

Received September 7, 2020, accepted September 20, 2020, date of publication September 25, 2020, date of current version October 7, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3026864

Forecasting of Wind Capacity Ramp Events Using Typical Event Clustering Identification

JIANG LI¹, (Member, IEEE), TIANYU SONG¹, BO LIU¹, HAOTIAN MA¹, JIKAI CHEN¹, (Member, IEEE), AND YUJIAN CHENG²

¹School of Electrical Engineering, Northeast Electric Power University, Jilin City 132012, China

²Nanjing Nanrui Jibao Electric Company, Ltd., Nanjing 021000, China

Corresponding authors: Bo Liu (liubo973@163.com) and Jiang Li (lijiang@neepu.edu.cn)

This work was supported in part by the National Nature Science Foundation of China under Grant 51977030.

ABSTRACT With the large-scale integration of wind generation into the power grid, violent wind speed fluctuation will cause wind power ramp events that can affect the safe and stable operation of power systems. In this article, a forecasting method for day-ahead ramp events is proposed based on wind speed event definition and profile analysis. Firstly, event-based K-means (EB-K) clustering is used to preprocess historical wind speed. Typical event indexes, such as change rate, amplitude, and time intervals are then extensively used to describe ramp event characteristics and decrease the computational burden for the following event identification within given intervals. Then, the similarity of wind power event set is obtained through empirical probability estimation of successive history ramp events. Typical event clustering identification (TECI) algorithm based on EB-K clustering, wind capacity events, and event cluster profiles is proposed to search the maximum occurrence probability for historical data with the similarity indicator. Finally, a case study on a practical farm in Hebei, China is used to verify the effectiveness and accuracy of wind capacity ramp event forecasting by using TECI.

INDEX TERMS Ramp events, event clustering identification, wind power uncertainty, event forecasting.

NOMENCLATURE

A. ACRONYMS

ARIMA	Autoregressive Integrated Moving Average	M	Number of occurring ramp events within the next 2 hours
ASD	Atomic Sparse Decomposition	N	Number of similar forecast events patterns corresponding to the key mode of the k -th cluster
BP	Back Propagation Neural Network	K	Number of clusters
DRE	Down-Ramp Event	v	Wind speed
EB-K	Event-Based K-means	α	Value of weight
EPDF	Empirical Probability Distribution Function	N_k	Number of wind capacity events in the K -th cluster
MAPE	Mean absolute percentage error	$m/m + n$	Lower / upper limit of the time interval
OpSDA	Optimized Swinging Door Algorithm	v_{thr}	Value of threshold
TECI	Typical Event Clustering Identification	Δv	Difference of wind speed in a time interval
URE	Up-Ramp Event	ΔQ	Difference of wind capacity in a time interval

B. VARIABLES

n_k	Number of k -th sample data sets	e_i	A wind capacity ramp event
Δt	Time interval	β_w	Pitch angle of wind turbine
ρ_w	Air density	λ_w	Blade tip speed ratio
A	Wind turbine area	$Q_{re-(t_i+\Delta t_i)}$	Wind capacity from time t_i to $t_i + \Delta t_i$
S	Unit area	v_{re}^t	Wind speed at time t
C_p	Wind energy utilization coefficient	$\{V_{re-set-p}\}$	Wind capacity event sets

The associate editor coordinating the review of this manuscript and approving it for publication was Shouguang Wang¹.

$\{W_{re-set-q}\}$	Ramp event sets
f_i	Ramp amplitude of the sample event
R_i	Change rate of the sample event
ρ_{pq}	Pearson correlation coefficient
f_{re-p}	Ramp amplitude of the key wind capacity event pattern
R_{re-p}	Change rate of the key wind capacity event pattern
f_{re-q}	Ramp amplitude of the wind capacity event sequence
R_{re-q}	Change rate of the wind capacity event sequence
$EP_{re-set-n_k}$	Relation between history actual ramp event sets and forecasting ramp event sets
$P_{re-i}^{t_0}$	Occurrence probability of ramp event set i
$P_{re}^{t_0}$	Probability of a ramp event at time t_0
$Q_{re}(p.u.)$	Amplitude of the ramp event during the time interval $(m, m+n)$
$Q_{re}^j(p.u.)$	Amplitude of the ramp event in the j -th similar wind capacity event pattern
$Q(\%)$	Percentage of the total installed capacity for a wind farm

I. INTRODUCTION

Wind energy has achieved rapid development in recent years worldwide [1]. However, due to the intermittency of wind resources, large-scale wind farm power output can be severely affected by weather and operation constraints. Thus, ramp events are defined to describe dramatic wind power fluctuations in the short term [2], [3]. They may seriously threaten the security of power systems, particularly steam turbines, and generators which have strict ramp constraints [4], [5]. In February 2008, the active power output of the Texas Wind Farm in the US dropped from 170MW to 0 MW within 15 minutes. A total of 1150MW power was cut off from the grid, and the frequency dropped to 59.85Hz [6], [7]. Therefore, it is necessary to forecast uncertain ramp events with high accuracy, which can reduce the uncertainty of wind power [8] and facilitate the use of predictive control [9], [10].

Wind power ramp events caused by violent wind speed changes [11]–[13] can generally be divided into the up-ramp event (URE) and down-ramp event (DRE) [14], [15]. When a strong low-pressure system (or cyclone), low-altitude jet, thunderstorm, gust, or similar extreme meteorological event occurs, the corresponding power output of a wind farm would fluctuate sharply in a short period. A URE occurs when wind speed changes from weak to strong, or the speed on a wind turbine is higher than the cut-out wind speed [16]. Oppositely, a DRE occurs when the wind power drops suddenly in a short time or is withdrawn from the main grid [17], [18].

At present, many researchers have studied the evaluation and detection methods of wind power ramp events. Reference [19] proposes a risk indicator that can assess ramp events and risk assessment based on the Monte Carlo method.

A high-risk ramp event detection method is proposed in [20] based on the boundary area, providing a standard for ramp event risk analysis to determine the boundary of the wind power security region. In [21], an optimized swinging door algorithm (OpSDA) is proposed to detect ramp events by a dynamic programming algorithm, which has advantages in calculating cost and performance.

The forecasting of ramp events could be generally divided into two categories: indirect forecasting and direct forecasting. For the former, wind power is usually forecasted according to traditional point-based forecasting methods, and the ramp events forecasting is achieved indirectly by defining and detecting from discrete forecasting power points. Reference [22] investigates wind power in different situations by building a multivariate time sequence model, in which the ramp change rate is used to indirectly forecast ramp events. In [17], an *Autoregressive Integrated Moving Average* (ARIMA) algorithm is proposed to forecast wind power ramp events through calculating wind speed and converting wind speed into power. Reference [23] adopts a wind ramp forecasting method based on empirical mode decomposition to improve forecasting accuracy, which does not distinguish between upward and downward ramp events. For the latter, ramp events are generally extracted by using both historical data and forecasted wind speed, and the forecasting of wind power is not necessary in this case. In [24], the optimized revolving door algorithm is proposed to improve forecasting accuracy. In [25], a data mining algorithm is proposed to classify and forecast ramp events. In [18], an extreme value analysis method is proposed to forecast ramp events, depending on data stability. Furthermore, a combined forecasting model based on the atomic sparse decomposition (ASD) and back propagation (BP) neural networks are proposed in [26] to avoid the impact of historical data instability on results.

Most existing research relies on deterministic analysis methods, while the uncertainties of ramp events are ignored. The probability analysis methods are powerful for dealing with uncertainty problems. Reference [27] proposes a probabilistic forecasting method to analyze characteristic parameters. Reference [28] adopts a data-driven probability inequality model to estimate the upper boundary of ramp events with a given probability. In the aforementioned work, most researchers focus on the detection and indirect forecasting of ramp events, viewed as point-based or indirect forecasting methods by using necessary power data, such as combined methods, sampling training, and interval analysis.

To guarantee forecasting accuracy for ramp events, reduce the computational burden, this article proposes a new typical event clustering identification (TECI) algorithm in event sets for the sequence of wind capacity ramp events, which is different from the point-based forecasting and single event analysis methods. The main contributions of this article are:

(i) Event-based K-means (EB-K) clustering algorithm can cluster ramp events by using empirical probability estimation and key pattern search, thus reducing the computational burden caused by the point-based iteration strategy.

(ii) Building event sets can effectively avoid the complexity of time series. According to the key event sets and event characteristics, the problems caused by the change of event correlation and time interval are solved.

The remainder of this article is organized as follows. The definition and statistical characteristics of ramp events are given in Section II. Typical event clustering identification includes history data processing, the key pattern of clustering, clustering number determination are proposed in Section III. The ramp events forecasting algorithm based on TECI is described by using the empirical probability estimation of event sequence in Section IV. The effectiveness of the proposed forecasting method is verified by using case studies in Section V. Section VI draws relevant conclusions.

II. DEFINITION AND STATISTICAL CHARACTERISTICS OF RAMP EVENTS

A. DEFINITION OF WIND CAPACITY RAMP EVENTS

Wind speed is the core parameter that determines the power output of a wind farm. The relationship between wind speed and wind power is shown in (1) [29].

$$P_m = \frac{1}{2} \rho_w A v^3 C_p(\lambda_w, \beta_w) \quad (1)$$

where ρ_w is the air density, A is the wind turbine area, v is wind speed, C_p is the wind energy utilization coefficient, which is related to pitch angle β_w and blade tip speed ratio λ_w .

The characteristic curve C_p has been described in detail in [30]. However, wind power is not always equal to (1) because of economic dispatching, wind curtailment, or cut-off operation. Thus, wind speed is the fundamental variable to directly define a ramp event in this article. The wind capacity can be obtained according to air volume at wind speed v , as shown in (2):

$$Q_{re-(t_i+\Delta t_i)} = \int_{t_i}^{t_i+\Delta t_i} S v(t) dt \quad (2)$$

where, $Q_{re-(t_i+\Delta t_i)}$ is wind capacity (m^3) from time t_i to $t_i + \Delta t_i$, $v(t)$ is wind speed (m/s), and S is the unit area (m^2).

A large wind power change in a short period is defined as a wind capacity ramp event or an energy event within the given interval, which is different from the wind power ramp event defined by instantaneous power. In general, if wind power difference in a short time interval changes significantly (say, 30% capacity in 10 minutes or 15 minutes), it is considered that a ramp event that affects the safety and stability of the power grid has occurred [31]. In this article, a wind capacity ramp event is considered to occur under similar conditions, as shown in Fig. 1.

Fig. 1(a) shows that a URE and a DRE based on wind speed occur at time t_1 and t_2 , respectively. Δv_1 is the wind speed change of the URE during the period $t_1 \sim t_1 + \Delta t$, and Δv_2 is the wind speed change of the DRE during the period $t_2 \sim t_2 + \Delta t$ [32]. In Fig. 1(b), the ordinate Q is wind capacity calculated by equation (2), it shows that a URE and

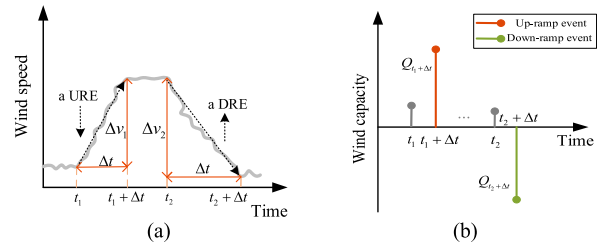


FIGURE 1. Definition of ramp event: (a) typical wind speed; (b) wind capacity ramp events.

a DRE based on wind capacity occur during $t_1 \sim t_2 + \Delta t$. Δt is the time interval of ramp events, $Q_{t_1+\Delta t}$ is the wind capacity of the URE at time t_1 , and $Q_{t_2+\Delta t}$ is the wind capacity of the DRE at time t_2 . According to [33], the definition of wind power ramp events in this article is shown as follows [34], [35]:

Definition: A wind capacity ramp event e_i can be defined as $e_i = \{t_i, \Delta t, v_t, \Delta v, \Delta Q\}$. Δt is the time interval, v_t is the wind speed at the time $t - \Delta t$, Δv is the difference of wind speed in a fixed time interval, ΔQ is the difference of wind capacity within the fixed time interval. When the absolute value of the difference between the start and end wind speeds within the fixed time interval Δt is bigger than the given threshold v_{thr} , a ramp event e_i occurs, which can be expressed as

$$\frac{|v_{re}^{(t+\Delta t)} - v_{re}^t|}{\Delta t} > v_{thr}, \quad \begin{cases} \text{if } v_{re}^{(t+\Delta t)} - v_{re}^t > 0, & \text{URE} \\ \text{if } v_{re}^{(t+\Delta t)} - v_{re}^t < 0, & \text{DRE} \end{cases} \quad (3)$$

where $v_{re}^{(t+\Delta t)}$ is the wind speed at the time $t + \Delta t$, v_{re}^t is the wind speed at time t , v_{thr} is the threshold value, Δt is the time duration of the ramp event whose reference value is given in [36]. Considering the sampling time of the case in this article, Δt is selected to be 15min. v_{thr} is the corresponding wind speed under which the energy reaches 30% of the total rated energy within the time interval (15min) [37].

B. STATISTICAL CHARACTERISTICS OF RAMP EVENTS

This article uses the ramp event definition in (3) to analyze the wind speed and wind capacity of a wind farm in Hebei, China in October 2017, where the recorded data is shown in Fig. 2.

Fig. 2 (a) is the actual wind speed within 48h through October 4~5, 2017. According to the definition, the actual wind capacity ramp event statistics in Fig. 2(b) is taken as an example. In Fig. 2, the time interval Δt is 15min and v_{thr} is the relevant threshold speed for 30% of total energy. Each red line indicates a URE, and each green line indicates a DRE. Fig. 2(b) can directly demonstrate the characteristics of wind capacity ramp events by using $Q_{re-(t_i+\Delta t_i)}$, which change in proportion to wind speed.

III. TYPICAL EVENT CLUSTERING IDENTIFICATION

This section preprocesses a large amount of historical data, analyzes key patterns of ramp events, and proposes a

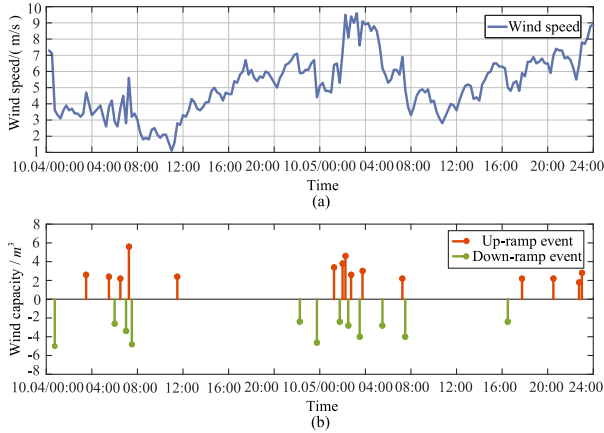


FIGURE 2. Ramp event characteristics: (a) actual wind speed (15min); (b) ramp events statistics (15min).

clustering number determination method to effectively reduce the computational burden.

A. PREPROCESSING HISTORY DATA

The forecasting of wind capacity ramp event relies on time sequence analysis of historical wind speed data. Firstly, the historical data is normalized to eliminate dimensions and build connections under different conditions. This article uses the z-score normalization to include the average and standard deviation for defined events in (3), as shown in (4):

$$z = \frac{h - \mu}{\sigma} \tag{4}$$

where, h represents historical data, such as wind speed, μ is the average value, and σ is the standard deviation. By implementing normalization, bad data can be taken out and removed, and the improved wind capacity event sets $\{V_{re-set-p}\}$ can be rebuilt to decrease forecasting errors.

B. KEY PATTERNS OF CLUSTERING

To reduce the calculation burden of historical data processing, key patterns can be generated by clustering algorithms [38], and the event-based K-means (EB-K) clustering algorithm is used to classify all historical events [39]. EB-K clustering algorithm uses an iterative strategy to divide the clustering features of ramp events into subsets of different categories.

The main concept of EB-K is to divide the sample event sets into subsets with different categories by using the iterative strategy, and the mean of each subset is viewed as the cluster center of the sample subset. The proposed algorithm aims to optimize the criterion equation for evaluating the clustering performance so that the clustering results can be compact and independent [40].

Algorithm 1:

- a) Use k to form a sample event matrix by equation (3).
- b) Randomly select k objects in all samples to achieve the initial clustering pattern.

c) Divide the sample events into k clusters using the smallest distance in equation (5) as the index.

$$\sum_{\bar{j} \in \{1,2,\dots,k\}} \min \|\theta_i - p_j\|^2$$

$$\theta_i = \alpha f_i + (1 - \alpha) R_i \tag{5}$$

where θ_i is the sample event vector, f_i is the ramp amplitude of the sample event, R_i is the change rate, α is the weight value of ramp amplitude. When α is close to 0, the matching degree of event amplitude is ignored. When α is close to 0.5, the matching degree is still low [41]. Through the training from historical data and experience, To balance between computation burden and matching degree, $\alpha = 0.85$ is reasonable by the training from historical data and experience. p_j is the Centroid, the subscript i is the sample number, and subscript \bar{j} is the cluster center number.

d) Calculate the mean of all events $\bar{\theta}_i$ for clusters G_i in equation (6), which is the new cluster center.

$$\bar{\theta}_i = \frac{1}{|G_i|} \sum_{x_i \in G_i} \theta_i \tag{6}$$

e) Finally, obtain the square error criterion equation in equation (7) is by using step (c) and step (d).

$$SSE = \sum_{i=1}^k \sum_{\theta_i \in G_i} |\theta_i - \bar{\theta}_i|^2 \tag{7}$$

Algorithm 1 is based on the EB-K clustering, which is used to generate a key pattern of wind capacity ramp events. The Pearson correlation coefficient of wind capacity events can be expressed as [42] (8), as shown at the bottom of the next page, where, ρ_{pq} is the Pearson correlation coefficient, defined as the quotient of the covariance and standard deviation, and \hat{N} is the size of the rolling window. $v_{re-p}(\hat{j})$ and $w_{re-q}(\hat{j})$ are comparison samples indexed by \hat{j} , which are the feature quantity of the key wind capacity event pattern and the feature quantity of the wind capacity event sequence. f_{re-p} is the ramp amplitude of the critical wind capacity event pattern, R_{re-p} is the change rate of the key wind capacity event pattern, f_{re-q} is the ramp amplitude of the wind capacity event sequence, R_{re-q} is the change rate of the wind capacity event sequence, and α is weight value, assumed to be 0.85. $\bar{v}_{re-p}(\hat{j})$ and $\bar{w}_{re-q}(\hat{j})$ are the sample averages of $v_{re-p}(\hat{j})$ and $w_{re-q}(\hat{j})$. In (8), $\rho_{pq} \in [-1, 1]$ and the larger the absolute value of ρ_{pq} , the higher the correlation between the two variables. The evaluation criteria of the Pearson correlation coefficient are shown in Table 1 [43].

The key wind capacity event pattern obtained by K-means clustering iteration via *Algorithm 1* should meet (9) and (10).

$$d = 1 - \rho_{pq} \tag{9}$$

$$K = \arg \min \sum_{k=1}^{K_0} \left(\frac{1}{N_k} \sum_{i=1}^{N_k} d_i^k \right) \tag{10}$$

TABLE 1. Correlation of pearson correlation coefficient.

Value range of $ \rho_{pq} $	Relevance
[0.8,1.0]	Extremely strong correlation
[0.6,0.8)	Strong correlation
[0.4,0.6)	Medium correlation
[0.2,0.4)	Weak correlation
[0.0,0.2)	Uncorrelation

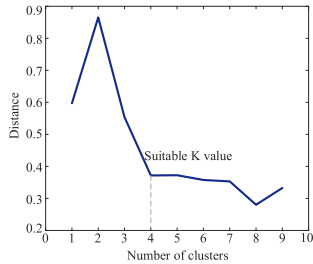


FIGURE 3. EB-K clustering number and distance.

where d_i^k is the Pearson correlation distance between the \hat{i} -th wind capacity event and the cluster center in cluster K , N_k is the number of wind capacity events in the K -th cluster, and K_0 is the total number of clusters.

In each cluster, the Pearson correlation distance between each wind capacity event and the unique key wind capacity event mode is summed and then the average distance is calculated. Similarly, the average distances of all clusters are calculated. In another way, the K calculated by the above equation (10) minimizes the final summation equation, then K is the most suitable cluster number.

C. CLUSTER NUMBER DETERMINATION

In equation (10), when K is smaller than the number of the key clusters, the aggregation degree of each cluster will increase as the cluster number k grows. However, when K reaches the key number of clusters, the degree of polymerization decreases rapidly and tends to be flat as cluster number k grows. When K is equal to 4 for history data from the given Hebei wind farm, the result of equation 10 is relatively smooth. When the cluster number is increased, the degree of polymerization tends to tiny decrease. That is, the cluster number 4 is the best cluster number to balance between computational burden and core parameters, shown in Fig. 3.

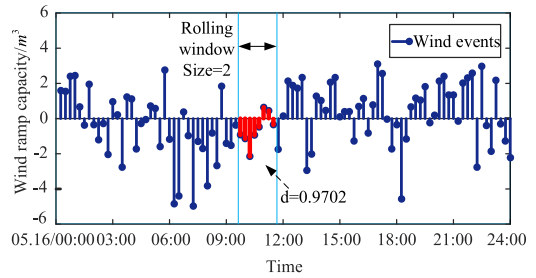


FIGURE 4. Rolling windows for event sets.

IV. EVENT FORECASTING METHOD USING TECI

A. SEARCHING FRAMEWORK IN HISTORY EVENT SETS

According to the above analysis in Section III, searching techniques can be used to extract and estimate key patterns of wind power ramp event sets. Ramp amplitude, ramp rate, and time interval of ramp events can all be used as the key characteristic during searching. Pearson correlation coefficient in (8) is used to measure the similarity between two ramp events.

The actual ramp events can be identified according to the wind capacity ramp event definition, and wind capacity sequence can be obtained from historical wind speed. A wind capacity forecasting pattern similar to a historical key wind capacity model (referred to as similar forecast events pattern) can be expressed according to wind capacity sequence when ramp event occurs. Considering the wind speed forecasting interval and the dispatching interval, a 2-hour time window is used to extract key patterns from the air wind capacity forecasting data, shown in Fig.4.

In Fig. 4, the rolling window is used to calculate the similarity between key event modes and wind capacity events pattern by using TECI. It iteratively searches the errors of key characteristics, such as ramp amplitude and ramp rate between history wind capacity event sets and forecast events. A similar forecast events pattern is obtained by using the search technique of rolling window event similarity. The Pearson correlation coefficient of a similar forecast pattern between the key event model is greater than the threshold. The threshold is set to 0.8 [44]. The schematic diagram is shown in Fig. 4, where the blue box is a similar forecast events pattern identified the Pearson correlation coefficient of 0.9702.

$$\rho_{pq} = \frac{\sum_{\hat{j}=1}^{\hat{N}} (v_{re-p}(\hat{j}) - \bar{v}_{re-p}(\hat{j})) (w_{re-q}(\hat{j}) - \bar{w}_{re-q}(\hat{j}))}{\sqrt{\sum_{\hat{j}=1}^{\hat{N}} (v_{re-p}(\hat{j}) - \bar{v}_{re-p}(\hat{j}))^2} \sqrt{\sum_{\hat{j}=1}^{\hat{N}} (w_{re-q}(\hat{j}) - \bar{w}_{re-q}(\hat{j}))^2}}$$

$$v_{re-p} = \alpha f_{re-p} + (1 - \alpha) R_{re-p}, \quad w_{re-q} = \alpha f_{re-q} + (1 - \alpha) R_{re-q} \tag{8}$$

Although ramp events can be represented by key characteristics, there is still no indicator to accurately capture the key modes of ramp event sets. Here, index equation I in (11) can be used to represent the correlation between historical ramp events and forecasting ramp events in event sets sequence.

$$I(m \leq \tilde{x} < m+n) = \begin{cases} 1, & m \leq \tilde{x} < m+n \\ 0, & \tilde{x} < m \\ 0, & \tilde{x} \geq m+n \end{cases} \quad (11)$$

where, m is the lower limit of the time interval, and $m+n$ is the upper limit of the time interval.

In [44], an optimization model is proposed to select a suitable window size for power forecasting. According to the updated weather forecast data period is 2h, so the size of the rolling time window is 2h and the probability resolution is 0.25h. In (11), $m = 0, 0.25, 0.5, \dots, 1.5, 1.75$ and $n = 0.25$ are candidate. When the minimum time interval between the actual ramp event and forecasting ramp event is less than m or more than $m+n$, the indicator equation is 0, which means forecasting events is irrelevant to historical ramp events. Otherwise, the indicator equation is 1, which means forecasting events is relevant to historical ramp events.

B. PROBABILITY FORECASTING OF RAMP EVENTS

This article applies an empirical probability distribution function (EPDF) based on a non-parametric model to estimate the uncertainty of the ramp event and reduce the error caused by the similarity searching algorithm. Through analyzing a large amount of historical data, the relation between history actual ramp event sets and forecasting ramp event sets can be defined as $EP_{re-set-n_k}$.

$$EP_{re-set-n_k}(m \leq \tilde{t}_i < m+n) = \frac{1}{n_k} \sum_{i=1}^{n_k} I(m \leq \tilde{t}_i < m+n) \quad (12)$$

where, n_k is the k -th sample data sets, and \tilde{t}_i is an independent variable, representing the minimum time interval.

Therefore, the occurrence probability of the i -th forecasting ramp event in the k -th cluster based on (12) can be expressed by equation (13)

$$p_{re,i}^{t_0} = EP_{re-set-n_k}(|\tilde{t}_i - t_0| \leq \tilde{x}_{t_i} < |\tilde{t}_i - t_0| + 0.25) \quad (13)$$

where, $p_{re,i}^{t_0}$ is the occurrence probability of ramp event set i , i.e., each ramp event in the rolling window can affect the occurrence probability of another ramp event at time t_0 .

The occurrence probability of a ramp event at time t_0 can be expressed by equation (14)

$$p_{re}^{t_0} = 1 - \prod_{i=1}^{\tilde{n}} (1 - p_{re,i}^{t_0}) \quad (14)$$

where t_0 is the time to analyze the occurrence probability of ramp event; and n is the total number of forecasted ramp events within the next 24 hours.

C. AMPLITUDE FORECASTING OF RAMP EVENTS

The forecasting method for ramp amplitude based on EPDF is proposed in [45] to quantitatively forecast ramp events. The similar forecasting patterns of wind capacity events are obtained from the sequence of the events. The j -th similar forecasting pattern of wind capacity events is in the k -th cluster. The event amplitude corresponding to the j -th pattern can be expressed based on the probability estimation method as

$$Q_{re}^j(p.u.) = \frac{\sum_{j=1}^N \rho_{re-set}^{t_j} \times Q_{re}(p.u.)}{\sum_{j=1}^N \rho_{re-set}^{t_j}} \quad |\tilde{t}_j - t_0| \leq \tilde{x}_{t_j} < |\tilde{t}_j - t_0| + 0.25 \quad (15)$$

where, $Q_{re}(p.u.)$ is the ramp event amplitude during the time interval $m < t < m+n$, $\rho_{re-set}^{t_j}$ is Pearson correlation between the searched wind capacity forecasting pattern and the key wind capacity event pattern, N is the number of similar forecast events pattern corresponding to the key mode of the k -th event cluster. $Q_{re}^j(p.u.)$ is the amplitude of the ramp event in the j -th similar wind capacity event pattern.

Using equation (15), the event amplitude $Q_{re}^{t_0}(p.u.)$ (16) at time t_0 in the rolling window can be obtained as

$$Q_{re}^{t_0}(p.u.) = \frac{\sum_{j=1}^M \rho_{re-set}^{t_j} \times Q_{re}^j(p.u.)}{\sum_{j=1}^M \rho_{re-set}^{t_j}} \quad (16)$$

where t_0 is the estimation time, and M is the total number of occurring ramp events within the next 2 hours.

D. RAMP EVENT FORECASTING ALGORITHM

The scheme for the forecasting algorithm of wind capacity ramp events proposed in this article is shown in Fig. 5. It includes historical data preprocessing, event sequence definition, the key pattern of clustering, cluster number determination, similarity search, and empirical probability estimation. The algorithm can forecast ramp events probabilistically and improve the forecasting accuracy by extracting event amplitude, ramp rate, and time interval as event characteristics. The specific steps in the flowchart are formed as follows.

Algorithm 2:

a) Wind capacity event sets $\{V_{re-set-p}\}$ at the time sequence are defined by equation (2) and ramp event sets $\{W_{re-set-q}\}$ at the time sequence are detected by equation (3).

b) The cluster number k is decided by equation (9) and (10). According to *Algorithm 1*, the key patterns of wind ramp events are obtained.

c) Through the rolling window, search for the sequence of ramp events from time t_0 , and then the similarity and key patterns of ramp events are estimated in the rolling search window. Obtain a similar wind capacity event forecasting

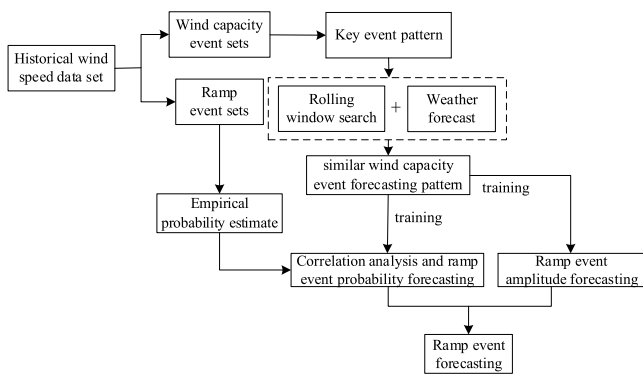


FIGURE 5. Ramp event forecasting scheme.

pattern by equation (8) with a Pearson coefficient bigger than the threshold 0.8.

d) The correlation $EP_{re-set-n_k}$ between historical and forecasting ramp events is estimated by using the empirical probability estimation equation (12).

e) The probability $p_{re-i}^{t_0}$ of a ramp event occurring in the i -th similar wind capacity event forecasting pattern is obtained by equation (13).

f) Finally, the occurrence probability $p_{re}^{t_0}$ of a ramp event is estimated and ranked by equation (14). A similar forecast events pattern is extracted by step (c). The event amplitude $Q_{re}^{t_0}(p.u.)$ can be obtained by equations (15) and (16).

V. CASE STUDY

In this article, the wind speed and wind power of a wind farm (21MW) in Hebei Province, China recorded in 2017-2018, are used to verify the effectiveness of the proposed algorithm in Section IV. The data set is divided into a training set (date in 2017) and a test set (date in 2018). The training set is used to define ramp events, extract key patterns, and empirical probability estimates. The test set is used to verify the effectiveness of TECI and forecasting methods.

A. PROBABILISTIC ANALYSIS OF ACTUAL RAMP EVENTS

Considering the power system dispatching period $\Delta t = 15$ minutes in equation (3), the time interval between ramp events is set as 15 minutes. The threshold v_{thr} is equal to or bigger than 30% of the maximum wind capacity value corresponding to the wind farm. $v_{thr(max)}$ is the upper limit threshold of the ramp event, and $v_{thr(min)}$ is the lower limit threshold of the ramp event.

Based on the statistical analysis on historical ramp events, a 2-hour rolling window is used to extract key patterns according to wind speed from January 1, 2017, to December 31, 2017. The wind speed and ramp events forecasting in October are shown in Fig. 6 (a) and Fig. 6 (b) respectively.

When the cluster number k is 8, the calculation burden may increase in the iterative process with a total calculation time 568.5265s. When the cluster number k is 4, the calculation time is 182.8045s and the decline rate of the Pearson correlation distance slows down significantly from Fig. 3. When the

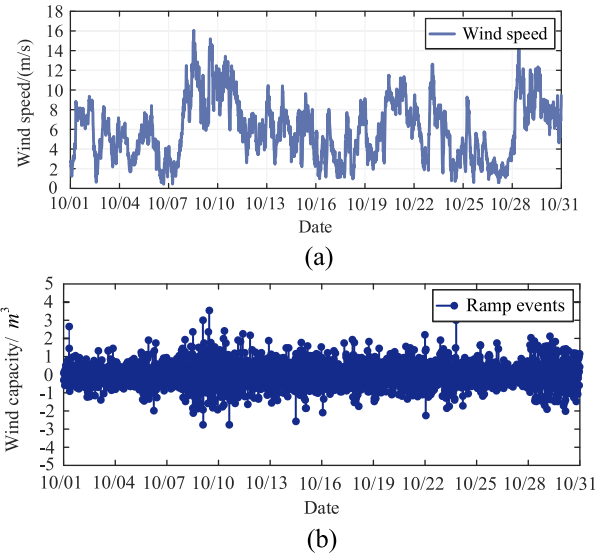


FIGURE 6. Wind speed and wind capacity events in October 2017 (a) wind speed; (b) wind capacity ramp events.

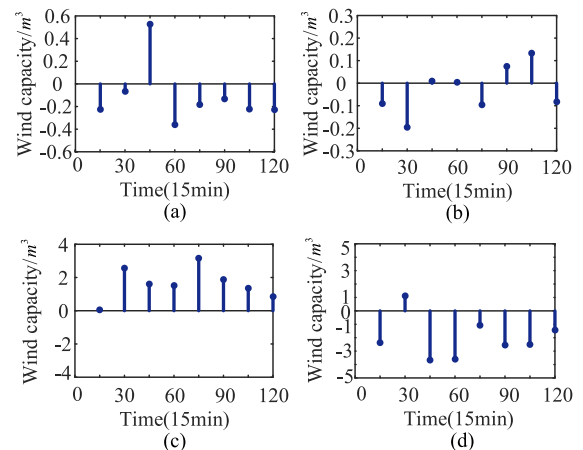


FIGURE 7. Key event pattern: (a) key mode 1;(b) key mode 2;(c) key mode 3;(d) key mode 4.

number of clusters is 50, it is the 12.29 times of the calculation time with 4 clusters. Thus, the EB-K clustering algorithm for sampling events can dramatically reduce the calculation burden by using a small cluster number. According to equation 10 and Fig. 3, 4 categories can balance between calculation burden and forecasting accuracy. Then, each cluster center is selected as a key pattern by using the similarity search technology. Thus, four key patterns are generated, as shown in Fig. 7.

According to the empirical probability estimation equation (12), the correlation of the occurrence probability in ramp event set $EP_{re-set-n_k}$ between history ramp events and forecasting wind capacity events can be estimated. It is then used to search wind ramp events of 2018. Firstly, the occurrence probability $p_{re-i}^{t_0}$ of the i -th forecasted ramp event set and the occurrence probability $p_{re}^{t_0}$ of a ramp event at time t_0 can be determined by using equations (13) and (14) respectively.

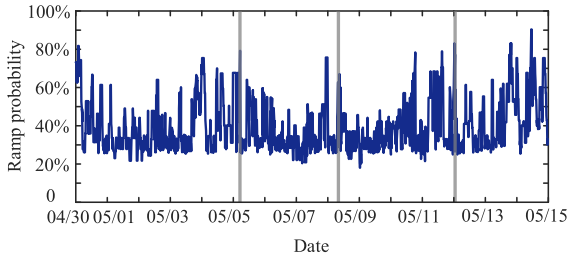


FIGURE 8. The occurrence probability of ramp events from April 30 to May 15, 2018.

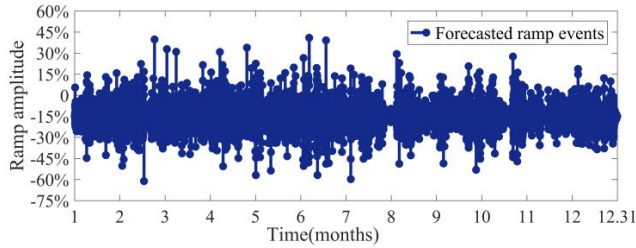


FIGURE 9. Forecasted ramp events results (2018).

Then, the wind capacity $Q_{re}^j(p.u.)$ of the ramp event can be obtained from equation (15). Finally, wind capacity $Q_{re}^{t_0}(p.u.)$ corresponding to the ramp event at t_0 can be found by using equation (16). In Fig. 7, the key mode 1 has a large URE, and the wind capacity of DRE is relatively low. In key mode 2, the power does not change after the DRE. In key mode 3, there are UREs, and in key mode 4, there are more DREs but the amplitude difference is small.

B. ACCURACY ANALYSIS OF FORECASTED RAMP EVENTS

According to the proposed TECI in Section?, this article selects 15-day probability forecasting results for display, shown in Fig. 8.

In Fig. 8, the blue line indicates the probability of the ramp event occurrence and the grey shaded part represents the time effective interval limited to $\pm 20\%$ error.

The ramp amplitude of forecasted wind power ramp events is shown in Fig. 9. To demonstrate the occurrence probability of ramp events, the ordinates of Fig. 9 and Fig. 10 use the ramp amplitude Q (%) at time t_0 as an indicator for ramp events, which is a percentage of the total installed capacity p_c for a wind farm, given in equation (17)

$$Q(\%) = \frac{Q_{re}^{t_0}(p.u.)}{p_c} \times 100\% \quad (17)$$

By using TECI, the forecasted ramp events are shown in Fig. 9. To compare the deviations between actual events and forecasted events, the actual ramp events from May 4 to 11, 2018 are taken as an example in Fig. 10. The blue line represents actual ramp events, the red line represents the forecasted ramp events obtained from proposed Algorithm 2, and the green line represents the ramp event using the traditional

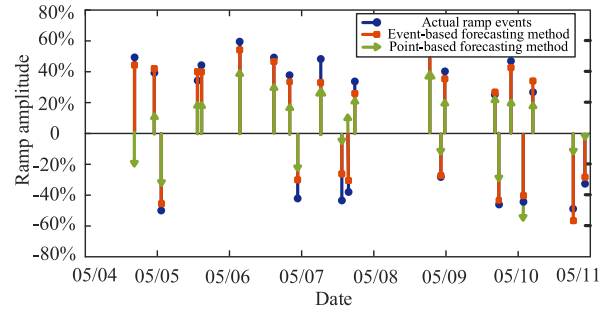


FIGURE 10. Comparison of actual and forecasted ramp events from May 4 to May 11, 2018.

TABLE 2. Forecasted events by using the event-based forecasting method and point-based forecasting method.

Event indicators	Event-based forecasting method	Point-based forecasting method
Total event times	312	312
Right times	253	162
False times	59	150
Invalid alarm times	173	263

point-based method. As seen, the maximum ramp amplitude of URE is 12.88MW, reaching 61.36% of the total installed capacity.

In Fig. 10, the forecasting accuracy of the ramp event-based forecasting method proposed in this article is significantly higher than that of is point-based forecasting method. The forecasting amplitude in Fig. 10 is generally lower than that of real ramp events. The main reason is that the lower average magnitude values during forecasting can get a better expectation value with a low deviation for event-based or point-based forecasting methods. Compared with the point-based forecasting method, the event-based forecasting method can result in the lower deviation values because of the accumulation effect of the uncertain data within the given time interval.

In Table 2, the forecasted results of ramp events by traditional point-based forecasting methods from January 1 to December 31, 2018, are compared with the forecasted results obtained by Algorithm 2. In Table 2, the *Right times* are defined as the number of events when ramp events occurrence time and amplitude are accurately forecasted. When the forecasted ramp event amplitude does not exceed $\pm 20\%$ the actual amplitude, it is assumed that the amplitude prediction is accurate. The *False times* represent the number of events that the amplitude prediction is wrong and the actual ramp event is not forecasted. The *Invalid alarm times* represent the number of events that the ramp event has not occurred but is forecasted wrongly.

To quantitatively evaluate the forecasting error of ramp events during 2018, this article uses mean absolute percentage error (MAPE), as shown in Table 3.

From Table 2, 312 ramp events occurred in 2018. The proposed method accurately forecast 253 ramp events, while the traditional point-based method can only forecast 162 ramp

TABLE 3. Forecasted amplitude results by using the event-based forecasting method and point-based forecasting method.

Months (2018)	Event-based forecasting method	Point-based forecasting method
	MAPE%	MAPE%
January	9.20	18.94
February	8.57	17.90
March	8.22	17.65
April	10.35	17.53
May	11.38	18.45
June	9.71	16.08
July	10.08	17.92
August	10.26	16.25
September	8.36	16.33
October	12.56	18.62
November	9.21	17.30
December	8.85	18.85
Average	9.73	17.65

events, leading to 150 forecasting false. In Table 2, the event-based forecasting method accuracy is 81.08% by using *Algorithm 2*, but the traditional point-based forecasting method is only 52%. From Table 3, an amplitude forecasting perspective, the average MAPE of the proposed event-based forecasting method is 9.73% and it is 17.65% for the traditional point-based forecasting method. Thus, the event-based forecasting method is better than the accuracy of the point-based forecasting method in the forecasting of ramp amplitude.

VI. CONCLUSION

This article proposes a TECI algorithm to forecast the probability and magnitude of wind capacity ramp events. Through an extensive case study, the following conclusions can be drawn:

- (i) The proposed event indicators can provide quantitative evaluation and analysis methods for discrete probability characteristics within the given time interval. Point-based forecasting methods are converted into event-based forecasting methods by wind capacity event definition and EB-K clustering algorithm, which can realize the energy event balance within a time interval in the future.
- (ii) Compared with the conventional scenario analysis methods, the TECI algorithm uses typical clustering and event set indicators to decrease the computational burden. Cluster number and rolling window scale can be preprocessed in TECI, which makes it possible to avoid the scenario dimension disaster problem in multiple scenarios.

In the future, the proposed wind capacity ramp event forecasting method can provide theoretical support for power system predictive control and uncertainty event predictive control from the point of energy balance.

REFERENCES

- [1] A. Couto, P. Costa, L. Rodrigues, V. V. Lopes, and A. Estanqueiro, "Impact of weather regimes on the wind power ramp forecast in portugal," *IEEE Trans. Sustain. Energy*, vol. 6, no. 3, pp. 934–942, Jul. 2015.
- [2] Y. Fujimoto, Y. Takahashi, and Y. Hayashi, "Alerting to rare large-scale ramp events in wind power generation," *IEEE Trans. Sustain. Energy*, vol. 10, no. 1, pp. 55–65, Jan. 2019.
- [3] A. Bossavy, R. Girard, and G. Kariniotakis, "Forecasting uncertainty related to ramps of wind power production," in *Proc. Eur. Wind Energy Conf. Exhib. (EWEC)*, vol. 2, Apr. 2010, pp. 1–9.
- [4] D. Zhang, H. Zhang, X. Zhang, X. Li, K. Ren, Y. Zhang, and Y. Guo, "Research on AGC performance during wind power ramping based on deep reinforcement learning," *IEEE Access*, vol. 8, pp. 107409–107418, Jun. 2020.
- [5] M. Cui, D. Ke, Y. Sun, D. Gan, J. Zhang, and B.-M. Hodge, "Wind power ramp event forecasting using a stochastic scenario generation method," *IEEE Trans. Sustain. Energy*, vol. 6, no. 2, pp. 422–433, Apr. 2015.
- [6] R. Sevlian and R. Rajagopal, "Detection and statistics of wind power ramps," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 3610–3620, Nov. 2013.
- [7] M. Cui, C. Feng, Z. Wang, and J. Zhang, "Statistical representation of wind power ramps using a generalized Gaussian mixture model," *IEEE Trans. Sustain. Energy*, vol. 9, no. 1, pp. 261–272, Jan. 2018.
- [8] Y. Gong, Q. Jiang, and R. Baldick, "Ramp event forecast based wind power ramp control with energy storage system," *IEEE Trans. Power Syst.*, vol. 31, no. 3, pp. 1831–1844, May 2016.
- [9] C. Li, H. Zhou, J. Li, and Z. Dong, "Economic dispatching strategy of distributed energy storage for deferring substation expansion in the distribution network with distributed generation and electric vehicle," *J. Cleaner Prod.*, vol. 253, Apr. 2020, Art. no. 119862.
- [10] Q. Wang, Z. Yu, R. Ye, Z. Lin, and Y. Tang, "An ordered curtailment strategy for offshore wind power under extreme weather conditions considering the resilience of the grid," *IEEE Access*, vol. 7, pp. 54824–54833, Apr. 2019.
- [11] G. Sideratos and N. D. Hatzigiorgiou, "Wind power forecasting focused on extreme power system events," *IEEE Trans. Sustain. Energy*, vol. 3, no. 3, pp. 445–454, Jul. 2012.
- [12] M. Lange, "On the uncertainty of wind power predictions—Analysis of the forecast accuracy and statistical distribution of errors," *J. Sol. Energy Eng.*, vol. 127, no. 2, pp. 177–184, May 2005.
- [13] Y. Zhang, S. Gao, J. Han, and M. Ban, "Wind speed prediction research considering wind speed ramp and residual distribution," *IEEE Access*, vol. 7, pp. 131873–131887, Sep. 2019.
- [14] B. Cao, L. Chang, X. Gong, J. L. C. Barrera, T. Levy, and R. Kilpatrick, "Probability forecasting of wind power ramp events using a time series similarity search algorithm," in *Proc. IEEE Energy Convers. Congr. Expo. (ECCE)*, Portland, OR, USA, Sep. 2018, pp. 972–976.
- [15] C. Kamath, "Understanding wind ramp events through analysis of historical data," in *Proc. IEEE PES Transmiss. Distrib. Conf. Expo.*, New Orleans, LA, USA, Apr. 2010, pp. 1–6.
- [16] M. Cui, J. Zhang, A. R. Florita, B.-M. Hodge, D. Ke, and Y. Sun, "An optimized swinging door algorithm for wind power ramp event detection," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Denver, CO, USA, Jul. 2015, pp. 1–5.
- [17] A. K. Nayak, K. C. Sharma, R. Bhakar, and J. Mathur, "ARIMA based statistical approach to predict wind power ramps," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Denver, CO, USA, Jul. 2015, pp. 1–5.
- [18] D. Ganger, J. Zhang, and V. Vittal, "Statistical characterization of wind power ramps via extreme value analysis," *IEEE Trans. Power Syst.*, vol. 29, no. 6, pp. 3118–3119, Nov. 2014.
- [19] H. Ma and Y. Liu, "Real-time recognition of wind power ramp events," in *Proc. 2nd IET Renew. Power Gener. Conf. (RPG)*, Beijing, China, 2013, pp. 1–4.
- [20] H. Wang, H. Gong, and H. Wang, "A region-based method for high-risk wind power ramp events detection," in *Proc. Int. Conf. Renew. Power Gener. (RPG)*, Beijing, China, 2015, pp. 1–5.
- [21] M. Cui, J. Zhang, A. R. Florita, B.-M. Hodge, D. Ke, and Y. Sun, "An optimized swinging door algorithm for identifying wind ramping events," *IEEE Trans. Sustain. Energy*, vol. 7, no. 1, pp. 150–162, Jan. 2016.
- [22] H. Zheng and A. Kusiak, "Prediction of wind farm power ramp rates: A data-mining approach," *J. Sol. Energy Eng.*, vol. 131, no. 3, pp. 031011–031018, Jul. 2009.
- [23] X. Qiu, Y. Ren, P. N. Suganthan, and G. A. J. Amaratunga, "Short-term wind power ramp forecasting with empirical mode decomposition based ensemble learning techniques," in *Proc. IEEE Symp. Ser. Comput. Intell. (SSCI)*, Honolulu, HI, USA, Nov. 2017, pp. 1–8.

[24] J. Zhang, M. Cui, B.-M. Hodge, A. Florita, and J. Freedman, "Ramp forecasting performance from improved short-term wind power forecasting over multiple spatial and temporal scales," *Energy*, vol. 122, pp. 528–541, Mar. 2017.

[25] H. Zareipour, D. Huang, and W. Rosehart, "Wind power ramp events classification and forecasting: A data mining approach," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Detroit, MI, USA, Jul. 2011, pp. 1–3.

[26] M. Cui, Y. Sun, and D. Ke, "Wind power ramp events forecasting based on atomic sparse decomposition and BP neural networks," *Autom. Electr. Power Syst.*, vol. 38, no. 12, pp. 6–11, Jun. 2014.

[27] J. Heckenbergerova, P. Musilek, and J. Marek, "Analysis of wind speed and power time series preceding wind ramp events," in *Proc. 15th Int. Sci. Conf. Electr. Power Eng. (EPE)*, Brno, Czech Republic, May 2014, pp. 279–283.

[28] Y. Cao, W. Wei, C. Wang, S. Mei, S. Huang, and X. Zhang, "Probabilistic estimation of wind power ramp events: A data-driven optimization approach," *IEEE Access*, vol. 7, pp. 23261–23269, Feb. 2019.

[29] J. Huo, "Dynamic equivalence and parameter identification for the doubly-fed wind farm," M.S. thesis, School Elect. Electron. Eng., North China Electr. Power Univ., Beijing, China, 2015.

[30] W. Huang and X. Zhen, "Feature analysis based equivalent modeling for large-scale wind farms," *Power Syst. Technol.*, vol. 37, no. 8, pp. 2271–2277, 2013.

[31] Y. Zhang, Y. Dai, X. Zhang, J. Zhang, Z. Wang, and L. Xue, "Review and prospect of research on wind power ramp events," *Power Syst. Technol.*, vol. 42, no. 6, pp. 1783–1792, Jun. 2018.

[32] J. Li, X. Ai, and J. Wen, "Constraints of wind power ramp event in robust unit commitment," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Boston, MA, USA, Jul. 2016, pp. 1–5.

[33] J. Li, G. Liu, and S. Zhang, "Smoothing ramp events in wind farm based on dynamic programming in energy Internet," *Energy*, vol. 12, no. 4, pp. 550–559, Dec. 2018.

[34] C. Gallego, A. Costa, and A. Cuerva, "Improving short-term forecasting during ramp events by means of regime-switching artificial neural networks," *Adv. Sci. Res.*, vol. 6, no. 1, pp. 55–58, Mar. 2011.

[35] T. Ouyang, X. Zha, L. Qin, and A. Kusiak, "Optimisation of time window size for wind power ramps prediction," *IET Renew. Power Gener.*, vol. 11, no. 8, pp. 1270–1277, Jun. 2017.

[36] C. W. Potter, E. Grimit, and B. Nijssen, "Potential benefits of a dedicated probabilistic rapid ramp event forecast tool," in *Proc. IEEE/PES Power Syst. Conf. Expo.*, Seattle, WA, USA, Mar. 2009, pp. 1–5.

[37] *Implementation Rules of Technical Regulations for Wind Farm Access to Power Grid*, State Grid Corp. China, Beijing, China, 2009.

[38] W. Liu, Y. Gong, G. Geng, and Q. Jiang, "Refined ramp event characterisation for wind power ramp control using energy storage system," *IET Renew. Power Gener.*, vol. 13, no. 10, pp. 1731–1740, Jul. 2019.

[39] E. Erdem, J. Shi, and Y. Peng, "Short-term forecasting of wind speed and power—A clustering approach," in *Proc. IIE Annu. Conf. Expo*, May 2014, pp. 1–8.

[40] Y. Li, C. Dai, T. Wang, Z. Zhou, S. Zhou, L. Cai, P. Musilek, and E. Lozowski, "Separate wind power and ramp predictions based on meteorological variables and clustering method," in *Proc. IEEE 6th Int. Conf. Power Syst. (ICPS)*, New Delhi, India, Mar. 2016, pp. 1–6.

[41] H. Yan, X. Jun, X. Ma, and D. Hui, "Wind power output schedule tracking control method of energy storage system based on ultra-short term wind power prediction," *Power Syst. Technol.*, vol. 39, pp. 432–439, Feb. 2015.

[42] X. Tang, Y. Dai, Q. Liu, X. Dang, and J. Xu, "Application of bidirectional recurrent neural network combined with deep belief network in short-term load forecasting," *IEEE Access*, vol. 7, pp. 160660–160670, Nov. 2019.

[43] Y. Xiao, Y. Zhao, and Z. Xu, "Topology checking method for low voltage distribution network based on improved pearson correlation coefficient," *Power Syst. Protection Control*, vol. 47, no. 11, pp. 37–43, 2019.

[44] T. Ouyang, X. Zha, L. Qin, Y. Xiong, and H. Huang, "Model of selecting prediction window in ramps forecasting," *Renew. Energy*, vol. 108, pp. 98–107, Aug. 2017.

[45] A. A. Farag, A. M. Ali, S. Elhajian, and A. A. Farag, "Probability density estimation by linear combinations of Gaussian kernels-generalizations and algorithmic evaluation," in *Proc. Int. Conf. Multimedia Technol.*, Hangzhou, China, Jul. 2011, pp. 6491–6494.



JIANG LI (Member, IEEE) received the B.S. degree from the Shanghai University of Electric Power, Shanghai, China, in 2003, the M.S. degree in electrical engineering from Northeast Electric Power University, Jilin City, China, in 2006, and the Ph.D. degree from North China Electric Power University, Beijing, China, in 2010.

He was with Cornell University, Ithaca, NY, USA, for one year, as a Visiting Scholar. He is currently a Professor with the School of Electrical Engineering, Northeast Electric Power University. His research interests include the analysis of uncertain events in power systems and intelligent control of distribution networks.



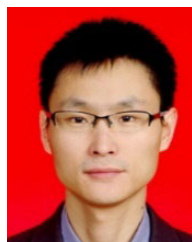
TIANYU SONG received the B.S. degree in electrical engineering from Northeast Electric Power University, Jilin City, China, in 2018, where she is currently pursuing the master's degree. Her research interest includes the analysis of distribution networks.



BO LIU was born in Heilongjiang, China, in 1985. She received the Ph.D. degree in control science and engineering from the Harbin Institute of Technology, Harbin, China, in 2016. She currently works with the School of Electrical Engineering, Northeast Electric Power University. She is a coauthor of nine publications (mainly in journals) in the field of motion control for ultrasonic motors. Her research interest includes the optimal operation of power systems.



HAOTIAN MA was born in Heilongjiang, China, in 1996. He received the B.S. degree from Heilongjiang University, Harbin, China, in 2018. He is currently pursuing the M.S. degree in electrical engineering with Northeast Electric Power University, Jilin City, China. His research interest includes modeling and analysis of the integrated electrical-thermal systems.



JIKAI CHEN (Member, IEEE) received the Ph.D. degree in electrical engineering from the Harbin Institute of Technology, Harbin, China, in 2011. He is currently an Associate Professor with the School of Electrical Engineering, Northeast Electric Power University. His research interests include HVDC control technologies, power quality, and analysis of power electronics dominated power systems.



YUJIAN CHENG received the B.S. degree in electrical engineering from Northeast Electric Power University, Jilin City, China, in 2018. He is currently pursuing the part-time master's degree with Nanjing Nanrui Jibao Electric Company, Ltd. His research interest includes power system control.

• • •