A Diminishing Influence Running Mean (DIRM) Model for Instantaneous Power Output of a Bergey Excel-s

by
Mel Tyree, Professor
Department of Renewable Resources
University of Alberta, Edmonton, Canada T6G 0C6

Introduction

During windy events, the Bergey Excel-s turbine made by Bergey Windpower Corp (BWC) tends to spend lots of time offline not producing power; see SWIEP reports R#40, R#41, R#42 and R#47, which can be downloaded from: http://www.ualberta.ca/~mtyree/SWIEP/Publications.html The Excel-s turbine sends power to a GridTek10 inverter that converts the variable-voltage, variable-frequency power produced by the turbine to 240 V AC 60 Hz power compatible with power requirements of home appliances. But the GridTek10 is mismatched to the Excel-s generator. The generator is too powerful for the inverter so the inverter has to shut down when it becomes too windy. Other turbine/inverters are better matched, for example see R#43 as an example of a system that works well for winds up to 20 m/s. A time-series measurement that illustrates these events for the BWC system is shown in Fig. 1 below.

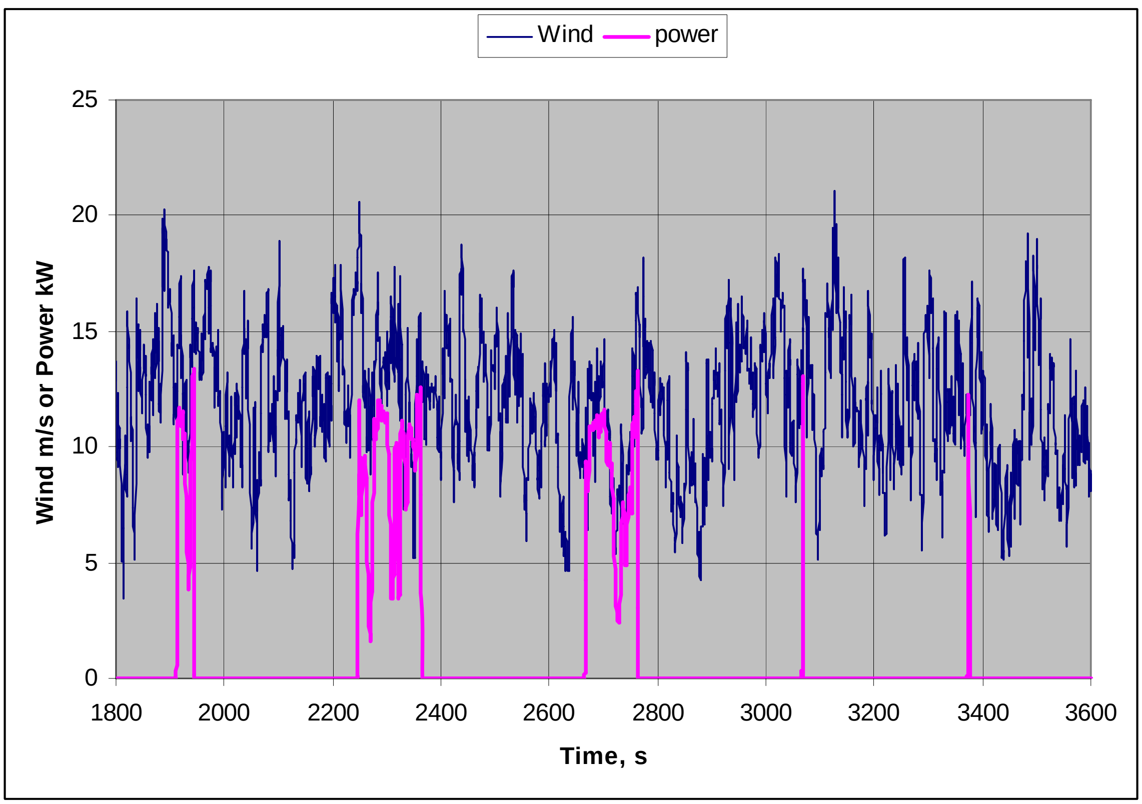


Figure 1: Wind speed (m/s) and power output (kW) was sampled every second during a wind-storm and the second half of 1-h of data is shown above. The inverter was frequently overloaded and spent most of its time offline. Generally speaking, the 'rule' is that if the inverter is online and the wind speed measured at the anemometer is above 17.5 m/s for more than 3-s then the inverter shuts down. The rule isn't perfect because the airfoils do not experience exactly the same wind regime as the anemometer at the same instant in time.

Generally speaking, whenever the wind speed exceeds about 17.5 m/s for more than 3 s, the inverter goes offline for 5-minutes then reconnects. But most of the time, when the inverter is off, the wind speed is below 17.5 m/s. The purpose of this report is to show how I developed a model to predict 1-s power readings from the wind-speed during times when the GridTek10 is offline. Michael Klemen strongly suggested that I give lots of details to people interested in small wind power systems, so they can be sure I am using appropriate and accurate methods. Hence, Michael's opinion has motivated me to write this report.

Methods

The methods I use for measuring wind-speed and power output have been covered in previous reports (R#48 and R#19 respectively). Briefly, I can measure wind-speed using an NRG cup anemometer for velocities between 0.6 and 37 m/s with a resolution of 0.01% to 0.1% and with an ultimate accuracy of about 1%, based on a factory calibration. I can measure power with a resolution of 0.1% and an accuracy of about 0.3%. Power and wind speeds are logged at 1 Hz (once per second). Wind speeds were measured on the same tower that supports the turbine at a height of 102 ft (31 m) and a distance of 8 ft (2.4 m) from the tower; the anemometer was supported by a boom pointing WNW from the tower upwind of the most common wind direction. The turbine hub was about 20 ft (6.1 m) above the anemometer and about 6 ft downwind (1.8 m), for the most common wind direction.

I am NOT using IEC/AWEA standards for measuring a power curve; rather I am measuring a power correlation curve (PCC) that will permit me to predict lost-power production when the GridTek10 inverter is off line. The method is to log many days of data of wind speed, u, and power production, P, for wind speeds between 0 and 20 m/s. From that I can obtain a function relationship $P = f(u)$, where $f(u)$ is a polynomial equation or any other equation that fits the data accurately. I obtain the PCC in 'bins' as in the IEC standard for measuring power curves, i.e., I compute 1-minute averages (60 1 Hz samples) of wind speed and power and compute means of each, then put them into wind-speed bins of 0.5 m/s increment. The average bin velocity is then plotted on the x-axis versus average bin power on the y-axis.

The usual approach (simple approach) would be to use $P = f(u)$ to compute 1-min average power losses from 1-min average wind speeds for periods when the GridTek10 inverter is offline, i.e., for periods when the pink line in Fig. 1 is at P = 0 in Fig. 1. However, I wanted to see if I could do something a little more elegant, i.e., come up with a reliable way to estimate 'instantaneous' P_{model} (1-s values) from some kind of 'mean' wind speed, $\bar{u}$, for a few seconds prior to the measured power $P_{measured}$. I am interested in the response time of turbines to rapidly changing wind speeds, because wind speeds ALWAYS change rapidly as shown in Fig. 1. So I wanted to see if I could model for that response time.

Wind-speed rises and falls in a time series and the turbine has a 'mechanical memory' of past wind events. You can think of a wind-gust as speeding up the rate of spinning, and even if the wind falls to a lower constant speed for the next few seconds it will take a while for the turbine to slow down again. The memory of past wind is stored as kinetic energy in the rotating mass of the blades and generator. It takes time at any given wind speed to store up the kinetic energy and hence increase the rotational speed.

When the wind speed falls it takes time for this kinetic energy to be lost and hence for the rotational speed to slow down. The rate of loss of kinetic energy will be determined by friction and the rate at which the GridTek10 inverter exports energy to the grid. If I can devise a model that matches $P_{measured}$ at any instant in time to P_{model} based on the recent history of the wind-speed time-series then I will have some rough idea of how quickly a wind turbine responds to wind. Hence I have set a more difficult task for me than if I just used the normal approach, but I thought I would learn something from the exercise.

To accomplish this objective, I need some quantitative way of evaluating how good my model is. This is done through time-series analysis of residuals and root-mean-square (RMS) deviations. For the specific case of wind power, our time series is a repeated series of measurements of time, *t*, plus the independent variable, *u*, and a dependent power variable $P_{measured}$. You then try to devise a model to predict power, P_{model}. A model is perfect if $P_{measured}$ always equals P_{model}, but models are never perfect. There are always residuals (errors). So the residual, *r*, is defined as $r = (P_{measured} - P_{model})$ at any given time, *t*. Lets say we measure 1 Hz values for 1 hour, so that is 3600 time-series readings and that *P* and *u* change a lot in that hour. A good model will have a sum of $r = 0$, i.e., P_{model} will be higher than $P_{measured}$ about as often as it is lower. We also want a model with residuals that are either a constant fraction of *P* or at least constant as *P* varies. We also want the absolute size of *r* to be as small as possible. That is where the *RMS* error comes in; E_{RMS} (root mean square Error) is defined as:

$$E_{RMS} = \sqrt{\frac{\sum r^2}{N}} = \sqrt{\frac{\sum (P_{measured} - P_{model})^2}{N}} \qquad (1)$$

Where $\sum r^2$ is the sum of the residuals squared for *N* readings ($N = 3600$ if we do 1 Hz readings for 1 h). The *RMS* is a kind of ‘average’ residual for a time series. Statistically the form of the equation is like a standard deviation, SD, where P_{model} takes the place of mean. A good model will have the lowest possible E_{RMS}-value.

Results and Analysis

Figure 2 shows my PCC measured from 12 days of data. I filtered out two 5 min periods when the GridTek10 was offline. The data are plotted as bin means of power versus bin means of wind-speed with error bars. The error bars are generally smaller than the points. The bars give SEM (standard error of the mean) on N values, which statistically is defined as

$$\text{SEM} = \frac{SD}{\sqrt{N-1}} \qquad (2)$$

Sometimes you see power curves published with SD values, but SEM is useful too. Here is the difference: Our objective is to get a reliable mean value, SD measures the variation of individual values in the mean. From a probability point of view, SD tells you that if you go out and measure ONE more sample there is a 63% chance this next value will fall within the SD error bars and a 37% chance it will fall outside. But SD does not tell you how accurate the mean is; generally we want to know if we have a ‘good’ mean, e.g., if we measure it again what is the chance that we will get the same value again? The SEM gives the probability of that. If we measure the same mean again with the same number of samples, then there is a 63% chance the mean will fall within the SEM error bars and a 37% chance it will fall out. Hence SEM is more valuable than SD, if you are going to report only one or the other. For example, let us say you want to try out some new airfoils on your old turbine and you compute a new power curve with your new blades. If you plot the two power curves on the same graph with means and SEM-error bars and the error bars are 'far apart' it means you really have made an improvement. There is a statistical test to work out the probability the differences are significant, but just glancing at the SEM-error bars gives you a good clue. I hope people will start reporting power curves with SEMs or SDs AND N so SEM can be calculated. It is a useful statistic for R&D.

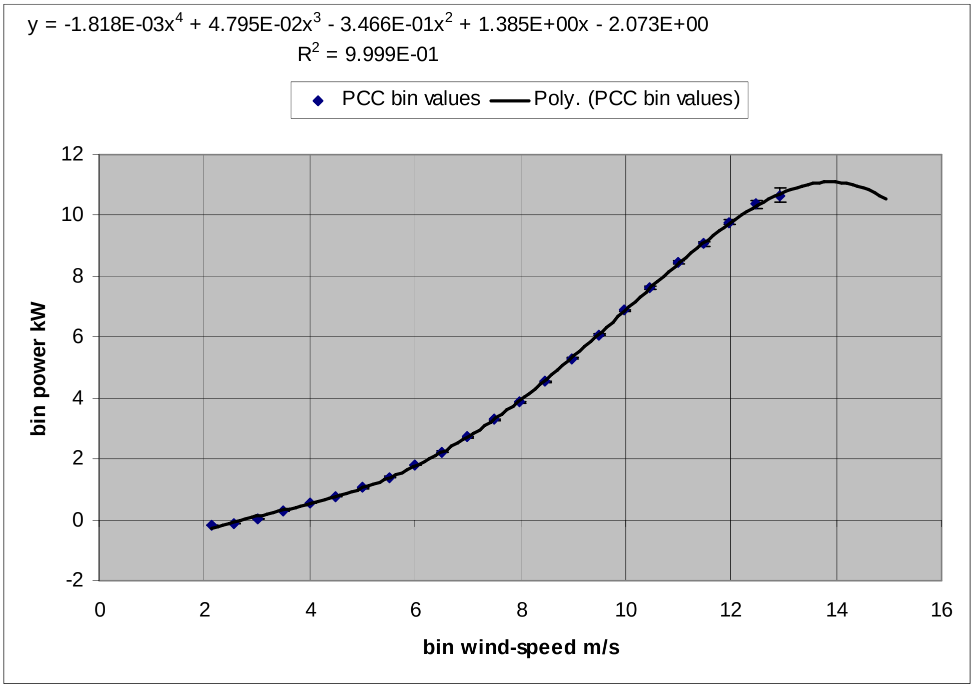


Figure 2: A power correlation curve measured on my BWC Excel-s turbine. Error bars are SEM values (see text for details).

I used the polynomial equation in Fig. 2 as my initial model to compute P_{model} from wind data. The main limitation of polynomial equations is that you cannot extrapolate beyond the measured data and except to get the right answer. I know from other data that the extrapolated line beyond the last point in at 13 m/s bin wind speed underestimates the actual 'instantaneous' performance of the GridTek10. If you collect 8-s bins then power can peak above 13 kW. So the equation above where $P_{model} = y$ and $\bar{u} = x$ is for 'starters' to examine how to formulate a model.

Again, the purpose of the model is to predict 'instantaneous' power ($P_{measured}$ from 1 Hz readings) from some kind of mean wind speed ($\bar{u}$) for a few seconds prior to the reading. Can this be done? The simplest kind of mean is a running mean of the previous k 1 Hz readings:

$$\bar{u}_{rm} = \frac{1}{k}\sum_{i=1}^{k} u_i \qquad (3)$$

where $\bar{u}_{rm}$ = the running mean of u_1 .. u_k where u_1 is the wind velocity at the same second that $P_{measured}$ is recorded and u_k is the wind speed k seconds earlier. And the model is:

$$P_{model} = a_4\bar{u}^4 + a_3\bar{u}^3 + a_2\bar{u}^2 + a_1\bar{u} + a_0 \qquad (4)$$

where a_0 .. a_4 are the coefficients in the equation in Fig. 2. What I have to do is optimize the model, i.e., compute E_{RMS} for different numbers of seconds ($t=k$) and figure out which value gives the smallest E_{RMS}. I did this for 1-hour worth of data collected around 2 PM on 1 Jan 2009. Figure 3 shows the results.

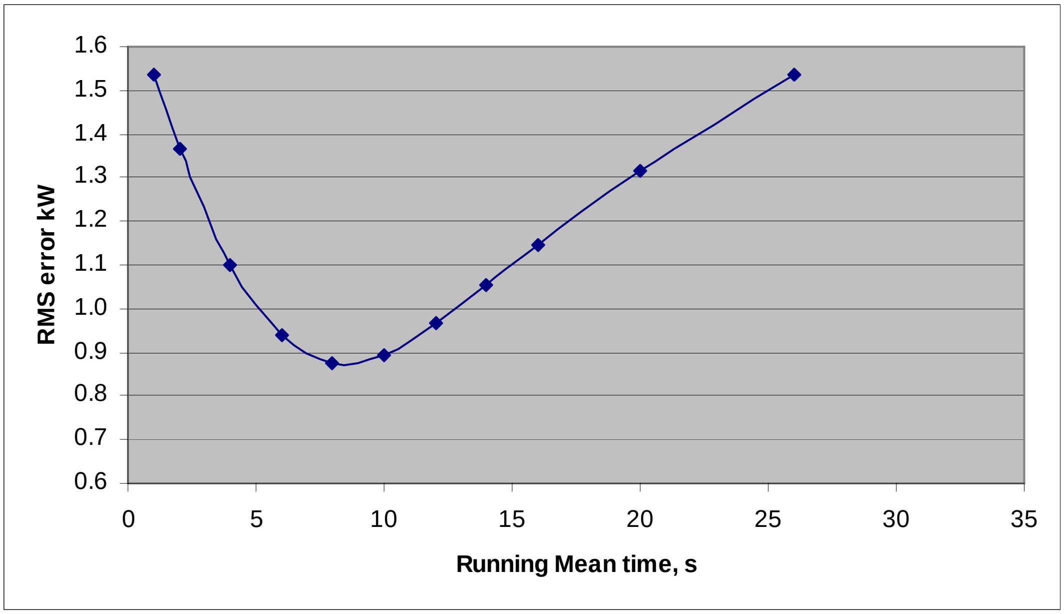


Figure 3: E_{RMS} versus the number of values used in the running mean. The minimum E_{RMS} is reached when I take the mean wind speed from the previous 8 s of data.

The curve in Fig. 3 means that if I compute the recent mean wind speed $\overline{u}$ from the previous 8 s of data, I will have the least error in computing the power at any given instant using Eq. (4) as my model. What does this mean in terms of a time series? The minimum E_{RMS} value is about ± 0.88 kW and what this means is that I can use the model to estimate 'instantaneous' kW values within ± 0.88 kW. Can we do better?

My idea was that a wind gust 8 s ago would have much less influence on $P_{meaured}$ currently than a wind gust 1 or 2 s ago. So what we need is a Diminishing Influence Running Mean or what is more commonly called a weighted running mean in which the weight applied to an old wind gust is less than the current wind gust. So the next thing I tired was to weight the mean using an exponential decay using:

$$\overline{u}_{dirm} = \frac{\sum w_i u_i}{\sum w_i} \tag{5}$$

where w_i is the weight applied = a number ≤ 1. The weighting coefficient I tried was from an exponential decay equation $w_i = \exp(-t/\tau)$ where t = the time *prior to* the last power reading and τ = a decay time constant. I found I could improve the fit slightly using $\tau = 5$, which results in the w_i coefficients in Fig. 5.

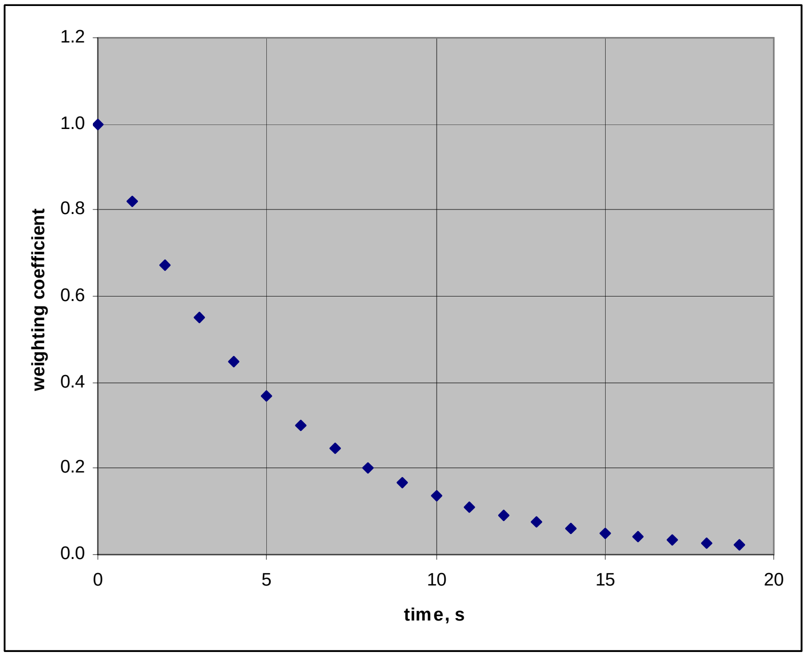


Figure 5: The weighting coefficients (w_i) used in my DIRM model = diminishing influence running mean model based on a mean of the previous 20 s of data.

I found that the DIRM model slightly reduced the E_{RMS} error from 0.88 to 0.85. What does this mean in terms of actual data? Appendix 1 gives the time series of the hour under analysis so that you can see for yourself. A close examination of the data in Appendix 1 reveals that for long periods of time the agreement between $P_{measured}$ and P_{model} is quite good and there are other periods of rather poor agreement. Why? I really do not know, but my guess is that the wind regime near my cup anemometer is not always the same as near my airfoils 7 m away. Wind speeds not only change rapidly with time (Fig. 1) but also between different locations within a range of 0 to 20 m. I think if I had measured wind speed from a met tower 2 to 4 blade-diameters away (14 to 28 m) the agreement might have been even worse mainly because of the time lag between wind reaching the cup and the blades which could be several seconds depending on the wind speed.

How robust is this model? To check robustness I ran the model on 10 different 1-h records of data. The summary results are shown in Table 1. Based on 10 h of model simulation, I can predict the actual energy produced to within 0.64% (= the mean error) of the measured production when the turbine/inverter is working properly. During windy events, the GridTek10 can be offline for anywhere from 5 min to 58 min per hour on Fault 1 codes (See R#46). Alternatively, sometimes the inverter can be offline for several hours until I perform a manual reset (Fault 2 R#46). I think this model can adequately predict energy loss during these fault conditions. Since 1 Sept until 31 Dec 2008 the inverter has been offline enough to account for 32% loss in energy production.

Table 1: Mean u = the mean wind speed for the hour, model RMS is the E_{RMS} for each hour. "Meas kWh" is the measured total kWh of energy production for the hours. "Pred kWh" is the hour-long predicted kWh value from the model. "Error %" is the percentage error between the measured and predicted energy production in the hour.

File #	Date/Time	mean u	model RMS	Meas kWh	Pred kWh	Error %
1478	1/9/2009 14:02	7.76	0.782	3.745	3.789	1.17%
1476	1/9/2009 12:02	8.75	0.856	5.073	5.178	2.07%
1462	1/8/2009 22:00	9.43	1.013	6.105	6.008	-1.60%
1455	1/8/2009 15:00	8.46	0.878	4.631	4.755	2.69%
1448	1/8/2009 8:00	7.87	0.809	3.885	4.020	3.46%
1330	1/3/2009 10:25	9.01	0.911	5.538	5.503	-0.63%
1328	1/3/2009 8:25	8.46	0.828	4.866	4.788	-1.61%
1275	1/1/2009 3:25	8.85	0.776	5.102	5.220	2.32%
1274	1/1/2009 2:35	9.07	0.847	5.538	5.568	0.55%
1273	1/1/2009 1:35	8.56	0.819	4.767	4.673	-1.98%

Given how fast the turbine responds to changes in wind speed and how fast the wind speeds changes, is there any reason to change the timing used for measuring power curves? I don't want to spend a lot of time on this issue, but I have looked into it and have decided the answer is 'no'. I have compiled PCC curves like those in Fig. 2 using time bins between 8 s and 90 s and the result is that the mean values are unchanged. The only advantage of shorter times in my case is that I can get the power curve up to 13 kW at wind speeds of 15 m/s without a Fault 1 condition occurring.

Conclusions

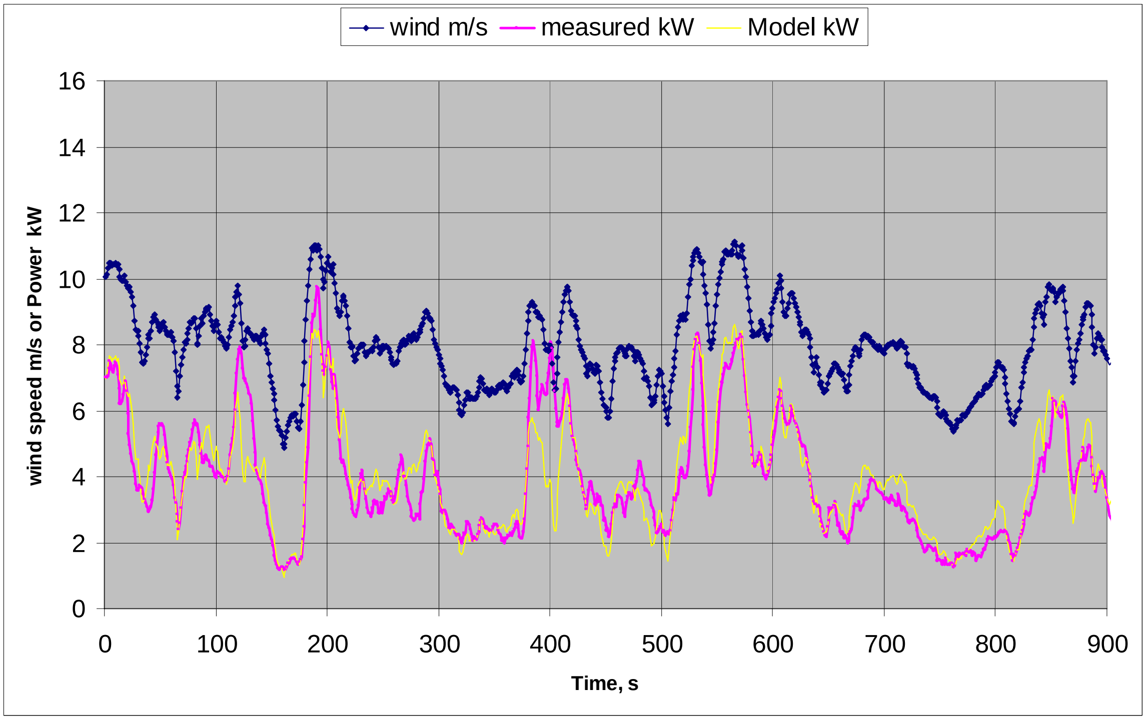
I think I have established that I have a reasonably good model for predicting lost energy production during periods when the GridTek10 inverter is overloaded and offline. I am not suggesting that my DIRM is better than the conventional way of estimating power production. The conventional way would be to measure mean wind speed for any given minute then use that average and a power curve (or PCC as in Fig. 2) to calculate mean power for that minute, $\overline{P}$, then power loss in kWh for the minute would equal $\overline{P}/60$. I think this method is about as accurate for long-term averages (= over periods of hours) as my DIRM model, but the advantage of the DIRM model is it gives me some background information on how fast the BWC Excel-s responds to changes in wind speed. I think it responds with a time constant near 5 s.

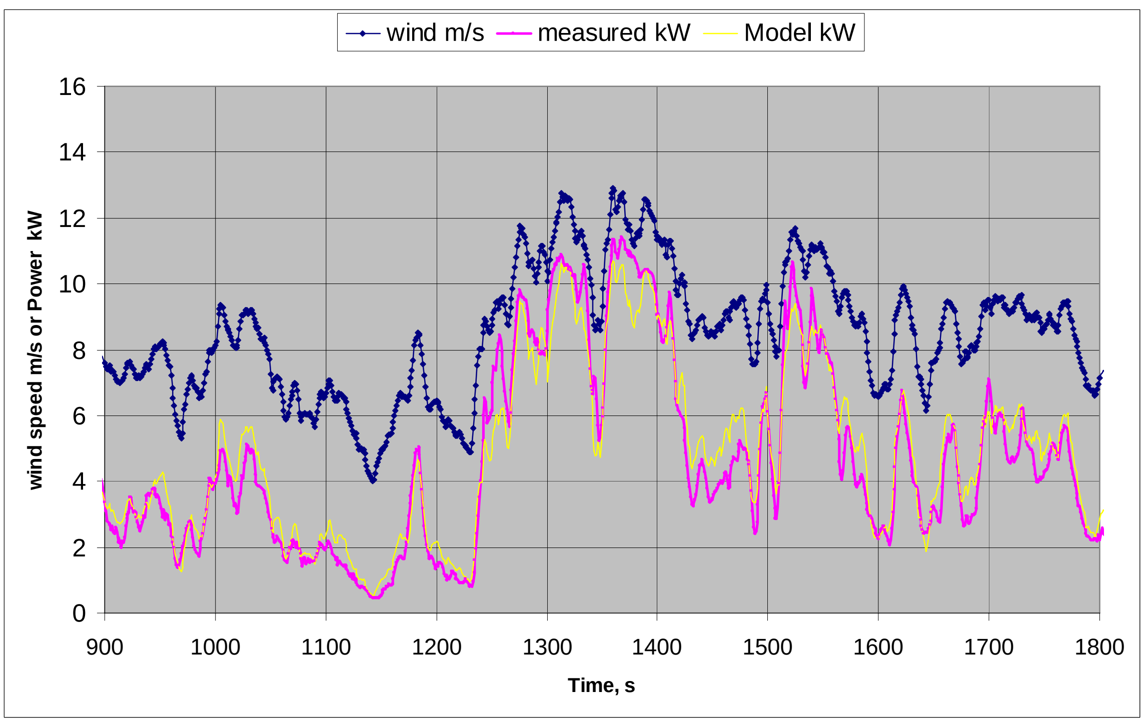
I welcome comments and suggestions for improvement to this and other SWIEP reports. I promise to correct mistakes promptly thru revised reports with acknowledgement to the person finding the mistake.

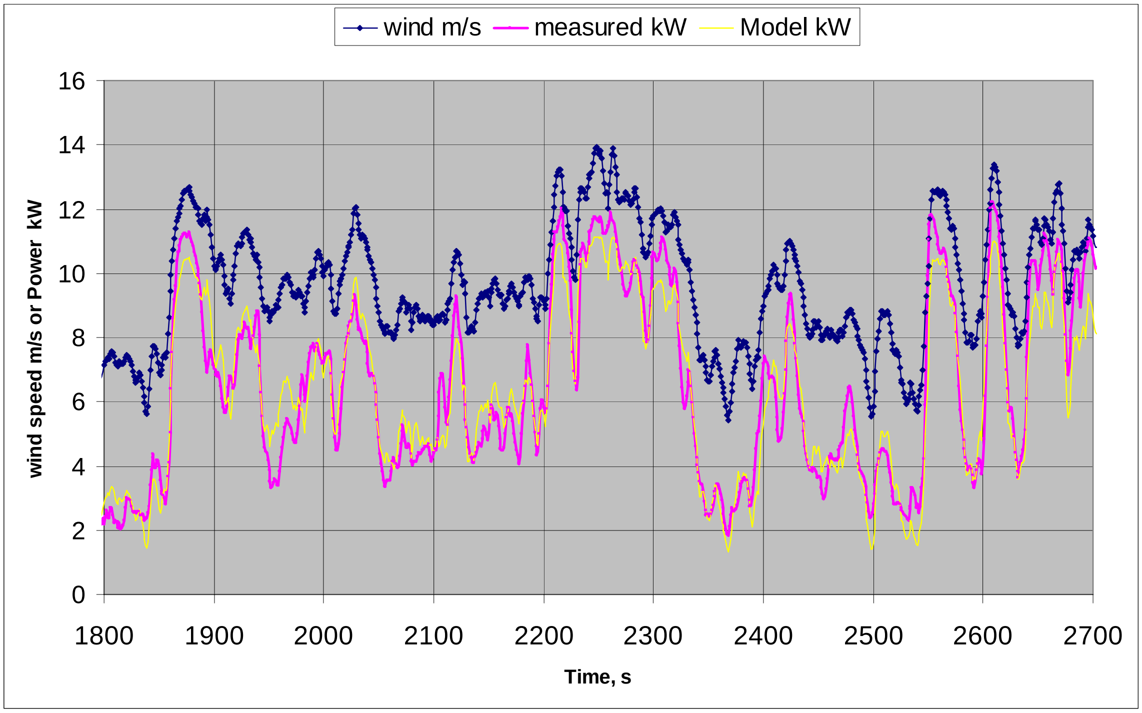
Mel Tyree

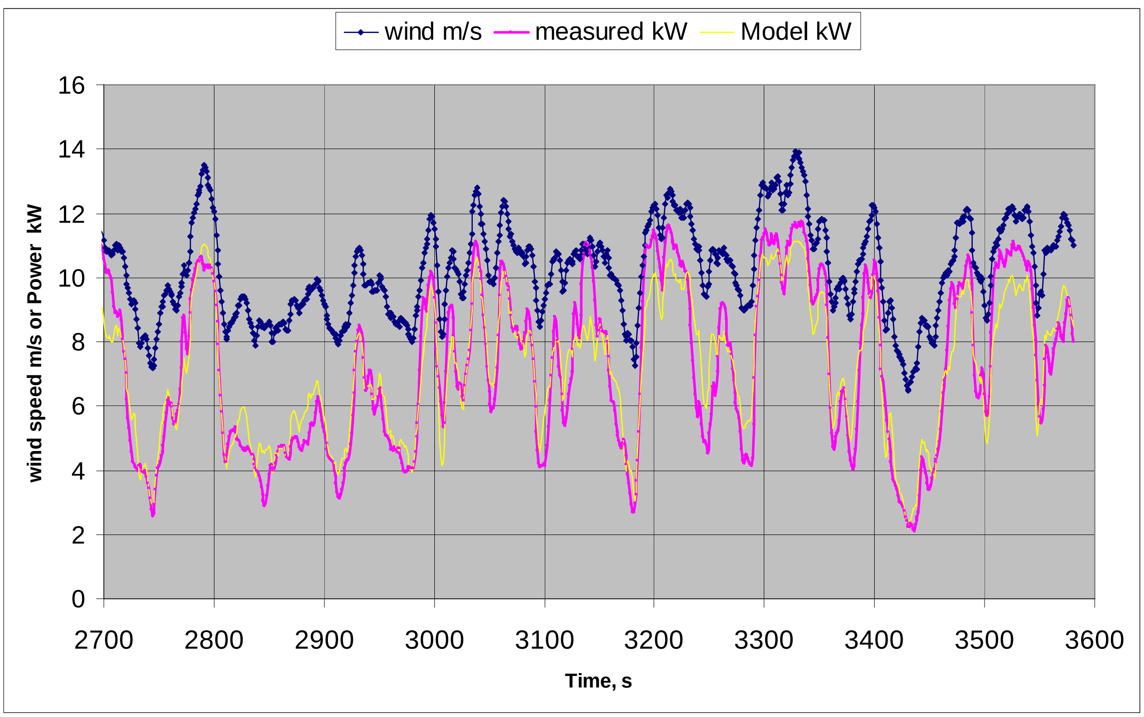
Appendix 1:

In the 4 graphs below wind speed is plotted as the running mean of 8 s of 1 Hz data just prior to the instantaneous measurement of power output measured on the GridTek10 inverter (pink line). The yellow line is the model value based on Eq. (5).






wind m/s
measured kW
Model kW
wind speed m/s or Power kW
0
2
4
6
8
10
12
14
16
1800
1900
2000
2100
2200
2300
2400
2500
2600
2700
Time, s


wind m/s
measured kW
Model kW
wind speed m/s or Power kW
0
2
4
6
8
10
12
14
16
2700
2800
2900
3000
3100
3200
3300
3400
3500
3600
Time, s