Research Towards Improved Models for Annual Production of Energy of Small Wind Turbines
Part 4
How Weibull Functions Affect Estimates.

by
Mel Tyree
Department of Renewable Resources
University of Alberta, Edmonton, Canada

Introduction

There is increasing evidence that the WindCad models (industry standard models) for estimation of Annual Production of Energy (Y_p) of small wind turbines consistently overestimates actual measured production (Y_m). The purpose of this series of opinion papers is to examine each element of theory and data involved in the estimation of Y_p in order to determine the possible cause(s) of the overestimation. Four previous SWIEP papers address this issue. The first is a sensitivity analysis of the WindCad model (R#22) and the second examines one aspect of C_v = Turbulence Intensity on Y_p (R#30). The third examines the impact of air density on the estimation of Y_p (R#33), and the fourth is Charles Ronse's analysis of the influence of turbulence on Y_p. See also (R#32) for an introduction on how models are formulated and validated. All reports can be downloaded from:
http://www.ualberta.ca/~mtyree/SWIEP/Publications.html

TIMAPE Part 4 examines the use of the Weibull Function to estimate Y_p. WindCad models (e.g., R#21) use an estimate of mean annual wind speed and a Weibull Function as the main elements to estimate Y_p. The Weibull Function is one in a class of probability density functions applied to wind speed distributions. On the y-axis is plotted the probability of finding wind speed at the wind speed value shown on the x-axis. Probability density functions have the property that the area under the curve must equal one, i.e., the sum of all probabilities over the x-axis range of wind speeds must equal one. See R#22 to see how all calculations are done. The version of the Weibull function that is usually used is:

$$\Gamma(v) = (b\frac{K}{\bar{v}})(b\frac{v}{\bar{v}})^{K-1}\exp(-[b\frac{v}{\bar{v}}]^{K}) \qquad \text{Eq.(1)}$$

where v = the binned wind speed, $\bar{v}$ = the mean annual wind speed at the proposed site of the wind turbine, b = a constant (shape factor) usually fixed at 0.89 but sometimes allowed to vary from site to site, and K = another constant that has K = 2 for inland sites, K = 3 for coastal sites, and K = 4 for island sites. The higher values of K make the distribution more symmetrical and low K values give distributions skewed toward higher velocities (See Fig 2 3 & 4 in R#22).

The computation of Y_p involves Eq(1) and a measured power curve $P(v)$ for a wind turbine. The power curve is basically a table of values of power output (kW) of the turbine's inverter versus binned values of wind speed, v. Usually power P is measured at 1 Hz (one sample per s) for 10 min averages that are stored together with average wind speed measured at 1 Hz for 10 min and averaged. The values are binned according to wind speed intervals of 0.5 or 1 m/s and averages taken of these binned values of P and 10-min-average wind speed to create $P(v)$. Y_p is computed from Eq. (2):

$$Y_p = 8766(1 - T_f)(1 - D_f)\sum_{bin} P(v)\Gamma(v)\,, \qquad \text{Eq(2)}$$

where the factor 8766 = the number of hr in a year, $(1\text{-}T_f)$ = a power de-rating for turbulence, $(1\text{-}D_f)$ = a power de-rating for air density. Although $P(v)$ can be accurately measured, $\Gamma(v)$ is an approximation of an actual probability density function.

The purpose of this opinion paper is to compare **measured** $\Gamma(v)$, measured from anemometers on towers, to **predicted** $\Gamma(v)$ based on Eq. (1). The purpose of this paper is to answer two questions:

1. How much random error is introduced into the estimate of Y_p based on using standard practices in WindCad models?
2. How much systematic error (if any) is introduced into the estimate of Y_p based on using standard practices in WindCad models?

The approach used is to compare Y_p calculated from measured $\Gamma(v)$ values to those computed from Weibull-derived $\Gamma(v)$ values.

Methods

The MTC (Massachusetts Technology Collaborative) has commissioned a large number of wind resource assessments in MA. The measurements have been carried out by the University of Massachusetts, Renewable Energy Research Laboratory (RERL) and these reports are free to the public and can be downloaded from:
http://www.ceere.org/rerl/rerl_resourcedata.html
I took five 1-year reports from this above web site. Data are in comma-separated-value (*.csv) format and I wrote a program to extract **measured** $\Gamma(v)$ data sets from these files. For every anemometer, data are published of 10 min means of v and standard deviation (SD) based on a 1 Hz sampling rate. Hence one can also extract turbulence intensity data which is defined:

$$TI = \frac{SD}{\bar{u}} \qquad \text{Eq.(3)}$$

where $\bar{u}$ = the 10-min mean v and SD = the standard deviation of the mean. Note that TI is statistically identical to C_v = coefficient of variation.

Small wind turbines tend to be mounted on towers 24 to 37 m tall (80 to 120 ft) and it is considered good practice to mount turbines at least 10 m above the average height of trees or buildings within 100 or 200 m of the turbine tower. RERL wind data are usually collected at 3 to 4 different heights at each study site, and I selected the two heights that bracketed the range of 24 to 37 m, whenever possible. The sites included in the study are show in Table 1 below:

Table 1	**Map Coordinates**		**Tower base**		
Site Name	**North**	**West**	**elevation,m**	**Topography**	**Description**
Barnstable, MA	41.6648	70.3045	16	U	Cape Cod
Falmouth, MA	41.606	70.621	42	SU, F	SW Cape Cod on small hill
Harwick, MA	41.6911	70.0639	16	U	Cape Cod
Mt Tom, MA	42.2498	72.6456	329	R, F	Western MA, on escarpment cliff to W
Savoy, MA	42.6034	72.9699	621	R, F	Western MA,

U = urban (mostly buildings & trees), SU = suburban (dispersed housing), R = rural, F = forested.

Results

The Mt. Tom site had the best agreement between the measured $\Gamma(v)$ and the Weibull-predicted $\Gamma(v)$ as shown in Figure 1a. This was in ‘inland’ site hence convention would suggest that K should be assigned a value of 2. But I also used the solver routine in Excel to obtain the best fit curve with an adjusted value of K = 2.19, see Figure 1b.

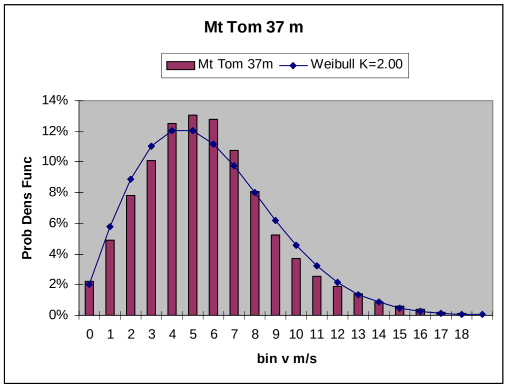

Fig 1a: Probability density functions: bars = measured; line+diamonds = conventional fit with K = 2.00.

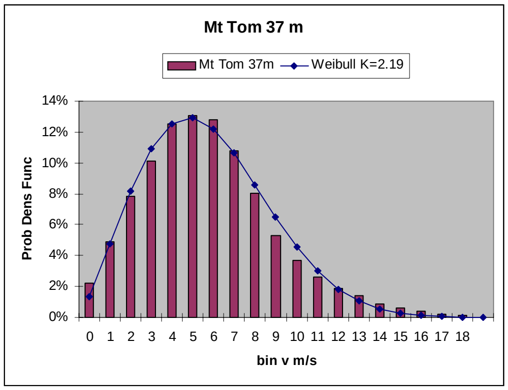

Fig 1b: Probability density functions: bars = measured; line+diamonds = best fit value of K.

Other sites had rather poor agreement between measured $\Gamma(v)$ and the Weibull-predicted $\Gamma(v)$ as shown in Figure 2a, where data are plotted from Harwich, MA. This was a Cape Cod site on a narrow peninsula, which might behave more like an island or coastal site so a best fit K was obtain with Excel’s solve. The best fit K (= 2.58) still does not give a great fit and would be the K-value associated with half way between inland and coastal.

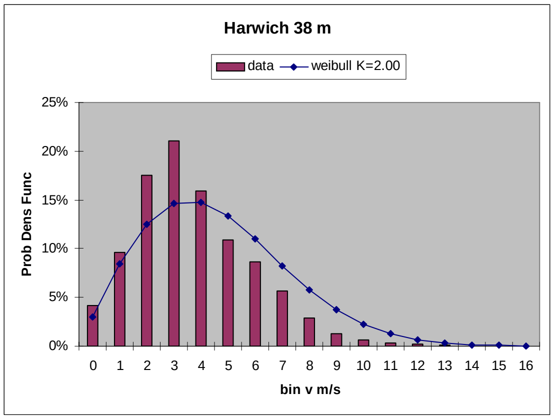

Fig 2a. Probability density functions: bars = measured; line+diamonds = conventional fit with K = 2.00

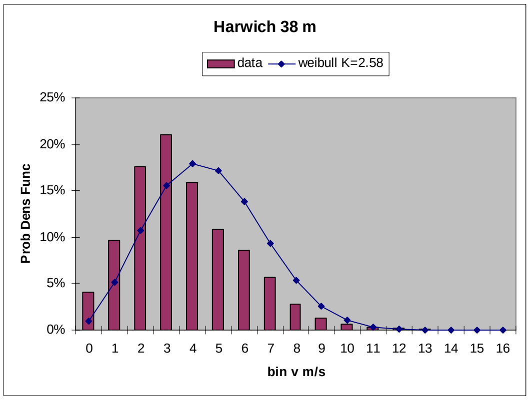

Fig 2b: Probability density functions: bars = measured; line+diamonds = best fit value of = 2.58.

No instructions are given in the WindCad spread sheets as to how close a turbine has to be to the shore to be considered ‘coastal K = 3’ nor is any information given on how small an island has to be to be an ‘island K = 4’. Hence I chose to use a default K = 2 for all sites initially and then later to obtain a best fit value of K to help decide when a turbine location is ‘coastal’ or ‘island’. My idea was that by comparing the best fit K values to Earth Google images of the anemometer sites, I might be able to figure out the ‘rules’. Unfortunately, 5 sites were not enough to discern the ‘rules’ if any. Best fit K values changed with anemometer height at any one site and changed between sites with no rhyme or reason. The range of best fit K values was from 1.9 to 3.07. Two of the lowest K values were at Savoy (inland west MA) and Barnstable (on Cap Cod), which both had K = 1.9. The highest value K = 3.07 was at SW Cape Cod (Falmouth MA). The Falmouth anemometer tower was 2.6 km from the coast versus 3.8 km for Barnstable. However, the Savoy tower was on a mountain top about 200 m above surrounding land, which might have made it behave like an ‘island’ in an inland site but appeared not to have that affect.

Table 2 (next page) summarizes the overall results of my initial study of 5 sites with two anemometer heights on each site. The ‘W ratio’ in the table gives the fractional disagreement between Y_p predicted from the Weibull fit and the Y_p calculated from actual data (= Probability Density Function derived from one-year of 10 min wind speed measurements). In the calculation of Y_p in Table 2, no correction was made for turbulence or air density (Eq. 2) because the primary purpose of the exercise was to compute the ratio. When the ratio is taken the factor = $8766(1-T_f)(1-D_f)$ in Eq. 2 is eliminated. The power curve used in Eq. 2 was that for the 10 kW Bergey as measured by NREL (see R#26). When using the default value of K = 2 then mean ratio was 1.139 ± 0.140 (mean ± SD) for the upper anemometers versus 1.080 ± 0.211 for the lower anemometers. The ratio (agreement) improved when best fit values of K were used; in this case the values were 0.991 ± 0.048 and 0.953 ± 0.084 for the upper and lower anemometers, respectively.

Discussion and Conclusions

The MTC in MA has promoted the installation of many small turbines in MA and has provided subsidies to the land-owners who installed them. Subsequently the MTC contracted the Cadmus Group to do a study of how well the measured annual energy production Y_m agreed with the predicted value, Y_p, i.e., the value that was predicted from $\bar{v}$ and Weibull functions. The Cadmus Report eliminated turbines with known electrical or mechanical problems and computed Y_m/Y_p for the remaining 19 units (R#17) and statistical analysis of these data (R#25) has revealed that Y_m/Y_p has a mean ± SD of 0.245 ± 0.127 and a $C_v = \pm 0.518$.

The purpose of this study in part was to see if errors in estimating Y_p from Weibull functions versus from real data might account for this low value of Y_m/Y_p. Table 2 reveals some random errors in the ratio and perhaps some systematic errors that could contribute to errors in Y_m/Y_p. The systematic errors, when K is assigned a default values = 2, are in the right direction but only account for a factor of 1.08 to 1.14 in overestimation of Y_p, whereas we need to account for a factor of 4. Also the random errors (SD = 0.05 to 0.2) are similar to the SD values computed from the Cadmus Report. I would expect that the SD on the ratio Y_m/Y_p to be determined by the annual variation in mean wind speed and by the errors associated with using the Weibull function.

Table 2. Summary of results: $\bar{v}$ = annual mean wind speed, h = height of anemometer above ground, shear coef = wind shear calculated between the high and low anemometer heights, TI = turbulence intensity, K = Weibull constant used to compute Y_p Weibull, K best fit = best fit value obtained by using Excel's solver routine, b = standard Weibull constant used in WindCad programs, Yp data = predicted annual energy production based on real data, Yp Weibull = Yp based on the Weibull fit, W ratio = ratio of Yp Weibull to Yp data.

Site Name	$\bar{v}$,m/s	**h, m**	**shear coef**	**TI**	**K**	**K best fit**	**b**	**Yp,data**	**Yp, Weibull**	**W ratio**
Barnstable, MA	5.22	39	0.304	0.200	2.00	2.37	0.89	8128	9821	1.208
Falmouth, MA	4.77	39	0.360	0.172	2.00	3.07	0.89	5724	7578	1.324
Harwick, MA	5.12	38	0.473	0.221	2.00	2.58	0.89	8027	9290	1.157
Mt Tom, MA	6.2	37	0.375	0.159	2.00	2.19	0.89	14996	15450	1.03
Savoy, MA	5.55	39	0.563	0.190	2.00	2.11	0.89	11932	11632	0.975
Barnstable, MA	5.22	39	0.304	0.200	2.37	2.37	0.89	8128	8695	1.07
Falmouth, MA	4.77	39	0.360	0.172	3.07	3.07	0.89	5724	5670	0.991
Harwick, MA	5.12	38	0.473	0.221	2.58	2.58	0.89	8027	7774	0.968
Mt Tom, MA	6.2	37	0.375	0.159	2.19	2.19	0.89	14996	14788	0.986
Savoy, MA	5.55	39	0.563	0.190	2.11	2.11	0.89	11932	11222	0.94
Barnstable, MA	4.82	30	0.304	0.21	2.00	1.9	0.89	6565	7792	1.187
Falmouth, MA	4.34	30	0.360	0.17	2.00	3.04	0.89	4184	5658	1.352
Harwick, MA	3.78	20	0.473	0.28	2.00	2.19	0.89	3376	3570	1.057
Mt Tom, MA	5.27	24	0.375	0.21	2.00	2.23	0.89	9843	10066	1.023
Savoy, MA	3.81	20	0.563	0.28	2.00	1.9	0.89	4703	3639	0.78
Barnstable, MA	4.82	30	0.304	0.21	1.90	1.9	0.89	6565	7009	1.068
Falmouth, MA	4.34	30	0.360	0.17	3.04	3.04	0.89	4184	4065	0.972
Harwick, MA	3.78	20	0.473	0.28	2.19	2.19	0.89	3376	3210	0.951
Mt Tom, MA	5.27	24	0.375	0.21	2.23	2.23	0.89	9843	9296	0.944
Savoy, MA	3.81	20	0.563	0.28	1.90	1.9	0.89	4703	3911	0.832

A final distressing point concerns the lack of guidance given on the selection of K which is supposed to be = 2 for inland sites, = 3 for coastal sites, and = 4 for island sites. The value of K selected does influence the computed value of Y_p. In a sensitivity analysis previously done the value of Y_p can increase by 30% of the value when K = 4 versus when K = 2.

Perhaps a more extensive study involving 30 or 40 sites might reveal some rules-of-thumb for the selection of K values. How far does a turbine have to be from the sea to be ‘inland’? How small does in island have to be to be an ‘island’ site for the purposes of selection of K = 4? Unfortunately the limited study shown in Table 2 provides no insight into what the rules might be.

Readers wishing to see other examples of how well the Weibull function (Eq.1) fits real data might like to download the interim reports at:

http://www.warwickwindtrials.org.uk/2.html

These reports apply to small wind turbines installed in urban setting and often in sites with low wind speeds. These reports show that improved fitting can be achieved if the coefficient b in Eq. (1) is allowed to change from the standard value of b =0.89. The Weibull function, which is defined in the Jan 08 interim report, is identical to Eq. (1) except the shape factor used is $c = \bar{v} / b$. I have also found improved fits if I allow b to change. However, this information is of little value for WindCad purposes, because usually estimates of Y_p are desired for situation in which measured probability density distributions are not available. If you already have the distribution then there is no need to evoke a Weibull function.

TIMAPE part 5 will address the issue of estimating mean annual wind speed. I cannot yet eliminate the possibility that this is the largest source of error although I have found several other sources: (1) The effect of turbulence on power curves R#30 & R#32, (2) The fact that temperature is not included in WindCad models R#33, and (3) the accuracy of Weibull functions to predict real-world probability density distributions of wind speed.

Comments and suggested improvements to this document are encouraged so that it can be improve in future revisions. I am still a novice in this field so welcome comments from experts and armatures alike.

Best regards,

Melvin T. Tyree, BA PhD (*Cantab*) LLD (*h.c.*) FRSC
Moderator of SWIEP