Mid-term Scenario Generation for Wind Power Using GAN with Temporal-correlation Enhancement Block

Yupu Ma[1,2,*], *Ming* Ma[3], *Ningbo* Wang[3], *and Ying* Qiao[1]

[1]State Key Lab of Control and Simulation of Power Systems and Generation Equipment (Dept. of Electrical Engineering, Tsinghua University), Haidian District, Beijing 100084, China
[2]School of Electrical Engineering, Northeast Electric Power University, Jilin 132012, Jilin Province, China
[3]Wind Power Technology Center of Gansu Electric Power Company, Lanzhou Province 730070, China

Abstract. Under the background of increasing renewable energy penetration and rigid demand for mid-term generation planning, the accurate expression of wind power uncertainty becomes more and more important. Firstly, this paper analyses the problem of scenario generation models based on traditional Generative Adversarial Networks(GAN), point that the fluctuation of scenarios that it generated usually deviates greatly from the real one. And further proposes a convolutional structure, that called Temporal-correlation Enhancement block (TE block), which can solve the aforementioned problem by enhance the temporal correlation perception ability of convolutional layers. Then, the problems of traditional conditional Wasserstein GAN-Gradient Penalty(WGAN-GP) in mid-scale scenario generation are discussed, and a conditional WGAN-GP scenario generation model that is suitable for mid-term scenario generation is presented. This model-free non-parametric model can generate a large number of realistic wind scenarios efficiently according to the given conditions. In order to validation, we use the data collected in a province in northern China to train the model in the Case Study part, and compare it with the traditional model. Result shows the fluctuation distribution deviation problem is improved obviously on the model with TE block, and in the comparison of auto-correlation coefficient, the proposed model also outperform the traditional model. This verifies the superiority of the proposed model in temporal expression ability compared with the traditional model, as well as the feasibility of the mid-term wind scenario generation.

1 Introduction

In recent years, the rapid development of installed capacity of renewable energy in China has brought great challenges to the accommodation of renewable energy. Wind power, as one of the most uncertain power sources, is more prominent. On the one hand, in China's new energy center, 72% of the units are large-scale thermal power units over 300 MW[1]. The minimum start-up and shutdown time of these units is long, and the start-up and shutdown cost is high, so the overall flexibility is insufficient. Therefore, it is necessary to formulate the start-stop plan of these units in the mid-term. On the other hand, in order to avoid the uncertainties brought by wind power, the guaranteed capacity of wind power is usually calculated according to the year-round statistics. [2] points out that when the guaranteed rate is 95%, the guaranteed capacity coefficient can even be reduced to less than 0.4%, which leads to wind power being almost ignored on the start-stop and maintenance plan, which will affect the wind accommodation seriously. Therefore, in order to solve the above contradictions, the uncertain modelling of wind power in the mid-term scale is becoming more and more important.

Scenario method is a commonly used method to express the uncertainty of wind power. To be more specific, it is to generate a series of wind scenarios according to the historical data distribution of wind power. Each scenario represents a possible hypothetical reality to realize the simulation of physical uncertainty. The generated scenarios can be used in different optimization models and prediction models. In [3], a new scenario generation method is proposed to deal with the uncertainty of wind power forecasting, and then the wind scenario is used into the stochastic security-constrained unit commitment (SCUC) model. In [4], neural networks are used to generate wind scenarios. Each possible scenario generated in this way contains wind power ramping information, and the distribution of specified Wind Power Ramp Event (WPRE) attributes can be obtained randomly according to the scenarios.

Current scenario generation methods mainly include parametric and model-based methods, such as [5], [6], and nonparametric method. These parametric methods are based on the assumption that wind power obeys a certain probability distribution, and it can express the temporal-spatial correlation of wind power to a certain extent. However, it is difficult to reproduce the high-order and high-nonlinear characteristics of wind power

---

* Corresponding author: 293596479@qq.com

by using these low-order models, which leads to inaccurate modeling of the spatial-temporal characteristics of real wind power distribution. Nonparametric methods do not require wind power to obey a distribution assumption. For example, in [7], radial basis function neural networks (RBF) algorithm is used to transform the generated scenarios into a series of quantiles to represent the uncertainty of wind power.

Recently, with the emergence of Deep Learning technology, more and more researches in the field of scenario generation has begun to focus on a powerful generative model, Generative Adversarial Networks (GAN)[8]. This is a novel unsupervised learning algorithm based on nonparametric method. The algorithm does not directly learn some explicit distribution, but generates synthetic data close to the real distribution of train dataset through the game of two neural networks. In [9] , for the first time, GAN with Wasserstein distance is introduced to generate renewable energy scenarios, which is a data-driven, model-free and non-parametric scenario generation method. It proves that GAN has high feasibility in expressing temporal-spatial correlation and uncertainty of wind power. In [10], the conditional WGAN-GP algorithm is introduced for the first time. The prediction error category is predicted by SVM with the wind power prediction curves, and the scenario with this kind of prediction error is generated.

The above scenario generation based on GAN algorithm can accurately simulate the distribution of wind power, but the convolutional layer is used as the feature extraction method in their models. This leads to the convolution-based GAN algorithm cannot accurately express the probability distribution of fluctuations in sequence generation. In this paper, we first introduce the Temporal-correlation Enhancement block (TE block) as a supplement to the inherent defects of convolutional layer. At the same time, the above scenario generation studies all limit the time scale to the short-term time scale which is predominant in the past, so this paper then analyses the problems that conditional WGAN-GP model may encounter in the mid-term scale, and gives a set of solutions. Finally, a series of case studies are given to demonstrate that the proposed TE block model improves the expression ability of temporal correlation in generating scenarios, and the specific effect of mid-term wind scenario generation is shown at last.

2 Temporal-correlation enhancement block

2.1 Problems in traditional GAN scenario generation structures

Previous scenario generation based on GAN, including WGAN model and WGAN-GP model, whose discriminators are all based on convolutional neural network. Convolutional network structure is more suitable for feature extraction with translational invariance. However, wind power curve is a kind of time series. The operation of moving summation in

convolutional layers will destroy the correlation between adjacent elements. From the result of the final generated curve, it can be seen that the generated curve and the original curve are in good agreement with the probability distribution of the wind power, but the fluctuation distribution of the correlation which reacts sequential correlation will deviate greatly. This inherent defect of convolutional networks will make the generated curves more fluctuating than the actual ones, which will have a negative impact on some occasions where the uncertainty of fluctuation needs to be accurately modeled. And when it comes to the Mid-term scale, this phenomenon will be more obvious with the increase of time scale and the fluctuation rate.

2.2 Explanation of the TE block

Since the moving summation of the traditional convolutional layer will destroy the temporal correlation, the temporal correlation is artificially reconstructed by the Temporal-correlation Enhancement (TE) block using different time scale differences in this section.
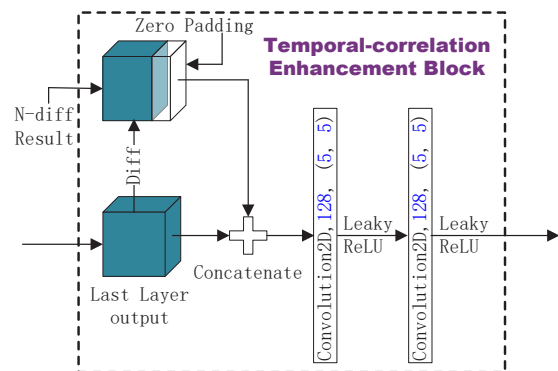

Fig. 1. Schematic of The Structure of TE Block.

The schematic of the proposed TE block is shown in Fig. 1. The TE block still adopts the convolutional structure, so its input is also the feature maps $F_{w \times h \times c}$ of the whole batch output of the last layer.

Firstly, the TE block reshapes the feature maps F to the series S as shown in (1).

$$F_{w \times h \times c} \rightarrow S_{m \times c} \qquad (1)$$

Where w , h , c and m are the width, height, number of channels of the F and the length of the reshaped series S , respectively.

Then 1-th to N-th order difference (N-diff) of reshaped S is performed N times as shown in (2).

$$s_{m_0,c_0}^N = s_{m_0+N,c_0} - s_{m_0,c_0}, m_0 \leq m-N \qquad (2)$$

where, s is a sample in the S , and is used to form the fluctuation series s^N on the time scale of N . After N-diff operation, the length of the these series is form $m-N$ to m , respectively, so it is necessary to do the zero-padding to let the short series be as long as the original one.

Finally these N-diff series is reshaped to $w \times h \times c$ and concatenated (or summated if the amount of channels is too large), by the channel dimension to form new feature map tensors, which will act as the input of the convolutional layer later. So the TE block could

output the convolutional feature maps reflecting the difference feature directly.

3 Mid-term wind scenario generation based on conditional WGAN-GP

3.1 Conditional WGAN-GP

Conditional WGAN-GP algorithm is a kind of generative model. Its training process can be regarded as a game between two neural networks. When the game reaches Nash equilibrium, one of the neural networks can transform noise into the conditional synthetic data. The two neural networks are connected sequentially and named generator and discriminator respectively. The training methods are as follows:

First, sampling produces noise $z \sim p(z)$, where $p(z)$ denotes the distribution of noises, which usually obey Gauss distribution. The difference of conditional GAN is that samples need to be labelled, a batch of samples (x, y) are taken out, the labels of samples y , usually one-hot classed, are linked together with z . The generated data $x_g = G(z \mid y, \theta^G)$ can be obtained by the generator, where $G(\cdot, \theta^G)$ represents the generator with current parameters. Then take x , x_g and y into the discriminator to get the discriminator outputs $V_T = D(x \mid y, \theta^D)$ and $V_F = D(x_g \mid y, \theta^D)$.

1) Discriminator Training: freeze θ^G while training θ^D , use Wasserstein-l as the loss function, and add an gradient penalty term additionally. The discriminator's loss function is:

$$L_D = \underset{x_g \sim \mathbb{P}_g}{\mathbb{E}} [D(x_g \mid y, \theta^D)]$$
$$- \underset{x \sim \mathbb{P}_r}{\mathbb{E}} [D(x \mid y, \theta^D)] \qquad (3)$$
$$+ \lambda \underset{x_{av} \sim \mathbb{P}_{av}}{\mathbb{E}} \left[\left(\left\| \nabla_{x_{av}} D(x_{av}) \right\|_2 - 1 \right)^2 \right]$$

Where x_{av} is the result of random weighted averaging of x and x_g . The loss function of WGAN-GP not only trains the discriminator to distinguish whether the two distributions are consistent, but also discriminates the size of their difference. The Wasserstein-l loss function can improve the gradient disappearance or explosion problems in traditional GAN training. [11] mentions that Wasserstein-l is founded on the premise that the neural network satisfies the 1-Lipschitz constraint, while the literature [12] introduces the gradient penalty (GP) term, which is the third term in (3),to meet that requirement.

2) Generator Training: When the discriminator is trained n_{critic} times, freeze θ^D training θ^G , similarly use y , z to generate V_F . Backpropagation based on the generator's loss function, enabling the generator to generate synthetic data that is sufficient to fool the discriminator. The specific loss function of the generator is as follows:

$$L_D = - \underset{x_g \sim \mathbb{P}_g}{\mathbb{E}} [D(x_g \mid y, \theta^D)] \qquad (4)$$

Overall, if we ignore the GP term, the training process is as follows:

$$\min_G \max_{D \in \mathcal{D}} \underset{x \sim \mathbb{P}_r}{\mathbb{E}} [D(x)] - \underset{x_g \sim \mathbb{P}_g}{\mathbb{E}} [D(x_g)] \qquad (5)$$

3.2 Mid-term wind scenario generation

3.2.1 The drawback of traditional conditions GAN in mid-term scale

In the past, [10] has used conditional GAN for one-day scenario generation with the day-ahead wind power prediction or clustering results as its one-hot conditions. This method is very suitable for single-day conditional scenario generation, but when the time scale comes to the mid-term, directly using this method may bring a series of problems:

First of all, the forecast error of wind power in the mid-term is larger than day-ahead, and even in some cases there is no point predictions for wind power, so the mid-scale scenario generation should use the energy forecast of wind power as the basis.

More importantly, the number of one-hot encoded classes will increase exponentially as time scale increases. Assuming that there are c categories for single-day wind power, n -day wind power requires at least c^n classes to have the same performance as single-day scenario generation. Because of this problem, when the time scale is relatively long, the Classification condition method will greatly reduce the feasibility of the algorithm.

In addition, if clustering result is used as conditions, the conditions will not have an obvious physical meaning, and cannot be continuously changed as needed, resulting in inflexibility in practical applications.

3.2.2 The model overview

For the above reasons, it is appropriate to use the interval of daily/total energy of wind power as the condition. The model overview is shown in Figure 2. The model uses the TE block to replace the convolutional layers in generator and discriminator respectively to ensure the fluctuation of different time scales could be reflected. The rest parts of model are consistent with the conditional WGAN-GP.
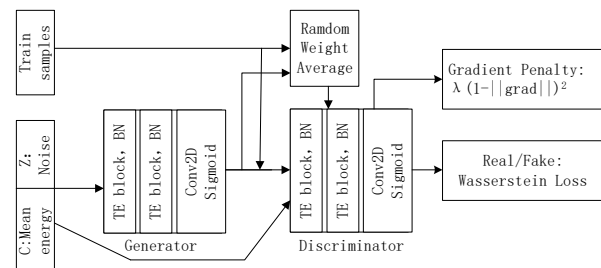

Fig. 2. The overview of Mid-term Scenario Generation Model.

4 Case study

Because this paper mainly discusses how to enhance the temporal correlation modeling ability of the generative models, the spatial correlation between multiple wind fields is not considered as the key point. Therefore, the dataset uses the annual total wind power data collected from a province in northern China. The total installed capacity is 7300MW. Moreover, the wind power of a larger spatial scale is more representative of the mid-term characteristics of wind power than the short-term. Since the time scale is mid-term, a resolution of 1h is considered sufficient. The length of the train samples in this paper is 5 days (120h), but as a matter of fact the proposed model structure can be applied to various non-extreme time scales.

4.1 Model details

The training samples are obtained for each successive 120h on the annual wind power curve and slide one hour every time to form a sample matrix, which is then normalized to enable the neural network to converge correctly. The model input is the concatenation of the condition C and noise Z, and the dimension of noise is also set to 120 here. The mode of the input condition is the interval category of the total energy of wind power per day/total time resolution. And it is repeated and extended to the same length as the noise to make it easier to concatenate. As shown in Fig. 3 (a) classified by daily time and (b) classified by total time. The wind energy interval is evenly divided according to the quantile of the daily/total-time energy of wind power. For the sake of brevity, this paper only divides the energy into three parts, and the threshold is (1970MW/3405MW) × Total time, respectively.
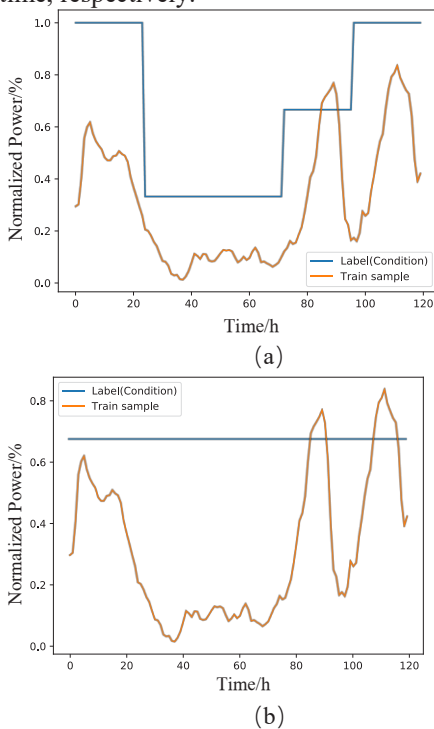

(a)

(b)

Fig. 3. The daily/total-time condition formation mode.

4.2 Comparison of learning effects between two models

In order to test whether the proposed model can improve the ability of uncertainty modelling of the generated scenarios' fluctuation, and ensure that the specific distribution generating ability of the.generative model does not decrease at the same time. The conditional probability distribution of the power and the fluctuation of the generated scenarios are compared in this section, to tell if the model is able to reproduce the real data distribution. In order to achieve a sufficient amount of data for each category, = total-time wind energy classification is used as its condition.

Figures Fig.4(a) and Fig.4(b) show the probability density distributions of the first and second classes of the original sample probability density distribution, the distribution of generated scenarios output by the WGAN-GP model without/containing the TE block with corresponding conditions. By comparing the three curves, we can see that the two models have better learning effects and can accurately reflect the distribution of wind power under different conditions.
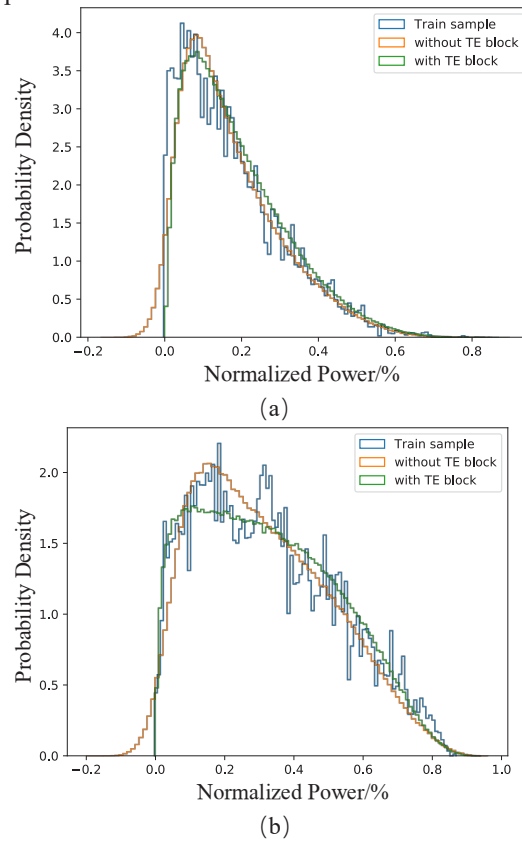

(a)

(b)

Fig. 4. Learning results of models on different classes of wind power distribution.

Fig. 5(a) and Fig. 5 (b) respectively show the conditional probability density distribution of the first and second class of original sample wind power fluctuations, and the fluctuation distribution of scenarios generated by the WGAN-GP model without/containing the TE block. It can be clearly seen that the green curve (denotes the model containing the TE block) is more close to the real distribution, and the wind power fluctuation distribution of the original WGAN-GP

generated scenarios deviates significantly from the real distribution. From the point of view of fluctuation distribution, it is proved that the proposed model is more capable of capturing temporal correlation.
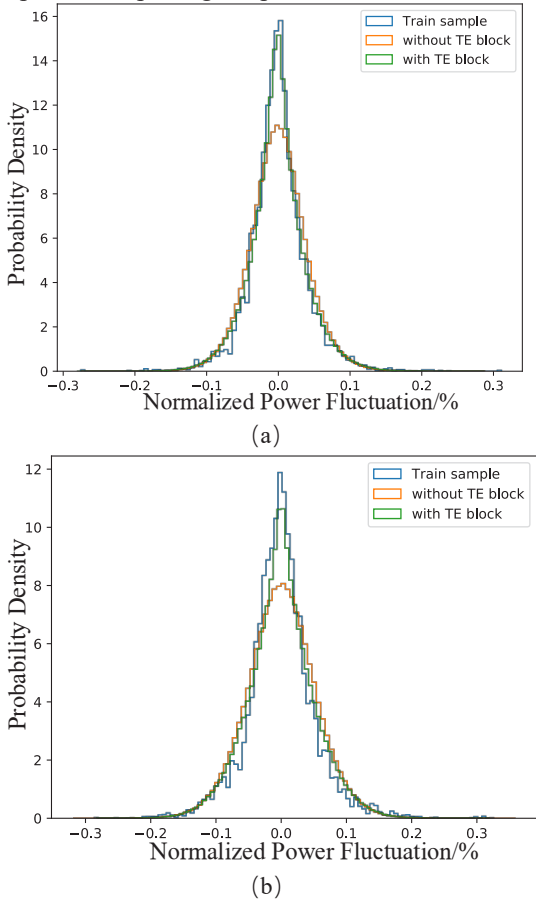

(a)

(b)

Fig. 5. Learning results of models on different classes of fluctuation distribution.

In addition, the auto-correlation coefficients which can represent the temporal correlation to some extent are also counted. By using traversal search to generate similar scenarios nearest to the train samples, the MAE distribution of the autocorrelation function is counted as shown in Fig. 6(a). It can be seen that the ACF function of the model with TE block is closer to the real situation. And Figure 6(b) gives a more specific case study.
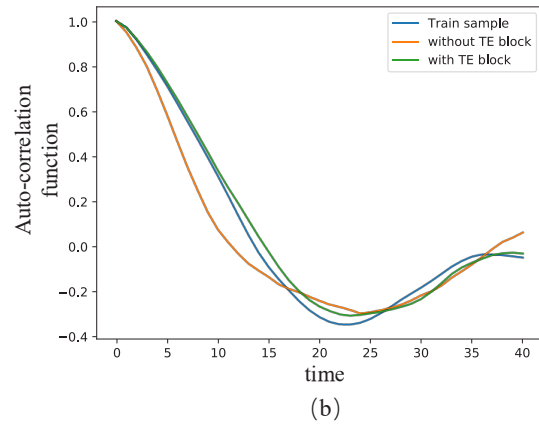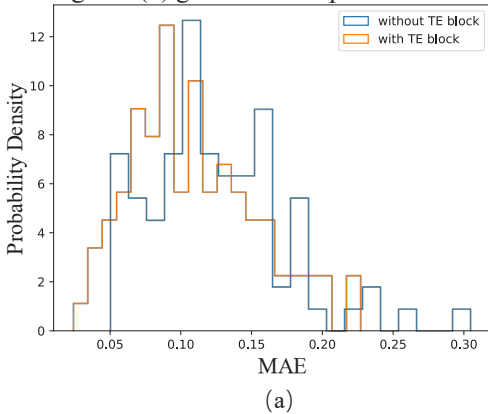

(a)

(b)

Fig. 6. Auto-correlation coefficients comparing with its nearest sample.

4.3 Specific scenarios generation

Using the condition of daily classification, scenario generation is performed on two different conditions of 100h and 8000h respectively to simulate the uncertainty of wind power. The training samples corresponding to 100h and 8000h (blue thick lines) are shown in fig. 7 (a) and fig. 7(b), and the input conditions are also shown for the 8000h samples.
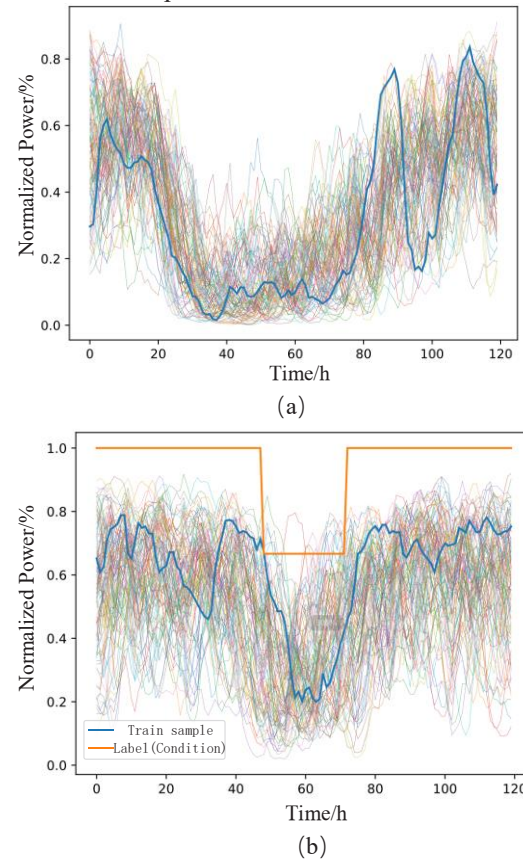

(a)

(b)

Fig. 7. A large number of scenarios generating with a given condition.

5 Conclusion

Scenario generation is an important tool for decision-making in the power system with high wind penetration rate. Firstly, this paper discusses the insufficient expression of temporal correlation in the common conditional GAN algorithms, and proposes a solution based on modified neural network structures in the model. Then, the possible problems of applying conditional WGAN-GP algorithm in the mid-term are studied, and a scenario generation model based on improved conditional WGAN-GP model for mid-term wind scenario generation is established. Finally, examples are given in Case Study to verify the feasibility of the proposed model in scenario generation. Compared with the model without TE block, the proposed method improves the learning ability in temporal correlation such as fluctuation, and a series of generated mid-term wind power scenarios is shown in the end.

References

1. State Grid Corporation. Gansu Power Grid 2018 Annual Operation Mode [R].2018.

2. Huang Honglin，Song Lili, Zhou Rongwei, et al. Characteristics Analysis of Wind Power Fluctuations for Large-scale Wind Farms [J]. Proceedings of the CSEE, 2017(06):26-37.

3. Quan H , Srinivasan D , Khosravi A . Incorporating Wind Power Forecast Uncertainties Into Stochastic Unit Commitment Using Neural Network-Based Prediction Intervals[J]. IEEE Transactions on Neural Networks and Learning Systems, 2014, 26(9):2123-2135.

4. Cui M , Ke D , Sun Y , et al. Wind Power Ramp Event Forecasting Using a Stochastic Scenario Generation Method[J]. IEEE Transactions on Sustainable Energy, 2015, 6(2):422-433.

5. Ma, Xi-Yuan, Yuan-Zhang Sun, and Hua-Liang Fang. "Scenario generation of wind power based on statistical uncertainty and variability." IEEE Transactions on Sustainable Energy 4.4 (2013): 894-904.

6. Morales J M, Mínguez R, Conejo A J. A methodology to generate statistically dependent wind speed scenarios[J]. Applied Energy, 2010, 87(3):843-855.

7. Sideratos G , Hatziargyriou N D . Probabilistic Wind Power Forecasting Using Radial Basis Function Neural Networks[J]. IEEE Transactions on Power Systems, 2012, 27(4):1788-1796.

8. Goodfellow I J , Pouget-Abadie J , Mirza M , et al. Generative Adversarial Networks[J]. Advances in Neural Information Processing Systems, 2014, 3:2672-2680.

9. Chen Y , Wang Y , Kirschen D S , et al. Model-Free Renewable Scenario Generation Using Generative Adversarial Networks[J]. IEEE Transactions on Power Systems, 2018:1-1.

10. Zhang, Yufan, et al. "Typical wind power scenario generation for multiple wind farms using conditional improved Wasserstein generative adversarial network." International Journal of Electrical Power & Energy Systems 114 (2020): 105388.

11. Arjovsky M, Chintala S, Bottou L. Wasserstein GAN[J]. 2017.

12. Gulrajani I, Ahmed F, Arjovsky M, et al. Improved Training of Wasserstein GANs[J]. 2017.