Probabilistic Modeling and Statistical Characteristics of Aggregate Wind Power

H. Louie and J. M. Sloughter

Abstract The stochasticity of the electrical power output by wind turbines poses special challenges to power system operation and planning. Increasing penetration levels of wind and other weather-driven renewable resources exacerbate the uncertainty and variability that must be managed. This chapter focuses on the probabilistic modeling and statistical characteristics of aggregated wind power in large electrical systems. The mathematical framework for probabilistic models—accounting for geographic diversity and the smoothing effect—is developed, and the selection and application of parametric models is discussed. Statistical characteristics from several real systems with high levels of wind power penetration are provided and analyzed.

Keywords Copulas · Correlation · Geographic diversity · Smoothing effect · Wind generators · Wind power modeling

1 Introduction

Wind turbines are classified as weather-driven renewable resources due to the dependency of their power output on local meteorological conditions [1]. These conditions are inherently transient and often erratic. Consequently, the power output by wind turbines—hereafter also simply referred to as "wind power"—is appropriately characterized as being *variable* and *uncertain*. Variability refers to

H. Louie (✉)
Department of Electrical and Computer Engineering, Seattle University, 901 12th Ave, Seattle WA 98122, USA
e-mail: louieh@seattleu.edu

J. M. Sloughter
Department of Mathematics, Seattle University, 901 12th Ave, Seattle WA 98122, USA
e-mail: sloughtj@seattleu.edu

J. Hossain and A. Mahmud (eds.), *Large Scale Renewable Power Generation*, Green Energy and Technology, DOI: 10.1007/978-981-4585-30-9_2,

the unintentional tendency for wind power to change—perhaps rapidly—from one moment to the next, whereas uncertainty refers to the wide range of unknown future values of wind power.

The stochasticity of wind power is a concern for system operators, as the legacy electric grid was designed to be operated with primarily deterministic sources [2, 3]. Although stochastic, wind power often exhibits identifiable patterns and quantifiable statistical distributions, which can be modeled and exploited to better manage the system. These models, whether mathematically formalized or tacitly understood, have applications in several areas, including wind power forecast systems, stochastic unit-commitment programs, risk analysis, and Monte Carlo-based simulations for resource planning and research [4–6].

This chapter focuses on the *aggregate* system-wide wind power, rather than the wind power from individual wind plants or turbines. We are motivated to take this macro-level view because for many system operators it is the aggregate—not individual—wind power that is of utmost concern. Our goal is to identify and develop probabilistic models of aggregate wind power and analyze its statistical characteristics. More specifically, we use parametric distributions—probability density functions (pdf) and cumulative distribution functions (cdf)—to model the instantaneous and moment-to-moment variations of aggregate wind power.

The remainder of this chapter is organized as follows. Section 2 describes the general characteristics of aggregate wind power. Section 3 formulates an idealized probabilistic model of wind power output from an individual wind plant. Aspects of geographic diversity including correlation, dependency structures, and practical considerations are discussed in Sect. 4, leading to probabilistic models for instantaneous and moment-to-moment wind power variation in Sect. 5. Aggregate wind power data from four large systems are analyzed and discussed in Sect. 6. The concluding remarks are given in Sect. 7.

2 General Characteristics of Aggregate Wind Power

Aggregate wind power is defined as the sum of the real power delivered by all wind plants in a system as measured at their point of interconnection with the grid. We are concerned with both the instantaneous and moment-to-moment variations of aggregate wind power. The statistical characteristics of instantaneous aggregate wind power provide information about the uncertainty, whereas the statistical characteristics of moment-to-moment variation of aggregate wind power provide information about the variability. Rather than formulating models in the power domain, it is more useful to do so in the normalized power domain. This facilitates easier comparison and allows the models to be scaled to the desired capacity level, increasing their applicability. The units in the normalized power domain are per-unit (p.u.), where the normalization is done with respect to the total capacity of the wind plants in the system.

The characteristics of aggregate wind power are derived from—but different than—the characteristics of wind power from individual wind plants. Aggregate wind power is strongly influenced by the *geographic diversity* of the wind plants in the system. Geographic diversity is a term describing the tendency for wind plants in different wind regimes and separated by large distances to exhibit low correlation in their instantaneous wind power and moment-to-moment variations.

Many aspects of geographic diversity have been widely explored in the literature [7–21]. All other things being equal, a system with high geographic diversity has lower variability and uncertainty than one with low geographic diversity. The reduction of variability caused by geographic diversity is also known as the *smoothing effect*. The benefits of high geographic diversity include less frequent occurrences of extremely high and low power output; less frequent ramp events; and improved accuracy of wind power forecasts. The results are greater economy and reliability, decreased reserve requirements, and more efficient commitment and dispatch of generators [22].

2.1 Uncertainty of Aggregate Wind Power

We first consider the uncertainty characteristics of aggregate wind power. The definition of uncertainty can be subjective, with several appropriate interpretations possible, depending on the application or situation. For example, a system operator may be concerned about the probability of extremely high or low aggregate wind power. In this case uncertainty is best measured using quantiles. Another operator may be interested in the general spread of potential values of aggregate wind power, in which case the standard deviation is an appropriate measure. Rather than strictly defining uncertainty, our approach is to recognize that the uncertainty information of aggregate wind power is contained in its probability density function, from which the quantiles, standard deviation, and other metrics of uncertainty can be measured or computed.

The shape of the probability density function can be approximated by constructing an empirical histogram of instantaneous aggregate wind power. Figure 1a shows a typical histogram of wind power from an individual wind plant and is provided for comparison purposes. Figure 1b–d shows aggregate wind power from three large systems. We will discuss the specific details of these systems and others in greater detail in Sect. 6. For now it suffices to know that each system has over 4 GW of installed wind capacity—a considerable amount. The computed standard deviation is provided with each plot. From inspection of Fig. 1, we make two important observations: (1) the modality of the wind power from an individual wind plant is different from that of wind power aggregated across a large system, and (2) different systems exhibit different uncertainty characteristics.

By most measures, the system corresponding to Fig. 1b has higher uncertainty than other systems. The standard deviation is greater than in other systems, and there are more frequent occurrences of zero and near rated (1.0 p.u.) power

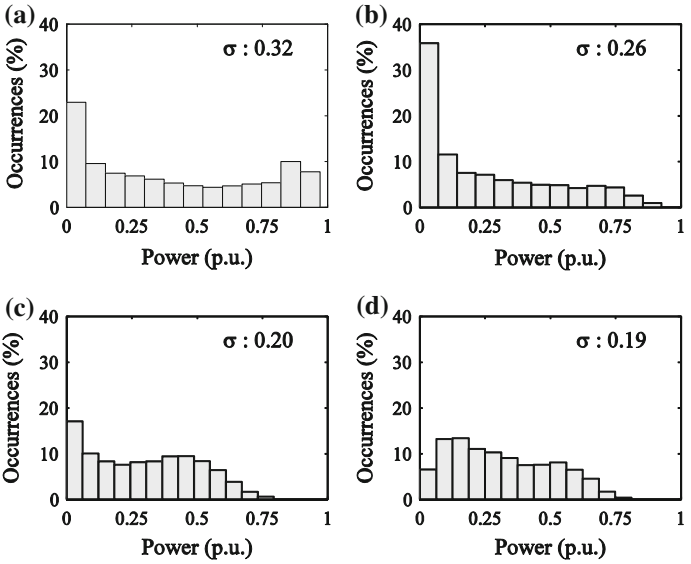

Fig. 1 Histograms of normalized wind power from an individual wind plant (**a**), and from large systems (**b–d**) with standard deviation displayed. **a** Single wind plant, **b** Lower diversity, **c** Medium diversity and **d** Higher diversity

production. These characteristics are similar to those of the individual wind plant, and as such the system can be described as having low geographic diversity. The wind plants in systems with low geographic diversity are often in close proximity—perhaps separated by 200 km or less—and are in the same or similar wind regime. The Bonneville Power Administration (BPA) is an example of a system with low geographic diversity.

The system corresponding to Fig. 1c exhibits less uncertainty than Fig. 1b. The standard deviation and occurrences of low output are reduced, and the maximum power output rarely exceeds 0.75 p.u. Figure 1d exhibits the lowest uncertainty of the systems, which is characteristic of a system with appreciable geographic diversity. Episodes of extremely low and high production are rare, and the standard deviation is lower than the others. Systems with this level of geographic diversity tend to have wind plants spread over very large territories. The Midwest ISO (MISO) and PJM systems are examples of systems with higher geographic diversity.

It is evident that there is no proto-typical histogram or probability density function of instantaneous aggregate wind power, and so the uncertainty will vary depending on system specifics—mainly the level of geographic diversity. Our approach, therefore, is to seek a flexible multi-parametric model capable of representing the commonly exhibited probability density function shapes by systems with various levels of geographic diversity.

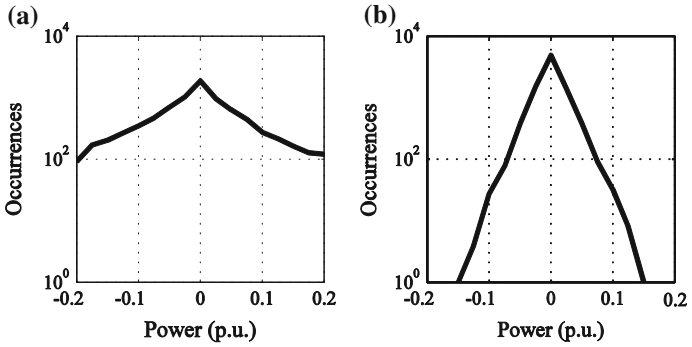

Fig. 2 Normalized hourly variability in wind power from an individual wind plant (**a**) and a large system (**b**) over the course of 1 year

2.2 Variability of Aggregate Wind Power

Instantaneous wind power values provide us with information on uncertainty, but we are also interested in wind power variability. The variability of wind power is examined through the statistical characteristics of wind power variation. Wind power variation is defined as the difference in instantaneous wind power at two points in time, usually 1 h: $\Delta P[t] = P[t] - P[t-q]$ where ΔP is the variation of wind power, t is the time, and q is the variation period. Similar to our approach with uncertainty, the variability of aggregate wind power is assessed by considering the probability density function of ΔP, as well as its statistical characteristics such as standard deviation and quantiles. We will see that variability in aggregate wind power is strongly influenced by geographic diversity.

Figure 2 shows typical normalized hour-to-hour wind power variation in an individual wind plant (a) and system (b) over a period of 1 year. Note that for clarity the ordinate is logarithmically scaled. The trace for the individual wind plant is much broader than for the system aggregate, indicating more frequent extreme variability. Aggregation, therefore, tends to smooth wind power variability. In Fig. 2a, b, the nearly linear decrease in occurrences on the logarithmically scaled ordinate suggests that an appropriate parametric model will have an exponential form. The slope of the decrease is influenced by the geographic diversity of the system, as well as the variation period, with shorter periods having steeper slopes.

In the above we have briefly described typical uncertainty and variability characteristics of aggregate wind power. These characteristics depend on the geographic diversity in a system, as well as the characteristics of the constituent individual wind plants. Therefore in order to thoughtfully propose aggregate wind power models, we must begin by modeling individual wind plants and then establishing the mathematical framework for geographic diversity's effect on uncertainty and variability.

3 Individual Wind Plant Model

The characteristics of aggregate wind power—especially at low levels of geographic diversity—depend on the characteristics of the system's constituent wind plants. We first derive an analytic model of wind plant power output under idealized conditions: that the wind speed follows a parametric probability density function and the energy conversion process is deterministic, among other assumptions. We conclude the section by considering non-idealities in wind plant power production.

The power delivered by a wind plant P_i is the sum of the real power produced by its constituent wind turbines, less collector system losses:

$$P_i = \sum_{j=1}^{M} P_{\mathrm{WT},j} - P_{L,i} \tag{1}$$

where M is the number of wind turbines in the wind plant, $P_{\mathrm{WT},j}$ is the real power generated by the jth wind turbine and $P_{\mathrm{L},i}$ is the ith wind plant's collector system losses at the current operating state.

Although (1) appears straightforward, wind turbines are nonlinear energy conversion devices whose power output is primarily dependent on wind speed, which is a random variable. This implies that P_i will be stochastic, and that a probabilistic model of P_i can be derived by transforming the probability density function of the wind speed.

3.1 Probabilistic Wind Speed Model

Let $\tilde{v}$ be a random variable representing the wind speed at a certain wind turbine with corresponding probability density function $f_v(\tilde{v})$. The presence of the tilde indicates that the variable is random. We can approximate $f_v(\tilde{v})$ by the repeated independent sampling of $\tilde{v}$. These samples can be binned into a histogram and scaled to approximate $f_v(\tilde{v})$. Two typical, yet specific, scaled histograms of wind speed are shown in Fig. 3.

Although histograms are helpful visual approximations of probability density functions, it is often desirable to represent them using a parametric function. Let $\hat{f}_v(\tilde{v}, \boldsymbol{\theta})$ be the model of $f_v(\tilde{v})$ with parameters arranged in the vector $\boldsymbol{\Theta}$. For the sake of brevity, we will suppress $\boldsymbol{\Theta}$ in all distributions hereafter, so that, for example, $\hat{f}_v(\tilde{v})$ represents the parametric model of $f_v(\tilde{v})$.

Returning to Fig. 3, we see that each histogram is asymmetric with a distinctly positive skewness, indicating that high wind speeds are less frequent than low wind speeds. These characteristics can be modeled using the three-parameter Generalized Gamma distribution [23]. However, estimating the parameters of this distribution can be difficult. Instead, the two-parameter Weibull distribution, which

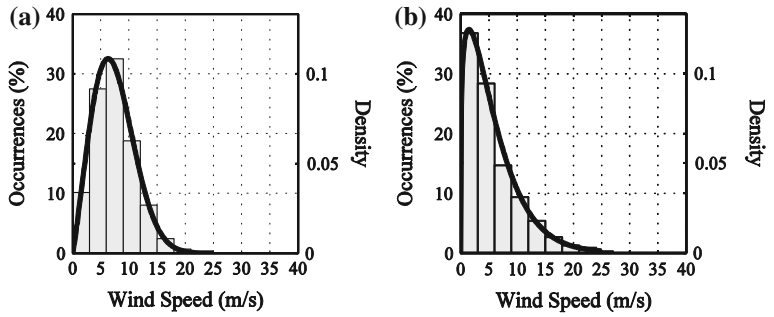

Fig. 3 Histograms of wind speed from two locations with fit Weibull probability density functions. **a** wind speed distribution for location 1, **b** wind speed distribution for location 2

is a special case of the Generalized Gamma distribution, is commonly used. The Weibull probability density function [24] is

$$\hat{f}_v(\tilde{v}) = \begin{cases} 0 & \tilde{v} < 0 \\ \frac{k}{\lambda}\left(\frac{\tilde{v}}{\lambda}\right)^{k-1} e^{-(\tilde{v}/\lambda)^k} & \tilde{v} \geq 0 \end{cases} \tag{2}$$

where k and λ are the shape and scale parameters, respectively. The parameters can be estimated from sampled data using the maximum likelihood estimation (MLE) method or the method of moments, though for the Weibull distribution these methods can be mathematically cumbersome [25, 26]. Examples of Weibull distributions with parameters estimated using MLE are shown as the solid traces in Fig. 3. In each case, the parametric function is a reasonable approximation to the data.

The two-parameter Weibull model can often be simplified to a single parameter model without appreciably sacrificing accuracy. For locations with wind regimes suitable for wind plant development, the estimated shape parameter of the Weibull distribution is often near 2.0. Therefore, the Weibull distribution can be reduced to the Rayleigh distribution [24]. A useful feature of the Rayleigh distribution is that its parameter can be estimated with the method of moments using only the mean of the wind speed, which is sometimes the only quantity available.

3.2 Idealized Wind Turbine Power Curve

We next examine the effect of the wind turbine in shaping the wind power probability density function. Wind turbines convert a portion of the kinetic energy into a mass of moving air to electrical energy by way of electric generator. The electrical power output by a wind turbine is computed from:

$$\tilde{P}_{\mathrm{WT}} = \frac{1}{2} C_{\mathrm{p}} A \rho \tilde{v}^3 \tag{3}$$

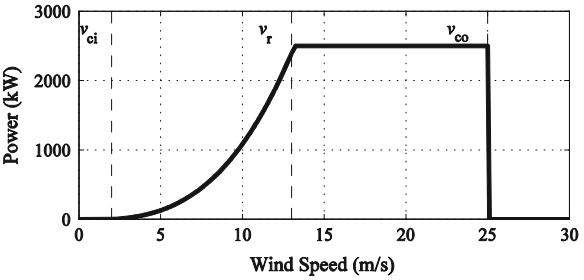

Fig. 4 Idealized *power curve* of a 2.5 MW wind turbine with cut-in speed of 2.5 m/s, rated wind speed of 3 m/s, and cut-out wind speed of 25 m/s

where C_p is the dimensionless power coefficient, A is the area swept by the rotor blades in m^2, and ρ and $\tilde{v}$ are the density and velocity of the air mass incident to the wind turbine, in kg/m^3 and m/s, respectively [27]. The power coefficient represents the overall efficiency of the energy conversion process, which depends on turbine design and operating state. If a constant C_p can be maintained, then the wind speed and wind power are cubically related. Under low or high wind speed conditions the wind turbine is operated such that C_p is zero, for reasons discussed later. Although (3) is useful, a more common and illustrative way to show the relationship between wind speed and wind turbine power output is with the power curve.

A power curve deterministically maps each wind speed to the corresponding power output of a wind turbine. An example of an idealized power curve is shown in Fig. 4. In general, there are four regions of operation.

3.2.1 Below Cut-in Wind Speed $(v < v_{ci})$

At low wind speeds no electrical power is produced and C_p is zero. The power in the wind is not enough to either overcome the friction of the drivetrain, or to result in positive net power production. The threshold wind speed at which power generation begins is known as the cut-in wind speed (v_{ci}).

3.2.2 Between Cut-in and Rated Wind Speed $(v_{ci} \leq v < v_r)$

When the wind speed is between the cut-in and rated wind speed (v_r), the wind turbine generates power. In this region, the turbine is designed or controlled to maximize C_p, and a nearly cubic wind speed-turbine power relationship is observed. This relationship can be approximated as

$$P_{WT} = av^3 - bP_r \tag{4}$$

where a and b are coefficients and P_r is the rated power of the wind turbine [27].

The coefficients can be determined by enforcing $P_r = av_r^3 - bP_r$ and $0 = av_{ci}^3 - bP_r$ and then numerically solving the resulting set of equations.

3.2.3 Between Rated and Cut-out Wind Speed $(v_r \leq v < v_{co})$

At wind speeds above rated and below cut-out (v_{co}), the wind turbine is controlled to maintain constant power production. As the wind speed increases over this region constant power is maintained by reducing C_p through active pitch control or passive stall design.

3.2.4 At and Above Cut-out Wind Speed $(v_{co} \leq v)$

At excessively high wind speeds, the wind turbine is in danger of mechanical failure. The turbine is aerodynamically slowed and stopped, and then mechanically locked into place to prevent rotation. C_p is zero over this region.

The effect of the varying power coefficient can be implicitly accounted for by expressing (3) as the piecewise function $P_{WT} = g(v)$, where $g(v)$:

$$P_{WT} = g(v) = \begin{cases} 0 & v < v_{ci} \\ av^3 - bP_r & v_{ci} \leq v < v_r \\ P_r & v_r \leq v < v_{co} \\ 0 & v_{co} \leq v \end{cases} \tag{5}$$

As previously mentioned, the power curve in Fig. 4 and expressed as (5) are idealized. Most utility-scale wind turbine manufacturers develop the power curve for a particular wind turbine model under carefully controlled conditions according to accepted standards [28]. Field performance of wind turbines can be inconsistent and appreciably differ from the manufacturer-supplied power curve. These non-idealities are discussed in detail in Sect. 3.4.

3.3 Idealized Wind Plant Model

A basic model of power from a wind plant is

$$\tilde{P}_i = M\tilde{P}_{WT} = Mg(\tilde{v}) \tag{6}$$

where $\tilde{P}_i$ is the total real power from the M wind turbines of wind plant i. This model makes several assumptions, such as all wind turbines experience the same wind speed and have the same power curve. We will discuss the reasonableness of these assumptions in Sect. 3.4. However, for the following, we will assume that (6) holds. Since $\tilde{P}_i$ is a random variable, we can characterize it with a probability density function or cumulative distribution function (cdf).

Let the cdf of the wind speed be $F_v(\tilde{v})$ and let the cdf of the wind power from the wind plant be $F_P(\tilde{P}_i)$. The cdf of the power from the wind plant is the piecewise function:

$$F_P(\tilde{P}_i) = \begin{cases} F_v(v_{\text{ci}}) + 1 - F_v(v_{\text{co}}) & \tilde{P}_i = 0 \\ F_v\left(g^{-1}\left(\tilde{P}_i/M\right)\right) + 1 - F_v(v_{\text{co}}) & 0 < \tilde{P}_i < MP_r \\ 1 & \tilde{P}_i = MP_r \end{cases} \tag{7}$$

where $g^{-1}(P_{\text{WT}})$ is the inverse power curve.

We can see that this gives us the following probabilities:

$$\begin{aligned} \Pr\{\tilde{P}_{\text{i}} = 0\} &= F_v(v_{\text{ci}}) + 1 - F_v(v_{\text{co}}) \\ \Pr\{0 < \tilde{P}_{\text{i}} < MP_{\text{r}}\} &= F_v(v_{\text{r}}) - F_v(v_{\text{ci}}) \\ \Pr\{\tilde{P}_{\text{i}} = MP_{\text{r}}\} &= F_v(v_{\text{co}}) - F_v(v_{\text{r}}) \end{aligned} \tag{8}$$

As there is a measurable probability of wind power being exactly equal to either 0 or MP_r, we do not have a purely continuous distribution function. We instead have a mixed discrete/continuous distribution function.

Our pdf for wind power, then, will need to use the Dirac delta function $\delta(\cdot)$ to address the probabilities of wind power being exactly equal to either 0 or MP_r. For all other values of wind power, we can calculate the pdf using the traditional change-of-variables method. Using (5), the inverse power curve for the region between 0 and MP_{r} is

$$v = g^{-1}(P_{\text{WT}}) = \left(\frac{P_{\text{WT}} + bP_r}{a}\right)^{1/3} \tag{9}$$

The corresponding probability density function of wind power for this region is found by taking the derivative of (7) with respect to $\tilde{P}_i$ so that:

$$\begin{aligned} f_P(\tilde{P}_i) &= \frac{\mathrm{d}F_P(\tilde{P}_i)}{\mathrm{d}\tilde{P}_i} = \frac{\mathrm{d}F_v\left(g^{-1}\left(\tilde{P}_i/M\right)\right)}{\mathrm{d}\tilde{P}_i} = \frac{\mathrm{d}F_v\left(g^{-1}\left(\tilde{P}_i/M\right)\right)}{\mathrm{d}g^{-1}\left(\tilde{P}_i/M\right)} \cdot \frac{\mathrm{d}g^{-1}\left(\tilde{P}_i/M\right)}{\mathrm{d}\tilde{P}_i} \\ &= \frac{\mathrm{d}F_v(\tilde{v})}{\mathrm{d}\tilde{v}} \cdot \frac{\mathrm{d}g^{-1}\left(\tilde{P}_i/M\right)}{\mathrm{d}\tilde{P}_i} \end{aligned} \tag{10}$$

We note that $\frac{\mathrm{d}F_v(\tilde{v})}{\mathrm{d}\tilde{v}}$ is just the pdf of the wind speed $\hat{f}_v(\tilde{v})$ evaluated at $\tilde{v} = g^{-1}\left(\tilde{P}_i/M\right)$ and that

$$\frac{\mathrm{d}g^{-1}\left(\tilde{P}_i/M\right)}{\mathrm{d}\tilde{P}_i} = \frac{\left(\frac{\tilde{P}_i/M + bPr}{a}\right)^{-2/3}}{3aM} \tag{11}$$

For the specific case of a Weibull wind speed distribution (10) becomes:

$$f_p(\tilde{P}_i) = \left(\frac{k}{\lambda} \left(\frac{g^{-1}\left(\tilde{P}_i/M\right)}{\lambda} \right)^{k-1} e^{-\left(\frac{g^{-1}\left(\tilde{P}_i/M\right)}{\lambda} \right)^{k}} \right) \frac{\left(\frac{\tilde{P}_i/M + bPr}{a} \right)^{-2/3}}{3aM}. \quad (12)$$

We now have closed-form algebraic expressions for the probability density function of wind power when $0 < \tilde{P}_i < MP_r$.

Applying (6), the idealized individual wind plant model for a Weibull distribution of wind speed is therefore:

$$f_p(\tilde{P}_i) = \left(\frac{k}{\lambda} \left(\frac{g^{-1}\left(\tilde{P}_i/M\right)}{\lambda} \right)^{k-1} e^{-\left(\frac{g^{-1}\left(\tilde{P}_i/M\right)}{\lambda} \right)^{k}} \right) \frac{\left(\frac{\tilde{P}_i/M + bPr}{a} \right)^{-2/3}}{3aM} + (F_v(v_{ci}) + 1 - F_v(v_{co}))\delta(\tilde{P}_i) + (F_v(v_{co}) - F_v(v_r))\delta(\tilde{P}_i - MP_r) \quad (13)$$

For $0 \le \tilde{P}_i \le MP_r$, and 0 everywhere else. For a Rayleigh distribution of wind speed:

$$f_p(\tilde{P}_i) = \left(\left(\frac{g^{-1}\left(\tilde{P}_i/M\right)}{\lambda^2} \right) e^{-\left(\frac{g^{-1}\left(\tilde{P}_i/M\right)^2}{2\lambda^2} \right)} \right) \frac{\left(\frac{\tilde{P}_i/M + bPr}{a} \right)^{-2/3}}{3aM} + (F_v(v_{ci}) + 1 - F_v(v_{co}))\delta(\tilde{P}_i) + (F_v(v_{co}) - F_v(v_r))\delta(\tilde{P}_i - MP_r) \quad (14)$$

again for $0 \le \tilde{P}_i \le MP_r$, and 0 everywhere else.

The models of individual wind plant power output in (13) and (14) were analytically derived and are dependent only on a small number of wind speed and power curve parameters. Figure 5 shows an example of the pdf and cdf of a wind plant. We note that the pdf shown is similar in appearance to that of Fig. 1a, which is reassuring. However, there are differences, particularly near rated power. Our analytic model overestimates the probability of rated power production from the wind plant. This discrepancy is an artifact of the idealized assumptions implied by (6). Despite the differences the derived model has utility—it serves as a reasonable starting point in the absence of additional data, and it allows for further analytic manipulation, as we will see later.

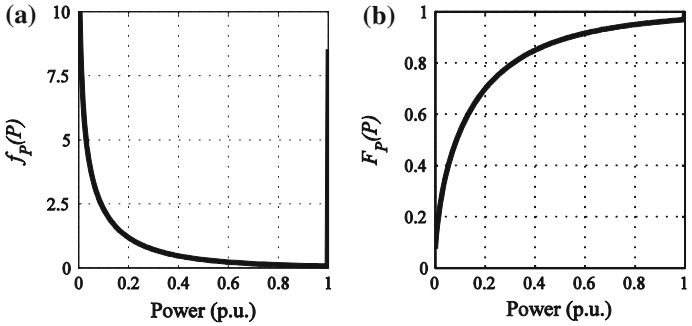

Fig. 5 Analytic pdf (**a**) and cdf (**b**) of wind power from an individual wind plant with *power curve* in Fig. 4 and a Rayleigh distribution of wind speed with scale parameter $\lambda = 5$

3.4 Non-idealized Wind Plant Modeling

The derived probability density functions in (14) and (15) are idealized models of wind power under several assumptions implied by (6), including:

1. the collector system is lossless;
2. all wind turbines are in service;
3. the power curves are deterministic for each wind turbine;
4. the wind speed and air density are the same at each wind turbine.

The first and second assumptions allow the wind plant to reach 100 % rated output. However, collector system losses are nonzero—usually between 1 and 5 % [29]—and wind turbines are routinely taken out of service due to scheduled maintenance or malfunction. A typical wind turbine requires regularly scheduled maintenance at 6 month intervals, although online condition monitoring may yield more optimal schedules [30]. Maintenance and malfunction-related outages can last several hours or longer, with offshore wind turbines generally requiring longer outages. For these reasons, a large wind plant may have one or more wind turbines out of service at any given time. Although rated power is not achieved, occurrences of near rated power may be common, as shown in Fig. 1a.

The third assumption, while greatly simplifying the model, is not substantiated by empirical data. Even after correcting for air density, power production strays from the manufacturer-supplied power curve for several reasons, including: dependency of power output on wind direction; forced curtailment; wind shear and turbulence; manufacturing and installation defects; efficiency degradation over time; and internal power consumption for control systems, lighting, pumps, and other functions. These effects are illustrated by an empirical power curve. An empirical power curve is constructed by plotting repeated in situ simultaneous measurements of wind speed and wind power, as shown in Fig. 6. It is clear from this figure that the idealized power curve model in (5) does not capture the true in situ relationship between wind speed and wind power.

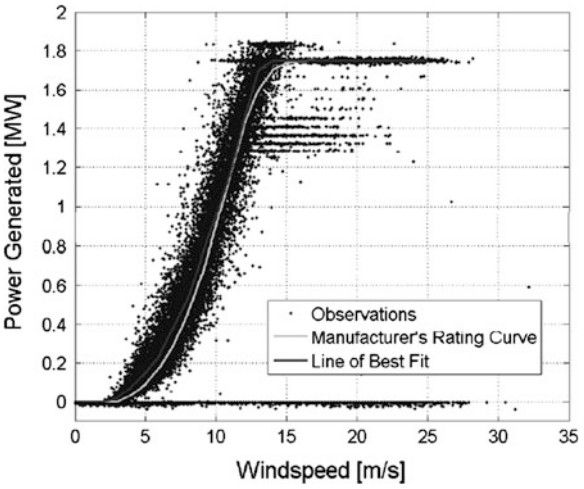

Fig. 6 Empirical *power curve* of a wind turbine showing effects of non-idealities © 2007 IEEE. Reprinted, with permission, from [31]

The fourth assumption implied by (6) is naïve, though often used. In reality, wind turbines can be separated by several kilometres and installed in varying terrain. The wind turbines almost never simultaneously experience the same wind speed. In other words, there is geographic diversity within each wind plant that must be accounted for. Over the years, several approaches to more realistically account for non-idealities in power curves have been proposed [32–34]. Some involve creating sub-groups of wind turbines, where each sub-group has its own wind speed. Other methods involve creating data-driven wind plant power curves using statistical methods. Provided they are well-defined, wind plant power curves can be used in place of (5) to analytically derive a probabilistic model of wind power using the same steps detailed in Sect. 3.3.

To summarize our progress, we started by identifying suitable parametric probabilistic models of wind speed: Rayleigh or Weibull. We then developed pdfs and cdfs for individual wind plant power output by transforming the wind speed models using an idealized power curve, as in (13) and (14). This method can be used to account for non-idealities using a wind plant power curve if desired. We are now poised to consider the aggregate power production from a fleet of wind plants in a common system. To do this, we must model the correlation and dependency of wind power among the wind plants. In other words, we must model the geographic diversity.

4 Geographic Diversity

The most important influencer of the uncertainty and variability of aggregate wind power is the geographic diversity of the system. A system's geographic diversity is a reference to the general level of dependence between the wind power from its

constituent wind plants. Dependence is often quantified using a correlation coefficient, with lower correlation generally resulting in decreased uncertainty and variability. The goal of this section is to formulate the mathematic basis for this phenomenon, discuss factors influencing correlation and dependency, as well as to comment on practical considerations.

4.1 Theoretical Basis

Aggregate wind power can mathematically be represented as

$$\tilde{P}_{\text{agg}} = \frac{1}{N}\sum_{i=1}^{N}\frac{\tilde{P}_i}{C_i} \tag{15}$$

where $\tilde{P}_{\text{agg}}$ is the normalized aggregate wind power, N is the number of wind plants considered, and C_i and P_{i} are the capacity and real power delivered by the ith wind plant.

Hereafter, subscripts are used to associate a variable with a specific wind plant, so that x_{n} pertains to the nth wind plant. We start with the idealized assumption that the wind speed at each wind plant $\tilde{v}_1, \ldots, \tilde{v}_N$ are independent random variables. Consequently, $\tilde{P}_1, \ldots, \tilde{P}_N$ will also be independent random variables, and are therefore uncorrelated.

As a result of the assumption of independence, the probability distribution of the aggregate power output is found through convolution and change-of-variables:

$$f_{P_{\text{agg}}}\left(\tilde{P}_{\text{agg}}\right) = \left(N\sum_{i=1}^{N} C_i\right) f_{P_1}\left(\tilde{P}_1\right) * \cdots * f_{P_N}\left(\tilde{P}_N\right) \tag{16}$$

where * is the convolution operator.

By the Central Limit Theorem, as N increases $f_{P_{\text{agg}}}\left(\tilde{P}_{\text{agg}}\right)$ will approach the Gaussian distribution so that:

$$f_{P_{\text{agg}}}\left(\tilde{P}_{\text{agg}}\right) = \frac{1}{\sigma_{\text{agg}}\sqrt{2\pi}} e^{-\frac{\left(\tilde{P}_{\text{agg}}-\mu_{\text{agg}}\right)^2}{2\sigma_{\text{agg}}^2}} N \to \inf \tag{17}$$

where μ_{agg} and σ_{agg} are the mean and standard deviation of $f_{P_{\text{agg}}}\left(\tilde{P}_{\text{agg}}\right)$. The transformation of the probability density function to a Gaussian distribution according to the Central Limit Theorem has important implications to the uncertainty of aggregate wind power.

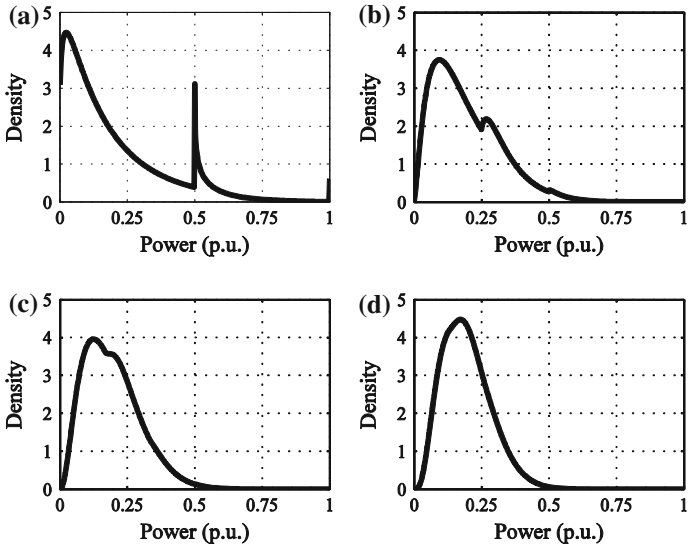

Fig. 7 Probability density functions of normalized aggregate wind power for systems with 2 (**a**), 4 (**b**), 6 (**c**) and 8 (**d**) wind plants under assumption of independence

4.2 Uncertainty and Variability Reduction

The normalized aggregate wind power's statistical variance will decrease as more wind plants are added to the system according to

$$\sigma_{\text{agg}}^2 = \frac{1}{N^2}\sum_{i=1}^{N}\sigma_i^2 \tag{18}$$

where σ_i^2 is the variance of the normalized power output of the *i*th wind plant. The decrease in variance causes the Gaussian distribution to contract, which can be interpreted as a decrease in the uncertainty of the wind power.

The evolution of $f_{P_{\text{agg}}}\left(\tilde{P}_{\text{agg}}\right)$ to a Gaussian distribution can be noticeable even for small values of *N*. This is illustrated in Fig. 7 where $f_{P_{agg}}\left(\tilde{P}_{\text{agg}}\right)$ for $N = 2, 4, 6,$ and 8 are plotted using numerical convolution based on the pdf in Fig. 5a and assuming independence in power output. Note how the distribution contracts as more wind plants are aggregated, resulting in decreased standard deviation and decreased probability of extremely high or low power output. In other words, the uncertainty decreases.

The effects of geographic diversity also apply to the moment-to-moment variation in aggregate wind power: under the assumption of independence, as more wind plants are aggregated the moment-to-moment variation approaches a Gaussian distribution with decreasing variance.

The preceding derivation showed the mathematical mechanism by which geographic diversity reduces uncertainty and variability in aggregate wind power. Its foundation is the assumption that the wind speed (and power) at all N wind plants are independent, and hence not correlated, but how realistic and restrictive is this assumption? It is intuitive that a pair of wind plants in close proximity will experience similar wind conditions. However, if they are separated by several hundred kilometres, the same claim becomes unreasonable. In other words, wind power tends to exhibit high spatial dependency of its correlation. It is these features that ultimately determine the geographic diversity of a system.

4.3 Correlation of Instantaneous Wind Power

The dependency of two random variables can be quantified by correlation coefficients. The correlation coefficient does not fully capture the underlying dependency structure between the variables as the joint distribution would, but nonetheless it is useful in quantifying the dependency. Among the most used correlation coefficients are Pearson's ρ and Kendall's τ. Pearson's ρ measures linear correlation, whereas Kendall's τ measures rank correlation [35]. There is no compelling reason to believe that wind power would be linearly correlated, and it has been argued that Kendall's τ is a more suitable metric for wind speed or power [36]. Nonetheless, many studies to date have focused strictly on Pearson's ρ. Whichever metric is used, it is widely recognized that there exists a strong relationship between separation distance and correlation coefficient [10–16].

Empirical studies have shown that the relationship between separation distance and correlation coefficient can be modeled as

$$\tau = e^{-\mathrm{rd}^s} \tag{19}$$

where r is the so-called decay constant, s is the stretching coefficient, and d is the distance between wind plants in kilometers [7, 16]. Figure 8 shows the computed Kendall's τ for various randomly sampled wind plant pairings obtained from the NREL Eastern Wind Integration Data Set [37]. Similar traces have been reported for Pearson's ρ [7, 14]. The solid trace is for hourly averaged data; the dashed trace is for daily-averaged data. The decay and stretching coefficients for the hourly averaged data are 0.0037 and 0.92, respectively; and 0.0096 and 0.81 for the daily averaged data, respectively. In general, longer averaging periods result in larger correlation coefficients. Zero correlation in wind power among wind plant pairs does occur but is rare, and when it does occur it is usually at separation distances greater than 1,000 km.

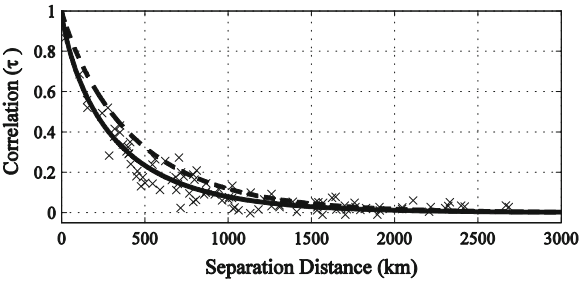

Fig. 8 Kendall's τ computed for 100 samples with hourly averaging (*solid*) and daily averaging (*dashed*). Only the data points for hourly averaging are shown for the sake of clarity

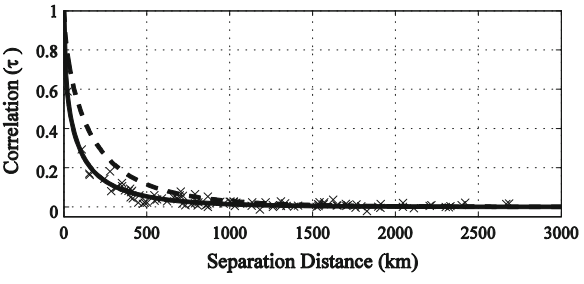

Fig. 9 Kendall's τ computed for 100 samples of aggregate wind power variation with hourly variation periods (*solid*) and 6-hourly variation periods (*dashed*). Only the data points for hourly variation periods are shown for the sake of clarity

4.4 Correlation of Wind Power Variation

Geographic diversity tends to have more pronounced effect on the moment-to-moment variation of wind power than on instantaneous wind power. Figure 9 shows the rank correlation coefficient for variation periods of 1 h (solid) and 6 h (dashed) for the same 100 wind plant pairs that were considered in the previous section. As before, an equation of the form of (19) suitably fits the data. The decay and stretching coefficients are 0.108 and 0.532 for the hourly variation, and 0.029 and 0.692 for the 6-hourly variation, respectively. From Fig. 9, and in general, longer variation periods tend to have higher correlation than short ones.

When compared to instantaneous power, the correlation coefficients of wind power variations are smaller and decay faster with distance. Near-zero correlation is exhibited at closer distances, around 750 km. The reduction of variation (smoothing effect), therefore, is noticeable in many systems.

4.5 Other Factors Influencing Correlation

Correlation of wind power is not solely dependent on separation distance. Other variables influencing correlation of wind power are:

- Terrain. Wind plants located close to each other but in different terrain may experience different wind regimes, which may decrease correlation.

- Averaging and variation period. Shorter periods tend to exhibit less correlation. Compare, for example, the solid and dashed lines in Figs. 8 and 9.
- Direction of the separation. For example, many North America wind plants with East–West separation have greater correlation than those with North–South separation [38].
- Number of wind turbines. Correlation between aggregate wind power in systems tends to be higher than the correlation between wind plants of individual wind turbines [16]. High frequency fluctuations tend to be uncorrelated and are filtered by dispersed wind turbines.

It is also notable that the correlation coefficient itself may exhibit erratic variation. Wind plant pairs may exhibit high correlation one day, and low correlation the next [15, 18].

4.6 Wind Power Dependency Structures

The correlation coefficient, while useful in quantifying dependence, does not provide sufficient information to construct the pdf of aggregate wind power from a pair of wind plants. Rather, information on the dependency structure of the wind power contained in the pair's joint distribution is needed.

The dependency structure of wind power is best explained visually and by considering the bi-variate case. Figure 10 shows typical, yet specific, contour plots of joint probability density functions for four different wind plant pairs, each with different rank correlation coefficients. The darker shading indicates greater density. The marginal histograms are shown on the top and right side of each plot.

Inspecting Fig. 10 shows how rank correlation tends to influence the joint distribution of wind power. At the lowest levels of correlation (Fig. 10a), mutually high power output is rare. Mutually low power output does occur, but that is an artifact of the marginal distributions having increased density at low power output.

At correlation levels in Fig. 10b—around 0.25—a concentration of higher density appears in the area of mutually high power output, but the density at mutually lower power remains. As correlation increases (Fig. 10c, d), the density begins to align along the diagonal, with areas of increased density at the extremes.

The normalized power output from each wind plant pair can be summed to examine the aggregate wind power in each of the four cases. The resulting histograms of aggregate power are shown in Fig. 11. Note that the power has been re-normalized by dividing the aggregate power by two, using a simplifying assumption that the capacities of each wind plant are identical.

As can be expected from its near-zero correlation, Fig. 11a resembles a system with higher geographic diversity. The standard deviation is lower than the others, and there is decreased probability of extreme power output. As correlation increases in Fig. 11b–d, so does the standard deviation and propensity for extreme power output. In Fig. 11d, the histogram appears similar to that of an individual wind plant, which is a signature of low geographic diversity.

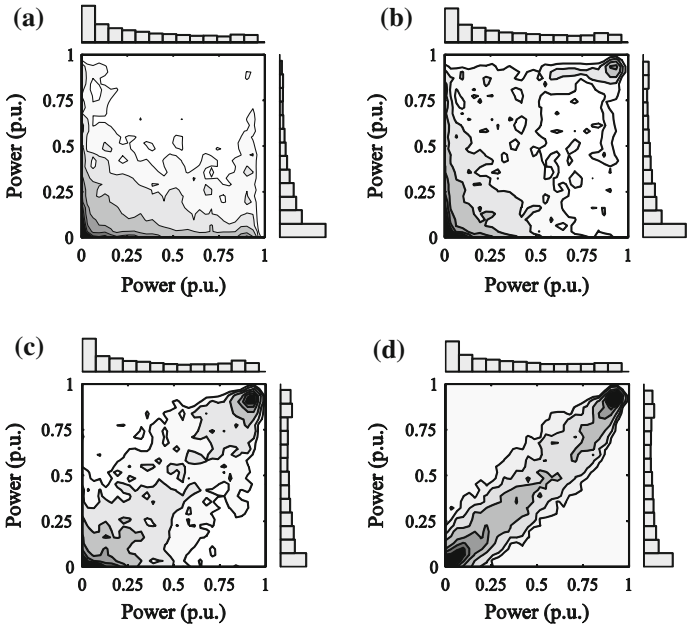

Fig. 10 Joint probability density function contours for wind plant pairs with increasing rank correlation. *Darker shading* indicates greater density. **a** $\tau = 0.01$, **b** $\tau = 0.25$, **c** $\tau = 0.52$ and **d** $\tau = 0.76$

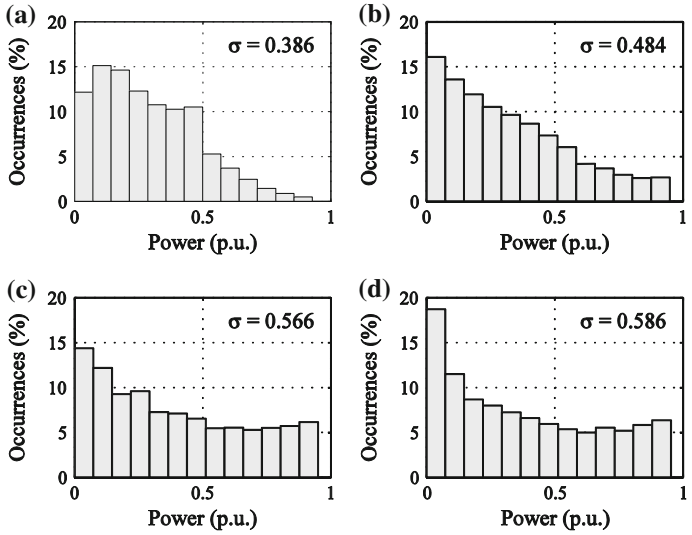

Fig. 11 Histograms of aggregate wind power for increasing rank correlation coefficients. **a** $\tau = 0.01$, **b** $\tau = 0.25$, **c** $\tau = 0.52$ and **d** $\tau = 0.76$

Figure 11 also illustrates that the wind plants in a system do not need be uncorrelated to realize the benefits of geographic diversity. Rather, uncertainty and variability are decreased even if the wind plants exhibit higher correlation—although not as noticeably as if the wind plants are uncorrelated.

4.7 Multivariate Models and Simulation

In the special case that aggregate and individual wind power need to be modeled or simulated, a multivariate model is required. For example, a Monte Carlo simulation can be performed by sampling from the joint distribution of the power from the wind plants. However, a proper joint distribution model must be identified.

Returning to Fig. 10, we see that the joint distributions do not obviously conform to any common parametric functions, and are certainly not Gaussian. One reason for this is that the marginal distributions add complexity to the overall structure. It is possible to decouple the influence of the marginal distributions from the joint by transforming the data from the wind power domain to the rank/uniform domain by way of the cdf of the individual wind plants. The result is uniform marginal distributions, with a joint distribution that is more amenable to parametric modeling. The resulting dependency structure can be modeled using copulas [39]. Though beyond the scope of this chapter, other works have shown that Gumbel and Gaussian copulas are appropriate for multi-variate wind power dependency modeling [40, 41]. The selection of a specific copula depends on the number of wind plants and the desired rank correlation.

One method of Monte Carlo simulation of wind power using copulas is: (1) identify the separation distances of the wind plants to be modeled; (2) using (19), compute the corresponding rank correlation matrix; (3) select an appropriate multivariate copula to model the dependency structure, using the correlation coefficient to determine the copula's parameters; (4) randomly draw the desired number of samples from the copula in the rank/uniform domain; (5) transform the samples to the wind power domain using the inverse cdf of (7). See [40, 41] for additional implementation details.

4.8 Practical Considerations

Geographic diversity can theoretically reduce uncertainty and variability of aggregate wind power. However, in many systems, the perceived benefits of geographic diversity have failed to materialize. There are four main reasons for this [16]:

1. In many electricity markets, the system-wide benefits of reduced uncertainty and variability do not directly translate into gains for individual wind plants—they do not have compelling economic reasons to seek geographic diversity.

2. Wind plants tend to have greater capacities, resulting in high concentrations of wind turbines.
3. There is limited access to suitable transmission limits the number of geographically diverse regions that can be interconnected.
4. There is only so much geographic diversity that can be had. In a given region, as wind plants are added, the benefits of geographic diversity saturate. Additionally, large-scale forces such as insolation similarly influence very wide areas (e.g., the continental United States).

In this section we have shown that geographic diversity, particularly in the separation distance between wind plants, reduces dependency and correlation. This in turn leads to a transformation of the probability density function of instantaneous and moment-to-moment variations in aggregate wind power toward a Gaussian distribution, with decreased standard deviation, and hence less uncertainty and variability. However, geographic diversity can and should not be viewed as a panacea for eliminating variability and uncertainty. Wind plants exhibit statistically significant correlation at large separation distances, and there are several practical reasons for geographic diversity not to occur. The prospects for variability reduction, however, are better as the correlation of wind power variability rapidly decreases with separation distance.

5 Aggregate Wind Power Models

In many cases, especially in the absence of transmission congestion, we are less concerned about the power from individual wind plants, and more concerned about the aggregate wind power. In these cases, we need not model the individual wind plants, and instead focus only on the aggregate wind power.

In this section we identify appropriate parametric models of the probability density functions of aggregate instantaneous wind power and wind power variation. Parametric approximation has several advantages including: fewer data points are needed to "fill out" the distribution; analytic calculations are more tractable; and the model can be expressed with fewer pieces of information—hopefully just one or two parameters.

5.1 Instantaneous Aggregate Wind Power Model

The shape of the histograms of aggregate wind power in Fig. 1 suggests that we seek a model with two parameters—one controlling the shape at low wind power, the other at high wind power.

One promising candidate is the Beta probability density function:

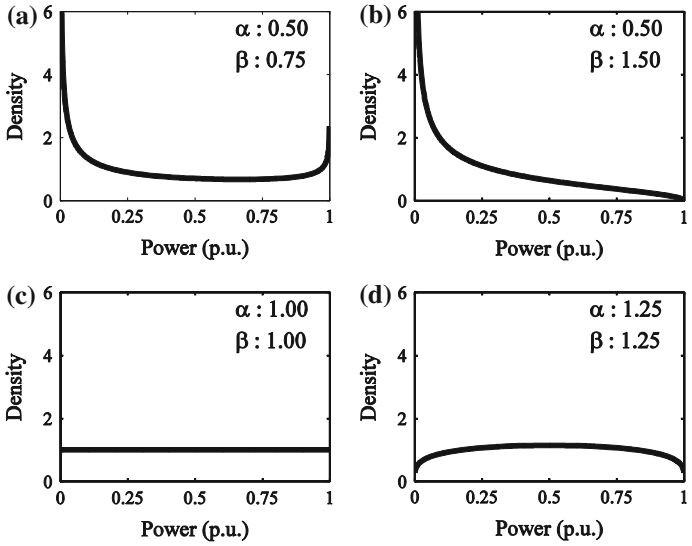

Fig. 12 Examples of Beta probability density functions with various α and β parameters

$$\hat{f}_P\left(\tilde{P}_{\text{agg}}\right) = \frac{\tilde{P}_{\text{agg}}^{\alpha-1} \times \left(1 - \tilde{P}_{\text{agg}}\right)^{\beta-1}}{B(\alpha, \beta)} \tag{20}$$

where $B(\alpha, \beta)$ is the Beta function:

$$B(\alpha, \beta) = \int_0^1 u^{\alpha-1} \times (1-u)^{\beta-1} \mathrm{d}u \tag{21}$$

and α and β are shape parameters and u is the variable of integration. The domain of the Beta pdf is [0, 1], thereby necessitating the use of normalized aggregate wind power values. A wide range of shapes are possible by careful selection of α and β, as shown in Fig. 12.

5.2 Beta Distribution Parameter Selection

An important advantage of using the Beta pdf to model aggregate wind power is that its parameters lend themselves to meaningful interpretation [42]. Values of α that are less than 1 indicate an increasing probability density of 0 or near-zero power output; whereas α values greater than 1 indicate decreasing probability density over this range. The β values have a similar interpretation. If β is greater

than 1, then the probability density decreases toward full power output; whereas if it is less than 1, there is increasing probability density.

In the absence of specific information, heuristic guides can be followed for modeling aggregate wind power. For low diversity, α and β are both greater than 1; for medium diversity $\alpha > 1$ and $\beta < 1$; and for higher diversity $\alpha < 1$, $\beta < 1$ or a Gaussian distribution can be used. Specific values for the parameters can be selected from estimations of the mean and variance.

The shape parameters are related to the mean as

$$\mu = \frac{\alpha}{\alpha + \beta} \tag{22}$$

The mean wind power considered over a length of time is also known as the capacity factor, which ranges between 20 and 40 % for most systems. Therefore the ratio of α to β from 1:1.5 to 1:4 is reasonable. The parameters are related to the variance as

$$\sigma^2 = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)} \tag{23}$$

Typical variance values range from about 0.02 to 0.10.

The parameters of the Beta pdf can be estimated using the method of moments according to

$$\hat{\alpha} = \hat{\mu}\left(\frac{\hat{\mu}(1 - \hat{\mu})}{\hat{\sigma}^2} - 1\right) \tag{24}$$

$$\hat{\beta} = (1 - \hat{\mu})\left(\frac{\hat{\mu}(1 - \hat{\mu})}{\hat{\sigma}^2} - 1\right) \tag{25}$$

where $\hat{\mu}$ and $\hat{\sigma}^2$ are the sampled mean and variance of the wind power data. There is no closed-form MLE of $\hat{\alpha}$ and $\hat{\beta}$, but numerical methods can be used to compute them [43].

Traces of the Beta pdf with parameters fit using MLE to aggregate wind power data from four large systems in different years are found in Figs. 13 and 14. In each case the Beta pdf is able to reasonably approximate the data. A rigorous evaluation of the fit of Beta distributions to aggregate wind power is found in [42].

5.3 Aggregate Wind Power Variation Model

Wind speed and wind power tend to exhibit large autocorrelation coefficients at short time lags. As such $\Delta\tilde{P}$ tends to be near-zero, with only rare occurrences of extreme deviations. The trend is that of symmetric and exponential decay.

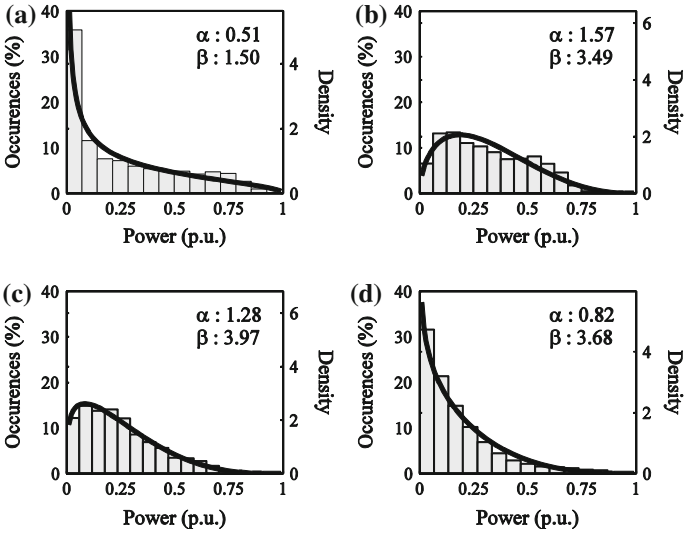

Fig. 13 Normalized aggregate wind power histograms in 2012 with fit Beta probability density functions. **a** BPA, **b** MISO, **c** PJM and **d** 50 Hz

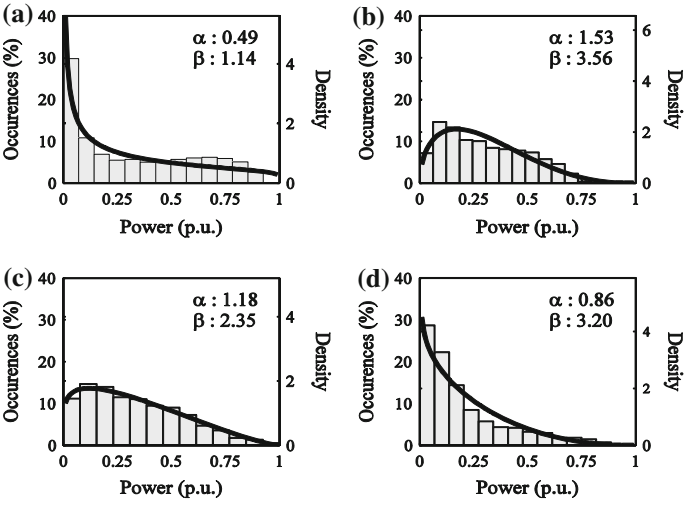

Fig. 14 Normalized aggregate wind power histograms in 2008 with fit Beta probability density functions. **a** BPA, **b** MISO, **c** PJM and **d** 50 Hz

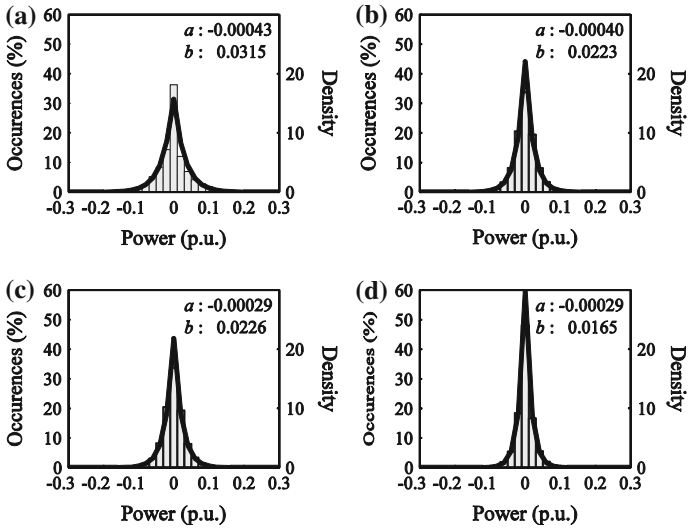

Fig. 15 Normalized aggregate wind power hourly variations in 2012 and fit Laplace probability density functions. **a** BPA, **b** MISO, **c** PJM and **d** 50 Hz

An appropriate distribution for these features is the Laplace distribution, also known as the double exponential distribution. The Laplace pdf is

$$\hat{f}_{\Delta\tilde{P}_{\text{agg}}}\left(\Delta\tilde{P}_{\text{agg}}\right) = \frac{1}{2b} e^{-\left|\Delta\tilde{P}_{\text{agg}} - a/b\right|} \tag{26}$$

where $\Delta\tilde{P}_{\text{agg}}$ is normalized aggregation power output variation and a and b are the location and scale parameters, respectively. The scale parameter is an indicator of the variability, where a larger scale parameter indicates greater variability.

5.4 Laplace Distribution Parameter Selection

The parameters of the Laplace distribution can be fit using MLE method where $\hat{a}$ is equal to the sample median, which is usually zero, and $\hat{b}$ is computed from

$$\hat{b} = \frac{1}{N}\sum_{i=1}^{N}\left|\Delta\tilde{P}_{agg,i} - \hat{a}\right| \tag{27}$$

The scale parameter is generally between 0.015 and 0.04 for 1 h variation periods. The scale parameter tends to increase with longer variation periods. Traces of the Laplace pdf with parameters fit using MLE to hourly aggregate wind power variation data are found in Figs. 15 and 16. In each case the Laplace pdf is

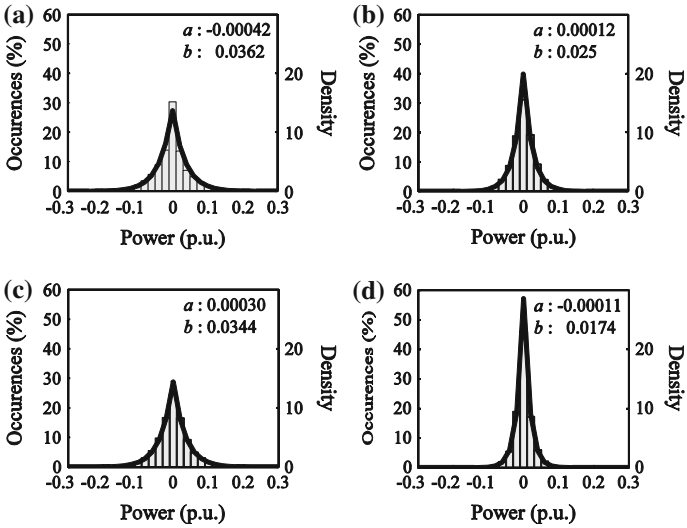

Fig. 16 Normalized aggregate wind power hourly variations in 2008 and fit Laplace probability density functions. **a** BPA, **b** MISO, **c** PJM and **d** 50 Hz

able to reasonably approximate the data. A rigorous evaluation of the fit of Laplace distributions to aggregate wind power is found in [44].

5.5 Influence of Variation Period

An important factor influencing the pdf of wind power variation is the variation period. As discussed in Sect. 4.4, the variation in wind power from wind plants exhibits greater correlation as the variation period increases. Additionally, there is also more time for the wind power to change, so there is greater potential for extreme variation. Figure 17 shows histograms and fit Laplace pdfs for hourly and 4-hourly variation periods. It is clear that a longer variation period results in a broader distribution, and hence, more variability that must be managed. The Laplace distribution is able to reasonably fit the data at both variation periods mainly by adjusting its scale parameter.

6 Statistical Characteristics of Aggregate Wind Power

It has only been within the last several years that aggregate wind power data sets have been made widely available to the research community. In this section we analyze several data sets with the goal of documenting and discussing statistical

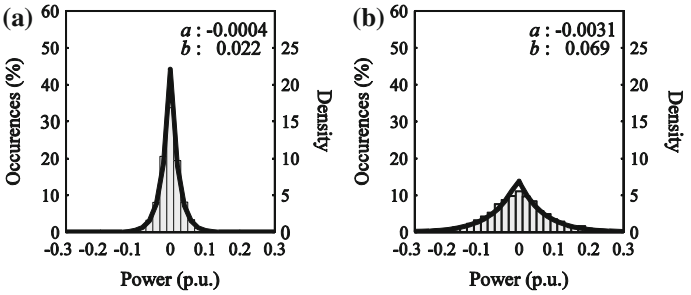

Fig. 17 Normalized aggregate wind power variation histograms and fit Laplace distributions for 1-h (**a**) and 4-h variation periods (**b**)

Table 1 Data set descriptions

System	2008 start capacity (MW)	2008 end capacity (MW)	2012 start capacity (MW)	2012 end capacity (MW)
BPA	1,301	1,671	4,131	4,711
MISO	2,462	4,327	10,514	12,270
PJM	1,150	1,277	5,318	6,457
50 Hz	9,091	9,493	11,570	12,420

characteristics of aggregate wind power, as well as seeking additional insight into geographic diversity.

6.1 Data Set Descriptions

Four data sets of aggregate wind power are considered: Bonneville Power Administration (BPA), Midwest ISO (MISO), PJM Interconnection (PJM), and 50 Hz [45–48]. BPA's territory is located in the Pacific Northwest of the United States, primarily in Washington State and Oregon. MISO has territory in 12 states in the midwest of the United States and in the Canadian province of Manitoba. PJM's territory covers all or parts of 13 states in the eastern portion of the U.S. 50 Hz's territory is in the northern and eastern portion of Germany. Each system has a large amount of wind plant capacity.

The data correspond to hourly averages for the years 2008 and 2012. The data have been normalized to reported system-wide wind capacity. However, these reports are infrequently issued. To overcome this, linear interpolation between reporting dates was used. This inherently introduces some error in the analysis, and so the reported statistics must be interpreted with this in mind. The capacities at the start and end of 2008 and 2012 are provided in Table 1.

Table 2 Instantaneous aggregate wind power statistical information for 2012

System	Mean (p.u.)	Standard deviation (p.u.)	Q(1) (p.u.)	Q(10) (p.u.)	Q(20) (p.u.)	Q(50) (p.u.)	Q(80) (p.u.)	Q(90) (p.u.)	Q(99) (p.u.)
BPA	0.26	0.26	0.00	0.01	0.02	0.16	0.52	0.68	0.85
MISO	0.31	0.19	0.02	0.08	0.13	0.28	0.51	0.58	0.72
PJM	0.25	0.17	0.01	0.05	0.09	0.21	0.39	0.50	0.68
50 Hz	0.18	0.17	0.00	0.02	0.04	0.12	0.28	0.41	0.78

Table 3 Instantaneous aggregate wind power statistical information for 2008

System	Mean (p.u.)	Standard deviation (p.u.)	Q(1) (p.u.)	Q(10) (p.u.)	Q(20) (p.u.)	Q(50) (p.u.)	Q(80) (p.u.)	Q(90) (p.u.)	Q(99) (p.u.)
BPA	0.32	0.29	0.00	0.01	0.03	0.25	0.64	0.76	0.88
MISO	0.30	0.19	0.02	0.08	0.12	0.27	0.49	0.58	0.71
PJM	0.33	0.22	0.01	0.07	0.12	0.30	0.53	0.64	0.89
50 Hz	0.20	0.20	0.01	0.03	0.05	0.13	0.35	0.52	0.80

6.2 Statistical Analysis of Uncertainty

Several statistical quantities of instantaneous aggregate wind power for each system for 2012 and 2008 are provided in Tables 2 and 3, respectively. In addition to the mean and standard deviation, several quantiles were computed, where, for example, Q(50) refers to the median.

The mean values range between 0.18 and 0.33 p.u., and the standard deviation ranges from 0.17 to 0.29 p.u. These typical ranges of mean and standard deviation can be used to construct models of aggregate wind power, as discussed in Sect. 5.2. The Q(99) quantiles indicates the magnitude of rare wind power events. For example, for BPA in 2012, in 99 % of the hours the aggregate wind power was less than or equal to 85 % of the rated capacity. This means that 1 % of year—approximately 88 h—the power was greater than 0.85 p.u. For all systems, the values corresponding to the 99 % quantile ranged between 0.68 and 0.88 p.u. Extremely low values were common: the 10 % quantile is 0.08 p.u. or less in all cases. In other words, in each system the power output is less than 8 % of rated capacity for nearly 876 h each year. The relatively high frequency of low power output has system-wide generation capacity implications. In general, the systems have very different levels of uncertainty that must be managed by their operators. By several measures, MISO, PJM, and 50 Hz have less uncertainty than BPA.

6.3 Statistical Analysis of Variability

The statistics for hour-to-hour variation for each system in 2012 and 2008 are provided in Tables 4 and 5, respectively. The variations of aggregate wind power

Table 4 Aggregate wind power hourly variation for 2012

System	Standard deviation (p.u.)	Kurtosis	Q(1) (p.u.)	Q(10) (p.u.)	Q(20) (p.u.)	Q(50) (p.u.)	Q(80) (p.u.)	Q(90) (p.u.)	Q(99) (p.u.)
BPA	0.048	7.0	−0.130	−0.053	−0.029	0.00	0.026	0.057	0.144
MISO	0.030	4.4	−0.79	−0.036	−0.022	0.00	0.021	0.036	0.083
PJM	0.032	28.2	−0.84	−0.035	−0.022	0.00	0.021	0.037	0.083
50 Hz	0.024	6.5	−0.66	−0.027	−0.015	0.00	0.014	0.027	0.072

Table 5 Instantaneous aggregate wind power hourly variation for 2008

System	Standard deviation (p.u.)	Kurtosis	Q(1) (p.u.)	Q(10) (p.u.)	Q(20) (p.u.)	Q(50) (p.u.)	Q(80) (p.u.)	Q(90) (p.u.)	Q(99) (p.u.)
BPA	0.054	8.2	−0.138	−0.061	−0.034	0.00	0.030	0.061	0.167
MISO	0.034	5.0	−0.093	−0.041	−0.024	0.00	0.024	0.040	0.089
PJM	0.049	7.3	−0.128	−0.055	−0.032	0.00	0.031	0.055	0.131
50 Hz	0.026	7.0	−0.075	−0.027	−0.015	0.00	0.015	0.029	0.073

in large systems tend to be symmetric, with near-zero mean and median. In most system, the variations are less than 0.06 p.u. in magnitude for over 80 % of the hours. Extreme outliers corresponding to changes of ± 0.08 to 0.17 p.u. do occur, but are rare. The standard deviations range from 0.024 to 0.054 p.u. In each system, the variations are leptokurtic, indicating a higher occurrence of "tail event" variations. In general, MISO, PJM, and 50 Hz have less variability than BPA.

6.4 Effect of Capacity on Uncertainty and Variability

The effect of capacity and capacity increases on aggregate wind power uncertainty and variability are often of interest as system operators anticipate and plan for increased amounts of wind plants in their system. The specifics of how the uncertainty and variability associated with aggregate wind power are related to capacity depend on the build-out of the wind plants in the system, and are difficult to analytically derive. However, based on the available data and past experience, there are several general observations that can be made:

1. Systems with larger installed capacities do not necessarily exhibit less uncertainty and variability in aggregate wind power than systems with smaller capacities.
2. Additions of wind plant capacity to a system tend reduce uncertainty and variability.
3. Within a given system, the benefits of geographic diversity can become saturated and insensitive to increases in capacity.

As an example of the first observation, we compare the statistics of instantaneous power in PJM and BPA. At the end of 2012, there was 4.7 GW of installed wind capacity in BPA. At the end of 2008 in PJM there was 1.3 GW of installed wind capacity. It would seem that BPA should have less uncertainty given its much higher installed capacity and therefore greater opportunity for geographic diversity. However, the standard deviation of PJM (0.22 p.u.) was less than that of BPA (26 p.u.). BPA also generally exhibited more occurrences of extremely low and high power production, as shown by the quantiles in Tables 2 and 3. A similar example for variability can be made by comparing 2008 MISO data with 2008 BPA data. These cases should not be misinterpreted as evidence that systems with higher capacities of wind plants have greater uncertainty and variability than those with lower capacities. Rather, they are simply examples that the contrary is not always the case. The important concept is that different systems experience different build-outs of wind plants—some lead to appreciable geographic diversity, others do not.

The second observation—that additions of wind plant capacity to a system tend reduce uncertainty and variability—is observed by comparing the statistics of each system in 2008–2012. In each system, the standard deviation of the instantaneous power and power variation decreased from 2008 to 2012. The exception is MISO, whose standard deviation of instantaneous power remained the same. Regardless, the data support the general notion that wind plant capacity additions decrease uncertainty and variability.

The final observation—that the benefits of uncertainty and variability reduction can become saturated—is supported by examining the sensitivity of uncertainty and variability to capacity addition. For BPA, PJM and 50 Hz, for every 1 GW of new wind plant installations, the standard deviation of instantaneous wind power decreased by just one percentage point, based upon year-end capacity values. For MISO, the standard deviation did not change, despite an increase in 8 GW of wind power. The quantiles in many systems also showed modest changes despite large capacity additions. The modest effects of capacity additions on uncertainty can be seen by inspecting Figs. 13 and 14. The histograms and approximated probability density functions of the system do not appear appreciably different in 2012 than they did in 2008, despite several thousand megawatts of wind plant installations in each system. The histograms and probability density functions of variations in Figs. 15 and 16, however, have more noticeable differences.

The sensitivity of variability to changes in capacity are mixed. The standard deviation of wind power variability in BPA and PJM decreased by approximately 0.25 % point for each 1 GW of new capacity. A change in percentage point of this magnitude is appreciable since the standard deviations of variation are all less than 6 %. However, both MISO and 50 Hz exhibited small changes to their standard deviation between 2008 and 2012.

7 Conclusions

Many power systems around the world now have total installed wind plant capacities in excess of several gigawatts. System operators must manage the inherent uncertainty and variability of the aggregate wind power in their systems in order to maintain reliability and economic efficiency. The characteristics of aggregate wind power can be quite different from individual wind plants, depending on the level of geographic diversity in the system. Tools such as wind power forecast systems, stochastic unit commitment, and resource planning require reasonable and practical probabilistic models of aggregate wind power and wind power variation as inputs.

This chapter demonstrated that the uncertainty and variability exhibited by aggregate wind power can be reasonably represented using parsimonious parametric models. More specifically, the two-parameter Beta distribution is well suited for modeling instantaneous aggregate wind power and the two-parameter Laplace distribution is well suited for modeling moment-to-moment variation.

Several aspects of geographic diversity were explored. Among the main conclusions are that geographic diversity should not be viewed as a panacea for the challenges of wind integration. Reduction in variability—the smoothing effect—is noticeable, but reduction in uncertainty requires exceedingly large geographic areas. Appreciable correlation of instantaneous power amongst wind plants can exist at distances approaching 1,000 km. Clustering and other practical considerations also limit the amount of geographic diversity that occurs in many systems. Analysis of several systems showed that the effects of geographic diversity, particularly on uncertainty, saturate when installations of wind plants reach several gigawatts in total.

References

1. Burton T, Sharpe D, Jenkins N, Bossanyi E (2001) Wind energy handbook. Wiley, West Sussex
2. Smith JC, Milligan M, DeMeo EA, Parsons B (2007) Utility wind integration and operating impact state of the art. IEEE Trans Power Syst 22:900–908. doi:10.1109/TPWRS.2007.901598
3. Smith JC, Thresher R, Zavadil R, DeMeo EA, Piwko R, Ernst B, Ackerman T (2009) A mighty wind. IEEE Power Energy Mag 7:41–51. doi:10.1109/MPE.2008.931492
4. Tuohy A, Meibom P, Denny E, O'Malley M (2009) Unit commitment for systems with significant wind penetration. IEEE Trans Power Syst 24:592–601. doi:10.1109/TPWRS.2009.2016470
5. Ruiz P, Philbrick CR, Zak E, Cheung K, Sauer P (2008) Applying stochastic programming to the unit commitment problem. In: Probabilistic methods applied to power systems; 2008
6. Pinson P, Kariniotakis G (2010) Conditional prediction intervals of wind power generation. IEEE Trans Power Syst 25:1845–1856
7. Hasche B (2010) General statistics of geographically dispersed wind power. Wind Energy 13:773–784. doi:10.1002/we.397

8. EnerNex Corporation (2011) Eastern wind integration and transmission study. Technical report NREL/SR-5500-47078, NREL, Golden, CO, USA
9. GE Energy (2010) Western wind and solar integration study. Technical report NREL/SR-550-47434, NREL, Golden, CO, USA 2010
10. McNerney G, Richardson R (1992) The statistical smoothing of power delivered to utilities by multiple wind turbines. IEEE Trans Energy Convers 7(4):644–647. doi:10.1109/60.182646
11. Archer C, Jacobson M (2003) Spatial and temporal distributions of U.S. winds and wind power at 80 m derived from measurements. J Geophys Res 108(D9):10–1–10–20
12. Wan Y (2004) Wind power plant behaviors: analyses of long-term wind power data. Technical report NREL/TP-500-36551
13. Holttinen H (2005) Hourly wind power variations in the Nordic countries. Wind Energy 8:173–195
14. Ernst B, Wan Y, Kirby B (1999) Short-term power fluctuation of wind turbines: analyzing data from the German 250-MW measurement program from the ancillary services viewpoint. Technical report NREL/CP-500-26722
15. Wan Y, Milligan M, Parsons B (2003) Output power correlation between adjacent wind power plants. J Sol Energy Eng 125:551–555
16. Louie H (2013) Correlation and statistical characteristics of aggregate wind power in large transcontinental systems. Wind Energy. doi:10.1002/we.1597
17. Tastu J, Pinson P, Kotwa E, Madsen H, Nielsen H (2011) Spatio-temporal analysis and modeling of short-term wind power forecast errors. Wind Energy 14:43–60. doi:10.1002/we.401
18. Nanahara T, Asari M, Maejima T, Sato T, Yamaguchi K, Shibata M (2004) Smoothing effects of distributed wind turbines. Part 2. Coherence among power output of distant wind turbines. Wind Energy 7:75–85. doi:10.1002/we.108
19. Krich A, Milligan M (2005) The impact of wind energy on hourly load following requirements: an hourly and seasonal analysis. Technical report NREL/CP-500-38061
20. Wan Y (2011) Analysis of wind power ramping behavior in ERCOT. Technical report NREL/TP-5500-49218 2011
21. Gibescu M, Brand A, Kling W (2008) Estimation of variability and predictability of large-scale wind energy in the Netherlands. Wind Energy 12:241–260. doi:10.1002/we.291
22. Milligan M (2000) Modelling utility-scale wind power plants. Part 2: Capacity credit. Wind Energy 3:167–206. doi:10.1002/we.36
23. Sloughter JM, Gneiting T, Raftery AE (2010) Probabilistic wind speed forecasting using ensembles and Bayesian model averaging. J Am Stat Assoc 105:25–35. doi:10.1198/jasa.2009.ap08615
24. Papoulis A, Pillai SU (2002) Probability, random variables and stochastic processes, 4th edn. McGraw-Hill, New York
25. Justus CG, Hargraves WR, Mikhail A, Graber D (1978) Methods for estimating wind speed frequency distributions. J Appl Meteorol 17:350–353
26. Tuzuner A, Yu Z (2008) A theoretical analysis on parameter estimation for the weibull wind speed distribution. IEEE PES General Meeting 2008
27. Twidell J, Weir T (2006) Renew Energy Res, 2nd edn. Taylor & Francis, London
28. International Electrotechnical Commission (2005), Power performance measurements of electricity producing wind turbines. Standard 61400-12-1
29. Camm EH, Behnke MR, Bolado O et al (2009) Wind power plant substation and collector system redundancy, reliability and economics. IEEE PES general meeting 2009
30. Fischer K, Besnard F, Bertling L (2012) Reliability-centered maintenance for wind turbines based on statistical analysis and practical experience. IEEE Trans Energy Convers 27(184):195. doi:10.1109/TEC.2011.2176129
31. Potter CW, Gil H, McCaa J (2007) Wind power data for grid integration studies. IEEE PES general meeting

32. Hayes B, Ilie I, Porpodas A, Djokic S, Chicco G. Equivalent power curve model of a wind farm based on field measurement data. In: IEEE PowerTech; 2011
33. Jin T, Tian Z (2010) Uncertainty analysis for wind energy production with dynamic power curves. In: Probabilistic methods applied to power systems; 2010
34. Collins J, Parkes J, Tindal A (2009) Forecasting for utility-scale wind farms—the power model challenge. In: CIGRE/IEEE joint symposium on integration of wide-scale renewable resources into the power delivery system 2009
35. Kendall M (1938) A new measure of rank correlation. Biometrika 30:81–89
36. Louie H (2012) Evaluating Archimedean copula models of wind speed for wind power modeling. Power Africa 2012:1–5. doi:10.1109/PowerAfrica.6498610
37. Wind integration datasets (2011) National Renewable Energy Laboratory. http://www.nrel.gov/wind/integrationdatasets. Accessed 1 July 2013
38. Osborn D, Hendersen M, Nickell B, Lasher W, Liebold C, Adams J, Caspary J (2011) Driving forces behind wind. Power Energy Mag 9:60–74
39. Nelsen R (2006) An introduction to copulas, 2nd edn. Springer, New York
40. Louie H (2012) Evaluation of bivariate Archimedean and elliptical copulas to model wind power dependency structures. Wind Energy. doi:10.1002/we.1571
41. Díaz G (2013) A note on the multivariate Archimedean dependence structure in small wind generation sites. Wind Energy. doi:10.1002/we.1633
42. Louie H (2010) Characterizing and modelling aggregate wind plant power output in large systems. IEEE PES general meeting; 2010, pp 1–8
43. Beckman RJ, Tietjen GL. Maximum likelihood estimation for the beta distribution. J Stat Comput Simul 7:253–258
44. Louie H (2010) Evaluation of probabilistic models of wind plant power output characteristics. In: Probabilistic methods applied to power systems; 2010, pp 442–447, doi: 10.1109/PMAPS.2010.5528963
45. McManus B (2013) Wind generation & total load in the BPA balancing authority. Bonneville Power Administration. http://www.transmission.bpa.gov/Business/Operations/Wind/default.aspx. Accessed 1 July 2013
46. Market reports (2013). MISO. https://www.midwestiso.org/Library/MarketReports/. Accessed 1 July 2013
47. Operational analysis (2013). PJM interconnection www.pjm.com/markets-and-operations/ops-analysis.aspx. Accessed 1 July 2013
48. Archive wind power (2013). 50 Hertz. http://www.50hertz.com/en/1983.htm. Accessed 1 July 2013

http://www.springer.com/978-981-4585-29-3

Large Scale Renewable Power Generation
Advances in Technologies for Generation, Transmission and Storage
Hossain, J.; Mahmud, A. (Eds.)
2014, XII, 462 p. 285 illus., 170 illus. in color.,
Hardcover
ISBN: 978-981-4585-29-3