Versatile distribution of wind power output for a given forecast value

Zhao-Sui Zhang, *Student Member*, *IEEE*, Yuan-Zhang Sun, *Senior Member*, *IEEE*, Jin Lin, *Student Member*, *IEEE*, Lin Cheng, *Member*, *IEEE*, and Guo-Jie Li, *Member*, *IEEE*

***Abstract*—The wind power output distribution for a given forecast value is often described by a theoretical distribution, such as Gaussian, Beta and Cauchy. However, the theoretical distributions could hardly simulate the actual wind power uncertainty for all situations. Moreover, the cumulative distribution function (CDF) or quantile function is more concerned by system operators than the probability density function (PDF). Nonetheless, the CDF of a theoretical distribution usually could not be expressed in a closed form and is commonly derived from the numerical integral of PDF. The estimation error of PDF may hence be cumulatively enlarged through the numerical integral process. Given this background, this paper presents a versatile distribution model that can simulate any shape of the actual distributions of wind power forecast errors. The CDF of the versatile distribution could be written as a closed form so it can be directly applied on fitting the actual CDF. The mathematical feature of the versatile distribution can also facilitate the dispatching decision-making and benefit power system analysis. The results demonstrate the feasibility and effectiveness of the proposed distribution model.**

***Index Terms*—Cumulative distribution function (CDF), forecast, probability density function (PDF), versatile distribution, wind power**

I. Introduction

Unlike load, wind power is difficult to be accurately predicted because it is strongly correlated to the climate, season and geography. The provision of only single-point forecast values makes little sense for power system operation, so it is very important to let the forecast uncertainty of wind power being offered as well. Nowadays, probabilistic forecasts are a highly valuable input to dispatching problems related the management or trading of wind generation. The distribution of wind power output for a given forecast, also defined as the distribution of forecast errors, is progressively followed with a great deal of interest over the academia.

Previous studies have examined the distribution of forecast errors for different forecast horizons and different spatial dispersions of wind power plants. Three theoretical distributions of Gaussian, Beta and Cauchy are found effectual as addressed below: a) For ultra-short-term (1min~60min) forecast, the wind power forecast error obeys Cauchy distribution [1]. The reason is that, compared with the longer time-scale forecast, ultra-short-term forecast is relatively accurate so that the probability density function (PDF) of wind power output for the forecast value appears a peak near the forecast value. b) For short-term (1h~48h) forecast, the forecast error can be assumed to follow Beta distribution [2], [3]. Since the nonlinear transformation from wind speeds to wind powers, for small or large wind speeds/powers, the Gaussian wind speed distributions are strongly deformed and no longer symmetric [4]. c) For the large number and geographical dispersion of wind power plants with short-term forecast, Gaussian distribution better describes the statistical feature of forecast errors. Like in the case of electricity demand, a large number of multi-location wind power plants permit the invocation of the central limit theorem to justify the normality assumption of forecast errors [5], [6].

However, these theoretical distributions have limitations as claimed subsequently: a) The PDF of wind power output for the forecast value is a conditional probability curve with the forecasted wind power as average. When the forecast value is close to 0 or the installed capacity, the PDF is no longer symmetric due to the nonlinear transformation from wind speeds to wind powers [7]. At this situation, it is unreasonable to assume forecast errors to follow Gaussian and Cauchy distributions. b) If the average and standard deviation are given, the shape parameters of Gaussian and Beta distributions are fixed. Nevertheless, for the wind power plants with different forecast horizons and spatial dispersions, the shapes of the actual distributions are impossibly always the same even though they have the same average and standard deviation. For instance, for the forecast horizon in minute level, the actual PDF appears a peak near the forecasted wind power which cannot be described by Gaussian and Beta distributions [1]. c) The forecast uncertainty, commonly represented by an interval with an associated confidence level [8], is usually derived from the inverse operation of the cumulative distribution function (CDF) or quantile function. Hence, system operators are more interested in CDF than PDF and it is preferable to directly

This work was supported by National High-Technology Research and Development Program ("863" Program) of China (2011AA05A103).

Z.-S. Zhang is with the State Key Laboratory of Power Systems, Department of Electrical Engineering, Tsinghua University, Beijing 100084, China and visiting the Earth Institute, Columbia University in the City of New York, NY 10025, USA. (e-mail: zhangzhaosui@gmail.com)

Y.-Z. Sun, J. Lin, and L. Cheng are with the State Key Laboratory of Power Systems, Department of Electrical Engineering, Tsinghua University, Beijing 100084, China. (e-mail: yzsun@tsinghua.edu.cn; tulinjin@gmail.com; chenglin@tsinghua.edu.cn)

G.-J. Li is with Key Laboratory of Control of Power Transmission and Transformation, Ministry of Education (Department of Electrical Engineering, Shanghai Jiao Tong University), Shanghai 200240, China. (e-mail: liguojie@sjtu.edu.cn)

978-1-4673-2729-9/12/$31.00 ©2012 IEEE

simulate the actual CDF with a pre-defined distribution. Unfortunately, the CDFs of Gaussian and Beta distributions usually do not have the closed forms. One can only deduce the CDF by the numerical integral of its corresponding PDF. However, the estimation error due to the PDF simulation would be further enlarged through the numerical integral. Also, numerical integral expands the computing scale and impacts the computing speed.

Given the above background, this paper presents a versatile distribution model to simulate the actual distribution of wind power output for a given forecast value. The proposed distribution could overcome the above three limitations and combine the advantages of Gaussian, Beta and Cauchy distributions. Section II raises the PDF and CDF formulas of the versatile distribution and outlines their mathematical properties. In Section III, the comparisons among the versatile, Gaussian, Beta and Cauchy distributions are studied to find the most suitable one that fit the actual distribution. Section IV proceeds with some examples of how this work can benefit power system analysis. Section V and VI finally discuss and conclude this paper.

II. Formulation of Versatile Distribution Model

The PDF and CDF formulas of versatile distribution are newly presented as shown in (1) and (2), respectively.

$$\mathrm{PDF}(x \mid a,b,c) = \frac{abe^{-a(x-c)}}{(1+e^{-a(x-c)})^{b+1}} \tag{1}$$

$$\mathrm{CDF}(x \mid a,b,c) = (1+e^{-a(x-c)})^{-b} \tag{2}$$

where x is the probabilistic wind power output and a, b, c are the shape parameters for a given forecast value. This novel distribution has very valuable properties revealed from the following two aspects:

a) The versatile distribution can be deformed owing to the adjustable shape parameters a, b and c. In other words, this distribution can well simulate the actual distribution of forecast errors for any forecast horizon and any spatial configuration through regulating its shape parameters. For a wind power plant with a certain forecast time-scale, one can generate the assorted shape parameters for each wind power forecast value (such as Table I). Therefore, the versatile distribution can significantly improve the accuracy of curve fitting.

b) In practical applications, the wind power forecast profiles with a corresponding confidence level are usually gained from the deduction of CDF. However, the CDFs of Gaussian and Beta distributions include the errors due to the numerical integral of PDF. By contrast, the versatile distribution CDF as shown in (2) is preferable because it has a closed form so that it can directly simulate the actual CDF. Thus, the proposed distribution considerably improves the computing efficiency and reduces the estimation errors.

III. Methodology and Results

This section compares the versatile, Gaussian, Beta and Cauchy distributions through analyzing the data of a real wind power plant. We firstly introduce the methodology how to generate the actual/simulated PDF and CDF for a specific wind power forecast. Then, the qualitative and quantitative estimations are investigated for the purpose of selecting the best distribution fitting for wind power forecast errors. The forecast confidence intervals, generated based on different distributions, are also validated at the end of this section.

A. Data processing

The 5-minute time series of wind power output in 2008, provided by a 150 MW wind power plant in the Inner Mongolia Autonomous Region, China, are used in this paper. The wind power is standardized as a fraction within [0, 1] of the installed capacity. The 15-minute-ahead, 1-hour-ahead and 4-hour-ahead wind power forecast values are generated with the persistence forecast method [9]. This study assumes that the parameter of forecast delay in the persistence forecast model is set as 15 minutes. The methodology for producing the actual/simulated PDF and CDF for a given wind power forecast value is summarized as the following steps [10]:

a) For a certain forecast horizon (15 minutes, 1 hour or 4 hours in this case), rearrange the data based on the wind power forecast values in an ascending order.

b) Divide the forecast data into N_f bins. In this study, N_f=25, so the width of each forecast bin is 1/25=0.04 p.u.. For example, the range of forecast value is within [0.48, 0.52] for No. 13 forecast bin and within [0.16, 0.2] for No. 5 forecast bin. Note that the size of N_f relies on the scale of the historical data. The more historical data are introduced, the more forecast bins can be built [11].

c) For each forecast bin, there is a group of associated wind power outputs. The actual PDF is equivalent to the possibility density histogram produced through the statistical analysis of these output values. The actual CDF is hence derived from the discrete integral of the histogram.

d) Calculate the average and standard deviation of the wind power outputs for each forecast bin. The corresponding parameters for Gaussian and Beta distributions are therefore derived as introduced in [3].

e) For the versatile and Cauchy distributions, the shape parameters can only be achieved from curve fitting for the actual PDF/CDF. Notice that the shape parameters obtained by simulating the actual PDF are different from those gained from simulating the actual CDF. With an example of 1-hour-ahead forecast based on CDF curve fitting approach, one can establish a look-up table of shape parameters for the simulated versatile distribution (see Table I).

B. Qualitative estimation for possible distributions of wind power forecast errors

In the qualitative estimation, two cases are studied to observe the distribution which best fits the actual stochastic characterization of wind power output for a given forecast. In Case 1, the forecast value is located within [0.48, 0.52] and the forecast horizon is 15 minutes. Fig.1 and Fig. 2 show the PDFs and CDFs of the wind power output for the 15-minute-ahead forecast within [0.48, 0.52], respectively. The shape parameters of the versatile and Cauchy distributions are determined based on PDF curve fitting method. As shown in Fig. 1, the actual PDF appears a peak when the wind power output reaches nearly 0.5 p.u.. This peak is much better outlined by the versatile and Cauchy distributions compared with Gaussian and Beta distributions. However, Fig.1 and Fig. 2 both illustrate that Cauchy distribution has fat tails which indicate large deviations at the ends of the distribution. Therefore, the versatile distribution achieves the best result for fitting the actual random feature of the forecast uncertainty. In fact, PDF curve fitting method is of little significance since system operators are more interested in CDF. The shape parameters of the versatile and Cauchy distributions are newly generated through the direct curve fitting for the actual CDF. As shown in Fig. 3, the newly generated Cauchy CDF moves closer to the actual CDF while the new versatile CDF almost performs a perfect simulation.

TABLE I
LOOK-UP TABLE OF SHAPE PARAMETERS OF VERSATILE DISTRIBUTION FOR THE WIND POWER PLANT WITH 1-HOUR-AHEAD FORECAST BASED ON CDF CURVE FITTING METHOD

Forecast bin No.	Forecast value	Shape parameters *a*	*b*	*c*
1	[0.00, 0.04]	30.62	344.06	-0.21
2	[0.04, 0.08]	29.91	487.38	-0.18
3	[0.08, 0.12]	25.25	5.04	0.01
4	[0.12, 0.16]	21.95	4.05	0.04
5	[0.16, 0.20]	19.26	3.45	0.09
6	[0.20, 0.24]	15.66	5.44	0.07
7	[0.24, 0.28]	16.59	2.35	0.17
8	[0.28, 0.32]	16.19	1.70	0.23
9	[0.32, 0.36]	15.82	1.69	0.26
10	[0.36, 0.40]	13.77	1.70	0.31
11	[0.40, 0.44]	15.80	1.29	0.37
12	[0.44, 0.48]	14.85	1.20	0.41
13	[0.48, 0.52]	14.93	1.28	0.43
14	[0.52, 0.56]	14.22	1.05	0.51
15	[0.56, 0.60]	18.28	0.90	0.56
16	[0.60, 0.64]	17.21	0.70	0.62
17	[0.64, 0.68]	20.79	0.63	0.66
18	[0.68, 0.72]	21.52	0.60	0.70
19	[0.72, 0.76]	22.85	0.62	0.72
20	[0.76, 0.80]	21.85	0.65	0.77
21	[0.80, 0.84]	26.19	0.37	0.83
22	[0.84, 0.88]	67.43	0.17	0.91
23	[0.88, 0.92]	110.01	0.18	0.93
24	[0.92, 0.96]	88.16	0.49	0.93
25	[0.96, 1.00]	327.23	0.16	0.98

In Case 2, the forecast value is located within [0.16, 0.2] and the forecast horizon is 4 hours. Having the forecast value as average, the PDF of wind power output distorts a little as shown in Fig. 4. However, Gaussian and Cauchy could not simulate this distortion. The Beta PDF does not fit the actual PDF as well since it over skews although towards the correct direction. Compared among all the distributions, the versatile distribution performs best for fitting the actual PDF and CDF as shown in Fig. 4 and Fig. 5. Similar to Case 1, the versatile distribution acts even better as shown in Fig. 6 if the shape parameters are generated based on the CDF curve fitting method.

In view of the importance of CDF in practical applications, the shape parameters of the versatile and Cauchy distributions will be designed based on CDF curve fitting method in the following sections.

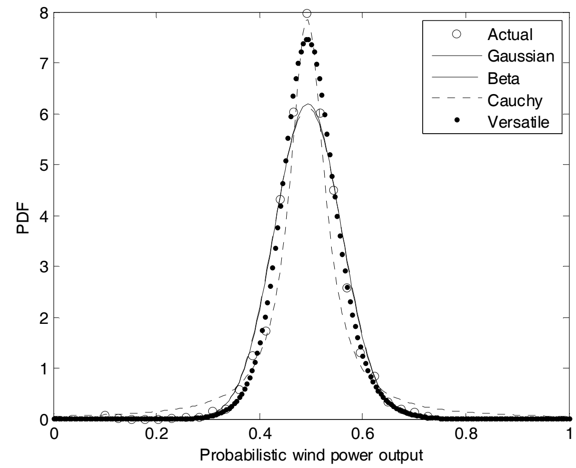

Fig. 1. PDFs of wind power output for the 15-minute-ahead forecast within [0.48, 0.52] (No. 13 forecast bin) based on PDF curve fitting method.

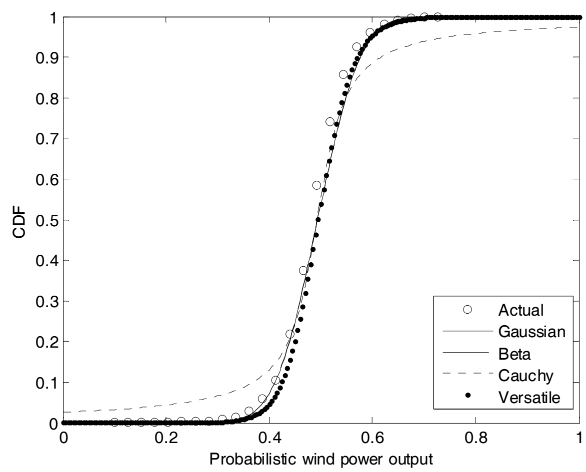

Fig. 2. CDFs of wind power output for the 15-minute-ahead forecast within [0.48, 0.52] (No. 13 forecast bin) based on PDF curve fitting method.

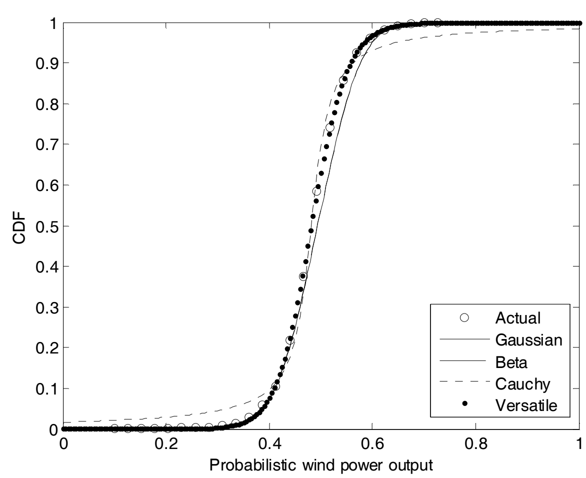

Fig. 3. CDFs of wind power output for the 15-minute-ahead forecast within [0.48, 0.52] (No. 13 forecast bin) based on CDF curve fitting method.

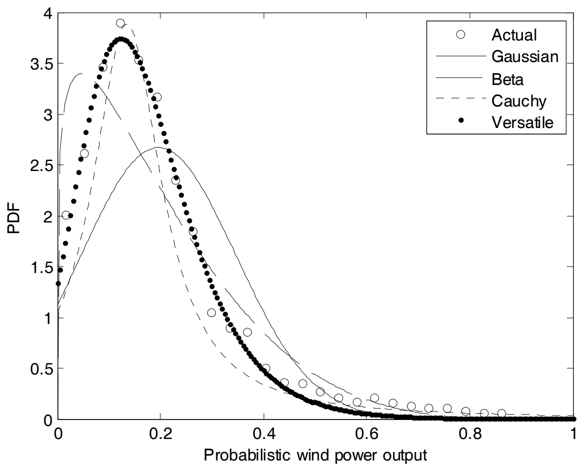

Fig. 4. PDFs of wind power output for the 4-hour-ahead forecast within [0.16, 0.2] (No. 5 forecast bin) based on PDF curve fitting method.

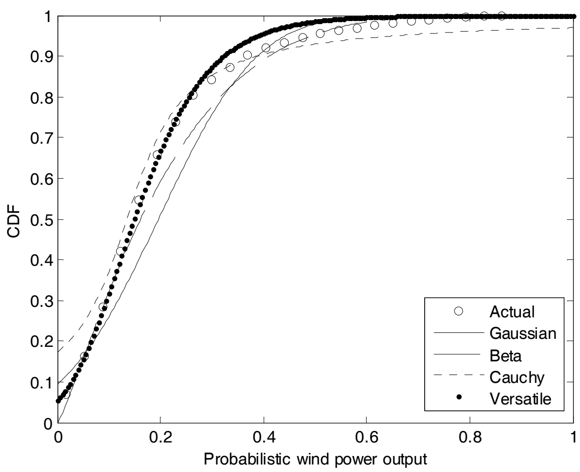

Fig. 5. CDFs of wind power output for the 4-hour-ahead forecast within [0.16, 0.2] (No. 5 forecast bin) based on PDF curve fitting method.

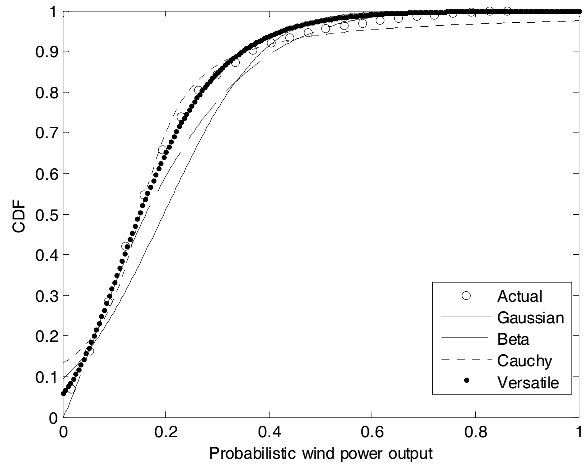

Fig. 6. CDFs of wind power output for the 4-hour-ahead forecast within [0.16, 0.2] (No. 5 forecast bin) based on CDF curve fitting method.

C. Quantitative estimation for possible distributions of wind power forecast errors

In order to facilitate the comparison of four possible distributions, the root-mean-square error (RMSE) between the actual CDF and simulated CDF is used for quantitative estimation. Similar to the forecast data, the wind power output is separated into N_o bins and here we suppose N_o=N_f=25. RMSE is defined as (3):

$$\mathrm{RMSE}_j = \sqrt{\frac{1}{N_o}\sum_{i=1}^{N_o}(\mathrm{CDF}_{act,j}(\frac{i}{N_o}) - \mathrm{CDF}_{sim,j}(\frac{i}{N_o}))^2} \qquad (3)$$

where j is the forecast bin number, $\mathrm{CDF}_{act,j}$ is the actual CDF for the jth forecast bin and $\mathrm{CDF}_{sim,j}$ is the simulated CDF for the jth forecast bin.

TABLE II
RMSEs FOR DIFFERENT FORECAST HORIZONS, DIFFERENT DISTRIBUTIONS AND DIFFERENT WIND POWER FORECAST VALUES

Forecast bin No.	Forecast value	RMSE (15-minute-ahead forecast)				RMSE (1-hour-ahead forecast)				RMSE (4-hour-ahead forecast)			
		Gaussian	Beta	Cauchy	Versatile	Gaussian	Beta	Cauchy	Versatile	Gaussian	Beta	Cauchy	Versatile
1	[0.00, 0.04]	0.068	0.024	0.025	0.007	0.080	0.026	0.015	0.002	0.088	0.021	0.016	0.007
2	[0.04, 0.08]	0.050	0.041	0.024	0.001	0.067	0.046	0.016	0.007	0.079	0.028	0.020	0.014
3	[0.08, 0.12]	0.041	0.033	0.035	0.007	0.056	0.038	0.027	0.008	0.070	0.030	0.024	0.012
4	[0.12, 0.16]	0.054	0.048	0.033	0.002	0.056	0.041	0.029	0.006	0.066	0.029	0.030	0.014
5	[0.16, 0.20]	0.049	0.044	0.030	0.008	0.054	0.042	0.029	0.008	0.069	0.037	0.028	0.015
6	[0.20, 0.24]	0.042	0.036	0.034	0.002	0.052	0.040	0.035	0.006	0.056	0.033	0.031	0.008
7	[0.24, 0.28]	0.046	0.042	0.031	0.005	0.045	0.036	0.039	0.006	0.065	0.039	0.032	0.012
8	[0.28, 0.32]	0.044	0.041	0.034	0.006	0.045	0.038	0.035	0.005	0.049	0.027	0.039	0.005
9	[0.32, 0.36]	0.042	0.040	0.033	0.004	0.040	0.035	0.036	0.006	0.048	0.033	0.040	0.011
10	[0.36, 0.40]	0.048	0.046	0.035	0.004	0.041	0.037	0.036	0.005	0.045	0.031	0.041	0.009
11	[0.40, 0.44]	0.043	0.042	0.034	0.006	0.040	0.038	0.033	0.008	0.041	0.031	0.039	0.007
12	[0.44, 0.48]	0.046	0.046	0.031	0.006	0.037	0.036	0.037	0.005	0.028	0.032	0.035	0.010
13	[0.48, 0.52]	0.035	0.035	0.034	0.004	0.037	0.037	0.037	0.005	0.028	0.031	0.038	0.009
14	[0.52, 0.56]	0.039	0.040	0.033	0.005	0.031	0.035	0.038	0.005	0.028	0.030	0.039	0.011
15	[0.56, 0.60]	0.034	0.035	0.038	0.008	0.029	0.032	0.042	0.008	0.023	0.027	0.044	0.010
16	[0.60, 0.64]	0.028	0.030	0.043	0.009	0.029	0.034	0.039	0.007	0.029	0.027	0.051	0.020
17	[0.64, 0.68]	0.033	0.036	0.041	0.006	0.025	0.029	0.043	0.008	0.026	0.034	0.036	0.008
18	[0.68, 0.72]	0.035	0.038	0.031	0.008	0.028	0.033	0.039	0.006	0.030	0.034	0.041	0.011
19	[0.72, 0.76]	0.041	0.045	0.031	0.006	0.029	0.034	0.038	0.007	0.040	0.041	0.042	0.009
20	[0.76, 0.80]	0.028	0.034	0.040	0.007	0.026	0.034	0.037	0.009	0.025	0.027	0.051	0.023
21	[0.80, 0.84]	0.031	0.037	0.043	0.012	0.030	0.031	0.038	0.006	0.052	0.032	0.062	0.029
22	[0.84, 0.88]	0.033	0.034	0.039	0.009	0.048	0.036	0.051	0.012	0.047	0.035	0.053	0.016
23	[0.88, 0.92]	0.037	0.040	0.034	0.004	0.051	0.045	0.037	0.014	0.084	0.075	0.056	0.034
24	[0.92, 0.96]	0.041	0.048	0.037	0.013	0.047	0.054	0.023	0.014	0.099	0.082	0.048	0.046
25	[0.96, 1.00]	0.037	0.039	0.046	0.028	0.070	0.065	0.033	0.023	0.075	0.052	0.043	0.023

Table II shows the comparisons of RMSEs based on different forecast horizons, different distributions and different wind power forecast values. It can be clearly found that every RMSE of the versatile distribution (in shadow) reaches minimum compared with that of Gaussian, Beta or Cauchy for the same condition. Therefore, the versatile can best approximate the actual distribution of wind power forecast errors.

D. Determination of forecast confidence intervals

The needs from many system operators in wind power prediction always contain an accurate forecast value with an associated confidence level [8]. The information of confidence interval for a given wind power forecast is very important for the end users to develop decision support tools in some specific areas, like trading, scheduling or reserve allocation.

Taking the 1-hour-ahead forecast as an example, Fig. 7 shows the relationships between the 90% confidence interval and wind power forecast. The upper bundle of curves implies the function of the upper boundary of confidence interval versus the wind power forecast value. The lower bundle indicates the function of the lower boundary of confidence interval with respect to the wind power forecast value. For instance, when the wind power forecast value (x-axis) equals 0.48, the confidence intervals (y-axis) for the actual distribution are 0.25 p.u. (lower bound) and 0.63 p.u. (upper bound), which can be conveniently looked up from Fig. 7. For the same forecast value, the confidence intervals span 0.25~0.62 p.u. for the versatile distribution, 0.26~0.64 p.u. for Gaussian distribution, 0.27~0.66 p.u. for Beta distribution and 0.11~0.75 p.u. for Cauchy distribution. In Fig. 7, the versatile, Gaussian, Beta distribution confidence intervals all meet the actual confidence intervals well. Similar results can also be found from the case of 95% confidence level as shown in Fig. 8. As one can observe, for most of the points in Fig.7 and Fig. 8, the confidence intervals generated based on the versatile distribution follow the actual intervals best. However, the estimations for 90% and 95% confidence intervals are both too optimistic for Cauchy distribution. The reason is that the PDF of Cauchy distribution has fat tails which leads to wide bounds for high percentage confidence intervals [1]. This characteristic may mislead the system operators to provide many unneeded spinning reserves for wind power plants. By contrast, for the case of 80% confidence level, the estimation for the confidence interval using Cauchy distribution performs much better as shown in Fig. 9. In this condition, the versatile distribution confidence interval almost coincides with the actual distribution confidence interval.

In summary, the system operators may make more or less inaccurate decision in generation scheduling or reserve allocating if the confidence intervals are derived based on Gaussian, Beta and Cauchy distributions. The confidence interval for each wind power forecast value can also be examined using the same method mentioned above for the other forecast horizons. Similar results can be found for the cases of 15-minute-ahead and 4-hour-ahead forecasts. Thereby, this paper will not further study these cases.

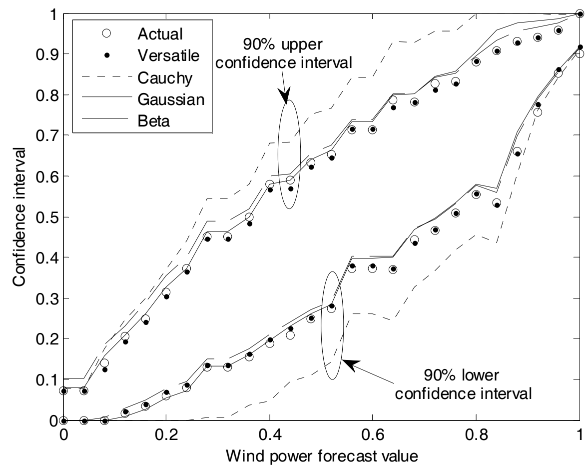

Fig. 7. 90% confidence interval against 1-hour-ahead wind power forecast. The upper/lower bundle of curves implies the function of the upper/lower boundary of confidence interval versus the wind power forecast value.

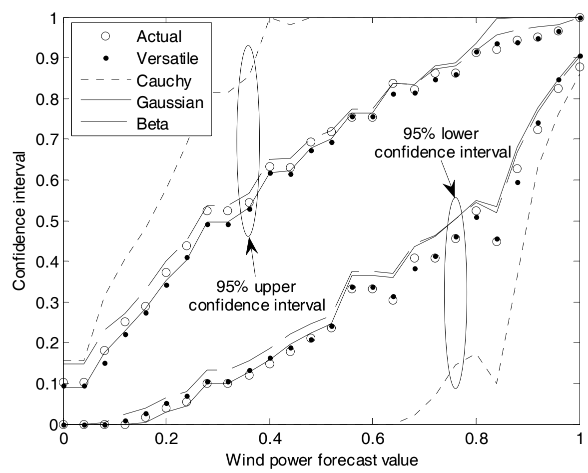

Fig. 8. 95% confidence interval against 1-hour-ahead wind power forecast. The upper/lower bundle of curves implies the function of the upper/lower boundary of confidence interval versus the wind power forecast value.

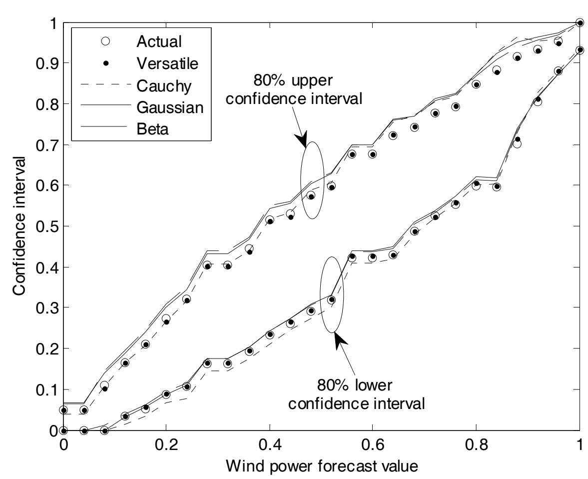

Fig. 9. 80% confidence interval against 1-hour-ahead wind power forecast. The upper/lower bundle of curves implies the function of the upper/lower boundary of confidence interval versus the wind power forecast value.

IV. Benefits for Power System Analysis Due to the Application of Versatile Distribution Model

Except the good features aforementioned, the application of versatile distribution model can also benefit the power system operation and control. The most important benefit is that the versatile distribution CDF has analytical inverse function like (4), which is deduced from (2).

$$\mathrm{CDF}^{-1}(y \mid a,b,c) = c - \frac{1}{a}\ln(y^{-\frac{1}{b}} - 1) \quad (4)$$

where y is the CDF value.

This property facilitates system operators to obtain the confidence interval in an analytical way, which provides a helpful auxiliary tool for power system analysis. Once the confidence level is given, the upper and lower confidence intervals as shown in (5) and (6) could be analytically deduced from the inverse function (4), respectively. By contrast, it is much less convenient to achieve the confidence interval for the theoretical distribution, such as Gaussian and Beta. The reason is that there are no closed forms for the CDF and its inverse function to these theoretical distributions. It should be claimed that the confidence intervals generated by Gaussian and Beta distributions appeared in Section III-D are derived through numerical calculation.

$$p_{f,up} = c - \frac{1}{a}\ln[(0.5+0.5*l)^{-\frac{1}{b}} - 1], \quad (p_{f,up} <= 1) \quad (5)$$

$$p_{f,low} = c - \frac{1}{a}\ln[(0.5-0.5*l)^{-\frac{1}{b}} - 1], \quad (p_{f,low} >= 0) \quad (6)$$

where $p_{f,up}$ and $p_{f,low}$ (in per-unit system) are the upper and lower boundaries of the confidence interval of wind power forecast, respectively, and l is the forecast confidence level. The only data supported by wind power forecast systems are the shape parameters of a, b and c.

The property of the existence of CDF inverse function is also able to facilitate the economic dispatch programming, where the wind power uncertainty is usually considered with a chance constrained formulation [12]. For example, we hope the allocated spinning reserve could compensate the bias caused by the practical wind power output deviating below the scheduled wind power at the scheduled time. This requirement could be described with a chance constraint under a specific confidence level which is given by (7).

$$\Pr\{r \geq x_s - x\} \geq \alpha \quad (7)$$

where r is the spinning reserve amount, x is the probabilistic wind power output, x_s is the scheduled wind power (dispatch instruction) and α is the security confidence level. Rewrite (7), we obtain:

$$\Pr\{r \geq x_s - x\} = \Pr\{x \geq x_s - r\} = 1 - \mathrm{CDF}(x_s - r) \geq \alpha \quad (8)$$

Then the chance constraint is transformed to (9).

$$x_s - r \leq \mathrm{CDF}^{-1}(1-\alpha) \quad (9)$$

where $\mathrm{CDF}^{-1}(\cdot)$ could be derived from (4). Therefore, the application of the versatile distribution enables a chance constraint to be transferred to a linear constraint. This greatly simplifies the economic dispatch problems incorporating wind power uncertainty.

V. Discussions

As proposed in [13], the actual distribution of wind power forecast errors, precisely expressed by a probability density histogram, is directly applied in the optimal power flow analysis. One may wonder why this paper does not use the actual distribution in a similar way. This reason is that the histogram is a discrete function represented with a set of pairs. Each pair contains a power value and its corresponding probability. Apparently, discrete computing techniques increase the quantity of computing variables and reduce the computing speed, especially when a high computing accuracy is required. Therefore, it is deemed as a future trend to fit the histogram bars with a mathematical distribution and transfer the discrete function to a continuous one [11].

Although this study results in very good characteristics of versatile distribution in power system analysis, some issues are still in pressing need to be concerned. For example, it is unreasonable to obtain the wind power forecast uncertainty using the historical power data simply from the electrical point of view. The wind power forecast uncertainty is correlated with many random elements and can hardly be estimated in such easy way. Thus, how to take all the factors, like climate, season and geography, into account needs to be further investigated. For another example, to multiple wind power plants, it is difficult to obtain the aggregated CDF and PDF based on the implementation of versatile distribution, particularly when these wind power plants are not independent.

VI. Conclusion

The commonly used theoretical distributions, such as Gaussian, Beta and Cauchy, are not able to describe the statistical characterization of wind power forecast errors for all situations. In order to overcome this shortcoming, this paper presents a versatile distribution which could simulate any shape of the actual distribution of wind power forecast errors. The qualitative and quantitative comparisons among the versatile, Gaussian, Beta and Cauchy distributions all demonstrate that the proposed distribution is the most suitable one to fit the actual PDF/CDF. The confidence intervals are also determined for a set of wind power forecast values. The results show that estimating the confidence intervals based on Gaussian, Beta and Cauchy distributions may lead to more or less inaccurate decision-making in power system scheduling.

The versatile distribution also shows another good mathematical property underlying that, its CDF and the inverse function of CDF both have analytical forms. This property facilitates the power system operators to derive the confidence level and simplifies the economic dispatch programming to some extent.

The proposed versatile distribution model is able to be applied not only on the description of the statistical feature of wind power forecast errors but also on the uncertainty analysis research in many scientific fields. The application of the versatile distribution in power system analysis will be further developed by the authors in future.

VII. References

[1] B-M. Hodge, and M. Milligan, “Wind power forecasting error distributions over multiple timescales,” National Renewable Energy Laboratory, Golden, CO, USA, NREL/CP-5500-50614, Mar. 2011.

[2] S. Bofinger, A. Luig, and H. G. Beyer, “Qualification of wind power forecasts,” in *Proc. Global Wind Power Conf.*, Paris, France, 2002.

[3] A. Fabbri, T. G. S. Roman, J. R. Abbad, and V. H. M Quezada, “Assessment of the cost associated with wind generation prediction errors in a liberalized electricity market,” *IEEE Trans. on Power Syst.*, vol. 20, no. 3, pp. 1440-1446, Aug. 2005.

[4] M. Lange, “On the uncertainty of wind power predictions—analysis of the forecast accuracy and statistical distribution of errors,” *Journal of Solar Energy Engineering*, vol. 127, pp. 177-184, May. 2005.

[5] R. Doherty, and M. O’Malley, “A new approach to quantify reserve demand in systems with significant installed wind capacity,” *IEEE Trans. on Power Syst.*, vol. 20, no. 2, pp. 587-595, May. 2005.

[6] F. Bouffard, and F. D. Galiana, “Stochastic security for operations planning with significant wind power generation,” *IEEE Trans. on Power Syst.*, vol. 23, no. 2, pp. 306-316, May. 2008.

[7] Z.-S. Zhang, Y.-Z. Sun, G.-J. Li, L. Cheng, and J. Lin, “A solution of the economic dispatch problem considering wind power uncertainty,” *Automation of Electric Power Systems*, vol. 35, no. 22, pp. 125-130, Nov. 2011.

[8] G. Giebel, R. Brownsword, G. Kariniotakis, M. Denhard, and C. Draxl. (2011, Jan.). The state-of-the-art in short-term prediction of wind power: A literature overview. Deliverable D-1.2 of ANEMOS project. [Online]. Available: http://130.226.56.153/zephyr/publ/GGiebelEtAl-StateOfTheArtInShortTermPrediction_ANEMOSplus_2011.pdf

[9] H. Bludszuweit, J. A. Domínguez-Navarro, and A. Llombart, "Statistical analysis of wind power forecast error," *IEEE Trans. on Power Syst.*, vol. 23, no. 3, pp. 983-991, Aug. 2008.

[10] A. T. Al-Awami, and M. A. El-Sharkawi, "Statistical characterization of wind power output for a given wind power forecast", in *Proc. North American Power Symposium*, Starkville, MS, USA, 2009.

[11] Z.-S. Zhang, Y.-Z. Sun, J. Lin, V. Modi, L. Cheng, and G. Chen, "Innovation for wind power: Adding smart processors between prediction systems and dispatch systems", in *Proc. IEEE PES Transmission and Distribution Conference and Exposition*, Orlando, Florida, USA, May. 2012.

[12] M. H. Alabadi, and E. F. El-Saadany, “Comparative study on impacts of wind profiles on thermal units scheduling costs,” *IET Renew. Power Gener.*, vol. 5, no. 1, pp. 26-35, 2011.

[13] R. A. Jabr, and B. C. Pal, “Intermittent wind generation in optimal power flow dispatching,” *IET Gener. Transm. Distrib.*, vol. 3, no. 1, pp. 66-74, 2009.

VIII. Biographies

Zhao-Sui Zhang received the B.S. degree in electrical engineering from Tsinghua University, Beijing, China in 2009. He then continues to pursue a Ph.D. degree at Tsinghua University. He is now sponsored by China Scholarship Council (CSC) to be a joint Ph.D. student at Columbia University in the City of New York. His main research interests are wind power, power systems, power economics and energy policy.

Yuan-Zhang Sun (M’99－SM’01) received the B.S. degree from Wuhan University of Hydro and Electrical Engineering, China, the M.S. degree from Electric Power Research Institute (EPRI), China, and Ph.D. degree from Tsinghua University, Beijing, China, in 1978, 1982 and 1988, respectively, all in electrical engineering. He is now a chair professor at Tsinghua University and dean of School of Electrical Engineering at Wuhan University. His main research interests are in the areas of power systems dynamics and control, voltage stability and control, and renewable energy.

Jin Lin received the B.S. degree and Ph.D. degree in electrical engineering from Tsinghua University, Beijing, China, in 2007 and 2012, respectively. He is now a post-doctorate at Tsinghua University. His main research interests are renewable energy and the stability analysis on power systems.

Lin Cheng (M'05) received the B.S. degree from Tianjin University, the Ph.D. degree from Tsinghua University, Beijing, China, in 1996, and 2001, respectively. Currently, he is an associate professor of Electrical Engineering Department at Tsinghua University. His main research interests are in the areas of power system reliability, power system dynamics and control, voltage stability and control.

Guo-Jie Li (M’09) received his B.S. and M.S. degrees in Electrical Engineering from Tsinghua University, Beijing, China in 1989 and 1993, respectively. He also received Ph. D. degree in the School of Electrical and Electronic Engineering, Nanyang Technological University Singapore in 1999. He was an associate professor in the Department of Electrical Engineering, Tsinghua University, Beijing, China. He is now a professor in the Department of Electrical Engineering, Shanghai Jiaotong University, Shanghai, China. His current research interests include power system analysis and control, wind power analysis and control, and FACTS.