A Copula Model of Wind Turbine Performance

Bruce Stephen, Stuart Galloway, David McMillan, David C. Hill and David Infield

***Abstract*-- The conventional means of assessing the performance of a wind turbine is through consideration of its power curve which provides the relationship between power output and measured wind speed. In this paper it is shown how the joint probability distribution of power and wind speed can be learned from data, rather than from examination of the implied function of the two variables. Such an approach incorporates measures of uncertainty into performance estimates, allows inter-plant performance comparison, and could be used to simulate plant operation via sampling. A preliminary model is formulated and fitted to operational data as an illustration.**

***Index Terms*-- Wind power generation, Energy conversion, Power generation reliability.**

I. Wind Turbine Performance: Power Curve Analysis

RECENT investments in renewable power generation have raised legitimate concerns over the security of supply from wind generation due to their greater dependence on environmental factors than conventional forms of generation. Whilst there will be some power generated so long as a wind turbine is operating this may well be significantly less than its rated power, and may vary with a range of factors such as wind turbulence levels and those related to the condition of the wind turbine itself [1]. It is thus important to quantify and monitor how much of the wind available on the site in question is actually converted into useable power. From a plant operators perspective it should be useful to track changes in the conversion efficiency as the plant ages, and also between plants.

II. A Generative Model of Wind Turbine Performance

The accepted measure of wind turbine performance is the power curve: the implied relation between power and wind speed. Calculation of the curve as prescribed by International Standards [2], follows a regression approach, binning the data, which can omit several important features. The probability distribution of the wind energy and its derivatives at a given site provides an assessment of the probability of generation at a given level with frequently occurring values resulting in more dense data regions on the curve, as may be seen in a scatter plot. There is a limited region over which wind can be used to generate power [3]; if wind speed is too low it falls below the turbines cut-in limit; too high and it exceeds its cut out limit; somewhere between these two values it reaches rated power (this wind speed is known as the rated wind speed). Between the rated wind speed and the cut-out wind speed, it operates at more or less a constant level, the rated power. However, between cut-in and rated, the power depends strongly on the wind speed, as shown in figure 1.

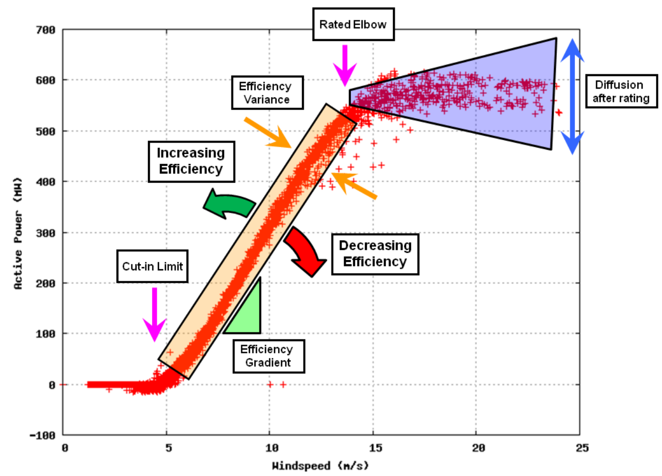

Fig. 1. Notable features of a typical power curve.

An alternative view may be obtained via a probabilistic modeling approach [4], in particular the density estimation of the joint probability of wind speed and active power. This representation is already implied by the scatter plot of active power and wind speed as shown in figure 1. If the data deviates significantly from the range of values that the model is expecting then this likelihood measure will be lower. In operational terms this may be related to the presence of an incipient problem. However, correctly estimating the form of the joint distribution is critical to this being achieved.

III. Addressing Limitations of Multivariate Dependency Modeling: Copulas

There are circumstances when dependent variables are related by more than linear dependence. Copulas [5] offer a means of relating the marginal densities of observed variables to a joint density with a complex dependency structure specified in a single function. The Archimedean family of Copulas uses only the cumulative marginal distributions F_i, i=1…d to obtain the cumulative joint probability density H of d variables using copula C, given as

$$H(x_1, x_2, \ldots, x_d) = C(F_1(x_1), F_2(x_2), \ldots, F_d(x_d)) \quad (1)$$

Dr. B. Stephen is a Research Fellow in the Advanced Electrical Systems Research Group, Institute of Energy and Environment, University of Strathclyde, Glasgow, G1 1XW (phone: +44 (0)141 548 5864, e-mail: bstephen@eee.strath.ac.uk)

One example of the Archimedean family is the Frank Copula which results in a Copula over cumulative marginal distributions $u=F_1(x_1)$ and $v=F_2(x_2)$ with a single parameter κ, given by

$$C(u,v) = -\frac{1}{\kappa}\ln\left[1+\frac{\left(e^{-u\kappa}-1\right)\left(e^{-v\kappa}-1\right)}{\left(e^{-\kappa}-1\right)}\right] \quad (2)$$

The joint probability density function, $P(x_1, x_2)$ of the Copula can then be obtained from the joint cumulative density using the following relation

$$P(x_1,x_2) = \frac{\partial^2 H(x_1,x_2)}{\partial F_1(x_1)\partial F_2(x_2)} \quad (3)$$

The cumulative distribution F for each marginal x_d can take the form of a linear mixture of cumulative Gaussian distributions that use the component parameters obtained from the fitting of a simple finite mixture of Gaussian distributions to each marginal

$$P(x_d) = \sum_{m=1}^{M}\alpha_m N(x_d;\mu_m,\sigma_m) \quad (4)$$

where the mixture component weights, α, means, μ, and variances, σ, of the M Gaussian (or Normal - N) distributions provide an arbitrary fit to the distribution of the variable x_d. This allows the marginal distributions to be approximated without the constraints implied by a classical distribution that may not permit skewed or multiple modes.

IV. Model of Operational Data

SCADA data was obtained from an operational wind park of 24 turbines in central Scotland from which a model was derived based on a single turbine month of 10 minute wind speed/power observations. The turbine used was a Bonus 600kW Mk IV machine of fixed pitch, stall control design. In this configuration of wind turbine the rotor speed is held nominally constant through the grid connection to the electrical generator of the turbine, and power is controlled by means of aerodynamic stall of the flow over the blades. This is in contrast to current designs in which power is controlled by varying the blade pitch. Figure 2 shows the joint density model for a single turbine based on data from a month of normal operation. The marginal distributions for wind speed and active power were fitted individually with M set to 3 in order to correspond to the operating regimes observed in the power curve (below (or near) cut-in; cut-in to rated; above rated). Although the Copula parameter κ was determined experimentally, the statistical foundation of the model allows M or κ to be estimated from data through optimization of some criteria such as maximum likelihood and employing the model selection and confidence measures that accompany them [4].

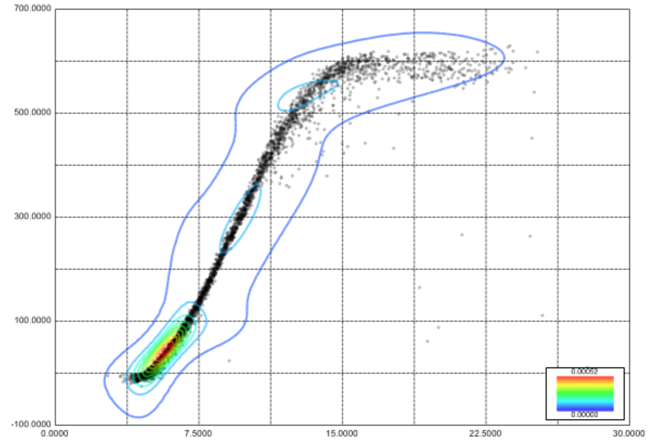

Fig. 2. Joint density of wind speed (x axis) and power (y axis) modeled using a Frank Copula with marginal densities approximated using Gaussian Mixtures.

V. Conclusions and Further Work

Capturing an accurate representation of the joint density function for power and wind speed in a parametric form is only a first step. The generative modeling approach outlined here permits model comparison through the same axioms of inference that are used to identify the model parameters themselves from data [4]. Such an approach could allow engineers to track inter and intra plant similarity. Since model changes should reflect actual physical changes to the turbine itself, including the turbine controller, these comparisons can be used to monitor the condition of the plant, an issue of increasing importance to wind farm operators as their fleets increase in size. Additionally, the models can be used to simulate plant behavior [1]: samples generated from a suitable wind profile can be used in conjunction with stochastic techniques such as Gibbs sampling [4] to provide time series of power generation representative of the turbines and the inherent uncertainties in their individual and by extension, their collective operation.

VI. References

[1] Kusiak, A. and Wenyan Li , "Virtual Models for Prediction of Wind Turbine Parameters,", *IEEE Transactions on Energy Conversion*, vol.25, no.1, pp.245-252, March 2010

[2] IEC 61400-12 Wind turbines - Part 12-1: Power performance measurements of electricity producing wind turbines, IEC 2005-12-16.

[3] T. Üstüntaş, and A.D. Şahin, "Wind turbine power curve estimation based on cluster center fuzzy logic modeling," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 96 (5), pp. 611-620. May 2008.

[4] D.J.C. MacKay. "Information Theory, Inference, and Learning Algorithms". Cambridge University Press, 2003.

[5] A. Patton. "Copula-Based Models for Financial Time Series", in T.G. Andersen, R.A. Davis, J.P. Kreiss and T. Mikosch (eds.) *Handbook of Financial Time Series*, Springer Verlag 2009.