Chapter 11
Wind Power and Ramp Forecasting for Grid Integration

Cong Feng and Jie Zhang

11.1 Introduction

Wind energy is a sustainable alternative to the conventional energy in relieving global warming and fuel energy shortage. Notable progress has been made in increasing the wind energy capacity. However, the uncertain and variable characteristics of the wind resource present challenges to wind integration, especially at large penetrations. Accurately forecasting the wind power generation and the extreme wind power changes would greatly help power system operators make better operation schedules, thereby improving the system economic and reliability performance.

Wind forecasting consists of wind speed forecasting and wind power forecasting (Ren et al. 2015). Significant improvements of the wind forecasting have been achieved by developing various forecasting models in the past decades. The wind forecasting models can be classified by different criteria. Based on the algorithm principles, they are generally divided into physical models, statistical models, and hybrid physical and statistical models (Feng et al. 2017a). Based on the forecasting horizons, wind forecasting models are grouped into very short-term models (intra-hour), short-term models (1-h to 6-h-ahead), midterm models (6-h to 1-week-ahead), and long-term models (over 1 week) (Chang 2014).

Different types of statistical models have been applied in the wind forecasting, including conventional time series models, machine learning models, and deep learning models. Conventional time series models include the autoregressive (AR) model (Poggi et al. 2003), the autoregressive moving average (ARMA) model (Erdem and Shi 2011), and the autoregressive integrated moving average (ARIMA)

C. Feng · J. Zhang (✉)
Department of Mechanical Engineering, The University of Texas at Dallas, Richardson, TX, USA
e-mail: jiezhang@utdallas.edu

W. Hu (ed.), *Advanced Wind Turbine Technology*,
https://doi.org/10.1007/978-3-319-78166-2_11

model (Liu et al. 2015). The most popular machine learning algorithms are artificial neural networks (ANNs) (Li and Shi 2010), support vector machine (SVM) (Chen and Yu 2014), random forest (RF) (Feng et al. 2017a), and gradient boosting machine (GBM) (Nagy et al. 2016). Compared with shallow machine learning models, deep learning models are able to capture the hidden invariant structures in the wind speed. The deep belief network algorithm (Wang et al. 2016) and the deep convolutional neural network (Wang et al. 2017) are also employed in the short-term wind forecasting.

Among many wind integration challenges, severe fluctuation incidents with large magnitudes and short durations, so-called ramping events, are a major concern of power system operators. Wind power ramping events (WPREs) are usually caused by complicated physical processes and atmospheric phenomena, such as thunderstorms, wind gusts, cyclones, and low-level jets (Freedman et al. 2008). The research on WPREs can be generally classified into three directions: WPRE detection, WPRE forecasting, and WPRE application. The WPRE detection uses a mathematical algorithm and wind power ramping definitions to extract all the wind power ramps from actual or forecasted wind power data. The WPRE detection method can be directly applied to historical measured wind power data to extract all historical ramping events. Statistical and machine learning methods can then be developed based on the historical ramping events to directly forecast WPREs. The accuracy of WPRE forecasting highly depends on the accuracy of WPRE detection.

This chapter reviews and discusses different types of models for short-term wind forecasting and ramp forecasting, including both individual and ensemble machine learning models and a recently developed optimized swinging door algorithm.

11.2 Wind Forecasting

The most popular short-term wind forecasting models include ANN, SVM, GBM, and RF machine learning models, which provide accurate forecasts with relatively low computational cost. The ensemble of individual machine learning models is another efficient way to improve the wind forecasting accuracy. Both individual and hybrid machine learning models are reviewed and discussed in this section.

11.2.1 Single Machine Learning Algorithm Models

ANN is a popular algorithm in speech recognition, target tracking, signal analysis, and nonlinear regression problems (such as time series forecasting). ANN mimics the structure of the human brain that consists interconnected neurons. Each neuron is a weighted sum of its inputs and is connected to the neurons in the next layer. The ANN architecture contains one input layer, one or more hidden layer(s), and one output layer. The configuration of the ANN model needs to be well designed to avoid

over-fitting issues. ANN can be classified into different types based on different activation functions and learning algorithms. Deep learning is also a configuration of ANN. The mathematical description of the ANN is expressed as:

$$y_i^{(n)} = f\left(\sum_{j=1}^{N} w_{ij}^{(n,n-1)} y_j^{(n-1)} + \theta_i^n\right) \tag{11.1}$$

where i is a neuron of the nth layer, w_{ij} is the weight from the neuron j in the layer $(n-1)$ to the neuron i in layer n, and θ_i^n is the threshold of the neuron i in layer n.

SVM is originally a supervised linear classifier proposed by Vapnik (1995). As one of the most popular classification methods, SVM has been applied in text categorization, image classification, and other recognition tasks. When dealing with linearly inseparable data, nonlinear mapping-based kernel methods, $K(x) : \mathbb{R}^n \longrightarrow \mathbb{R}^{n_h}$, are used to map the nonlinear data into the high-dimensional feature space. Then, a linear hyperplane is found by maximizing the distance between support vectors and the hyperplane. The SVM algorithm can also be applied in regression problems, which is called support vector regression (SVR). The performance of the SVR was reported to be better than other algorithms (e.g., ANN) in the literature. However, the compute and storage requirements increase significantly with the data dimension. The hyperplane function, also called the SVR function, is described as (Feng et al. 2017b):

$$f(x) = \omega^T K(x) + b \tag{11.2}$$

where ω and b are variables solved by minimizing the empirical risk, which is given by:

$$R(f) = \frac{1}{n}\sum_{i=1}^{n} \Theta(y_i, f(x)) \tag{11.3}$$

where $\Theta_\varepsilon(y_i, f)$ is the ε-insensitive loss function, expressed as:

$$\Theta_\varepsilon(y_i, f) = \begin{cases} \|f - y\| - \varepsilon, \text{ if } \|f - y\| \geq \varepsilon \\ 0, \text{ otherwise} \end{cases} \tag{11.4}$$

Then the optimal hyperplane is found by solving the inequality-constrained quadratic optimization problem.

GBM is a highly customizable learning algorithm widely used in the regression and classification fields. A GBM model relies on the combination of "weak learners" to create an accurate learner therefore, is able to generate both deterministic and probabilistic results in the time series forecasting. The combination is achieved by adding the weighted base learner to the previous model iteratively (Kaur et al. 2014).

The principle of GBM is illustrated by the pseudo-code in Algorithm 11.1. In each iteration, the negative gradient of the chosen loss function is calculated and used to estimate the split variables a by Eqs. (11.5) and (11.6). Then the multiplier β is optimized by Eq. (11.7). The weak learner $\beta h(x; a)$ is added to the previous model, where $h(x; a)$ is a learning function.

Algorithm 11.1 Gradient boosting machine (GBM)

1 Initialize $f_0(x)$ to be a constant, $f_0(x) = \arg\min_\rho \sum_{i=1}^{n} \Psi(y_i, \rho)$

2 for $i = 1$ to M do

3 Compute the negative gradient of the loss function:

$$\overline{y_i} = -\left[\frac{\partial \Psi(y_i, F(x_i))}{\partial F(x_i)}\right]_{f(x)} = f_{i-1}(x), i = \{1, 2, \ldots, n\} \tag{11.5}$$

4 Fit a model to $\overline{y}$ by least-squares to get a_t:

$$a_t = \arg\min_{\alpha,\beta} \sum_{i=1}^{n} \left[\overline{y_i} - \beta h(x_i, a)\right]^2 \tag{11.6}$$

5 Calculate β_t by:

$$\beta_t = \arg\min_{\beta} \sum_{i=1}^{n} \Psi(y_i, f_{t-1}(x_i) + \beta h(x_i, a_t)) \tag{11.7}$$

6 Update the model by:

$$f_t(x) = f_{t-1}(x) + \beta_t h(x; a_t) \tag{11.8}$$

7 end for

8 Output $\widehat{f}(x) = f_T(x)$

RF is another supervised ensemble learning method that consists of many single classification and regression trees (CARTs):

$$T = \left\{t\left(X, s_{\Lambda_1}\right), t\left(X, s_{\Lambda_2}\right), \ldots, t\left(X, s_{\Lambda_n}\right)\right\} \tag{11.9}$$

where T is a set of CARTs, t is a single CART, X is the input to the RF model, and s_{Λ_i} is a random vector to extract bootstrap samples which are determined by the bagging algorithm. The robustness of RF models is enhanced by randomness of the bagging algorithm and the best split search process. Since RF is a combination of various different regressions, the model is generally free from over-fitting (Ibarra-Berastegi et al. 2015).

11.2.2 Hybrid Machine Learning Models

Due to the nonlinear and nonstationary characteristics of wind speed, it is challenging to develop a generic model based on a single machine learning algorithm that can produce the best forecasts at different spatial and temporal scales. Hybriding several single machine learning models can make the forecasting more robust. A hybrid example is described in this section as shown in Fig. 11.1. This hybrid model

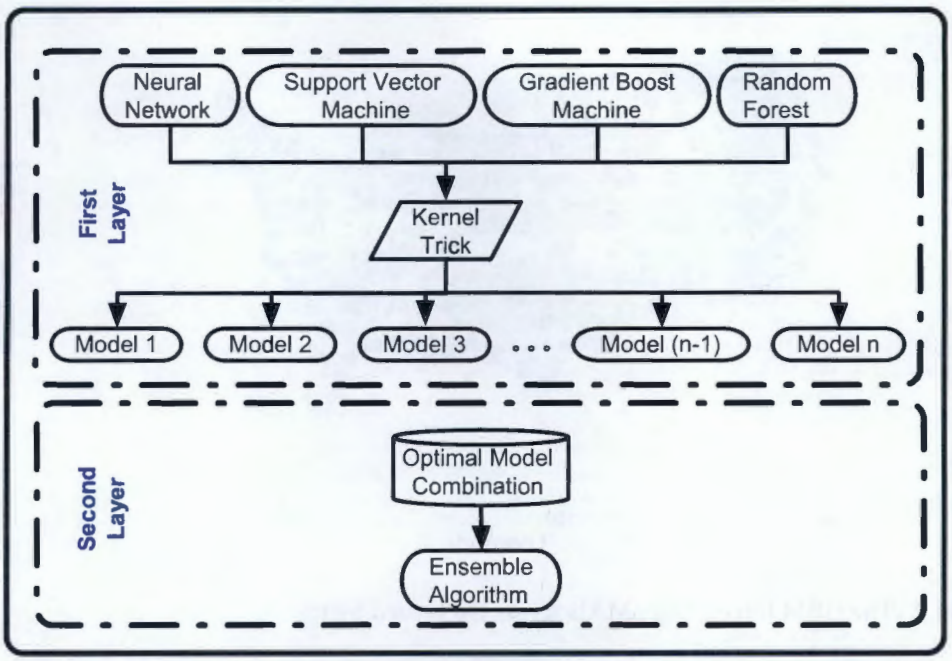

Fig. 11.1 Overall framework of the ensemble forecasting model

has a two-layer forecasting structure (Feng et al. 2017a). The first layer machine learning models are built based on the inputs, such as historical data. These models forecast wind speed or wind power as the output. A blending model is developed in the second layer to combine the forecasts produced by different algorithms from the first layer and to generate both deterministic and probabilistic forecasts. This blending model is expected to integrate the advantages of different algorithms by canceling or smoothing the local forecasting errors. The mathematical description is shown as:

$$y_i = f_i\left(x_1, x_2, \ldots, x_p\right) \tag{11.10}$$

$$\widehat{y} = \Phi\left(y_1, y_2, \ldots, y_m\right) \tag{11.11}$$

where $f_i(*)$ is the ith algorithm and y_i is the wind speed forecasted by $f_i(*)$. $\Phi(*)$ is the second-layer blending algorithm.

11.2.3 Deterministic Results of the Multi-model Forecasting

The performance of the single-algorithm and hybrid machine learning models is evaluated in this section. Two evaluation metrics are utilized to evaluate the forecasting accuracy (Feng et al. 2017a): the normalized mean absolute error

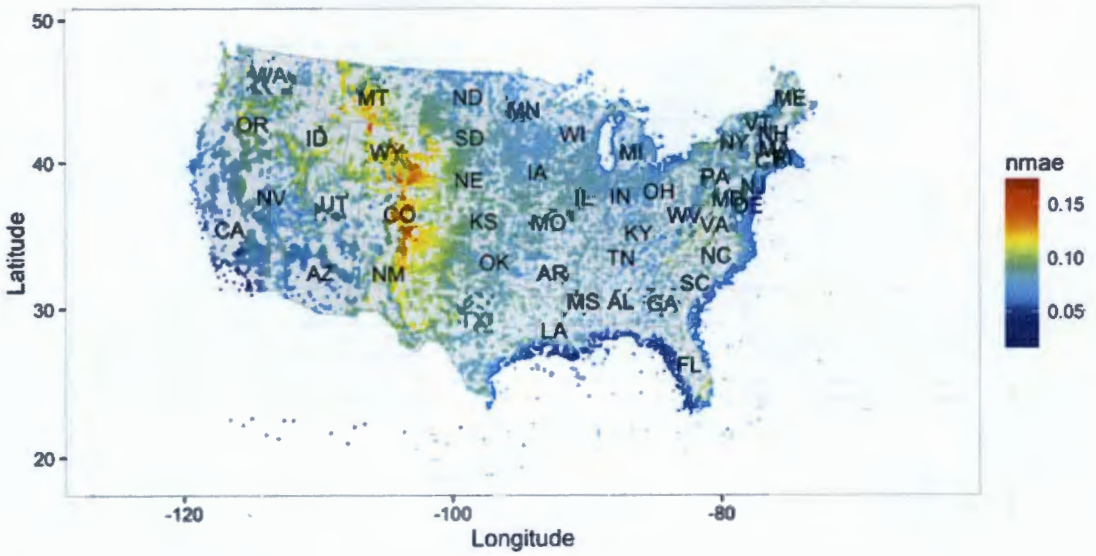

Fig. 11.2 The GBM forecasting nMAE across the United States

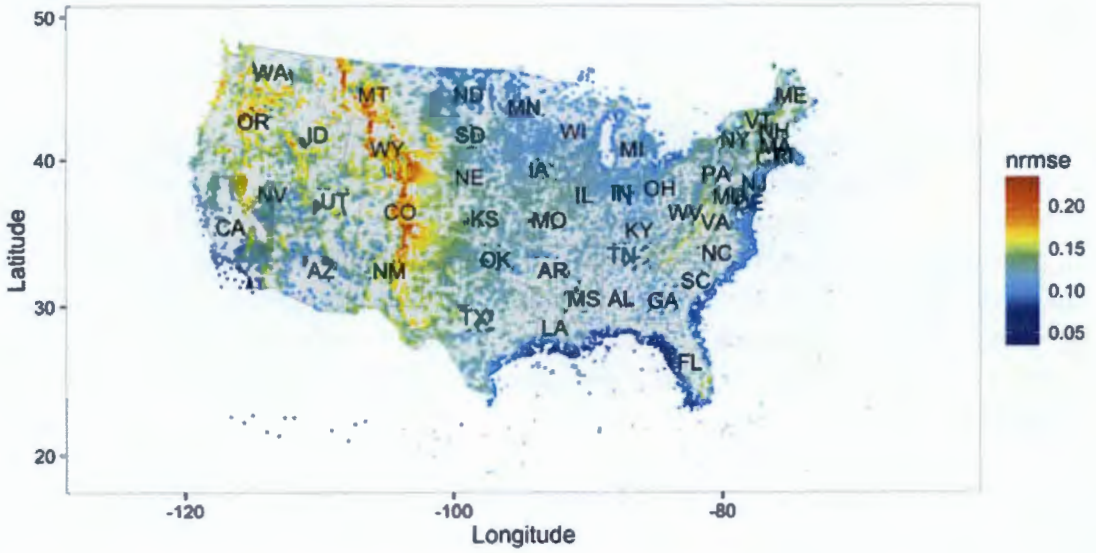

Fig. 11.3 The GBM forecasting nRMSE across the United States

(nMAE) and the normalized root mean square error (nRMSE). The wind forecasting results provided by the GBM models (selected for the robustness and free of preprocessing) of more than 126,000 wind farms over the entire United States are shown in Figs. 11.2 and 11.3. The Wind Integration National Dataset (WIND) Toolkit. Toolkit (Draxl et al. 2015) data was used for the wind forecasting. It is seen from Figs. 11.2 and 11.3 that, the offshore locations, such as the Gulf of Mexico and the East Coast, present relatively high forecasting accuracy; the mountain areas, such as Colorado and New Mexico, have relatively low forecasting accuracy. By comparing Figs. 11.2 and 11.3, regions such as Washington and Oregon present small nMAE but large nRMSE. The variation in the forecasting accuracy across the

Table 11.1 SURFRAD locations

Name	State	Lat.	Long.	Elev. (m)
Bondville (BND)	IL	40.05	−88.37	230
Boulder (TBL)	CO	40.12	−105.24	1689
Desert Rock (DRA)	NV	36.62	−116.02	1007
Fort Peck (FPK)	MT	48.31	−105.24	634
Goodwin Creek (GCM)	MS	34.25	−89.87	98
Penn. State Univ. (PSU)	PA	40.72	−77.93	375
Sioux Falls (SXF)	SD	43.73	−96.62	473

Table 11.2 The nMAE of 1-h-ahead forecasts

Models		BND	TBL	DRA	FPK	GCM	PSU	SXF
SAM	P	4.05	4.27	5.25	4.28	4.13	5.78	3.91
	SVR_li	5.26	5.04	6.65	5.18	5.42	7.13	4.93
	SVR_poly	5.04	4.90	6.17	4.93	5.06	6.86	4.86
	ANN	5.35	5.96	6.23	5.29	5.65	6.90	4.73
	GBM_g	4.95	4.82	6.02	4.80	4.82	6.68	4.78
	GBM_l	5.01	4.80	6.23	4.94	4.96	6.67	4.93
	RF	5.32	4.93	6.51	5.31	5.58	7.51	5.25
MMF	SVR_li	4.32	5.28	5.44	4.45	6.04	6.03	4.05
	SVR_poly	4.20	4.54	5.36	4.31	5.14	5.84	4.01
	GBM	4.26	4.58	5.49	4.37	5.81	6.11	4.19
	RF	4.26	4.60	5.66	4.33	5.34	6.09	4.22

United States is affected by a number of factors. For example, the terrain roughness and the climatic characteristics of the states like Washington and Oregon make wind series more chaotic and less forecastable.

Both the single-algorithm models (SAM) and hybrid multi-model framework (MMF) are applied to the data collected from the Surface Radiation Network (SURFRAD), which includes seven stations (as shown in Table 11.1) with diverse climates. Tables 11.2 and 11.3 list the nMAE and nRMSE, respectively. The multi-model framework includes multiple individual models in the first layer and also several models in the second layer. Different algorithms are tested in both layers, which include SVR with the linear (SVR_li) and polynomial (SVR_poly) kernels, ANN with feed-forward back-propagation learning function and the sigmoid activation function, the GBM models with Gaussian (GBM_g) and Laplacian (GBM_l) loss functions, and the random forest (RF).

As shown in Tables 11.2 and 11.3, none of the SAM models performs better than the persistence method (which assumes that the conditions at the time of the forecast will not change). Without considering the persistence model, no SAM model is always most accurate at all seven locations. Comparing SAM models and MMF models, the MMF with different blending algorithms outperforms the SAM models. The two-layer models have improved the accuracy of the component models by up to 23.8% based on nMAE and 25.6% based on nRMSE. For the blending algorithms,

Table 11.3 The nRMSE of 1-h-ahead forecasts

Models		BND	TBL	DRA	FPK	GCM	PSU	SXF
SAM	P	5.65	6.60	7.36	5.91	5.68	8.27	5.42
	SVR_li	7.76	8.37	9.88	7.92	8.09	9.90	6.95
	SVR_poly	7.05	7.62	8.58	6.81	6.72	9.33	6.51
	ANN	7.27	8.09	8.47	6.94	7.05	9.37	6.30
	GBM_g	6.78	7.77	8.06	6.59	7.01	9.24	6.37
	GBM_l	6.79	7.71	8.86	6.68	6.67	9.42	6.52
	RF	7.36	7.21	9.10	7.35	7.46	10.04	7.11
MMF	SVR_li	6.20	8.96	7.51	6.29	9.21	8.52	5.61
	SVR_poly	5.77	7.22	7.36	6.05	7.08	8.16	5.49
	GBM	5.95	7.29	7.58	6.00	8.23	8.48	5.72
	RF	5.85	7.52	7.63	5.92	7.53	8.46	5.74

the models with nonlinear blending algorithms have better performance than the models with linear blending algorithms. This shows that the forecasts produced from the first-layer models exhibit a nonlinear relationship with the actual wind speed. The model with the polynomial-kernel SVM algorithm is the most accurate model among all the MMF models.

11.2.4 Probabilistic Results of the Multi-model Forecasting

In addition to deterministic forecasts, the multi-model methodology can also produce probabilistic forecasts. Figure 11.4 provides an example of the deterministic forecasts along with the confidence intervals in the form of fan chart, at BND. The confidence bands are calculated based on the component models. The colors of the intervals fade with the increasing confidence level, ranging from 10% to 90% in a 10% increment. The intervals are symmetric around the deterministic forecasting curves with a changing width. When the wind speed fluctuates within a small range, the confidence bands are narrow, as shown by hours 0–10. When there is a significant ramp, the uncertainty of the forecasts is increased and the bands tend to be broader, as shown by hours 150–170. This further proves the necessity of probabilistic forecasting.

11.2.5 The Value of Wind Forecasting

The accurate deterministic and probabilistic wind forecasting could benefit power system operators, energy traders, and wind plant owners by (i) assisting utilities to reduce the backup, therefore achieving savings; (ii) minimizing the production costs by optimizing the slow- and quick-start unit capacity; (iii) providing schedules for

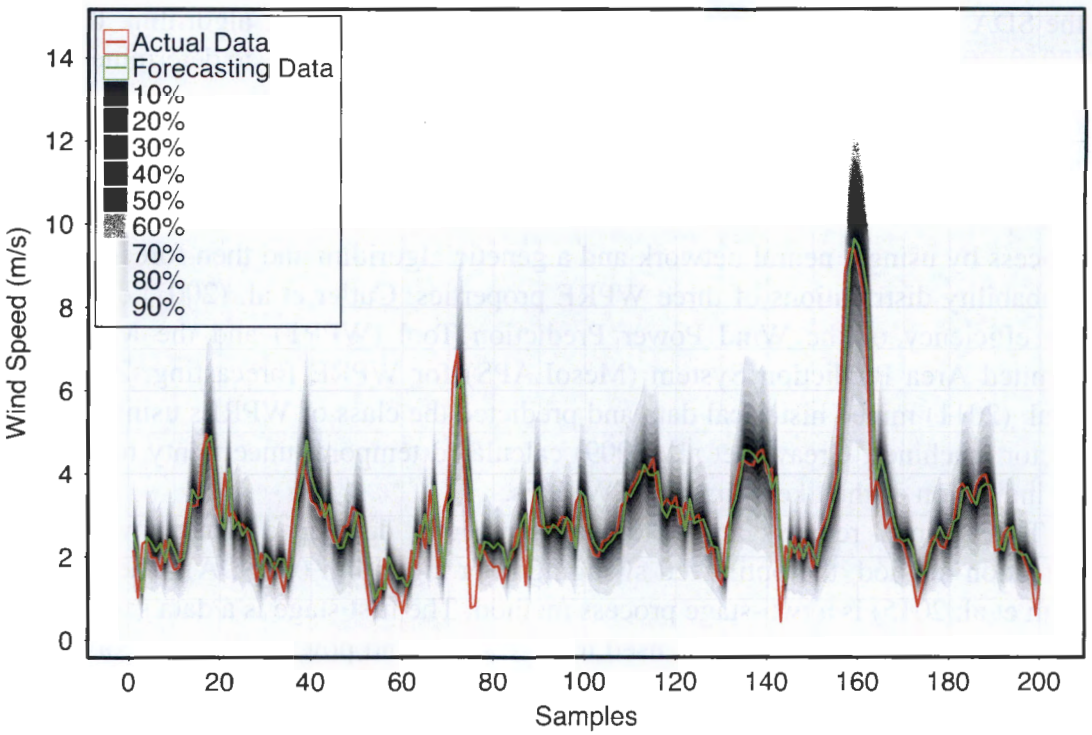

Fig. 11.4 Deterministic forecasting from the MMF_FS with confidence intervals at BND

wind turbine starting up/shutting down in the response to fluctuations; (iv) helping wind farm operators, especially offshore wind farm operators, to better schedule wind turbine maintenance; and (v) reducing curtailment of the wind generation. Overall, the improved wind forecasts could be helpful in reducing the operation costs and increasing the system reliability. The forecasts can also be used to determine the charge and discharge schedule of energy storage in a micro-grid system with distributed wind generators and energy storage.

11.3 Wind Power Ramp Event Detection

Wind power ramps significantly affect the regulation of traditional generators for better managing and dispatching the wind power. Therefore, better detecting and forecasting ramp events are very helpful for power system operators to make operational decisions. Regarding wind power ramp detection, Sevlian and Rajagopal (2012, 2013) proposed an optimal detection technique to identify all WPREs by defining a family of scoring functions associated with any ramping rules and using recursive dynamic programming. Zhang et al. (2014) adopted the swinging door algorithm (SDA) to extract ramp events from actual and forecasted wind power time series. Cui et al. (2016) developed an optimized swinging door algorithm (OpSDA) to improve ramp detection performance, by segregating wind power time series with

the SDA and merging all ramps with a dynamic programming algorithm. Kamath (2010, 2011) used feature selection techniques from data mining to determine ramps in wind power generation.

A number of statistical and machine learning methods have been developed in the literature to forecast wind power ramps at multiple forecasting horizons. For example, Cui et al. (2016, 2017) modeled the wind power generation as a stochastic process by using a neural network and a genetic algorithm and then forecasted the probability distributions of three WPRE properties. Cutler et al. (2007) compared the efficiency of the Wind Power Prediction Tool (WPPT) and the Mesoscale Limited Area Prediction System (MesoLAPS) for WPRE forecasting. Zareipour et al. (2011) mined historical data and predicted the class of WPREs using support vector machines. Greaves et al. (2009) calculated temporal uncertainty to provide an indication of the likely timing of WPREs.

This chapter reviews and discusses a recently developed wind power ramp detection method, the optimized swinging door algorithm (OpSDA). The OpSDA (Cui et al. 2015) is a two-stage process method. The first stage is a data segregation process based on SDA. SDA is used to segregate wind power signals according to the user-specified definition of a ramp. The second stage is an optimization process based on a dynamic programming algorithm. Dynamic programming is used to merge adjacent segments that are segregated with the same ramp changing direction in the first stage.

11.3.1 Swinging Door Algorithm (SDA)

The SDA algorithm (Bristol 1990; Barr 1994) is based on the concept of a "swinging door" with a "hinge" or "pivot point" whenever the next point in the time series causes any intermediate point to fall outside the area partitioned by the up and down segment bounds. The segment bounds are defined by the door width, $\pm\varepsilon$, which is the only tunable parameter in the SDA. More detailed descriptions of the SDA can be found in Florita et al. (2013) and Makarov et al. (2009). After segregating the wind power signal by SDA, wind power ramping events (WPREs) are extracted according to the user-specified definition of a significant ramp.

11.3.2 Optimized Swinging Door Algorithm (OpSDA)

The objective of the optimization in the SDA is to minimize the number of individual ramps whereas still approximating the wind power signal as a ramp. Therefore, adjacent segments that have the same slope (e.g., up-ramps) can be merged into one segment. Toward this end, an optimization process is applied to the original segments (from the SDA) using a dynamic programming algorithm. Dynamic programming is a method for solving a complex problem by breaking it down into

a collection of simpler subproblems. Every subinterval (subproblem) of the ramp detection problem complies with the same ramp rules. First, the subintervals that satisfy the ramp rules are rewarded by a score function; otherwise, their score is set to zero. Next, the current subinterval is retested as above after being combined with the next subinterval. This process is performed recursively to the end of the dataset. Finally, the significant ramp with the maximum score is extracted. More detailed formulations of the dynamic programming algorithm used in this work are shown in Eqs. (11.12), (11.13), and (11.14).

In this chapter, an increasing length score function, S, is designed based on the length of the interval segregated by the SDA. The optimization problem seeks to maximize the length score function, which corresponds to a ramp event. Given a time interval, (i, j), of all discrete time points and an objective function, J, of the dynamic programming algorithm, a WPRE is detected by maximizing the objective function:

$$J(i, j) = \max_{i<k\leq j} [S(i, k) + J(k, j)], \quad i < j \tag{11.12}$$

subject to:

$$S(i, j) > S(i, k) + S(k + 1, j), \quad \forall i < k < j \tag{11.13}$$

$$S(i, j) = (j - i)^2 \times R(i, j) \tag{11.14}$$

where J(i, j) can be computed as the maximum over (j−i) subproblems. The term of S(i, k) is a positive score value corresponding to the interval, (i, k), which conforms to a super-additivity property in Eq. (11.13). There is a family of score functions satisfying Eq. (11.13), and the score function presented in (Sevlian and Rajagopal 2013) is adopted in this research, expressed as Eq. (11.14). R(i, j) represents a ramp within the time interval (i, j). Significant wind power ramps can be defined based on the power change magnitude, direction, and duration. Three definitions proposed in (Zhang et al. 2014) are investigated in this research:

(i) Significant ramp definition 1—the change in wind power output is greater than 20% of the installed wind capacity without constraining the ramping duration.
(ii) Significant ramp definition 2—the change in wind power output is greater than 20% of the installed wind capacity within a time span of 4 h or less.
(iii) Significant ramp definition 3—a significant up-ramp is defined as the change in wind power output greater than 20% of wind capacity within a time span of 4 h or less; a significant down-ramp is defined as the change in wind power output greater than 15% of the installed wind power capacity within a time span of 4 h or less.

If R(i, j) conforms to the threshold of ramp definitions, R(i, j) is 1; otherwise, R(i, j) is 0. Since the process of detecting down-ramps is the opposite process of

detecting up-ramps, note that up-ramp detection is taken as an example to illustrate the specific detecting process.

When optimizing ramps, one of the more interesting findings was the presence of small ramps (non-WPREs), which are termed "bumps" in this paper and set as B(i, j) in the formulations below. The key characteristic of a bump is the changing direction (e.g., a down-bump between two up-ramps or an up-bump between two down-ramps), which makes the iteration of the dynamic programming to break abruptly due to the strict super-additivity property in Eq. (11.13). When a bump occurs, it breaks one integrated WPRE into two discrete ramps, which affects the performance of WPRE detection. To address this issue, the dynamic programming process is improved so that it can also merge ramps and bumps with different changing directions. If B(i, j) conforms to the threshold of bump definitions, B(i, j) is assigned to be 1; otherwise, B(i, j) is assigned to be 0. During the recursion, bumps are also considered and merged into the WPRE.

11.3.3 Experimental Results

In this section, the OpSDA is applied to two case studies. We present various statistics to analyze the detected WPREs and parameterize the WPRE process. The total wind power generation is taken from a balancing area in the northwestern region of the United States. The dataset contains 7,884,012 samples sampled every 4 s spanning from October 1, 2012, to September 30, 2013. In this case, we use the maximum power output, 123 MW, as the base benchmark capacity. The 4-s dataset is averaged to obtain wind power data at different timescales: 1-min, 5-min, 15-min, 30-min, 1-h, and 2-h. A total of 2,089 ramps within 1-min timescale (1,941 ramps within 5-min timescale, 1,701 ramps within 15-min timescale, 1,340 ramps within 30-min timescale, 1,009 ramps within 1-h timescale, and 705 ramps within 2-h timescale) are detected and utilized to generate the probability density distributions. Figure 11.5 shows the ramp feature statistics and seasonal ramp counts of each timescale over the course of a whole year.

Figure 11.5a indicates that along with the increasing timescale (from 1-min to 2-h), for ramp durations, the peak duration value and probability density rise from 50 min with 0.03–400 min with 0.17. For the distribution of ramp change rate in Fig. 11.5b, the peak change rate value decreases from 0.004 p.u./min to 0.001 p.u./min, whereas the corresponding probability density rises from 80 to 820. For the distribution of ramp magnitude in Fig. 11.5c, the peak magnitude value rises from 0.21 p.u. to 0.33 p.u., whereas the corresponding probability density decreases from 9 to 2. Figure 11.5d illustrates that the seasonal ramp counts also decrease along with the increasing timescale in each season. There are relatively fewer ramp events occurring in winter and spring, whereas there are relatively more ramp events occurring in summer and fall. This can be partially attributed to the higher wind generation in summer and fall as shown in Fig. 11.5e. It is seen from Figs. 11.5d and 11.5e that seasonal ramp counts increase along with the increasing wind generation.

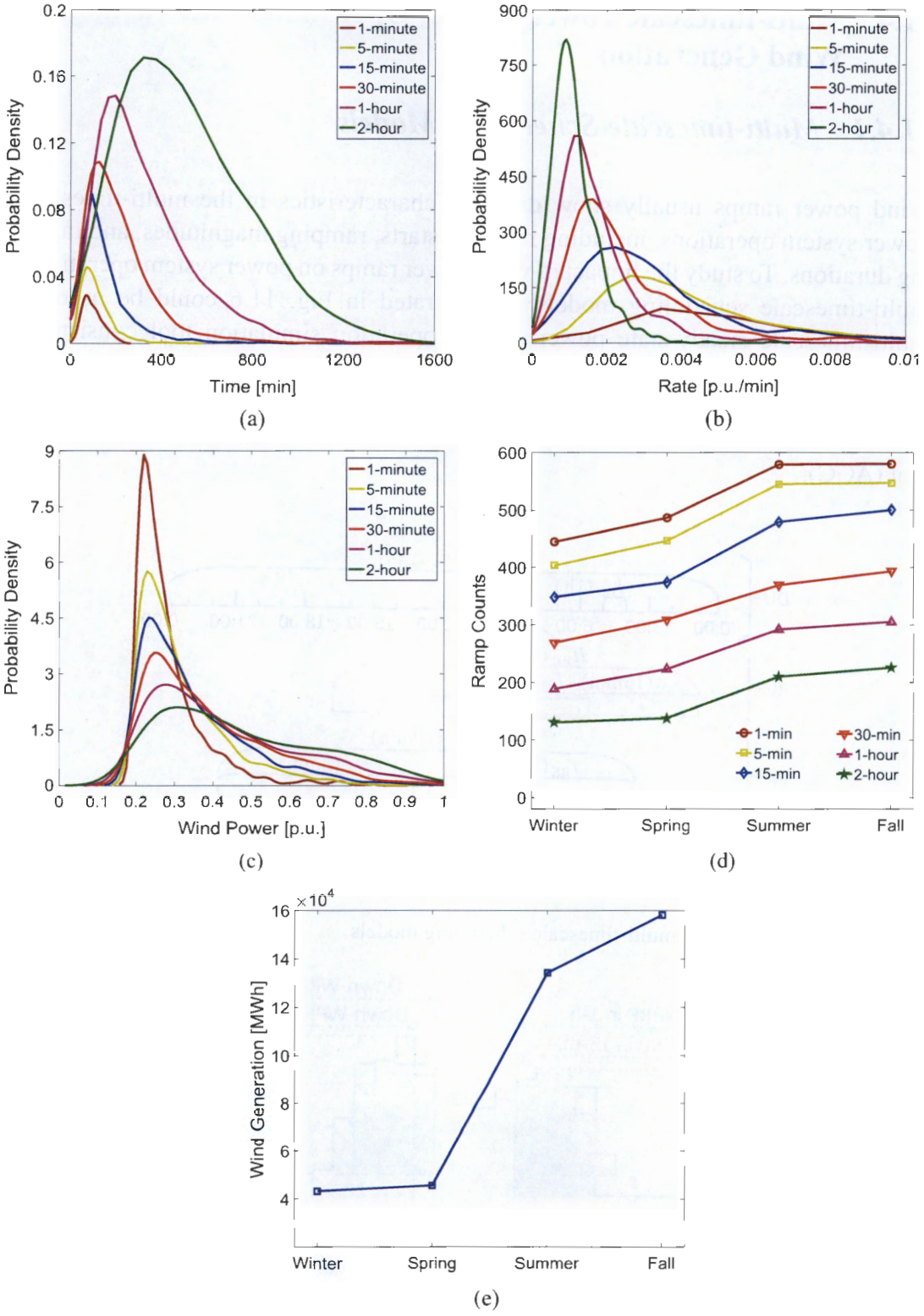

Fig. 11.5 Probability density distributions of ramp features of six timescales (1-min, 5-min, 15-min, 30-min, 1-h, and 2-h) and seasonal ramp counts over a whole year for Case II. (**a**) Ramp duration. (**b**) Ramp change rate. (**c**) Ramp magnitude. (**d**) Seasonal ramp counts. (**e**) Seasonal wind generation

11.4 Multi-timescale Power System Operations with Variable Wind Generation

11.4.1 Multi-timescale Scheduling Models

Wind power ramps usually show different characteristics in the multi-timescale power system operations, including ramping starts, ramping magnitudes, and ramping durations. To study the impact of wind power ramps on power system operations, multi-timescale scheduling models as illustrated in Fig. 11.6 could be used. A multi-timescale steady-state power system operation simulation tool consists of different sub-models, such as day-ahead security-constrained unit commitment (DASCUC), real-time security-constrained unit commitment (RTSCUC), real-time security-constrained economic dispatch (RTSCED), and automatic generation control (AGC).

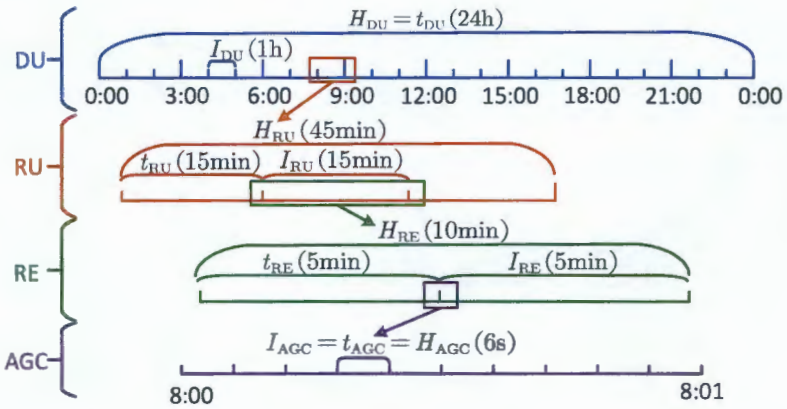

Fig. 11.6 Timeframes of multi-timescale scheduling models

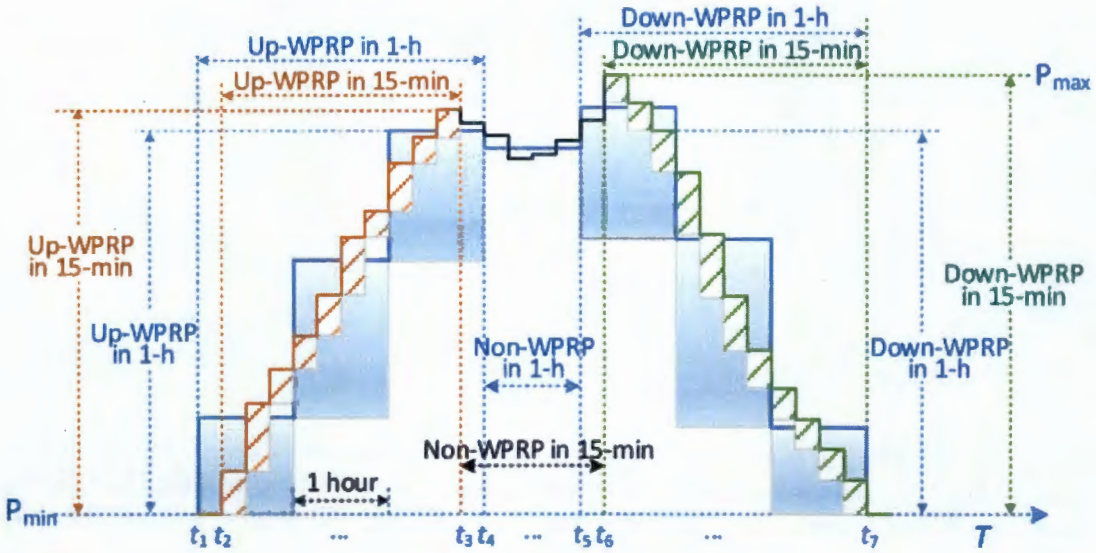

Fig. 11.7 WPRP performance in the multi-timescale operations using the stairstep graph

Figure 11.7 exhibits an example of multi-timescale wind power ramping product (WPRP) detection. The rectangle represents the ramping product that can be provided by wind power at the current time. The first three blue rectangles consist of one up-WPRP with the ramping start time t1 and ramping end time t4 in the 1-h timescale. The up-WPRP in the 15-minute-timescale model starts at time t2 (=t1 + 15 min) and terminates at time t3 (=t4-15 min). Moreover, the ramping capacity in the 15-min-timescale model is much less than that in the 1-h-timescale model, according to the areas of blue and yellow rectangles. The same phenomenon can also be found in the down-WPRP (time t5 t7 and time t6 t7). Under this circumstance, it is essential to characterize and consider WPRP features in a multi-timescale fashion.

11.5 Conclusions

In this chapter, several widely used models for the short-term wind forecasting and ramp forecasting were reviewed and discussed. The 1-h-ahead wind power forecasts at over 126,000 wind sites in the United States were generated using a gradient boosting machine model. We also found that the ensemble machine learning models have improved the wind forecasting accuracy, compared with the single-algorithm models. A recently developed wind power ramp detection method was introduced in this chapter. The results showed that the OpSDA successfully identified wind power ramps and performed significantly better than the SDA. The accurate wind power forecasts and ramp detection could benefit power system operators, energy traders, and wind plant owners.

References

Barr DC (1994) The use of a data historian to extend plant life. In Proceedings of international conference life management of power plants, Heriot-Watt University, Dec. 12–14. Edinburgh, UK

Bristol EH (1990) Swinging door trending: adaptive trend recording? In Proceedings of ISA national conference, pp 749–756

Chang WY (2014) A literature review of wind forecasting methods. *Journal of Power and Energy*. Engineering 2(04):161

Chen K, Yu J (2014) Short-term wind speed prediction using an unscented Kalman filter based state-space support vector regression approach. Appl Energy 113:690–705

Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20(3):273–297

Cui M, Zhang J, Florita AR, Hodge BM, Ke D, Sun Y (2015) An optimized swinging door algorithm for wind power ramp event detection. IEEE Power Energy Society General Meeting, Denver

Cui M, Zhang J, Florita AR, Hodge BM, Ke D, Sun Y (2016) An optimized swinging door algorithm for identifying wind ramping events. IEEE Trans Sustainable Energy 7(1):150–162

Cui MJ, Zhang J, Feng C, Florita AR, Sun YZ, Hodge BM (2017) (2017). Characterizing and analyzing ramping events in wind power, solar power, load, and netload. Renew Energy 111:227–244

Cutler N, Kay M, Jacka K, Nielsen TS (2007) Detecting categorizing and forecasting large ramps in wind farm power output using meteorological observations and WPPT. Wind Energy 10(5):453–470

Draxl C, Clifton A, Hodge B-M, McCaa J (2015) The WIND integration national dataset (WIND) toolkit. Appl Energy 151:355–366

Erdem E, Shi J (2011) ARMA based approaches for forecasting the tuple of wind speed and direction. Appl Energy 88(4):1405–1414

Feng C, Cui M, Hodge BM, Zhang J (2017a) A data-driven multi-model methodology with deep feature selection for short-term wind forecasting. Appl Energy 190:1245–1257

Feng C, Cui MJ, Lee M, Zhang J, Hodge BM, Lu SY, Hamann HF (2017b) Short-term global horizontal irradiance forecasting based on sky imaging and pattern recognition. IEEE Power Energy Society General Meeting, Chicago

Florita A, Hodge BM, Orwig K Identifying wind and solar ramping events. in Proceedings of IEEE 5th Green Technologies Conference, Denver, CO, 4–5, 2013

Freedman J, Markus M, Penc R (2008, Jan. 28) Analysis of West Texas Wind Plant Ramp-Up and Ramp-Down Events [Online]. Available: http://interchange.puc.state.tx.us/WebApp/Interchange/Docu ments/33672_1014_580034.pdf

Greaves B, Collins J, Parkes J, Tindal A (2009) Temporal forecast uncertainty for ramp events. Wind Eng 33(4):309–320

Ibarra-Berastegi G, Saénz J, Esnaola G, Ezcurra A, Ulazia A (2015) Short-term forecasting of the wave energy flux: analogues, random forests, and physics-based models. Ocean Eng 104:530–539

Kamath C (2010) Understanding wind ramp events through analysis of historical data. Proceedings transmission and distribution conference exposition. New Orleans, LA

Kamath C (2011) Associating weather conditions with ramp events in wind power generation. Proceedings transmission and distribution conference exposition. Phoenix, AZ

Kaur A, Pedro HT, Coimbra CF (2014) Ensemble re-forecasting methods for enhanced power load prediction. Energy Convers Manag 80:582–590

Li G, Shi J (2010) On comparing three artificial neural networks for wind speed forecasting. Appl Energy 87(7):2313–2320

Liu H, Tian HQ, Li YF (2015) An EMD-recursive ARIMA method to predict wind speed for railway strong wind warning system. J Wind Eng Ind Aerodyn 141:27–38

Makarov YV, Loutan C, Ma J, Mello P (2009) Operational impacts of wind generation on California power systems. IEEE Trans Power Syst 24(2):1039–1050

Nagy GI, Barta G, Kazi S, Borbély G, Simon G (2016) GEFCom2014: probabilistic solar and wind power forecasting using a generalized additive tree ensemble approach. Int J Forecast 32(3):1087–1093

Poggi P, Muselli M, Notton G, Cristofari C, Louche A (2003) Forecasting and simulating wind speed in Corsica by using an autoregressive model. Energy Convers Manag 44(20):3177–3196

Ren Y, Suganthan PN, Srikanth N (2015) Ensemble methods for wind and solar power forecasting—a state-of-the-art review. Renew Sust Energ Rev 50:82–91

Sevlian R, Rajagopal R (2012) Wind power ramps: detection and statistics. IEEE Power Energy Society General Meeting, San Diego

Sevlian R, Rajagopal R (2013) Detection and statistics of wind power ramps. IEEE Trans Power Syst 28(4):3610–3620

Wang HZ, Wang GB, Li GQ, Peng JC, Liu YT (2016) Deep belief network based deterministic and probabilistic wind speed forecasting approach. Appl Energy 182:80–93

Wang HZ, Li GQ, Wang GB, Peng JC, Jiang H, Liu YT (2017) Deep learning based ensemble approach for probabilistic wind power forecasting. Appl Energy 188:56–70

Zareipour H, Huang D, Rosehart W (2011) Wind power ramp events classification and forecasting: a data mining approach. IEEE Power Energy Society Generation Meeting, San Diego

Zhang J, Florita AR, Hodge BM, Freedman J (2014) Ramp forecasting performance from improved short-term wind power forecasting. In: ASME International Design Engineering Technical Conference & Computers and information in engineering conference (IDETC/CIE 2014), Buffalo