Probabilistic wind power generation model: Derivation and applications

Henry Cheng, Yunhe Hou, and Felix Wu

Abstract—Wind power output is always uncertain but, in a sufficiently long time interval, the output exhibits statistical behavior that is meaningful enough to be characterized by probability distribution. The aim of this paper is to develop a model for probabilistic wind power generation. In particular, we successfully derive the analytical expression and statistics up to the fourth order of the wind power density function. The work also extends the modeling of wind power output up to a regional scale by Gram-Charlier series. Model results are checked by empirical power data and Monte Carlo simulation. This paper discusses some applications of the wind power statistics such as probabilistic production costing and reliability evaluation in power system literature.

Keywords— Gram-Charlier series, production cost, reliability, wind power cumulant/moment

I. Introduction

ELECTRICAL energy is increasingly produced from renewable sources worldwide for a number of reasons such as environmental concern, elevating fossil fuel prices, political motive, etc. Wind power is one of the most popular forms of renewable energy. Yet wind power is intermittent and virtually uncontrollable. Its large-scale deployment would influence power system in unprecedented ways. High penetration wind power poses a need of refinement to a number of existing methodologies for power system analysis.

Power system analysis has a lot to do with probabilistic evaluation, notably, production costing and reliability. In order to incorporate wind power generation into existing analytical framework, probabilistic wind power model is highly desirable. Such model shall represent wind power generator as a multi-state (capacity) unit. Early attempt did not consider failure and repair characteristics of wind turbine [1]. It was improved to incorporate wind turbine availability by discrete convolution [2][3]. In [4], wind speed is modeled by Markov chain, then power curve and availability of individual wind turbines are all taken into account to form the probability distribution of wind farm output capacity. Building the wind power generation model using the Markov method theoretically requires a stochastic process for the wind speed. From a different perspective, wind energy production in a sufficiently long time interval shall be able to be statistically characterized if the wind speed distribution and power curve representation are reasonably accurate. Say, in a year, wind power generation should follow certain probability density function (PDF). Simulated wind power PDF was first noted in [5]. Mathematically, analytical expression of the wind power PDF based on wind speed distribution and linearized power curve can be derived; reference [6] and [7] works on Weibull and Rayleigh wind speed respectively, and both are able to provide first order statistic (mean) of the wind power PDF. More complicated works based on Weibull wind speed and nonlinear power curve simultaneously are also available [8][9]. Consideration of wind directions in the wind power distribution is reported in [10].

The objective of this work is also to build a probabilistic wind power model. Polished from our earlier paper [11], this paper aligns with previous references, but works in a simplified domain that permits relatively straightforward analytical derivation of higher order statistics of the wind power PDF. The wind power statistics have important applications in conjunction with production costing and reliability evaluation. Production costing of conventional generators can be performed by cumulant method [12][13]. Those higher order statistics can supplement the cumulant method to determine expected cost of the conventional generation net of wind power. The wind power statistics can also be utilized in reliability evaluation, which is similar to the handling of generator's forced outage rate by cumulant method [14][15].

Previous works tend to be confined by modeling the output of single wind turbine or wind farm, but are not readily generalized to wind power output in regional scale, which is needed for production costing and reliability evaluation in a national perspective. The wind power statistics make the modeling of regional wind power possible by furnishing analytical approximation of the total wind power PDF using Gram-Charlier series. If wind power output is treated as a random variable, then any total regional wind power is obtained by recursively adding all individual distributions. Correlations due to intra-farm and inter-farm proximity are all efficiently rendered by the correlated cumulant method [12]. The resultant regional wind power distribution is anticipated to be bell-shaped suggested by the Central Limit Theorem.

This paper is organized as follow. Section II offers the proposed wind power model, with most of its derivation kept in the Appendix. Section III highlights the data sets of wind

This work is supported by Research Grant Council, Hong Kong SAR, under grant HKU7186/08E. The first author is also supported by the CLP Fellowship in Electrical Engineering, awarded by the CLP Holdings Limited.

Henry Cheng, Yunhe Hou and Felix Wu are with the Department of Electrical and Electronics Engineering, The University of Hong Kong, Hong Kong, China (email: mkcheng@eee.hku.hk; yhhou@eee.hku.hk; ffwu@eee.hku.hk)

speed and power. Pre-processing of data is carefully explained in Section IV. Section V of this paper is a collection of simulated and empirical results.

II. Ideal Wind Power Model

Most probabilistic wind generation models start with discussion on the wind speed distribution. Several distributions other than Gaussian had been early suggested [16]. Among them, the Weibull and its special case Rayleigh, have generally been recognized and are employed in engineering literature [17][18]. On the other hand, the wind turbine power curve is often made linear for simplicity. A generic wind generation model based on those simplifying assumptions is always nice, yet ideal. It may not be able to address a number of issues such as empirical distribution mismatch, power curve actual performance, wake effect when scaling up power from single turbine to wind farm, and some minor problems like wind direction and turbine outage.

Aerodynamic principle states that the power in airflow follows a cubic relationship with the free wind speed w:

$$P_{air} = \frac{1}{2}\rho A w^3 \tag{1}$$

where ρ is the air density and A is the area swept by the rotor. Not all of this power is extractable by a wind turbine as mechanical power, apart from the power loss during conversion to electrical power. In fact if the turbine output power P_{wt} is measured and divided by P_{air}, the ratio is referred as power coefficient C_P:

$$C_P = \frac{P_{wt}}{P_{air}} \tag{2}$$

It is observed that no matter how the wind turbine is designed, there is a maximum value for C_P, known as the Betz Limit and C_{PMax} equals 0.593. In other words one can never extract more than 60% of the power in airflow. Furthermore, power coefficient is a nonlinear function of tip speed ratio γ [19], in which γ is defined as

$$\gamma = \frac{\omega r}{w} \tag{3}$$

where r is the radius of the blade and ω is the angular speed of the rotor. To achieve the maximum power coefficient, the tip speed ratio has to remain constant at a particular value. It implies that the desirable γ value occurs at a particular wind speed for fixed speed wind generator. For variable speed wind generator its rotational speed can be adjusted to track the wind speed in order to operate the wind turbine at optimal tip speed ratio, hence preserving maximum power coefficient [20].

Knowing that the power in airflow follows a cubic relationship with free wind speed, the ascending segment of the power curve should better be modelled as a curve. However, it does not necessarily mean the output power is proportional to w^3 because the actual performance is complicated by the power coefficient. The power curve should theoretically be represented by general nonlinear function. So in order for analytical formulation to be carried on, this work proposes to break down the ascending segment into many smaller linear pieces, as represented by the following expression:

$$\tilde{g} = \begin{cases} 0 & 0 \le w \le w_{in} \quad or \quad w \ge w_{out} \\ a_1 w + b_1 & w_{in} < w \le w_1 \\ a_2 w + b_2 & w_1 < w \le w_2 \\ \vdots & \vdots \\ a_n w + b_n & w_{n-1} < w < w_r \\ g_m & w_r \le w < w_{out} \end{cases} \tag{4}$$

where

$$a_1 = \frac{g_1 - g_0}{w_1 - w_{in}}, b_1 = \frac{g_0 w_1 - g_1 w_{in}}{w_1 - w_{in}}$$
$$a_2 = \frac{g_2 - g_1}{w_2 - w_1}, b_2 = \frac{g_1 w_2 - g_2 w_1}{w_2 - w_1},$$
$$\vdots$$
$$a_n = \frac{g_m - g_{n-1}}{w_r - w_{n-1}}, b_n = \frac{g_{n-1} w_r - g_m w_{n-1}}{w_r - w_{n-1}}$$

and $g_0=0$, g_1, g_2, …, g_{n-1}, g_m correspond to the wind turbine output at w_{in}, w_1, w_2, …, w_{n-1}, w_r respectively. $\tilde{g}$ is the random wind power, defined by constants a_i and b_i $\{i \in [1,n]\}$, g_m is the maximum net output, and w_{in}, w_r and w_{out} are the cut-in, rated and cut-out wind speed respectively. Without loss of generality, consider breaking the ascending segment into only two segments, with (w_1, g_1) a knee point freely chosen on the curve part.

On the other hand, the probability density function (PDF) and cumulative distribution function (CDF) of Rayleigh distribution of wind speed $\tilde{w}$ are respectively

$$f_{\tilde{w}}(w) = \frac{2w}{\lambda^2} e^{-(\frac{w}{\lambda})^2} \quad w \ge 0 \tag{5}$$

and

$$P\{\tilde{w} \le w\} = \int_0^w \frac{2v}{\lambda^2} e^{-(\frac{v}{\lambda})^2} dv = 1 - e^{-(\frac{w}{\lambda})^2} \tag{6}$$

where λ is the scale parameter. By integration with respect to each wind speed partition, CDF of the wind power is determined. Upon differentiating the CDF, the wind power PDF $f_{\tilde{g}}(x)$ is obtained as follow.

$$f_{\tilde{g}}(x) = \begin{cases} (1 - e^{-(\frac{w_{in}}{\lambda})^2} + e^{-(\frac{w_{out}}{\lambda})^2})\delta(x) & x = 0 \\ 2(\frac{x - b_1}{a_1^2 \lambda^2}) e^{-(\frac{x-b_1}{a_1 \lambda})^2} & 0 < x \le g_1 \\ 2(\frac{x - b_2}{a_2^2 \lambda^2}) e^{-(\frac{x-b_2}{a_2 \lambda})^2} & g_1 < x < g_m \\ (e^{-(\frac{w_r}{\lambda})^2} - e^{-(\frac{w_{out}}{\lambda})^2})\delta(x - g_m) & x = g_m \end{cases} \tag{7}$$

where $\delta(x)$ is a delta function to maintain the derivative-integral relation between CDF and PDF.

Based on the analytical wind power PDF, it is tempting to ask for its statistics. Derivation of formulae of higher order statistics is tedious and lengthy; therefore results of the four statistics, i.e. mean, variance, skewness and kurtosis are left in the Appendix. Effectively, the four statistics are cumulants of wind power PDF and they complete the statistical characterization of the wind power generation model.

III. Wind Speed and Power Data

Overall this work is based on two sources of hourly wind data. The first one is a multi-decade, speed-only database of tens of Dutch locations [21]. The second one [22] is less extensive, but with corresponding output from a real wind turbine, which is most valuable. Brief descriptions of them are as follow.

A. Royal Netherlands Meteorological Institute

There are more than 50 measuring stations, onshore or offshore, each has a name and ID. Every station measures its hourly wind speed, with figures of measuring height and roughness length. Data entry may be void (missing) or negative (faulty). Out of all 31 stations are chosen that are most complete in the past 19 years (1991-2009). Void and negative data entries are assigned a very small positive value. All 29th Feb in leap year are ignored for easy programming. Effectively, each location has 8,760 x 19 equals 166,440 hour wind speed.

B. Vermont Small-scale Wind Energy Demonstration Program

In this dataset, wind speed and corresponding output of medium scale wind turbine (nominal 10kW) are available from a few sites. Data are recorded at hourly intervals. Within each hour, average wind speed and total energy production are known. Maximum and minimum wind speed and output power per hour are also given. It provides better insight because the average wind speed of an hour is weak to show any intra-hour highs and lows. Most importantly, those speed-power pairs can be used to determine empirically the wind turbine power curve. Separately, specifications of the actual wind turbine are found in [23].

IV. Data Pre-processing

A. Wind speed measuring height

Normally, wind speed is measured at a convenient height above ground. It has to be interpolated to wind turbine hub level for power conversion. In general wind profile follows a logarithmic relation with its height above ground [25]:

$$w_h \propto \ln(\frac{h}{z_0}) \tag{8}$$

where h is the elevation, w_h is the wind speed at the elevation concerned and z_0 is so-called the roughness length. Roughness length indicates surface friction and is defined as the height drops to where the mean wind speed becomes zero. It is necessary to up-scale measured wind speed to the hub level. All Dutch wind speeds in section III. A are scaled to 100m high.

B. Wind speed partitions and the parameter lambda

Cut-in, rated and cut-out speeds of a wind turbine are designed to suit the prospective wind regime where it is located. The rule of thumb of design is based on the annual average wind speed $\bar{w}$ onsite: w_{in} equals 0.6 $\bar{w}$; w_r equals 1.5-1.75 $\bar{w}$ and w_{out} equals 3 $\bar{w}$ [20]. Next, value of the scale parameter λ is needed for generating Rayleigh random variates. Such value can be estimated from a set of historical wind speed data. Alternatively, if the long-term average wind speed is known, the lambda could be determined passively by the following equation

$$\bar{w} = \frac{\lambda\sqrt{\pi}}{2}, \tag{9}$$

which relates λ and the Rayleigh mean.

C. Power curve

Power curve measures the conversion ability of a wind turbine. A simplified power curve consists of linear segments is described in (4). In general the ramp segment should rise nonlinearly because aerodynamic principle states that the power in airflow follows a cubic relationship with free wind speed. Meanwhile, its actual performance depends on generator type (fixed or variable speed) and control method (pitch or stall controlled). In practice, power curve data are empirical and given by wind turbine manufacturer. It consists of power values (and losses) across a range of operating wind speed, e.g. in pp. 68 of [18]. Hence an empirical power curve may not have a proper mathematical form.

Even wind speeds at different hours are the same, the energies produced could be different. Arguably, the speed is only average of an hour and intra-hour fluctuations could be significantly different. Wind turbine performance varies with wind gust, turbulence and wind directions to different extents. It is reported in commercial software WAsP [26] that wind speed within a particular directional sector does not normally show Weibull distribution, fortunately, improvement made to the estimation of power output by considering directional effect is little in most cases.

The approach here is to run regression on numerous real speed-power data to determine an empirical power curve [27]. The empirical power curve has theoretically captured any long run site characteristic, e.g. terrain, wind direction etc. Prior to the regression process, it is necessary to have a hypothesis of what the underlying function of the power curve is. And since a power curve can be reasonably partitioned into an ascending segment and a plateau, it is logical to have two underlying functions: cubic function and constant function of speed respectively. Then we use linear regression subroutine built in Matlab to obtain the fitted curves.

We make reference to the data source mentioned in section III. B. The site picked for demonstration is called RicLin. The span of data is typically a year of hourly speed-power data points, with omission of faulty measurements. Fig. 1 shows the power curve under trial regression.

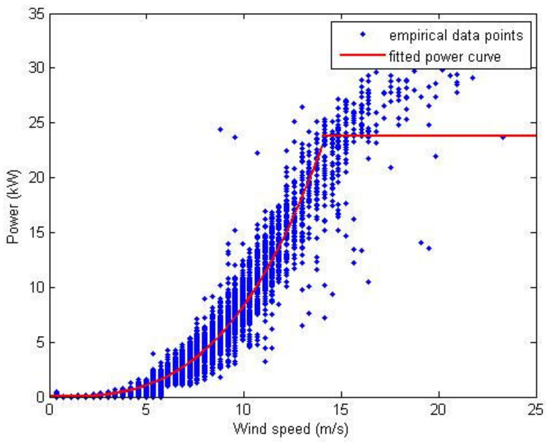

Fig. 1 Empirical power curve determined by regressing real data

D. Scaling and wake effect

When many wind turbines are assembled in an array, downstream wind speed is reduced by upstream wind turbines. This is called wake effect. Hence the total power of wind farm is less than the proportional scaling of single wind turbine power. To account for the speed deficit, a wake effect model consolidated in [28] is applied. Let w_x denote the reduced wind speed x metre behind the upstream turbine, w_0 denote the undisturbed wind speed, then

$$w_x = w_0[1-(1-\sqrt{1-C_T})(\frac{r}{r+kx})^2], \quad (10)$$

where r is the upstream turbine rotor radius, k is the so-called wake decay constant and C_T is the thrust coefficient. k and C_T depends on wind farm terrain and wind turbine type respectively. Downwind distance x is normally $12r$ to $24r$. Though waked wind speed is different from the undisturbed one in magnitude, it still demonstrates a reasonable shaped Rayleigh distribution. Fig. 2 shows a sample waked wind speed profile for wind turbine with typical C_T value [26].

We believe existing methods of how to incorporate wake effect into wind farm expected output could be improved. Our approach is conceptually very simple. If wind power is treated as a random variable, then any total wind power is obtained by recursively adding all individual output distributions. There is no need to distinguish wind turbine out of a wind farm in the summation process for the reason that, given any downwind turbine speed reduced by wake effect, the waked speed profile could be modeled by a new, separate distribution. Correlations due to intra-farm and inter-farm proximity are rendered by the correlated cumulant method.

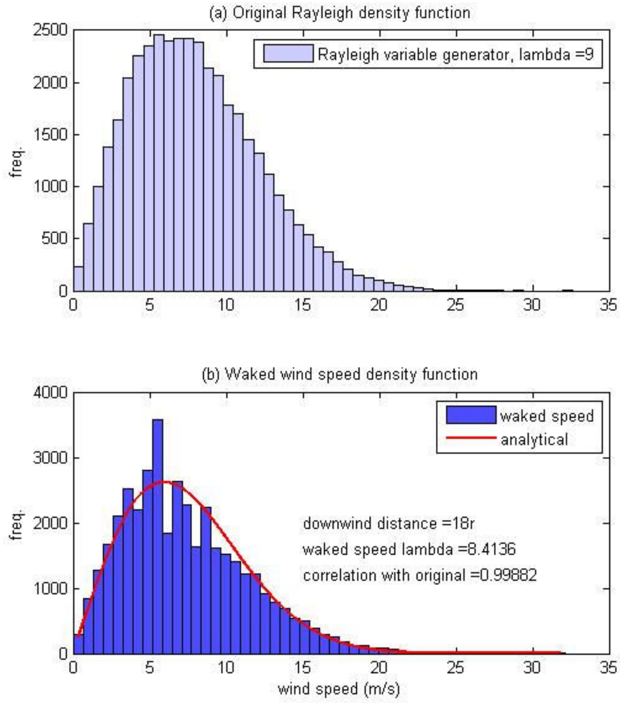

Fig. 2 Waked wind speed density function (b) compared with its original Rayleigh source (a)

V. Model Results

This paper does not preclude the assumptions behind the generic wind generation model, rather, it investigates how much different the model prediction is compared with empirical result. Empirical result is based on real data of wind speed and power. Availability of comprehensive real data has been a difficult task in course of the research. In particular, wind speed and corresponding power data of large wind farm are very difficult to find in open source. This work makes use of the data as mentioned in Section III. We tested locations for their analytical wind power distributions, which fit the same set of hourly empirical wind data satisfactorily. Results followed are divided into three parts.

A. Verification of the statistical formulae by Monte Carlo simulation

The formulae of four wind power statistics are checked by Monte Carlo simulation. Result variations due to power curve (three or four segments) and simulation runs are demonstrated in TABLE I. It can be seen that deviations between corresponding analytic and simulated values for the case of 3-segment power curve is larger than those of 4-segment power curve, indicating the latter is a more accurate model.

Computation is implemented in Matlab 7.0 on a common desktop computer with 2.66GHz Intel® Core™ 2 processor and 2GB RAM.

TABLE I
Checking wind power statistics by Monte Carlo simulation

Rated power = 3000kW		$\bar{g}$	σ	skewness	kurtosis excess
3-segment power curve	Monte Carlo 10000 runs	1190.7	1131.9	0.4209	-1.3377
	Analytic	1181.7	1127.7	0.4358	-1.316
	Deviations (%)	1.86	1.07	5.73	3.26
4-segment power curve	Monte Carlo 2500 runs	1069	1117.8	0.6963	-1.0321
	Monte Carlo 10000 runs	1041	1106.2	0.7297	-0.964
	Analytic	1053.1	1103.4	0.7094	-0.9861
	Deviations (%)	1.15	0.26	2.86	2.24

Wind power PDFs are also plotted for visual inspection. Fig. 3 shows the PDF of the case of 3-segment (simple) power curve, which is smooth in the middle. Fig. 4 shows the case of 4-segment (improved) power curve in which the PDF shows a jump. Theoretically, the model accuracy increases as the power curve is partitioned more.

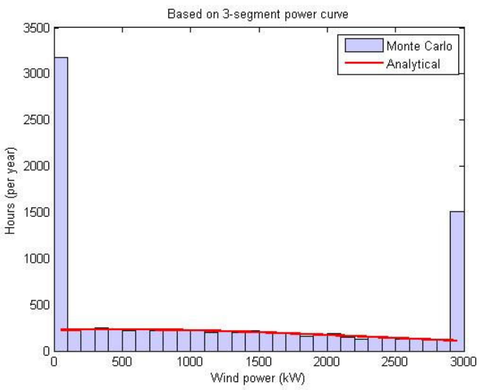

Fig. 3 Wind power PDF built from simple power curve

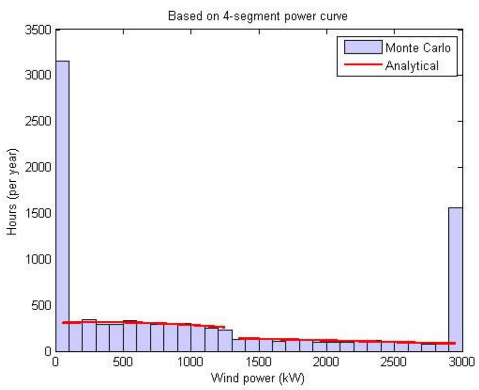

Fig. 4 Wind power PDF built from improved power curve

B. *Comparison between analytical and empirical wind power density functions*

The fact that analytical and simulated results in Section V. A are matched only means the formulae are correctly derived, it does not guarantee the same for analytical and any empirical data. Analytical model uses only parameter of wind speed distribution and simplified power curve to calculate wind power. Empirical data are what actually be measured from the wind turbine under the same speed distribution. Accuracy of the analytical model lies on the fitness of distribution envelope and precision of power curve. The key effort of this part is to substantiate the proposed wind power model by benchmarking with some real speed-power data. Annual data of the Harvest Hill Farm, another site from the Vermont dataset, are used. Its analytical and simulated wind power PDFs are presented in Fig. 5(a) as control, the critical point is whether the analytical PDF models the empirical power data well in Fig. 5(b). Compared with Fig. 4, the empirical power distribution does not show any frequency at the right end of the horizontal axis, although the empirical power curve in Fig. 1 clearly shows occurrence of high power output. It is because the empirical power curve is made up of maximum values of speed and power within an hour whereas the wind power PDF is derived from hourly average values. It suggests that only persistently large wind speed can generate high power output.

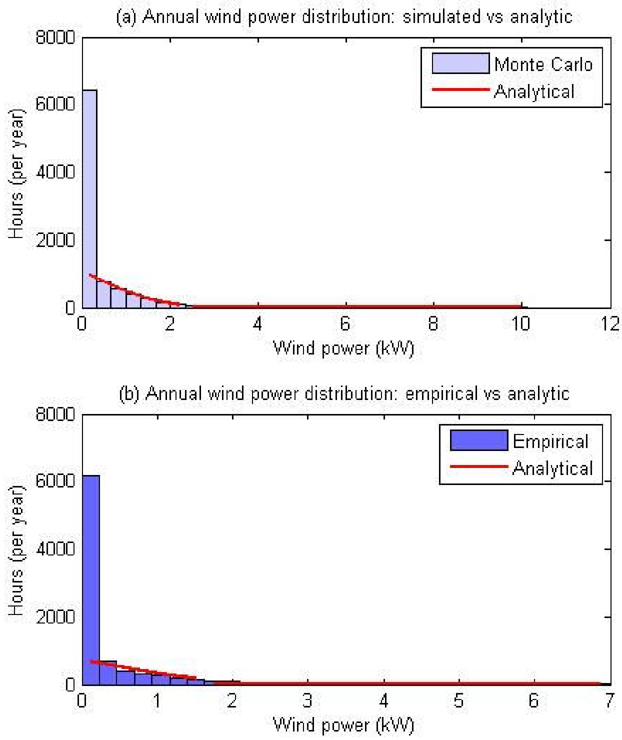

Fig. 5 Modeling empirical wind power by analytical PDF

The wind power distributions in terms of CDF for the case of RicLin and Harvest Hill Farm are also plotted in Fig. 6 and Fig. 7 respectively. It can be seen that the empirical and analytical distributions for concurrent year data are reasonably matched. More exhaustive empirical work would reinforce the robustness of the result. In terms of chronological simulation, [24] reported good tracking ability of its testing wind turbine for varying wind speed condition and anticipated power output.

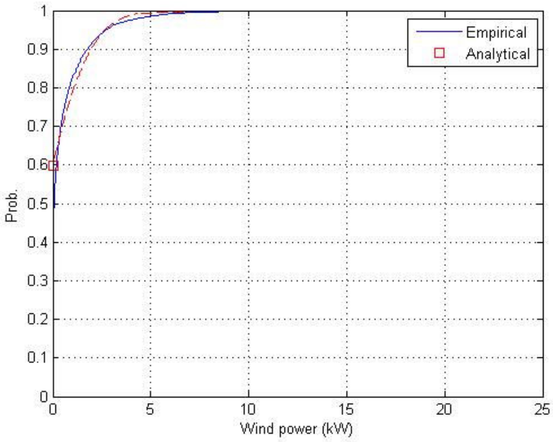

Fig. 6 Wind power distribution CDFs for RicLin (Oct 07 – Sept 08)

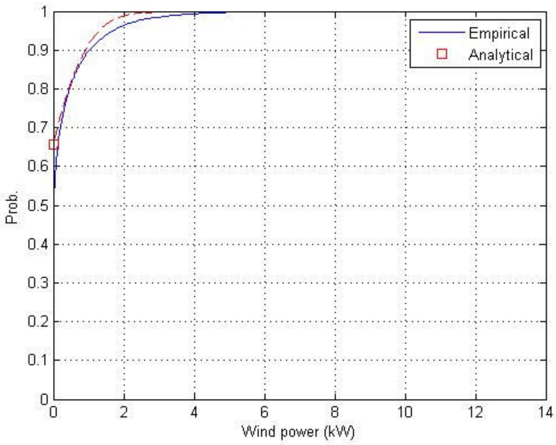

Fig. 7 Wind power distribution CDFs for Harvest Hill Farm (Mar 07 – Feb 08)

It is beyond the scope of this paper to tell any dependence of inter-year wind speed, hence wind power distributions for each site. The research would involve enormous amount of empirical data, though availability of multi-year real wind power may be difficult to obtain. Apparently, inter-year wind speed distributions have distribution parameters in closed range, suggested by some trial work on the Dutch wind speed database, but assertion cannot be made until significant future work is done. Perhaps the annual average is the attribute that is the easiest to check for any long-term trend. We plot the annual average wind power, converted from hourly wind speed data of about two decades, for a Dutch location as in Fig. 8.

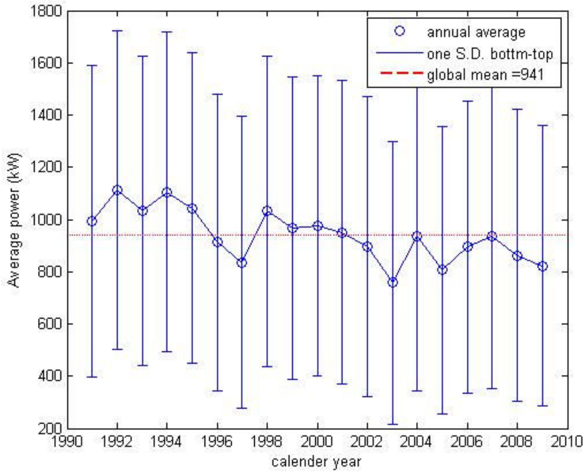

Fig. 8 Trend of multi-year annual average wind power for Valkenburg

C. *Regional wind power distribution*

An extended numerical example of this work is to synthesize the probability distribution of regional wind power output for production costing and reliability evaluation in large scale. As worked out, individual wind turbine output is a random variable characterized by probability distribution. Mathematically, any total wind power is the convolution sum of wind turbine random outputs. But obviously it is too overwhelming if thousands of wind turbines, each with numerous capacity states, are convoluted to yield the resultant distribution. We seek to employ the Gram-Charlier series to approximate the resultant wind power density function, but soon the substantial correlation among wind speeds is realized and the independence assumption of Gram-Charlier approach is potentially violated. Fortunately, the correlated cumulant method is readily applicable, which is the blueprint of this section. The cumulant method requires only one-off calculation of cumulants, and the value of analytical statistics derived earlier comes in place. No matter how individual wind turbines are different in terms of availabilities, ratings and wind profiles, etc., each wind turbine possesses a set of characteristic cumulants for the summation process.

From Fig. 3 to Fig. 5, it is shown that the power density function of one wind turbine has concentration at zero output. When many of such PDFs of different capacities are added together, one would anticipate by Central Limit Theorem the resultant density function to be bell-shaped. We verify this belief by enumeration of real data. The resultant wind power density function can be obtained by two ways. On one hand, it is done by successive convolution of individual outputs, which is rendered by Matlab subroutine. On the other hand, the Gram-Charlier series gives an analytical approximation to the resultant wind power PDF by processing the cumulants of all individual outputs. The two approaches should match in outcome when sufficient quantities of wind turbine are grouped together. However, extensive real wind power data could be difficult to obtain as an open source. For research trial and illustration, we build a hypothetical scenario with reference to real data as much as possible.

Denmark is used as the backdrop of our hypothetical scenario. Denmark has over 3,000 MW wind power installed capacity in 2008 [29], shared by slightly more than 5,000 wind turbines in 2009, with breakdown by ratings as grouped in TABLE II [30]. The ideal exposition of this numerical example would be to synthesize a power density function of all wind turbines using empirical wind power data. As a compromise, wind power data are converted by wind speed, and hourly wind speed records of tens of Dutch locations for any single year are available. But obviously there are not enough wind speed profiles compared with the number of wind turbines, and by no means these wind speed profiles are matched with the actual wind turbine locations. Therefore we only conduct test example.

TABLE II
Wind turbine breakdown by capacities in Denmark 2009

Ratings (kW)	0-225	226-499	500-999	1,000+	Total
Numbers	1427	306	2596	749	5078

Specific formulations of the test example follow. Thirty one wind speed profiles for a particular year from the Dutch database are extracted. Each profile of hourly wind speed is tagged with a wind turbine of randomised rating to enumerate the respective annual wind power distribution. All these power distributions exhibit randomness to various degrees, and are mildly correlated due to wind speed correlation. It is interesting and meaningful to check how closed to bell shape the resultant density function would be when individual wind power PDFs are convoluted successively. To visualize the improvements of large number, Fig. 9 plots the wind power PDFs generated by an occasion of six successive convolutions stage by stage. It is demonstrated that the resultant density function approaches bell shape when the number of sites convoluted increases. The corresponding cumulative wind power capacity increases along the horizontal axis.

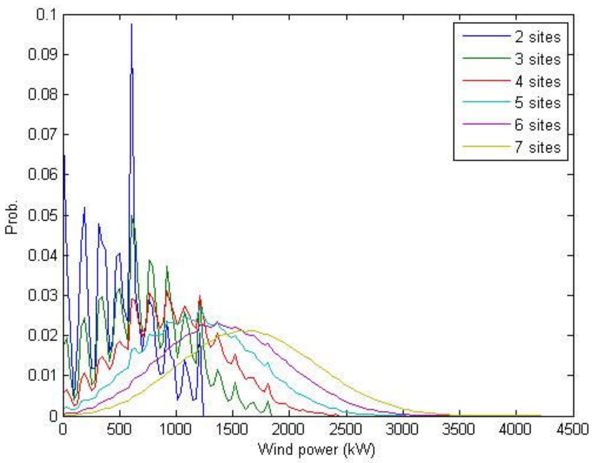

Fig. 9 Successive convolution of individual wind turbine outputs

Apart from numerical convolution, the resultant wind power, in terms of standardized density function, can be approximated by the correlated cumulant method. Fig. 10 shows the analytical approximation of the resultant wind power PDF to the same setting as by convolution. The envelope is laid on a histogram of simulation of a standard normal random variable for visual checking. It could be seen that the analytical PDF and the Monte Carlo simulation result are similar, but not yet matched for the fact that only seven sets of cumulants are available to synthesize the curve. The result is improved considerably when the number of wind turbines increases, say up to 31, as shown in Fig. 11. Then the analytical envelope is very much alike a standard normal PDF.

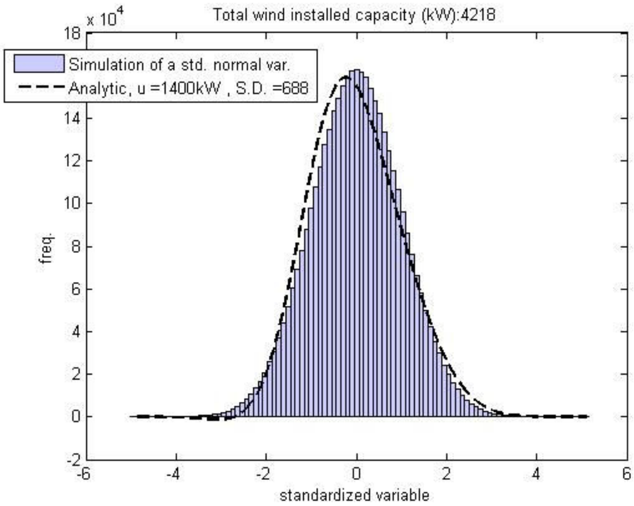

Fig. 10 Standardized PDF of Gram-Charlier series of 7 variables

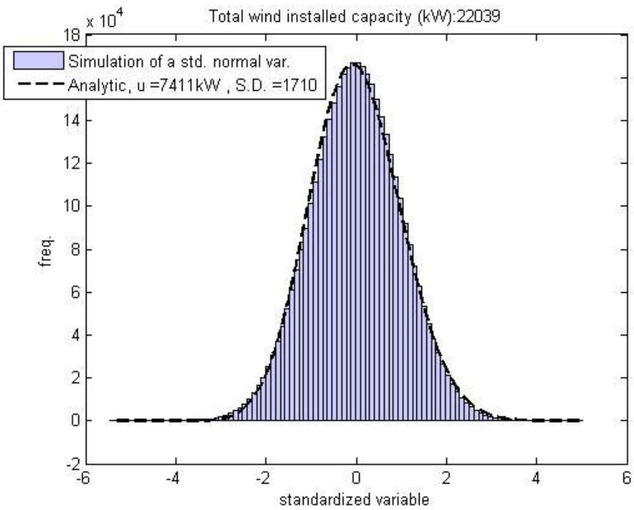

Fig. 11 Standardized PDF of Gram-Charlier series of 31 variables

As an attempt we have compared resultant wind power PDF by numerical convolution and by analytical approximation. Some possible reasons of mismatch are aware of. Apparently, faulty wind speed data explain partially. The fact that only cumulants up to the fourth order are used also leads to less precision.

VI. Conclusion

This paper has provided a probabilistic wind power model based on Rayleigh wind speed distribution and linearized power curve. Formulae of wind power statistics up to the fourth order are completely derived. The model is not only checked by simulation, but also benchmarked with empirical data, and shows satisfactory performance. The model can help evaluate production costing and reliability with wind power.

VII. APPENDIX

Suppose $f(x)$ is the wind power density function, its r[th] moment (about zero) and r[th] central moment are given respectively as

$$m_r = \int_{-\infty}^{\infty} x^r f(x)dx \quad \text{and} \qquad \text{(VII.1)}$$

$$c_r = \int_{-\infty}^{\infty} (x-m_1)^r f(x)dx \qquad \text{(VII.2)}$$

In particular, m_1 is the mean and $c_1=0$. For completeness we may also define $m_0=c_0=1$. By binomial theorem, any r[th] moment can be expressed in terms of the r[th] and lower order central moments [31]. Therefore,

$$m_r = \sum_{j=0}^{r} \binom{r}{j} c_{r-j} m_1^j \qquad \text{(VII.3)}$$

or vice versa,

$$c_r = \sum_{j=0}^{r} \binom{r}{j} m_{r-j} (-m_1)^j \qquad \text{(VII.4)}$$

They can be deduced interchangeably and it is more convenient to start with the mean m_1 and then all higher order moments can be generated. Consider (VI.1),

$$\begin{aligned}
\int_{-\infty}^{\infty} x^r f(x)dx &= \int_{-0}^{+0} 0^r (1 - e^{-(\frac{w_{in}}{\lambda})^2} + e^{-(\frac{w_{out}}{\lambda})^2})\delta(x)dx \\
&+ \int_{g_0}^{g_1} x^r 2(\frac{x-b_1}{a_1^2\lambda^2}) e^{-(\frac{x-b_1}{a_1\lambda})^2} dx + \int_{g_1}^{g_m} x^r 2(\frac{x-b_2}{a_2^2\lambda^2}) e^{-(\frac{x-b_2}{a_2\lambda})^2} dx \\
&+ \int_{-g_m}^{g_m} g_m^r (e^{-(\frac{w_r}{\lambda})^2} - e^{-(\frac{w_{out}}{\lambda})^2})\delta(x-g_m)dx \\
&= \int_{g_0}^{g_1} x^r 2(\frac{x-b_1}{a_1^2\lambda^2}) e^{-(\frac{x-b_1}{a_1\lambda})^2} dx + \int_{g_1}^{g_m} x^r 2(\frac{x-b_2}{a_2^2\lambda^2}) e^{-(\frac{x-b_2}{a_2\lambda})^2} dx \\
&+ g_m^r (e^{-(\frac{w_r}{\lambda})^2} - e^{-(\frac{w_{out}}{\lambda})^2})
\end{aligned} \qquad \text{(VII.5)}$$

The first two terms are the same except with different parameters for different power curve segments, implying that the formula derived for one segment should work for others. For simplicity, we abuse the notation of a and b by omitting their subscripts. Let $y = \frac{x-b}{a\lambda}$, then

$$\begin{aligned}
\int_{g_1}^{g_m} x^r 2(\frac{x-b}{a^2\lambda^2}) e^{-(\frac{x-b}{a\lambda})^2} dx &= \int_{\frac{g_1-b}{a\lambda}}^{\frac{g_m-b}{a\lambda}} 2y(ya\lambda+b)^r e^{-y^2} dy \\
&= \int_{k_1}^{k_2} 2y(ya\lambda+b)^r e^{-y^2} dy
\end{aligned} \qquad \text{(VII.6)}$$

where the lower and upper limits of the integration domain are denoted by k_1 and k_2 respectively for convenience (for the segment g_0 to g_1, let the limits be l_1 and l_2 respectively). For an integral of the form $\int y^n e^{-y^2} dy$, denote it by I_n. Let $g = -\frac{y^{n-1}}{2}, \frac{dh}{dx} = -2ye^{-y^2}$ and consider integration by parts, then recursively

$$I_n = -\frac{y^{n-1}}{2} e^{-y^2} + \frac{n-1}{2} I_{n-2} \quad for\ n \ge 2 \qquad \text{(VII.7)}$$

For $n=1$, it is observed that

$$I_1 = -\frac{1}{2} e^{-y^2} \qquad \text{(VII.8)}$$

For $n=0$,

$$(I_0)^2 = \iint_R e^{-(u^2+v^2)} dudv = \iint_P e^{-r^2} r dr d\theta \qquad \text{(VII.9)}$$

After considering the integration domain,

$$I_1 = [-\frac{1}{2} e^{-y^2}]_{k_1}^{k_2} = \frac{1}{2}(e^{-k_1^2} - e^{-k_2^2}) \qquad \text{(VII.10)}$$

$$I_0 = \int_{k_1}^{k_2} e^{-y^2} dy = \int_{k_1}^{0} e^{-y^2} dy + \int_{0}^{k_2} e^{-y^2} dy \qquad \text{(VII.11)}$$

Hence I_0 can be expressed in terms of error function *erf(.)*:

$$I_0 = \frac{\sqrt{\pi}}{2} \{erf(k_2) - erf(k_1)\} \qquad \text{(VII.12)}$$

Although an error function does not have closed form solution, one can expand the integrand e^{-y^2} by Taylor series and integrate term by term to obtain an approximation. In modern computer programs, it is readily generated by numerical integration.

By now we have obtained a recursive formula of any higher order of the integral I_n once I_1 and I_0 are generated. Let us show the derivation up to I_3 for the calculation of the mean and variance.

A. Mean

The mean of the random wind power is obtained by setting r=1 in (VI.1). From (VI.6),

$$\begin{aligned}
&\int_{k_1}^{k_2} 2y(ya\lambda+b)e^{-y^2} dy \\
&= 2a\lambda \int_{k_1}^{k_2} y^2 e^{-y^2} dy + 2b \int_{k_1}^{k_2} ye^{-y^2} dy \\
&= 2a\lambda I_2 + 2bI_1 \\
&= a\lambda([-ye^{-y^2}]_{k_1}^{k_2} + I_0) + b(e^{-k_1^2} - e^{-k_2^2}) \\
&= \frac{a\lambda\sqrt{\pi}}{2} \{erf(k_2) - erf(k_1)\} + g_1 e^{-k_1^2} - g_m e^{-k_2^2}
\end{aligned} \qquad \text{(VII.13)}$$

Adding the results of two segments together, the mean or first order moment m_1 is

$$\begin{aligned}
\bar{x} = m_1 &= \int_{-\infty}^{\infty} xf(x)dx \\
&= \frac{a_1\lambda\sqrt{\pi}}{2} \{erf(l_2) - erf(l_1)\} + g_0 e^{-l_1^2} - g_1 e^{-l_2^2} \\
&+ \frac{a_2\lambda\sqrt{\pi}}{2} \{erf(k_2) - erf(k_1)\} + g_1 e^{-k_1^2} - g_m e^{-k_2^2} + g_m (e^{-(\frac{w_r}{\lambda})^2} - e^{-(\frac{w_{out}}{\lambda})^2})
\end{aligned} \qquad \text{(VII.14)}$$

B. Variance

The variance of $\tilde{x}$, or the second order central moment, is given by (VI.2). Conveniently, from (VI.4) we can use

$$\sigma^2 = c_2 = m_2 - m_1^2 \qquad \text{(VII.15)}$$

Therefore we need m_2 first. Setting $r=2$ in (VII.6),

$$\int_{k_1}^{k_2} 2y(ya\lambda+b)^2 e^{-y^2} dy = \int_{k_1}^{k_2} (2a^2\lambda^2 y^3 + 4ab\lambda y^2 + 2b^2 y)e^{-y^2} dy$$
$$= 2a^2\lambda^2 I_3 + 4ab\lambda I_2 + 2b^2 I_1$$
$$= a^2\lambda^2([-y^2e^{-y^2}]_{k_1}^{k_2} + 2I_1) + 2ab\lambda([-ye^{-y^2}]_{k_1}^{k_2} + I_0) + 2b^2 I_1$$
$$= (a^2\lambda^2 k_1^2 + a^2\lambda^2 + 2ab\lambda k_1 + b^2)e^{-k_1^2}$$
$$-(a^2\lambda^2 k_2^2 + a^2\lambda^2 + 2ab\lambda k_2 + b^2)e^{-k_2^2} + 2ab\lambda I_0$$
$$= \{g_1^2 + (a\lambda)^2\}e^{-k_1^2} - \{g_m^2 + (a\lambda)^2\}e^{-k_2^2} + ab\lambda\sqrt{\pi}\{erf(k_2) - erf(k_1)\} \qquad \text{(VII.16)}$$

Adding the results of two segments together, then m_2:

$$m_2 =$$
$$\{g_1^2 + (a_2\lambda)^2\}e^{-k_1^2} - \{g_m^2 + (a_2\lambda)^2\}e^{-k_2^2} + a_2b_2\lambda\sqrt{\pi}\{erf(k_2) - erf(k_1)\}$$
$$+\{g_0^2 + (a_1\lambda)^2\}e^{-l_1^2} - \{g_1^2 + (a_1\lambda)^2\}e^{-l_2^2} + a_1b_1\lambda\sqrt{\pi}\{erf(l_2) - erf(l_1)\}$$
$$+g_m^2(e^{-(\frac{w_r}{\lambda})^2} - e^{-(\frac{w_{out}}{\lambda})^2}) \qquad \text{(VII.17)}$$

The formula of variance follows when m_1 and m_2 are known.

C. Skewness and Kurtosis

It becomes clear that any higher order moments or central moments can be built for better characterization of the PDF. Also eventually any orders of cumulant can be calculated by (VI.5) and those cumulants could readily be used in the Gram-Charlier series. Without showing the steps, the third and fourth cumulants are stated as follow.

Third cumulant:

$$\kappa_3 = c_3 = m_3 - 3m_2m_1 + 2m_1^3 \qquad \text{(VII.18)}$$

where

$$m_3 = \{g_1^3 + \frac{3}{2}a_2^2\lambda^2(g_1+b_2)\}e^{-k_1^2}$$
$$-\{g_m^3 + \frac{3}{2}a_2^2\lambda^2(g_m+b_2)\}e^{-k_2^2} + 3a_2\lambda(\frac{a_2^2\lambda^2}{2} + b_2^2)\frac{\sqrt{\pi}}{2}\{erf(k_2) - erf(k_1)\}$$
$$+\{g_0^3 + \frac{3}{2}a_1^2\lambda^2(g_0+b_1)\}e^{-l_1^2}$$
$$-\{g_1^3 + \frac{3}{2}a_1^2\lambda^2(g_1+b_1)\}e^{-l_2^2} + 3a_1\lambda(\frac{a_1^2\lambda^2}{2} + b_1^2)\frac{\sqrt{\pi}}{2}\{erf(l_2) - erf(l_1)\}$$
$$+g_m^3(e^{-(\frac{w_r}{\lambda})^2} - e^{-(\frac{w_{out}}{\lambda})^2}) \qquad \text{(VII.19)}$$

Fourth cumulant:

$$\kappa_4 = c_4 - 3c_2^2 = (m_4 - 4m_3m_1 + 6m_2m_1^2 - 3m_1^4) - 3(m_2 - m_1^2)^2 \qquad \text{(VII.20)}$$

where

$$m_4 = \{g_1^4 + 2(a_2\lambda)^2(g_1^2 + b_2g_1 + a_2^2\lambda^2 + b_2^2)\}e^{-k_1^2}$$
$$-\{g_m^4 + 2(a_2\lambda)^2(g_m^2 + b_2g_m + a_2^2\lambda^2 + b_2^2)\}e^{-k_2^2}$$
$$+(3a_2^3b_2\lambda^3 + 2a_2b_2^3\lambda)\sqrt{\pi}\{erf(k_2) - erf(k_1)\}$$
$$+\{g_0^4 + 2(a_1\lambda)^2(g_0^2 + b_1g_0 + a_1^2\lambda^2 + b_1^2)\}e^{-l_1^2} \qquad \text{(VII.21)}$$
$$-\{g_1^4 + 2(a_1\lambda)^2(g_1^2 + b_1g_1 + a_1^2\lambda^2 + b_1^2)\}e^{-l_2^2}$$
$$+(3a_1^3b_1\lambda^3 + 2a_1b_1^3\lambda)\sqrt{\pi}\{erf(l_2) - erf(l_1)\}$$
$$+g_m^4(e^{-(\frac{w_r}{\lambda})^2} - e^{-(\frac{w_{out}}{\lambda})^2})$$

By definition, we can also find skewness and kurtosis excess respectively as:

$$\gamma_1 = \frac{c_3}{\sigma^3} \qquad \text{(VII.22)}$$

$$\gamma_2 = \frac{c_4}{\sigma^4} - 3 \qquad \text{(VII.23)}$$

VIII. References

[1] R. G. Deshmukh and R. Ramakumar, "Reliability analysis of combined wind-electric and conventional generation systems," *Solar Energy*, Vol. 28, No. 4, pp. 345-352, 1982.

[2] P. Giorsetto and K. F. Utsurogi, "Development of a new procedure for reliability modeling of wind turbine generators," *IEEE Trans. PAS.*, Vol. PAS-102, No. 1, pp. 134-143, Jan 1983.

[3] X. Wang, H. Dai, and R. J. Thomas, "Reliability modeling of large wind farms and electric utility interface system," *IEEE Trans. PAS.*, Vol. PAS-103, No. 3, pp. 569-575, Mar 1984.

[4] F. Castro Sayas and R. N. Allan, "Generation availability assessment of wind farms," *IEE Proc.-Gen. Trans. Distrib.,* Vol. 134, No. 5, pp. 507-518, Sept. 1996.

[5] G. Papaefthymiou, P. H. Schavemaker, L. van der Sluis, W. L. Kling, D. Kurowicka and R. M. Cooke, "Integration of stochastic generation in power systems," *International Journal of Electrical Power & Energy Systems.*, Vol. 28, No. 9, pp. 655-667, 2006.

[6] Ciprian Nemes and Florin Munteanu, "The wind energy system performance overview: capacity factor vs. technical efficiency," International Journal of Mathematical Models and Methods in Applied Sciences, Vol. 5, No. 1, pp. 159-166, 2011.

[7] Hui Zhou, Yunhe Hou, Yaowu Wu, Haiqiong Yi, Chengxiong Mao, and Gonggui Cheng, "Analytical assessment of wind power generation asset in restructured electricity industry," *in Proc. the Universities Power Engineering Conference*, UPEC, 2007, pp. 1086-1092.

[8] M. H. Albadia and E. F. El-Saadanyb, "New method for estimating CF of pitched-regulated wind turbines," *Electric Power Systems Research*, 80, pp. 1182-1188, 2010.

[9] D. Villanueva and A. Feijoo, "Wind power distributions: A review of their applications," *Renewable and Sustainable Energy Reviews*, 14, pp. 1490-1495, 2010.

[10] Mario Ar Talaia, "Analysis of wind power potential of a region of Aveiro, Portugal," *WSEAS Transactions on Environment and Development*, Vol. 6, No. 7, pp. 519-528, Jul 2010.

[11] Henry M. K. Cheng, Yunhe Hou, and Felix Wu, "Probability distribution of the output power of wind turbine," *in Proc. The 16th International Conference on Electrical Engineering, Busan, Korea*, 11-14 Jul, 2010.

[12] J. P. Stremel, R. T. Jenkins, R. A. Babb, and W. D. Bayless, "Production costing using the cumulant method of representing the equivalent load curve," *IEEE Trans. on PAS.*, vol. PAS-99, no. 5, pp. 1947-1956, Sept/Oct 1980.

[13] Jeremy A. Bloom, "Probabilistic production costing with dependent generating sources," *IEEE Trans. Power App. Syst.*, Vol. PAS-104, No. 8, pp. 2064-2071, Aug. 1985.

[14] C. Singh and A. Lago-Conzalez, “Reliability modeling of generation systems including unconventional energy sources,” *IEEE Trans. PAS.*, Vol. PAS-104, No. 5, pp. 1049-1056, May 1985.
[15] C. Singh and Y. Kim, “An efficient technique for reliability analysis of power systems including time dependent sources,” *IEEE Trans. Power Syst.*, Vol. 3, No. 3, pp. 1090-1096, Aug 1988.
[16] Barbara G. Brown, Richard W. Katz, and Allan H. Murphy, “Time series models to simulate and forecast wind speed and wind power,” *Journal of Climate and Applied Meteorology*, Vol. 23, pp. 1184-1195, May 1984.
[17] L. L. Freris, *Wind Energy Conversion Systems*, NJ: Prentice-Hall, 1990.
[18] David A. Spera, *Wind Turbine Technology, Fundamental Concepts of Wind Turbine Engineering*, NY: ASME Press, 1994.
[19] Lata Gidwani and Harpal Tiwari, “Dynamic analysis of grid connected wind turbine with a permanent magnet synchronous generator during fault conditions,” *WSEAS Transactions on Power Systems*, Vol. 5, No. 4, pp. 287-296, Oct 2010.
[20] Brendan Fox, Damian Flynn, Leslie Bryans, Nick Jenkins, David Miborrow, Mark O’Malley, Richard Watson, and Olimpo Anaya-Lara, *Wind Power Integration, Connection and System Operation Aspects*, London: IET Power and Energy Series, 2007.
[21] Royal Netherlands Meteorological Institute. [Online]. Available: http://www.knmi.nl/klimatologie/onderzoeksgegevens/potentiele_wind/
[22] Vermont Small-scale Wind Energy Demonstration Program. [Online]. Available: http://www.vtwindprogram.org/
[23] Specifications of a Bergey xls wind turbine. . [Online]. Available: http://www.windenergydirect.org/Windgenerators-bergy-xls10kw.php
[24] Jeff Huang, Zulati Litifu, and Ken Nagasaka, “Installation of wind power into a weak power system considering system fluctuation and limitation,” *WSEAS Transactions on Environment and Development*, Vol. 4, No. 1, pp. 33-528, Jan 2008.
[25] Lars Landberg, “A mathematical look at a physical power prediction model,” *Wind Energy*, 1, pp. 23-28, 1998.
[26] Wind Atlas Analysis and Application Program. [Online]. Available: http://www.wasp.dk/index.htm
[27] Cameron W. Potter, Hugo A. Gil, and Jim McCaa, “Wind power data for grid integration studies,” *in Proc. 2007 IEEE PES General Meeting, Tampa, 24-28 June 2007.*
[28] Julija Matevosyan and Lennart Soder, “Minimization of imbalance cost trading wind power on the short-term power market,” *IEEE Trans. on Power Systems*, vol. 21, no.3, pp. 1396-1404, Aug 2006.
[29] World Wind Energy Report 2008. World Wind Energy Association, Bonn, Germany. [Online]. Available: http://www.wwindea.org/home/images/stories/worldwindenergyreport2008_s.pdf
[30] Wind Turbines in Denmark. Danish Energy Agency, Amaliegade, Copenhagen, Nov 2009.
[31] A. Stuart and J. K. Ord, *Kendall’s Advanced Theory of Statistics: Vol. 1. Distribution Theory*, sixth edition, London: Edward Arnold, 1994.

Henry Cheng (M’ 05) received his B. Eng. degree (2003) from the Department of Electrical and Electronic Engineering, and master degree in economics (2006), both at the University of Hong Kong. He is currently a Ph.D. candidate in the Center of Electrical Energy Systems at the same university.

Yunhe Hou (M’ 07) received his B.E. (1999), M.E. (2002) and Ph.D. (2005) degrees from the Huazhong University of Science and Technology, China. He worked as postdoctoral research fellow at the Tsinghua University from 2005 to 2007. He is now a research assistant professor with the Department of Electrical and Electronic Engineering at the University of Hong Kong.

Felix Wu (M’74-F’89) is the Philip Wong Wilson Wong Professor in Electrical Engineering at the University of Hong Kong, Hong Kong, where he served as Pro Vice Chancellor (Vice President) from 1997 to 2001. He is also a Professor Emeritus at the University of California, Berkeley, where he has been on the Faculty since 1974.