NREL
NATIONAL RENEWABLE ENERGY LABORATORY

Ramp Forecasting Performance from Improved Short-Term Wind Power Forecasting

Preprint

J. Zhang, A. Florita, and B.-M. Hodge
National Renewable Energy Laboratory

J. Freedman
AWS Truepower

To be presented at the ASME 2014 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference (IDETC/CIE 2014)
Buffalo, New York
August 17–20, 2014

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Conference Paper
NREL/CP-5D00-61730
May 2014

Contract No. DE-AC36-08GO28308

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Available electronically at http://www.osti.gov/scitech

Available for a processing fee to U.S. Department of Energy and its contractors, in paper, from:

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831-0062
phone: 865.576.8401
fax: 865.576.5728
email: mailto:reports@adonis.osti.gov

Available for sale to the public, in paper, from:

U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
phone: 800.553.6847
fax: 703.605.6900
email: orders@ntis.fedworld.gov
online ordering: http://www.ntis.gov/help/ordermethods.aspx

Cover Photos: (left to right) photo by Pat Corkery, NREL 16416, photo from SunEdison, NREL 17423, photo by Pat Corkery, NREL 16560, photo by Dennis Schroeder, NREL 17613, photo by Dean Armstrong, NREL 17436, photo by Pat Corkery, NREL 17721.

Printed on paper containing at least 50% wastepaper, including 10% post consumer waste.

Proceedings of the ASME 2014 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference
IDETC/CIE 2014
August 17-20, 2014, Buffalo, New York, USA

DETC2014-34775

RAMP FORECASTING PERFORMANCE FROM IMPROVED SHORT-TERM WIND POWER FORECASTING

Jie Zhang[*]
National Renewable Energy Laboratory
Golden, Colorado 80401
Email: jie.zhang@nrel.gov

Anthony Florita[#]
National Renewable Energy Laboratory
Golden, Colorado 80401
Email: anthony.florita@nrel.gov

Bri-Mathias Hodge[†]
National Renewable Energy Laboratory
Golden, Colorado 80401
Email: bri.mathias.hodge@nrel.gov

Jeffrey Freedman[‡]
AWS Truepower
Albany, New York 12205
Email: jfreedman@awstruepower.com

ABSTRACT

The variable and uncertain nature of wind generation presents a new concern to power system operators. One of the biggest concerns associated with integrating a large amount of wind power into the grid is the ability to handle large ramps in wind power output. Large ramps can significantly influence system economics and reliability, on which power system operators place primary emphasis. The Wind Forecasting Improvement Project (WFIP) was performed to improve wind power forecasts and determine the value of these improvements to grid operators. This paper evaluates the performance of improved short-term wind power ramp forecasting. The study is performed for the Electric Reliability Council of Texas (ERCOT) by comparing the experimental WFIP forecast to the current short-term wind power forecast (STWPF). Four types of significant wind power ramps are employed in the study; these are based on the power change magnitude, direction, and duration. The swinging door algorithm is adopted to extract ramp events from actual and forecasted wind power time series. The results show that the experimental short-term wind power forecasts improve the accuracy of the wind power ramp forecasting, especially during the summer.

Keywords: Wind forecasting, grid integration, ramp forecasting, performance diagram, swinging door algorithm

[*]Postdoctoral Researcher, Transmission and Grid Integration Group, ASME Member, Corresponding author.
[#]Research Engineer, Transmission and Grid Integration Group, ASME Member.
[†]Senior Research Engineer, Transmission and Grid Integration Group.
[‡] Lead Research Scientist.

INTRODUCTION

Wind energy is becoming an increasingly important source of renewable energy in the electric power system. Currently, wind power meets 3.78% of U.S. electricity demand [1], with some systems (such as in Texas) having instantaneous wind power penetrations up to 28% [2]. Wind power has a maximum upper availability limit that is both variable and uncertain at multiple timescales. More generally, the variability and uncertainty of wind generation presents a primary concern to system operators. Thus, wind power forecasting plays an important part in power system operations given the increasing amount of wind power integrated in the electrical system. Although the current power system is capable of handling small amounts of uncertainty and variability, ramp (or extreme) events—sudden and large changes in wind power—are a critical issue. Improving the accuracy of wind power forecasting is expected to reduce the discrepancy between the forecast and actual wind power output, thereby enhancing the performance of wind power ramp forecasting and reducing wind integration costs.

Overview of Wind Forecasting

Wind forecast models can be broadly divided into two categories [3]: (i) forecasting based on the analysis of historical time series of wind; and (ii) forecasting based on numerical weather prediction (NWP) models. The first type of forecast model generally provides reasonable results in the estimation of long-term horizons, such as mean monthly, quarterly, and annual wind speed. In addition, statistical and machine learning techniques that utilize historical data have been shown to work well for forecast horizons less than one hour [4, 5]. For short-

term horizons more than one hour (daily or hourly forecasts), the impact of atmospheric dynamics becomes more important and NWP models—such as the Advanced Regional Prediction System (ARPS) [6, 7], the Weather Research & Forecasting Model (WRF) [8], and Global Forecast System (GFS) [9]—become more suitable. Short-term wind power forecasting (between 1 hour and 72 hours) is uniquely helpful in power system planning for the unit commitment and economic dispatch process. A variety of topics on short-term wind power forecasting have been studied in the literature, including distributions of wind power forecast errors [10-13], uncertainties in wind forecasting [13-15], and wind power ramp forecasting [16-21]. The focus of this paper is to investigate the impacts of improved wind power forecasting on wind ramp forecasting performance.

Overview of Ramp Forecasting

One of the biggest concerns associated with integrating a large amount of wind power into the grid is the ability to handle large ramps in wind power output. For example, a large down ramp event occurred in the Electric Reliability Council of Texas (ERCOT) system on February 26, 2008, that caused a system emergency [22]. Different time and geographic scales influence wind ramps, and there can be both up and down ramps with varying levels of severity. There are two main ways in which inaccurate forecasting of ramp events can lead to large errors: ramp magnitude and timing errors. The magnitude error is defined as an event that is forecasted to occur at an expected time but with significantly different magnitude. In ramp timing errors, the actual ramp in power significantly leads/lags the forecasted ramp time.

Based on three key characteristics—direction, duration, and magnitude—ramps can be defined, characterized, and identified. Ferreria et al. [16] provided an overview of different ramp definitions and approaches in ramp event forecasting. Greaves et al. [17] defined a ramp as a change in wind power output that is at least 50% of the installed wind capacity and occurs within a time span of 4 hours or less. Zheng and Kusiak [19] employed the rate of change of wind power output over a 10-minute interval to define ramps. Potter et al. [23] defined a ramp event as the change in power between two consecutive hours that is greater than or equal to 10% of the installed wind capacity. In a report by AWS Truewind [24], the up and down ramps were differently defined: (i) a down ramp occurs if the power changes at least 15% of total capacity within one hour; and (ii) an up ramp occurs if the power changes at least 20% of total capacity within 1 hour. AWS Truewind also analyzed the ERCOT system, which defines a ramp event as a change of 20% or more of the rated capacity in any 30-minute period [25].

Research Motivation and Objectives

The U.S. Department of Energy funded the Wind Forecasting Improvement Project (WFIP) [26], which was performed to improve short-term wind power forecasts and determine the value of these improvements to grid operators. Large ramps can significantly influence system economics and reliability, on which power system operators place primary emphasis. This paper evaluates the performance of wind ramp forecasting based on improved short-term wind power forecasting. The analysis is performed for ERCOT by comparing the experimental WFIP forecast to the current short-term wind power forecast (STWPF).

The following topics are discussed in the remainder of the paper: (i) the WFIP analyzed regions and forecasting system; (ii) a statistical analysis comparing the performance of the WFIP to the current STWPF for ERCOT; (iii) the ramp forecasting methodology and the metrics for assessing ramp forecasting; and (iv) results and discussion of the ramp event forecasting through improved WFIP wind forecasts.

WIND FORECASTING IMPROVEMENT PROJECT

WFIP encompassed two study regions: the northern study region and the southern study region [26]. In this paper, the performance of wind power ramp forecasting based on the improved wind forecasting is analyzed for the southern study region.

WFIP Southern Study Region

The WFIP southern study region covers most of the ERCOT service area, as shown in Fig. 1. As of March 2012, ERCOT had 9,838 MW of wind capacity installed. The WFIP study region included 8,296 MW of wind capacity spread throughout 84 wind power plants. Day-ahead (DA) and 1- to 6-hour-ahead (1HA to 6HA) wind power forecasts were generated by using both the WFIP and STWPF systems for a nearly 12-month period from October 2011 to mid-September 2012. Actual and forecast load data at the same wind forecast timescales for the 12-month period were obtained from ERCOT.

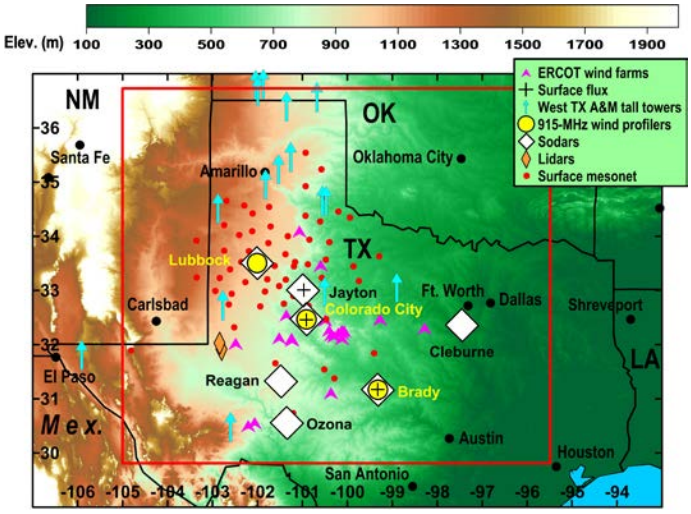


FIGURE 1. WFIP: SOUTHERN STUDY REGION IN ERCOT [26]

Wind Forecasting System

The current wind forecasting system used by ERCOT is the baseline for this study. The WFIP experimental forecast system consists of an ensemble of high-resolution rapid-update NWP models. Each of these ensemble members incorporates a variety of model configurations, physics parameterizations, and data

assimilation techniques. The purpose of integrating all of these ensemble members into one system is to construct an optimized composite forecast able to predict forecast uncertainty and assess the relative performance of different modeling approaches. Figure 2 shows the overall framework of the wind power forecasting system. The ensemble members include [27]:

i. The National Oceanic and Atmospheric Administration's 3-km High-Resolution Rapid Refresh (HRRR) updated hourly;
ii. Nine NWP models updated every 2 hours on a 5-km grid:
   (a) Three configurations of the Advanced Regional Prediction System (ARPS);
   (b) Three configurations of the Weather Research & Forecasting (WRF) model;
   (c) Three configurations of the Mesoscale Atmospheric Simulations System (MASS); and
iii. Oklahoma University's version of ARPS updated every 6 hours on a 2-km grid.

The data from additional sensors deployed for this project, as well as the data from a set of participating wind power plants within Texas, were assimilated into most of the ensemble members; however, the data from the project sensors were withheld from some ensemble members to gauge their impact on the forecasts [27].

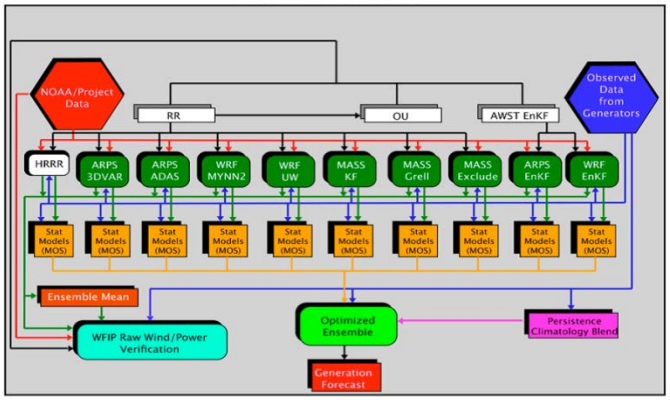


FIGURE 2. OVERALL FRAMEWORK OF THE WIND POWER FORECASTING SYSTEM [27]

A model output statistics (MOS) procedure was applied to the forecasts from each NWP system. The MOS is designed to correct systematic errors of relevant NWP meteorological variables (e.g., wind speed and direction) at forecast sites. The MOS output for the individual NWP systems was then used as input to an optimized ensemble model (OEM), which created a composite deterministic or probabilistic forecast from the set of MOS-adjusted NWP forecasts. In addition to the NWP forecasts, statistical predictions based purely on recent observational data were also included in the ensemble. Two OEM training strategies were tested: one was based on a rolling sample of the last 30 days; a second approach was based on a customized "analog" training sample. The training sample was constructed by matching key weather parameters of the current forecast period to those of cases in a historical archive. The objective of the regime-based approach was to weight the individual members of the ensemble according to their performance in weather patterns that were similar to the one expected during the forecast period. More details of the forecasting system can be found in [27].

WIND POWER FORECASTING PERFORMANCE COMPARISON OF WFIP AND STWPF

Before analyzing the wind power ramp forecasting performance of WFIP versus STWPF, a statistical analysis was performed to understand the forecast improvements provided by WFIP.

Statistical Metrics for Comparing WFIP to STWPF

Three statistical metrics were adopted to compare the forecasting performance of WFIP and STWPF: (i) Pearson's correlation coefficient, (ii) root mean square error (RMSE), (iii) mean absolute error (MAE), and mean bias error (MBE).

Formulations of Statistical Metrics Pearson's correlation coefficient is a measure of the correlation between two variables (or sets of data), expressed as:

$$\rho = \frac{\mathrm{cov}(p, \hat{p})}{\sigma_p \sigma_{\hat{p}}} \tag{1}$$

where p and $\hat{p}$ represent the actual and forecasted wind power output, respectively. Pearson's correlation coefficient is a global error measure metric; a larger value of Pearson's correlation coefficient indicates a wind power forecast that more closely matches reality (i.e., has smaller errors).

The RMSE also provides a global error measure during the entire forecasting period. It is given by

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{p}_i - p_i)^2} \tag{2}$$

where p_i represents the actual wind power generation at the i^{th} time step, $\hat{p}_i$ is the corresponding wind power forecast, and N is the number of points estimated in the forecasting period.

The MAE has been widely used in regression problems and by the renewable energy industry to evaluate forecasting performance. It is given by

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|\hat{p}_i - p_i\right| \tag{3}$$

The MAE metric, unlike the RMSE metric, weights all values equally and thus does not add additional weight to extreme forecasting events.

The MBE is expressed as

$$MBE = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{p}_i - p_i\right) \tag{4}$$

The MBE metric intends to indicate average forecasting bias. Understanding the overall forecasting bias (over- or underforecasting) would allow power system operators to better allocate resources to compensate for forecasting errors in the dispatch process.

A Comparison of Forecasts Using Statistical Metrics The values of different metrics used to evaluate the STWPF and WFIP wind power forecasts at multiple forecasting horizons are shown in Table 1. The relative RMSE and MAE (the RMSE and MAE divided by the wind power capacity 8,296 MW) are shown in the table. As expected and inferable from the metrics of correlation coefficient, RMSE, and MAE, (i) the 1HA forecast performed best and the DA forecast performed worst; and (ii) the experimental WFIP provided better results than the current STWPF for all forecasting timescales. Table 2 shows a comparison of the performances of the 6HA forecasts in WFIP and STWPF for each month. It was observed that 6HA WFIP forecasts had smaller RMSE and MAE values for all 12 months. According to the MBE metric, the experimental WFIP forecast tended to overforecast in most months; and the current STWPF forecast tended to underforecast in most months. Although the WFIP forecast had less bias than the STWPF forecast in most months, relatively more bias was shown in the WFIP forecast in December and September.

TABLE 1. METRICS VALUES FOR EVALUATING STWPF AND WFIP IN A YEAR

Metric	Correlation Coefficient		RMSE/Capacity		MAE/Capacity	
	STWPF	WFIP	STWPF	WFIP	STWPF	WFIP
1HA	0.96	0.99	7.24%	3.66%	5.43%	2.70%
2HA	0.94	0.97	8.89%	6.60%	6.70%	4.91%
3HA	0.93	0.95	9.95%	8.23%	7.53%	6.16%
4HA	0.92	0.94	10.65%	9.22%	8.09%	6.91%
5HA	0.91	0.93	11.12%	9.88%	8.46%	7.44%
6HA	0.91	0.92	11.44%	10.38%	8.72%	7.82%
DA	0.90	0.90	11.96%	11.68%	9.01%	8.64%

TABLE 2. METRICS VALUES OF 6HA STWPF AND WFIP IN EACH MONTH

Month	RMSE/Capacity		MAE/Capacity		MBE/Capacity	
	STWPF	WFIP	STWPF	WFIP	STWPF	WFIP
Oct.-11	11.44%	9.25%	8.73%	6.88%	-1.03%	-0.17%
Nov.-11	10.30%	9.07%	7.84%	6.97%	1.09%	0.36%
Dec.-11	11.86%	11.41%	9.27%	8.71%	-0.04%	1.58%
Jan.-12	11.33%	10.38%	8.73%	7.98%	-1.76%	-1.59%
Feb.-12	13.56%	13.05%	10.12%	9.76%	-1.16%	0.47%
Mar.-12	12.82%	10.87%	9.99%	8.46%	-2.89%	0.03%
Apr.-12	12.24%	11.22%	9.12%	8.27%	-0.59%	0.28%
May.-12	12.51%	10.48%	9.93%	8.39%	-0.38%	0.44%
Jun.-12	11.49%	11.39%	8.91%	8.42%	-1.21%	0.21%
Jul.-12	9.36%	8.76%	6.92%	6.57%	0.52%	-0.32%
Aug.-12	9.56%	8.56%	7.53%	6.75%	1.40%	1.09%
Sep.-12	9.20%	8.75%	7.01%	6.35%	-0.34%	-1.31%

Distributions of Wind Power Forecasting Errors

Forecasts are important considerations in committing and dispatching generating units. The estimation of forecast confidence intervals can be calculated using an assumed error distribution on the point forecast and based on historical data. An important resulting characteristic of the WFIP forecasts was a preferable distribution of wind power forecasts, as described in more detail in the next sections.

Method for Estimating Error Distributions Multiple distribution types have been analyzed in the literature to quantify the distribution of wind power forecasting errors, including the hyperbolic distribution, kernel density estimation (KDE), the normal distribution, and the Weibull and beta distributions [28]. KDE was adopted in this paper to model the distribution of wind power forecasting errors for different forecasting scenarios. KDE is a nonparametric approach to estimate the probability density function of a random variable. It has been widely used in the wind energy community for wind speed distribution and wind power forecasting errors [13, 29, 30]. In this paper, the Gaussian kernel is considered throughout. The mean integrated squared error, the most commonly used optimality criterion, is used in this article for bandwidth selection.

Comparing the Results of Error Distributions Figure 3 and Fig. 4 show the distributions of wind power forecasting errors at different forecasting horizons by using the STWPF and WFIP, respectively. The wind power forecasting errors were normalized by the WFIP region wind capacity of 8,296 MW in the analysis. Among DA and 1HA to 6HA forecasts, it was observed that the 1HA horizon performed the best for both the STWPF and WFIP forecasts, as indicated by the narrower error distribution curves. The DA forecasts in both the STWPF and WFIP performed the worst, as expected. As shown in Fig. 3 and Fig. 4, the 1HA forecasts had the largest probability when the forecasting error was smaller (i.e., pronounced peak), and the DA forecasts had the largest probability when the forecasting error was larger (i.e., fat tails). The expected observations could be partially attributed to the atmospheric dynamics that can be more accurately predicted fewer hours ahead. A comparison of the STWPF to the WFIP showed that the 1HA forecasts from the WFIP performed significantly better than the 1HA forecasts from the STWPF.

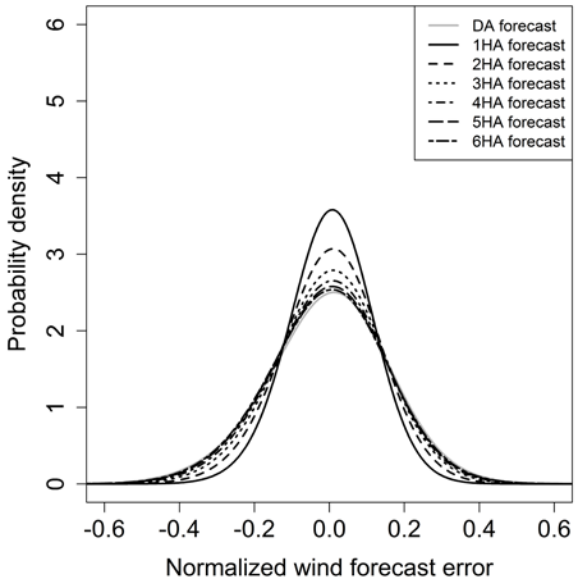


FIGURE 3. DISTRIBUTIONS OF STWPF FORECAST ERRORS AT DIFFERENT TIMESCALES

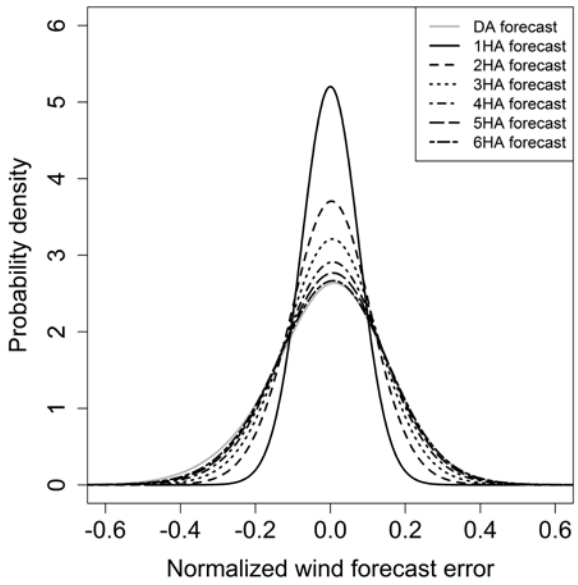


FIGURE 4. DISTRIBUTIONS OF WFIP FORECASTING ERRORS AT DIFFERENT TIMESCALES

RAMP FORECASTING METHODOLOGY AND PERFORMANCE EVALUATION METRICS

The study employed four definitions of significant wind power ramps based on the power change magnitude and duration. Ramp events need to be extracted from actual and forecast wind power series, and the swinging door algorithm was adopted for this purpose. A suite of metrics were proposed to evaluate the performance of ramp forecasting compared to different types of wind forecasts.

Definitions of Significant Wind Power Ramps

Four types of significant ramps are defined based on (i) ramp magnitude only; (ii) ramp magnitude and duration; (iii) ramp change rate; and (iv) ramp direction, magnitude, and duration.

Significant Ramp Definition 1: Ramp Magnitude Only The first definition of a significant ramp is based on the magnitude of wind power change. In this paper, a significant ramp is defined as the change in wind power output that is greater than 30% of the installed wind capacity, expressed as

$$\left|P\left(t+\Delta_t\right)-P\left(t\right)\right| > P_{val} \tag{5}$$

where $P\left(t\right)$ is the wind power output at time t; Δ_t is the duration of the ramp, which is not specified in significant ramp definition 1; and P_{val} is the predefined threshold value, which is 30% of the wind capacity.

Significant Ramp Definition 2: Ramp Magnitude and Duration The second definition defines a significant ramp based on both the magnitude and duration of wind power change. In the study, the significant ramp is defined as the change in wind power output that is greater than 25% of the installed wind capacity and occurs within a time span of 4 hours or less, which is also expressed in Eq. (5). In the equation, the duration of the ramp, Δ_t, is less than or equal to 4 hours; and the threshold value, P_{val}, is 25% of the installed wind capacity.

Significant Ramp Definition 3: Ramp Change Rate The third definition of significant ramps is based on the change rate of wind power. In this paper, a significant ramp rate is defined as the change rate in wind power output that is greater than 10% of the installed wind capacity, expressed as:

$$\frac{\left|P\left(t+\Delta_t\right)-P\left(t\right)\right|}{\Delta_t} > R_{val} \tag{6}$$

where R_{val} is the predefined threshold value of change rate in wind power output.

Significant Ramp Definition 4: Ramp Direction, Magnitude, and Duration The fourth definition of significant ramps is based on the change direction, magnitude, and duration of wind power output. In the paper, a significant up ramp is defined as the change in wind power that is greater than 20% of wind capacity within a time span of 4 hours or less; and a significant down ramp is defined as the change in wind power that is greater than 15% of wind capacity within a time span of 4 hours or less.

$$P\left(t+\Delta_t\right)-P\left(t\right) > P_{val}^{u} \tag{7}$$

$$P\left(t+\Delta_t\right)-P\left(t\right) < -P_{val}^{d} \tag{8}$$

where P_{val}^{u} and P_{val}^{d} represent the up and down ramp threshold values, respectively.

Ramp Extraction Using the Swinging Door Algorithm

Ramps are extracted through a linear piecewise approximation to the original time series of data, actual, or forecasted wind power in this study. To determine any significant ramp or ramp rate as defined by definitions 1 to 4, the start and end points of all ramps in a given time series of wind power need to be identified. Toward this end, the swinging door algorithm is adopted to extract ramp periods in a series of power signals by identifying the start and end points of each ramp.

The swinging door algorithm allows for the consideration of a threshold parameter influencing the algorithm's sensitivity to ramp variations. The only tunable parameter in the algorithm is the width of a "door," represented by ε in Fig. 5. The parameter ε directly characterizes the threshold sensitivity to noise and/or insignificant fluctuations to be specified. With a smaller ε value, many small ramps will be identified; with a larger ε value, only a few large ramps will be identified. It is important to note that the scale in Fig. 5 is arbitrary for the purpose of explanation, and in general the signal magnitude is much larger than the scale of the threshold bounds. A detailed description of the swinging door algorithm can be found in [18].

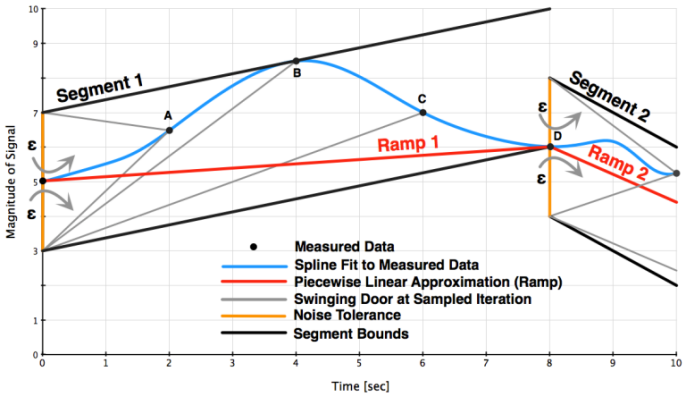


FIGURE 5. THE SWINGING DOOR ALGORITHM FOR THE EXTRACTION OF RAMPS IN POWER FROM THE TIME SERIES [18]

Figure 6 show a typical example in the extraction of ramps from actual wind power generation at ERCOT during a 100-hour period. The hourly wind power data is recorded in the WFIP region, which is an aggregation of outputs from 84 wind power plants. The tolerance value, ε, is set at 2.5% of installed wind capacity. In the figure, the solid and dashed lines represent the actual wind power and the piecewise linear approximation (generated by the swinging door algorithm), respectively. An accurate piecewise linear approximation to the actual wind power profile is obtained as shown in Fig. 6. The figure presents the nature of up and down ramps with large, medium, and insignificant changes in power. The extracted ramps in the actual wind power generation are visualized by rise-run distributions. Figure 7 shows the bivariate distribution of all wind power ramps. It is observed that the distribution spreads within the more immediate ramp region.

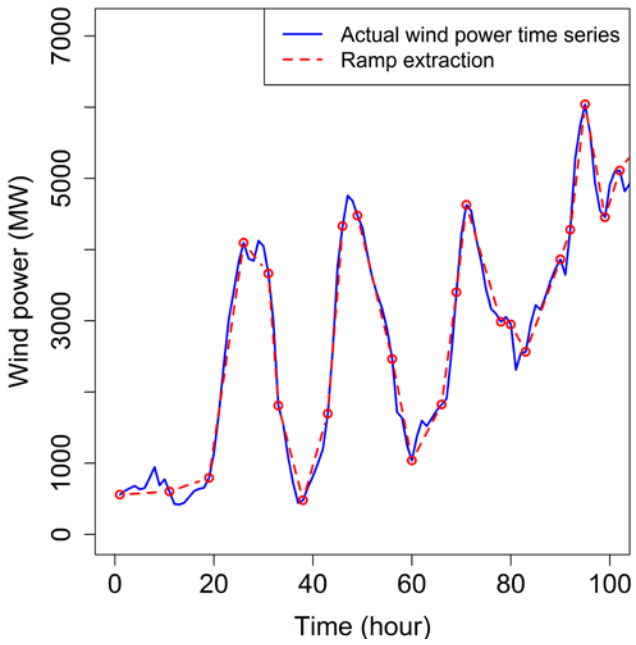


FIGURE 6. RAMP EXTRACTION FROM THE ACTUAL WIND POWER GENERATION

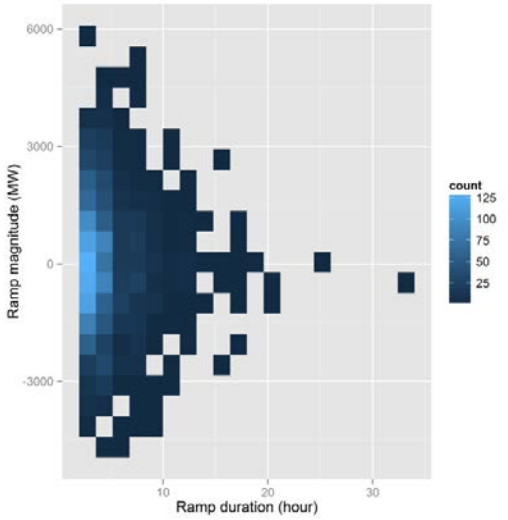


FIGURE 7. BIVARIATE DISTRIBUTION OF ACTUAL WIND POWER RAMPS

Based on the four ramp definitions, significant ramps can be identified. Each significant ramp is characterized by the (i) start and end hours; (ii) wind power at the start and end points; and (iii) direction of the ramp.

Metrics of Evaluating Significant Ramp Forecasting Performance

A suite of event detection metrics are used to evaluate the performance of ramp forecasting, including a contingency table, categorical statistics, and performance diagrams.

Contingency Table The contingency table provides a measure of skill for forecasts of notable events [16]. To evaluate ramp forecasting, all forecasts and observations of significant wind ramps are grouped into four categories based on whether the forecast accurately predicts the actual ramps. Table 3 is a contingency table that summarizes the results of a ramp forecasting system. *True positive* (TP) represents the number of forecasted ramps (*forecast YES*) that are actually observed in the actual power output (*observed YES*); *false positive* (FP) is the number of forecasted ramps that are not observed in the actual wind power (*observed NO*); *false negative* (FN) represents the number of observed ramps (*observed YES*) that are not predicted by the wind forecasting system (*forecast NO*); *true negative* (TN) is the number of non-occurring events for both observed and forecast results; and *N* is the total number of events.

TABLE 3. CONTINGENCY TABLE FOR RAMP EVENT OBSERVATION AND FORECAST

	Observed YES	Observed NO	Total
Forecast YES	*TP (hits)*	*FP (false alarm)*	*TP+FP*
Forecast NO	*FN (misses)*	*TN*	*FN+TN*
Total	*TP+FN*	*FP+TN*	*N= TP+FP+ FN+TN*

Categorical Statistics and Performance Diagram

Categorical statistics provide measures of accuracy and skill for forecasts of notable events, such as ramps in power, detrimental temperatures, or rainfall. Based on the contingency table, a suite of metrics can be derived for ramp forecasting performance evaluation, given as follows.

Probability of detection (POD) is defined as the ratio between the number of true positives and the number of observed positives, which indicates the fraction of observed YES events that are actually forecasted.

$$POD = \frac{TP}{TP + FN} \tag{9}$$

Critical success index (CSI) is used to measure the fraction of observed and/or forecasted events that are correctly predicted, given by

$$CSI = \frac{TP}{TP + FN + FP} \tag{10}$$

The value of CSI is between 0 and 1, with 1 representing perfect prediction. The CSI considers only observed and forecasted ramps, excluding true negative events.

Frequency bias score (FBIAS) measures the ratio of the frequency of forecasted YES events to the frequency of observed YES events.

$$FBIAS = \frac{TP + FP}{TP + FN} \tag{11}$$

The ramp forecast system tends to underforecast when FBIAS<1, and it tends to overforecast when FBIAS>1.

False alarm ratio (FAR) measures the fraction of predicted YES events that did not occur, given by

$$FAR = \frac{FP}{FP + TP} \tag{12}$$

The metric *success ratio* (SR) is calculated from FAR by subtracting it from 1. SR measures the fraction of predicted YES events that occurred.

The relationship among the POD, CSI, FBIAS, and FAR can be visualized on a *performance diagram* [31] based on

$$CSI = \frac{1}{\frac{1}{POD} + \frac{1}{1 - FAR} - 1} \tag{13}$$

$$FBIAS = \frac{POD}{1 - FAR} \tag{14}$$

EVALUATING SIGNIFICANT RAMP FORECASTING PERFORMANCE BASED ON IMPROVED WIND POWER FORECASTS

This section shows the effects of forecast improvement on wind power ramp forecasting. Greaves et al. [17] defined a true positive forecast as a forecast ramp with a measured ramp of the same direction (either up or down) within ±12 hours of the time of the forecast ramp. In this study, a true positive forecast is defined as a forecast ramp with a measured ramp of the same direction within ±6 hours of the time of the forecast ramp. This ±6-hours range could provide sufficient data for temporal uncertainty analysis and maintain a realistic connection between forecast and measured significant wind ramp events.

Annual Ramp Forecasting Performance

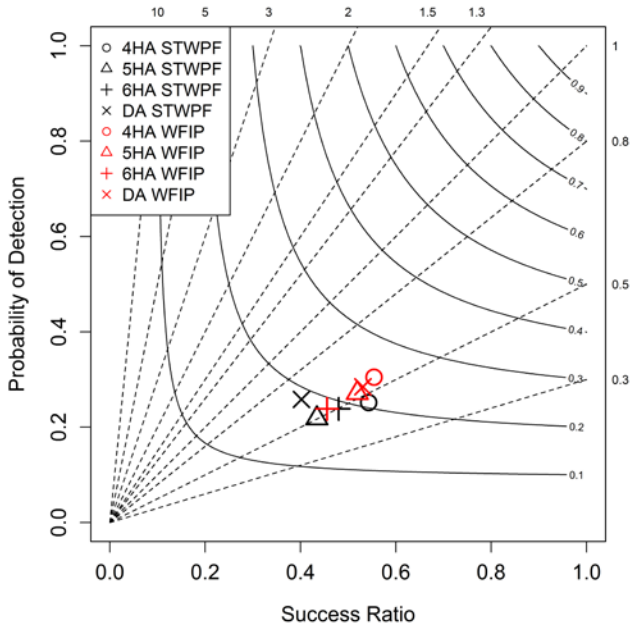
Table 4 lists the number of significant up and down ramps in actual wind power throughout the whole year. There were more up ramps than down ramps for all four ramp definitions. There are more significant ramps of definition 4 than of the other three criteria.

TABLE 4. NUMBER OF OBSERVED RAMPS

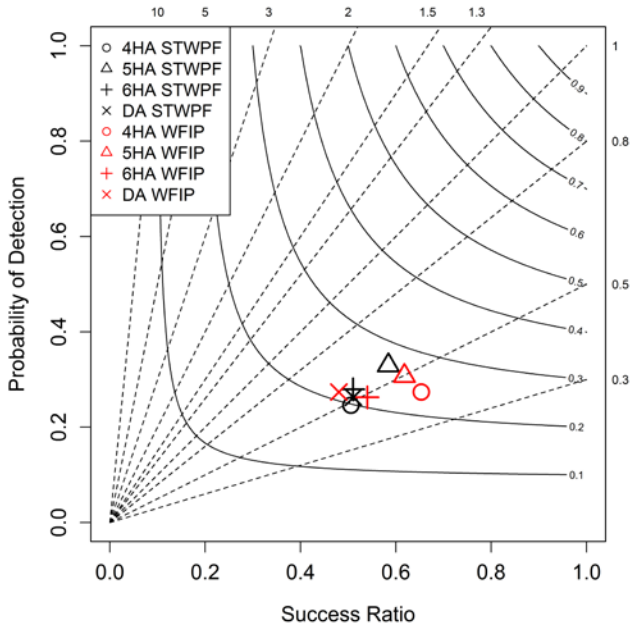
Ramp Type	Ramp Def. 1	Ramp Def. 2	Ramp Def. 3	Ramp Def. 4
Up Ramps	87	105	86	233
Down Ramps	64	74	54	223

A performance diagram can be used to understand whether the wind power forecasting is improved. Figure 8 shows the annual ramp forecasting performance between STWPF and WFIP for different forecast timescales and significant ramp definitions. In the performance diagrams, (i) the left axis represents the value of POD; (ii) the bottom axis represents the success ratio; (iii) the diagonal dashed lines represent FBIAS with the values shown on the right and top axes; and (iv) the solid curves show CSI with the values on the right-inside graph border. The figure shows the 4HA, 5HA, 6HA, and DA forecasts. The performance diagram shows the ramp forecast performance space—that is, as the forecast moves toward the upper right of the diagram, the overall ramp forecast metrics improve.

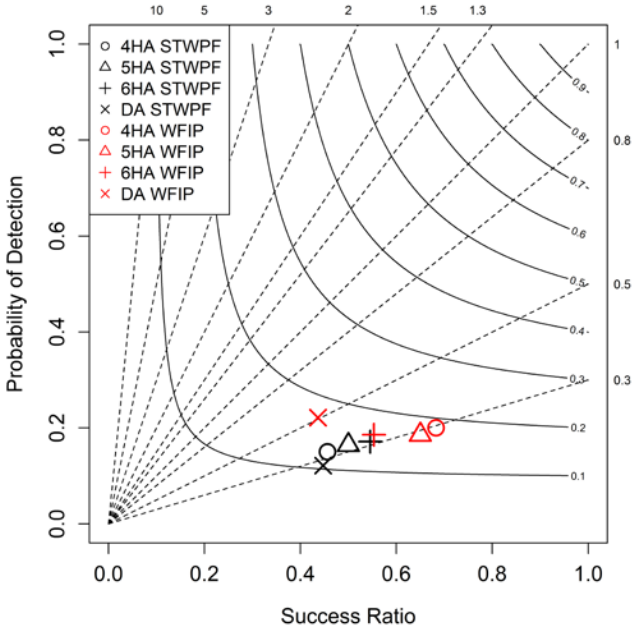
The performance diagram in Fig. 8 compares the annual ramp forecasting performance of the STWPF to the WFIP for different forecast timescales and ramp definitions. The WFIP ramp forecasting performance is evaluated using four forecasting timescales: 4HA, 5HA, 6HA, and DA. It is observed in Fig. 8(a) that (i) the 4HA WFIP has a larger success ratio, POD, and CSI values than the 4HA STWPF; (ii) the 5HA WFIP also has a larger success ratio, POD, and CSI values than the 5HA STWPF; (iii) the ramp forecasting performance of the 6HA WFIP and STWPF is similar; and (iv) the DA WFIP also has a larger success ratio, POD, and CSI values than the DA STWPF. Similar results are also observed from the other three definitions of significant ramps in Figs. 8(b)-8(d). Overall, the ramp forecasting based on the improved wind power forecasts is more accurate. The FBIAS values are smaller than one for all cases in Fig. 8, showing that the WFIP and STWPF ramp forecasting tend to underforecast.



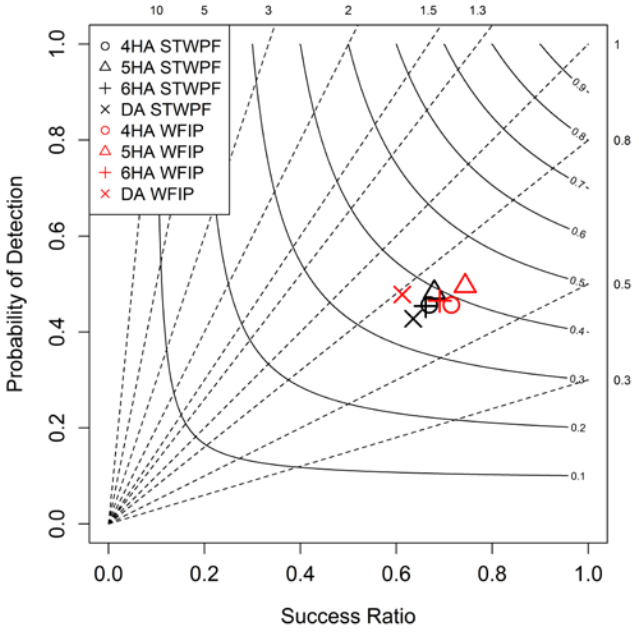
(a) Ramp Definition 1: Ramp magnitude only



(b) Ramp Definition 2: Ramp magnitude and duration



(c) Ramp Definition 3: Ramp change rate

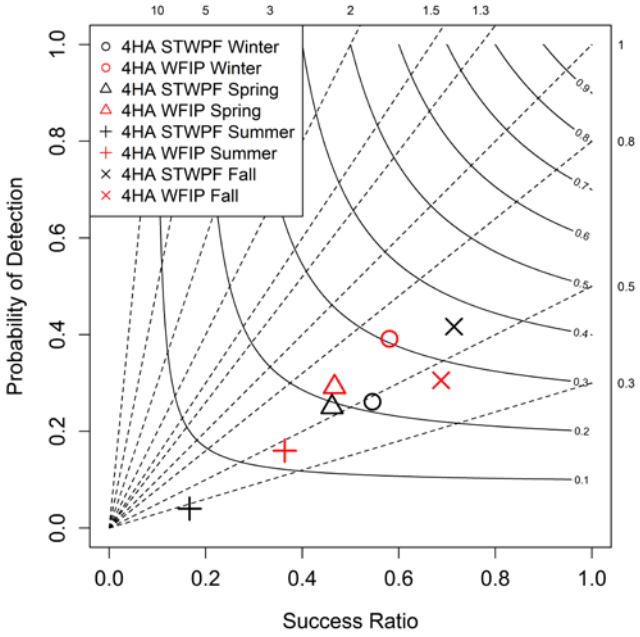


(d) Ramp Definition 4: Ramp direction, magnitude, and duration

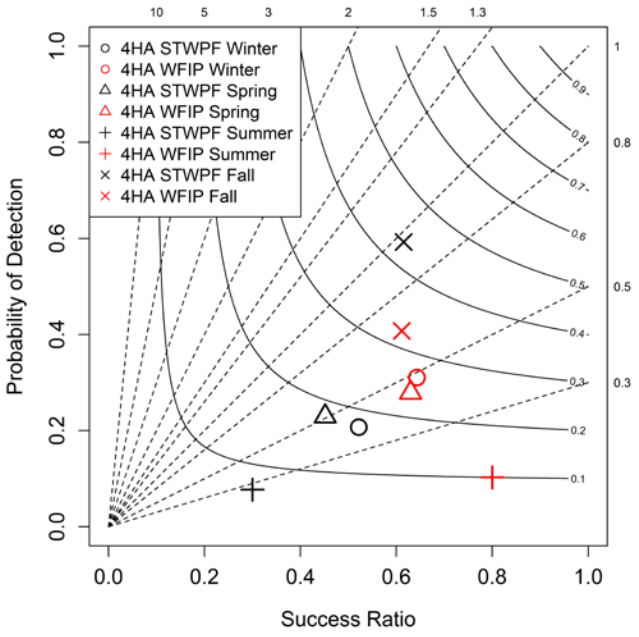
FIGURE 8. COMPARING ANNUAL RAMP FORECASTING PERFORMANCE OF STWPF AND WFIP FOR DIFFERENT FORECAST TIMESCALES AND RAMP DEFINITIONS

Seasonal Ramp Forecasting Performance

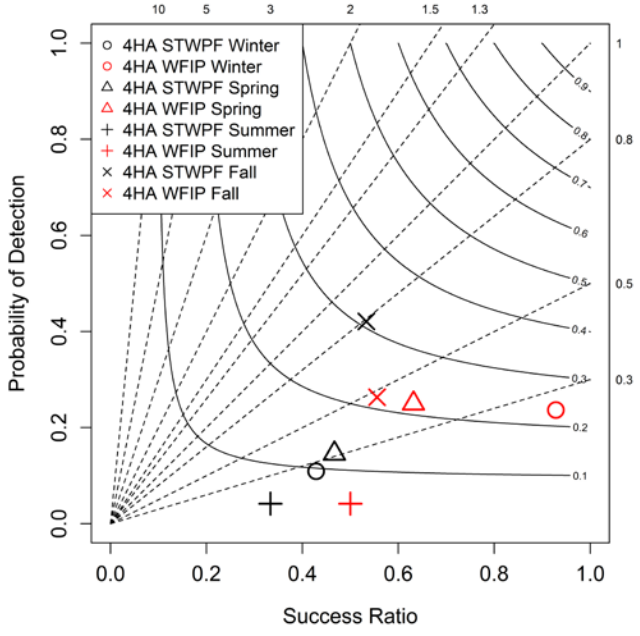
Figure 9 compares seasonal ramp forecasting performance of the 4HA STWPF to the 4HA WFIP forecasts. For all four significant ramp definitions in Figs. 9(a)-9(d), the ramp forecasting performs relatively better in fall and relatively worse in summer for both the WFIP and STWPF. This can be partially attributed to the features responsible for ramps. The summer tends to be more convective (mesoscale), hence it is more difficult to forecast (especially on the 4HA time scale) than the larger synoptic scale (such as fronts) features that cause ramps in the colder seasons. It is observed in Figs. 9(a) and 9(b) that during summer the ramp forecasting based on WFIP performs significantly better than that based on STWPF. However, during fall the ramp forecasting based on STWPF performs slightly. This significant improvement of ramp forecasting performance in summer based on the improved wind forecasts could play an important part in enhancing system economics and reliability because high electric demand is generally expected in ERCOT during the summer period. After comparing the ramp forecasting performance among the four different significant ramp definitions, it is evident that ramp forecasts generally have a larger success ratio, POD, and CSI values with Definition 4, as shown in Fig. 9(d). The FBIAS value is also closer to one with Definition 4.



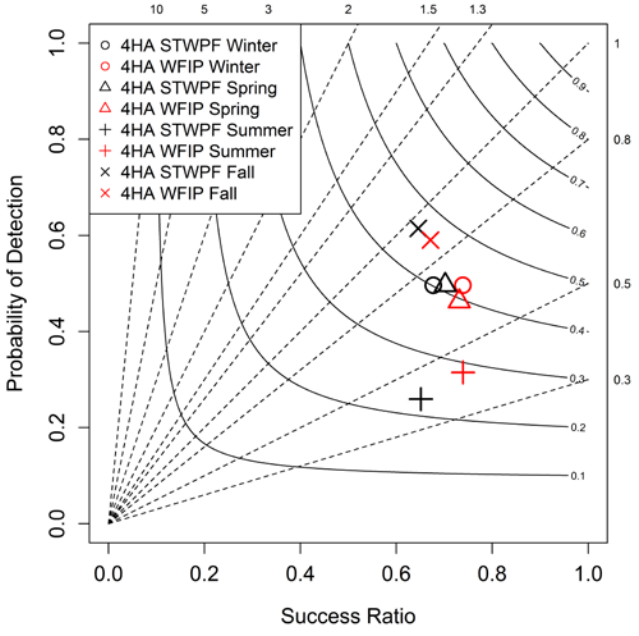
(a) Ramp Definition 1: Ramp magnitude only



(b) Ramp Definition 2: Ramp magnitude and duration



(c) Ramp Definition 3: Ramp change rate

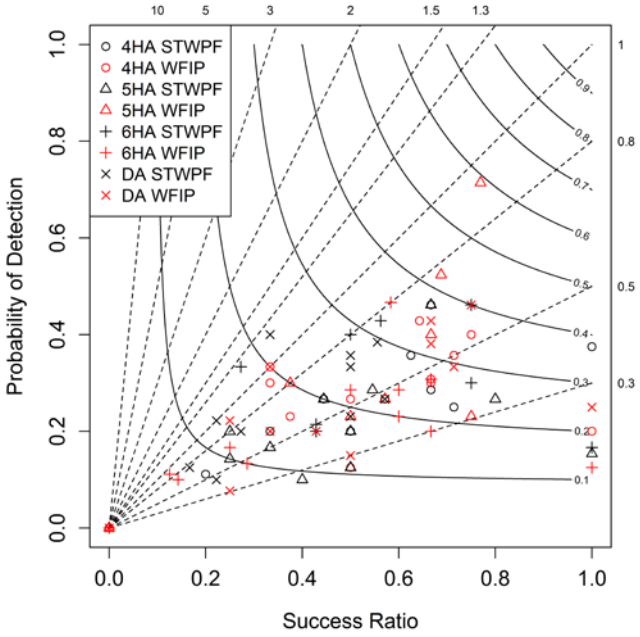


(d) Ramp Definition 4: Ramp direction, magnitude, and duration

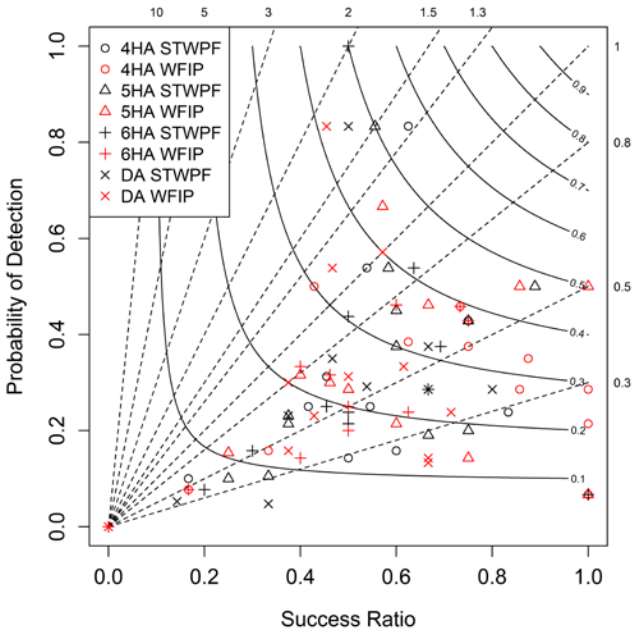
FIGURE 9. COMPARING SEASONAL RAMP FORECASTING PERFORMANCE OF STWPF TO WFIP FOR 4HA FORECASTS

Monthly Ramp Forecasting Performance

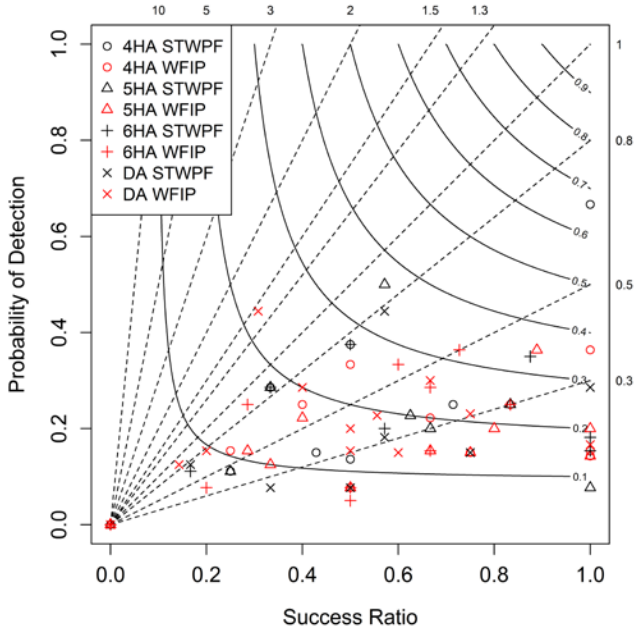
Figure 10 compares monthly ramp forecasting performance among the 4HA to 6HA and DA STWPF and WFIP forecasts. The 12 points represented by each symbol (e.g., circle) indicate the ramp forecasting performance in each month. It is observed that the values of POD and success ratio could reach one during some months, such as 4HA WFIP, 5HA WFIP, and 6HA STWPF, as shown in Fig. 10(b). It is also observed that with different forecasting timescales and months, the variation in the ramp forecasting performance based on the significant ramp Definition 4 is relatively less than that based on the other three significant ramp definitions. Although the annual ramp forecasting tends to underforecast (FBIAS<1 as shown in Fig. 8), the system tends to overforecast in a few months as shown in Fig. 10 in the cases of FBIAS>1.



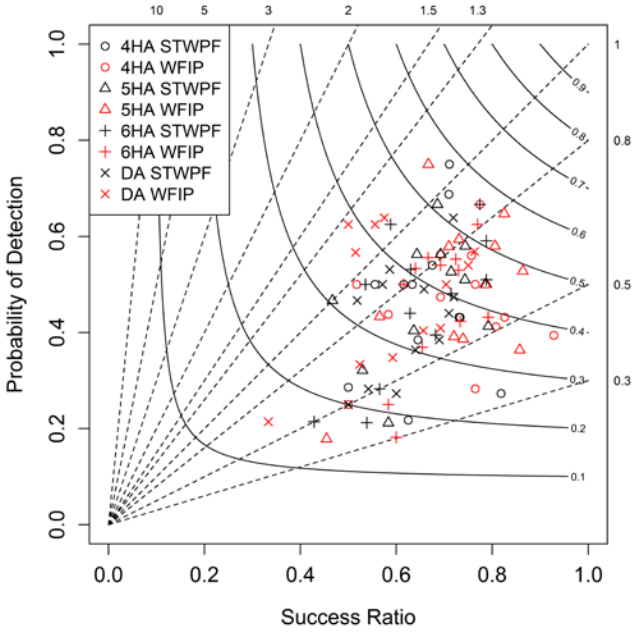
(a) Ramp Definition 1: Ramp magnitude only



(b) Ramp Definition 2: Ramp magnitude and duration



(c) Ramp Definition 3: Ramp change rate



(d) Ramp Definition 4: Direction, magnitude, and duration

FIGURE 10. COMPARING MONTHLY RAMP FORECASTING PERFORMANCE OF STWPF TO WFIP FOR DIFFERENT FORECAST TIMESCALES AND RAMP DEFINITIONS

CONCLUSION

This paper characterized ramp forecasting performance by using the experimental forecasts from the Wind Forecasting Improvement Project (WFIP) in place of the current Electric Reliability Council of Texas (ERCOT) short-term wind power forecast (STWPF). A suite of statistical metrics were used to evaluate the overall improvement of the WFIP in short-term wind power forecasting accuracy for different forecasting horizons. Statistical analyses of the results showed that in most seasons/months, the experimental WFIP provided better performance than the current STWPF for all forecasting horizons. An evaluation of the ramp forecasting improvement was performed based on four types of significant ramp definitions. The wind power ramps were extracted using the swinging door algorithm.

The results showed that improved wind power forecasts could also improve the accuracy of wind power ramp forecasting, especially during the summer period in ERCOT. The results also showed that the ramp forecasting for both the WFIP and STWPF tended to underforecast during the whole year. The ramp forecasting performed relatively better with the

ramp Definition 4 based on the direction, magnitude, and duration of the ramps.

ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Energy under Contract No. DE-AC36-08-GO28308 with the National Renewable Energy Laboratory.

REFERENCES

[1] U.S. Energy Information Administration, 2013, "Electric Power Monthly," Technical Report, U.S. Department of Energy, Washington, D.C.

[2] Electric Reliability Council of Texas, Inc., http://www.ercot.com/.

[3] Foley, A. M., Leahy, G. L., Marvuglia, A., and Mckeogh, E. J., 2012, "Current Methods and Advances in Forecasting of Wind Power Generation," *Renewable Energy*, 37, pp. 1–8.

[4] Kariniotakis, G. N., Stavrakakis, G. S., and Nogaret, E. F., 1996, "Wind Power Forecasting Using Advanced Neural Networks Models," *IEEE Transactions on Energy Conversion*, 1(14), pp. 762–767.

[5] Monteiro, C., Bessa, R., Miranda, V., Botterud, A., Wang, J., and Conzelmann, G., 2009, "Wind Power Forecasting: State-of-the-Art 2009," Technical Report No. ANL/DIS-10-1, Argonne National Laboratory, Lemont, IL.

[6] Xue, M., Droegemeier, K., and Wong, V., 2000, "The Advanced Regional Prediction System (ARPS)—A Multi-Scale Nonhydrostatic Atmospheric Simulation and Prediction Model. Part I: Model Dynamics and Verification," *Meteorology and Atmospheric Physics*, 75(3–4), pp. 161–193.

[7] Xue, M., Droegemeier, K., Wong, V., Shapiro, A., Brewster, K., Carr, F., Weber, D., Liu, Y., and Wang, D., 2001, "The Advanced Regional Prediction System (ARPS)—A Multi-Scale Nonhydrostatic Atmospheric Simulation and Prediction Model. Part II: Model Physics and Applications," *Meteorology and Atmospheric Physics*, 76(3–4), pp. 143–165.

[8] Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Wang, W., and Powers, J. G., 2005, "A Description of the Advanced Research WRF Version 2," Technical Note No. 468+STR, National Center for Atmospheric Research, Boulder, CO.

[9] Kanamitsu, M., 1989, "Description of the NMC Global Data Assimilation and Forecast System," *Weather and Forecasting*, 4(3), pp. 335–342.

[10] Bludszuweit, H., Domínguez-Navarro, J. A., and Llombart, A., 2008, "Statistical Analysis of Wind Power Forecast Error," *IEEE Transactions on Power Systems*, 23(3), pp. 983–991.

[11] Hodge, B.-M., and Milligan, M., 2011, "Wind Power Forecasting Error Distributions Over Multiple Timescales," *IEEE Power & Energy Society General Meeting*, Detroit, MI.

[12] Zhang, Z., Sun, Y., Gao, D., Lin, J., and Cheng, L., 2013, "A Versatile Probability Distribution Model for Wind Power Forecast Errors and Its Application in Economic Dispatch," *IEEE Transactions on Power Systems*, 28(3), pp. 3,114–3,125.

[13] Zhang, J., Hodge, B.-M., and Florita, A., 2013, "Investigating the Correlation Between Wind and Solar Power Forecast Errors in the Western Interconnection," *ASME 7th International Conference on Energy Sustainability*, Minneapolis, MN.

[14] Hodge, B.-M., Orwig, K., and Milligan, M., 2012, "Examining Information Entropy Approaches as Wind Power Forecasting Performance Metrics," *12th International Conference on Probabilistic Methods Applied to Power Systems*, Istanbul, Turkey.

[15] Bessa, R. J., Miranda, V., Botterud, A., and Wang, J., 2011, "'Good' or 'Bad' Wind Power Forecasts: A Relative Concept," *Wind Energy*, 14, pp. 625–636.

[16] Ferreira, C., Gama, J., Matias, L., Botterud, A., and Wang, J., 2010, "A Survey on Wind Power Ramp Forecasting," Technical Report No. ANL/DIS-10-13, Argonne National Laboratory, Lemont, IL.

[17] Greaves, B., Collins, J., Parkes, J., and Tindal, A., 2009, "Forecasting Ramps of Wind Power Production with Numerical Weather Prediction Ensembles," *Wind Engineering*, 33(4), pp. 309–320.

[18] Florita, A., Hodge, B.-M., and Orwig, K., 2013, "Identifying Wind and Solar Ramping Events," *IEEE 5th Green Technologies Conference*, Denver, CO.

[19] Zheng, H. and Kusiak, A., 2009, "Prediction of Wind Farm Power Ramp Rates: A Data-Mining Approach," *Journal of Solar Energy Engineering*, 131, pp. 031011.

[20] Cutler, N., Kay, M., Jacka, K., and Nielsen, T. S., 2007, "Detecting, Categorizing and Forecasting Large Ramps in Wind Farm Power Output Using Meteorological Observations and WPPT," *Wind Energy*, 10, pp. 453–470.

[21] Bossavy, A., Girard, R., and Kariniotakis, G., 2013, "Forecasting Ramps of Wind Power Production with Numerical Weather Prediction Ensembles," *Wind Energy*, 16, pp. 51–63.

[22] Ela, E., and Kirby, B., 2008, "ERCOT event on February 26, 2008: lessons learned," Technical Report No. NREL/TP-5500-43373, National Renewable Energy Laboratory, Golden, CO.

[23] Potter, C. W., Grimit, E., and Nijssen, B., 2009, "Potential Benefits of a Dedicated Probabilistic Rapid Ramp Event Forecast Tool," *IEEE Power Systems Conference and Exposition*, Seattle, WA.

[24] AWS Truewind, 2008, "AWS Truewind's final report for the alberta forecasting pilot project," Technical Report, Alberta Electric System Operator, Alberta, Canada.

[25] Freedman, J., Markus, M., and Penc, R., 2008, "Analysis of West Texas Wind Plant Ramp-Up and Ramp-Down Events," Technical Report, AWS Truewind, Albany, NY, pp. 250–278.

[26] Orwig, K., Clark, C., Cline, J., Benjamin, S., Wilczak, J., Marquis, M., Finley, C., Stern, A., and Freedman, J., 2012, "Enhanced Short-Term Wind Power Forecasting and Value to Grid Operations," *11th Annual Int. Workshop on Large-Scale Integration of Wind Power into Power Systems*, Lisbon, Portugal.

[27] Zack, J. W., Manobianco, J., Natenberg, E. J., Young, S. H., Melino, T., and Freedman, J., 2013, "Analysis of the Forecast Sensitivity and Predictability of Wind Ramp Events During the Field Campaign of the Southern Study Region of the Wind

Forecasting Improvement Project (WFIP)," *4th Conference on Weather, Climate and the New Energy Economy*, Austin TX.

[28] Zhang, J., Hodge, B.-M., and Florita, A., 2014, "Joint Probability Distribution and Correlation Analysis of Wind and Solar Power Forecast Errors in the Western Interconnection," *Journal of Energy Engineering*, DOI: 10.1061/(ASCE)EY.1943-7897.0000189.

[29] Zhang, J., Chowdhury, S., Messac, A., and Castillo, L., 2013, "A Multivariate and Multimodal Wind Distribution Model," *Renewable Energy*, 51, pp. 436–447.

[30] Juban, J., Siebert, N., and Kariniotakis, G. N., 2007, "Probabilistic Short-Term Wind Power Forecasting for the Optimal Management of Wind Generation," *IEEE Lausanne Powertech*, Lausanne, Switzerland, pp. 683–688.

[31] Roebber, P. J., 2009, "Visualizing Multiple Measures of Forecast Quality," *Weather and Forecasting*, 24, pp. 601–608.