Estimation of Energy Loss due to Variability on the 1-Minute Scale

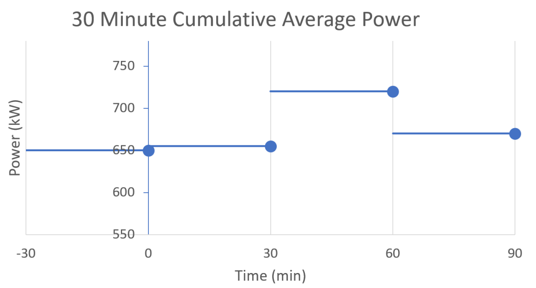
Chae Gordon : Lake Parime — chae.gordon@lakeparime.com

January 2020

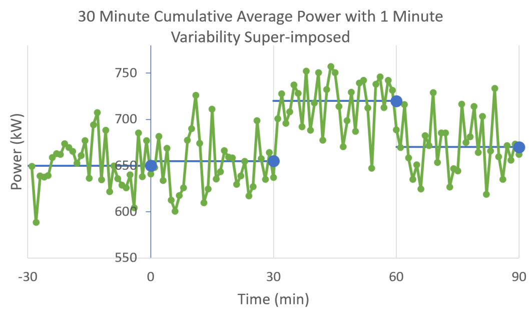
Abstract

A methodology to estimate energy losses resulting from renewable source variability at frequencies higher than available data resolutions was developed. The methodology is expected to be only required for sites of capacity less than 1MW, where data of frequency one Hz is hard to come by. This is because the inertia constant of wind turbines is greater than one second. The losses were found in the range 4-14 % depending on the frequency of the data available (10 and 30 minutes respectively). The error in this range was estimated using arguments based on power spectral density and found to be ± 7%.

Contents



(a) Example of available data resolution at $\leq$ 1MW sites.



(b) Hypothetical example of 1-minute data.

Figure 1: Figure compares the data available (a) with a hypothetical example of 1-minute data (b) to demonstrate that the available data resolution affects the apparent variability of wind at the site by averaging out turbulence effects.

1 Introduction

1.1 Problem Statement

For smaller sites (generation < 1MW) data of a resolution finer than 10-30 minutes may be hard to come by. The question then is, given how variable renewable energy (RE) sources can be, how much energy will we actually have available to IT in a 30-minute period? The problem arises because the PowerBox has a maximum power draw and if the RE source power fluctuates above this, energy will be "spilled". We predict this problem will be worst for wind energy and will disappear as we scale the business to sites above 1MW (where data at 1Hz is readily available). Data of frequency higher than 1Hz will not be required because the inertia constant of wind turbines is between 2-9s [18]. This problem is considered separately to the PowerBox's response time and the losses incurred from this. This is important because server response behaviour will change drastically as Lake Parime's *PowerDrive* is constantly updated and we expect *PowerDrive's* reaction time to be comfortably within the inertia constant of a turbine.

1.1.1 Problem Regimes

The problem manifests itself in three ways.

Firstly, when the 30-minute average power is lower than the maximum IT but some 1-minute points vary above the max IT load. In this case, this energy is spilled and can therefore not compensate the points varying below the average power. Therefore, the actual power available to the PowerBox is lower than the 30-minute average.

Secondly, when the 30-minute average power is above that of the max IT load but 1-minute points vary beneath the max IT load. This is a loss because from the 30-minute data one would expect that you could run the PowerBox as max IT load for the entire 30-minute period but for those points beneath the max IT load this would not be possible.

Thirdly, if the average 30-minute power is sufficiently far from the max IT load, then no points will vary to cross the max IT load from the average point. In this regime no loss is incurred.

The two first regimes are depicted in figure 2.

1.2 Literature Review

The problem sits at the intersection of the cutting edge of meteorology, grid balancing and turbine failure research. However, the specifics of the PowerBox project force us to combine these fields in a novel way for our solution. The paper by Sørensen et al. (2008) examines the variability of power from a large wind farm from the time scale of 1-minute to 1-hour [22]. Additionally Watson (2014) [26] examines the quantification of wind variability. A common theme in this research is that wind varies on different scales in a hierarchy with seasonal effects giving way to weather fronts and then finally to turbulence. Another, is that when dealing with wind extreme values must be considered and therefore the use of a Weibull distribution is necessary and indeed the median has to be used instead of the mean [11], [3], [15].

In Cadogan et. al (2000) [13] a transition matrix is described which quantifies the probability of transitioning from one percentage of rated capacity to another in any one minute period. The matrix put forward in the paper was for a 10 MW site in the American mid-west. The paper shows that in most

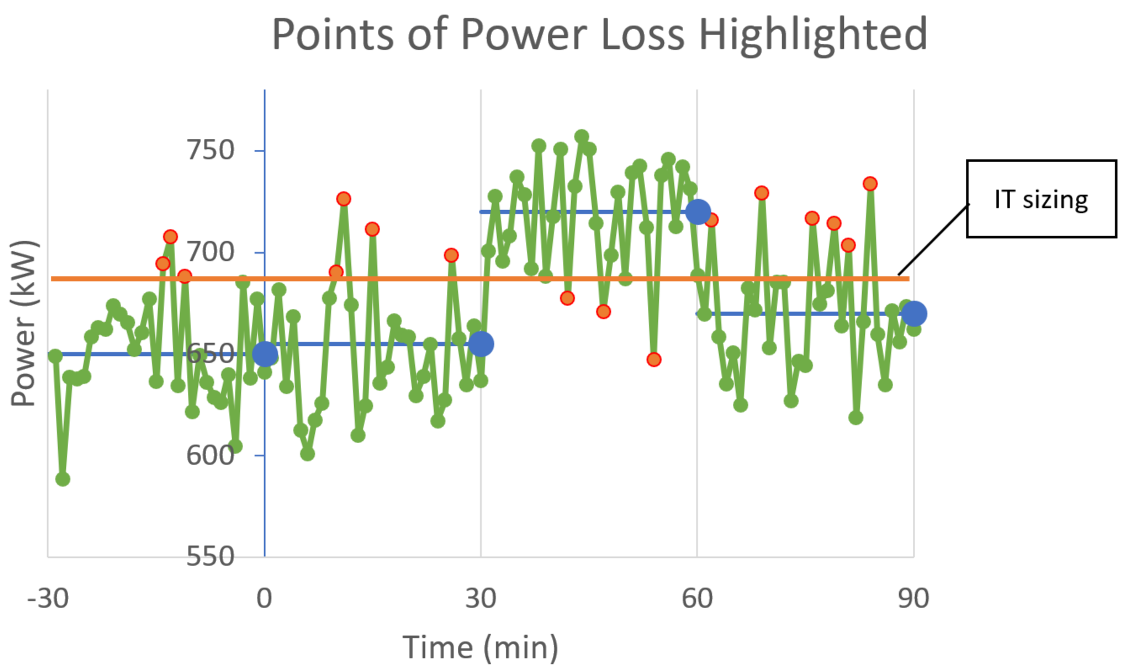


Figure 2: Figure depicts the different regimes in which loss occurs, points of loss are highlighted in orange. The first two 30 minutes depicted (-30 to 30 minutes) show loss occurring due to power varying above the IT sizing. The period between 30 to 60 minutes shows the loss occurring when the average power is above the IT sizing. In this period a naïve analysis would suggest that the PowerBox could be run at full capacity for the entire period. Therefore, when the power varies beneath the IT sizing (highlighted orange) we have a loss compared to our expected available power.

cases (> 90%) the power will not vary on this time scale. Another seminal paper in the field is that of Van der Hozen (1957) [24], putting forward the first thorough research of wind variation across multiple frequency regimes. The research examined the power spectral density of wind up to frequencies of 8 Hz. It showed that there were peaks in the spectral density around the periods of 12 hours, 10 minutes and 1 minutes. Indeed, the approach of examining the spectral reponse and inertia of turbines is the most common topic of research in this area [4], [19], [8], [2].

The problem of wind variability is often studied in the context of extreme events and load failure of turbines [6], [1], [20], [10], [23], [14], [7]. These articles, while interesting, tend to deal either solely with maximum values or with hourly data. Therefore, while the methods can inform the present study they unfortunately are not directly relevant. A useful result is found in Moore (2012) [18] where it is stated that the inertia constant of wind turbines is between 2-9s. Therefore variations beneath the 1s time scale need not be considered. This is useful as we expect to receive 1Hz data for all large deployments. Meaning we should out-grow the need for the more approximate methods outlined below, favouring more empirical methods.

Uncertainties in fitting extreme value distributions for wind data have been previously examined in Vanem (2015) [25] and uncertainties when dealing with wind data are examined in the papers Rose et al., Yang and Mata [21], [16], [17].

2 Method

2.1 Data Requirements

In all scenarios the more data provided by the client, the better the estimate. With between 5-10 years being optimal, data of at least 1 year will be required to account for seasonal variability [12]. In the absence of site-specific data for this period conservative approximations should be made.

2.2 Scenario 1: Client has 10-minute SCADA data with max, min and average power readings.

It is possible to envision a worst-case waveform for losses due to variability. This worst case scenario is a square wave alternating between the maximum capacity of the turbine and zero power, with the time at each power set so as to conserve the average power in the power (shown in figure 3).

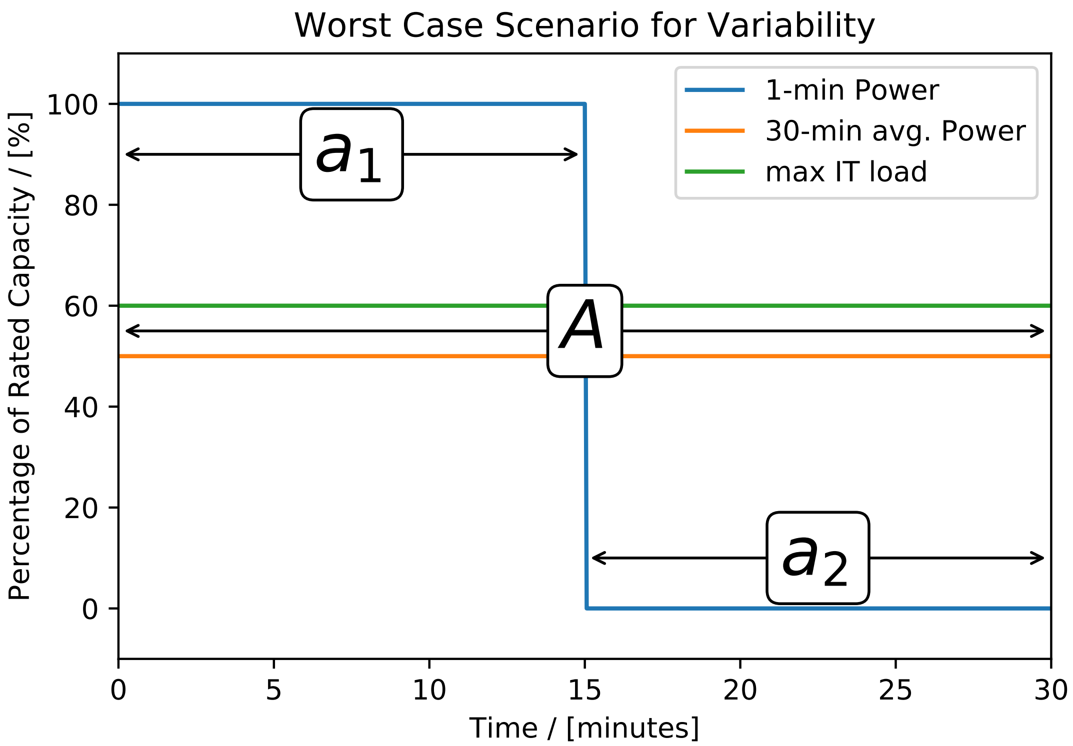


Figure 3: Plot explains the definitions of a_1, a_2 and A. With the A being the total time over which the 1-minute data-points are averaged (either 10 or 30 minutes), a_2 is the time spent at min power and a_1 is the time spent at max power. Although in this case $a_1 = a_2$, this is not always the case and is only a quirk of the average power being 50% of the rated capacity.

In this worst case scenario there are essentially two regimes in which loss occurs. Firstly, when the max IT load is above or equal to the 30-minute average (depicted in figure 4). In this case when the power is above the max IT load it is not utilised and the lost energy (l) is equal to the spilled power l integrated with respect to time. For a square wave we can just multiply l by the time at max power (a_1 as defined in figure 3). The equation for the loss in this case is:

$$E_l = l(a_i), i = 1, 2 \tag{1}$$

with i=1.

Secondly, the average 30-minute power could be above the max IT load but the variability could cause the power to at some times be beneath the max IT load. This constitutes spill because naïve to the extent of the variability we would assume that we could run the PowerBox at maximum load for the entire 30 minute period. However, when the power dips below the max IT load this is not the case (illustrated in figure 5). In this case the lost energy is given again by equation 1 with i=2.

In order to quantify the probable loss % from energy on a given site due to wind variability we can take max, min and mean figures for power to calculate which regime we are in and then just evaluate equation 1 for that time period, provided we know the time a_1 or a_2. There are two main equations we can solve for this unknown:

Firstly, that the time in the measurement period is conserved:

$$A = a_1 + a_2 \tag{2}$$

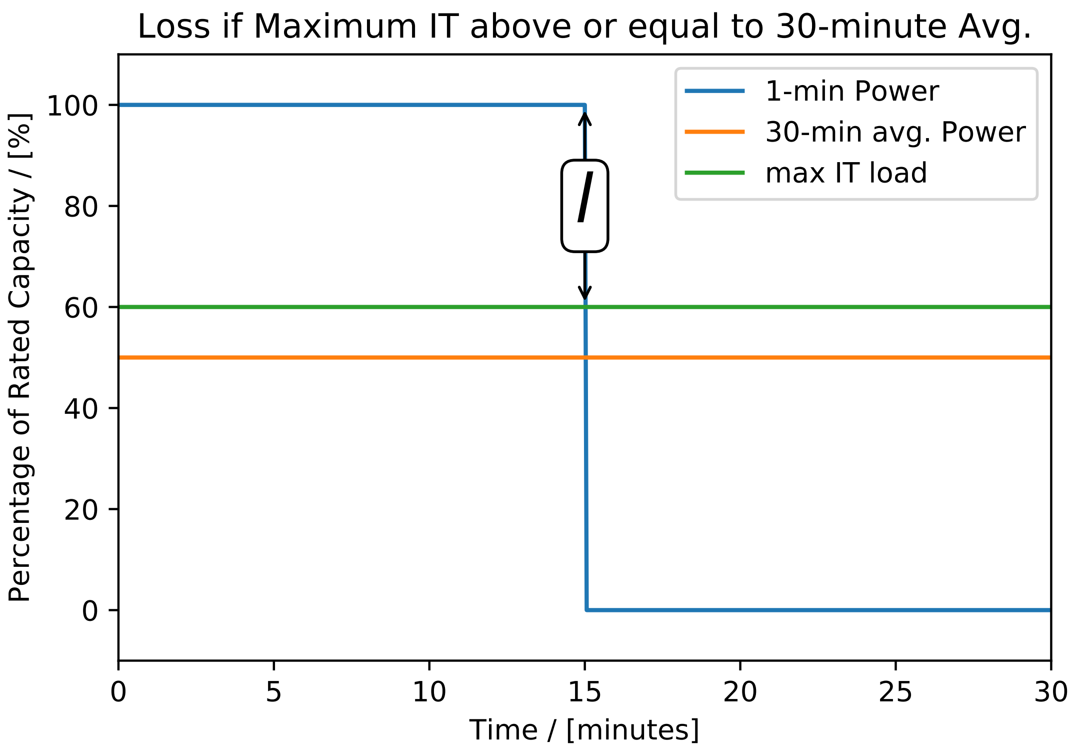


Figure 4: Plot explains the loss l when the max IT load is above the 30-minute average power. When the generated power is above the max IT load this excess power, l, is not utilised. We can therefore integrate this power with respect to time to get the 'spilled' energy.

Secondly, that the energy is conserved:

$$A\bar{p} = a_1 p_1 + a_2 p_2 \tag{3}$$

where $\bar{p}$ is the average power over the measurement period.

We can then easily solve this system of equations for a_i:

$$a_i = A\frac{\bar{p} - p_j}{p_j + p_i} \tag{4}$$

where $i = 1, 2$ and all the other symbols have their pre-defined meaning.

If we have a_i in units of minutes and power in units of kW then we need only multiply their product by $\frac{1}{60}$ (sixty minutes in an hour) to get the loss in units of kWh.

Once we have collated the list of worst case losses, one could calculate a simple mean; however, we could also make use of the Fisher-Tippett-Gnedenko theorem to use an extreme value theory distribution. This would have the benefit of mitigating errors in the mean deriving from small sample size. The theorem states that all extreme distributions tend to either the Gumbel distribution, the Fréchet distribution or the Weibull distribution. The Gumbel distribution can be ruled out because loss is positive definite, the Weibull distribution is selected because research shows that power and wind speed follow Weibull distributions and it has one less parameter to fit. Calculating the median to account for outliers and skewness:

$$\text{median} = \lambda(\ln 2)^{1/k} \tag{5}$$

Here the symbols λ and k are the parameters of the Weibull's probability distribution:

$$f(x) = \left\{ \begin{array}{ll} \frac{k}{\lambda}(\frac{x}{\lambda})^{k-1}e^{-(\frac{x}{\lambda})^k}, & \text{for } 0 \leq x \\ 0, & \text{for } x < 0 \end{array} \right\} \tag{6}$$

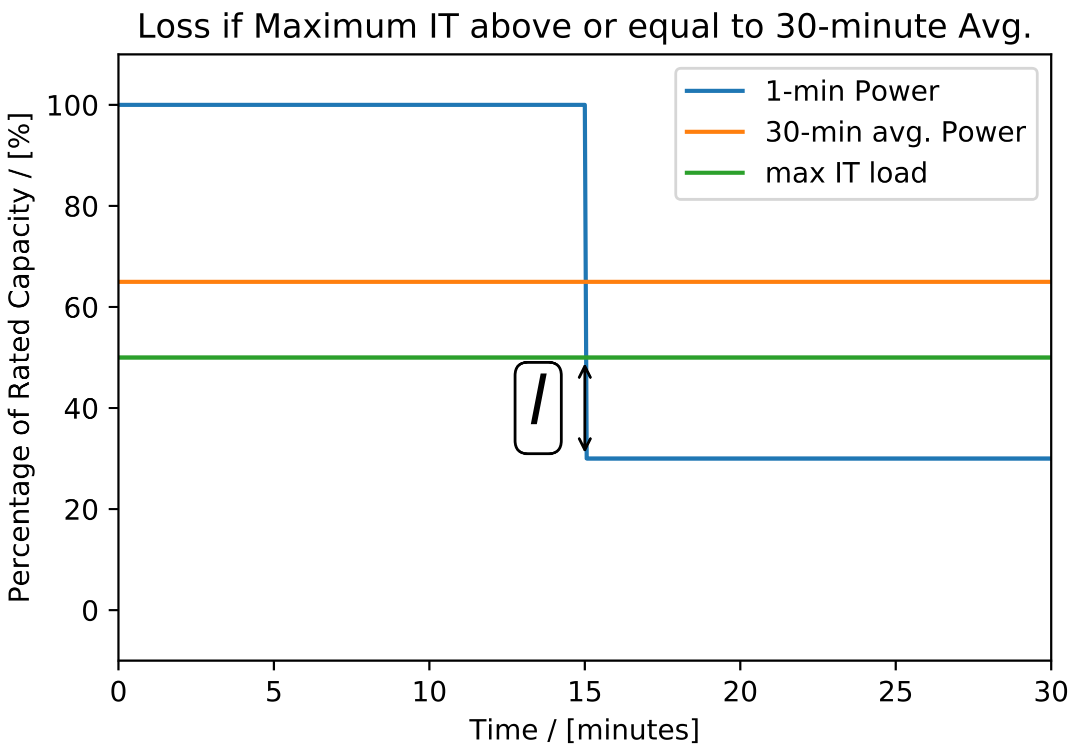


Figure 5: Plot explains the loss l when the max IT load is below the 30-minute average power. When the generated power is below the max IT load we cannot run the PowerBox at maximum load, this then means we have a power deficit of l, compared to our expected load. We can therefore integrate this power deficit with respect to time to get the 'spilled' energy.

2.3 Scenario 2: Client has only 30-minute generation data or 10-minute SCADA data without maximum and minimum power readings.

In the absence of site-specific ≤ 1 minute resolution data we use the transition matrix (figure 6) from Cadogan et al. (2000) [13]. The matrix was calculated by the National Renewable Energy Laboratory (NREL) using 1-minute data from a Mid-West American wind farm of 10MW. This is likely to be slightly less variable than a single turbine site due to the general trend that variability reduces as capacity of site increases. This is because the coherence length of turbulence fluctuations is smaller than the turbine spacing in larger sites. However, 10MW is not so large that this effect will be too disruptive to the methods outlined above.

Each row in the transition matrix depicts the probability that the turbine will transition from the percentage of rated capacity that is the row heading to the rated capacity percentage of a column heading.

The transition matrix allows us to fit transition probability distributions for each 10% increment of the rated capacity. It is then possible to simulate interpolating points in the 30-minute period. The key here is to understand the problem we are aiming to address. We are not looking to create accurate power wave-forms and can ignore the need to have power continuous at all points.

The method is as follows:

- For each cumulative average power data point we interpolate to 1-minute data points
- This is done by starting on the cumulative average power for the first data point and adding a random amount of power variation (drawing from the distributions given by figure 6). Each consecutive point is then centered on the previous one and some random power fluctuation added, until we arrive at the end of the cumulative average period. The next average period is then interpolated in the same way.
- For extreme values (eg. 0% or 100 %) for the whole average period no fluctuation is applied because fluctuations would not conserve the average power produced in this period.

- By having points centered on the previous point we don't reduce the occurrence of extreme fluctuations and if the transition distributions are fairly symmetrical we should still see average power conserved on average.

	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
10%	0.9905	0.0094	0.0001	0	0	0	0	0	0	0
20%	0.0183	0.9592	0.0220	0.0002	0.0001	0.0001	0	0	0	0
30%	0.0007	0.0247	0.9448	0.0295	0.0001	0.0000	0	0	0	0
40%	0.0004	0.0003	0.0339	0.9302	0.0350	0.0001	0	0	0	0
50%	0.0004	0.0002	0.0002	0.0408	0.9151	0.0428	0.0004	0.0001	0	0
60%	0.0004	0.0001	0.0005	0.0003	0.0455	0.9066	0.0465	0.0001	0	0
70%	0.0001	0.0004	0.0001	0.0003	0.0005	0.0465	0.9118	0.0399	0.0002	0
80%	0.0004	0	0.0001	0.0002	0.0001	0.0002	0.0399	0.9192	0.0400	0
90%	0.0004	0	0.0003	0.0001	0	0	0.0001	0.0431	0.9328	0.0231
100%	0.0001	0.0003	0.0001	0.0001	0.0001	0.0001	0.0003	0.0003	0.0364	0.9620

Figure 6: Table depicts percentage probability of a transition in a one minute period. Each row depicts the probability of row header percentage of capacity transitioning to the column header percentage [13].

2.4 Scenario 3: Client has 1 minute data or even 1Hz data

1-minute data (or higher resolution if possible) can be used to assess the variability of power. If we have sufficient 1-minute data for general sites but none for the site in question, we could calculate the probability density function of transitions from the current percentage of rated capacity to another percentage of rated capacity using data from sites similar in capacity and location. Then Scenario 2's method could be used, this could be useful in cases where we have lots of 1-minute data for similar sites but none for the actual site under consideration.

Otherwise, if we have actual 1-minute or even 1 Hz data from the client it is trivial to calculate the loss using equation 1. In this case the time a_i would be set to the time between data measurements. However, this value should be compared against values for other sites if the period of data is small and the more conservative value chosen, or a different method considered if there is a large discrepancy.

3 Results

3.1 Northern Isle of Scotland— 900kW— 10-minute data with max and min (Scenario 1)

For our site in a Scottish Northern Isle we are in possession of 10-minute SCADA data with maximum and minimum readings. Therefore, we apply the method for scenario one. When the PowerBox max IT load is equal to the average annual power of the turbine the loss percentage of useable energy was 4.52%. The loss percentage showed two regimes: while the PowerBox was sized beneath the average power the loss percentage stayed fairly constant between 4-5%; when the PowerBox sizing exceeds the average power of the turbine the loss percentage drops rapidly to zero. These regimes are seen in figure 7.

3.2 Western Isle of Scotland—670kW—30-minute data (Scenario 2)

For our site in the Western Isles of Scotland we have only 30-minute generation data due to the SCADA system at the site being nonoperational. We therefore apply the method for scenario 2 (figure 8). The loss when the PowerBox was sized equal to the annual average energy was (14±7)%. The loss again showed a n-parabolic relationship to the IT sizing of the PowerBox with the loss plateauing around the annual average power (figure 9). The loss's relationship to the variability of wind was weak, the loss increased by only 25% for an increase of the coefficient of variance (%COV) of 3000%.

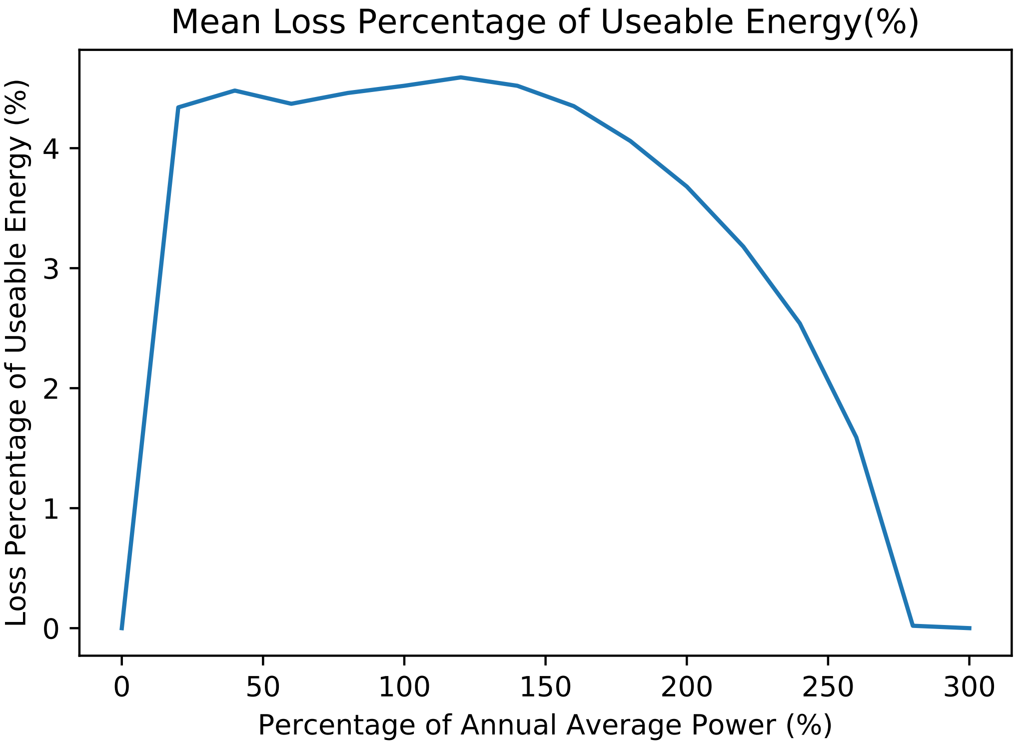


Figure 7: Figure depicts the parabolic relationship between PowerBox IT sizing and the percentage energy loss using our data from the Northern Isles of Scotland. The linear approximation (plateau) coincides with IT sized at the annual average power of the turbine.

3.2.1 Energy Conservation in the Interpolation

Energy was conserved fairly well by the interpolation method with the energy conservation violation only 0.82% after five years of 30-minute data is interpolated.

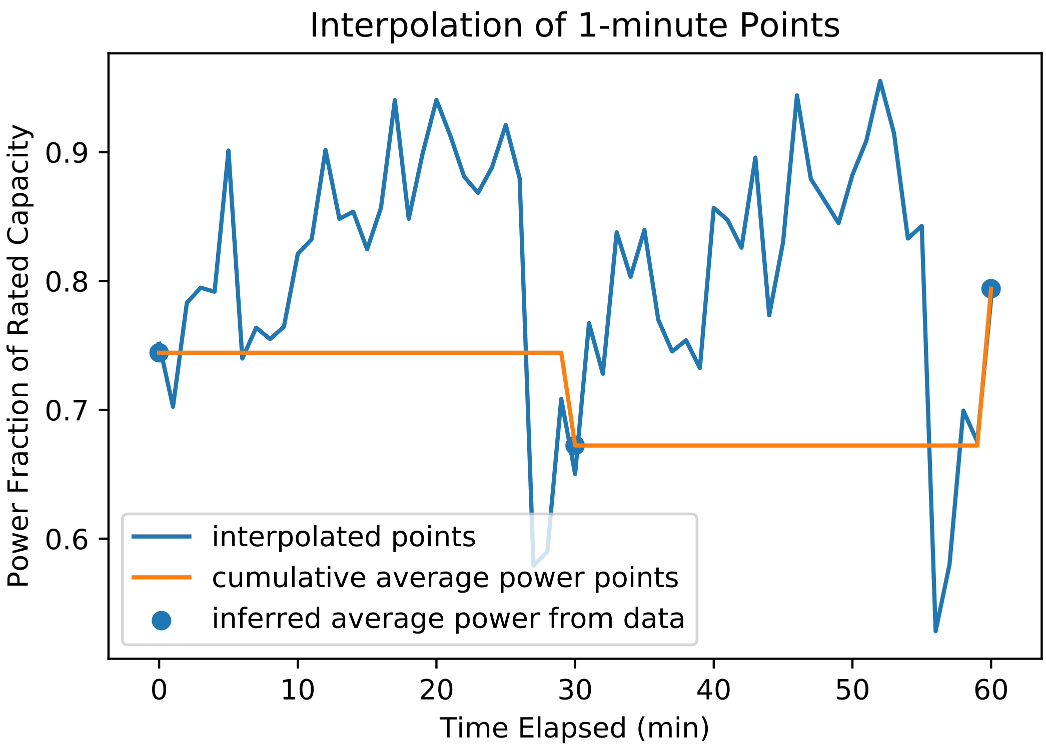


Figure 8
Figure shows the interpolated power points with the cumulative average power super-imposed.

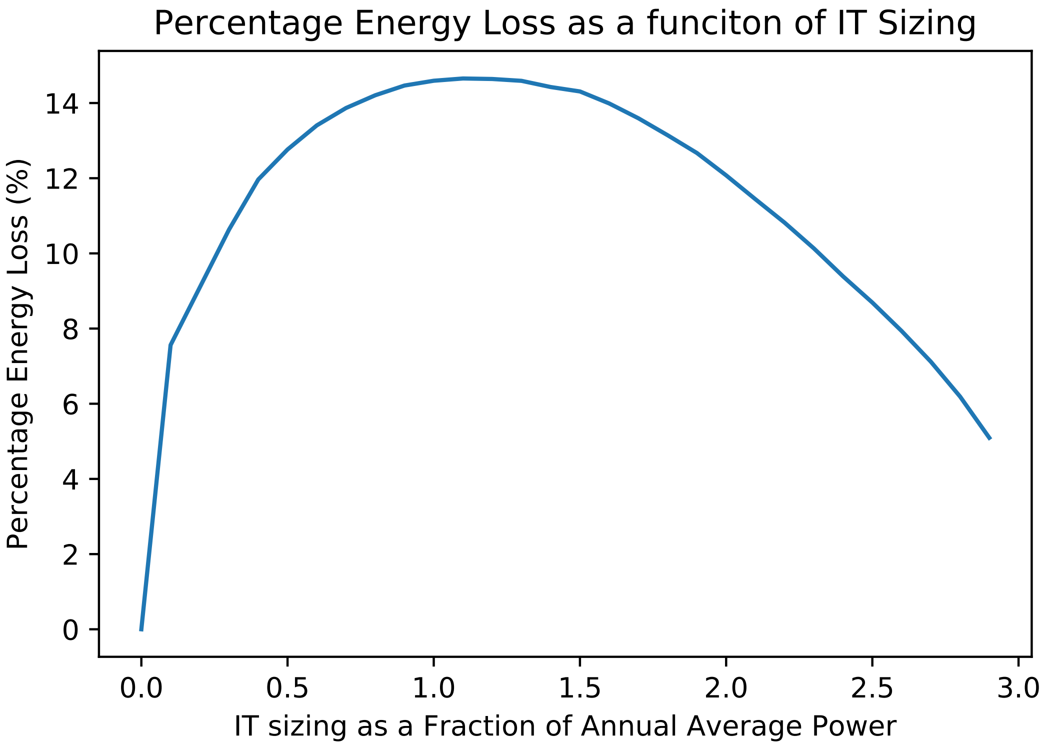


Figure 9
Figure depicts the parabolic relationship between PowerBox IT sizing and the percentage energy loss using data from our site in the Western Isles. The linear approximation (plateau) coincides with IT sized at the annual average power of the turbine.

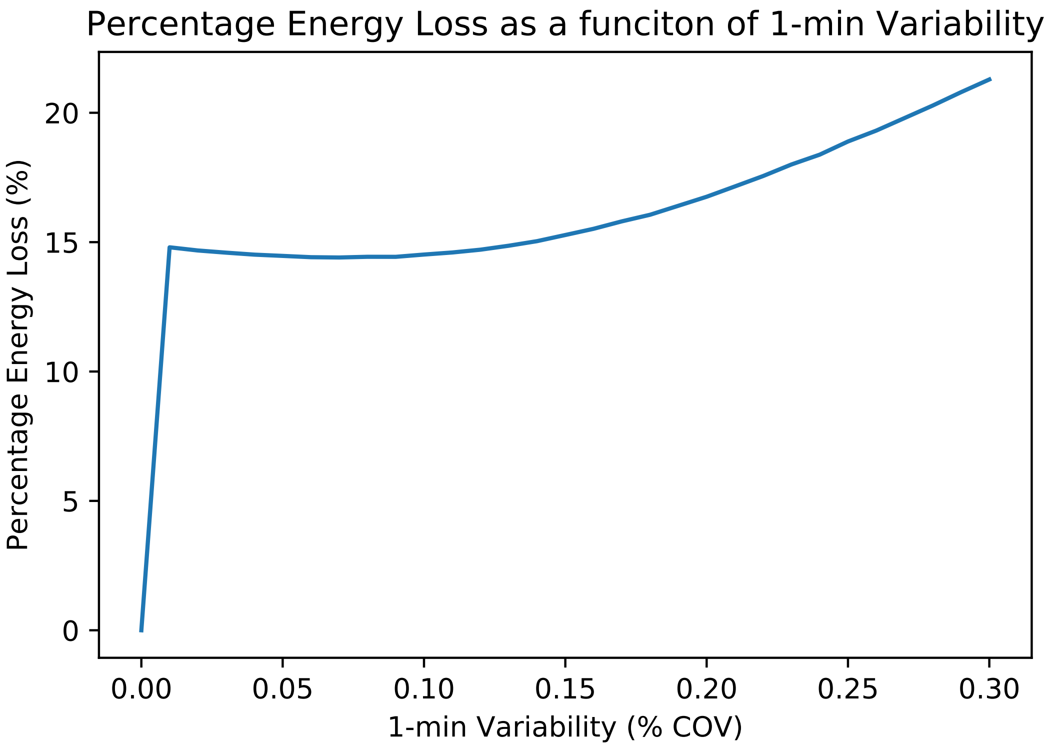


Figure 10: The energy loss surprisingly shows a weak dependence on the coefficient of variance of the 1-minute wind variability.

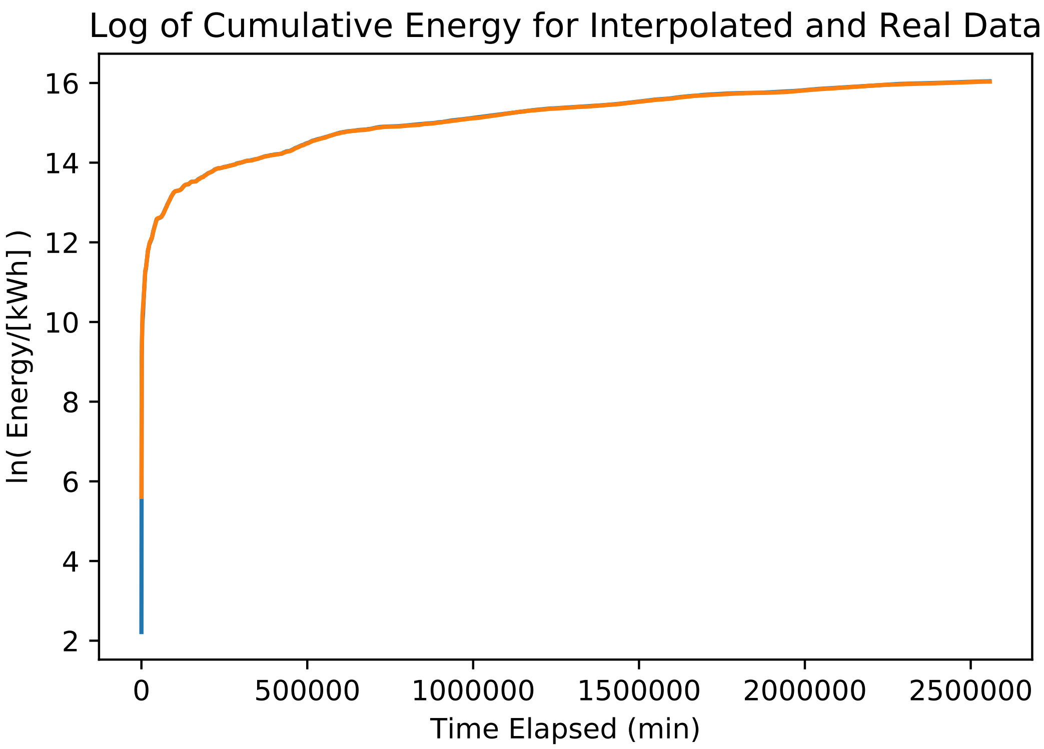


Figure 11: Energy is conserved to within 0.82% after 5 years' worth of 30-minute data is interpolated.

4 Evaluation

4.1 Approximate Uncertainty from Power Spectral Density

There is a wealth of research suggesting that the power spectral density and hence the variability is largest on the di-urnal scale of 12 hours [24]. With this in mind, for scenario 2 it is possible to estimate the uncertainty in the loss. This is done by first considering that σ cannot exceed the value that is obtained for 12 hours' aggregate data. This is trivially calculated and for the Western Isles of Scotland was found to be $\sigma = 0.28$. Using this value it is possible to calculate the loss as 21.4%. It is highly likely that the variability and hence the loss cannot exceed this value, therefore this represents the so-called "3-sigma" of the statistical distribution of loss. Therefore, the uncertainty in loss is a third of this value $\pm 7\%$. Despite scenario 1 having its loss calculated differently we expect the uncertainty to be less because the data is of higher frequency. Therefore, we can conservatively use the same value for scenario 1's uncertainty.

The effect of altering the IT sizing of the PowerBox was shown to have a weak effect around the average annual power (where losses are highest). This means that 1-minute variability is unlikely to affect the optimal IT sizing of the PowerBox. This is because the optimal sizing of the PowerBox usually resides in the range of $80\% < \text{sizing} < 200\%$. The loss was also shown to have a weak relationship to the variability of wind for 30-minute data, this unfortunately means that the loss (compared to 30-minute data) will likely not drop much below <10%. However, conversely it reduces the uncertainty of the methodology outlined due to site-specific variability.

When fitting a distribution, errors can come in the form of either parameter estimation or frequency measurement [17]. In future methods we would wish to relax the Gaussian approximation and try other more extreme distributions. This would mitigate distribution selection error. It would also make the method more robust to outliers. Parameter fitting errors are likely to be negligible in comparison to the 7% uncertainty outlined above.

Furthermore, in later versions it would be preferable to make use of the entirety of the transition matrix and have the random distributions skew and spread adjusted to the current power level. This would improve the conservation of energy which at present is violated because the supposed symmetrical Gaussian is skewed by the truncating limits of 0 and 100%. This is because the truncation combines a mean-conserving contraction combined with a mean-altering rigid shift. Therefore, the variance of the truncated distribution is smaller than the original normal distribution's. Additionally, because the power distribution is not symmetric around half the rated capacity, we have more points below 50% capacity where the Gaussian is skewed upwards. Therefore, these effects cause energy conservation to be violated (by 0.82%) by the current interpolation method.

5 Conclusions

The methodology outlined in the present work estimates losses due to variability at scales smaller than 30-minutes. The method will require validation before the uncertainties in it can be quantified properly. However, the average loss percentage from scenarios when the max IT load of the PowerBox equals the average annual power generated was estimated. Scenario 1, taking data from one of our sites in the Northern Isles of Scotland, had a loss of approximately $(4.5 \pm 7)\%$. However, Scenario 2 (using data from a site in the Western Isles of Scotland) had a loss of $(14 \pm 7)\%$. It seems with high certainty that losses due to 1-minute variability will be $< 21\%$. A large reduction in loss between the two scenarios is attributed to the fact that scenario 1 had data of 10-minute resolution. It seems therefore that for 10-minute SCADA data we can with high levels of probability say that this 21 % threshold should be much lower, of the order of 10%.

6 Truncated Gaussian: Back of the Envelope Method

By applying the identically independent distribution (i.i.d) approximation to data we can get an analytical estimate for the loss. The independence of points assumption is non-conservative because it reduces the likelihood of extreme events. However, we chose the distribution from the transition matrix with the highest spread in an attempt to be as conservative as possible. We then calculate the expectation value of the normal distribution centred on the average power and truncated by the max IT load. When running this analysis it is convenient to work in percentages of the rated capacity of the turbine. We can then calculate the expected magnitude of power beyond the IT load for any given average 30-minute power. However, again in an attempt to be conservative we take the worst case regime where the loss is greatest and the average 30-minute power is equal to the max IT load. We are then able to calculate the loss in this scenario and can say that the loss in general should be $\leq$ to this value if the initial i.i.d assumption partially holds on the 30-minute time scale, because in most cases the power of the turbine should be further away from the IT sizing and because of the exponential damping in the truncated Gaussian the loss should therefore decrease.

For each row in the transition matrix (figure 6) we transform coordinates so that the Gaussian is centred on zero, we can then calculate standard deviation of the distribtuion as:

$$\sigma = \sqrt{\frac{\Sigma f x^2}{\Sigma f}} \tag{7}$$

where f is the probability of measurement x (in new coordinates), and we have made use of the fact that our coordinate shift sets the mean $\mu = 0$. This result is further simplified by the fact that $\Sigma f = 1$ for a normal distribution.

We can now make use of general results for a truncated Gaussian.

The expectation value of a truncated Gaussian is given by:

$$\alpha = \frac{|p_{IT} - \bar{p}|}{\sigma p_{ratedcapacity}} \tag{8}$$

$$E < x >= \sigma \frac{\phi(\alpha)}{\Phi(\alpha)} \tag{9}$$

where the functions ϕ and Φ are defined as:

$$\phi(\alpha) = \frac{1}{\sqrt{2\pi}} \exp -\frac{1}{2}\alpha^2 \tag{10}$$

$$\Phi(\alpha) = \frac{1}{2}(1 + \text{erf}(\frac{\alpha}{\sqrt{2}})) \tag{11}$$

where erf(x) is the well known error function.

This expectation value is the most likely variation above the IT sizing in % of the rated capacity. Therefore to convert this to the % loss of the useable energy we need only divide by the load factor:

$$\text{loss} = \frac{\sigma}{\text{load factor}} \frac{\phi(\alpha)}{\Phi(\alpha)} \tag{12}$$

If we take the limit of $p_{IT} = \bar{p}$ which is the worst case for loss then this result simplifies to:

$$\text{loss} = \frac{\sigma}{\text{load factor}} \sqrt{\frac{2}{\pi}} \tag{13}$$

When evaluated based on the transition table in figure 6, we get a loss of approximately 2.4% multiplied by one over the load factor of a site:

$$\text{loss} = \frac{1}{\text{load factor}(2.4\%)} \tag{14}$$

for most sites the load factor is between 20 and 40 % and so this gives a range of loss between 6-12 %. Therefore, because this range agrees well with the results of our methodology, despite the rough and ready approximation made along the way we see that this value offers a good indicator of loss.

7 Further Work

7.1 Extreme Value Theory

The Gaussian distribution is known to underestimate the occurrence of extreme events when fitting with small sample sizes. As a first “quick fix” we could replace the Gaussian for a Weibull distribution. The Weibull is both a limiting case of the extreme value theory and the statistical distribution most closely resembling the distribution of wind speeds at a site generally.

7.2 Kaimal Filters

A method employing Kaimal filters to white noise in order to estimate turbulence effects is outlined in Gavriluta et al. (2012) [5]. This method should be investigated in the future to see if the additional complexity garners any better results.

7.3 Markov & Machine Learning Approaches

As we gather more data, machine learning and Markov chain models might become worth considering [9]. However, it is possible that by this point we will have outgrown the problem: when we have enough data to employ such techniques, we can calculate the losses directly.

Appendices

A Error Calculations

A.1 Validating the Method: Calculating the *Actual* Error

The methods outlined should be used to calculate losses and then compared with the losses calculated empirically on a 30-minute basis. The medians of the losses should also be compared to see if they are a similar and a frequency distribution fitted to the losses to estimate the 95% confidence level and compare the nature of the simulations distribution to that of the actual distribution.

When empirically calculating the loss E_l the following formula should be used:

$$E_l = \left\{ \begin{array}{ll} t(p - p_{IT}), & \text{for } \bar{p} \leq p_{IT} < p \\ t(p_{IT} - p), & \text{for } p < p_{IT} < \bar{p} \end{array} \right\} \tag{15}$$

where t is the time spacing between measurements and the other symbols are as defined earlier in the report.

References

[1] T. Ishihara A. Yamaguchi, M. B. Solomon. An effect of the averaging time on maximum mean wind speeds during tropical cyclone. 2012.

[2] Jay Apt. The spectrum of power from wind turbines. *Journal of Power Sources*, 169(2):369–374, 2007.

[3] Abul Kalam Azad, Mohammad Golam Rasul, and Talal Yusaf. Statistical diagnosis of the best weibull methods for wind power assessment for agricultural applications. *Energies*, 7(5):3056–3085, 2014.

[4] Radian Belu and Darko Koracin. Statistical and Spectral Analysis of Wind Characteristics Relevant to Wind Energy Assessment Using Tower Measurements in Complex Terrain. *Journal of Wind Energy*, 2013:1–12, 2013.

[5] I. Mosincat C. Citro I. Candela P. Rodriguez C. Gavriluta, S. Spataru. Complete methodology on generating realistic wind speed profiles based on measurements. 2012.

[6] C. Kluppelberg C. Steinkohl, R. A. Davis. Extreme value analysis of multivariate high frequency wind speed data. *Journal of Statistical Theory and Practice*, 2010.

[7] P Ceppi, C Appenzeller, and C Appenzeller. Extreme Value Analysis of Wind Speed Observations over Switzerland. 3.

[8] R. M. Endlich, R. C. Singleton, and J. W. Kaufman. Spectral Analysis of Detailed Vertical Wind Speed Profiles, 1969.

[9] Lee Fawcett and David Walshaw. Markov chain models for extreme wind speeds. *Environmetrics*, 17(8):795–809, 2006.

[10] Petra Friederichs, Martin Göber, Sabrina Bentzien, Anne Lenz, and Rebekka Krampitz. A probabilistic analysis of wind gusts using extreme value statistics. *Meteorologische Zeitschrift*, 18(6):615–629, 2009.

[11] Emilio Gómez-Lázaro, María C. Bueso, Mathieu Kessler, Sergio Martín-Martínez, Jie Zhang, Bri Mathias Hodge, and Angel Molina-García. Probability density function characterization for aggregated large-scalewind power based on Weibull mixtures. *Energies*, 9(2):1–15, 2016.

[12] J. K. Lundquist J. C. Y. Lee, M. J. Fields. Assessing variability of wind speed: comparison and validation of 27 methodologies. *Wind Energy Science*, 3(2):845–868, 2018.

[13] Y. Wan B. Kirby J. Cadogan, M. Milligan. Short-term output variations in wind farms — implications for ancillary services in the united states. 2000.

[14] Alexander Kislov and Tatyana Matveeva. An Extreme Value Analysis of Wind Speed over the European and Siberian Parts of Arctic Region. *Atmospheric and Climate Sciences*, 06(02):205–223, 2016.

[15] Joseph C. Y. Lee, M. Jason Fields, and Julie K. Lundquist. Assessing variability of wind speed: comparison and validation of 27 methodologies. *Wind Energy Science*, 3(2):845–868, 2018.

[16] Y. Cui Q. Yang B. Huang M. Yang, L. Zhang. The impact of wind field spatial heterogeneity and variability on short-term wind power forecast errors. *Journal of Renewable Sustainable Energy*, 11, 2019.

[17] A. Mata. Parameter uncertainty for extreme value distributions. 2000.

[18] Ian Moore. *Inertial Response from Wind Turbines*. PhD thesis, 2012.

[19] E. Muljadi, V. Gevorgian, M. Singh, and S. Santoso. Understanding inertial and frequency response of wind power plants. *PEMWA 2012 - 2012 IEEE Power Electronics and Machines in Wind Applications*, (July), 2012.

[20] J. P. Palutikof, B. B. Brabson, D. H. Lister, and S. T. Adcock. A review of methods to calculate extreme wind speeds. *Meteorological Applications*, 6(2):119–132, 1999.

[21] Stephen Rose and Jay Apt. Quantifying sources of uncertainty in reanalysis derived wind speed. *Renewable Energy*, 94:157–165, 2016.

[22] Poul Sørensen, Nicolaos Antonio Cutululis, Antonio Vigueras-Rodríguez, Henrik Madsen, Pierre Pinson, Leo E. Jensen, Jesper Hjerrild, and Martin Donovan. Modelling of power fluctuations from large offshore wind farms. *Wind Energy*, 11(1):29–43, 2008.

[23] Christina Steinkohl, Richard A. Davis, and Claudia Klüppelberg. Extreme value analysis of multivariate high-frequency wind speed data. *Journal of Statistical Theory and Practice*, 7(1):73–94, 2013.

[24] Isaac Van der Hoven. Power spectrum of horizontal wind speed in the frequency range from 0.0007 to 900 cycles per hour. *Journal of Meteorology*, 14(2):160–164, 1957.

[25] Erik Vanem. Uncertainties in extreme value modelling of wave data in a climate change perspective. *Journal of Ocean Engineering and Marine Energy*, 1(4):339–359, 2015.

[26] Simon Watson. Quantifying the variability of wind energy. *Wiley Interdisciplinary Reviews: Energy and Environment*, 3(4):330–342, 2014.