I. Abstract

A method for synthesising wind speed time series (WSTS) from limited data is required that can be used for reliability examination of wind farms and maintenance strategies for a range of wind speed scenarios. Key characteristics of the wind resource need to be captured, including energy availability and maintenance weather windows. 4 WSTS simulators were used to produce synthetic WSTS based on benchmark data from a meteorological mast data at the offshore Egmond aan Zee wind farm in the Netherlands.

These synthetic WSTS were compared with test criteria to determine their suitability for reliability analysis. This included comparing the synthetic WSTS to the benchmark data in terms of the energy availability in the wind and from a typical turbine, residence time at wind speeds, number of transitions between 1m/s wind speed bins, replication of seasonal characteristics including weather windows, and underlying statistical properties.

Based on the chosen criteria, the most appropriate WSTS simulator was the modified Markov process. However, no modelling technique performed best against all criteria and none capture the autocorrelation function (ACF) as closely as desired. Therefore, there is scope for a more advanced technique for wind speed modelling for reliability analysis which combines the best aspects of the models used in this work.

Keywords – Wind speed time series, reliability analysis, weather windows, evaluation criteria.

II. Introduction

To meet EU renewable energy targets for 2020 and beyond, the Levelised Cost of Energy (LCoE) of offshore wind needs to be reduced from the current £140/MWh to below £100/MWh [1]. As Operation and Maintenance (O&M) accounts for around 30% of LCoE [2], researchers have carried out reliability studies on offshore wind to explore the root causes of these costs.

A unique characteristic of wind generation that impacts O&M is the stochastic nature of the fuel. This adds complexity to turbine operation and impacts the available maintenance opportunities, known as weather windows [3]. A weather window is defined as a point in time where a maintenance team could be dispatched to repair a component as the wind speed is below a threshold for a sufficiently long time period to carry out the repair (Figure 1).

Weather windows are particularly critical for offshore O&M due the higher sustained wind speeds leading to fewer opportunities for maintenance. This is compounded by longer lead times due to the large distances from shore. As such there is a need to ensure that these weather windows are accurately represented in O&M research.

The wind resource also adds uncertainty to the expected loss of generation revenue during turbine downtime. Therefore reliability studies need to

Figure 1: Example of a weather window. The weather window (shaded area) is the only time where the wind speed is below the threshold (red line) for long enough to carry out the repair.
accurately capture the elements of the wind resource that impact turbine and farm O&M.

To carry out this reliability analysis a site specific wind speed dataset for the lifetime of the farm is essential. Typically this data is collected for a year prior to turbine construction using a meteorological mast, with future data collection affected due to farm wake effects. Therefore a method for synthesising wind speed time series (WSTS) from limited data is required.

The use of synthetic WSTS facilitates detailed reliability studies and maintenance strategy evaluation under a range of wind speed scenarios. For example, a proposed maintenance strategy could be evaluated against a range of scenarios using a Monte-Carlo simulation where the weather windows are not fixed each time, but their occurrence is allowed to vary randomly with a probability determined by the original recorded data. Due to the complexity of wind speed data this is not a trivial task and all characteristics of the wind speed may not be captured using one modelling technique. To ensure a suitable model is chosen a list of reliability analysis specific criteria is needed.

This paper formalises the desirable attributes of synthetic WSTS for use in reliability analysis of offshore wind farms by formulating a list of key criteria. WSTS simulators are used to produce synthetic WSTS which are compared with these criteria to determine the most suitable simulator. From this work the following learning objectives are aimed to be fulfilled:

- To understand what characteristics of the wind speed are important for reliability analysis of offshore wind turbines.
- To provide a testing procedure for ensuring WSTS simulators for synthetic WSTS are fit for purpose.
- To determine the most promising modelling techniques of synthetic WSTS for reliability analysis.

The paper is organised as follows; Section III details past wind speed modelling techniques and why an evaluation criteria is needed, Section IV details the benchmark WSTS data used, the evaluation criteria developed and the wind speed simulators that will be evaluated including a new model methodology, Section V details the results of this evaluation, and Section V concludes the paper.

### III. Past Wind Speed Modelling Techniques

A number of wind farm reliability studies have aimed to capture the stochastic nature of the wind resource, with a particular focus on a site’s mean wind speed. As local wind speed data for a particular site is limited by the length of time a meteorological mast has been installed at the site, these long term reliability studies have required the wind speed to be simulated based on this limited data. This modelling has been carried out in a number of ways:

- Mathematical expectation calculations based on probability distribution functions [4-10].
- WSTS produced by randomly sampling from probability distributions [12-14].
- WSTS produced from Markov processes [15].
- WSTS produced using an auto-regressive moving average (ARMA) model [16-18].

By far the most common of these methods is the use of probability distribution functions for expectation calculations [4-10]. This calculation uses a probability distribution function to represent the wind resource at a site. This is typically using the Weibull distribution [8, 9], or a special case of the distribution with a shape factor of 2 known as the Rayleigh distribution [4, 5, 7, 10]. This probability distribution function can also be simplified to a smaller number of points to provide simpler computation [6]. These methods are used to calculate the expected energy generated or power output, which are used to derive index values such as Availability and annual energy production (AEP). These methods can be used to quickly assess the performance of a wind farm configuration, but cannot be used to detail the effects of factors such as weather windows and maintenance delays. Therefore, this model type is not suitable for the reliability analysis that this research is targeting.

The production of WSTS is more appropriate for investigating these factors. For example, the use of a WSTS can more accurately represent the effect of a turbine failure by giving the range of impacts that may occur, along with their respective likelihood. This can give a confidence level in the energy production produced by the model. Therefore, the production of WSTS is of interest.

One way to generate this time series is to use hindcasting. This involves using the WSTS data available from the site in question directly in the model for producing WSTS for future years. The simplest method is use the original WSTS directly from an anemometer at a site [11] and repeat this data for each simulated year. Alternatively, trends in the data to predict the amplitude variation whilst maintaining the same seasonal variation. Hindcasting is a popular method in system adequacy studies for demand [19] and ensures that the data is realistic, but the stochastic element of the wind resource is not captured.

A common method used is to generate an average farm wind speed is to randomly sample from a probability distribution. Generally, this is done from a Weibull distribution [12-14]. A variation on this is done in [20], where the daily mean wind speed was produced by randomly sampling from a Weibull distribution, but the hourly wind speed is generated from a normal distribution using this mean daily wind
speed. By using random sampling from a distribution, it is ensured that the wind speeds generated are random and based on the site specific details. A large number of times series can be generated, which is required for Monte-Carlo simulation. However, whilst this method can be modified to allow for seasonal variations, it is difficult to include short term trends such as weather windows. These short term trends are important for forecasting to allow accurate maintenance dispatch characteristics.

A more novel approach is the use of a continuous Markov process to generate the WSTS [15]. This method uses the transitions between wind speeds within a real time series to dictate the likelihood of transitions occurring in the generated time series. This is then converted into transition times between states to give the time at a state before moving to the next state [21]. Like the random sampling from a probability density function, the use of probabilities retains the stochastic nature of the wind, but the integration of other characteristics such as seasonality is simpler. However, the calculation of transition rates is more complex, and the wind speed transitions in the model can only occur if they exist in the original data unless transition rate estimations are used.

Finally, analytical relationships derived from original wind speed data can be used. A common method used is known as ARMA. This was developed in [16] for reliability studies and has been used in subsequent work [17, 18]. These models can accurately replicate the original WSTS and be used to generate a number of time series. However, the accuracy of these methods are reliant on the quantity of data available. They also assume that the wind speed follows an analytical relationship across the year, rather than a probabilistic relationship. Whilst noise can be introduced to increase the stochastic characteristics, it still makes the wind speed synthesis predictable.

A number of the above methods are suitable for time series generation for Monte-Carlo simulation. Whilst each method has qualitative benefits and drawbacks that may determine which method is used, no work has been done to quantify the adequacy of these models for reliability analysis. For example factors such as energy available in the wind have not been considered when choosing the approach. This quantification is needed to justify the use of these wind speed models. This will become especially important if methods are introduced that are modifications of the above methods.

IV. Evaluation Strategy

This section outlines the field WSTS data used for benchmarking, the measurement criteria used to assess synthetic WSTS quality, and the WSTS models that are assessed.

a. Benchmark Data

To assess the quality of WSTS simulators, a field WSTS dataset is needed. The data was taken from the meteorological mast at the Egmond aan Zee offshore wind farm in the Netherlands [22]. The data has been validated by Mierij Meteo before it was made publicly available on the NoordzeeWind web pages [23].

Though there are several years of data they are not directly comparable. From July 2006 construction work began on the wind farm and soon afterwards the wind farm began operating. This has affected the wind speed readings from the direction of the wind farm due to wake effects. The directions that have been impacted are detailed in [24]. Therefore, for the purposes of this study only data from 01/07/2005-30/06/2006 inclusive was used.

The data reported is the raw 10-minute supervisory control and data acquisition (SCADA) average wind data at 3 heights (26, 70 and 116m) and on three booms facing north-east (NE), north-west (NW) and south (S). Depending on the wind direction each of these booms are shadowed by the mast at some point. Therefore to produce an accurate benchmark WSTS the data needed to be pre-processed.

The operator has recommended steps for derived wind speed data acquisition [25]. However there were a number of assumptions that have had to be made. The following steps were made to derive this wind speed:

1. The first step was to determine the derived wind direction in order to choose the correct boom to record the wind speed from. The derived wind direction is produced by averaging over two weather vanes for a given time stamp based on a reference wind direction (Table 1). The operator does not state which direction should be used as a reference so the NW boom was used. If the wind direction for the NW boom is missing then the other two booms are checked. If no data is available or if one vane data point are used for averaging is unavailable the derived wind direction is labelled as missing. No statement has been made on what to do when data missing derived wind direction data. It has been assumed that the wind direction has not changed since the last healthy derived wind direction data point. If it is the first time step, a default of 180° is used.

<table>
<thead>
<tr>
<th>Reference Wind Direction (D)</th>
<th>Derived D Vanes</th>
</tr>
</thead>
<tbody>
<tr>
<td>330° ≤ D &lt; 30°</td>
<td>Average of NW &amp; NE vanes</td>
</tr>
<tr>
<td>30° ≤ D &lt; 90°</td>
<td>Average of S &amp; NW vanes</td>
</tr>
<tr>
<td>90° ≤ D &lt; 150°</td>
<td>Average of S &amp; NE vanes</td>
</tr>
<tr>
<td>150° ≤ D &lt; 210°</td>
<td>Average of NW &amp; NE vanes</td>
</tr>
<tr>
<td>210° ≤ D &lt; 270°</td>
<td>Average of NW &amp; S vanes</td>
</tr>
<tr>
<td>270° ≤ D &lt; 330°</td>
<td>Average of NE &amp; S vanes</td>
</tr>
</tbody>
</table>

Table 1: Weather vanes used for derived wind direction based on current wind direction.
Table 2: Cup anemometers used for derived wind speed based on derived wind direction

<table>
<thead>
<tr>
<th>Derived Wind Direction ($\Delta \theta$)</th>
<th>Anemometer</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0^\circ \leq \Delta \theta &lt; 120^\circ$</td>
<td>NE cup anemometer</td>
</tr>
<tr>
<td>$120^\circ \leq \Delta \theta &lt; 240^\circ$</td>
<td>S cup anemometer</td>
</tr>
<tr>
<td>$240^\circ \leq \Delta \theta &lt; 360^\circ$</td>
<td>NW cup anemometer</td>
</tr>
</tbody>
</table>

2. Using the derived wind direction, a derived wind speed can be produced. This wind speed is taken from the cup anemometer which is chosen from the current derived wind direction as detailed in Table 2.

There were some missing wind speed data points. The WSTS simulators used in this work require complete datasets and therefore the missing data was substituted rather than omitted.

A number of data cleaning methods were attempted but the following algorithm was found to provide the best results.

1. If the mean wind speed value from a different height was available the logarithmic height law (1) was used to estimate the wind speed at that time period.

$$U_z = \frac{\ln \frac{Z}{Z_0}}{\ln \frac{Z}{Z_{ref}}} U_{z ref} \quad (1)$$

Where $U_z$ is the wind speed (m/s) at the required height $Z$ (m), $Z_0$ is the roughness length (m), and $U_{z ref}$ is the wind speed (m/s) at the anemometer height $Z_{ref}$ (m/s).

2. If both alternative wind speed heights are available, the alternative height data with the highest correlation coefficient to the healthy height data in question was used.

3. If no alternative data was available but the data point is isolated, the value was linearly interpolated between previous and future data points at that the height in question.

4. If no alternative data was available as a reference and there were a large number of corresponding missing data points, data from another year was substituted in. This only represented 0.75% of the data.

Following this procedure a cleaned derived benchmark WSTS was produced at all 3 boom heights. For this study the data at 70m was used at this had the highest data recovery rate before cleaning and is the hub height of the wind turbines found at the Egmond aan Zee wind farm. Figure 2 displays this benchmark WSTS.

To note a key assumption of using this data is that this one year is representative of the wind speed for all years the wind farm will be operating. In order to produce more data measure-correlate-predict (MCP) could be implemented. MCP is the process of using past data from nearby meteorological stations to predict what the wind speed was at the farm site. This can be done by comparing the known data at the wind farm site with the same time period of data from the meteorological stations to produce an analytical relationship between the datasets. This relationship can be then used to predict the data to give a much longer WSTS (in some cases decades). This would be an interesting extension but has not been implemented here.

To summarise, wind speeds from a meteorological mast at the Egmond aan Zee wind farm were used. This wind speed had to be derived from the raw SCADA data and cleaned to produce a complete dataset. The derived mean wind speed data was taken at 10 minute intervals at 70m above sea level for the year 01/07/2005 to 30/06/2006.

---

Figure 2: Cleaned derived 10-minute average wind speed data at 70m height used as the benchmark WSTS.
### b. Desirable Criteria for Synthetic WSTS

To determine the suitability of synthetic WSTS for reliability analysis a list of criteria has been created for comparison with the benchmark WSTS. These criteria have been chosen as the results of any reliability analysis using synthetic WSTS will be sensitive to the accuracy of these parameters. The synthetic WSTS should replicate the benchmark data's:

1. Total energy availability in the wind resource.
2. Energy availability from a typical turbine.
3. Cumulative time at all wind speeds.
4. Number of transitions between wind speed states.
5. Longer term seasonal trends and occurrence of sustained low and high wind conditions.
6. The underlying statistical relationships, determined by the sample auto-correlation function (ACF), in the benchmark WSTS.

To make these criteria measurable the calculation for the benchmark and synthetic WSTS needed to be defined. Note the wind speed data was discretised into 1 m/s states.

1. To quantify energy availability in the wind, the expected power densities \( E(p) \) of benchmark and synthetic WSTS were computed (2). \( E(p) \) was used to remove any un-required information such as turbine size and turbine life span.

\[
E(p) = \sum_{u=0}^{u_{\text{max}}} 0.5 \rho u^3 F(u) \tag{2}
\]

Where \( \rho \) is air density (kg/m\(^3\)), \( u \) is wind speed (m/s), \( u_{\text{max}} \) is the maximum wind speed (m/s) and \( F(u) \) is the probability distribution function of the wind speeds.

2. The expected power from a typical turbine \( E(p_t) \) was computed similarly (3), but with limits to represent the turbine power curve (4).

\[
E(p_t) = \sum_{u=0}^{u_{\text{max}}} 0.5 C_p u_t^3 F(u) \tag{3}
\]

\[
u_t = \begin{cases} 
0, & u < u_{\text{in}} \\
u, & u_{\text{in}} \leq u < u_{\text{rated}} \\
u_{\text{rated}}, & u_{\text{rated}} \leq u < u_{\text{out}} \\
0, & u \geq u_{\text{out}}
\end{cases} \tag{4}
\]

Where \( C_p \) is the coefficient of performance, \( u \) is the equivalent wind speed for a wind turbine (m/s), \( u_{\text{in}} \) is the cut-in wind speed (m/s), \( u_{\text{rated}} \) is the rated wind speed (m/s), and \( u_{\text{out}} \) is the cut-out wind speed (m/s).

The turbine data is given in Table 3. In reality the \( C_p \) would vary to main the power extraction, rather than the wind speed itself changing, and would not be at the Betz limit.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \rho )</td>
<td>1.225 kg/m(^3)</td>
</tr>
<tr>
<td>( C_p )</td>
<td>0.593</td>
</tr>
<tr>
<td>( u_{\text{in}} )</td>
<td>4 m/s</td>
</tr>
<tr>
<td>( u_{\text{rated}} )</td>
<td>13 m/s</td>
</tr>
<tr>
<td>( u_{\text{out}} )</td>
<td>25 m/s</td>
</tr>
</tbody>
</table>

Table 3: Wind and turbine parameters.

3. A plot was used to visually compare the cumulative time at all wind speeds between benchmark and synthetic WSTS.
4. The number of transitions between wind speed states for the benchmark and synthetic WSTS was recorded for comparison.
5. To assess the quality of replicating seasonal characteristics two tests were carried out. Firstly, the frequency spectrum of the synthetic WSTS was compared to the benchmark WSTS by using a Fourier Transform. In the spectrum, the frequencies for both WSTS should have similar amplitudes if the seasonal variation has been modelled successfully.

The second test quantified the occurrence of weather windows. This was computed by calculating the percentage of time a maintenance team could be dispatched. The length of the weather window and the wind speed threshold is dependent on the maintenance type and the travel distance. For this work a wind speed threshold of 10 m/s and a time of 48 hours was taken as a weather window, similar to those found for a jack-up vessel in [3].

6. To assess whether the underlying statistical properties of the wind speed were captured the ACF was computed for an exemplar synthetic WSTS for each modelling technique and the benchmark WSTS.

These 6 criteria produced 7 measurements to be used as a metric to assess synthetic WSTS quality. These characteristics and measurements are summarised in Table 4.

### a. Synthetic WSTS Simulation

As discussed in Section III there are a number of WSTS simulators that have been used for synthesising WSTS. This section outlines the 4 modelling techniques that have been assessed, including one modified approach developed for this work.

**Random sampling from probability distributions (PDF).** The model used was developed in [20] and randomly samples from both Weibull and normal distributions to produce the WSTS. Figure 3 outlines this method.
### WSTS Desirable Criteria

<table>
<thead>
<tr>
<th>Criterion No.</th>
<th>Measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Energy availability in the wind resource Comparison of expected power density (E(\rho)) in wind</td>
</tr>
<tr>
<td>2</td>
<td>Energy availability from a typical turbine Comparison of expected power density (E(\rho t)) from a turbine</td>
</tr>
<tr>
<td>3</td>
<td>Cumulative time at all wind speeds Comparison of plotted time at each wind speed</td>
</tr>
<tr>
<td>4</td>
<td>Number of transitions between wind speed states Comparison of transitions between 1m/s wind speed bins</td>
</tr>
<tr>
<td>5</td>
<td>Longer term seasonal trends and occurrence of sustained wind conditions 1. Comparison of Frequency spectrums from a Fourier transform 2. Comparison of % of time a maintenance team could be dispatched (weather window); below 10m/s for 48 hrs</td>
</tr>
<tr>
<td>6</td>
<td>Same underlying statistical relationships Comparison of sample auto-correlation functions (ACF)</td>
</tr>
</tbody>
</table>

**Table 4: Summary of desirable criteria and corresponding measurement strategy.**

---

**Continuous Markov Process:** The Markov process uses the transitions between wind speeds in the original WSTS to dictate the transitions occurring in the synthetic WSTS. The next wind speed state is determined by the current state, and the shortest transition time from the current state (Figure 4).

A Markov process approach assumes that the system is memory-less; future random behaviour is dependent on the current state, and the process is stationary; the behaviour of the system is time independent. As such, the transition rates must be constant. The state residence times are assumed to follow an exponential distribution [26]. The transition rates are calculated using (4). The transitions times from the current state \(i\) to state \(j\) are calculated using (5).

\[
\lambda_{ij} = \frac{N_{ij}}{T_i} \quad \text{(4)}
\]

\[
L_{ij} = h_{yr} \cdot \frac{\ln(R_i)}{\lambda_{ij}} \quad \text{(5)}
\]

Where \(N_{ij}\) is the number of transitions between states \(i\) and \(j\), \(T_i\) is the total time at state \(i\) (years), \(L_{ij}\) is the transition time (hours), \(h_{yr}\) is the number of hours in a year (8760), and \(R_i\) is a uniform random number between 0 and 1.

The Markov process model used is similar to that found in [15], with a number of variations:

- The initial state is determined by randomly sampling from a Weibull distribution.
- The input data is based on 10 minute average, rather than hourly average, wind speed data.

---

**Figure 3:** Flow diagram of PDF sampling model.

**Figure 4:** Schematic diagram of Markov process. \(\lambda_{ij}\) is the transition rate between states \(i\) and \(j\) (occurrences/year).
• Missing transition rates due to having a small dataset were filled using linear interpolation.
• Transitions can occur to any other state from the current state, unlike the birth-and-death model used in [15]. The next state is determined by the smallest transition time calculated (5) and states are represented by 1 m/s bins.
• The benchmark data was sampled at 10 minutes, so transitions times that are below 10 minutes are rejected. As a continuous process is used, transition times can be a non-integer multiple of 10 minutes to produce a continuous-time WSTS.

Modified Markov Process: This model uses the continuous Markov but consists of higher order processes that do not determine the output of the model, but do dictate which lower order states are available. In this paper the high order processes were used to produce periods of higher and lower sustained wind speeds. The process is outlined in Figure 5.

Auto-Regressive (AR(X)) model: An AR(X) model uses the weighted value of the previous X observations in a time series, alongside a randomly generated error term, to determine the next value in the series. In this case a simple third order model (AR(3)) was used. This model was used as follows:

1. The mean of the benchmark WSTS was subtracted from each data point in the benchmark WSTS.
2. MATLAB’s ar function was used to fit the AR(3) parameters.
3. The noise at each time step \( e(t) \) was calculated (6).

\[
e(t) = \sigma_e R_n(t)
\]

Where \( \sigma_e \) is the standard deviation of the noise from the ar fit and \( R_n \) is a normally distributed random number between 0 and 1.

4. The synthetic WSTS is generated using (7).

\[
y(t) = a_1 y(t-1) + a_2 y(t-2) + a_3 y(t-3) + e(t)
\]

Where \( y \) is the synthetic WSTS and \( a_{1-3} \) are the AR coefficients. \( y(1-3) \) is taken from the benchmark WSTS.

Table 5 gives the values used for this AR(3) model.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma_e )</td>
<td>0.683 m/s</td>
</tr>
<tr>
<td>( a_1 )</td>
<td>0.8794</td>
</tr>
<tr>
<td>( a_2 )</td>
<td>0.0047</td>
</tr>
<tr>
<td>( a_3 )</td>
<td>0.1044</td>
</tr>
</tbody>
</table>

Table 5: Parameters for AR(3) model.

To summarise, four synthetic WSTS simulators have been modelled for comparison with the benchmark WSTS. They are a PDF model, a continuous Markov process, a modified Markov process to include weather windows, and an AR(3) model.

**V. Results and Discussion**

This section summarises the results. From Table 6 all models perform well against criterion 1 (within 3.2% of benchmark), with AR(3) performing marginally best (underestimated by 1.6%). There is slightly more variation in the results for criterion 2 (up to 3.8%), but this time the modified Markov performs best with the benchmark \( \mathbb{E}(p_i) \) falling within confidence bounds of the modified Markov result. From criteria 1 and 2 none of the models can be ruled out.

Results from criterion 3 (Figure 7) provide more insight. Both the Markov and modified Markov models replicate the mean bin times very accurately. This is to be expected as the models replicate the probability at being at each state in the calculation of \( \lambda_j \). Their slight overestimation for criteria 1 (2.8% and 1.9% respectively) and, to a lesser extent, 2 (1.3% and 0.3% respectively) may be due to the linear interpolation of \( \lambda_j \) for higher wind speeds, increasing
how regularly these wind speeds occur. As the energy available in the wind follows a cubic relationship, a very slight overestimation in residence times of higher wind speeds can have a noticeable impact in the results.

The PDF model performed worst for criteria 1 and 2, overestimating by 3.2% for criterion 1 and underestimating by 3.8% for criterion 2. This overestimation is due to high wind speeds increasing the $E(p)$, whilst the distribution skew to the left in the turbine operating region (Figure 7) causes the underestimation for criterion 2. The PDF also produces a different distribution to that of the benchmark WSTS, showing that the normal distribution was not a suitable estimation for the daily wind speed distribution.

The AR(3) model performs better than the PDF model against criterion 3 as it does not have the same skew to the left (Figure 7) and this is reflected in the better performance against criterion 1 and 2 (Table 6). The AR(3) model is poor at low wind speeds, with an overestimation of 0m/s bin time (Figure 7). This is not apparent in the criterion 1 and 2 results as the amount of energy is negligible in the low speed region. This overestimation may be due to the normally distributed noise variance assumption. Therefore examining the distribution of noise variance and modelling accordingly may provide a more accurate model.

None of the models accurately replicate the number of transitions in the wind speed (criterion 4, Table 1). The PDF model has only 21.5% of transitions due to the over-suppression of variation from the filter. The AR(3) model has 135% the number of transitions as the model works on a fixed time step basis with the wind speed value varying at nearly every step. The lower number of transitions for the Markov process is due to a skew towards longer transitions times from the rejection of time steps lower than 10 minutes. The discrepancy between that and the modified Markov is likely due to the elongated periods when the model is at higher wind speeds where the transition rates are much higher, causing shorter transition times than the normal Markov. This would highlight that, whilst the modified Markov performs best against criterion 4, the distribution of transitions across wind speeds needs further investigation to verify if this is desirable.

The modified Markov performs best against criterion 6 (Table 6). This would be expected as the model is deliberately set up to replicate weather windows, but this result verifies that it has performed this successfully. The Markov and AR(3) models do not capture these weather window opportunities, showing the need to explicitly model these windows. The PDF approach overestimates the amount of time available for maintenance by 63.5%, again due to the filter. This could have a significant impact on the results of any reliability analysis that considers weather windows.

Figure 6 reveals that the seasonality of the benchmark WSTS (represented by peaks at low frequencies) was not captured by the Markov process. The PDF model

### Table 6: Results of evaluation. Best results are given in bold.

<table>
<thead>
<tr>
<th>Evaluation Criteria</th>
<th>Benchmark WSTS</th>
<th>PDF</th>
<th>Markov</th>
<th>Modified Markov</th>
<th>AR(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E(p)$ (W/m$^2$) [1]</td>
<td>657.3</td>
<td>678.3 ±9.4</td>
<td>675.8 ± 6.9</td>
<td>669.6 ± 9.2</td>
<td>646.9 ± 8.9</td>
</tr>
<tr>
<td>$E(p_t)$ (W/m$^2$) [2]</td>
<td>510.9</td>
<td>491.6 ± 4.8</td>
<td>517.5 ± 3.8</td>
<td>512.4 ± 5.4</td>
<td>529.2 ± 5.5</td>
</tr>
<tr>
<td>Number of Wind Speed Transitions [4]</td>
<td>19378</td>
<td>4169 ± 18.0</td>
<td>11624 ± 26.9</td>
<td>15838 ± 29.3</td>
<td>26139 ± 32.0</td>
</tr>
<tr>
<td>Weather Window Opportunities (%) [6]</td>
<td>14.8</td>
<td>24.2 ± 0.6</td>
<td>4.5 ± 0.2</td>
<td>13.5 ± 0.5</td>
<td>6.8 ±0.3</td>
</tr>
</tbody>
</table>

![Figure 7: Wind speed mean residence times of benchmark and synthetic WSTS (criterion 3).](image7.png)

![Figure 6: Frequency spectrum of benchmark and selected models. The arrow indicates 1 cycle/year.](image6.png)
used mean speeds for each month, and therefore replicated the high amplitudes at low frequencies of the benchmark WSTS with accuracy. The AR(3) and modified Markov results are not shown here as they perform almost identically to the Markov model results. Seasonality should be added to any WSTS simulator.

None of the models replicate the ACF of the benchmark WSTS (criterion 7, Figure 8). The AR(3) model is closest, which is expected as it uses the ACF to fit the model coefficients. This would indicate that a higher order AR model is required. The Markov and modified Markov perform the worst. This is because the Markov processes only using the previous time step to dictate the next state. Therefore a higher order AR model is likely to perform better against the ACF than any of the models represented here.

![Figure 8: ACF of WSTS (criterion 7).](image)

In summary:

- The AR(3) performed best for criterion 1 and 7, but performed poorly against criterion 3 and 6.
- The modified Markov performed best for criteria 2-4, and 6 though the result for criterion 4 needs further investigation. It performed poorly against criterion 7.
- Only the PDF model captured the seasonal variations (criterion 5).
- The PDF model performed the worst for all criteria other than criteria 5 and 7 and therefore is not suitable for this kind of reliability analysis without modification.

Arguably the wind speed residence times (criterion 3) and weather windows (criterion 5) are most important for reliability analysis as they dictate both the energy availability and maintenance opportunities, which are likely to have the highest impact on LCoE. Therefore, based on the analysis presented the most appropriate WSTS simulator is the modified Markov process.

However, no modelling technique performed best against all criteria and none captured the ACF as closely as desired. Therefore, there is scope for a more advanced technique for WSTS simulation for reliability analysis. Based on the results detailed, a combination of a more complex AR model for short term modelling and elements of the modified Markov process for short and long term weather conditions could prove the most appropriate modelling technique.

**VI. Conclusion**

This work has outlined 6 desirable criteria of synthetic WSTS for use in offshore wind farm reliability analysis, with 7 measures to quantify how well synthetic WSTS matched these criteria. These included considering the time spent at all wind speeds and opportunities for maintenance via weather windows. This evaluation criteria can be used to evaluate any WSTS simulators developed specifically for sequential reliability studies.

4 WSTS simulators were used and were compared with these criteria. These were based on models produced in previous reliability studies and a modified Markov model developed in this work. Based on the evaluation criteria the most appropriate WSTS simulator for reliability analysis was the modified Markov process. This is as it closely replicated a number of criteria including the weather window opportunities and the wind speed distribution. No modelling technique performed against all criteria and none captured the ACF closely. Therefore there is scope for a more advanced technique. A combination of an AR model and elements of the modified Markov process could produce promising results.

**References**


