Wave Energy Scotland

Operations and Maintenance Simulation Tool

Weather Simulation Report

This report details the logic, algorithms and validation process of a Markov-based model which has been created to provide an Operations and Maintenance (O&M) simulation tool with a realistic time series of weather conditions. This enables the tool to assess weather windows suitable for marine operation and calculate revenue generated by a wave energy array. The work has stemmed from an Engineering Doctorate sponsored by Wave Energy Scotland in partnership with the Industrial Doctoral Centre for Offshore Renewable Energy.

Related documents

- *WES 2017. Operations and Maintenance Simulation Tool. MS Excel (Macro-Enabled)*
- *WES 2017. Operations and Maintenance Simulation Tool - Functionality Report. PDF*
- *WES 2017. Operations and Maintenance Simulation Tool - User Guide. PDF*
- *WES 2017. Operations and Maintenance Simulation Tool - Future Upgrades. PDF*

Revision 1

CONTENTS

1 INTRODUCTION

1.1 BACKGROUND OF THE O&M TOOL

The seas around Scotland are some of the most powerful and inhospitable on the planet, which makes them ideal for deploying wave energy converters (WECs) – devices that use wave action to generate electricity. Wave Energy Scotland (WES) was set up by the Scottish Government in 2014 to fund and support innovative solutions to the technical challenges of harnessing energy from the waves. One aspect making the commercialisation of wave energy devices difficult is the uncertainty surrounding lifetime costs of wave energy arrays, particularly during the operations and maintenance (O&M) phase. Having a reliable means of estimating these costs in as realistic a way as possible is therefore hugely important in the development of the wave energy sector.

Wave Energy Scotland has released an O&M simulation tool to analyse the lifetime logistics of a wave energy array. The tool originates from a research project sponsored by WES through the Industrial Doctoral Centre for Offshore Renewable Energy (www.idcore.ac.uk). The tool was initially developed in collaboration with Pelamis Wave Power, one of the world's leading wave energy technology companies at the time, with an emphasis on commercial-scale WECs rated up to 1MW. The tool's methodology was then applied to the much smaller off-grid WECs being designed by Albatern, another Scottish wave energy developer. These collaborations have enabled WES to produce an O&M simulation tool which can be applied to an array of any wave energy converter.

An O&M simulation model is an extremely useful tool for three primary reasons:

1. At early stages of development, an O&M model can help identify critical components which would have the biggest impact on array performance, thus providing feedback into the device design
2. The tool provides estimates of array availability, revenue and operational expenditure which helps to refine Levelised Cost of Energy (LCOE) calculations
3. As device development moves towards real sea testing, the tool can assist in planning aspects of the O&M strategy for future arrays

The tool has been created using Microsoft Excel and the associated VBA programming language. It uses the Monte Carlo method to simulate the occurrence of faults on each WEC in an array by utilising failure rate data. All the components of the device are represented by fault categories, assigned following a Failure Modes and Effects Analysis (FMEA) of the device. The user can choose whether certain faults can be repaired whilst the WEC is offshore, or if all faults require the device to be towed to the safety of a sheltered quayside or onshore O&M base for repair. This 'reactive' maintenance modelling is coupled with the option to include modelling of 'proactive' routine servicing of WECs. Maintenance parameters, such as repair times and parts costs, are defined by the user, as are other aspects of the O&M strategy, such as weather limits for marine operations. The model utilises a time series of weather conditions in order to assess windows of accessibility and calculate revenue generated by the array at each time step. The model simulates the array lifetime as realistically as possible by enforcing logistical constraints, including technician availability and quayside access. A full breakdown (per device and per year) of outputs including availability, revenue and operational expenditure is presented, as well as a table attributing costs to each fault category.

1.2 WEATHER SIMULATION

The time series of weather conditions provided as an input to the O&M model is generated from a hindcast dataset for a particular site. Hindcast datasets of weather conditions can be created using software packages such as SWAN (Simulating Waves Nearshore, Delft University of Technology) and usually have the resolution of one, three or six hours. They contain a number of parameters defining the weather conditions at a particular location, most notably significant wave height (Hs), wave peak period (Tp), wave energy period (Te) and wind speed (U). As described in the 'Functionality Report' (WES 2017a), the hindcast dataset is not used directly to define weather windows and calculate revenue in the O&M model. Instead, a multivariate Markov Chain Model (MCM) has been developed to simulate realistic weather conditions for use in the O&M tool (Gray 2017). This enables the model to be used to undertake sensitivity studies into variations in weather conditions at the same site and for different project lifetimes. This Markov Chain method has been shown to produce realistic weather window simulations for use in offshore wind O&M tools (Hagen et al. 2013, Scheu et al. 2012).

The key principle of the Markov Chain method is that the weather conditions at a given time step depends solely on the weather conditions at the previous time step. The process involves breaking down the hindcast dataset into months and identifying the probability of transitioning from one set of weather conditions (referred to as a 'sea state') to another within that month. In essence, a matrix for each month is created, as shown in Figure 1.1.

		Sea State$_{t+1}$							
		1	2	3	4	5	6	...	n
Sea State$_t$	1	P_{11}	P_{12}	P_{13}	P_{14}	P_{15}	P_{16}	...	P_{1n}
	2	P_{21}	P_{21}	P_{21}	P_{21}	P_{21}	P_{21}	...	P_{2n}
	3	P_{31}	P_{31}	P_{31}	P_{31}	P_{31}	P_{31}	...	P_{3n}
	4	P_{41}	P_{41}	P_{41}	P_{41}	P_{41}	P_{41}	...	P_{4n}
	5	P_{51}	P_{51}	P_{51}	P_{51}	P_{51}	P_{51}	...	P_{5n}
	6	P_{61}	P_{61}	P_{61}	P_{61}	P_{61}	P_{61}	...	P_{6n}
	...	...	...	...	...	...	...	...	...
	n	P_{n1}	P_{n2}	P_{n3}	P_{n4}	P_{n5}	P_{n6}	...	P_{nn}

Figure 1.1. Probability matrix representative of the Markov Chain Model process

As a result, the MCM generates a time series which will contain different values from the observed ones in the hindcast dataset, whilst retaining similar statistical metrics and seasonal trends. This report aims to clearly explain the logic (section 3) and code (section 4) of the MCM, as well as providing evidence of validation (section 5). The code of the model is written in VBA (Visual Basic for Applications), thereby utilising data contained within Microsoft Excel spreadsheets. The 'Functionality Report' (WES 2017a) should be referred to for an explanation of how the Markov Chain-generated time series' are incorporated into the Operations and Maintenance software tool.

2 TERMINOLOGY

2.1 GENERAL

CDF - Cumulative Distribution Function

EMEC - European Marine Energy Centre

Hs - Significant Wave Height

IDCORE - Industrial Doctoral Centre for Offshore Renewable Energy

MCM - Markov Chain Model

NaN - Not a Number

O&M - Operations and Maintenance

Sea state - a combination of parameters defining weather conditions

Te - Wave Energy Period

Tp - Peak Wave Period

U - Wind speed

U10 - Wind Speed at 10m above sea level

VBA - Visual Basic for Applications

WEC - Wave Energy Converter

WES - Wave Energy Scotland

Weather window - a period where weather conditions remain accessible for marine operations

2.2 VBA TERMINOLOGY

Argument - a variable sent to a procedure for use

Call - a procedure is 'called' by another one in order for it to undertake its action

Cells - an in-built VBA function used to refer to an Excel cell

Class module - secondary object in VBA programming

Const - a term used to assign a variable a certain value at the beginning of an object or procedure

Data type - format in which data is stored in a variable

Dim - a term used to define a variable as a particular data type

Function - a procedure that performs an action and can return values

Module - primary object in VBA programming

Procedure - a set of programming instructions to perform some action

Subroutine - a procedure that performs an action

Variable - a temporary holder of information. Data can be in numerous types, such as String (words), Integers (whole number between -32,768 and 32,767), Long (whole number up to 2 billion) or Double (a decimal number)

Worksheet - VBA term for an Excel spreadsheet

3 METHODOLOGY

3.1 SELECTING HINDCAST DATA

The three parameters chosen to represent weather conditions in the O&M tool are significant wave height (Hs), wave period and wind speed. As stated in the 'Functionality Report' (WES 2017a), the wave period can be represented by either wave peak period (Tp) or wave energy period (Te), but must match the format of the power matrix used to calculate revenue generated by a wave energy converter (WEC) in the O&M tool. The probabilistic nature of the Markov Chain Model (MCM) means that it is important to obtain a hindcast dataset containing these parameters which is of a sufficient length and has an appropriate resolution. A dataset with a minimum length of 10 years (with a resolution of 1, 3 or 6 hours) can be deemed suitable for this method, as it represents sufficient inter-annual variation of the resource (Equimar, 2011).

3.2 'BINNING' DATA

It is necessary to group the hindcast values of significant wave height, wave period and wind speed into 'bins'. The sizes of the Hs and wave period bins must match the values of the power matrix for the O&M tool to function correctly (see 'Functionality Report', WES 2017a). A bin is represented by the midpoint value (e.g. 2.25m Hs denotes the bin 2m ≤ Hs < 2.5m. The minimum and maximum constraints of each parameter (e.g. a maximum of 50kts wind speed) must be selected in order to include at least 95% of possible observations from the hindcast dataset. In the rare cases where observations lie outside of these limits, they are placed in the closest bin. An <u>example</u> of appropriate bin ranges and resolutions is as follows:

- Significant wave height (Hs) ranging from 0.25m to 9.75m, in steps of 0.5m.
- Wave period ranging from 3s to 15s, in steps of 2s.
- Wind speed ranging from 2.5kts to 47.5kts, in steps of 5kts.

3.3 SEA STATE ID

It is necessary to have a system of assigning IDs to sea states (i.e. a combination of all three 'binned' parameters) that is consistent across all months as this enables monthly transitions to be easily carried out (see section 3.5). Whilst all three parameters in the model are vital, significant wave height is viewed as the most important due to its dominant role in defining weather windows. As a result, the sea state IDs are assigned in order of Hs first, followed by wind speed, then wave period. Therefore, the actual ID number a sea state is given is trivial and doesn't have any significant meaning in terms of the parameters (other than sea states with high ID numbers containing large Hs values). However, it is vital to group combinations of the three parameters together in this manner to enable the MCM to easily calculate transitions from one time step to the next.

3.4 MONTHLY DATA

In order to account for seasonal variability, the original dataset (once binned) is broken up into months. The monthly data includes the last five days of the previous month and first five days of the next month. For each time step, the sea state ID of the next interval is recorded, thus providing the possible transitions with which to carry out the probabilistic calculations of the MCM. This is fundamental to the Markov property, as the modelled sea state at any given interval is determined solely by the sea state at the previous interval.

3.5 TRANSITION PROBABILITIES

The occurrence of each sea state within the monthly data is identified, and the possible transitions to the next interval are listed. From this point, it is possible to calculate the probabilities of each of the possible transitional sea states occurring using the formula below:

$$p_{ij} = \frac{N_{ij}}{N_i}$$

Equation 3.1

Where:

p_{ij} = probability of transitioning from sea state *i* to state *j* during this month
N_{ij} = number of observed transitions from sea state *i* to state *j* in monthly dataset
N_i = number of occurrences of sea state *i* in monthly dataset

3.5.1 Starting Probabilities

The sea state occurring at the very first interval of the modelled data has to be chosen probabilistically. To achieve this, the following formula is applied to every sea state within each monthly dataset:

$$p_i = \frac{N_i}{N}$$

Equation 3.2

Where:

p_i = probability of starting at sea state *i* during this month
N_i = number of occurrences of sea state *i* in monthly dataset
N = total number of intervals in monthly dataset

By applying this universally, it means that the modelled dataset can begin at any month of the users choosing, creating a more versatile model. In addition, it provides a fail-safe option for selecting monthly transitions (see next section).

3.5.2 Monthly Transitions

A consistent sea state ID system has been used to enable the MCM to calculate transitions from month to month. In some situations, the final state from the previous month may not occur in the dataset for the next month. This will mean that the next state cannot be chosen probabilistically using Equation 3.1. To account for this, a three tier hierarchical system of determining the sea state at the first interval in the next month is as follows:

Monthly transition 1: In most cases, the final state from the previous month will occur in the next month. If so, the sea state at the first interval in the next month is selected using Equation 3.1.

Monthly Transition 2: If the final state from the previous month does not occur in the next month, then the possible next states (from the previous month) are considered. Any sea states from this list which do not appear in the next month dataset are deleted. If one or more states remain, then one is chosen to become an intermediate state. This is done probabilistically using a modified version of Equation 3.1:

$$p_{i1j2} = \frac{N_{i1j1}}{new\ N_{i1}}$$

Equation 3.3

Where:

p_{i1j2} = probability of selecting state *j* in next month from state *i* in previous month

N_{i1j1} = number of observed transitions from sea state *i* to state *j* in previous month

new N_{i1} = number of occurrences of sea state *i* in monthly dataset, once non-applicable states have been deleted

The intermediate state is then treated as if it were the final state in the previous month. Equation 3.1 can once again be used to determine the sea state at the first interval in the next month. This monthly transition option is effectively like skipping one state.

Monthly transition 3: In very rare situations, none of the possible next states from the previous month will exist in the next month dataset. In these cases, the starting probabilities calculated previously (section 3.5.1) are used to select the new sea state. Although this may result in larger jumps (in terms of Hs for example) from one interval to the next, it has been deemed acceptable due to the fact that this option is required for less than 0.5% of monthly transitions for any given modelled time series (see Appendix A, section 8.1).

The three tier monthly transition hierarchy is shown graphically in Figure 3.1 below:

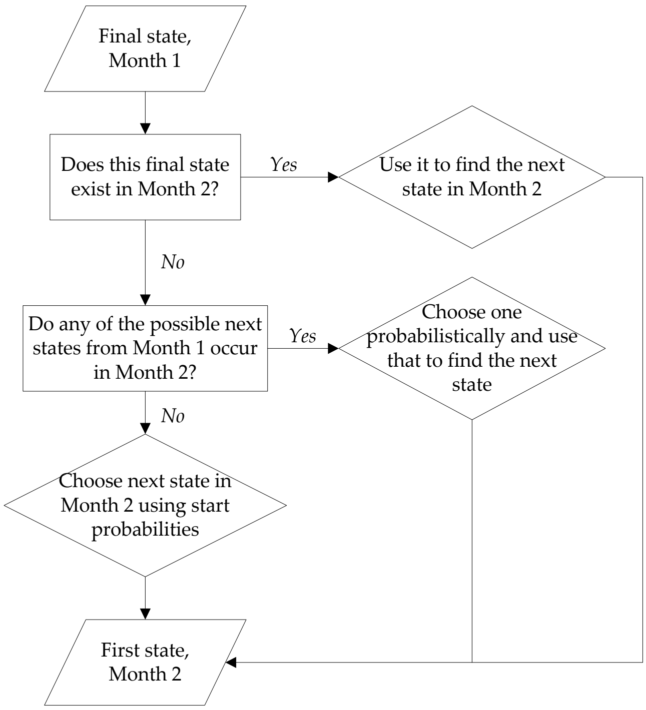

Figure 3.1. Decision flowchart for monthly transitions (Gray 2017)

3.6 MODELLED DATASET

Following the steps outlined here, the MCM can generate a time series of weather conditions, the length of which (i.e. number of years) is chosen by the user. The modelled dataset, providing values for significant wave height, wave period and wind speed, will have similar statistical parameters as the original hindcast dataset. A thorough validation process has been undertaken to ensure this is the case (section 5). The generated dataset can then be placed in the O&M tool in order to identify weather windows and WEC power output. However, it is useful to generate several datasets of the same length as this enables weather-based sensitivity studies to be undertaken. Therefore, datasets of the same length must be stored in the same Excel workbook, whose name corresponds to the particular wave energy site represented by those weather conditions. Each spreadsheet within that workbook must be named appropriately. For example, the first 20 year time series for the Farr Point wave energy site should be named FPT-020-001 (i.e. site acronym - project lifetime - ID of time series). This naming convention is defined by the VBA code contained within the O&M tool (refer to the 'Functionality Report', WES 2017a). A new workbook must be created when a new project lifetime length is to be assessed.

4 VBA ALGORITHMS

The Markov Chain Model (MCM) has been built in Microsoft Excel and uses a large amount of the associated Visual Basic for Applications (VBA) coding language. This section details the pieces of code that are used and attempts to explain the algorithms. It is assumed that the reader has some prior knowledge of VBA. It is strongly recommended that the reader has access to the MCM whilst studying this section of the document. VBA procedures and variables are given italic script.

4.1 FUNCTIONS

A number of useful VBA Functions have been developed to assist in the MCM process. The majority of these are generalised and could be used by any program. However, there are some which are specific to this MCM, and should therefore be used with care.

4.1.1 Number of Rows

The *num_rows* function is used extensively throughout the MCM. This calculates the number of rows in an Excel worksheet which are not empty. This is vital as there are several loops throughout the MCM where each row needs to be considered. Calling procedures specify the worksheet and the maximum number of rows to search through. In later versions, there is also an option to select a start row other than the first. This is useful when the first or second rows are initially left blank in order to print headers at a later date.

4.1.2 Insert Sheet

Insert_sheet checks if a worksheet with a certain name already exists in the Excel file. If it does not exist then it is inserted at the end of the workbook.

4.1.3 Delete Empty Rows

Frequently in the MCM process, there are many rows which get deleted or, in the case of creating the monthly data sheets, newly created worksheets don't contain as many rows as the data source (i.e. worksheet with all the original data). In these situations, it can leave scrolling down a worksheet difficult as the scroll bar does not get resized automatically. The *delete_empties* function solves this by deleting all those rows after the last filled row of a worksheet. The workbook must be saved before the scroll bar is resized.

4.1.4 Join Text

A function, *JoinText,* is a vital addition to the model as a way of tackling multiple parameters. This takes a range of values and joins them as text, separated by a character of the users choosing. This is used to combine all three parameters (Hs, wave period and wind speed) in order to assign a sea state ID.

4.1.5 Assign Sea States

All the functions mentioned previously have the ability to be used in any Excel file. *Assign_states* is the first one which is workbook-specific. This is because it requires the three parameters to be in columns in the correct order. *Assign_states* loops through every row (i.e. interval) in a specified worksheet and uses *JoinText* (see section 4.1.4) to group the three parameters at that interval. From here, the function can search through the sea state ID worksheet (discussed in section 3.3) and identify the correct state for those parameters. This is used to complete the 'Reference' worksheet which contains the binned values of the original hindcast dataset.

4.1.6 Number of Days per Month

The number of days in each month is required at several stages throughout both the Markov process of generating a new time series and the validation procedure. These values are stored in the function *num_month_days*. It is worth noting that the number of days in February has been set as 28. The decision to ignore leap years and delete any data for 29th February has been seen as a valid assumption for the MCM and the O&M tool.

4.1.7 Month ID and Month Name

The function *get_month_ID* takes a month name in the format "Jan", "Feb" etc. and returns an integer 1 to 12. *get_month_name* carries out the reverse process. This is useful for printing and calculations throughout the MCM.

4.1.8 Timer

A timer has been installed so that a message box appears after selected subroutines are run, telling the user how long it took. This is a useful addition to the model.

4.1.9 Delete Groups of Worksheets

The function *delete_yearly_sheets* is used to delete all worksheets containing modelled data. This is achieved by deleting worksheets that have 'Year' as the first four characters of their name. It is necessary to make Excel stop displaying alerts for the duration of this function. *delete_most_sheets* is also available, which deletes every sheet apart from 'Original' (containing the hindcast observed data).

4.1.10 User Defined Years

When generating a new modelled time series, it is necessary for the user to specify the number of years they would like to create. The function *user_defined_years* displays an input box requesting a value and also indicates the expected run time. The default value is 20 years. A form of error handling is included which ensures that the requested number of years is between 1 and 120.

4.1.11 Count Modelled Years

A method of obtaining the number of years of modelled data is required. This is not used for the MCM calculations, but is useful during the validation stage. It is achieved using the function *count_year_shts*.

4.2 DATA SORTING

The first step is to get the original data into the format required by the MCM calculations.

4.2.1 Organise Original Data

Original data is copied into the workbook from another file. Wind speed in knots is converted from the m/s equivalent using Excel. The subroutine *covert_and_sort* is then used to carry out three key tasks: i) delete any duplicate rows that may exist, ii) delete any data for 29th February, and iii) calculate wave energy period (Te) from wave peak period (Tp), if required. This dataset is stored in a worksheet named 'Original'. Care should be taken here as Tp needs to be in a certain column.

4.2.2 Set Up ID Reference

The reference worksheet for the sea state IDs is created using the subroutine *setup_ID_ref*. Firstly, the worksheet 'ID_ref' is created and cleared. The following formula is then used to fill an array for each of the three parameters with the mid-bin values defined during the 'binning' process (section 3.2):

$$Array\ value\ (i) = \left(i \times \frac{k}{2}\right) + \left(\frac{k}{2} \times (i-1)\right)$$

Where:

i = position in array

k = resolution of parameter

These values are then printed to the 'ID_ref' worksheet and each possible combination is assigned a sea state ID, as described in section 3.3. Wind speed is in column A, significant wave height in column B and wave period in column C. This order of parameters is consistent throughout the model and must match the convention of the O&M tool (refer to the 'Functionality Report', WES 2017a). The state ID is displayed in column D.

4.2.3 Creating Reference Sheet

The subroutine *create_ref_sheet* is used to group the hindcast data into bins. The worksheet 'Reference' is created and all the data from the hindcast sheet is copied across. For each interval and each parameter, the following process is carried out:

- Round each value up to the upper level of the relevant bin (e.g. round 2.32m Hs up to 2.5m) using the formula (note: *k* = resolution of parameter):

$$New\ Value = k\ \times RoundUp\left(\frac{Old\ Value}{k}, [to\ 0\ d.p.]\right)$$

- Place the new value into the correct bin based upon certain conditions to conform to the ranges defined during the 'binning' process (section 3.2):
    - If the old value is below the minimum constraint, then add *k/2* to the new value
    - For any parameter, if the Old Value was greater than the maximum bin, then place the New Value in this maximum bin
    - Otherwise, subtract *k⁄2* from the New Value to place it in the correct bin

The newly binned values are printed in columns to the right of the copied data. As a result, it is necessary to delete the copied data and replace it with the new values. Once complete, the *create_ref_sheet* subroutine calls the *Assign_states* function (see section 4.1.5) to allocate each interval a sea state ID based on the parameter values at that period. The 'Reference' worksheet is then ready to be used by the remaining procedures of the MCM, as well as during the validation process.

4.3 CREATE MONTHLY DATA

It is then necessary to break the binned data down into months in order to account for seasonal variability.

4.3.1 Copy Months

The first time the variable *MonthArray* is seen is in the *copy_months* subroutine. This is an array containing all month names, in the format 'Jan', 'Feb' etc., and will be used extensively throughout the rest of the model. It is particularly useful in the validation process, where seasonal variability is scrutinised closely. The *copy_months* subroutine is used to create sheets containing all relevant data for each month. The same process is carried out for every month name in a loop, 'Jan' to 'Dec'.

Firstly, a new worksheet is created which the same name as the month. All data from the 'Reference' worksheet is copied into every monthly sheet, including the sea state ID. An extra column is added which is used to contain the sea state ID for the next interval. This information exists for almost all of the data. However, the final value in the 'Reference' dataset will not have a next state. Similarly, the dataset might not end precisely on 31st December; therefore the final entry in the first 'block' (post-expansion) may also not have a next state. 'N/A' is printed in the relevant cell for both these cases. Next, the function *get_month_id* is used to enable all the data from months other than the one under consideration, as well as the two either 'side', to be deleted. In other words:

- When the loop is looking at 'Jan', data outside 'Jan', 'Feb' and 'Dec' is deleted
- When the loop is looking at 'Dec', data outside 'Dec', 'Jan' and 'Nov' is deleted
- For other months: e.g. when considering 'Apr', data outside 'Apr', 'May' and 'Mar' is deleted

As discussed in section 3.4, the monthly datasets will include values from the last five days of the previous month and first five days from the next month. To achieve this, *get_month_id* is once again used, with 'Jan' and 'Dec' being treated as special cases as before. In addition, the function *num_month_days* is used to calculate the day IDs of the last five days of a month. Any values in days that lie outside the desired limits are deleted.

For every monthly dataset, as a form of error handling, the last interval of the fifth day in the next month is considered using the additional subroutine *account_for_mismatch*. In very rare occasions, it is possible that the next state identified here will not be present throughout the rest of the dataset. If that is the case then the next state will be replaced with 'N/A'.

The *delete_empties* function then resizes the scroll bar, completing the month worksheet.

4.3.2 Create Monthly State Sheets

The key worksheets which will be used in generating a synthetic time series are the monthly state sheets. These are created and populated using the subroutine *create_states_sheets* and associated procedures. Once again, this is carried out for every month defined in *MonthArray*.

Initially, the month state worksheet is created and named in the format '*month*_States' (e.g. 'Jan_States', 'Feb_States' etc.). The parameter values and associated sea state IDs are copied across from the relevant monthly dataset. The dates are no longer required and are therefore ignored. All duplicated data is then deleted using VBAs in-built function *RemoveDuplicates*, leaving only unique sea states in the worksheet. At this point, the empty rows are deleted using the *delete_empties* function in order to resize the scroll bar. The remaining table is then reorganised in order of the sea state ID. This isn't necessary for the upcoming calculations, but it is intuitive to the user. A new column is added with the header 'Occurrence'. A loop then counts the number of appearances of each unique sea state in the monthly dataset and prints this value.

After the occurrence of each unique sea state has been printed, the *create_states_sheets* subroutine then calls the following procedures in order:

- *print_possible_states*
- *insert_rows_below*
- *calc_probabilities*
- *calc_start_probs*

4.3.3 Print Possible States

The function *print_possible_states* is used to print a list of the possible sea states that could occur immediately after the current state. A nested loop is used so that every row (i.e. interval) in the month sheet is considered for each row in the monthly state sheet, as follows.

Initially, the current state and its occurrence are identified in the month state worksheet. Then, every interval in the full monthly dataset (contained within the month worksheet) is looked at. When a matching sea state is found, the state ID for the next interval (printed during the *copy_months* subroutine) is considered. As described in section 4.3.1, some next state cells will contain 'N/A' as a form of error handling. In these cases, the next state is ignored and the total occurrence initially identified is decreased accordingly. However, in the majority of situations, the next state will be successfully identified. An array, *State_ID_Store*, is created to keep track of the next state IDs. The monthly sheet loop is exited once the size of *State_ID_Store* has reached the total occurrence value in order to avoid unnecessary looping. A final step is to print all the possible next states for each unique sea state in the month state worksheet. These are printed in columns along the same row.

4.3.4 Insert Rows Below

At this point, the data contained within the month state sheets is as follows:

- Each unique sea state ID, along with its associated wind speed, Hs and wave period values
- The total occurrence of the sea state in the monthly dataset (modified in some cases)
- A list of possible states for the next interval

The remaining information required is the transitional and starting probabilities outlined in section 3.5. To fit this data onto the month state worksheets, a new line is needed below each row. This is achieved via the *insert_rows_below* function.

4.3.5 Calculate Transition Probabilities

The function *calc_probabilities* calculates the probability of transitioning from a sea state at one time step to a particular sea state at the next interval, as defined by Equation 3.1. A loop is set up for every unique state in the month state worksheet. It is important to note that the step size of the loop is now 2, due to the newly inserted rows. The number of possible next states (i.e. N_i in Equation 3.1) is identified and the IDs are collected in the range *SrchRng*. Excel's in-built function *CountIf* is then used to count the number of observed transitions for each next state (i.e. N_{ij} in Equation 3.1). The calculated probability (i.e. p_{ij}) is then printed in the cell directly below the associated next state.

Another nested loop is included after this step to delete repeated next states. It is unnecessary for the month state worksheets to store repeated values. To achieve this, each possible next state is considered from left to right (note: IDs have been printed in columns) in a 'Do' loop. The position of the sea state under consideration is identified with the variable *ref_col*. For each of these states, every other ID is looked at from right (i.e. final possible state) to left (not including the state under consideration) as a nested loop, identified by *srch_col*. A Boolean variable *Repeat* is then used to identify whether the 'search' state is the same as the 'reference' state under consideration. In these cases, the 'search' state and associated probability are both deleted. Otherwise, the nested loop continues onto the next possible 'search' state to the left. Once this nested loop has searched through all the other possibilities, a new 'reference' state is selected by the 'Do' loop. In situations when states are deleted, the first few *srch_col* cells (search cells) of the next 'reference' may be empty. In these cases, the loop will effectively ignore them and move onto the next 'search' state. This process has been thoroughly tested and operates as desired.

4.3.6 Calculate Starting Probabilities

As discussed in section 3.5.1, the first sea state of the modelled time series will be chosen probabilistically. The probabilities required to achieve this are printed via the *calc_start_probs* function. This makes use of Equation 3.2.

The total number of intervals in the monthly dataset (once any 'N/A' instances have been taken into account) is calculated using Excel's in-built *Sum* function. This becomes *N* in Equation 3.2. A loop is then carried out to calculate the starting probability for each unique sea state. These starting probabilities are printed on every month state sheet, as this provides a third option for the monthly transitions discussed in section 3.5.2.

4.4 Generate a Time Series

The state sheets (one for each month) contain all the necessary information with which to generate a synthetic time series with the same statistical parameters as the original dataset. The entire process is carried out using the *generate_time_series* subroutine. As an error handling measure, if a problem occurs during the running of this procedure then a message box will appear directing the user to the cause. This addition was used extensively throughout the development of the procedure and has been left in as a failsafe.

A number of key steps are taken before *generate_time_series* goes through the main looping processes. It is necessary to use VBAs in-built *Randomize* function in order to produce different time series' each time the model is run. Initially, the user is requested to enter the number of desired years they wish to generate. This is controlled with the *user_defined_years* function discussed in section 4.1.10. A variable, named *ThisState*, is used to store the ID of a sea state at a given interval of the modelled time series. It is necessary to initialise *ThisState* to be zero so that residual state IDs from the previously modelled time series do not interfere with the new run. The function *delete_yearly_sheets* (see section 4.1.9) is called to avoid confusion between previously generated time series'. A temporary worksheet, named 'temp', is created in order to store temporary values useful for later calculations. *MonthArray* is again used here to distinguish between different months. However, where before it stored the month names in order from 'Jan' to 'Dec', the *MonthArray* used in *generate_time_series* is ordered 'Dec', 'Jan'...'Nov', as shown in Figure 4.1. This is because the O&M tool requires weather data in this format in order to yield seasonal-based results (e.g. the meteorological definition of winter is December to February inclusive).

```
MonthArray = Array("Dec", "Jan", "Feb", "Mar", "Apr", "May", "Jun", "Jul", "Aug", "Sep", "Oct", "Nov")
```

Figure 4.1. *MonthArray* used in *generate_time_series*

4.4.1 Set Up Yearly Sheets

The main looping process involves going through each year, each month and then through every interval. The top loop runs from year 1 up to the desired number of years specified by the user. The *setup_year* function is used to create each year sheet, labelled 'Year X' (where *X* is the year number). The headers are printed, along with the year number, month ID, and day and hour values. The month ID is obtained using the *get_month_id* function. The day number is calculated by using Excels in-built *RoundUp* function, dividing the position of each interval by the number of hour intervals per day.

Throughout the monthly loop, a variable named *Month_start_row* defines the correct row in the modelled year worksheets for printing. For every year, this is initialised to be 2 in order to account for the headers printed in the first row.

4.4.2 Select Starting State

The monthly loop now begins, running for every month identified in *MonthArray*. The first task is to identify which sea state begins that particular month. The first month ('Dec') in year 1 is a special case. As previously discussed in section 3.5.1, this initial state is selected using starting probabilities. Firstly, the function *sort_start_table* is used to extract the starting probabilities from the 'Dec_States' worksheet. These probabilities are printed to the 'temp' worksheet alongside the correct sea state ID. The values need to be in a certain format for a probabilistic selection to take place. As a result, the state is printed in the second column whilst the first column contains the cumulative sum of the probabilities (starting at 0 for the first state). Once *sort_start_table* has completed this function, a random number between 0 and 1 is generated back in the *generate_time_series* monthly loop. Excel's in-built function *VLookup* is a vital part of the Markov Chain Model. By applying the newly generated random number and the *VLookup* function to the table in the 'temp' worksheet, it is possible to probabilistically select the first sea state of the modelled data. The state ID is stored in the variable *ThisState*. The associated values of wind speed, significant wave height and wave period are obtained from the month state worksheet and printed to the modelled year worksheet by the *print_start_state* function.

4.4.3 Choosing a Monthly Transition

The very first modelled sea state is unique as it does not involve the three-tier monthly transition hierarchy outlined in section 3.5.2. For every other month in every year, the *state_to_start_new_month* function is required. Every sea state will have been modelled for all previous intervals when *state_to_start_new_month* is called. The Boolean variables *Month_Trans_1* and *Month_Trans_2* are used to determine whether a monthly transition option has been successful. These parameters are initialised to be False (i.e. unsuccessful). The sea state from the last interval of the previous month is recognised as *PrevState*.

Monthly Transition 1

In monthly transition 1, every row (i.e. interval) in the state worksheet of the new month is looked at. If *PrevState* is found to exist in that dataset then it is selected as the reference state with which to probabilistically choose the next sea state. It is likely that most monthly transitions will successfully be selected in this manner. If that is the case then *Month_Trans_1* is changed to be *True* and the loop is exited in order to avoid unnecessary calculations. The ID of the selected sea state is stored in the variable *ThisState*.

Monthly Transition 2

If *Month_Trans_1* is returned as *False* (i.e. monthly transition 1 has not been a success) then monthly transition option 2 is required. Initially, the previous month name is identified by using the *get_month_id* and *get_month_name* functions. A new function named *get_this_states_row* is then used to obtain the position of *PrevState* within the previous month's state worksheet. With this information, the number of possible next states from the previous month dataset can be identified. This is vital as the function of monthly transition 2 is to check if any of these states exist in the new month's dataset, delete any which do not, and then probabilistically select one of the remaining states.

A form of error handling is present in the code to ensure that if the possible number of next states is zero (i.e. the state has no options) then *state_to_start_new_month* will move onto monthly transition option 3. However, in almost all cases, possible next states will be evident. First, a variable *num_deltd* is initialised to zero to indicate that no states have been ruled out at this point. The 'temp' worksheet is then cleared to avoid confusion with previous probability tables (such as that set up to select the starting probability). A difference from when the 'temp' worksheet was used for the starting probability is that the occurrence is also required here. Each of the possible next states is printed to the 'temp' worksheet in the second column. The total occurrence of *PrevState* in the previous month is identified. The occurrence of each possible next state is then calculated by multiplying its probability by the total occurrence. This step is required because of the fact that one or more of the possible next states will get deleted if they do not exist in the new month's dataset. This intermediate step ensures that a modified probability can be easily obtained.

A new loop is then required to consider each of the possible next states in reverse order (i.e. from bottom to top). This is required to enable row deletion to operate correctly. A Boolean variable, *This_State_Exists*, is initialised to be *False* (i.e. assumes the state does not exist unless found). A nested loop then looks at each row in the new month's state worksheet. If the sea state in any row in new month dataset is found the be the same as the possible next state under consideration, then *This_State_Exists* is changed to *True* and the nested loop is exited. If the nested loop finishes and *This_State_Exists* is still *False* (i.e. the possible next state does not exist in the new month) then the row in the 'temp' worksheet is deleted. Also, the variable *num_deltd* is updated every time a state is deleted.

The new number of possible next states is then calculated (i.e. the final value of *num_deltd* is subtracted from the original number of possible next states). This is stored in the variable *new_num_poss_states*. If all of the possible next states have been deleted because they don't exist in the new month's dataset, then *new_num_poss_states* will equal zero. If this is the case then *Month_Trans_2* is chosen to be *False*, forcing the procedure to move onto the third monthly transition option. If, however, one or more possible next states do exist in the new month's dataset then monthly transition 2 continues.

Month_Trans_2 is changed to *True* and the new total occurrence is calculated by using Excel's in-built *Sum* function on the remaining states in the 'temp' worksheet. The formatted table required for *VLookup* is then set up. This is achieved by calculating the cumulative probabilities for the remaining sea states using the individual and total occurrences. The cumulative probabilities are printed in the first column. As with the selection of the starting probabilities, a random number between 0 and 1 is generated. Excels' in-built function *VLookup* can then be applied to the table and one of the sea states is selected. The variable *ThisState* is then updated to store the ID of the selected state. How this sea state is used in discussed in section 4.4.4.

Monthly Transition 3

If both *Month_Trans_1* and *Month_Trans_2* are *False* then the procedure will move onto monthly transition option 3. This follows the exact same methodology as the selection of the very first sea state of the modelled time series (section 4.4.2):

- The function *sort_start_table* prints a table to the 'temp' worksheet containing all the sea states and associated starting probabilities (cumulative) for that month
- A random number between 0 and 1 is generated
- The in-built function *Vlookup* is used to select one of the states (stored as *ThisState*)

4.4.4 Model Every Interval

Whilst within the monthly loop (which is nested within the yearly loop), an additional loop is required in order to model the sea state at every interval of the month under consideration. Before this step, it is necessary to use the *num_month_days* function to identify the number of days within each month, and therefore calculate the total number of intervals. It should be noted that the first month ('Dec') in year 1 is a special case as the printing begins at the second interval (the starting sea state will have already been printed).

Every row (i.e. interval) within that month is then considered, beginning at *Month_start_row* (a variable which will be updated at the end of the monthly loop). The variable *ThisState* stores the ID of the previous interval's sea state, as discussed in the monthly transition section previously. By using the function *get_this_states_row*, the position of *ThisState* within the monthly state worksheet is obtained, thus providing access to all the necessary information with which to select a new sea state. The number of possible states for the next interval is identified using Excel's in-built function. The function *sort_each_state_table* (similar to *sort_start_table*) sets up a table in the 'temp' worksheet containing the IDs and probabilities of each of the possible new states in the necessary format. Once again, a random number is generated so that the in-built function *VLookup* can select one of the new states probabilistically. The ID of this new sea state becomes the new *ThisState*. The associated values of wind speed, significant wave height and wave period are then printed to the modelled year worksheet via the function *print_new_state*.

After the sea state for every interval of that month has been modelled and printed, the variable *Month_start_row* is updated to ensure printing continues correctly for the remaining months of that year. At this stage, the monthly and yearly loops both end, resulting in one worksheet of modelled data for every year requested by the user.

4.4.5 Create Full Dataset

The final step in generating the time series is to compile all the modelled yearly datasets onto one worksheet. This is achieved with the function *create_full_modelled_dataset*. A worksheet named 'Modelled_Full' is created to store the data. A loop then runs for every year and the number of rows is calculated using the *num_rows* function. The full dataset in the year worksheet is then copied. For year 1 this includes the headers. A variable *paste_row* identifies the correct row in the 'Modelled_Full' worksheet and the copied range is pasted in. This ensures that there is a fully modelled dataset flowing smoothly from one year to the next.

5 VALIDATION

It is vital that the previously described Markov Chain Model (MCM) is validated before the generated time series' are used as inputs to the Operations and Maintenance (O&M) simulation tool. The method has been validated using hindcast data for the Farr Point wave energy site (available on request from Wave Energy Scotland) as described by Gray, Johanning & Dickens (2015). A series of validation tests have been performed in order to ensure that the generated time series' have similar statistical parameters as the hindcast dataset, as well as showing the same seasonal trends. The case study of the Pelamis P2 wave energy converter (WEC) has been used for this validation process (Gray 2017).

5.1 ORIGINAL DATASET

The MCM assessed in this study has been developed using hindcast data for Farr Point, a possible site for a wave energy array located off the North coast of Scotland. This original dataset is for an 18 year period from 1/1/1992 through to 31/8/2010 with 3 hourly intervals; providing a total of 54504 data points. The three parameters chosen to represent the weather conditions as the site are significant wave height (Hs), wave energy period (Te, in order to align with the P2 power matrix) and wind speed (U, for defining weather windows).

As previously described, the hindcast values must be placed into 'bins' to enable the probabilistic Markov Chain Model to function correctly. In this case study, these resolutions were chosen to be compatible with the power matrix of the P2 WEC, and were also based upon how engineers make decisions about marine operations in real life. The range and resolution of each parameter is as follows:

- Significant wave height (Hs) ranges from 0.25m to 9.75m, in steps of 0.5m.
- Wave energy period (Te) ranges from 3s to 15s, in steps of 2s.
- Wind speed (U) ranges from 2.5kts to 47.5kts, in steps of 5kts.

The modified hindcast dataset (i.e. once values are rounded and placed in bins) does not statistically differ from the observed collection of values, as shown for significant wave height in Figure 5.1, Figure 5.2 and Figure 5.3. The values for wave energy period and wind speed follow this pattern also (see Appendix B and Appendix C). These graphs show the mean, maximum and minimum values of the corresponding time steps across all years in the hindcast dataset (18 years).

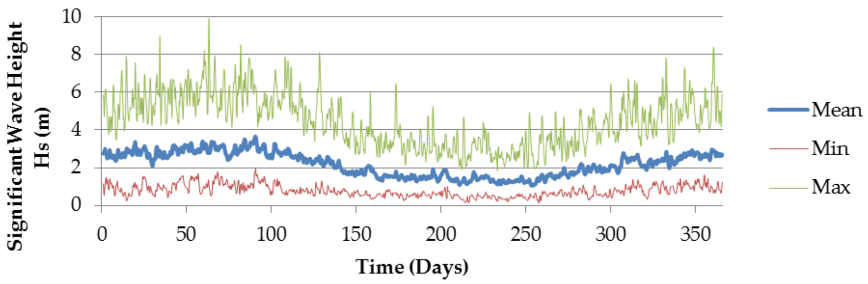

Figure 5.1. Observed original statistical Hs values

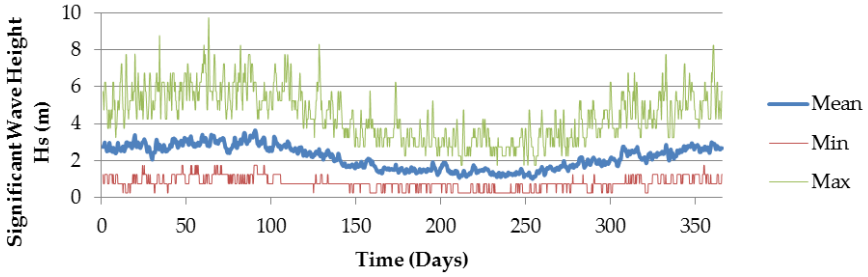

Figure 5.2. Modified original statistical Hs values

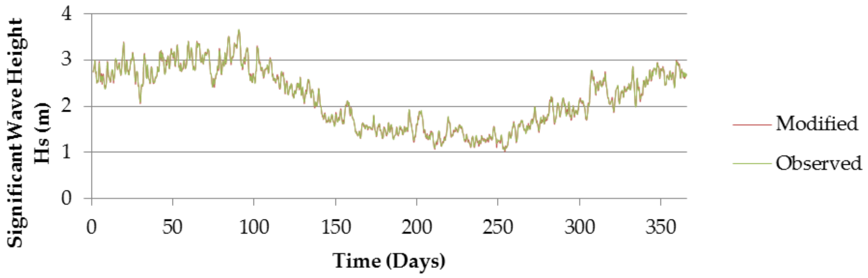

Figure 5.3. Mean Hs comparison between modified and observed original datasets

5.2 MODELLED DATASET

A 100 year modelled dataset has been generated by the MCM for this validation case study. When used by the O&M tool, it is unlikely that this many years will be required. A more suitable time scale would be 15 to 20 years, the typical design lifetime of a wave energy array. However, a 100 year time series provides an extensive dataset with which to confidently assess all statistical parameters of the MCM. An initial validation step was to compare the modelled average values for all three parameters (Hs, Te and U). Figure 5.4 shows the mean, minimum and maximum significant wave heights of the corresponding intervals across all 100 years of the modelled dataset.

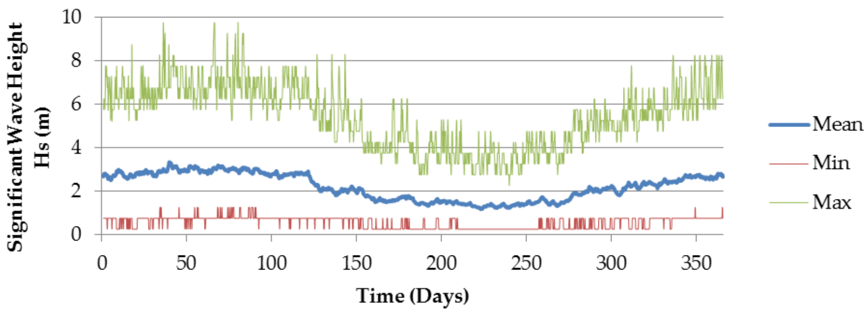

Figure 5.4. Modelled statistical Hs values

When these mean Hs values are plotted against the observed and modified original datasets, it can be seen that the modelled time series is different enough to provide variance, yet clearly follows the same seasonal trends (see Figure 5.5). This is also true for wave energy period and wind speed (see Appendix B and Appendix C).

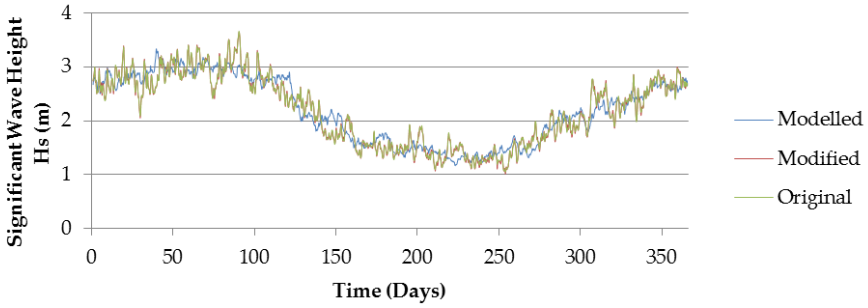

Figure 5.5. Mean Hs comparison between observed original, modified original and modelled datasets

From Figure 5.5, it can be seen that the mean values of the modelled time series follow the same monthly trends as the hindcast dataset. There is less variance in this modelled time series due to the length being 100 years (rather than 18 years for the hindcast dataset). This initial analysis shows that the Markov Chain Model is successful at modelling a time series which has similar statistical parameters to the original hindcast dataset, whilst offering a degree of variation.

5.3 PARAMETER CORRELATION

It is important that the relationships between each of the three parameters (Hs, Te and U) are successfully replicated in order to show that the MCM can deal with multiple variants. Correlations have been analysed to assess the ability of the MCM to achieve this. This method has been used to validate a Markov-based model for use in offshore wind farm O&M simulations (Scheu, Matha & Muskulus, 2012). Significant wave height and wind speed (U given at 10m height above sea level, denoted as u10) is the most obvious relationship, as higher wind speeds tend to lead to greater wave heights. Therefore, the relationship is assumed to be approximately linear. Figure 5.6, Figure 5.7 and Figure 5.8 illustrate this relationship graphically. It is important to consider the modified original dataset here, as well as the observed values, in order to fully assess the capability of the MCM. The correlation has been quantified using Pearson's correlation coefficient (R), which is effectively a ratio. This means that the difference between the R values can be expressed as a percentage, as shown in Table 5.1.

Table 5.1. Pearson's correlation coefficients for all three datasets for Hs vs u10

	Original Observed	Modified Observed	Modelled
R	0.639	0.623	0.621
% Difference from Original	-	2.54%	2.86%
% Difference from Modified	-	-	0.32%

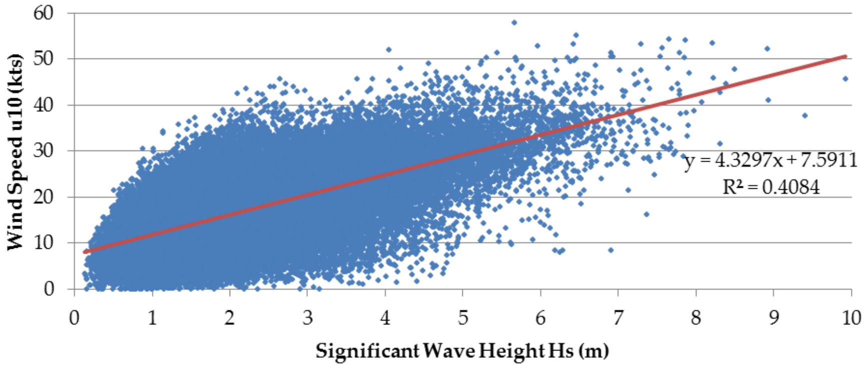

Figure 5.6. Wave height and wind speed correlation for the original observed dataset

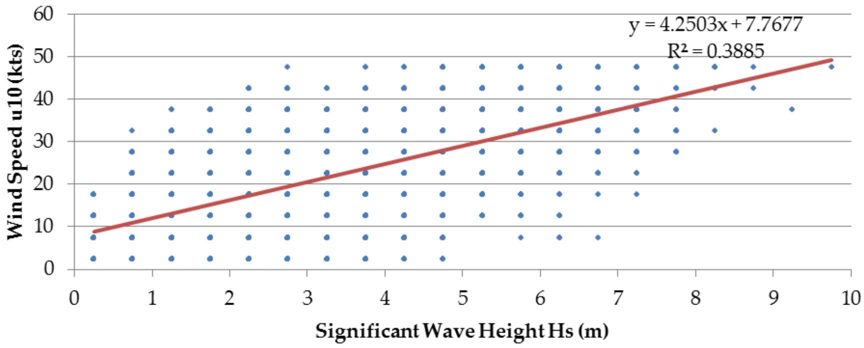

Figure 5.7. Wave height and wind speed correlation for the modified observed dataset

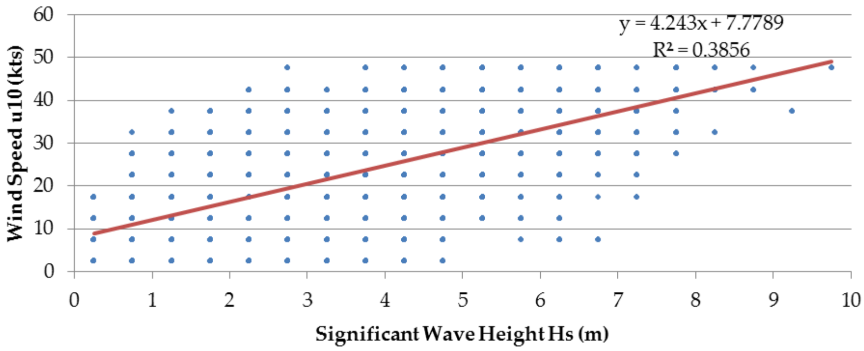

Figure 5.8. Wave height and wind speed correlation for the modelled observed dataset

It is clear that binning the original values has some effect on the correlation. It is expected that this stems from the resolution of the bins, as well as from the method of pulling values that lie outside the relevant constraints into the nearest bin, rather than being ignored. Yet, the percentage difference in the R value seen by rounding is approximately 2.5%. This correlation could be improved if the bin resolution was increased (i.e. resulting in a greater number of bins), however, it is deemed acceptable due to fact that a lower resolution provides more data points with which to undertake the probabilistic Markov method. The modelled dataset clearly shows a similar wind and wave correlation to the original values, with less than 0.5% difference from the modified data. A similar pattern was found by assessing the correlation between significant wave height and wave energy period (see Appendix D).

5.4 PERSISTENCE OF WEATHER

The primary reason for building such a detailed and extensive weather simulation model is to represent realistic access windows. This is a hugely important consideration for O&M. The length of time a weather window remains closed for is determined by the persistence of the conditions (i.e. the amount of time the weather conditions exceed given limits). Seasonal variability is the best way to analyse this. It is vital to check that the persistence of weather conditions in the synthetic time series does not differ significantly from the original dataset. Figure 5.9, Figure 5.10 and Figure 5.11 show that there is little difference when considering the cumulative distribution functions (CDFs) of varying significant wave heights during winter. The results are similar for Spring, Summer and Autumn (see Appendix E). Note: It has been proven that the binning process does not affect the original dataset significantly. As a result, all 'original' values now refer to the modified dataset.

A useful way of interpreting these persistence graphs is that they identify the probability of having to wait less than certain number of days for a given weather window. For example, it can be deduced from Figure 5.9 that there is approximately a 40% probability of having to wait less than 5 days for a weather window of 1.5m Hs during the winter months.

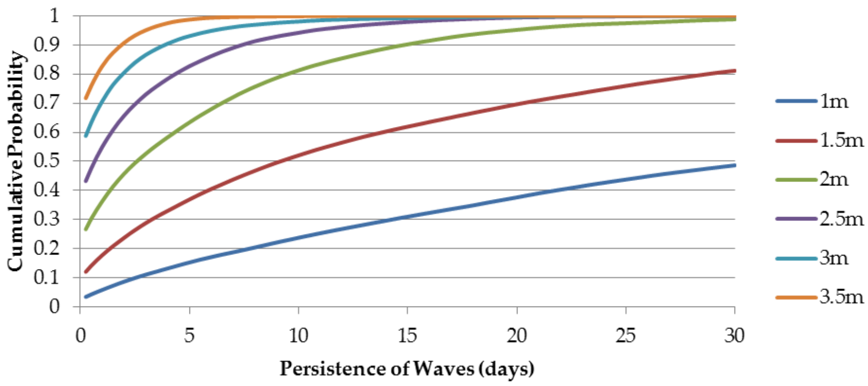

Figure 5.9. Winter (Dec-Feb) persistence of Hs for the original dataset

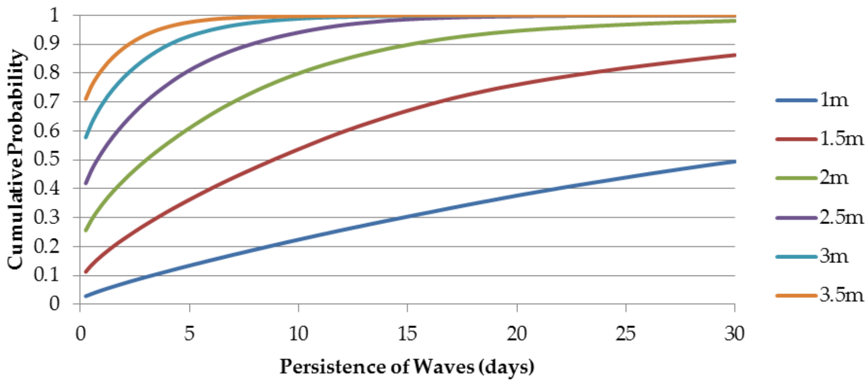

Figure 5.10. Winter (Dec-Feb) persistence of Hs for the modelled dataset

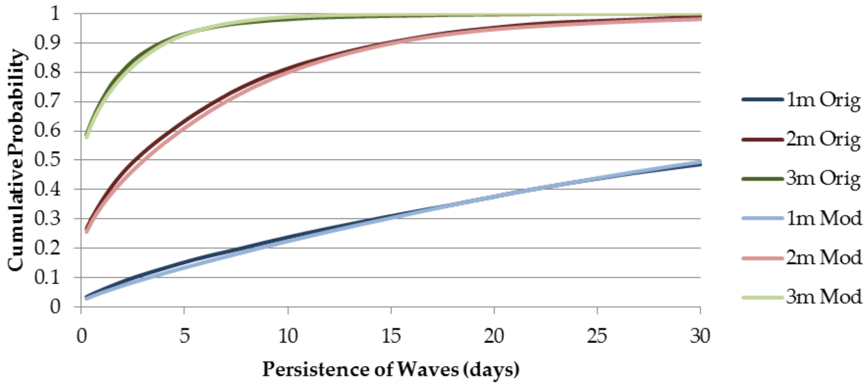

Figure 5.11. Winter (Dec-Feb) persistence of Hs comparison for original and modelled datasets

However, this validation step can go much further. The multivariate nature of the MCM allows persistence to be defined in greater detail. A number of constraints were identified during the P2 testing programme undertaken between 2010 and 2014. Firstly, no marine operations are carried out when the wind speed is greater than 20 knots. Also, a removal can be carried out in rougher seas than an install. In addition, the maximum significant wave height allowed for marine operations depends on the wave energy period. These constraints are shown in Figure 5.12. A weather window is open if the sea state is below the relevant line in Figure 5.12, and the wind speed is below 20 knots, for a given length of time.

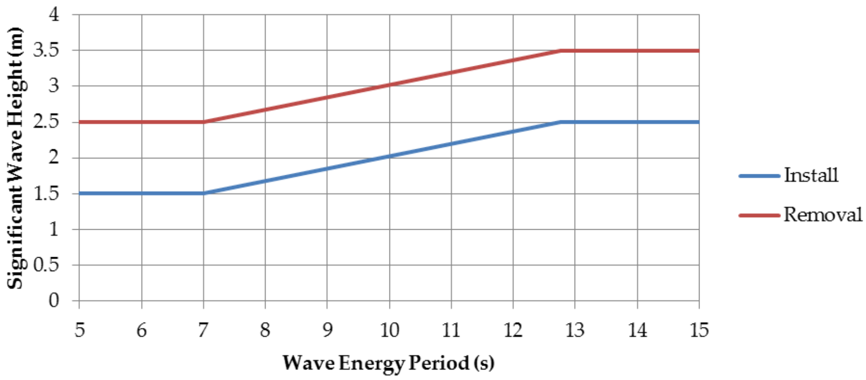

Figure 5.12. Maximum wave conditions allowed for a P2 install and removal

The CDFs created when using these operational limits to define persistence also show that there is little difference between the original and modelled time series' (see Figure 5.13 for winter, see Appendix F for remaining months). As assessing weather windows for accessibility is a key aspect of the O&M tool, this close correspondence is a vital step in validating the Markov Chain Method.

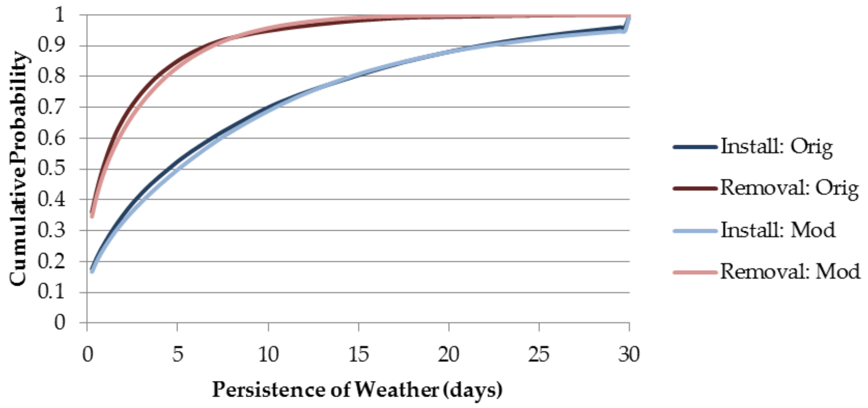

Figure 5.13. Winter (Dec-Feb) persistence of non-accessible weather conditions. Comparison of original and modelled datasets for install and removal constraints

This information has been quantified using seasonal mean wait times (i.e. the average time spent waiting for an open weather window) with 95% confidence intervals applied (Figure 5.14). From this validation step, it can be said that the statistical metrics in the modelled time series are not significantly different from the original dataset in terms of accessibility, with 95% confidence. The percentages of open weather windows also show very little difference between the two datasets (see Appendix G).

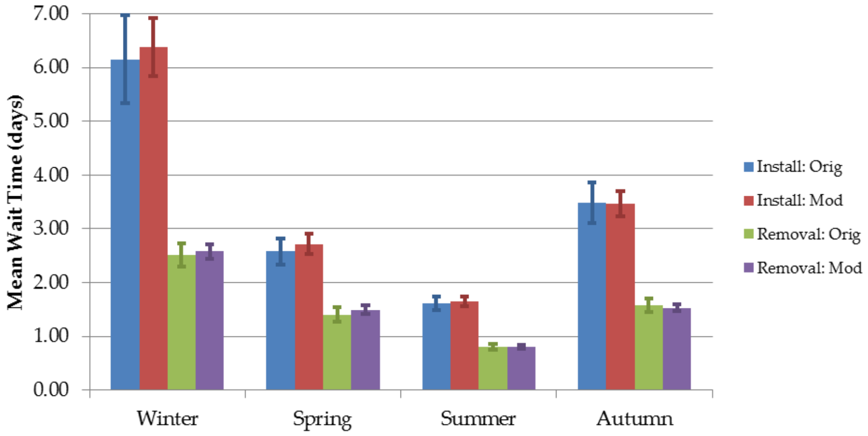

Figure 5.14. Mean time to wait for a weather window for each season, using install & removal weather constraints. Comparison of original and modelled datasets with 95% confidence intervals applied

5.5 POWER CAPTURE

The other key reason for developing a detailed weather model is to gain more realistic estimations of power generation for the wave farm. The binned values of significant wave height and wave energy period can be compared to the values in the P2 power matrix (more specifically, the O&M contract agreed target table, Gray 2017). Figure 5.15 compares the average power output of the original and 100 year modelled datasets, with 95% confidence intervals shown. As with the weather persistence analysis, the modified dataset is used here to represent the 'original' values.

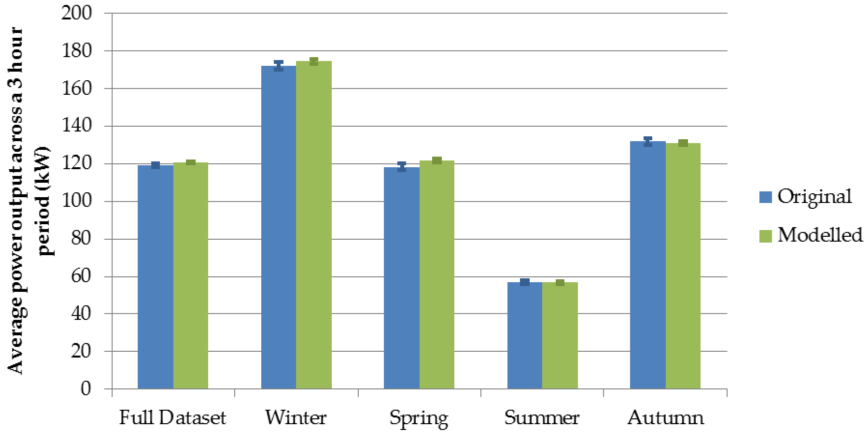

Figure 5.15. Average power output over a 3 hour period for the full dataset, and for each season. Comparison of original and modelled datasets with 95% confidence intervals applied

Figure 5.15 shows that there is seasonal consistency in terms of estimated power output between the original and modelled datasets. However, the 95% confidence intervals do not overlap for the full dataset average, nor do they overlap for the season of spring. It is expected that this anomaly stems from the power matrix used. It was not possible to interpolate all the values from the power target table. This includes situations where the power had to be assumed to be zero, even though there would clearly be some power output in reality. Although these instances would be very rare, they may have accounted for the slight discrepancies seen here. Nevertheless, the average power outputs estimated from the modelled dataset are realistic and show the expected seasonal variability.

6 CONCLUSIONS

6.1 DISCUSSION

The Markov Chain Model (MCM) described in section 3 is used to generate a time series of weather conditions from a hindcast dataset which shows the same seasonal variability, yet produces a unique time series. The user needs to obtain a hindcast dataset for a given wave energy site with at least 10 years of data of an appropriate resolution (e.g. 1, 3 or 6 hours). The method involves binning the hindcast data in a way that the values match up with a power matrix for the WEC, whilst corresponding to accessibility limits for marine operations. The VBA code described in section 4 needs to be modified accordingly to ensure the MCM functions as designed. Once the code has been set up for the specific hindcast dataset used as the input, the user can select the required number of years to generate in the modelled time series. The synthetic time series' created by this method must be stored in an Excel workbook with an appropriate name in order to be recognised by the O&M tool, as described in the 'Functionality Report' (WES 2017a).

The Markov Chain Model (MCM) undergone a rigorous validation process using the case study of a Pelamis P2 device located at Farr Point, off the North coast of Scotland. Significant wave height, wave energy period and wind speed are all generated at 3 hourly intervals. A 100 year modelled time series, produced using an 18 year hindcast dataset for the Farr Point site, has been utilised for the validation process described in section 5.

It has been shown that the modelled time series successfully replicates the seasonal variability of the original hindcast dataset. It was found that the method of treating multiple parameters was suitable, as there was little difference in variable correlation between the two datasets. Realistic representation of weather windows and power capture is of vital importance to the O&M simulation tool. It was proven that the statistical metrics of the 100 year modelled time series were not significantly different from the original dataset in terms of accessibility and estimated power capture. Whilst showing the expected trends throughout each year, the modelled time series differed from the hindcast dataset sufficiently to justify utilising such a method.

In conclusion, the validation phase has shown that the Markov Chain Model is suitable for use in the O&M tool described in the 'Functionality Report' (WES 2017a). However, the method does have some limitations which may need to be addressed in future work.

6.2 LIMITATIONS OF THE METHOD

At present, the MCM only considers significant wave height, wind speed and wave period. This means that the effects of tidal velocity and elevation are not taken into account, even though these may be limiting factors for marine operations, particularly in terms of port access.

The Markov Chain method of generating time series' from a hindcast dataset means that extreme events are only included in the O&M tool if such events occurred in the hindcast dataset. This method also has an anomaly where an unnatural 'jump' between two months can occur if a suitable sea state cannot be found in the next month's dataset. Although this anomaly rarely occurs (~1% of the monthly transitions, see Appendix A), an alternative method could be to find the next 'closest' sea state, ideally in terms of significant wave height due to its importance in defining weather windows.

Even if a hindcast dataset with more than 10 years' worth of values is used, there is no guarantee that storm events will have been recorded. As such, these events will not appear in any of the generated time series'.

In general, suitable hindcast datasets of weather conditions are obtained with resolutions of 3 or 6 hours. This means that the O&M tool is limited to this resolution, meaning the estimates for power generation are averaged across these periods. Better estimations of power, as well as more realistic representations of weather windows, could be achieved if the model resolution was increased. One factor limiting this development at present is the quality of office computer processors. This constraint will become less of an issue as the technology advances.

7 REFERENCES

EQUIMAR. (2011) Protocols for the Equitable Assessment of Marine Energy Converters. The Institute of Energy Systems, The University of Edinburgh. Edited by: Ingram, D., Smith., G., Bittencourt-Ferreira, C. & Smith, H.

GRAY, A. M. (2017) *Modelling Operations and Maintenance Strategies for Wave Energy Arrays*. Thesis (EngD). College of Engineering, Mathematics and Physical Sciences, University of Exeter.

Gray, A., Johanning, L. & Dickens, B. (2015). *A Markov Chain Model to Enhance the Weather Simulation Capabilities of an Operations and Maintenance Tool for a Wave Energy Array*. IN: 11th European Wave and Tidal Energy Conference. 6-11 September 2015. Nantes, France.

HAGEN, B., SIMONSEN, I., HOFMANN, M & MUSKULUS, M. (2013) *A multivariate Markov Weather Model for O&M Simulation of Offshore Wind Parks*. IN: 10th Deep Sea Offshore Wind R&D Conference, DeepWind'2013. 24-25 January 2013. Trondheim, Norway. Energy Procedia, 35, pp137-147.

SCHEU, M., MATHA, D., HOFMANN, M & MUSKULUS, M. (2012) *Maintenance strategies for large offshore wind farms*. IN: 9th Deep Sea Offshore Wind R&D Seminar, DeepWind'2012. 19-20 January 2012. Trondheim, Norway. Energy Procedia, 24, pp281-288.

SCHEU, M., MATHA, D. & MUSKULUS, M. (2012) *Validation of a Markov-based Weather Model for Simulation of O&M for Offshore Wind Farms*. IN: Proceedings of the Twenty-second (2012) International Offshore and Polar Engineering Conference. 17-22 June 2012. Rhodes, Greece.

WAVE ENERGY SCOTLAND (WES). (2017a) *Operations and Maintenance Simulation Tool - Functionality Report.*

WAVE ENERGY SCOTLAND (WES). (2017b) *Operations and Maintenance Simulation Tool - User Guide.*

WAVE ENERGY SCOTLAND (WES). (2017c) *Operations and Maintenance Simulation Tool - Future Upgrades.*

8 APPENDICES

8.1 Appendix A

The figure below shows the percentage occurrence of each monthly transition option in a nominal 200 year modelled dataset, generated using hindcast data for the Farr Point wave energy site.

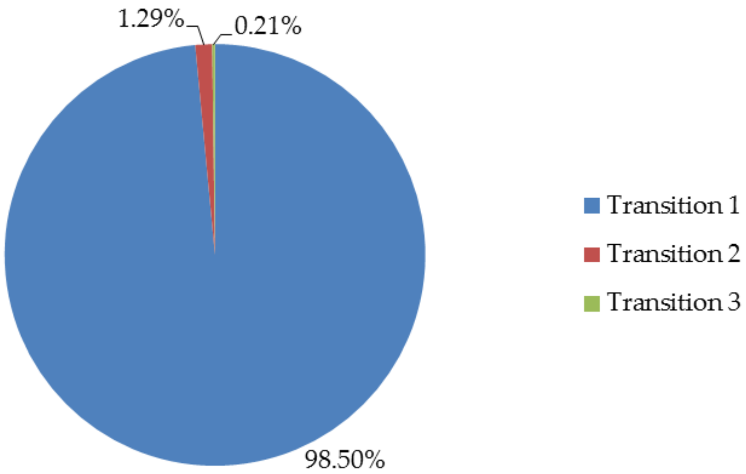

Figure 8.1. Percentage of monthly transitions

8.2 Appendix B

Average wave energy period (Te) comparisons of hindcast, modified and modelled datasets.

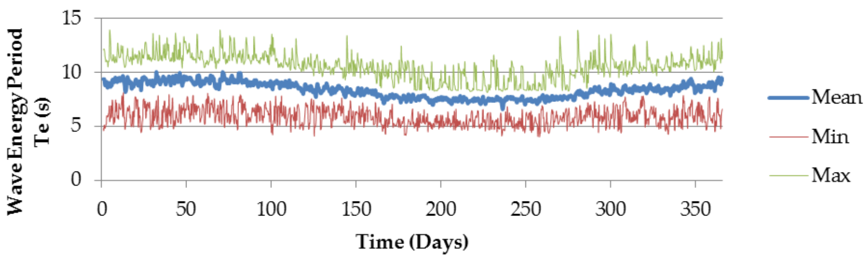

Figure 8.2. Observed original statistical Te values

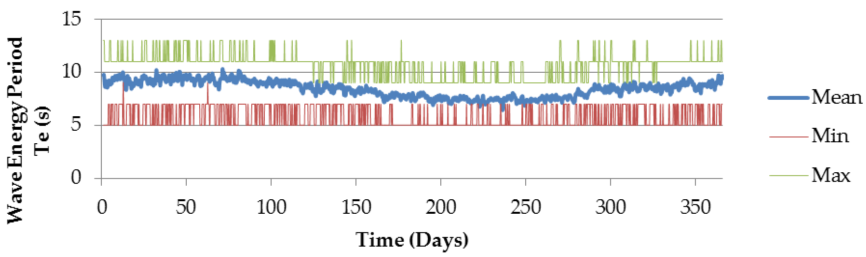

Figure 8.3. Modified original statistical Te values

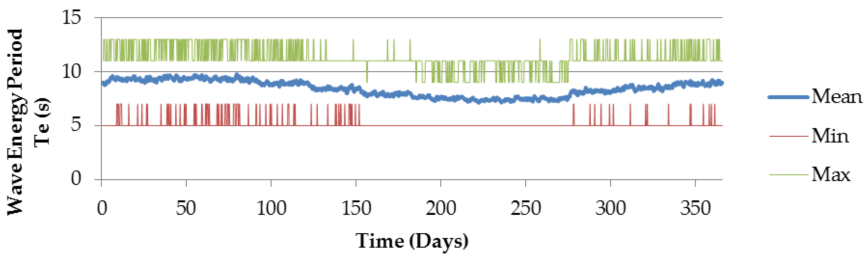

Figure 8.4. Modelled statistical Te values

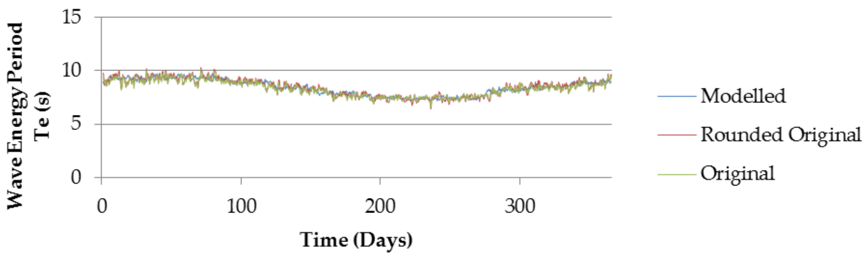

Figure 8.5. Mean Te comparison between modified and observed original datasets

8.3 Appendix C

Average wind speed (U, u10 at 10m) comparisons of hindcast, modified and modelled datasets.

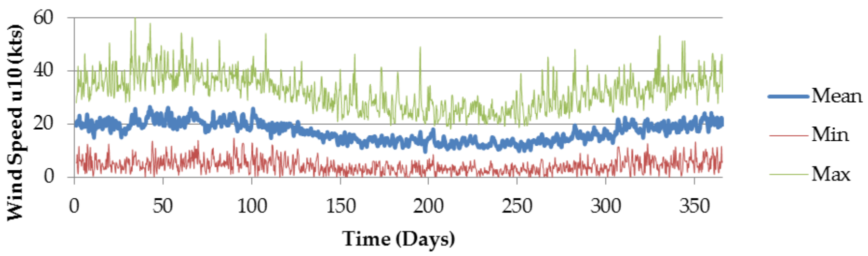

Figure 8.6. Observed original statistical u10 values

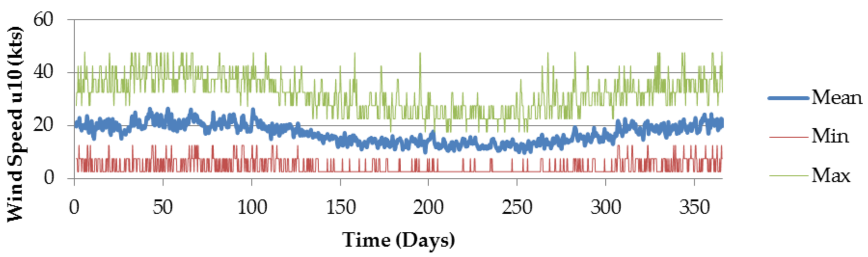

Figure 8.7. Modified original statistical u10 values

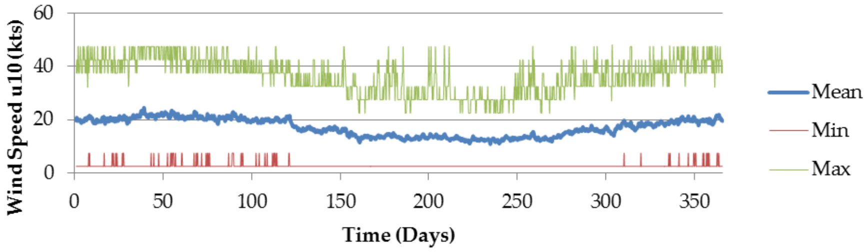

Figure 8.8. Modelled statistical u10 values

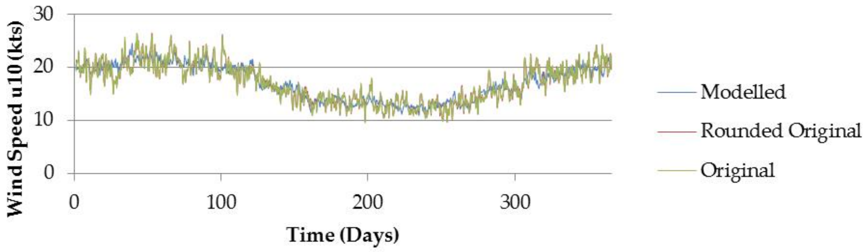

Figure 8.9. Mean u10 comparison between modified and observed original datasets

8.4 Appendix D

Significant wave height (Hs) vs wave energy period (Te) correlation.

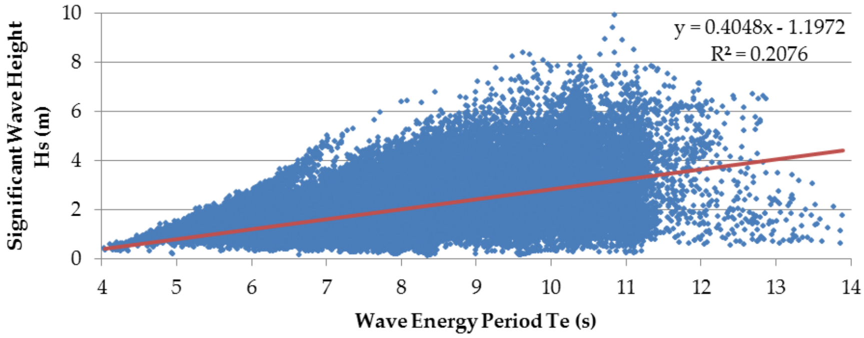

Figure 8.10. Wave height and period correlation for the original observed dataset

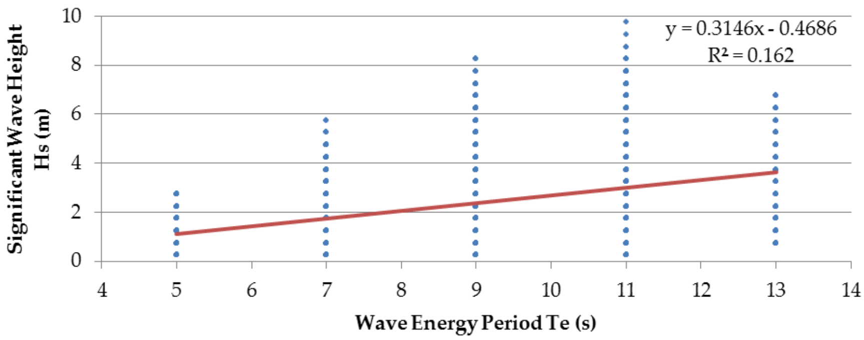

Figure 8.11. Wave height and period correlation for the modified original dataset

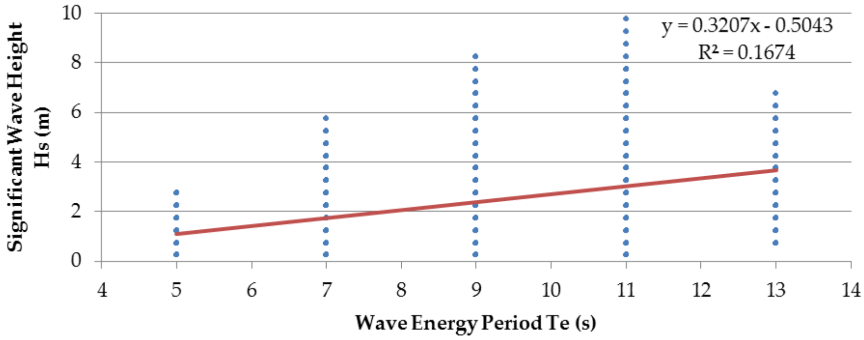

Figure 8.12. Wave height and period correlation for the modelled dataset

Table 8.1. Pearson's correlation coefficients for all three datasets for Hs vs Te

	Original Observed	**Modified Observed**	**Modelled**
R	0.456	0.402	0.409
% Difference from Original	-	12.38%	10.75%
% Difference from Modified	-	-	1.64%

8.5 Appendix E

Persistence of Hs for Spring, Summer and Autumn.

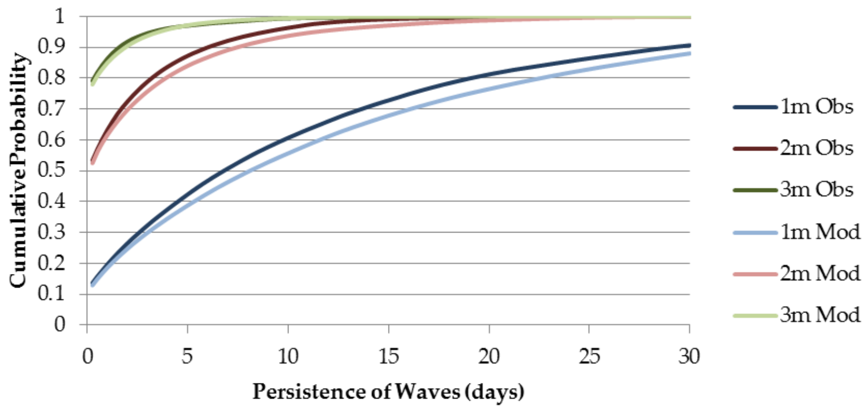

Figure 8.13. Spring (Mar-May) persistence of Hs comparison for original and modelled datasets

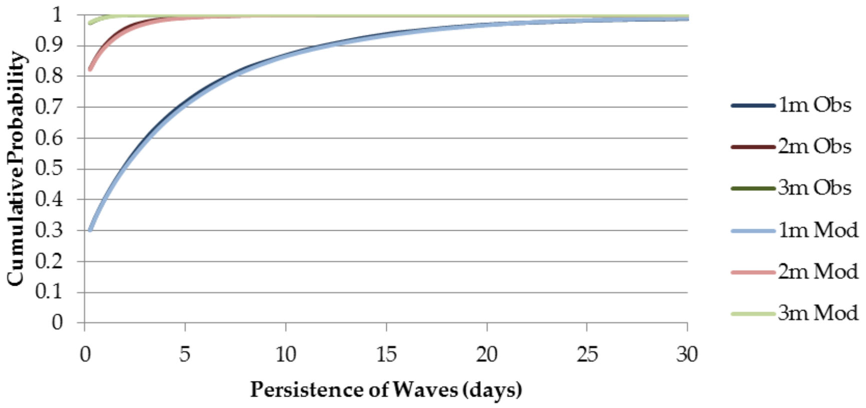

Figure 8.14. Summer (Jun-Aug) persistence of Hs comparison for original and modelled datasets

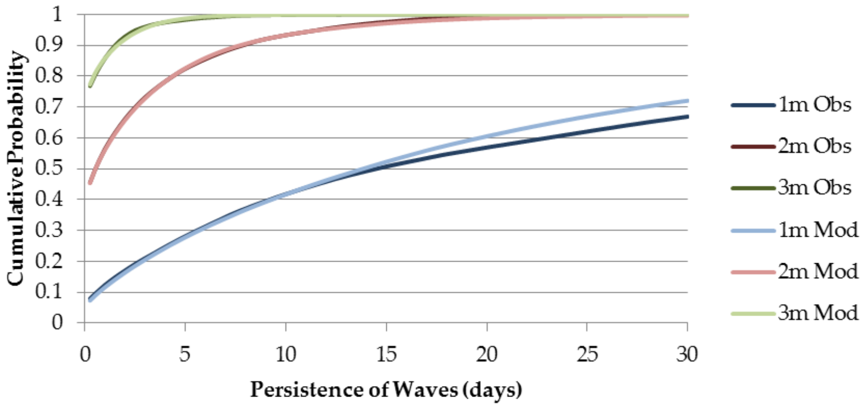

Figure 8.15. Autumn (Sep-Nov) persistence of Hs comparison for original and modelled datasets

8.6 Appendix F

Persistence of install & removal weather limits for Spring, Summer and Autumn

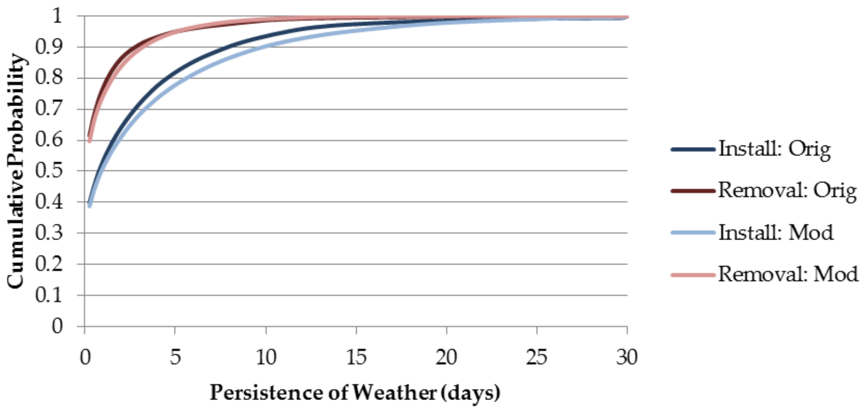

Figure 8.16. Spring (Mar-May) persistence of non-accessible weather conditions. Comparison of original and modelled datasets for install and removal constraints

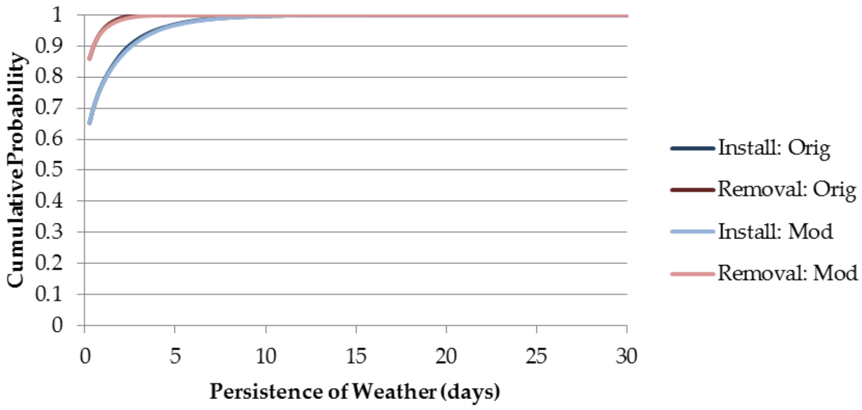

Figure 8.17. Summer (Jun-Aug) persistence of non-accessible weather conditions. Comparison of original and modelled datasets for install and removal constraints

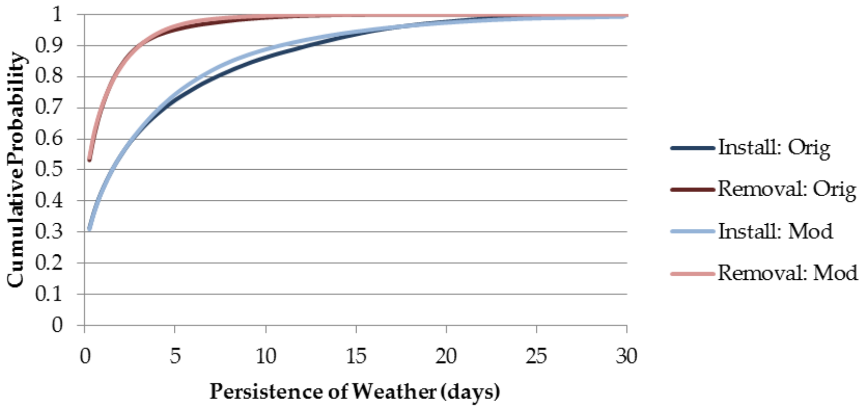

Figure 8.18. Autumn (Sep-Nov) persistence of non-accessible weather conditions. Comparison of original and modelled datasets for install and removal constraints

8.7 Appendix G

Table 8.2. Percentage of open weather windows using install & removal weather limits

	Percentage of open weather windows				*Difference (%)*	
	Observed		Modelled			
	Install	Removal	Install	Removal	*Install*	*Removal*
Full dataset	38.8	59.5	38.0	58.7	*-0.8*	*-0.8*
Winter	17.6	36.1	16.7	34.7	*-0.9*	*-1.4*
Spring	39.8	61.5	38.7	59.7	*-1.1*	*-1.8*
Summer	65.2	86.0	65.3	86.1	*0.1*	*0.1*
Autumn	31.3	53.1	30.9	54.0	*-0.4*	*0.8*

8.8 List of Figures

8.9 List of Tables