Applied Energy 280 (2020) 115992

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Review of wind power scenario generation methods for optimal operation of renewable energy systems

Jinghua Li *, Jiasheng Zhou, Bo Chen

School of Electrical Engineering, Guangxi University, Nanning 530004, People's Republic of China

HIGHLIGHTS

- The state-of-the-art scenario generation methods are classified and reviewed comprehensively.
- An evaluation framework for scenario generation methods is established.
- The applications of scenario generation methods are summarized and discussed.
- Limitations and challenges of scenario generation methods are discussed.

ARTICLE INFO

Keywords:
Scenario generation
Stochastic programming
Wind power
Uncertainty
Application strategy

ABSTRACT

Scenario generation is an effective method for addressing uncertainties in stochastic programming for energy systems with integrated wind power. To comprehensively understand scenario generation and optimize solutions for uncertainties, the various methods and applications of scenario generation are classified and discussed in this work. First, the basic concepts are presented and scenario generation methods for addressing stochastic programming problems are discussed. Second, three categories of scenario generation methods are briefly introduced, along with their derived methods, advantages, and disadvantages. Third, an evaluation framework for these methods is established. Subsequently, applications of the scenario generation methods in power systems are discussed to identify the properties of these methods. Further, a comparative analysis and discussion are presented to show the suitability of each scenario generation method and to help choose the appropriate methods for different practical situations. Finally, the current limitations and future works with regard to scenario generation for stochastic programming in wind-power-integrated systems are highlighted and discussed. The results of this study are expected to provide references for applying scenario generation methods to the optimal operation of renewable energy systems.

1. Introduction

The intermittent nature of wind power is a major barrier to its large-scale integration into power grids, which may lead to uncertainty and variability of usage. The variability of wind power necessitates extra capacity for the power system, which lowers the economy and stability of the system. Moreover, wind power may have stochastic variations, whose features cannot be accurately formulated using tractable mathematical equations [1]. Owing to the properties of wind power, the operators and dispatchers require large quantities of accurate information about its characteristics to arrive at corrective and effective decisions for power systems [2,3].

Wind power is one of the main sources of renewable energy. Its stochastic nature is a challenge to operational decision-making in energy systems, including probabilistic energy flow, unit commitment, and economic dispatch. Therefore, stochastic programs for wind power need to be examined in depth. Solving a stochastic programming problem involves making decisions that perform well under the maximum possible number of scenarios generated by the problem [4]. In such cases, the stochastic programs can be solved with discrete distributions of a limited number of scenarios that represent stochastic features, which is a common method of handling the stochastic programs of energy systems.

In recent years, several methods have been proposed to achieve scenario generation (SG) for wind power. The current SG methods can

* Corresponding author.
E-mail address: happyjinghua@163.com (J. Li).

https://doi.org/10.1016/j.apenergy.2020.115992
Received 4 June 2020; Received in revised form 1 October 2020; Accepted 3 October 2020
Available online 14 October 2020

Nomenclature

Indicates

t	index of time periods from 1 to T
s	index of generated scenarios from 1 to S
m	index of observed scenarios from 1 to M
φ	area index
r	thermal generation unit index
k	index of stochastic variables from 1 to K

Parameters

η	confidence level
ρ_S	the probability of each scenario
ν_i	i-th statistical feature of observed scenarios
ω_i	weight of i-th statistical feature
$\zeta_{s,t}$	s-th scenario at time period t
β_m	m-th scenario in observed scenarios
ξ	renewable power installed capacity
θ	represents the ramp event and the persistent characteristic
$\lambda_s(t)$	cost coefficient at time period t
$G_s^{\varphi}(t)$	wind power generation of area φ at time period t in scenario s
$p_{s,t}^w$	wind power output at time period t in scenario s
$p_{min,t}^w$,$p_{max,t}^w$	min/max wind power output at time period t in all scenario S
$P_r^{\varphi}(t)$	power output of thermal unit r over area φ at time period t
a_r,b_r,c_r	coefficients of cost for thermal unit r
α_t	white noise sequence at time period t
$f(x)$	objective function
g_j	given constraint mapping functions
$\mathbf{I}(\cdot)$	an indicator function that takes 0 or 1
F	probability density function
E	mean function
d	distance function
C_k	cumulative distribution function

Variables

$C_r^{\varphi}(t)$	generating cost of thermal unit r over area φ at time period t
$C_M(s)$	unit operation cost in scenario s

Abbreviation

SG	scenario generation
MC	monte carlo
LHS	latin hypercube sampling
ARMA	auto-regressive moving average
MM	moment matching
PDF	probability distribution function
SAA	sample average approximation
ARIMA	autoregressive integrated moving average
RBFNNs	radial basis function neural networks
ANNs	artificial neural networks
LSTM	long short-term memory
GANs	generative adversarial networks
BR	backward reduction
FS	forward selection
MCMC	markov chain monte carlo
PVMC	persistence and variation-monte carlo
PSO	particle swarm optimization
WGAN	wasserstein generative adversarial network
GDFM	generalized dynamic factor model
SMT	scenarios mapping technology
MAE	mean absolute error
MAPE	mean absolute percentage error
RMSE	root mean square error
PPMCC	pearson product-moment correlation coefficient
TC	tail correlation
KT	kendall's tau
SR	spearman's rho
GC	gini correlation
IES	integrated energy systems

be divided into three main classes: sampling-based methods [5], forecasting-based methods [6,7], and optimization-based methods [8,9]. This paper describes, discusses in detail, and summarizes these SG methods. Generally speaking, sampling-based methods use historical data as samples for mining the stochastic features of wind power signals and subsequent sampling to generate scenarios that satisfy these features; some examples of this class are the Monte Carlo (MC) method [10], Latin hypercube sampling (LHS) method [11], and copula function sampling method [12,13]. These methods usually involve a priori assumptions and large scenario sets. Without considering statistical assumptions, the forecasting-based methods train prediction models to generate scenarios using large quantities of historical observation data, such as the auto-regressive moving average (ARMA) method [14,15] and machine-learning algorithms [16,17].

Alternatively, optimization-based methods tend to generate scenarios by reducing the observed historical data as optimization techniques can substantially reduce the scale of the scenarios. Current optimization-based methods for SG mainly include distance matching [18] and moment matching (MM) [19], and these methods commonly use distance indexes to eliminate scenarios that are not representative or fail to meet the error constraints. However, this method will eliminate small-probability but high-risk scenarios.

Although several reviews of SG methods have been reported [20], they lack systematic classifications and discussions. Furthermore, the applications of SG methods have not been considered in depth. There is also a distinct lack of systematic discussion on the evaluation indicators for SG methods. Therefore, this paper presents a comprehensive review of the available SG methods and provides classifications as well as detailed discussions. To the best of the author's knowledge, a comprehensive review of state-of-the-art SG methods has not been reported in the context of integrated renewable energy systems.

The main contributions of this paper are as follows:

1) A detailed bibliometric analysis was conducted to provide the sources and scope of references for renewable energy research and application.
2) A comprehensive analysis of the advantages and disadvantages of wind power SG methods is provided to serve as a guideline for power systems with integrated wind power.
3) An evaluation framework for SG methods to systematically assess the quality of generation scenarios is presented.
4) The applications of SG methods are summarized, which can provide reference for the selection of solutions to address stochastic programming problems.
5) Limitations and challenges of the SG methods are discussed, including their theoretical methods and practical applications.

The remainder of this paper is organized as follows: Section 2 presents descriptions of the basic ideas of SG, including the definition and classification of SG. Section 3 presents an overview of the available SG methods. Section 4 classifies the evaluation methods for SGs, and Section 5 describes their practical applications. Moreover, a systematic summary analysis and discussion of the different SG methods, along with the optimization strategies for power systems, are provided.

Finally, the conclusions of this study and prospects for future work are proposed in Section 6.

2. Description of scenario generation

2.1. Definition of SG methods

In most practical applications of stochastic programming, the probability distributions of the stochastic variables are approximated by discrete distributions ((ζ_s,ρ_s) , $s = 1,2,\cdots,S$, where ζ_s is a scenario and ρ_s is its probability) called SG, which provide feasible approaches by simplifying large quantities of data to approximate the probability distributions. As shown in Fig. 1, a continuous probability distribution function (PDF) is discretized. In this example, the original PDF is represented by five scenarios as rectangular bars, with the height of each rectangular bar representing its corresponding probability.

2.2. Classification of SG methods for wind-power scenarios

To provide an overview of existing literature on SG methods, several keywords, such as wind power, scenario, scenario generation, scenario reduction, scenario tree generation, stochastic programming, optimal problem, and evolutionary programming, were used to search relevant literature. In this study, 125 articles were found; most of them were published during 2015–2020. The proportion of publications corresponding to these keywords is shown in Fig. 2. A detailed bibliometric analysis was conducted based on the web of science databases, which can provide the sources and scope of references for renewable energy research and application, and the statistical results are as follows. The countries from which these documents were published are shown in Fig. 3, while Fig. 4 depicts the journals ranked by the number of relevant papers published in them.

A wide range of SG methods have been proposed to approximate the stochastic features of wind power. In this work, the existing methods were classified into three categories: sampling-based [5], forecasting-based [6,7], and optimization-based [8,9] methods, as shown in Fig. 5. All of these methods are further discussed in Section 3.

Theoretically, sampling-based methods generate discrete scenarios based on sampling the probabilistic distributions and deploying the outputs as the generated scenarios. Examples of sampling-based methods include the MC, LHS, copula function sampling, Markov chain MC, sample average approximation (SAA), persistence and variation-MC, and combined LHS and Nataf transformation methods.

Forecasting-based methods are used to train models with historical data to generate scenarios without considering any assumptions

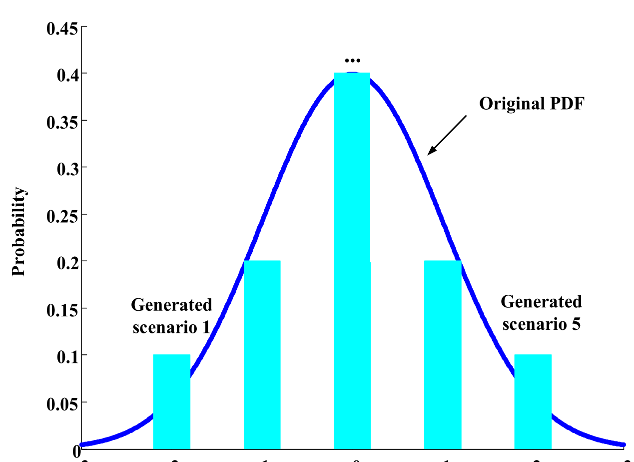

Fig. 1. Discrete probabilities of the continuous probability distribution function.

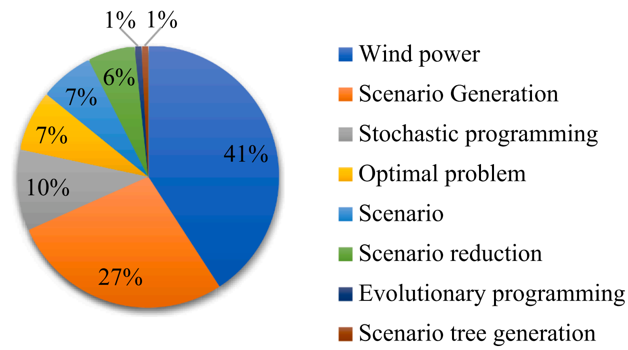

Fig. 2. Frequency of publication of different keywords.

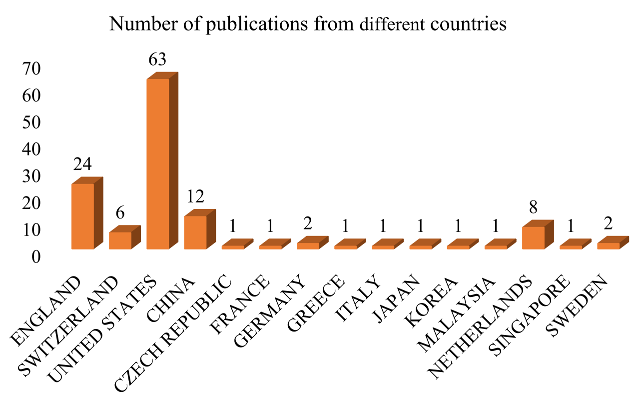

Fig. 3. Number of publications from different countries.

regarding the distribution functions. Examples of forecasting-based methods include the ARMA method, auto-regressive integrated moving average (ARIMA) method, state-space method, radial basis function neural networks (RBFNNs), artificial neural networks (ANNs), long short-term memory (LSTM) networks, and generative adversarial networks (GANs).

Optimization-based methods differ from the above two methods by focusing on reducing the number of scenarios given a large set of all possible scenarios. Usually, the optimal model enables minimization of the distances between the reduced scenarios and all possible scenarios. According to the approximate index, the optimization-based methods include the clustering, backward reduction (BR), forward selection (FS), and MM methods, as well as other combinations of SG methods. The clustering, BR, and FS methods use spatial distances (such as Euclidean distance) for the approximate target. Alternatively, the MM method uses statistical indexes (such as mean and variance) for the approximate target. In addition, some references have proposed a combination method of sampling-based and optimization-based for SG as well. Fig. 6 shows the number of publications for each SG method. It can be observed from this figure that the ARMA method has the most amount of published material. However, with the development of artificial intelligence, machine learning has gradually become a popular SG method.

3. SG methods in renewable power systems

This section briefly introduces the different categories of SG methods; the commonly used methods are presented along with the derived methods. The advantages and disadvantages of the different SG methods are enumerated with their applicable conditions and limitations.

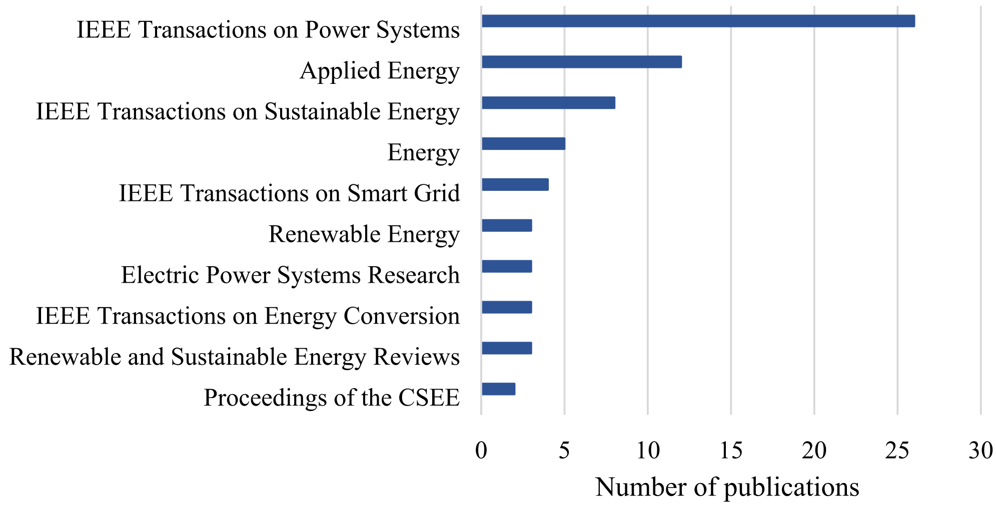

Fig. 4. Number of publications in the nine most popular journals for these keywords.

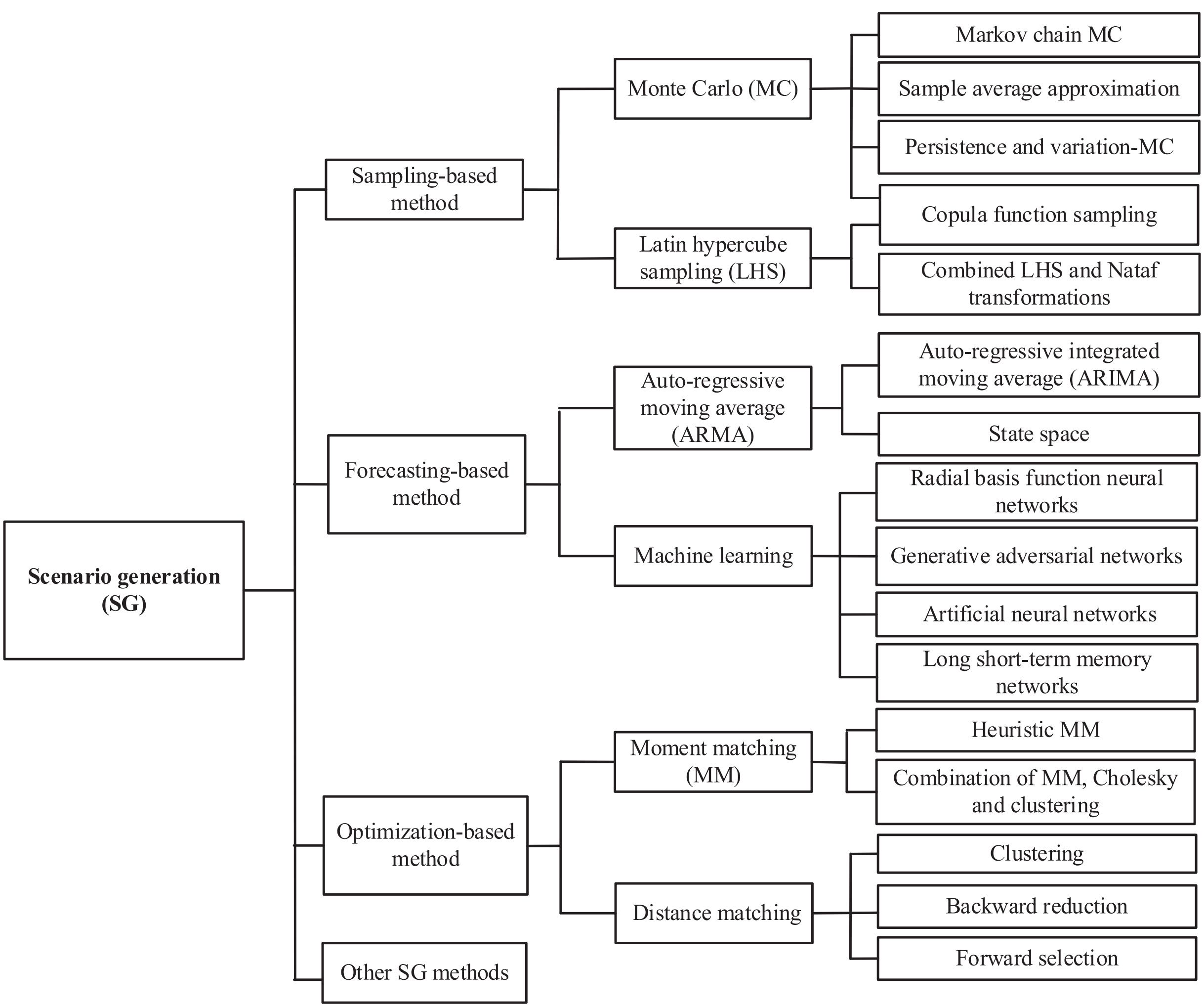

Fig. 5. Classification of SG methods.

3.1. Sampling-based methods

3.1.1. MC method

(1) Brief introduction to the MC method

Sampling is the most common method for generating scenarios, and the MC-based method is a typical classic method to determine the underlying function condition [21,22]. The main steps of the MC method are shown in Fig. 7.

The MC method has the advantages of simplicity and rapidity,

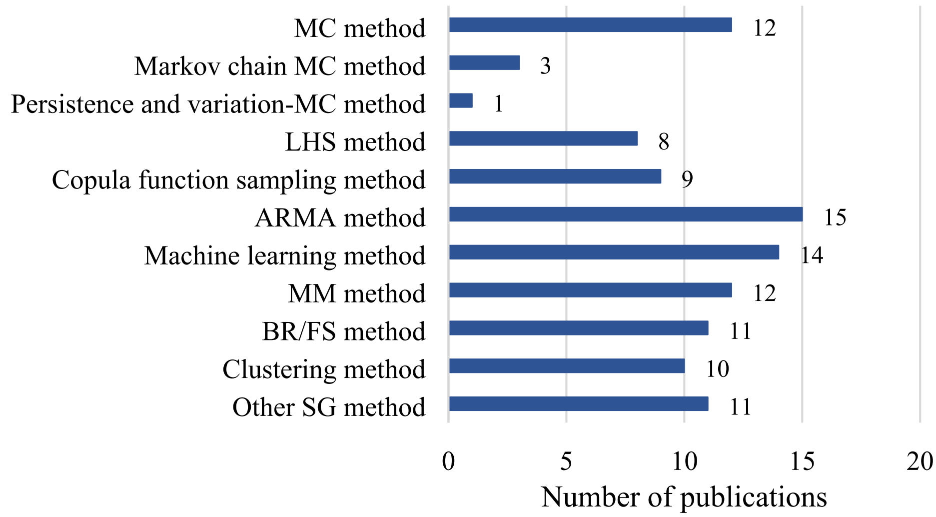

Fig. 6. Number of publications for SG methods.

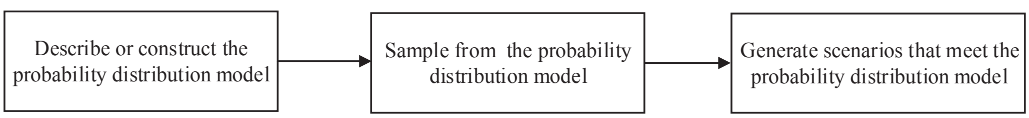

Fig. 7. Steps of the Monte Carlo generated scenarios.

thereby avoiding the complicated process of mathematical deduction. However, the probability distribution of the stochastic variables must be known [23]. When stochastic variables do not obey a normal or other general distribution, the PDF is often difficult to obtain. Moreover, the MC method not only requires a large number of samples to match the statistical features of the observed scenarios, resulting in heavy computational burden, but also satisfactorily captures the autocorrelation features of a time series. Therefore, some improvements to this method are necessary.

The SAA is a typical MC-based method that replaces the actual distribution using an empirical distribution corresponding to the Monte Carlo samples. The SAA method is popular for solving the chance-constrained problems of energy systems operations. Usually, the SAA method generates a finite set of scenarios and then changes the chance constraint to an approximate deterministic constraint based on the sample scenarios. The main procedure of SAA is as follows [24]:

Consider the chance-constrained problems of Eq. (1) are to be solved:

$$\begin{aligned} & \min f(x) \\ s.t. \quad & Pr\{g_j(x,\zeta)\geqslant 0, j=1,\dots,m\}\geqslant 1-\eta \\ & x\in X \end{aligned} \tag{1}$$

where $f: R^n \to R$ is the objective function, $X \subset R^n$ is the set of deterministic constraints, $\zeta: \Omega \to R^d$ is the stochastic variable with distribution F, g_j is a constraint mapping function, and η is a confidence parameter.

The SAA method chooses a subset of the sampled constraints that are to be satisfied. Accordingly, the SAA formulation of the chance-constrained problems is given by Eq. (2):

$$\begin{aligned} & \min f(x) \\ s.t. \quad & \frac{1}{S}\sum_{s=1}^{S} \boldsymbol{I}\{g_j(x,\zeta_s)\geqslant 0, j=1,\dots,m\}\geqslant 1-\eta \\ & x\in X \end{aligned} \tag{2}$$

where scenarios $\zeta_1, \cdots, \zeta_S$ are independent and sampled from the distribution F based on the MC method. $\boldsymbol{I}(\cdot)$ is an indicator function that takes the value of one when the expression within the curly brackets is true and zero otherwise. Thus, the intractable chain constraint in Eq. (1) is changed to a tractable deterministic constraint in Eq. (2) using SAA methods. The main advantage of the SAA method is that there is no restriction on the distribution F and only scenarios that can be obtained are assumed.

(2) Derived MC methods

Based on the fact that the MC method does not satisfactorily capture the autocorrelation features of a time series, a Markov-chain Monte Carlo (MCMC) [25] method was proposed considering both the PDF and autocorrelation function of the generated wind-power time series. The MCMC method employs Markov chains to characterize the transition features of the scenario. Using the MCMC scheme, the scenario at a future time can be obtained with the transition probabilities instead of the historical PDF, which captures the autocorrelation of wind power. In addition, the MCMC method has advantages in high-dimensional joint distribution sampling, and a dynamic SG method based on Gibbs sampling was proposed by considering multiple renewable energy power plants [26]. Since the MCMC technique fails to provide a quantified description of the time-domain features and to take both statistical and time-domain features into consideration, an improved MCMC method, called the persistence and variation-Monte Carlo method (PVMC), was proposed [27]. The PVMC considers the persistence and variation characteristics of wind power generation. Herein, the generated scenarios are closer to the statistical characteristics of wind power. Considering the efficiency of SG, a combination of roulette-wheel mechanism and the MC method was proposed for optimal scheduling of virtual power plants [28]. In [29], the SAA method was presented to solve the stochastic optimization problem.

3.1.2. LHS method

(1) Brief introduction to the LHS method

The general LHS method consists of two steps: sampling and permutation [11]. The objective of sampling is to generate representative samples that reflect the distribution of each random input variable. Permutation aims at reducing the correlations between samples of different stochastic input variables for multiple input problems [30]. The mathematical process of the LHS is as follows:

Let $\{X_k\}(k = 1, 2, \cdots, K)$ be the stochastic variables in a probabilistic problem. The cumulative probability function of X_k can be denoted as

$$C_k = F_k(X_k) \tag{3}$$

Assume C_k as shown in Fig. 8; the range of C_k is divided into S non-overlapping intervals of equivalent length. One sampling value is chosen from each interval by choosing the midpoint or random selection. Here, the midpoint value is used. Then, the sampled values of X_k are calculated with the inverse function of Eq. (3) The s-th sampling value can be calculated as

$$x_{ks} = F_k^{-1}\left(\frac{s - \frac{1}{2}}{S}\right) \tag{4}$$

where x_{ks} represents the value of the k element of the s row vector, K represents the number of stochastic variables, and S represents the number of generated scenarios. The initial $K \times S$ sampling matrix $X_{K\times S}$ can be obtained by assembling the sampled values of each stochastic variable in a row of the sampling matrix. After permutation, each row of the $K \times S$ sampling matrix represents the generated scenarios.

Compared with the typical sampling method, the LHS scheme possesses several advantages. First, the LHS is a stratified sampling method that requires a small number of samples and avoids repetition. Moreover, the LHS can reflect theoretical distribution and guarantee that the sampling points cover all sampling areas.

(2) Derived LHS methods

It is important to note that when the LHS scheme solves multivariate input random problems, the simulation accuracy is affected not only by the sample values but also by the correlations between the samples of different input random variables. Based on the LHS, Cai *et al.* included the correlations between random variables based on the LHS to establish their probability distribution models [30]. Chen *et al.* proposed a combination of the LHS and Nataf transformations to cope with probabilistic load flow [31].

3.1.3. Copula function sampling method

The copula function sampling method involves scenarios that are obtained by sampling the copula function.

Based on the Sklar theorem [32], assume $F(X_1, X_2, \cdots, X_K)$ is the joint distribution of k-dimensional random variables, and $F_1(X_1), F_2(X_2), \cdots F_K(X_K)$ are its marginal distribution function. Then, the dependent structure of the variables can be expressed by a k-dimensional copula function C as

$$F(X_1, X_2, \cdots, X_K) = C(F_1(X_1), F_2(X_2), \cdots F_K(X_K)) \tag{5}$$

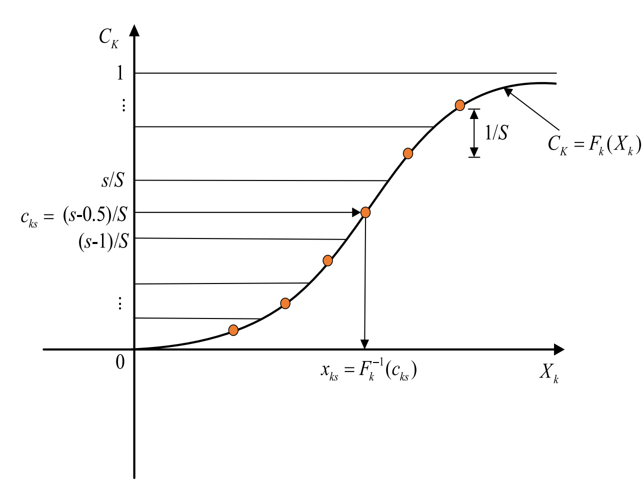

Fig. 8. Sampling step by LHS method.

After obtaining the copula function C, sampling from the copula function C to obtain discrete scenarios. Then, the obtained Copula function discrete scenarios are used to calculate the original joint distribution function scenarios through the inverse operation of the marginal distribution function. Finally, joint scenarios of multi-wind farm power are generated. The main procedure of the copula function sampling method is shown in Table. 1.

The copula function is employed to model the dependences between random variables in a multivariate structure. Moreover, the copula function has no limitations for marginal distribution and is able to obtain the nonlinear and asymmetrical relationships between variables. Li and Lan proposed methods for SG based on the copula function [33], which was applied to generate multi-wind power outages for time- and space-dependent scenarios in [34]. When three or more renewable energy sources are involved, a more flexible and accurate SG method based on the vine copula was proposed in [35,36]. In [37], the authors present a multivariate copula approach to generate correlated scenarios of renewable generation. The copula function sampling method can be used to effectively avoid constructing a joint probability distribution. The resulting scenarios can better capture the dependency structure, and the copula function methods can accurately simulate renewable energy power.

3.1.4. Summary of sampling-based method

There are numerous sampling-based methods for SGs. Each of these has different characteristics and properties for different applications, so that different methods might be suitable for specific situations. Table 2 presents a comparative analysis of the sampling-based methods of SG.

3.2. Forecasting-based methods

The ARMA method and machine learning algorithms are two typical forecasting-based methods for generating wind power scenarios. Here, both methods are briefly introduced, including derived methods.

3.2.1. ARMA method

(1) Brief introduction to the ARMA method

Table 1
Steps of the copula function sampling method.

Algorithm: Copula function sampling method	
Step1:	Determine the marginal distributions of the random variables.
Step2:	According to the correlation characteristics of the random variables, select the appropriate copula function. Commonly used functions mainly include the normal function, t-copula function, Gumbel copula function, Clayton copula function, and Frank copula function.
Step3:	According to the selected copula function, estimate the unknown parameters in the copula model.
Step4:	Generate $N \times K$-dimensional data samples that satisfy the copula function distribution, where N is the total number of samples, and K is the dimension of random variables.
Step5:	Determine the number of scenarios S, use clustering method to classify $N \times K$-dimensional data samples into S categories, and use various centers (the average of all samples in this category) $\boldsymbol{u}^s = [u_1^s, u_2^s, \cdots u_d^s]$ as the quantiles of the scenarios; count the ratio of samples falling in this category to the total number of samples and use it as the probability value $\rho_s(s = 1,2,\cdots, S)$ of each category.
Step6:	Use the formula $X_i^s = F_i^{-1}(u_i^s)(i = 1,2,\cdots, d)$ to convert $\boldsymbol{u}^s = [u_1^s, u_2^s, \cdots u_d^s]$ into the original joint distribution function scenario to obtain the scenario. The corresponding probability value of the scenario is $\rho_s(s = 1,2,\cdots, S)$.

Table 2
Comparison of the advantages and disadvantages of sampling-based SG methods.

Sampling-based method	Advantages	Disadvantages	Applications
MC method	The convergence rate is independent of the dimension of the problem	Low efficiency, with certain assumptions	Single state
MCMC method	Capture the autocorrelation features	Possible inaccuracy for complicated problems	Multiple states
PVMC method	Consideration of persistence and variation characteristics, suitable for generating wind power scenarios over several years	Computational complexity	Multiple states
LHS method	High efficiency to reflect theoretical distribution of random variables	With certain assumptions, computational complexity	Single state
Copula function sampling method	No limitation for marginal distribution, description of non-linear correlation between random variables	Dependency on existing copula function	Single/ multiple states

ARMA [38] is another common parametric method used for wind power SG to capture the linear correlation of wind power series. The general expression of ARMA is as follows:

$$y_t = \sum_{i=1}^{A} \phi_i y_{t-i} + \alpha_t - \sum_{j=1}^{B} \theta_j \alpha_{t-j} \tag{6}$$

where y_t represents the sequence value at time t, y_{t-i} represents the sequence value at time $t-i$, $\phi_i (i = 1, 2, \cdots, A)$ and $\theta_j (j = 1, 2, \cdots, B)$ represents the auto-regressive and moving average parameters, respectively; α_t represents the white noise sequence at time t, and α_{t-j} represents a white noise sequence at time $t-j$.

Theoretically, ARMA generates scenarios based on the states of the system in the past and on the memory of invading noise. The value of a sequence at a certain time can be expressed by the linear combination of the historical and white noise sequences [39]. In [40,41], the ARMA time-series model was used to simulate hourly wind speeds. Based on this, an ARMA wind speed model for forecasting wind power output was proposed [42]. In [43], the ARMA model was used to generate day-ahead market price scenarios. In [44], ARMA does not need to directly consider the correlations of other related random variables, but it needs to assume that the stochastic process has a multivariate Gaussian distribution, and therefore, smoothing the non-stationary sequences is required.

(2) Derived ARMA methods

For nonstationary and non-Gaussian stochastic processes, the direct application of ARMA is not feasible. Based on autoregressive systems, Chen *et al.* proposed a stochastic wind power model based on ARIMA [45]. The proposed model takes the nonstationary characteristics of wind power generation into account; thus, it does not rely on quantization or suffer from quantization errors. Díaz *et al.* transformed the ARMA model into its state-space form to simulate spatiotemporal dependent wind power scenarios [46]. However, it is difficult to grasp the complex nonlinear relationships of multiple wind sites.

3.2.2. Brief introduction to machine learning algorithms

Recently, machine learning algorithms, or non-parametric methods, have become popular for SG. However, there is currently no general scheme for machine learning. This section introduces a popular machine learning method, GANs [17], whose main steps for SG methodology are shown in Fig. 9.

In [47], ANNs were combined with historical observations and exogenous variables (e.g., ambient temperature, wind speed, and solar radiation) to generate scenarios for various stochastic variables used as input parameters. Based on neural network theory [48], wind power scenarios are simulated as stochastic processes that can carry the properties of ramping events. Sideratos *et al.* proposed a probabilistic wind power prediction model that combined RBFNNs with a particle swarm optimization (PSO) algorithm [49] to generate scenarios [16]. The authors in [50] proposed an SG approach based on LSTM networks to characterize stochasticity in electricity markets. In [51], a data-driven method based on GANs is utilized to generate scenarios of renewable resources. The proposed generative model not only requires no fitting of the probabilistic models of stochastic variables but also avoids manually labeling data sets. An improved GAN for generating wind power scenarios was proposed in [17]. Moreover, considering special meteorological conditions, Zhang et al. applied a conditional improved Wasserstein generative adversarial network (WGAN) to generate wind power scenarios that capture the spatiotemporal relationships of multiple wind farms [52].

In summary, machine learning methods are generally based on data-driven methods, and the quality of the generated scenarios depend on historical observation samples. Moreover, it captures the full diversity of renewable resources and the complex nonlinear relationship of the time series.

3.2.3. Summary of forecasting-based methods

A comparison of the forecasting-based SG methods and their corresponding advantages, disadvantages, and applications are summarized in Table 3.

3.3. Optimization-based SG methods

This section presents a brief introduction to two types of optimization-based SG methods, including the MM and distance matching methods. Commonly used methods are also presented along with the derived methods.

3.3.1. MM method

(1) Brief introduction to the MM methods

The main principle of moment matching (MM) involves matching the statistical features (i.e., mean, variance, skewness, and kurtosis) of the generated scenarios with those of the observed scenarios. MM can generate scenarios that match the stochastic features of wind power [53,54]. This method is based on the moments of the discrete probability distributions. The basic approach for generating scenarios by MM is to solve a nonlinear model, which is given as

$$\begin{aligned} & \min_{s,p} \sum_{i \in \boldsymbol{R}} \omega_i [f_i(s,\rho) - \nu_i]^2 \\ & s.t. \ \sum_{s=1}^{S} \rho_s = 1, \rho_s \geqslant 0 \end{aligned} \tag{7}$$

where $\boldsymbol{R}$ represents the set of statistical features, ω_i represents the weight of the statistical feature i, S represents the set of all scenarios, ν_i represents the known value of the statistical feature i, and ρ_s represents the probability of the s-th scenario. The process of an improved MM method is as shown in Fig. 10:

Step1: Based on a sample of historical observations, estimate the statistical characteristics of the historical samples, i.e., the mean, standard deviation, partial state, peak state, and covariance, which are used as the target moments to be matched.

Step2: Historical samples are clustered using the clustering center as

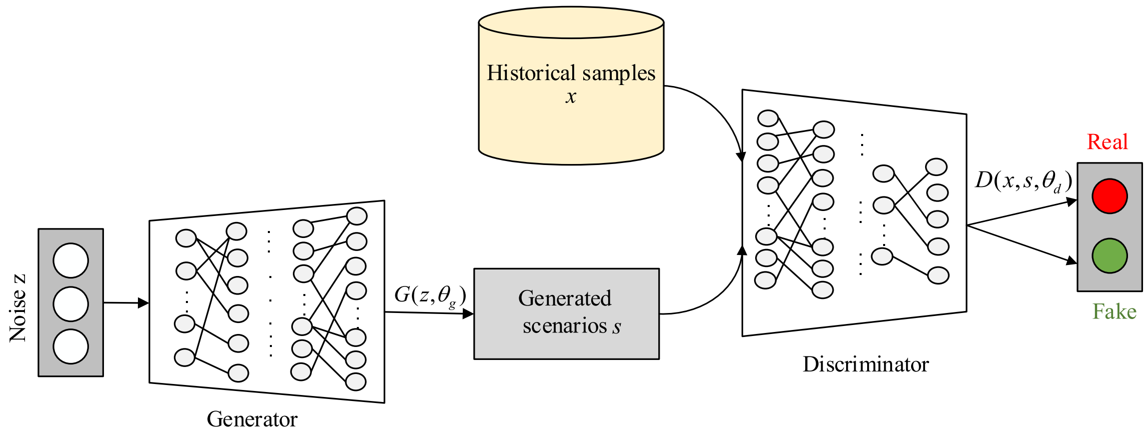

Fig. 9. Flowchart of the scenario generation algorithm based on GANs.

Table 3
Comparison of the advantages and disadvantages of forecasting-based SG methods.

Forecasting-based method	Advantages	Disadvantages	Applications
ARMA method	Good at capturing the linear correlations of wind power	Susceptible to inaccurate transformation	Multiple stages
Machine learning method	Capture the spatiotemporal relationships of renewable energy, the full diversity of scenarios, distribution-free	Strict requirements on input data, careful selection of training parameters	Multiple stages

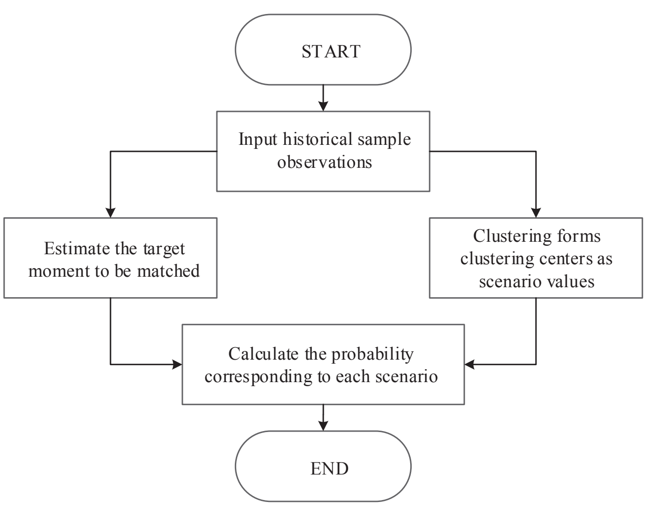

Fig. 10. Illustration of the steps of the MM algorithm.

the scenario value for that stochastic variable.

Step3: After Step1 and Step2, the estimated mean, standard deviation, partial state, peak state, covariance, and generated values of each scenario are applied to Eq. (7) to solve the corresponding probability of each scenario.

In addition, the MM method has been successful in stochastic programming as it better describes the statistical features of wind power generation and does not require a specific marginal distribution. However, the MM method uses non-convex optimization to generate scenarios that match the stochastic properties of the original scenarios, which usually impose a large computational burden [53].

(2) Derived MM methods

The MM method is challenging to solve nonlinear non-convex optimization problems [54]; this method also fails to consider the distance between the generated and observed scenarios. Hence, it may result in an obvious bias between the generated and observed scenarios.

Regarding the heuristic for MM, Høyland *et al.* proposed a combination of Cholesky decomposition and various transformations to achieve the first four moments and correlations, thus reducing the computational burden [55]. Based on this, Ehsan *et al.* applied this method to distribution network planning under uncertainties [56,57]. Li and Zhu proposed an improved MM method [8] that combines the conventional MM method, clustering method, and Cholesky decomposition. As a result, the minimum distance can be obtained, and the stochastic features between the generated and observed scenarios can be minimized. Rubasheuski *et al.* developed a multistage SG method by combining MM and scenario reduction, significantly reducing the complexity of matching problems and avoiding massive nonlinear problems [58].

3.3.2. Distance matching method

(1) Brief introduction to the distance methodsThe distance method refers to a category of optimization-based SG methods that reduce scenarios by measuring the differences between the generated and observed scenarios. The objective of the distance matching method is to minimize the distance between the generated and observed scenarios. The general optimization objective is defined as follows:

$$\sum_{s=1}^{S} \min_{m\in\{1,2,\cdots,M\}} d(\zeta_s - \beta_m) \tag{8}$$

where ζ_s denotes the generated scenarios, β_m denotes the observed scenarios, S is the number of generated scenarios, M is the number of observed scenarios, and $d(\zeta_s - \beta_m)$is the distance between scenario ζ_s and scenario β_m.

According to distinct distance measures, these can be further divided into the clustering, BR, and FS methods. The BR and FS methods are two typical optimal SG methods based on distance matching [59]. Both methods are used to reduce scenarios that model stochastic data processes in stochastic programs of energy systems. They impose no requirements on the stochastic data processes or the structures of the scenarios. The processes of the BR and FS methods are detailed as follows:

1) BR method

Fig. 11 shows the schematic process of the BR method. For each

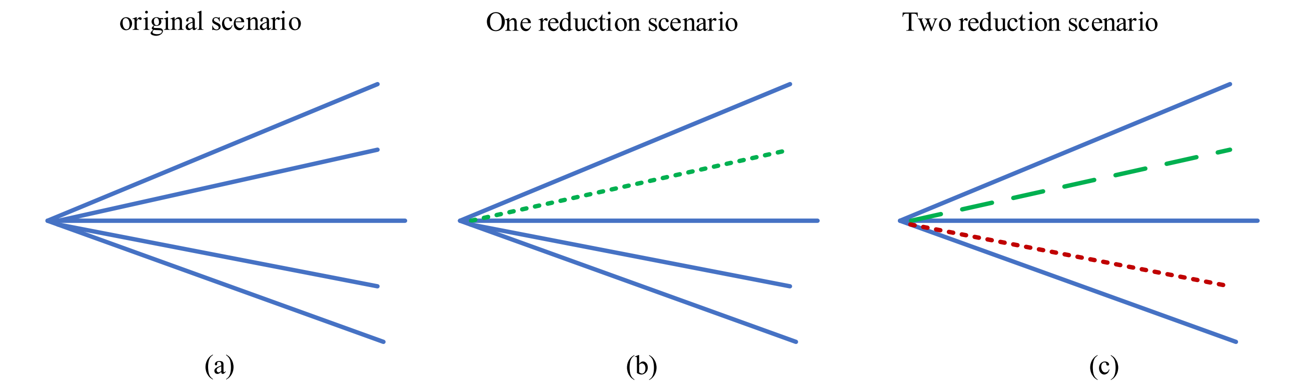

Fig. 11. Schematic process of the BR method.

reduction, one scenario was deleted from the original set of scenarios. As shown in Fig. 11(a), there are five scenarios in the original set. After the first reduction, there are four remaining scenarios, as shown in Fig. 11 (b), and so on.

2) FS method

Fig. 12 shows the schematic process of the FS method. For each iteration, one scenario was selected from the original scenarios. As shown in Fig. 12(a), the initial scenario set has one scenario. After the first selection, two scenarios are obtained, as shown in Fig. 12(b), and so on.

The clustering method [60] is a classic and simple approach for generating representative scenarios. Clustering involves the process of dividing the observed scenarios into different classes, with scenarios in the same class bearing great similarities, which are measured by their Euclidean distance [61]. Although the basic clustering methods are similar, there are various approaches for implementing them. Guan *et al.* proposed a method for generating daily wind power time series scenarios based on a longitudinal-horizontal clustering strategy [62]. K-means clustering, which was proposed by Pranevicius *et al.* for scenario tree construction and scenario reduction, is a classic clustering method based on distances [63]. Sutiene *et al.* combined the k-means clustering and copula functions to generate individual scenarios in the multivariate structure [64].

Based on the calculated distances, The BR and FS methods eliminate the scenarios that fail to meet certain constraints [65,66]. Basically, the BR and FS methods are repetitions of the reduction for a single scenario until the prescribed reduction amount is satisfied. In addition, the BR method generates a limit set of scenarios reflecting the stochastic characteristics of the original scenarios, providing a more efficient wind power dispatch plan [66]. Although it is easily implemented, the BR and FS methods are computationally inefficient for large problems.

(2) Derived distance methods

The optimization strategies for various distance reduction methods can be different, and several derived methods have been proposed. Li *et al.* proposed a two-dimensional optimization method for SGs [67]. The basic concept of this method is to reduce the observed scenarios from two directions, vertically and horizontally. Representative scenarios of every time period can be obtained by reducing the historical daily wind power series values in the longitudinal direction. The Tabu search algorithm was used for choosing a scenario for each time period to form a representative wind power time series scenario. Sumaili *et al.* used a set of representative scenarios to formulate the uncertainty of wind power forecasting, and these scenarios represent the PDFs of power forecasting [68]. A realistic scenario generator based on clustering was proposed and applied to futures market trading in electricity markets [69]. An optimal reduction method was proposed in [70] to improve the reduction accuracy. Without any assumptions on the probability distribution function, the optimal reduction can consider both spactial and stochastic feature distances.

3.3.3. *Summary of optimization-based method*

A comparison of the optimization-based SG methods and their corresponding advantages, disadvantages, and applications are

Table 4
Comparison of the advantages and disadvantages of optimization-based SG methods.

Optimization-based method	Advantages	Disadvantages	Applications
MM method	Good performance of capturing the stochastic features, capture the relevance of renewable energy	Involves severe NP-hard problem, computationally inefficient for large problems	Multiple stages
Distance method	High calculation efficiency, simplicity	The reduced number of scenarios is predefined, not suitable for medium-scale and large-scale situations, extreme even extreme scenario be hard to capture	Single stage

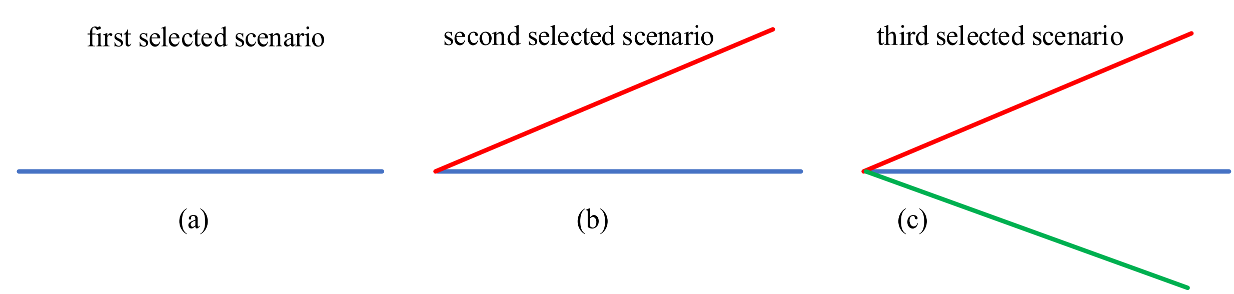

Fig. 12. Schematic process of the FS method.

summarized in Table 4.

3.4. Other SG methods

Several other methods have been proposed for SGs. Goyal *et al.* proposed a decision tree algorithm approach for generating scenarios with better suitabilities [71]. Ma *et al.* assumed wind power fluctuation over a unit interval following a *t*-location-scale distribution and characterized the wind power forecasting error via empirical distributions to generate a large number of wind power scenarios [72]. Duehee *et al.* resented a generalized dynamic factor model (GDFM) that could represent the load and wind power as the sum of the periodic, idiosyncratic noise, and common components to generate load and wind power scenarios [4]. GDFM has the advantage of preserving the correlation structure between the load and wind power. A finite element method that could provide asymptotic convergence for multistage stochastic programming was proposed by Casey *et al.* [73]. Pflug *et al.* presented the dynamic generation of scenario trees for multistage stochastic optimization [74]. In their method, the structure of the tree was not determined beforehand, but was dynamically adapted to meet a distance criterion, which measures the quality of the generated scenarios. With this improvement, the dynamic tree method can be applied to stochastic programs with continuous-state scenarios and higher-dimensional state spaces. Based on scenarios tree, Du *et al.* proposed a scenario mapping technology (SMT) for generating representative renewable energy scenarios [75]. SMT has good performance in characterizing the uncertainty and variability of wind power and limits the model complexity. Yang *et al.* proposed a bootstrapped sampling method. Bootstrapping is a resampling method, also known as a self-help method, with the advantage that the original sample data can be used directly to copy the observation information for statistical inference without assuming the overall distribution or imposing any restrictions [76]. To improve the efficiency of the optimization algorithm, Das *et al.* proposed a tent mapping method to generate load and renewable energy scenarios [77]. Wang *et al.* presented a hybrid SG method combining copula sampling and BR [78], with another hybrid SG method combining LHS and FS proposed in [79]. Jamali *et al.* utilized a roulette-wheel mechanism to generate scenarios for wind power generation and market prices using the Kantorovich distance index to reduce the number of scenarios [80]. This method in [81] has also been applied to establish the uncertainty model of wind power and load demand.

4. Evaluation of SG methods

At present, there are several studies on SG methods, but there exist very few systematic evaluations of SG quality. To improve the accuracy and maximize the economic viability of SG methods, it is necessary to evaluate the quality of generated scenarios to assess whether they reflect the actual characteristics of situations.

We divided the evaluation methods into three groups: output-based, distribution-based, and event-based evaluations [82]. For SG whose outcomes are in the form of expected values, these are classified into the output-based evaluation group. For SG whose outcomes are in the form of discrete probability distributions, classification is achieved by considering the distribution-based evaluation group. Besides these two evaluations of SG outcomes, the event-based grouping evaluates the SG at the overall level. In the following subsection, the three groups are introduced separately.

4.1. Output-based evaluation

Differences exist between the generated and actual scenarios. The mean absolute error (MAE) is the average of the absolute values of deviations of all individual observations from the arithmetic mean [83]. The MAE can reflect the actual prediction errors from the absolute value of the deviations. The root mean square error (RMSE) [84] is the square root of the ratio of the sum of squares of deviations between the observed and true values, and the square root of the ratios of the number of observations. The RMSE is a numerical index for measuring the accuracy of the generated scenarios. The mean absolute percentage error (MAPE) [85] can reflect the degree of generated scenarios deviating from actual scenarios. The formulas for MAE, RMSE, and MAPE are shown in [82,85].

4.2. Distribution-based evaluation

In addition to measuring the errors between the outputs of the generated and actual scenarios, sometimes the generated scenarios can be expressed as discrete probability distributions. In this case, the generated scenarios need to be evaluated using distribution-based evaluation methods.

4.2.1. Distance

The distance, or probability metric, can reflect the similarity between two probability distributions. A commonly used distance is the Wasserstein or Kantorovich–Rubinstein distance, which is a distance function defined between the probability distributions on a given metric space. For instance, the Wasserstein distance is defined as follows:

$$W_\gamma(F_1, F_2) = inf\mathbf{E}[d(X, Y)^\gamma] \tag{9}$$

where E is a mean function, whose infimum is taken over all joint distributions of random variables X and Y with marginal distributions F_1 and F_2, respectively, $d(X, Y)$ is a distance function, γ and is the exponent of the distance function. Euclidean distance is another commonly used distance measurement for scenario pairs. The formulas for the Euclidean distance are shown in [74].

4.2.2. Energy score

The energy score [86] can measure the differences between the discrete probability distribution and actual output. The formula for the energy score is given as

$$E_s = \frac{1}{S}\sum_{i=1}^{S}\|\zeta_i - P\|_2 - \frac{1}{2S^2}\sum_{i=1}^{S}\sum_{j=1}^{S}\|\zeta_i - \zeta_j\|_2 \tag{10}$$

where ζ_i represents the sequence of generated scenarios, P represents the actual output of wind power, and $\|\cdot\|_2$represents the Euclidean norm. From Eq. (10), E_s reflects the difference between the mean of the generated scenarios and the actual value. As a result, for a smaller E_s, smaller differences are observed between the generated and actual scenarios.

4.2.3. Quantile score

For each discrete probability distribution, let q_i represent the mean of the generated scenarios at quantile i, let P_t^w represent the actual value of wind power, and let N represent the number of quantiles. Calculating the score L_i using a pinball loss function yields the following formulas:

$$L_i(q_i, P_t^w) = \begin{cases} \left(1 - \frac{i}{N+1}\right)\cdot(q_i - P_t^w), P_t^w < q_i \\ \frac{i}{N+1}\cdot(P_t^w - q_i),\ P_t^w \geqslant q_i \end{cases} \tag{11}$$

Quantile scoring [87] is defined as the average of L_i scores and is given as

$$\mathrm{QS} = \frac{1}{N}\sum_{i=1}^{N} L_i(q_i, P_t^w) \tag{12}$$

From Eq. (11) and Eq. (12), quantile scoring characterizes the degree of distribution corresponding to the actual values of wind power. Smaller quantile scores indicate that the distribution is closer to the actual value; in other words, the actual power can be obtained from a narrower interval with higher probability. Therefore, a smaller quantile score indicates higher accuracy.

4.3. Event-based evaluation

The event-based evaluation process considers the generated scenarios as entire events instead of individual data. Therefore, the evaluation methods in this section focus on the overall performances of the generated scenarios.

4.3.1. Coverage rate

The coverage rate defines the probability of a generated scenario containing the observed scenarios. The formula for the coverage rate [82] is given as

$$c = \frac{1}{T}\sum_{t=1}^{T} \mathrm{I}\left(p_{min,t}^{w} \leqslant P_{t}^{w} \leqslant p_{max,t}^{w}\right) \tag{13}$$

$$p_{min,t}^{w} = min\left(p_{s,t}^{w}\right), \; s = 1, 2, \cdots, S \tag{14}$$

$$p_{max,t}^{w} = max\left(p_{s,t}^{w}\right), \; s = 1, 2, \cdots, S \tag{15}$$

where c represents the coverage rate, and $\mathbf{I}(\cdot)$ represents a binary variable: if the condition within the brackets is satisfied, the value of the binary variable is 1 and 0 otherwise. $p_{s,t}^{w}$ represents the s-th scenario value in t time period, where S represents the number of scenarios. $p_{min,t}^{w}$,$p_{max,t}^{w}$ represent the maximum and minimum scenario values in time period t, respectively. The coverage rate represents the probability of the observed scenarios being present within the set of generated scenarios. A higher coverage rate indicates that the generated scenarios are more likely to represent observed scenarios, which means that the generated scenarios are more accurate and reliable.

4.3.2. Brier score

The Brier score [82,88] is score function that measures the accuracy of probabilistic predictions. It is applicable to tasks in which predictions must assign probabilities to a set of mutually exclusive discrete outcomes. In this case, the Brier score indicates the differences between certain events occurring in generated and observed scenarios. The smaller the Brier score, the stronger is the ability of the generated scenario to reflect a certain event.

The generated scenarios should reflect certain events in the observed scenarios; therefore, an event-based evaluation is proposed herein. A large change in wind power output in the short term may affect the transient stability and frequency modulation, among others. Hence, two involved events are presented: ramp event and persistent characteristic.

The function of the ramp event can be described as

$$g(\boldsymbol{U}; t, h, \xi) = \mathbf{1}\left(\left(\max_{i\in I} u_i - \min_{i\in I} u_i\right) \geqslant \xi\right) \cdot sign\left(u_{t+h/2} - u_{t-h/2}\right) \tag{16}$$

$$I = \{t - h/2, \cdots, t + h/2\} \tag{17}$$

where h represents the width of a time window centered at time t, u_i represents the i-th element in the sequence $\boldsymbol{U}$, ξrepresents the assigned threshold, which is the percentage of renewable power installed capacity, and $sign(\cdot)$ represents the symbol of the value within the brackets, where positive means upward ramp and negative means downward ramp.

The functional persistence characteristic can be described as:

$$\tilde{g}(\boldsymbol{U}; t, h, \xi) = \prod_{i=t-h/2}^{i=t+h/2} \mathbf{1}(u_i \geqslant \xi) \tag{18}$$

In summary, the general forms of Eqs. (16), (17), and (18) are denoted by $g(\boldsymbol{U};\theta)$, where θ represents a certain event.

The most common formulation of Brier score is given as

$$\boldsymbol{B}_s = \frac{1}{S}\sum_{s=1}^{S} \left[g(\boldsymbol{\zeta}_s;\theta) - g(P;\theta)\right]^2 \tag{19}$$

where $\boldsymbol{B}_s$ is a multi-dimensional vector.

4.3.3. Correlation coefficients

A correlation coefficient is a numerical measure that quantitatively describes the linear statistical relationship between two variables; it can be used to measure the degree of similarity between two variables. In the SG, the correlation coefficients can be used to calculate the similarities between observed and generated scenarios. The most commonly used correlation coefficients include the Pearson product-moment correlation coefficient (PPMCC), tail correlation (TC), Kendall's tau (KT), Spearman's rho (SR), and Gini correlation (GC). The PPMCC is a measure of the strength and direction of the linear correlation between two variables that is defined as the covariance of those variables divided by the product of their standard deviations [89]. The TC is a measure of the probability that random variables increase or decrease at the same time, which mainly reflects the extreme variation probability of occurrence; KT is used to measure whether the trends of random variables are consistent, and the SR is used to measure the correlations among the variable correlation structural variables. Lastly, the GC can be used to measure the consistency of direction and degree of change of the random variables [90].

5. Comprehensive analyses of SG methods in energy systems applications

This section provides a comprehensive analysis of the SG methods used in practical applications involving stochastic programming of wind-power integrated systems. By analyzing the practical application of each SG method in the energy systems, their strengths and weaknesses can be obtained from the practical requirements of the scenarios. In the following section, several practical applications of the SG methods are introduced. The corresponding representative SG methods applied to these situations are analyzed to provide useful information for power systems with integrated wind power.

In current research, the practical applications of SG methods mainly include economic dispatch, unit commitment, wind power forecasting, and strategic planning with uncertainty. A discussion of each application is presented below.

5.1. Comprehensive analysis of SG methods in economic dispatch

In an economic dispatch, the uncertainties of wind power generation, energy price, and system demands need to be simulated to ensure a practical approach. Thus, SG can be employed to help characterize the stochastic processes and contribute to dispatch decision-making. The authors in [89] introduced a novel economic dispatch model based on the theoretical scenario for improving the dispatch performance. In [91], the authors presented a Lagrangian relaxation using the incremental proximal method to ensure efficiency of the economic dispatch problem with a large number of scenarios. In [92], the authors studied the dynamic economic emission dispatch problems for wind-power-integrated systems and employed the roulette-wheel mechanism to generate the load and wind power forecast error scenarios. The authors in [93] proposed an approach to generate multi-wind-farm output scenarios based on the copula correlation theory and applied it to economic dispatch. A novel distance-based mixed-integer quadratic programming model was applied to economic dispatch [94], and the results indicated that the complexity of the economic dispatch problem was reduced.

In general, the target of economic dispatch is to reasonably allocate the outputs of the generating units and wind power generation to minimize the total generation cost in the dispatch period. For instance, for a typical multi-area wind power system, the objective function for the total cost can be defined as

$$TC = \sum_{s,t,\varphi} \rho_s \cdot G_s^{\varphi}(t) \cdot \lambda_s(t) + \sum_{r,t,\varphi} C_r^{\varphi}(t) \tag{20}$$

where ρ_s is the probability of the scenario s. $G_s^{\varphi}(t)$ is the wind power generation at time t, scenario s, and area φ. $\lambda_s(t)$ is the cost coefficient at time t for scenario s. $C_r^{\varphi}(t)$ is the generation cost of r, the thermal unit located at time t and area φ, which can be defined as follows:

$$C_r^{\varphi}(t) = a_r \left(P_r^{\varphi}(t)\right)^2 + b_r P_r^{\varphi}(t) + c_r \tag{21}$$

where a_r, b_r, and c_r are the cost coefficients of the thermal unit r. $P_r^{\varphi}(t)$ is the power produced by the thermal unit r over area φ at time t.

It can be seen from Eq. (20) and Eq. (21) that the economic dispatch problem depends on the probabilities of the scenarios; different methods correspond with different economic dispatch results. Meanwhile, the economic dispatch problem also places requirements on the properties of the SG methods. For instance, the economic dispatch problem is often dynamic and requires probability distributions for each period of study. Therefore, it is necessary to combine the properties of SG methods and the practical requirements of economic dispatch when arriving at relevant decisions. Table 5 lists several common properties of the relevant SG methods that are related to practical situations based on current research on SG methods as applied to economic dispatch.

These comparative results answer several important questions: whether multiple-stage scenarios, dynamic generation, and probability function assumptions are required for SG, and whether the generated scenarios can handle developing situations in events. The observed results are as follows:

1) The multiple-stage scenarios can reflect the sequential features of the temporal adjacent scenarios. Thus, the generation of multiple stages reflects whether practical applications are important to the features of scenarios, such as changing trends.
2) For dynamic generation, the calculation efficiency required is relatively higher than that for non-dynamic generation, which determines the demand for simplicity and calculation efficiency in SG methods.
3) From the theoretical principles of the SG methods, a probability function assumption may help improve the calculation efficiency and ease implementation. However, current assumptions cannot fully fit actual wind power generation situations, thereby leading to inaccuracies in the mathematical model.
4) In economic dispatches, specific events should be considered, such as the short-term large-scale variation of wind power outputs. Therefore, the development of events should be included to reduce their influence.

Table 5
Comparison of SG methods in economic dispatch.

Economic dispatch	Multiple-stage scenarios	Dynamic generation	Function assumption	Grasp the developing situation of the events
MC method	×	×	√	×
MCMC method	√	×	√	√
PVMC method	√	×	√	√
LHS method	×	√	√	×
Copula function sampling	√	√	×	√
ARMA method	√	√	×	√
Machine learning method	√	√	×	√
Distance method	×	√	×	×
MM method	×	×	×	×

5.2. Comprehensive analysis of SG methods in unit commitment

A common practice for representing system uncertainties is by using a large number of scenarios, which results in a large deterministic equivalent problem. In such cases, SG methods can be utilized to simulate the features of the wind power outputs, which then helps dispatchers allocate unit commitment.

Du *et al.* proposed a novel scenario map technique to address the commitment problem with uncertain wind power generation [75]. Ummels *et al.* utilized a new simulation-based method to assess the impact of wind power on unit commitment [95]. In [96], the LHS method was implemented to simulate the possible scenarios for representing intermittency and volatility of wind power generation. It was demonstrated through simulations that the security of power system operation could be effectively solved by simulating wind power generation scenarios. In [97], the authors proposed a two-stage stochastic programming model for unit commitment with large amounts of wind power. And wind power scenarios were generated from autoregressive methods.

In [98], the authors presented a summary of the scenario reduction method for unit commitment. The uncertainty model used in unit commitment consists of a set of scenarios; each scenario represents a possible realization of the uncertainty and has a certain associated probability. Wang *et al.* proposed a probability-based scenario to solve the unit commitment problem. Wind power forecasting values and errors could be employed in sampling and generating large-scale scenarios to simulate the fluctuations of wind power outputs. Then, these scenarios could be reduced to representative scenarios, whose probability information can be efficiently used in unit commitment by the lambda-iteration method [99]. In [100], a two-stage scenario-based stochastic programming model was presented for unit commitment to take into account the wind power uncertainties, by considering optimal scheduling and minimizing the total cost. The authors of [101] presented a risk-constrained bidding strategy to solve the price-based unit commitment problem, where the scenarios of market price uncertainties were simulated using Monte Carlo simulations. In [102], the authors proposed an approach combining Markov and interval optimizations [103] to solve the transmission-constrained unit commitment problem with uncertain wind power generation.

For scenario-based unit commitment, the general objective function can be defined as

$$minE(C_M) = \sum_{s=1}^{S} \rho_s C_M(s) \tag{22}$$

This minimizes the total expected cost in S representative scenarios, where s are scenarios from 1 to S. $C_M(s)$ is the operating cost in the s-th scenario, which includes the startup cost, fuel consumption cost, wind curtailment penalty term, and load shedding penalty term. The probability of each scenario is ρ_s.

It can be seen from Eq. (22) that the solution to the scenario-based unit commitment problem depends on the probability of each scenario. Table 6 lists several common properties of relevant SG methods that are related to practical situations.

These comparative results answer several important questions: whether multiple-stage scenarios, dynamic generation, and probability function assumptions are required for SG. The observed results are as follows:

Table 6
Comparison of SG methods in unit commitment.

Unit commitment	Multiple-stage scenarios	Dynamic generation	Function assumption
MC method	×	×	√
MCMC method	√	×	√
PVMC method	√	×	√
LHS method	×	√	√
Copula function sampling	√	√	×
ARMA method	√	√	×
Machine learning method	√	√	×
Distance method	×	√	×
MM method	×	×	×

1) For unit commitment, introducing large-scale wind power scenarios in the stochastic unit commitment model will result in large amounts of computation. Hence, optimization-based methods are often applied because they can help simplify the calculations.
2) Unit commitment is often a dynamic problem that requires the operator to allocate constantly. Therefore, dynamic generation, which is dependent on the complexity and calculation efficiency of SG methods, is significant.
3) Since the relevant reduction-based methods extract scenarios that meet constraints from actual observations, a probability function assumption is not necessary for reduction.

5.3. Comprehensive analysis of SG methods in wind power forecasting

SG methods can help characterize the stochastic features of wind power. In [104], the authors proposed an approach to generate scenarios to construct the prediction intervals of extreme learning machine based regression. The proposed bootstrap-based extreme learning machine method is a generalized framework for probabilistic forecasting of wind power as an efficient and meaningful online tool for power system applications, including wind-power forecasting. In [105,106], the authors proposed a novel approach to directly formulate the prediction intervals of wind power generation based on the extreme learning machine and PSO; this approach was proven to be highly efficient and reliable through preliminary case studies using real-world wind-farm data, indicating high potential for practical applications. In [107], the authors proposed an ensemble prediction method to forecast the wind power density using generalized autoregressive conditional heteroskedasticity models for wind speed series; thus, the wind power generator series was obtained approximately from the relationship between wind speed and wind power. Moreover, with increasing penetration of wind energy, wind power ramp events are now drawing more attention as they considerably affect power systems operations. Here, forecasting is no longer confined to output forecasting, and can also be used for event forecasting, such as wind power ramp event forecasting. In [108], a probabilistic ramp forecasting method based on an ANN was proposed. The numerical simulation results showed that the developed probabilistic wind power ramp forecasting method had higher accuracy than the conventional methods.

Table 7 lists several common properties of the relevant representative SG methods related to practical situations in current research.

The significant results from the comparative analyses of SG methods related to wind power forecasting are as follows:

1) Multiple-stage scenarios in wind power forecasting refer to predicting multiple temporal adjacent scenarios simultaneously. For a forecasting type category, such as day-ahead forecasting, continuous generation helps power system allocations and storage.
2) For real-time forecasting and dispatch situations, dynamic forecasting plays a significant role owing to its fast corresponding characteristics; however, it has higher data requirements.
3) It has been proven that wind power prediction errors do not obey the existing type distributions [3], which indicate that modeling for wind power outputs with probability function assumptions can lead to inaccuracies in reflecting the properties of wind power. Because wind power forecasting results are required to depict the detailed properties of wind power, inaccuracies caused by probability function assumptions should be avoided as much as possible.

5.4. Comprehensive analysis of SG methods in strategic planning

The purpose of strategic planning for power systems is to optimize the financial incentives to achieve the desired goals under various environments. Owing to the uncertainties of modern power systems with respect to wind-power integration, SG methods can be applied for these planning purposes. Specifically, planning can be divided into network and storage capacity planning. These two aspects have different emphases: network planning focuses on the transmission expansion of the network [109], whereas storage capacity planning is oriented towards a single power system [110].

In [111], the authors presented a multi-object framework for energy storage planning and modeled the uncertainty of wind power by scenario analysis. In [112], the authors proposed a two-stage stochastic programming approach for a home energy management system, including battery energy storage and electric vehicles, based on an existing method [6]. In [113], the authors proposed a method to solve the transmission network expansion planning problem with multiple generation scenarios. In [114], the authors used a clustering method to generate typical scenarios, so as to improve the efficiency of solving the transmission network expansion planning problem and obtain the optimal investment plan.

Table 8 lists several common properties of relevant representative SG methods that are related to practical situations according to current research on SG methods in strategic planning.

From Table 8, the following observations are made:

1) For strategic planning, the planning strategies are often determined in advance; the sequential properties matter less in strategic planning.

Table 7
Comparison of SG methods in wind power forecasting.

Wind power forecasting	Multiple-stage scenarios	Dynamic generation	Function assumption	Point prediction	Interval prediction	Probability prediction	Event prediction
MC method	×	×	√	×	×	√	×
MCMC method	√	×	√	×	×	√	×
LHS method	×	√	√	×	×	√	×
Copula function sampling	√	√	×	×	×	√	×
ARMA method	√	√	×	√	×	×	×
Machine learning method	√	√	×	√	×	×	√

Table 8
Comparison of SG methods in strategic planning.

Network planning, Storage capacity planning	Multiple-stage scenarios	Dynamic generation	Function assumption
MC method	×	×	√
MCMC method	√	×	√
PVMC method	√	×	√
LHS method	×	√	√
Copula function sampling	√	√	×
ARMA method	√	√	×
Machine learning method	√	√	×
Distance method	×	√	×
MM method	×	×	×

2) Dynamic generation is typically used to solve problems with constant allocation; however, the existence of dynamic factors is also related to the dynamic generation of SG methods.
3) Strategic planning attaches much importance to the future state of a power system. With a probability function assumption, the state and features are easier to analyze quantitatively; however, without the probability function assumption, model inaccuracies can be minimized.

5.5. Comprehensive analysis of SG methods in integrated energy systems

The uncertainties of renewable generation and energy demand have presented serious challenges to integrated energy systems (IES) operation schedules [115,116]. To handle such multiple uncertainties in renewable energy generation and energy demand, a scenario-based method was employed in the stochastic optimal operation of the IES [117]. Considering the correlations between multiple energy sources, a multi-dimensional correlation SG approach to solve the multi-energy power flow of the IES was proposed in [118]. In [119], the authors proposed an optimization-based SG method to generate typical scenarios for the IES, in which the Wasserstein distance was used as a metric. The work in [120] explored the optimal allocations of regional integrated energy based on typical scenarios of cluster generation.

5.6. Summary of SG methods

A detailed literature review and discussion of the SG methods for stochastic programming was presented above. As noted in Section 3, several reported works have been devoted to wind-power SG methods. Table 9 summarizes the existing SG methods, SG problem of each work, and the properties of the generated scenarios. It can be observed from the table that the main problems associated with SG include optimal power flow, unit commitment, reliability analysis, transmission

Table 9
Summary the application of SG methods.

Scenario generation methods	Problem under consideration	The properties of generated scenarios						
		Single	Multiple	Univariate	Multivariable	Assumption	Temporal	Spatial
Sampling-based method								
MC simulation	Probabilistic power flow [121]	√	×	√	×	√	×	×
	Stochastic economic viability analysis [122,123]	√	×	√	×	√	×	×
MCMC method	Generate synthetic wind power time series [25]	×	√	√	×	×	√	×
PVMC method	Generate synthetic wind power time series [27]	×	√	√	×	√	√	×
LHS method	Reliability analysis [124]	√	×	√	×	√	×	×
	Probabilistic load flow [30,78]	√	×	√	×	√	×	×
	Security-constrained unit commitment with wind power [96]	√	×	√	×	√	×	×
Combination LHS and Nataf transformations	Probabilistic load flow [31]	√	×	√	√	√	×	√
T-copula sampling	Stochastic optimal power flow [33]	√	×	√	√	×	×	√
D-vine copulas sampling	Transmission expansion planning [35]	√	×	√	√	×	×	√
Mixture vine copulas sampling	Generate multi-wind power scenarios [36]	√	×	√	√	×	×	√
Forecasting-based method								
ARMA method	Economic feasibility analysis of V2G [42]	×	√	√	×	×	√	×
	Power system operation and planning [44]	×	√	√	√	×	√	√
	Reliability assessment of power systems [45]	×	√	√	×	×	√	×
	Reliability analysis [125]	×	√	√	√	×	√	√
State space	Analysis of wind power generation capacity [46]	×	√	√	√	×	√	√
ANNs	Simulation the load, photovoltaic and wind production [47]	×	√	√	×	×	√	×
	Forecast wind power ramp events [48]	×	√	√	×	×	√	×
RBFNNs	Stochastic unit commitment [49]	×	√	√	×	×	√	×
GANs	Renewables scenario generation [51]	×	√	√	√	×	√	√
Improved GANs	Multiple wind power generation scenarios [52]	×	√	√	√	×	√	√
LSTM	Electricity markets [50]	×	√	√	√	×	√	√
Optimization-based method								
MM method	Financial portfolio optimization [53]	√	×	√	×	×	×	×
Heuristic MM	Distributed generation investment planning [56]	√	×	√	√	×	×	√
K-mean clustering	Wind power investment problem [61]	√	×	√	×	×	×	×
Affinity propagation clustering	Stochastic optimal power flow [62]	√	×	√	×	×	×	×
BR method	Power system optimization decision [66]	√	×	√	×	√	×	×
Submodularity-based scenario reduction	Transmission expansion planning [18]	√	×	√	×	×	×	×
A two-dimensional optimization	Optimal power flow [67]	√	×	√	×	×	×	×
Optimal model	Generate wind power time series [70]	√	×	√	×	×	×	×

expansion planning, and renewable power generation. Moreover, the generated scenarios vary by several properties, such as single stage, multiple stages, univariate, multivariable, function assumption, temporal relationship, and spatial relationship.

As mentioned above, SG methods have been widely used in energy systems, but it is very important to choose an appropriate method in practical scenarios. Hence, a discussion on the abilities of SG methods in different applications was presented, and the corresponding evaluation indicators were provided as guidance for researchers. A comprehensive comparison of the performance of the SG methods for different applications of energy systems was provided, and a comprehensive reference for researchers to solve the uncertainties of wind-power-integrated systems is provided, as shown in Table 10.

6. Conclusion and future works

6.1. Conclusions regarding SG methods

Research and studies on SGs have contributed significantly to improving and optimizing stochastic programs for power systems. Further, with the highly accurate generated scenarios, "passive reply" in power systems can be converted to "active regulation" to greatly aid the decision-making capabilities of operators and dispatchers of power systems. Hence, research on SG has great practical significance. The features and challenges of applying the aforementioned three types of SG methods are summarized as follows:

(1) Sampling-based methods have the advantages of simplicity and rapidity, thereby avoiding the complicated process of mathematical deduction. Monte Carlo, Latin hypercube sampling, and sample average approximation are some typical sampling-based methods, which are suitable for generating single-stage and single-variable scenarios. Sampling-based methods are widely applied to calculate probabilistic power flows, transmission expansion planning, and economic dispatch. However, because they are not suitable for capturing the temporal and spatial correlations of variables, they cause limitations in the multiple-stage and multiple-variable scenarios. Furthermore, the probabilistic distribution of renewable energy power needs to be assumed, which causes bias between the generated scenarios and their true probabilistic distributions.

(2) Forecasting-based methods are good at capturing the characteristics of variables, including correlation, complex nonlinear relationships of the time series, persistence, and variation features. Furthermore, without any assumption of the distribution function, the forecasting-based methods train models using historical data to generate scenarios. Auto-regressive moving average and machine learning are some of the typical forecasting-based methods. These methods are widely applied in multiple-stage and multiple-variable problems, for example, forecasting of power from renewable energy sources, economic dispatch, unit commitment, reliability assessment, and capacity optimization. However, forecasting-based methods are generally based on data-driven methods, and the qualities of the generated scenarios depend on historical observation samples.

(3) Optimization-based methods have high accuracies in scenario approximations. Moment matching, backward reduction, and forward selection are some of the typical optimization-based methods. These methods are suitable for single-/multiple-stage and single-/multiple-variable problems and are widely applied in operation optimization, investment problems, planning, and power system decisions, among others. However, optimization-based methods involve a severe NP-hard problem, which solves complex difficulties but difficult to apply in large-scale power systems.

6.2. Future work

Scenario generation of power from renewable energy sources is a popular topic that has recently been studied extensively. However, more research needs to be conducted in the near future for better understanding:

(1) Enhance the quality of historical data for generation: Most of the current SG methods depend on large numbers of samples. However, such large data requirements result in poor adaptability. Therefore, efficient analyses of the quality of historical data must be explored to widen the range of applications.

(2) Propose new efficient SG methods: Owing to renewable energy integration, uncertainty has become an inevitable problem in analysis. Currently, many SG methods describe the uncertain information concerning wind power by assuming probability distribution functions, which may not precisely describe the characteristics of wind power. Therefore, other methods without probability models have been proposed, such as machine learning. It is worth exploring how the characteristics of wind power can be minimized by accurately describing the assumptions to adapt to specific situations.

(3) Set up functions to describe stochastic distribution features more accurately: Future SG methods should demand more stringent distribution feature requirements. In the future, the demands of the generated distributions are not only constrained by availability but also focus on accurate distribution features, such as sharpness and thick tail.

(4) Build a reliable and universal system for identifying practical applications for SG methods: Currently, SG methods have been increasingly applied in power systems. To reduce the computational complexities associated with these methods, the number of generated scenarios is typically limited. In fact, a large number of scenarios are needed to describe the characteristics of wind power. Hence, the computational efficiency is low, and methods to effectively use these strategies in operation and planning are extremely significant in future research.

(5) Improve the method for dynamic scenarios with multiple renewable resources: Most existing SG methods can be applied only to the generation of single-site or system aggregation total output scenarios. With large-scale renewable energy integration

Table 10
Comparison of the performance of SG methods in different applications.

Applications	Scenario generation methods							Evaluation indicator
	MC	LHS	Copula function	ARMA	Machine learning	MM	Distance matching	
Economic dispatch	*	***	***	*	*	*	***	Energy score
Unit commitment	***	***	***	***	*	**	***	Brier score
Forecasting	*	*	**	***	***	*	*	Brier score
Strategic planning	***	***	**	*	*	***	***	Coverage rate
IES	*	*	***	*	***	*	***	Distance

Note: The number of *, **, ***represents the low, medium, and high suitability of SG methods.

systems, the renewable energy outputs have complex spatio-temporal distributions. The SG of multi-regional renewable resources is a difficult problem that must be studied in depth because of the space-time distribution characteristics.

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.

Acknowledgement

This work is supported by the National Key Research and Development Program of China (2016YFB0900100) and National Natural Science Foundation of China (51977042).

References

[1] Li JH, Ye L, Zeng Y, Wei H. A Scenario-Based Robust Transmission Network Expansion Planning Method for Consideration of Wind Power Uncertainties. CSEE J Power Energy Syst 2016;2(1):11–8.
[2] Xue Y, Lei X, Xue F, Yu C, Dong Z, Wen F, et al. A Review on Impacts of Wind Power Uncertainties on Power Systems. Proceedings of the CSEE 2014;34(29):5029–40.
[3] Yan J, Liu Y, Han S, Wang Y, Feng S. Reviews on uncertainty analysis of wind power forecasting. Renew Sustain Energy Rev 2015;2015(52):1322–30.
[4] Duehee L, Ross B. Load and Wind Power Scenario Generation Through the Generalized Dynamic Factor Model. IEEE Trans Power Syst 2017;32(1):400–10.
[5] Kaut M, Wallace SW. Evaluation of Scenario-Generation Methods for Stochastic Programming. Pacific J Optim 2003;3(2):14–2003.
[6] Kaut M. Forecast-based scenario-tree generation method. Optimization. Online 2017.
[7] Lucheroni C, Boland J, Ragno C. Scenario generation and probabilistic forecasting analysis of spatio-temporal wind speed series with multivariate autoregressive volatility models. Appl Energy 2019;239:1226–41.
[8] Li J, Zhu D. Combination of moment-matching, Cholesky and clustering methods to approximate discrete probability distribution of multiple wind farms. IET Renew Power Gener 2016;10(9):1450–8.
[9] Lin C, Fang C, Chen Y, Liu S, Bie Z. Scenario generation and reduction methods for power flow examination of transmission expansion planning. In: 2017 IEEE 7th International Conference on Power and Energy Systems (ICPES); 2017. p. 90–5.
[10] Hart EK, Jacobson MZ. A Monte Carlo approach to generator portfolio planning and carbon emissions assessments of systems with large penetrations of variable renewables. Renew Energy 2011;36(8):2278–86.
[11] Yu H, Chung CY, Wong KP, Lee HW, Zhang JH. Probabilistic Load Flow Evaluation With Hybrid Latin Hypercube Sampling and Cholesky Decomposition. IEEE Trans Power Syst 2009;24(2):661–7.
[12] Becker R. Generation of time-coupled wind power infeed scenarios using pair-copula construction. IEEE Trans Sustain Energy 2017;9(3):1298–306.
[13] Haghi HV, Lotfifard S. Spatiotemporal modeling of wind generation for optimal energy storage sizing. IEEE Trans Sustain Energy 2014;6(1):113–21.
[14] Boone A. Simulation of short-term wind speed forecast errors using a multi-variate ARMA (1, 1) time-series model. Master thesis, KTH Roy. Inst. Technol, Stockholm, Sweden, 2005.
[15] Matevosyan J, Soder L. Minimization of imbalance cost trading wind power on the short-term power market. IEEE Trans Power Syst 2006;21(3):1396–404.
[16] Sideratos G, Hatziargyriou ND. Probabilistic wind power forecasting using radial basis function neural networks. IEEE Trans Power Syst 2012;27(4):1788–96.
[17] Jiang C, Mao Y, Chai Y, Yu M, Tao S. Scenario generation for wind power using improved generative adversarial networks. IEEE Access 2018;6:62193–203.
[18] Park SW, Xu Q, Hobbs BF. Comparing scenario reduction methods for stochastic transmission planning. IET Gener Transm Distrib 2019;13(7):1005–13.
[19] Mehrotra S, Papp D. Generating moment matching scenarios using optimization techniques. SIAM J Optim 2013;23(2):963–99.
[20] Lin J, Cheng L, Chang Y, Zhang K, Shu B, Liu G. Reliability based power systems planning and operation with wind power integration: A review to models, algorithms and applications. Renew Sustain Energy Rev 2014;31(Complete):921–34.
[21] Dong G, Chen Z, Wei J. Sequential Monte Carlo filter for state of charge estimation of lithium-ion batteries based on auto regressive exogenous model. IEEE Trans Ind Electron 2019;66(11):8533–44.
[22] Xie Z, Ji T, Li M, Wu Q. Quasi-Monte Carlo based probabilistic optimal power flow considering the correlation of wind speeds using copula function. IEEE Trans Power Syst 2017;33(2):2239–47.
[23] Rakipour D, Barati H. Probabilistic optimization in operation of energy hub with participation of renewable energy resources and demand response. Energy 2019;173:384–99.
[24] Cheng J, Gicquel C, Lisser A. Partial sample average approximation method for chance constrained problems. Optim Lett 2019;13(4):657–72.
[25] Papaefthymiou G, Klockl B. MCMC for Wind Power Simulation. IEEE Trans Energy Convers 2008;23(1):234–40.
[26] Tang C, Wang Y, Xu J, Sun Y, Zhang B. Efficient scenario generation of multiple renewable power plants considering spatial and temporal correlations. Appl Energy 2018;221:348–57.
[27] Li J, Li J, Wen J, Cheng S, Xie H, Yue C. Generating wind power time series based on its persistence and variation characteristics. Sci China Technol Sci 2014;57(12):2475–86.
[28] Alahyari A, Ehsan M, Mousavizadeh M. A hybrid storage-wind virtual power plant (VPP) participation in the electricity markets: A self-scheduling optimization considering price, renewable generation, and electric vehicles uncertainties. J Storage Mater 2019;25:100812.
[29] Swamy C, Shmoys DB. Sampling-based approximation algorithms for multi-stage stochastic optimization. IEEE Symposium on Foundations of Computer Science. IEEE; 2012.
[30] Cai D, Shi D, Chen J. Probabilistic load flow computation with polynomial normal transformation and Latin hypercube sampling. IET Gener Transm Distrib 2013;7(5):474–82.
[31] Chen Y, Wen J, Cheng S. Probabilistic Load Flow Method Based on Nataf Transformation and Latin Hypercube Sampling. IEEE Trans Sustain Energy 2013;4(2):294–301.
[32] Sklar A. Fonctions de repartition à n dimensions et leurs marges. Publication de l'Institut de Statistique de l'Université de Paris 1959;8:229–31.
[33] Li J, L F. Copula-Based Monte Carlo Scenarios Generation Method for STOPF Problem. Electricity, 2014(z1):41–50.
[34] Hoeltgebaum H, Fernandes C, Street A. Generating Joint Scenarios for Renewable Generation: The Case for Non-Gaussian Models With Time-Varying Parameters. IEEE Trans Power Syst 2018;33(6):7011–9.
[35] 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.
[36] Qiu Y, Li Q, Pan Y, Yang H, Chen W. A scenario generation method based on the mixture vine copula and its application in the power system with wind/hydrogen production. Int J Hydrogen Energy 2019;44(11):5162–70.
[37] 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.
[38] Meibom P, Barth R, Hasche B, Brand H, Weber C, O'Malley M. Stochastic Optimization Model to STudy the Operational Impacts of High Wind Penetrations in Ireland. IEEE Trans Power Syst 2011;26(3):1367–79.
[39] Broersen PMT, De Waele S. Finite sample properties of ARMA order selection. IEEE Trans Instrum Meas 2004;53(3):645–51.
[40] Gao Z, Mao A, Chen D, Song Y. A Wind Farm Capacity Credibility Calculation Method Based on Parabola. Appl Mech Mater 2014;472:953–7.
[41] Wangdee W, Billinton R. Probing the Intermittent Energy Resource Contributions From Generation Adequacy and Security Perspectives. IEEE Trans Power Syst 2012;27(4):2306–13.
[42] Ghofrani M, Arabali A, Etezadi-Amoli M, Fadail MS. Smart Scheduling and Cost-Benefit Analysis of Grid-Enabled Electric Vehicles for Wind Power Integration. IEEE Trans Smart Grid 2014;5(5):2306–13.
[43] Abbasi M, Taki M, Rajabi A, Li L, Zhang J. Coordinated operation of electric vehicle charging and wind power generation as a virtual power plant: A multi-stage risk constrained approach. Appl Energy 2019;239:1294–307.
[44] Morales J, Mínguez R, Conejo AJ. A methodology to generate statistically dependent wind speed scenarios. Appl Energy 2010;87(3):843–55.
[45] Chen P, Pedersen T, Bak-Jensen B, Chen Z. ARIMA-Based Time Series Model of Stochastic Wind Power Generation. IEEE Trans Power Syst 2010;25(2):667–76.
[46] Díaz G, Gómez-Aleixandre J, Coto J. Wind power scenario generation through state-space specifications for uncertainty analysis of wind power plants. Appl Energy 2016;162:21–30.
[47] Vagropoulos SI, Kardakos EG, Simoglou CK, Bakirtzis AG, Catalao JP. ANN-based scenario generation methodology for stochastic variables of electric power systems. Electr Power Syst Res 2016;134:9–18.
[48] Cui M, Ke D, Sun Y, Gan D, Zhang J, Hodge BM. Wind Power Ramp Event Forecasting Using a Stochastic Scenario Generation Method. IEEE Trans Sustain Energy 2015;6(2):422–33.
[49] Pappala VS, Erlich I, Rohrig K, Dobschinski J. A stochastic model for the optimal operation of a wind-thermal power system. IEEE Trans Power Syst 2009;24(2):940–50.
[50] Stappers B, Paterakis NG, Kok K, Gibescu M. A Class-Driven Approach Based on Long Short-Term Memory Networks for Electricity Price Scenario Generation and Reduction. IEEE Trans Power Syst 2020.
[51] Chen Y, Wang Y, Kirschen D, Zhang B. Model-free renewable scenario generation using generative adversarial networks. IEEE Trans Power Syst 2018;33(3):3265–75.
[52] Zhang Y, Ai Q, Xiao F, Hao R, Lu T. Typical wind power scenario generation for multiple wind farms using conditional improved Wasserstein generative adversarial network. Int J Electr Power Energy Syst 2020;114:105388.
[53] Ponomareva K, Roman D, Date P. An algorithm for moment-matching scenario generation with application to financial portfolio optimisation. Eur J Oper Res 2015;240(3):678–87.
[54] Xu D, Chen Z, Yang L. Scenario tree generation approaches using K-means and LP moment matching methods. J Comput Appl Math 2012;236(17):4561–79.

[55] Høyland K, Kaut M, Wallace SW. A heuristic for moment-matching scenario generation. Comput Optim Appl 2003;24(2–3):169–85.
[56] Ehsan A, Yang Q. Scenario-based investment planning of isolated multi-energy microgrids considering electricity, heating and cooling demand. Appl Energy 2019;235:1277–88.
[57] Ehsan A, Cheng M, Yang Q. Scenario-based planning of active distribution systems under uncertainties of renewable generation and electricity demand. CSEE J Power Energy Syst 2019;5(1):56–62.
[58] Rubasheuski U, Oppen J, Woodruff DL. Multi-stage scenario generation by the combined moment matching and scenario reduction method. Oper Res Lett 2014;42(5):374–7.
[59] Growe-Kuska N, Heitsch H, Romisch W. Scenario reduction and scenario tree construction for power management problem. In: Power Tech Conference Proceedings, 2003 IEEE Bologna. IEEE; 2004.
[60] Li B, Sedzro K, Fang X, Hodge BS. A clustering-based scenario generation framework for power market simulation with wind integration. J Renew Sustain Energy 2020;12(3):036301.
[61] Baringo L, Conejo AJ. Correlated wind-power production and electric load scenarios for investment decisions. Appl Energy 2013;101:475–82.
[62] Guan L, Wen B, Zhan X, Zhou B, Zhao W. Scenario Generation of Wind Power Based on Longitudinal-Horizontal Clustering Strategy. In: 2018 IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia), Singapore, 934–9; 2018.
[63] Pranevicius H, Sutiene K. Scenario tree generation by clustering the simulated data paths. 2007.
[64] Sutiene K and Pranevicius H. Scenario Generation Employing copulas. Proceeding of the world congress on engineering. UK: London; 2007.
[65] Henrion R, Kuchler C, Romisch W. Discrepancy distances and scenario reduction in two-stage stochastic mixed-integer programming. J Indus Manage Optim 2017;4(2):363–84.
[66] Razali NMM, Hashim AH. Backward reduction application for minimizing wind power scenarios in stochastic programming. Power Engineering and Optimization Conference (PEOCO), 4th International. 2010.
[67] Li, J, Sun H, Wen J, Cheng S, Luo W, Ge w, et al. A Two-dimensional Optimal Technology for Constructing Wind Power Time Series Scenarios. Proceedings of the CSEE 2014;34(16):2544–551.
[68] Sumaili J, Keko H, Miranda M, Zhou Z, Botterud A, Wang J. Finding representative wind power scenarios and their probabilities for stochastic models. International Conference on Intelligent System Application to Power Systems. IEEE; 2011.
[69] Silva F, Teixeira B, Brígida Pinto T, Santos G, Vale Z, Praca I. Generation of realistic scenarios for multi-agent simulation of electricity markets. Energy 2016;116:128–39.
[70] Li J, Lan F, Wei H. A Scenario Optimal Reduction Method for Wind Power Time Series. IEEE Trans Power Syst 2016;31(2):1657–8.
[71] Goyal MK, Ojha CSP. Evaluation of Rule and Decision Tree Induction Algorithms for Generating Climate Change Scenarios for Temperature and Pan Evaporation on a Lake Basin. J Hydrol Eng 2014;19(4):828–35.
[72] Ma X, Sun Y, Fang H. Scenario Generation of Wind Power Based on Statistical Uncertainty and Variability. IEEE Trans Sustain Energy 2013;4(4):894–904.
[73] Casey MS, Sen S. The Scenario Generation Algorithm for Multistage Stochastic Linear Programming. Mathem Oper Res 2005;30(3):615–31.
[74] Pflug G, Pichler A. Dynamic generation of scenario trees. Comput Optim Appl 2015;62(3):641–68.
[75] Du E, Zhang N, Kang C, Xia Q. Scenario Map Based Stochastic Unit Commitment. IEEE Trans Power Syst 2018;33(5):4694–705.
[76] Yang X, Xue M, Guo F, Zhang H, Zhang J. Wind power probability interval prediction based on Bootstrap quantile regression method. Chinese Automation Congress (CAC) 2017:1504–9.
[77] Das S, Basu M. Day-ahead optimal bidding strategy of microgrid with demand response program considering uncertainties and outages of renewable energy resources. Energy 2020;190:116441.
[78] Wang C, Liu C, Tang F, Liu D, Zhou Y. A scenario-based analytical method for probabilistic load flow analysis. Electr Power Syst Res 2020;181:106193.
[79] Daneshvar M, Mohammadi-Ivatloo B, Zare K, Asadi S. Two-stage stochastic programming model for optimal scheduling of the wind-thermal-hydropower-pumped storage system considering the flexibility assessment. Energy 2020;193:116657.
[80] Jamali A, Aghaei J, Esmaili M, Nikoobakht A, Niknam T, Shafie-kha M, et al. Self-scheduling approach to coordinating wind power producers with energy storage and demand response. IEEE Trans Sustain Energy 2019;11(3):1210–9.
[81] Biswas P, Suganthan PN, Mallipeddi R, Amaratunga GAJ. Optimal reactive power dispatch with uncertainties in load demand and renewable energy sources adopting scenario-based approach. Appl Soft Comput 2019;75:616–32.
[82] Wang X, Hu Z, Zhang M. Research on Establishment of Quality Evaluation Framework of Short-Term Wind Power Scenarios. Power Syst Technol 2017;5:33.
[83] Prosper MA, Otero-Casal C, Canoura Fernandez F, Miguez-Macho G. Wind power forecasting for a real onshore wind farm on complex terrain using WRF high resolution simulations. Renew Energy 2019;135(674–686).
[84] Gao Y, Xue F, Yang W, Yang Q, Sun Y, Sun Y, et al. Optimal operation modes of photovoltaic-battery energy storage system based power plants considering typical scenarios. Protection Control Modern Power Syst 2017;2(1):36.
[85] Ouyang T, Zha X, Qin L. A combined multivariate model for wind power prediction. Energy Convers Manage 2017;144:361–73.
[86] Pinson P, Girard R. Evaluating the quality of scenarios of short-term wind power generation. Appl Energy 2012;96:12–20.
[87] Pinson P, Kariniotakis G. Conditional prediction intervals of wind power generation. IEEE Trans Power Syst 2010;25(4):1845–56.
[88] Hyndman RJ, Koehler AB. Another look at measures of forecast accuracy. Int J Forecast 2006;22(4):679–88.
[89] Ming H, Xie L, Campi M, Garatti S, Kumar PR. Scenario-based economic dispatch with uncertain demand response. IEEE Trans Smart Grid 2017;10(2):1858–68.
[90] Ma R, Xu W, Liu S, Zhang Y, Xiong J. Asymptotic mean and variance of Gini correlation under contaminated Gaussian model. IEEE Access 2016;4:8095–104.
[91] Tang C, Xu J, Tan Y, Sun Y, Zhang B. Lagrangian relaxation with incremental proximal method for economic dispatch with large numbers of wind power scenarios. IEEE Trans Power Syst 2019;34(4):2685–95.
[92] Aghaei J, Niknam T, Azizipanah-Abarghooee R, Arroyo JM. Scenario-based dynamic economic emission dispatch considering load and wind power uncertainties. Int J Electr Power Energy Syst 2013;47:351–67.
[93] Xie M, Xiong J, Ke S, Liu M. Two-Stage Compensation Algorithm for Dynamic Economic Dispatching Considering Copula Correlation of Multi-wind Farms Generation. IEEE Trans Sustain Energy 2016;8(2):1.
[94] Pourakbari-Kasmaei M, Rider MJ, Mantovani JRS. An Unambiguous Distance-Based MIQP Model to Solve Economic Dispatch Problems with Disjoint Operating Zones. IEEE Trans Power Syst 2016;31(1):825–6.
[95] Ummels BC, Gibescu M, Pelgrum E, Kling W, Brand AJ. Impacts of Wind Power on Thermal Generation Unit Commitment and Dispatch. IEEE Trans Energy Convers 2007;22(1):44–51.
[96] Wang J, Shahidehpour M, Li Z. Security-Constrained Unit Commitment With Volatile Wind Power Generation. IEEE Trans Power Syst 2008;23(3):1319–27.
[97] Papavasiliou A, Oren SS, O'Neill RP. Reserve Requirements for Wind Power Integration: A Scenario-Based Stochastic Programming Framework. IEEE Trans Power Syst 2011;6(4):2197–206.
[98] Dvorkin Y, Wang Y, Pandzic H, Kirschen D. Comparison of scenario reduction methods for the stochastic unit commitment. In: Pes General Meeting | Conference & Exposition. IEEE, 214;1–5.
[99] Ji D, Wang H. A scenario probability based method to solve unit commitment of large scale energy storage system and thermal generation in high wind power penetration level system. In: Power and Energy Engineering Conference. IEEE; 2016. p. 84–8.
[100] Wang X, Hu Z, Zhang M, Hu M. Two-stage stochastic optimization for unit commitment considering wind power based on scenario analysis. In: China International Conference on Electricity Distribution. IEEE, 2016.
[101] Li T, Shahidehpour M, Li Z. Risk-Constrained Bidding Strategy With Stochastic Unit Commitment. IEEE Trans Power Syst 2007;22(1):449–58.
[102] Yu Y, Luh PB, Litvinov E, Zheng T, Zhao J, Zhao F. Grid Integration of Distributed Wind Generation: Hybrid Markovian and Interval Unit Commitment. IEEE Trans Smart Grid 2015;6(6):1.
[103] Wu L, Shahidehpour M, Li Z. Comparison of Scenario-Based and Interval Optimization Approaches to Stochastic SCUC. IEEE Trans Power Syst 2012;27(2):913–21.
[104] Wan C, Xu Z, Pinson P, Dong ZY, Wong KP. Probabilistic Forecasting of Wind Power Generation Using Extreme Learning Machine. IEEE Trans Power Syst 2014;29(3):1033–44.
[105] Wan C, Xu Z, Pinson P. Direct Interval Forecasting of Wind Power. IEEE Trans Power Syst 2013;28(4):4877–8.
[106] Wan C, Xu Z, Pinson P, Dong ZY, Wong KP. Optimal Prediction Intervals of Wind Power Generation. IEEE Trans Power Syst 2014;29(3):1166–74.
[107] Taylor JW, Mcsharry PE, Buizza R. Wind Power Density Forecasting Using Ensemble Predictions and Time Series Models. IEEE Trans Energy Convers 2009;24(3).
[108] Cui M, Feng C, Wang Z, Zhang J, Wang Q, Florita A, et al. Probabilistic wind power ramp forecasting based on a scenario generation method. In: IEEE Power & Energy Society General Meeting. IEEE; 2017. p. 1.
[109] Arabpour A, Besmi MR, Maghouli P. Transmission expansion and reactive power planning considering wind energy investment using a linearized AC model. J Electr Eng Technol 2019;14(3):1035–43.
[110] Hashemi S, Ostergaard J, Yang G. A Scenario-Based Approach for Energy Storage Capacity Determination in LV Grids With High PV Penetration. IEEE Trans Smart Grid 2014;5(3):1514–22.
[111] Saber H, Moeini-Aghtaie M, Ehsan M, Fotuhi-Firuzabad M. A scenario-based planning framework for energy storage systems with the main goal of mitigating wind curtailment issue. Int J Electr Power Energy Syst 2019;104:414–22.
[112] Zeynali S, Rostami N, Ahmadian A, Elkamel A. Two-stage stochastic home energy management strategy considering electric vehicle and battery energy storage system: An ANN-based scenario generation methodology. Sustain Energy Technol Assess 2020;39:100722.
[113] Freitas PFS, Macedo LH. Romero RA strategy for transmission network expansion planning considering multiple generation scenarios. Electr Power Syst Res 2019;172:22–31.
[114] Sun M, Teng F, Konstantelos I, Strbac G. An objective-based scenario selection method for transmission network expansion planning with multivariate stochasticity in load and renewable energy sources. Energy 2018;145:871–85.
[115] Chen C, Sun H, Shen X, Guo Y, Guo Q, Xia, Tian. Two-stage robust planning-operation co-optimization of energy hub considering precise energy storage economic model. Appl Energy 2019;252:113372.
[116] Qiao X, Zou Y, Li Y, Chen Y, Liu F, Jiang L, et al. Impact of uncertainty and correlation on operation of micro-integrated energy system. Int J Electr Power Energy Syst 2019;112:262–71.

[117] Li Y, Zou Y, Tan Y, Cao Y. Optimal Stochastic Operation of Integrated Low-Carbon Electric Power, Natural Gas, and Heat Delivery System. IEEE Trans Sustain Energy 2018;9(1):273–83.
[118] Lei Y, Wang D, Jia H, Chen J, Li J, Song L, 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.
[119] Fu X, Guo Q, Sun H, Pan Z, Xiong W, Wang L. Typical scenario set generation algorithm for an integrated energy system based on the Wasserstein distance metric. Energy 2017;135:153–70.
[120] Shi T, Huang RM, Ding CB. Research on the Optimal Configuration of Regional Integrated Energy System Based on Production Simulation. Processes 2020;8(8):892.
[121] Yu H, Rosehart B. Probabilistic power flow considering wind speed correlation of wind farms. In: 17th Power Systems Computation Conf. Stockholm, Sweden, 1–7 August; 2011.
[122] Rocha LCS, Junior PR, Paiva AP, Oliveira PE, Aquila G, Balestrassi PP. A stochastic economic viability analysis of residential wind power generation in Brazil. Renew Sustain Energy Rev. 2018;90:412–9.
[123] Aquila G, Rocha LCS, Junior PR, Oliveira PE, Queiroz AR, Paiva AP. Wind power generation: An impact analysis of incentive strategies for cleaner energy provision in Brazil. J Cleaner Prod 2016;137:1100–8.
[124] Shu Z, Jirutitijaroen P. Latin Hypercube Sampling Techniques for Power Systems Reliability Analysis With Renewable Energy Sources. IEEE Trans Power Syst 2011;26(4):2066–73.
[125] Billinton R, Wangdee W. Reliability-Based Transmission Reinforcement Planning Associated With Large-Scale Wind Farms. IEEE Trans Power Syst 2007;22(1).