



# Data-driven scenario generation of renewable energy production based on controllable generative adversarial networks with interpretability

Wei Dong, Xianqing Chen, Qiang Yang\*

*College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China*

## HIGHLIGHTS

- A controllable GAN model with interpretability is proposed for renewable scenario generation.
- Interpretable features with physical meanings are designed on latent manifold space.
- Mutual information maximization and imitation learning sampling are developed.
- Scenario characteristics can be manually controlled to generate new patterns.

## ARTICLE INFO

### Keywords:

Renewable scenario generation  
Generative adversarial networks  
Interpretable latent space  
Controllable scenario generation

## ABSTRACT

Efficient and reliable scenario generation is of paramount importance in the modeling of uncertainties and fluctuations of wind and solar based renewable energy production for power system planning and operation in the presence of highly penetrated renewable sources. This paper proposes a data-driven method for renewable scenario creation by embedding interpretable manifold space in controllable generative adversarial networks (GAN). Without the need for laborious probabilistic modeling and sampling procedures, the proposed machine learning-based model can adaptively understand the inherent stochastic and dynamic characteristics of renewable resources. The generation of renewable patterns can be deliberately modified by embedding characteristic features with interpretability in latent input space. To address the controllable generation, the mutual information maximization and imitation learning sampling techniques are developed and incorporated into the existing GAN networks. The proposed approach is verified by the real-time series data of wind and solar energy generation profiles. The numerical results demonstrate that the proposed solution can achieve the controllable generation of scenarios covering various statistical characteristics and even create new generation patterns that are different from previous scenarios.

## 1. Introduction

Climate change and environmental pollution, along with the conventional energy crisis, are driving the development of more sustainable, clean, and efficient energy systems with increasing renewable energy sources (RESs) [1]. However, the fluctuation and intermittent nature of renewable energy power have a significant impact on the stable operation and economic management of power systems. Low integrated energy efficiency and insufficient interaction between supply and demand are the common problems existing in renewable energy systems [2]. Thus, building flexible and intelligent energy systems to efficiently use multiple renewable energy sources is a focus of applied

energy research. For example, the regional integrated energy system (RIES) involving multi-energy sources with their complementarity can improve the utilization efficiency with a high proportion of distributed generations (DGs) [3]. The virtual power plant (VPP) can coordinate the aggregated plants by integrating flexible consumers or energy storage systems (ESS) from static storages and electric vehicles (EVs) to participate in ancillary services [4,5].

However, for these renewable energy generation systems, the uncertainty and variability characteristic of high-level RESs pose huge challenges in operation (such as economic dispatch [6], unit commitment [7] and management of ESS [8]) and planning (such as RIES planning [9] and transmission expansion planning [10]). The operators and planners require large quantities of accurate information about their

\* Corresponding author.

E-mail address: [qyang@zju.edu.cn](mailto:qyang@zju.edu.cn) (Q. Yang).

<b>Nomenclature</b>		$D_{KL}$	Kullback-Leibler divergence
<i>Sets and Indices</i>		$Q(\cdot)$	Auxiliary distribution to approximate posterior distribution
$G$	Generator network of generative adversarial networks	$E$	Expectation of the probability distribution
$D$	Discriminator network of generative adversarial networks	$S$	Skewness of the probability distribution
$Q$	Auxiliary network of controllable generative adversarial networks	$K$	Kurtosis of the probability distribution
		$R$	Autocorrelation coefficient of the time series samples
		$\mu$	Mean value of the stochastic variable
		$\sigma$	Standard deviation of the stochastic variable
		$\tau$	Time-lag of the time series samples
<i>Parameters</i>		<i>Abbreviation</i>	
$\alpha$	Learning rate for training process	GAN	Generative adversarial networks
$\beta_1, \beta_2$	Exponential decay rate of adaptive moment estimation algorithm	RESs	Renewable energy sources
$m$	Batch size of stochastic gradient descent process	RIES	Regional integrated energy system
$N_{loop}$	Number of iterations for discriminator per generator iteration	DGs	Distributed generations
$N_{iter}$	Number of iterations for training process	VPP	Virtual power plant
<i>Variables and Functions</i>		ESS	Energy storage systems
$P_t$	Renewable power generation at slot $t$	EVs	Electric vehicles
$P_{ave}$	Average power of renewable generation with a period of time $T$	MCS	Monte Carlo sampling
$P_{max}$	Maximum power of renewable generation with a period of time $T$	LHS	Latin Hypercube sampling
$P_{min}$	Minimum power of renewable generation with a period of time $T$	ARMA	Autoregressive moving average
$P_{fluct}$	Average fluctuating power of renewable generation with a period of time $T$	SS	State space
$P_{climb}$	Maximum climbing power of renewable generation with a period of time $T$	ANN	Artificial neural network
$P_G(\cdot)$	Generator distribution of Generative Adversarial Networks	RBFNN	Radial basis function neural networks
$P_{data}(\cdot)$	Real distribution of historical data	PSO	Particle swarm optimization
$P_{noise}(\cdot)$	Gaussian distribution of random noise variable	NWP	Numerical weather predictions
$x$	Real samples from historical data	cGAN	Conditional generative adversarial networks
$c$	Control vector of latent space	SGAN	Semi-supervised generative adversarial networks
$\tilde{c}$	Characteristic indicators close to control vector	MI	Mutual information
$\tilde{x}$	Generated samples according to control vector	VAE	Variational auto-encoder
$I(\cdot)$	Mutual information function	BNN	Bayesian neural networks
$H(\cdot)$	Entropy function	Adam	Adaptive moment estimation
		DCGAN	Deep convolutional neural network
		MLP	Multilayer perceptron layers
		FC	Fully connected layer
		KDE	Kernel density estimation

forecasting trajectories and probabilistic distributions to arrive at effective decisions for energy systems [11,12]. The decisions are generally derived from a stochastic optimization that relies on scenarios of aggregated RES production to take uncertainties into account. Thus, the quality of the scenarios and the selection of uncertainty representation have a direct impact on the performance of the optimization model [13]. To ensure the efficiency and reliability of increasingly intelligent energy systems, the inherent characteristics of the intermittent and stochastic renewables should be more accurately and effectively described to provide support for more advanced operation and planning methods.

In literature, the widely adopted technique for capturing the uncertainties in RESs is to employ a set of time-series scenarios [14]. To obtain an accurate representative distribution, many generation models have been broadly implemented to reflect the inherent patterns for renewable power time-series. The models should generate new scenarios with enough diversity that are similar but not identical to what has already been encountered [15]. The commonly used approaches for renewable scenario generation can be split into two categories: parametric and non-parametric methods.

In general, the parametric model-based scenario generation follows a two-step approach: first obtaining the probabilistic distribution or forecast error of renewable power, and then scenarios are generated from the statistical distribution using a sampling method such as Monte

Carlo sampling (MCS) or Latin Hypercube sampling (LHS) [16]. Regarding the first step, several families of probabilistic models have been suggested to characterize the probability distribution in previous research. For example, the empirical cumulative distribution function, e.g., Gaussian distributions [17] and beta distributions [18], can be used to model the uncertainties and forecast error [19]. Alternatively, copula techniques are often used to model the parameterized distribution to capture the dependence structures in stochastic processes. To capture the time dependence for wind prediction error series, the copula function was adopted in [20]. In [21], a Gaussian copula was adopted to analyze the uncertainty of large-scale wind power integration.

Another common kind of parametric method is one that exploits time series to characterize interdependence between time steps. In [14], the autoregressive moving average (ARMA) model is used to generate spatiotemporal scenarios with given power generation profiles by assuming wind samples have linear dependence. The authors in [22] investigated the ARIMA model in state-space (SS) and then extended the SS model to include correlated wind speeds across different locations. The method was enhanced in reference [23] by using an artificial neural network (ANN) approximated by a normal distribution to capture the non-linear dependencies and create representative scenarios. Although easy to implement, with simple statistical assumptions, the parametric model-based approaches are prone to over-fitting and misidentification of patterns.

Non-parametric approaches are distribution-free and have been extensively studied in recent years. The moment matching [24] and distance matching [25] methods were applied to generate scenarios by minimizing the distance indexes between the expected properties and outcomes. However, this method will eliminate small-probability but high-risk scenarios [16]. In recent years, several machine learning algorithms are also proposed for scenario generation. In [26], a radial basis function neural networks (RBFNN) algorithm is combined with the particle swarm optimization (PSO) method to generate scenarios using numerical weather predictions as input (NWP). The study in [27] trained the neural network to model either stochastic process or ramp events probability of wind power. The machine learning-based algorithms may characterize the nonlinear dynamics of renewable generating scenarios better than copula or time series techniques. However, the performance of these methods firmly relies on the careful selection of input features, and hence are not sufficiently flexible and reliable in practice [15].

Another typical type of nonparametric method is based on generative adversarial networks, GAN [28], which are known as a set of innovative generative models for renewable scenarios generation and have gained a lot of attention in recent years. By learning the distribution of historical samples, reference [15] develops a GAN-based method for renewable scenario generation considering spatiotemporal correlations of multiple sites. In [29], Bayesian GAN is developed and trained to capture different patterns of historical renewable data. In [30], an unsupervised method for scenario forecasting based on GAN is suggested to create a set of future realizations. To enforce the Lipschitz constraint, the GAN is improved in [31] with an additional gradient penalty term. Deep generative models can directly learn to capture the uncertainties and the spatiotemporal relationships from historical data in a semi-supervised manner without cumbersome manual labeling. Compared with other generative models, these GAN-based methods can provide a more accurate generation process by reflecting the dynamic characteristics of renewable resources with a full diversity of patterns.

GAN-based renewable scenario generation remains a challenging problem, although it can be easily applied to diverse sets of renewable profiles. The patterns of the generated scenarios cannot be controlled because the generative network uses random noise as input. The relationship between the latent space and the output results is not interpretable and weakly correlated. In practice, we often cannot obtain a complete and sufficient amount of historical power data of renewable energy. Seasonal differences cause current data to fail to cover uncaptured scenario patterns. However, information such as geography, climate and local experience can depict some key characteristics of renewable energy profiles. Taking northern China with continental monsoon climate in spring as an example, wind power generation has the characteristics of large average output, obvious peak-valley difference and strong fluctuation. These interpretable features can be further used as transfer knowledge for the scenario generation with unknown patterns.

The conditional GAN (cGAN) [32] can overcome the above problem to some extent. However, the classified feature cannot be tuned continuously within each condition. The number of the conditions cannot be set as too many due to the limitation of the discrete classification. The whole GAN model has to be retrained if new tunable features are added to the generation process. The input feature of the existing GAN-based model cannot directly guide the generation of patterns because the latent space is not interpretable.

To the authors' best knowledge, the data-driven scenario generation solution that leverages interpretable features to directly control the generating patterns has not been reported yet in the existing literature. In this paper, a novel generative adversarial network with a latent controllable vector in manifold space is proposed for scenario generation of renewable energy production. The main contributions of this paper can be summarized as follows:

- 1) *Distribution-free Scenario Generation.* The GAN-based model with deep neural networks is data-driven, which can adaptively learn the inherent stochastic distribution and dynamic correlation hiding in historical data of renewable scenarios. Thus, it can be easily scaled up to various sets of renewable profiles.
- 2) *Tunable Features with Interpretability.* The construction of the network input space contains interpretable features as control vector on the manifold. These tunable features have real meanings and can directly reflect the characteristics of scenario patterns.
- 3) *Controllable Scenario Generation Process.* Mutual information maximization and imitation learning sampling techniques are developed to address the controllable generation process. The generated scenarios can be deliberately modified for statistical characteristics and generate new patterns unknown to the historical data.

The rest of the work is organized as follows: Section 2 presents the preliminaries of interpretable characteristics and generative adversarial networks; Section 3 presents the proposed controllable scenario generation solution in details; Section 4 evaluates the proposed solution through simulation experiments and presents a set of numerical results; finally, the conclusive remarks are given in Section 5.

## 2. Preliminaries

### 2.1. Renewable time-series characteristic interpretable analysis

Renewable scenario generation is generally considered as the generation of time series that represents the possible output patterns of renewable energy sources over a period of time (e.g., one day). Therefore, it is important to make a time-series analysis from the existing historical samples. Currently, the main approach to describe the output of renewable energy profile is to design the appropriate statistical indicators. However, for data-driven generation methods, statistical indicators (e.g., Expectation, Variance, Skewness, Kurtosis, and Correlation) are usually used for model validation and evaluation after the scenario generation process. The model needs to iteratively adjust the parameters or structure of the network to obtain reasonable generation scenarios. This method cannot directly pre-control the characteristics of the generation scenarios, so it is difficult to guide the generation preference of different patterns.

To enable the expected scenario patterns to be pre-depicted with interpretability before generation, the characteristic indicators should be designed to have clear and comprehensible realistic meanings, which can be used to guide the preference of generations through expert experience. Using these characteristic indicators as control vector in latent space, the results of generation can be controlled with interpretability. This work selects five typical characteristic indicators that can comprehensively describe scenarios over different patterns. For renewable resources generation with a period of time  $T$  (usually one day), the continuous out power sequence is denoted as  $\{P_t\}$ ,  $t = 1, 2, \dots, T$ .

- 1) Average power:  $P_{ave} = \sum_{t=1}^T P_t / T$ , it directly reflects the average power generation level of renewable energy resources under this scenario pattern.
- 2) Maximum power:  $P_{max} = \max\{P_t\}$ ,  $t = 1, 2, \dots, T$ , it reflects the maximum power generation over a period of time.
- 3) Minimum power:  $P_{min} = \min\{P_t\}$ ,  $t = 1, 2, \dots, T$ , it reflects the minimum power generation, and then the peak-valley difference of power over a period of time can be easily calculated.
- 4) Average fluctuating power:  $P_{fluct} = \sum_{t=2}^T (|P_t - P_{t-1}|) / T - 1$ , the average fluctuating power is the average absolute value of the climbing power in a period of time, which reflects the fluctuation of renewable energy power generation.
- 5) Maximum climbing power:  $P_{climb} = \max\{|P_t - P_{t-1}|, \forall t \in \{2, 3, \dots, T\}\}$ , the climbing power reflects the fluctuating quantity

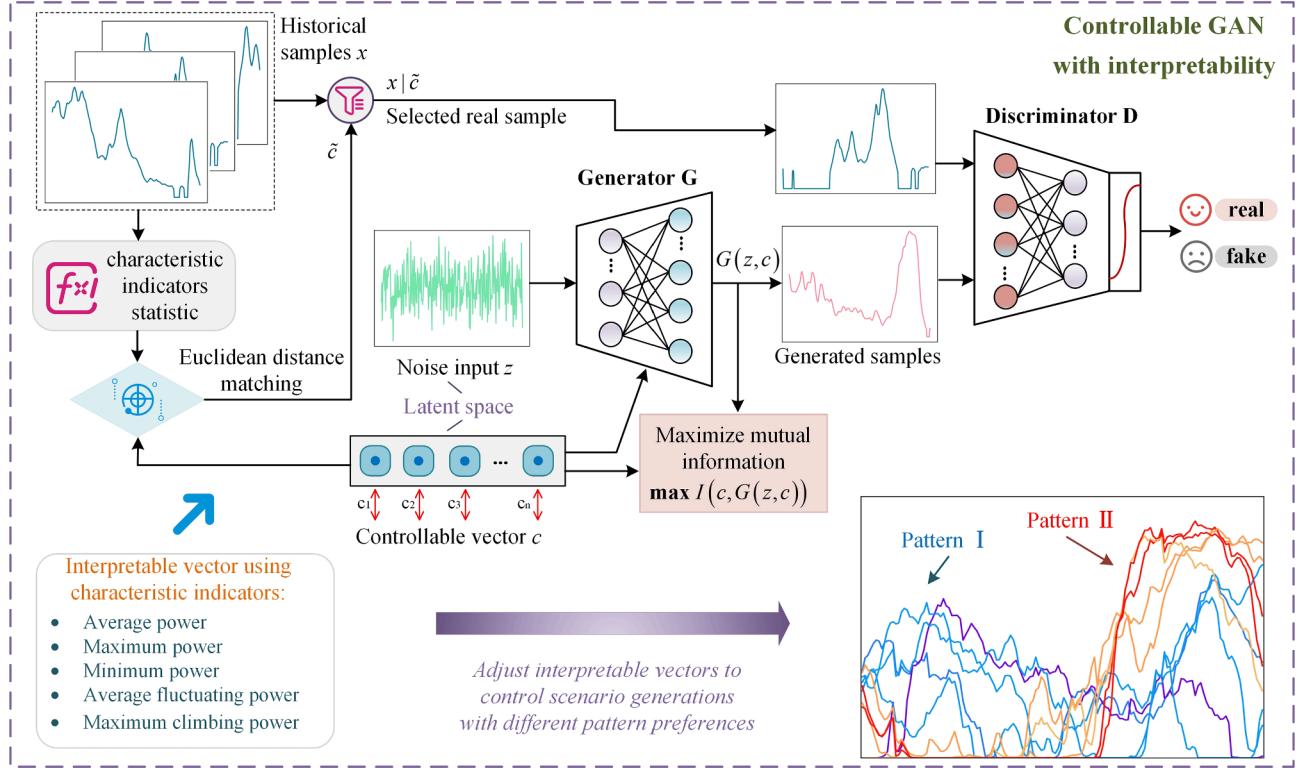


Fig. 1. Framework of controllable GAN with interpretability for scenario generation.

of renewable power generation in adjacent periods, and the maximum climbing power is the maximum absolute value of the climbing power in a period of time. This characteristic indicator provides important reference information for setting the system reserves and making the scheduling plans.

## 2.2. Basic generative adversarial networks

Generative Adversarial Networks (GAN), introduced by Goodfellow et al. [28], is a famous generative framework for training deep networks based on minimax game theory. It aims to learn a generator distribution  $P_G(x)$  that matches the real data distribution  $P_{data}(x)$ . Rather than attempting to assign the probability to each  $x$  explicitly in the data distribution, GAN trains a generator network  $G$  that generates samples from the generator distribution  $P_G$  by transforming a random noise variable  $z \sim P_{noise}(z)$  into a sample  $G(z)$ . This generator is trained by competing against an adversarial discriminator network  $D$ , which attempts to differentiate between samples from the real data distribution  $P_{data}$  and samples from the generator's distribution  $P_G$ . As a result, the optimal discriminator for a given generator is  $D(x) = P_{data}(x)/(P_{data}(x) + P_G(x))$ . The well-trained generator network  $G$  can deceive the discriminator into regarding its outputs as true data.

The process is to iteratively train these two networks with high dimensional parameters. The game can finally reach a Nash equilibrium [33,34] between the generator network  $G$  and the discriminator network  $D$ . The notable application of Nash equilibrium in the field of neural networks is GAN [28] and their variants, e.g., cGAN [32], Semi-supervised GAN (SGAN) [35]. The minimax game in GAN is given in the expression (1):

$$\min_{G} \max_{D} V(D, G) = \mathbb{E}_x P_{data} [\log D(x)] + \mathbb{E}_z P_{noise} [\log (1 - D(G(z)))] \quad (1)$$

## 3. Controllable GAN with interpretability

### 3.1. Framework and inspiration

Typical generative adversarial networks mostly belong to unsupervised learning, which generates corresponding high-dimensional data through latent vectors (i.e. random noise data). The advantage of unsupervised learning is the ability to categorize the disordered data at the machine-level, which is very valuable for data analysis. An example of unsupervised learning is the auto-encoder, which is trained to encode high-dimensional data through neural networks so that the latent space of low-dimensional coding can extract the true information from the original data. Experiments on existing works [36,37] have revealed that there are two good characteristics in the latent space: (1) Most points in the manifold can produce reasonable data; and (2) Due to the continuous property, the interpolation between two points on the manifold can still lead to a smooth transition of the output data [36].

It should be highlighted that unsupervised learning solutions may not be effective for certain tasks. The pattern preferences cannot be implemented during the scenario generation based on unsupervised learning as the data representation in latent space lacks of interpretability. As a result, the characteristics of the generated scenarios cannot be manipulated as the scenario patterns only rely on the distribution of the available historical data [15,30]. However, the studies of power system planning and management need to be carried out considering extensive operational scenarios (e.g., worst-case scenarios). Thus, the scenario generation is expected to incorporate certain preferences or generate new patterns based on the expert knowledge, which requires a controllable scenario generation approach with interpretability. The solution proposed in this work can pre-describe the characteristics of the expected scenarios through interpretable features, so as to control the preference of the scenario patterns and further generate unknown distributions using the transferred domain knowledge.

To make the input features interpretable and to utilize the controllability of the latent space on the manifold, the traditional GAN random

input is divided into two parts: the first part is  $z$ , which represents random noise. The second part  $c$ , representing the controllable latent vector consisting of a set of interpretable features that can be selected as shown in Section 2.1. To ensure that GAN does not ignore the controllable latent coding information  $c$  on the manifold, referring to the existing work [37], the concept of mutual information (MI) is introduced to evaluate the relationship between latent vector  $c$  and generated distribution  $G(z, c)$ . Maximizing the MI between them can reduce the information loss of controllable latent vector  $c$  in the generation process. Meanwhile, discriminator  $D$  needs to distinguish the truth and the falsehood to ensure that the generated data is very close to the real data.

The aforementioned unsupervised learning framework combined maximization of MI only ensures that latent space on the manifold can guide and impact the generation of results. To realize the precise and quantitative guidance of the controllable vector to generate different patterns, the model needs to further clarify the implicit mapping between the interpretable features and the corresponding patterns of the generation scenarios through adding a supervised learning process. The actual physical meaning of the control vector comes from the characteristic indicators in Section 2.1, which can also be used as the interpretable features for selecting historical samples. Thus, the selecting real data  $x$  and generation data  $G(z, c)$  will be used together for discrimination.  $\tilde{c}$  is the characteristic indicators most similar to the control vector  $c$  (i.e., the smallest Euclidean distance). Selecting training samples through the interpretability of control vector, the principle of this fashion comes from imitation learning, which can avoid the limited number of classification labels like the condition GAN. The framework of controllable GAN with interpretability is shown in Fig. 1.

### 3.2. Mathematics and algorithm

Mutual information between  $X$  and  $Y$  in information theory quantifies the “amount of information” derived through knowledge of random variable  $Y$  about the other random variable  $X$ . Mutual information is defined as the difference between two entropy terms in (2):

$$I(X; Y) = H(X) - H(X|Y) \\ = \left[ - \sum_{i=1}^N p(x_i) \log(p(x_i)) \right] - \left[ - \sum_{i=1}^N \sum_{j=1}^M p(x_i, y_j) \log(p(x_i|y_j)) \right] \quad (2)$$

where  $H(X)$  is the Shannon entropy [38] of a discrete random variable  $X = (x_1, x_2, \dots, x_N)$ ;  $x_i$  refers to the possible values that  $X$  can take for discrete variable or the possible value range for continuous variable;  $p(x_i)$  is the probability mass function; the conditional entropy  $H(X|Y)$  is the amount of uncertainty left in  $X$  when a variable  $Y = (y_1, y_2, \dots, y_M)$  is introduced, so it is less than or equal to the entropy of both variables.

This definition is straightforward to comprehend:  $I(X, Y)$  is the decrease of uncertainty in  $X$  when  $Y$  is observed. If  $X$  and  $Y$  are independent, then  $I(X, Y) = 0$ . Conversely, if  $X$  and  $Y$  are linked by a deterministic relationship, mutual information should be maximized.

Based on the basic GAN described in Section 2.2, and similar mutual information inspired objectives have been considered in InfoGAN [37]. The minimax game of the proposed controllable GAN with interpretability is given in the expression (3):

$$\min_{G} \max_{D} V(D, G) = \mathbb{E}_x P_{data} [\log D(x|c)] + \mathbb{E}_z P_{noise} [\log(1 - D(G(z, c)))] - \lambda I(c; G(z, c)) \quad (3)$$

where approximate features  $\tilde{c} = \text{argmin}\{|F - c|\}$  are obtained by matching the characteristic indicators  $F$  of the historical samples with the control vector  $c$  according to Euclidean distance. The characteristic indicators  $F = \{P_{ave}, P_{max}, P_{min}, P_{fluct}, P_{climb}\}$  with interpretability are described in Section 2.1.

The mutual information term  $I(c; G(z, c))$  is difficult to maximize

directly since it requires access to the posterior  $P(c|\tilde{x})$ ,  $\tilde{x} \sim G(z, c)$ . In practice, it is feasible to get a lower bound of it by defining an auxiliary distribution  $Q(c|\tilde{x})$  to approximate  $P(c|\tilde{x})$ , as shown in (4):

$$I(c; G(z, c)) = H(c) - H(c|G(z, c)) \\ = \mathbb{E}_{x \sim G(z, c)} [\mathbb{E}_{\tilde{c} \sim P(c|x)} [\log P(c|\tilde{x})]] + H(c) \\ = \mathbb{E}_{x \sim G(z, c)} [D_{KL}(P(\cdot|x)||Q(\cdot|x))] + \mathbb{E}_{\tilde{c} \sim P(c|x)} [\log Q(c|\tilde{x})] + H(c) \quad (4)$$

Kullback-Leibler (KL) divergence [39] is used to measure the loss of information when a real distribution is represented by an approximate distribution, and its applications include variational auto-encoder (VAE) [40], Bayesian neural networks (BNN) [41], and so forth. AS KL divergence  $D_{KL}(P(\cdot|\tilde{x})||Q(\cdot|\tilde{x})) \geq 0$ , the lower bound of mutual information is obtained, known as Variational Information Maximization [42] in (5):

$$I(c; G(z, c)) = \mathbb{E}_{x \sim G(z, c)} [D_{KL}(P(\cdot|x)||Q(\cdot|x))] + \mathbb{E}_{\tilde{c} \sim P(c|x)} [\log Q(c|\tilde{x})] + H(c) \quad (5)$$

According to the lemma derived in reference [37], it can be obtained:  $\mathbb{E}_{x \sim X, y \sim Y|x} [f(x, y)] = \mathbb{E}_{x \sim X, y \sim Y|x} [f(\tilde{x}, y)]$ . By using this lemma, we can define a variational lower bound  $L_I(G, Q)$  of the mutual information  $I(c; G(z, c))$  in (6):

$$L_I(G, Q) = \mathbb{E}_{c \sim P(c), x \sim G(z, c)} [\log Q(c|x)] + H(c) \\ = \mathbb{E}_{x \sim G(z, c)} [\mathbb{E}_{\tilde{c} \sim P(c|x)} [\log Q(c|\tilde{x})] + H(c)] \\ \leq I(c; G(z, c)) \quad (6)$$

where  $L_I(G, Q)$  is easy to approximate with Monte Carlo simulation and can be maximized w.r.t.  $Q$  directly and w.r.t.  $G$  via the reparametrization trick. In practice, reference [37] proposes that the auxiliary distribution  $Q$  can be parametrized as a neural network. As an outcome, the infoGAN is defined as the minimax game with a variational regularization of mutual information and a hyperparameter  $\lambda$  in (7):

$$\min_{G, Q} \max_D V(D, G, Q) = \mathbb{E}_{x \sim P_{data}} [\log D(x|c)] + \mathbb{E}_{z \sim P_{noise}} [\log(1 - D(G(z, c)))] - \lambda L_I(G, Q) \quad (7)$$

The proposed model is described in the following steps:

- 1) The wind or photovoltaic power historical data are prepared as the training set. The latent space on the manifold is constructed, in which the noise variables  $z$  are sampled from the Gaussian distribution, and the control vector  $c$  are the characteristic indicators with interpretable physical meanings.
- 2) The control vector are set manually in batches. The real sample whose characteristic indicators were closest to the control vector was selected by matching minimum distance.
- 3) The generator  $G$ , discriminator  $D$  and auxiliary network  $Q$  are trained in turns for each epoch until the criterion is reached. The objective of the minimax game combined mutual information is defined in (7).
- 4) In the generation process, the controllable vector are adjusted along one or more interpretable features to achieve the scenario generation of predetermined patterns.

The aforementioned networks in the proposed model were trained by backpropagation through the *Adaptive Moment Estimation* (Adam) algorithm [43]. The Adam-based optimization algorithm tracks the average of the exponential decay of the past gradient and its square, as shown in Algorithm 1:

---

**Algorithm 1. Adam: Adaptive Moment Estimation [43].**


---

**Require:**  $\alpha$ : Learning rate (or called stepsize)  
**Require:**  $\beta_1, \beta_2 \in [0, 1]$ : Exponential decay rates for the moment estimates  
**Require:**  $f(\theta)$ : Stochastic objective function with parameters  $\theta$   
**Require:**  $\theta_0$ : Initial parameter vector  
 $m_0 \leftarrow 0$ (Initialize 1st moment vector)  
 $v_0 \leftarrow 0$ (Initialize 2nd moment vector)  
 $t \leftarrow 0$ (Initialize timestep)

(continued on next page)

(continued)

**Algorithm 1.** Adam: Adaptive Moment Estimation [43].

---

```

while  $\theta_t$  not converged do
     $t \leftarrow t + 1$ 
     $g_t \leftarrow \nabla_{\theta} f_t(\theta_{t-1})$  (Get gradients w.r.t. stochastic objective at timestep  $t$ )
     $m_t \leftarrow \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t$  (Update biased first moment estimate)
     $v_t \leftarrow \beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot g_t^2$  (Update biased second raw moment estimate)
     $\hat{m}_t \leftarrow m_t / (1 - \beta_1^t)$  (Compute bias-corrected first moment estimate)
     $\hat{v}_t \leftarrow v_t / (1 - \beta_2^t)$  (Compute bias-corrected second raw moment estimate)
     $\theta_t \leftarrow \theta_{t-1} - \alpha \cdot \hat{m}_t / (\sqrt{\hat{v}_t} + \epsilon)$  (Update parameters)
end while
return  $\theta_t$  (Resulting parameters)

```

---

The control vector  $c$  is made up of the selected interpretable features, as described in Section 2.1. In the training process, due to the need for a large amount of input data (including the noise variable  $z$  and the control vector  $c$ ), the control vector values can be obtained by calculating the characteristic indicators from the historical samples. Once the model training is completed, the generation of renewable patterns can be deliberately modified by adjusting the control vector  $c$  of the generator  $G$  input space. The training process of the proposed solution is given in Algorithm 2.

**Algorithm 2.** Training Process of the Proposed Scenario Generation Model.

---

```

Require: Learning rate  $\alpha$ , batch size  $m$ , Adam hyperparameters  $\beta_1$  and  $\beta_2$ , number of iterations for discriminator per generator iteration  $N_{loop}$ , number of iterations  $N_{iter}$ .
Require: Initial weights  $\theta^{(D)}$  for discriminator,  $\theta^{(G)}$  for generator and  $\theta^{(Q)}$  for auxiliary network  $Q$ 

for  $N_{iter}$  training iterations do
    for  $t = 0, 1, \dots, N_{loop}$  do
        # Update parameter for Discriminator
        Sample data  $x$  from historical data  $P_{data}$ ; Noise variable  $z$  from Gaussian distribution  $P_{noise}$ ; Controllable vector  $c$  from characteristic indicators  $F$  of training sample data  $x$ ; Selected features  $\tilde{c}$  from distance matching.
        # Update parameter for Discriminator
         $\theta^{(D)} \leftarrow Adam \left( \nabla_{\theta^{(D)}} \left[ -\frac{1}{m} \sum_{i=1}^m D(x^{(i)}, \tilde{c}^{(i)}) + \frac{1}{m} \sum_{i=1}^m D(G(z^{(i)}, c^{(i)})) \right], \alpha, \beta_1, \beta_2 \right)$ 
        # Update parameter for auxiliary network  $Q$ 
         $\theta^{(Q)} \leftarrow Adam \left( \nabla_{\theta^{(Q)}} \left[ -\frac{1}{m} \sum_{i=1}^m Q(c^{(i)} | G(z^{(i)}, c^{(i)})) \right], \alpha, \beta_1, \beta_2 \right)$ 
    end for
    # Update parameter for Generator
     $\theta^{(G)} \leftarrow Adam \left( \nabla_{\theta^{(G)}} \left[ \frac{1}{m} \sum_{i=1}^m D(G(z^{(i)}, c^{(i)})) - \frac{1}{m} \sum_{i=1}^m L_I(G(z^{(i)}, c^{(i)}), Q) \right], \alpha, \beta_1, \beta_2 \right)$ 
end for

```

---

Furthermore, to verify the control generation ability of the proposed model for unknown scenario patterns, K-fold cross-validation is used for model training and generation. Interpretable features can represent the pattern distribution of renewable scenarios. Therefore, the historical samples can be grouped according to the characteristic indicators  $\{F_1, F_2, \dots, F_N\}$ ,  $F_i \in \mathbb{R}^D$  of the historical time series sample  $\{x_1, x_2, \dots, x_N\}$ ,  $x_i \in \mathbb{R}^T$ . In this work, the K-means clustering algorithmic is used to partition the dataset into  $k$  cohesive groups through an unsupervised learning process [44]. After determining the feature cluster center  $\{a_1, a_2, \dots, a_k\} \in \mathbb{R}^D$ , the corresponding series sample  $\{x_1, x_2, \dots, x_N\}$  can be grouped into the subsets of the nearest cluster according to the close characteristic indicators. Scenarios in each subset have similar patterns. Then, each pattern is selected in turn as the validation set, and the other four patterns are used as the training set to train against the proposed model. In such manner, the simulation of the scenario generation for new patterns can be implemented.

**Table 1**

The proposed GAN model structure.

	Generator $G$	Discriminator $D$	Auxiliary network $Q$
Input	50 + 5	12*12	12*12
Layer 1	MLP, 1024	Conv, 64	Conv, 64
Layer 2	MLP, 512	Conv, 128	Conv, 128
Layer 3	MLP, 128	MLP, 1024	FC, 5
Layer 4	Deconv, 128	MLP, 128	–
Layer 5	Deconv, 64	Flatten, MLP, 1	–

## 4. Performance evaluation and numerical result

### 4.1. Data description and model parameters

The wind power data are collected from a real wind farm: Anzishan wind farm (capacity of 45 MW, Henan, China, hub height of 70 m) [45]. The profiles of solar generations (with a rated capacity of 60 MW) are adopted from NREL Solar Integration Datasets [46], located in the State of California, USA. The samples data of one year with the 10-minute resolution are used to train the models.

The structure of proposed model is inspired by the deep convolutional neural network (DCGAN) [47], as shown in Table 1. The input layer parameters of the model are determined by the case generation

task: for generator  $G$ , 50 represents the dimension of noise variable  $z$ , 5 represents the dimension of control vector  $c$  (five interpretable features selected in Section 2.1). For discriminator  $D$ , a one-day time series sample with a 10-minute resolution can be converted into a  $12 \times 12$  pixel image. The structure of hidden layers and the number of neurons in each layer were set according to reference [15]. It is observed that two convolution layers are adequate to represent the daily dynamics for the training set and is efficient for training [15]. The generator  $G$  structure includes 3 multilayer perceptron layers (MLP) and 2 de-convolutional layers with a stride size of  $2 \times 2$  for upsampling. The discriminator  $D$  has a reversed architecture with 2 convolutional layers with stride size of  $2 \times 2$  and 3 MLP layers. The auxiliary distribution  $Q$  is parameterized as a neural network, which shares all convolutional layers of discriminator  $D$ . It only needs one final fully connected (FC) layer to output parameters for the conditional distribution  $Q(c|\tilde{x})$  [37]. Batch normalization is adopted before the Leaky-ReLU activation in every layer of the generator  $G$  while the spectral normalization technique is used in discriminator  $D$ . The models are trained using the Adam optimizer with a mini-batch size of 64.

**Table 2**

Comparison of probability statistics for historical samples and generated samples.

Test cases	Expectation (MW)	Variance (MW <sup>2</sup> )	Skewness	Kurtosis
Historical scenarios of wind	257.68	261.42 <sup>2</sup>	0.95	2.88
Generated scenarios of wind	261.92	254.40 <sup>2</sup>	0.95	2.97
Historical scenarios of PV	10.94	14.75 <sup>2</sup>	1.05	2.62
Generated scenarios of PV	11.03	15.01 <sup>2</sup>	1.04	2.56

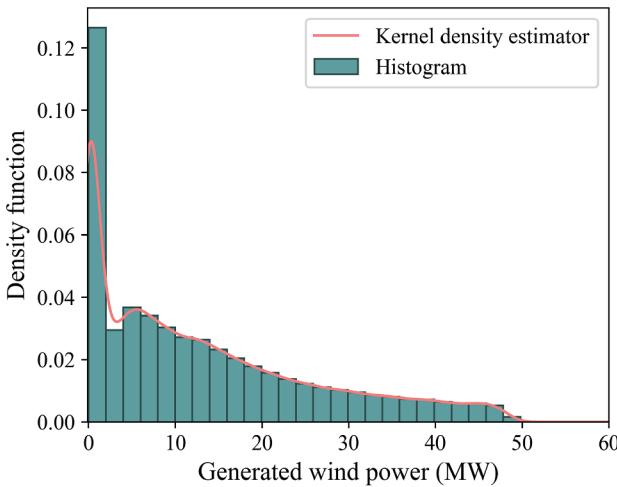
The computing tools for the programming and training of the above network are based on standard open-sourced Tensorflow [48] in Python platform (<https://www.python.org>). The implemented hardware parameters are 3.00 GHz Intel(R) Core (TM) i5-7400U CPU, Nvidia GTX 1650 GPU and 8.00 GB RAM.

#### 4.2. Performance of scenario generation based on statistical resemblance

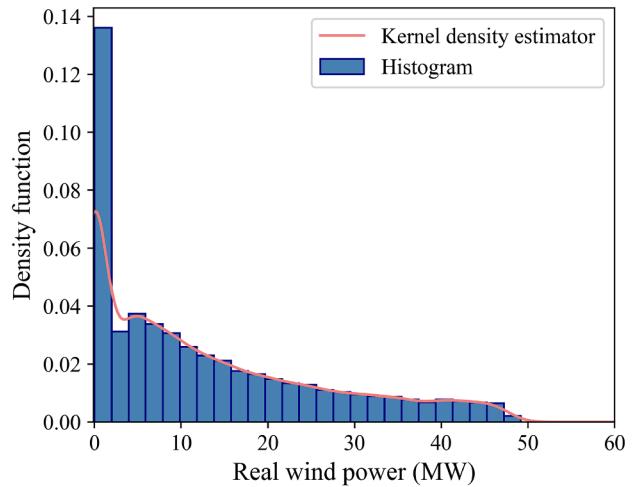
In this case study, we first verify the basic scenario generation capability of the proposed method. Without imposing specific control preferences, the renewable scenarios with different resource forms (wind and solar) are generated adaptively in a data-driven manner.

To evaluate the performance of proposed scenario generation, statistical properties of each moment and correlation are often used to measure the accuracy of scenarios simulation. According to the theory of probability and statistics, a group of infinite moments can uniquely determine a probability distribution [49]. Therefore, by calculating and comparing the multi-order moments of two probability distributions, the similarity between the generated probability distribution and the actual distribution can be measured. The first four order moments (i.e., Expectation, Variance, Skewness and Kurtosis) are considered sufficient to be used as the evaluation indexes, as suggested in [50].

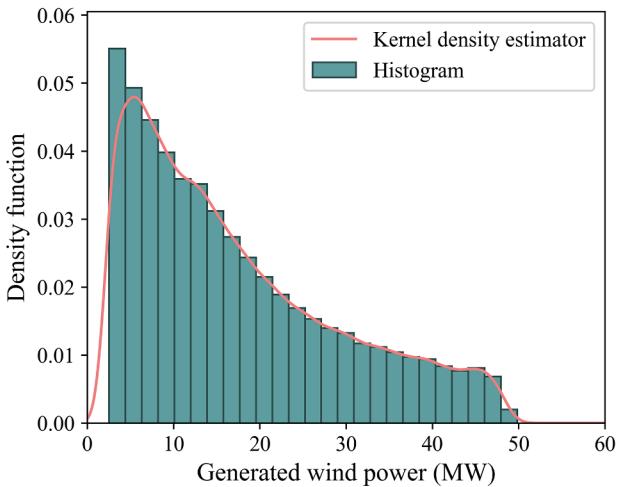
Skewness reflects the symmetry information of the overall distribution function curve. The closer the Skewness is to 0, the more symmetrical the distribution is. If the Skewness is negative, the sample follows the left biased distribution, otherwise following the right biased distribution. The calculation formula of Skewness is:  $S = E(X - \mu)^3 / \sigma^3$ , where  $\mu$  is the mean value of  $X$ ;  $\sigma$  is the standard deviation of  $X$ ;  $E$  is the



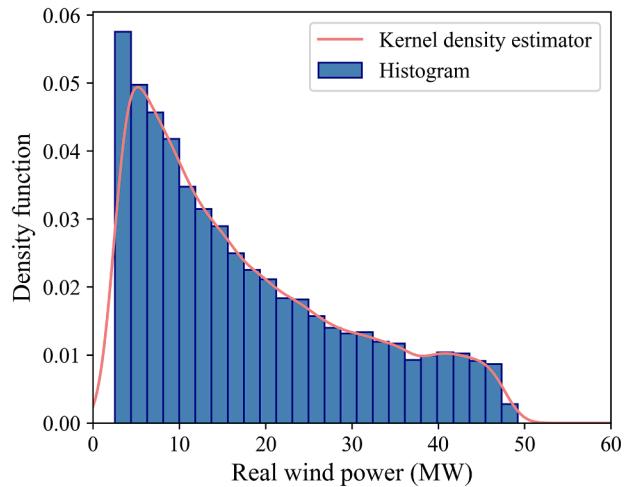
(a)



(b)

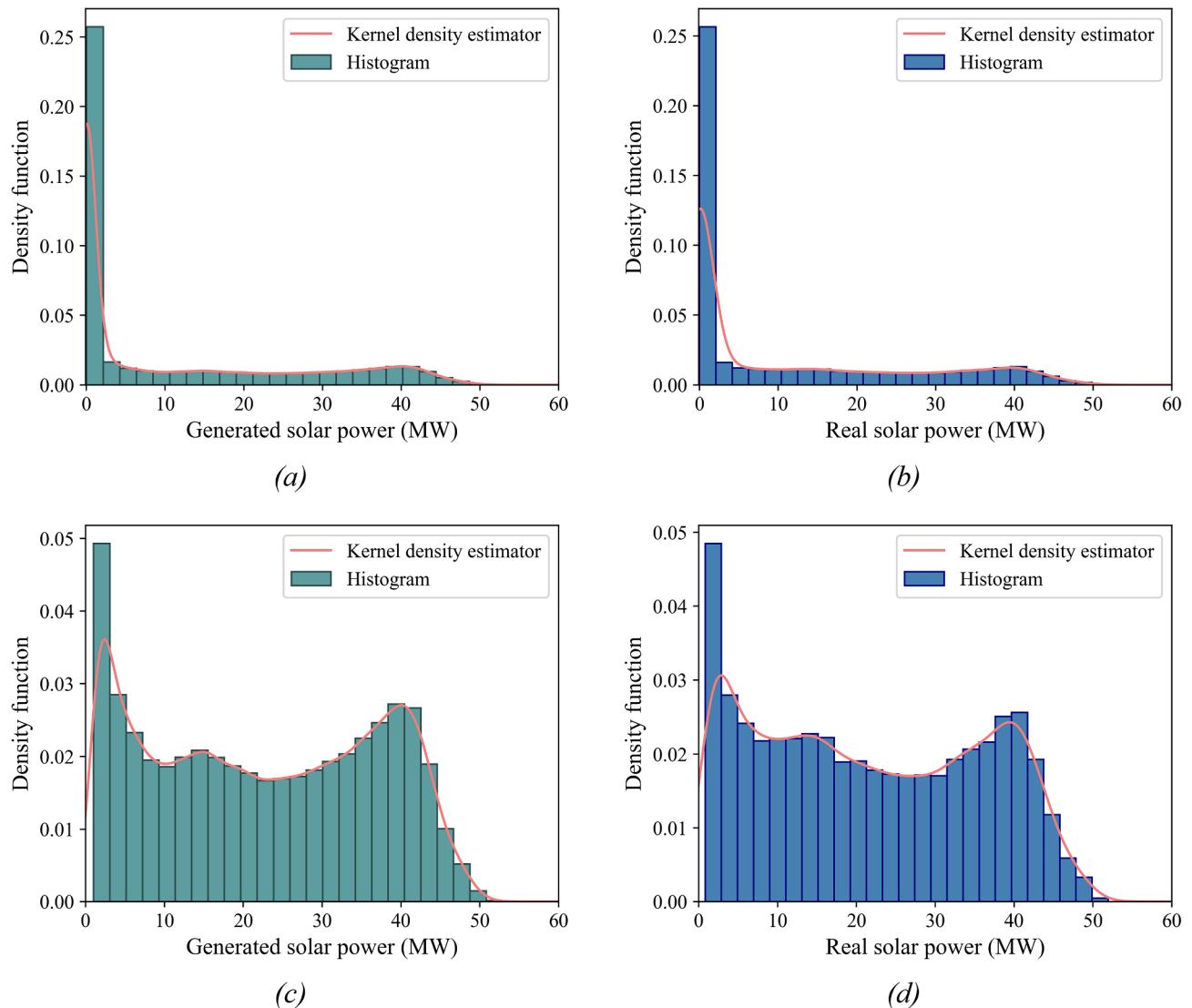


(c)

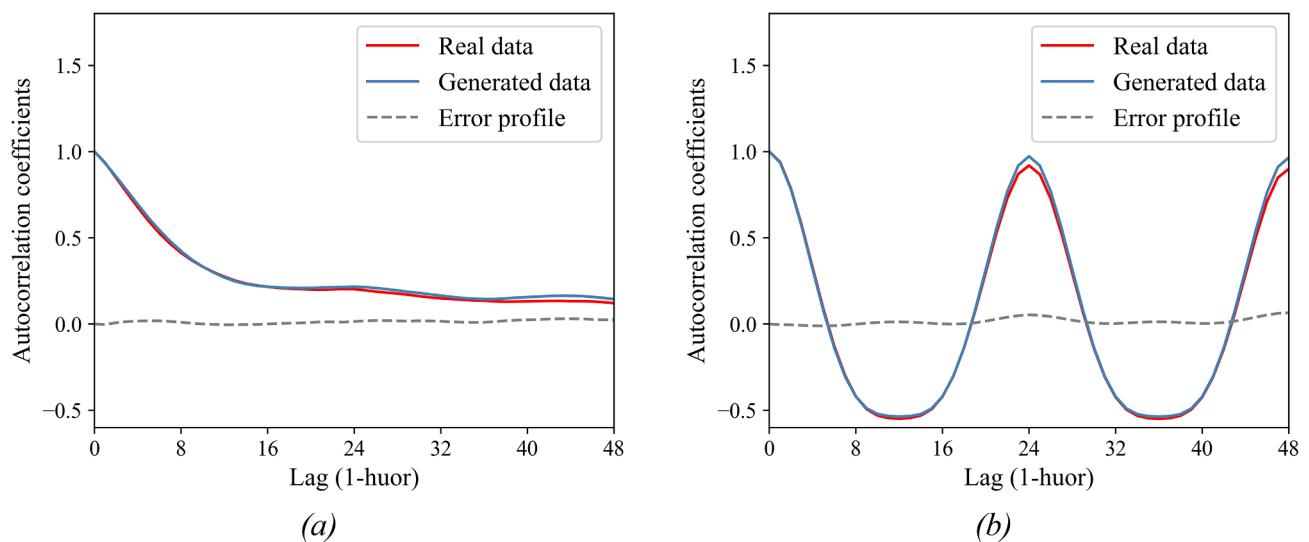


(d)

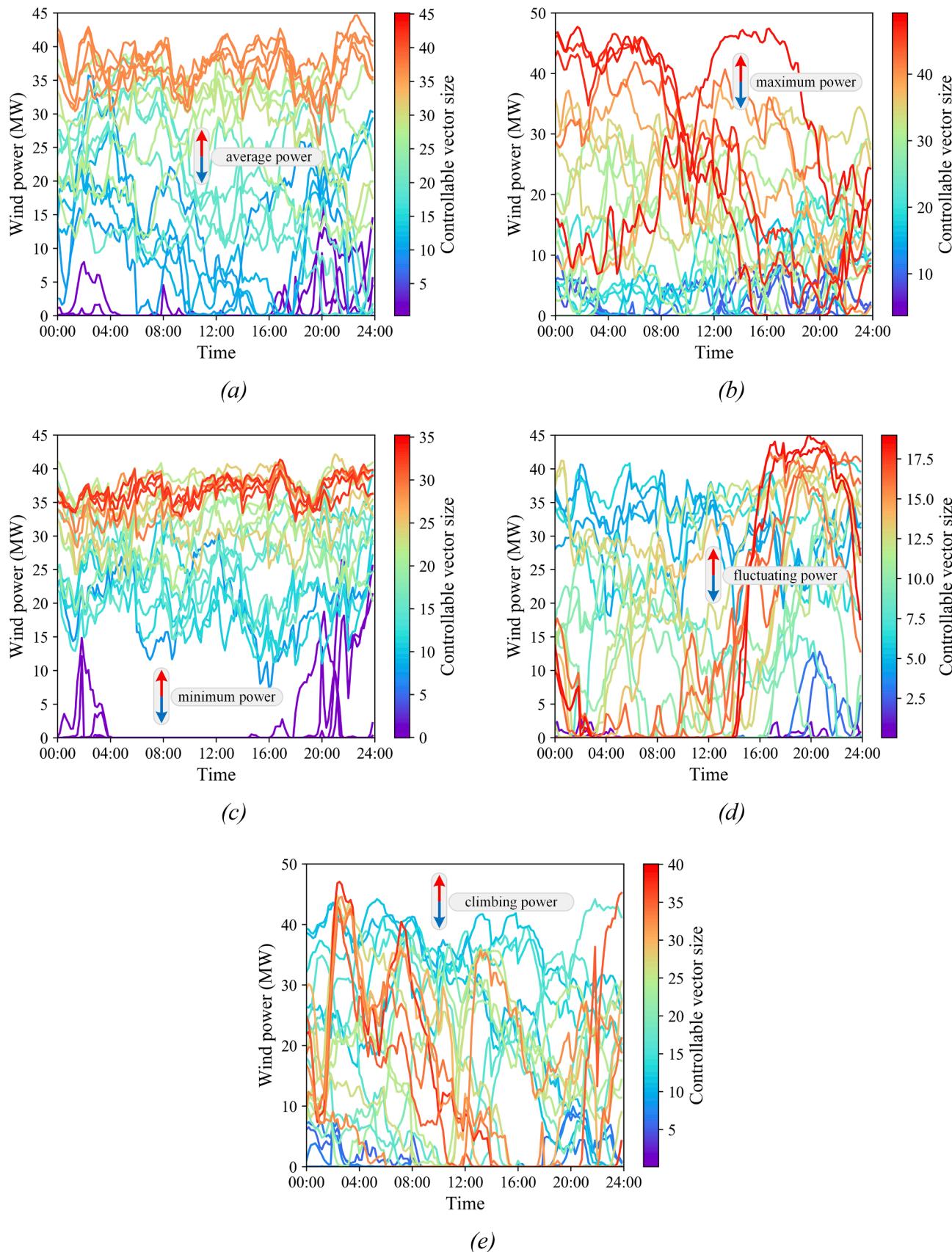
**Fig. 2.** Distribution of wind power generation: (a) all real samples; (b) all generated samples; (c) real samples under effective wind speed; (d) generated samples under effective wind speed.



**Fig. 3.** Distribution of solar power generation: (a) all real samples; (b) all generated samples; (c) real samples under sufficient radiation; (d) generated samples under sufficient radiation.



**Fig. 4.** Autocorrelation: (a) wind power samples; (b) solar power samples.



**Fig. 5.** Generation results of proposed GAN model tuned with different features: (a) average power; (b) maximum power; (c) minimum power; (d) average fluctuating power; and (e) maximum climbing power.

**Table 3**

Comparison between the preset controllable vector and the real statistical indicator for each interpretable feature.

Interpretable features	Comparison	Variation tendency of characteristic indicators					
		1.0	10.0	20.0	28.0	36.0	
Average power (MW)	controllable vector	1.0	10.0	20.0	28.0	36.0	
	real indicator	1.0	10.9	19.5	28.0	36.9	
Maximum power (MW)	controllable vector	10.0	20.0	28.0	36.0	46.0	
	real indicator	9.7	19.2	30.4	36.3	46.0	
Minimum power (MW)	controllable vector	0.0	10.0	16.0	24.0	32.0	
	real indicator	0.0	10.8	16.9	23.3	31.1	
Average fluctuating power (MW)	controllable vector	2.0	5.0	10.0	13.0	17.0	
	real indicator	1.4	5.3	9.5	12.9	17.1	
Maximum climbing power (MW)	controllable vector	5.0	12.0	20.0	28.0	36.0	
	real indicator	6.7	12.7	18.7	26.5	34.2	

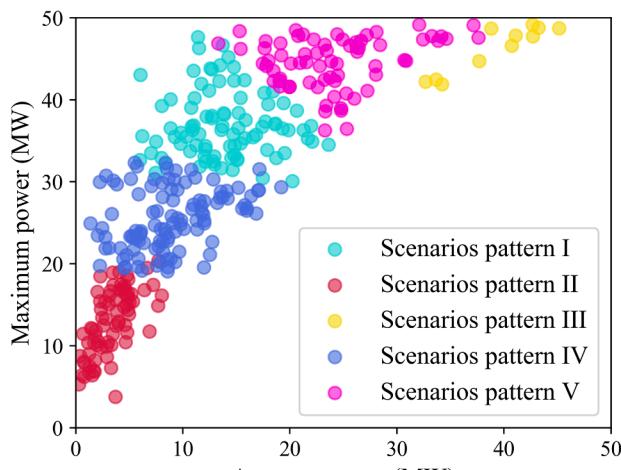
mathematical expectation.

Kurtosis reflects the steepness of the overall distribution curve near its peak. The Kurtosis of the normal distribution is 3. If the Kurtosis of the sample is greater than 3, the density curve near the peak is steeper than

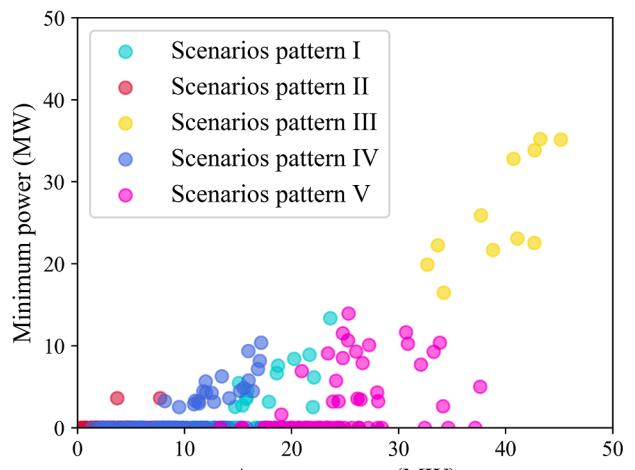
the normal distribution, otherwise is smoother. The calculation formula of Kurtosis is:  $K = E(X - \mu)^4 / \sigma^4$ .

Probability statistics were conducted on all sample points for the real historical data and generated data of renewable energy according to the above statistical measures, and the fourth-order statistical moments of the real distribution and generated distribution were obtained respectively, as shown in Table 2.

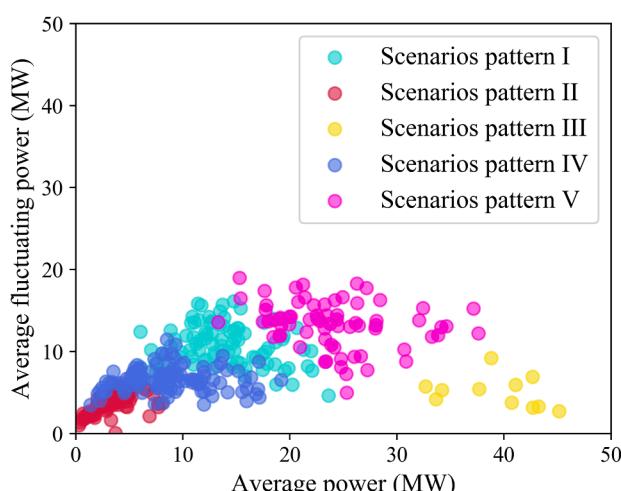
Further, the frequency histograms of all sample points from historical and generated renewable power data are plotted, and the corresponding kernel density estimation curves are depicted. Kernel density estimation (KDE) can reflect the empirical distribution function of renewable energy resources, which can be close to the probability density function of the actual distribution of renewable output power. Fig. 2 and Fig. 3 respectively show the distribution of wind and solar power. It can be seen that the distributions of wind and solar power generation are strongly right-skewed when the zero power samples are counted in Fig. 2 (a)-(b) and Fig. 3 (a)-(b). These samples of zero power are due to the fact that wind or solar power generators cannot generate electricity when the wind speed is below or above a certain value (i.e. outside the effective wind speed range [51]) or the solar radiation is insufficient. To more clearly compare the resemblance of the non-zero power generation, Fig. 2 (c)-(d) and Fig. 3 (c)-(d) presents the real and generated distributions of wind and solar power under the effective



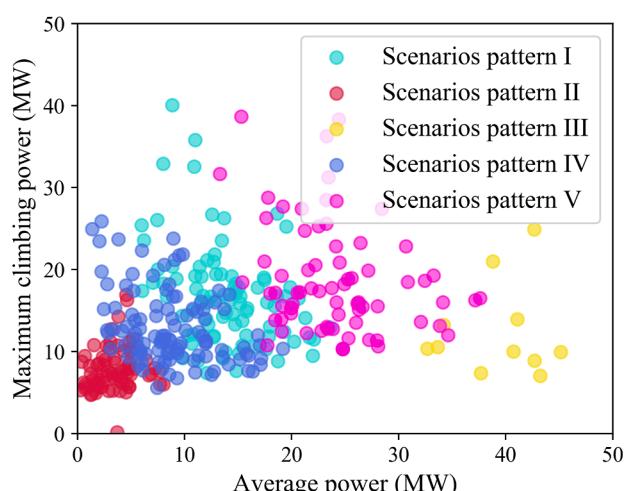
(a)



(b)



(c)



(d)

**Fig. 6.** Patterns distribution of real data clustering by the K-means model.

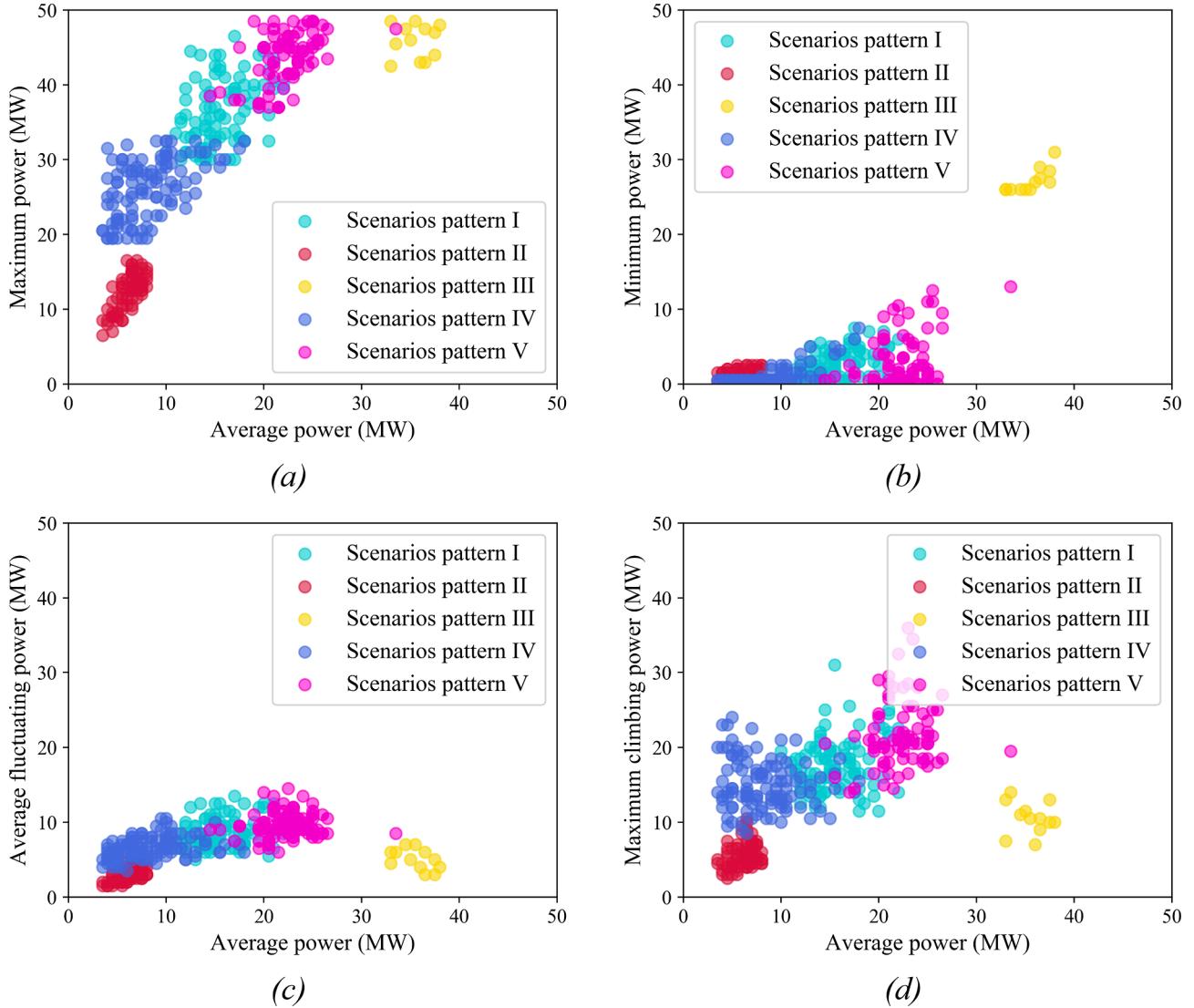


Fig. 7. Patterns distribution of controllable generated scenarios in K-fold cross-validation.

wind speed and sufficient radiation conditions, respectively. By comparing the historical and generated samples frequency distribution histogram and kernel density estimation curve, it can be seen that the distribution of generation scenarios based on the data-driven GAN method are similar to the real probability distribution of renewable energy resources.

Additionally, we validate that generated scenarios contain the same temporal characteristics as historical data. Autocorrelation is a statistical measure of temporal correlation at a renewable energy source, and capturing the proper temporal behavior is essential for power system operations. The sample autocorrelation coefficients  $R(\tau) = E[(X_t - \mu)(X_{t+\tau} - \mu)]/\sigma^2$  regarding the time-lag  $\tau$  are calculated and compared for the original and generated samples, as presented in Fig. 4. The results verify that the proposed GAN is capable of generating samples with similar temporal behaviors compared to the real data.

The above case study demonstrates that the proposed GAN-based model is effective for scenario generation for both wind and solar. The approach presented is data-driven and model-free. It utilizes the capabilities of deep neural networks and historical data to carry out the task of automatically generating scenarios that correspond to the same distribution as the real data, without sophisticatedly modeling the specific distribution.

#### 4.3. Controllable scenario generation by tuning interpretable features

In this section, the renewable scenarios are generated using the proposed controllable GAN whose latent space can be tuned along the interpretable features to control the characteristics of the profiles. Limited by space, this case study takes the generation of wind power scenario as an example to verify the effectiveness of the controllable generation process.

During the generation process, each interpretable feature (i.e., average power, maximum power, minimum power, average fluctuating power and maximum climbing power) of the controllable vector is adjusted in turn, and then several modulated scenarios are generated for each controllable vector. The variation tendency of the controllable generated power scenarios is intuitively shown in Fig. 5, and the color of the curve reflects the size of the corresponding characteristic indicator of the generated samples. The comparison between the expected values of controllable vector and the real statistics of generated samples for each characteristic indicator is shown in Table 3. The numerical results demonstrate that under the reasonable correlation of each feature, the proposed controllable model can independently control the specified characteristic with accurate quantity, and hence can implement the generation of interpretable preference patterns.

#### 4.4. New pattern generation verified by K-fold cross-validation

In this section, K-fold cross-validation is used for training and generation to verify the control generation ability of the proposed model for unknown scenario patterns. The characteristics of real samples with a two-dimensional scatterplot are shown in Fig. 6. The trained model generates new patterns by assigning the control vector of the target characteristics. The number of generated scenarios is consistent with the number of profiles contained in the validation set. The results of 5-fold cross-validation are shown in Fig. 7.

Comparing the results in Fig. 6 and Fig. 7, it is shown that the generated patterns coincide well with the target patterns, although the distribution of characteristics does not exactly match. In the case where the training data set does not contain the characteristics of the generated scenarios, the controllable GAN model proposed in this paper can still generate the pattern distribution of the expected preference. The interpretable feature of the new target pattern is embedded into the GAN model generation process as the control vector, which can be used for the new scenario generation task without sufficient data. Due to the fact that the statistical characteristic of the generated scenarios can be interpreted by using the latent vector on manifold space, the proposed model can achieve the generation task that the traditional GAN network cannot do.

#### 5. Conclusions

In this paper, a controllable Generative Adversarial Networks (GAN) model with interpretability is proposed for the renewable scenario generation task. The presented data-driven method can capture the stochastic characteristics of renewable resources without any statistical assumptions. It leverages the capabilities of deep learning and data-driven techniques to address the task of directly generating scenarios that correspond to the same distribution of historical data.

Through the adoption of the MI maximization between the latent space vector and the generated data and imitation learning sampling techniques, the generation process can be controlled using the latent vector on manifold space. The interpretable features of the renewable scenarios, such as average power, maximum power, minimum power, average fluctuating power and maximum climbing power, are manually adjusted as control vector. Experimental results demonstrate that the proposed model can not only achieve controllable scenario generation of individual features but also generate new scenario patterns by presetting excepted characteristics.

The proposed scenario generation solution is based on the data-driven technique that assumes the consistent distribution of historical training samples, and hence its performance is limited by the diversity of the patterns obtained from the available data. Although the proposed controllable model with interpretability can generate scenarios with unknown patterns, it may not lead to satisfactory result under certain circumstances, e.g., extreme weather conditions. Thus, the assessment of renewable generation uncertainties due to the frequent extreme climate change still remains a challenge to be addressed.

In future work, the controllable GAN model with interpretability proposed in this work can be further validated for the scenario generation of different forms of operational uncertainties, e.g., electricity, heat and cooling demand, dynamic pricing schemes. Also, the learning-based decision-making and control strategies of power systems based on the proposed scenario generation can be further investigated considering the high penetration of renewable energy sources.

#### CRediT authorship contribution statement

**Wei Dong:** Conceptualization, Methodology, Data curation, Writing – original draft. **Xianqing Chen:** Visualization, Investigation. **Qiang Yang:** Supervision, Writing – review & editing.

#### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgment

This work is supported in part by The National Natural Science Foundation of China (52177119).

#### References

- [1] Keirstead J, Jennings M, Sivakumar A. A review of urban energy system models: Approaches, challenges and opportunities. *Renew Sustain Energy Rev* 2012;16(6):3847–66.
- [2] Long S, Marjanovic O, Parisio A. Generalised control-oriented modelling framework for multi-energy systems. *Appl Energy* 2019;235:320–31.
- [3] Ibrahim MS, Dong W, Yang Q. Machine learning driven smart electric power systems: Current trends and new perspectives. *Appl Energy* 2020;272:115237. <https://doi.org/10.1016/j.apenergy.2020.115237>.
- [4] Rubino L, Capasso C, Veneri O. Review on plug-in electric vehicle charging architectures integrated with distributed energy sources for sustainable mobility. *Appl Energy* 2017;207:438–64.
- [5] Dallinger D, Wietschel M. Grid integration of intermittent renewable energy sources using price-responsive plug-in electric vehicles. *Renew Sust Energ Rev* 2012;16(5):3370–82.
- [6] Dong W, Yang Q, Li W, Zomaya AY. Machine-learning-based real-time economic dispatch in islanding microgrids in a cloud-edge computing environment. *IEEE Internet Things J* 2021;8(17):13703–11.
- [7] Powell WB, Meisel S. Tutorial on stochastic optimization in energy—Part I: Modeling and policies. *IEEE Trans Power Syst* 2016;31(2):1459–67.
- [8] Zhao C, Guan Y. Unified stochastic and robust unit commitment. *IEEE Trans Power Syst* 2013;28(3):3353–61.
- [9] Lei Y, Wang D, Jia H, Chen J, Li J, Song Y, et al. Multi-objective stochastic expansion planning based on multi-dimensional correlation scenario generation method for regional integrated energy system integrated renewable energy. *Appl Energy* 2020;276:115395. <https://doi.org/10.1016/j.apenergy.2020.115395>.
- [10] Sun M, Cremer J, Strbac G. A novel data-driven scenario generation framework for transmission expansion planning with high renewable energy penetration. *Appl Energy* 2018;228:546–55.
- [11] Dong W, Yang Q, Fang X, Ruan W. Adaptive optimal fuzzy logic based energy management in multi-energy microgrid considering operational uncertainties. *Appl Soft Comput* 2020;98(1):106882.
- [12] Yan J, Liu Y, Han S, Wang Y, Feng S. Reviews on uncertainty analysis of wind power forecasting. *Renew Sustain Energy Rev* 2015;52:1322–30.
- [13] Camal S, Teng F, Michiorri A, Kariniotakis G, Badesa L. Scenario generation of aggregated wind, photovoltaics and small hydro production for power systems applications. *Appl Energy* 2019;242:1396–406.
- [14] Morales JM, Mínguez R, Conejo AJ. A methodology to generate statistically dependent wind speed scenarios. *Appl Energy* 2010;87(3):843–55.
- [15] Chen Y, Wang Y, Kirschen D, Zhang B. Model-free renewable scenario generation using generative adversarial networks. *IEEE Trans Power Syst* 2017;33(99):3265–75.
- [16] Li J, Zhou J, Chen B. Review of wind power scenario generation methods for optimal operation of renewable energy systems. *Appl Energy* 2020;280:115992. <https://doi.org/10.1016/j.apenergy.2020.115992>.
- [17] Yang M, Fan S, Lee W-J. Probabilistic short-term wind power forecast using componential sparse Bayesian learning. *IEEE Trans Ind Appl* 2013;49(6):2783–92.
- [18] Bludszuweit H, Dominguez-Navarro JA, Lombart A. Statistical analysis of wind power forecast error. *IEEE Trans Power Syst* 2008;23(3):983–91.
- [19] Ma X-Y, Sun Y-Z, Fang H-L. Scenario generation of wind power based on statistical uncertainty and variability. *IEEE Trans Sustain Energy* 2013;4(4):894–904.
- [20] Pinson P, Madsen H, Nielsen HA, Papaefthymiou G, Klöckl B. From probabilistic forecasts to statistical scenarios of short-term wind power production. *Wind Energy* 2009;12(1):51–62.
- [21] Papaefthymiou G, Kurowicka D. Using copulas for modeling stochastic dependence in power system uncertainty analysis. *IEEE Trans Power Syst* 2009;24(1):40–9.
- [22] Diaz G, Gómez-Alexandre J, Coto J. Wind power scenario generation through statespace specifications for uncertainty analysis of wind power plants. *Appl Energy* 2016;162:21–30.
- [23] Vagropoulos SI, Kardakos EG, Simoglou CK, Bakirtzis AG, Catalão JPS. ANN-based scenario generation methodology for stochastic variables of electric power systems. *Electr Power Syst Res* 2016;134:9–18.
- [24] Hoyleland K, Kaut M, Wallace SW. A heuristic for momentmatching scenario generation. *Comput Optim Appl* 2003;24(2):169–85.
- [25] Park S, Xu Q, Hobbs BF. Comparing scenario reduction methods for stochastic transmission planning. *IET Gener Transm Distrib* 2019;13(7):1005–13.
- [26] Sideratos G, Hatzigergiou ND. Probabilistic wind power forecasting using radial basis function neural networks. *IEEE Trans Power Syst* 2012;27(4):1788–96.

- [27] Cui M, Ke D, Sun Y, Gan D, Zhang J, Hodge B-M. Wind power ramp event forecasting using a stochastic scenario generation method. *IEEE Trans Sustainable Energy* 2015;6(2):422–33.
- [28] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial nets. *Adv Neural Inf Process Syst* 2014:2672–80.
- [29] Chen Y, Li P, Zhang B. Bayesian renewables scenario generation via deep generative networks. In: 2018 52nd Annual Conference on Information Sciences and Systems (CISS); 2018. p. 1–6. <https://doi.org/10.1109/CISS.2018.8362314>.
- [30] Chen Y, Wang X, Zhang B. An unsupervised deep learning approach for scenario forecasts. In: 2018 Power Systems Computation Conference (PSCC); 2018. p. 1–7. <https://doi.org/10.23919/PSCC.2018.8442500>.
- [31] Jiang C, Mao Y, Chai Yi, Yu M, Tao S. Scenario generation for wind power using improved generative adversarial networks. *IEEE Access* 2018;6:62193–203.
- [32] Mirza M, Osindero S. Conditional Generative Adversarial Nets. *Computer. Science* 2014:2672–80.
- [33] Nash J. Non-cooperative games. *Ann. Mathe.* 1951;54:286–95. <https://doi.org/10.2307/1969529>.
- [34] Carmona G, Podczeck K. On the existence of pure-strategy equilibria in large games. *J Econ Theory* 2009;144(3):1300–19. <https://doi.org/10.1016/j.jet.2008.11.009>.
- [35] Odena A. Semi-supervised learning with generative adversarial networks. arXiv preprint arXiv:1606.01583, 2016.
- [36] Qiao J, Pu T, Wang X. Renewable scenario generation using controllable generative adversarial networks with transparent latent space. *CSEE J Power Energy* 2021;7 (1):66–77.
- [37] Chen X, Duan Y, Houthooft R, Schulman J, Sutskever I, Abbeel P. InfoGAN: interpretable representation learning by information maximizing generative adversarial nets. In: Proceedings of the 30th International Conference on Neural Information Processing Systems. 2016, p. 2180–188.
- [38] Shannon CE. A mathematical theory of communication. *Bell Syst Tech J* 1948;27 (3):379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>.
- [39] Kullback S, Leibler RA. On information and sufficiency. *Ann Stat* 1951;22(1): 79–86. <https://doi.org/10.1214/aoms/1177729694>.
- [40] Kingma D P, Welling M. Auto-encoding variational bayes. arXiv preprint arXiv: 1312.6114, 2013.
- [41] Blundell C, Cornebise J, Kavukcuoglu K, et al. Weight uncertainty in neural network. In: International Conference on Machine Learning. PMLR; 2015. p. 1613–22.
- [42] Barber D, Agakov F. The IM algorithm: A variational approach to information maximization. *Adv Neural Inf Process Syst* 2004;16(320):201.
- [43] Kingma D P, Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, 2014.
- [44] Celebi ME, Kingravi HA, Vela PA. A comparative study of efficient initialization methods for the k-means clustering algorithm. *Expert Syst Appl* 2013;40(1): 200–10.
- [45] Dong W, Yang Q, Fang X. Multi-step ahead wind power generation prediction based on hybrid machine learning techniques. *Energies* 2018;11(8):1975.
- [46] <https://www.nrel.gov/grid/solar-integration-data.html>.
- [47] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434, 2015.
- [48] Abadi M, Agarwal A, Barham P, et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467, 2016.
- [49] Lasserre JB. *Moments, positive polynomials and their applications*. World Scientific; 2009.
- [50] Fleishman AI. A method for simulating non-normal distributions. *Psychometrika* 1978;43(4):521–32.
- [51] Jena D, Rajendran S. A review of estimation of effective wind speed based control of wind turbines. *Renew Sust Energ Rev* 2015;43:1046–62.