Copyright © and Moral Rights for this thesis are retained by the author and/or other copyright owners. 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 copyright holder/s. The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given e.g.

AUTHOR (year of submission) "Full thesis title", University of Southampton, name of the University School or Department, PhD Thesis, pagination

http://eprints.soton.ac.uk
BEYOND SIGNIFICANT WAVE HEIGHT: A NEW APPROACH FOR VALIDATING NUMERICAL WAVE MODELS

Edgar Peter Dabbi

A dissertation submitted in partial fulfillment of the degree of

MSc in Engineering in the Coastal Environment

by instructional course.

September 2013
Summary

Spectral wave data are required in many engineering applications such as the design of coastal defence and offshore platforms. As such, numerical wave models have been developed to estimate these wave data at locations where observed records are not available. In the published literature, the model results are often validated using sea-state parameters like significant wave height, peak wave period, mean wave period and mean wave direction. However, in some cases these parameters are not sufficient to describe the entire wave spectrum. In theory, the sea-state values could have a good agreement while the wave spectrums diverge from each other. Therefore, the main aim of this research work is to develop a new, robust approach for validating wave models by applying new parameterisation to the frequency wave spectrum.

A series of parameters from wave mechanics and other disciplines have been reviewed to better define wave spectrums. These parameters are tested over a range of JONSWAP wave spectrum idealized scenarios to analyse their sensitivity and performance. The result shows that a family of seven parameters including significant wave height, peak frequency, peak energy density, squared Euclidean distance, skewness, kurtosis and mean width deviation are required to best describe the characteristic differences between an observed and predicted spectrum. Parallel analysis of the parameters reveals more qualitative information about the two spectrums, in contrast to the individual assessment of each parameter.

The feasibility of the new approach developed here has been proven through the validation of a hindcast spectral wave model at three nearshore sites around the UK. The result shows that the model performance varies in both the temporal and spatial domains. Two-dimensional validation matrices have also been applied to illustrate the relationship between the various parameters with the relative magnitude, shape and position of the wave spectrums.
Acknowledgement

I would like to express my utmost gratitude to my supervisor, Dr. Ivan Haigh for his generous support and excellent mentorship throughout the duration of this project. Also thanks go to the staff of Associated British Ports Marine Environmental Research Ltd (ABPmer), in particular Dr. David Lambkin and Jamie Hernon, my two external co-supervisors for their guidance and technical support. The project would not have been possible without the permission to utilize their hindcast spectral wave models. I would like to thank Travis Mason from the Channel Coastal Observatory as well for sharing her knowledge on spectral wave partitioning.

Moreover, I am grateful to the Commonwealth Scholarship Commission for providing me with the scholarship to undergo my postgraduate Masters degree at the University of Southampton, England. Lastly, my sincere thanks go to my beloved parents, grandmother, sisters and friends, especially Cal, for their endless support and motivation in the pursuit of my dreams. Cheers to everyone and God bless!
# Table of Contents

Summary ................................................................................................................................. i
Acknowledgement ....................................................................................................................... ii
Table of Contents ......................................................................................................................... iii
List of Figures ............................................................................................................................. v
List of Tables .............................................................................................................................. vii
List of Abbreviations and Symbols ............................................................................................. ix

1 Introduction ............................................................................................................................. 1
  1.1 Project Background .............................................................................................................. 1
  1.2 Aim and Objectives .............................................................................................................. 2
  1.3 Overview of Dissertation .................................................................................................. 3

2 Literature Review ...................................................................................................................... 5
  2.1 Concept of Frequency Wave Spectrum ........................................................................... 5
  2.2 Application of Wave Spectrum in Engineering Applications ..................................... 8
  2.3 Wave Measurements in UK Waters ............................................................................... 9
  2.4 Development of Numerical Wave Model ..................................................................... 10
  2.5 Methodology for Comparison of Frequency Wave Spectrums .................................... 11
    2.5.1 Fitting of JONSWAP Shape Parameters ................................................................. 12
    2.5.2 Spectral Width and Spectral Peakedness Parameter ............................................... 14
    2.5.3 Skewness and Kurtosis (Central Moments of Wave Spectrum) ......................... 14
    2.5.4 Measuring Distance between Two Spectrums ....................................................... 15
    2.5.5 Spectral Partitioning for Modal Identification ....................................................... 17
  2.6 Summary ............................................................................................................................ 19

3 Sensitivity Analysis on Parameters ......................................................................................... 20
  3.1 Methodology ..................................................................................................................... 20
    3.1.1 Wave Spectrum with Unimodal Distribution ......................................................... 20
    3.1.2 Wave Spectrum with Bimodal and Trimodal Distribution .................................... 22
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2</td>
<td>Result and Discussion ...............................................................27</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Wave Spectrum with Unimodal Distribution ...........................................27</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Wave Spectrum with Bimodal and Trimodal Distribution .............................38</td>
</tr>
<tr>
<td>3.2.3</td>
<td>Relationship between Validation Parameters ...........................................45</td>
</tr>
<tr>
<td>4</td>
<td>Mean Width Deviation ........................................................................50</td>
</tr>
<tr>
<td>4.1</td>
<td>Methodology ....................................................................................50</td>
</tr>
<tr>
<td>4.2</td>
<td>Result and Discussion ........................................................................53</td>
</tr>
<tr>
<td>5</td>
<td>Example of Application .......................................................................58</td>
</tr>
<tr>
<td>5.1</td>
<td>Methodology ......................................................................................58</td>
</tr>
<tr>
<td>5.1.1</td>
<td>Numerical Model Setup .........................................................................58</td>
</tr>
<tr>
<td>5.1.2</td>
<td>Spectral Data Collection ....................................................................60</td>
</tr>
<tr>
<td>5.1.3</td>
<td>Numerical Model Validation ................................................................61</td>
</tr>
<tr>
<td>5.2</td>
<td>Result and Discussion ..........................................................................61</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Time Series Analysis of Validation Parameters .......................................61</td>
</tr>
<tr>
<td>5.2.2</td>
<td>Statistical Performance of Primary Parameters ......................................67</td>
</tr>
<tr>
<td>5.2.3</td>
<td>Qualitative Assessment of Unimodal Spectrums .......................................69</td>
</tr>
<tr>
<td>5.2.4</td>
<td>Qualitative Assessment of Bimodal Spectrums ........................................79</td>
</tr>
<tr>
<td>6</td>
<td>Conclusions and Recommendations ........................................................80</td>
</tr>
<tr>
<td>References</td>
<td>............................................................................................................82</td>
</tr>
<tr>
<td>Appendix A1</td>
<td>.................................................................................................87</td>
</tr>
<tr>
<td>Appendix A2</td>
<td>.................................................................................................123</td>
</tr>
<tr>
<td>Appendix A3</td>
<td>.................................................................................................129</td>
</tr>
</tbody>
</table>
List of Figures

Figure 1: Definition of the JONSWAP shape parameters ..............................................6
Figure 2: Example of unimodal (12 a.m.), bimodal (2 a.m.) and trimodal (1 a.m.)
spectrums at West of Hebrides on 15th March 2009 (WaveNet 2013). ...............7
Figure 3: Comparison between JONSWAP and P-M spectrums with the JONSWAP
spectrums generated using $\gamma = 3.3$, $\sigma_a = 0.07$ and $\sigma_b = 0.09$ ...............13
Figure 4: The Kolmogorov-Smirnov distance (black dashed line) between two
cumulative spectral energy density curves .........................................................17
Figure 5: Spectral partition (red dashed line) between swell and sea components ....18
Figure 6: Sensitivity tests for Scenario 1 to Scenario 8 ........................................21
Figure 7: Sensitivity tests for Scenario 9 to Scenario 12 .........................................23
Figure 8: Flow diagram of the spectral partitioning method (forward scan) ..........25
Figure 9: Flow diagram of the spectral partitioning method (backward scan) .......26
Figure 10: Normalized spectral energy density versus frequency for baseline spectrum
and predicted spectrums in Scenario 4 to Scenario 6 ........................................29
Figure 11: Wave spectrums that have identical skewness and kurtosis .................31
Figure 12: Squared Euclidean distance and Kullback-Leibler divergence for Scenario
1 to Scenario 4 ......................................................................................................32
Figure 13: Squared Euclidean distance and Kullback-Leibler divergence for Scenario
5 to Scenario 8 ......................................................................................................33
Figure 14: Scaling of frequency wave spectrums for the Kolmogorov-Smirnov test .35
Figure 15: Cumulative energy and the Kolmogorov-Smirnov distance for Scenario 1
to Scenario 8 ......................................................................................................36
Figure 16: Functions of the key validation parameters ...........................................37
Figure 17: Example of partitioned wave spectrum in Scenario 9 .........................40
Figure 18: Two-dimensional validation matrix between $\Delta E_{\text{max}}$ and $\Delta H_s$ ....47
Figure 19: Two-dimensional validation matrix between $\Delta E_{\text{max}}$ and $\Delta f_p$ .......48
Figure 20: Concept of absolute width parameter ....................................................50
Figure 21: Flow diagram for solving the mean width deviation in MATLAB® 201252
Figure 22: Mean width deviation for Scenario 9 ...................................................54
Figure 23: Mean width deviation for Scenario 10 ....................................................54
Figure 24: Mean width deviation for Scenario 11 ....................................................55
Figure 25: Mean width deviation for Scenario 12 ....................................................55
Figure 26: Regional view of the model bathymetry in UTM-30 projection. The view is distorted towards the west of the North Atlantic Ocean because of the projection system ..........................................................................................58

Figure 27: Local view of the model bathymetry in UTM-30 projection around the UK Waters ........................................................................................................................................59

Figure 28: Location of the three nearshore sites on Google Earth ........................................60

Figure 29: Primary validation parameters at Hastings in 2009 ..............................................62

Figure 30: Secondary validation parameters at Hastings in 2009 ........................................63

Figure 31: Primary validation parameters at Hastings in November 2009 ..........................64

Figure 32: Secondary validation parameters at Hastings in November 2009 .......................65

Figure 33: Predicted versus observed significant wave heights at Hastings .........................67

Figure 34: Predicted versus observed peak frequencies at Hastings .....................................68

Figure 35: Predicted versus observed peak energy densities at Hastings ..............................68

Figure 36: Validation with modified matrix ($\Delta E_{\text{max}}/\text{Observed } E_{\text{max}}$ versus $\Delta H_s/\text{Observed } H_s$) for unimodal spectrums at Hastings .................................................................73

Figure 37: Validation with modified matrix ($\Delta E_{\text{max}}/\text{Observed } E_{\text{max}}$ versus $\Delta H_s/\text{Observed } H_s$) for unimodal spectrums at Moray Firth .................................................................74

Figure 38: Validation with modified matrix ($\Delta E_{\text{max}}/\text{Observed } E_{\text{max}}$ versus $\Delta H_s/\text{Observed } H_s$) for unimodal spectrums at West of Hebrides ..................................................74
List of Tables

Table 1: Description of sensitivity tests for unimodal distribution ................................................................. 22
Table 2: Significant wave height, peak frequency and peak energy density for baseline and predicted spectrums in scenario 1 to 8 (Figure 6) ................................................................. 27
Table 3: Spectral width parameter, spectral peakedness parameter, skewness and kurtosis for baseline and predicted spectrums in Scenario 1 to 8 (Figure 6) ............................................. 28
Table 4: Squared Euclidean distance, Kullback-Leibler divergence and Kolmogorov-Smirnov distance for Scenario 1 to 8 (Figure 6) ......................................................................................... 31
Table 5: Validation parameters for Scenario 9 to Scenario 12 (Figure 7) without spectral partitioning ................................................................................................................................. 38
Table 6: Number of modes and partitioning frequencies reported by the spectral partitioning algorithm .......................................................................................................................... 39
Table 7: Validation parameters for Scenario 9 with spectral partitioning ......................................................... 40
Table 8: Validation parameters for Scenario 10 (Figure 7) with spectral partitioning ........................................ 41
Table 9: Validation parameters for Scenario 11 (Figure 7) with spectral partitioning ........................................ 42
Table 10: Validation parameters for Scenario 12 (Figure 7) with spectral partitioning ........................................ 43
Table 11: Classification of primary, secondary and rejected validation parameters ........................................... 45
Table 12: Description of regions in the two-dimensional matrix between $\Delta E_{\text{max}}$ and $\Delta H_s$ .................................................................................................................................................. 47
Table 13: Description of regions in the two-dimensional matrix between $\Delta E_{\text{max}}$ and $\Delta f_p$ .................................................................................................................................................. 48
Table 14: Mean width and mean width deviation in Scenario 1 to Scenario 8 (Figure 6) ........................................ 53
Table 15: Example of two hypothetical spectrums with zero mean width deviation ........................................... 56
Table 16: Family of final validation parameters .................................................................................................. 57
Table 17: Sites for model validation and availability of spectral data ............................................................... 60
Table 18: Top 10 squared Euclidean distances at Hastings in November 2009 ................................................. 66
Table 19: Boundaries of region 1 to region 9 for two-dimensional matrix between $\Delta E_{\text{max}}$ and $\Delta H_s$ .................................................................................................................................................. 69
Table 20: Validation with $\Delta E_{\text{max}}$ versus $\Delta H_s$ matrix for unimodal spectrums at Hastings .................................................................................................................................................. 70
Table 21: Validation with $\Delta E_{max}$ versus $\Delta H_s$ matrix for unimodal spectrums at Moray Firth

Table 22: Validation with $\Delta E_{max}$ versus $\Delta H_s$ matrix for unimodal spectrums at West of Hebrides

Table 23: Boundaries of region 1 to region 9 for two-dimensional matrix between $\Delta E_{max}$ and $\Delta f_p$

Table 24: Validation with $\Delta E_{max}$ versus $\Delta f_p$ matrix for unimodal spectrums at Hastings

Table 25: Validation with $\Delta E_{max}$ versus $\Delta f_p$ matrix for unimodal spectrums at Moray Firth

Table 26: Validation with $\Delta E_{max}$ versus $\Delta f_p$ matrix for unimodal spectrums at West of Hebrides
# List of Abbreviations and Symbols

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Aw_x$</td>
<td>Absolute width at $x^{th}$ percent of a threshold level</td>
</tr>
<tr>
<td>$C_o$</td>
<td>Cumulative spectral energy density (observed)</td>
</tr>
<tr>
<td>$C_p$</td>
<td>Cumulative spectral energy density (predicted)</td>
</tr>
<tr>
<td>$D_{KL}$</td>
<td>Kullback-Leibler divergence</td>
</tr>
<tr>
<td>$D_{KS}$</td>
<td>Kolmogorov-Smirnov distance</td>
</tr>
<tr>
<td>$D_{SE}$</td>
<td>Squared Euclidean distance</td>
</tr>
<tr>
<td>$E$</td>
<td>Spectral energy density</td>
</tr>
<tr>
<td>$E_{max}$</td>
<td>Peak spectral energy density</td>
</tr>
<tr>
<td>$E_o$</td>
<td>Spectral energy density (observed)</td>
</tr>
<tr>
<td>$E_p$</td>
<td>Spectral energy density (predicted)</td>
</tr>
<tr>
<td>$f$</td>
<td>Spectral frequency</td>
</tr>
<tr>
<td>$f_m$</td>
<td>Mean spectral frequency</td>
</tr>
<tr>
<td>$f_p$</td>
<td>Peak spectral frequency</td>
</tr>
<tr>
<td>$f_{sd}$</td>
<td>Standard deviation of spectral frequency</td>
</tr>
<tr>
<td>$g$</td>
<td>Gravitational acceleration $(9.81 \text{ m}^2/\text{s})$</td>
</tr>
<tr>
<td>$H_s$</td>
<td>Significant wave height</td>
</tr>
<tr>
<td>$K$</td>
<td>Kurtosis</td>
</tr>
<tr>
<td>$m_n$</td>
<td>$n^{th}$ moment of the frequency wave spectrum</td>
</tr>
<tr>
<td>$Mw$</td>
<td>Mean width</td>
</tr>
<tr>
<td>$N$</td>
<td>Wave action density</td>
</tr>
<tr>
<td>$N_{aw}$</td>
<td>Number of absolute width measurements</td>
</tr>
<tr>
<td>$Q_p$</td>
<td>Spectral peakedness parameter</td>
</tr>
</tbody>
</table>
$R$ Pearson product-moment correlation

$Sk$ Skewness

$t$ Time

$\bar{T}_c$ Mean wave crest period

$T_m$ Mean wave period

$T_p$ Peak wave period

$\bar{T}_z$ Mean wave zero up-crossing period

$x$ Horizontal axis on Cartesian coordinate

$y$ Vertical axis on Cartesian coordinate

$\alpha$ Phillips constant (0.0081)

$\gamma, \sigma_a, \sigma_b$ JONSWAP shape parameters

$\Delta E_{max}$ Difference between observed and predicted peak energy densities

$\Delta f_p$ Difference between observed and predicted peak frequencies

$\Delta H_s$ Difference between observed and predicted significant wave heights

$\Delta Mw$ Mean width deviation

$\epsilon$ Spectral width parameter

$\theta$ Wave direction

$\theta_m$ Mean wave direction

$\bar{v}$ Propagation velocity of wave group

$\omega$ Angular frequency
1 Introduction

1.1 Project Background

Wind-generated waves are capable of inducing strong, destructive forces in the coastal and marine environment. To assess the influence of these wave loadings, wave observations are required in the design of coastal defences, platforms, pipelines and offshore wind farms (Goda 1985). These wave measurements are also needed to analyse the local sea state and estimate the operational downtime for ports, jetties, marine vessels, barges and dredgers. In short, the analysis of wave data is imperative in many engineering applications.

Historical wave records from ship-mounted instruments and waverider buoys are normally scarce in the offshore but are more widely available in the coastal areas. Most of the existing records only cover a short period, typically less than 10 years. Evaluation of design wave conditions for extreme storm events however require longer periods of wave measurement (e.g. more than 30 years).

Satellite altimeter has also been used to measure wave heights remotely in recent years. However, these data are not suitable for local studies because of their intermittent records and limited spatial resolution (Krogstad 1999). To solve these problems, numerical models have been developed to compute the spectral growth and transformation of swell waves and wind waves at sea. Spectral output can be extracted from the models at any specific location of interest, where measured wave records may not be available.

Consequently, there has been a surging demand for predicted spectral wave data in the oil and gas and energy industry. These simulated datasets can provide better input for the design of structural members that are subjected to repeated wave loadings. These datasets are also used to study the motion response of vessels at sea (Shaw 1999). Within the energy sector, the spectral energy distribution can provide valuable information on the amount of renewable energy yield that can be efficiently farmed from an offshore site.

In more recent years, numerical wave models have also been applied for multi-decadal analysis of wave climate using hindcast wind data from meteorological
For example, Weisse and Günther (2007) found a positive trend in the 99th percentile wave height at the southern North Sea region, between 1958 and early 1990s. In the North-East Atlantic Ocean, Dodet et al. (2010) found a significant increase in the significant wave height from 1953 to 2009. Bosserelle et al. (2012) also found a positive trend in the annual mean wave height between 1970 and 2009 in the southwest region of Western Australia.

The IPCC (2012) report has concluded that in general, there is low confidence in the future wave height projections with regards to effects of climate change. This is due to the small number of studies conducted, inconsistent wind projections between the wave models and limitations in their ability to simulate extreme wind conditions. Nonetheless, any future positive or negative trend changes in the wave climate would reflect changes of the wind conditions as well.

Typically, these wave models are validated against measured wave records to establish the accuracy and reliability of their results. Sea-state parameters like the significant wave height, peak wave period and mean wave direction are often used to show the validity of the model predictions (Holthuijsen 2007; Palmer 2011). However, within the context of validation, these values are insufficient to describe the entire wave spectrum especially for multimodal wave distributions. In theory, it is possible for the sea-state parameters to have good agreement, even when the observed and predicted spectral shapes diverge from one another. A misrepresentation of the wave spectrums could lead to an engineering failure, with the potential to result in huge financial loss and threat to human lives.

Hence, there is a growing need for a more thorough validation of wave models to be developed, in which the comparison between wave spectrums is better defined, to ensure the predictions are well representative of reality.

1.2 Aim and Objectives

Based on the above, the overarching aim of this research work is to develop a robust approach for validating numerical wave models by applying new parameterisation to the frequency wave spectrums. To achieve this aim, the following four key objectives were defined:
• To review existing parameters for defining wave spectrums in wave mechanics and continuous distributions in other related disciplines;
• To assess the performance of the parameters, both individually and collectively for describing the characteristic differences between two wave spectrums;
• To develop a new additional parameter, if required, for a better representation of the features in wave spectrums; and
• To demonstrate the application of the new approach through the validation of a real wave model output.

The desired outcome of this research is to provide a better framework for both modellers and researchers to help understand the physical wave processes that are being well or poorly modelled. More importantly, it would bring significant benefits to coastal, marine and offshore engineers who are the primary end users of these wave datasets, mainly for engineering design purposes.

The scope of the study covers the analysis of frequency wave spectrums only. Directional wave spectrums are excluded because of their higher level of complexity. A brief description on their differences is presented in the literature review.

1.3 Overview of Dissertation

The main body of this dissertation is divided into six chapters. The four following chapters address each of the four study objectives in turn.

Chapter 2 begins with a literature review on the concept of frequency wave spectrums and their related engineering applications. Following that, the history of wave records around the UK Waters and the development of numerical wave models are presented. The main section of Chapter 2 is the literature review on the existing methods and parameters that could be potentially incorporated into the new validation approach. A brief summary is then given at the end of the chapter.

In Chapter 3, a series of sensitivity tests have been carried out to evaluate the performance and robustness of the reviewed parameters on ideal JONSWAP wave spectrums, in order to fulfill the requirement of the second objective. Based on the results, the parameters are classified into primary, secondary or rejected parameters. Two-dimensional validation matrices have also been developed to illustrate the
relationship between the primary parameters with the magnitude, shape and position of the wave spectrums.

Chapter 4 is a follow-up to the findings from the previous chapter, in which the author develops a new additional parameter, known as the mean width deviation, to complement the other parameters and fulfill the third objective. In Chapter 5, the application of the new approach is demonstrated through validation of a hindcast wave model at three nearshore sites around the UK Waters.

Lastly, Chapter 6 presents the conclusion and further recommendations for future studies.
2 Literature Review

2.1 Concept of Frequency Wave Spectrum

Wind-generated waves are usually described by a wave spectrum that defines the energy density in the frequency or direction domain. A frequency wave spectrum shows the energy distribution with respect to the frequency only, regardless of the wave direction. On the other hand, a directional wave spectrum provides the energy distribution for both frequency and direction (Goda 1985). A wave spectrum in principle describes the stochastic process occurring on the sea surface under the prevailing conditions at the time of observation (Holthuijsen 2007).

From the early 1960s, mathematical models of the frequency wave spectrum have been developed. Pierson and Moskowitz (1964) suggested the following formula for fully-developed wind waves, now known as the P-M model:

\[ E_{PM}(f) = \alpha g^2 (2\pi)^{-4} f^{-5} \exp\left[ -\frac{5}{4} \left( \frac{f}{f_p} \right)^{-4} \right] \]  \hspace{1cm} (2.1)

Where \( \alpha \) is the Phillips constant (equal to 0.0081), \( g \) is gravitational acceleration, \( f \) is frequency and \( f_p \) is the frequency corresponding to the peak energy density (hereafter referred to as peak frequency). Subsequently, following a wave observation program in the North Sea, Hasselmann et al. (1973) proposed the JONSWAP model:

\[ E_{JONSWAP}(f) = \alpha g^2 (2\pi)^{-4} f^{-5} \exp\left[ -\frac{5}{4} \left( \frac{f}{f_p} \right)^{-4} \right] \exp\left( -\frac{(f-f_p)^2}{2\sigma^2 f_p^2} \right) \]  \hspace{1cm} (2.2)

where \( \sigma \{ \begin{cases} a & \text{for } f \leq f_p \\ b & \text{for } f > f_p \end{cases} \) \)

The original P-M model is essentially modified into Equation 2.2 with the inclusion of the peak enhancement factor:

\[ \frac{\exp\left( -\frac{(f-f_m)^2}{2\sigma^2 f_p^2} \right)}{\gamma} \]

Where \( \gamma \) is the ratio of the peak energy density between the JONSWAP and P-M model while \( \sigma_a \) and \( \sigma_b \) are the left and right-sided widths of the spectral peak.
Hasselmann et al. (1973) also evaluated the average values of the three shape parameters at the North Sea, as follows:

\[ \gamma = 3.3 \quad \text{varies between 1 and 7} \]

\[ \sigma_a = 0.07 \quad \text{considered fix} \]

\[ \sigma_b = 0.09 \quad \text{considered fix} \]

![Figure 1: Definition of the JONSWAP shape parameters](image)

Other forms of the frequency wave spectrum have been previously proposed by Bretschneider (1969), Liu (1971), Mitsuyasu (1972) and Ochi and Hubble (1976). Nonetheless, the P-M spectrum has been regarded as one of the most representative for waters around the world. As such, it has been extensively applied in offshore engineering problems. The JONSWAP model however is more popular for studies in the North Sea. Design storm waves for extreme events are often represented by the peakier JONSWAP wave spectrum (Chakrabarti 1987).

Actual wave spectrums can be quite different from these standard forms, especially in the presence of both swell and sea waves (Goda 1985). Swell are waves
that are generated outside of the local fetch and have travelled for long distances while sea waves are waves that are generated by the local wind (Kamphuis 2000). Consequently, a secondary or tertiary peak can be observed in the spectrum, where the position of the peaks correspond to the dominant frequency of the swell or wind waves (Goda 1985).

Palmer (2011) found three different kinds of sea states in the English Channel that result in unimodal or multimodal frequency wave spectrum:

- **Unimodal**: Single peak in the spectrum, associated with the existence of swell or wind sea only;
- **Bimodal**: Two distinct peaks in the spectrum, usually related to the existence of both swell waves and one local wind sea;
- **Trimodal**: Three distinct peaks in the spectrum. Due to a sudden change in the wind direction, a new “growing” wind sea could happen at the same instance as the older “dying” wind sea, resulting in two different peaks. Presence of swell waves would then form a third peak in the spectrum.

![West of Hebrides (57.292° N, 7.914° E)](image_url)

Figure 2: Example of unimodal (12 a.m.), bimodal (2 a.m.) and trimodal (1 a.m.) spectrums at West of Hebrides on 15th March 2009 (WaveNet 2013).
2.2 Application of Wave Spectrum in Engineering Applications

Sea-state parameters are typically derived from the wave spectrum for various coastal engineering applications (Palmer 2011). These include the significant wave height, $H_s$; peak wave period, $T_p$; mean wave period, $T_m$ and mean wave direction, $\theta_m$ (Holthuijsen 2007; Palmer 2011).

The concept of significant wave height was introduced by Sverdrup and Munk (1947). It is defined as the mean height of the highest one-third of all observed waves within a record. It can also be approximated from the zeroth moment of the wave spectrum (Chakrabarti 1987). Generally, the $n^{th}$ moment of the frequency wave spectrum is defined as:

$$m_n = \int_0^\infty f^n E(f) df$$  \hspace{1cm} (2.3)

The significant wave height is calculated with the following formula:

$$H_s = 4\sqrt{m_0}$$  \hspace{1cm} (2.4)

Where $m_0$ is the zeroth moment, equivalent to the cumulative area under the wave spectrum (e.g. Chakrabarti 1987).

However, in some cases these parameters are insufficient to describe the entire wave spectrum, especially in the presence of both energetic swell waves and sea waves that are manifested in a multimodal spectrum. For these cases, the effect of swell waves would tend to be overlooked in the design of coastal flood defence. Long period swell waves (e.g. 15-20 seconds) result in higher run-up than the shorter period wind sea (Palmer 2011). Hence, it is important to analyse the spectral data as a whole, rather than relying on the integrated parameters only.

The safe operations of pipelines, platforms, marine vessels, dredgers, and barges are highly dependent on local wave conditions (Hardisty 1990). Using spectral wave data, the amount of wave energy at specific frequencies and directions can be directly extracted. This is useful for designing, constructing and operating structural members that are both safer and more economical in the deep sea. Moreover, motion response algorithms can be coupled with spectral wave forecast to predict the roll, pitch and heave of ships and assess the likely hazards (Shaw 1999).
2.3 Wave Measurements in UK Waters

This section provides a brief overview of wave measurements in UK Waters, including the WaveNet Nearshore Wave Recording Network commissioned under the Centre for Environment, Fisheries and Aquaculture Science (Cefas) and the UK Met Office.

According to Hardisty (1980), the first electronic records of coastal waves were made in Cornwall in 1947 (Darbyshire 1962) while the first offshore measurements were made in the mid-1950’s at Morecambe Bay Light Vessel (Draper 1968). In shallow coastal waters, long-term measurements of wave height and wave period are acquired through sensors which are mounted on the seabed or fixed structures close to the shoreline. However, the same method is not feasible for deeper waters where the effects of wave orbitals are greatly reduced with depth. Instead, shipborne recorders and waverider buoys are more common for measuring wave conditions in the offshore (Hardisty 1980).

The shipborne recorder system consists of a pressure transducer and a vertical accelerometer which are both closely-mounted on the ship hull. The pressure transducer measures the water depth above the pressure port. To compensate for the ship motion, the accelerometer also determines the location of the port relative to still water level. Variation between the two signals gives the “true wave height”. A waverider buoy works by recording its own acceleration on the sea surface. The surface profile of the waves is then computed from the double integral of the acceleration measurements (Hardisty 1980).

More recently, satellite radar altimeter and Synthetic Aperture Radar (SAR) have been used to approximate wave heights remotely. Unlike in-situ wave buoys that are able to capture the sea-state parameters at regular, short-period intervals (e.g. every hourly or three hourly), satellite measurements are usually more intermittent. However, as satellite records have a larger spatial coverage, they could be used to provide wave statistics at regional and global scales (Krogstad 1999). Woolf et al. (2003) for example investigated the wave climate seasonality around the British Isles in the late 20th century using data from two satellites, TOPEX/Poseidon and ERS-1/ERS-2.
Historically, wave measurements in the UK were carried out on a “short-term ad hoc basis with limited periods in connection with specific projects” (Hawkes 2001). There was no strategic planning in the data collection and archiving. Therefore, the WaveNet project was implemented as a national network for continuous monitoring of wave conditions around the nearshore UK Waters (defined as 20-30 km offshore with depth of between 10-20 m) beginning from 2002 (Hawkes 2001; WaveNet Implementation Plan 2013). It was mainly driven by the increasing demand for reliable wave data in coastal defence studies, flood forecasting, shoreline management plans and climate change monitoring. Numerical wave models could be validated using real-time data collected from the network of both new and existing waverider buoys (Hawkes 2001).

2.4 Development of Numerical Wave Model

This section presents a brief description of the development of numerical wave models, followed by a general background on the DHI MIKE 21 Spectral Wave (SW) model which has been used in this research work.

In general, there have been three generations of numerical wave models. The 1st generation wave model was based on the spectral transport equation derived by Gelci et al. (1957). Non-linear wave interactions were not considered. The 2nd generation wave models estimate these interactions, but free evolution of the wave spectrum was restricted. This led to the formation of the Wave Modelling (WAM) Group that developed the first 3rd generation wave model known as the WAM model. It was able to integrate the spectral transport equations without prior assumptions on the wave spectrum form (The WAMDI Group 1988).

The MIKE 21 SW model is a more recent, commercial 3rd generation wave model that simulates the spectral growth, transformation and decay of both swell waves and wind sea (DHI Water & Environment 2007). It uses an Eulerian approach to solve the governing wave action balance equation (Komen et al. 1994; Young 1999) on an unstructured triangular mesh. The high degree of flexibility of the irregular mesh allows control of the grid resolution to the appropriate physical scales. Less computational effort is required compared to the traditional nested structured grid models (Sorensen et al. 2004). The governing equation for wave action is given as:
\[ \frac{\partial N}{\partial t} + \nabla \cdot (\mathbf{v} N) = \left( E_{in} + E_{nl} + E_{ds} + E_{bot} + E_{surf} \right) \frac{1}{\omega} \]  

(2.5)

Where \( N \) is the wave action density, \( t \) is time, \( \mathbf{v} \) is propagation velocity of the wave group and \( \nabla \) is the differential operator in the four phase spaces: the Cartesian coordinates, \( x \) and \( y \); the angular frequency, \( \omega \) and direction of wave propagation, \( \theta \). Spatial discretization is based on a “cell-centered finite volume” approach while integration in time is solved using an explicit Euler scheme (Sorensen et al. 2004).

Presently, in many engineering and oceanography studies, wave model results are validated with the sea-state parameters \( H_s, T_p, T_m \) or \( \theta_m \) (e.g. Weisse and Günther 2007; Dodet et al. 2010). However, within the context of model validation, these parameters are not adequate to describe the entire wave spectrum. In addition, they do not provide much information on how accurate the spectral shape is being modelled, aside from the total area below the spectrum (see Equation 2.4).

### 2.5 Methodology for Comparison of Frequency Wave Spectrums

To achieve the main aim of this project, the new approach should be able to compare the differences between two frequency wave spectrums in terms of their characteristics (e.g. magnitude and shape). It should be robust, in that the application should not be limited for specific spectral forms only. The output should provide a quick, easy and better understanding on how well the energy distribution has been predicted.

Perhaps the most basic yet effective method is a direct visual comparison between the observed and predicted spectrums on a same graph. This is feasible for comparing wave spectrums for a short simulation period (e.g. 24 hours wave model). However, over long period simulations (e.g. few months or years), there is a need to describe the frequency wave spectrum with a single parameter, or a family of parameters that best represent the changing characteristics in a time series plot.

Based on the above, this section explores a series of methodologies and parameters found in wave mechanics and other disciplines that could be potentially applied for wave spectrum validation.
2.5.1 Fitting of JONSWAP Shape Parameters

The first approach is based on the premise that given a set of sufficient data points on the wave spectrum, the JONSWAP shape parameters could be estimated from them. To ensure there are enough data points to resolve the spectral shape, a high resolution discretization is required for the frequency bins. Equation 2.2 can be rewritten as following:

\[ \frac{E_{\text{JONSWAP}}(f)}{E_{\text{PM}}(f)} = \gamma \exp \frac{-(f-f_p)^2}{2\sigma^2 f_p^2} \]  \hspace{1cm} (2.6)

\[ \log_{10} \left( \frac{E_{\text{JONSWAP}}(f)}{E_{\text{PM}}(f)} \right) = \log_{10} \left( \gamma \exp \frac{-(f-f_p)^2}{2\sigma^2 f_p^2} \right) \]

\[ \log_{10} \left( E_{\text{JONSWAP}}(f) \right) - \log_{10} \left( E_{\text{PM}}(f) \right) = \exp \frac{-(f-f_p)^2}{2\sigma^2 f_p^2} \cdot \log_{10} (\gamma) \]

\[ \frac{\log_{10} \left( E_{\text{JONSWAP}}(f) \right) - \log_{10} \left( E_{\text{PM}}(f) \right)}{\log_{10} (\gamma)} = \exp \frac{-(f-f_p)^2}{2\sigma^2 f_p^2} \]

Let \( \beta \) be equal to the expression on the left side of the equation. Hence:

\[ \beta = \frac{\log_{10} \left( E_{\text{JONSWAP}}(f) \right) - \log_{10} \left( E_{\text{PM}}(f) \right)}{\log_{10} (\gamma)} \]  \hspace{1cm} (2.7)

\[ \ln(\beta) = \ln \left( \exp \frac{-(f-f_p)^2}{2\sigma^2 f_p^2} \right) \]

\[ \ln(\beta) = \frac{-(f-f_p)^2}{2\sigma^2 f_p^2} \]

\[ \sigma = \sqrt{\frac{-(f-f_p)^2}{2\ln(\beta)f_p^2}} \]  \hspace{1cm} (2.8)

Mathematically, the shape parameter \( \sigma \) can be solved to determine the spectral peak widths with Equation 2.8. However, there are two key limitations to this:
1. \( \beta \) must be a positive value for Equation 2.8 to be solved. The JONSWAP energy density must be larger than the P-M energy density, given that \( \gamma \) is more than 1. As such, the equation can be solved for fully-developed wind sea only, which requires the observed and predicted wave spectrums to have larger energy densities than the P-M spectrum at all frequency bins.

2. \( \ln (\beta) \) must be negative for Equation 2.8 to be solved. This can be achieved with \( \beta \) falling between 0 and 1 only.

![Figure 3: Comparison between JONSWAP and P-M spectrums with the JONSWAP spectrums generated using \( \gamma = 3.3 \), \( \sigma_a = 0.07 \) and \( \sigma_b = 0.09 \)](image)

Another main limitation of the approach is relating \( \gamma \) with the physical scale of the wave spectrum. Two JONSWAP spectrums that have identical value of \( \gamma \) could be entirely different from each other (Figure 3). For two wave components having equal wave heights, the shorter period wave contains less spectral energy than the higher period wave. Therefore, the maximum attainable value for the peak energy density would gradually decrease, as the position of the peak frequency moves to the right side of the spectrum. It is difficult to make a meaningful interpretation of \( \gamma \) unless both the observed and predicted wave spectrums have the same peak frequency.
2.5.2 Spectral Width and Spectral Peakedness Parameter

The spectral width parameter, $\varepsilon$ was originally defined as the root mean square (RMS) width of the frequency wave spectrum, ranging from 0 to 1. As the parameter $\varepsilon$ approaches to 1, the spectrum is described to be “broad-banded”. On the other hand, the spectrum is described to be “narrow-banded” for value of $\varepsilon$ closer to 0 (Chakrabarti 1987). It is computed with the following equation:

$$
\varepsilon = \sqrt{1 - \left(\frac{T_c}{T_z}\right)^2}
$$

(2.9)

Where $T_c$ is the mean wave crest period and $T_z$ is the mean zero up-crossing period. Cartwright and Longuet-Higgins (1956) suggested the following relationship for computing the spectral width parameter from the moments of the frequency wave spectrum:

$$
\varepsilon = \sqrt{\frac{m_0m_4 - m_2^2}{m_0m_4}}
$$

(2.10)

Where $m_0$, $m_2$ and $m_4$ are the zeroth, second and fourth moment respectively. Due to the presence of the higher-order moments in Equation 2.10, noise in the higher frequency region of the spectrum would be amplified and this could severely affect the estimate of $\varepsilon$ (Chakrabarti 1987).

As an alternative to the spectral width parameter, Goda (1974) has proposed the spectral peakedness parameter, $Q_p$. It is defined as following:

$$
Q_p = \frac{2}{m_0^2} \int_0^\infty f E^2(f) df
$$

(2.11)

A higher value of the spectral peakedness parameter indicates a sharper, narrower peak and vice versa. Based on shipborne wave records on the coasts of India, Prasada Rao (1988) has found that estimates of $Q_p$ are more robust and independent of the wave spectral form, in comparison to $\varepsilon$.

2.5.3 Skewness and Kurtosis (Central Moments of Wave Spectrum)
In the study of geology and soil mechanics, Blott and Pye (2004) have used sample statistics like mean, standard deviation, skewness and kurtosis to analyse the particle size distribution of a given sample. The sample statistics were based on the central moments of the distribution (not to be confused with the term moment in the previous sections), where the skewness and kurtosis are equal to the standardized third and fourth central moment respectively. Both measures of skewness and kurtosis are location and scale independent (MacGillivray & Balanda 1988).

A similar approach can be adopted to quantify these statistics in a wave energy density distribution:

\[
Skewness, Sk = \frac{\sum f E(f)(f - f_m)^3}{f_{sd}^3}
\]  \hspace{1cm} (2.12)

\[
Kurtosis, K = \frac{\sum f E(f)(f - f_m)^4}{f_{sd}^4}
\]  \hspace{1cm} (2.13)

Where \(f_m\) is the mean spectral frequency and \(f_{sd}\) is the standard deviation.

For an ideal, unimodal fully-developed wind sea, the wave spectrum is expected to have a positive skew with short left tail and long right tail (e.g. Figure 3). Kurtosis can be interpreted as the “movement of probability mass from the shoulders of a distribution into its center and tails” (Balanda & MacGillivray 1988). Three general terms are usually used for describing kurtosis:

- **Platykurtic** for “flat-topped” distribution, where \(K < 3\)
- **Leptokurtic** for “sharply peaked” distribution where \(K > 3\)
- **Mesokurtic** for a normal distribution where \(K = 3\)

A platykurtic distribution would be synonymous to a broad-banded spectrum, where energetic waves are found over a broad range of frequencies. On the other hand, a leptokurtic distribution would indicate a narrow-banded spectrum, where most of the spectral energy is focused on a single frequency bin. However, it is false to assume that an increase in kurtosis would always imply a peakier distribution (MacGillivray & Balanda 1988).

2.5.4 Measuring Distance between Two Spectrums
In acoustics and signal processing, Helén and Virtanen (2007) investigated the similarity between two audio samples by calculating the squared Euclidean distance \( D_{SE} \) between their probability density functions. The same method can be used to quantify the distance between two wave spectrums:

\[
D_{SE} = \int_0^\infty [E_o(f) - E_p(f)]^2 df \quad \text{(2.14)}
\]

Where \( E_o \) and \( E_p \) are the spectral energy density of the observed and predicted wave data respectively.

Another measure of distance that would be further tested is the Kullback-Leibler divergence (Kullback & Leibler 1951) which has its origins in the field of information theory. It is a non-symmetric measure of distance between \( p \) and \( q \), where \( p \) is the “real” probability distribution function while \( q \) is the “approximating” probability distribution model. The Kullback-Leibler divergence denotes the amount of lost information when \( q \) is used to predict \( p \) (Burnham & Anderson 2002), as presented below:

\[
D_{KL}(p,q) = \int_0^\infty \ln \left( \frac{p(x)}{q(x)} \right) p(x) dx \quad \text{(2.15)}
\]

Equation 2.15 can then be modified to determine the Kullback-Leibler divergence between an observed and predicted wave spectrum:

\[
D_{KL}(E_o,E_s) = \int_0^\infty \ln \left( \frac{E_o(f)}{E_p(f)} \right) E_o(f) df \quad \text{(2.16)}
\]

The natural logarithm of a float-type variable would grow exponentially to infinity as the variable approaches to zero. Hence, to ensure the stability of the mathematical model, the minimum threshold for the spectral energy density would be fixed at 0.001 m\(^2\)/Hz (any lower value would not be included in the calculation of Equation 2.16).

Lastly, the third measure of distance that would be evaluated is the Kolmogorov-Smirnov test for goodness of fit (Massey Jr 1951). Originally, the test computed the maximum distance between a cumulative frequency distribution function and the observed cumulative step-function. From this definition, the test can
be applied to estimate the maximum deviation between the cumulative spectral energy of the observed and predicted spectrums, in units of percentage:

\[ D_{KS} = \text{Max}(|C_o(f) - C_p(f)|) \quad (2.17) \]

Where \( C_o \) and \( C_p \) are the cumulative spectral energy of the observed and predicted spectrums respectively.

![Kolmogorov-Smirnov Test](image)

Figure 4: The Kolmogorov-Smirnov distance (black dashed line) between two cumulative spectral energy density curves

Other advanced measures of distance between spectral densities include the Itakuro-Saito distance, Likelihood Ratio distance and COSH distance (e.g. Wei & Gibson 2001). However, these are not covered in the research work, mainly because of the difficulty in making physical sense from their values.

### 2.5.5 Spectral Partitioning for Modal Identification

The concept of spectral partitioning was initially suggested by Gerling (1992). It was then developed by Hasselmann et al. (1994), Voorrips et al. (1997) and Hanson and Phillips (2001) to assess how well a numerical wave model is able to produce the different wave components of a spectrum.
Hanson and Phillips (2001) applied a “steepest ascent” algorithm for partitioning the directional wave spectrums. However, because the WaveNet (2013) archive provides data on the frequency wave spectrum only, the main focus of this study is on validating models through the frequency wave spectrum.

To the knowledge of the author, the first published work on spectral partitioning with the frequency wave spectrum was carried out by Mason et al. (2008). Conditions for the spectral partitioning algorithm were derived from a coastal engineering perspective, where the goal was to partition a wave spectrum with “well-defined peaks” in both swell and sea components, given that they have sufficient energy to initiate sediment transport.

The conditions were found to be arbitrary but should be applicable at any location (Mason T. 2013, personal communication). A bimodal sea state is defined with the following criteria:

- Minimum total energy of the wave spectrum is equal to $H_s = 0.5$ m
- The smaller peak energy is at least one-third of the larger peak energy
- The smaller peak energy is at least $0.4 \text{ m}^2/\text{Hz}$ (equal to $H_s \sim 0.2$ m at 10 s)
- Energy at the minimum trough is less than half of the smaller peak energy
However, Palmer (2011) found two main limitations of the algorithm. First, it was not able to separate a trimodal spectrum because more than one partition would be required. Second, as wind speed and wind direction change, the wind sea would tend to overlap in the frequency domain, resulting in less-defined peaks in the spectrum. In this study, a new partitioning algorithm is tested to detect the number of modes in a spectrum. It is able to identify up to three modes in the spectrum.

2.6 Summary

Numerical wave models are often validated using sea-state parameters such as $H_s$, $T_p$, $T_m$ and $\theta_m$. However, in some cases these sea-state parameters are insufficient to describe the entire wave spectrum. Since 2002, the WaveNet project under Cefas and the UK Met Office has provided real-time data, including the observed frequency wave spectrum for continuous monitoring of wave conditions around the UK Waters (Hawkes 2001; WaveNet Implementation Plan 2013). The availability of these datasets provides opportunity for wave models to be better validated through the magnitude and shape of the frequency wave spectrum itself.

As such, a list of possible methods and parameters has been reviewed for comparison between an observed and predicted wave spectrum, including:

(a) Fitting of JONSWAP shape parameters: $\gamma$ and $\sigma$
(b) Spectral width parameter, $\varepsilon$ and spectral peakedness parameter, $Q_p$
(c) Skewness, $Sk$ from third central moment of wave spectrum
(d) Kurtosis, $K$ from fourth central moment of wave spectrum
(e) Squared Euclidean distance, $D_{SE}$
(f) Kullback-Leibler divergence, $D_{KL}$
(g) Kolmogorov-Smirnov distance, $D_{KS}$
(h) Spectral partitioning for modal identification

Each of these has their own advantages and disadvantages. Mathematically, the first method, involving the fitting of JONSWAP shape parameters, is very restricted in its application in comparison to the other methods. Therefore, only parameters (b) to (g) and method (h) are considered suitable to be brought forward to the next section of this research work – where sensitivity tests are carried out to gauge their performance and robustness.
3 Sensitivity Analysis on Parameters

3.1 Methodology

3.1.1 Wave Spectrum with Unimodal Distribution

The following parameters (hereafter referred to as validation parameters) are investigated over a range of scenarios to analyse their sensitivity with respect to the variation in magnitude and shape of two unimodal frequency wave spectrums:

(a) Significant wave height, $H_s$ (Equation 2.4)
(b) Spectral width parameter, $\varepsilon$ (Equation 2.10)
(c) Spectral peakedness parameter, $Q_p$ (Equation 2.11)
(d) Skewness, $Sk$ from third central moment of wave spectrum (Equation 2.12)
(e) Kurtosis, $K$ from fourth central moment of wave spectrum (Equation 2.13)
(f) Squared Euclidean distance, $D_{SE}$ (Equation 2.14)
(g) Kullback-Leibler divergence, $D_{KL}$ (Equation 2.16)
(h) Kolmogorov-Smirnov distance, $D_{KS}$ (Equation 2.17)

In addition to the above, the author proposes the inclusion of the peak frequency, $f_p$ and the peak energy density, $E_{max}$ parameters into the model validation approach (see section 2.1). Due to their basic mathematical definitions, a direct comparison on the spectral maxima can be made with these two parameters.

A baseline spectrum is generated using the JONSWAP model (Equation 2.2) with the values of $f_p = 0.167$ Hz, $\gamma = 3.3$, $\sigma_a = 0.07$ and $\sigma_b = 0.09$. It is then compared with eight other artificial spectrums, representing the likely scenarios that could be encountered between an observed and a predicted wave spectrum. All the spectrums are discretized with a frequency interval of 0.01 Hz.

For each case, the baseline spectrum is denoted by a blue line while the predicted spectrum is denoted by a red line (Figure 6). In reality, the wave spectrum would be noisier, characterized by abrupt spikes and troughs in the frequency domain. Nonetheless, the smooth spectral curves have been used in the analysis to reduce the uncertainties. The features of the wave spectrum are systematically changed one at a time to determine their influence on the validation parameters.
Figure 6: Sensitivity tests for Scenario 1 to Scenario 8
Table 1: Description of sensitivity tests for unimodal distribution

<table>
<thead>
<tr>
<th>Scenario</th>
<th>JONSWAP Parameters</th>
<th>Notes on Predicted Spectrum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$f_p$ (Hz)</td>
<td>$\gamma$</td>
</tr>
<tr>
<td>1</td>
<td>0.167</td>
<td>3.3</td>
</tr>
<tr>
<td>2</td>
<td>0.167</td>
<td>3.3</td>
</tr>
<tr>
<td>3</td>
<td>0.167</td>
<td>3.3</td>
</tr>
<tr>
<td>4</td>
<td>0.167</td>
<td>3.3</td>
</tr>
<tr>
<td>5</td>
<td>0.167</td>
<td>2.0</td>
</tr>
<tr>
<td>6</td>
<td>0.167</td>
<td>2.0</td>
</tr>
<tr>
<td>7</td>
<td>0.143</td>
<td>2.0</td>
</tr>
<tr>
<td>8</td>
<td>0.200</td>
<td>6.0</td>
</tr>
</tbody>
</table>

Physically, all the wave spectrums are realistic except for the Scenario 3 predicted spectrum. The upper limit for the peak energy density should be lower for shorter period waves. However, the main purpose of this scenario is to demonstrate the effect of skewness on the validation parameters, rather than representing the true scale of the wave spectrum. At lower energy regime, a negative skew distribution could possibly develop (e.g. “young” growing wind waves).

### 3.1.2 Wave Spectrum with Bimodal and Trimodal Distribution

For this section, the sensitivity tests are conducted on smooth bimodal and trimodal wave spectrums. These are generated by the superposition of more than one JONSWAP wave models. The baseline spectrums are not necessarily identical for all scenarios (Figure 7) unlike in the previous section.
Two methods have been tested to solve the validation parameters. The first approach is similar to the previous section, where all the parameters are computed directly from the spectrums. For the second approach, a new spectral partitioning algorithm is implemented to identify the number of modes and separate them, before the modes are analysed independently from each other.

The spectral partitioning algorithm was developed with the primary objective of identifying and separating the modes, rather than partitioning the swell from the wind waves. For most bimodal events however, segregation of the modes would often coincide with separation of the swell components from the wind sea components. The script was written in MATLAB® 2012 and contained the following procedures:

**First Forward Scan**

1. The position of the peak energy density, $E_{max}$ is detected and assigned as the reference peak, $E_f$;
2. A forward scan is carried out (starting from $E_{\text{max}}$ to the high frequency region) to locate the next peak;
3. If a new peak is found, it is assigned as the reference peak $E_2$;
4. The minimum point between $E_1$ and $E_2$ is assigned as $T$;
5. Three conditional statements are then executed:
   - $E_2$ must be at least one-third of $E_{\text{max}}$;
   - $E_2$ must be at least 0.4 m$^2$/Hz;
   - $T$ must be less than or equal to 70% of $E_2$;

These conditions are slightly modified from the Mason et al. (2008) paper. If all three conditions are satisfied, the point $T$ will be assigned as a partitioning frequency. Else, the forward scan is resumed at the next frequency bin;
6. Steps 3 to 5 are looped until the last frequency bin is reached.

**Second Forward Scan**

7. The position of $E_1$ is reassigned to the largest peak detected from the first forward scan but the position of $E_{\text{max}}$ is maintained. If no peaks were found from the first forward scan, skip to step 9;
8. Steps 2 to 6 are looped again.

**Backward Scan**

9. Steps 1 to 8 are repeated but the forward scans are replaced by backward scans (starting from $E_{\text{max}}$ to the lowest frequency bin).

The partitioning frequencies are identified from the baseline dataset (or for real application, the observed dataset). Equal width compartments are then assigned to both the baseline and predicted spectrums. The forward and backward scans are carried out twice to ensure that successful partitioning of bimodal and trimodal spectrums would always be achieved, regardless of the relative position of the peaks and troughs in the frequency domain, provided that all conditions in step 5 are fulfilled. These conditions are incorporated to prevent noise (e.g. minor spikes) from being regarded as modes in the spectrum. Figure 8 and Figure 9 illustrate the flow diagrams of the overall script.
Figure 8: Flow diagram of the spectral partitioning method (forward scan)
Figure 9: Flow diagram of the spectral partitioning method (backward scan)
3.2 Result and Discussion

The validation parameters are evaluated in this section. For each parameter, not all scenarios are discussed, but only those that could provide specific evidence on their strengths or weaknesses.

3.2.1 Wave Spectrum with Unimodal Distribution

The first set of parameters consists of $H_s,f_p$ and $E_{max}$ (Table 2).

Table 2: Significant wave height, peak frequency and peak energy density for baseline and predicted spectrums in scenario 1 to 8 (Figure 6)

<table>
<thead>
<tr>
<th>Spectrum</th>
<th>Significant Wave Height, $H_s$ (m)</th>
<th>Peak Frequency, $f_p$ (Hz)</th>
<th>Peak Energy Density, $E_{max}$ (m²/Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>1.7</td>
<td>0.17</td>
<td>3.49</td>
</tr>
<tr>
<td>1</td>
<td>1.7</td>
<td>0.14</td>
<td>3.49</td>
</tr>
<tr>
<td>2</td>
<td>1.7</td>
<td>0.11</td>
<td>3.49</td>
</tr>
<tr>
<td>3</td>
<td>1.7</td>
<td>0.24</td>
<td>3.49</td>
</tr>
<tr>
<td>4</td>
<td>2.1</td>
<td>0.17</td>
<td>3.61</td>
</tr>
<tr>
<td>5</td>
<td>1.6</td>
<td>0.17</td>
<td>2.15</td>
</tr>
<tr>
<td>6</td>
<td>1.7</td>
<td>0.17</td>
<td>2.19</td>
</tr>
<tr>
<td>7</td>
<td>2.1</td>
<td>0.14</td>
<td>4.63</td>
</tr>
<tr>
<td>8</td>
<td>1.4</td>
<td>0.20</td>
<td>2.69</td>
</tr>
</tbody>
</table>

As a function of the area below the curves, the $H_s$ are equal for Scenario 1 to Scenario 3 because the predicted spectrums are merely transposed or flipped from the initial baseline spectrum. A frequency offset or a reversed skew distribution could be identified with the following properties:

- Baseline $H_s$ is equal to predicted $H_s$ (or slightly different from each other with the uncertainties from noise in a true spectrum);
- Baseline $f_p$ is different from predicted $f_p$;
- Baseline $E_{max}$ is exactly equal to predicted $E_{max}$.
A more interesting observation is the identical $H_s$ and $f_p$ (or reciprocal of $T_p$) between the baseline and predicted spectrums in Scenario 6, despite an evident visual difference between their shapes. This proves the notion on how validation with these two sea-state parameters could be insufficient, even for unimodal spectrums.

A better engineering judgment can be made with the addition of the $E_{max}$ parameter. For Scenario 6, it would be reasonable to assume that the predicted spectrum has a wider base, because $E_{max}$ is approximately 63% lower than the baseline spectrum, given the knowledge that they are both unimodal. A broader base would compensate for the loss area under the peak, hence resulting in equal wave energies.

The three parameters in Table 2 are also found to be sufficient in explaining the variance in Scenario 7 and Scenario 8, where the predicted spectrums have failed to estimate the correct scale and position of the peak energy density. Nonetheless, all of the spectrums would be better described with parameters that could define their characteristic shape (e.g. spectral width, skewness and kurtosis).

Table 3: Spectral width parameter, spectral peakedness parameter, skewness and kurtosis for baseline and predicted spectrums in Scenario 1 to 8 (Figure 6)

<table>
<thead>
<tr>
<th>Spectrum</th>
<th>Spectral Width Parameter, $\varepsilon$</th>
<th>Spectral Peakedness Parameter, $Q_p$</th>
<th>Skewness, $Sk$</th>
<th>Kurtosis, $K$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>0.27</td>
<td>3.26</td>
<td>1.61</td>
<td>5.59</td>
</tr>
<tr>
<td>1</td>
<td>0.35</td>
<td>2.69</td>
<td>1.61</td>
<td>5.59</td>
</tr>
<tr>
<td>2</td>
<td>0.45</td>
<td>2.13</td>
<td>1.61</td>
<td>5.59</td>
</tr>
<tr>
<td>3</td>
<td>0.12</td>
<td>4.51</td>
<td>-1.61</td>
<td>5.59</td>
</tr>
<tr>
<td>4</td>
<td>0.25</td>
<td>2.98</td>
<td>1.73</td>
<td>6.77</td>
</tr>
<tr>
<td>5</td>
<td>0.27</td>
<td>2.60</td>
<td>1.34</td>
<td>4.61</td>
</tr>
<tr>
<td>6</td>
<td>0.26</td>
<td>2.58</td>
<td>1.44</td>
<td>5.25</td>
</tr>
<tr>
<td>7</td>
<td>0.33</td>
<td>2.53</td>
<td>1.62</td>
<td>5.86</td>
</tr>
<tr>
<td>8</td>
<td>0.19</td>
<td>4.70</td>
<td>1.68</td>
<td>5.83</td>
</tr>
</tbody>
</table>
Unfortunately, attempts to distinguish the spectral shapes with $\varepsilon$ and $Q_p$ are difficult (Table 3). Although they have identical shapes, the values of $\varepsilon$ and $Q_p$ in Scenario 1 and Scenario 2 are greatly different. Both parameters are found to be extremely sensitive to the position of the peak frequencies.

From the definitions in section 2.5.2, it is expected that $\varepsilon$ would increase while $Q_p$ would decrease for a broader-band spectrum. However, visual comparisons between the baseline and predicted spectrums show an inconsistent trend from Scenario 4 to Scenario 6. It is suspected that the values of $\varepsilon$ and $Q_p$ represent the widths of the normalized peaks ($E/E_{\text{max}}$), rather than the absolute peaks (Figure 10).

![Normalized Spectral Energy Density vs Frequency](image)

Figure 10: Normalized spectral energy density versus frequency for baseline spectrum and predicted spectrums in Scenario 4 to Scenario 6

As the position of the peak frequency influences their estimates, it would be hard to make a meaningful interpretation from the parameters unless the peak frequencies are exactly equal. As such, both parameters are considered to be poor in robustness.
On the other hand, skewness and kurtosis are independent of the location of the peak frequencies (e.g. Scenario 1 to Scenario 3). This is in agreement with what MacGillivray & Balanda (1988) have described on the properties of these central moment parameters. Therefore, frequency offsets and reversed skew conditions can be distinguished from each other with the following properties:

- A predicted spectrum would have a frequency offset if $H_s$, $E_{max}$, $Sk$ and $K$ are identical (or almost identical) with the baseline spectrum but $f_p$ is different;
- A predicted spectrum would be the reversed skew distribution of the baseline spectrum if $H_s$, $E_{max}$, and $K$ are identical (or almost identical) but $f_p$ is different. $Sk$ would be equal in magnitude but opposite in sign.

The measure of skewness can still be quite difficult to interpret in various instances. For example, although the peaks have the same position in Scenario 5, the skewness for the predicted spectrum is smaller than the baseline spectrum. Hence, it would be false to assume that smaller values of skewness are immediately associated with a longer left tail and shorter right tail in the distribution, or vice versa.

Similarly, it is hard to analyse the relative values of kurtosis because it could be related to either the width or height of the peaks, as demonstrated in Scenario 4 and Scenario 5 respectively. For the former scenario, a larger kurtosis is associated with a broader spectrum. However, for the latter scenario, a smaller kurtosis is associated with a shorter peak, rather than a narrower spectrum.

Extra care should also be taken when making inferences from these parameters because of their scale independence. Two spectrums with equal skewness and kurtosis values could have significant difference in their magnitude (Figure 11). Due to these uncertainties, no conclusion should be made from the relative measures of skewness and kurtosis only, unless they are aided by visual reference of the wave spectrums.
Figure 11: Wave spectrums that have identical skewness and kurtosis

The final parameters that would be investigated are measures of distance between two spectrums, consisting of the squared Euclidean distance, Kullback-Leibler divergence and Kolmogorov-Smirnov distance.

Table 4: Squared Euclidean distance, Kullback-Leibler divergence and Kolmogorov-Smirnov distance for Scenario 1 to 8 (Figure 6)

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Squared Euclidean Distance, $D_{SE}$ (m$^2$/Hz)</th>
<th>Kullback-Leibler Divergence, $D_{KL}$ (m$^2$)</th>
<th>Kolmogorov-Smirnov Distance, $D_{KS}$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.28</td>
<td>0.10</td>
<td>14.64</td>
</tr>
<tr>
<td>2</td>
<td>0.46</td>
<td>0.20</td>
<td>18.80</td>
</tr>
<tr>
<td>3</td>
<td>0.42</td>
<td>0.16</td>
<td>16.47</td>
</tr>
<tr>
<td>4</td>
<td>0.08</td>
<td>-0.06</td>
<td>5.42</td>
</tr>
<tr>
<td>5</td>
<td>0.03</td>
<td>0.04</td>
<td>4.86</td>
</tr>
<tr>
<td>6</td>
<td>0.03</td>
<td>0.00</td>
<td>7.41</td>
</tr>
<tr>
<td>7</td>
<td>0.33</td>
<td>-0.02</td>
<td>12.32</td>
</tr>
<tr>
<td>8</td>
<td>0.26</td>
<td>0.24</td>
<td>18.50</td>
</tr>
</tbody>
</table>
Figure 12: Squared Euclidean distance and Kullback-Leibler divergence for Scenario 1 to Scenario 4
**Figure 13**: Squared Euclidean distance and Kullback-Leibler divergence for Scenario 5 to Scenario 8

- **Scenario 5**
  - $D_{SE} = 0.03$
  - $D_{KL} = 0.04$

- **Scenario 6**
  - $D_{SE} = 0.03$
  - $D_{KL} = 0.00$

- **Scenario 7**
  - $D_{SE} = 0.33$
  - $D_{KL} = -0.02$

- **Scenario 8**
  - $D_{SE} = 0.26$
  - $D_{KL} = 0.24$

Legend:
- Gray: Squared Euclidean Distance
- Yellow: Kullback-Leibler Divergence
The main aim of measuring the integrated distance between two frequency wave spectrums is to quantify their overall fitness in terms of magnitude, shape and/or position. It does not provide explicit information on how the spectrums are different. However, it could be used alongside other parameters (e.g. $f_p$ and $E_{max}$) to determine which parts of the spectrum are being poorly modelled or which parts are being well modelled.

The first parameter is the squared Euclidean distance, which measures the magnitude difference between two spectrums at all frequency bins. This is represented by the grey-shaded area in Figure 12 and Figure 13. It is found to be robust because all kinds of divergence (e.g. magnitude and spectral shape difference) would add up to its integrated distance. With the second power term in Equation 2.14, a greater weightage is placed on larger distances. However, due to this same term, the distance would always remain positive and thus, it could not be ascertained whether the baseline spectrum is under- or overestimated.

On the other hand, the Kullback-Leibler divergence measures the amount of spectral energy “lost” from estimating the baseline spectrum with the predicted spectrum. A positive value would suggest a net energy loss (underestimation) while a negative value would suggest a net energy gain (overestimation) but only relative to the baseline spectrum. This is found to be the main drawback of the parameter.

In Scenario 2, the yellow-shaded area between 0.10-0.14 Hz is smaller than the area between 0.15-0.20 Hz. From visual inspection though, the vertical distance between the two spectrums is almost identical at these regions. This is not reflected by the Kullback-Leibler divergence. At range of frequencies where the baseline spectrum is close to zero, the value would tend to be minimal, despite of any considerable difference between the distributions.

Another problem with using the Kullback-Leibler divergence is the prevalent issue of making a qualitative interpretation. A zero divergence could either mean that the two spectrums have a good agreement or the sum of energy gain has cancelled out the sum of energy loss, such as the one shown in Scenario 6. Therefore, a zero or small divergence value could easily be misconstrued. Based on the above arguments, the squared Euclidean distance is found to be more superior to the Kullback-Leibler divergence for computing distance between wave spectrums.
Lastly, the Kolmogorov-Smirnov distance has been adopted from statistical studies to check the goodness of fit between the baseline and predicted spectrums. Prior to the test, the spectral energy densities must be scaled to ensure that their cumulative values add up to a hundred percent.

![Spectral Energy Density (%) versus Frequency](image)

**Figure 14:** Scaling of frequency wave spectrums for the Kolmogorov-Smirnov test

On one hand, all of the wave spectrums have equal energies. As a result, the difference in the absolute energy becomes irrelevant to the Kolmogorov-Smirnov distance. With this unique property, its value could be automatically attributed to divergence in either the spectral shape or position only. On the other hand, the scaling itself distorts the distributions (Figure 14). Rather than reporting the largest deviation from the original spectrums, the Kolmogorov-Smirnov test measures the maximum divergence from the distorted spectrums.

Hence, it would be misleading to make any direct comparative and qualitative interpretation from its value. It does appear that larger distances are better associated with frequency offsets than other conditions, but this is inconclusive.
Figure 15: Cumulative energy and the Kolmogorov-Smirnov distance for Scenario 1 to Scenario 8
The Kolmogorov-Smirnov test should never be analysed on its own. It may be possible to use it with other validation parameters to make better inferences on the spectral shape variation. Otherwise, it should be avoided altogether because of the distortion issue.

In summary, the key findings from this section are:

- $f_p$ and $E_{max}$ are two robust parameters that define the position of the spectral maxima in the frequency domain. These parameters should be used in parallel with significant wave height, $H_s$, to achieve a better model validation.
- $\varepsilon$ and $Q_p$ are sensitive to the location of peak frequencies. Their values are inappropriate for comparison between wave spectrums.

![Figure 16: Functions of the key validation parameters](image)

- Skewness and kurtosis are scale and location independent. They provide a rough indicator on the characteristic shape of the spectrums. However, the scale of their values should not be associated with the degree of skewness, peakiness or broadness of the spectrum, unless aided by visual references.
- The squared Euclidean distance is considered better than the Kullback-Leibler divergence for measuring distance between two spectrums.
The Kolmogorov-Smirnov distance is difficult to interpret because of the distortion effects but theoretically, its value is related to either shape variation or frequency offsets.

Aside from $\epsilon$, $Q_p$ and $D_{KL}$, the other parameters are brought forward to the next section for analysis on bimodal and trimodal distributions.

### 3.2.2 Wave Spectrum with Bimodal and Trimodal Distribution

This section presents the sensitivity analysis on bimodal and trimodal wave spectrums. The first approach involves direct computation of the parameters from the respective wave spectrums without spectral partitioning (Table 5).

Table 5: Validation parameters for Scenario 9 to Scenario 12 (Figure 7) without spectral partitioning

<table>
<thead>
<tr>
<th>Validation Parameter</th>
<th>Scenario 9</th>
<th>Scenario 10</th>
<th>Scenario 11</th>
<th>Scenario 12</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
<td>Predicted</td>
<td>Baseline</td>
<td>Predicted</td>
</tr>
<tr>
<td>$H_s$ (m)</td>
<td>2.0</td>
<td>1.7</td>
<td>2.2</td>
<td>1.7</td>
</tr>
<tr>
<td>$f_p$ (Hz)</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td>$E_{max}$ (m$^2$/Hz)</td>
<td>3.49</td>
<td>3.49</td>
<td>3.50</td>
<td>3.49</td>
</tr>
<tr>
<td>$Sk$</td>
<td>0.90</td>
<td>1.61</td>
<td>0.62</td>
<td>1.61</td>
</tr>
<tr>
<td>$K$</td>
<td>4.42</td>
<td>5.59</td>
<td>3.28</td>
<td>5.59</td>
</tr>
<tr>
<td>$D_{SE}$ (m$^2$/Hz)</td>
<td>0.06</td>
<td></td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>$D_{KS}$ (%)</td>
<td>23.60</td>
<td></td>
<td>19.28</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Baseline</td>
<td>Predicted</td>
<td>Baseline</td>
<td>Predicted</td>
</tr>
<tr>
<td>$H_s$ (m)</td>
<td>2.2</td>
<td>1.9</td>
<td>2.2</td>
<td>2.0</td>
</tr>
<tr>
<td>$f_p$ (Hz)</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td>$E_{max}$ (m$^2$/Hz)</td>
<td>3.50</td>
<td>3.50</td>
<td>3.50</td>
<td>3.50</td>
</tr>
<tr>
<td>$Sk$</td>
<td>0.62</td>
<td>1.18</td>
<td>0.62</td>
<td>0.70</td>
</tr>
<tr>
<td>$K$</td>
<td>3.28</td>
<td>4.02</td>
<td>3.28</td>
<td>3.59</td>
</tr>
<tr>
<td>$D_{SE}$ (m$^2$/Hz)</td>
<td>0.06</td>
<td></td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>$D_{KS}$ (%)</td>
<td>19.98</td>
<td></td>
<td>4.04</td>
<td></td>
</tr>
</tbody>
</table>
For all scenarios, $f_p$ and $E_{max}$ have a good agreement between the baseline and predicted spectrums. $H_s$, $Sk$ and $K$ are considerably different but these data alone are insufficient to imply the existence of a multimodal frequency wave spectrum. For instance, it could be possible that the same values are obtained from two unimodal spectrums that are varied in shape, as demonstrated in section 3.2.1. Without referring to the plotted spectrums, it would be impossible to distinguish between unimodal, bimodal and trimodal distributions from these integrated parameters.

The second approach involves computation of the validation parameters from partitioned wave spectrums. Each compartment represents a mode that has been identified from the baseline spectrum using the new spectral partitioning algorithm. The partitioning frequencies form the boundaries between them (Table 6).

Table 6: Number of modes and partitioning frequencies reported by the spectral partitioning algorithm

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Modes from Baseline Spectrum</th>
<th>Partitioning Frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>2</td>
<td>0.12 Hz</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>0.12 Hz, 0.22 Hz</td>
</tr>
<tr>
<td>11</td>
<td>3</td>
<td>0.12 Hz, 0.22 Hz</td>
</tr>
<tr>
<td>12</td>
<td>3</td>
<td>0.12 Hz, 0.22 Hz</td>
</tr>
</tbody>
</table>

A question that might be raised is: Why is the initial algorithm implemented on the baseline spectrum but not the predicted spectrum? It is because it is based on the analogy of validating a wave model – the objective is to determine the number of modes from a measured wave record and then assess how the wave model performs at each partition. It is more useful and important to evaluate how the model can replicate the physical processes in reality, rather than the opposite.
Table 7: Validation parameters for Scenario 9 with spectral partitioning

<table>
<thead>
<tr>
<th>Scenario 9</th>
<th>Mode</th>
<th>1 (0.00 – 0.12 Hz)</th>
<th>2 (0.12 – 0.40 Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spectrum</td>
<td></td>
<td>Baseline</td>
<td>Predicted</td>
</tr>
<tr>
<td>$H_s$ (m)</td>
<td></td>
<td>1.0</td>
<td>0.2</td>
</tr>
<tr>
<td>$f_p$ (Hz)</td>
<td></td>
<td>0.10</td>
<td>0.12</td>
</tr>
<tr>
<td>$E_{max}$ (m$^2$/Hz)</td>
<td></td>
<td>1.48</td>
<td>0.18</td>
</tr>
<tr>
<td>$Sk$</td>
<td></td>
<td>-0.82</td>
<td>-1.96</td>
</tr>
<tr>
<td>$K$</td>
<td></td>
<td>3.47</td>
<td>6.05</td>
</tr>
<tr>
<td>$D_{SE}$ (m$^4$/Hz)</td>
<td></td>
<td></td>
<td>0.06</td>
</tr>
<tr>
<td>$D_{KS}$ (%)</td>
<td></td>
<td></td>
<td>67.10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scenario 9</th>
<th>Mode</th>
<th>2 (0.12 – 0.40 Hz)</th>
<th>2 (0.12 – 0.40 Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spectrum</td>
<td></td>
<td>Baseline</td>
<td>Predicted</td>
</tr>
<tr>
<td>$H_s$ (m)</td>
<td></td>
<td>1.8</td>
<td>1.7</td>
</tr>
<tr>
<td>$f_p$ (Hz)</td>
<td></td>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td>$E_{max}$ (m$^2$/Hz)</td>
<td></td>
<td>3.49</td>
<td>3.49</td>
</tr>
<tr>
<td>$Sk$</td>
<td></td>
<td>1.57</td>
<td>1.62</td>
</tr>
<tr>
<td>$K$</td>
<td></td>
<td>5.60</td>
<td>5.60</td>
</tr>
<tr>
<td>$D_{SE}$ (m$^4$/Hz)</td>
<td></td>
<td></td>
<td>0.01</td>
</tr>
<tr>
<td>$D_{KS}$ (%)</td>
<td></td>
<td></td>
<td>7.23</td>
</tr>
</tbody>
</table>

Figure 17: Example of partitioned wave spectrum in Scenario 9
Table 8: Validation parameters for Scenario 10 (Figure 7) with spectral partitioning

<table>
<thead>
<tr>
<th>Mode</th>
<th>1 (0.00 – 0.12 Hz)</th>
<th>2 (0.12 – 0.22 Hz)</th>
<th>3 (0.22 – 0.40 Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spectrum</strong></td>
<td><strong>Baseline</strong></td>
<td><strong>Predicted</strong></td>
<td><strong>Baseline</strong></td>
</tr>
<tr>
<td>$H_s$ (m)</td>
<td>1.0</td>
<td>0.2</td>
<td>1.6</td>
</tr>
<tr>
<td>$f_p$ (Hz)</td>
<td>0.10</td>
<td>0.12</td>
<td>0.17</td>
</tr>
<tr>
<td>$E_{max}$ (m$^2$/Hz)</td>
<td>1.48</td>
<td>0.18</td>
<td>3.50</td>
</tr>
<tr>
<td>$Sk$</td>
<td>-0.82</td>
<td>-1.96</td>
<td>0.15</td>
</tr>
<tr>
<td>$K$</td>
<td>3.47</td>
<td>6.05</td>
<td>2.65</td>
</tr>
<tr>
<td>$D_{SE}$ (m$^4$/Hz)</td>
<td></td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>$D_{KS}$ (%)</td>
<td></td>
<td>67.10</td>
<td></td>
</tr>
</tbody>
</table>

Scenario 10
Table 9: Validation parameters for Scenario 11 (Figure 7) with spectral partitioning

<table>
<thead>
<tr>
<th>Mode</th>
<th>1 (0.00 – 0.12 Hz)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Spectrum</td>
<td>Baseline</td>
<td>Predicted</td>
</tr>
<tr>
<td></td>
<td>$H_s$ (m)</td>
<td>1.0</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>$f_p$ (Hz)</td>
<td>0.10</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>$E_{max}$ (m$^2$/Hz)</td>
<td>1.48</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>$Sk$</td>
<td>-0.82</td>
<td>-1.96</td>
</tr>
<tr>
<td></td>
<td>$K$</td>
<td>3.47</td>
<td>6.05</td>
</tr>
<tr>
<td></td>
<td>$D_{SE}$ (m$^4$/Hz)</td>
<td></td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>$D_{KS}$ (%)</td>
<td></td>
<td>67.10</td>
</tr>
<tr>
<td></td>
<td>2 (0.12 – 0.22 Hz)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Spectrum</td>
<td>Baseline</td>
<td>Predicted</td>
</tr>
<tr>
<td></td>
<td>$H_s$ (m)</td>
<td>1.6</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td>$f_p$ (Hz)</td>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>$E_{max}$ (m$^2$/Hz)</td>
<td>3.50</td>
<td>3.50</td>
</tr>
<tr>
<td></td>
<td>$Sk$</td>
<td>0.15</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td>$K$</td>
<td>2.65</td>
<td>2.83</td>
</tr>
<tr>
<td></td>
<td>$D_{SE}$ (m$^4$/Hz)</td>
<td></td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>$D_{KS}$ (%)</td>
<td></td>
<td>8.07</td>
</tr>
<tr>
<td></td>
<td>3 (0.22 – 0.40 Hz)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Spectrum</td>
<td>Baseline</td>
<td>Predicted</td>
</tr>
<tr>
<td></td>
<td>$H_s$ (m)</td>
<td>1.2</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>$f_p$ (Hz)</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>$E_{max}$ (m$^2$/Hz)</td>
<td>1.28</td>
<td>1.02</td>
</tr>
<tr>
<td></td>
<td>$Sk$</td>
<td>1.19</td>
<td>1.15</td>
</tr>
<tr>
<td></td>
<td>$K$</td>
<td>3.74</td>
<td>3.61</td>
</tr>
<tr>
<td></td>
<td>$D_{SE}$ (m$^4$/Hz)</td>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>$D_{KS}$ (%)</td>
<td></td>
<td>1.37</td>
</tr>
</tbody>
</table>
Table 10: Validation parameters for Scenario 12 (Figure 7) with spectral partitioning

<table>
<thead>
<tr>
<th>Mode</th>
<th>1 (0.00 – 0.12 Hz)</th>
<th>2 (0.12 – 0.22 Hz)</th>
<th>3 (0.22 – 0.40 Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Scenario 12</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Spectrum</strong></td>
<td><strong>Baseline</strong></td>
<td><strong>Predicted</strong></td>
<td><strong>Baseline</strong></td>
</tr>
<tr>
<td>$H_s$ (m)</td>
<td>1.0</td>
<td>0.8</td>
<td>1.6</td>
</tr>
<tr>
<td>$f_p$ (Hz)</td>
<td>0.10</td>
<td>0.10</td>
<td>0.17</td>
</tr>
<tr>
<td>$E_{max}$ (m$^2$/Hz)</td>
<td>1.48</td>
<td>1.04</td>
<td>3.50</td>
</tr>
<tr>
<td>$Sk$</td>
<td>-0.82</td>
<td>-0.82</td>
<td>0.15</td>
</tr>
<tr>
<td>$K$</td>
<td>3.47</td>
<td>3.46</td>
<td>2.65</td>
</tr>
<tr>
<td>$D_{SE}$ (m$^4$/Hz)</td>
<td>0.01</td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>$D_{KS}$ (%)</td>
<td>1.09</td>
<td>1.09</td>
<td>2.08</td>
</tr>
</tbody>
</table>
It is apparent that analysis of the individual modes gives a better picture on the similarity between the baseline and predicted spectrums:

- **Scenario 9 (Table 7):** The parameters have a poor agreement at Mode 1 but good agreement at Mode 2. This is supported by the squared Euclidean distance at Mode 1, which is 600% larger than at Mode 2. The data also show the failure of the predicted spectrum to capture the energetic long period waves (~ 8 s) found in the baseline spectrum, as confirmed in Figure 17.

- **Scenario 10 (Table 8):** Unlike the previous scenario, a trimodal distribution has been identified in Scenario 10. The parameters have a good agreement at Mode 2, but poor agreement at Mode 1 and Mode 3. However, what is more intriguing to observe is the close value of Kolmogorov-Smirnov distance between Mode 2 and Mode 3, despite the fact that no peak has been found in the 0.22-0.40 Hz region of the predicted spectrum. One would expect that the absence of the peak would result in a larger shape deviation. However, this is not reflected by the Kolmogorov-Smirnov distance. This could be due to the distortion issue again. It further proves the unreliability of the parameter in this study.

- **Scenario 11 (Table 9):** Similar to the previous scenario, but with improved validation result at Mode 3. The peak at Mode 3 has been captured at the correct frequency but its energy density is underestimated by roughly 20%. Values for the skewness and kurtosis have also converged, indicating more similarity in their characteristic shapes.

- **Scenario 12 (Table 10):** Similar to the previous scenario, but with improved agreement between the baseline and predicted spectrum at Mode 1. A shorter peak has been estimated by the predicted spectrum.

Another recurring observation is the overlapping of partitioning frequencies with the peak frequencies (e.g. Scenario 9 Mode 1, Scenario 10 Mode 1 and Mode 2). It is found to be the most important relationship in analyzing frequency spectrums with bimodal or trimodal distributions. The absence of peaks is implied when these two values overlap.
The spectral partitioning approach provides more reliable information about the two spectrums than direct computation of the validation parameters. The spectral partitioning algorithm should be implemented at all times to ensure that distinct modes are separated for the validation process. In the event that a unimodal spectrum is encountered, no partitions would be formed and the spectrum would be analysed as a whole unit.

### 3.2.3 Relationship between Validation Parameters

From the preceding discussion, it can be concluded that a family of validation parameters would be required to best describe the characteristics of the wave spectrums for validation purposes. The main outcome from the sensitivity tests is a systematic classification of the parameters into three categories (Table 11):

- **Primary parameters**: In terms of hierarchy, these are the basic but most important parameters in the validation process. A primary parameter should be straight-forward and easy to interpret. Visual references are not required to explain the meaning of their values. It should also be robust and consistent, in the sense that it would provide concrete evidence on the difference between the spectrums, regardless of what the spectral conditions are.

- **Secondary parameters**: These are parameters that could provide more detailed information on divergence between the two spectrums, but their values are also more difficult to interpret and subjective. Visual references of the spectrums are recommended to support any inferences made from their values.

- **Rejected parameters**: These are parameters that are very difficult to interpret and unsuitable for validation purposes. Problems such as over-sensitivity to more than one variable and spectral distortion are associated with them. These parameters have been dismissed from the validation approach.

<table>
<thead>
<tr>
<th>Primary Parameters</th>
<th>Secondary Parameters</th>
<th>Rejected Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_s$</td>
<td>$Sk$</td>
<td>$\varepsilon$</td>
</tr>
<tr>
<td>$f_p$</td>
<td>$K$</td>
<td>$Q_p$</td>
</tr>
<tr>
<td>$E_{max}$</td>
<td></td>
<td>$D_{KL}$</td>
</tr>
<tr>
<td>$D_{SE}$</td>
<td></td>
<td>$D_{KS}$</td>
</tr>
</tbody>
</table>
By further observing the relationship between two parameters, a qualitative inference can be made on the similarity between an observed and predicted frequency wave spectrum. Based on the findings from previous sections, the author proposes the usage of conceptual two-dimensional matrices (Figure 18 and Figure 19) to assist modellers in evaluating the magnitude, shape and position of the wave spectrums. The three delta terms found in the figures are calculated with the following equations:

\[ \Delta H_s = \text{Observed } H_s - \text{Predicted } H_s \quad (3.1) \]

\[ \Delta f_p = \text{Observed } f_p - \text{Predicted } f_p \quad (3.2) \]

\[ \Delta E_{\text{max}} = \text{Observed } E_{\text{max}} - \text{Predicted } E_{\text{max}} \quad (3.3) \]

Each matrix is divided into nine separate regions, where the 9\textsuperscript{th} region is the origin of both axes (0, 0). Values to the right and above the origin are positive.

These matrices should be applied directly for comparisons between unimodal spectrums only. For bimodal or trimodal distributions, the main focus is to establish the presence or absence of peak in each partition – based on whether the partitioning frequencies and peak frequencies are overlapping, as explained in section 3.2.2. The two matrices could then be adopted for comparison between individual modes.

In the strictest sense of their definitions, the values within region 5 to region 8 should fall on the x- or y-axes (in other words, it should actually be a line, rather than an area). However, the delta terms are expected to be slightly offset from zero because of the difference in precision between the observed and predicted datasets. As such, the extent of the blue-dashed boundaries could vary from one case study to another.
Table 12: Description of regions in the two-dimensional matrix between $\Delta E_{\text{max}}$ and $\Delta H_s$

<table>
<thead>
<tr>
<th>Region</th>
<th>Description of Predicted Wave Spectrum</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Larger spectral area but with lower peak energy. Broader spectrum, where energetic waves are distributed over a larger frequency range.</td>
</tr>
<tr>
<td>2</td>
<td>Smaller spectral area with lower peak energy. Inconclusive for width.</td>
</tr>
<tr>
<td>3</td>
<td>Smaller spectral area but with higher peak energy. Narrower spectrum, where energetic waves are focused on a smaller frequency range.</td>
</tr>
<tr>
<td>4</td>
<td>Larger spectral area with higher peak energy. Inconclusive for width.</td>
</tr>
<tr>
<td>5</td>
<td>Equal spectral area but with lower peak. Broader spectrum.</td>
</tr>
<tr>
<td>6</td>
<td>Smaller spectral area but with equal peak energy. Narrower spectrum.</td>
</tr>
<tr>
<td>7</td>
<td>Equal spectral energy but with higher peak. Narrower spectrum.</td>
</tr>
<tr>
<td>8</td>
<td>Larger spectral area but with equal peak energy. Broader spectrum.</td>
</tr>
<tr>
<td>9</td>
<td>Excellent agreement but inconclusive for position, skewness and kurtosis.</td>
</tr>
</tbody>
</table>

Notes:
1. The term “spectral area” refer to the integrated area below the spectral curves
2. The terms “broader” and “narrower” can refer to either the base and/or peak of the spectrums

Figure 18: Two-dimensional validation matrix between $\Delta E_{\text{max}}$ and $\Delta H_s$
Table 13: Description of regions in the two-dimensional matrix between $\Delta E_{\text{max}}$ and $\Delta f_p$.

<table>
<thead>
<tr>
<th>Region</th>
<th>Description of Predicted Wave Spectrum</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Lower peak energy with underestimated peak wave period.</td>
</tr>
<tr>
<td>2</td>
<td>Lower peak energy with overestimated peak wave period.</td>
</tr>
<tr>
<td>3</td>
<td>Higher peak energy with overestimated peak wave period.</td>
</tr>
<tr>
<td>4</td>
<td>Higher peak energy with underestimated peak wave period.</td>
</tr>
<tr>
<td>5</td>
<td>Lower peak energy but peak wave period is precise.</td>
</tr>
</tbody>
</table>
| 6      | Equal peak energy but peak wave period is underestimated. When $\Delta H_s \sim 0$:  
  - If $Sk$ and $K$ are equal, frequency offset occurs;  
  - If $Sk$ are not equal, most likely the spectrum is too left-skewed. |
| 7      | Higher peak energy but peak wave period is precise. |
| 8      | Equal peak energy but peak wave period is overestimated. When $\Delta H_s \sim 0$:  
  - If $Sk$ and $K$ are equal, frequency offset occurs;  
  - If $Sk$ are not equal, most likely the spectrum is too right-skewed. |
| 9      | Excellent agreement between the positions of peak energy densities. |

Figure 19: Two-dimensional validation matrix between $\Delta E_{\text{max}}$ and $\Delta f_p$. 
In conclusion, a total of six parameters \((H_s, f_p, E_{max}, D_{SE}, Sk, K)\) have been found to be suitable for the new validation approach. These are further divided into primary and secondary parameters, depending on their level of complexity and ease of interpretation. A parallel analysis of these parameters could reveal more information about the relative spectral conditions, in comparison to individual assessment of the same parameters.

In the next chapter, a new seventh parameter will be introduced. The new parameter would provide a measure on the spectral width difference, which is still lacking from the existing family of primary and secondary parameters.
4 Mean Width Deviation

In the previous section, six parameters were found to be suitable for the new proposed validation approach. In this section, a new seventh parameter is proposed by the author, known as the mean width deviation for the validation approach. It replaces $\varepsilon$ and $Q_p$ to provide a measure on the difference in the spectral widths between two spectrums.

4.1 Methodology

The absolute width, $Aw_x$, is defined as the actual width of the wave spectrum at $x^{th}$ percent of a threshold level, which is equivalent to the smaller spectral maxima between an observed and predicted wave spectrum. It is measured in the unit of Hertz. This concept is illustrated in Figure 20.

![Figure 20: Concept of absolute width parameter](image)

The peak energy of the predicted spectrum (2.7 m$^2$/Hz) is smaller than the observed spectrum (3.6 m$^2$/Hz). As a result, the threshold level is equal to 2.7 m$^2$/Hz. The absolute widths at the 33rd and 66th percentage level would correspond to the widths at 0.9 m$^2$/Hz and 1.8 m$^2$/Hz respectively.
The main purpose of using a common, threshold level is to ensure that the widths are measured at equal heights, rather than relative heights, which would allow a practical comparison between them. $x$ is an arbitrary number and $A_{w_x}$ can be measured at multiple levels (e.g. every one, two, five, ten or twenty percent). The mean width, $M_w$ is the average value of these multi-level measurements:

\[
M_w = \frac{1}{N_{aw}} \sum A_{w_x} \quad (4.1)
\]

Where $N_{aw}$ is the number of measurements. The mean width deviation between an observed and predicted spectrum is defined with the following relationship:

\[
\Delta M_w = Observed \ M_w - Predicted \ M_w \quad (4.2)
\]

To assess its reliability and robustness, sensitivity tests were carried out using the same 12 scenarios from Chapter 3 (Figure 6 and Figure 7). The absolute widths are computed at every one-percent interval to ensure that the spectral widths are well resolved. Figure 21 illustrates the procedure for solving Equations 4.1 and 4.2. For bimodal and trimodal distributions, the mean width deviation is calculated at each partition, rather than for the whole spectrum.
Figure 21: Flow diagram for solving the mean width deviation in MATLAB® 2012
4.2 Result and Discussion

The mean width is considered to be a dynamic parameter. Its value could change from one situation to another because of changes to the threshold level. Hence, the main focus should be on interpreting the mean width deviation, rather than evaluating the mean width values per se.

Table 14: Mean width and mean width deviation in Scenario 1 to Scenario 8 (Figure 6)

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Mean Width, $M_w$ (Hz)</th>
<th>Mean Width Deviation, $\Delta M_w$ (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
<td>Predicted</td>
</tr>
<tr>
<td>1</td>
<td>0.052</td>
<td>0.052</td>
</tr>
<tr>
<td>2</td>
<td>0.052</td>
<td>0.052</td>
</tr>
<tr>
<td>3</td>
<td>0.052</td>
<td>0.052</td>
</tr>
<tr>
<td>4</td>
<td>0.052</td>
<td>0.073</td>
</tr>
<tr>
<td>5</td>
<td>0.074</td>
<td>0.070</td>
</tr>
<tr>
<td>6</td>
<td>0.074</td>
<td>0.085</td>
</tr>
<tr>
<td>7</td>
<td>0.052</td>
<td>0.061</td>
</tr>
<tr>
<td>8</td>
<td>0.063</td>
<td>0.042</td>
</tr>
</tbody>
</table>

A consistent trend has been observed for the mean width deviation, where a positive value implies a broader-band baseline spectrum while a negative value indicates a narrower-band baseline spectrum, relative to the predicted spectrum. Between Scenario 1 to Scenario 3 (frequency offsets and reversed skew conditions), the value is zero. This proves to be the main advantage of using mean width deviation rather than $\varepsilon$ or $Q_p$ – it is insensitive to the position of the peak frequencies.

In Scenario 5, the mean width deviation is reported to be 0.005 Hz. This is attributed to the top one-third of the predicted spectrum peak, where the left and right-hand sides of the spectrum converge to its maximum point. At these levels, the absolute widths are smaller than the baseline spectrum. For the bottom two-thirds however, the absolute widths are equal. This shows that mean width deviation could be biased towards either the positive or negative direction in certain conditions.

For multimodal distributions, mean width deviation is found to be larger at partitions where peaks are missing in the predicted spectrum (Figure 22 to Figure 25).
Figure 22: Mean width deviation for Scenario 9

Scenario 9

Mode 1:
\[ \Delta M_w = 0.068 \text{ Hz} \]

Mode 2:
\[ \Delta M_w = 0.005 \text{ Hz} \]

Figure 23: Mean width deviation for Scenario 10

Scenario 10

Mode 1:
\[ \Delta M_w = 0.068 \text{ Hz} \]

Mode 2:
\[ \Delta M_w = 0.006 \text{ Hz} \]

Mode 3:
\[ \Delta M_w = 0.036 \text{ Hz} \]
Figure 24: Mean width deviation for Scenario 11

Figure 25: Mean width deviation for Scenario 12
In Scenario 10, the mean width deviation at Mode 1 has been derived from the absolute widths between 0.00 and 0.18 m²/Hz. It does not suggest whether the actual peak on the baseline spectrum is broad or narrow, because the peak itself is located above the threshold level. However, it should be emphasized that the new parameter has been developed with the specific goal of quantifying the difference in the mean widths, measured from the spectral base to the smaller maxima. It is not designed to provide explicit information on the width of the whole mode or the spectrum.

The mean width deviation has its own limitation as well. Hypothetically, two different wave spectrums could have equal mean widths and thus, a zero mean width deviation (Table 15). Spectrum A has a broader base but narrower peak compared to Spectrum B. The parameter would fail to describe this spectral condition.

Table 15: Example of two hypothetical spectrums with zero mean width deviation

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Spectrum A</th>
<th>Spectrum B</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_w_{80}$</td>
<td>0.010 Hz</td>
<td>0.020 Hz</td>
</tr>
<tr>
<td>$A_w_{60}$</td>
<td>0.010 Hz</td>
<td>0.020 Hz</td>
</tr>
<tr>
<td>$A_w_{40}$</td>
<td>0.045 Hz</td>
<td>0.040 Hz</td>
</tr>
<tr>
<td>$A_w_{20}$</td>
<td>0.095 Hz</td>
<td>0.080 Hz</td>
</tr>
<tr>
<td>$M_w$</td>
<td>0.040 Hz</td>
<td>0.040 Hz</td>
</tr>
<tr>
<td>$\Delta M_w$</td>
<td></td>
<td>0.000 Hz</td>
</tr>
</tbody>
</table>

Nevertheless, the standard deviation of the absolute widths could be used to compare the statistical variation of the data from the mean values. In this example, Spectrum A has a standard deviation of 0.040 Hz while Spectrum B has a standard deviation of 0.028 Hz. From these values, it can be inferred that the absolute widths in Spectrum A are more dispersed than in Spectrum B.

The mean width deviation can be used in conjunction with Figure 18, the two-dimensional matrix between $\Delta E_{max}$ and $\Delta H_s$ to redefine Region 2 and Region 4:

- Region 2: Smaller spectral area with lower peak energy. If $\Delta M_w$ is positive, the predicted spectrum is broader, and vice versa;
- Region 4: Larger spectral area with higher peak energy. If $\Delta M_w$ is negative, the predicted spectrum is narrower, and vice versa.
Due to its intricate definition, the mean width deviation is categorized as a secondary parameter. A visual reference may not be required but still recommended. In short, seven parameters have been shortlisted for the validation approach. Each of them provides a specific or related detail to the frequency wave spectrums.

Table 16: Family of final validation parameters

<table>
<thead>
<tr>
<th>Primary Parameter</th>
<th>Detail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Significant wave height, $H_s$</td>
<td>Integrated area below the curve</td>
</tr>
<tr>
<td>Peak frequency, $f_p$</td>
<td>Position of peak on the x-axis</td>
</tr>
<tr>
<td>Peak energy density, $E_{max}$</td>
<td>Position of peak on the y-axis</td>
</tr>
<tr>
<td>Squared Euclidean distance, $D_{SE}$</td>
<td>Integrated distance between two spectrums</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Secondary Parameter</th>
<th>Detail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean width deviation, $\Delta M_{w}$</td>
<td>Spectral width difference</td>
</tr>
<tr>
<td>Skewness, $Sk$</td>
<td>Positive or negative skewed spectrum</td>
</tr>
<tr>
<td>Kurtosis, $K$</td>
<td>“Flat-topped” or “sharply peaked” spectrum</td>
</tr>
</tbody>
</table>

The next chapter demonstrates the application of the new approach using real numerical model output and wave records from around the UK Waters. It is important to establish the feasibility of the new approach for real, practical applications.
5 Example of Application

This chapter presents an example of a hindcast wave model validation at three nearshore sites around the UK Waters, based on measured spectral data from the WaveNet website (www.cefas.defra.gov.uk/our-science/observing-and-modelling/monitoring-programmes/wavenet.aspx)

5.1 Methodology

5.1.1 Numerical Model Setup

For this study, one of the prototype wave models developed by ABPmer (www.abpmer.co.uk) as part of their SEASTATES project has been used to assess the feasibility of the new validation approach. The model was developed using the DHI MIKE 21 SW module. More details about the project can be found on the official SEASTATES website (www.seastates.net).

![Regional view of the model bathymetry in UTM-30 projection. The view is distorted towards the west of the North Atlantic Ocean because of the projection system](image)

Figure 26: Regional view of the model bathymetry in UTM-30 projection. The view is distorted towards the west of the North Atlantic Ocean because of the projection system
The model bathymetry is resolved using an unstructured triangular mesh that extends from 80° W to 24° E, covering most parts of the North Atlantic Ocean. The finer mesh elements near to the coast have a spatial resolution of approximately 5 km.

Figure 27: Local view of the model bathymetry in UTM-30 projection around the UK Waters

The model is solved using a fully spectral, instationary formulation at every one hour time step interval between 29th Dec 2008 and 31st Dec 2009. Swell and sea components are not separated in the model. A non-equidistant, logarithmic type discretization has been used in the frequency domain:

- Number of frequencies: 25 bins;
- Minimum frequency: 0.033 Hz;
- Frequency factor: 1.1.

Wave growth and transformation are driven by one-hourly wind fields derived from the NCEP Global Forecast System (GFS) Reanalysis II project. Bottom friction is applied using a constant Nikuradse roughness of 0.04 meters. In addition, energy dissipation by white capping is defined by two parameters in the model, the Cdis and DELTAdis coefficients. A two-dimensional grid series has been used to specify the spatially-varying coefficients across the model domain. No variations in the water level and current conditions are applied in the simulation. Effects from diffraction and ice coverage are also not included.
The model results are saved every one hour. Spectral outputs are saved between 0 Hz and 0.3404 Hz, at every 0.001702 Hz interval.

5.1.2 Spectral Data Collection

Observed spectral data at three selected nearshore sites have been retrieved from the WaveNet (2013) archive to validate the hindcast wave model (Table 17). These sites have been selected to demonstrate the spatial variation in the wave climate at the English Channel, North Sea and Sea of Hebrides. The data have been recorded at every 30 minutes interval.

Table 17: Sites for model validation and availability of spectral data

<table>
<thead>
<tr>
<th>Site</th>
<th>Data Period</th>
<th>Latitude</th>
<th>Longitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hastings</td>
<td>26/11/2002 to present</td>
<td>50° 44.79’ N</td>
<td>0° 45.32’ E</td>
</tr>
<tr>
<td>Moray Firth</td>
<td>29/08/2008 to present</td>
<td>57° 57.98’ N</td>
<td>3° 19.99’ W</td>
</tr>
<tr>
<td>West of Hebrides</td>
<td>23/02/2009 to present</td>
<td>57° 17.54’ N</td>
<td>7° 54.86’ W</td>
</tr>
</tbody>
</table>

Figure 28: Location of the three nearshore sites on Google Earth
The wave spectrums are resolved at 14 frequency bins that vary from one time step to another. It is suspected that the spectral data have been compressed because of limited transmission capacity between the waverider buoys and the satellite network (Mason T. 2013, personal communication). As a result, the data are only stored in 14 frequency bins that could best highlight the characteristic peaks and troughs for each single time step.

No reference was found regarding how these data have been calculated. However, it is suspected that a Fast Fourier Transform (FFT) algorithm has been applied to transform the observed datasets from the temporal domain into the frequency domain.

5.1.3 Numerical Model Validation

Results from the model have been validated with the WaveNet spectral data using a series of MATLAB® 2012 functions developed by the author. See Appendix A1 for more details.

5.2 Result and Discussion

5.2.1 Time Series Analysis of Validation Parameters

The primary and secondary parameters can be plotted against time to identify temporal trends in the datasets and to analyse the general agreement between the observed and predicted parameters (Figure 29 and Figure 30).
Figure 29: Primary validation parameters at Hastings in 2009
Figure 30: Secondary validation parameters at Hastings in 2009

The distances between the observed and predicted spectrums at Hastings are found to be the largest in November 2009. It seems that the larger values could be linked with the occurrences of bimodal events (acquired from the spectral partitioning algorithm). Swell waves could have propagated up into the English Channel during these periods but were not captured by the model. Several trends are also observed in the same month:

- Higher $H_s$, lower $f_p$ and higher $E_{max}$. All of these indicate the presence of longer period, high energetic waves in the channel.
- Smaller fluctuations in mean width deviation, relative to the rest of the year. It fluctuates at about zero, implying a good agreement in the spectral widths.
- A more consistent, positive skewness between the mid to the end of November 2009. This is attributed to the existence of more dominant, longer period wind sea than short period wind sea in the spectral distributions.
A more detailed analysis can be conducted by zooming into the respective time series at shorter time periods.

Figure 31: Primary validation parameters at Hastings in November 2009
Figure 32: Secondary validation parameters at Hastings in November 2009

Two data windows have been selected in Figure 31 and Figure 32 for the purpose of demonstration. In both cases, the spikes in the squared Euclidean distances are due to the overestimation of spectral wave energies in the simulation. This is reflected by the higher values of $H_s$ and $E_{max}$.

For window 1, there is a slight offset of about 0.02-0.03 Hz in the position of peak frequencies. For window 2, the offsets are more negligible, which explains the smaller spikes in the squared Euclidean distance. The observed and predicted spectrums have also shown good agreement with the other validation parameters (e.g. skewness and kurtosis). As such, it can be concluded that the differences between the spectrums are quite similar to Scenario 7 (Figure 6) from the previous sensitivity tests.

To determine whether there is a relationship between the model performance and the occurrence of bimodal events, the top ten squared Euclidean distances in November 2009 have been ranked.
Table 18: Top 10 squared Euclidean distances at Hastings in November 2009

<table>
<thead>
<tr>
<th>Date and Time</th>
<th>Squared Euclidean Distance, $D_{SE}$ (m$^2$/Hz)</th>
<th>Modes</th>
</tr>
</thead>
<tbody>
<tr>
<td>14/11/2009 16:00</td>
<td>45.88</td>
<td>1</td>
</tr>
<tr>
<td>14/11/2009 15:00</td>
<td>39.91</td>
<td>1</td>
</tr>
<tr>
<td>14/11/2009 14:00</td>
<td>37.44</td>
<td>2</td>
</tr>
<tr>
<td>25/11/2009 04:00</td>
<td>26.42</td>
<td>1</td>
</tr>
<tr>
<td>14/11/2009 17:00</td>
<td>26.06</td>
<td>1</td>
</tr>
<tr>
<td>25/11/2009 03:00</td>
<td>20.23</td>
<td>1</td>
</tr>
<tr>
<td>25/11/2009 05:00</td>
<td>17.95</td>
<td>1</td>
</tr>
<tr>
<td>14/11/2009 13:00</td>
<td>17.68</td>
<td>1</td>
</tr>
<tr>
<td>23/11/2009 09:00</td>
<td>17.27</td>
<td>1</td>
</tr>
<tr>
<td>14/11/2009 12:00</td>
<td>15.00</td>
<td>2</td>
</tr>
</tbody>
</table>

Initially, it was suspected that the numerical wave model performs poorly in predicting the bimodal distribution in the observed spectrum. However, it turns out that only 2 out of the 10 largest squared Euclidean distances are related to bimodal spectrums. In fact, it appears that 90% of these distances are related to two separate storm events occurring on the 14$^{th}$ and 25$^{th}$ November 2009, which are predominantly unimodal (Table 18). Therefore, it can be concluded that there is no direct connection between the model performance and the occurrence of bimodal events.

Time series plots for the other two locations, namely Moray Firth and West of Hebrides can be found in Appendix A2. Spatial trends and variation can be analysed from the same set of validation parameters. Because of its more exposed location to the North Atlantic Ocean, larger waves are observed at West of Hebrides (e.g. maximum observed $H_s \sim 12$ m), relative to Moray Firth (e.g. maximum observed $H_s \sim 4$ m). The occurrence of bimodal and trimodal events are also more frequent at West of Hebrides and Moray Firth, compared to Hastings. For all three sites, high energy waves are found to be more common in the autumn and winter periods.

Perhaps the most important observation is the spikes in the squared Euclidean distances, which largely increase with the magnitude of the wave spectrums, based on the $H_s$ and $E_{max}$. At West of Hebrides, the maximum squared Euclidean distance is
about 18 times larger than at Moray Firth. While the value tends to increase with the scale of the spectrums, it is false to assume any direct relationship between them. Two large but identical wave spectrums would still show a zero value for the parameter.

This proves to be the main challenge in establishing the significance from the squared Euclidean distances. What is the range of distances that would be considered acceptable or not acceptable in a numerical wave model, given that the value has an inclination to grow with larger spectrums? At present, there is no objective answer to the problem. While the new validation approach could provide a more qualitative comparison between the observed and predicted spectrums, no specific criteria could be established on what defines an excellent or poor agreement between them. It would depend entirely on the personal discretion of the modeller. Due to this limitation, it is acknowledged that the new approach is more suited for relative comparison of wave conditions in the temporal and spatial domains only.

5.2.2 Statistical Performance of Primary Parameters

A significance level (p-value) of 0.05 has been selected to determine whether the correlation of the primary parameters is statistically significant in this study.

![Figure 33: Predicted versus observed significant wave heights at Hastings](image)

- **RMSE:** 0.31
- **Bias:** -0.001
- **Scatter Index:** 31.0
- **Correlation, R:** 0.94
- **p-value:** 0.00
The equations for the root mean square error (RMSE), bias, scatter index and correlation can be found in Appendix A3. The relationships between the observed and predicted parameters at Hastings are found to be statistically significant with the p-value falling below 0.05 (Figure 33 to Figure 35). There is good correlation (above 80%) for $H_s$ and $E_{max}$. On the other hand, the correlation for $f_p$ is considerably poor.

The horizontal binning pattern that emerges in Figure 34 is attributed to the discretization of spectral frequencies in the MIKE 21 SW model setup. A better correlation might be achieved by adding the number of frequency bins and reducing...
the discretization log factor. However, the tradeoff for this is a longer computational time for the simulation.

The statistical analysis for Moray Firth and West of Hebrides are listed in Appendix A2. Once again, the relationships between the observed and predicted parameters are proven to be statistically significant because all of the p-values are less than 0.05. In terms of spatial variation, it appears that the primary parameters have the best agreement at West of Hebrides (R = 0.96 for $H_s$, R = 0.68 for $f_p$ and R = 0.89 for $E_{max}$). The correlation of $f_p$ at West of Hebrides is about forty-percent higher than at Hastings and Moray Firth.

It is suspected that the DELTAdis coefficient for defining white-capping could be better calibrated at Hastings and Moray Firth, in order to achieve better agreement between the observed and predicted $f_p$. A coefficient value of less than 0.5 would result in smaller wave periods while a coefficient value of more than 0.5 would result in larger wave periods (DHI Water & Environment 2007). Sensitivity tests are recommended to confirm whether this would improve the results.

5.2.3 Qualitative Assessment of Unimodal Spectrums

The first assessment involves the application of the two-dimensional matrix between $\Delta E_{max}$ and $\Delta H_s$ in Figure 18. The precision of the WaveNet records has been considered in defining the boundaries of region 1 to region 9.

**Table 19: Boundaries of region 1 to region 9 for two-dimensional matrix between $\Delta E_{max}$ and $\Delta H_s$**

<table>
<thead>
<tr>
<th>Region</th>
<th>$\Delta H_s$ (m)</th>
<th>$\Delta E_{max}$ (m$^2$/Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$&lt; -0.05$</td>
<td>$&gt; 0.01$</td>
</tr>
<tr>
<td>2</td>
<td>$&gt; 0.05$</td>
<td>$&gt; 0.01$</td>
</tr>
<tr>
<td>3</td>
<td>$&gt; 0.05$</td>
<td>$&lt; -0.01$</td>
</tr>
<tr>
<td>4</td>
<td>$&lt; -0.05$</td>
<td>$&lt; -0.01$</td>
</tr>
<tr>
<td>5</td>
<td>-0.05 to 0.05</td>
<td>$&gt; 0.01$</td>
</tr>
<tr>
<td>6</td>
<td>$&gt; 0.05$</td>
<td>-0.01 to 0.01</td>
</tr>
<tr>
<td>7</td>
<td>-0.05 to 0.05</td>
<td>$&lt; -0.01$</td>
</tr>
<tr>
<td>8</td>
<td>$&lt; -0.05$</td>
<td>-0.01 to 0.01</td>
</tr>
<tr>
<td>9</td>
<td>-0.05 to 0.05</td>
<td>-0.01 to 0.01</td>
</tr>
</tbody>
</table>
Table 20: Validation with $\Delta E_{\text{max}}$ versus $\Delta H_s$ matrix for unimodal spectrums at Hastings

![Graph showing the relationship between $\Delta E_{\text{max}}$ and $\Delta H_s$ with data points for Hastings region.]

<table>
<thead>
<tr>
<th>Region</th>
<th>No. of Timesteps</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>341</td>
<td>4.2</td>
</tr>
<tr>
<td>2</td>
<td>3339</td>
<td>41.1</td>
</tr>
<tr>
<td>3</td>
<td>79</td>
<td>1.0</td>
</tr>
<tr>
<td>4</td>
<td>2585</td>
<td>31.8</td>
</tr>
<tr>
<td>5</td>
<td>734</td>
<td>9.0</td>
</tr>
<tr>
<td>6</td>
<td>119</td>
<td>1.5</td>
</tr>
<tr>
<td>7</td>
<td>531</td>
<td>6.5</td>
</tr>
<tr>
<td>8</td>
<td>73</td>
<td>0.9</td>
</tr>
<tr>
<td>9</td>
<td>323</td>
<td>4.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>8124</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

**Region 2 Summary**

<table>
<thead>
<tr>
<th>Broader spectrum</th>
<th>No. of Timesteps</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broader spectrum</td>
<td>66</td>
<td>0.8</td>
</tr>
<tr>
<td>Narrower spectrum</td>
<td>3273</td>
<td>40.3</td>
</tr>
</tbody>
</table>

**Region 4 Summary**

<table>
<thead>
<tr>
<th>Broader spectrum</th>
<th>No. of Timesteps</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broader spectrum</td>
<td>2523</td>
<td>31.1</td>
</tr>
<tr>
<td>Narrower spectrum</td>
<td>62</td>
<td>0.8</td>
</tr>
</tbody>
</table>
Table 21: Validation with $\Delta E_{\text{max}}$ versus $\Delta H_s$ matrix for unimodal spectrums at Moray Firth

![Diagram](image)

<table>
<thead>
<tr>
<th>Region</th>
<th>No. of Timesteps</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>201</td>
<td>2.6</td>
</tr>
<tr>
<td>2</td>
<td>4688</td>
<td>61.1</td>
</tr>
<tr>
<td>3</td>
<td>209</td>
<td>2.7</td>
</tr>
<tr>
<td>4</td>
<td>1507</td>
<td>19.7</td>
</tr>
<tr>
<td>5</td>
<td>402</td>
<td>5.2</td>
</tr>
<tr>
<td>6</td>
<td>83</td>
<td>1.1</td>
</tr>
<tr>
<td>7</td>
<td>416</td>
<td>5.4</td>
</tr>
<tr>
<td>8</td>
<td>25</td>
<td>0.3</td>
</tr>
<tr>
<td>9</td>
<td>136</td>
<td>1.8</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>7667</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

**Region 2 Summary**

<table>
<thead>
<tr>
<th>No. of Timesteps</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broader spectrum</td>
<td>57</td>
</tr>
<tr>
<td>Narrower spectrum</td>
<td>4631</td>
</tr>
</tbody>
</table>

**Region 4 Summary**

<table>
<thead>
<tr>
<th>No. of Timesteps</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broader spectrum</td>
<td>1425</td>
</tr>
<tr>
<td>Narrower spectrum</td>
<td>82</td>
</tr>
</tbody>
</table>
Table 22: Validation with $\Delta E_{\text{max}}$ versus $\Delta H_s$ matrix for unimodal spectrums at West of Hebrides

<table>
<thead>
<tr>
<th>Region</th>
<th>No. of Timesteps</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>703</td>
<td>11.4</td>
</tr>
<tr>
<td>2</td>
<td>3241</td>
<td>52.4</td>
</tr>
<tr>
<td>3</td>
<td>221</td>
<td>3.6</td>
</tr>
<tr>
<td>614</td>
<td>1316</td>
<td>21.3</td>
</tr>
<tr>
<td>5</td>
<td>392</td>
<td>6.3</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>0.2</td>
</tr>
<tr>
<td>7</td>
<td>275</td>
<td>4.4</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>0.2</td>
</tr>
<tr>
<td>9</td>
<td>14</td>
<td>0.2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>6182</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

**Region 2 Summary**

<table>
<thead>
<tr>
<th>Broader spectrum</th>
<th>No. of Timesteps</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>213</td>
<td>3.4</td>
</tr>
<tr>
<td>Narrower spectrum</td>
<td>3028</td>
<td>49.0</td>
</tr>
</tbody>
</table>

**Region 4 Summary**

<table>
<thead>
<tr>
<th>Broader spectrum</th>
<th>No. of Timesteps</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1249</td>
<td>20.2</td>
</tr>
<tr>
<td>Narrower spectrum</td>
<td>67</td>
<td>1.1</td>
</tr>
</tbody>
</table>
In general, most of the points are distributed in region 2 (Moray Firth 61.1%, West of Hebrides 52.4%, Hastings 41.1%) where the predicted wave spectrums have lower peak energy densities and smaller area below the curves, in comparison to the observed wave spectrums. Points are also clustered in region 4 (Hastings 31.8%, West of Hebrides 21.3%, Moray Firth 19.7%) where the peak energy densities and the area below the curves are over predicted. With its greater exposure to waves propagating from the North Atlantic Ocean, the points at West of Hebrides seem to be more scattered than at Hastings and Moray Firth.

Due to this inherent bias, the scatter does not imply that the model results are poorer at West of Hebrides. This is contrary to the findings from the previous section. The original validation matrices are used to illustrate the absolute scale of difference between the primary parameters – but they are not statistical measures like correlation or the p-value between parameters.

Nonetheless, to complement the above analysis, the x- and y-axes can also be normalized to remove the scale factor in modified validation matrices.

Figure 36: Validation with modified matrix ($\Delta E_{\text{max}}$/Observed $E_{\text{max}}$ versus $\Delta H_s$/Observed $H_s$) for unimodal spectrums at Hastings
Figure 37: Validation with modified matrix ($\Delta E_{\text{max}}/\text{Observed } E_{\text{max}}$ versus $\Delta H_s/\text{Observed } H_s$) for unimodal spectrums at Moray Firth

Figure 38: Validation with modified matrix ($\Delta E_{\text{max}}/\text{Observed } E_{\text{max}}$ versus $\Delta H_s/\text{Observed } H_s$) for unimodal spectrums at West of Hebrides
With the modified validation matrix, the maximum attainable value on both axes is +1, which indicates that the observed $H_s$ or $E_{\text{max}}$ is underestimated by a hundred-percent. On the other hand, a negative value would indicate how many times the above parameters are being overestimated. It can be observed that the points at West of Hebrides are the least dispersed (Figure 36 to Figure 38). At Hastings and Moray Firth, there are instances where the $H_s$ are under predicted by more than 50%.

However, because the two sites are also more sheltered, it could be possible that these points correspond to calmer sea conditions. From an engineering viewpoint, a 10 m wave record that is under predicted by 30% would be a bigger problem than a 0.4 m wave record that is underestimated by 80%, especially in calculating the design wave conditions for extreme storm events (e.g. 100 or 1000 years wave return period) from a short period measurement. Perhaps a more practical approach would be to select the top 5-10% of the most energetic waves only for the analysis. A qualitative comparison can then be made on the differences in magnitude and shape between the observed and predicted spectrums using the matrices.

The second assessment involves the application of the two-dimensional matrix between $\Delta E_{\text{max}}$ and $\Delta f_p$ in Figure 19.

Table 23: Boundaries of region 1 to region 9 for two-dimensional matrix between $\Delta E_{\text{max}}$ and $\Delta f_p$

<table>
<thead>
<tr>
<th>Region</th>
<th>$\Delta f_p$ (Hz)</th>
<th>$\Delta E_{\text{max}}$ (m$^2$/Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>&lt; -0.001</td>
<td>&gt; 0.01</td>
</tr>
<tr>
<td>2</td>
<td>&gt; 0.001</td>
<td>&gt; 0.01</td>
</tr>
<tr>
<td>3</td>
<td>&gt; 0.001</td>
<td>&lt; -0.01</td>
</tr>
<tr>
<td>4</td>
<td>&lt; -0.001</td>
<td>&lt; -0.01</td>
</tr>
<tr>
<td>5</td>
<td>-0.001 to 0.001</td>
<td>&gt; 0.01</td>
</tr>
<tr>
<td>6</td>
<td>&gt; 0.001</td>
<td>-0.01 to 0.01</td>
</tr>
<tr>
<td>7</td>
<td>-0.001 to 0.001</td>
<td>&lt; -0.01</td>
</tr>
<tr>
<td>8</td>
<td>&lt; -0.001</td>
<td>-0.01 to 0.01</td>
</tr>
<tr>
<td>9</td>
<td>-0.001 to 0.001</td>
<td>-0.01 to 0.01</td>
</tr>
</tbody>
</table>
Table 24: Validation with $\Delta E_{\text{max}}$ versus $\Delta f_p$ matrix for unimodal spectrums at Hastings

<table>
<thead>
<tr>
<th>Region</th>
<th>No. of Timesteps</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1767</td>
<td>21.8</td>
</tr>
<tr>
<td>2</td>
<td>2554</td>
<td>31.4</td>
</tr>
<tr>
<td>3</td>
<td>2406</td>
<td>29.6</td>
</tr>
<tr>
<td>4</td>
<td>722</td>
<td>8.9</td>
</tr>
<tr>
<td>5</td>
<td>93</td>
<td>1.1</td>
</tr>
<tr>
<td>6</td>
<td>324</td>
<td>4.0</td>
</tr>
<tr>
<td>7</td>
<td>67</td>
<td>0.8</td>
</tr>
<tr>
<td>8</td>
<td>184</td>
<td>2.3</td>
</tr>
<tr>
<td>9</td>
<td>7</td>
<td>0.1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>8124</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>
Table 25: Validation with $\Delta E_{\text{max}}$ versus $\Delta f_p$ matrix for unimodal spectrums at Moray Firth

<table>
<thead>
<tr>
<th>Region</th>
<th>No. of Timesteps</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2336</td>
<td>30.5</td>
</tr>
<tr>
<td>2</td>
<td>2820</td>
<td>36.8</td>
</tr>
<tr>
<td>3</td>
<td>1473</td>
<td>19.2</td>
</tr>
<tr>
<td>4</td>
<td>605</td>
<td>7.9</td>
</tr>
<tr>
<td>5</td>
<td>135</td>
<td>1.8</td>
</tr>
<tr>
<td>6</td>
<td>135</td>
<td>1.8</td>
</tr>
<tr>
<td>7</td>
<td>54</td>
<td>0.7</td>
</tr>
<tr>
<td>8</td>
<td>107</td>
<td>1.4</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>0.0</td>
</tr>
<tr>
<td>Total</td>
<td>7667</td>
<td>100</td>
</tr>
</tbody>
</table>
Table 26: Validation with $\Delta E_{\text{max}}$ versus $\Delta f_p$ matrix for unimodal spectrums at West of Hebrides

<table>
<thead>
<tr>
<th>Region</th>
<th>No. of Timesteps</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2239</td>
<td>36.2</td>
</tr>
<tr>
<td>2</td>
<td>1710</td>
<td>27.7</td>
</tr>
<tr>
<td>3</td>
<td>1035</td>
<td>16.7</td>
</tr>
<tr>
<td>4</td>
<td>675</td>
<td>10.9</td>
</tr>
<tr>
<td>5</td>
<td>387</td>
<td>6.3</td>
</tr>
<tr>
<td>6</td>
<td>15</td>
<td>0.2</td>
</tr>
<tr>
<td>7</td>
<td>102</td>
<td>1.6</td>
</tr>
<tr>
<td>8</td>
<td>17</td>
<td>0.3</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>0.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>6182</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

At all three sites, a cross-like pattern emerges from the validation matrix, with most points clustered in the first three quadrants and very close to the axes. Along the horizontal axis, the points are more dispersed at Hastings and Moray Firth, representing larger offsets in the position of the peak frequencies than at West of Hebrides. This in agreement with the findings from section 5.2.2.
The cross-like pattern brings out a new perspective on the model performance. It shows that when the peak energy densities are severely under- or overestimated, the position of the peak frequencies are often close between the observed and predicted spectrums. It is suspected that these conditions are associated with long period, high energy wave events. On the other hand, when the peak frequencies are further apart from each other, the differences in the peak energy densities would tend to be smaller. These could be associated with shorter period, low energy wave events during calm sea conditions. In general, the pattern is a good indicator of the model performance – at least one of the two parameters has a good agreement at each time step.

In conclusion, a joint analysis of the primary parameters has been carried out with the two-dimensional validation matrices to provide better, qualitative assessment of the unimodal spectrums. The matrices should be used to complement the statistical measures of correlation and p-value in evaluating the model results.

5.2.4 Qualitative Assessment of Bimodal Spectrums

Further evaluation of the bimodal and trimodal spectrums has not been carried out in this example. A qualitative but separate assessment of the first, second and third modes can still be done using the same two-dimensional matrices, in addition to the predicted versus observed parameter plots. The range of frequencies where the peaks are not well represented can then be identified from these figures (e.g. the first mode would consist of peaks in lower frequency region than the subsequent modes).
6 Conclusions and Recommendations

A new approach for validating numerical wave models has been developed by applying new parameterisation to the frequency wave spectrum. This new method provides more details about the wave spectrums, in comparison to the conventional approach of using only sea-state parameters, such as significant wave height, peak wave period, mean wave period and/or mean wave direction. It is a robust approach, in the sense that its application is not limited to any specific spectral conditions.

A series of parameters and methodologies from wave mechanics and other disciplines have been reviewed to better define the wave spectrum. Instead of a single parameter, a family of six parameters including the significant wave height, peak frequency, peak energy density, squared Euclidean distance, skewness and kurtosis are required to best describe the characteristic differences between an observed and predicted wave spectrum. A seventh parameter, which is the mean width deviation has been developed specifically by the author to complement these six parameters. The mean width deviation quantifies the spectral width difference between an observed and predicted wave spectrum.

The parameters have been categorized into primary and secondary parameters, based on their level of detail and ease of interpretation. Two-dimensional validation matrices have then been developed to illustrate the relationship between these primary parameters with the magnitude, shape and position of the wave spectrums. Parallel analysis of the parameters reveals more qualitative information about the spectrums, than when an individual assessment of each parameter is undertaken in isolation.

The feasibility of the new approach has been proven through the validation of a hindcast spectral wave model at three nearshore sites in UK Waters. Analysis with time series plots has shown variation in the model performance, in both the temporal and spatial domains. In addition, the validation matrices have been used to complement the statistical measures of correlation and p-value in evaluating the model results.
One of the main limitations of the new validation approach is the absence of criteria for defining whether the model results should be accepted or not accepted. The main problem that has been encountered is establishing significance from the squared Euclidean distances. Therefore, future studies are recommended to determine whether objective conditions can be established from the relationship between the distances and other input or output parameters (e.g. the distance should be less than X for wind speed ranging from 0 to 5 m/s and less than Y for wind speed above 5 m/s). Future studies are also suggested to better assess the application of the new approach for analyzing bimodal and trimodal wave spectrums. More comparisons should be made on regional or global sites with different wave characteristics.

Lastly, the new approach could also be potentially developed into a calibration tool for modellers. Certain spectral conditions could be associated with specific model parameters (e.g. bottom friction or energy dissipation from white-capping). A more comprehensive sensitivity analysis should be carried out in order to establish a better understanding on how the characteristics of the predicted wave spectrums are influenced by these input parameters.
References


Hardisty, J. 1990. The British Seas: An Introduction to the Oceanography and Resources of the North-West European Continental Shelf. Routledge, UK.


Mason, T. Personal communication, 30 July 2013.


Appendix A1

The following MATLAB® 2012 functions have been developed to compare between the observed and simulated spectrums in Chapter 5:

- **main.m**: Defines the name of the input files (ASCII format) and optional output file (Excel spreadsheet). Calls other functions to read, process and write data.
- **timestep.m**: Fills the gaps in the WaveNet data with “NaN”. Downsampling the data frequency to the same time step interval (1 hour) as the model results.
- **modality.m**: Identifies the number of modes from the WaveNet spectral data (e.g. unimodal, bimodal or trimodal distribution).
- **partition.m**: A sub-function of modality.m that executes the spectral partitioning algorithm (Figure 8 and Figure 9)
- **spectrastat.m**: Interpolates the WaveNet spectral data to the same frequency interval as the wave model result. Calculates the seven validation parameters and other statistical parameters (e.g. root mean square error, scatter index, bias, correlation coefficient, p-value) for comparison between the observed and predicted data.
- **bimodal.m**: Similar to spectrastat.m but computes the validation parameters for two separate modes in a bimodal distribution.
- **trimodal.m**: Similar to spectrastat.m but computes the validation parameters for three separate modes in a trimodal distribution.

In spectrastat.m, a linear interpolation method has been used to interpolate the WaveNet data. A cubic spline interpolation is not recommended because the peaks in the spectrum could be smoothened or flattened out.
spectrastat.m

clc; clear all;

%===============================================================================
%  INPUT DATA
%===============================================================================
% WaveNetID = Name of WaveNet Spectral Wave Data
% M21ID = Name of M21 Spectral Wave Result
% xlsID = Name of Output File.
% writexls = 0 to suppress the write output section.
% = Other values to execute the write output section.
% NOTE: START and END time for both files must be identical

WaveNetID = 'WN_Hastings_20090101to20091231.txt';
M21ID = 'M21_Hastings_20090101to20091231.txt';
xlsID = 'ver6_Hastings_20090101to20091231'; writexls = 1;
M21NumOfFreq = 201; M21FreqInterval = 0.001702;

tic
%===============================================================================
% READ AND PROCESS
%===============================================================================

[DateArray,f,E]=timestep(WaveNetID);
[ModeID,Bimodal,Trimodal,ModeSummary]=modality(f,E);
[Euclid,deltaMw,ParaM21,ParaWaveNet,HsStat,PeakStat,RegionStat,RegionStat2,RegionStat4]=spectrastat(f,E,M21ID,M21NumOfFreq,M21FreqInterval,ModeID,ModeSummary);
[Bi_ParaWaveNet,Bi_ParaM21,Bi_deltaMw,Bi_Euclid]=bimodal(f,E,M21ID,M21NumOfFreq,M21FreqInterval,Bimodal);
[Tri_ParaWaveNet,Tri_ParaM21,Tri_deltaMw,Tri_Euclid]=trimodal(f,E,M21ID,M21NumOfFreq,M21FreqInterval,Trimodal);

fprintf('Read and Process: ');
toc;

if writexls ~= 0
    tic
    warning('off','MATLAB:xlswrite:AddSheet');
%===============================================================================
% WRITE OUTPUT
%===============================================================================
    % Sheet 1
    xlswrite(xlsID, {'Year'}, 1, 'A1');
    xlswrite(xlsID, {'Month'}, 1, 'B1');
    xlswrite(xlsID, {'Day'}, 1, 'C1');
    xlswrite(xlsID, {'Hour'}, 1, 'D1');
    xlswrite(xlsID, {'Min'}, 1, 'E1');
    xlswrite(xlsID, {'Sec'}, 1, 'F1');
    xlswrite(xlsID, DateArray, 1, 'A2');

    xlswrite(xlsID, {'Hs (WaveNet)'}, 1, 'H1');
    xlswrite(xlsID, ParaWaveNet(:,1), 1, 'H2');
    xlswrite(xlsID, {'Hs (M21)'}, 1, 'I1');
    xlswrite(xlsID, ParaM21(:,1), 1, 'I2');
    xlswrite(xlsID, {'fp (WaveNet)'}, 1, 'J1');
    xlswrite(xlsID, ParaWaveNet(:,2), 1, 'J2');
    xlswrite(xlsID, {'fp (M21)'}, 1, 'K1');
    xlswrite(xlsID, ParaM21(:,2), 1, 'K2');
    xlswrite(xlsID, {'Emax (WaveNet)'}, 1, 'L1');
    xlswrite(xlsID, ParaWaveNet(:,3), 1, 'L2');
    xlswrite(xlsID, {'Emax (M21)'}, 1, 'M1');
    xlswrite(xlsID, ParaM21(:,3), 1, 'M2');
xlswrite(xlsID, {'Squared Euclid'}, 1, 'N1');
xlswrite(xlsID, Euclid(:,1), 1, 'N2');
xlswrite(xlsID, {'Modality'}, 1, 'O1');
xlswrite(xlsID, ModeID, 1, 'O2');
xlswrite(xlsID, {'Hs Stat'}, 1, 'Q2');
xlswrite(xlsID, {'RMSE'}, 1, 'Q3');
xlswrite(xlsID, {'Bias'}, 1, 'Q4');
xlswrite(xlsID, {'Scatter Index'}, 1, 'Q5');
xlswrite(xlsID, {'Correlation'}, 1, 'Q6');
xlswrite(xlsID, {'p-value'}, 1, 'Q7');
xlswrite(xlsID, HsStat(:,1), 1, 'R3');
xlswrite(xlsID, {'Emax Stat'}, 1, 'T2');
xlswrite(xlsID, {'RMSE'}, 1, 'T3');
xlswrite(xlsID, {'Bias'}, 1, 'T4');
xlswrite(xlsID, {'Scatter Index'}, 1, 'T5');
xlswrite(xlsID, {'Correlation'}, 1, 'T6');
xlswrite(xlsID, {'p-value'}, 1, 'T7');
xlswrite(xlsID, PeakStat(:,1), 1, 'U3');
xlswrite(xlsID, {'fp Stat'}, 1, 'W2');
xlswrite(xlsID, {'RMSE'}, 1, 'W3');
xlswrite(xlsID, {'Bias'}, 1, 'W4');
xlswrite(xlsID, {'Scatter Index'}, 1, 'W5');
xlswrite(xlsID, {'Correlation'}, 1, 'W6');
xlswrite(xlsID, {'p-value'}, 1, 'W7');
xlswrite(xlsID, PeakStat(:,2), 1, 'X3');

temp=(1:9)';
xlswrite(xlsID, {'Region Summary (delEmax vs delHs)'}, 1, 'Q10');
xlswrite(xlsID, temp, 1, 'Q11');
xlswrite(xlsID, RegionStat(1,:)', 1, 'R11');
xlswrite(xlsID, RegionStat(2,:)', 1, 'S11');
xlswrite(xlsID, {'Region Summary (delEmax vs delfp)'}, 1, 'U10');
xlswrite(xlsID, temp, 1, 'U11');
xlswrite(xlsID, RegionStat(3,:)', 1, 'V11');
xlswrite(xlsID, RegionStat(4,:)', 1, 'W11');
xlswrite(xlsID, {'Region 2 Summary'}, 1, 'Q23');
xlswrite(xlsID, {'Broader'}, 1, 'Q24');
xlswrite(xlsID, {'Narrower'}, 1, 'Q25');
xlswrite(xlsID, RegionStat2(1,:)', 1, 'R24');
xlswrite(xlsID, RegionStat2(2,:)', 1, 'S24');
xlswrite(xlsID, {'Region 4 Summary'}, 1, 'Q27');
xlswrite(xlsID, {'Broader'}, 1, 'Q28');
xlswrite(xlsID, {'Narrower'}, 1, 'Q29');
xlswrite(xlsID, RegionStat4(1,:)', 1, 'R28');
xlswrite(xlsID, RegionStat4(2,:)', 1, 'S28');
xlswrite(xlsID, {'Modal Distribution'}, 1, 'U23');
xlswrite(xlsID, {'Unimodal'}, 1, 'U24');
xlswrite(xlsID, {'Bimodal'}, 1, 'U25');
xlswrite(xlsID, {'Trimodal or More'}, 1, 'U26');
xlswrite(xlsID, {'Missing Data'}, 1, 'U27');
xlswrite(xlsID, ModeSummary(1,:)', 1, 'V24');
xlswrite(xlsID, ModeSummary(2,:)', 1, 'W24');

% Sheet 2
xlswrite(xlsID, {'Year'}, 2, 'A1');
xlswrite(xlsID, {'Month'}, 2, 'B1');
xlswrite(xlsID, {'Day'}, 2, 'C1');
xlswrite(xlsID, {'Hour'}, 2, 'D1');
xlswrite(xlsID, {'Min'}, 2, 'E1');
xlswrite(xlsID, {'Sec'}, 2, 'F1');
xlswrite(xlsID, DateArray, 2, 'A2');
xlswrite(xlsID, {'Mean Width Dev'}, 2, 'H1');
xlswrite(xlsID, deltaMw(:,1), 2, 'H2');
xlswrite(xlsID, {'Sk (WaveNet)'}, 2, 'I1');
xlswrite(xlsID, ParaWaveNet(:,5), 2, 'I2');
xlswrite(xlsID, {'Sk (Mike21)'}, 2, 'J1');
xlswrite(xlsID, ParaM21(:,5), 2, 'J2');
xlswrite(xlsID, {'K (WaveNet)'}, 2, 'K1');
xlswrite(xlsID, ParaWaveNet(:,6), 2, 'K2');
xlswrite(xlsID, {'K (Mike21)'}, 2, 'L1');
xlswrite(xlsID, ParaM21(:,6), 2, 'L2');

% Sheet 3
xlswrite(xlsID, {'Bimodal_WaveNet'}, 3, 'A1');
xlswrite(xlsID, {'Year'}, 3, 'A3');
xlswrite(xlsID, {'Month'}, 3, 'B3');
xlswrite(xlsID, {'Day'}, 3, 'C3');
xlswrite(xlsID, {'Hour'}, 3, 'D3');
xlswrite(xlsID, {'Min'}, 3, 'E3');
xlswrite(xlsID, {'Sec'}, 3, 'F3');
xlswrite(xlsID, DateArray, 3, 'A4');
xlswrite(xlsID, {'Partition'}, 3, 'H3');
xlswrite(xlsID, Bimodal, 3, 'H4');
xlswrite(xlsID, {'Mode 1'}, 3, 'J1');
xlswrite(xlsID, {'Hs'}, 3, 'J3');
xlswrite(xlsID, {'fp'}, 3, 'K3');
xlswrite(xlsID, {'Emax'}, 3, 'L3');
xlswrite(xlsID, {'Sk'}, 3, 'M3');
xlswrite(xlsID, {'K'}, 3, 'N3');
xlswrite(xlsID, {'Mode 2'}, 3, 'P1');
xlswrite(xlsID, {'Hs'}, 3, 'P3');
xlswrite(xlsID, {'fp'}, 3, 'Q3');
xlswrite(xlsID, {'Emax'}, 3, 'R3');
xlswrite(xlsID, {'Sk'}, 3, 'S3');
xlswrite(xlsID, {'K'}, 3, 'T3');
xlswrite(xlsID, Bi_ParaWaveNet, 3, 'J4');

% Sheet 4
xlswrite(xlsID, {'Bimodal_M21'}, 4, 'A1');
xlswrite(xlsID, {'Year'}, 4, 'A3');
xlswrite(xlsID, {'Month'}, 4, 'B3');
xlswrite(xlsID, {'Day'}, 4, 'C3');
xlswrite(xlsID, {'Hour'}, 4, 'D3');
xlswrite(xlsID, {'Min'}, 4, 'E3');
xlswrite(xlsID, {'Sec'}, 4, 'F3');
xlswrite(xlsID, DateArray, 4, 'A4');
xlswrite(xlsID, {'Partition'}, 4, 'H3');
xlswrite(xlsID, Bimodal, 4, 'H4');
xlswrite(xlsID, {'Mode 1'}, 4, 'J1');
xlswrite(xlsID, {'Hs'}, 4, 'J3');
xlswrite(xlsID, {'fp'}, 4, 'K3');
xlswrite(xlsID, {'Emax'}, 4, 'L3');
xlswrite(xlsID, {'Sk'}, 4, 'M3');
xlswrite(xlsID, {'K'}, 4, 'N3');
xlswrite(xlsID, {'Mode 2'}, 4, 'P1');
xlswrite(xlsID, {'Hs'}, 4, 'P3');
xlswrite(xlsID, {'fp'}, 4, 'Q3');
xlswrite(xlsID, {'Emax'}, 4, 'R3');
xlswrite(xlsID, {'Sk'}, 4, 'S3');
xlswrite(xlsID, {'K'}, 4, 'T3');
xlswrite(xlsID, Bi_ParaM21, 4, 'J4');

% Sheet 5
xlswrite(xlsID, {'Bimodal'}, 5, 'A1');
xlswrite(xlsID, {'Year'}, 5, 'A3');
xlswrite(xlsID, {'Month'}, 5, 'B3');
xlswrite(xlsID, {'Day'}, 5, 'C3');
xlswrite(xlsID, {'Hour'}, 5, 'D3');
xlswrite(xlsID, {'Min'}, 5, 'E3');
xlswrite(xlsID, {'Sec'}, 5, 'F3');
xlswrite(xlsID, DateArray, 5, 'A4');
xlswrite(xlsID, {'Mode 1'}, 5, 'H1');
xlswrite(xlsID, {'Mean Width Dev'}, 5, 'H3');
xlswrite(xlsID, {'Squared Euclid'}, 5, 'I3');
xlswrite(xlsID, Bi_deltaMw(:,1), 5, 'H4');
xlswrite(xlsID, Bi_Euclid(:,1), 5, 'I4');
xlswrite(xlsID, {'Mode 2'}, 5, 'K1');
xlswrite(xlsID, {'Mean Width Dev'}, 5, 'K3');
xlswrite(xlsID, {'Squared Euclid'}, 5, 'L3');
xlswrite(xlsID, Bi_deltaMw(:,2), 5, 'K4');
xlswrite(xlsID, Bi_Euclid(:,2), 5, 'L4');

% Sheet 6
xlswrite(xlsID, {'Trimodal_WaveNet'}, 6, 'A1');
xlswrite(xlsID, {'Year'}, 6, 'A3');
xlswrite(xlsID, {'Month'}, 6, 'B3');
xlswrite(xlsID, {'Day'}, 6, 'C3');
xlswrite(xlsID, {'Hour'}, 6, 'D3');
xlswrite(xlsID, {'Min'}, 6, 'E3');
xlswrite(xlsID, {'Sec'}, 6, 'F3');
xlswrite(xlsID, DateArray, 6, 'A4');
xlswrite(xlsID, {'Partition'}, 6, 'H3');
xlswrite(xlsID, Trimplodal, 6, 'H4');
xlswrite(xlsID, {'Mode 1'}, 6, 'K1');
xlswrite(xlsID, {'Hs'}, 6, 'K3');
xlswrite(xlsID, {'fp'}, 6, 'L3');
xlswrite(xlsID, {'Emax'}, 6, 'M3');
xlswrite(xlsID, {'Sk'}, 6, 'N3');
xlswrite(xlsID, {'K'}, 6, 'O3');
xlswrite(xlsID, {'Mode 2'}, 6, 'Q1');
xlswrite(xlsID, {'Hs'}, 6, 'Q3');
xlswrite(xlsID, {'fp'}, 6, 'R3');
xlswrite(xlsID, {'Emax'}, 6, 'S3');
xlswrite(xlsID, {'Sk'}, 6, 'T3');
xlswrite(xlsID, {'K'}, 6, 'U3');
xlswrite(xlsID, {'Mode 3'}, 6, 'W1');
xlswrite(xlsID, {'Hs'}, 6, 'W3');
xlswrite(xlsID, {'fp'}, 6, 'X3');
xlswrite(xlsID, {'Emax'}, 6, 'Y3');
<table>
<thead>
<tr>
<th>Sheet 7</th>
<th>Sheet 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>xlswrite(xlsID, {'Sk'}, 6, 'Z3');</td>
<td>xlswrite(xlsID, {'Sk'}, 6, 'Z3');</td>
</tr>
<tr>
<td>xlswrite(xlsID, {'K'}, 6, 'AA3');</td>
<td>xlswrite(xlsID, {'K'}, 6, 'AA3');</td>
</tr>
<tr>
<td>xlswrite(xlsID, Tri_ParaWaveNet, 6, 'K4');</td>
<td>xlswrite(xlsID, Tri_ParaM21, 7, 'K4');</td>
</tr>
</tbody>
</table>

% Sheet 7
xlswrite(xlsID, {'Trimodal M21'}, 7, 'A1');
xlswrite(xlsID, {'Year'}, 7, 'A3');
xlswrite(xlsID, {'Month'}, 7, 'B3');
xlswrite(xlsID, {'Day'}, 7, 'C3');
xlswrite(xlsID, {'Hour'}, 7, 'D3');
xlswrite(xlsID, {'Min'}, 7, 'E3');
xlswrite(xlsID, {'Sec'}, 7, 'F3');
xlswrite(xlsID, DateArray, 7, 'A4');

xlswrite(xlsID, {'Partition'}, 7, 'H3');
xlswrite(xlsID, Trimodal, 7, 'H4');
xlswrite(xlsID, {'Mode 1'}, 7, 'K1');
xlswrite(xlsID, {'Hs'}, 7, 'K3');
xlswrite(xlsID, {'fp'}, 7, 'L3');
xlswrite(xlsID, {'Emax'}, 7, 'M3');
xlswrite(xlsID, {'Sk'}, 7, 'N3');
xlswrite(xlsID, {'K'}, 7, 'O3');
xlswrite(xlsID, {'Mode 2'}, 7, 'Q1');
xlswrite(xlsID, {'Hs'}, 7, 'Q3');
xlswrite(xlsID, {'fp'}, 7, 'R3');
xlswrite(xlsID, {'Emax'}, 7, 'S3');
xlswrite(xlsID, {'Sk'}, 7, 'T3');
xlswrite(xlsID, {'K'}, 7, 'U3');
xlswrite(xlsID, {'Mode 3'}, 7, 'W1');
xlswrite(xlsID, {'Hs'}, 7, 'W3');
xlswrite(xlsID, {'fp'}, 7, 'X3');
xlswrite(xlsID, {'Emax'}, 7, 'Y3');
xlswrite(xlsID, {'Sk'}, 7, 'Z3');
xlswrite(xlsID, {'K'}, 7, 'AA3');
xlswrite(xlsID, Tri_ParaM21, 7, 'K4');

% Sheet 8
xlswrite(xlsID, {'Trimodal'}, 8, 'A1');
xlswrite(xlsID, {'Year'}, 8, 'A3');
xlswrite(xlsID, {'Month'}, 8, 'B3');
xlswrite(xlsID, {'Day'}, 8, 'C3');
xlswrite(xlsID, {'Hour'}, 8, 'D3');
xlswrite(xlsID, {'Min'}, 8, 'E3');
xlswrite(xlsID, {'Sec'}, 8, 'F3');
xlswrite(xlsID, DateArray, 8, 'A4');

xlswrite(xlsID, {'Mode 1'}, 8, 'H1');
xlswrite(xlsID, {'Mean Width Dev'}, 8, 'H3');
xlswrite(xlsID, {'Squared Euclid'}, 8, 'I3');
xlswrite(xlsID, Tri_deltaMw(:,1), 8, 'H4');
xlswrite(xlsID, Tri_Euclid(:,1), 8, 'I4');

xlswrite(xlsID, {'Mode 2'}, 8, 'K1');
xlswrite(xlsID, {'Mean Width Dev'}, 8, 'K3');
xlswrite(xlsID, {'Squared Euclid'}, 8, 'L3');
xlswrite(xlsID, Tri_deltaMw(:,2), 8, 'K4');
xlswrite(xlsID, Tri_Euclid(:,2), 8, 'L4');

xlswrite(xlsID, {'Mode 3'}, 8, 'N1');
xlswrite(xlsID, {'Mean Width Dev'}, 8, 'N3');
clear all;

fprintf('Write: '); toc;
end

fprintf('
Operation done
');
timestep.m

function [DateArray,f,E]=timestep(WaveNetID)

[Year,Month,Day,Hr,Min,Sec,f,E] = textread(WaveNetID,'%f-%f-%f
%f:%f:%f %f %*f %*f %*f %*f','delimiter',','); %#ok<*REMFF1>

DateInput=datenum(Year,Month,Day,Hr,Min,Sec);
NumDays=ceil(daysact(DateInput(1),DateInput(length(DateInput))));
DateInput=datevec(DateInput);

% 1. Create an empty date array
DateArrayTemp=datenum(Year(1),Month(1),Day(1),0,0:30:NumDays*30*48-1,0)';
DateArrayTemp=datevec(DateArrayTemp);

for j=1:length(DateArrayTemp)
    for k=0:13
        DateArray1(14*j-k,:)=DateArrayTemp(j,:);
    end
end

clear DateArrayTemp;

% 2. Assign frequency and spectral energy to the new date array
k = 0;
for i=1:length(DateInput);
    if DateArray1(i+k,:) == DateInput(i,:)
        f1(i+k)=f(i);
        E1(i+k)=E(i);
    else
        for j=1:14
            f1(i+k)=0; %Fill gaps with '0' value
            E1(i+k)=0; %Fill gaps with '0' value
            k=k+1;
        end
        f1(i+k)=f(i);
        E1(i+k)=E(i);
    end
end
f1=f1'; E1=E1';

% 3. Downsampling to one hour intervals
DateArrayTemp=datenum(Year(1),Month(1),Day(1),0:1:NumDays*24-1,0,0)';
DateArrayTemp=datevec(DateArrayTemp);

for j=1:length(DateArrayTemp)
    for k=0:13
        DateArray2(14*j-k,:)=DateArrayTemp(j,:);
    end
end

k = 0;
for i=1:length(DateArray2);
    if DateArray2(i,:) == DateArray1(i+k,:)
        f2(i)=f1(i+k);
E2(i) = E1(i+k);

else
    k = k+14;
    f2(i) = f1(i+k);
    E2(i) = E1(i+k);
end

end

f2 = f2'; E2 = E2';

DateArray = DateArrayTemp;
f = f2; E = E2;

clearvars -except f E DateArray DateArrayTemp;
modality.m

function [ModeID,Bimodal,Trimodal,ModeSummary]=modality(f,E)

MaxTS = length(f)/14; Bimodal=NaN(MaxTS,1); Trimodal=NaN(MaxTS,2);

for i=1:MaxTS
    [f1]=f((i*14)-13:i*14);
    [E1]=E((i*14)-13:i*14);
    [part]=partition(f1,E1);

    % Unimodal Distribution
    if isempty(part)==1;
        ModeID(i,1)=1;
    end

    % More than one modal distribution
    if isempty(part)==0 & isnan(part)==0;
        ModeID(i,1)=length(part)+1;
    end
    if ModeID(i,1)==2;
        Bimodal(i)=part;
    end
    if ModeID(i,1)==3;
        if part(1) < part(2)
            Trimodal(i,1)=part(1);
            Trimodal(i,2)=part(2);
        else
            Trimodal(i,1)=part(2);
            Trimodal(i,2)=part(1);
        end
    end

end

% Missing data
if isempty(part)==0 & isnan(part)==1;
    ModeID(i,1)=NaN;
end

end

ModeSummary(1,1) = sum(ModeID==1); % Unimodal Distribution
ModeSummary(1,2) = sum(ModeID==2); % Bimodal Distribution
ModeSummary(1,3) = sum(ModeID==3); % Trimodal Distribution
ModeSummary(1,4) = sum(isnan(ModeID)); % Missing data
ModeSummary(2,1) = ModeSummary(1,1)/MaxTS*100;
ModeSummary(2,2) = ModeSummary(1,2)/MaxTS*100;
ModeSummary(2,3) = ModeSummary(1,3)/MaxTS*100;
ModeSummary(2,4) = ModeSummary(1,4)/MaxTS*100;
partition.m

function [part] = partition(f,E)

if sum(f) == 0 || sum(E) == 0 || length(unique(f(:)))~ length(f(:))
    part = NaN;
else
    me = max(E);

%====================================================================
% FORWARD SCHEME
%====================================================================

counter = max(size(f));
for i=1:counter
    if E(i) == me
        fwdf = f(i:counter);
        fwdE = E(i:counter);
    end
end

% 1. Measures the slope between two successive points
counter = length(fwdE);
slope = NaN(counter-1,1);
for i=1:counter-1
    slope(i)=((fwdE(i+1)-fwdE(i))/(fwdf(i+1)-fwdf(i)));
end

% 2. Assign value=1 for peaks
peak = zeros(counter,1);
for i=1:counter-2
    if slope(i)>0 && slope(i+1)<0
        peak(i+1)=1;
    end
end

clear slope;

% 3a Identify partition relative to first maxima
crest=[]; trough=[]; k=0;
for i=1:counter
    if peak(i) == 1
        minE = min(fwdE(1:i));
        templ = fwdE == minE;
        minf = fwdf(templ); clear templ;

        %Condition 1: Trough less than 70% of the smaller peak
        if minE <= 0.7*fwdE(i)

        %Condition 2: Smaller peak is at least 33% of max peak
        if fwdE(i) >= 0.33*fwdE(1)

        %Condition 3: Smaller peak is at least 0.4m2/Hz
        if fwdE(i) >= 0.4

    end
end
k=k+1;
crest(k,1)=fwdf(i); crest(k,2)=fwdE(i);
trough(k,1)=minf(1); trough(k,2)=minE;
end

clear minf minE;
end
end
end
end

% 3b Identify partition relative to second maxima
if isempty(crest) == 0
  
  for j=1:k
    if crest(j,2)== max(crest(:,2))
      max_crest = crest(j,:);
    end
  end

  temp2 = find(fwdf == max_crest(:,1));

  for i=1:counter
    if peak(i) == 1
      if fwdf(i) < max_crest(1,1)
        minE = min(fwdE(i:temp2));
        temp3 = fwdE == minE;
        minf = fwdf(temp3); clear temp3;
      else
        minE = min(fwdE(temp2:i));
        temp3 = fwdE == minE;
        minf = fwdf(temp3); clear temp3;
      end

      %Condition 1: Trough less than 70% of the smaller peak
      if minE <= 0.7*max_crest
        %Condition 2: Smaller peak is at least 33% of max peak
        if fwdE(i) >= 0.33*fwdE(1)
          %Condition 3: Smaller peak is at least 0.4m2/Hz
          if fwdE(i) >= 0.4
            k=k+1;
            crest(k,1)=fwdf(i); crest(k,2)=fwdE(i);
            trough(k,1)=minf(1); trough(k,2)=minE;
          end
        end
      end
    end
  end
end
end
fwd_part = unique(trough(:,1));
else
fwd_part = [];
end

clearvars -except fwd_part f E me;

%====================================================================
% BACKWARD SCHEME
%====================================================================

E = flipdim(E,1);
f = flipdim(f,1);

counter = max(size(f));
for i=1:counter
if E(i) == me
    bwdf = f(i:counter);
    bwdE = E(i:counter);
end
end

% 1. Measures the slope between two successive points
counter = length(bwdE);
slope = NaN(counter-1,1);
for i=1:counter-1
    slope(i)=((bwdE(i+1)-bwdE(i))/(bwdf(i)-bwdf(i+1)));
end

% 2. Assign ID=1 for peaks
peak = zeros(counter,1);
for i=1:counter-2
    if slope(i)>0 && slope(i+1)<0
        peak(i+1)=1;
    end
end

clear slope;

% 3a Identify bi- or trimodal distribution relative to first maxima
crest=[]; trough=[]; k=0;
for i=1:counter
    if peak(i) == 1
        minE = min(bwdE(1:i));
        templ = bwdf == minE;
        minf = bwdf(templ);
        clear templ;

        %Condition 1: Trough less than 70% of the smaller peak
        if minE <= 0.7*bwdE(1)

        %Condition 2: Smaller peak is at least 33% of max peak
    end
end

99
if bwdE(i) >= 0.33*bwdE(1)

% Condition 3: Smaller peak is at least 0.4m2/Hz
  if bwdE(i) >= 0.4
    k=k+1;
    crest(k,1)=bwdf(i); crest(k,2)=bwdE(i);
    trough(k,1)=minf(1); trough(k,2)=minE;
  end
end

clear minf minE;
end
end
end

% 3b Identify bi- or trimodal distribution relative to second maxima
if isempty(crest) == 0
  for j=1:k
    if crest(j,2)== max(crest(:,2))
      max_crest = crest(j,:);
    end
  end
  temp2 = find(bwdf == max_crest(:,1));
  for i=1:counter
    if peak(i) == 1
      if bwdf(i) > max_crest(1,1)
        minE = min(bwdE(i:temp2));
        temp3 = bwdE == minE;
        minf = bwdf(temp3); clear temp3;
      else
        minE = min(bwdE(temp2:i));
        temp3 = bwdE == minE;
        minf = bwdf(temp3); clear temp3;
      end

      % Condition 1: Trough less than 70% of the smaller peak
      if minE <= 0.7*max_crest

      % Condition 2: Smaller peak is at least 33% of max peak
      if bwdE(i) >= 0.33*bwdE(1)

      % Condition 3: Smaller peak is at least 0.4m2/Hz
      if bwdE(i) >= 0.4
        k=k+1;
        crest(k,1)=bwdf(i); crest(k,2)=bwdE(i);
        trough(k,1)=minf(1); trough(k,2)=minE;
      end
end
end
end
end

clear minf minE;
bwd_part = unique(trough(:,1));
else
    bwd_part = [];
end

part = vertcat(fwd_part,bwd_part);
end
clearvars -except part
spectrastat.m


for i=1:M21NumOfFreq
    M21f(i)=M21FreqInterval*(i-1);
end

[M21E] = textread(M21ID,'%f');
M21TimeStep = length(M21E)/M21NumOfFreq;
M21E = transpose(reshape(M21E,M21NumOfFreq,M21TimeStep));
f = transpose(reshape(f,14,M21TimeStep));
E = transpose(reshape(E,14,M21TimeStep));

ETemp = NaN(M21TimeStep,M21NumOfFreq);
for i=1:M21TimeStep
    if length(unique(f(i,:)))==length(f(i,:))
        ETemp(i,:)=interp1(f(i,:),E(i,:),M21f(1,:),'linear');
        for k=1:M21NumOfFreq
            if isnan(ETemp(i,k)) == 1
                ETemp(i,k) = 0;
            end
        end
    end
end
E = ETemp; clear ETemp;
N=M21TimeStep-ModeSummary(1,4);

%====================================================================
% 1. Squared Euclidean Distance

Euclid=NaN(M21TimeStep,2);
EuclidTemp = NaN(M21TimeStep,M21NumOfFreq);

for i=1:M21TimeStep
    if isnan(E(i,:)) == 0
        for k=1:M21NumOfFreq
            EuclidTemp(i,k) = ((M21E(i,k) -
            E(i,k))^2)*M21FreqInterval;
        end
        Euclid(i,1) = sum(EuclidTemp(i,:));
        Euclid(i,2) = ModeID(i);
    end
end

%====================================================================
% 2. Significant Wave Height

HsWaveNet=NaN(M21TimeStep,1); HsM21=NaN(M21TimeStep,1);

for i=1:M21TimeStep
    if isnan(E(i,:)) == 0
        HsWaveNet(i)=4*(sqrt(sum(E(i,:)))*M21FreqInterval));
end

102
\[ HsM21(i) = 4 \times (\sqrt{\text{sum}(M21E(i,:)) \times M21FreqInterval}) \]

\[ \text{if} \hspace{1cm} \text{isnan}(HsWaveNet(i)) = 0 \text{ and } \text{isnan}(HsM21(i)) = 0 \]
\[ \text{HsRMSETemp}(i,1) = (HsM21(i) - HsWaveNet(i))^2; \]
\[ \text{HsBiasTemp}(i,1) = (HsM21(i) - HsWaveNet(i))/N; \]

\[ \text{end} \]

\[ \text{end} \]

\[ HsStat(1,1) = \sqrt{nansum(HsRMSETemp)/N}; \]
\[ HsStat(2,1) = nansum(HsBiasTemp); \]
\[ HsStat(3,1) = (HsStat(1,1)/(nansum(HsWaveNet)/N))*100; \]

\% 2c Calculate R and p-value with corrcoef

\[ \text{HsWaveNetCropped} = HsWaveNet(0 == \text{sum(isnan(HsWaveNet), 2), :}); \]
\[ \text{HsM21Cropped} = HsM21(0 == \text{sum(isnan(HsM21), 2), :}); \]

\[ [r \text{ p}] = \text{corrcoef}(\text{HsWaveNetCropped}(:,1), \text{HsM21Cropped}(:,1)); \]
\[ \text{HsStat}(4,1) = r(1,2); \]
\[ \text{HsStat}(5,1) = p(1,2); \]

\% 3. Peak Frequency and Peak Energy Distribution

\[ \text{PeakWaveNet} = \text{NaN(M21TimeStep, 2);} \]
\[ \text{PeakM21} = \text{NaN(M21TimeStep, 2);} \]
\[ \text{PeakRMSETemp} = \text{NaN(M21TimeStep, 2);} \]
\[ \text{PeakBiasTemp} = \text{NaN(M21TimeStep, 2);} \]

\[ \text{for } i = 1:M21TimeStep \]
\[ \hspace{1cm} \text{if} \hspace{1cm} \text{isnan(E(i,:))} = 0 \text{ and } \text{isnan(M21E(i,:))} = 0 \]
\[ \hspace{2cm} \text{PeakWaveNet}(i,1) = \text{max(E(i,:))}; \]
\[ \hspace{2cm} \text{PositionWaveNet} = \text{find(E(i,:) == max(E(i,:)))}; \]
\[ \hspace{2cm} \text{PeakWaveNet}(i,2) = \text{M21f(PositionWaveNet(1))}; \]
\[ \hspace{2cm} \text{PeakM21}(i,1) = \text{max(M21E(i,:))}; \]
\[ \hspace{2cm} \text{PositionM21} = \text{find(M21E(i,:) == max(M21E(i,:)))}; \]
\[ \hspace{2cm} \text{PeakM21}(i,2) = \text{M21f(PositionM21(1))}; \]
\[ \text{end} \]

\% 3b Calculate RMSE, Bias and Scatter Index

\[ \text{if} \hspace{1cm} \text{isnan(PeakWaveNet(i,:))} = 0 \text{ and } \text{isnan(PeakM21(i,:))} = 0 \]
\[ \hspace{2cm} \text{PeakRMSETemp}(i,1) = (\text{PeakM21}(i,1) - \text{PeakWaveNet}(i,1))^2; \]
\[ \hspace{2cm} \text{PeakRMSETemp}(i,2) = (\text{PeakM21}(i,2) - \text{PeakWaveNet}(i,2))^2; \]
\[ \hspace{2cm} \text{PeakBiasTemp}(i,1) = (\text{PeakM21}(i,1) - \text{PeakWaveNet}(i,1))/N; \]
\[ \hspace{2cm} \text{PeakBiasTemp}(i,2) = (\text{PeakM21}(i,2) - \text{PeakWaveNet}(i,2))/N; \]
\[ \text{end} \]

\[ \text{PeakStat}(1,1) = \sqrt{nansum(PeakRMSETemp(:,1))/N}; \]
\[ \text{PeakStat}(1,2) = \sqrt{nansum(PeakRMSETemp(:,2))/N}; \]
\[ \text{PeakStat}(2,1) = nansum(PeakBiasTemp(:,1)); \]
\[ \text{PeakStat}(2,2) = nansum(PeakBiasTemp(:,2)); \]
\[ \text{PeakStat}(3,1) = (\text{PeakStat}(1,1)/(\text{nansum(PeakWaveNet(:,1))/N})*100; \]
\[ \text{PeakStat}(3,2) = (\text{PeakStat}(1,2)/(\text{nansum(PeakWaveNet(:,2))/N})*100; \]

\% 3c Calculate R and p-value with corrcoef
PeakWaveNetCropped = PeakWaveNet(0 == sum(isnan(PeakWaveNet), 2), :);
PeakM21Cropped = PeakM21(0 == sum(isnan(PeakM21), 2), :);

[r p] = corrcoef(PeakWaveNetCropped(:,1), PeakM21Cropped(:,1));
PeakStat(4,1) = r(1,2); PeakStat(5,1) = p(1,2);
[r p] = corrcoef(PeakWaveNetCropped(:,2), PeakM21Cropped(:,2));
PeakStat(4,2) = r(1,2); PeakStat(5,2) = p(1,2);

% 4. Mean Width Deviation

x = 0;
AwWaveNet = NaN(M21TimeStep, 99); AwM21 = NaN(M21TimeStep, 99);
for m = 1:99
    x = x + 0.01;
    for i = 1:M21TimeStep
        if isnan(PeakWaveNet(i,:)) == 0 & isnan(PeakM21(i,:)) == 0
            if PeakWaveNet(i,1) <= PeakM21(i,1);
                refPeak = PeakWaveNet(i,1);
            else
                refPeak = PeakM21(i,1);
            end
            refPeak = refPeak * x;

            counterM21 = 0; counterWaveNet = 0;
            for k = 1:M21NumOfFreq
                if E(i,k) >= refPeak;
                    counterWaveNet = counterWaveNet + 1;
                end
                if M21E(i,k) >= refPeak;
                    counterM21 = counterM21 + 1;
                end
            end
            Aw1 = counterWaveNet * M21FreqInterval;
            Aw2 = counterM21 * M21FreqInterval;
            AwWaveNet(i,m) = Aw1; AwM21(i,m) = Aw2;
        end
    end
end

for i = 1:M21TimeStep;
    MwWaveNet(i) = nanmean(AwWaveNet(i,:));
    MwM21(i) = nanmean(AwM21(i,:));
    deltaMw(i) = MwWaveNet(i) - MwM21(i);
end

deltaMw = deltaMw';

% 5. Skewness and Kurtosis
MomentWaveNet = NaN(M21TimeStep,5); MomentM21 = NaN(M21TimeStep,5);

for i=1:M21TimeStep
    SumE = sum(E(i,:)); SumM21E = sum(M21E(i,:));
    if isnan(E(i,:)) == 0 & isnan(M21E(i,:)) == 0
        MomentWaveNet(i,1) = sum(E(i,2:M21NumOfFreq).*M21f(1,2:M21NumOfFreq))/SumE;
        MomentWaveNet(i,2) = sum(E(i,2:M21NumOfFreq).*((M21f(1,2:M21NumOfFreq)-MomentWaveNet(i,1)).^2))/SumE;
        MomentWaveNet(i,3) = sqrt(MomentWaveNet(i,2));
        MomentWaveNet(i,4) = sum(E(i,2:M21NumOfFreq).*((M21f(1,2:M21NumOfFreq)-MomentWaveNet(i,1)).^3))/SumE/(MomentWaveNet(i,3)^3);
        MomentWaveNet(i,5) = sum(E(i,2:M21NumOfFreq).*((M21f(1,2:M21NumOfFreq)-MomentWaveNet(i,1)).^4))/SumE/(MomentWaveNet(i,3)^4);
    end
end

%====================================================================
% 6. Return value
ParaWaveNet=NaN(M21TimeStep,6); ParaM21=NaN(M21TimeStep,6);

ParaWaveNet(:,1)=HsWaveNet;
ParaWaveNet(:,2)=PeakWaveNet(:,2);
ParaWaveNet(:,3)=PeakWaveNet(:,1);
ParaWaveNet(:,4)=MwWaveNet;
ParaWaveNet(:,5)=MomentWaveNet(:,4);
ParaWaveNet(:,6)=MomentWaveNet(:,5);

ParaM21(:,1)=HsM21;
ParaM21(:,2)=PeakM21(:,2);
ParaM21(:,3)=PeakM21(:,1);
ParaM21(:,4)=MwM21;
ParaM21(:,5)=MomentM21(:,4);
ParaM21(:,6)=MomentM21(:,5);

%====================================================================
% 7. Relationship between parameters for unimodal distributions

del_Hs=NaN(M21TimeStep,1); del.fp=del_Hs; del.Emax=del_Hs;
del_Sk=del_Hs; del.K=del_Hs;

for i=1:M21TimeStep
del_Hs(i)=ParaWaveNet(i,1)-ParaM21(i,1);
del_fp(i)=ParaWaveNet(i,2)-ParaM21(i,2);
del_Emax(i)=ParaWaveNet(i,3)-ParaM21(i,3);
del_Sk(i)=ParaWaveNet(i,5)-ParaM21(i,5);
del_K(i)=ParaWaveNet(i,6)-ParaM21(i,6);
end

% 7b Between del_Hs and del_Emax
region=NaN(M21TimeStep,1); RegionStat=zeros(4,9);
RegionStat2=zeros(2,2); RegionStat4=RegionStat2;
for i=1:M21TimeStep
    if isnan(del_Hs(i))==0 & isnan(del_Emax(i))==0 & ModeID(i)==1;
        if del_Hs(i)<-0.05 & del_Emax(i)>0.01 & deltaMw(i)<0
            region(i)=1;
            RegionStat(1,1)=RegionStat(1,1)+1;
        elseif del_Hs(i)>0.05 & del_Emax(i)>0.01 & deltaMw(i)< 0
            region(i)=2;
            RegionStat(1,2)=RegionStat(1,2)+1;
        elseif del_Hs(i)>0.05 & del_Emax(i)<-0.01
            region(i)=3;
            RegionStat(1,3)=RegionStat(1,3)+1;
        elseif del_Hs(i)<-0.05 & del_Emax(i)<-0.01
            region(i)=4;
            RegionStat(1,4)=RegionStat(1,4)+1;
        elseif abs(del_Hs(i))<=0.05 & del_Emax(i)>0.01
            region(i)=5;
            RegionStat(1,5)=RegionStat(1,5)+1;
        elseif del_Hs(i)>0.05 & abs(del_Emax(i))<=0.01
            region(i)=6;
            RegionStat(1,6)=RegionStat(1,6)+1;
        elseif abs(del_Hs(i))<=0.05 & del_Emax(i)<-0.01
            region(i)=7;
            RegionStat(1,7)=RegionStat(1,7)+1;
        elseif del_Hs(i)<-0.05 & abs(del_Emax(i))<=0.01
            region(i)=8;
            RegionStat(1,8)=RegionStat(1,8)+1;
        elseif abs(del_Hs(i))<=0.05 & abs(del_Emax(i))<=0.01
            region(i)=9;
            RegionStat(1,9)=RegionStat(1,9)+1;
    end
end
for i=1:9
    RegionStat(2,i)=(RegionStat(1,i)/sum(RegionStat(1,:)))*100;
end
for i=1:2
    RegionStat2(2,i)=(RegionStat2(1,i)/sum(RegionStat2(1,:)))*100;
RegionStat4(2,i)=(RegionStat4(1,i)/sum(RegionStat(1,:)))*100;
end

% 7c Between del_fp and del_Emax
region=NaN(M21TimeStep,1);

for i=1:M21TimeStep
  if isnan(del_fp(i))==0 & isnan(del_Emax(i))==0 & ModeID(i)==1;
    if del_fp(i)<-0.001 & del_Emax(i)>0.01
      region(i)=1;
    RegionStat(3,1)=RegionStat(3,1)+1;
  elseif del_fp(i)>0.001 & del_Emax(i)>0.01
    region(i)=2;
    RegionStat(3,2)=RegionStat(3,2)+1;
  elseif del_fp(i)>0.001 & del_Emax(i)<-0.01
    region(i)=3;
    RegionStat(3,3)=RegionStat(3,3)+1;
  elseif del_fp(i)<-0.001 & del_Emax(i)<-0.01
    region(i)=4;
    RegionStat(3,4)=RegionStat(3,4)+1;
  elseif abs(del fp(i))<=0.001 & del_Emax(i)>0.01
    region(i)=5;
    RegionStat(3,5)=RegionStat(3,5)+1;
  elseif del_fp(i)>0.001 & abs(del_Emax(i))<=0.01
    region(i)=6;
    RegionStat(3,6)=RegionStat(3,6)+1;
  elseif abs(del fp(i))<=0.001 & del_Emax(i)<-0.01
    region(i)=7;
    RegionStat(3,7)=RegionStat(3,7)+1;
  elseif del_fp(i)<-0.001 & abs(del_Emax(i))<=0.01
    region(i)=8;
    RegionStat(3,8)=RegionStat(3,8)+1;
  elseif abs(del fp(i))<=0.001 & abs(del_Emax(i))<=0.01
    region(i)=9;
    RegionStat(3,9)=RegionStat(3,9)+1;
  else
    RegionStat(3,10)=RegionStat(3,10)+1;
  end
end

for i=1:9
  RegionStat(4,i)=(RegionStat(3,i)/sum(RegionStat(3,:)))*100;
end
function \[Bi_{\text{ParaWaveNet}},Bi_{\text{ParaM21}},Bi_{\text{deltaMw}},Bi_{\text{Euclid}}\]=bimodal(f,E,M21NumOfFreq,M21FreqInterval,Bimodal)

for i=1:M21NumOfFreq
    M21f(i)=M21FreqInterval*(i-1);
end

[M21E] = textread(M21ID,'%f');
M21TimeStep = length(M21E)/M21NumOfFreq;
M21E = transpose(reshape(M21E,M21NumOfFreq,M21TimeStep));
f = transpose(reshape(f,14,M21TimeStep));
E = transpose(reshape(E,14,M21TimeStep));

ETemp = NaN(M21TimeStep,M21NumOfFreq);

for i=1:M21TimeStep
    if length(unique(f(i,:)))==length(f(i,:))
        ETemp(i,:)=interp1(f(i,:),E(i,:),M21f(1,:),'linear');
        for k=1:M21NumOfFreq
            if isnan(ETemp(i,k)) == 1
                ETemp(i,k) = 0;
            end
        end
    end
end

E = ETemp; clear ETemp;

Bi_{\text{HsWaveNet}}=NaN(M21TimeStep,2); Bi_{\text{HsM21}}=NaN(M21TimeStep,2);
Bi_{\text{PeakWaveNet}}=NaN(M21TimeStep,4); Bi_{\text{PeakM21}}=NaN(M21TimeStep,4);

%====================================================================
% 1. Significant Wave Height, Peak Frequency and Peak Energy

for i=1:M21TimeStep
    if isnan(Bimodal(i))==0
        [t,bin]=histc(Bimodal(i),M21f);
        if abs(Bimodal(i)-M21f(bin)) > abs(Bimodal(i)-M21f(bin+1))
            bin=bin+1;
        end
        clear t;
        M21f_Model = M21f(1,1:bin); M21f_Mode2 = M21f(1,bin:M21NumOfFreq);
        M21E_Model = M21E(1,1:bin); M21E_Mode2 = M21E(1,bin:M21NumOfFreq);
        E_Model = E(1,1:bin); E_Mode2 = E(1,bin:M21NumOfFreq);

        % 1a. Significant Wave Height
        Bi_{\text{HsWaveNet}}(i,1)=4*sqrt((sum(E_Model))*M21FreqInterval);
        Bi_{\text{HsWaveNet}}(i,2)=4*sqrt((sum(E_Mode2))*M21FreqInterval);
        Bi_{\text{HsM21}}(i,1)=4*sqrt((sum(M21E_Model))*M21FreqInterval);
    end
Bi_HsM21(i,2)=4*sqrt((sum(M21E_Mode2))*M21FreqInterval);

% 1b. Peak Frequency and Peak Energy Distribution

Bi_PeakWaveNet(i,1) = max(E_Mode1);
PositionWaveNet_Mode1 = find(E_Mode1 == max(E_Mode1));
Bi_PeakWaveNet(i,2) = M21f_Mode1(PositionWaveNet_Mode1(1));
Bi_PeakWaveNet(i,3) = max(E_Mode2);
PositionWaveNet_Mode2 = find(E_Mode2 == max(E_Mode2));
Bi_PeakWaveNet(i,4) = M21f_Mode2(PositionWaveNet_Mode2(1));

Bi_PeakM21(i,1) = max(M21E_Mode1);
PositionM21_Mode1 = find(M21E_Mode1 == max(M21E_Mode1));
Bi_PeakM21(i,2) = M21f_Mode1(PositionM21_Mode1(1));
Bi_PeakM21(i,3) = max(M21E_Mode2);
PositionM21_Mode2 = find(M21E_Mode2 == max(M21E_Mode2));
Bi_PeakM21(i,4) = M21f_Mode2(PositionM21_Mode2(1));

end

end

%====================================================================
% 2. Mean Width Deviation

AwWaveNet_Mode1 = NaN(M21TimeStep, 99);
AwWaveNet_Mode2 = NaN(M21TimeStep, 99);
AwM21_Mode1 = NaN(M21TimeStep, 99);
AwM21_Mode2 = NaN(M21TimeStep, 99);

x = 0;
for m = 1:
    x = x + 0.01;
    for i = 1:M21TimeStep
        if isnan(Bimodal(i)) == 0
            [t, bin] = histc(Bimodal(i), M21f);

            if abs(Bimodal(i) - M21f(bin)) > abs(Bimodal(i) -
                M21f(bin+1))
                bin = bin + 1;
            end
        clear t;

        M21E_Mode1 = M21E(i, 1:bin); M21E_Mode2 =
            M21E(i, bin:M21NumOfFreq);
        E_Mode1 = E(i, 1:bin); E_Mode2 = E(i, bin:M21NumOfFreq);
        dim_Model = bin; dim_Mode2 = M21NumOfFreq - bin + 1;

        if Mode 1
            if Bi_PeakWaveNet(i, 1) <= Bi_PeakM21(i, 1);
                refPeak = Bi_PeakWaveNet(i, 1);
            else
                refPeak = Bi_PeakM21(i, 1);
            end
        else
            refPeak = Bi_PeakM21(i, 1);
        end

        refPeak = refPeak * x;
        counterM21_Model = 0; counterWaveNet_Model = 0;
for k=1:dim_Mode1
    if E_Mode1(k)>= refPeak;
        counterWaveNet_Mode1=counterWaveNet_Mode1+1;
    end

    if M21E_Mode1(k)>= refPeak;
        counterM21_Mode1=counterM21_Mode1+1;
    end
end

AwWaveNet_Mode1(i,m)=counterWaveNet_Mode1*M21FreqInterval;
AwM21_Mode1(i,m)=counterM21_Mode1*M21FreqInterval;

% Mode 2
if Bi_PeakWaveNet(i,3) <= Bi_PeakM21(i,3);
    refPeak2=Bi_PeakWaveNet(i,3);
else
    refPeak2=Bi_PeakM21(i,3);
end

refPeak2=refPeak2*x;
counterM21_Mode2=0; counterWaveNet_Mode2=0;
for k=1:dim_Mode2
    if E_Mode2(k)>= refPeak2;
        counterWaveNet_Mode2=counterWaveNet_Mode2+1;
    end

    if M21E_Mode2(k)>= refPeak2;
        counterM21_Mode2=counterM21_Mode2+1;
    end
end

AwWaveNet_Mode2(i,m)=counterWaveNet_Mode2*M21FreqInterval;
AwM21_Mode2(i,m)=counterM21_Mode2*M21FreqInterval;
end
end
delimit

for i=1:M21TimeStep;

    % Mode 1
    MwWaveNet_Mode1(i)=nanmean(AwWaveNet_Mode1(i,:));
    MwM21_Mode1(i)=nanmean(AwM21_Mode1(i,:));
    Bi_deltaMw(i,1)=MwWaveNet_Mode1(i)-MwM21_Mode1(i);

    % Mode 2
    MwWaveNet_Mode2(i)=nanmean(AwWaveNet_Mode2(i,:));
    MwM21_Mode2(i)=nanmean(AwM21_Mode2(i,:));
    Bi_deltaMw(i,2)=MwWaveNet_Mode2(i)-MwM21_Mode2(i);
end
% 3. Skewness and Kurtosis

MomentWaveNet_Mode1=NaN(M21TimeStep,5);
MomentWaveNet_Mode2=NaN(M21TimeStep,5);
MomentM21_Mode1=NaN(M21TimeStep,5);
MomentM21_Mode2=NaN(M21TimeStep,5);

for i=1:M21TimeStep
    if isnan(Bimodal(i))==0
        [t,bin]=histc(Bimodal(i),M21f);
        if abs(Bimodal(i)-M21f(bin)) > abs(Bimodal(i)-M21f(bin+1))
            bin=bin+1;
        end
        clear t;
        M21f_Mode1 = M21f(1,1:bin); M21f_Mode2 = M21f(1,bin:M21NumOfFreq);
        M21E_Mode1 = M21E(i,1:bin); M21E_Mode2 = M21E(i,bin:M21NumOfFreq);
        E_Mode1 = E(i,1:bin); E_Mode2 = E(i,bin:M21NumOfFreq);
        dim_Mode1 = bin; dim_Mode2 = M21NumOfFreq - bin+1;

        % Mode 1
        SumE_Mode1 = sum(E_Mode1);
        if isnan(E(i,:)) == 0
            MomentWaveNet_Mode1(i,1) = sum(E_Mode1(1:dim_Mode1).*M21f_Mode1(1,1:dim_Mode1))/SumE_Mode1;
            MomentWaveNet_Mode1(i,2) = sum(E_Mode1(1:dim_Mode1).*((M21f_Mode1(1,1:dim_Mode1)-MomentWaveNet_Mode1(i,1)).^2))/SumE_Mode1;
            MomentWaveNet_Mode1(i,3) = sqrt(MomentWaveNet_Mode1(i,2));
            MomentWaveNet_Mode1(i,4) = sum(E_Mode1(1:dim_Mode1).*((M21f_Mode1(1,1:dim_Mode1)-MomentWaveNet_Mode1(i,1)).^3))/SumE_Mode1/(MomentWaveNet_Mode1(i,3)^3);
            MomentWaveNet_Mode1(i,5) = sum(E_Mode1(1:dim_Mode1).*((M21f_Mode1(1,1:dim_Mode1)-MomentWaveNet_Mode1(i,1)).^4))/SumE_Mode1/(MomentWaveNet_Mode1(i,3)^4);
        end
        SumM21E_Mode1 = sum(M21E_Mode1);
        if isnan(M21E(i,:)) == 0
            MomentM21_Mode1(i,1) = sum(M21E_Mode1(1:dim_Mode1).*M21f_Mode1(1,1:dim_Mode1))/SumM21E_Mode1;
            MomentM21_Mode1(i,2) = sum(M21E_Mode1(1:dim_Mode1).*((M21f_Mode1(1,1:dim_Mode1)-MomentM21_Mode1(i,1)).^2))/SumM21E_Mode1;
            MomentM21_Mode1(i,3) = sqrt(MomentM21_Mode1(i,2));
            MomentM21_Mode1(i,4) = sum(M21E_Mode1(1:dim_Mode1).*((M21f_Mode1(1,1:dim_Mode1)-MomentM21_Mode1(i,1)).^3))/SumM21E_Mode1/(MomentM21_Mode1(i,3)^3);
        end
    end
end
MomentM21_Mode1(i,5) = 
  sum(M21E_Mode1(1:dim_Mode1).*((M21f_Mode1(1,1:dim_Mode1) - 
  MomentM21_Mode1(i,1)).^4))/SumM21E_Mode1/(MomentM21_Mode1 
  (i,3)^4);

end

% Mode 2
SumE_Mode2 = sum(E_Mode2);
if isnan(E(i,:)) == 0
  MomentWaveNet_Mode2(i,1) = 
    sum(E_Mode2(1:dim_Mode2).*M21f_Mode2(1,1:dim_Mode2))/SumE 
    _Mode2;
  MomentWaveNet_Mode2(i,2) = 
    sum(E_Mode2(1:dim_Mode2).*((M21f_Mode2(1,1:dim_Mode2) - 
    MomentWaveNet_Mode2(i,1)).^2))/SumE_Mode2;
  MomentWaveNet_Mode2(i,3) = 
    sqrt(MomentWaveNet_Mode2(i,2));
  MomentWaveNet_Mode2(i,4) = 
    sum(E_Mode2(1:dim_Mode2).*((M21f_Mode2(1,1:dim_Mode2) - 
    MomentWaveNet_Mode2(i,1)).^3))/SumE_Mode2/(MomentWaveNet 
    Mode2(i,3)^3);
  MomentWaveNet_Mode2(i,5) = 
    sum(E_Mode2(1:dim_Mode2).*((M21f_Mode2(1,1:dim_Mode2) - 
    MomentWaveNet_Mode2(i,1)).^4))/SumE_Mode2/(MomentWaveNet 
    Mode2(i,3)^4);
end

SumM21E_Mode2 = sum(M21E_Mode2);
if isnan(M21E(i,:)) == 0
  MomentM21_Mode2(i,1) = 
    sum(M21E_Mode2(1:dim_Mode2).*M21f_Mode2(1,1:dim_Mode2))/S 
    umE_Mode2;
  MomentM21_Mode2(i,2) = 
    sum(M21E_Mode2(1:dim_Mode2).*((M21f_Mode2(1,1:dim_Mode2) - 
    MomentM21_Mode2(i,1)).^2))/SumM21E_Mode2;
  MomentM21_Mode2(i,3) = sqrt(MomentM21_Mode2(i,2));
  MomentM21_Mode2(i,4) = 
    sum(M21E_Mode2(1:dim_Mode2).*((M21f_Mode2(1,1:dim_Mode2) - 
    MomentM21_Mode2(i,1)).^3))/SumM21E_Mode2/(MomentM21_Mode2 
    (i,3)^3);
  MomentM21_Mode2(i,5) = 
    sum(M21E_Mode2(1:dim_Mode2).*((M21f_Mode2(1,1:dim_Mode2) - 
    MomentM21_Mode2(i,1)).^4))/SumM21E_Mode2/(MomentM21_Mode2 
    (i,3)^4);
end

end

%====================================================================
% 4. Squared Euclidean Distance

Bi_Euclid=NaN(M21TimeStep,2);
for i=1:M21TimeStep
  if isnan(Bimodal(i))==0
    [t,bin]=histc(Bimodal(i),M21f);
    if abs(Bimodal(i)-M21f(bin)) > abs(Bimodal(i)-M21f(bin+1))
      bin=bin+1;
  end

end

%====================================================================
% 4. Squared Euclidean Distance

Bi_Euclid=NaN(M21TimeStep,2);
for i=1:M21TimeStep
  if isnan(Bimodal(i))==0
    [t,bin]=histc(Bimodal(i),M21f);
    if abs(Bimodal(i)-M21f(bin)) > abs(Bimodal(i)-M21f(bin+1))
      bin=bin+1;
  end

end
clear t;

M21E_Mode1 = M21E(i,1:bin); M21E_Mode2 = M21E(i,bin:M21NumOfFreq);
E_Mode1 = E(i,1:bin); E_Mode2 = E(i,bin:M21NumOfFreq);
dim_Model = bin; dim_Mode2 = M21NumOfFreq-bin+1;

EuclidTemp_Mode1 = NaN(dim_Mode1);
EuclidTemp_Mode2 = NaN(dim_Mode2);

% Mode 1
for k=1:dim_Mode1
    EuclidTemp_Mode1(k) = ((M21E_Mode1(k) - 
        E_Mode1(k))^2)*M21FreqInterval;
end
Bi_Euclid(i,1) = nansum(EuclidTemp_Mode1(:));

% Mode 2
for k=1:dim_Mode2
    EuclidTemp_Mode2(i,k) = ((M21E_Mode2(k) - 
        E_Mode2(k))^2)*M21FreqInterval;
end
Bi_Euclid(i,2) = nansum(EuclidTemp_Mode2(:));

end

% 5. Return value

Bi_ParaWaveNet=NaN(M21TimeStep,11);
Bi_ParaM21=NaN(M21TimeStep,11);

Bi_ParaWaveNet(:,1)=Bi_HsWaveNet(:,1);
Bi_ParaWaveNet(:,2)=Bi_PeakWaveNet(:,2);
Bi_ParaWaveNet(:,3)=Bi_PeakWaveNet(:,1);
Bi_ParaWaveNet(:,4)=MomentWaveNet_Mode1(:,4);
Bi_ParaWaveNet(:,5)=MomentWaveNet_Mode1(:,5);
Bi_ParaWaveNet(:,6)=NaN;
Bi_ParaWaveNet(:,7)=Bi_HsWaveNet(:,2);
Bi_ParaWaveNet(:,8)=Bi_PeakWaveNet(:,4);
Bi_ParaWaveNet(:,9)=Bi_PeakWaveNet(:,3);
Bi_ParaWaveNet(:,10)=MomentWaveNet_Mode2(:,4);
Bi_ParaWaveNet(:,11)=MomentWaveNet_Mode2(:,5);

Bi_ParaM21(:,1)=Bi_HsM21(:,1);
Bi_ParaM21(:,2)=Bi_PeakM21(:,2);
Bi_ParaM21(:,3)=Bi_PeakM21(:,1);
Bi_ParaM21(:,4)=MomentM21_Mode1(:,4);
Bi_ParaM21(:,5)=MomentM21_Mode1(:,5);
Bi_ParaM21(:,6)=NaN;
Bi_ParaM21(:,7)=Bi_HsM21(:,2);
Bi_ParaM21(:,8)=Bi_PeakM21(:,4);
Bi_ParaM21(:,9)=Bi_PeakM21(:,3);
Bi_ParaM21(:,10)=MomentM21_Mode2(:,4);
Bi_ParaM21(:,11)=MomentM21_Mode2(:,5);
trimodal.m

function [Tri_ParaWaveNet,Tri_ParaM21,Tri_deltaMw,Tri_Euclid]=trimodal(f,E,M21ID,M21NumOfFreq,M21FreqInterval,Trimodal)

for i=1:M21NumOfFreq
    M21f(i)=M21FreqInterval*(i-1);
end

[M21E] = textread(M21ID,'%f');
M21TimeStep = length(M21E)/M21NumOfFreq;
M21E = transpose(reshape(M21E,M21NumOfFreq,M21TimeStep));
f = transpose(reshape(f,14,M21TimeStep));
E = transpose(reshape(E,14,M21TimeStep));

ETemp = NaN(M21TimeStep,M21NumOfFreq);
for i=1:M21TimeStep
    if length(unique(f(i,:)))==length(f(i,:))
        ETemp(i,:)=interp1(f(i,:),E(i,:),M21f(1,:),'linear');
        for k=1:M21NumOfFreq
            if isnan(ETemp(i,k)) == 1
                ETemp(i,k) = 0;
            end
        end
    end
end
E = ETemp; clear ETemp;

Tri_HsWaveNet=NaN(M21TimeStep,3); Tri_HsM21=NaN(M21TimeStep,3);
Tri_PeakWaveNet=NaN(M21TimeStep,6); Tri_PeakM21=NaN(M21TimeStep,6);

%====================================================================
% 1. Significant Wave Height, Peak Frequency and Peak Energy

for i=1:M21TimeStep
    if isnan(Trimodal(i,:))==0
        [t,bin1]=histc(Trimodal(i,1),M21f);
        if abs(Trimodal(i,1)-M21f(bin1)) > abs(Trimodal(i,2)-M21f(bin1+1))
            bin1=bin1+1;
        end
        clear t; 
        [t,bin2]=histc(Trimodal(i,2),M21f);
        if abs(Trimodal(i,2)-M21f(bin2)) > abs(Trimodal(i,2)-M21f(bin2+1))
            bin2=bin2+1;
        end
        clear t;
        M21f_Model1 = M21f(1,1:bin1); M21f_Model2 = M21f(1,bin1:bin2);
M21f_Mode3 = M21f(1,bin2:M21NumOfFreq);
M21E_Mode1 = M21E(i,1:bin1); M21E_Mode2 = M21E(i,bin1:bin2);
M21E_Mode3 = M21E(i,bin2:M21NumOfFreq);
E_Mode1 = E(i,1:bin1); E_Mode2 = E(i,bin1:bin2);
E_Mode3 = E(i,bin2:M21NumOfFreq);

% 1a. Significant Wave Height
Tri_HsWaveNet(i,1)=4*sqrt((sum(E_Mode1))*M21FreqInterval);
Tri_HsWaveNet(i,2)=4*sqrt((sum(E_Mode2))*M21FreqInterval);
Tri_HsWaveNet(i,3)=4*sqrt((sum(E_Mode3))*M21FreqInterval);
Tri_HsM21(i,1)=4*sqrt((sum(M21E_Mode1))*M21FreqInterval);
Tri_HsM21(i,2)=4*sqrt((sum(M21E_Mode2))*M21FreqInterval);
Tri_HsM21(i,3)=4*sqrt((sum(M21E_Mode3))*M21FreqInterval);

% 1b. Peak Frequency and Peak Energy Distribution
Tri_PeakWaveNet(i,1)= max(E_Mode1);
PositionWaveNet_Mode1=find(E_Mode1 == max(E_Mode1));
Tri_PeakWaveNet(i,2)= M21f_Mode1(PositionWaveNet_Mode1(1));
Tri_PeakWaveNet(i,3)= max(E_Mode2);
PositionWaveNet_Mode2=find(E_Mode2 == max(E_Mode2));
Tri_PeakWaveNet(i,4)= M21f_Mode2(PositionWaveNet_Mode2(1));
Tri_PeakWaveNet(i,5)= max(E_Mode3);
PositionWaveNet_Mode3=find(E_Mode3 == max(E_Mode3));
Tri_PeakWaveNet(i,6)= M21f_Mode3(PositionWaveNet_Mode3(1));

Tri_PeakM21(i,1)= max(M21E_Mode1);
PositionM21_Mode1=find(M21E_Mode1 == max(M21E_Mode1));
Tri_PeakM21(i,2)= M21f_Mode1(PositionM21_Mode1(1));
Tri_PeakM21(i,3)= max(M21E_Mode2);
PositionM21_Mode2=find(M21E_Mode2 == max(M21E_Mode2));
Tri_PeakM21(i,4)= M21f_Mode2(PositionM21_Mode2(1));
Tri_PeakM21(i,5)= max(M21E_Mode3);
PositionM21_Mode3=find(M21E_Mode3 == max(M21E_Mode3));
Tri_PeakM21(i,6)= M21f_Mode3(PositionM21_Mode3(1));

end
end

%====================================================================%
% 2. Mean Width Deviation

AwWaveNet_Mode1=NaN(M21TimeStep,99);
AwWaveNet_Mode2=NaN(M21TimeStep,99);
AwWaveNet_Mode3=NaN(M21TimeStep,99);
AwM21_Mode1=NaN(M21TimeStep,99);
AwM21_Mode2=NaN(M21TimeStep,99);
AwM21_Mode3=NaN(M21TimeStep,99);
x=0;
for m=1;
x=x+0.01;
for i=1:M21TimeStep
    if isnan(Trimodal(i,:)) == 0
        [t,bin]=histc(Trimodal(i,1),M21f);
        if abs(Trimodal(i,1)-M21f(bin1)) > abs(Trimodal(i,2)-
M21f(bin1+1))
    bin1=bin1+1;
end

clear t;

[t,bin2]=histc(Trimodal(i,2),M21f);

if abs(Trimodal(i,2)-M21f(bin2)) > abs(Trimodal(i,2)-M21f(bin2+1))
    bin2=bin2+1;
end

clear t;

M21E_Mode1 = M21E(i,1:bin1); M21E_Mode2 = M21E(i,bin1:bin2);
M21E_Mode3 = M21E(i,bin2:M21NumOfFreq);
E_Mode1 = E(i,1:bin1); E_Mode2 = E(i,bin1:bin2);
E_Mode3 = E(i,bin2:M21NumOfFreq);
dim_Mode1 = bin1; dim_Mode2 = bin2-bin1+1;
dim_Mode3 = M21NumOfFreq-bin2+1;

% Mode 1
if Tri_PeakWaveNet(i,1) <= Tri_PeakM21(i,1);
    refPeak=Tri_PeakWaveNet(i,1);
else
    refPeak=Tri_PeakM21(i,1);
end
refPeak=refPeak*x;
counterM21_Mode1=0; counterWaveNet_Mode1=0;
for k=1:dim_Mode1
    if E_Mode1(k)>= refPeak;
        counterWaveNet_Mode1=counterWaveNet_Mode1+1;
    end
    if M21E_Mode1(k)>= refPeak;
        counterM21_Mode1=counterM21_Mode1+1;
    end
end
AwWaveNet_Mode1(i,m)=counterWaveNet_Mode1*M21FreqInterval;
AwM21_Mode1(i,m)=counterM21_Mode1*M21FreqInterval;

% Mode 2
if Tri_PeakWaveNet(i,3) <= Tri_PeakM21(i,3);
    refPeak2=Tri_PeakWaveNet(i,3);
else
    refPeak2=Tri_PeakM21(i,3);
end
refPeak2=refPeak2*x;
counterM21_Mode2=0; counterWaveNet_Mode2=0;
for k=1:dim_Mode2
if E_Mode2(k) >= refPeak2;
    counterWaveNet_Mode2 = counterWaveNet_Mode2 + 1;
end

if M21E_Mode2(k) >= refPeak2;
    counterM21_Mode2 = counterM21_Mode2 + 1;
end

daWaveWaveNet_Mode2(i,m) = counterWaveNet_Mode2*M21FreqInterval;
daWMM21_Mode2(i,m) = counterM21_Mode2*M21FreqInterval;

% Mode 3
if Tri_PeakWaveNet(i,5) <= Tri_PeakM21(i,5);
    refPeak3 = Tri_PeakWaveNet(i,5);
else
    refPeak3 = Tri_PeakM21(i,5);
end

refPeak3 = refPeak3*x;
counterM21_Mode3 = 0; counterWaveNet_Mode3 = 0;

for k=1:dim_Mode3
    if E_Mode3(k) >= refPeak3;
        counterWaveNet_Mode3 = counterWaveNet_Mode3 + 1;
    end

    if M21E_Mode3(k) >= refPeak3;
        counterM21_Mode3 = counterM21_Mode3 + 1;
    end
end

daWaveWaveNet_Mode3(i,m) = counterWaveNet_Mode3*M21FreqInterval;
daWMM21_Mode3(i,m) = counterM21_Mode3*M21FreqInterval;

end
end
end

for i=1:M21TimeStep;
    % Mode 1
    MwWaveNet_Mode1(i) = nanmean(AwWaveNet_Mode1(i,:));
    MwM21_Mode1(i) = nanmean(AwM21_Mode1(i,:));
    Tri_deltaMw(i,1) = MwWaveNet_Mode1(i) - MwM21_Mode1(i);

    % Mode 2
    MwWaveNet_Mode2(i) = nanmean(AwWaveNet_Mode2(i,:));
    MwM21_Mode2(i) = nanmean(AwM21_Mode2(i,:));
    Tri_deltaMw(i,2) = MwWaveNet_Mode2(i) - MwM21_Mode2(i);

    % Mode 3
    MwWaveNet_Mode3(i) = nanmean(AwWaveNet_Mode3(i,:));
    MwM21_Mode3(i) = nanmean(AwM21_Mode3(i,:));
    Tri_deltaMw(i,3) = MwWaveNet_Mode3(i) - MwM21_Mode3(i);
% 3. Skewness and Kurtosis

MomentWaveNet_Mode1=NaN(M21TimeStep,5);
MomentWaveNet_Mode2=NaN(M21TimeStep,5);
MomentWaveNet_Mode3=NaN(M21TimeStep,5);
MomentM21_Mode1=NaN(M21TimeStep,5);
MomentM21_Mode2=NaN(M21TimeStep,5);
MomentM21_Mode3=NaN(M21TimeStep,5);

for i=1:M21TimeStep
    if isnan(Trimodal(i,:))==0
        [t,bin1]=histc(Trimodal(i,1),M21f);
        if abs(Trimodal(i,1)-M21f(bin1)) > abs(Trimodal(i,2)-M21f(bin1+1))
            bin1=bin1+1;
        end
        clear t;
        [t,bin2]=histc(Trimodal(i,2),M21f);
        if abs(Trimodal(i,2)-M21f(bin2)) > abs(Trimodal(i,2)-M21f(bin2+1))
            bin2=bin2+1;
        end
        clear t;
        M21f_Mode1 = M21f(1,1:bin1); M21f_Mode2 = M21f(1,bin1:bin2);
        M21f_Mode3 = M21f(1,bin2:M21NumOfFreq);
        M21E_Mode1 = M21E(i,1:bin1); M21E_Mode2 = M21E(i,bin1:bin2);
        M21E_Mode3 = M21E(i,bin2:M21NumOfFreq);
        E_Mode1 = E(i,1:bin1); E_Mode2 = E(i,bin1:bin2);
        E_Mode3 = E(i,bin2:M21NumOfFreq);
        dim_Mode1 = bin1; dim_Mode2 = bin2-bin1+1;
        dim_Mode3 = M21NumOfFreq-bin2+1;

    % Mode 1
    SumE_Model = sum(E_Model);
    if isnan(E(i,:)) == 0
        MomentWaveNet_Model(i,1) =
            sum(E_Model(1:dim_Model).*M21f_Model(1,1:dim_Model))/SumE_Model;
        MomentWaveNet_Model(i,2) =
            sum(E_Model(1:dim_Model).*((M21f_Model(1,1:dim_Model) -
                                          MomentWaveNet_Model(i,1)).^2))/SumE_Model;
        MomentWaveNet_Model(i,3) =
            sqrt(MomentWaveNet_Model(i,2));
        MomentWaveNet_Model(i,4) =
            sum(E_Model(1:dim_Model).*((M21f_Model(1,1:dim_Model) -
                                          MomentWaveNet_Model(i,1)).^3))/SumE_Model/(MomentWaveNet_Model(i,3)^3);
        MomentWaveNet_Model(i,5) =
            sum(E_Model(1:dim_Model).*((M21f_Model(1,1:dim_Model) -
                                          MomentWaveNet_Model(i,1)).^4))/SumE_Model/(MomentWaveNet_Model(i,3)^4);
    end
end
if isnan(M21E(i,:)) == 0
    MomentM21_Mode1(i,1) = sum(M21E_Mode1(1:dim_Mode1).*M21f_Mode1(1,1:dim.MODE1))/SumE_Mode1;
    MomentM21_Mode1(i,2) = sum(M21E_Mode1(1:dim_Mode1).*((M21f_Mode1(1,1:dim_Mode1) - MomentM21_Mode1(i,1)).^2))/SumM21E_Mode1;
    MomentM21_Mode1(i,3) = sqrt(MomentM21_Mode1(i,2));
    MomentM21_Mode1(i,4) = sum(M21E_Mode1(1:dim_Mode1).*((M21f_Mode1(1,1:dim_Mode1) - MomentM21_Mode1(i,1)).^3))/SumM21E_Mode1/(MomentM21_Mode1(i,3)^3);
    MomentM21_Mode1(i,5) = sum(M21E_Mode1(1:dim_Mode1).*((M21f_Mode1(1,1:dim_Mode1) - MomentM21_Mode1(i,1)).^4))/SumM21E_Mode1/(MomentM21_Mode1(i,3)^4);
end

% Mode 2
SumE_Mode2 = sum(E_Mode2);
if isnan(E(i,:)) == 0
    MomentWaveNet_Mode2(i,1) = sum(E_Mode2(1:dim_Mode2).*M21f_Mode2(1,1:dim_Mode2))/SumE_Mode2;
    MomentWaveNet_Mode2(i,2) = sum(E_Mode2(1:dim_Mode2).*((M21f_Mode2(1,1:dim_Mode2) - MomentWaveNet_Mode2(i,1)).^2))/SumE_Mode2;
    MomentWaveNet_Mode2(i,3) = sqrt(MomentWaveNet_Mode2(i,2));
    MomentWaveNet_Mode2(i,4) = sum(E_Mode2(1:dim_Mode2).*((M21f_Mode2(1,1:dim_Mode2) - MomentWaveNet_Mode2(i,1)).^3))/SumE_Mode2/(MomentWaveNet_Mode2(i,3)^3);
    MomentWaveNet_Mode2(i,5) = sum(E_Mode2(1:dim_Mode2).*((M21f_Mode2(1,1:dim_Mode2) - MomentWaveNet_Mode2(i,1)).^4))/SumE_Mode2/(MomentWaveNet_Mode2(i,3)^4);
end

SumM21E_Mode2 = sum(M21E_Mode2);
if isnan(M21E(i,:)) == 0
    MomentM21_Mode2(i,1) = sum(M21E_Mode2(1:dim_Mode2).*M21f_Mode2(1,1:dim_Mode2))/SumE_Mode2;
    MomentM21_Mode2(i,2) = sum(M21E_Mode2(1:dim_Mode2).*((M21f_Mode2(1,1:dim_Mode2) - MomentM21_Mode2(i,1)).^2))/SumM21E_Mode2;
    MomentM21_Mode2(i,3) = sqrt(MomentM21_Mode2(i,2));
    MomentM21_Mode2(i,4) = sum(M21E_Mode2(1:dim_Mode2).*((M21f_Mode2(1,1:dim_Mode2) - MomentM21_Mode2(i,1)).^3))/SumM21E_Mode2/(MomentM21_Mode2(i,3)^3);
    MomentM21_Mode2(i,5) = sum(M21E_Mode2(1:dim_Mode2).*((M21f_Mode2(1,1:dim_Mode2) - MomentM21_Mode2(i,1)).^4))/SumM21E_Mode2/(MomentM21_Mode2(i,3)^4);
end
% Mode 3
SumE_Mode3 = sum(E_Mode3);
if isnan(E(i,:)) == 0
    MomentWaveNet_Mode3(i,1) = sum(E_Mode3(1:dim_Mode3).*M21f_Mode3(1,1:dim_Mode3))/SumE_Mode3;
    MomentWaveNet_Mode3(i,2) = sum(E_Mode3(1:dim_Mode3).*((M21f_Mode3(1,1:dim_Mode3)-MomentWaveNet_Mode3(i,1)).^2))/SumE_Mode3;
    MomentWaveNet_Mode3(i,3) = sqrt(MomentWaveNet_Mode3(i,2));
    MomentWaveNet_Mode3(i,4) = sum(E_Mode3(1:dim_Mode3).*((M21f_Mode3(1,1:dim_Mode3)-MomentWaveNet_Mode3(i,1)).^3))/SumE_Mode3/(MomentWaveNet_Mode3(i,3)^3);
    MomentWaveNet_Mode3(i,5) = sum(E_Mode3(1:dim_Mode3).*((M21f_Mode3(1,1:dim_Mode3)-MomentWaveNet_Mode3(i,1)).^4))/SumE_Mode3/(MomentWaveNet_Mode3(i,3)^4);
end

SumM21E_Mode3 = sum(M21E_Mode3);
if isnan(M21E(i,:)) == 0
    MomentM21_Mode3(i,1) = sum(M21E_Mode3(1:dim_Mode3).*M21f_Mode3(1,1:dim_Mode3))/SumE_Mode3;
    MomentM21_Mode3(i,2) = sum(M21E_Mode3(1:dim_Mode3).*((M21f_Mode3(1,1:dim_Mode3)-MomentM21_Mode3(i,1)).^2))/SumM21E_Mode3;
    MomentM21_Mode3(i,3) = sqrt(MomentM21_Mode3(i,2));
    MomentM21_Mode3(i,4) = sum(M21E_Mode3(1:dim_Mode3).*((M21f_Mode3(1,1:dim_Mode3)-MomentM21_Mode3(i,1)).^3))/SumM21E_Mode3/(MomentM21_Mode3(i,3)^3);
    MomentM21_Mode3(i,5) = sum(M21E_Mode3(1:dim_Mode3).*((M21f_Mode3(1,1:dim_Mode3)-MomentM21_Mode3(i,1)).^4))/SumM21E_Mode3/(MomentM21_Mode3(i,3)^4);
end

Tri_Euclid=NaN(M21TimeStep,3);
for i=1:M21TimeStep
    if isnan(Trimodal(i,:)) == 0
        [t,bin1]=histc(Trimodal(i,1),M21f);
        if abs(Trimodal(i,1)-M21f(bin1)) > abs(Trimodal(i,2)-M21f(bin1+1))
            bin1=bin1+1;
        end
        clear t;
    end
[t,bin2]=histc(Trimodal(i,2),M21f);

if abs(Trimodal(i,2)-M21f(bin2)) > abs(Trimodal(i,2)-M21f(bin2+1))
    bin2=bin2+1;
end

clear t;

M21E_Mode1 = M21E(i,1:bin1); M21E_Mode2 = M21E(i,bin1:bin2);
M21E_Mode3 = M21E(i,bin2:M21NumOfFreq);
E_Mode1 = E(i,1:bin1); E_Mode2 = E(i,bin1:bin2);
E_Mode3 = E(i,bin2:M21NumOfFreq);
dim_Model = bin1; dim_Mode2 = bin2-bin1+1;
dim_Mode3 = M21NumOfFreq-bin2+1;

EuclidTemp_Mode1 = NaN(dim_Model);
EuclidTemp_Mode2 = NaN(dim_Mode2);
EuclidTemp_Mode3 = NaN(dim_Mode3);

% Mode 1
for k=1:dim_Model
    EuclidTemp_Mode1(k) = ((M21E_Mode1(k)–E_Mode1(k))^2)*M21FreqInterval;
end
Tri_Euclid(i,1) = nansum(EuclidTemp_Mode1(:));

% Mode 2
for k=1:dim_Mode2
    EuclidTemp_Mode2(k) = ((M21E_Mode2(k)–E_Mode2(k))^2)*M21FreqInterval;
end
Tri_Euclid(i,2) = nansum(EuclidTemp_Mode2(:));

% Mode 3
for k=1:dim_Mode3
    EuclidTemp_Mode3(k) = ((M21E_Mode3(k)–E_Mode3(k))^2)*M21FreqInterval;
end
Tri_Euclid(i,3) = nansum(EuclidTemp_Mode3(:));

end

%%%%%%%%%%%%%%%%%%%%%%%%%%%
% 5. Return value

Tri_ParaWaveNet=NaN(M21TimeStep,13);
Tri_ParaM21=NaN(M21TimeStep,13);

Tri_ParaWaveNet(:,1)=Tri_HsWaveNet(:,1);
Tri_ParaWaveNet(:,2)=Tri_PeakWaveNet(:,2);
Tri_ParaWaveNet(:,3)=Tri_PeakWaveNet(:,1);
Tri_ParaWaveNet(:,4)=MomentWaveNet_Mode1(:,4);
Tri_ParaWaveNet(:,5)=MomentWaveNet_Mode1(:,5);
Tri_ParaWaveNet(:,6)=NaN;
Tri_ParaWaveNet(:,7)=Tri_HsWaveNet(:,2);
Tri_ParaWaveNet(:,8)=Tri_PeakWaveNet(:,4);
Tri_ParaWaveNet(:,9)=Tri_PeakWaveNet(:,3);
Tri_ParaWaveNet(:,10)=MomentWaveNet_Mode2(:,4);
Tri_ParaWaveNet(:,11)=MomentWaveNet_Mode2(:,5);
Tri_ParaWaveNet(:,12)=NaN;
Tri_ParaWaveNet(:,13)=Tri_HsWaveNet(:,3);
Tri_ParaWaveNet(:,14)=Tri_PeakWaveNet(:,6);
Tri_ParaWaveNet(:,15)=Tri_PeakWaveNet(:,5);
Tri_ParaWaveNet(:,16)=MomentWaveNet_Mode3(:,4);
Tri_ParaWaveNet(:,17)=MomentWaveNet_Mode3(:,5);

Tri_Param21(:,1)=Tri_HsM21(:,1);
Tri_Param21(:,2)=Tri_PeakM21(:,2);
Tri_Param21(:,3)=Tri_PeakM21(:,1);
Tri_Param21(:,4)=MomentM21_Mode1(:,4);
Tri_Param21(:,5)=MomentM21_Mode1(:,5);
Tri_Param21(:,6)=NaN;
Tri_Param21(:,7)=Tri_HsM21(:,2);
Tri_Param21(:,8)=Tri_PeakM21(:,4);
Tri_Param21(:,9)=Tri_PeakM21(:,3);
Tri_Param21(:,10)=MomentM21_Mode2(:,4);
Tri_Param21(:,11)=MomentM21_Mode2(:,5);
Tri_Param21(:,12)=NaN;
Tri_Param21(:,13)=Tri_HsM21(:,3);
Tri_Param21(:,14)=Tri_PeakM21(:,6);
Tri_Param21(:,15)=Tri_PeakM21(:,5);
Tri_Param21(:,16)=MomentM21_Mode3(:,4);
Tri_Param21(:,17)=MomentM21_Mode3(:,5);
Appendix A2

The following illustrate the time series and performance statistics plots at Moray Firth and West of Hebrides respectively.

Figure A2.1: Primary validation parameters at Moray Firth in 2009
Figure A2. 2: Secondary validation parameters at Moray Firth in 2009

- **RMSE**: 0.45
- **Bias**: -0.16
- **Scatter Index**: 41.76
- **Correlation, R**: 0.86
- **p-value**: 0.00

Figure A2. 3: Predicted versus observed significant wave heights at Moray Firth
Figure A2. 4: Predicted versus observed peak frequencies at Moray Firth

Figure A2. 5: Predicted versus observed peak energy densities at Moray Firth
Figure A2. 6: Primary validation parameters at West of Hebrides in 2009
Figure A2. 7: Secondary validation parameters at West of Hebrides in 2009

Figure A2. 8: Predicted versus observed significant wave heights at West of Hebrides

West Of Hebrides

$\text{RMSE}: 0.48$

$\text{Bias}: -0.13$

$\text{Scatter Index}: 16.81$

$\text{Correlation, R}: 0.96$

$\text{p-value}: 0.00$
Figure A2. 9: Predicted versus observed peak frequencies at West of Hebrides

Figure A2. 10: Predicted versus observed peak energy densities at West of Hebrides
Appendix A3

Equations for performance statistics:

\[
RMSE = \sqrt{\frac{1}{N_i} \sum_{i=1}^{N_i} (P_i - O_i)^2}
\]

\[
Bias = \sum_{i=1}^{N_i} \frac{1}{N_i} (P_i - O_i)
\]

\[
Scatter\ Index = \frac{RMSE}{\frac{1}{N_i} \sum_{i=1}^{N_i} O_i} \times 100
\]

\[
Correlation = \frac{\sum_{i=1}^{N_i} (P_i - \bar{P}) (O_i - \bar{O})}{\sqrt{\sum_{i=1}^{N_i} (P_i - \bar{P})^2} \sqrt{\sum_{i=1}^{N_i} (O_i - \bar{O})^2}}
\]

Where \(N_i\) is the number of data, \(P_i\) is the predicted parameter and \(O_i\) is the observed parameter.

Reference: