A Markov Chain Model to Enhance the Weather Simulation Capabilities of an Operations and Maintenance Tool for a Wave Energy Array

Anthony Gray¹, Lars Johanning², Beth Dickens³

¹Industrial Doctoral Centre for Offshore Renewable Energy, Edinburgh, UK
Author for correspondence (anthony.gray@ed.ac.uk)

²College of Engineering, Mathematics and Physical Sciences, University of Exeter, Penryn, UK

³Pelamis Wave Power, Edinburgh, UK

Abstract— Operations and maintenance is a vital area of research in the push to make wave energy a commercial reality. A tool has previously been developed by Pelamis Wave Power to obtain reliable estimates for operational expenditure and ensure smooth running of wave energy arrays. Wave Energy Scotland is now tasked with the future development of this operations and maintenance tool. One of its key inputs is the wave and wind data used to simulate weather windows suitable for marine access. This paper details the creation and validation of a Markov Chain Model to enhance the weather simulation capabilities of the tool. This will ensure that the operations and maintenance strategy of wave energy arrays is modelled more realistically, resulting in an increased confidence in cost estimates and logistical arrangements.

Keywords— O&M Modelling, Markov Chains, Pelamis Wave Power, Wave Energy Scotland, Weather Windows

I. INTRODUCTION

Operational expenditure (Opex) is a significant cost of any offshore development. This cost needs to be minimised for a wave energy array to become commercially viable. Opex also needs to be modelled and analysed in order to obtain realistic estimates for levelised cost of energy. A research partnership between the Industrial Doctoral Centre for Offshore Renewable Energy (IDCORE) and Wave Energy Scotland (WES) seeks to address this challenge. The work had previously been undertaken in cooperation with Pelamis Wave Power (PWP), the company behind the world’s first commercial scale wave energy converter to generate electricity to a national grid. The operations and maintenance (O&M) strategy of the Pelamis technology is focused on rapid installation and removal of devices, using multicat vessels to bring them into the safety of a sheltered harbour for repair and inspection.

The research partnership centres on the review and upgrade of an O&M tool, originally created by Pelamis Wave Power in 2007. The Monte Carlo-based tool uses reliability data to simulate the occurrence of faults during machine operation [1]. This ‘reactive’ maintenance approach is accompanied by a ‘proactive’ routine service on each machine, resulting in a complete O&M strategy. The work will result in fully optimised O&M strategies for a series of different wave energy sites, ensuring smooth operation and maximising revenue.

Fig. 1 The use of inexpensive, readily available, multipurpose workboats is fundamental to Pelamis Wave Power’s O&M strategy.

Aside from reliability statistics, the other key input to the O&M tool is weather data. Weather windows (periods of accessibility) were previously calculated based solely on the probability of exceeding a certain significant wave height. The reality, however, is much more complex. Other parameters, wave period and wind speed in particular; play a part in defining the weather windows suitable for marine operations.

To address this complexity, a multivariate Markov Chain Model (MCM) has been developed to simulate realistic weather conditions for use in the O&M tool. It can also be used to obtain better estimates of power capture, and thus provide more realistic estimations of revenue. A similar methodology has been used for O&M tools in the offshore wind industry [2],[3].

This paper is split into two key sections. The first looks at the logic and methodology of the Markov Chain Model. The second discusses how the model has been validated to ensure maximum confidence in the output.

II. METHODOLOGY

A. Parameter Selection

Three variables are vital in defining a weather window suitable for the installation or removal of a Pelamis device:
- Significant Wave Height, Hs
- Wave Energy Period, Te
- Wind Speed, U10 (i.e. at 10m above sea level)

The limitations on installation and removal operations have been assessed over the course of thorough testing of two
750kW P2 devices, operated by Pelamis Wave Power from 2008 to 2014. The quick-release latch mechanism on the P2 machines allows marine operations to be carried out in rougher seas than earlier devices. The limitations can now be primarily attributed to the boat specifications. Further testing, under the remit of Wave Energy Scotland, will lead to an increased confidence in installation and removal techniques, resulting in the current P2 weather window constraints being expanded. For the development of the Markov Chain Model detailed in this paper, the current P2 operations limits will be used:

- No marine operations can be carried out in wind speeds of over 20 knots.
- Installation operations are limited to 1.5m significant wave height, though this can rise to 2.5m depending on the corresponding wave energy period.
- Removal operations are limited to 2.5m significant wave height, though again this can rise to 3.5m depending on the wave energy period.

The Hs-Te relationship described here, and shown graphically in figure 2, has been developed using the vast experience of Pelamis vessel engineers.

![Fig. 2 Current installation and removal limits for the P2 devices.](image)

Tidal currents are not included in the weather window simulations at this stage. The effect that tidal currents have on marine operations very much depends on wave site location in relation to the O&M base. As such, it may be useful to include tidal currents in future versions of the weather simulation model. However, this was not undertaken for the validation study described in this paper due to the requirement for flexibility in the Markov Chain Model.

**B. Input Data**

The Markov Chain Model (MCM) has been created and validated using hindcast data for Farr Point, one of the sites previously under development by Pelamis Wave Power, located off the north coast of Scotland. This dataset is over an 18 year period, from 1/1/1992 to 31/8/2010, with 3 hour intervals. Future studies will consider other sites such as the European Marine Energy Centre (EMEC) in Orkney and Wave Hub off the Cornish coast.

**C. Model Resolution and Data Expansion**

The Markov Chain Model uses a probabilistic approach, whereby step changes are calculated based upon the occurrence of ‘sea states’ in the original dataset. It is therefore important that the input data is grouped and expanded where possible to maximise the capability of the model. In previous versions of the O&M tool a weather window was defined as a 24 hour period where the significant wave height remains below a certain level. In addition to introducing the two new parameters into this definition, the resolution of the O&M tool has been changed to 6 hours. In other words, a weather window is now defined as a 6 hour period where the wave and wind conditions lie within the relevant constraints. By increasing the resolution of this model in this way, the changeable nature of weather conditions are better represented. It also allows the inclusion of more detailed permutations of boat operations (i.e. the number of installations and removals that are logistically possible over a certain period of time).

A large dataset collected over at least 20 years is preferable. As a result, a method of expanding the 18 year hindcast data (with 3 hour intervals) into a 36 year dataset with 6 hour intervals was developed. This involved obtaining 6 hourly averages at staggered intervals and using the alternate averages in newly created years. This method has been tested, showing that the statistical metrics of the expanded dataset, particularly weather wait times, are extremely similar to the hindcast data. The 36 year expanded dataset will be referred to as the ‘original dataset’ for the remainder of this paper.

**D. Sea States**

Due to the probabilistic nature of the Markov Chain process, it is necessary to group the original data values of significant wave height (Hs), wave energy period (Te) and wind speed (U10) into bins. These bins will be used to identify sea states containing all three parameters.

Wider ranging data bins (i.e. low resolution of values) will result in fewer of bins and therefore provide the MCM with more data points with which to calculate transition probabilities. This is because there will be a greater number of data points falling into each bin. Data points in this sense refer to the occurrence of a specific value in the original dataset. However, if the resolution of values is too low (i.e. very few bins) then the MCM loses the ability to produce analytical results. Conversely, if the resolution of the binned values is too high (i.e. greater number of bins) then the limited number of possible transitions makes the MCM less able to produce a realistic time series.

The three parameters (Hs, Te and U10) have been assigned resolutions based upon how experienced wave energy engineers make decisions about marine operations in real life. The resolutions have also been chosen to align with a power matrix used for analysis of the P2 devices, thus enabling power capture calculations to be carried out. A bin is represented by the midpoint value (i.e. 2.25m Hs denotes the bin $2m \leq Hs < 2.5m$).
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 (U10) ranges from 2.5kts to 47.5kts, in steps of 2s.

The minimum and maximum values here have been selected to be inclusive of at least 99% of possible values. In the rare cases where observations lie outside of these limits, they are placed in the closest bin.

These bins provide the basis for identifying sea states (combinations of all three parameters). It is necessary to have a system of assigning IDs to sea states that is consistent across all months as this enables monthly transitions to be easily carried out (this will be discussed in section F). 1400 different sea states have been assigned using the possible combinations of the binned values above. Whilst all three parameters in the model are vital, significant wave height is viewed as the most important when making decisions about marine operations. As a result, the sea state IDs are assigned in order of Hs first, followed nominally by U10 then Te. 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 will contain 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 6 hour period to the next.

**E. Monthly Data**

Following the grouping process described previously, the original dataset was broken up into months to account for seasonal variability. As a further means of expanding the dataset, the monthly data includes the last five days of the previous month and first five days of the next month. For each 6 hour interval, 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.

**F. Transitional Properties**

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

\[
p_{ij} = \frac{N_{ij}}{N_i}
\]

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

1) **Starting Probabilities:**

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

\[
p_i = \frac{N_i}{N}
\]

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 failsafe option for selecting monthly transitions.

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 the original equation. To account for this, a three tier hierarchical system of determining the sea state at the first 6 hour interval in the next month has been developed:

**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 6 hour interval in the next month is selected using the original equation.

**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’s dataset are deleted. If one or more states remain, then one is chosen to become an intermediate state. This is achieved using a modified version of the original equation:

\[
p_{1j2} = \frac{N_{i2j}}{n_{new_i}}
\]

Where:
- \(p_{1j2}\) = probability of selecting state \(j\) in next month from state \(i\) in previous month
- \(N_{i2j}\) = number of observed transitions from sea state \(i\) to state \(j\) in previous month
- \(n_{new_i}\) = 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. The original equation can then once again be used to determine the sea state at the first 6 hour interval in the next month. Monthly transition option 2 can therefore be thought of as skipping one sea state.

**Monthly transition 3**: In very rare situations, none of the possible next states from the previous month will exist in the next month’s dataset. In these cases, the starting probabilities described previously are used to select the new sea state. Although this may result in larger jumps (in terms of Hs for example) from one 6 hour 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.

---

**III. Validation**

**A. Data Expansion**

The first validation stage required is to confirm that the data expansion and ‘binning’ process described previously is acceptable for use. The selected method for this task is to compare the average values for all three parameters. The Farr Point hindcast data is referred to as the ‘observed data’, whilst the expanded and ‘binned’ data is labelled ‘modified’. Figure 4 shows that there is no significant statistical difference between the observed and modified values when analysing significant wave height. The same is true for wave energy period and wind speed, as shown in figures 6 and 7.

![Fig. 4 Mean Hs comparison between modified and observed datasets.](image)

**B. Modelled Dataset**

A 100 year modelled dataset has been generated for the remainder of the validation procedure. 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 design lifetime of a Pelamis wave farm. However, a 100 year time series provides an extensive dataset with which to confidently assess all statistical parameters of the MCM.

The initial validation step was extended to compare the modelled average values for all three parameters (Hs, Te, U10) against the modified and observed values. Figure 5 shows that the modelled time series for Hs differs enough from the original dataset to provide variance, yet clearly follows the same seasonal trends. Figures 6 and 7 demonstrate that this is also true for wave energy period and wind speed.

![Fig. 5 Mean Hs comparison between modelled, modified and observed datasets.](image)
C. Parameter Correlation

It is vital that the relationships between each of the three parameters are successfully replicated in order to show that the Markov Chain Model (MCM) can deal with multiple variants. Correlations have been analysed to assess the ability of the MCM to achieve this. This method has previously been used to validate a Markov-based model for use in offshore wind farm O&M simulations [4]. Significant wave height and wind speed 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. Figures 8-10 illustrate this relationship graphically. It is important to consider the ‘modified’ (i.e. expanded and ‘binned’) dataset here, as well as the ‘observed’ (Farr Point hindcast) values, in order to fully assess the capability of the MCM. The correlation has been quantified using Pearson’s correlation coefficient (R) (table 1). This means that the difference between the R values can be expressed as a percentage.

<table>
<thead>
<tr>
<th></th>
<th>Original</th>
<th>Modified</th>
<th>Modelled</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>R</strong></td>
<td>0.639</td>
<td>0.623</td>
<td>0.621</td>
</tr>
<tr>
<td>% Difference from Original</td>
<td>-</td>
<td>2.54%</td>
<td>2.86%</td>
</tr>
<tr>
<td>% Difference from Modified</td>
<td>-</td>
<td>-</td>
<td>0.32%</td>
</tr>
</tbody>
</table>

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%, which is acceptable. A better correlation could be obtained if the number of bins was increased, though this is unnecessary due to the benefits of selecting these resolutions (as described previously). 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.

D. Weather Waits

The primary reason for building such a detailed and extensive weather 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 weather conditions. 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 11 shows that there is little difference when considering the cumulative distribution functions (CDFs) of varying significant wave heights during the three winter months (December, January, February). The best way to interpret this graph is by thinking in terms of weather wait times. For example, from figure 11 it can be determined that there is approximately a 62% probability of having to wait less than 5 days for a 2m Hs weather window during the winter months.

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.

This validation step can go much further. The multivariate nature of the MCM allows persistence to be defined in greater detail. As stated previously, Pelamis vessel engineers have a number of constraints to consider before deciding to install or remove a machine from site. These conditions include that no marine operations are to be 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. The current constraints for the P2 device were shown in figure 1 (page 2). A weather window is deemed ‘open’ if the sea state’s wave height and period values lie below the relevant line in figure 1 (and if the wind speed is below 20 knots) for a given period of time (i.e. 6 hours in this analysis). It should be noted that these limits are expected to increase as installation and removal techniques are improved.

The persistence CDFs created when using these operational limits to identify wait times also show that there is little difference between the original and modelled time series’ (see figures 12-15)
This information can be quantified using seasonal mean wait times (i.e. the average time spent waiting for an open weather window) with 95% confidence intervals applied (figure 16). 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 (table 2).

<table>
<thead>
<tr>
<th>Percentage of open weather windows (%)</th>
<th>Difference (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full dataset</td>
<td></td>
</tr>
<tr>
<td>Winter</td>
<td></td>
</tr>
<tr>
<td>Spring</td>
<td></td>
</tr>
<tr>
<td>Summer</td>
<td></td>
</tr>
<tr>
<td>Autumn</td>
<td></td>
</tr>
</tbody>
</table>

E. 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). The power matrix has been modified slightly to ensure consistency with the binned data. Figure 17 compares the 6 hourly average power output of the original and 100 year modelled datasets, with 95% confidence intervals shown. As with the weather persistence analysis, the modified (i.e. expanded and ‘binned’) dataset is used here to represent the ‘original’ values.

IV. CONCLUSIONS

A Markov Chain Model has been used to generate a time series of synthetic weather data for use in the Pelamis-based O&M tool. This method was selected over other options described by Monbet, Alliot and Prevsto [5] due to its wide ranging and proven use. Markov models have been used for simulating weather sea states for some time [6],[7]. The model is required to provide more realistic data for assessing weather windows and power capture than is currently available in the O&M tool. Significant wave height, wave energy period and wind speed are all generated at 6 hourly intervals. A 100 year modelled time series, produced using an 18 year hindcast dataset from the Farr Point site, has been analysed in order to validate the process.

There are several limitations and improvements that could be made to the model. Firstly, the resolution of 6 hours has been chosen as it is suitable for assessing marine operations within the O&M tool. However, when used to probabilistically determine sea states, this resolution may be too large. By averaging values of wave height and period over lengths of 6 hours it is likely that swells are not accounted for. This is particularly relevant when using wave energy period; a parameter calculated from the spectral moment. Another issue identified is with the monthly transitions when a new sea state cannot be found. An alternative method for the third monthly transition stage could be to find the next ‘closest’ sea state, ideally in terms of significant wave height. This would avoid the potential for substantial ‘jumps’ at the beginning of a month.
However, even without these potential improvements, it has been shown that the modelled time series is successful at replicating the seasonal variability of the original 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. It has been proven that the statistical metrics of the 100 year modelled time series are not significantly different from the original dataset in terms of accessibility and estimated power capture.

In conclusion, this validation phase has shown that the Markov chain model is suitable for use in the Pelamis-based O&M tool, although improvements could be made. A vast database of modelled time series, of varying lengths, has been created. As a result the O&M tool is capable of searching through, say the 20 year collection, and choosing one at random with which to carry out its simulations. This is necessary in order for the O&M tool to maintain the function of statistical analysis between different weather scenarios.

V. FUTURE WORK

The Markov Chain Model has been developed as a collection of modules using the coding language Visual Basic for Applications (VBA). As a result, it is flexible and can be easily manipulated to work with any size and resolution of input data. Future work of this IDCORE project will involve using the MCM to process weather data for a number of sites around the UK suitable for wave energy devices. By also identifying logistical bases of operations for each site, it will be possible to analyse the O&M strategies for each, and thus compare them in terms of accessibility and operational expenditure. To the author’s knowledge, this level of detailed analysis specific to O&M has not yet, at the time of writing, been undertaken for wave energy devices.

ACKNOWLEDGMENT

The author would like to thank the academic supervisors of this IDCORE project for their contributions, advice and support. Similar thanks must go to the engineers, past and present, who have dedicated their time to the Pelamis project. The industrial supervisor deserves a special mention, without her experience and guidance this study would not have been a success. The author would also like the IDCORE programme and its funding bodies, in particular the ETP (Energy Technology Partnership), for their support.

REFERENCES