencies and 36 directions is likely to improve the accuracy of the wave output. This may be necessary to improve predictions of wave period, observed in Figure 4.22 especially for the North Sea.

This thesis tested the hypothesis that long-term climatic trends exist in the wave climate of the North East Atlantic and North Sea as a consequence of greenhouse gas induced climate change. It was found that the wave climate does exhibit long-period variability, however, it was found that it is unlikely to be linked with changing levels of atmospheric greenhouse gases.

The research further used projections of wave conditions for different scenarios to estimate the state of the future wave climate and found an overall decrease in the significant wave height and energy period from the present state. A change in the mean wave direction was also observed. However, the statistical significance of these trends could not be established.

Finally, statistical methods were applied to historical data to obtain extreme wave height estimates for the 21th Century. In addition to demonstrating an application of the data resource produced by the research project, this also provides insight into the state of the extreme climate based on the near-historical climate to aid in the design and deployment of marine structures and offshore energy installations.

Bibliography

A. Agarwal, V. Venugopal, and G. P. Harrison. The assessment of extreme wave analysis methods applied to potential marine energy sites using numerical model data. *Renewable and Sustainable Energy Reviews*, 27:244–257, 2013.

C. Amante and B. W. Eakins. ETOPO 1 1 arc-minute Global Relief Model: Procedures, Data Sources and Analysis. Technical Report March, National Geophysicacl Data Center, NOAA, Boulder, Colorado, 2009.

C. W. Anderson, D. Carter, and P. Cotton. Wave Climate Variability and Impact on Offshore Design Extremes. Technical report, For Shell International, 2001.

A. H.-S. Ang and W. H. Tang. *Probability concepts in Engineering*. Wiley, 2nd editio edition, 2007. ISBN 978-0471720645.

J. R. Angel, W. E. Easterling, and S. W. Kirsch. Towards Defining Appropriate Averaging Periods for Climate Normals. *Climatological Bulletin*, 27(2):29–44, 1993.

F. Ardhuin and A. D. Jenkins. On the Interaction of Surface Waves and Upper Ocean Turbulence. *Journal of Physical Oceanography*, 36(3):551–557, March 2006. ISSN 0022-3670. doi: 10.1175/JPO2862.1. URL `http://journals.ametsoc.org/doi/abs/10.1175/JPO2862.1`.

F. Ardhuin, W. O'Reilly, T. Herbers, and P. Jensen. Swell Transformation across the Continental Shelf. Part I : Attenuation and Directional Broadening. *Journal of Physical Oceanography*, 33:1921–1939, 2003.

F. Ardhuin, L. Bertotti, J.-R. Bidlot, L. Cavaleri, V. Filipetto, J.-M. Lefevre, and P. Wittmann. Comparison of wind and wave measurements and models in the Western Mediterranean Sea. *Ocean Engineering*, 34(3-4):526–541, March 2007. ISSN 00298018. doi: 10.1016/j.oceaneng.2006.02.008. URL `http://linkinghub.elsevier.com/retrieve/pii/S0029801806001193`.

F. Ardhuin, E. Rogers, A. V. Babanin, J.-F. Filipot, R. Magne, A. Roland, A. van der Westhuysen, P. Queffeulou, J.-M. Lefevre, L. Aouf, and F. Collard. Semiempirical Dissipation Source Functions for Ocean Waves. Part I: Definition, Calibration, and Validation. *Journal of Physical Oceanography*, 40(9):1917–1941, September 2010. ISSN 0022-3670. doi: 10.1175/2010JPO4324.1. URL `http://journals.ametsoc.org/doi/abs/10.1175/2010JPO4324.1`.

A. Arguez and R. S. Vose. The Definition of the Standard WMO Climate Normal: The Key to Deriving Alternative Climate Normals. *Bulletin of the American Meteorological Society*, 92(6):699–704, June 2011. ISSN 0003-0007. doi: 10.1175/2010BAMS2955.1. URL `http://journals.ametsoc.org/doi/abs/10.1175/2010BAMS2955.1`.

S. Bacon and D. J. T. Carter. A Connection between Mean Wave Height and Atmospheric Pressure Gradient in the North Atlantic. *International Journal of Climatology*, 13:423–436, 1993.

S. Bacon and D. Carter. Wave climate changes in the North Atlantic and North Sea. *International Journal of Climatology*, 11(5):545–558, 1991.

J. Battjes and J. Janssen. Energy Loss and Set-up due to Breaking of Random Waves. *Coastal Engineering*, (1):569–587, 1978.

J. V. Beck and K. J. Arnold. *Parameter Estimation in Engineering and Science*. John Wiley & Sons, Inc., New York, USA, 1977. ISBN 0-471-06118-2.

P. Berrisford, D. Dee, K. Fielding, M. Fuentes, and P. Kå llberg. The ERA-Interim archive. Technical report, European Centre for Medium-Range Weather Forecasts, 2009.

X. Bertin, E. Prouteau, and C. Letetrel. A significant increase in wave height in the North Atlantic Ocean over the 20th century. *Global and Planetary Change*, 106:77–83, July 2013. ISSN 09218181. doi: 10.1016/j.gloplacha.2013.03.009. URL `http://linkinghub.elsevier.com/retrieve/pii/S092181811300088X`.

K. Beven and A. Binley. The Future of Distributed Models: Model Calibration and Uncertainty Prediction. *Hydrological Processes*, 6:279–298, 1992.

J.-R. Bidlot, P. Wittmann, M. Fauchon, H. Chen, J. M. Lefèvre, T. Bruns, D. Wetterdienst, D. Greenslade, and S. Park. Inter-comparison of operational wave forecasting systems. In *10th International workshop on Wave Hindcasting and Forecasting & Coastal Hazard Assessment*, pages 1–22, Oahu, Hawaii, 2002.

J.-R. Bidlot, P. Janssen, and S. Abdalla. A revised formulation of ocean wave dissipation and its impact. Technical Report January, European Centre for Medium-Range Weather Forecasts, 2007. URL `http://old.ecmwf.int/publications/library/ecpublications/_pdf/tm/501-600/tm509.pdf`.

J. Blunden and D. S. Arndt. State of the climate in 2012. *Bulletin of the American Meteorological Society*, 94(8):S1–S258, 2013. doi: http://dx.doi.org/10.1175/2013BAMSStateoftheClimate.1.

N. Booij and L. H. Holthuijsen. Propagation of ocean waves in discrete spectral wave models. *Journal of Computational Physics*, 68(2):307–326, February 1987. ISSN 00219991. doi: 10.1016/0021-9991(87)90060-X. URL http://linkinghub.elsevier.com/retrieve/pii/002199918790060X.

F. P. Bretherton and C. J. R. Garrett. Wavetrains in Inhomogeneous Moving Media. *Proceedings of the Royal Society of London*, 302(1471):529–554, 1969.

K. Briffa, P. D. Jones, and P. Kelly. Principal Component Analysis of the Lamb Catalogue of daily weather types: Part 2, Seasonal frequencies and update to 1987. *International Journal of Climatology*, 10:549–563, 1990.

B. Cahill and T. Lewis. Wave energy resource characterisation of the Atlantic Marine Energy Test Site. *International Journal of Marine Energy*, 1:3–15, April 2013. ISSN 22141669. doi: 10.1016/j.ijome.2013.05.001. URL http://linkinghub.elsevier.com/retrieve/pii/S2214166913000027.

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

R. Carnell. lhs: Latin Hypercube Samples, 2012. URL http://cran.r-project.org/package=lhs.

D. Carter. Estimating Extreme Wave Heights in the NE Atlantic from Geosat Data. Technical report, Health and Safety Executive, London, 1993.

D. Carter and P. G. Challenor. Return Wave Heights at Seven Stones and Famita Estimated from Monthly Maxima. Technical report, Institute of Oceanographic Sciences, 1978.

D. Carter and L. Draper. Has the north-east Atlantic become rougher. *Nature*, 1988.

T. R. Carter and Task Group on Data and Scenario Support for Impact and Climate Assessment. IPCC-TGICA 2007: General Guidelines on the Use of Scenario Data for Climate Impact and Adaptation Assessment. Version 2. Technical report, Intergovernmental Panel on Climate Change (IPCC), 2007.

E. Castillo and J. M. Sarabia. Engineering Analysis of Extreme Value Data: Selection of Models. *Journal of Waterway, Port, Coastal Ocean Engineering*, 118(2):129–146, 1992.

L. Cavaleri. Wave Modeling - Missing the Peaks. *Journal of Physical Oceanography*, 39 (11):2757–2778, November 2009. ISSN 0022-3670. doi: 10.1175/2009JPO4067.1. URL http://journals.ametsoc.org/doi/abs/10.1175/2009JPO4067.1.

L. Cavaleri and L. Bertotti. The improvement of modelled wind and wave fields with increasing resolution. *Ocean Engineering*, 33(5-6):553–565, April 2006. ISSN 00298018. doi: 10.1016/j.oceaneng.2005.07.004. URL http://linkinghub.elsevier.com/retrieve/pii/S0029801805001782.

M. A. Cialone and J. M. Smith. Wave Transformation Modeling with Bottom Friction applied to South East Oahu reefs. In *10th International workshop on Wave Hindcasting and Forecasting & Coastal Hazard Assessment*, number 2, pages 1–12, 2007.

T. J. Cleophas, A. H. Zwinderman, and T. F. Cleophas. *Statistics Applied to Clinical Trials*. Springer, 2006. ISBN 9781402042294.

S. Coles. *An Introduction to Statistical Modeling of Extreme Values*. Springer Series in Statistics, 2001.

G. P. Compo, J. S. Whitaker, P. D. Sardeshmukh, N. Matsui, R. J. Allan, X. Yin, B. E. Gleason, R. S. Vose, G. Rutledge, P. Bessemoulin, S. Brönnimann, M. Brunet, R. I. Crouthamel, a. N. Grant, P. Y. Groisman, P. D. Jones, M. C. Kruk, a. C. Kruger, G. J. Marshall, M. Maugeri, H. Y. Mok, O. Nordli, T. F. Ross, R. M. Trigo, X. L. Wang, S. D. Woodruff, and S. J. Worley. The Twentieth Century Reanalysis Project. *Quarterly Journal of the Royal Meteorological Society*, 137(654):1–28, January 2011. ISSN 00359009. doi: 10.1002/qj.776. URL `http://doi.wiley.com/10.1002/qj.776`.

G. P. Compo, J. S. Whitaker, and P. D. Sardeshmukh. Feasibility of a 100-Year Reanalysis Using Only Surface Pressure Data. *Bulletin of the American Meteorological Society*, 87(2): 175–190, February 2006. ISSN 0003-0007. doi: 10.1175/BAMS-87-2-175. URL `http://journals.ametsoc.org/doi/abs/10.1175/BAMS-87-2-175`.

Council of the European Union. Energy and Climate Change - Elements of the final compromise, 2008.

L. C. Cradden. *The Impact of Climate Change on Wind Energy Generation in the UK*. PhD thesis, University of Edinburgh, 2009. URL `https://www.era.lib.ed.ac.uk/handle/1842/4034`.

H. Cramér. *Mathematical methods of statistics*. Princeton University Press, 1946.

R. C. de Winter, A. Sterl, J. W. de Vries, S. L. Weber, and G. Ruessink. The effect of climate change on extreme waves in front of the Dutch coast. *Ocean Dynamics*, 62 (8):1139–1152, June 2012. ISSN 1616-7341. doi: 10.1007/s10236-012-0551-7. URL `http://link.springer.com/10.1007/s10236-012-0551-7`.

D. P. Dee, S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, a. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kå llberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut, and F. Vitart. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656):

553–597, April 2011. ISSN 00359009. doi: 10.1002/qj.828. URL http://doi.wiley.com/10.1002/qj.828.

Department for Environment Food and Rural Affairs. Observing and Modelling - WaveNet, 2013. URL http://www.cefas.defra.gov.uk/our-science/observing-and-modelling/monitoring-programmes/wavenet.aspx.

Department of Energy and Climate Change. National Renewable Energy Action Plan. Technical Report 0, HM Government, 2011.

Det Norske Veritas. Recommended Practice DNV-RP-C205: Environmental Conditions and Environmental Loads. Technical report, Det Norske Veritas, 2010.

J. L. Deutsch and C. V. Deutsch. Latin hypercube sampling with multidimensional uniformity. *Journal of Statistical Planning and Inference*, 142(3):763–772, March 2012. ISSN 03783758. doi: 10.1016/j.jspi.2011.09.016. URL http://linkinghub.elsevier.com/retrieve/pii/S0378375811003776.

K. W. Dixon and M. D. Shulman. A Statistical Evaluation of the Predictive Abilities of Climatic Averages. *Journal of Climate and Applied Meteorology*, 23, 1984.

L. Draper. Extreme Wave Conditions in British and adjacent waters. In *13th Coastal Engineering Conference*, New York, 1973.

L. Draper. Waves at Dowsing Light Vessel, North Sea. Technical report, Institute of Oceanographic Sciences, 1976.

L. Draper. Wave measurements at North Atlantic weather stations India and Juliett in the nineteen seventies. Technical report, Institute of Oceanographic Sciences, 1986.

D. Dupuis. Exceedances over High Thresholds: A Guide to Threshold Selection. *Extremes*, 1 (3):251 –261, 1998.

EquiMar. *Protocols for the Equitable Assessment of Marine Energy Converters*. Institute for Energy Systems, School of Engineering, University of Edinburgh, 1st edition, 2011. ISBN 9780950892016. URL http://www.see.ed.ac.uk/~shs/WaveEnergy/Equimarprotocols.pdf.

European Centre for Medium-Range Weather Forecasts. ERA-Interim, Daily fields, 2012. URL http://apps.ecmwf.int/datasets/data/interim_full_daily/.

European Centre for Medium-Range Weather Forecasts. ERA-20C, 2014. URL http://ecmwf.int/en/research/climate-reanalysis/era-20c.

J. A. Ewing, T. J. Weare, and B. A. Worthington. A Hindcast Study of Extreme Wave Conditions in the North Sea. *Journal of Geophysical Research*, 84(C9), 1979.

J. Ferreira and C. Guedes Soares. An Application of the Peaks Over Threshold Method to Predict Extremes of Significant Wave Height. *Journal of Offshore Mechanics and Arctic Engineering*, 120:165–176, 1998.

C. B. Field, V. Barros, T. F. Stocker, Q. Dahe, D. Dokken, K. Ebi, M. Mastrandrea, K. Mach, G.-K. Plattner, S. Allen, M. Tignor, and P. Midgley, editors. *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, 2012. ISBN 9781139177245. doi: 10.1017/CBO9781139177245. URL `http://ebooks.cambridge.org/ref/id/CBO9781139177245`.

R. A. Fisher. *Statistical Methods for Research Workers*. Oliver and Boyd, 1925.

Fisheries and Oceans Canada. Data and Products - Waves, 2013. URL `http://isdm.gc.ca/isdm-gdsi/waves-vagues/index-eng.asp`.

G. Flato, J. Marotzke, B. Abiodun, P. Braconnot, S. Chou, W. Collins, P. Cox, F. Driouech, S. Emori, V. Eyring, C. Forest, P. Gleckler, E. Guilyardi, C. Jakob, V. Kattsov, C. Reason, and M. Rummukainen. Evaluation of Climate Models. In T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley, editors, *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, 2013.

C. Folland, T. Karl, J. Christy, R. Clarke, G. Gruza, J. Jouzel, M. Mann, J. Oerlemans, M. Salinger, and S.-W. Wang. Observed Climate Variability and Change. In *Climate Change 2001: Working Group I: The Scientific Basis*. Cambridge University Press, 2001.

S. Fronzek, T. R. Carter, and K. Jylha. Representing two centuries of past and future climate for assessing risks to biodiversity in Europe. *Global Ecology and Biogeography*, 21:19–35, 2012.

Y. Goda. On the Methodology of selecting Design Wave Height. *Coastal Engineering*, pages 899–913, 1988.

Y. Goda. *Random Seas and Design of Maritime Structures*. World Scientific Publishing Co Pte Ltd, 3rd editio edition, 2010. ISBN 9789814282406.

I. A. Golinkin, D. D. Ruff, E. P. Kvam, G. P. McCabe, and A. F. Grandt. Application of Analysis of Variance (ANOVA) Statistical Methods to Breaking Load Corrosion Test. *Journal of Testing and Evaluation*, 25(6), 1997.

S. K. Gulev. Assessment of the reliability of wave observations from voluntary observing ships: Insights from the validation of a global wind wave climatology based on voluntary observing ship data. *Journal of Geophysical Research*, 108(C7):3236, 2003. ISSN 0148-0227. doi: 10.1029/2002JC001437. URL `http://doi.wiley.com/10.1029/2002JC001437`.

S. K. Gulev and V. Grigorieva. Last century changes in ocean wind wave height from global visual wave data. *Geophysical Research Letters*, 31(24):L24302, 2004. ISSN 0094-8276. doi: 10.1029/2004GL021040. URL `http://doi.wiley.com/10.1029/2004GL021040`.

S. K. Gulev, V. Grigorieva, K. Selemenov, and O. Zolina. Evaluation of Ocean Winds and Wind Waves from Voluntary Observation Ship data. In *Advances in the Applications of Marine Climatology - JCOMM technical report N0. 13*, pages 53–67. 2003.

H. Gunther. WAM Cycle 4.5. Technical report, Institute for Coastal Research, GKSS Research Centre Geesthacht, 2002.

H. Gunther, W. Rosenthal, M. Stawarz, J. C. Carretero, M. Gomez, I. Lozano, O. Serrano, and M. Reistad. The wave climate of the Northeast Atlantic over the period 1955-1994: the WASA wave hindcast. Technical report, DNMI, 1997.

M. E. Harr. Probabilistic estimates for multivariate analyses. *Applied Mathematical Modelling*, 13:313–318, 1989.

G. P. Harrison and R. A. Wallace. Climate Sensitivity of Marine Energy. *IEEE Transactions on Energy Conversion*, 20(4):870–877, 2005.

D. Hartmann, A. Klein Tank, M. Rusticucci, L. Alexander, S. Brönnimann, Y. Charabi, F. Dentener, E. Dlugokencky, D. Easterling, A. Kaplan, B. Soden, P. Thorne, M. Wild, and P. Zhai. Observations: Atmosphere and Surface. In T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley, editors, *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate ChangeI to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, UK and New York, USA, 2013.

D. Hasselmann, M. Dunckel, and J. Ewing. Directional wave spectra observed during JONSWAP 1973. *Journal of Physical Oceanography*, 10:1264–1280, 1980.

K. Hasselmann, T. Barnett, E. Bouws, H. Carlson, D. Cartwright, K. Enke, J. Ewing, H. Gienapp, D. Hasselmann, P. Kruseman, A. Meerburg, P. Muller, D. Olbers, K. Richter, W. Sell, and H. Walden. Measurements of Wind-Wave Growth and Swell Decay during the Joint North Sea Wave Project (JONSWAP). Technical Report 8 0, Deutsches Hydrographisches Institut, Hamburg, 1973.

S. Hasselmann, K. Hasselmann, J. H. Allender, and T. Barnett. Computations and parameterizations of the nonlinear energy transfer in a gravity-wave spectrum. Part II: Parameterizations of the nonlinear energy transfer for application in wave models. *Journal of Physical Oceanography*, 15:1378–1391, 1985.

S. Haver and K. Nyhus. A Wave Climate Description for Long Term Response Calculations. In *Fifth International Offshore Mechanics and Arctic Engineering Symposium*, 1986.

K. R. Hayes. Uncertainty and uncertainty analysis methods. Technical Report February, CSIRO, 2011.

G. C. Hegerl, F. W. Zwiers, P. Braconnot, N. P. Gillett, Y. Luo, J. A. M. Orsini, N. Nicholls, J. E. Penner, and P. A. Stott. Understanding and Attributing Climate Change. In *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. 2007.

G. Hegerl, O. Hoegh-Guldberg, G. Casassa, M. Hoerling, R. Kovats, C. Parmesan, D. Pierce, and P. Stott. Good Practice Guidance Paper on Detection and Attribution Related to Anthropogenic Climate Change. In T. Stocker, C. Field, D. Qin, V. Barros, G.-K. Plattner, M. Tignor, P. Midgley, and K. Ebi, editors, *Meeting Report of the Intergovernmental Panel on Climate Change Expert Meeting on Detection and Attribution of Anthropogenic Climate Change*. IPCC Working Group I Technical Support Unit, Bern, Switzerland, 2009.

HM Government. The UK Renewable Energy Strategy. Technical report, 2009.

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

J. Huang, H. M. van den Dool, and A. Barnston. Longlead seasonal temperature prediction using optimal climate normals. *Journal of Climate*, 9, 1996.

IEA. Key World Energy Statistics. Technical report, International Energy Agency, 2014.

IPCC. *Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Intergovernmental Panel on Climate Change, Geneva, Switzerland, 2007.

IPCC. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, 2013. URL `http://www.ipcc.ch/report/ar5/wg1/`.

IPCC. Carbon Dioxide: Projected emissions and concentrations, 2014. URL `http://www.ipcc-data.org/observ/ddc_co2.html`.

H. Janssen. Monte-Carlo based uncertainty analysis: Sampling efĩňĄciency and sampling convergence. *Reliability Engineering & System Safety*, 109:123–132, January 2013. ISSN 09518320. doi: 10.1016/j.ress.2012.08.003. URL `http://linkinghub.elsevier.com/retrieve/pii/S0951832012001536`.

P. A. E. M. Janssen. Quasi-linear theory of wind wave generation applied to wave forecasting. *Journal of Physical Oceanography*, 21:1631–1642, 1991.

E. Kaas, U. Andersen1, R. A. Flather, J. A. Williams, D. L. Blackman, P. Lionello, F. Dalan, E. Elvini, A. Nizzero, P. Malguzzi, A. Pfizenmayer, H. von Storch, D. Dillingh, M. Philippart, J. de Ronde, M. Reistad, K. H. Midtbø, O. Vignes, H. Haakenstad, B. Hackett, I. Fossum, and L. Sidselrud. Regional storm , wave and surge scenarios for the 2100 century. Technical report, Danish Meteorological Institude, 2001.

E. C. Kent and D. I. Berry. Quantifying random measurement errors in Voluntary Observing Ships' meteorological observations. *International Journal of Climatology*, 25(7):843–856, June 2005. ISSN 0899-8418. doi: 10.1002/joc.1167. URL `http://doi.wiley.com/10.1002/joc.1167`.

J. Kiecolt-Glaser, W. Garner, C. Speicher, G. Penn, J. Holliday, and R. Glaser. Psychosocial modifiers of immunocompetence in medical students. *Psychosomatic medicine*, 46(1):7–14, 1984.

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

S. Kuhnt and N. Rudak. Simultaneous Optimization of Multiple Responses with the R Package JOP. *Journal of Statistical Software*, 54(9):1–23, 2013. URL `http://www.jstatsoft.org/v54/i09/`.

P. J. Lamb and S. A. J. Changnon. On the "Best" Temperature and Precipitation Normals: The Illinois Situation. *Journal of Applied Meteorology*, 20(12), 1981.

G. Lavidas. Personal communication, 2014.

J. H. Lei and W. Schilling. Preliminary uncertainty analysis - a prerequisite for assessing the predictive uncertainty of hydrologic models. *Water Science and Technology*, 33(2):79–90, 1996.

R. V. Lenth. Response-Surface Methods in R, Using rsm. *Journal of Statistical Software*, 32 (7):1–17, 2009. URL `http://www.jstatsoft.org/v32/i07/`.

B. P. Leonard. A stable and accurate convective modelling procedure based on quadratic upstream interpolation. *Computer Methods in Applied Mechanics and Engineering*, 19: 59–98, 1979.

B. Leonard. The ULTIMATE conservative difference scheme applied to unsteady one-dimensional advection. *Computer Methods in Applied Mechanics and Engineering*, 88 (1):17–74, June 1991. ISSN 00457825. doi: 10.1016/0045-7825(91)90232-U. URL `http://linkinghub.elsevier.com/retrieve/pii/004578259190232U`.

I. P. Lewin. *Relating Statistics and Experiment Design*. SAGE University Papers Series, 1999.

R. E. Livezey, K. Y. Vinnikov, M. M. Timofeyeva, R. Tinker, and H. M. van den Dool. Estimation and Extrapolation of Climate Normals and Climatic Trends. *Journal of Applied Meteorology and Climatology*, 46(11):1759–1776, November 2007. ISSN 1558-8424. doi: 10.1175/2007JAMC1666.1. URL http://journals.ametsoc.org/doi/abs/10.1175/2007JAMC1666.1.

D. P. Loucks, E. van Beek, J. R. Stedinger, J. P. Dijkman, and M. T. Villars. Model Sensitivity and Uncertainty Analysis. In *Water resources systems planning and management: an introduction to methods, models and applications*, pages 255–290. UNESCO, 2005. URL https://ecommons.library.cornell.edu/bitstream/1813/2804/12/09_chapter09.pdf.

J. A. Lowe, T. Howard, A. Pardaens, J. Tinker, G. Jenkins, J. Ridley, J. Leake, J. Holt, S. Wakelin, J. Wolf, K. Horsburgh, T. Reeder, G. Milne, and S. Bradley. UK Climate Projections science report: Marine and coastal projections. Technical report, Met Office Hadley Centre, Exeter UK, 2009.

E. B. Mackay, A. S. Bahaj, and P. G. Challenor. Uncertainty in wave energy resource assessment. Part 2: Variability and predictability. *Renewable Energy*, 35(8):1809–1819, August 2010. ISSN 09601481. doi: 10.1016/j.renene.2009.10.027. URL http://linkinghub.elsevier.com/retrieve/pii/S0960148109004534.

Marine Institute. Weather Buoy Network, 2013. URL http://data.marine.ie/Dataset/Details/20972.

M. Mathiesen, Y. Goda, P. J. Hawkes, E. Mansard, M. J. Martin, E. Peltier, E. F. Thompson, and G. Van Vledder. Recommended Practice for Extreme Wave Analysis. *Journal of Hydraulic Research*, 32(6):803–814, 1994.

C. J. Mcadie, C. W. Landsea, C. J. Neumann, J. E. David, E. S. Blake, and G. R. Hammer. *Tropical cyclones of the North Atlantic Ocean, 1851-2006.* National Climatic Data Center, National Oceanic and Atmospheric Administration, Asheville, sixth revi edition, 2009.

M. D. Mckay, R. J. Beckman, and W. J. Conover. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics*, 42(1):55–61, 2000.

G. A. Meehl, C. Covey, K. E. Taylor, T. Delworth, R. J. Stouffer, M. Latif, B. McAvaney, and J. F. B. Mitchell. THE WCRP CMIP3 Multimodel Dataset: A New Era in Climate Change Research. *Bulletin of the American Meteorological Society*, 88(9):1383–1394, September 2007. ISSN 0003-0007. doi: 10.1175/BAMS-88-9-1383. URL http://journals.ametsoc.org/doi/abs/10.1175/BAMS-88-9-1383.

R. H. Moss, J. a. Edmonds, K. a. Hibbard, M. R. Manning, S. K. Rose, D. P. van Vuuren, T. R. Carter, S. Emori, M. Kainuma, T. Kram, G. a. Meehl, J. F. B. Mitchell, N. Nakicenovic, K. Riahi, S. J. Smith, R. J. Stouffer, A. M. Thomson, J. P. Weyant, and T. J. Wilbanks. The next generation of scenarios for climate change research and assessment. *Nature*, 463 (7282):747–56, February 2010. ISSN 1476-4687. doi: 10.1038/nature08823. URL `http://www.ncbi.nlm.nih.gov/pubmed/20148028`.

R. H. Myers and D. C. Montgomery. *Response Surface Methodology*. Wiley Series in Probability and Statistics, second edi edition, 2002. ISBN 0-471-41255-4.

N. Nakicenovic, J. Alcamo, G. Davis, B. de Vries, J. Fenhann, S. Gaffin, K. Gregory, A. Grubler, T. Y. Jung, T. Kram, E. L. La Rovere, L. Michaelis, S. More, T. Morita, W. Pepper, H. Pitcher, L. Price, K. Riahi, A. Roehrl, H.-H. Rogner, A. Saskovski, M. Schlesinger, P. Shukla, S. Smith, R. Swart, S. van Rooijen, N. Victor, and Z. Dadi. Special Report on Emission Scenarios. Technical report, Intergovernmental Panel on Climate Change (IPCC), Geneva, Switzerland, 2000. URL `www.ipcc.ch/ipccreports/sres/emission/index.php?idp=0`.

National Data Buoy Center. Historical NDBC Data, 2013. URL `www.ndbc.noaa.gov/historical_data.shtml`.

National Geophysical Data Center. ETOPO1 Global Relief Model, 2008. URL `http://www.ngdc.noaa.gov/mgg/global/global.html`.

National Weather Service - Environmental Modelling Center (NOAA). WAVEWATCH III Model - Model Distribution, 2012. URL `http://polar.ncep.noaa.gov/waves/wavewatch/wavewatch.shtml`.

S. Neelamani, K. Al-Salem, and K. Rakha. Extreme Water Waves in the UAE Territorial Waters. *Emirates Journal for Engineering Research*, 11(2):37–46, 2006.

H. J. A. Neu. Interannual Variations and Longer-Term Changes in the Sea State of the North Atlantic from 1970 to 1982. *Journal of Geophysical Research*, 89(4):6397–6402, 1984.

M. Ochi. New approach for estimating the Severest Sea State from statistical data. In *Coastal Engineering*, 1992.

C. J. Paciorek, J. S. Risbey, V. Ventura, and R. D. Rosen. Multiple Indices of Northern Hemisphere Cyclone Activity, Winters 1949 âĂŞ 99. *Journal of Climate*, 15:1573–1590, 2002.

V. Panchang, L. Zhao, and Z. Demirbilek. Estimation of extreme wave heights using GEOSAT measurements. *Ocean Engineering*, 26(3):205–225, 1998.

P. Parysow, G. Gertner, and J. Westervelt. Efficient approximation for building error budgets for process models. *Ecological Modelling*, 135(2-3):111–125, December 2000. ISSN 03043800. doi: 10.1016/S0304-3800(00)00347-1. URL http://linkinghub.elsevier.com/retrieve/pii/S0304380000003471.

B. Pfaff and A. McNeil. evir: Extreme Values in R, 2012. URL http://cran.r-project.org/package=evir.

I. Polyak. *Computational statistics in climatology*. Oxford University Press, Oxford, New York, 1996.

D. Randall, R. Wood, S. Bony, R. Colman, T. Fichefet, J. Fyfe, V. Kattsov, A. Pitman, J. Shukla, J. Srinivasan, R. Stouffer, A. Sumi, and K. Taylor. Climate Models and Their Evaluation. In S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor, and H. Miller, editors, *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, 2007.

N. Rangel-Buitrago and G. Anfuso. Winter wave climate, storms and regional cycles: the SW Spanish Atlantic coast. *International Journal of Climatology*, 33(9):2142–2156, July 2013. ISSN 08998418. doi: 10.1002/joc.3579. URL http://doi.wiley.com/10.1002/joc.3579.

D. Reeve, Y. Chen, S. Pan, V. Magar, D. Simmonds, and a. Zacharioudaki. An investigation of the impacts of climate change on wave energy generation: The Wave Hub, Cornwall, UK. *Renewable Energy*, 36(9):2404–2413, September 2011. ISSN 09601481. doi: 10.1016/j.renene.2011.02.020. URL http://linkinghub.elsevier.com/retrieve/pii/S096014811100111X.

T. Reichler and J. Kim. How Well Do Coupled Models Simulate Today's Climate? *Bulletin of the American Meteorological Society*, 89(3):303–311, March 2008. ISSN 0003-0007. doi: 10.1175/BAMS-89-3-303. URL http://journals.ametsoc.org/doi/abs/10.1175/BAMS-89-3-303.

Renewable Policy Team. Framework for the development and deployment of renewables in Scotland. Technical report, Scottish Government, 2008.

J. A. Rice. *Mathematical Statistics and Data Analysis*. Wadsworth Publishing Co Inc., 1994. ISBN 0-534-20934-3.

E. Roeckner, G. Bäuml, L. Bonaventura, R. Brokopf, M. Esch, M. Giorgetta, S. Hagemann, I. Kirchner, L. Kornblueh, E. Manzini, A. Rhodin, U. Schlese, U. Schulzweida, and A. Tompkins. The atmospheric general circulation model ECHAM5. Part I: Model description. Rep. No. 349. Technical report, Max-Planck-Institut für Meteorologie, Hamburt, 2003.

V. K. Rohatgi and A. K. M. E. Saleh. *An Introduction to Probability and Statistics*. Wiley Series in Probability and Statistics, 2nd editio edition, 2011.

S. Saha, S. Moorthi, H.-L. Pan, X. Wu, J. Wang, S. Nadiga, P. Tripp, R. Kistler, J. Woollen, D. Behringer, H. Liu, D. Stokes, R. Grumbine, G. Gayno, J. Wang, Y.-T. Hou, H.-Y. Chuang, H.-M. H. Juang, J. Sela, M. Iredell, R. Treadon, D. Kleist, P. V. Delst, D. Keyser, J. Derber, M. Ek, J. Meng, H. Wei, R. Yang, S. Lord, H. V. D. Dool, A. Kumar, W. Wang, C. Long, M. Chelliah, Y. Xue, B. Huang, J.-K. Schemm, W. Ebisuzaki, R. Lin, P. Xie, M. Chen, S. Zhou, W. Higgins, C.-Z. Zou, Q. Liu, Y. Chen, Y. Han, L. Cucurull, R. W. Reynolds, G. Rutledge, and M. Goldberg. The NCEP Climate Forecast System Reanalysis. *Bulletin of the American Meteorological Society*, 91(8):1015–1057, April 2010. ISSN 0003-0007. doi: 10.1175/2010BAMS3001.1. URL `http://dx.doi.org/10.1175/2010BAMS3001.1`.

H. Sahai and M. I. Ageel. *The Analysis of Variance - Fixed, Random and Mixed Models*. Birkhauser Boston, 2000.

Salas-Mélia, F. C. D., M. Déqué, H. Douville, J. Gueremy, P. Marquet, S. Planton, J. Royer, and S. Tyteca. Description and validation of the CNRM-CM3 global coupled model, CNRM working note 103. Technical report, Centre national de recherches météorologiques, Météo-France, 2005.

A. Saltelli, K. Chan, and E. M. Scott, editors. *Sensitivity analysis*. Saltelli, Andrea Chan, Karen Scott, E. Marian, 2000. ISBN 0471998923.

J. Slingo, S. Belcher, A. Scaife, M. McCarthy, A. Saulter, K. McBeath, A. Jenkins, C. Huntingford, T. Marsh, J. Hannaford, and S. Parry. The Recent Storms and Floods in the UK. Technical Report February, UK Met Office, 2014.

D. M. Spindler and H. L. Tolman. Example of WAVEWATCH III for the NE Pacific. Technical Report 259, National Oceanic and Atmospheric Administration, 2008. URL `http://polar.ncep.noaa.gov/mmab/papers/tn259/mmab259.pdf`.

V. R. Swail, A. T. Cox, and V. J. Cardone. Analysis of wave climate trends and variability. In *CLIMAR II - Second JCOMM Workshop on Advances in the Applications of Marine Climatology*, pages 217–226, 2003.

A. Tancredi, C. Anderson, and A. O'Hagen. Accounting for threshold uncertainty in extreme value estimation. *Extremes*, 9:87–106, 2006.

M. a. C. Teixeira and S. E. Belcher. On the distortion of turbulence by a progressive surface wave. *Journal of Fluid Mechanics*, 458:229–267, May 2002. ISSN 0022-1120. doi: 10.1017/S0022112002007838. URL `http://www.journals.cambridge.org/abstract_S0022112002007838`.

The Crown Estate. Energy and Infrastructure: Wave and Tidal - Portfolio, 2014. URL http://www.thecrownestate.co.uk/energy-and-infrastructure/wave-and-tidal/our-portfolio/.

The WAMDI Group. The WAM Model - A Third Generation Ocean Wave Prediction Model. *Journal of Physical Oceanography*, 18:1775–1810, 1988.

H. L. Tolman. A Third-Generation Model for Wind Waves on Slowly Varying, Unsteady and Inhomogenous Depths and Currents. *Journal of Physical Oceanography*, 21:782–797, 1991.

H. L. Tolman. On the selection of propagation schemes for a spectral wind-wave model. Technical Report November, National Ocean and Atmospheric Administration, 1995. URL http://polar.ncep.noaa.gov/mmab/papers/tn113/OMB113.pdf.

H. L. Tolman. Numerics in wind wave models. In *ECMWF Workshop on Ocean Wave Forecasting*, pages 5–16, 2001.

H. L. Tolman. Alleviating the Garden Sprinkler Effect in wind wave models. *Ocean Modelling*, 4:269–289, 2002a.

H. L. Tolman. Distributed Memory Concepts in the wave model WAVEWATCH III. *Parallel Computing*, 28:35–52, 2002b.

H. L. Tolman and D. Chalikov. Source Terms in a Third-Generation Wind Wave Model. *Journal of Physical Oceanography*, 26(11):2497–2518, November 1996. ISSN 0022-3670. doi: 10.1175/1520-0485(1996)026<2497:STIATG>2.0.CO;2. URL http://dx.doi.org/10.1175/1520-0485(1996)026<2497:STIATG>2.0.CO2.

H. L. Tolman, J.-H. G. Alves, and Y. Y. Chao. Operational Forecasting of Wind-Generated Waves by Hurricane Isabel at NCEP. *Weather and Forecasting*, 20:544–557, 2005. URL http://journals.ametsoc.org/doi/pdf/10.1175/WAF852.1.

H. L. Tolman, M. Accensi, H. Alves, F. Ardhuin, J. Bidlot, N. Booij, A.-C. Bennis, D. V. Chalikov, T. Campbell, A. Chawla, J.-F. Filipot, M. Foreman, P. Janssen, F. Leckler, J.-G. Li, K. Lind, M. Orzech, R. PadillaâĂŞHernÂt'andez, W. E. Rogers, A. Rawat, A. Roland, M. D. Sikiric, M. Szyszka, B. Tracy, G. P. van Vledder, A. van der Westhuysen, and S. Zieger. User manual and system documentation of WAVEWATCH III version 4.18. Technical Report 316, National Oceanic and Atmosphere Administration, National Centers for Environmental Prediction (NOAA/NCEP), 2014a. URL http://polar.ncep.noaa.gov/waves/wavewatch/manual.v4.18.pdf.

H. L. Tolman, A. Chawla, and H. Alves. Personal communication, 2014b.

F. Triefenbach. *Design of Experiments: The D-optimal Approach and its Implementation as a Computer Algorithm*. PhD thesis, Umea University / South Westphalia University of Applied Sciences, 2008.

M. J. Tucker. *Waves in Ocean Engineering*. Ellis Horwood, 1991.

A. Underwood. Techniques of analysis of variance in experimental marine biology and ecology. *Oceanography and Marine Biology an Annual Review*, 19:513–605, 1981.

A. Van der Westhuysen. Modeling nearshore wave processes. In *ECMWF Workshop on Ocean Waves*, number 298, pages 25–27, Reading, 2012. European Centre for Medium-Range Weather Forecasting.

A. J. van der Westhuysen, M. Zijlema, and J. a. Battjes. Nonlinear saturation-based whitecapping dissipation in SWAN for deep and shallow water. *Coastal Engineering*, 54 (2):151–170, February 2007. ISSN 03783839. doi: 10.1016/j.coastaleng.2006.08.006. URL `http://linkinghub.elsevier.com/retrieve/pii/S037838390600127X`.

G. van Vledder, M. Zijlema, and L. H. Holthuijsen. Revisiting the JONSWAP Bottom Friction Formulation. In *International Conference on Coastal Engineering*, pages 1–8, 2010.

G. P. van Vledder. The WRT method for the computation of non-linear four-wave interactions in discrete spectral wave models. *Coastal Engineering*, 53(2-3):223–242, February 2006. ISSN 03783839. doi: 10.1016/j.coastaleng.2005.10.011. URL `http://linkinghub.elsevier.com/retrieve/pii/S0378383905001377`.

H. Verbeeck, R. Samson, F. Verdonck, and R. Lemeur. Parameter sensitivity and uncertainty of the forest carbon flux model FORUG: a Monte Carlo analysis. *Tree physiology*, 26(6): 807–17, June 2006. ISSN 0829-318X. URL `http://www.ncbi.nlm.nih.gov/pubmed/16510397`.

J. Vinoth and I. R. Young. Global Estimates of Extreme Wind Speed and Wave Height. *Journal of Climate*, 24(6):1647–1665, March 2011. ISSN 0894-8755. doi: 10.1175/2010JCLI3680.1. URL `http://journals.ametsoc.org/doi/abs/10.1175/2010JCLI3680.1`.

H. von Storch and H. Reichardt. A Scenario of Storm Surge Statistics for the German Bight at the Expected Time of Doubled Atmospheric Carbon Dioxide Concentration. *Journal of Climate*, 10(10):2653–2662, October 1997. ISSN 0894-8755. doi: 10.1175/1520-0442(1997)010<2653:ASOSSS>2.0.CO;2. URL `http://journals.ametsoc.org/doi/abs/10.1175/1520-0442%281997%29010%3C2653%3AASOSSS%3E2.0.CO%3B2`.

X. L. Wang, F. W. Zwiers, and V. R. Swail. North Atlantic Ocean Wave Climate Change Scenarios for the Twenty-First Century. *Journal of Climate*, 17:2368–2383, 2004.

X. L. Wang, V. R. Swail, and F. W. Zwiers. Climatology and Changes of Extratropical Cyclone Activity : Comparison of ERA-40 with NCEP âĂŞ NCAR Reanalysis for 1958 âĂŞ 2001. *Journal of Climate*, 19:3145–3166, 2006.

WASA Group. Changing Waves and Storms in the Northeast Atlantic ? *Bulletin of the American Meteorological Society*, (January), 1998.

G. Watson, J. Halliday, J. Palutikof, T. Holt, R. Barthelmie, J. Coelingh, E. van Zuylen, and J. Cleijne. Predicting Offshore Wind Energy Resources (POWER). Technical report, 2001.

J. S. Whitaker, G. P. Compo, X. Wei, and T. M. Hamill. Reanalysis without Radiosondes Using Ensemble Data Assimilation. *Monthly Weather Review*, 132:1190–1200, 2004.

G. B. Whitham. A general approach to linear and non-linear dispersive waves using a Lagrangian. *Journal of Fluid Mechanics*, 22(02):273, March 2006. ISSN 0022-1120. doi: 10.1017/S0022112065000745. URL `http://www.journals.cambridge.org/abstract_S0022112065000745`.

M. O. Williams. *Wave mapping in UK waters.* HSE Books, rr621 edition, 2008.

W. Wimmer, P. Challenor, and C. Retzler. Extreme wave heights in the North Atlantic from Altimeter Data. *Renewable Energy*, 31:241–248, 2006.

D. K. Woolf, P. G. Challenor, and P. D. Cotton. Variability and predictability of the North Atlantic wave climate. *Science*, 275(5297):181–184, 2002. URL `http://dx.doi.org/10.1029/2001JC001124`.

World Meteorological Organization. Calculation of Monthly and Annual 30-Year Standard Normals. WCDP-No. 10, WMO-TD/No. 341. Technical report, World Meteorological Organization, 1989.

I. R. Young, S. Zieger, and a. V. Babanin. Global trends in wind speed and wave height. *Science*, 332(6028):451–5, April 2011. ISSN 1095-9203. doi: 10.1126/science.1197219. URL `http://www.ncbi.nlm.nih.gov/pubmed/21436400`.

W. Zhao, S. S. Chen, C. A. Blain, and J. Tian. Effects of winds, tides, and storm surges on ocean surface waves in the Japan/East Sea. In *Fifth Conference on Coastal Atmospheric and Oceanic Prediction and Processes*, pages 2–6, Seattle, 2003.

Appendix A

Numerical modelling - Figures

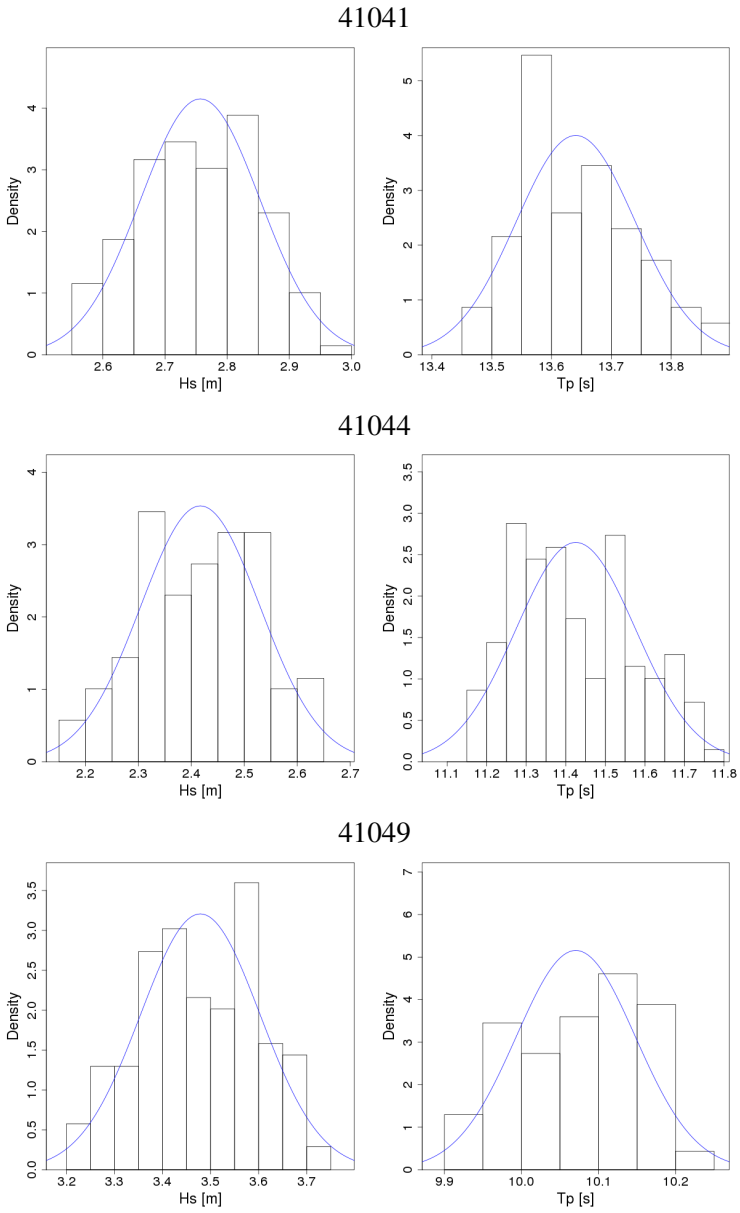

Figure A.1: Histograms of output from behavioural trials for 0000hrs January 6, 2010 (a)

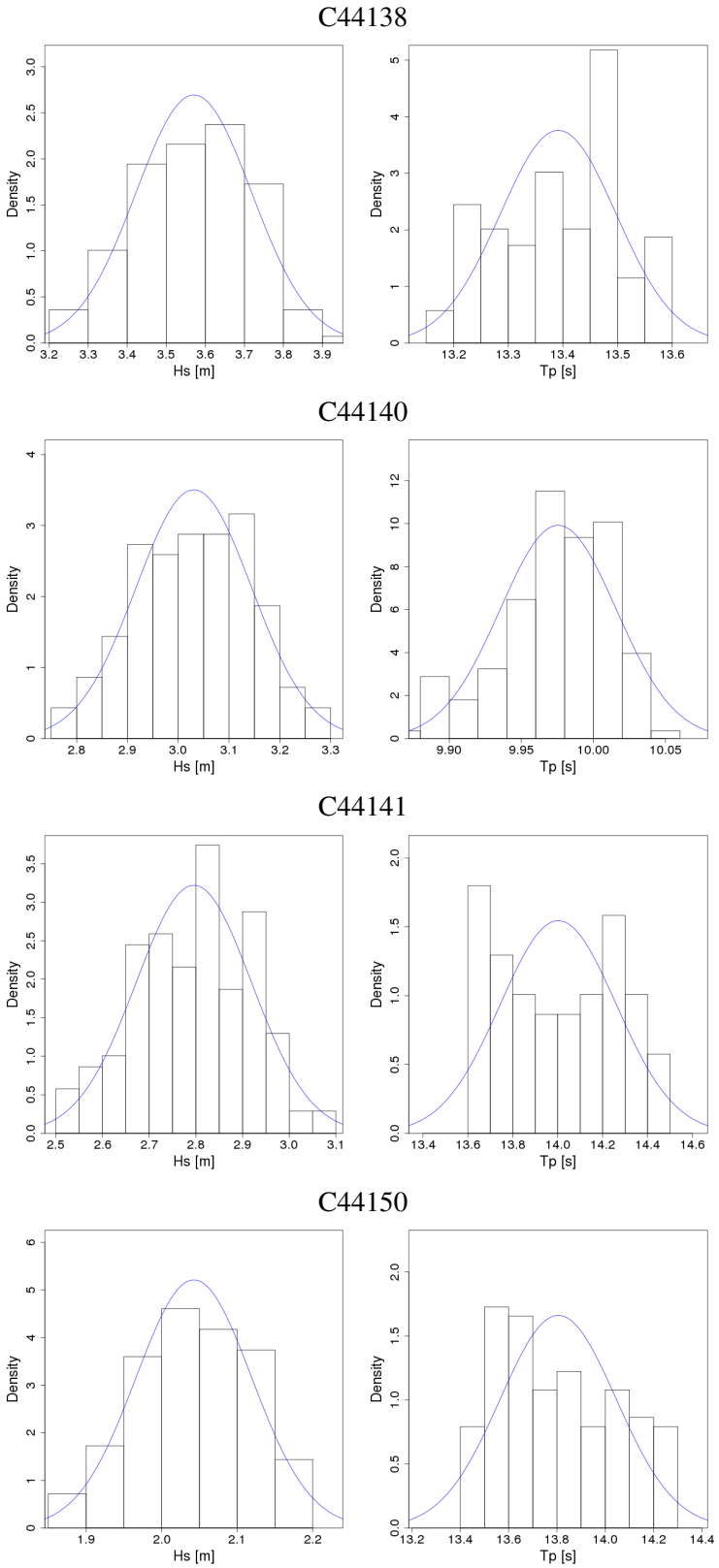

Figure A.2: Histograms of output from behavioural trials for 0000hrs January 6, 2010 (b)

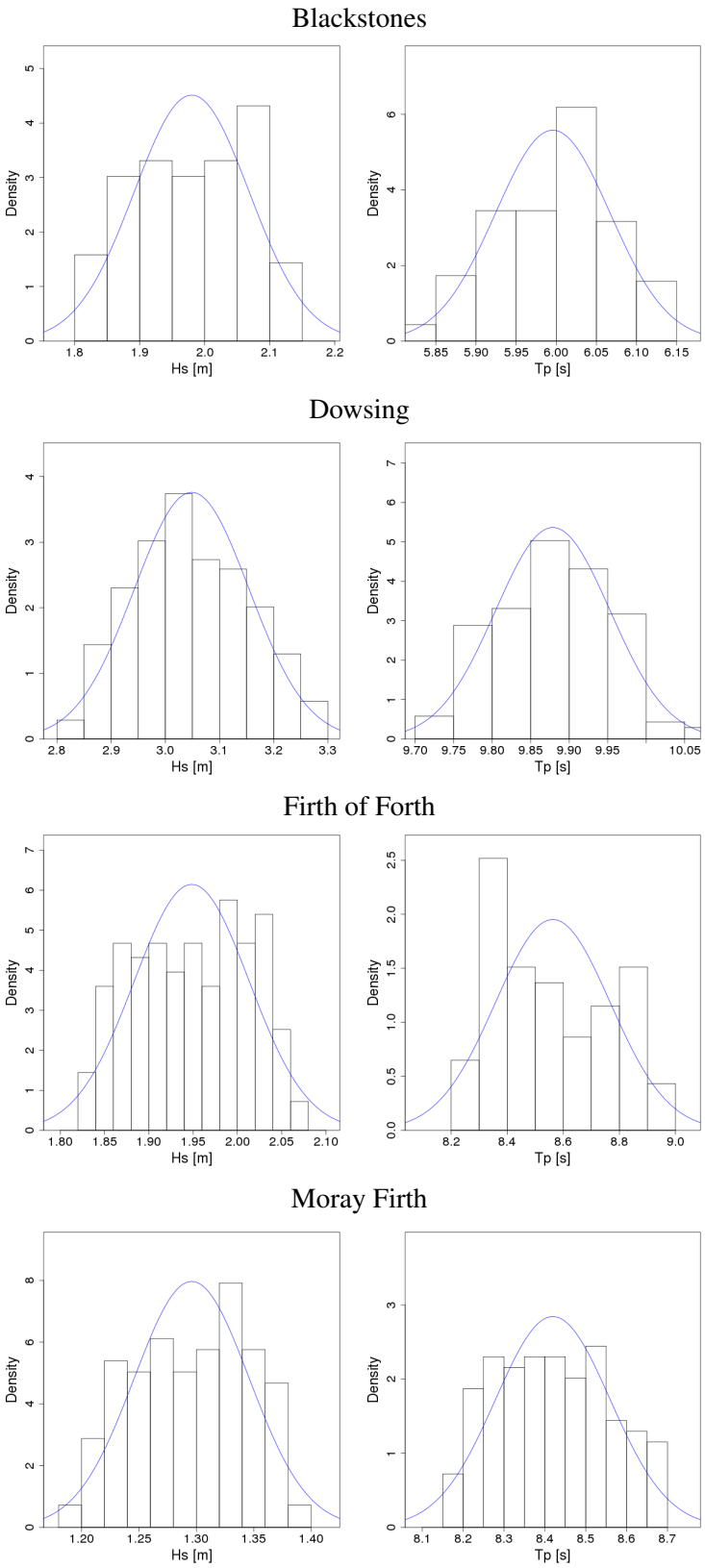

Figure A.3: Histograms of output from behavioural trials for 0000hrs January 6, 2010 (c)

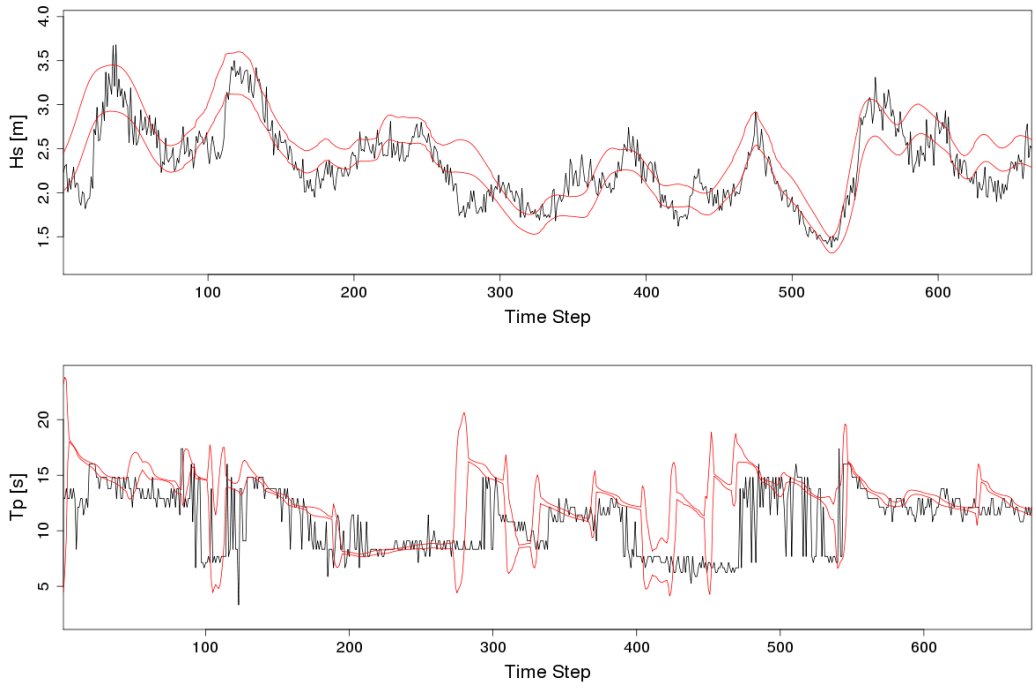

Figure A.4: Time-series plots for January 2010 along with buoy measurements for the location of NDBC buoy 41041

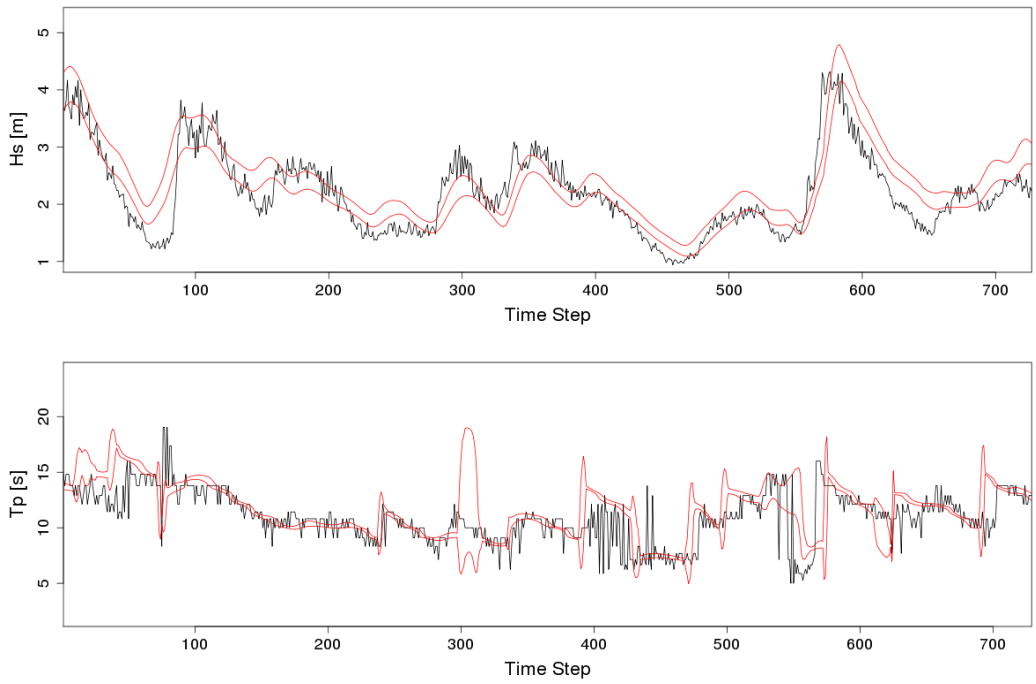

Figure A.5: Time-series plots for January 2010 along with buoy measurements for the location of NDBC buoy 41044

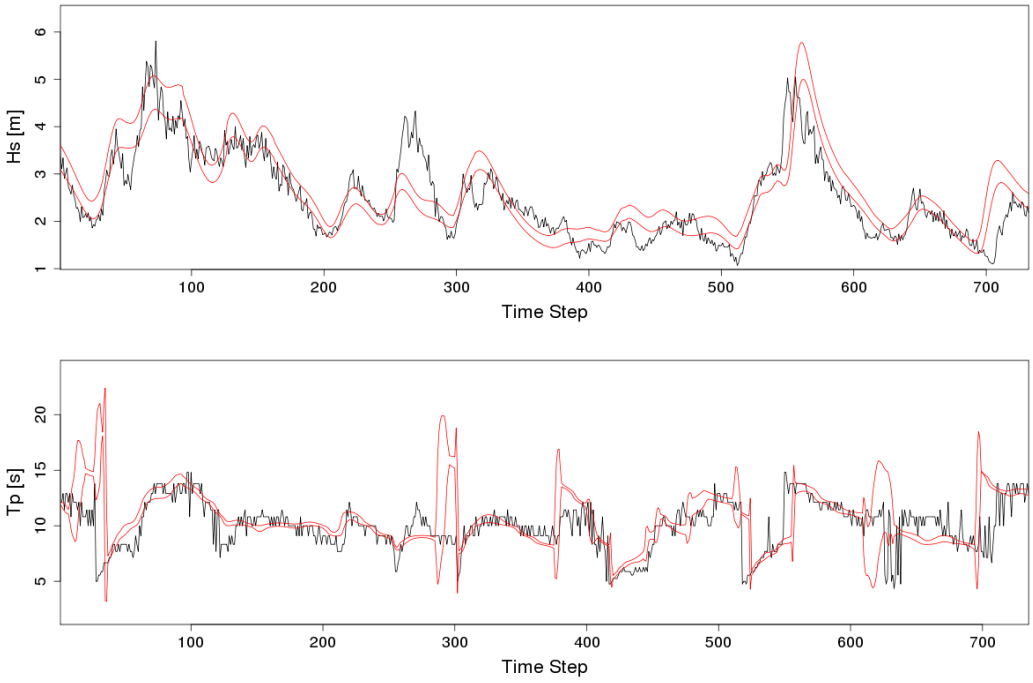

Figure A.6: Time-series plots for January 2010 along with buoy measurements for the location of NDBC buoy 41049

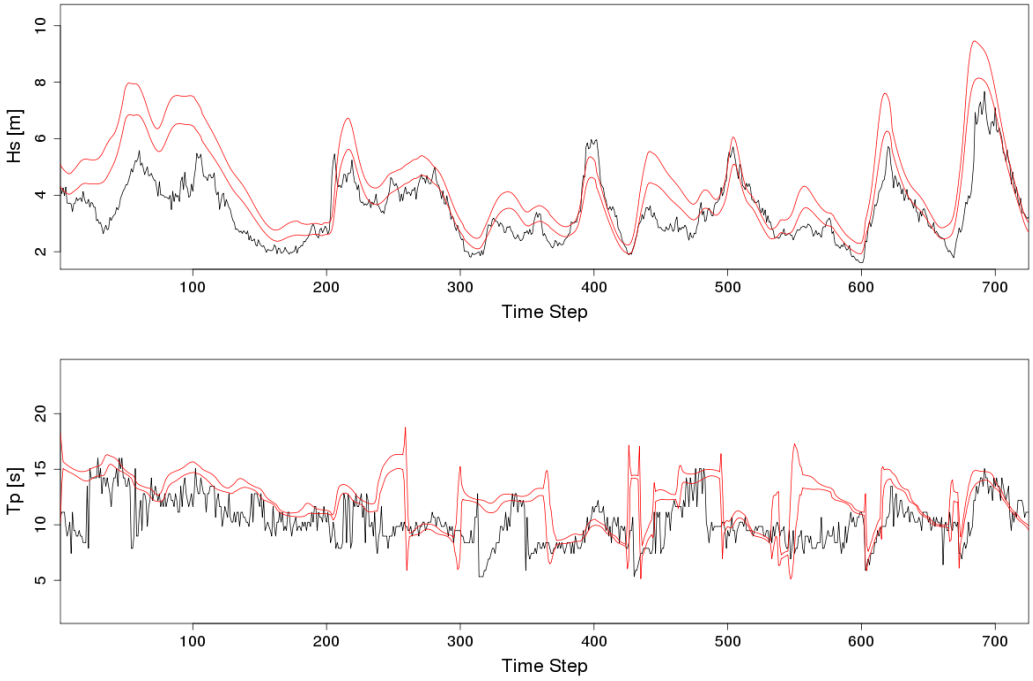

Figure A.7: Time-series plots for January 2010 along with buoy measurements for the location location of Canadian buoy C44138

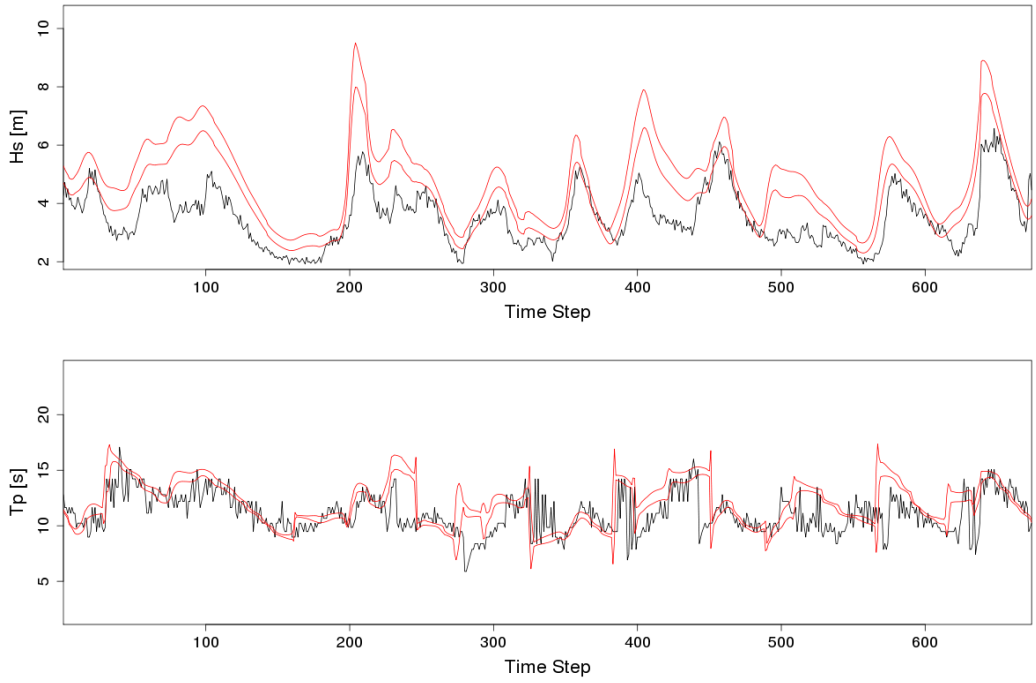

Figure A.8: Time-series plots for January 2010 along with buoy measurements for the location of Canadian buoy C44140

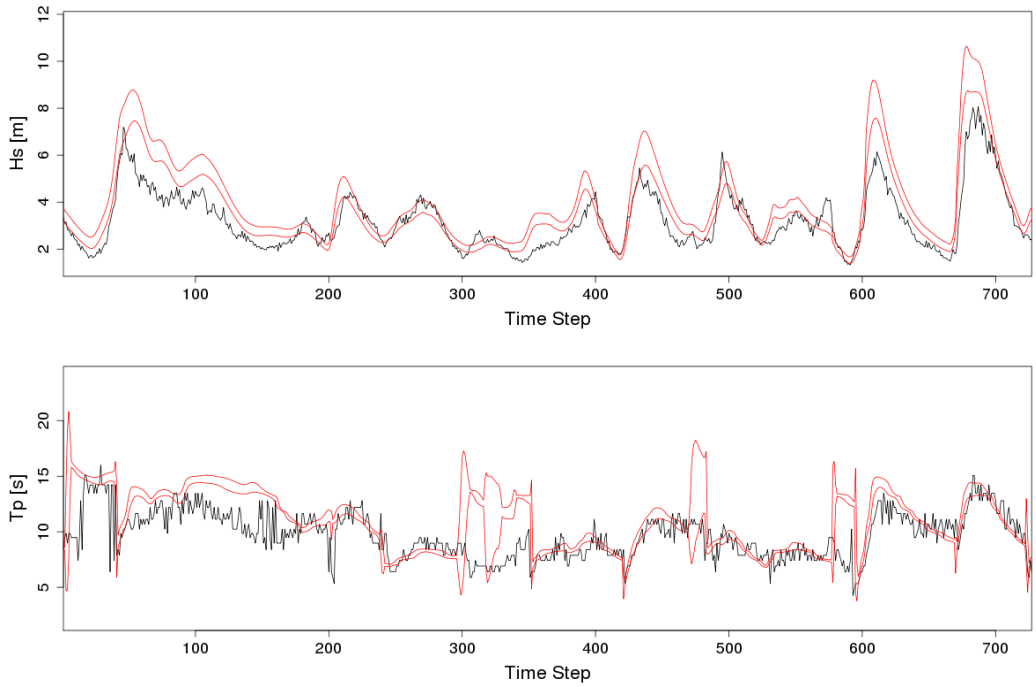

Figure A.9: Time-series plots for January 2010 along with buoy measurements for the location of Canadian buoy C44141

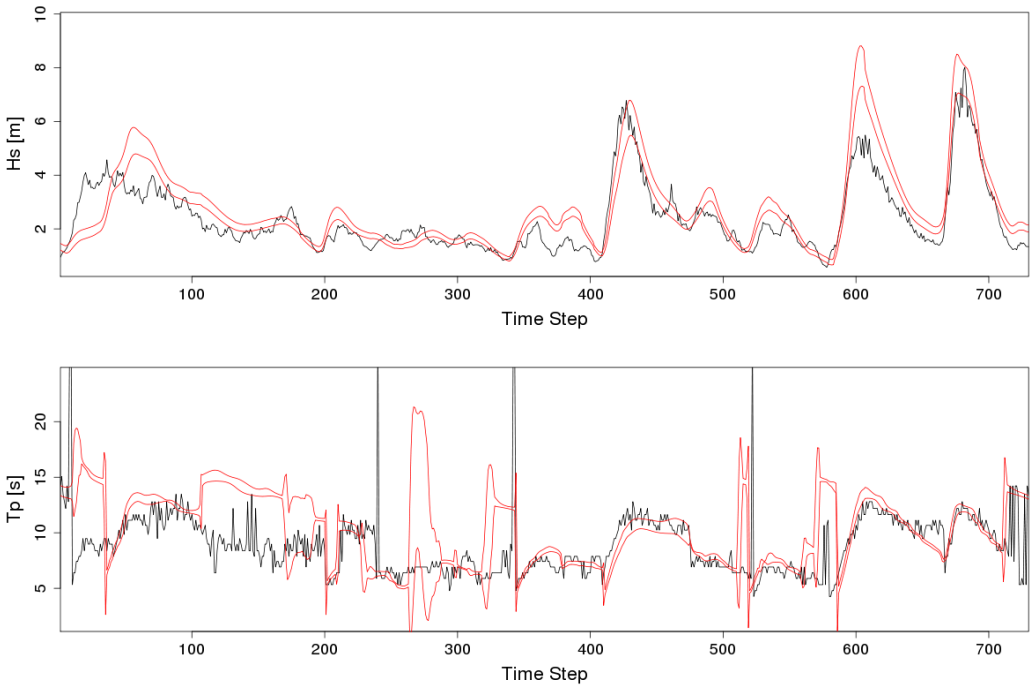

Figure A.10: Time-series plots for January 2010 along with buoy measurements for the location of Canadian buoy C44150

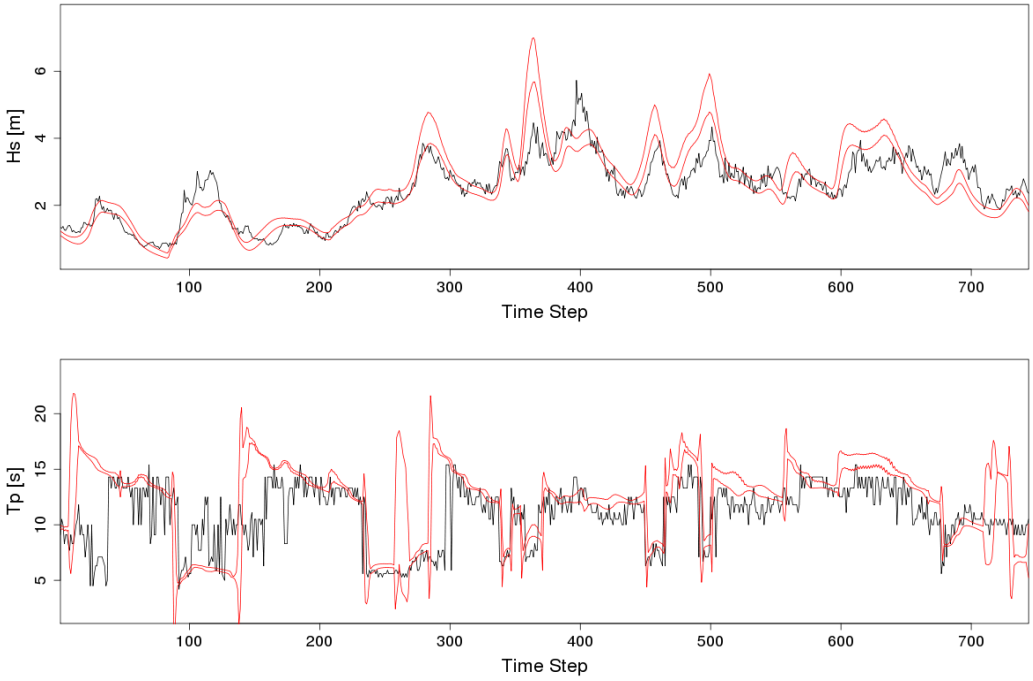

Figure A.11: Time-series plots for January 2010 along with buoy measurements for the location of CEFAS buoy at Blackstones

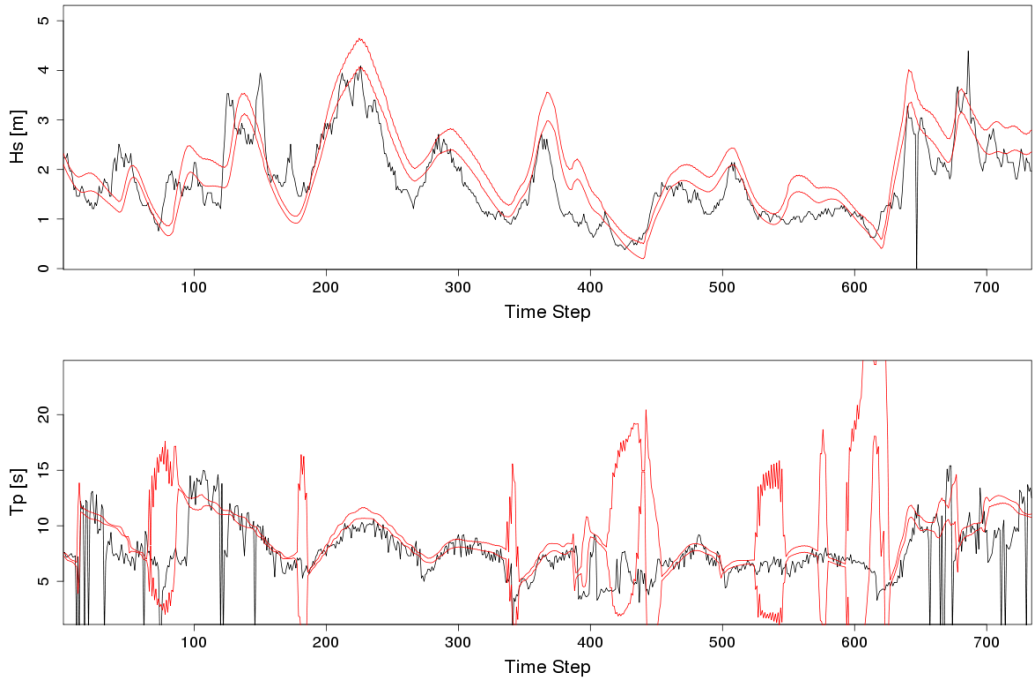

Figure A.12: Time-series plots for January 2010 along with buoy measurements for the location of CEFAS buoy at Dowsing

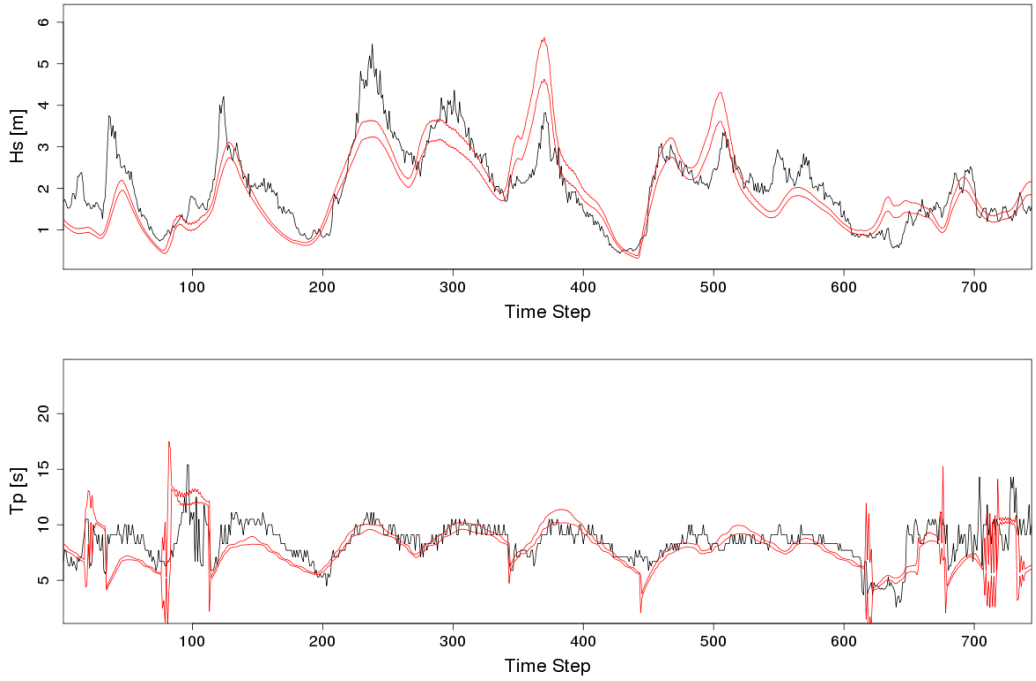

Figure A.13: Time-series plots for January 2010 along with buoy measurements for the location of CEFAS buoy at Firth of Forth

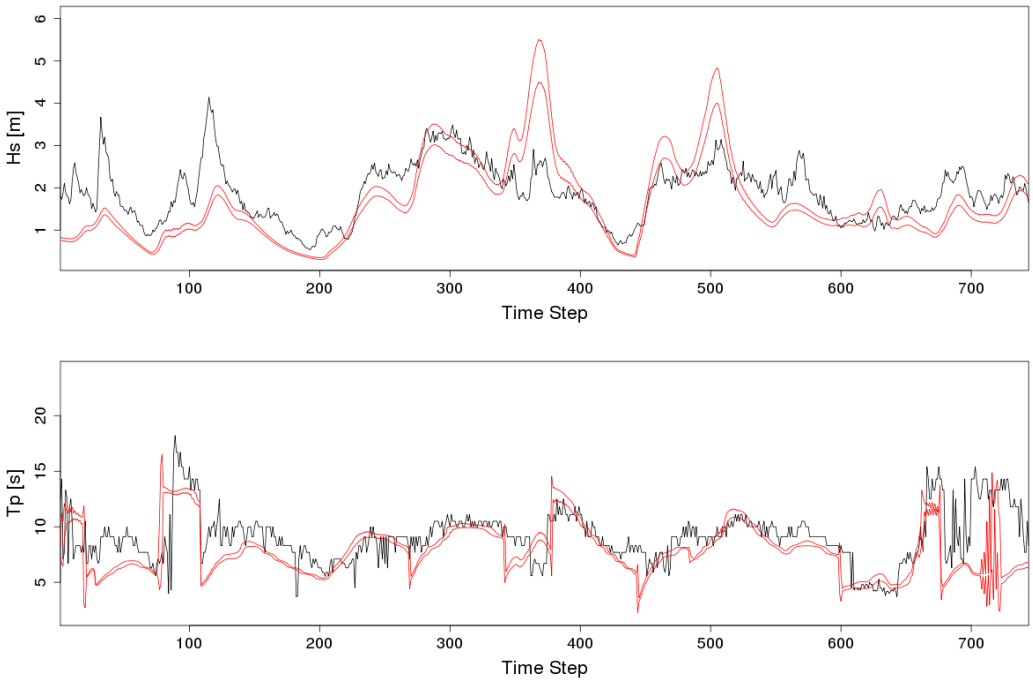

Figure A.14: Time-series plots for January 2010 along with buoy measurements for the location of CEFAS buoy at Moray Firth

Appendix B

Analysis of hindcast wave climate - Figures

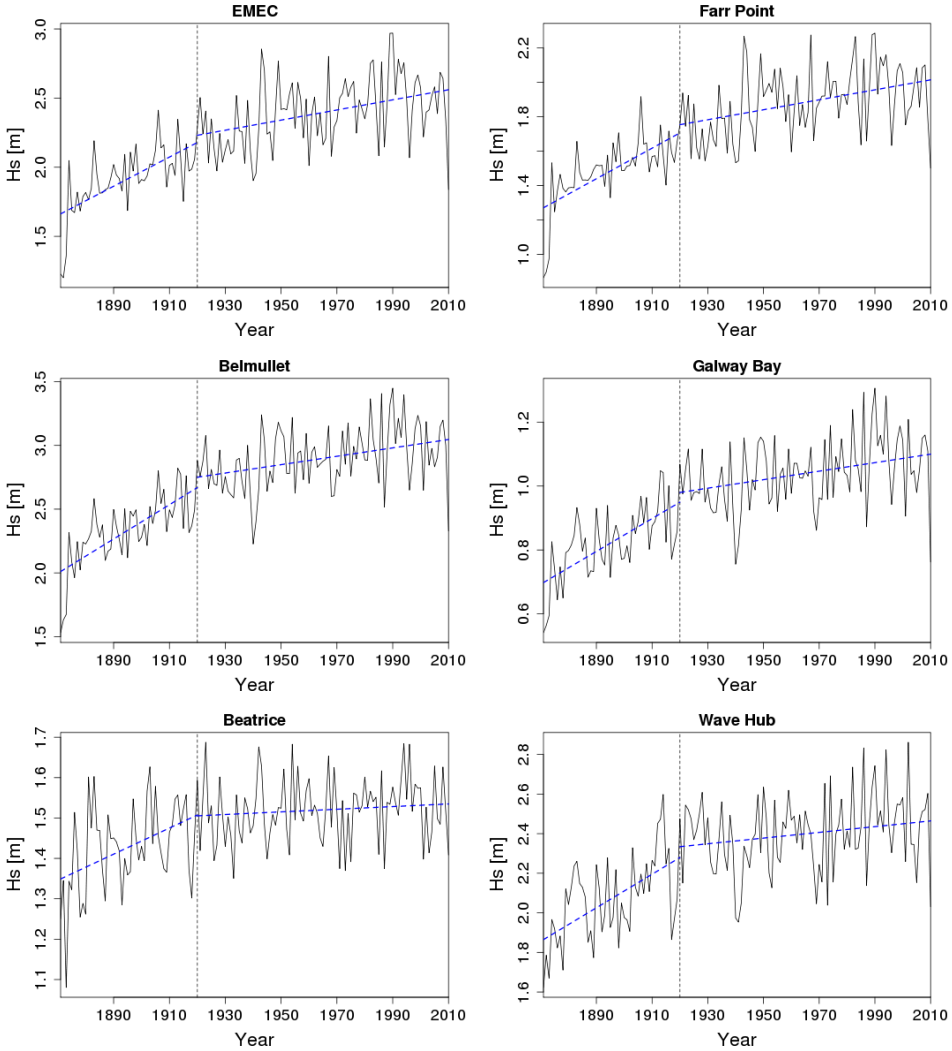

Figure B.1: 140 year time-series of annual mean significant wave height for selected locations with the hinge fit trend superimposed (data from WAVEWATCH III)

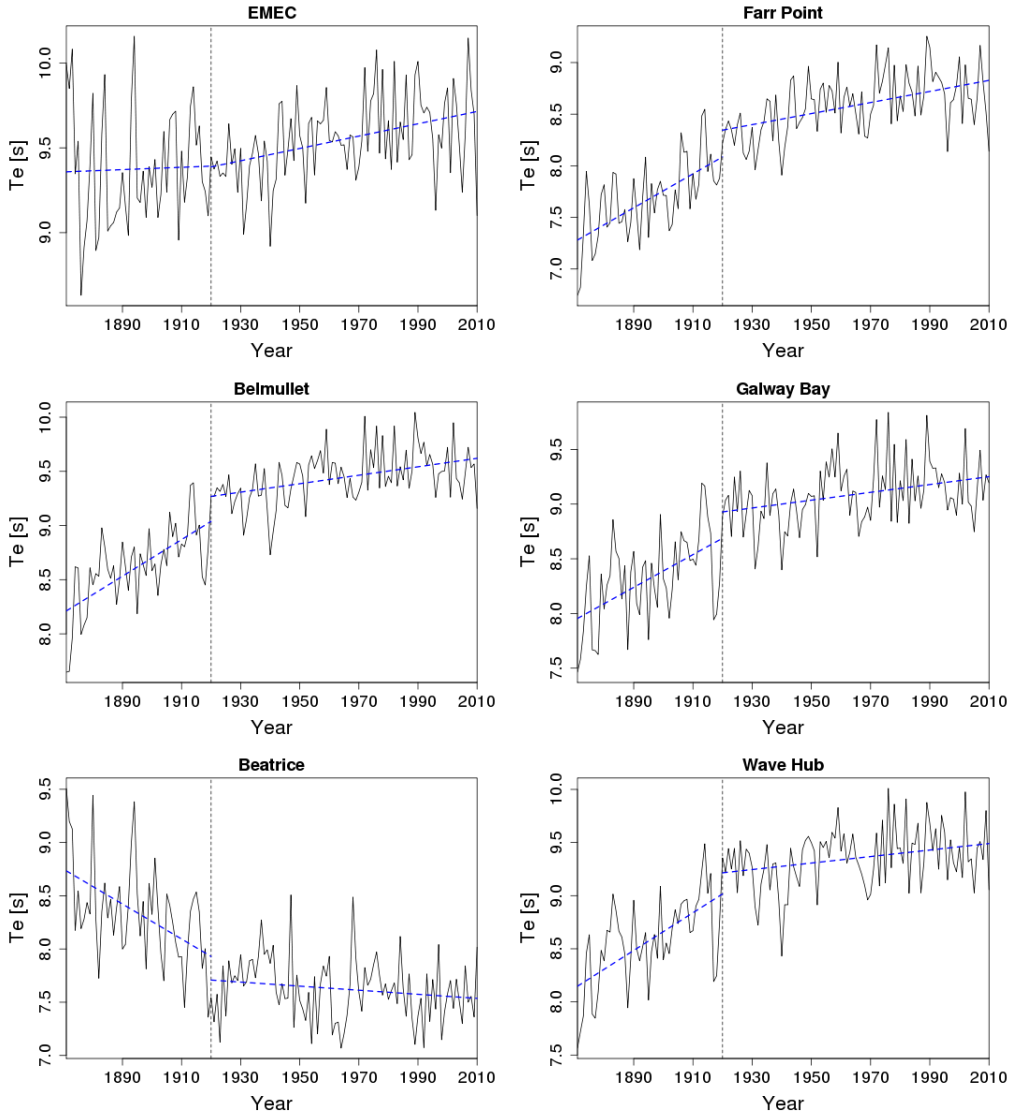

Figure B.2: 140 year time-series of annual mean energy period for selected locations with the hinge fit trend superimposed (data from WAVEWATCH III)

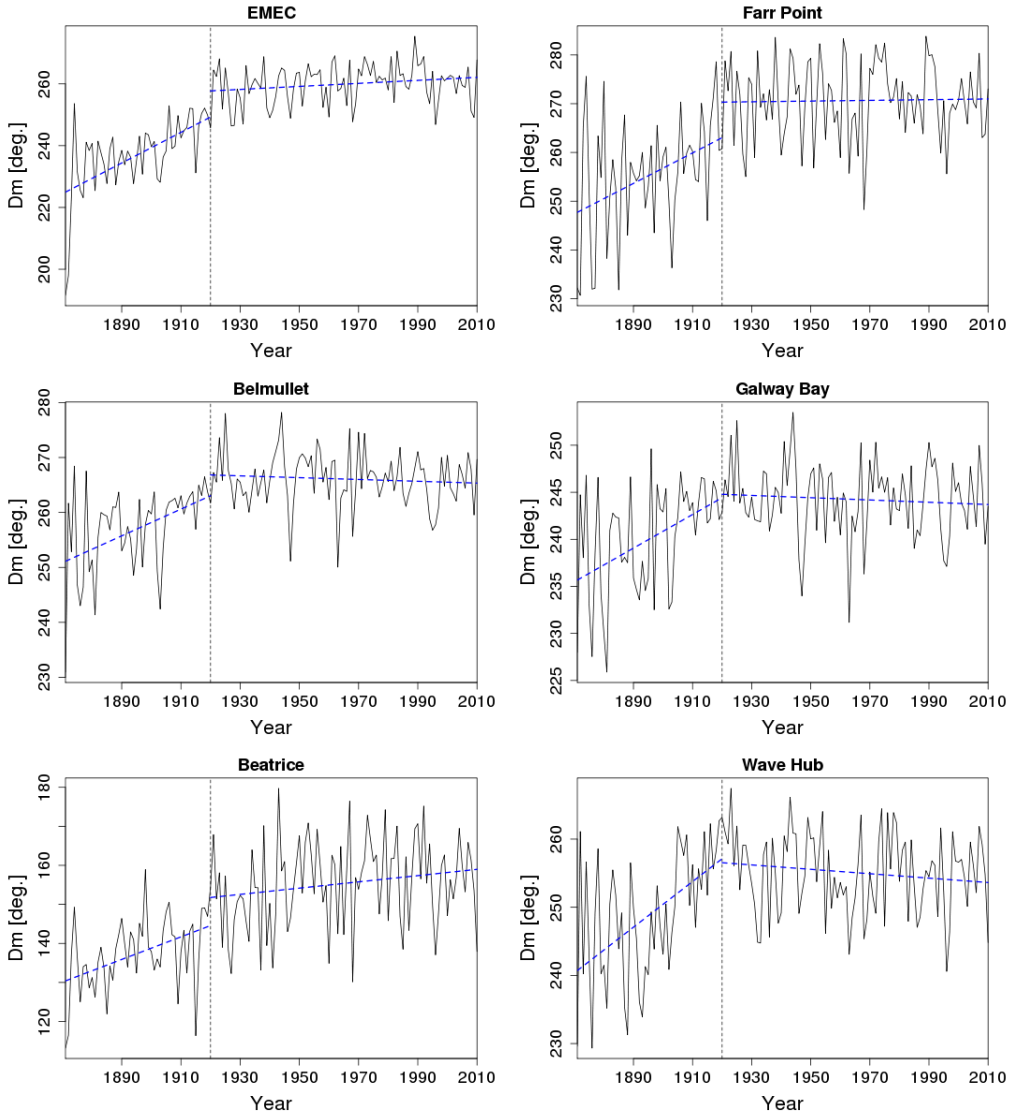

Figure B.3: 140 year time-series of annual mean wave direction for selected locations with the hinge fit trend superimposed (data from WAVEWATCH III)

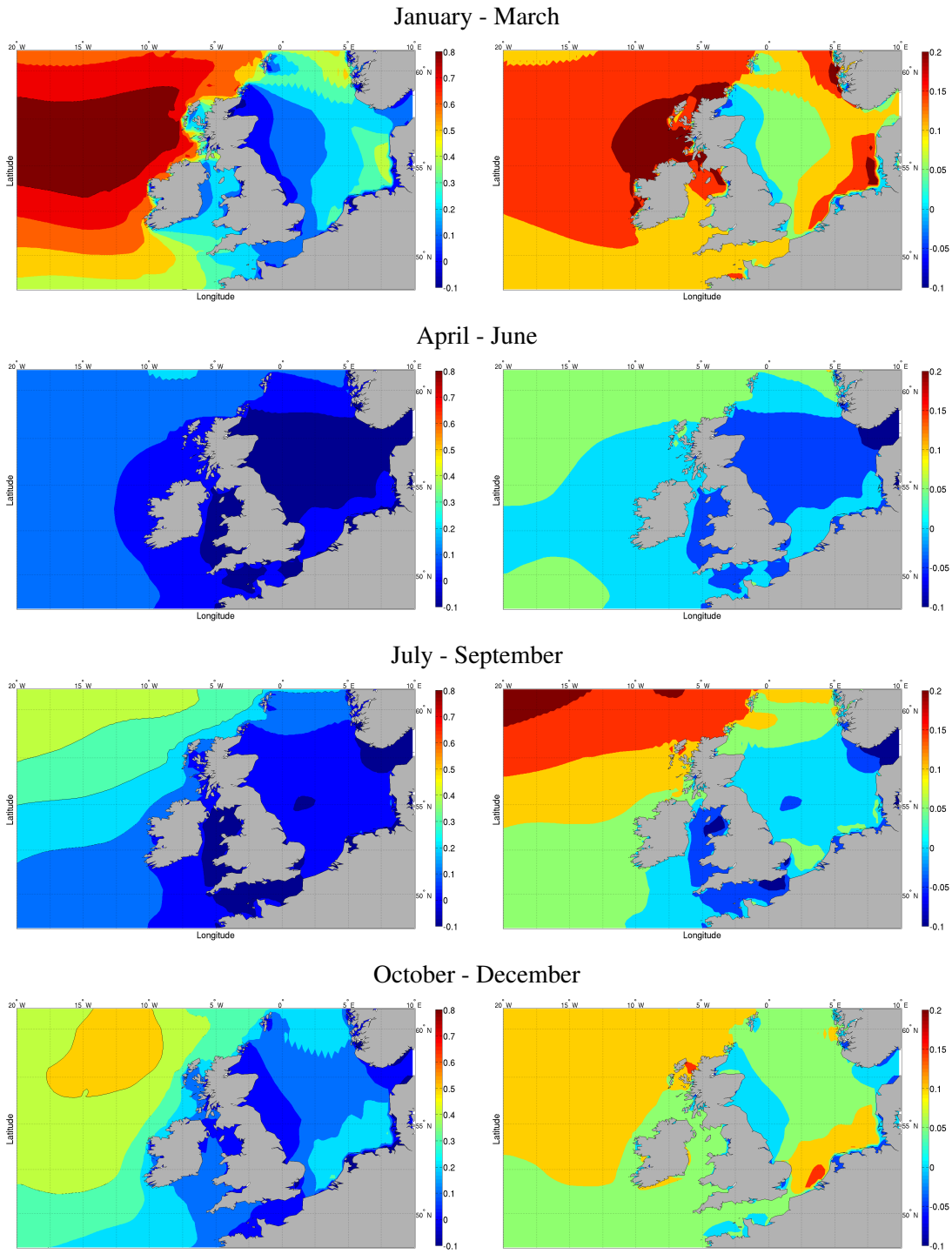

Figure B.4: Changes in the seasonal mean significant wave height computed between 1921 and 2010. Left - absolute change (cm per year); right - relative change (% per year).

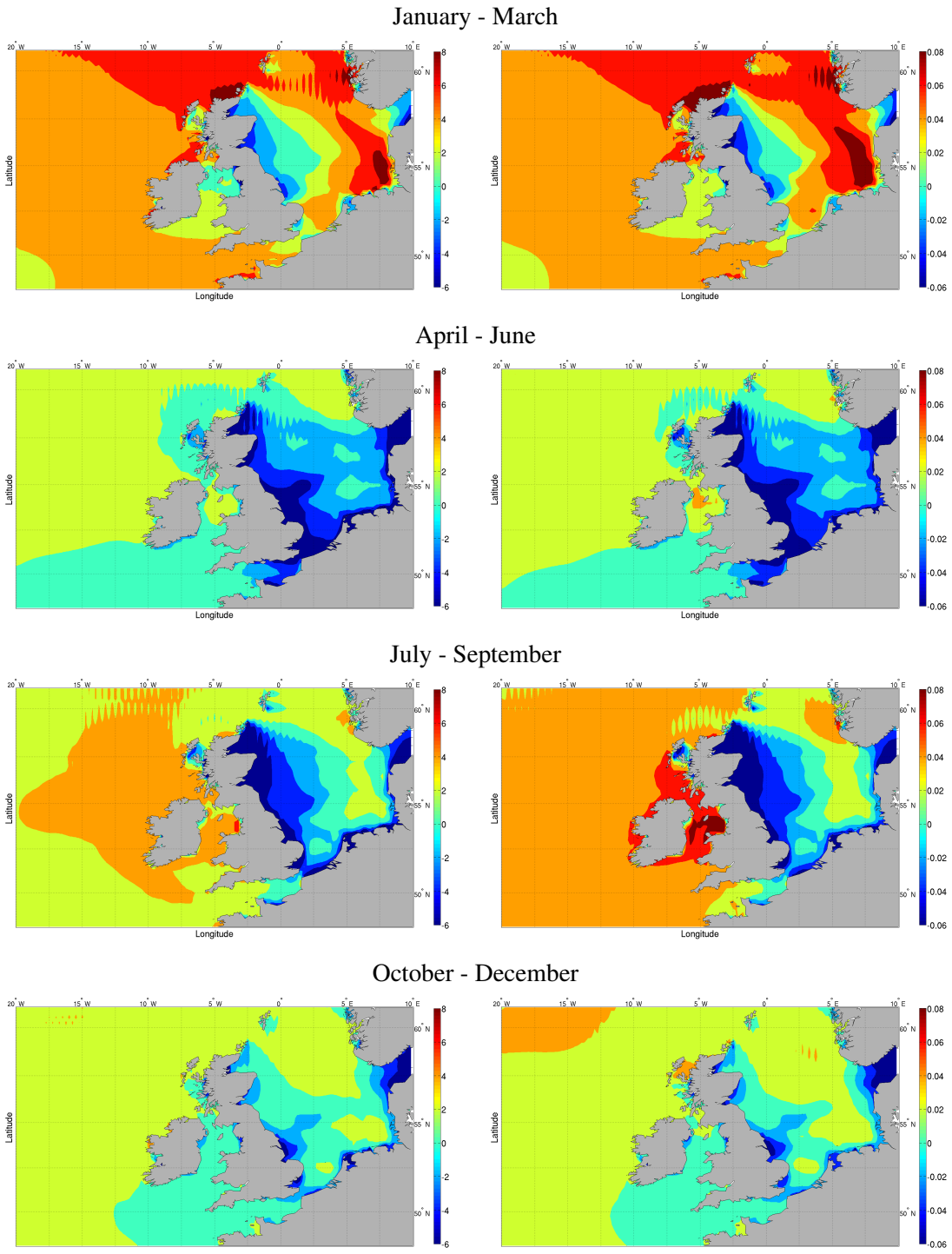

Figure B.5: Changes in the seasonal mean energy period computed between 1921 and 2010. Left - absolute change (x10^{-3}s per year); right - relative change (% per year).

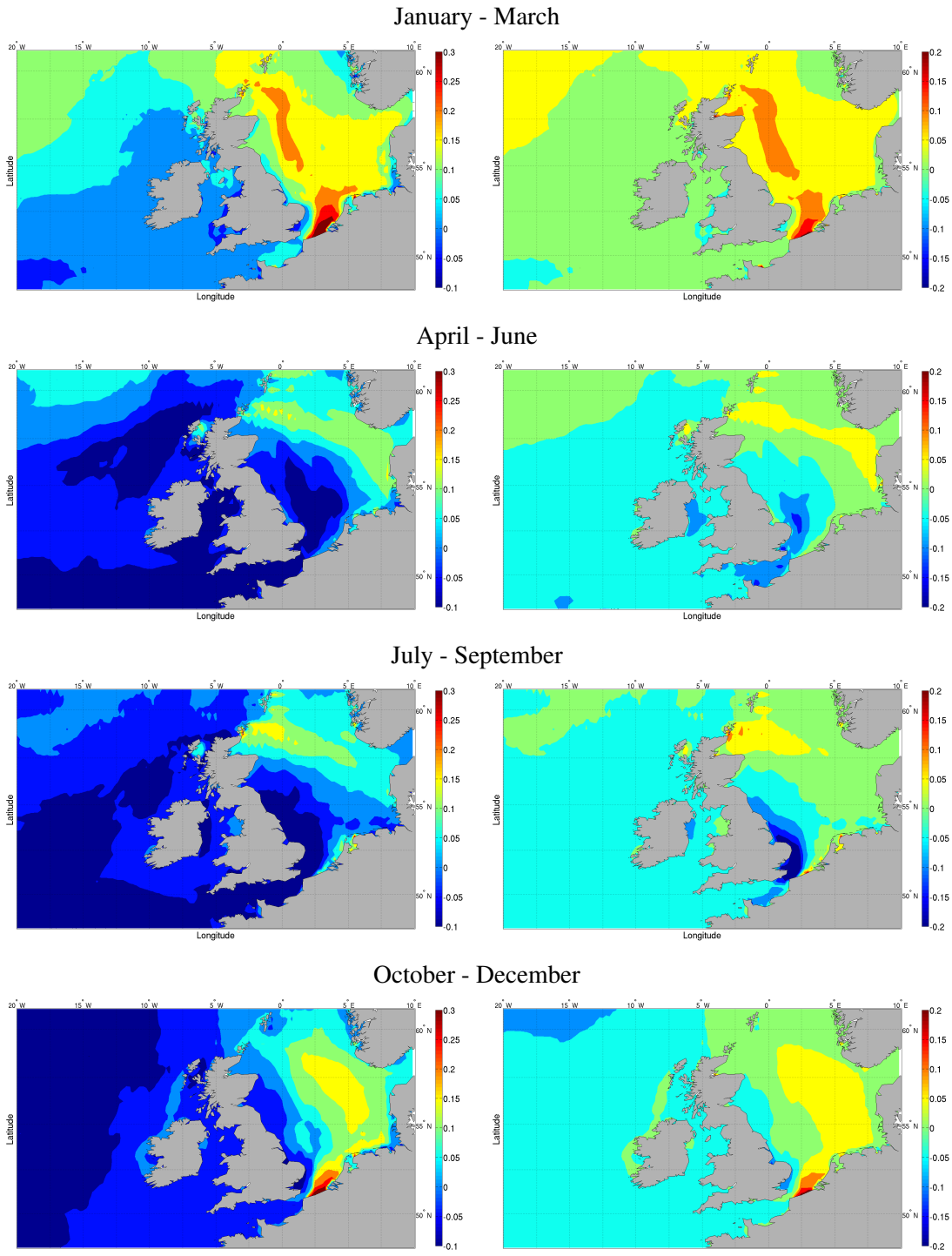

Figure B.6: Changes in the seasonal mean wave direction computed between 1921 and 2010. Left - absolute change (deg. per year); right - relative change (% per year).

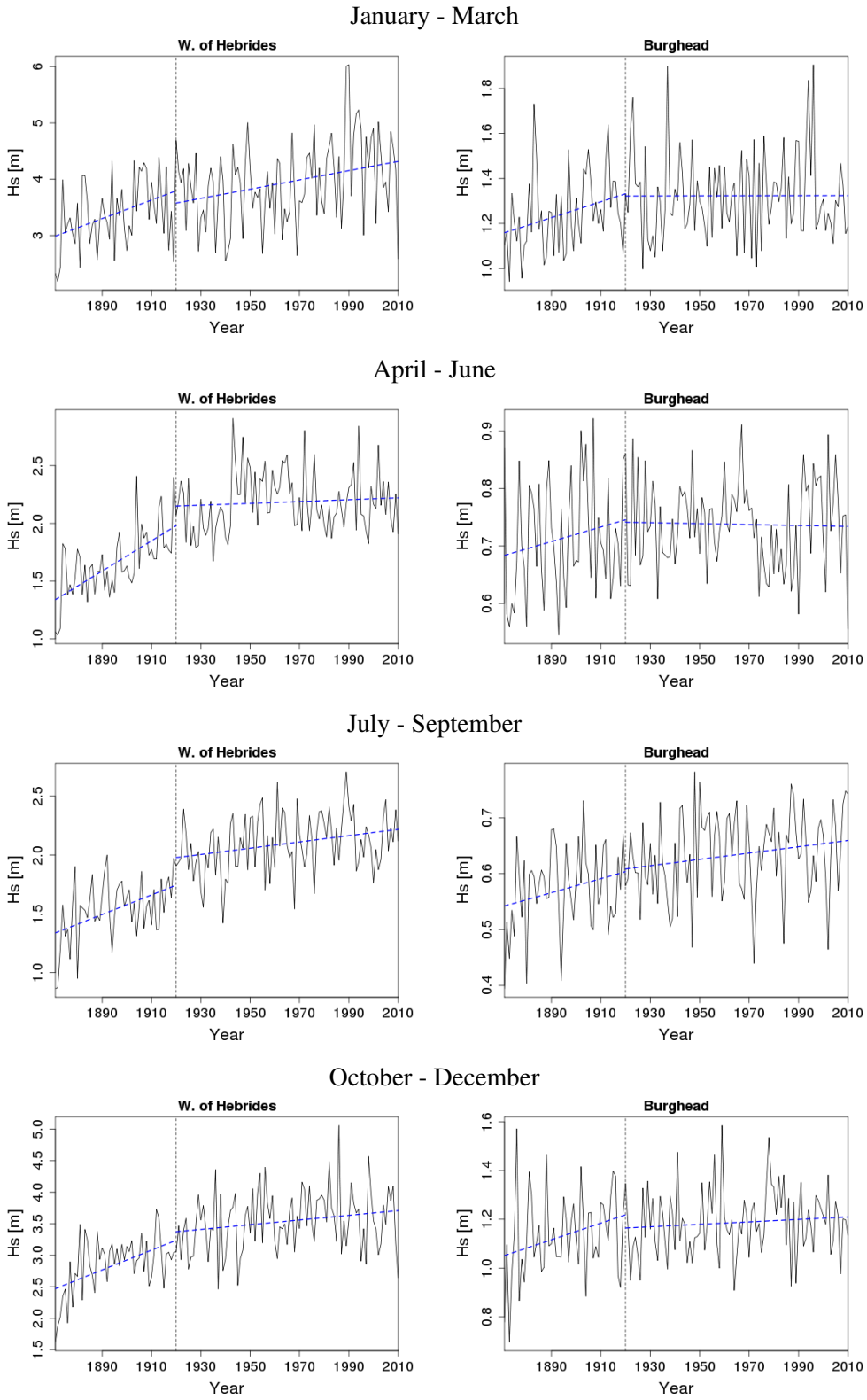

Figure B.7: 140 year time-series of seasonal mean significant wave height for West of Hebrides and Burghead with the hinge fit trend superimposed (data from WAVEWATCH III)

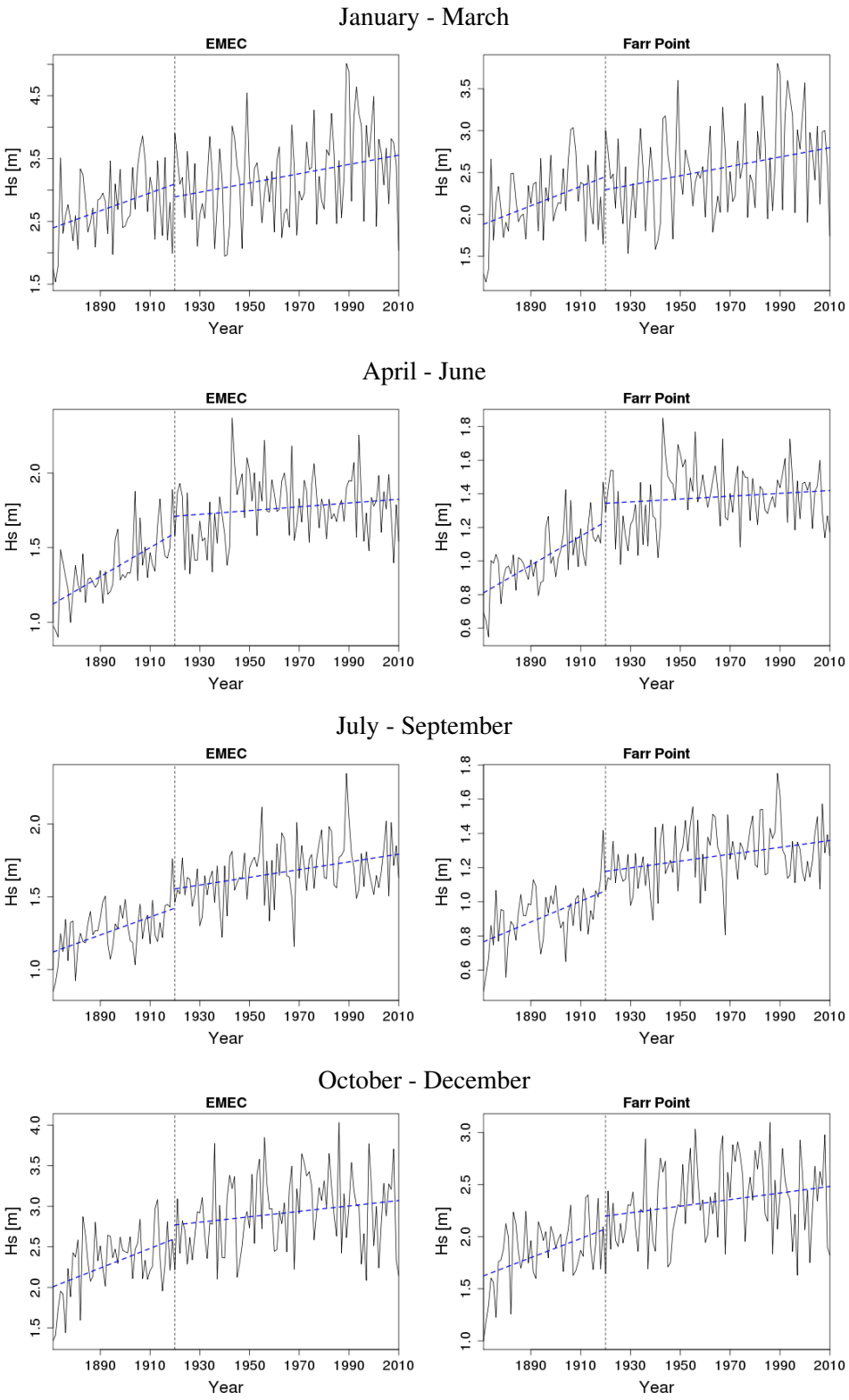

Figure B.8: 140 year time-series of seasonal mean significant wave height for EMEC and Farr Point with the hinge fit trend superimposed (data from WAVEWATCH III)

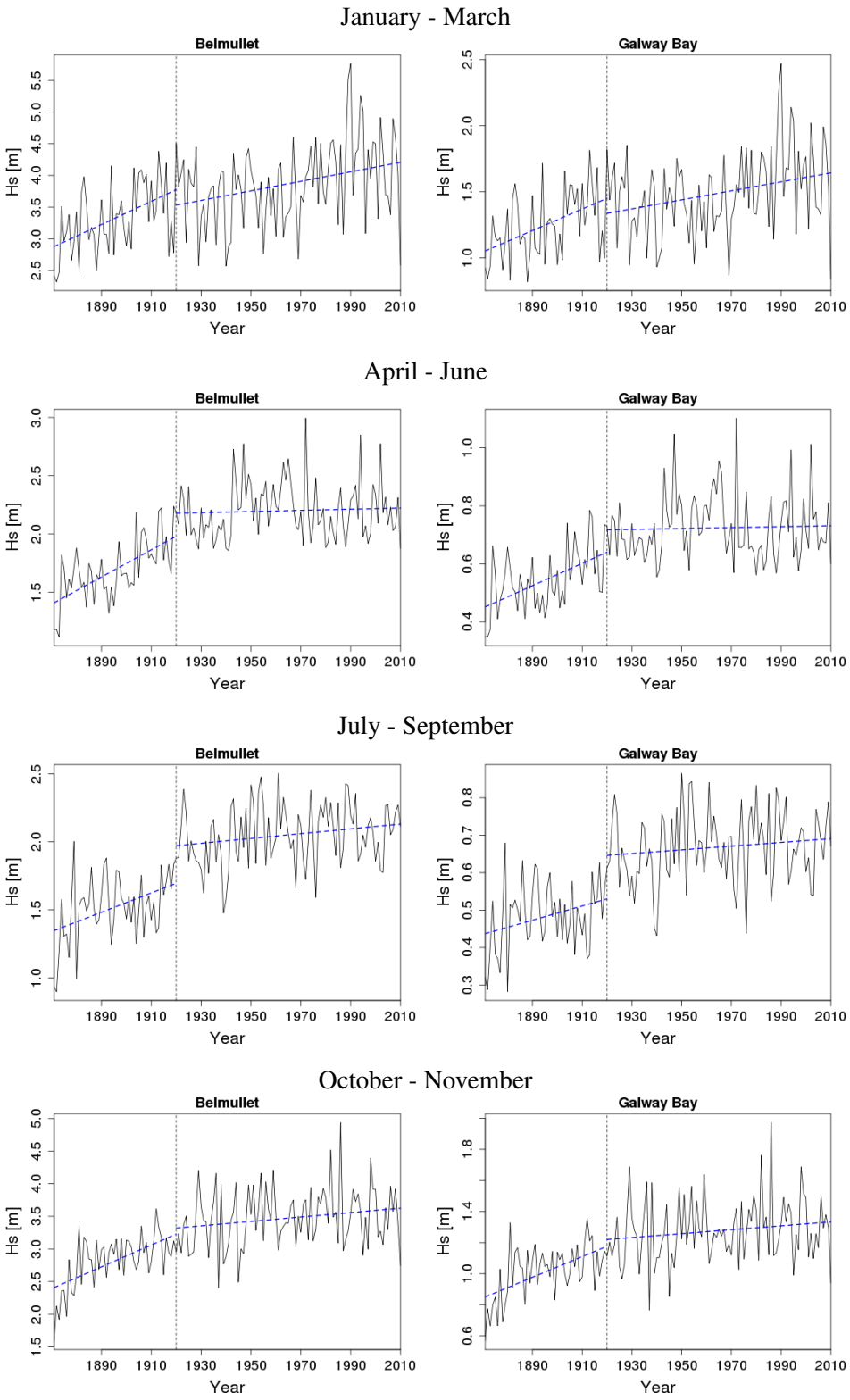

Figure B.9: 140 year time-series of seasonal mean significant wave height for Belmullet and Galway Bay with the hinge fit trend superimposed (data from WAVEWATCH III)

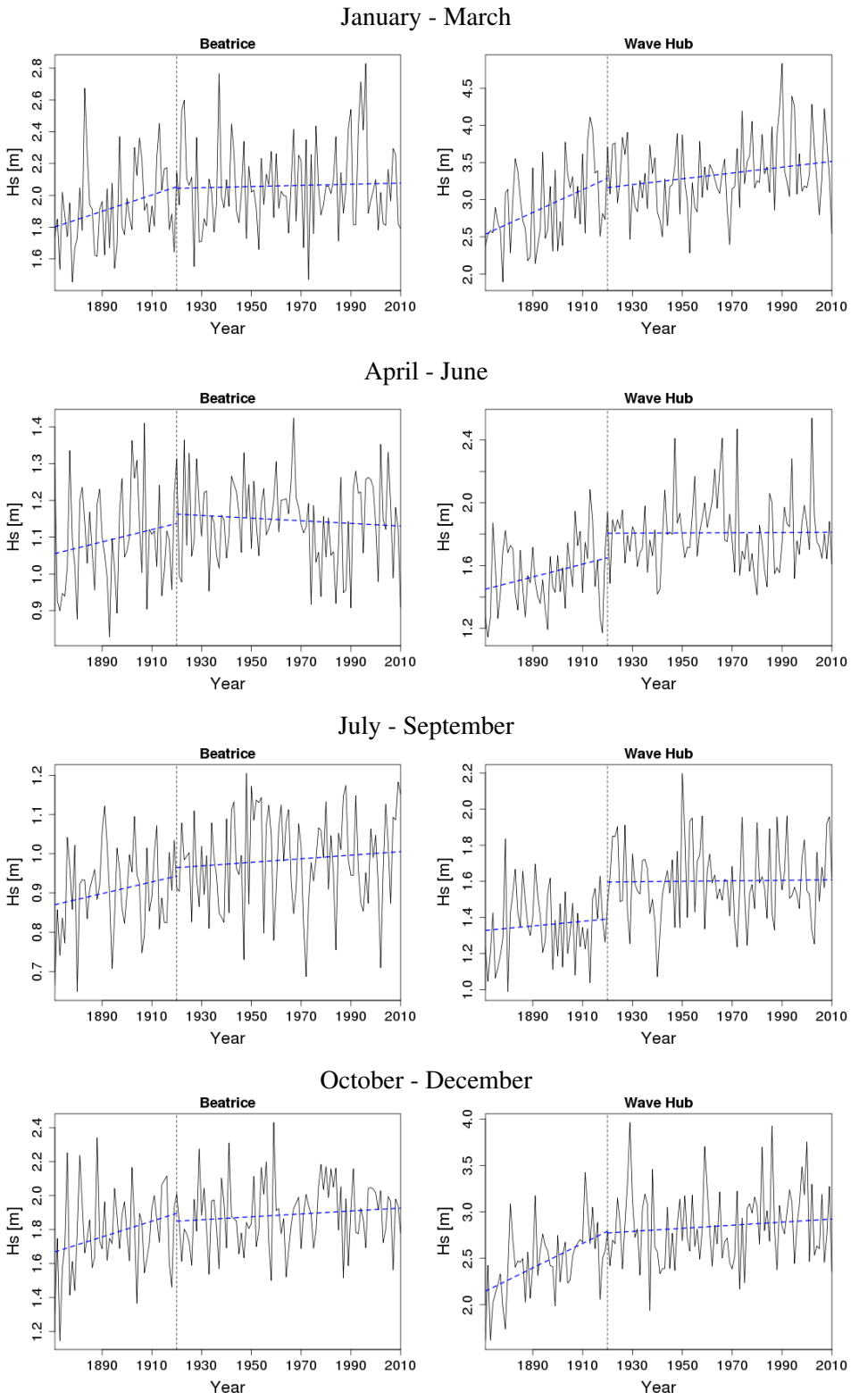

Figure B.10: 140 year time-series of seasonal mean significant wave height for Beatrice and Wave Hub with the hinge fit trend superimposed (data from WAVEWATCH III)

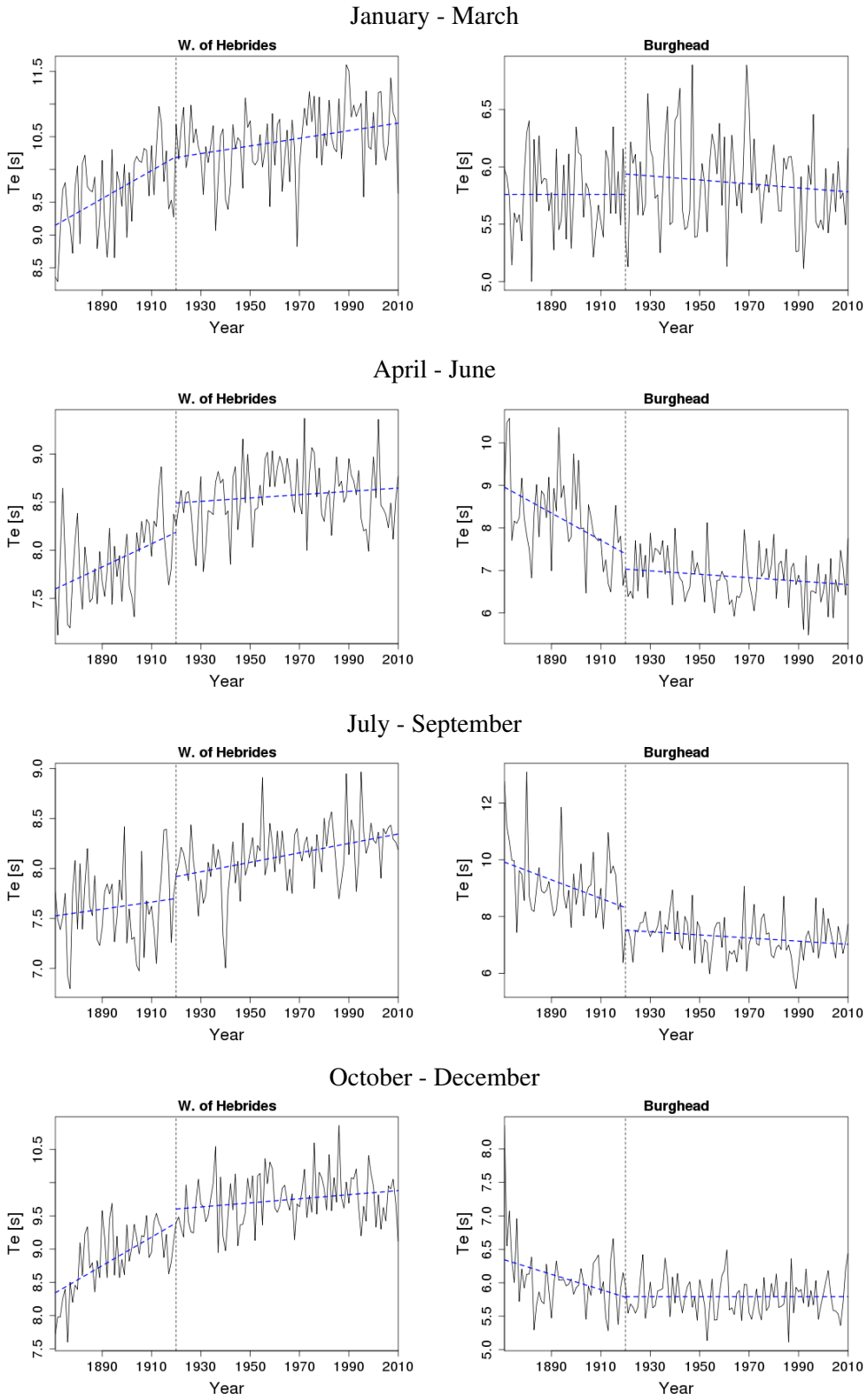

Figure B.11: 140 year time-series of seasonal mean energy period for West of Hebrides and Burghead with the hinge fit trend superimposed (data from WAVEWATCH III)

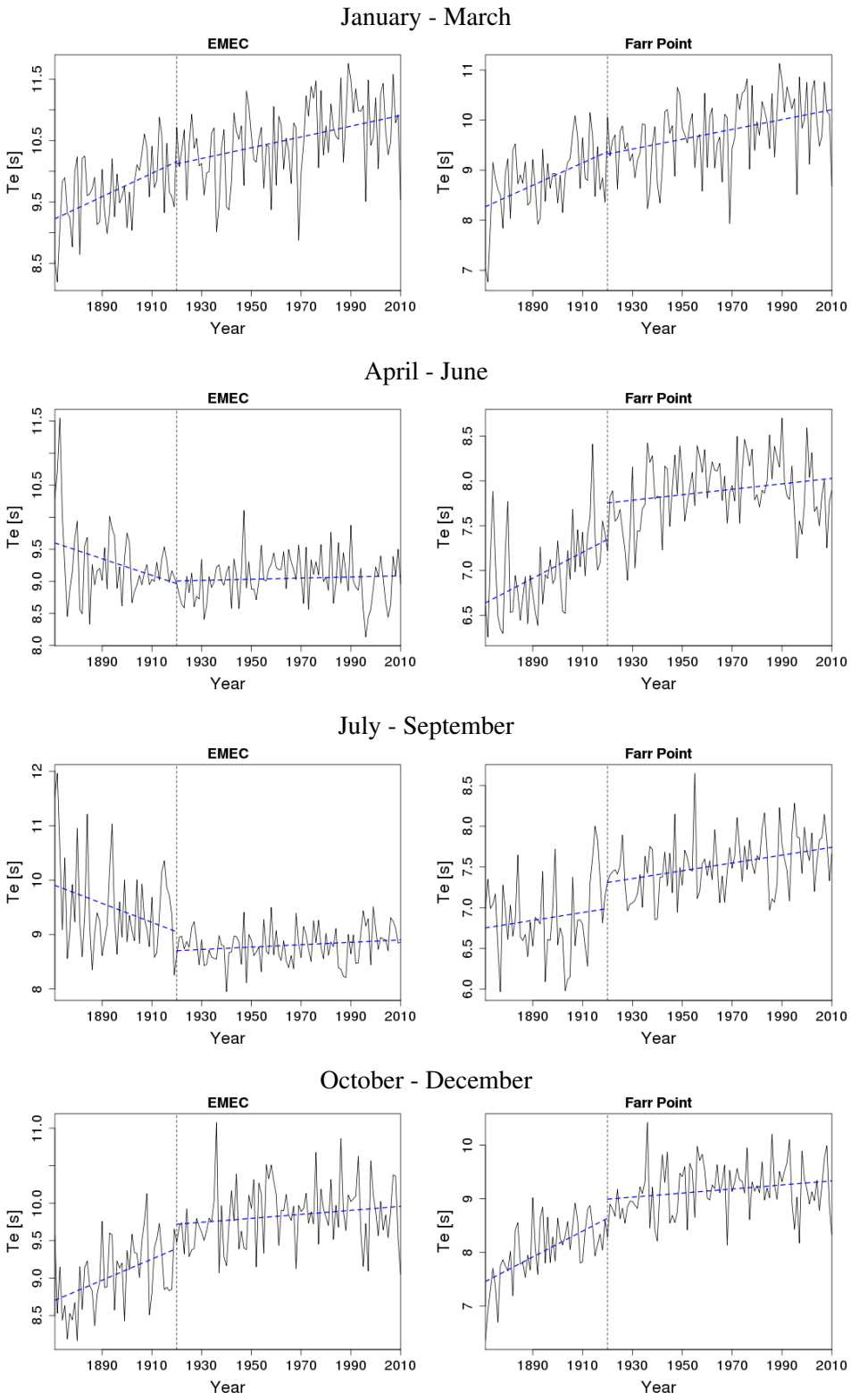

Figure B.12: 140 year time-series of seasonal mean energy period for EMEC and Farr Point with the hinge fit trend superimposed (data from WAVEWATCH III)

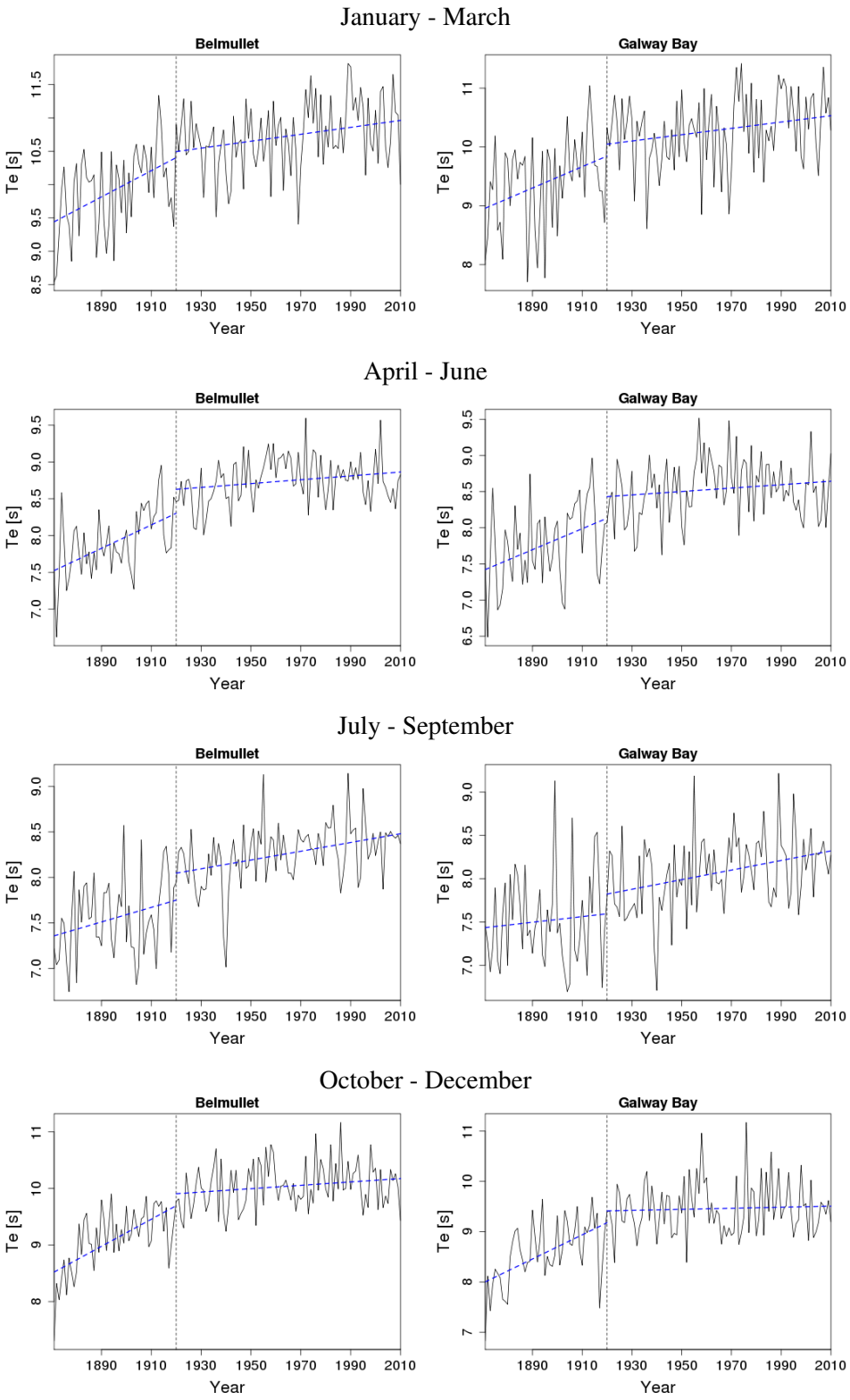

Figure B.13: 140 year time-series of seasonal mean energy period for Belmullet and Galway Bay with the hinge fit trend superimposed (data from WAVEWATCH III)

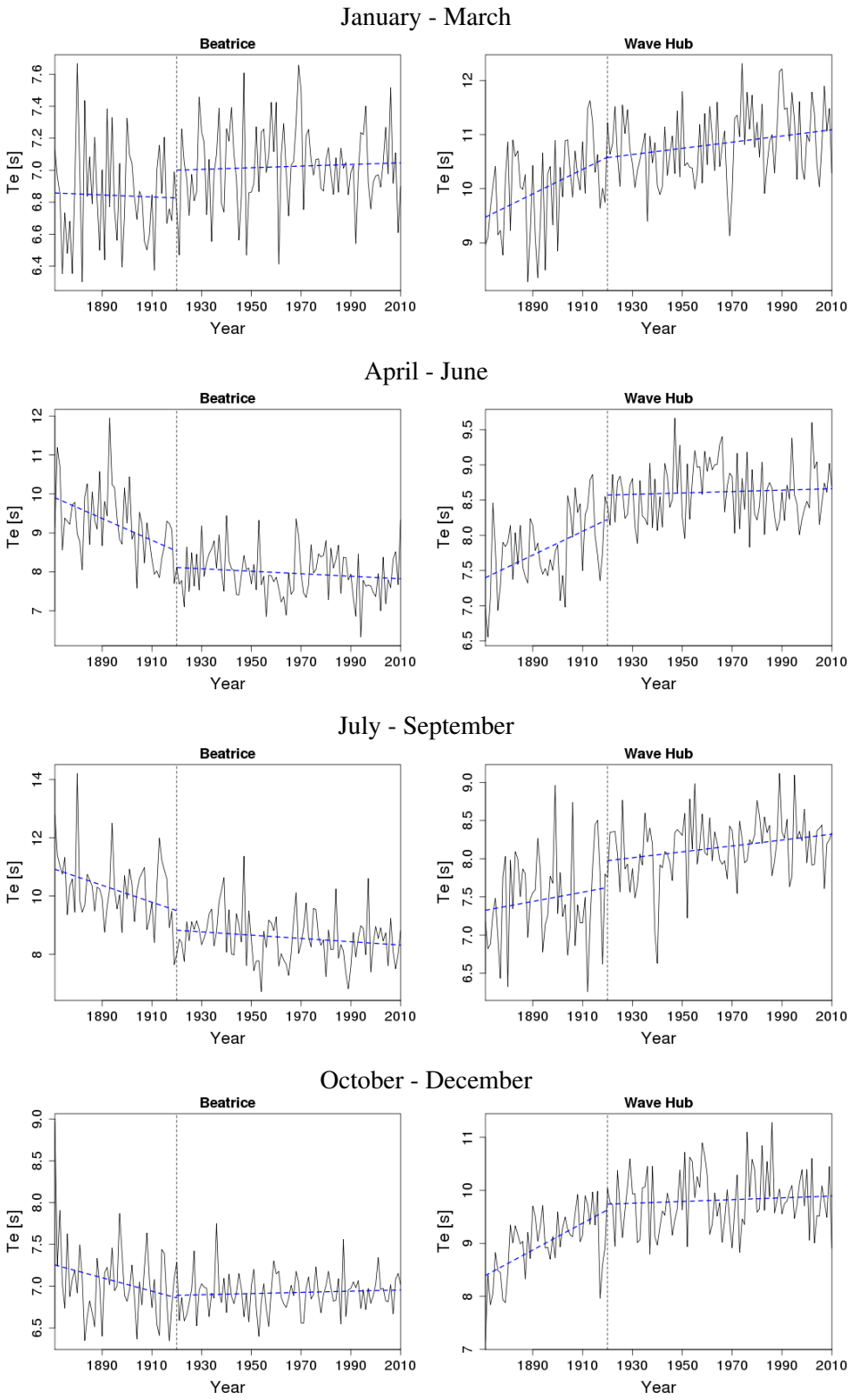

Figure B.14: 140 year time-series of seasonal mean energy period for Beatrice and Wave Hub with the hinge fit trend superimposed (data from WAVEWATCH III)

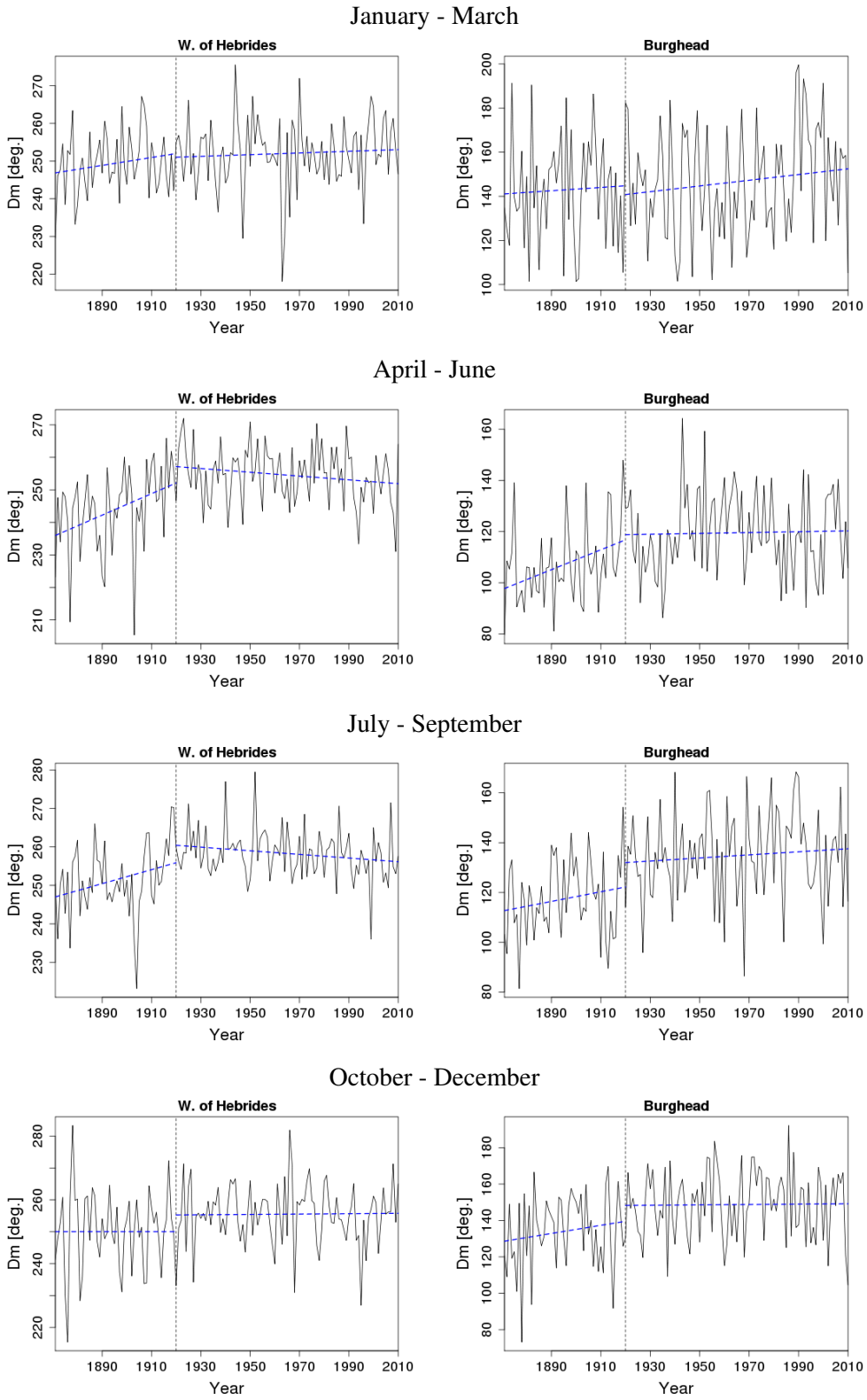

Figure B.15: 140 year time-series of seasonal mean wave direction for West of Hebrides and Burghead with the hinge fit trend superimposed (data from WAVEWATCH III)

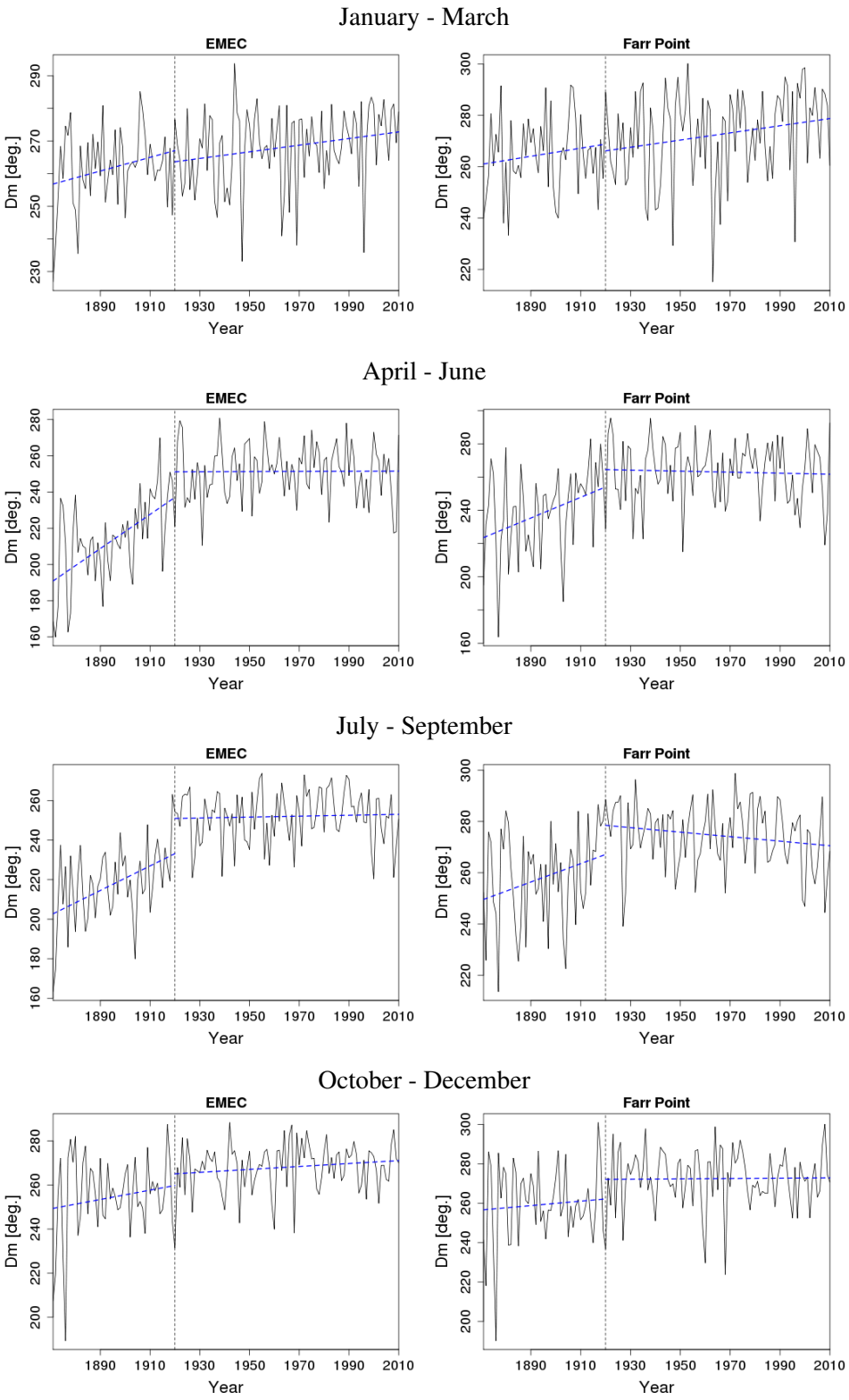

Figure B.16: 140 year time-series of seasonal mean wave direction for EMEC and Farr Point with the hinge fit trend superimposed (data from WAVEWATCH III)

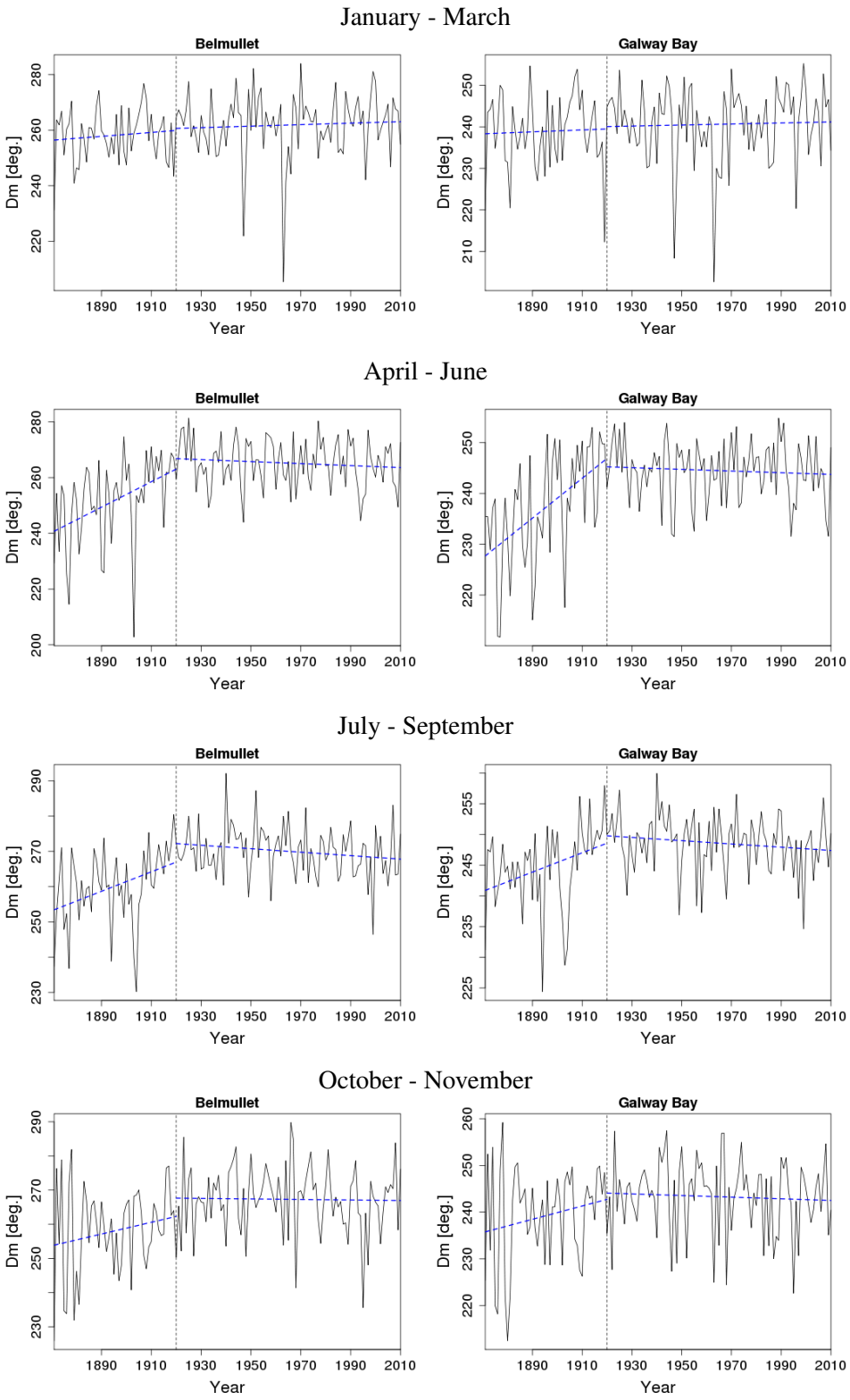

Figure B.17: 140 year time-series of seasonal mean wave direction for Belmullet and Galway Bay with the hinge fit trend superimposed (data from WAVEWATCH III)

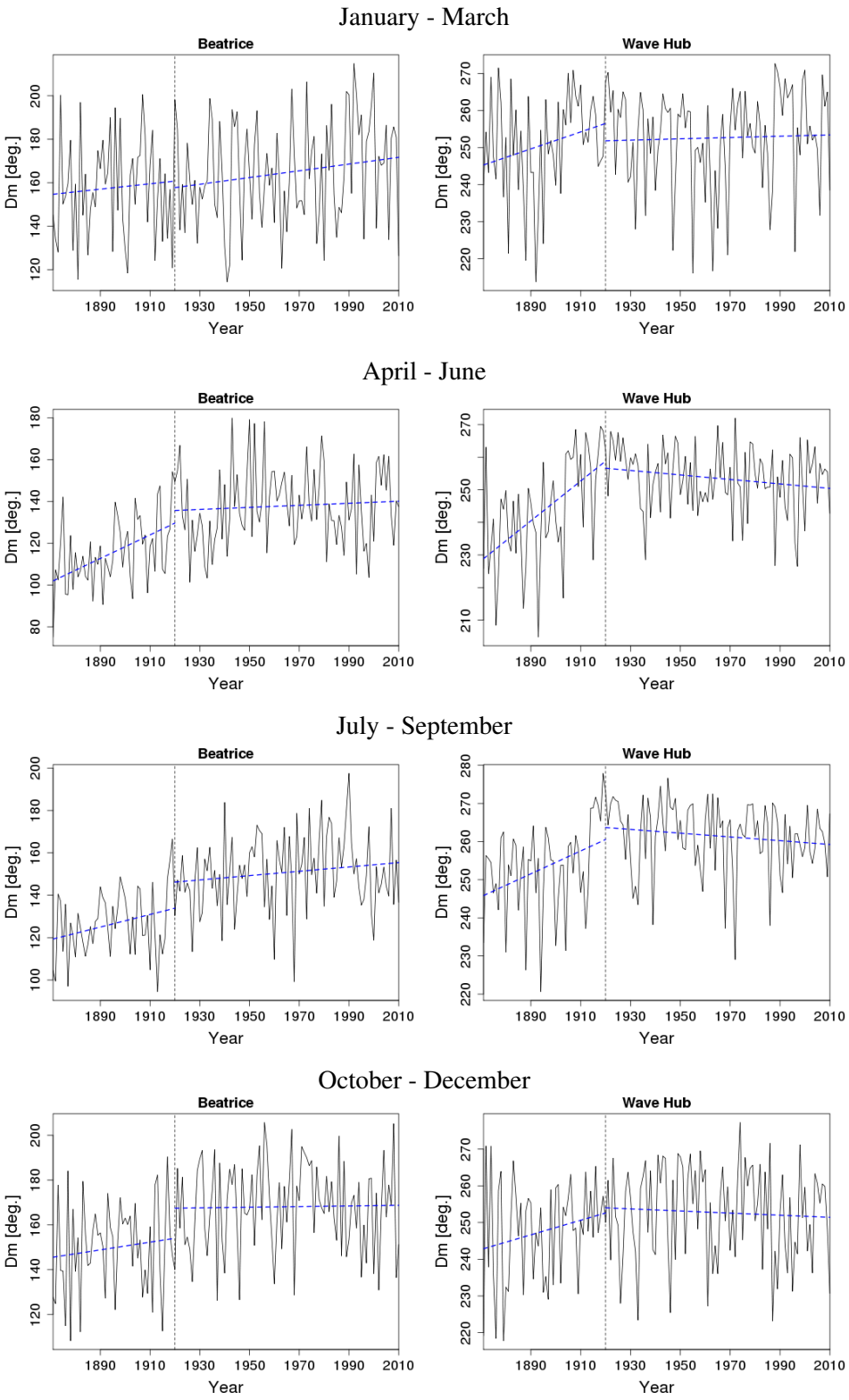

Figure B.18: 140 year time-series of seasonal mean wave direction for Beatrice and Wave Hub with the hinge fit trend superimposed (data from WAVEWATCH III)

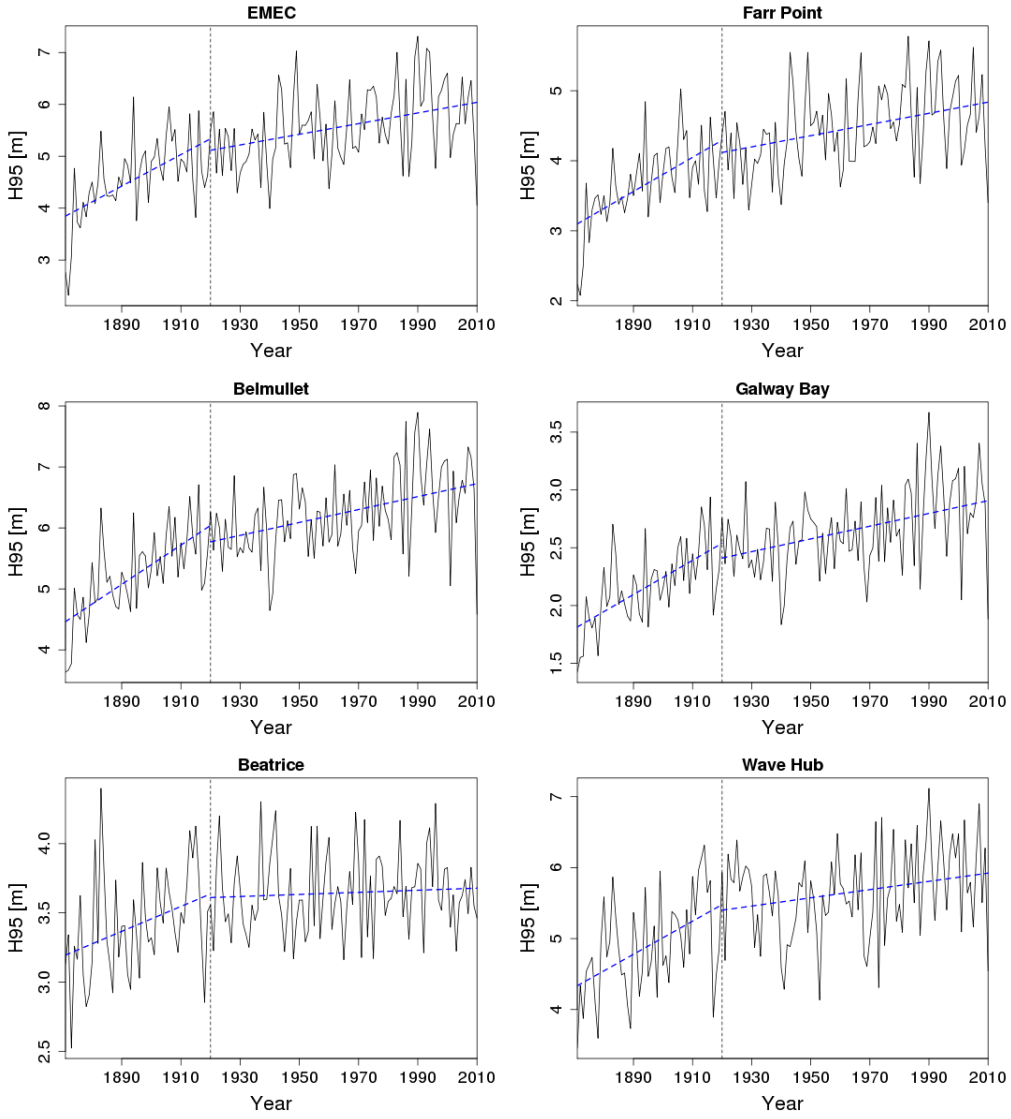

Figure B.19: 140 year time-series of annual 95th percentile significant wave height for selected locations with the hinge fit trend superimposed (data from WAVEWATCH III)

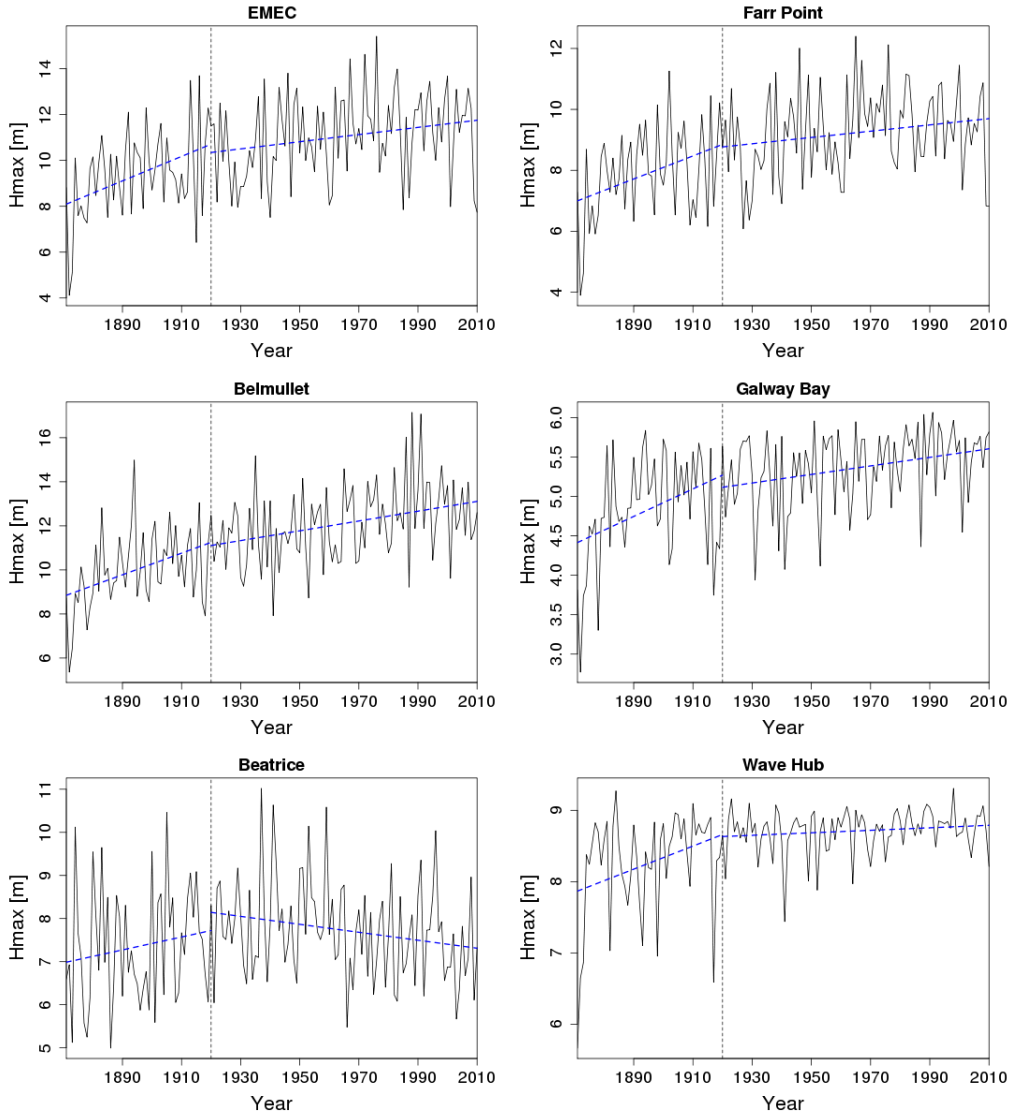

Figure B.20: 140 year time-series of annual maximum significant wave height for selected locations with the hinge fit trend superimposed (data from WAVEWATCH III)

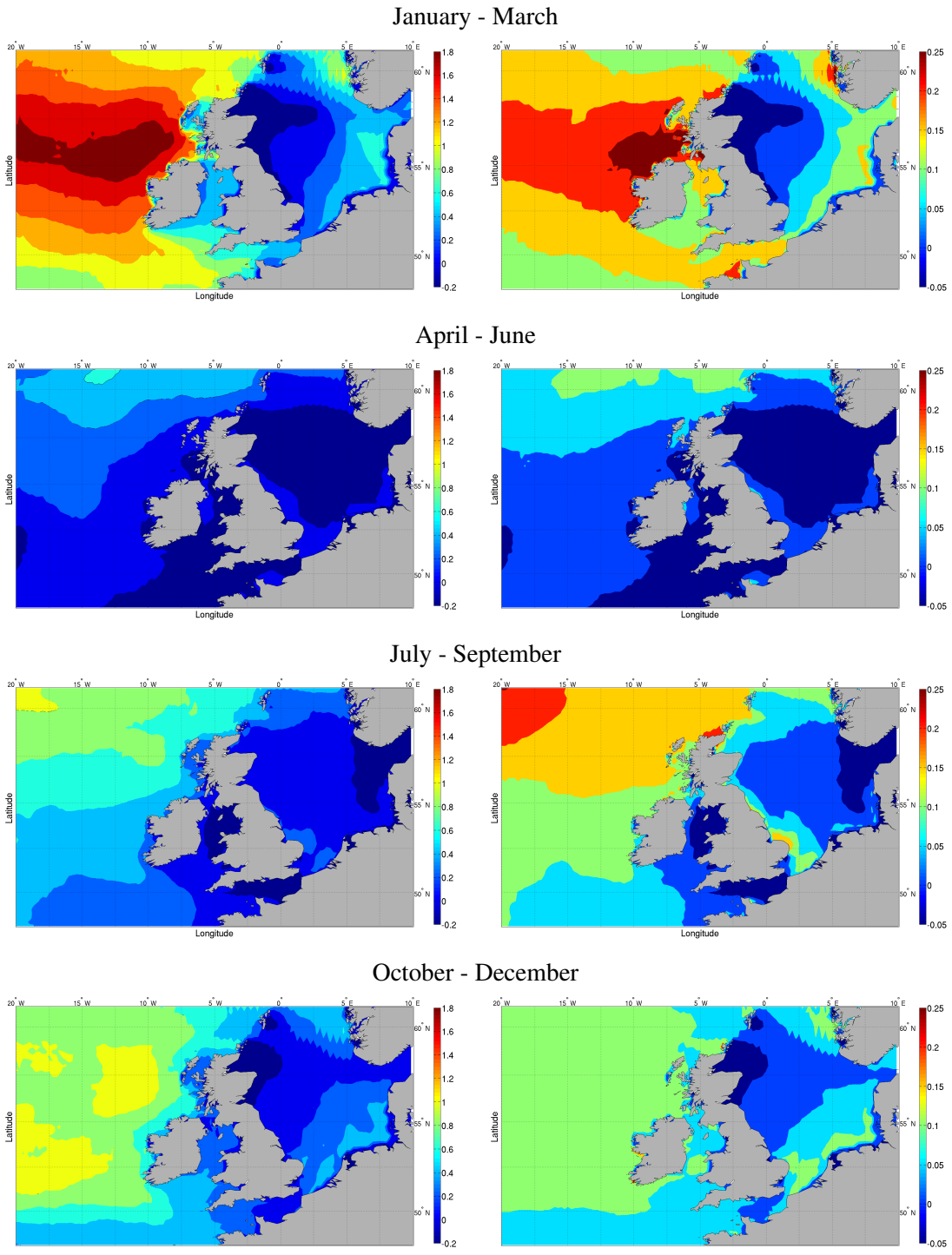

Figure B.21: Changes in the seasonal 95th percentile significant wave height computed between 1921 and 2010. Left - absolute change (cm per year); right - relative change (% per year).

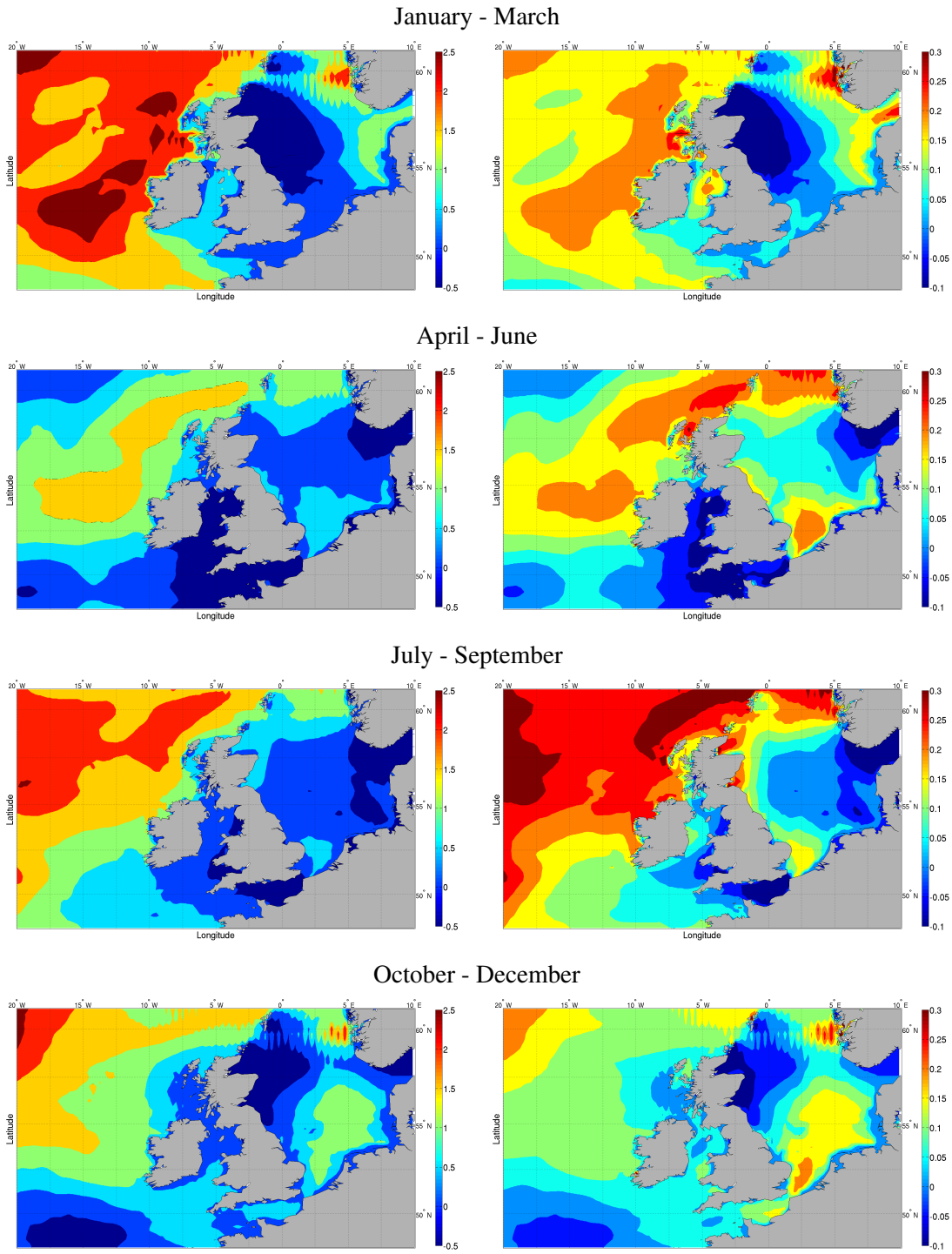

Figure B.22: Changes in the seasonal maximum significant wave height computed between 1921 and 2010. Left - absolute change (cm per year); right - relative change (% per year).

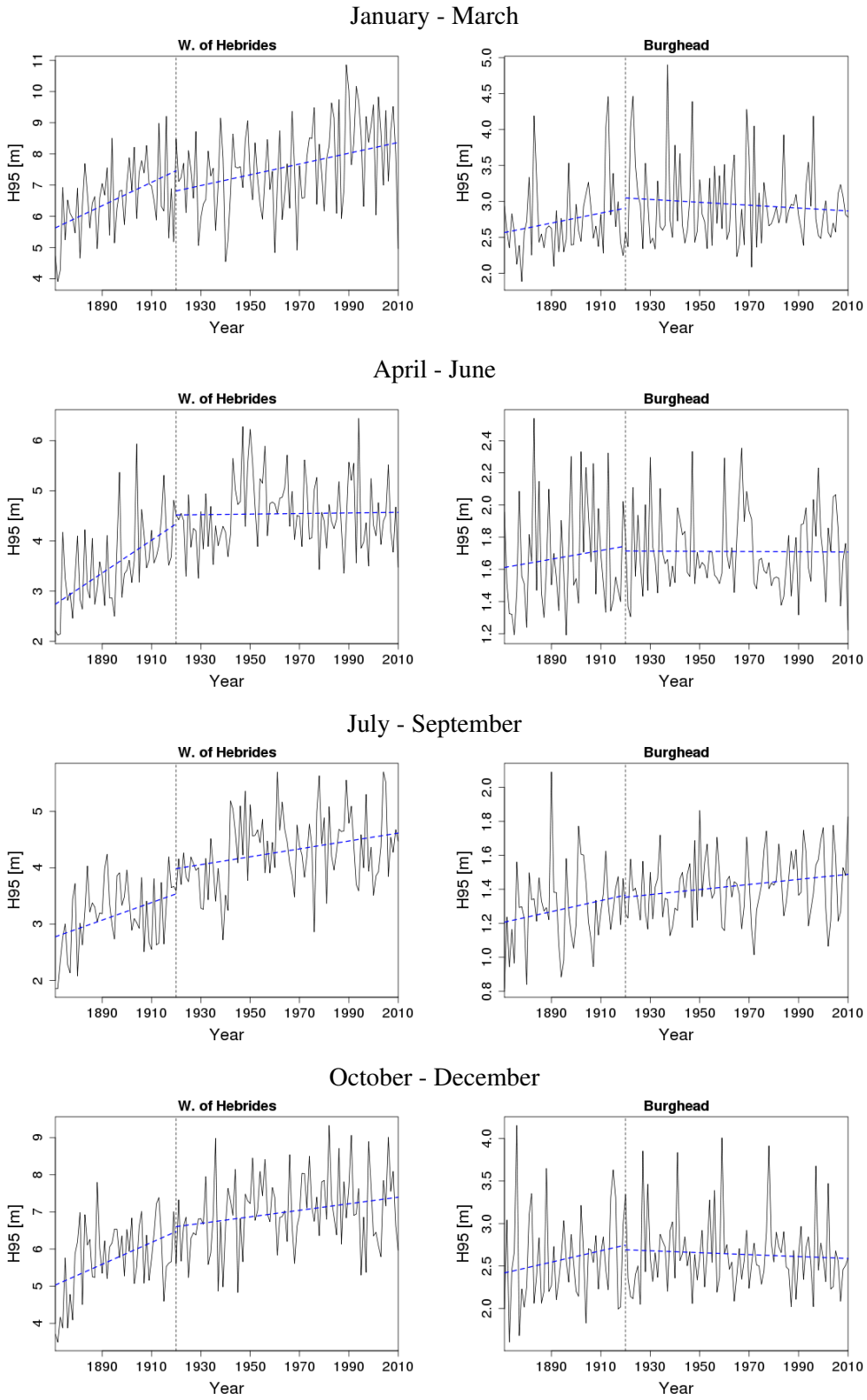

Figure B.23: 140 year time-series of seasonal 95th percentile significant wave height for West of Hebrides and Burghead with the hinge fit trend superimposed (data from WAVEWATCH III)

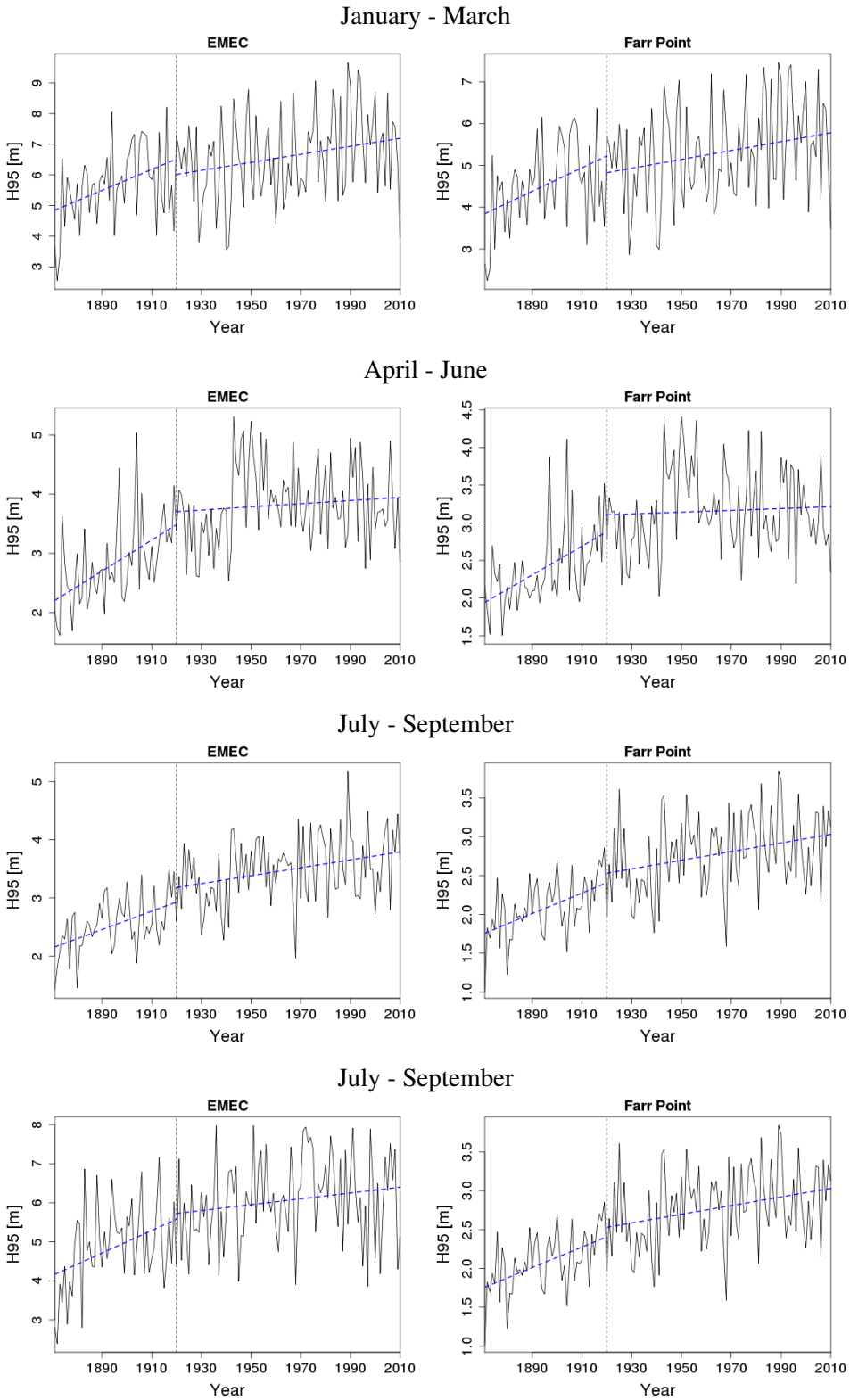

Figure B.24: 140 year time-series of seasonal 95th percentile significant wave height for EMEC and Farr Point with the hinge fit trend superimposed (data from WAVEWATCH III)

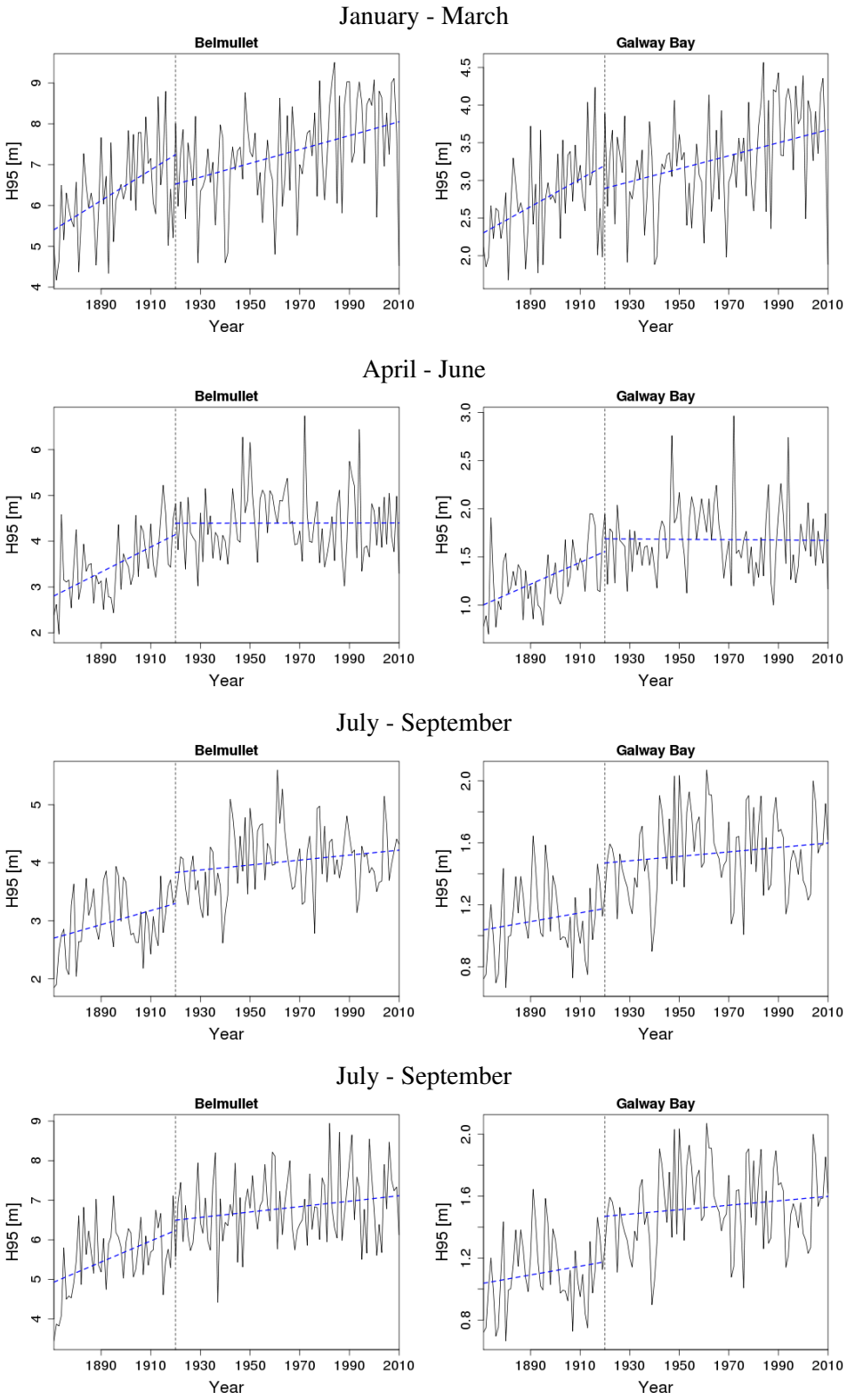

Figure B.25: 140 year time-series of seasonal 95th percentile significant wave height for Belmullet and Galway Bay with the hinge fit trend superimposed (data from WAVEWATCH III)

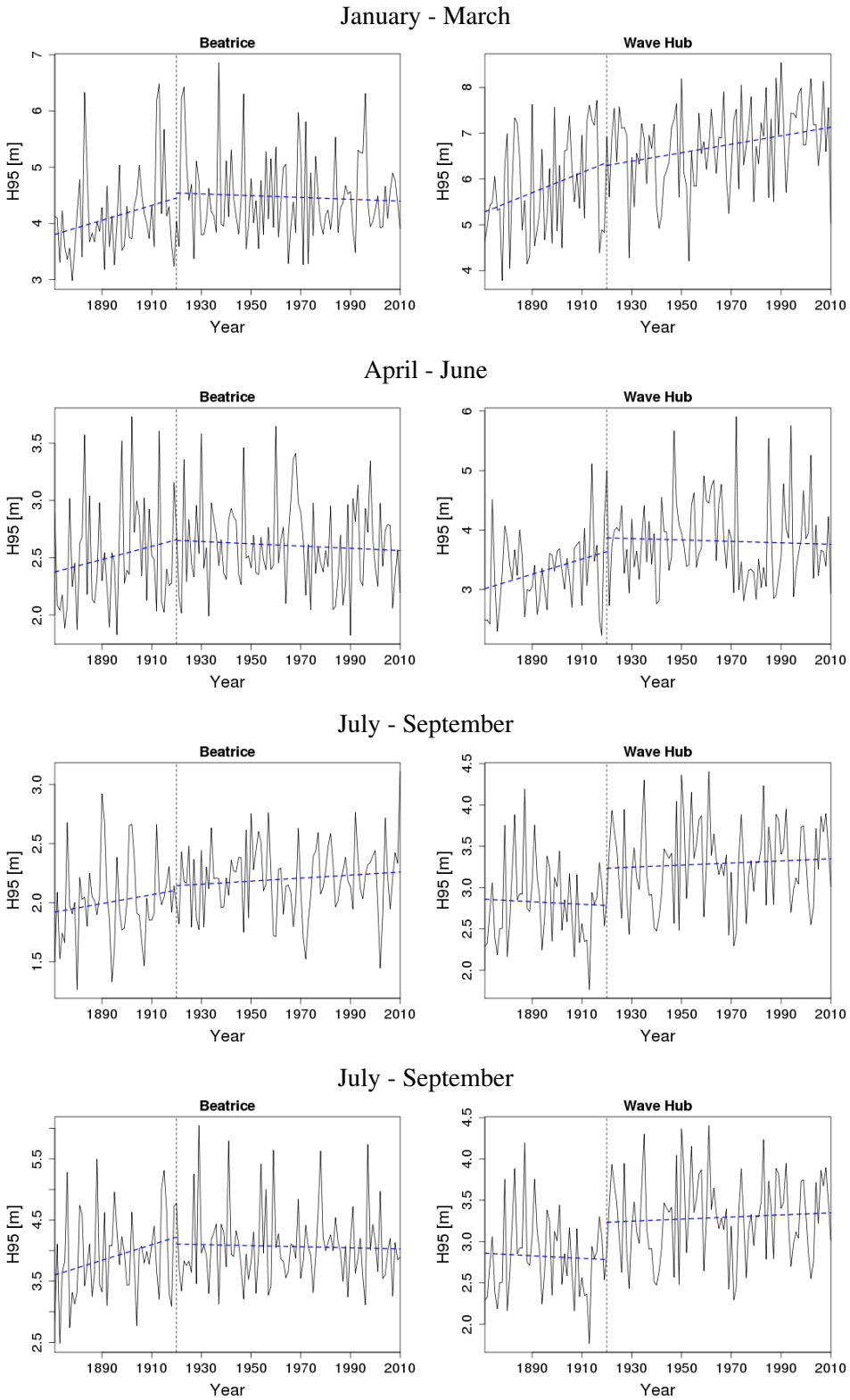

Figure B.26: 140 year time-series of seasonal 95th percentile significant wave height for Beatrice and Wave Hub with the hinge fit trend superimposed (data from WAVEWATCH III)

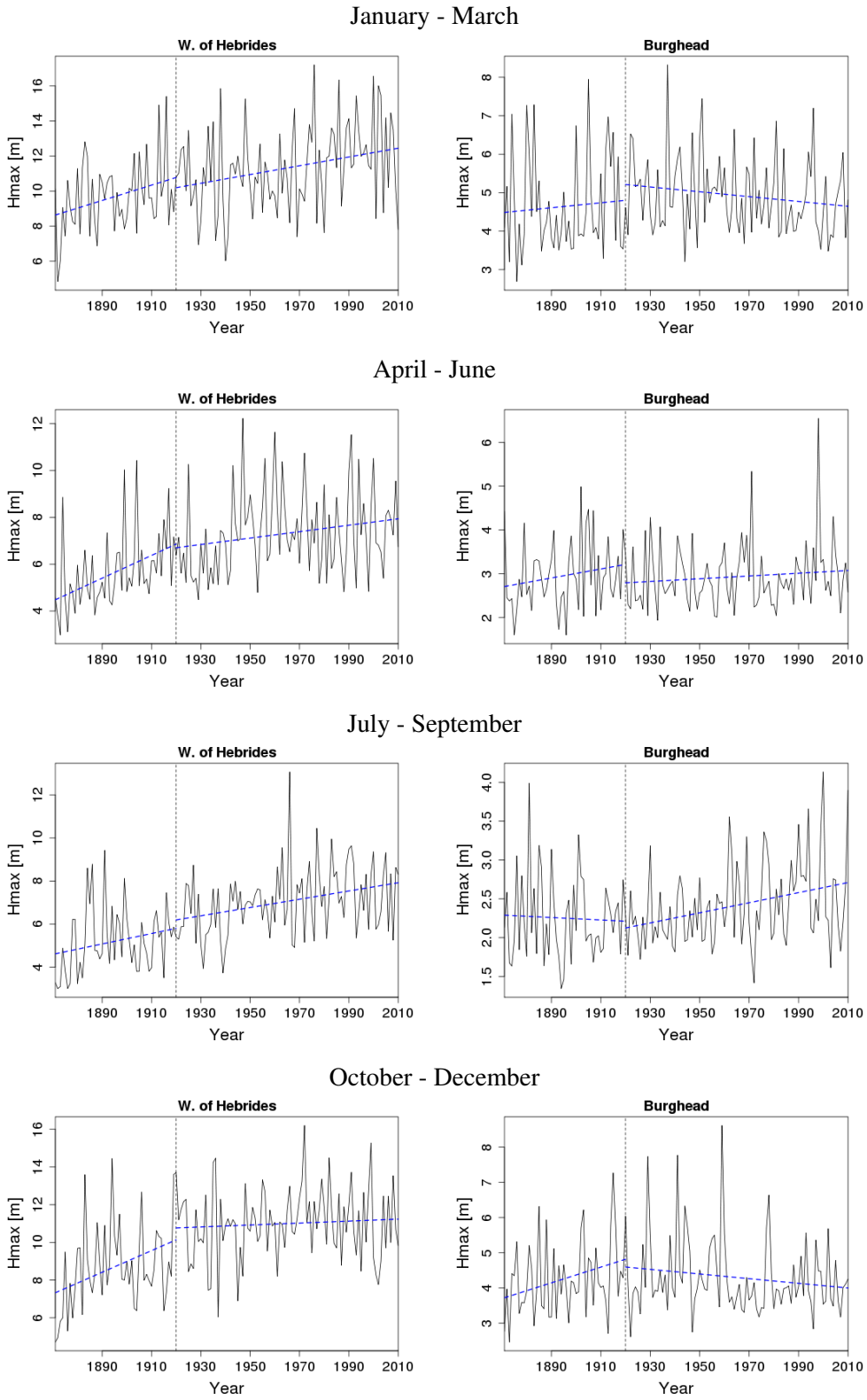

Figure B.27: 140 year time-series of seasonal maxima significant wave height for West of Hebrides and Burghead with the hinge fit trend superimposed (data from WAVEWATCH III)

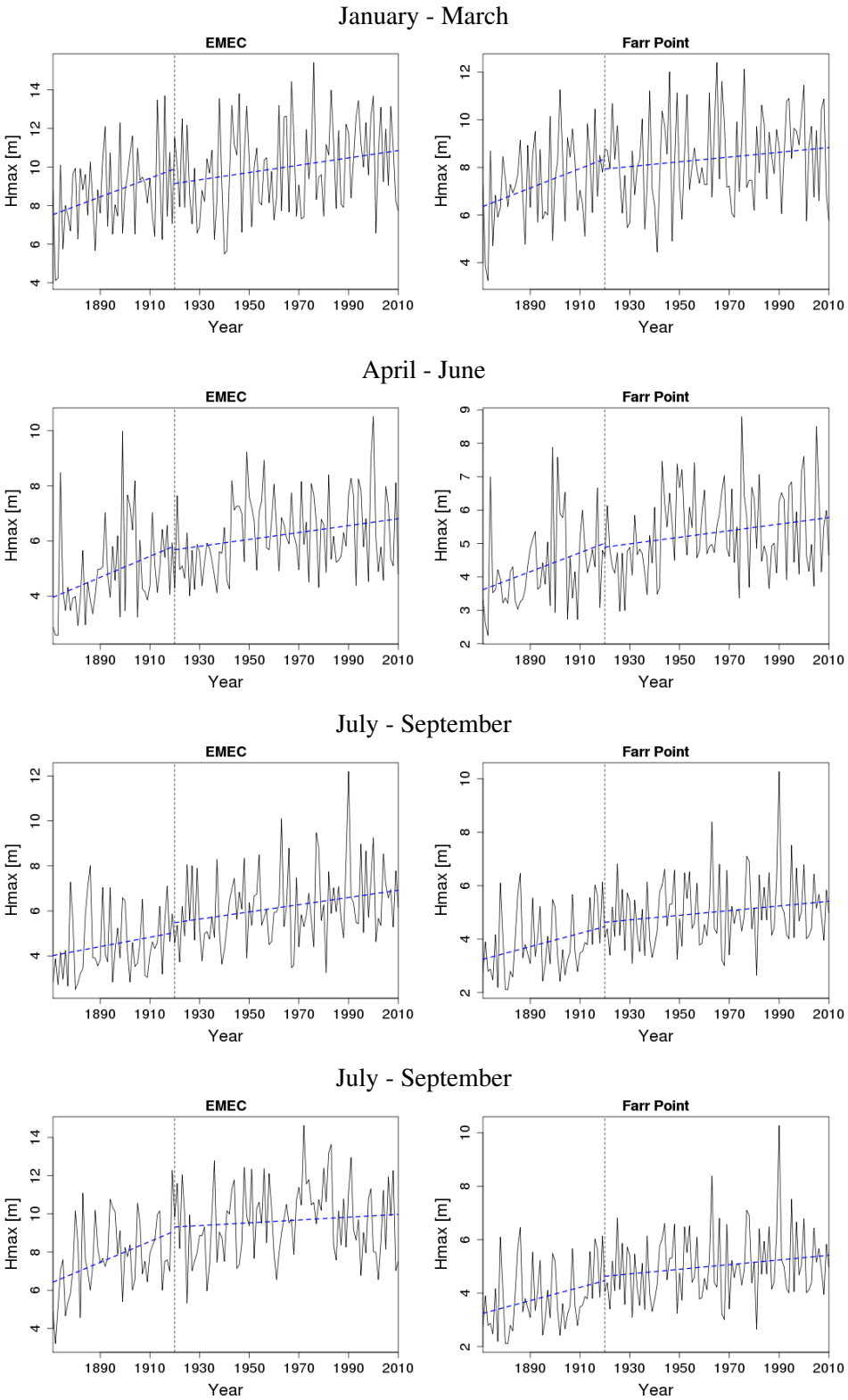

Figure B.28: 140 year time-series of seasonal maxima of significant wave height for EMEC and Farr Point with the hinge fit trend superimposed (data from WAVEWATCH III)

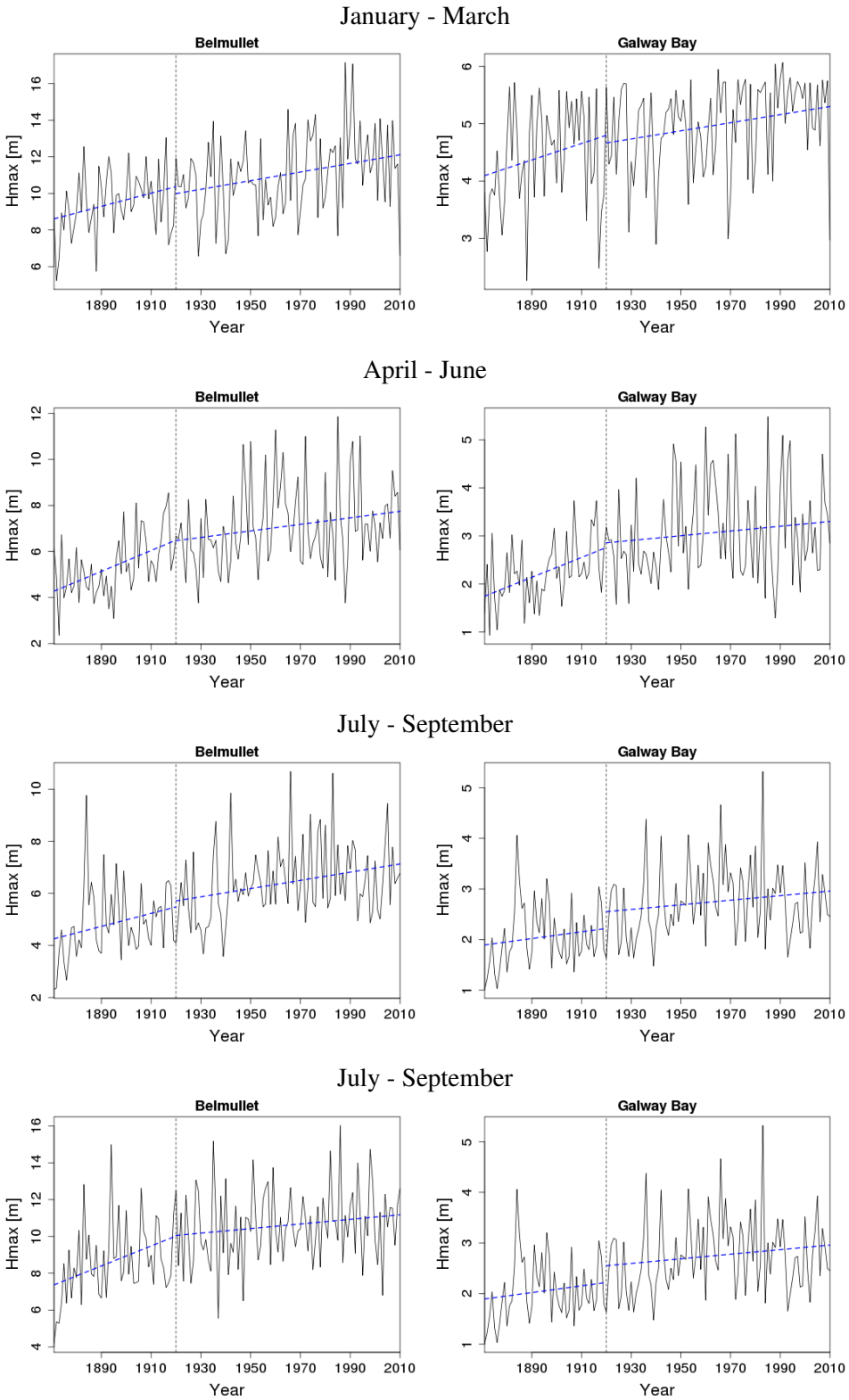

Figure B.29: 140 year time-series of seasonal maxima significant wave height for Belmullet and Galway Bay with the hinge fit trend superimposed (data from WAVEWATCH III)

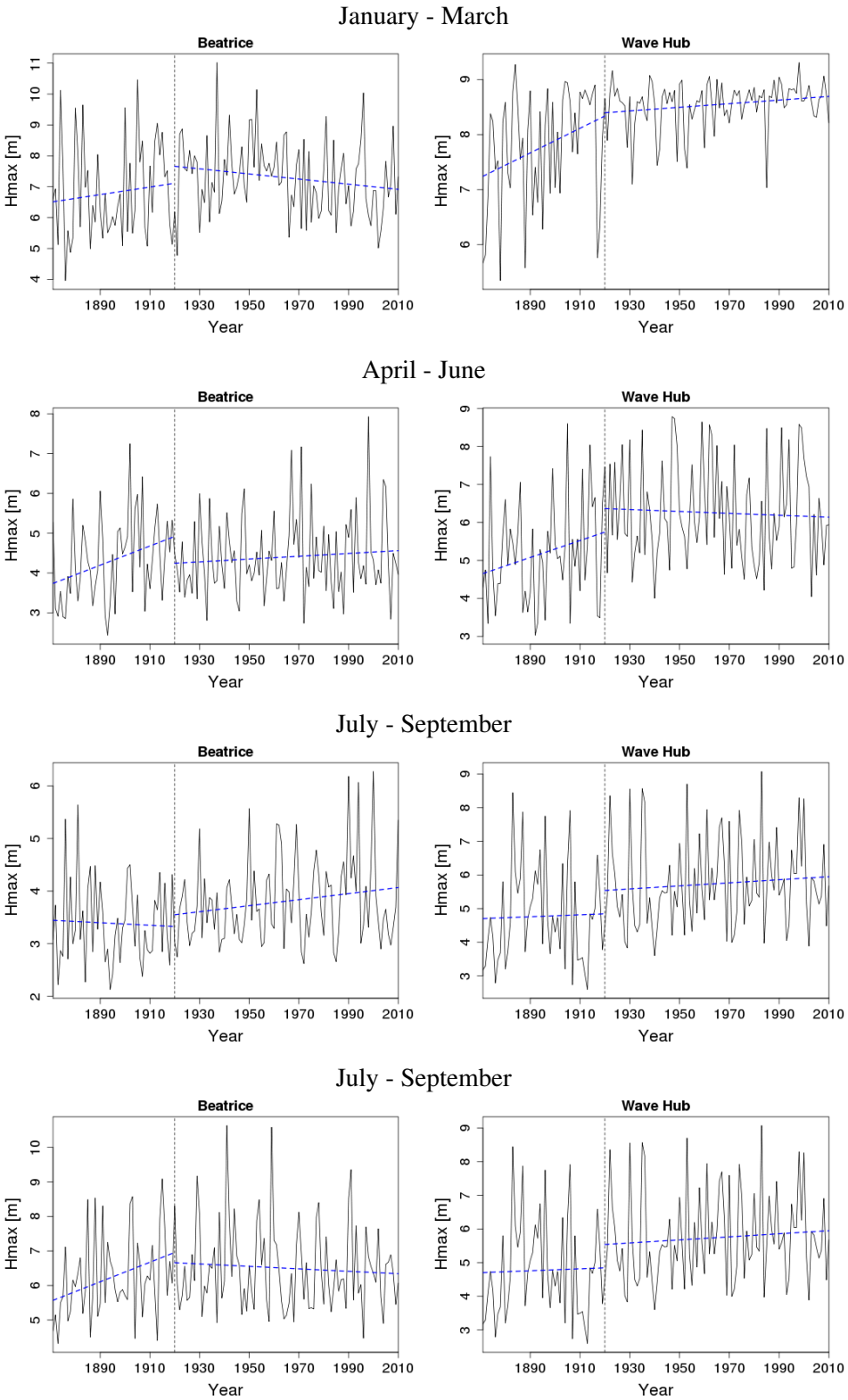

Figure B.30: 140 year time-series of seasonal maxima significant wave height for Beatrice and Wave Hub with the hinge fit trend superimposed (data from WAVEWATCH III)

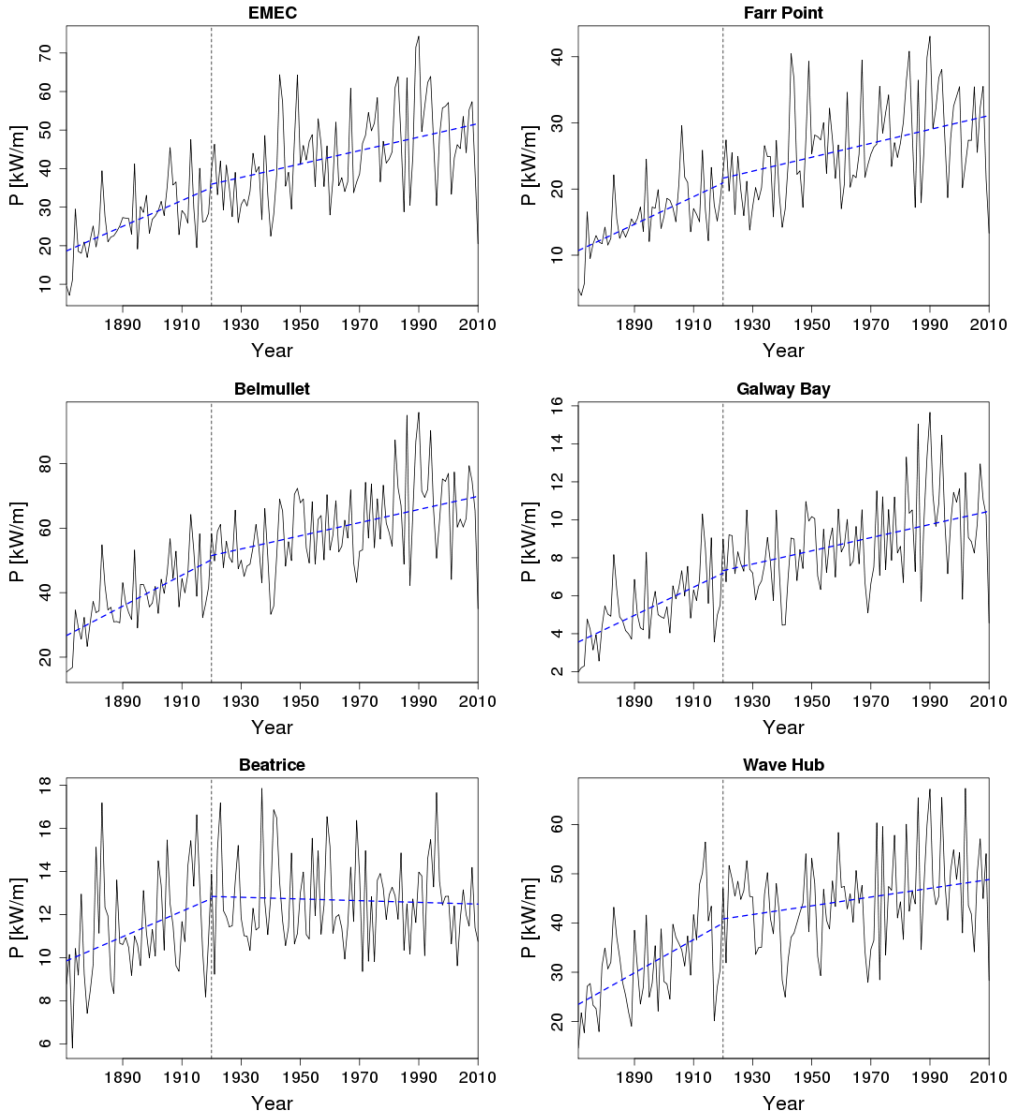

Figure B.31: 140 year time-series of annual mean wave power density for selected locations with the hinge fit trend superimposed (data from WAVEWATCH III)

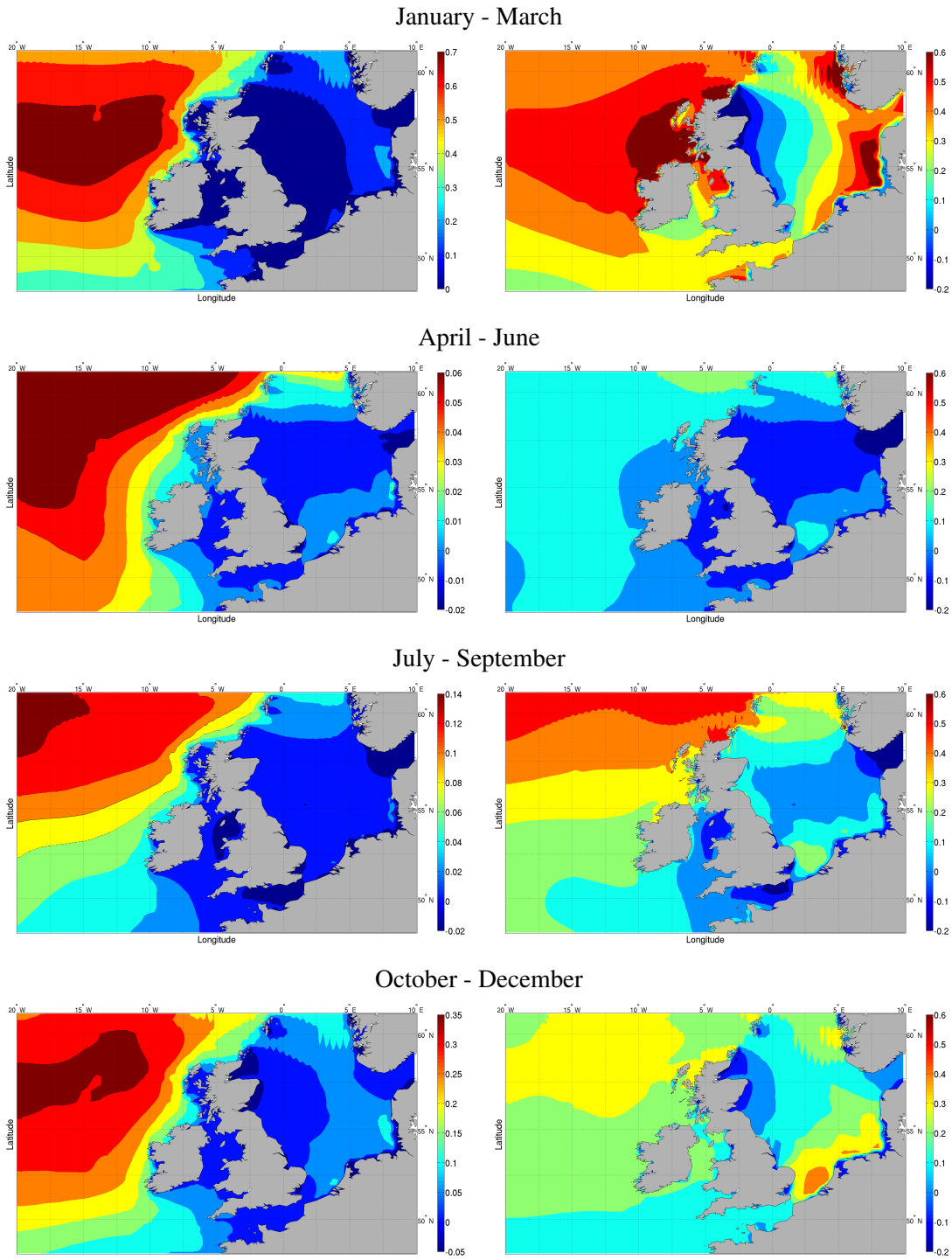

Figure B.32: Changes in the seasonal mean wave power density computed between 1921 and 2010. Left - absolute change (kW per year); right - relative change (% per year).

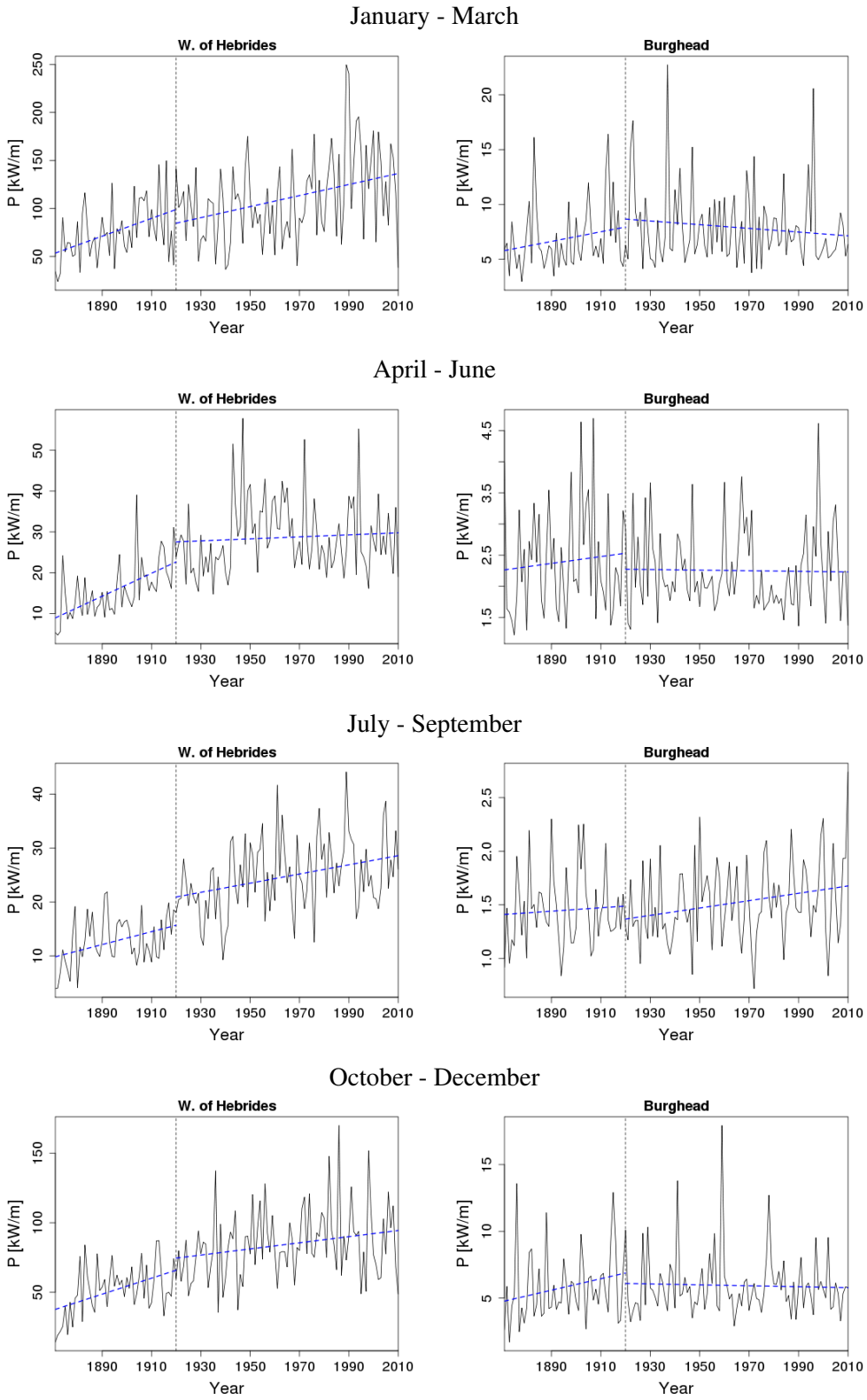

Figure B.33: 140 year time-series of seasonal mean wave power density for West of Hebrides and Burghead with the hinge fit trend superimposed (data from WAVEWATCH III)

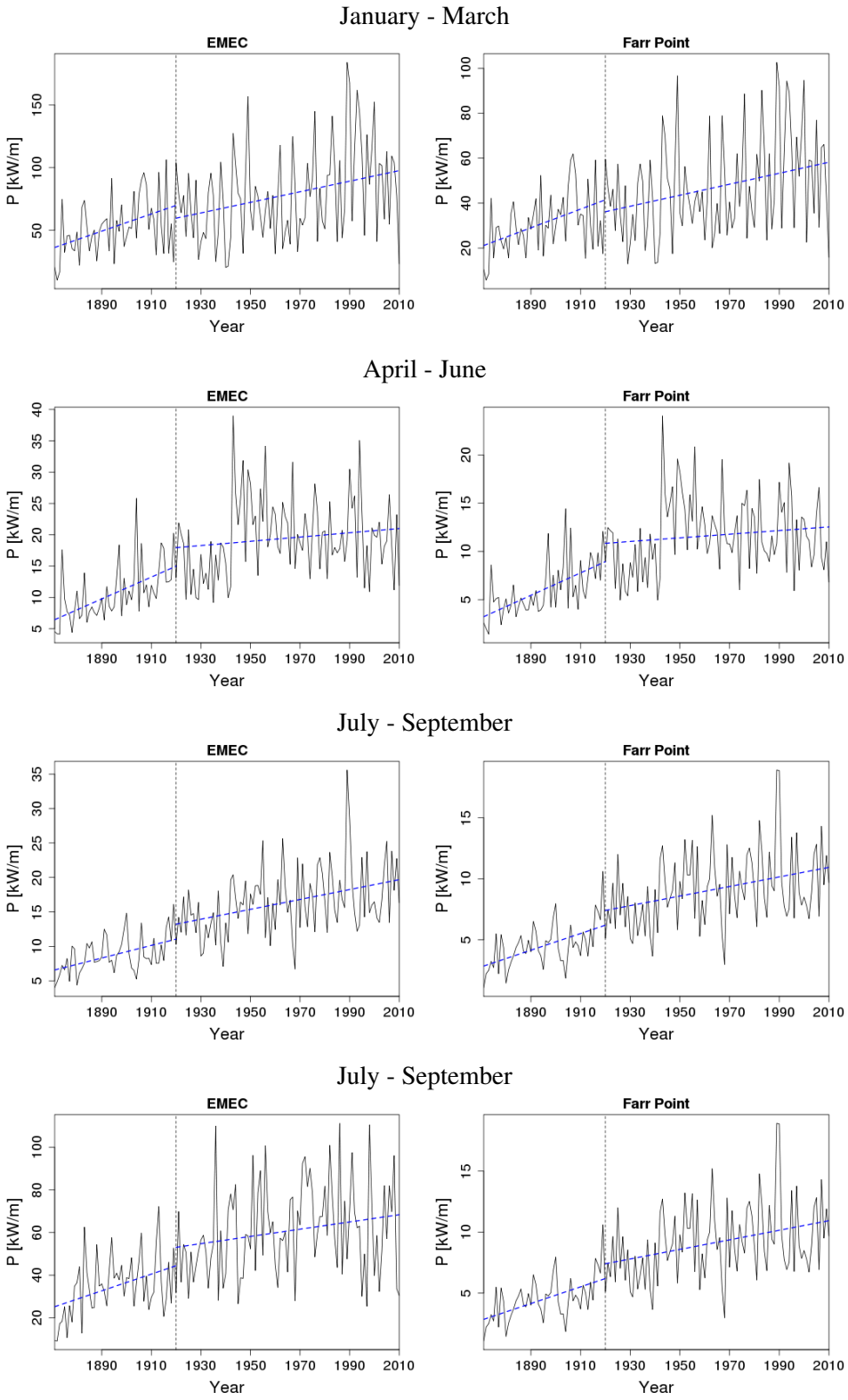

Figure B.34: 140 year time-series of seasonal mean wave power density for EMEC and Farr Point with the hinge fit trend superimposed (data from WAVEWATCH III)

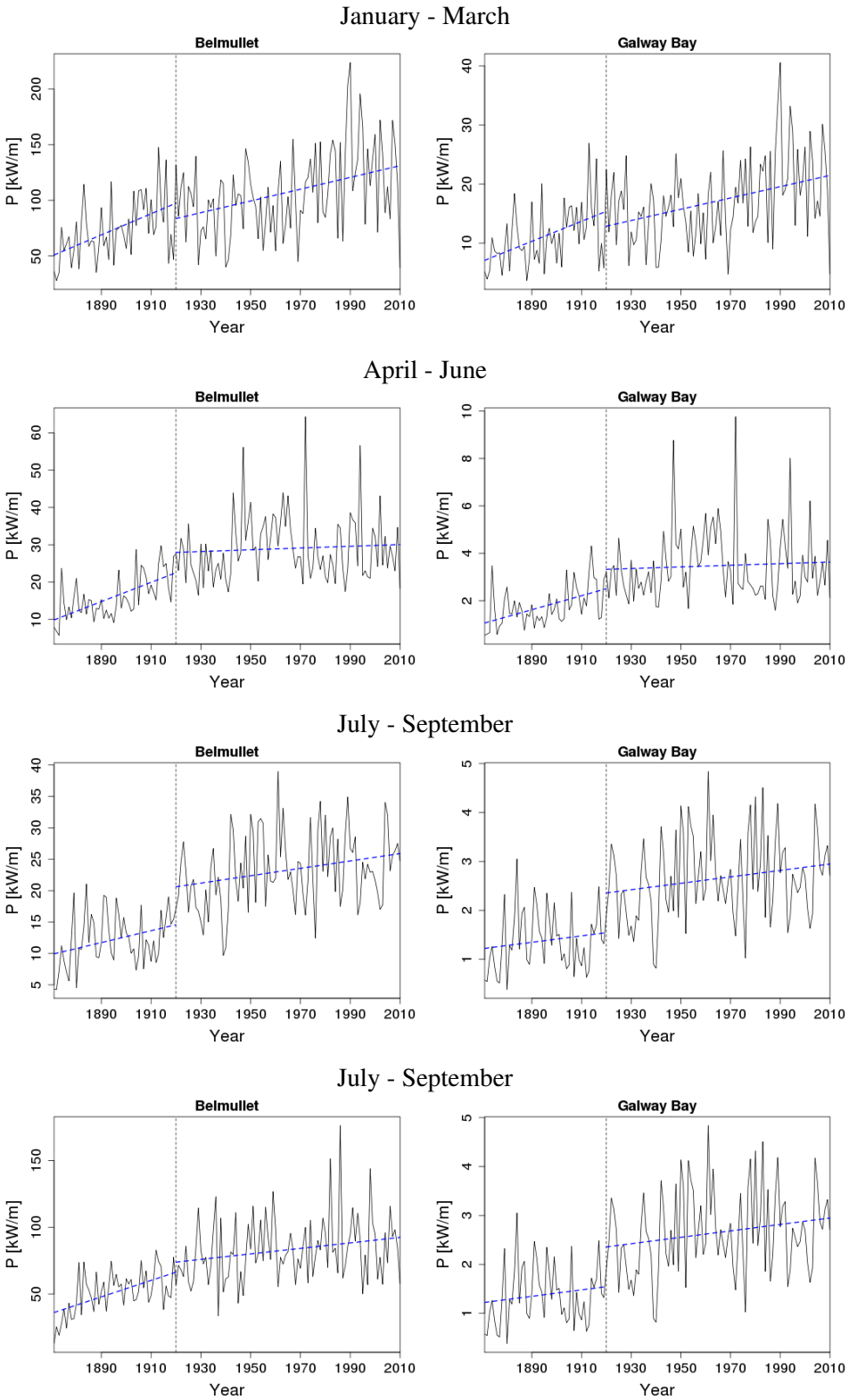

Figure B.35: 140 year time-series of seasonal mean wave power density for Belmullet and Galway Bay with the hinge fit trend superimposed (data from WAVEWATCH III)

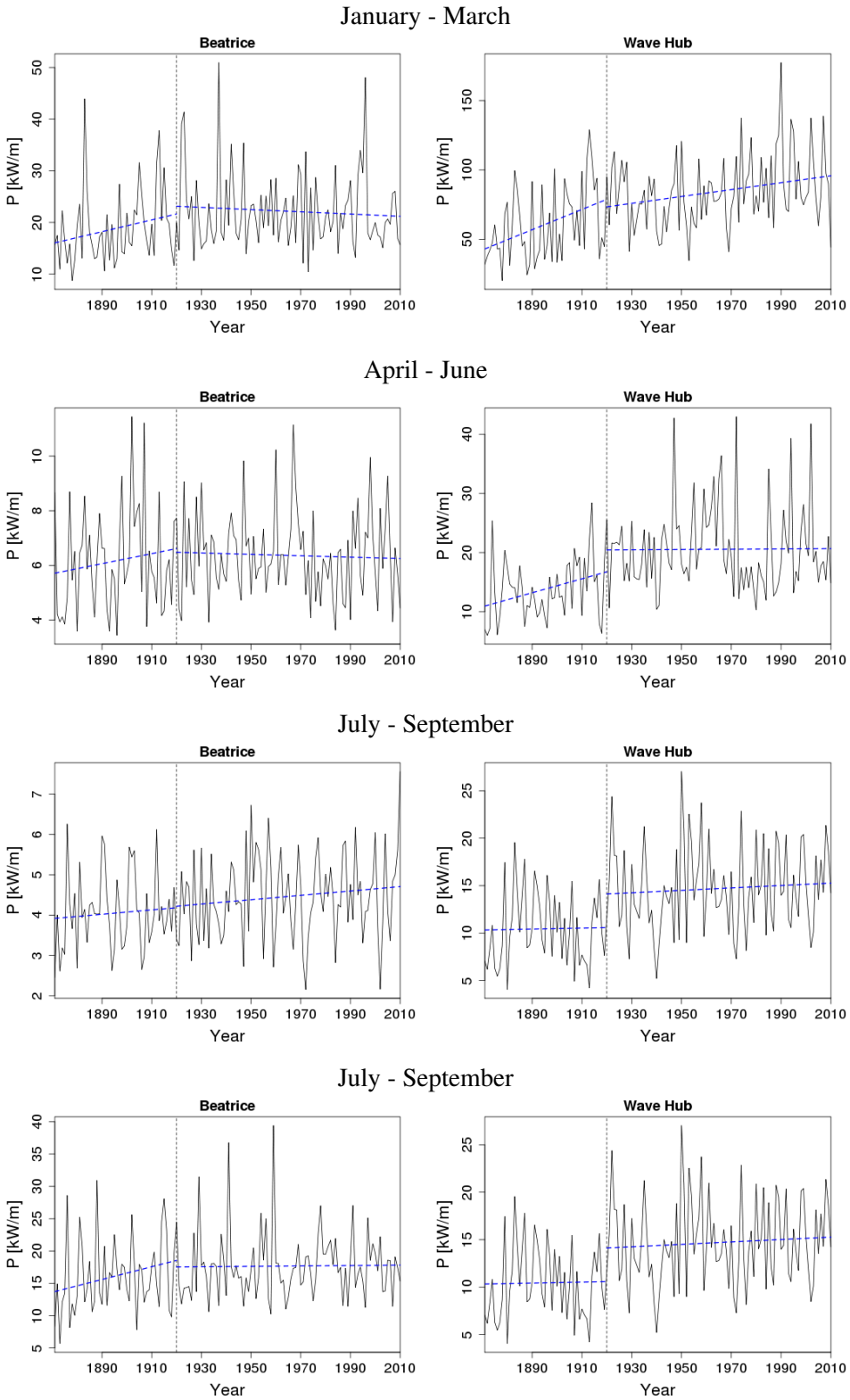

Figure B.36: 140 year time-series of seasonal mean wave power density for Beatrice and Wave Hub with the hinge fit trend superimposed (data from WAVEWATCH III)

Appendix C

Analysis of projected climate - Figures and Tables

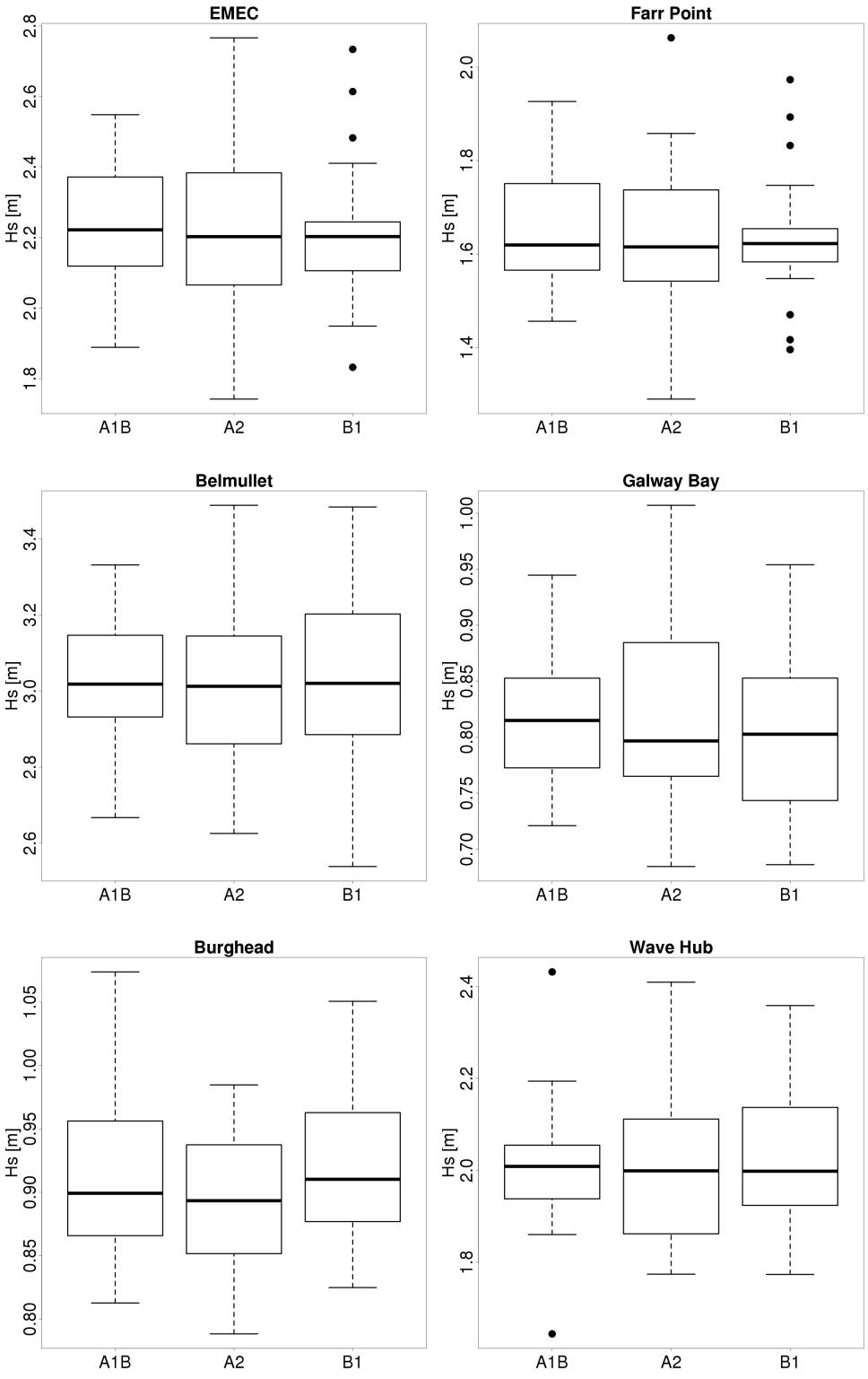

Figure C.1: Boxplots of samples of significant wave height used in the ANOVA test.

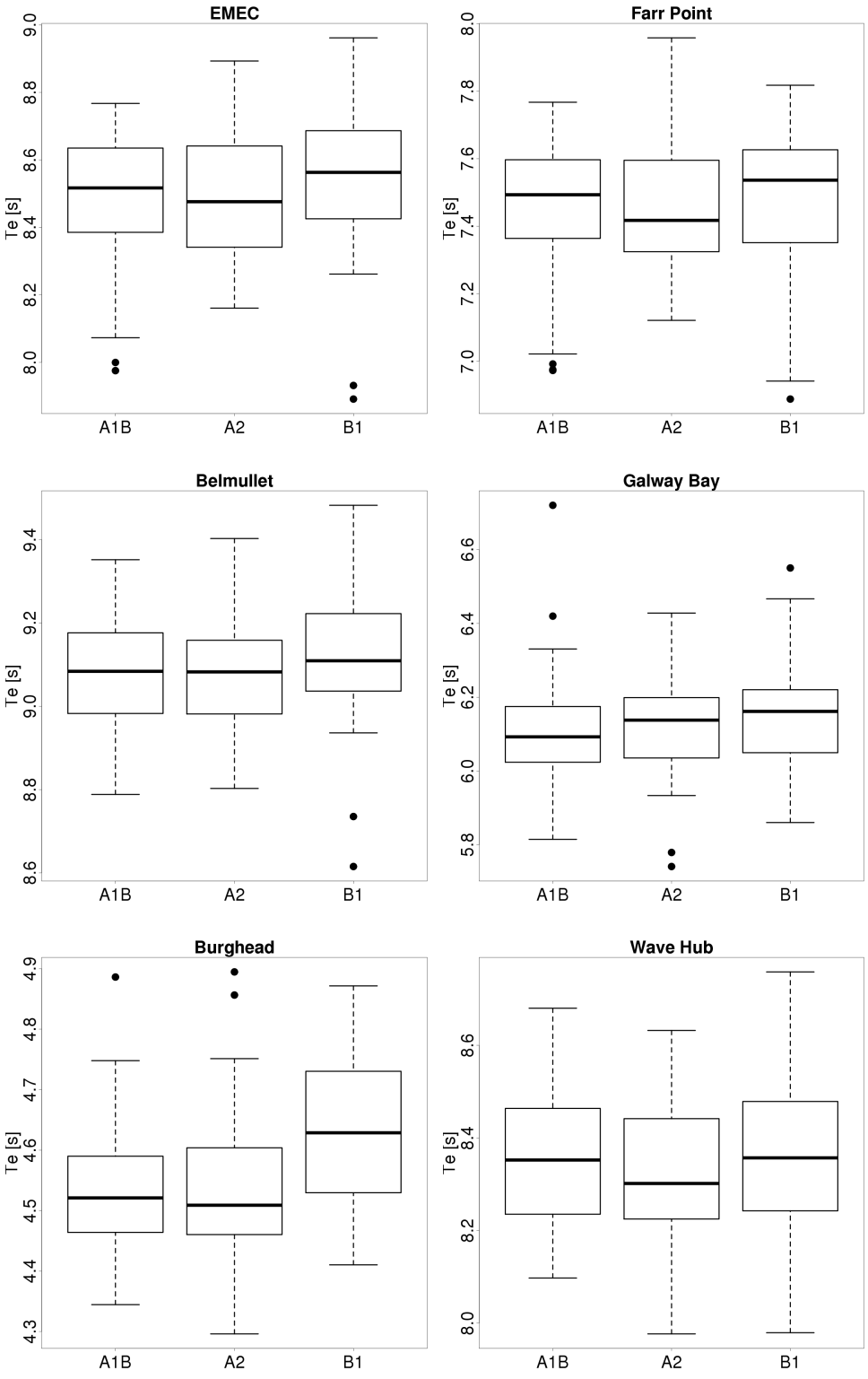

Figure C.2: Boxplots of samples of mean energy period used in the ANOVA test.

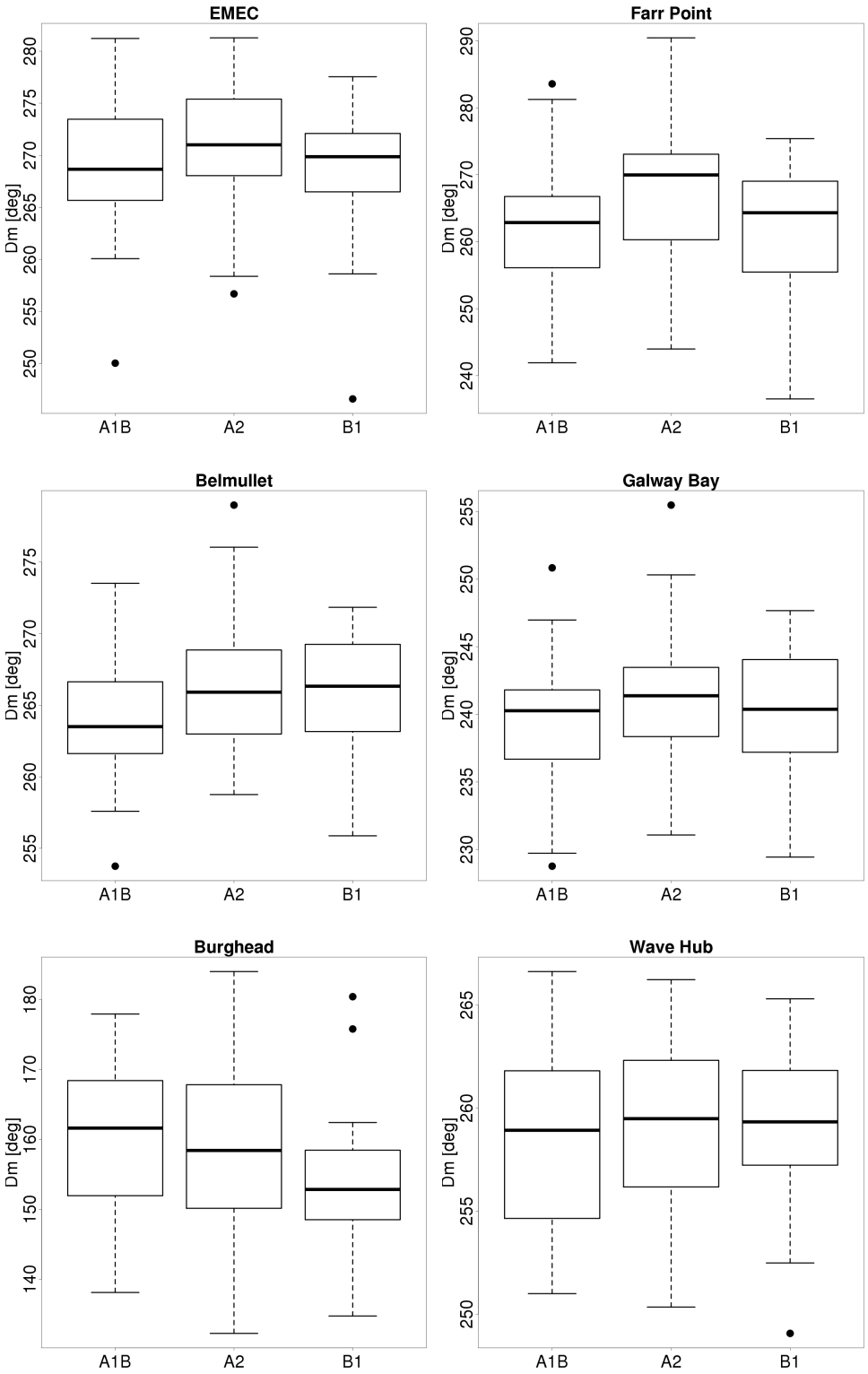

Figure C.3: Boxplots of samples of mean wave direction used in the ANOVA test.

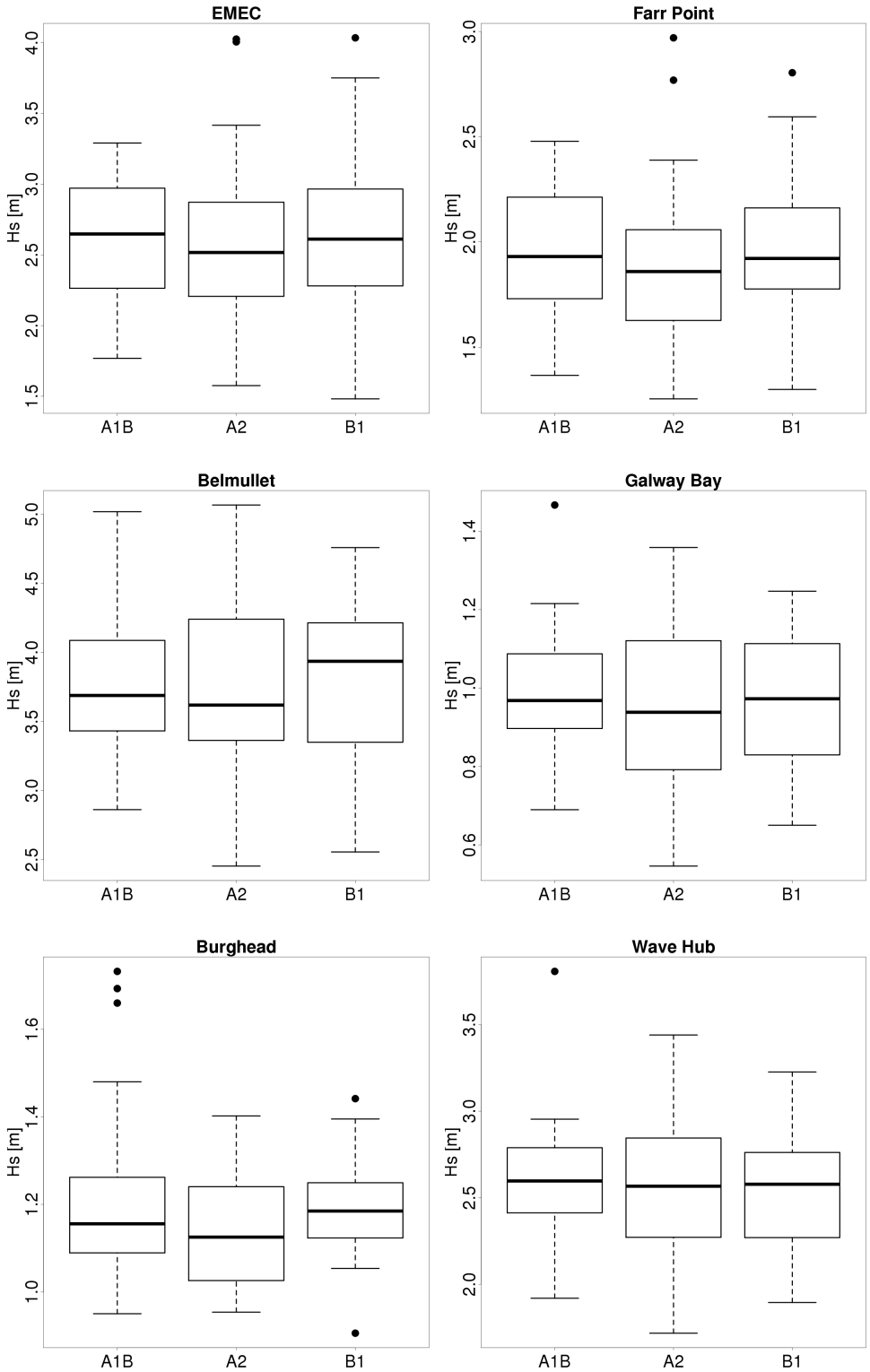

Figure C.4: Boxplots of samples of significant wave height for the months January through March as used in the ANOVA test.

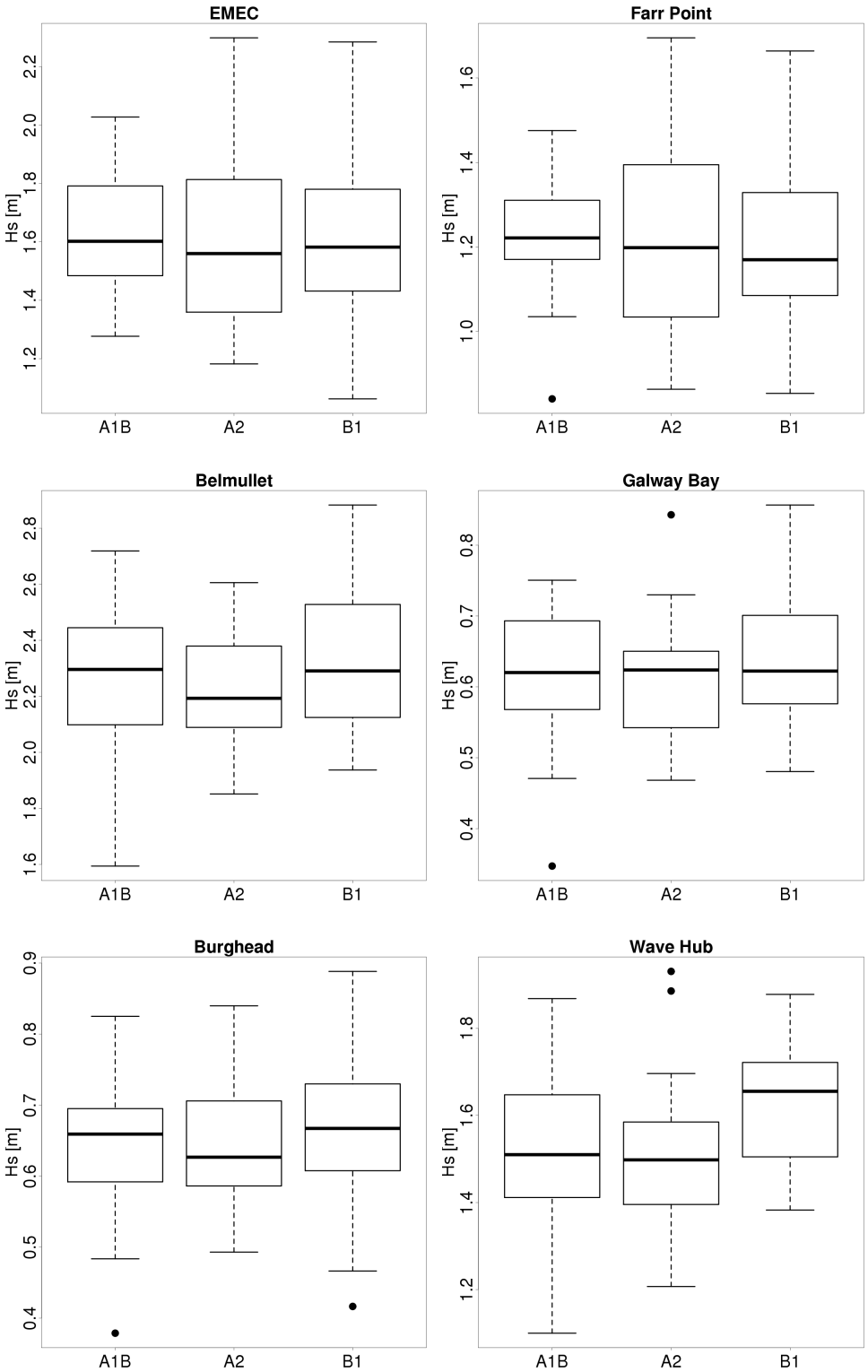

Figure C.5: Boxplots of samples of significant wave height for the months April through June as used in the ANOVA test.

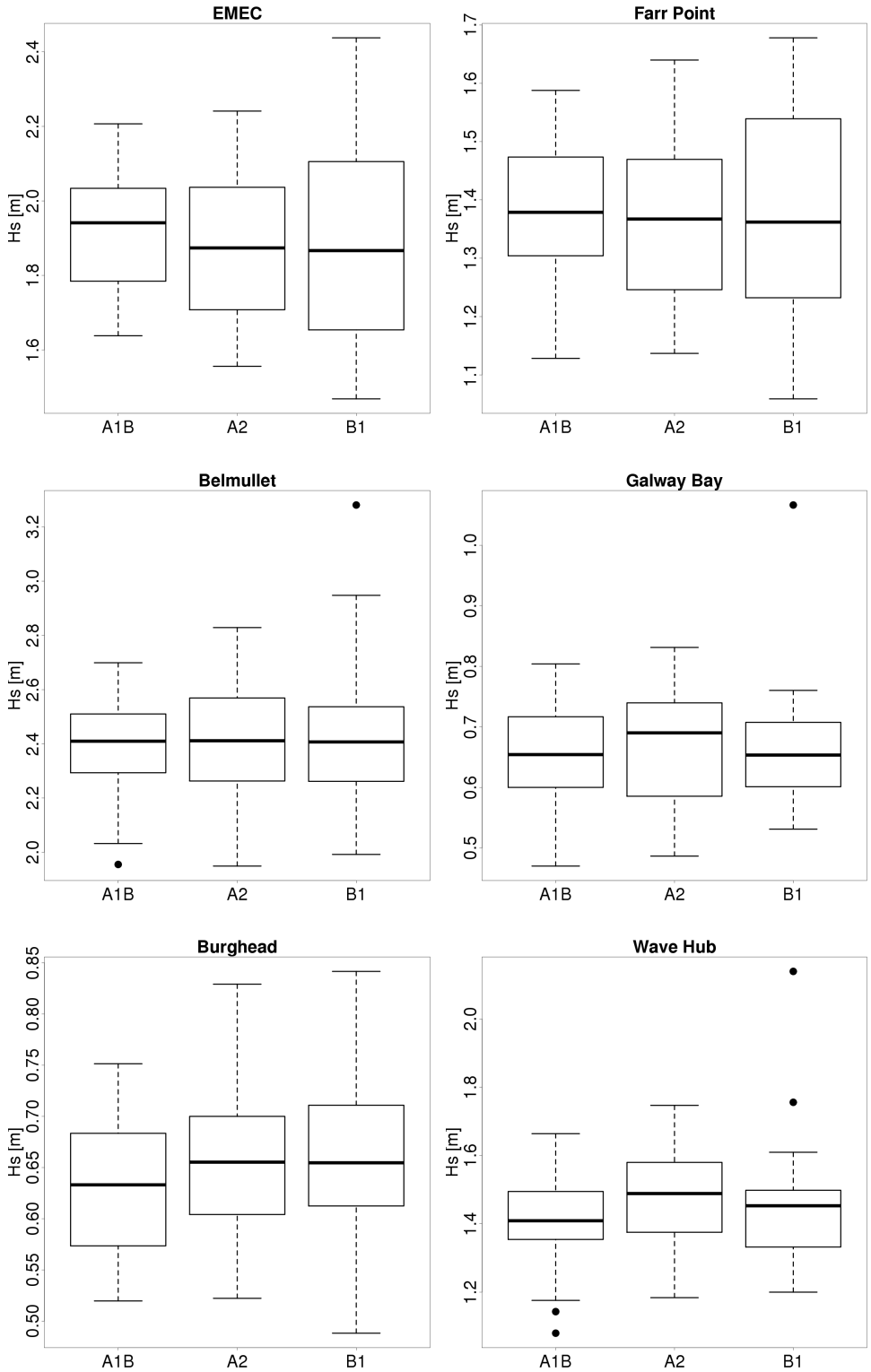

Figure C.6: Boxplots of samples of significant wave height for the months July through September as used in the ANOVA test.

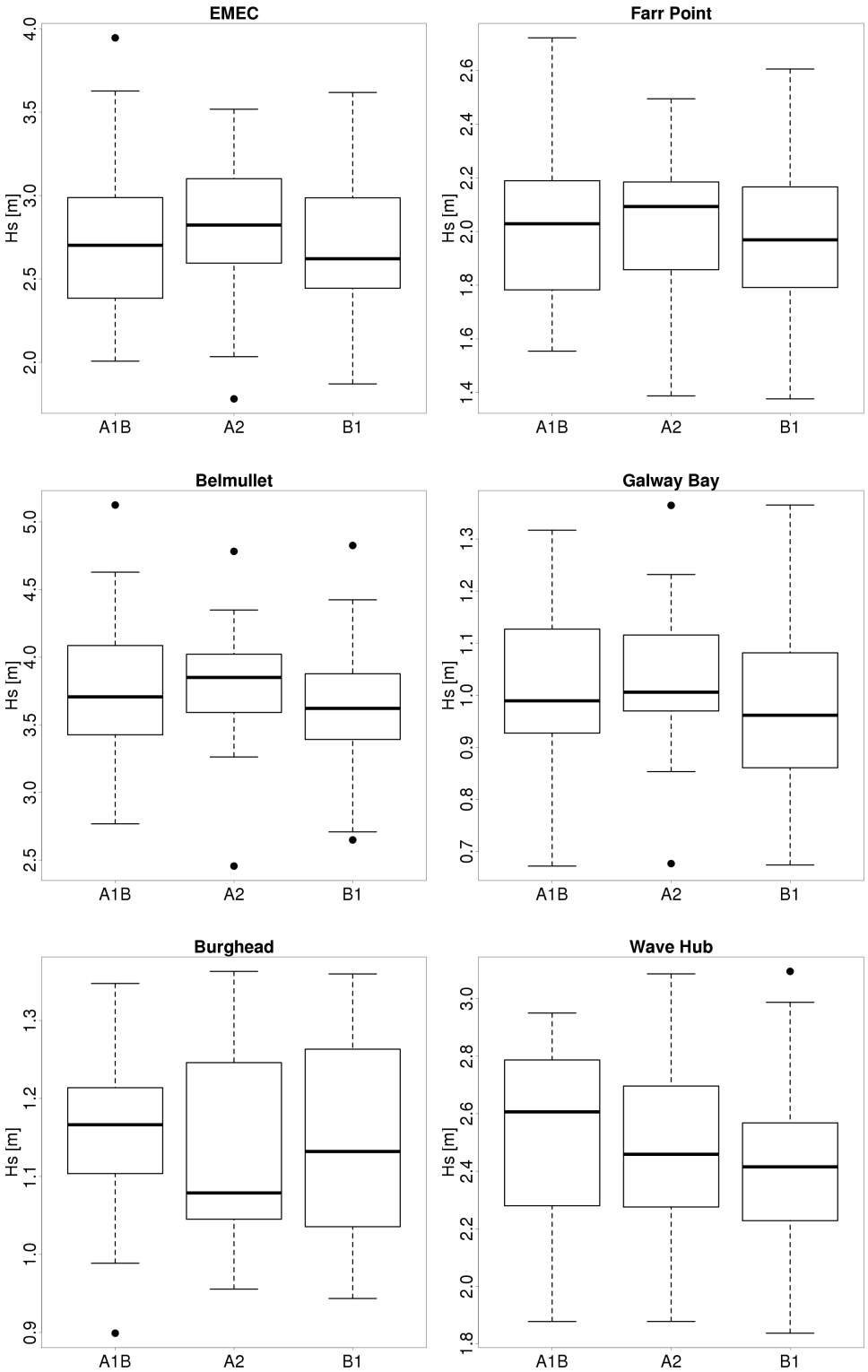

Figure C.7: Boxplots of samples of significant wave height for the months October through December as used in the ANOVA test.

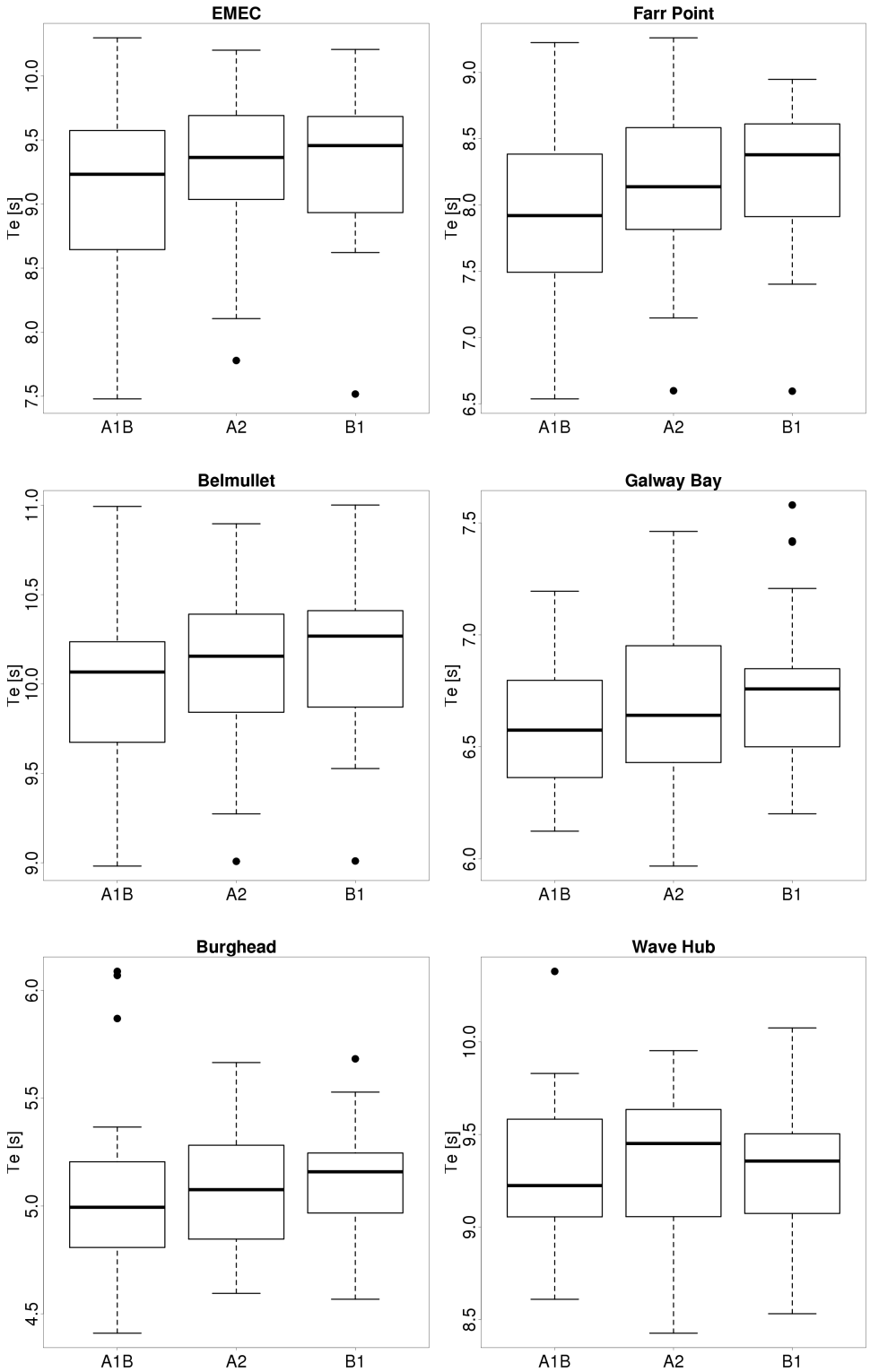

Figure C.8: Boxplots of samples of mean energy period for the months January through March as used in the ANOVA test.

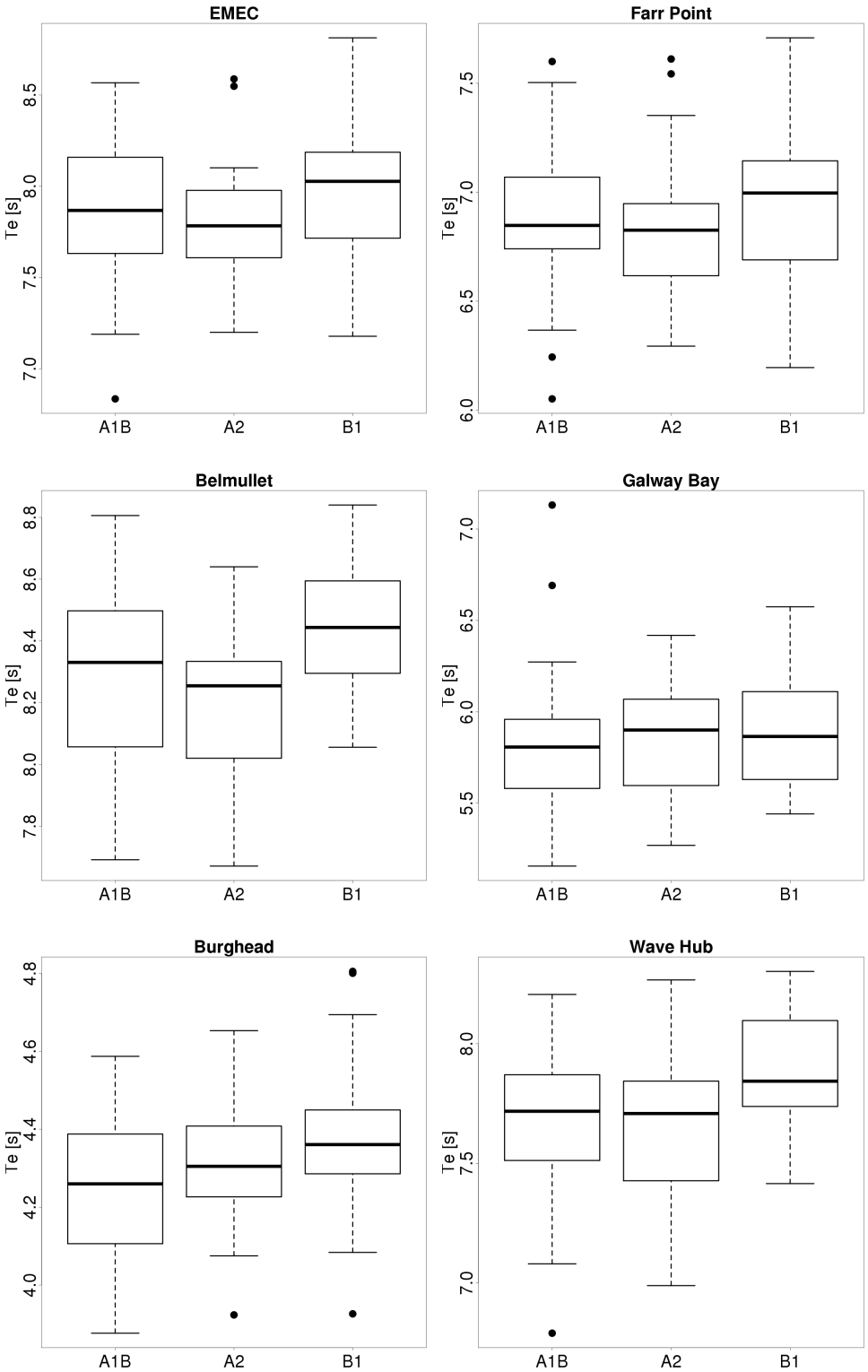

Figure C.9: Boxplots of samples of mean energy period for the months April through June as used in the ANOVA test.

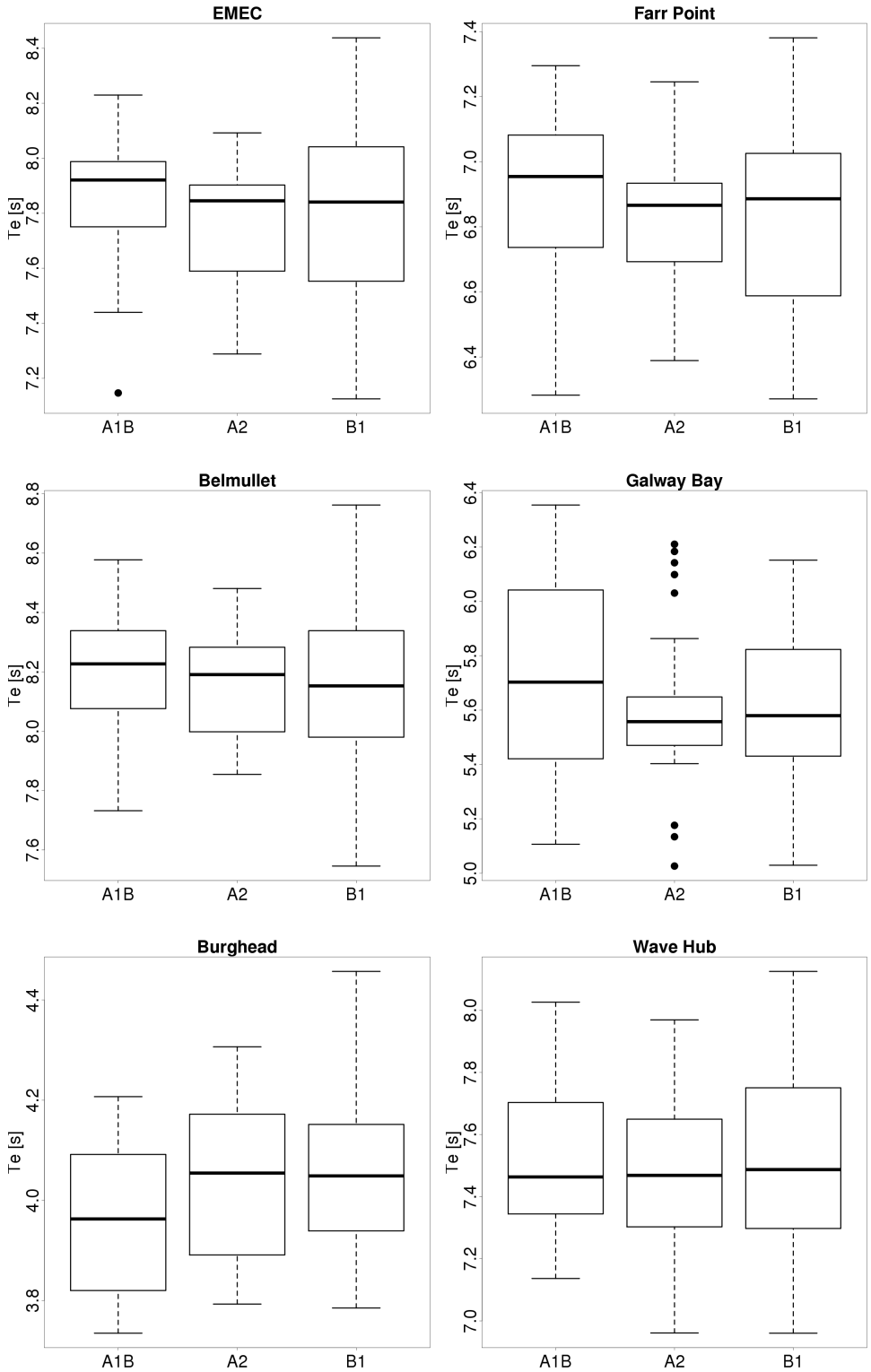

Figure C.10: Boxplots of samples of mean energy period for the months July through September as used in the ANOVA test.

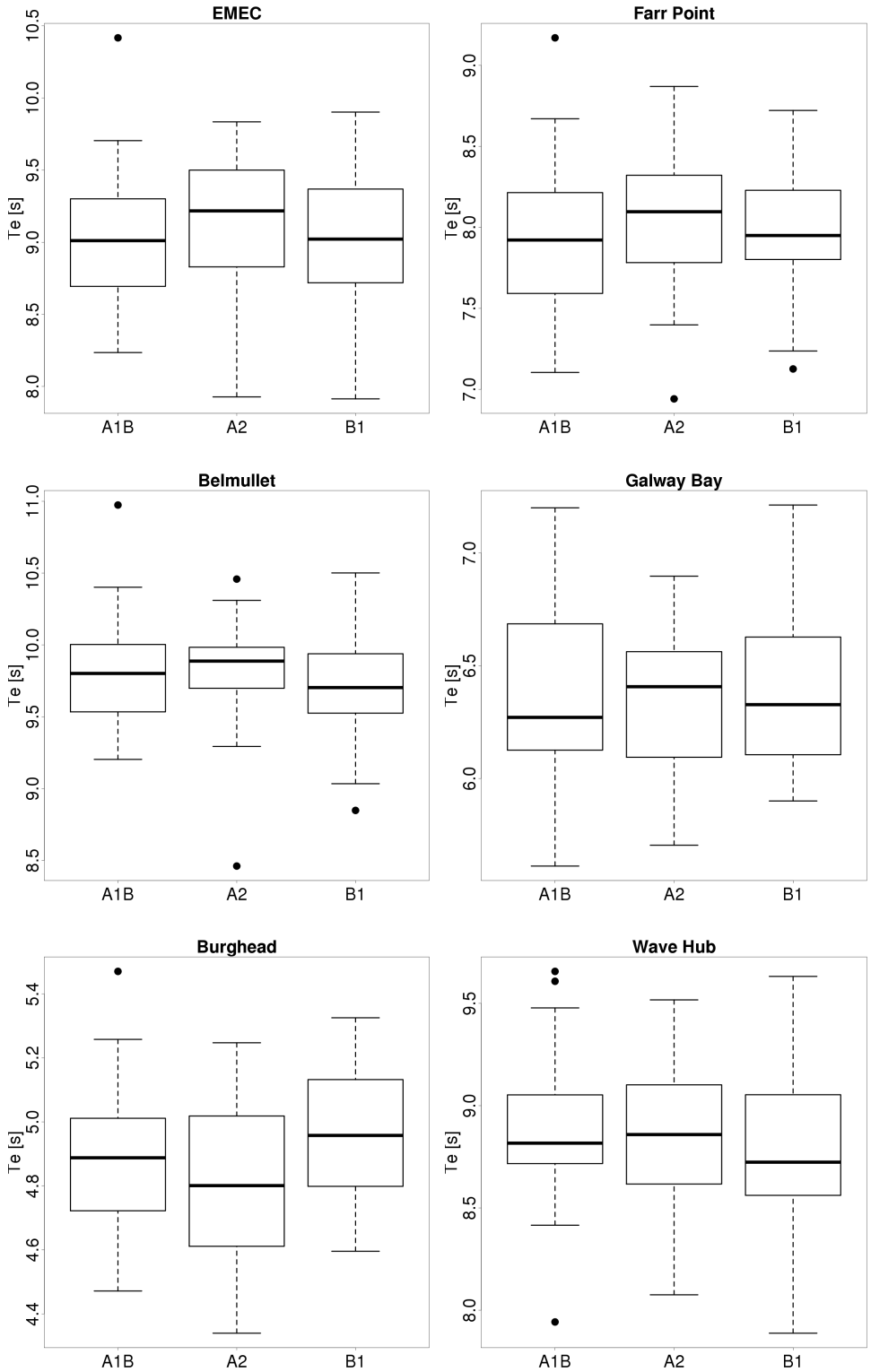

Figure C.11: Boxplots of samples of mean energy period for the months October through November as used in the ANOVA test.

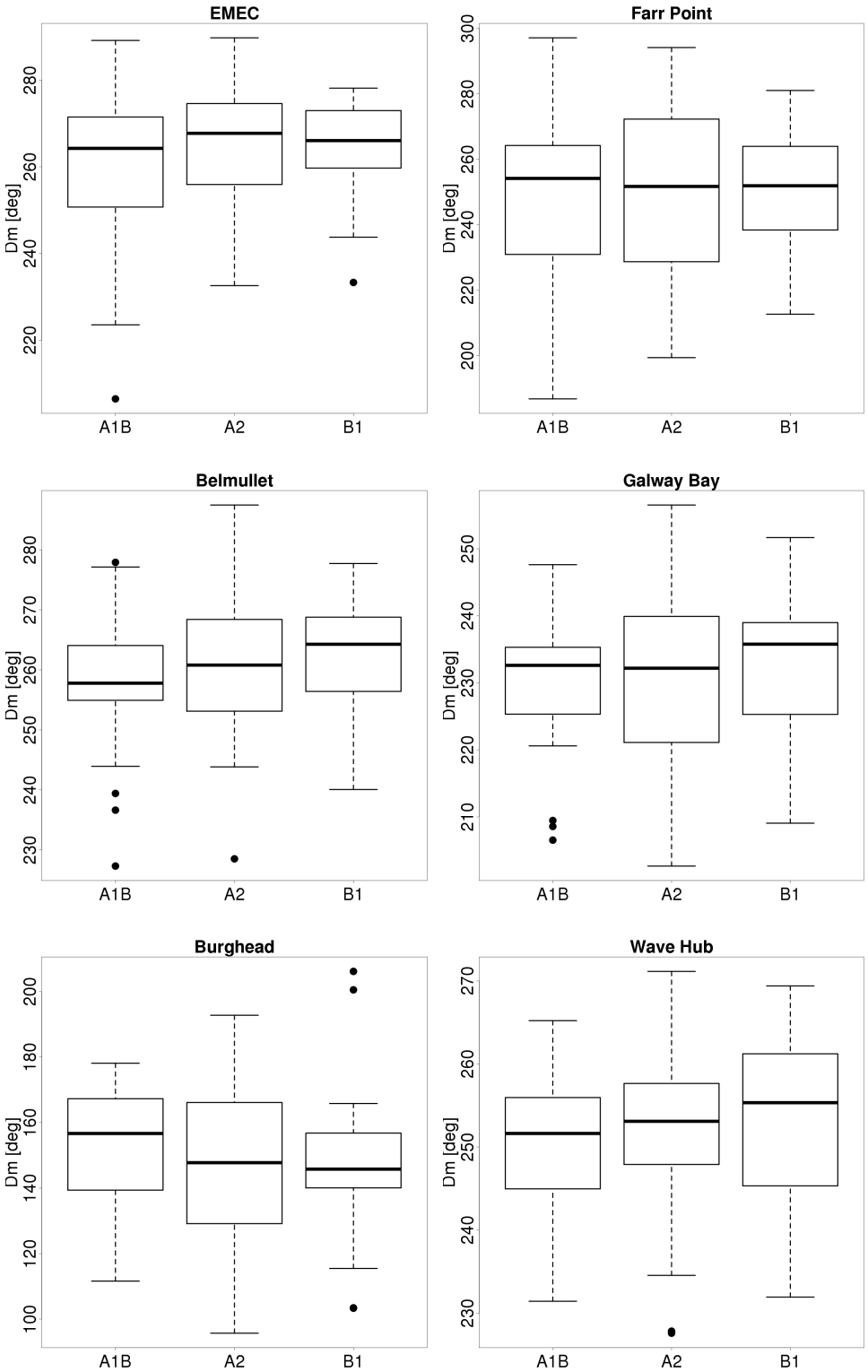

Figure C.12: Boxplots of samples of mean wave direction for the months January through March as used in the ANOVA test.

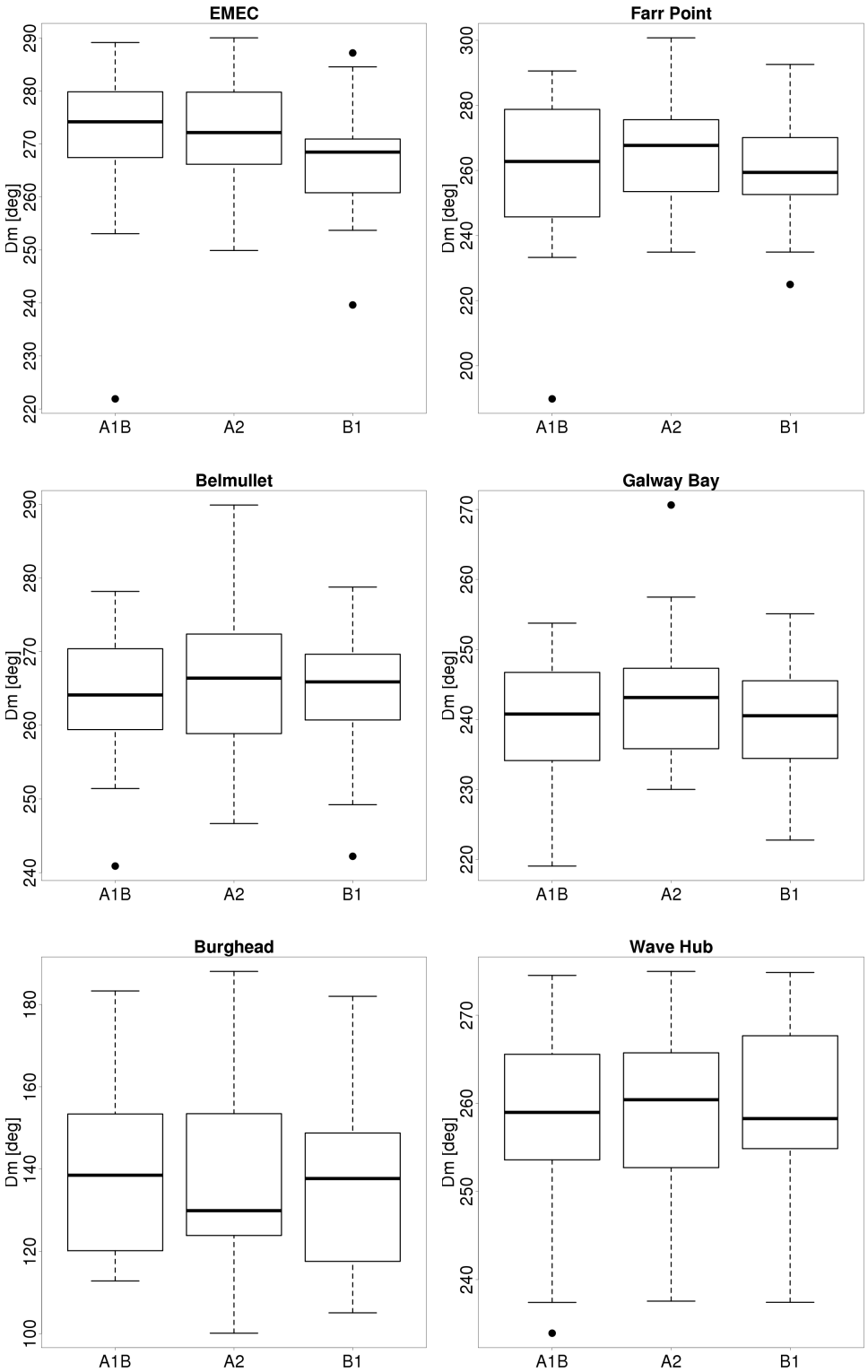

Figure C.13: Boxplots of samples of mean wave direction for the months April through June as used in the ANOVA test.

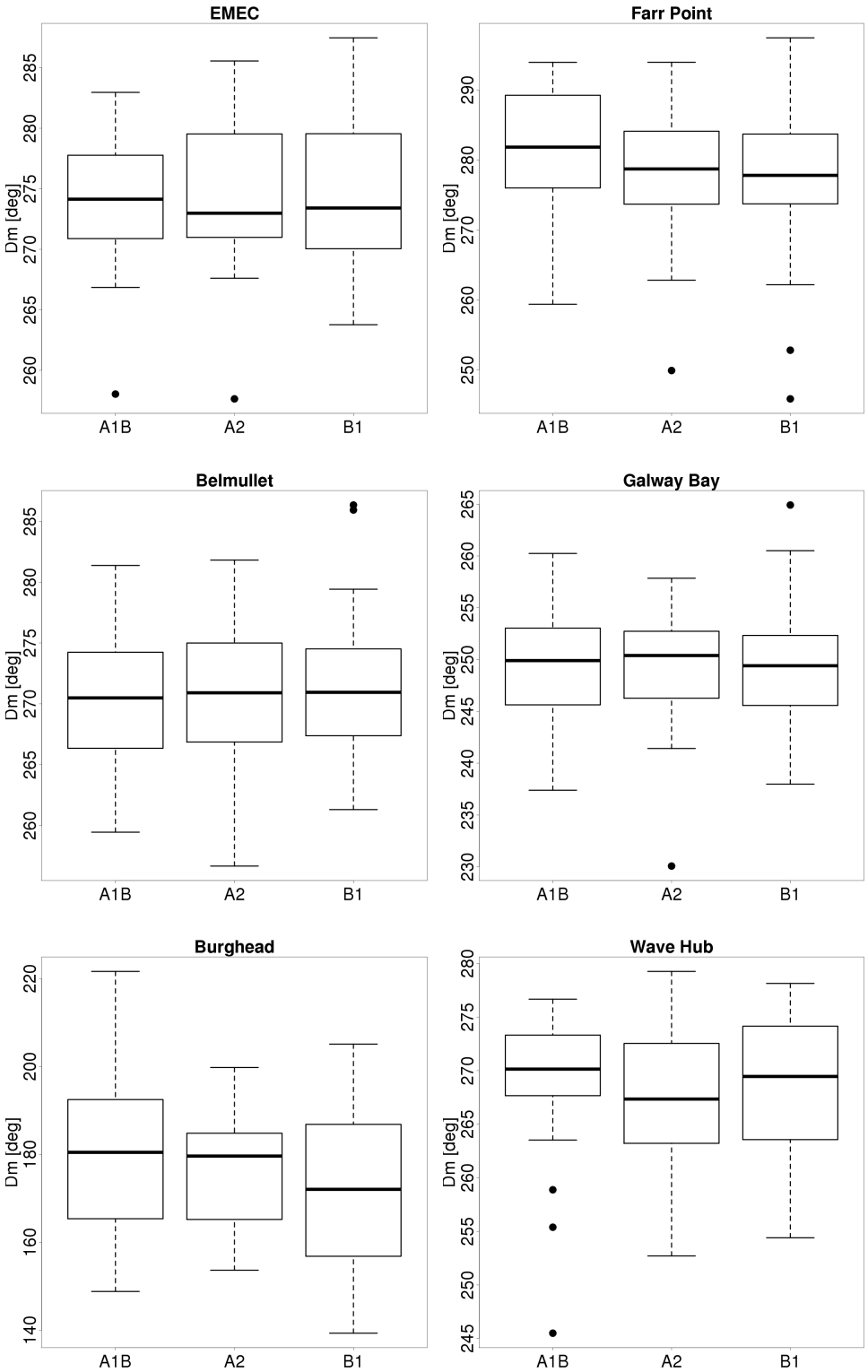

Figure C.14: Boxplots of samples of mean wave direction for the months July through September as used in the ANOVA test.

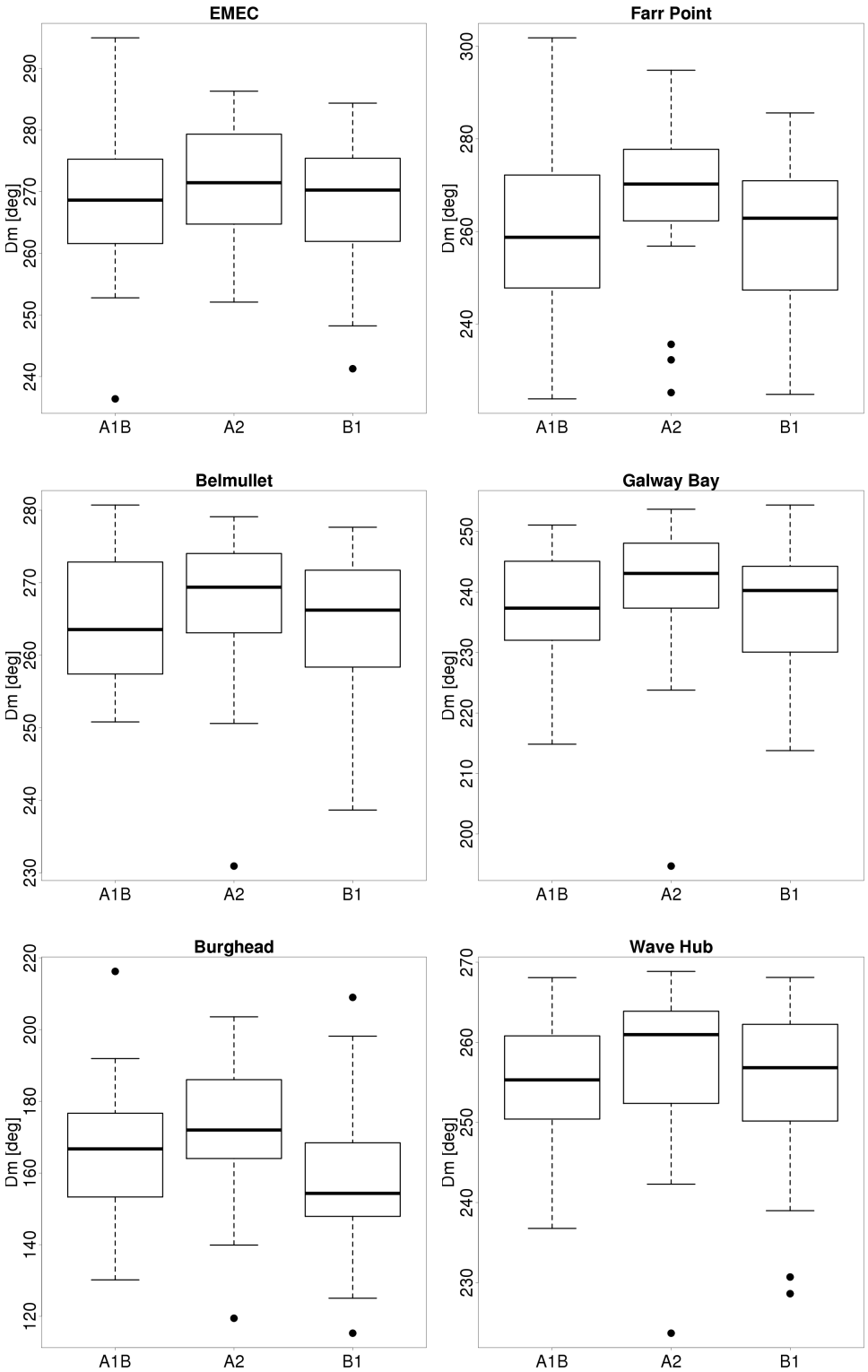

Figure C.15: Boxplots of samples of mean wave direction for the months October through December as used in the ANOVA test.

	H_s		T_e		D_m	
	F-value	P-value	F-value	P-value	F-value	P-value
Farr Point	0.2880	0.7505	1.2352	0.2959	0.2282	0.7964
EMEC	0.2078	0.8128	1.1530	0.3205	1.0613	0.3505
West of Hebrides	0.1804	0.8353	0.8019	0.4518	0.5023	0.6069
Galway Bay	0.2761	0.7594	1.7298	0.1834	0.4701	0.6265
Belmullet	0.1939	0.8241	0.7429	0.4788	1.7108	0.1868
Beatrice	1.9932	0.1424	0.5830	0.5604	0.6423	0.5286
Wave Hub	0.0053	0.9947	0.0016	0.9984	0.4792	0.6209
Burghead	2.0007	0.1414	0.1778	0.8374	0.5952	0.5537

Table C.1: Computed F-values and p-vales from the ANOVA method applied to samples of mean wave parameter data from A2, A1B and B1 scenarios for the months January through March

	H_s		T_e		D_m	
	F-value	P-value	F-value	P-value	F-value	P-value
Farr Point	0.2108	0.8104	0.5461	0.5811	1.1634	0.3172
EMEC	0.0657	0.9369	1.6555	0.1969	1.5598	0.2160
West of Hebrides	0.8621	0.4258	6.2768	0.0028	1.3260	0.2708
Galway Bay	1.0645	0.3494	0.2759	0.7595	0.7742	0.4642
Belmullet	1.8821	0.1584	7.0883	0.0014	0.5358	0.5871
Beatrice	0.4240	0.6558	1.8822	0.1584	0.7999	0.4526
Wave Hub	5.0284	0.0085	4.7924	0.0106	0.2374	0.7892
Burghead	1.8821	0.5606	4.9741	0.0090	0.3980	0.6729

Table C.2: Computed F-values and p-vales from the ANOVA method applied to samples of mean wave parameter data from A2, A1B and B1 scenarios for the months April through June

	H_s		T_e		D_m	
	F-value	P-value	F-value	P-value	F-value	P-value
Farr Point	0.0387	0.9621	1.3449	0.2659	1.2734	0.2850
EMEC	0.2624	0.7698	1.3083	0.2755	0.0767	0.9262
West of Hebrides	0.0634	0.9386	0.4822	0.6191	0.3292	0.7204
Galway Bay	0.3108	0.7336	1.5276	0.2228	0.0459	0.9551
Belmullet	0.0327	0.9679	0.6361	0.5318	0.1399	0.8696
Beatrice	1.7645	0.1774	0.9344	0.3967	2.3656	0.0999
Wave Hub	0.9443	0.3929	0.5727	0.5661	0.5424	0.5833
Burghead	0.0327	0.1917	4.0251	0.0212	2.1763	0.1198

Table C.3: Computed F-values and p-vales from the ANOVA method applied to samples of mean wave parameter data from A2, A1B and B1 scenarios for the months July through September

	H_s		T_e		D_m	
	F-value	P-value	F-value	P-value	F-value	P-value
Farr Point	0.2452	0.7831	0.4928	0.6126	3.3537	0.0395
EMEC	1.0917	0.3402	1.0528	0.3534	1.1450	0.3230
West of Hebrides	1.7676	0.1768	1.1411	0.3242	0.4471	0.6410
Galway Bay	1.4311	0.2446	0.1106	0.8954	1.0430	0.3567
Belmullet	1.4077	0.2502	1.0526	0.3534	0.7686	0.4668
Beatrice	1.0460	0.3557	1.7386	0.1818	1.7715	0.1762
Wave Hub	0.7144	0.4923	0.6059	0.5479	0.7933	0.4556
Burghead	1.4077	0.3308	2.8987	0.0604	3.4833	0.0350

Table C.4: Computed F-values and p-vales from the ANOVA method applied to samples of mean wave parameter data from A2, A1B and B1 scenarios for the months October through December

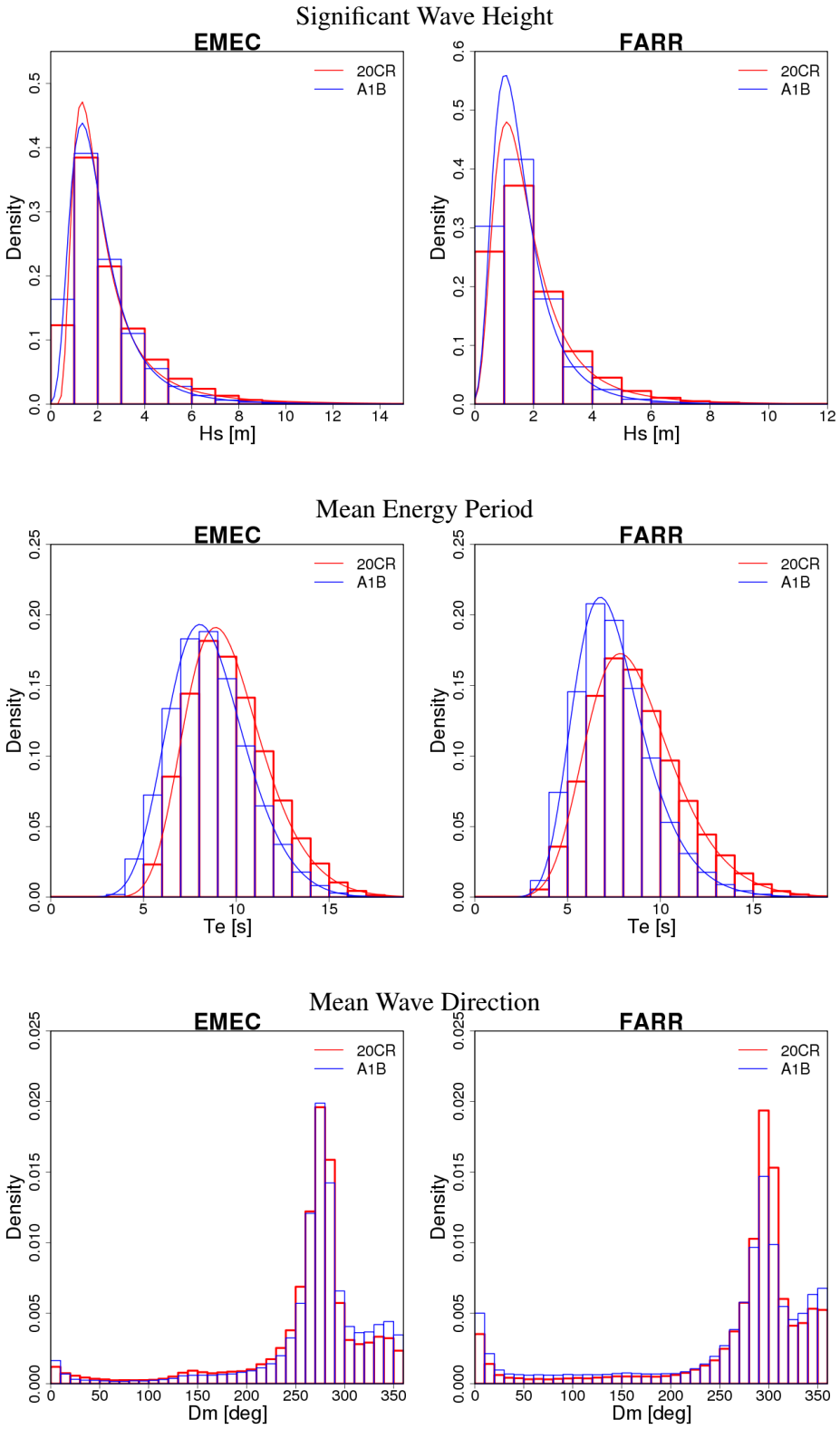

Figure C.16: Histograms of wave data from model runs forced with reanalysis (20CR) and GCM (20C3M+A1B) data describing the climate between 1981 and 2010 for EMEC and Farr Point

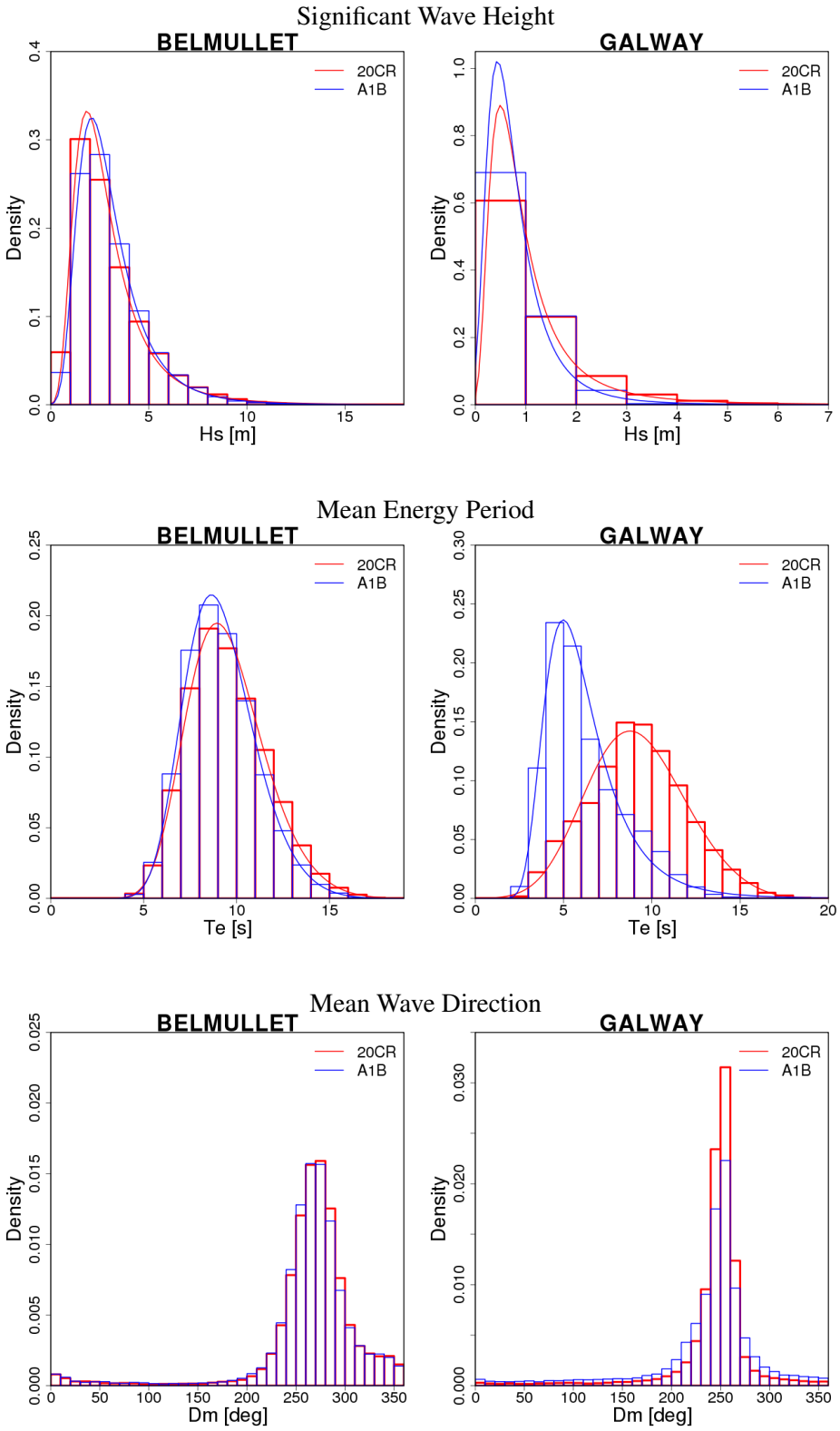

Figure C.17: Histograms of wave data from model runs forced with reanalysis (20CR) and GCM (20C3M+A1B) data describing the climate between 1981 and 2010 for Belmullet and Galway Bay

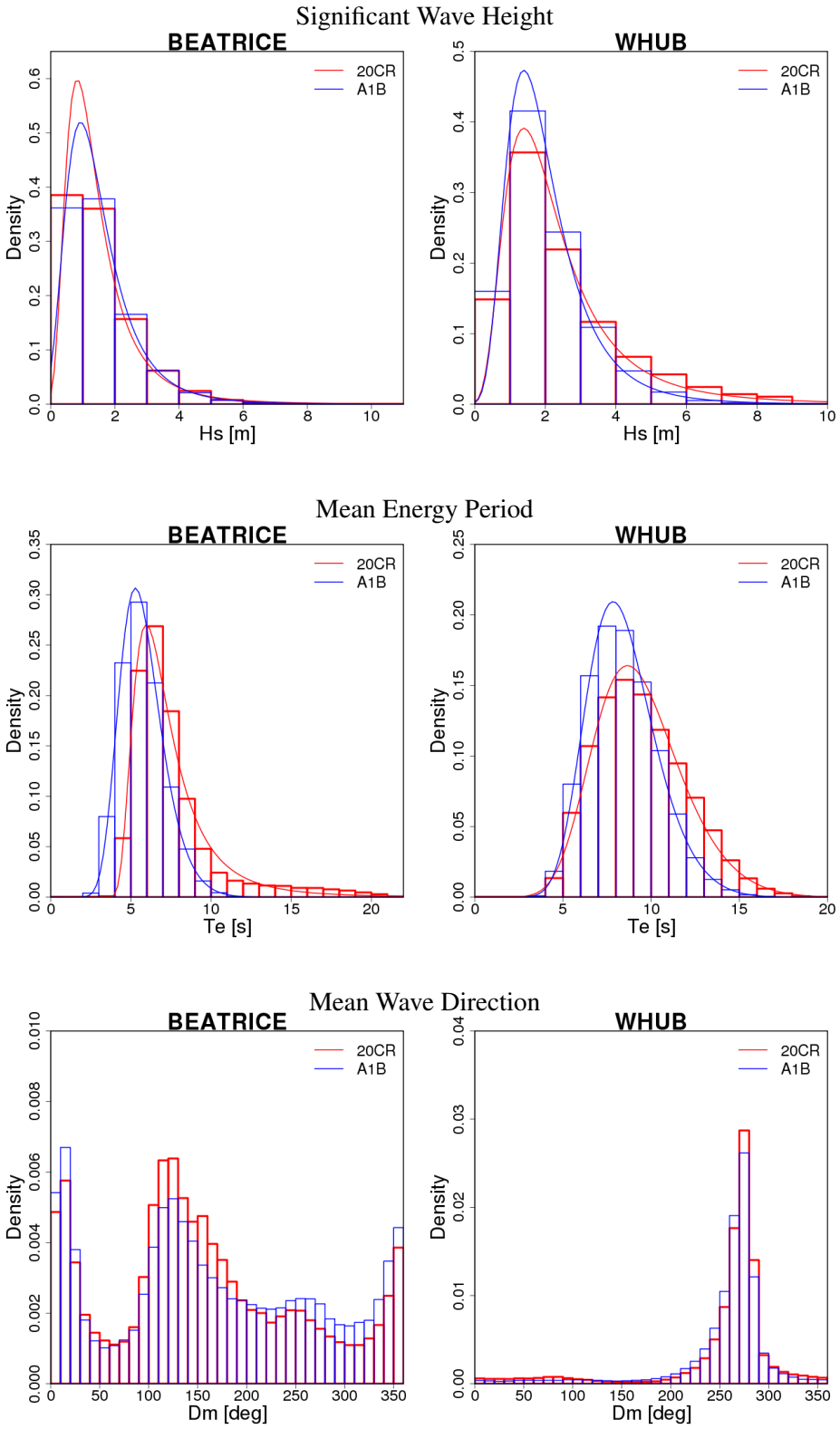

Figure C.18: Histograms of wave data from model runs forced with reanalysis (20CR) and GCM (20C3M+A1B) data describing the climate between 1981 and 2010 for Beatrice and Wave Hub

Site	20CR			20C3M+A1B		
	ξ	μ	σ	ξ	μ	σ
Farr Point	0.2622	1.2629	0.7911	0.1786	1.1358	0.6656
EMEC	0.4147	1.5886	0.8413	0.2232	1.5023	0.8590
West of Hebrides	0.2293	2.0606	1.1924	0.1841	2.0455	1.0977
Galway Bay	0.3725	0.6266	0.4398	0.2237	0.5126	0.3685
Belmullet	0.2210	2.0606	1.1317	0.1497	2.2441	1.1435
Beatrice	0.2639	0.9717	0.6348	0.1346	1.0204	0.7139
Wave Hub	0.2611	1.6078	0.9705	0.1337	1.4868	0.7838
Burghead	0.2726	0.6085	0.4113	0.1574	0.5782	0.4725

Table C.5: Parameter values of probability density functions fit to the 30-year H_s data for observed (20CR) and simulated (20C3M+A1B) climate

Site	20CR			20C3M+A1B		
	ξ	μ	σ	ξ	μ	σ
Farr Point	-0.0877	7.6357	2.1398	-0.0630	6.6526	1.7343
EMEC	-0.1017	8.6789	1.9347	-0.1461	7.7275	1.9246
West of Hebrides	-0.1369	8.4523	1.8817	-0.1422	8.3347	1.7707
Galway Bay	-0.2196	8.1216	2.6569	0.1046	5.1332	1.5635
Belmullet	-0.1382	8.6591	1.9076	-0.1331	8.3946	1.7280
Beatrice	0.2570	6.2674	1.4024	-0.0914	5.1668	1.2046
Wave Hub	-0.1263	8.3567	2.2607	-0.1249	7.6072	1.7709
Burghead	0.3335	5.0257	1.3080	-0.0313	4.1145	1.0516

Table C.6: Parameter values of probability density functions fit to the 30-year T_e data for observed (20CR) and simulated (20C3M+A1B) climate

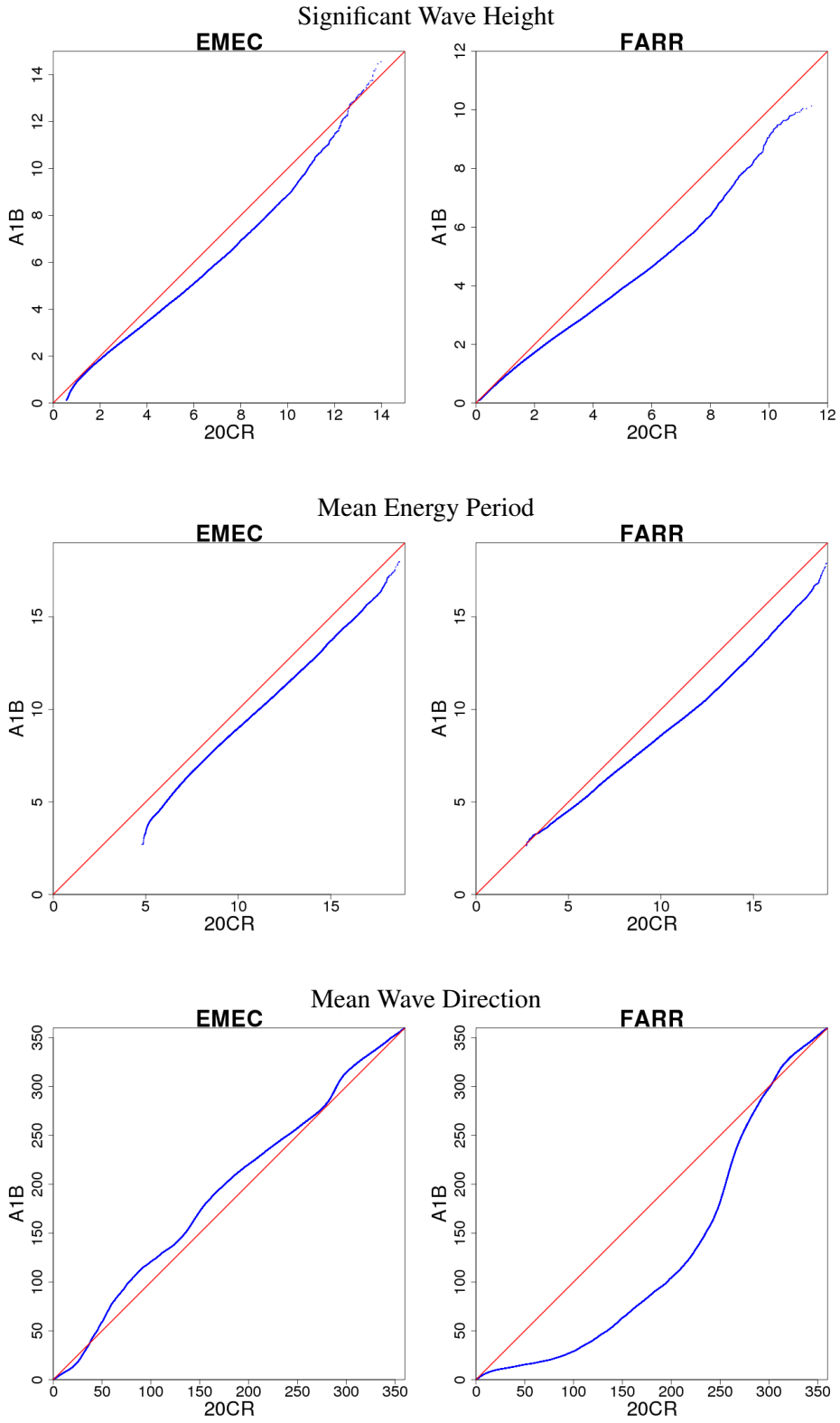

Figure C.19: Quantile plots of ordered wave data from model runs forced with reanalysis (20CR) and GCM (20C3M+A1B) data describing the climate between 1981 and 2010 for EMEC and Farr Point

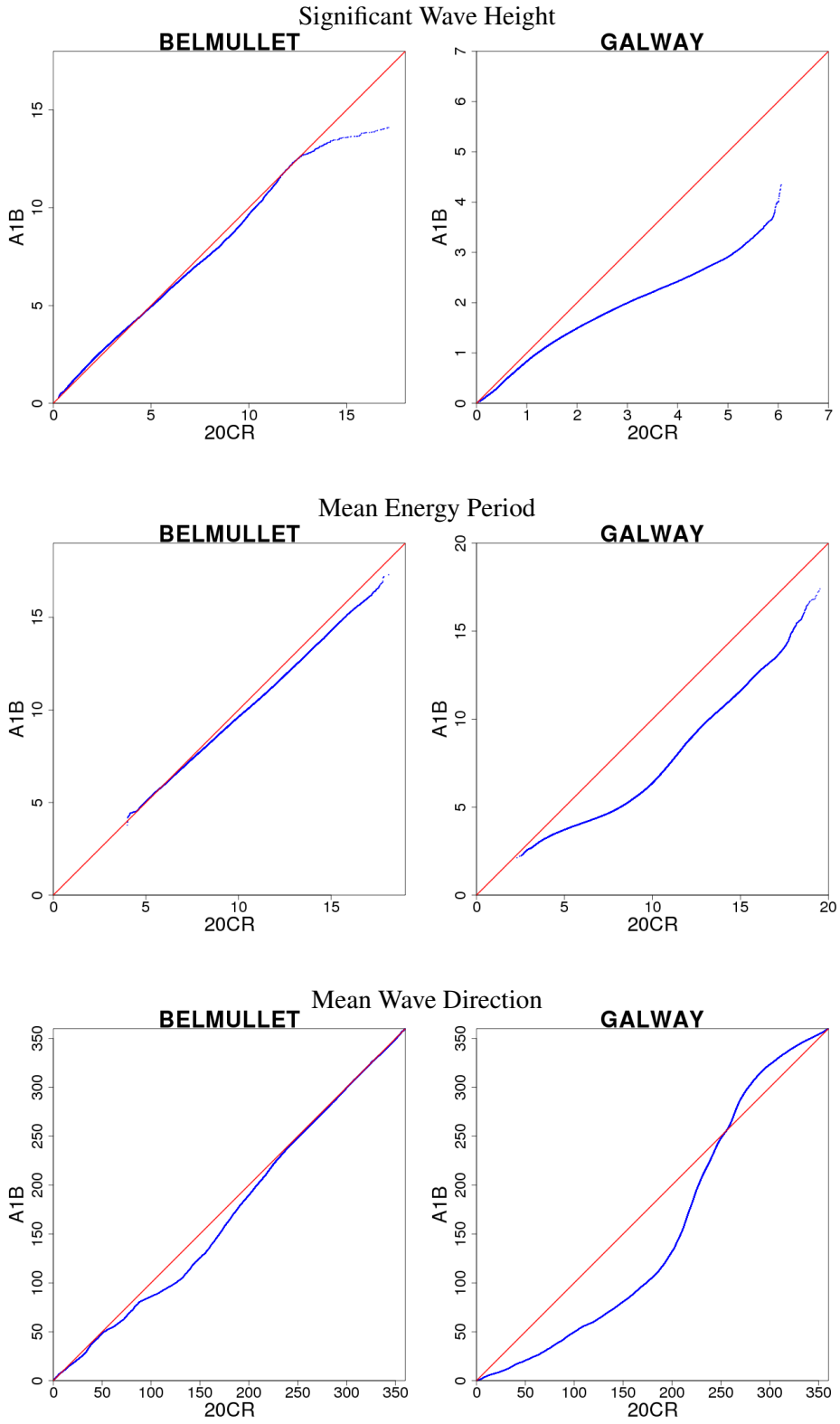

Figure C.20: Quantile plots of ordered wave data from model runs forced with reanalysis (20CR) and GCM (20C3M+A1B) data describing the climate between 1981 and 2010 for Belmullet and Galway Bay

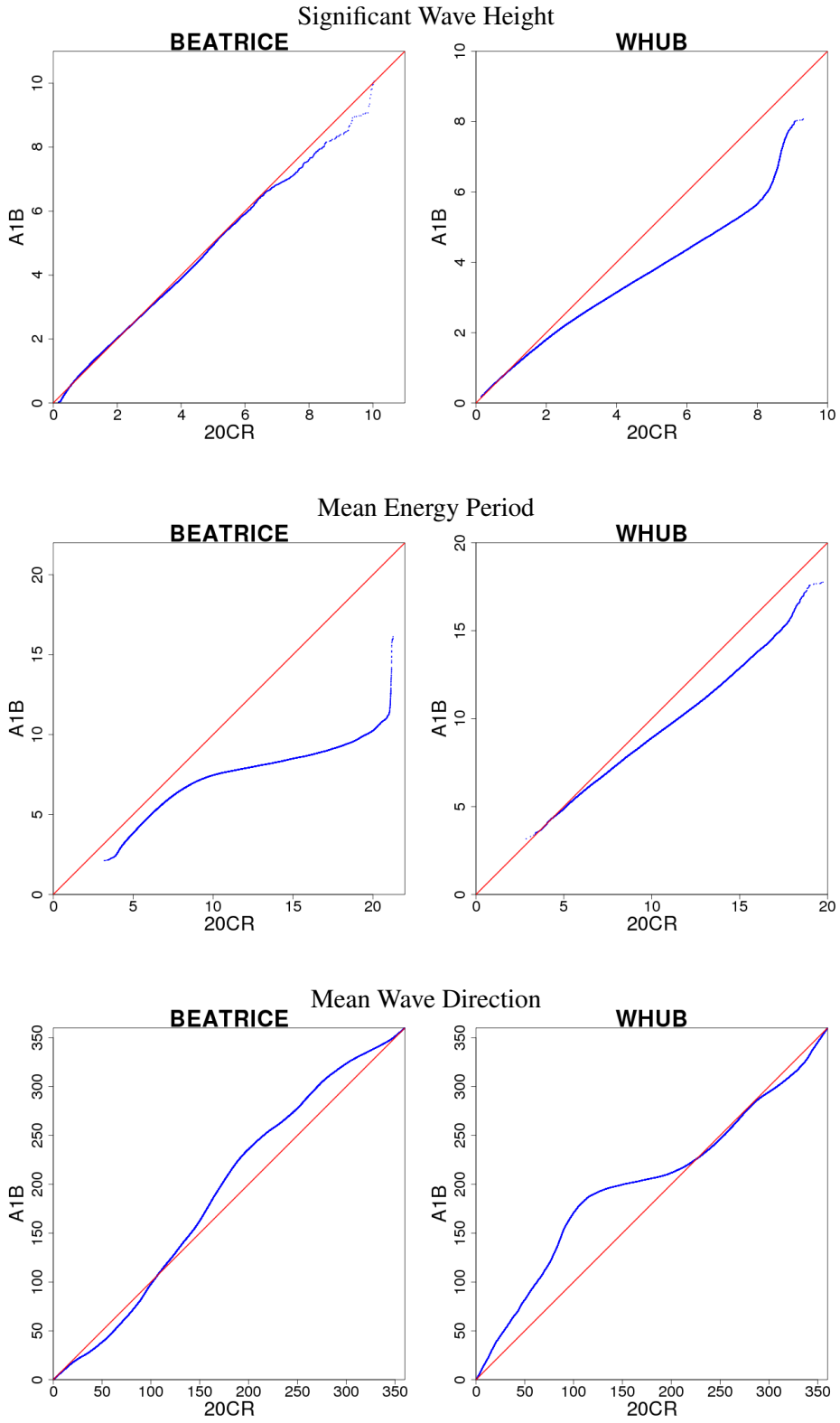

Figure C.21: Quantile plots of ordered wave data from model runs forced with reanalysis (20CR) and GCM (20C3M+A1B) data describing the climate between 1981 and 2010 for Beatrice and Wave Hub

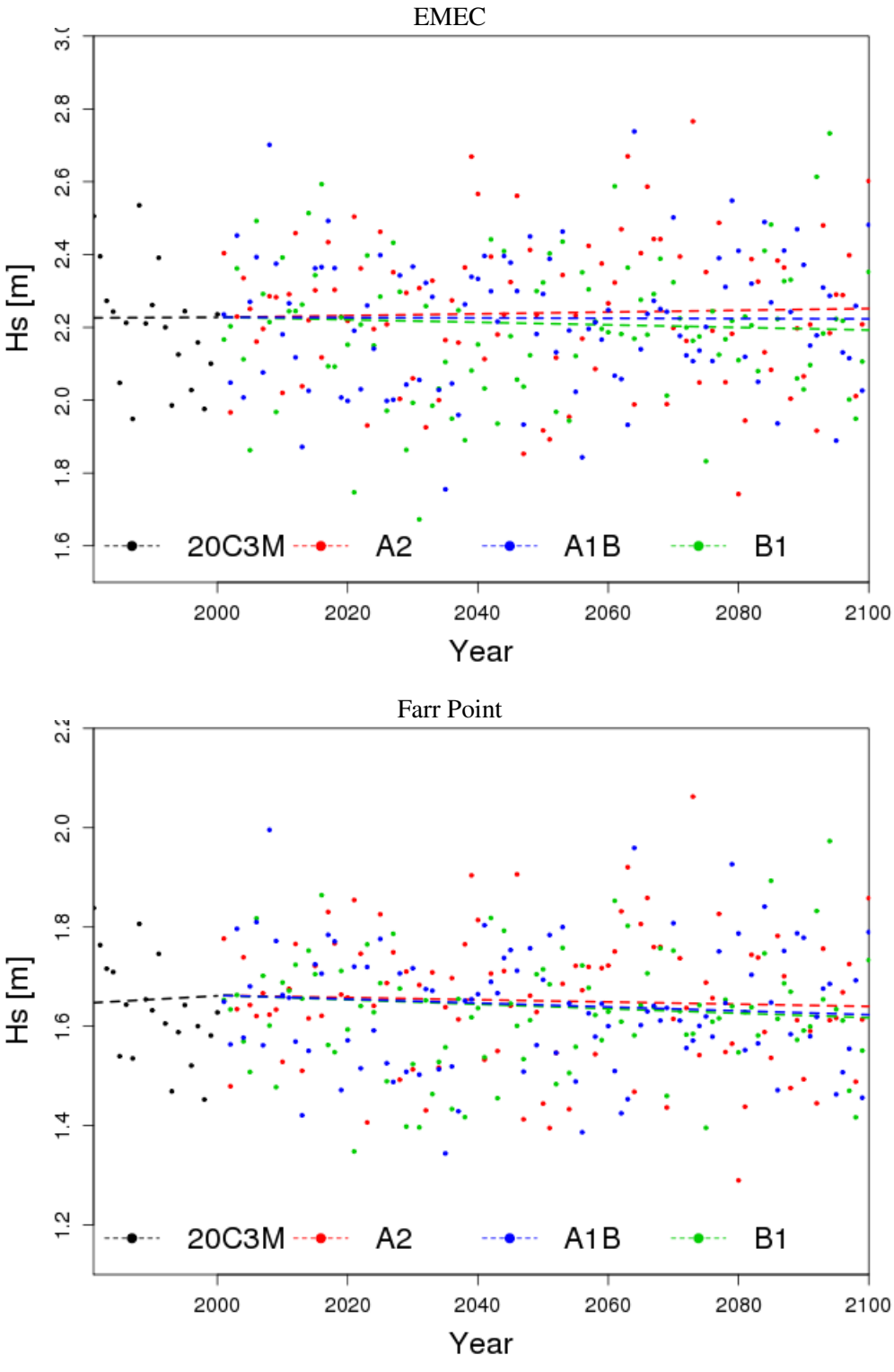

Figure C.22: Projections of annual mean H_s from the high (A2), medium (A1B) and low (B1) emissions scenarios for EMEC and Farr Point

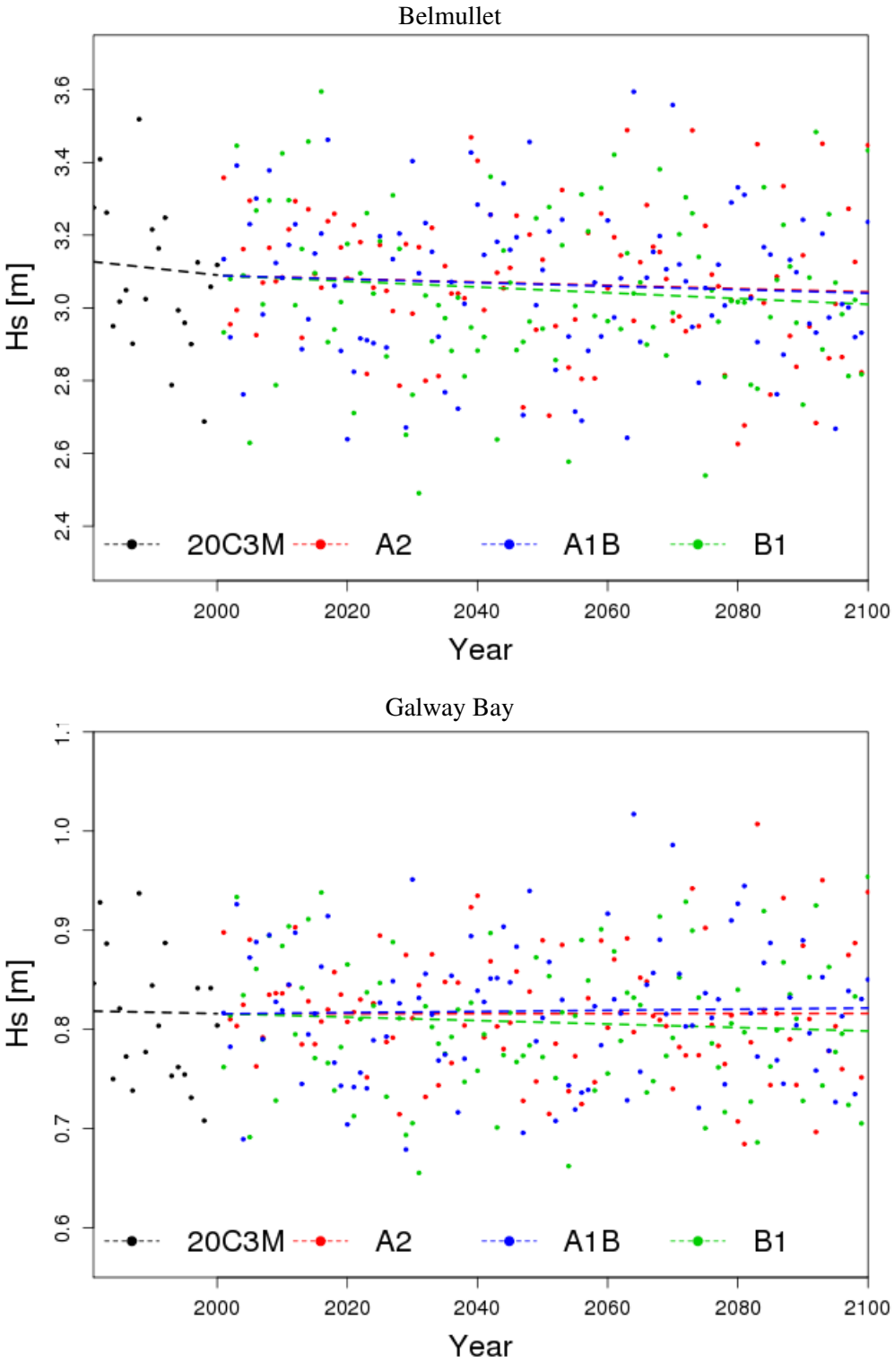

Figure C.23: Projections of annual mean H_s from the high (A2), medium (A1B) and low (B1) emissions scenarios for Belmullet and Galway Bay

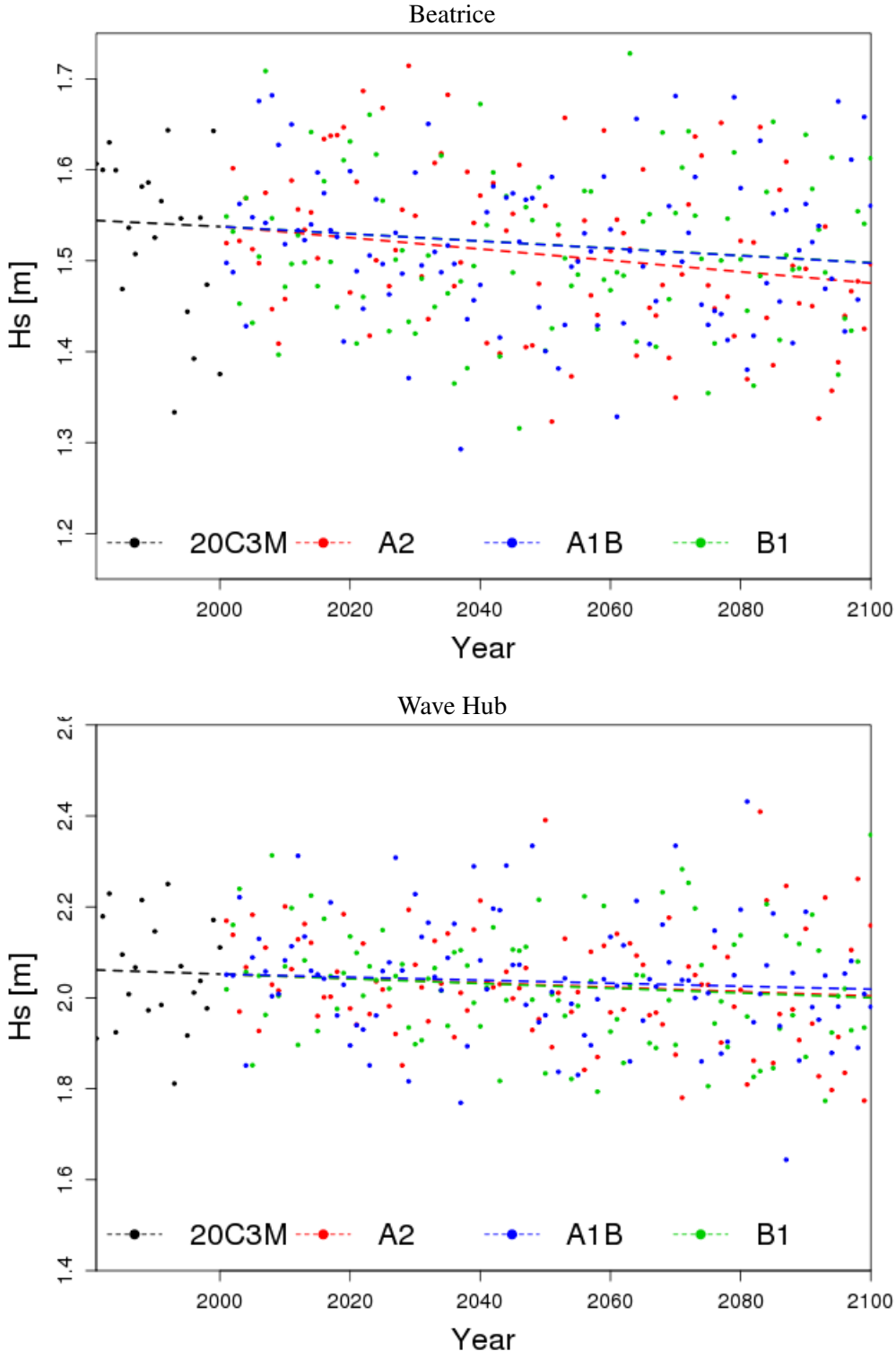

Figure C.24: Projections of annual mean H_s from the high (A2), medium (A1B) and low (B1) emissions scenarios for Beatrice and Wave Hub

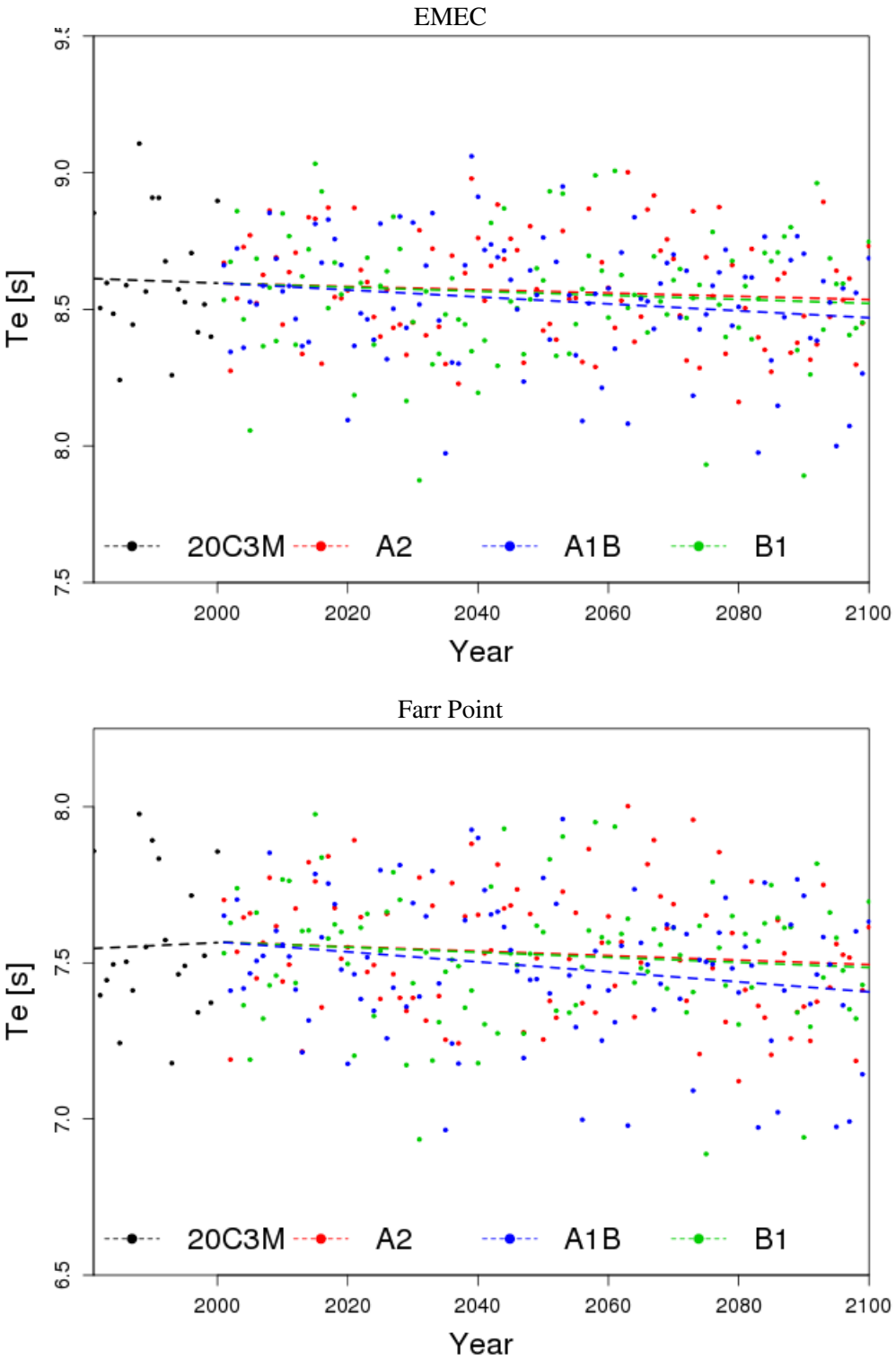

Figure C.25: Projections of annual mean T_e from the high (A2), medium (A1B) and low (B1) emissions scenarios for EMEC and Farr Point

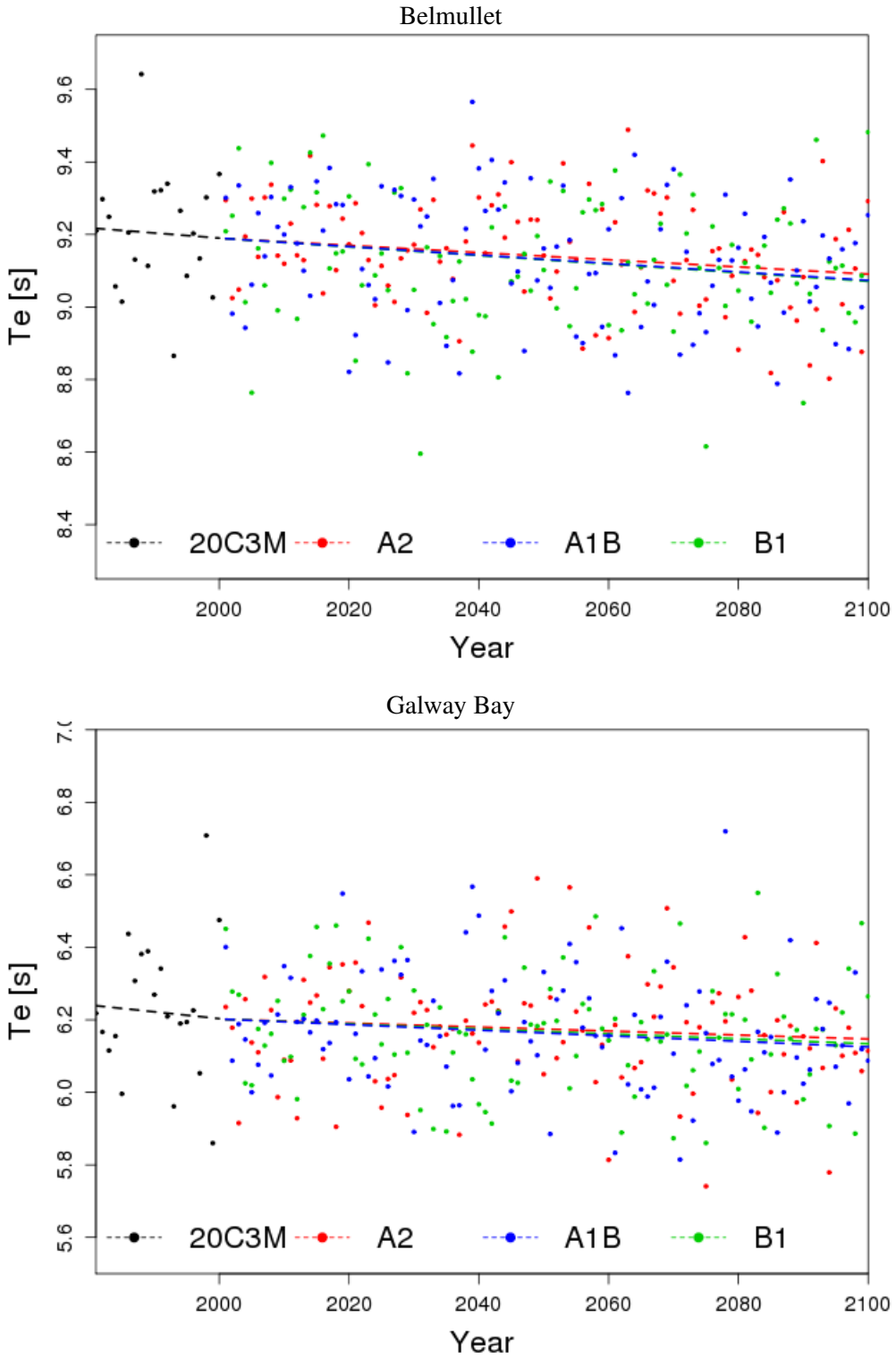

Figure C.26: Projections of annual mean T_e from the high (A2), medium (A1B) and low (B1) emissions scenarios for Belmullet and Galway Bay

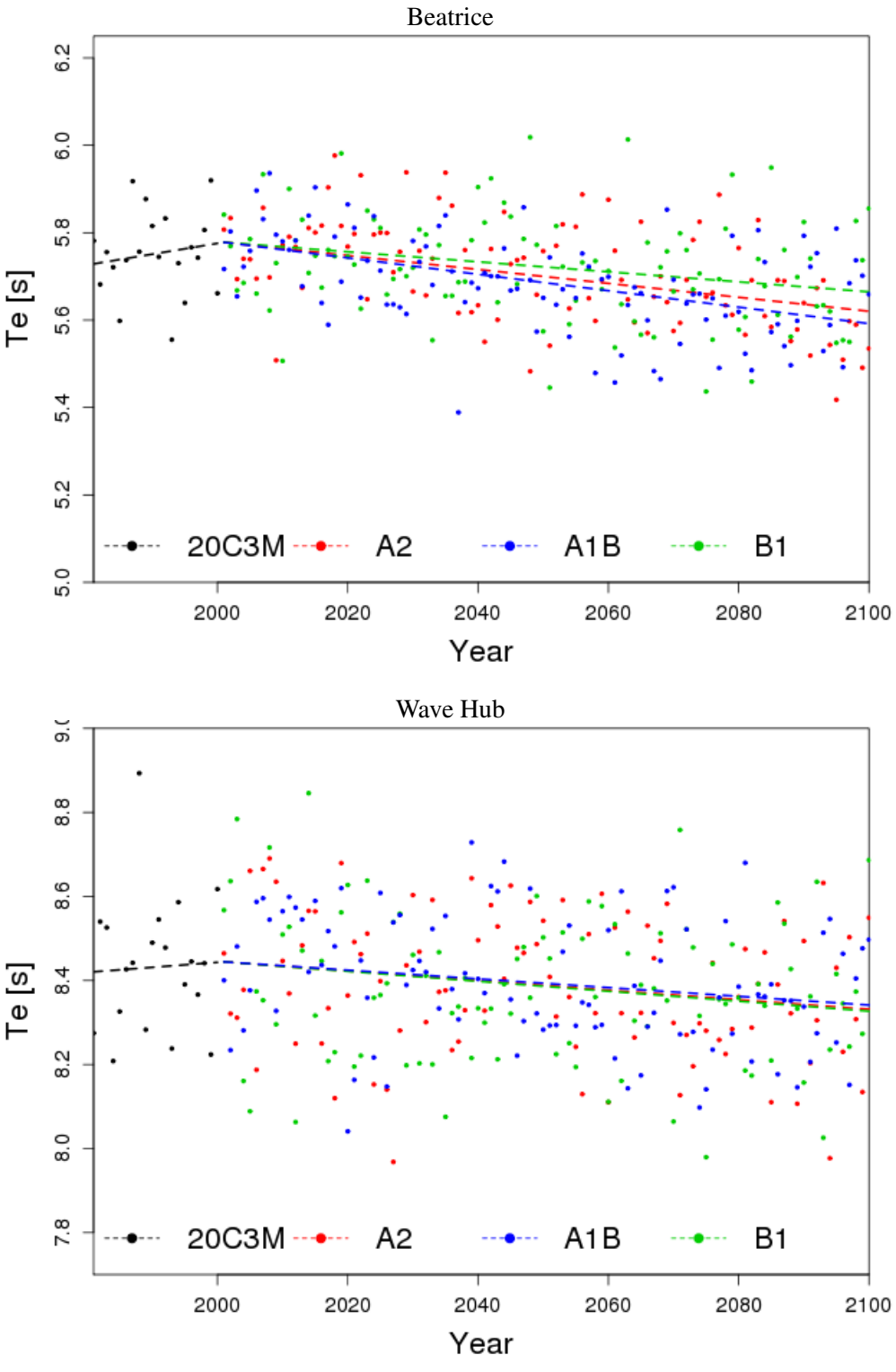

Figure C.27: Projections of annual mean T_e from the high (A2), medium (A1B) and low (B1) emissions scenarios for Beatrice and Wave Hub

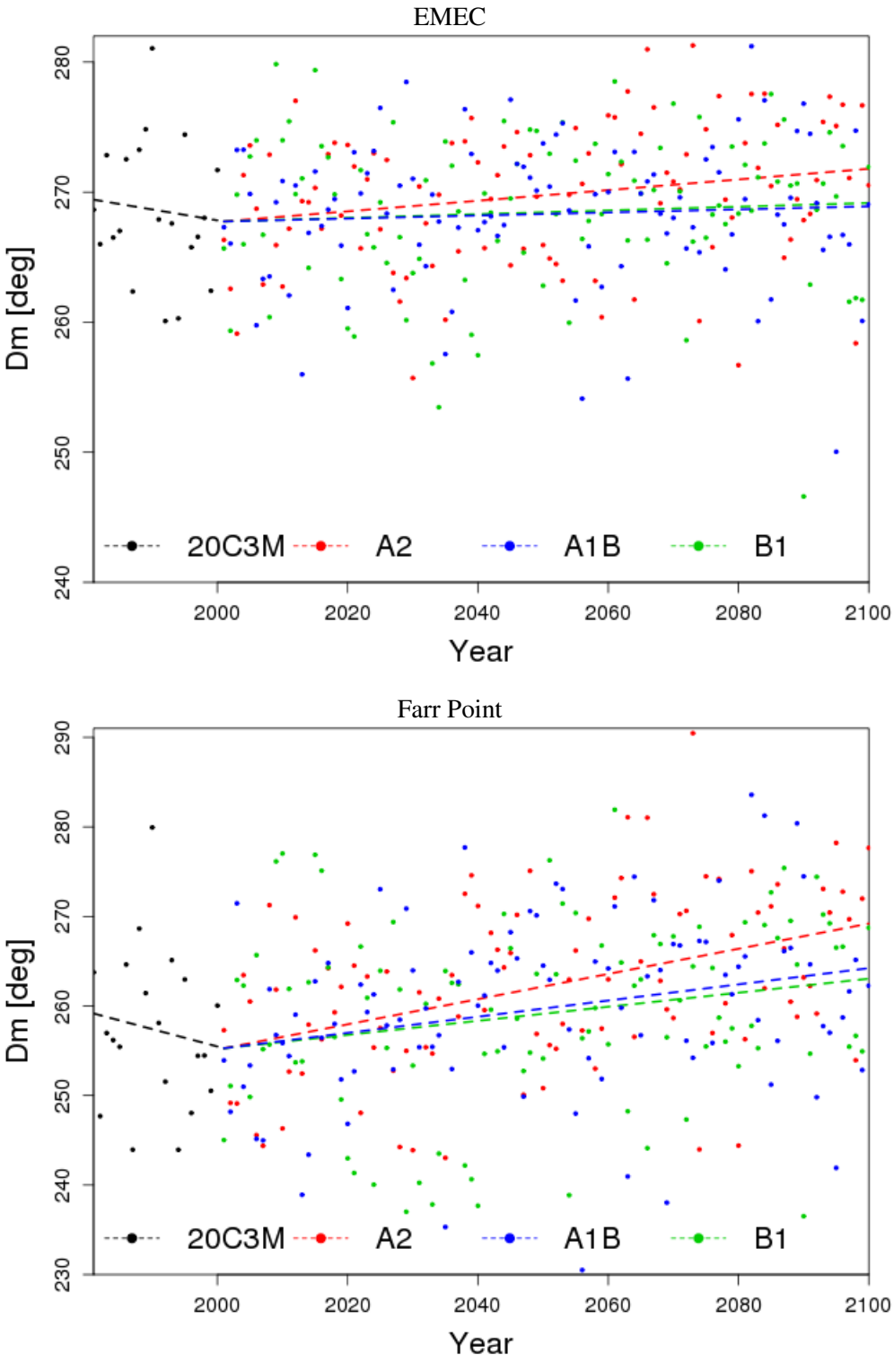

Figure C.28: Projections of annual mean D_m from the high (A2), medium (A1B) and low (B1) emissions scenarios for EMEC and Farr Point

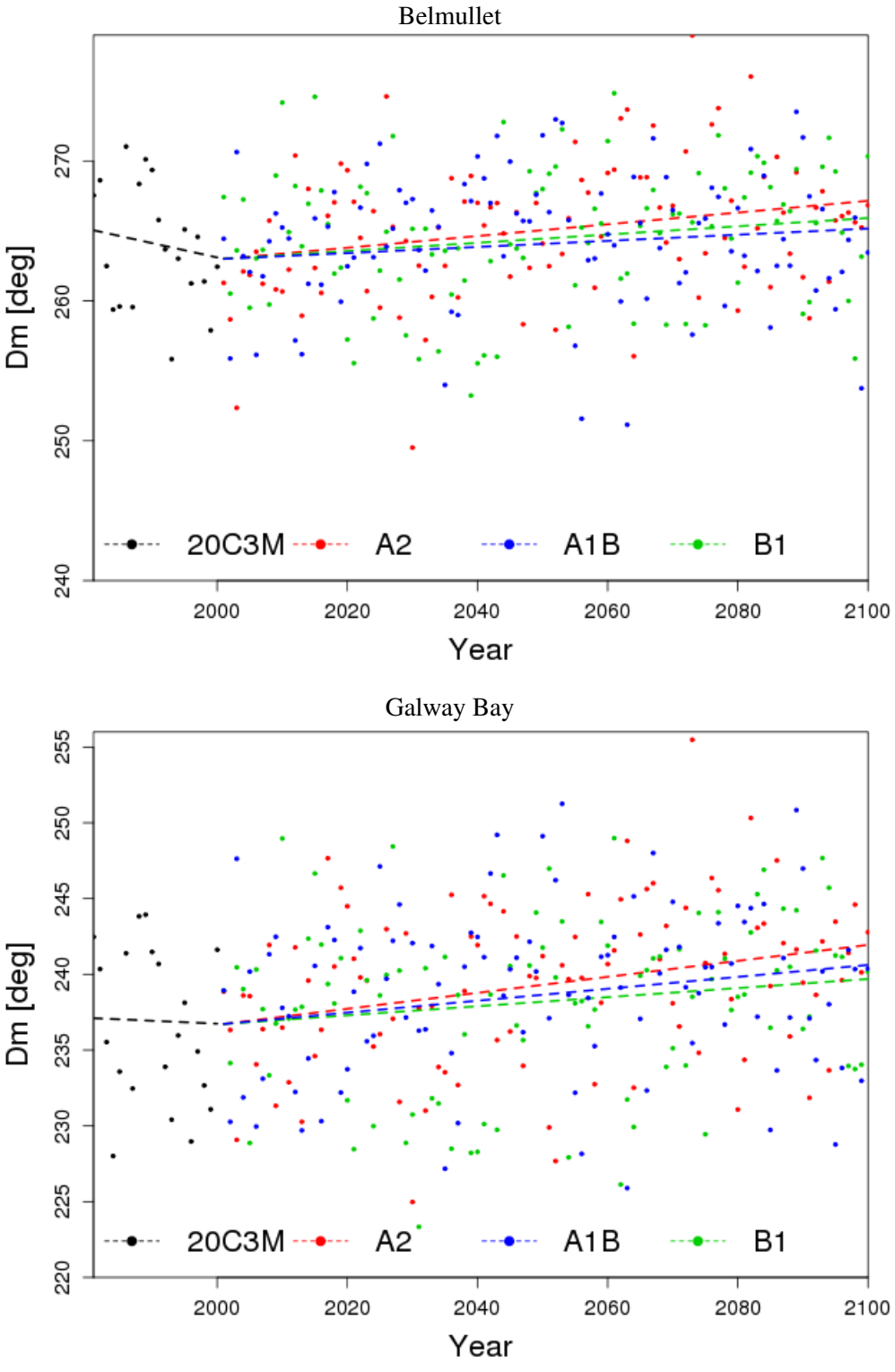

Figure C.29: Projections of annual mean D_m from the high (A2), medium (A1B) and low (B1) emissions scenarios for Belmullet and Galway Bay

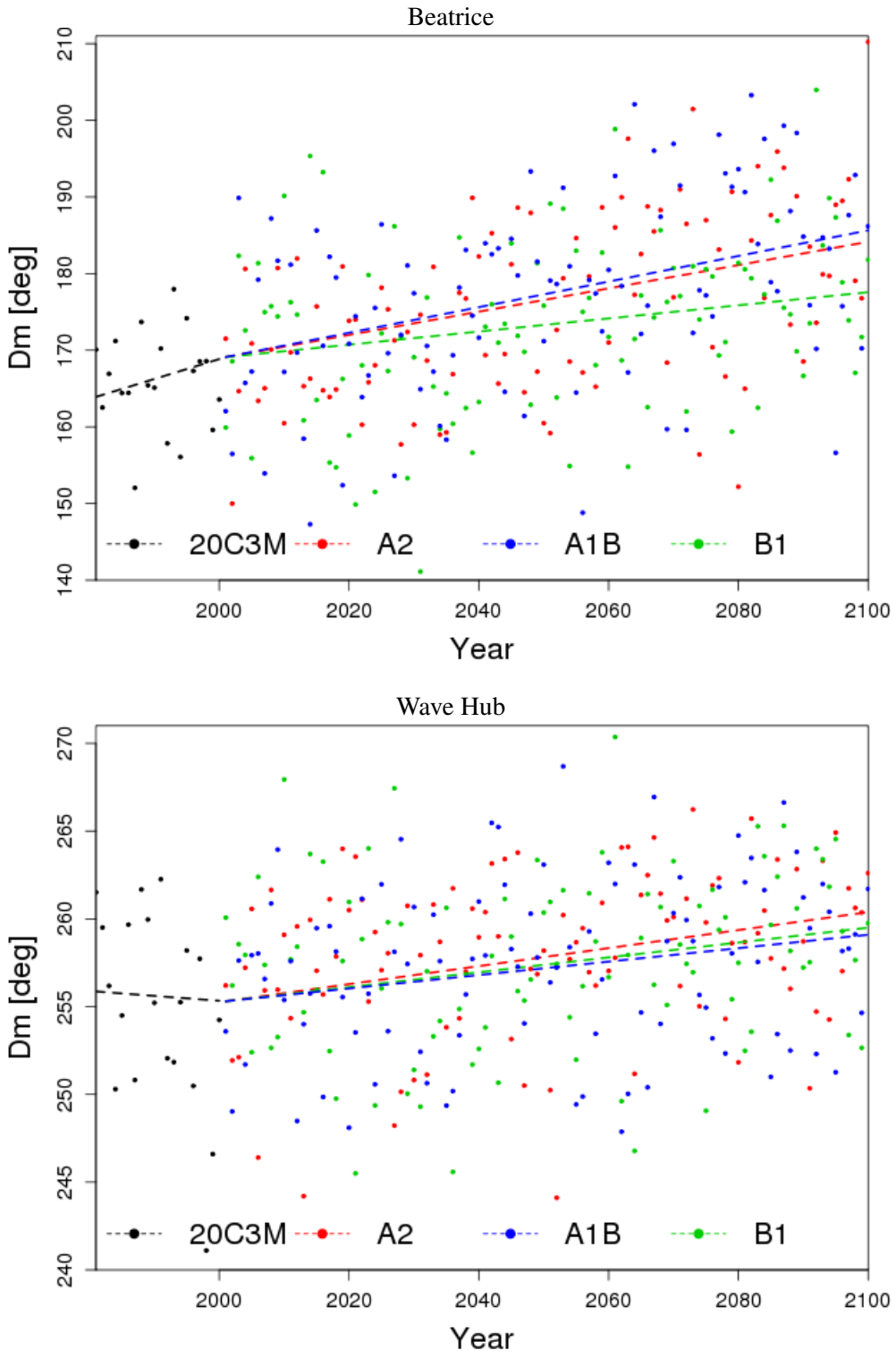

Figure C.30: Projections of annual mean D_m from the high (A2), medium (A1B) and low (B1) emissions scenarios for Beatrice and Wave Hub

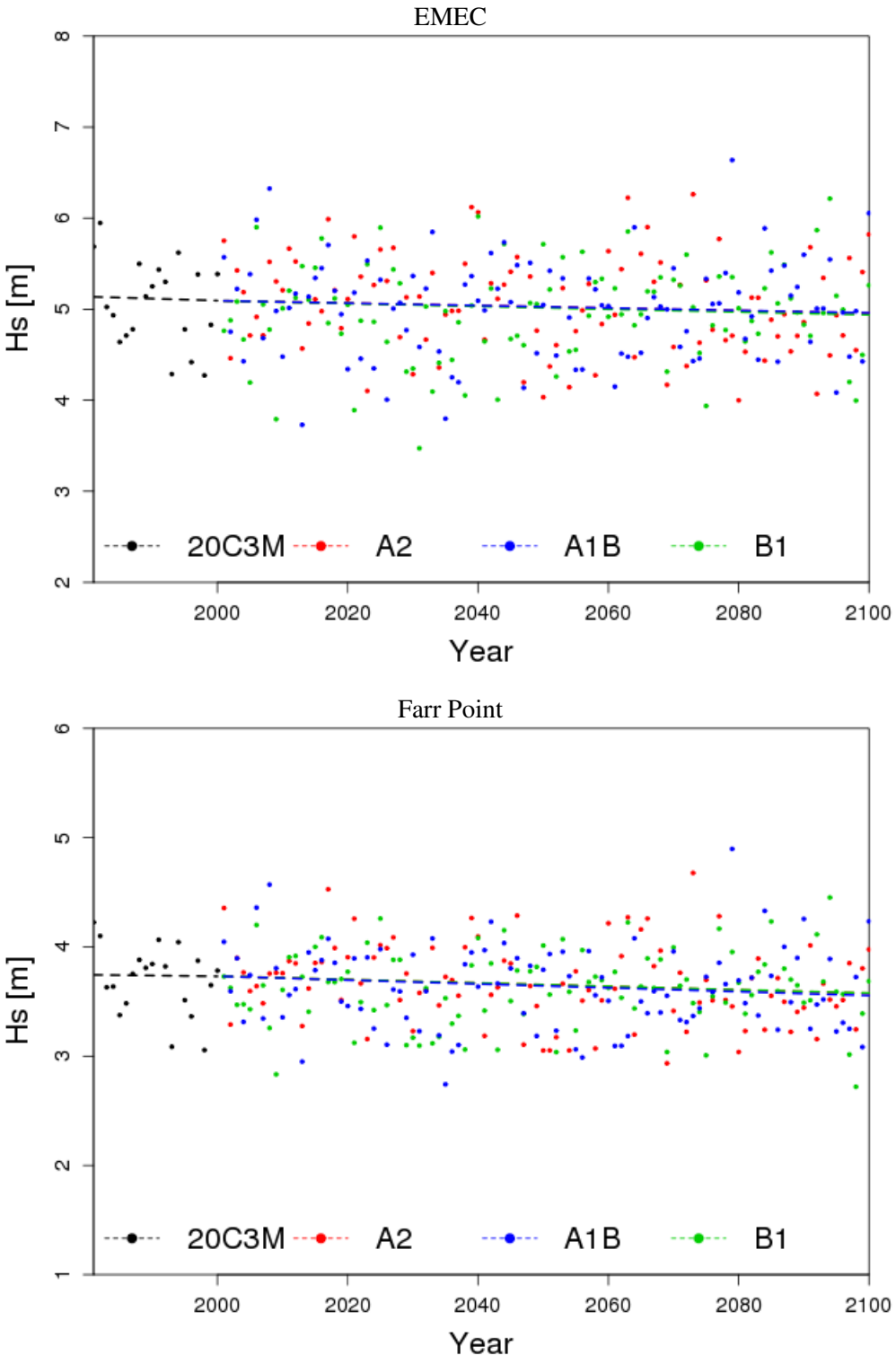

Figure C.31: Projections of annual 95th percentile H_s from the high (A2), medium (A1B) and low (B1) emissions scenarios for EMEC and Farr Point

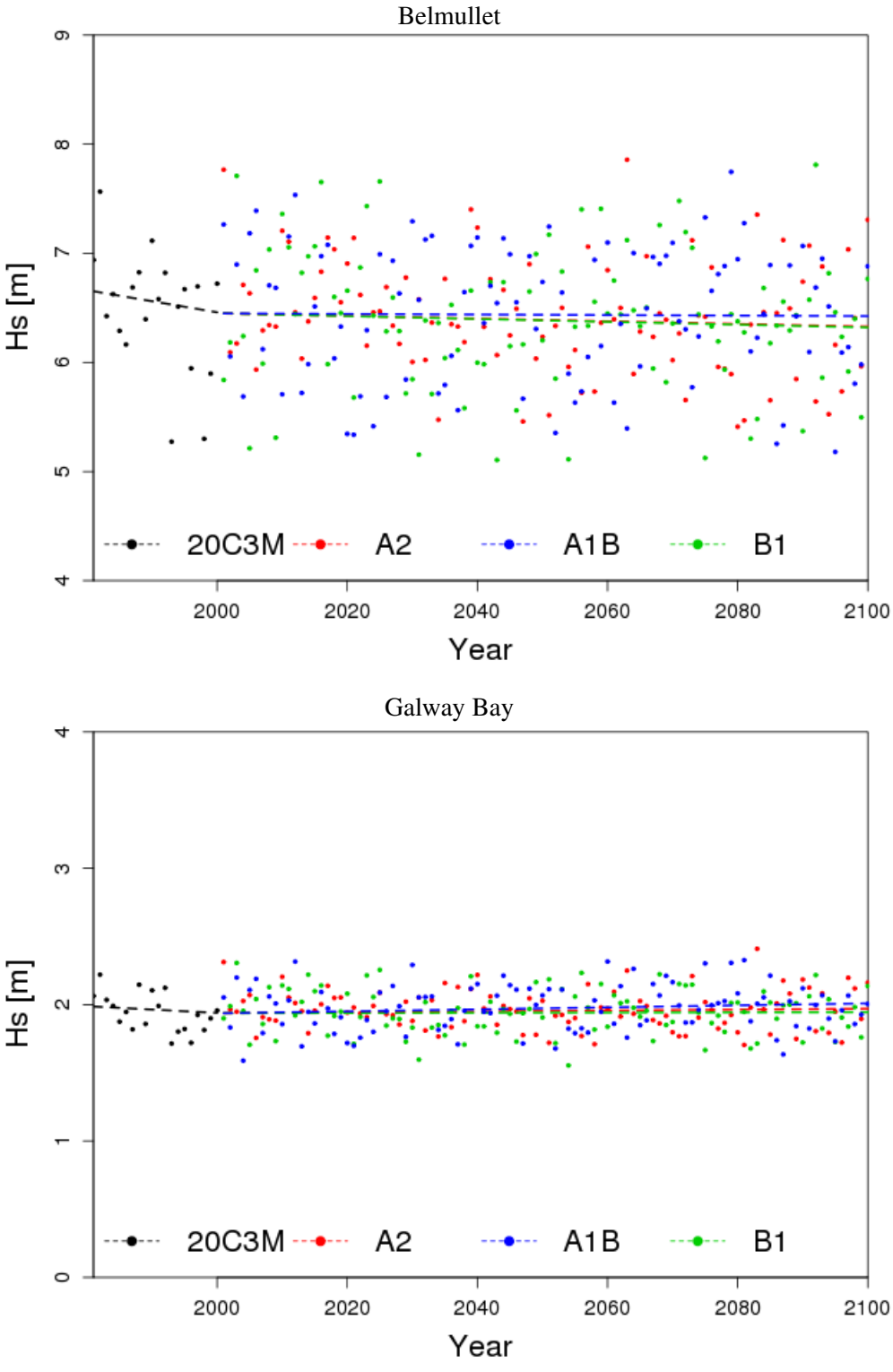

Figure C.32: Projections of annual 95th percentile H_s from the high (A2), medium (A1B) and low (B1) emissions scenarios for Belmullet and Galway Bay

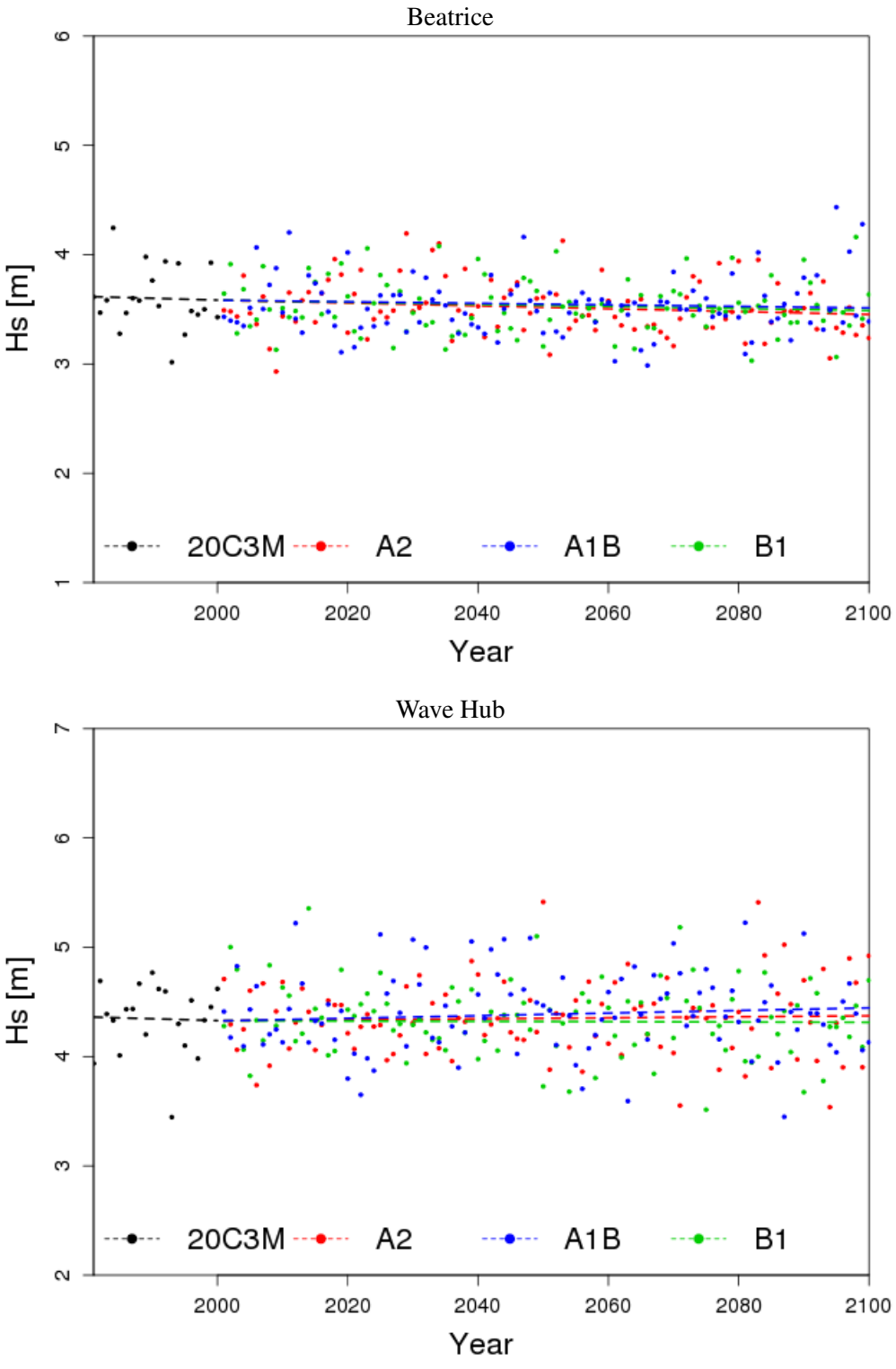

Figure C.33: Projections of annual 95th percentile H_s from the high (A2), medium (A1B) and low (B1) emissions scenarios for Beatrice and Wave Hub

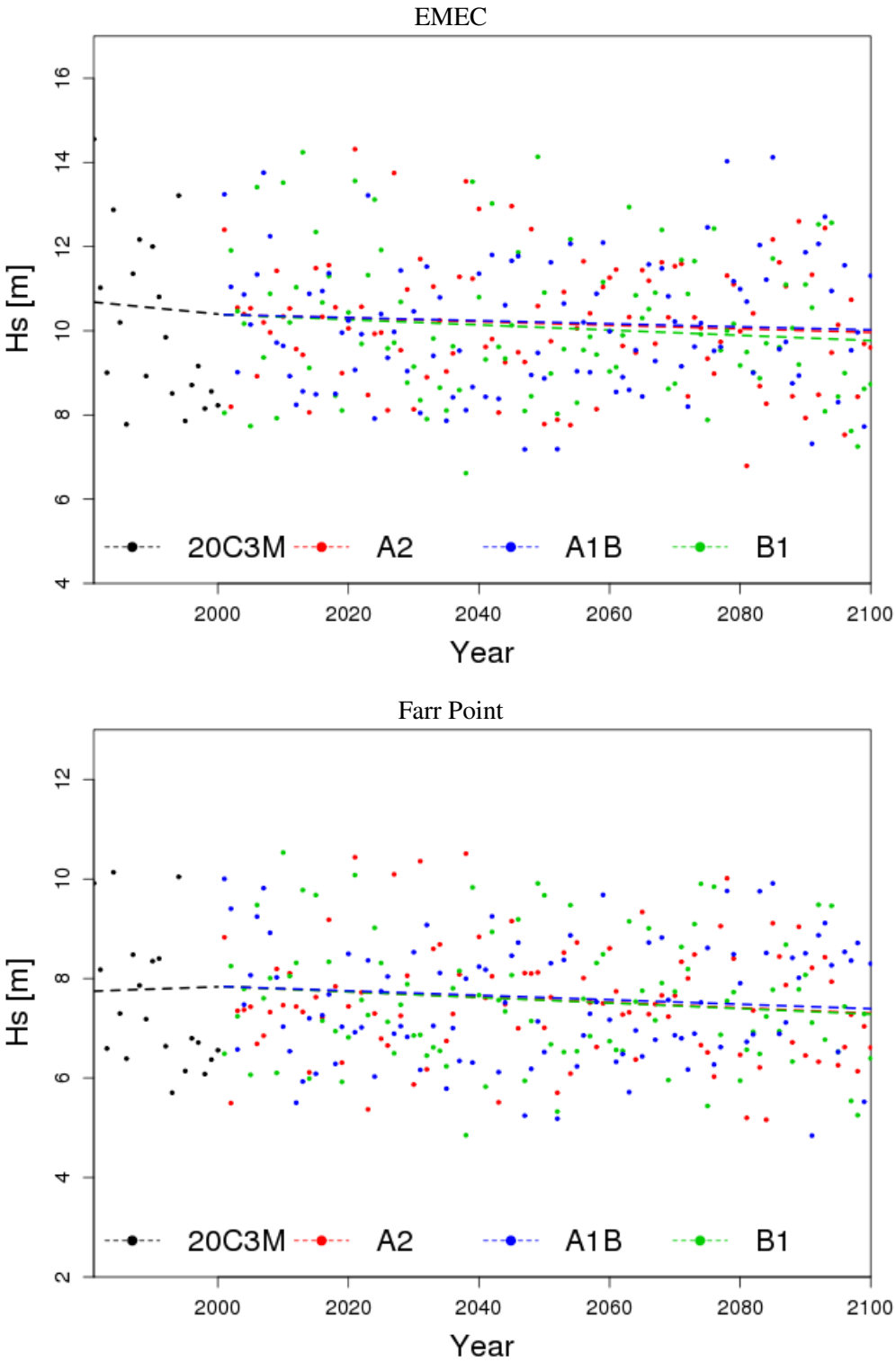

Figure C.34: Projections of annual maximum H_s from the high (A2), medium (A1B) and low (B1) emissions scenarios for EMEC and Farr Point

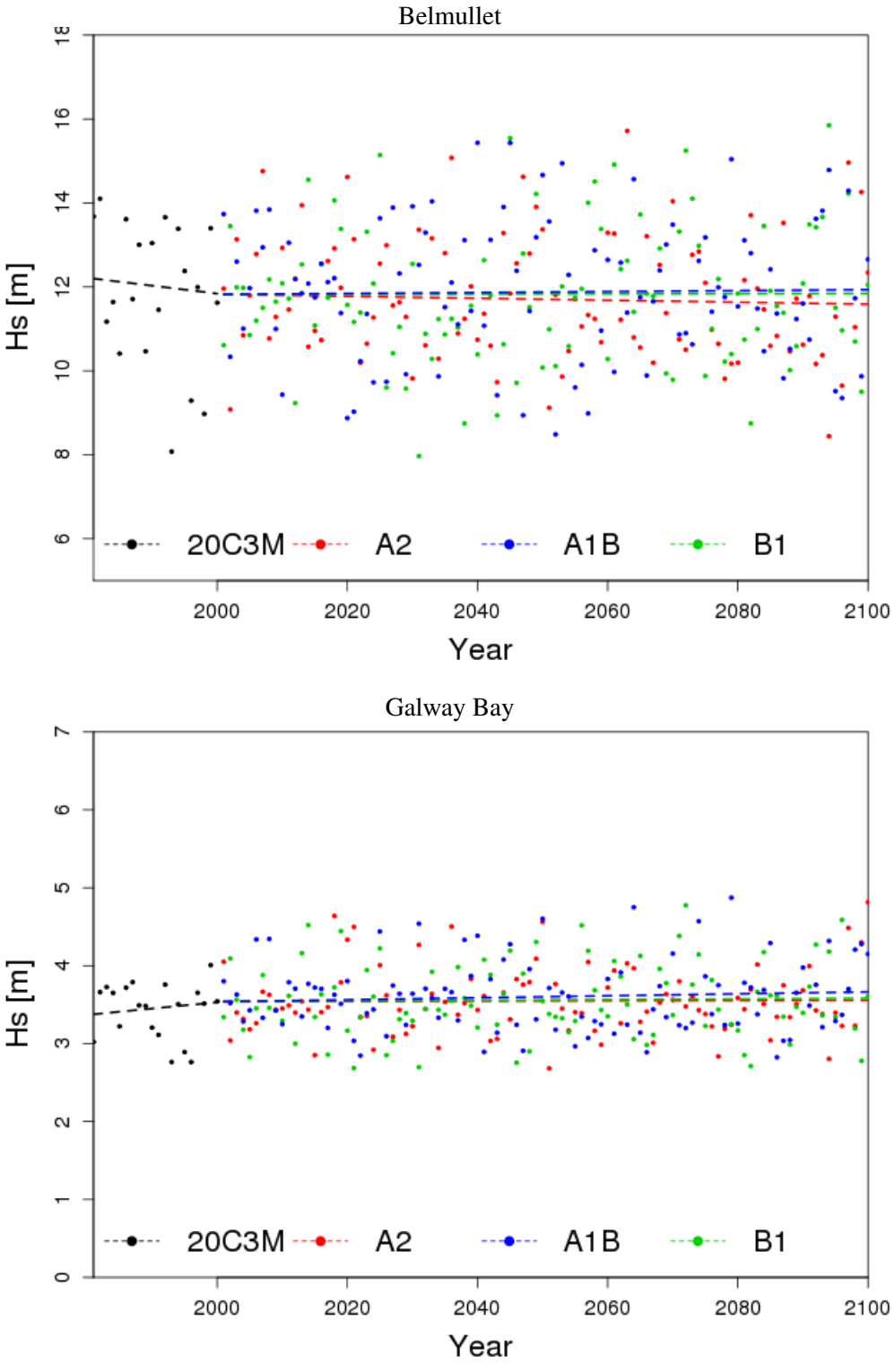

Figure C.35: Projections of annual maximum H_s from the high (A2), medium (A1B) and low (B1) emissions scenarios for Belmullet and Galway Bay

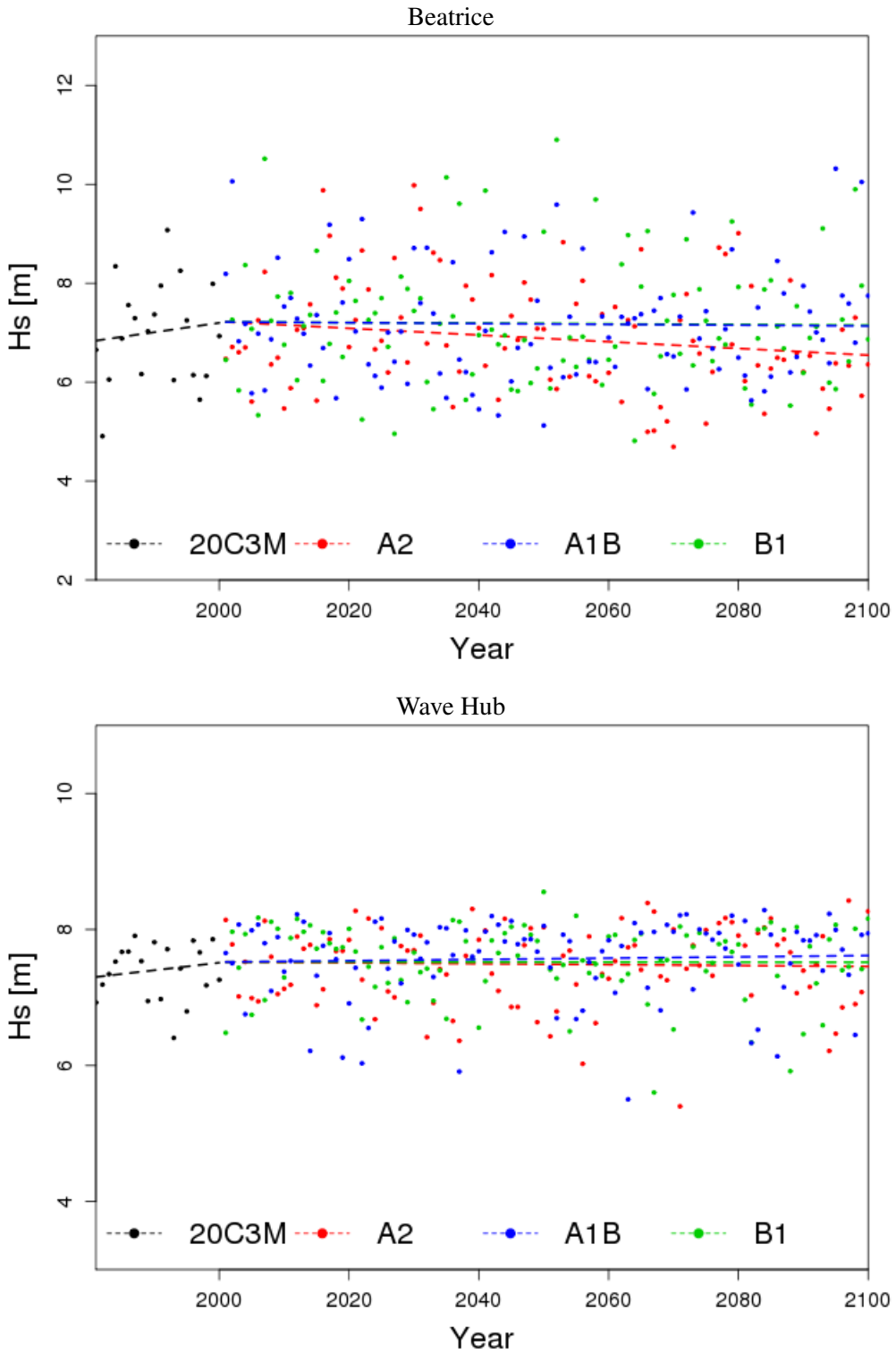

Figure C.36: Projections of annual maximum H_s from the high (A2), medium (A1B) and low (B1) emissions scenarios for Beatrice and Wave Hub