

## SPECIAL ISSUE RESEARCH ARTICLE

# Day-ahead renewable scenario forecasts based on generative adversarial networks

Congmei Jiang<sup>\*1,2</sup> | Yongfang Mao<sup>\*1,2</sup> | Yi Chai<sup>1,2</sup> | Mingbiao Yu<sup>3</sup>

<sup>1</sup>School of Automation, Chongqing University, Chongqing 400044, China

<sup>2</sup>Key Laboratory of Complex System Safety and Control, Ministry of Education, Chongqing 400044, China

<sup>3</sup>School of Instrument Science and Engineering, Southeast University, Nanjing 210096, China

## Correspondence

Congmei Jiang and Yongfang Mao, School of Automation, Chongqing University, Chongqing, China.  
Email: jiangcongmei@gmail.com; yfm@cqu.edu.cn

## Present Address

No. 174 Shazhengjie, Shapingba, Chongqing, 400044, China

## Summary

With the increasing penetration of renewable resources such as wind and solar, especially in terms of large-scale integration, the operation and planning of power systems are faced with great risks due to the inherent stochasticity of natural resources. Although this uncertainty can be anticipated, the timing, magnitude, and duration of fluctuations cannot be predicted accurately. In addition, the outputs of renewable power sources are correlated in space and time, and this brings further challenges for predicting the characteristics of their future behavior. To address these issues, this paper describes an unsupervised distribution learning method for renewable scenario forecasts that considers spatiotemporal correlation based on generative adversarial network (GAN), which has been shown to generate realistic time series for stochastic processes. We first utilize an improved GAN to learn unknown data distributions and model the dynamic processes of renewable resources. We then generate a large number of forecasted scenarios using stochastic constrained optimization. For validation, we use power generation data from the National Renewable Energy Laboratory wind and solar integration datasets. The simulation results show that the generated trajectories not only reflect the future power generation dynamics, but also correctly capture the temporal, spatial, and fluctuant characteristics of the real power generation processes. The experimental comparisons verify the superiority of the proposed method and indicate that it can reduce at least 50% of the training iterations of the generative model for scenario forecasts.

## KEYWORDS:

artificial intelligence, unsupervised learning, generative adversarial network, renewable energy sources, scenario generation

## 1 | INTRODUCTION

To protect the environment and reduce the consumption of conventional energy resources, renewable energy will become progressively more important in the future. However, the intermittent and volatile nature of renewable power generation brings negative impacts for managing the operation of power systems, and additional reserves and facilities are required to accommodate the resulting power imbalances and aid network transmission, especially in the case of large-scale integration<sup>1,2</sup>. Scenario generation plays an important role in characterizing the uncertainties inherent in the use of renewable resources, and this technique is widely used in time-dependent and multistage decision-making problems such as the optimal unit commitment, electricity

## Nomenclature

### A. Acronyms

AI	Artificial intelligence	CDF	Cumulative distribution function
ML	Machine learning	SO	Stochastic optimization
DL	Deep learning	PI	Prediction interval
GAN	Generative adversarial network	U	Upper bound
WGAN	Wasserstein generative adversarial network	L	Lower bound
ARMA	Auto regressive moving average	GDO	Gradient descent optimization
VAE	Variational autoencoder	NWP	Numerical weather prediction
ANN	Artificial neural network	WD	Wasserstein distance
DNN	Deep neural network		
CNN	Convolutional neural network		

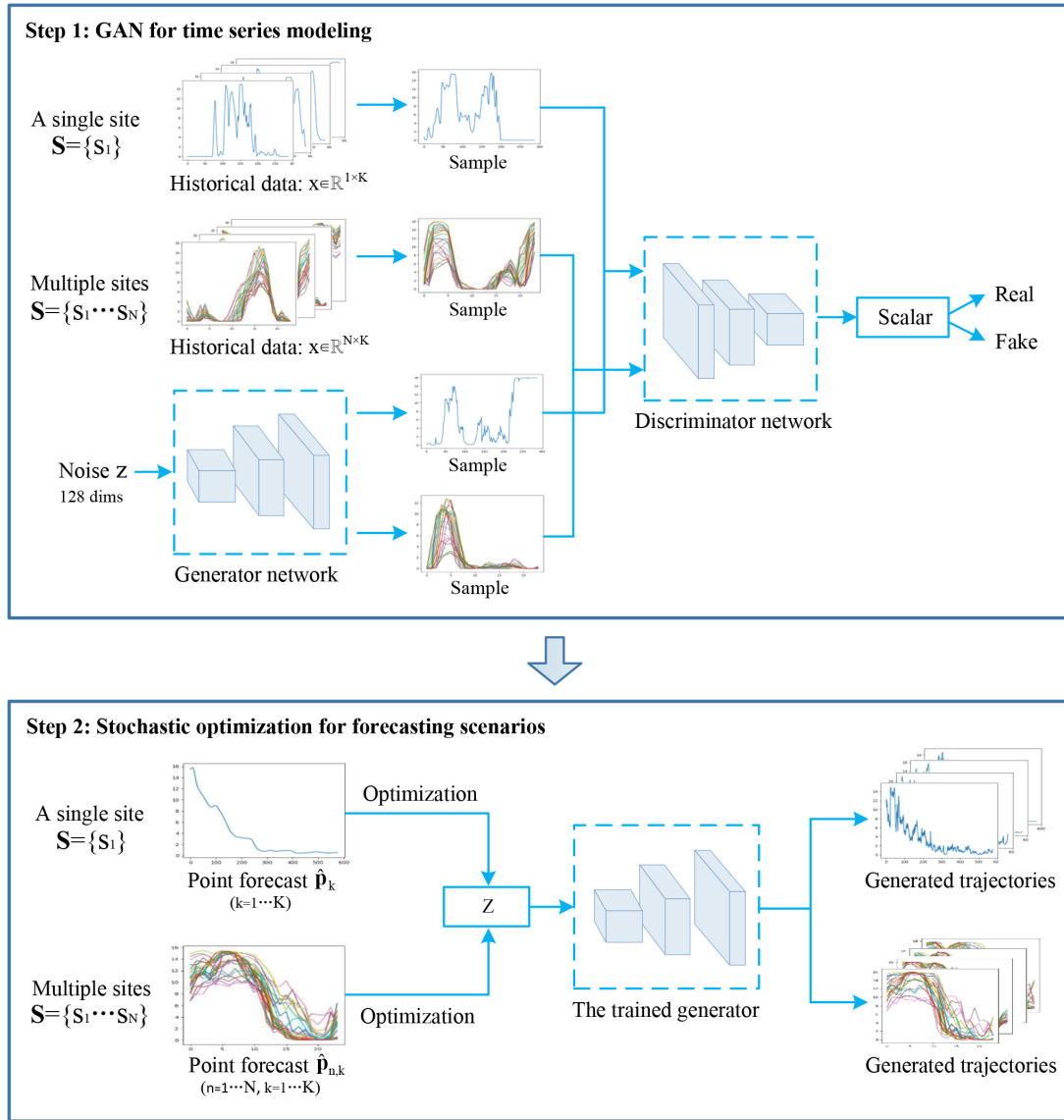
WGANGP	Wasserstein generative adversarial network with gradient penalty	<i>B. Parameters</i>	
WGANCT	Wasserstein generative adversarial network with a consistency term	N	Number of sites
		K	Length of forecast horizon
		$\theta$	Bound parameter for controlling the forecast range
$P_r$	Probability distribution of real time series data	$\tau, v$	Weight parameters of stochastic optimization objective
$P_G$	Probability distribution of generated time series data	$\eta$	Match parameter of optimizing generator and discriminator
NREL	National renewable energy laboratory	<i>C. Indices</i>	
RES	Renewable energy sources	n	Site index
PV	Photovoltaic	k	Time index
G	Generator/generative network	h	Lag time index
D	Discriminator/discriminative network		

market trading, and energy storage sitting and sizing<sup>3,4,5</sup>. For power system planning and operation, accurate modeling of the output of renewable energy sources (RES) is therefore key to enforcing reliability criteria and increasing economic benefits.

Extensive research has been conducted into scenario generation. Gaussian copula has been proposed for the generation of statistical scenarios, and the interdependence structure of forecast errors for various look-ahead times is considered for wind power<sup>6,7</sup>. Moment-matching scenario generation from multivariate random variables with specified moments and correlations was introduced by Hoyland et al.<sup>8</sup>, and Meibom et al. used a method based on auto regressive moving average (ARMA) model to produce wind and load scenarios<sup>9</sup>. An artificial neural network (ANN) was presented by Vagropoulos et al. to create more representative scenarios for electric loads, photovoltaic (PV) and wind production<sup>10</sup>. These methods have been applied to a single site, and some of them may be extended to capture spatial correlation for multisite datasets. Copula-based methods have been used to produce spatially correlated scenarios across different geographical sites<sup>11,12,13</sup>, and time-series models have been applied to characterize the stochastic processes representing spatial-temporal information from multisite wind data<sup>14,15</sup>. Each of these scenarios embodies time dependencies and is spatially dependent on other stochastic wind processes.

Probabilistic models are usually based on statistical assumptions, and it is a challenge to capture the intrinsic patterns underlying stochastic power generation dynamics using these models. The generated scenarios can be used to represent future uncertainty but cannot correctly describe all the salient features in the power output from renewable energy sources. Furthermore, the complex spatiotemporal behavior of the sources and the use of sampling from high-dimension joint distributions make it difficult to apply the probabilistic methods in practice.

Deep learning (DL) is one of the newest trends in artificial intelligence (AI) and machine learning (ML) research, with unprecedented results across various application domains<sup>16,17,18,19,20</sup>. In the task of unsupervised learning, generative models are one of the most promising technologies. Two popular examples of generative model approaches are generative adversarial network (GAN)<sup>21</sup> and variational autoencoder (VAE)<sup>22</sup>. Since generative models can learn the underlying realistic distribution of stochastic processes, they can generate data that exhibit similar characteristics to the original data. The GAN-based methods have been adopted for generating a scenario set that can capture the intrinsic patterns of volatile power production<sup>23,24,25</sup>. The



**FIGURE 1** Framework for the GAN used for different forecasting tasks.

1 simulated scenarios can correctly characterize the dynamic and time-varying characteristics of historical observations. Varia-  
 2 tional inference was also incorporated into GAN by Hu et al. to produce wind and solar scenarios that capture different salient  
 3 characteristics in the data <sup>26</sup>. In a previous work, we utilized an improved GAN to create scenarios for wind power by applying  
 4 alternative training techniques to enforce Lipschitz constraints on the discriminator <sup>27</sup>. As another popular framework for gen-  
 5 erative models, VAE learns the data distribution by calculating the mean square error between samples. Compared with GAN,  
 6 it has features that do not need to achieve Nash equilibrium and can directly compare the generated samples with the original  
 7 samples. A scenario generation method based on a variational autoencoder was proposed by Pan et al. to generate stochastic  
 8 load profiles for electric vehicles <sup>28</sup>. Zhang et al. designed the variational autoencoder to capture the fluctuant characteristics  
 9 and spatiotemporal correlation of solar and wind power <sup>29</sup>. Recently, the flow-based generative models <sup>30,31</sup> have also been used  
 10 for the generation of realistic scenarios. Zhang et al. proposed a flow-based conditional generative model to provide reliable and  
 11 sharp scenarios for residential load <sup>32</sup>. To accurately capture the potential behavior of real samples, a generative network based  
 12 on nonlinear independent component estimation was proposed by Ge et al. to model the daily load profiles <sup>33</sup>.

13 These data-driven models do not rely on probabilistic assumptions and can characterize the stochastic processes with a full  
 14 diversity of behaviors. However, although these methods can create realistic scenarios that correctly characterize the salient  
 15 features of training data, most of these methods do not consider forecast information to generate scenarios that can characterize

the future dynamics. Unlike uncertainty modeling that generates time series with the same statistical characteristics as historical data, scenario forecasts can generate a group of time series that represent the possible future realizations and can therefore be used for a wide range of downstream tasks in power systems. Previous works have made substantial progress in using generative models to model the stochastic dynamics of renewable power generation processes, but how to use these powerful models to forecast future renewable scenarios remains a challenging problem. Chen et al. proposed an approach based on generative models to generate future scenarios representing dynamic and time-varying characteristics for a particular location<sup>34</sup>. Due to the similarity of meteorological dynamics, the power output of renewable energy in different locations will have a natural correlation. Uncertainty estimation considering spatial dependence is essential for power system operation and planning studies. To tackle this issue, we propose a data-driven method based on generative adversarial network to forecast renewable scenarios for multiple geographical locations. Furthermore, due to the use of weight clipping, the GAN has problems such as unstable training and underused network capacity in the existing works<sup>24,34</sup>. We further adopted improved training techniques to improve the model performance for scenario forecasts.

The framework for the GAN used for renewable scenario forecasts is shown in Figure 1. In this work, we adapt GAN to efficiently generate future scenarios for a single or multiple renewable power generation processes. We first use GAN to learn and create realistic trajectories with the same statistic properties as the training data, then incorporate forecast information to generate a group of future scenarios for different forecasting tasks.

Specifically, the main contributions of this research can be listed as follows:

- Introduction of spatiotemporal scenario forecasts for wind and solar power by utilizing deep generative models.
- Scenario forecasts of different forecasting tasks via the same GAN model and network structure.
- High flexibility of the length of forecast horizons and the number of power generation sites via structure and parameter adjusting.
- Assessment of the improved performance for scenario forecasts via weights, scenario quality and convergence speed.

## 2 | GENERATIVE ADVERSARIAL NETWORK

GAN offers a new framework for drawing realistic samples from an unknown data distribution, and it is therefore currently one of the most active research areas in artificial intelligence. Since the introduction by Goodfellow in 2014, GAN has become a powerful method for a variety of applications<sup>35,36,37,38</sup>. However, the GAN can be remarkably difficult to train. Wasserstein GAN (WGAN)<sup>39</sup> has made progress in model performance, but still suffers from instability in training and may generate low-quality samples<sup>40</sup>. In this section, we first introduce the WGAN and the improved training techniques, and then present the training effect for capturing the data distribution of renewable resources.

### 2.1 | Wasserstein GAN

The idea of GAN is to estimate generative models as a competing game between two networks: a discriminator network  $D$  and a generator network  $G$ . Consider a historical dataset  $x$  with a data distribution denoted by  $P_r$ . To train the generator  $G$  over data  $x$ , we define an noise variable  $z$  (usually sampled from a simple distribution, such as Gaussian distribution or uniform distribution) and map it to a data space represented as  $G(z)$ . The generator  $G$  tries to output plausible samples by minimizing  $-D(G(z))$ , which means that the generator is trained to learn the data distribution of real data. The discriminator  $D$  is alternately updated by the generator, and can be trained by maximizing between  $D(G(z))$  and  $D(x)$ , which reflects that the discriminator is good at distinguishing the difference between input samples. In a general form, the loss functions for  $D$  and  $G$  can be expressed as

$$\begin{aligned} L_D &= -\mathbb{E}_{x \sim P_r}[D(x)] + \mathbb{E}_{z \sim P_z}[D(G(z))], \\ L_G &= -\mathbb{E}_{z \sim P_z}[D(G(z))]. \end{aligned} \quad (1)$$

With the above objectives defined, the adversarial networks try to minimize each objective function in each training iteration. Formally, the game between the discriminator  $D$  and the generator  $G$  is represented by the minimax objective

$$\min_G \max_D V(G, D) = \mathbb{E}_{x \sim P_r}[D(x)] - \mathbb{E}_{z \sim P_z}[D(G(z))]. \quad (2)$$

The value function (2) can be interpreted as Wasserstein distance (WD, also called Earth-Mover distance) <sup>41</sup>. Compared with other distances and divergences (e.g., Jensen-Shannon divergence <sup>21</sup>), this distance is shown a more reasonable objective function when used for the training of generative models. The Wasserstein distance is defined as

$$W(P_r, P_G) = \inf_{\psi \in \Pi(P_r, P_G)} \mathbb{E}_{(x,y) \sim \psi} [\|x - y\|], \quad (3)$$

where  $\Pi(P_r, P_G)$  denotes the set of all joint distributions  $\psi(x, y)$  whose marginals are  $P_r$  and  $P_G$ . Intuitively,  $\psi(x, y)$  indicates the transport cost of converting distribution  $P_G$  into distribution  $P_r$ . However, it is impractical to achieve the objective function in such a formula using neural networks. According to the Kantorovich–Rubinstein duality <sup>41</sup>, this distance can also be expressed as

$$W(P_r, P_G) = \sup_{\|D\|_L \leq 1} \mathbb{E}_{x \sim P_r} [D(x)] - \mathbb{E}_{z \sim P_z} [D(G(z))], \quad (4)$$

1 where  $D$  is the set of 1-Lipschitz functions.

## 2.2 | Improved training techniques

The Wasserstein distance has better theoretical properties that make generator optimization easier. However, it can still produce low-quality samples as a result of particular inputs or fail to converge in some settings. The clipping parameter  $c$  will affect the optimization process, which may lead to the problems of vanishing or exploding gradients in the training of generative models. To overcome these issues, an alternative to clipping weights has been proposed to improve the performance of WGAN <sup>40</sup>. In particular, a gradient penalty  $GP$  is introduced in the discriminator loss function.

$$GP|_{\hat{x}} = \mathbb{E}_{\hat{x} \sim P_{\hat{x}}} [(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2], \quad (5)$$

3 where  $\hat{x}$  represents uniform samples along straight lines between pairs of points from  $P_G$  and  $P_r$ .

4 It is better to utilize the gradient penalty for enforcing a Lipschitz constraint. Clear advantages of this when compared to weight  
5 clipping are its improved training speed and improved sample quality. Since the well-trained discriminator  $D$  has gradients with  
6 norm at 1 almost everywhere, enforcing this along straight lines between a pair of data points sampled from  $P_G$  and  $P_r$  seems  
7 feasible and results in good performance <sup>40</sup>.

The gradient penalty  $GP$  can penalize at the sampled data point  $\hat{x}$ , but it cannot penalize everywhere with a limited number of training iterations. In particular, in the early training stage, data point  $\hat{x}$  is distant from the manifold of the real data  $x$ . The Lipschitz continuity over the manifold is not enforced until data distribution  $P_G$  becomes close to data distribution  $P_r$ . Therefore, a consistency term ( $CT$ ) <sup>42</sup> is introduced to enhance the performance of GAN model by resorting to the basic definition of the Lipschitz continuity. The data manifold and surrounding regions are checked by using perturbed data near any real data  $x$ . The next-to-last layer  $D_-$  of  $D$  is also utilized to improve the training of the generative model. The final additional term can be expressed as

$$CT|_{x', x''} = \mathbb{E}_{x \sim P_r} [\max(0, d(D(x'), D(x'')) + 0.1 \cdot d(D_-(x'), D_-(x'')) - M')], \quad (6)$$

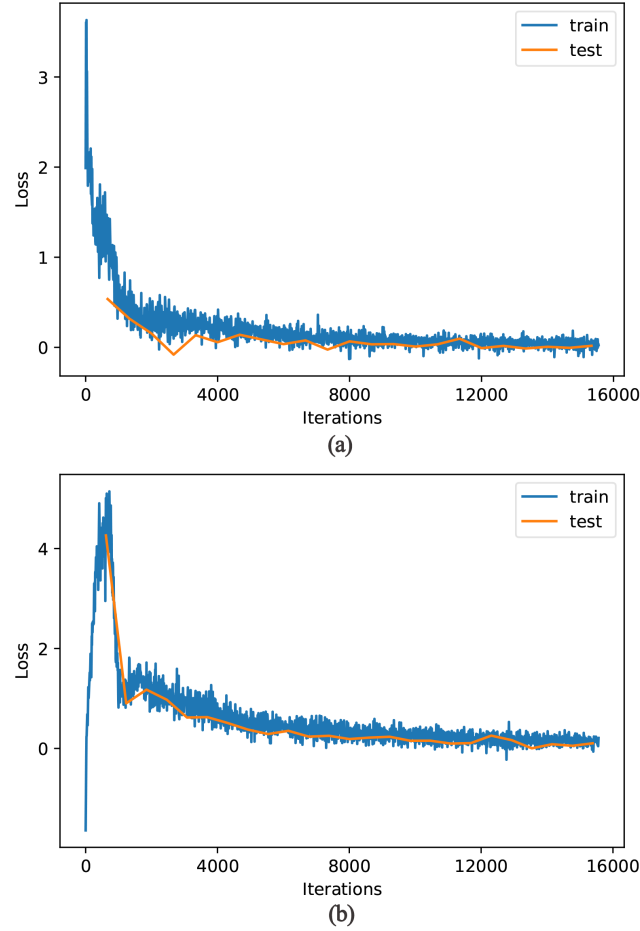
8 where  $d$  denotes the  $\ell_2$  metric on an input space,  $M'$  is a constant, and  $x'$  and  $x''$  are perturbed data points.

This new consistent regularization effectively complements and improves the gradient penalty used in the training of WGAN. Based on these improvements, the discriminator loss in (1) can be reformulated as

$$L_D = \mathbb{E}_{z \sim P_z} [D(G(z))] - \mathbb{E}_{x \sim P_r} [D(x)] + \lambda_1 GP|_{\hat{x}} + \lambda_2 CT|_{x', x''}. \quad (7)$$

9 With the redefined function  $L_D$ , we can train the GAN model to learn the real distribution of renewable data. An important  
10 benefit of WGAN is that its value function is correlated with sample quality, which provides a useful metric for training the gener-  
11 ator to optimality. To show the training effect of the generative model, we used solar and wind data from National Renewable  
12 Energy Laboratory (NREL) to train the model, and the convergence curves of the generative model are shown in Figure 2. By  
13 training the two adversarial networks to an equilibrium, we can see that the two loss curves for wind and solar power gradually  
14 converge to minima and remain stable. To further check whether the discriminator overfits and provides an inaccurate estimate  
15 of a training point at which all bets are correlated with sample quality, we explored the behavior of the loss curves from the test  
16 sets. It can be seen that the training and testing losses have almost the same trend for these two data sets, which indicates that  
17 the model can be well trained for renewable resources.

18 In order to show the effect of the improved techniques on the use of discriminator's weights for renewable integration datasets,  
19 we plotted the comparison of the weights in Figure 3. When the discriminator is trained using weight clipping on a wind or solar  
20 dataset, as Figure 3(a) shows, the weights for both datasets are pushed towards two sides of extreme values, which match the

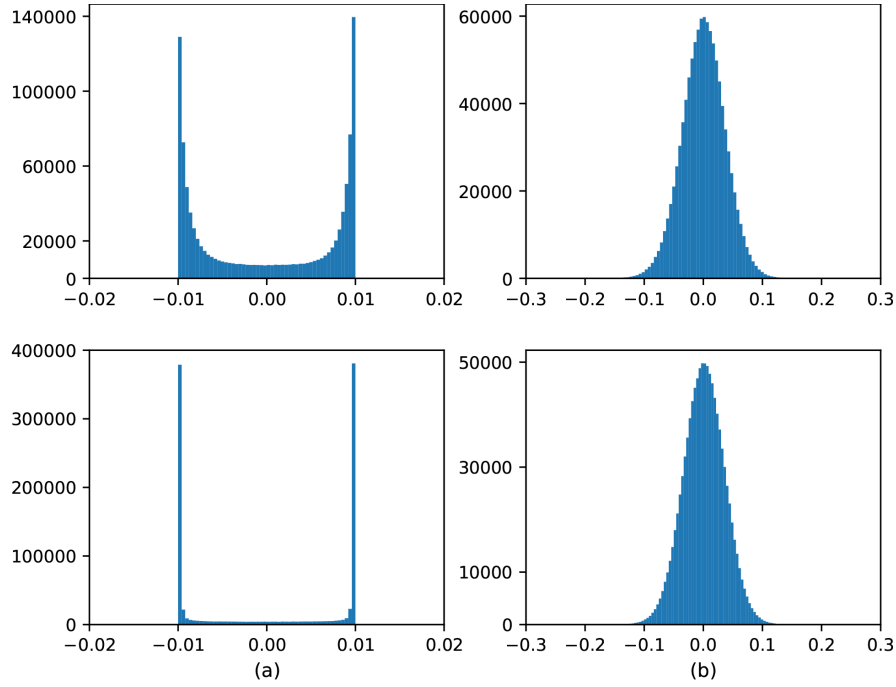


**FIGURE 2** Convergence curves of the generative model trained on (a) solar data and (b) wind data.

extremes of the clipping range. It shows that the network capacity of the discriminator is underused during the training. Besides, the optimization process is difficult because of interactions between the weight constraint and the cost function, which may result in either vanishing or exploding gradients, depending on the value of the clipping threshold. In contrast, as Figure 3(b) shows, the GAN trained using improved training techniques does not suffer from these behaviors. The weights of the discriminator trained on both datasets are controlled within a symmetric distribution. It is clear from these experimental results that the improved generative model can achieve a more stable training process for scenario forecasts.

### 3 | SCENARIO FORECASTS USING PRE-TRAINED GAN

Using a pre-trained GAN, we can generate realistic data that reflect the intrinsic patterns of renewable power generation dynamics. However, one more interesting problem is the generation of time series trajectories that can represent future uncertainty. Figure 1 shows the framework for the GAN used for scenario forecasts, which comprises two steps that proceed as follows. In step 1, based on deep learning, the GAN can learn and create samples with the same statistic properties as the training set. The generator network uses random noise to imitate real samples and the discriminator network tries to distinguish between real samples and generated samples. Once the training in step 1 is completed, an optimal generator is obtained that is able to generate realistic samples. Step 2 can further help us to incorporate forecast information to generate a group of future scenarios. The generated trajectories should be able to reflect the future power generation dynamics and capture the temporal, spatial, and fluctuant characteristics of the real power generation processes.



**FIGURE 3** Weight comparison of the discriminator trained on (top) a wind dataset and (bottom) a solar dataset using (a) weight clipping and (b) improved training techniques, respectively.

Since the data of a single site is a special case of the data of multiple sites, we can formulate the scenario forecasting problem for typical multiple renewable energy sources. We assume that at time node  $t$ , we are provided with the following point forecast for each look-ahead time and each power generation site.

$$\hat{p}_{pred} = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,K} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,K} \\ \vdots & \vdots & \ddots & \vdots \\ p_{N,1} & p_{N,2} & \cdots & p_{N,K} \end{bmatrix}, \quad (8)$$

where  $K$  represents the forecast horizon and  $N$  represents the number of sites.

The unpredictability and variability of renewable energy sources is one of the fundamental issues associated with its integration into the power system. Point forecasts are able to provide deterministic information about uncertain future changes. Since we focus on the scenario forecasting problem, this information can be provided by any method, e.g., information from numerical weather prediction (NWP).

Based on the pre-trained GAN model, the forecasted trajectories should reflect the future dynamics and be volatile within a certain range. We can describe the fluctuation by defining a prediction interval (PI)<sup>34,43</sup>, which can be controlled by a parameter  $\theta$ :

$$\begin{aligned} L_{\theta}(\hat{p}_{pred}) &= \frac{1}{\theta} \hat{p}_{pred} \\ U_{\theta}(\hat{p}_{pred}) &= \theta \hat{p}_{pred}, \end{aligned} \quad (9)$$

where  $L_{\theta}(\hat{p}_{pred})$  represents a lower bound and  $U_{\theta}(\hat{p}_{pred})$  represents an upper bound.

As introduced in Section 2,  $z$  is a pre-defined random variable that is mapped to a data space  $G(z)$ . If the variable  $z$  is not constrained, the trained generator  $G$  can generate any scenarios that may exceed the upper and lower bounds in (9). Since the central point forecast will not fall outside the prediction interval, we can solve the following problem to obtain an initial point  $z$

for producing time series.

$$\begin{aligned} \min_z & \left\| \mathbb{P}_{pred}(G(z)) - p_{init} \right\|_2 \\ \text{s.t.} & L_\theta(\hat{p}_{pred}) \leq p_{init} \leq U_\theta(\hat{p}_{pred}) \\ & z \in Z, \end{aligned} \quad (10)$$

where the constraint is element-wise inequality and  $p_{init}$  is sampled from an initial interval  $[L_\theta(\hat{p}_{pred}), U_\theta(\hat{p}_{pred})]$ .

Note that our goal is to produce trajectories that can capture the intrinsic characteristics of stochastic processes and represent the future uncertainty of variable power generation. A larger discriminator output for the generated samples indicates more realistic samples. Meanwhile, the fluctuation of the generated trajectories can be controlled by the pre-determined confidence interval  $\theta$ . Using all these objectives, with the pre-trained GAN and the element-wise inequality constraint of the forecast range, we can formulate the scenario forecasting problem to be a stochastic optimization (SO) problem where

$$\begin{aligned} \min_z & -D(G(z)) \\ \text{s.t.} & L_\theta(\hat{p}_{pred}) \leq \mathbb{P}_{pred}(G(z)) \leq U_\theta(\hat{p}_{pred}) \\ & z \in Z. \end{aligned} \quad (11)$$

To start with a good initial  $z$  incurring  $G$  to generate a forecasted scenario, the parameter  $\theta$  in (10) can be set to be smaller than that in the optimization (11). This can help  $G(z)$  to generate time series that fluctuate within an appropriate range. Since the inequality constraint is not convenient for solving the stochastic optimization (11), we can use two log barriers to replace it into the main objective. The optimization problem can then be expressed as

$$\begin{aligned} \min_z & -\left\| \tau \log(\mathbb{P}_{pred}(G(z)) - L_\theta(\hat{p}_{pred})) + v \log(U_\theta(\hat{p}_{pred}) - \mathbb{P}_{pred}(G(z))) \right\|_1 - D(G(z)) \\ \text{s.t.} & z \in Z, \end{aligned} \quad (12)$$

where  $\tau$  and  $v$  are weighting parameters.

Since the two adversarial networks  $G$  and  $D$  are highly nonconvex, there are many local optima in (11). By solving the stochastic optimization in (12) with different initial  $z$ , we can generate realistic forecasted trajectories that can characterize the intrinsic dynamics and different uncertainties of volatile power generation.

## 4 | EXPERIMENTAL

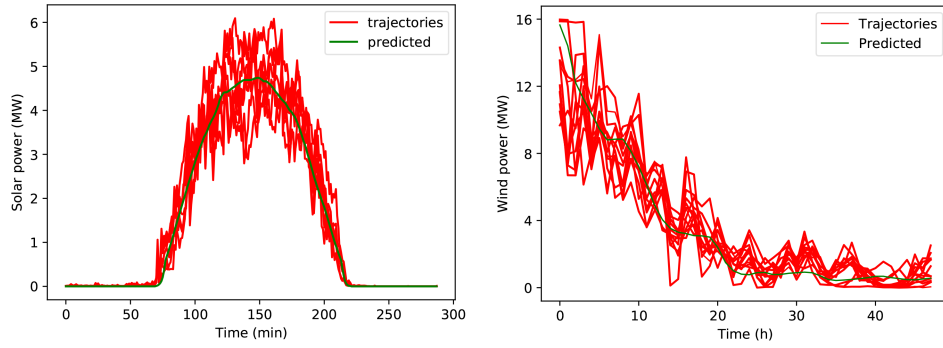
In this section, we discuss in detail the proposed method for capturing the characteristics of renewable resources. We show that the method can create realistic trajectories that are able to represent the future uncertainty for a single or multiple power generation locations. The experimental results validate that our proposed method is an effective tool for renewable scenario forecasts.

### 4.1 | Data description

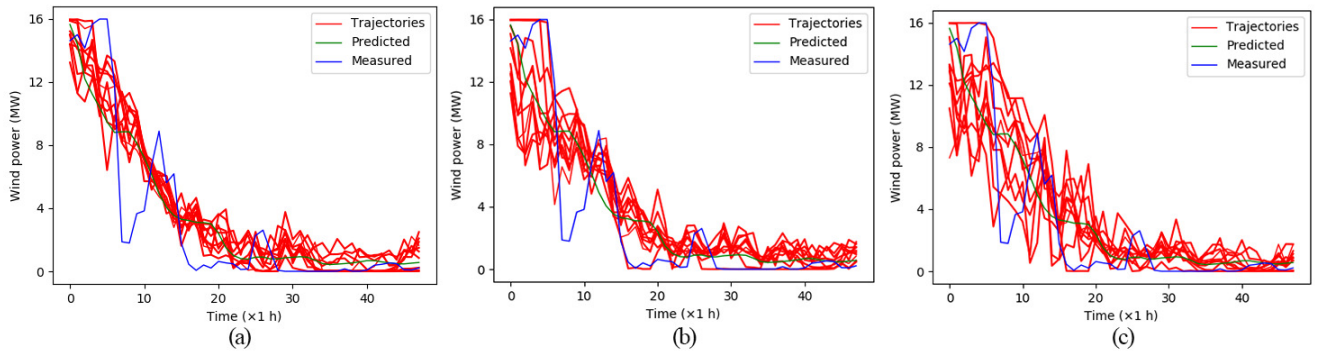
To test the performance of the proposed method, we used actual data from the previous publication<sup>34</sup> as the input. These data were collected from public datasets provided by NREL<sup>44</sup>, which provides tools and data for the analysis of grid technologies and strategies, including power system models and renewable datasets. This includes seven years' historical data with a temporal resolution of 5 minutes. More details about the relevant information can be found in the paper<sup>44</sup>. Wind data from 20 sites and solar data from 32 sites, all located in Washington State, were chosen to construct input data for different modeling tasks. We randomly selected 80% of the data to train the generative model, and the remaining 20% was used as testing data. The public datasets not only provide historical data, but also contain corresponding point forecasts. Along with the input data, this information was used for stochastic optimization based on the pre-trained GAN. The chosen sites have the same rated capacity for wind or solar power and the corresponding data were normalized into the range  $[0, 1]$ .

### 4.2 | Model structure and optimization algorithm

The network structure of GAN for scenario forecasts is based on our previous work<sup>27</sup>. Due to the strong feature extraction abilities, we use convolutional neural networks (CNNs) in the GAN model. In this system, a discriminator  $D$  is used to judge



**FIGURE 4** A group of generated trajectories for renewable resources: (left) solar power and (right) wind power.



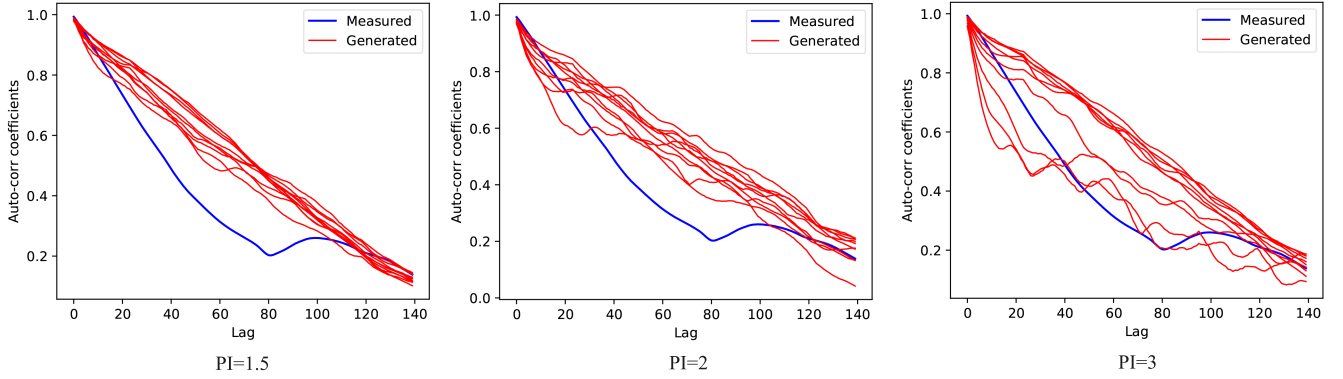
**FIGURE 5** Forecasted trajectories for wind power with a  $\theta$  value of (a) 1.5, (b) 2, and (c) 3.

whether a generator  $G$  can generate realistic samples. We use three convolutional layers with kernel size of  $4 \times 4$  and stride size of  $2 \times 2$  to down-sample the input data matrix  $x$  or  $G(z)$ . The generator  $G$  with a reverse structure is used to generate plausible samples. We use three deconvolutional layers with kernel size of  $4 \times 4$  and stride size of  $2 \times 2$  to do up-sampling from a random noise  $z$  to a data space  $G(z)$ . The choice of the kernel and stride is determined by experiment. In order to match the “game” process between the generator and the discriminator, the match parameter  $\eta$  is set to 5 so that the model is optimized between 1 step of  $G$  and 5 steps of  $D$ . The discriminator use multi-layer perceptron to output a scalar to distinguish the input and the samples are updated with a mini-batch size of 64. LeakyReLU and ReLU are used as activation functions for the discriminator  $D$  and the generator  $G$ , respectively. Stochastic dropout is adopted to the hidden layers of  $D$ . Batch normalization has been used in the network structures for previous time series modeling<sup>24,34</sup>. But it changes the mapping form between the input and output of the gradient penalty  $GP$ . Therefore, it can be simply omitted in the generative model.

In order to train the GAN model to the optimum and incorporate the prediction information to generate trajectories, different gradient descent optimization (GDO) algorithms can be used to achieve these stochastic optimization objectives. The Momentum and RMSprop optimization algorithms have long been popular for different deep learning structures, but the Adam algorithm combines the advantages of these two methods and is suitable for application to a wide range of non-convex optimization problems<sup>45</sup>. Therefore, we use Adam for the gradient based stochastic optimization. The generator  $G$  and the discriminator  $D$  are deep neural networks (DNNs), which can be implemented by the open-source AI learning framework TensorFlow<sup>46</sup>. All the training of the generative model is accelerated by an NVIDIA GeForce RTX GPU.

### 4.3 | Model training and trajectory generation

The proposed method for scenario forecasts consists of two steps as shown in Figure 1. To verify that our algorithm can generate the desired trajectories for renewable energy sources, we first trained a generative model to model the time series of renewable



**FIGURE 6** Autocorrelation plots for measured value and generated trajectories.

data, and then incorporated point forecasts to generate future trajectories. Repeatedly inputting real historical samples allowed the generative model to automatically learn the distribution of historical data. Figure 2 shows the training curves for the model trained on wind and solar data. At the early stage, the losses have large fluctuations because the model has not yet captured the distribution of power generation dynamics. The discriminator and the generator are continuously updated, and the generator gradually captures the intrinsic underlying patterns in the historical samples. As can be seen from the figure, the training evolution is reduced to a minimum and remains basically stable after about 8,000 iterations [see Figure 2(a)] for solar data, while it takes nearly about 13,000 iterations [see Figure 2(b)] for wind data to reach the same level. Wind power can be generated at any time during the day or night, while solar energy can only be generated during the day and cannot be generated under certain conditions (such as cloudy or rainy days). Therefore, the complexity of solar data relative to wind data is remarkably reduced. This may be an important reason why the generative model takes less time to learn the data distribution for solar data. When the generative model is trained to converge, the optimal generator can generate realistic time series representing the stochastic dynamics of renewable resources.

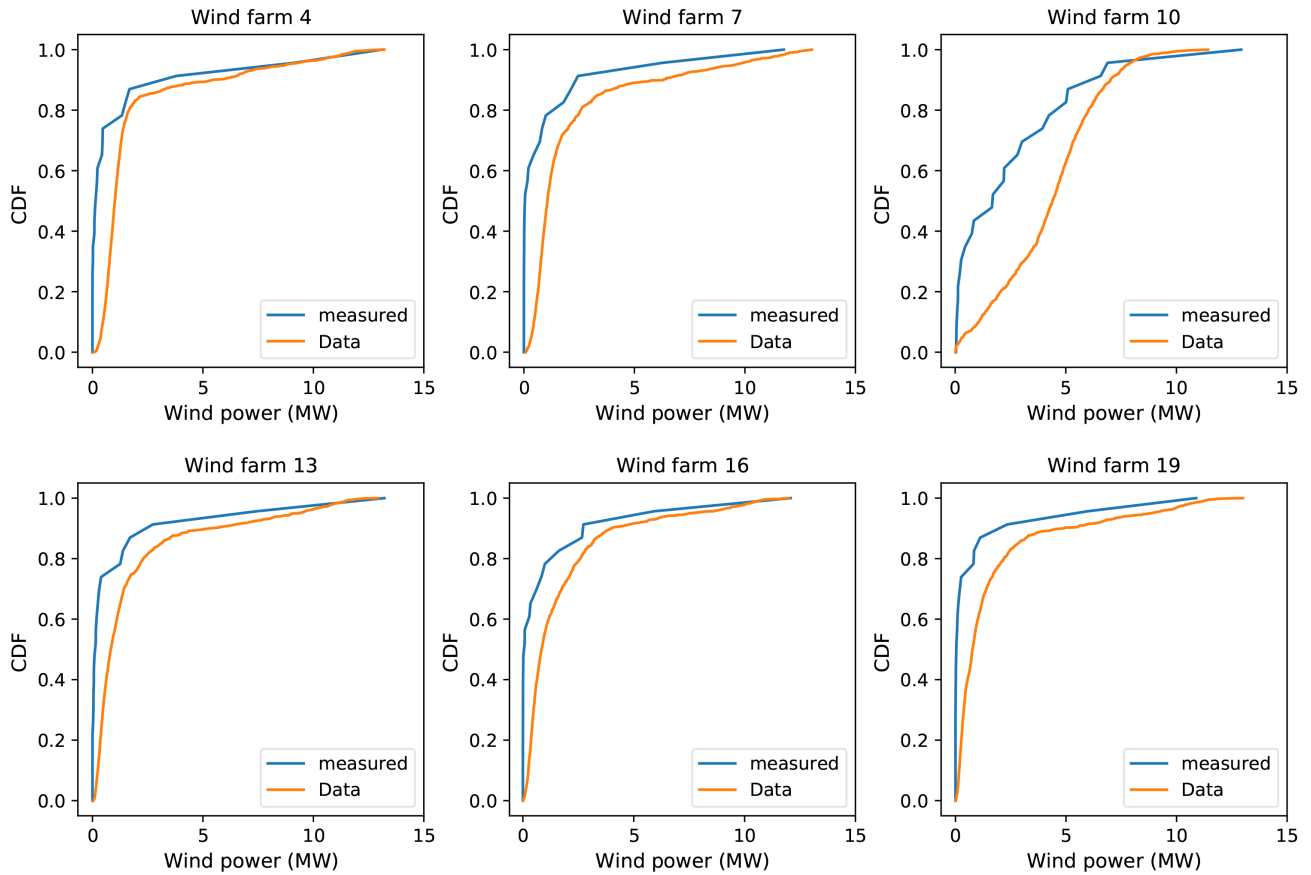
After the model training is completed, we can easily generate a large number of time series trajectories using the stochastic optimization objectives in (10) and (12). First, an appropriate initial  $z$  needs to be found according to the point forecast  $\hat{p}_{pred}$ . This  $z$  is then fed to the trained generator to generate future trajectories according to the parameter  $\theta$ . Figure 4 shows a group of generated trajectories for solar and wind power, respectively. It can be seen that the generated trajectories can reflect the nonlinear time dynamics of variable power generation. The volatility characteristics (peak-valleys, ramp events and daily patterns) of the predicted data can be correctly captured by the proposed method. In addition, wind energy usually changes quickly at any time and any place, while solar energy changes relatively slowly during its occurrence. Therefore, from Figure 4, we can see that wind power is more volatile than solar power.

#### 4.4 | Forecast uncertainty and evaluation

Figure 1 shows the framework of using GAN to generate future uncertain trajectories for renewable resources. The same model structure can be used to forecast scenarios for a single or multiple sites. Since wind power has more complex uncertainty and volatility characteristics than solar power, we can use wind power data for the following experimental demonstrations.

##### 4.4.1 | Scenario forecasts for a single site

To verify that a group of generated trajectories is representative of future uncertainty, the scenarios should cover the actual values of real power generation. The historical data provided have a resolution of 5 minutes. We used samples composed of two days of data as the training set. Training samples were randomly and repeatedly input to train the generative model for scenario forecasts. Figure 5 shows the simulation results when the prediction interval  $\theta$  is 1.5, 2, and 3. By choosing different parameters  $\theta$ , we can see that the resulting trajectories can be used to represent different degrees of uncertainties. When the parameter  $\theta$  is 1.5, the trajectories are relatively concentrated, but they cannot cover the actual power measurement. When  $\theta$  is 3, the time series have a larger degree of volatilities and can cover the measurement but are less concentrated. The range of the prediction



**FIGURE 7** Comparison of the cumulative distribution function of measured and generated data for multiple sites.

- 1 confidence can be chosen according to the accuracy of point forecast information and the required level of risk management.
- 2 At the same time, the weight parameters  $\tau$  and  $\nu$  in (12) can be adjusted so that the generated trajectories can better meet actual
- 3 needs.

To verify the temporal statistical characteristics of the created trajectories, we used an autocorrelation coefficient  $R(h)$  to measure the degree of correlation of a trajectory and a lagged copy of itself. This coefficient at lag time  $h$  can be calculated from

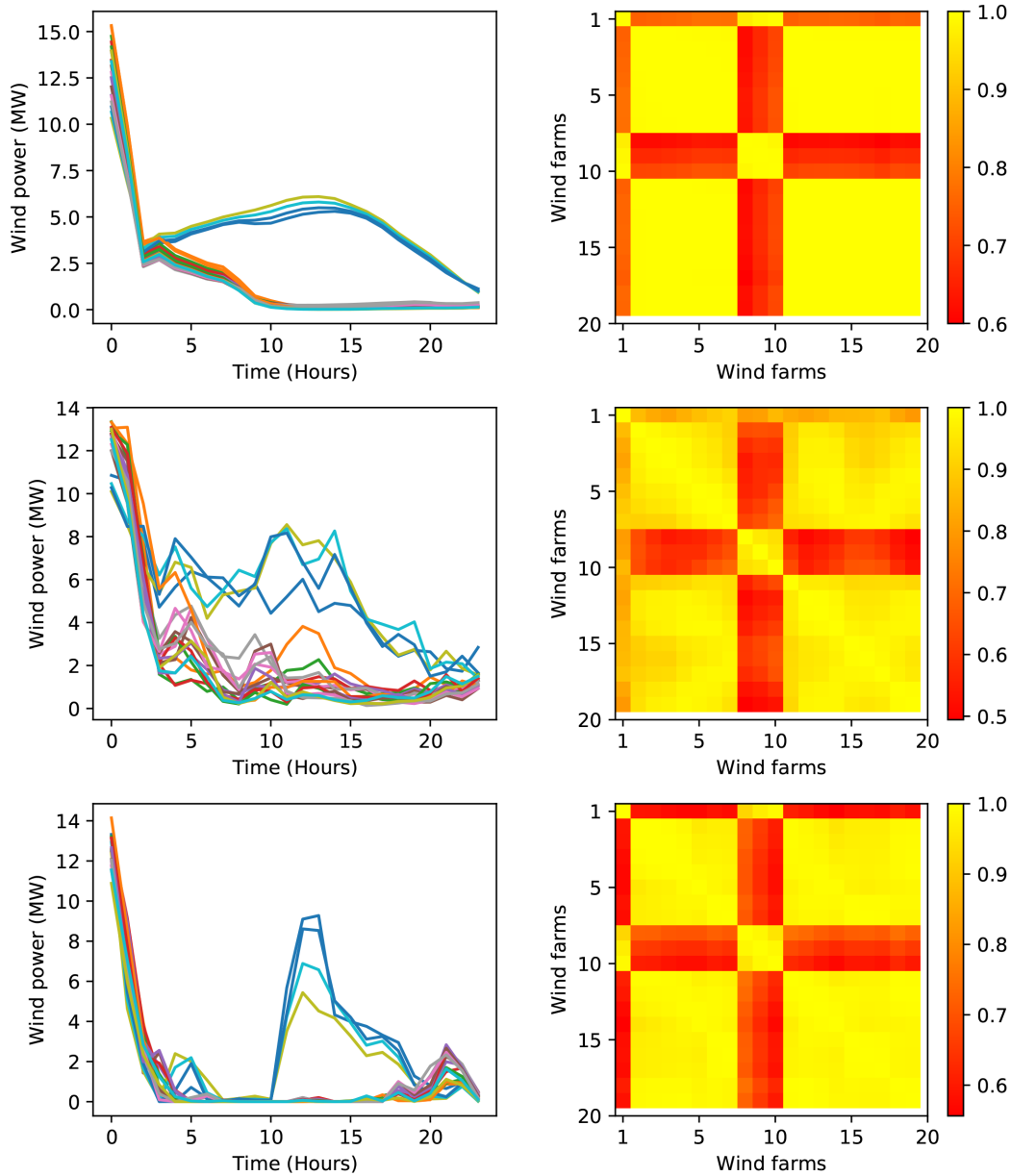
$$R(h) = \frac{E[(s_i - \mu)(s_{i+h} - \mu)]}{\sigma^2}, \quad (13)$$

- 4 where  $s$  denotes a time series with mean  $\mu$  and variance  $\sigma^2$ .

- 5 We computed and compared the autocorrelation for the measured value and generated trajectories when different prediction
- 6 interval values  $\theta$  were selected. As can be seen from Figure 6, the forecasted trajectories can represent different variations of
- 7 temporal correlation. When the parameter  $\theta$  is 3, the autocorrelation plots of the generated trajectories are able to cover the
- 8 range of the measured data, which means that the generated trajectories can represent the temporal correlation of variable power
- 9 generation.

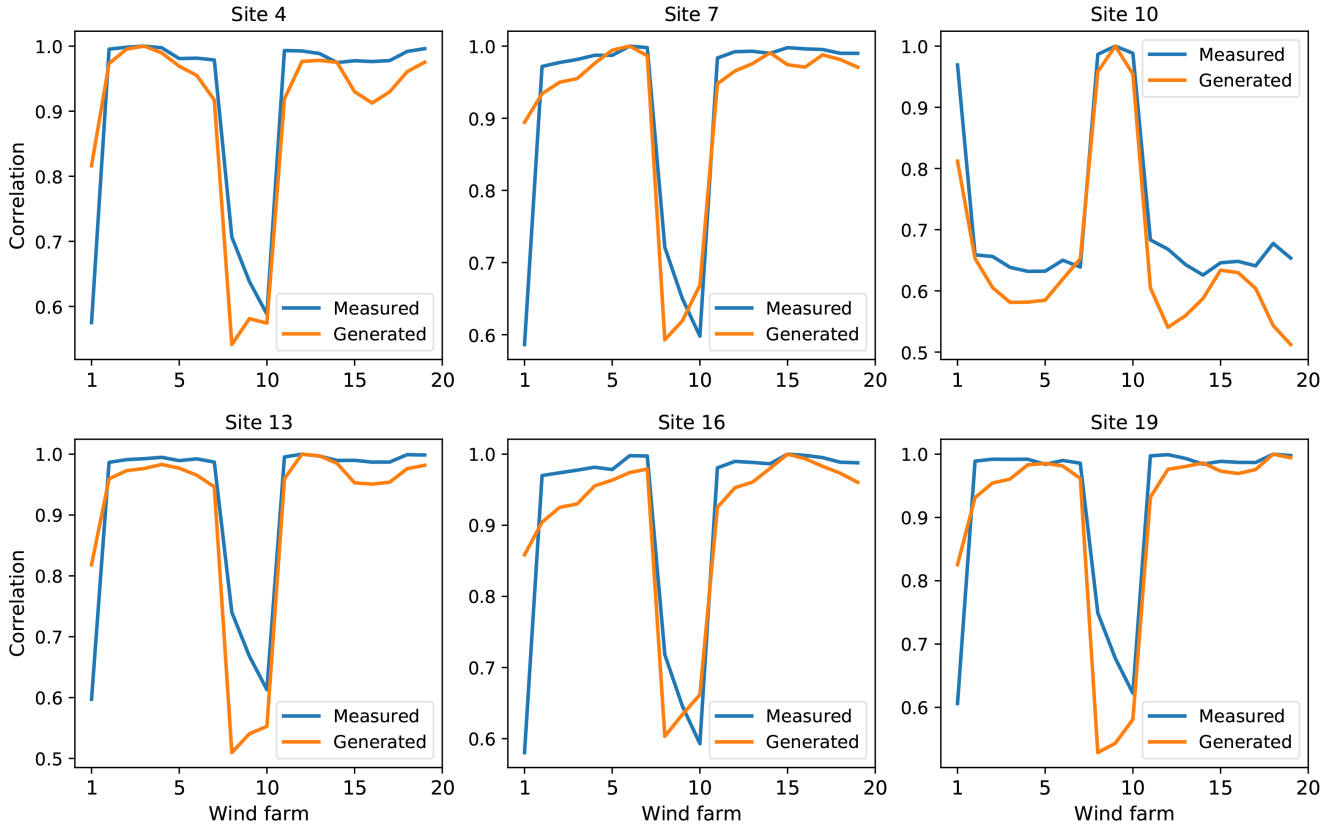
#### 10 4.4.2 | Scenario forecasts for multiple sites

- 11 To examine forecasting scenarios for multiple sites, we used a data matrix  $\{x^{(i)}\}$  of size  $N \times K$  as the historical samples input
- 12 to the generative model. Specifically, we selected  $N = 20$ ,  $K = 24$  with a resolution of one hour. After training the model to
- 13 converge, the available point forecasts can be easily incorporated to generate a large number of future scenarios. We first verified
- 14 the statistical properties for individual locations. In probability and statistics, the cumulative distribution function (CDF) of a
- 15 random variable  $X$  is the probability of getting a value less than a specific value  $x$ . We randomly generated 50 samples and



**FIGURE 8** The results of scenario forecasts for multiple wind-site data. Left: a group of predicted (top), generated (middle) and measured (bottom) data; right: their respective spatial correlation coefficient matrices.

- 1 computed the CDFs of these samples, and some of the results are plotted in Figure 7. For different wind farms, it can be found
- 2 that the generated data correctly maintain the marginal distributions of the measured values.
- 3 To show the correlation between individual locations, a group of predicted, generated and measured data is shown in Figure 8.
- 4 By visual inspection, it can be seen that they have similar behaviors and characteristics for the 20 wind farms. The spatio-temporal
- 5 correlation of the measured data is correctly preserved by the created trajectories. From the visualizations on the right of the
- 6 figure, it can be seen that the cross-site correlation coefficients of the three samples have relatively large values. This indicates
- 7 that all sites have a relatively high correlation for the plotted samples and the similarity among the upper, middle and lower
- 8 correlations suggests that the proposed method can correctly capture the stochastic dynamics for multiple power generation sites.
- 9 Besides, we can observe that there are four wind farms with different behaviors and dynamics in Figure 8. Their point forecasts
- 10 and measurements are not very consistent. This can be an explanation of why the CDF of wind farm 10 in Figure 7 is somewhat
- 11 different from those of other wind farms.



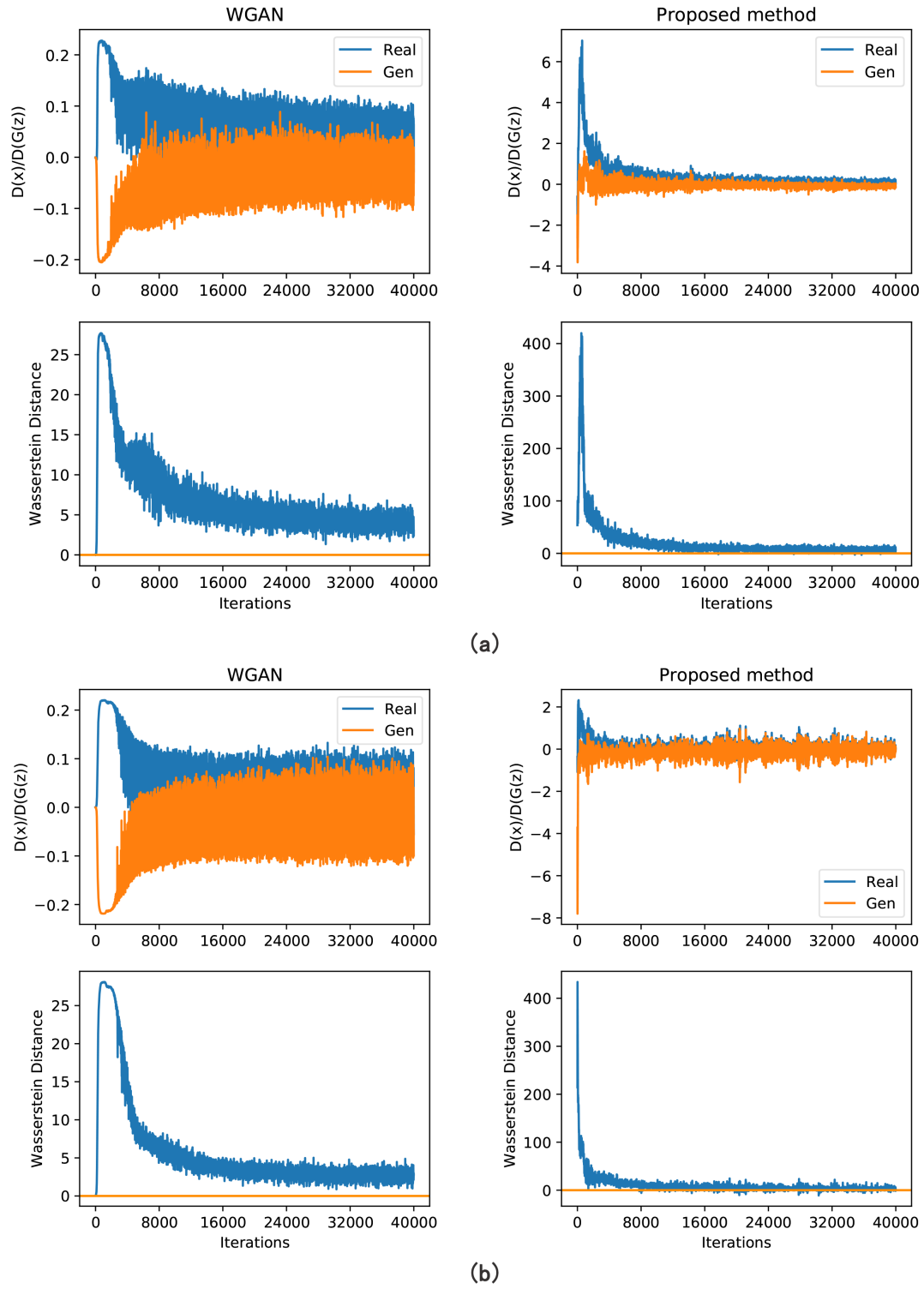
**FIGURE 9** Cross-correlation between a given site and all 20 sites for wind power data.

To further verify the spatial dynamics of multiple sites, we compared the correlation between each site and all other sites. The correlation coefficient matrix for each of the samples was computed for all different sites. In this matrix, the diagonal positions represent the auto-correlation of a site, and the other positions represent the cross-correlation between sites. Each row or column of a matrix represents the correlation between that particular site and the other sites. Figure 9 illustrates some of the elements of the correlation matrix for multiple power generation sites. It is clear that the correlation curves of the generated data are basically consistent with the measured values, demonstrating that the spatial correlation in stochastic power generation can be correctly captured.

## 4.5 | Experimental Comparison

Wasserstein distance is a meaningful index for measuring different data distributions. It can be used not only for debugging and hyperparameter searching, but also for judging the quality of generated trajectories. We assume that the Wasserstein distance can be continuously estimated during GAN training. When the model converges to the optimum, the closer the index is to zero, the higher the quality of generated trajectories.

To verify the superiority of the proposed method, we compared our method with WGAN on both wind and solar data. As shown in Figure 10, the training curves of the two methods for these two data sets have large difference at the initial stage of training. This is because the samples generated at this stage are completely different from the real samples. As the training progresses, the generator gradually learns the intrinsic characteristics of the real samples and the outputs  $D(x)$  and  $D(G(z))$  of the discriminator are close to each other. When the training converges, it can be seen from Figure 10(a) and Figure 10(b) that the Wasserstein distance of our method is closer to zero than that of WGAN on either wind or solar data, suggesting that our method can better capture the underlying patterns of real data. Therefore, the proposed method can improve the quality of generated trajectories for various prediction intervals and forecast horizons.



**FIGURE 10** Comparison of training evolution and Wasserstein distance between WGAN and the proposed method on renewable data: (a) wind power and (b) solar power.

The WGAN uses weight clipping to enforce a Lipschitz constraint in neural network settings so that the weights of the discriminator will be limited to a certain range. Since the capacity of the DNN is not fully utilized in the training process, this will slow down the training speed of the neural network. Wasserstein distance has been shown to be effective in judging the quality of generated trajectories, so it can conveniently help to indicate the training process of the model. Owing to the use of improved training strategies, from Figure 10, we can see that our method can achieve faster convergence for renewable energy sources. When trained on wind power data, as Figure 10(a) shows, our method reaches a minimum after about 13,000 iterations and the WGAN starts to stabilize after about 32,000 iterations. When trained on solar power data, as Figure 10(b) shows, our method can reach a minimum after about 8,000 iterations, while the WGAN takes nearly 24,000 iterations to reach convergence.

From the above experimental results, we can see that our method can not only improve the quality of created trajectories, but also accelerate the training speed of the generative model. It is clear that our proposed method can significantly improve the performance of the generative model for scenario forecasts.

## 5 | CONCLUSION AND FUTURE WORK

This paper presents an unsupervised distribution learning method to generate future scenarios that can characterize the temporal and spatial correlation of stochastic power generation processes. The improved training techniques detailed here can better utilize the network capacity and achieve more stable training for the generative model. The proposed method is able to characterize the uncertainty and variability characteristics of a single or multiple power generation locations. It can not only create a large number of future trajectories without relying on sampling techniques but can also capture the inherent dynamics of the stochastic power generation processes. Comprehensive case studies were used to validate the feasibility and validity of our proposed method for generating scenarios within a stochastic programming framework. The marginal distribution for each stochastic process was preserved by the data that were produced. The temporal correlation was examined at each site and the cross-correlation was verified at different geographical sites. Also, compared with the existing method for scenario forecasts, the proposed method can improve the quality of generated trajectories for various prediction intervals and forecast horizons, and achieve faster convergence for the generative model on both wind and solar data.

Since the generative model can adapt the input data of stochastic power generation dynamics to different regions only by adjusting the network structure and a few parameters, it has high flexibility in terms of the length of the forecast horizons and the number of power generation locations. With its high reliability and flexibility, our method provides a meaningful tool for characterizing the future behavior of renewable energy. Therefore, for further research, it would be interesting to apply the forecasted trajectories to the operation and planning studies on a wide range of decision-making problems in sustainable energy systems.

## References

1. Halamay D A, Brekken T K A, Simmons A, et al. Reserve requirement impacts of large-scale integration of wind, solar, and ocean wave power generation. *IEEE Trans Sustain Energy*. 2011;2(3): 321-328.
2. Ajeigbe O A, Munda J L, Hamam Y, et al. Towards maximising the integration of renewable energy hybrid distributed generations for small signal stability enhancement: A review. *Int J Energy Res*. 2020; 44(4): 2379-2425.
3. Wang J, Shahidehpour M, Li Z. Security-constrained unit commitment with volatile wind power generation. *IEEE Trans Power Syst*. 2008;23(3): 1319-1327.
4. Wang Y, Dvorkin Y, Fernandez-Blanco R, et al. Look-ahead bidding strategy for energy storage. *IEEE Trans Sustain Energy*. 2017; 8(3): 1106-1117.
5. Rathor S K, Saxena D. Energy management system for smart grid: An overview and key issues. *Int J Energy Res*. 2020; 44: 4067– 4109.
6. Pinson P, Madsen H, Nielsen H A, et al. From probabilistic forecasts to statistical scenarios of short-term wind power production. *Wind Energy*. 2009; 12(1): 51-62.

7. Pinson P, Girard R. Evaluating the quality of scenarios of short-term wind power generation. *Appl Energy*. 2012; 96: 12-20.
8. Hoyland K, Kaut M, Wallace S W. A heuristic for moment-matching scenario generation. *Comput Optim Appl*. 2003; 24(2): 169-185.
9. Meibom P, Barth R, Hasche B, et al. Stochastic optimization model to study the operational impacts of high wind penetrations in Ireland. *IEEE Trans Power Syst*. 2011; 26(3): 1367-1379.
10. Vagropoulos S I, Kardakos E G, Simoglou C K, et al. ANN-based scenario generation methodology for stochastic variables of electric power systems. *Electr Power Syst Res*. 2016; 134: 9-18.
11. Tastu J, Pinson P, Madsen H. Space-time scenarios of wind power generation produced using a Gaussian copula with parametrized precision matrix. Technical University of Denmark DTU Compute Technical Report 14.2013.
12. Villanueva D, Feijóo A, Pazos J L. Simulation of correlated wind speed data for economic dispatch evaluation. *IEEE Trans Sustain Energy*. 2012; 3(1): 142-149.
13. Hu W, Yong M, Zhou Y, et al. Wind power forecasting errors modelling approach considering temporal and spatial dependence. *J Mod Power Syst Clean Energy*. 2017; 5(3): 489-498.
14. Morales J, Mínguez R, Conejo A. A methodology to generate statistically dependent wind speed scenarios. *Appl Energy*. 2010; 87(3): 843-855.
15. Le D D, Gross G, Berizzi A. Probabilistic modeling of multisite wind farm production for scenario-based applications. *IEEE Trans Sustain Energy*. 2015; 6(3): 748-758.
16. Mirzapour F, Lakzaei M, Varamini G, et al. A new prediction model of battery and wind-solar output in hybrid power system. *J Amb Intel Hum Comp*. 2019; 10(1): 77-87.
17. Radaideh M I, Kozlowski T. Combining simulations and data with deep learning and uncertainty quantification for advanced energy modeling. *Int J Energy Res*. 2019; 43(14): 7866-7890.
18. Yao Y, Wang J, Long P, Xie M, and Wang J. Small-batch-size convolutional neural network based fault diagnosis system for nuclear energy production safety with big-data environment. *Int J Energy Res*. 2020; 44: 5841- 5855.
19. Xuan Q, Chen Z, Liu Y, Huang H, Bao G, and Zhang D. Multiview generative adversarial network and its application in pearl classification. *IEEE T Ind Electron*. 2019; 66(10): 8244-8252.
20. Ungurean L, Micea M V, Cârstoiu G. Online state of health prediction method for lithium-ion batteries, based on gated recurrent unit neural networks. *Int J Energy Res*. 2020; 44: 6767– 6777.
21. Goodfellow I J, Pouget-Abadie J, Mirza M, et al. Generative adversarial networks. *Adv Neural Inf Process Syst*. 2014; 3: 2672-2680.
22. Kingma D P, Welling M. Stochastic gradient VB and the variational auto-encoder. In: *Proceedings of the 2nd International Conference on Learning Representations*. Banff, AB; Canada; 2014.
23. Liang J, Tang W. Sequence Generative Adversarial Networks for Wind Power Scenario Generation. *IEEE J Sel Area Comm*. 2020; 38(1): 110-118.
24. Chen Y, Wang Y, Kirschen D S, et al. Model-Free Renewable Scenario Generation Using Generative Adversarial Networks. *IEEE T Power Syst*. 2018; 33(3): 3265-3275.
25. Zhang Y, Ai Q, Xiao F, et al. Typical wind power scenario generation for multiple wind farms using conditional improved Wasserstein generative adversarial network. *Int J Elec Power*. 2020; 114.
26. Hu W, Zhang H, Dong Y, et al. Short-term optimal operation of hydro-wind-solar hybrid system with improved generative adversarial networks. *Appl Energy*. 2019: 389-403.

27. Jiang C, Mao Y, Chai Y, et al. Scenario generation for wind power using improved generative adversarial networks. IEEE Access. 2018; 6: 62193-62203.
28. Pan Z, Wang J, Liao W, et al. Data-Driven EV Load Profiles Generation Using a Variational Auto-Encoder. Energies. 2019; 12(5).
29. Zhang H, Hu W, Yu R, et al. Optimized operation of cascade reservoirs considering complementary characteristics between wind and photovoltaic based on variational auto-encoder. MATEC Web Conf. 2018; 246: 01077.
30. Dinh L, Krueger D, Bengio Y, et al. NICE: Non-linear Independent Components Estimation. arXiv preprint. 2014; 1410.8516.
31. Kingma D P, Dhariwal P. Glow: Generative flow with invertible 1\*1 convolutions. in: Proc Adv Neural Inf Process Syst. 2018; 10236–10245.
32. Zhang L, Zhang B. Scenario forecasting of residential load profiles. IEEE J Sel Areas Commun. 2020; 38(1): 84-95.
33. Ge L, Liao W, Wang S, Bak-Jensen B and Pillai J R. Modeling Daily Load Profiles of Distribution Network for Scenario Generation Using Flow-Based Generative Network. IEEE Access. 2020; 8: 77587-77597.
34. Chen Y, Wang X, Zhang B. An unsupervised deep learning approach for scenario forecasts. In: Proceedings of 2018 Power Systems Computation Conference. Dublin, Ireland; 2018; 1-7.
35. Du C, Li J, Huang L, et al. Brain Encoding and Decoding in fMRI with Bidirectional Deep Generative Models. Engineering. 2019; 5(5): 948-953.
36. Creswell A, White T, Dumoulin V, et al. Generative adversarial networks: an overview. IEEE Signal Process Mag. 2018; 35(1): 53-65.
37. Liu H, Zhou J, Xu Y, et al. Unsupervised fault diagnosis of rolling bearings using a deep neural network based on generative adversarial networks. Neurocomputing. 2018; 315: 412-424.
38. Cheng L, Yu T. A new generation of AI: A review and perspective on machine learning technologies applied to smart energy and electric power systems. Int J Energy Res. 2019; 43(6): 1928-1973.
39. Arjovsky M, Chintala S, Bottou L. Wasserstein GAN. arXiv preprint. 2017; 1701.07875.
40. Gulrajani I, Ahmed F, Arjovsky M, et al. Improved training of Wasserstein GANs. arXiv preprint. 2017; 1803.01541.
41. Villani C. Optimal transport: old and new. Springer Science & Business Media. 2008; 338.
42. Wei X, Gong B, Liu Z, et al. Improving the improved training of Wasserstein GANs: a consistency term and its dual effect. arXiv preprint. 2018; 1803.01541.
43. Wan C, Xu Z, Pinson P, et al. Probabilistic forecasting of wind power generation using extreme learning machine. IEEE Trans Power Syst. 2014; 29(3): 1033-1044.
44. Draxl C, Clifton A, Hodge B-M, et al. The wind integration national dataset (wind) toolkit. Appl Energy. 2015; 151: 355-366.
45. Kingma D P, Ba J L. Adam: a method for stochastic optimization. In: Proceedings of International Conference on Learning Representations. San Diego, CA; USA; 2015; 1-13.
46. Abadi M, Agarwal A, Barham P, et al. Tensorflow: large-scale machine learning on heterogeneous distributed systems. arXiv preprint. 2016; 1603.04467.

