Joint Probability Distribution and Correlation Analysis of Wind and Solar Power Forecast Errors in the Western Interconnection

Jie Zhang[1]; Bri-Mathias Hodge[2]; and Anthony Florita[3]

Abstract: Wind and solar power generation differ from conventional power generation because of the variable and uncertain nature of their power output. This can have significant impacts on grid operations. Short-term forecasting of wind and solar power generation is uniquely helpful for planning the balance of supply and demand in the electric power system because it allows for a reduction in the uncertainty associated with their output. As a step toward assessing the simultaneous integration of large amounts of wind and solar power, this article investigates the spatial and temporal correlation between wind and solar power forecast errors. The forecast and actual data analyzed are obtained from one of the world's largest regional variable generation integration studies to date. Multiple spatial and temporal scales (day ahead, 4 h ahead, and 1 h ahead) of forecast errors for the Western Interconnection in the United States are analyzed. A joint probability distribution of wind and solar power forecast errors is estimated using kernel density estimation. The Pearson's correlation coefficient and mutual information between wind and solar power forecast errors are also evaluated. The results show that wind and solar power forecast errors are inversely correlated, and the correlation between wind and solar power forecast errors becomes stronger as the geographic size of the analyzed region increases. The absolute value of the correlation coefficient is generally less than 0.1 in the case of small geographic regions, while it is generally between 0.15 and 0.6 in the case of large geographic regions. The forecast errors are less correlated on the day-ahead timescale, which influences economic operations more than reliability, and more correlated on the 4-h-ahead timescale, where reliability is more impacted by the forecasts. It is also found that the correlation between wind and solar power forecast errors in summer (July) is relatively stronger than in winter (January). The inverse correlation implies that in systems with high penetrations of both wind and solar power, reserves that are held to accommodate the variability of wind or solar power can be at least partially shared. In addition, interesting results are found through time and seasonal variation analyses of wind and solar power forecast errors, and these insights may be uniquely useful to operators who maintain the reliability of the electric power system. **DOI: 10.1061/(ASCE)EY.1943-7897.0000189.**

Author keywords: Wind forecasting; Solar forecasting; Smart grids; Uncertainty; Spatial and temporal correlation.

Introduction

Variable renewable energy resources, such as wind and solar power, are becoming increasingly important sources of energy on the electric power system. It has been suggested that the United States can produce 20% of its electric power needs from wind power plants by the year 2030 (Lindenberg 2008). The *Utility Solar Assessment* study reported that solar power could provide 10% of U.S. power needs by 2025 (Pernick and Wilder 2008). At such high levels of renewable energy penetration, wind and solar power forecasting becomes very important for electric power system operations. One of the critical challenges of integrating wind and solar power generation into electric power system operations is the variable and uncertain nature of such resources. Because grid operators must continuously balance supply and demand to maintain the reliability of the electric power system, forecast inaccuracies can result in substantial economic losses and reliability risks. Although forecasting systems are improving, they are not perfect, and to some extent, wind and solar power forecast errors will always be present. To maximize the economic benefits of integrating variable renewable generation, it is crucial that grid operators understand the patterns and relationships of wind and solar power forecast errors. A better understanding of the correlation between wind and solar power forecast errors will lead to reduced electric power system production costs through superior operational planning.

Overview of Wind Forecasting

Wind power forecast models can be broadly divided into two categories (Foley et al. 2012): (1) data-driven models, such as forecasting based on the analysis of recent time series of wind; and (2) first-principle models, such as forecasting based on numerical weather prediction (NWP) models. The first type of forecast models generally use statistical approaches to provide reasonable results in the estimation of long-term horizons, such as mean monthly, quarterly, and annual wind speed. The second type of forecast models generally use explanatory variables (mainly hourly mean wind speed and direction) derived from a meteorological model of the wind dynamics to predict wind power. Statistical and machine learning techniques that utilize historical data have been shown

[1]Postdoctoral Researcher, National Renewable Energy Laboratory, Golden, CO 80401. E-mail: jie.zhang@nrel.gov

[2]Senior Research Engineer, National Renewable Energy Laboratory, Golden, CO 80401 (corresponding author). E-mail: bri.mathias.hodge@nrel.gov

[3]Research Engineer, National Renewable Energy Laboratory, Golden, CO 80401. E-mail: anthony.florita@nrel.gov

Note. This manuscript was submitted on May 31, 2013; approved on January 17, 2014; published online on May 15, 2014. Discussion period open until October 15, 2014; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Energy Engineering***, © ASCE, ISSN 0733-9402/B4014008(13)/$25.00.

to work well for forecast horizons of less than 1 h (Kariniotakis et al. 1996; Monteiro et al. 2009). For short-term horizons greater than 1 h, the impact of atmospheric dynamics becomes more important, and NWP models become more suitable. Short-term wind power forecasting (between 1 and 72 h) is uniquely helpful in power system operations because of the role that it can play in the unit commitment and economic dispatch process, which is also a focus of this paper. Analyzing the statistical properties of short-term forecasts is of interest for uncovering error patterns likely caused by local geographical features not considered in coarser-resolution NWP models.

To understand the impact of wind power forecast errors on wind integration, it is important to characterize wind power forecast errors at all timescales of interest, especially extreme errors. For a more thorough review of wind power forecasting methods, please see Monteiro et al. (2009) and Foley et al. (2012). A variety of topics on wind power forecast errors and their impacts on system operations have been studied in the literature. Distributions of wind power forecast errors were analyzed in order to demonstrate how operational forecasts are currently performing (Bludszuweit et al. 2008; Hodge and Milligan 2011; Bruninx et al. 2013; Zhang et al. 2013b, c). Multiple types of distribution methods have been utilized to characterize the wind power forecast error distribution, including the normal distribution (Methaprayoon et al. 2007; Castronuovo and Lopes 2004), Weibull distribution (Dietrich et al. 2009), Beta distribution (Bludszuweit et al. 2008), and hyperbolic distribution (Hodge et al. 2011, 2012b). The shape of the distribution was found to change significantly with the length of the forecasting timescale (Hodge and Milligan 2011). Uncertainties in wind forecasting were quantified and analyzed in Bessa et al. (2011) and Hodge et al. (2012c). An information entropy approach was proposed by Bessa et al. (2011) for assessing wind forecasting methods; and Rényi entropy was adopted in Zhang et al. (2013b) and Hodge et al. (2012c) to quantify the uncertainty in wind power forecasting.

Overview of Solar Forecasting

Solar irradiance variations are caused primarily by cloud movement, formation, and dissipation. In the literature, researchers have developed a variety of methods for solar power forecasting, such as the use of NWP models (Marquez and Coimbra 2011; Mathiesen and Kleissl 2011; Chen et al. 2011), tracking cloud movements from satellite images (Hammer et al. 1999; Perez et al. 2007), and tracking cloud movements from direct ground observations with sky cameras (Crispim et al. 2008; Chow et al. 2011; Marquez and Coimbra 2013). NWP models are the most popular method for forecasting solar irradiance several hours or days in advance. Mathiesen and Kleissl (2011) analyzed the global horizontal irradiance in the continental United States forecasted by three popular NWP models: the North American model, the Global Forecast System, and the European Centre for Medium-Range Weather Forecasts. Chen et al. (2011) developed an advanced statistical method for solar power forecasting based on artificial intelligence techniques. Crispim et al. (2008) used total sky imagers (TSIs) to extract cloud features using a radial basis function neural network model for time horizons from 1 to 60 min. Chow et al. (2011) also used TSIs to forecast short-term global horizontal irradiance. The results suggested that TSIs were useful for forecasting time horizons up to 15 to 25 min. Marquez and Coimbra (2013) presented a method using TSI images to forecast 1-min averaged direct normal irradiance at the ground level for time horizons between 3 and 15 min. As discussed previously, different solar irradiance forecast methods have been developed for various timescales. Loren et al. (2007) showed that cloud movement–based forecasts likely provide better results than NWP forecasts for forecast timescales of 3 to 4 h or less. Beyond that, NWP models tend to perform better. For the analysis of the solar power forecast error distribution, Hodge et al. (2011) analyzed solar ramping distributions at different timescales and for different weather patterns.

Research Motivation and Objectives

Wind and solar power forecast errors are generally important factors in variable renewable generation integration studies. The accuracy of wind and solar power forecast error distributions can have a significant impact on the uncertainty associated with wind and solar power forecasting and, hence, with the amount of reserves carried to accommodate these errors. The uncertainty can be estimated based on an assumed error distribution around the point forecasts. Understanding the correlation between wind power forecast errors and solar power forecast errors at different spatial and temporal scales can provide a better understanding of the flexibility requirements and reliability impacts of wind and solar integration on the grid. Therefore, the overall objective of this paper is to comprehensively analyze wind and solar power forecast errors using the following approaches:

1. Develop a model to represent and analyze the joint distribution of wind and solar power forecast errors;
2. Investigate the correlation between wind and solar power forecast errors at multiple temporal scales, specifically, day-ahead, 4-h-ahead, and 1-h-ahead wind and solar power forecast errors;
3. Investigating the correlation between wind and solar power forecast errors at multiple spatial scales, including one electric bus that consisted of both wind and solar power generations, a group of electric buses, and all wind power plants and solar power plants in an interconnection area;
4. Analyze variations in wind and solar power forecast errors at different times of day and in different seasons.

The remainder of the paper is organized as follows. The next section describes the methodology used to characterize the joint probability distribution of forecast errors and the correlation between wind and solar power forecast errors. The following section summarizes the data analyzed in the paper. Then the results and discussion of the case study based on bus numbers are presented. Finally, concluding remarks and ideas on areas for future exploration are given.

Spatial and Temporal Correlation between Wind and Solar Power Forecast Errors

Wind and Solar Power Forecast Errors

The distributions of wind and solar power forecast errors at multiple spatial and temporal scales were investigated in this work. The three forecasting timescales analyzed in this study were day ahead, 4 h ahead, and 1 h ahead. The forecast errors were calculated using the following equations:

$$e_w = P_{wf} - P_{wa} \tag{1}$$

$$e_s = P_{sf} - P_{sa} \tag{2}$$

where e_w and e_s = wind and solar power forecast errors, respectively; P_{wf} and P_{wa} = forecast and actual wind power generations, respectively; and P_{sf} and P_{sa} = forecast and actual solar power generations, respectively.

Correlation between Wind and Solar Power Forecast Errors

To evaluate the correlation between wind and solar power forecast errors, Pearson's correlation coefficient was used. The mutual information between wind and solar power forecast errors was also evaluated, which is a generalization of correlation for all moments. Distributions of forecast errors were evaluated with the kernel density estimation (KDE) method, and the Rényi entropy was used to characterize the uncertainties in wind and solar power forecast error distributions. Daily and seasonal variations in wind and solar power forecast errors were also analyzed.

The lack of solar photovoltaic energy generation at night is one concern with high penetrations of solar energy. Because of this lack of generation, a large portion of solar power forecast errors are equal to zero. These zero-magnitude solar power forecast errors do not reflect the accuracy or ability of the forecast methods and thus were removed in evaluations of the distribution of solar power forecast errors. For the distribution of wind power forecast errors, the original data set was used. When evaluating the joint distribution of wind and solar power forecast errors, to match the wind and solar data set, wind power forecast errors that corresponded to times of zero solar power output were also removed from the evaluated data set.

Kernel Density Estimation

Kernel density estimation (KDE), also known as the Parzen-Rosenblatt window method (Rosenblatt 1956; Parzen 1962), is a nonparametric approach used to estimate the probability density function of a random variable. KDE has been widely used in the wind energy community for wind speed distribution characterization (Zhang et al. 2013a, 2011; Chowdhury et al. 2013; Qin et al. 2011), wind power density estimation (Jeon and Taylor 2012), and wind power forecasting (Juban et al. 2007). For an independent and identically distributed sample, $x_1, x_2, \ldots, x_n$, drawn from some distribution with an unknown density f, the KDE is defined as (Jones et al. 1996)

$$\hat{f}(x;h) = \frac{1}{n}\sum_{i=1}^{n} K_h(x - x_i) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right) \tag{3}$$

In the equation, $K_h(\cdot) = (1/h)K(\cdot\ /h)$ has a kernel function K (often taken to be a symmetric probability density) and a bandwidth h (the smoothing parameter). For a d-variate random sample X_1, $X_2, \ldots, X_n$ drawn from a density f, the multivariate KDE is defined as

$$\hat{f}(x;H) = \frac{1}{n}\sum_{i=1}^{n} K_H(x - X_i) \tag{4}$$

where $x = (x_1, x_2, \ldots, x_d)^T$, $X_i = (X_{i1}, X_{i2}, \ldots, X_{id})^T$, and $i = 1, 2, \ldots, n$. Here, $K(x)$ is the kernel that is a symmetric probability density function, H is the bandwidth matrix that is symmetric and positive-definite, and $K_H(x) = |H|^{-1/2}K(H^{-1/2}x)$. The choice of K is not crucial to the accuracy of KDEs (Epanechnikov 1969). In this paper, the Gaussian kernel, $K(x) = (2\pi)^{-d/2}\exp(-1/2x^T x)$, is considered throughout. In contrast, the choice of H is crucial in determining the performance of $\hat{f}$ (Duong and Hazelton 2003). For dependent data sequences, which is bound to be the case with our data set, Hall et al. (1995) proved that the asymptotically optimal bandwidth for independent data was a good choice and suggested the plug-in empirical bandwidth selector based on this observation. The mean integrated squared error, the most commonly used optimality criterion (Duong and Hazelton 2003), is used in this article.

Pearson's Correlation Coefficient

Pearson's correlation coefficient is a measure of the correlation between two variables (or sets of data) (Rodgers and Nicewander 1988). In this article, the Pearson's correlation coefficient between wind and solar power forecast errors is evaluated. Pearson's correlation coefficient, ρ, is defined as the covariance of wind and solar power forecast error variables divided by the product of their standard deviations, which is mathematically expressed as

$$\rho = \frac{\mathrm{cov}(e_w, e_s)}{\sigma_{e_w}\sigma_{e_s}} \tag{5}$$

where e_w and e_s represent the wind and solar power forecast errors, respectively.

Rényi Entropy and Mutual Information

Forecasting metrics such as root-mean-square error and mean absolute error are unbiased only if the error distribution is Gaussian. Therefore, new metrics have been proposed based on the use of concepts from information theory that use all of the information present in the forecast error distributions. An information entropy approach was proposed in the literature (Bessa et al. 2011; Hodge et al. 2012c) for assessing wind forecasting methods. Entropy in information theory is a measure of the uncertainty in a random variable; a smaller information entropy value indicates less uncertainty in the forecasting. This information entropy approach based on Rényi entropy is adopted here to quantify the uncertainty in wind and solar power forecasting. The Rényi entropy is defined as

$$H_\alpha(X) = \frac{1}{1-\alpha}\log_2\sum_{i=1}^{n} p_i^\alpha \tag{6}$$

where α = a parameter that allows the creation of a spectrum of Rényi entropies; and p_i = probability density of the ith discrete section of the distribution. Large values of α favor high-probability events, whereas small values of α weight all of the instances more evenly (Hodge et al. 2012c).

The mutual information of two random variables measures the mutual dependence of the two variables, which is defined as (Cover and Thomas 2006)

$$I(X;Y) = \sum_{y\in Y}\sum_{x\in X} p(x,y)\log\left[\frac{p(x,y)}{p(x)p(y)}\right] \tag{7}$$

where $p(x,y)$ = joint probability distribution function of X and Y; and $p(x)$ and $p(y)$ = marginal probability distribution functions of X and Y, respectively. Mutual information is a quantitative measurement of how much the wind (or solar) power forecast error tells us about the solar (or wind) power forecast error. The more mutual information between wind and solar power forecast errors, the less uncertainty there is in wind (or solar) power forecast errors when the solar (or wind) power forecast errors are known.

Data Summary

The data used in this work were obtained from the *Western Wind and Solar Integration Study Phase 2* (WWSIS-2), which is one of the world's largest regional variable generation integration studies to date (Lew et al. 2010, 2013). Five scenarios were created in the WWSIS-2 (Lew et al. 2013); the high-mix scenario—16.5% wind and 16.5% solar—was adopted in this article for the correlation analysis. A brief summary of the wind and solar data sets is given in the following sections.

Wind Data Sets

For the WWSIS, wind speeds were synthesized using a NWP model on a 10-min, 2-km interval. Simulated wind plant power output for the years 2004–2006 was generated, referred to here as the actuals (Potter et al. 2008). Each wind power plant was assumed to consist of 10 3-MW turbines. In this article, the 60-min wind plant power output for 2006 was used as the actual data.

The day-ahead wind power forecasts were synthesized using the same NWP model as the actuals with a different input data set and at a different geographic resolution. The details of the data can be found in the WWSIS Phase 1 report (Lew et al. 2010) and the accompanying journal article on the data-set creation (Potter et al. 2008). The day-ahead wind power forecast errors from WWSIS Phase 1 were analyzed for the Public Service Company of Colorado (Denver, CO), the California Independent System Operator (Folsom, CA), and the Electric Reliability Council of Texas (Taylor, TX). The distributions of wind power forecast errors from WWSIS Phase 1 were adjusted to match the measured wind power forecast error distributions through statistical corrections, based on the first four statistical moments (mean, variance, skewness, and kurtosis). Although some of the bias was adjusted out through this process, some bias remained in the year-to-year error distribution as a remnant of the original forecasts. The details of the adjustment process can be found in Hodge et al. (2012a). The 4-h-ahead forecasts were synthesized using a 2-h-ahead persistence approach since a 2-h-ahead persistence approach most accurately resembled the forecast error distribution produced by state-of-the-art forecasting systems at a 4-h horizon. The WWSIS-2 report (Lew et al. 2013) examined 1-h, 2-h, 3-h, and 4-h persistence error distributions and found that the 2-h-ahead persistence distribution matched the operational 4-h-ahead forecast the best, especially at the tails of the distribution, which were the critical operational impact events. The 1-h-ahead forecasts were synthesized using a 1-h-ahead persistence approach. More information can be found in the WWSIS-2 report (Lew et al. 2013).

Solar Data Sets

The solar data were synthesized using the algorithm developed by Hummon et al. (2012). The algorithm generated synthetic global horizontal irradiance values based on a 1-min interval using satellite-derived, 10×10-km gridded, hourly irradiance data. In this article, the 60-min solar power plant output for 2006 was used as the actual data. The solar power output data include distributed-generation rooftop photovoltaic, utility-scale photovoltaic, and solar power concentrated with thermal storage.

Day-ahead solar forecasts were taken from the WWSIS Phase 1 solar forecasts conducted by 3Tier (Seattle, Washington) based on NWP simulations (Lew et al. 2010). No regional operational solar forecast error data were available for use to adjust these forecasts for current forecast methodologies, so these forecasts were unchanged from the WWSIS Phase 1. The 4-h-ahead forecasts were synthesized using a 2-h-ahead persistence of cloudiness approach. The 1-h-ahead forecasts were synthesized using a 1-h-ahead persistence of cloudiness approach. In this method, the solar power index (SPI) is first calculated, which represents the ratio between actual power (P) and clear sky power (P_{CS}). Then the solar forecast power is estimated by modifying the current power output by the expected change in clear sky output. For the 1-h-ahead persistence of cloudiness approach, the forecast solar power at time $t+1$ can be calculated as follows:

$$P(t+1) = P(t) + \mathrm{SPI}(t) \times [P_{CS}(t+1) - P_{CS}(t)] \quad (8)$$

where $P_{CS}(t+1)$ and $P_{CS}(t)$ = clear sky solar power at time $t+1$ and t, respectively; $P(t)$ = actual solar power output at time t; and $\mathrm{SPI}(t)$ = clear sky index at time t.

It is important to note that for shorter-term forecasts, the persistence approach is one of the most popularly used forecasting methods. Day-ahead forecasts are typically used for the unit commitment problem, and NWP models are often the basis for these forecasts. Economic dispatch can utilize forecasts from 10 min up to 1 or 2 h ahead, and at these timescales most models (weather physics–based or statistical) only provide slight improvements over the persistence method. A persistence approach of the 1- and 2-h-ahead forecasting also indicates the variability of the forecasts, and in this way the correlation in variability could also be quantified.

Analysis Based on Bus Enumeration

In this section, wind and solar power forecast error correlations are analyzed based on their grouping according to bus numbering. Each bus may aggregate multiple wind and solar power plants. In total, wind power and solar power were aggregated into 115 and 424 buses, respectively. The wind power capacity on each bus varied from 30 to 2,400 MW; the solar power capacity on each bus varied from 0.5 to 700 MW. Among the 115 sets of wind data and 424 sets of solar data, there were 35 pairs of data sets that had the same bus number. Three different cases were analyzed based on the bus numbers of the wind and solar power plants. The first case found all the buses that had both wind and solar power generation and investigated the correlation between wind and solar power forecast errors for each pair of wind and solar outputs. The second case analyzed wind and solar power forecast error correlation for all 35 pairs of wind and solar outputs considering wind power forecast error aggregation and solar power forecast error aggregation. The third case investigated the correlation of the aggregation of all wind power plant forecast errors and aggregated solar power plant forecast errors within the Western Interconnection of the United States. It is important to note that in figures thoughout the article, a positive error indicates an overforecast event, and a negative error indicates an underforecast event.

Case I: Results and Discussion

The first case examined wind and solar power plants located on the same bus and investigated the correlation between wind and solar power forecast errors for each pair of wind and solar outputs. The wind and solar power forecast errors within each bus are normalized by the average actual wind power and the average clear sky solar power of the analyzed time period, respectively.

Distribution of Wind and Solar Power Forecast Errors

All 35 pairs of data were analyzed in the first case. For brevity, the results for two pairs of wind and solar outputs are provided to show the diversity of distribution behavior. Figs. 1 and 2 show two typical types of joint distributions of wind and solar power forecast errors. Figs. 1(a–c) illustrate the distributions for day-ahead, 4-h-ahead, and 1-h-ahead forecast errors, respectively; Figs. 1(d and e) show the univariate distribution of wind and solar power forecast errors, respectively. For the 35 pairs of data sets, 105 joint distributions of wind and solar power forecast errors, including day-ahead, 4-h-ahead, and 1-h-ahead forecasts, were obtained. Among the 105 distributions, only 4 were found to be multimodal, and 1 multimodal example distribution is shown in

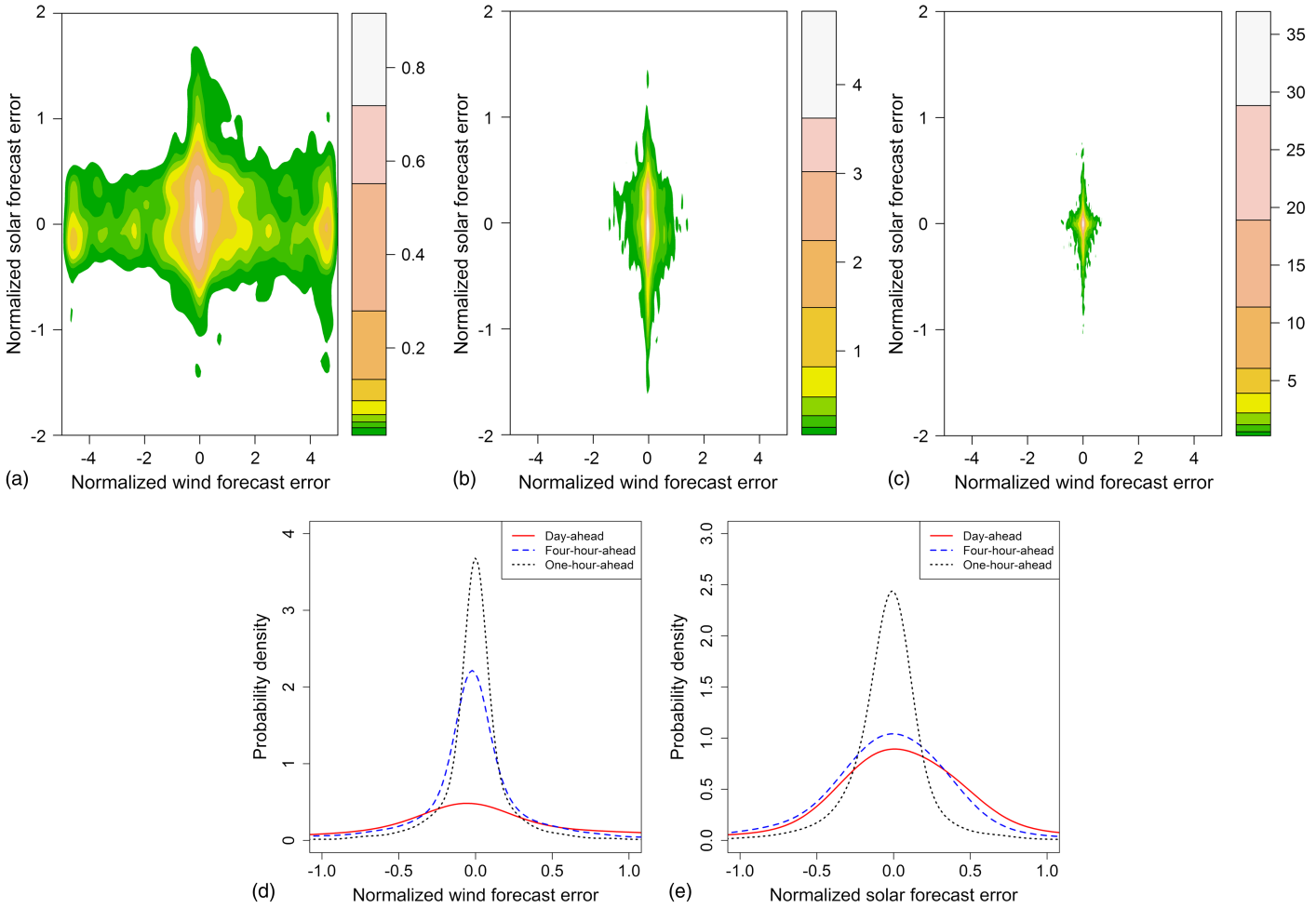

Fig. 1. Probability distributions of wind and solar power forecast errors of pair 11: (a) Day-ahead joint distribution; (b) 4-h-ahead joint distribution; (c) 1-h-ahead joint distribution; (d) univariate distribution of wind power forecast error; (e) univariate distribution of solar power forecast error

Fig. 1(a). A further investigation found that for the pair-11 bus, one major mode and two small modes were present in the univariate distribution of wind power forecast errors, and the univariate distribution of solar power forecast errors was unimodal [shown in Fig. 1(e)]. For the pair-35 bus shown in Fig. 2, the univariate distributions of both wind and solar power forecast errors are unimodal [shown in Figs. 2(d and e)]. As shown in Fig. 1(a), there is one major mode in the joint distribution, and the other modes are relatively smaller; therefore, the estimated joint distribution of wind and solar power forecast errors could be treated practically as unimodal. Univariate distributions of day-ahead, 4-h-ahead, and 1-h-ahead wind power forecast errors are compared, and the results of pairs 11 and 35 are shown in Figs. 1(d) and 2(d), respectively. The results show that 1-h-ahead wind forecasting performs the best for all the 35 buses. In addition, univariate distributions of day-ahead, 4-h-ahead, and 1-h-ahead solar power forecast errors are also compared, and the results of pairs 11 and 35 are shown in Figs. 1(e) and 2(e), respectively. The results show that 1-h-ahead solar forecasting performs the best for 35 buses, as would be expected.

The examination of these joint distributions provides important information for solar and wind power integration. The peak of each of the distributions is centered around zero, showing that the most likely occurrence is both a small wind power forecast error and a small solar power forecast error. Additionally, the spread of the distribution is always in the cardinal directions (i.e., due north–south or east–west). This means that when there is a large forecast error for either wind or solar, it is rare that there is also a large error for the other technology. This is a fortuitous result, as a diagonal spread of the distribution sloping upward (along a southwest to northeast axis) would indicate that at a time of high system stress (large wind or solar power forecast error), the other forecast would compound the problems experienced.

Pearson's Correlation Coefficient

The Pearson's correlation coefficient between wind and solar power forecast errors was computed for each of the 35 pairs of wind and solar power outputs. The minimum Pearson's correlation coefficients of the day-ahead, 4-h-ahead, and 1-h-ahead forecasts were estimated to be –0.09, –0.15, and –0.05, respectively; the maximum Pearson's correlation coefficients of the day-ahead, 4-h-ahead, and 1-h-ahead forecasts were estimated to be 0.06, 0.002, and 0.02, respectively. In addition, the average Pearson's correlation coefficients over all 35 sites for day-ahead, 4-h-ahead, and 1-h-ahead forecasts were estimated to be −0.01, −0.06, and −0.02, respectively. Therefore, the correlation between wind and solar power forecast errors on a single bus is actually slightly inverse.

Case II: Results and Discussion

The second case analyzed wind and solar forecast correlation considering the aggregated wind power and solar power forecast errors for all 35 paired bus locations. The 35 buses are dispersed

J. Energy Eng.

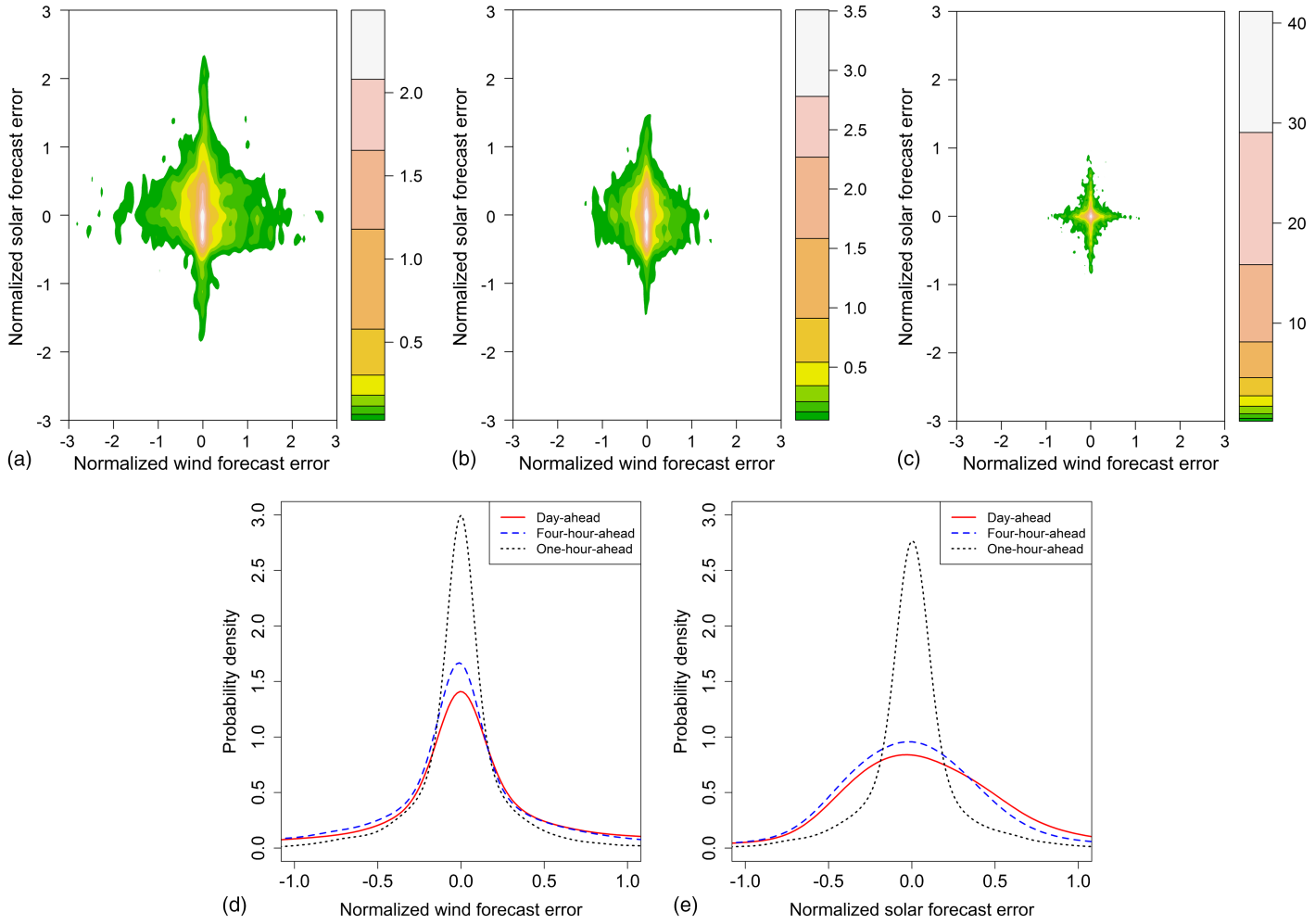

Fig. 2. Probability distributions of wind and solar power forecast errors of pair 35: (a) day-ahead joint distribution; (b) 4-h-ahead joint distribution; (c) 1-h-ahead joint distribution; (d) univariate distribution of wind power forecast error; (e) univariate distribution of solar power forecast error

throughout the entire Western Interconnection, including 4 buses in New Mexico, 4 buses in Arizona, 3 buses in Nevada, 9 buses in California, 2 buses in Montana, 2 buses in Utah, 6 buses in Colorado, 1 bus in Idaho, 1 bus in Wyoming, and 3 buses in the Pacific Northwest (Oregon and Washington). The aggregated wind and solar power forecast errors are normalized by the average actual aggregated wind power and the average aggregated clear sky solar power of the year, respectively. The average actual aggregated wind power and the average aggregated clear sky solar power are 3,910 and 1,209 MW, respectively.

Distribution of Wind and Solar Power Forecast Errors

Figs. 3(a and b) show the distributions of the aggregated wind and solar power forecast errors, respectively. In Fig. 3, the solid, dashed, and dotted lines represent the distribution of day-ahead, 4-h-ahead, and 1-h-ahead forecast errors, respectively. It is observed that the distributions of both wind and solar power forecast errors are unimodal. As shown in Fig. 3(a), (1) the 1-h-ahead wind power forecast error distribution has a larger probability density than the day-ahead and 4-h-ahead forecast error distributions when the forecast error is smaller (approximately –25 to 25% of average actual wind power generation); and (2) the 1-h-ahead forecast error distribution has a smaller probability density than the day-ahead and 4-h-ahead forecast error distributions when the forecast error is larger (approximately less than –25% and more than 25% of average actual wind power generation). Similar results among day-ahead, 4-h-ahead, and 1-h-ahead solar power forecast errors are shown in Fig. 3(b). These observations indicate that 1-h-ahead wind and solar forecasts are generally more accurate than day-ahead and 4-h-ahead forecasts. It is also important to note the positive skewness of the day-ahead solar forecasts, which lead to an abnormally long overforecasting tail. This tendency to overforecast is important in that the system actions taken to correct for underforecasting and overforecasting events are not equal. This overforecasting tendency could lead to a less than optimal number of large thermal units being committed, which need to be corrected through the starting of more expensive, but faster starting, units in the dispatch process.

The joint distribution of wind and solar power forecast errors is illustrated in Fig. 4. It is observed that the joint distributions for day-ahead, 4-h-ahead, and 1-h-ahead power forecast errors are unimodal. As shown in Fig. 4, the area of the contour region in 1-h-ahead forecasting is relatively smaller than that in day-ahead and 4-h-ahead forecasting, which also indicates that 1-h-ahead forecasts are generally more accurate than day-ahead and 4-h-ahead forecasts. It is important to note that there is a far larger spread of the joint distributions when aggregated over all the buses than when viewed from the perspective of an individual bus. This reflects the higher correlations observed between the two technologies' errors when considering larger geographic and time scales.

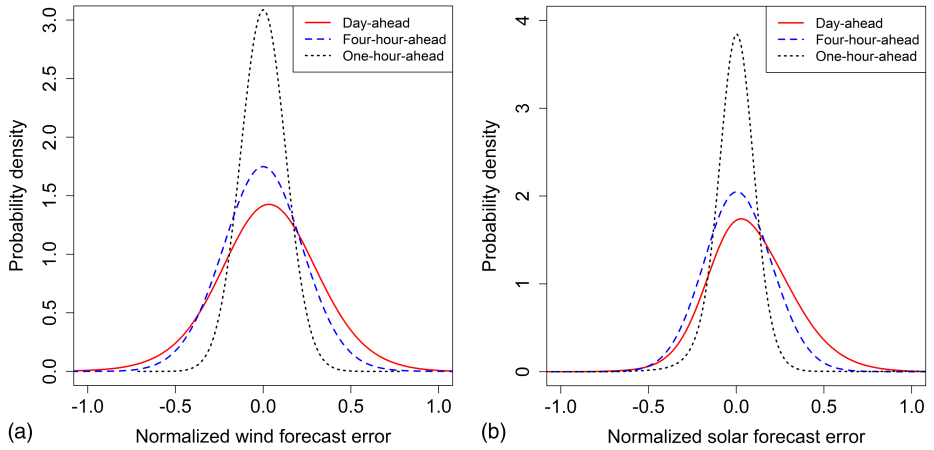

Fig. 3. Univariate distributions of wind and solar power forecast errors (Case II): (a) wind power forecast error distribution; (b) solar power forecast error distribution

The Pearson's correlation coefficients of day-ahead, 4-h-ahead, and 1-h-ahead forecast errors aggregated throughout all 35 sites were estimated to be –0.08, –0.29, and –0.10, respectively. The absolute values of the correlation coefficients were significantly larger than those in Case I, especially for the 4-h-ahead forecasting. An important point is that the aggregated forecast errors are less correlated at the day-ahead timescale, which influences economic operations more than reliability, and more correlated at the short-term timescale, where reliability is more impacted by the forecasts.

Hourly Variation of Wind and Solar Power Forecast Errors

Fig. 5 shows the wind and solar power forecast errors at different times of day, including (1) the day-ahead wind power forecast error at each hour in a day; (2) the day-ahead solar power forecast error at each hour in a day; (3) the 1-h-ahead wind power forecast error at each hour in a day; and (4) the 1-h-ahead solar power forecast error at each hour in a day. It is observed that (1) at each hour in a day, wind power forecast errors are approximately symmetric with respect to the zero error, as shown in Figs. 5(a and c); (2) the day-ahead solar forecasting tends to overforecast between 4:00 and 8:00 p.m. every day, as shown in Fig. 5(b); (3) the 1-h-ahead solar forecasting presents large overforecast errors between 6:00 and 8:00 a.m., as shown in Fig. 5(d); and (4) the 1-h-ahead solar forecasting tends to underforecast between 9:00 and 11:00 a.m., and to overforecast between 3:00 and 5:00 p.m. These observations indicate that, in a certain time period, the pattern of solar power forecast errors is more clear than that of wind power forecast errors. Thus, power system operators hold more information regarding solar forecasting in that time period, which helps the corresponding decision making on flexible reserves and operational reliability.

Figs. 6(a–c) show heat maps of the mean hourly day-ahead, 4-h-ahead, and 1-h-ahead solar power forecast error per month, respectively. For the 4-h-ahead solar forecasting in Fig. 6(b), it is observed that (1) solar forecasting tends to underforecast in the morning (between 7:00 and 11:00 a.m.) during the whole year; (2) solar forecasting tends to overforecast in the afternoon (between 2:00 and 6:00 p.m.); and (3) the period of overforecasting events is relatively longer in summer, which is between 1:00 and 7:00 p.m. It is also interesting to note that some of the largest mean forecast errors occur around noon during the winter months, indicating that the NWP models may be underforecasting the amount of cloud cover during these times. Since most underforecasting and overforecasting events occur in the morning and afternoon, respectively, the power system operators could mitigate the effects of solar

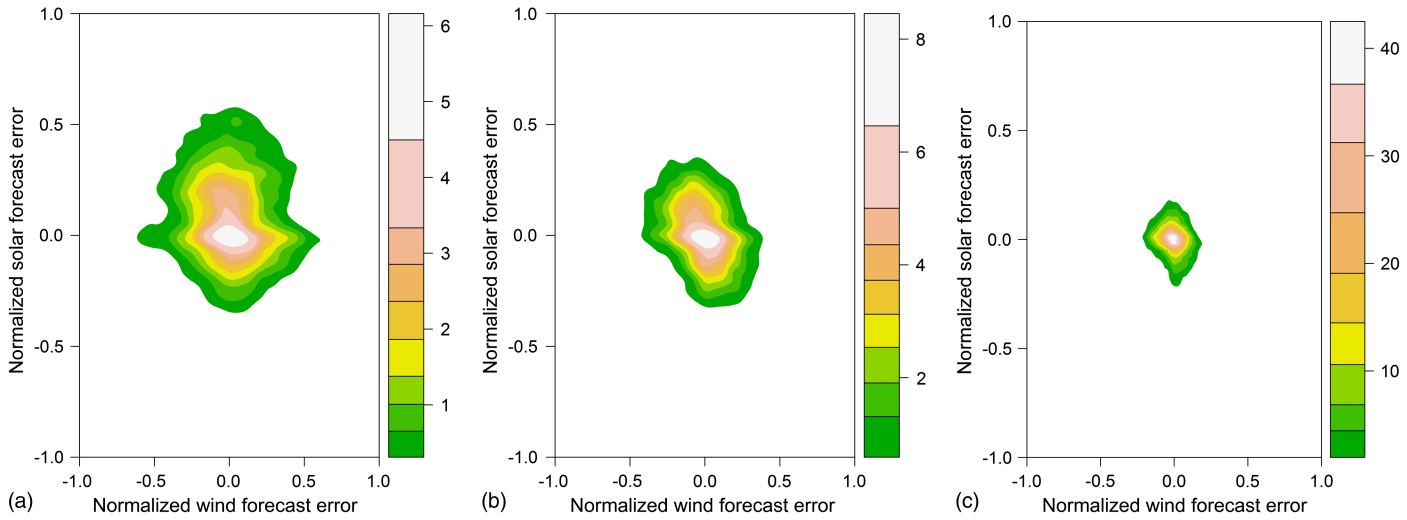

Fig. 4. Joint probability distributions of wind and solar power forecast errors (Case II): (a) day-ahead; (b) 4-h-ahead; (c) 1-h-ahead

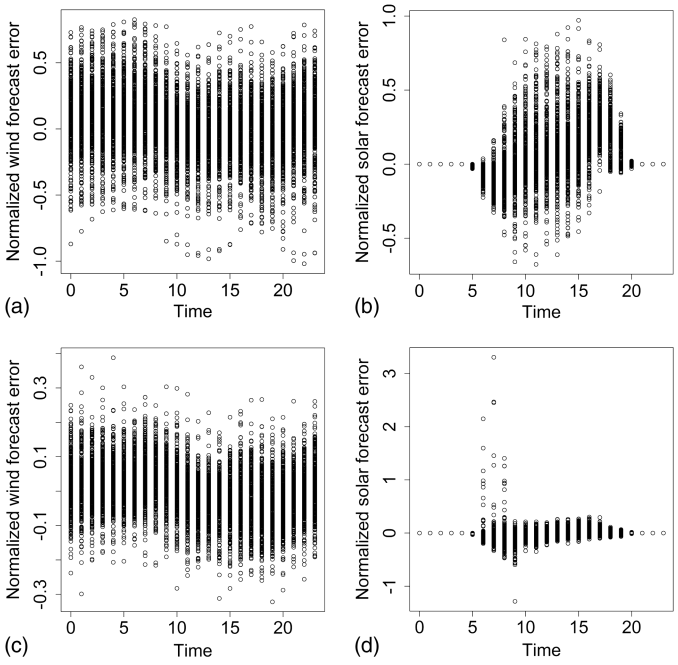

Fig. 5. Day-hour-ahead and 1-h-ahead wind and solar power forecast errors at each hour of the day: (a) day-ahead wind; (b) day-ahead solar; (c) 1-h-ahead wind; (d) 1-h-ahead solar

integration on power system reliability by adopting appropriate statistical corrections of solar power forecast errors. For the day-ahead solar forecasting in Fig. 6(a) and the 1-h-ahead solar forecasting in Fig. 6(c), it is obserseved that there are more overforecasting events than underforecasting events, which indicates that the required amount of "down" reserves to accommodate solar power forecast errors may be relatively less than the amount of "up" reserves. In Fig. 6(a), there are relatively more overforecasting events in winter months than in summer months, which partially accounts for the positive skewness of the day-ahead solar forecasts shown in Fig. 3(b).

Seasonal Variation of Wind and Solar Forecasting

To analyze the seasonal variation of wind and solar power forecasting, we analyzed the correlation between wind and solar power forecast errors for two typical months, January and July, as being indicative of the weather conditions in winter and summer, respectively. The results of forecast errors at different times of day are illustrated in Fig. 7; day-ahead and 1-h-ahead forecast errors are shown in Figs. 7(a–d) and Figs. 7(e–h), respectively.

It is observed in the winter month (January) that (1) wind power forecast errors are approximately symmetric with respect to the zero error, as shown in Figs. 7(a and e); (2) there are significantly more overforecasting events in the day-ahead solar forecasting, as showin in Fig. 7(b); and (3) there are relatively more 1-h-ahead solar underforecasting events between 9:00 a.m. and 12:00 p.m., as shown in Fig. 7(f). In the summer month (July), it is observed [Figs. 7(c and g)] that the day-ahead and 1-h-ahead wind forecasting tends to be overforecasted between 12:00 a.m. and 12:00 p.m., and there are relatively more wind underforecasting events between 3:00 and 8:00 p.m. This can be partially attributed to the tendency for wind to pick up in the evening and die out in the afternoon. Interestingly, the 1-h-ahead solar forecast in Fig. 7(h) tends to be underforecasted in the morning and overforecasted in the afternoon; and most extreme day-ahead solar forecast events happen between 5:00 and 6:00 p.m., which is shown in Fig. 7(d). Similar observations for local plants can be used to guide grid operator policies.

Case III: Results and Discussion

Case III investigated the correlation of the forecast errors arising from the aggregated power output of all 115 wind buses and 424 solar buses in the Western Interconnection of the United States. The aggregated wind and solar power forecast errors are normalized by the average actual aggregated wind power and the average aggregated clear sky solar power of the year, respectively. The average actual aggregated wind power and the average aggregated clear sky solar power are 15,421 and 10,129 MW, respectively.

Distribution of Wind and Solar Power Forecast Errors

Figs. 8(a and b) show the distributions of wind and solar power forecast errors, respectively. Both distributions of wind and solar power forecast errors also present a unimodal characteristic. It is again observed that both the 1-h-ahead wind and solar power forecast error distributions have relatively larger probability densities than day-ahead and 4-h-ahead forecast error distributions when forecast errors are smaller, and vice versa.

The joint distribution of wind and solar power forecast errors is illustrated in Fig. 9. As shown, the joint distributions for the day-ahead, 4-h-ahead, and 1-h-ahead forecast errors are unimodal. The area of the contour region in Fig. 9(c) is relatively smaller than that in Figs. 9(a and b), which also indicates that 1-h-ahead forecasts are generally more accurate than day-ahead and 4-h-ahead forecasts.

Pearson's Correlation Coefficient

Pearson's correlation coefficients of day-ahead, 4-h-ahead, and 1-h-ahead forecasts were estimated to be –0.19, –0.34, and –0.13, respectively. Table 1 lists the correlation coefficients for all three cases. Note that the correlation values of Case I in the table are average values of the 35 pairs. It is observed that (1) wind and solar power forecast errors are weakly inversely correlated; (2) the absolute value of the correlation coefficient between wind and solar power forecast errors increases with the size of the analyzed region, when analyzing the data of a whole year; (3) the absolute value of the correlation coefficient of 4-h-ahead forecast errors is generally greater than that of the day-ahead and 1-h-ahead forecast errors for the data of a whole year; and (4) the correlation between wind and solar power forecast errors in July is relatively stronger than that in January and during the whole year. For the 4-h-ahead forecasting, –0.57 and –0.63 are observed respectively for Cases II and III in July. A large inverse correlation would be preferable because a large positive wind forecasting event would be more likely to be offset by a negative solar forecasting event. The wind and solar power forecast errors are not practically correlated in Case I according to the small Pearson's correlation coefficients. However, the small negative values of correlation coefficients in Case I still indicate that wind and solar forecast errors can offset each other to a certain extent. To further investigate the large inverse correlation coefficient of 4-h-ahead wind and solar power forecast errors, the skewness values of wind and solar power forecast errors are analyzed for Cases II and III, as shown in Table 2. In Table 2, the difference in skewness is equal to the skewness of wind power forecast errors minus the skewness of solar power forecast errors. It is observed that the skewness difference of 4-h-ahead forecasts is significantly smaller than that of day-ahead and 1-h-ahead forecasts, which partially results in a relatively higher correlation between 4-h-ahead wind and solar power forecast errors.

The inverse correlation between wind and solar power forecast errors is an important finding for power systems operations.

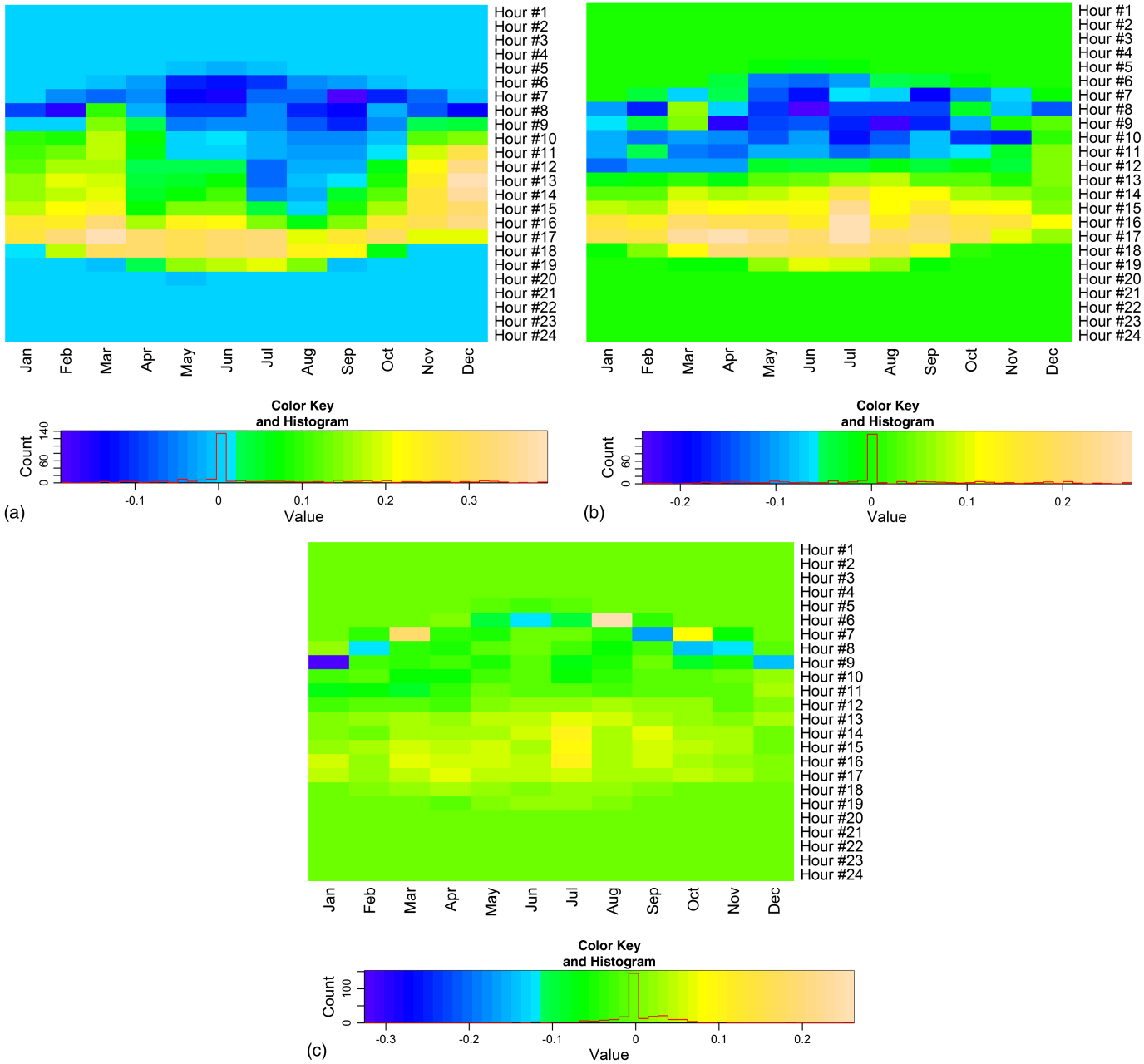

Fig. 6. Heat maps of mean solar power forecast errors for each hour of the day per month of year: (a) day-ahead; (b) 4-h-ahead; (c) 1-h-ahead

It implies that in systems with high penetrations of both wind and solar power, reserves that are held to accommodate the variability of wind or solar power can be at least partially shared. Typically in wind and solar integration studies, solar and wind reserves are calculated separately and then combined (Lew et al. 2013). Assuming that reserves necessary to compensate wind and solar power forecast errors are represented by using hourly time steps and 95th confidence intervals of forecast errors, the total reserves can be calculated by

$$R_t = \sqrt{(\Phi_{95}^{\text{wind}})^2 + (\Phi_{95}^{\text{solar}})^2} \tag{9}$$

where Φ_{95}^{wind} = 95th percentile of wind power forecast errors; Φ_{95}^{solar} = 95th percentile of solar power forecast errors; and R_t = reserve requirement calculated using a traditional method (Lew et al. 2013). Based on the inverse correlation finding, it would be more economically efficient to aggregate the wind and solar power forecast errors and then to calculate the 95th confidence intervals of aggregated forecast errors. The total reserve requirement can be calculated by $R_n = \Phi_{95}^{\text{wind+solar}}$. This is the equivalent of basing the reserve levels on the net load (load–wind–solar) instead of holding separate reserves for load, wind, and solar power forecasting errors independently. With the –0.19 correlation coefficient in July of Case II, the reserves of R_t and R_n are calculated to be 1,698 and 1,492 MW, respectively. The reserve amount is reduced by 12% because of the inverse correlation between wind and solar power forecast errors.

The inverse correlation is also somewhat surprising since many of the same meteorological phenomena that impact solar power output also impact wind power output, though at different timescales. This can be partially attributed to the fact that maximum

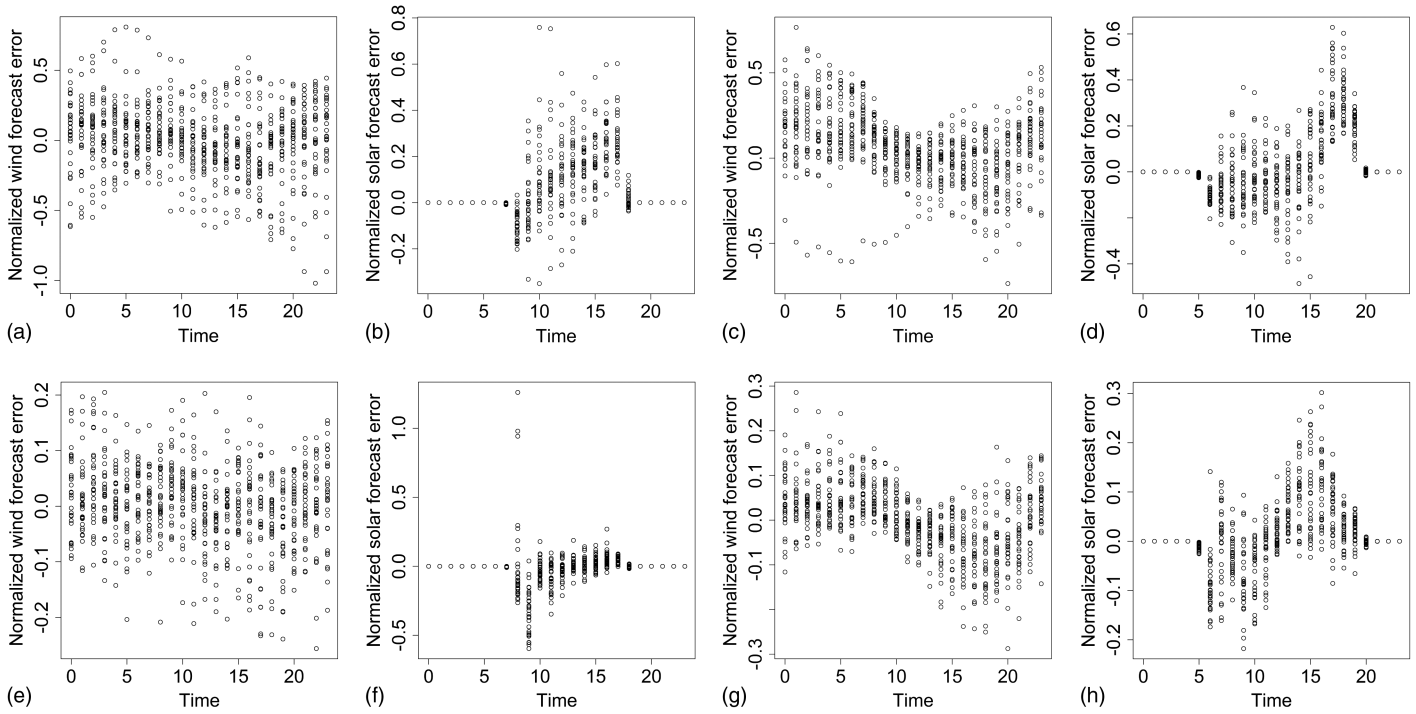

Fig. 7. Day-ahead and 1-h-ahead wind and solar power forecast errors in winter (January) and summer (July): (a) day-ahead wind (January); (b) day-ahead solar (January); (c) day-ahead wind (July); (d) day-ahead solar (July); (e) 1-h-ahead wind (January); (f) 1-h-ahead solar (January); (g) 1-h-ahead wind (July); (h) 1-h-ahead solar (July)

wind and solar power forecast errors are not likely to occur at the same time since the maximum wind and solar power generations are not likely to occur at the same time.

Rényi Entropy and Mutual Information

Tables 3–5 compare the Rényi entropy of wind power forecast error distribution, the Rényi entropy of solar power forecast error distribution, and the mutual information, respectively, for different cases, timescales, and seasons. For the calculation of Rényi entropy, the number of bins for probability estimation and the value of α are set at 100 and 2, respectively. In Case I, Tables 3–5 show the average value of Rényi entropy (or mutual information) for the 35 pairs of wind and solar power outputs. Assuming that a smaller information entropy value is better, it is observed that in Table 3, the uncertainty in day-ahead wind power forecast errors in January and in July increases with the size of the analyzed region; the uncertainty in 4-h-ahead and 1-h-ahead wind power forecast errors during the whole year and in July increases with the size of the analyzed region; in addition, in Table 4, the uncertainty in 1-h-ahead solar power forecast errors during the whole year and in January decreases with the size of the analyzed region for all timescales. When considering different timescales, as shown in Table 3, it is observed that the uncertainty in long-term wind forecast (day-ahead) is more than that in short-term wind forecast timescales 1)-h-ahead and 4-h-ahead) during the whole year; in addition, as shown in Table 4, for solar forecasting, the uncertainty in forecast error of the year and January decreases with the timescale.

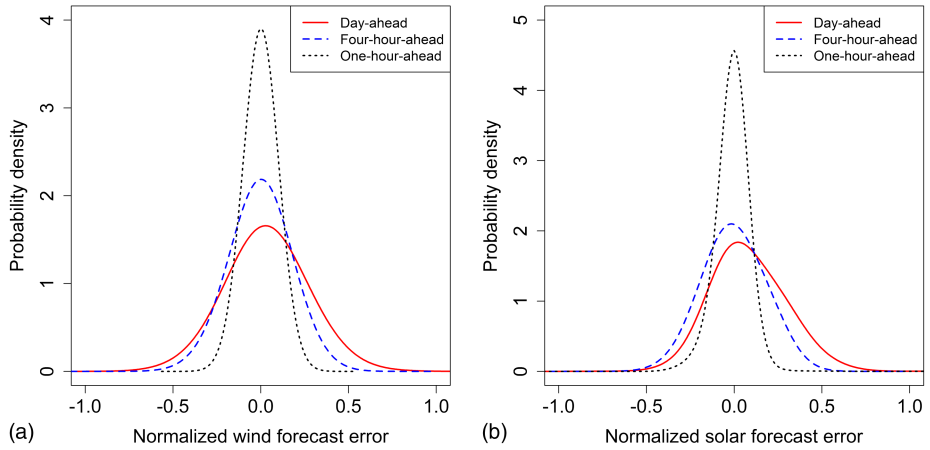

Fig. 8. Univariate distributions of wind and solar power forecast errors (Case III): (a) wind power forecast error distribution; (b) solar power forecast error distribution

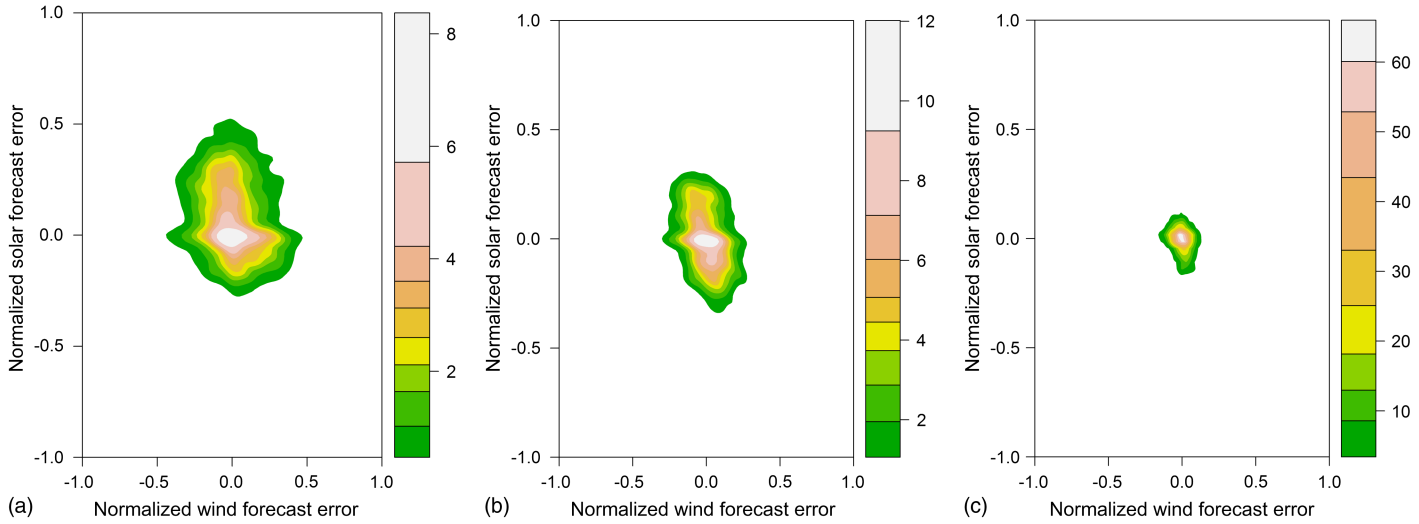

Fig. 9. Joint probability distributions of wind and solar power forecast errors (Case III): (a) day ahead; (b) 4 h ahead; (c) 1 h ahead

Table 1. Pearson's Correlation Coefficients

	Day ahead			4 h ahead			1 h ahead		
Cases	Year	January	July	Year	January	July	Year	January	July
Case I	−0.01	−0.04	−0.01	−0.06	−0.03	−0.15	−0.02	−0.02	−0.04
Case II	−0.08	−0.03	−0.19	−0.29	−0.10	−0.57	−0.10	−0.09	−0.36
Case III	−0.19	−0.21	−0.30	−0.34	−0.18	−0.63	−0.13	−0.06	−0.34

Table 2. Skewness of Wind and Solar Power Forecast Errors for Cases II and III

	Case II			Case III		
Skewness	Day ahead	4 h ahead	1 h ahead	Day ahead	4 h ahead	1 h ahead
Skewness (wind forecast error)	−0.10	0.03	−0.04	0.23	−0.12	−0.11
Skewness (solar forecast error)	0.48	−0.04	8.36	0.44	−0.09	10.00
Skewness difference	−0.58	0.07	−8.40	−0.21	0.03	−10.11

Table 3. Rényi Entropy of Distribution of Wind Power Forecast Errors

	Day ahead			4 h ahead			1 h ahead		
Cases	Year	January	July	Year	January	July	Year	January	July
Case I	4.24	4.45	4.19	3.88	4.28	3.93	3.24	3.80	3.35
Case II	5.37	5.42	5.38	5.29	5.69	5.46	5.16	5.69	5.42
Case III	5.36	5.86	5.43	5.30	5.57	5.80	5.31	5.54	5.69

Table 4. Rényi Entropy of Distribution of Solar Power Forecast Errors

	Day ahead			4 h ahead			1 h ahead		
Cases	Year	January	July	Year	January	July	Year	January	July
Case I	3.11	3.33	3.17	3.03	3.26	3.19	2.13	2.86	2.80
Case II	3.60	3.49	3.69	3.17	3.46	3.86	1.61	2.47	3.70
Case III	3.25	3.61	3.54	3.20	3.65	3.66	1.37	1.59	3.49

Table 5. Mutual Information between Wind and Solar Power Forecast Errors

	Day ahead			4 h ahead			1 h ahead		
Cases	Year	January	July	Year	January	July	Year	January	July
Case I	0.56	0.53	0.52	0.56	0.50	0.53	0.55	0.50	0.51
Case II	0.57	0.55	0.56	0.61	0.49	0.70	0.59	0.50	0.62
Case III	0.60	0.54	0.55	0.62	0.51	0.71	0.60	0.50	0.62

The mutual information between wind and solar power forecast errors is shown in Table 5. Mutual information is a quantitative measurement of how much the wind (or solar) power forecast error tells us about the solar (or wind) power forecast error, basically correlation for n-moments. The more mutual information between wind and solar power forecast errors, the less uncertainty there is in wind (or solar) power forecast errors when solar (or wind) power forecast errors are known. Table 5 shows that wind and solar power forecast errors share relatively more mutual information in Cases II and III with 4-h-ahead forecasts in July.

Conclusions

This paper investigated the correlation between wind and solar power forecast errors. Multiple spatial and temporal scales of forecast errors (day-ahead, 4-h-ahead, and 1-h-ahead) for the Western Interconnection of the United States were analyzed. The correlation between wind and solar power forecast errors was analyzed based on the bus numbers of the wind and solar power plants. The uncertainty in wind and solar forecasting was also quantified through the application of the Rényi entropy metric.

The results showed that both the wind and solar power forecast error distributions were generally unimodal. The results also found that wind and solar power forecast errors exhibited a weakly inverse correlation and that the correlation between wind and solar power forecast errors became stronger when the size of the analyzed region increased. The inverse correlation indicated that in systems with high penetrations of both wind and solar power, reserves that were held to accommodate the variability of wind or solar power could be shared. The reserve case study of Case II in July showed a 12% reduction in the reserve requirements. The seasonal variation analysis showed that the correlation between wind and solar power forecast errors in July was relatively stronger than that in January and during the whole year. The analysis of forecast errors at different times of day investigated the pattern of wind and solar power forecast errors, which could provide important information for grid operators to utilize in planning their operational strategies to integrate wind and solar, such as flexibility reserves.

Wind and solar power forecast errors depend primarily on the accuracy of the underlying forecasting methodologies. For 1-h-ahead or subhourly forecasts, the persistence approach is one of the most popularly used forecasting methods. The analysis results of shorter-term (1-h-ahead) forecasts in this study would be applicable for most power systems with high penetrations of wind and solar energy. The results provide a good indication of the correlation in the natural variability of wind and solar power since with the persistence forecasting method uncertainty is equal to variability. It is important to note that the general shape of the forecasting error distributions is consistent for both wind and solar power forecasts, using both the persistence and NWP forecasting methods. Day-ahead wind and solar power forecast errors could potentially be reduced through advanced forecasting methodologies. With improved wind and solar forecasting, the analysis results of a single bus might change in some geographic regions due to the site-specific weather tendencies. However, because of the large geographic area and diversity of the entire Western Interconnection, the correlation and uncertainty analysis results are not expected to be significantly sensitive to improved forecasting methods.

Future studies will (1) determine the clusters of wind and solar power forecast errors based on important features, e.g., time of a day, weather conditions, and scales of power plants, and (2) quantify the impacts of the correlation between wind and solar power forecast errors when assessing the ability of authority areas to integrate wind and solar power generation, e.g., examine the difference in reserve levels necessary when considering one renewable flexibility reserve product versus separate wind and solar flexibility reserve products.

Acknowledgments

This work was supported by the U.S. DOE under Contract DE-AC36-08-GO28308 with the National Renewable Energy Laboratory.

References

Bessa, R. J., Miranda, V., Botterud, A., and Wang, J. (2011). "'Good' or 'bad' wind power forecasts: A relative concept." *Wind Energy*, 14(5), 625–636.

Bludszuweit, H., Domínguez-Navarro, J. A., and Llombart, A. (2008). "Statistical analysis of wind power forecast error." *IEEE Trans. Power Syst.*, 23(3), 983–991.

Bruninx, K., Delarue, E., and D'haeseleer, W. (2013). "Statistical description of the error on wind power forecasts via a Lévy α-stable distribution." *Technical Rep. EUI Working Paper RSCAS 2013/50*, European Univ. Institute, Florence, Italy.

Castronuovo, E., and Lopes, J. (2004). "On the optimization of the daily operation of a wind-hydro power plant." *IEEE Trans. Power Syst.*, 19(3), 1599–1606.

Chen, C., Duan, S., Cai, T., and Liu, B. (2011). "Online 24-h solar power forecasting based on weather type classification using artificial neural network." *Solar Energy*, 85(11), 2856–2870.

Chow, W. C., et al. (2011). "Intra-hour forecasting with a total sky imager at the UC San Diego solar energy test bed." *Solar Energy*, 85(11), 2881–2893.

Chowdhury, S., Zhang, J., Messac, A., and Castillo, L. (2013). "Optimizing the arrangement and the selection of turbines for wind farms subject to varying wind conditions." *Renew. Energy*, 52, 273–282.

Cover, T. M., and Thomas, J. A. (2006). *Elements of information theory*, Wiley, New York.

Crispim, E. M., Ferreira, P. M., and Ruano, A. E. (2008). "Prediction of the solar radiation evolution using computational intelligence techniques and cloudiness indices." *Int. J. Innov. Comput. Inf. Control*, 4(5), 1121–1133.

Dietrich, K., Latorre, J., Olmos, L., Ramos, A., and Perez-Arriaga, I. (2009). "Stochastic unit commitment considering uncertain wind production in an isolated system." *Proc., 4th Conf. on Energy Economics and Technology*, Dresden Univ. of Technology, Dresden, Germany.

Duong, T., and Hazelton, M. (2003). "Plug-in bandwidth matrices for Bivariate Kernel density estimation." *J. Nonparametric Stat.*, 15(1), 17–30.

Epanechnikov, V. (1969). "Non-parametric estimation of a multivariate probability density." *Theor. Prob. Appl.*, 14(1), 153–158.

Foley, A. M., Leahy, G. L., Marvuglia, A., and Mckeogh, E. J. (2012). "Current methods and advances in forecasting of wind power generation." *Renew. Energy*, 37(1), 1–8.

Hall, P., Lahiri, S. N., and Truong, Y. K. (1995). "On bandwidth choice for density estimation with dependent data." *Ann. Stat.*, 23(6), 2241–2263.

Hammer, A., Heinemann, D., Lorenz, E., and Ckehe, B. L. (1999). "Short-term forecasting of solar radiation: A statistical approach using satellite data." *Solar Energy*, 67(1–3), 139–150.

Hodge, B.-M., et al. (2012b). "Wind power forecasting error distributions: An international comparison." *Proc., 11th Annual Int. Workshop on Large-Scale Integration of Wind Power into Power Systems as well as on Transmission Networks for Offshore Wind Power Plants*, Energynautics GmbH, Germany.

Hodge, B.-M., Ela, E., and Milligan, M. (2012a). "Characterizing and modeling wind power forecast errors from operational systems for use in wind integration planning studies." *Wind Eng.*, 36(5), 509–524.

Hodge, B.-M., Hummon, M., and Orwig, K. (2011). "Solar ramping distributions over multiple timescales and weather patterns." *Proc., 10th Int. Workshop on Large-Scale Integration of Wind Power into Power Systems as well as on Transmission Networks for Offshore Wind Power Plants*, Energynautics GmbH, Germany.

Hodge, B.-M., and Milligan, M. (2011). "Wind power forecasting error distributions over multiple timescales." *Proc., IEEE Power and Energy Society General Meeting*, IEEE, Piscataway, NJ.

Hodge, B.-M., Orwig, K., and Milligan, M. (2012c). "Examining information entropy approaches as wind power forecasting performance metrics." *Proc., 12th Int. Conf. on Probabilistic Methods Applied to Power Systems*, IEEE, Piscataway, NJ.

Hummon, M., Ibanez, E., Brinkman, G., and Lew, D. (2012). "Sub-hour solar data for power system modeling from static spatial variability analysis." *Proc., 2nd Int. Workshop on Integration of Solar Power in Power Systems*, Energynautics GmbH, Germany.

Jeon, J., and Taylor, J. W. (2012). "Using conditional kernel density estimation for wind power density forecasting." *J. Am. Stat. Assoc.*, 107(497), 66–79.

Jones, M., Marron, J., and Sheather, S. (1996). "A brief survey of bandwidth selection for density estimation." *J. Am. Stat. Assoc.*, 91(433), 401–407.

Juban, J., Siebert, N., and Kariniotakis, G. N. (2007). "Probabilistic short-term wind power forecasting for the optimal management of wind generation." *2007 IEEE Lausanne Power Tech*, IEEE, Piscataway, NJ.

J. Energy Eng.

Kariniotakis, G. N., Stavrakakis, G. S., and Nogaret, E. F. (1996). "Wind power forecasting using advanced neural networks models." *IEEE Trans. Energy Convers.*, 11(4), 762–767.

Lew, D., et al. (2010). "The western wind and solar integration study." *Technical Rep. NREL/SR-550-47434*, National Renewable Energy Laboratory, Golden, CO.

Lew, D., et al. (2013). "The western wind and solar integration study phase 2." *Technical Rep. NREL/TP-5500-55888*, National Renewable Energy Laboratory, Golden, CO.

Lindenberg, S. (2008). "20% wind energy by 2030: Increasing wind energy contribution to U.S. electricity supply." *Technical Rep. No. DOE/GO-02008-2567*, U.S. DOE: Energy Efficiency & Renewable Energy, Washington, DC.

Lorenz, E., Heinemann, D., Wickramarathne, H., Beyer, H. G., and Bofinger, S. (2007). "Forecast of ensemble power production by grid-connected PV systems." *Proc., 20th Eur. PV Conf.*, European Commission, Ispra, Italy.

Marquez, R., and Coimbra, C. F. M. (2011). "Forecasting of global and direct solar irradiance using stochastic learning methods, ground experiments and the NWS database." *Solar Energy*, 85(5), 746–756.

Marquez, R., and Coimbra, C. F. M. (2013). "Intra-hour DNI forecasting based on cloud tracking image analysis." *Solar Energy*, 91, 327–336.

Mathiesen, P., and Kleissl, J. (2011). "Evaluation of numerical weather prediction for intra-day solar forecasting in the continental United States." *Solar Energy*, 85(5), 967–977.

Methaprayoon, K., Yingvivatanapong, C., Lee, W. J., and Liao, J. R. (2007). "An integration of ANN wind power estimation into unit commitment considering the forecasting uncertainty." *IEEE Trans. Ind. Appl.*, 43(6), 1441–1448.

Monteiro, C., Bessa, R., Miranda, V., Botterud, A., Wang, J., and Conzelmann, G. (2009). "Wind power forecasting: State-of-the-art 2009." *Technical Rep. ANL/DIS-10-1*, Argonne National Laboratory, Lemont, IL.

Parzen, E. (1962). "On estimation of a probability density function and mode." *Ann. Math. Stat.*, 33(3), 1065–1076.

Perez, R., Moore, K., Wilcox, S., Renné, D., and Zelenka, A. (2007). "Forecasting solar radiation—Preliminary evaluation of an approach based upon the national forecast database." *Solar Energy*, 81(6), 809–812.

Pernick, R., and Wilder, C. (2008). "Utility solar assessment study reaching ten percent solar by 2025." *Technical Rep.*, Clean Edge, and Co-op America Foundation, Washington, DC.

Potter, C. W., Lew, D., McCaa, J., Cheng, S., Eichelberger, S., and Grimit, E. (2008). "Creating the dataset for the western wind and solar integration study (USA)." *Wind Energy*, 32(4), 325–338.

Qin, Z., Li, W., and Xiong, X. (2011). "Estimating wind speed probability distribution using Kernel density method." *Electr. Power Syst. Res.*, 81(12), 2139–2146.

Rodgers, J. L., and Nicewander, W. A. (1988). "Thirteen ways to look at the correlation coefficient." *Am. Stat.*, 42(1), 59–66.

Rosenblatt, M. (1956). "Remarks on some nonparametric estimates of a density function." *Ann. Math. Stat.*, 27(3), 832–837.

Zhang, J., Chowdhury, S., Messac, A., and Castillo, L. (2011). "Multivariate and multimodal wind distribution model based on kernel density estimation." *Proc., 5th Int. Conf. on Energy Sustainability*, ASME, New York.

Zhang, J., Chowdhury, S., Messac, A., and Castillo, L. (2013a). "A multivariate and multimodal wind distribution model." *Renew. Energy*, 51, 436–447.

Zhang, J., Hodge, B.-M., and Florita, A. (2013b). "Investigating the correlation between wind and solar power forecast errors in the western interconnection." *Proc., 7th Int. Conf. on Energy Sustainability*, ASME, New York.

Zhang, Z., Sun, Y., Gao, D., Lin, J., and Cheng, L. (2013c). "A versatile probability distribution model for wind power forecast errors and its application in economic dispatch." *IEEE Trans. Power Syst.*, 28(3), 3114–3125.