Methods



Figure 1: Representations of (top) the transition probability matrix (TPM) and (bottom) the cumulative transition probability matrix (CPM), which were used to generate simulated data from real data. In the TPM, a dark square in the grid indicates a low transition probability from the current wave height state to the next state, while a light colour indicates a high probability. In the CPM, the values along each row indicate the cumulative transition probabilities from the current state to all next states up to that point in the row. The two figures together show that the wave state at any time is overwhelmingly likely to be similar to its previous state.

To get a better idea of how the Wells turbine would perform over a year, it was decided to generate 100 annual sets of hourly wave data from the single annual set of real data that we had. These simulated data sets had to have the same descriptive statistics as the real data, and the same distribution of wave heights. There is a well-known technique for doing this based on use of Markov chains, and this is the approach we used.

First a transition probability matrix (TPM) is constructed from the ral data set that finds the probability that the wave height will take any value, given its previous value. From this, a cumulative set of probabilities are constructed for each wave height state, that give the cumulative probability that the next state will be below some value. These are illustrated in Figure 1.

To generate the simulated data, one starts at some random wave height then generates a random number between 0 and 1. The next state is found using the CPM. In the row of the current state, it is the first next state where the cumulative probability exceeds the random value. In this way, it is likely to be similar to the current state, just as in real data. This process is repeated for one year’s worth of values.

100 such simulated sets of data were generated in this way.

The population of the CPM and TPM, the generation from them of the simulated data and the analysis of that data was all carried out using R.

To calculate the power output of the Wells turbine for any given wave height, we devised the following scheme:

<add in the theory I showed you>

This requires knowledge of the lift and drag coefficients of the turbine as a function of angle of attack. These were taken from publicly available values and digitised. They are shown in Figure 3

The output power at any time is a function of the angle of attack. This is affected by the chosen rpm (revs per minute, revolutionary speed) of the turbine. We found the mean power over one year for one simulated data set as a function of rpm to identify the optimum rpm to use.

Given this rpm, we then found the mean and maximum power output of the turbine for each of the 100 simulated data sets.

We concerned to see how stable the power output was as a function of time of year.



From this figure we see that the lift to drag ratio is maximised for an angle of attack in the range 14-17.

The output power of the Wells turbine was calculated for each set of simulated data, for each hour of the year

This was implemented in Python.

**Results**

The statistical similarity of real and simulated data are shown in Figure 2, where histograms of real and one set of simulated data are shown together, along with a time-series plot of the two. The histograms of both the real and simulated data are both well modelled by a Weibull distribution, as is the case for wind speed data. This distribution is characterised by two parameters, the shape parameter, which governs the width of the distribution, and the scale parameter, which governs the location of the peak.

For the real data, the shape parameter is 1.85 +/-0.014 and the scale parameter is 2.05 +/- 0.09. The mean value for the shape parameter for the simulated data is 2.06 +/- 0.04 while that for the scale parameter is 2.01 +/- 0.15. This means that the simulated data agree with real data as to scale parameter and so have a similar average value, but slightly overestimate the shape parameter, which means that the scatter in their wave height values is slightly larger than for real data.



Figure 2: (top) Probability distributions for wave heights for real and one set of simulated data, together with Weibull distribution fits ot the data. (bottom) Sample time series of real data together with the same simulated data as was used for the histogram.

For comparison, the real data set is shown together with three complete simulated sets in Figure 4



Figure 4: (top) Whole year of real wave height data together with three sets of simulated wave height data.

The mean power output of the turbine as a function of rpm is shown in Figure 5 below:



Figure 5. Mean power output as a function of rpm.

This mean power output is clearly sensitively dependent on rpm, with an rpm=20 value being seen to be optimal. Interestingly, if the power output over one year for any of the simulated data sets is compared for this rpm value with what it would be for a much higher rpm value, say 200 rpm, the variability from month to month seems to be much less, as shown in Figure 6.



Figure 6: (red) Variation over one year of power output when rpm=200 and (blue) when rpm=20.

Finally the monthly variation of mean and max power are plotted in Figure 7 below



Figure 7 (top) monthly variation of mean power (bottom) monthly variation of maximum power.