Fu *et al.*
Protection and Control of Modern Power Systems (2022) 7:41
https://doi.org/10.1186/s41601-022-00262-x

Protection and Control of
Modern Power Systems

ORIGINAL RESEARCH

Open Access

Planning of distributed renewable energy systems under uncertainty based on statistical machine learning

Xueqian Fu[1,2*], Xianping Wu[1], Chunyu Zhang[1], Shaoqian Fan[1] and Nian Liu[3]

Abstract

The development of distributed renewable energy, such as photovoltaic power and wind power generation, makes the energy system cleaner, and is of great significance in reducing carbon emissions. However, weather can affect distributed renewable energy power generation, and the uncertainty of output brings challenges to uncertainty planning for distributed renewable energy. Energy systems with high penetration of distributed renewable energy involve the high-dimensional, nonlinear dynamics of large-scale complex systems, and the optimal solution of the uncertainty model is a difficult problem. From the perspective of statistical machine learning, the theory of planning of distributed renewable energy systems under uncertainty is reviewed and some key technologies are put forward for applying advanced artificial intelligence to distributed renewable power uncertainty planning.

Keywords: Distributed renewable energy systems, Statistical machine learning, Uncertainty planning, Renewable energy network

1 Introduction

In the context of the rapid development of renewable energy power generation, photovoltaic (PV) and wind power (WP) outputs fluctuate greatly and have strong randomness. This bring a series of problems in controlling, scheduling and planning. To address the randomness of distributed renewable energy (DRE), artificial intelligence (AI) has moved to the mainstream of renewable energy forecasting and prediction [1], while energy storage can also address the uncertainty of renewable energy systems [2]. Frequency control is one of the keys to WP integration [3], while multi-objective optimization is always a problem in a power grid with WP generation [4].

This paper is concerned with the influence of uncertain DRE output on traditional renewable energy network (REN) planning. Developing planning methods suitable for large-scale grid-connected renewable energy has important engineering significance. However, the main theoretical problem of uncertainty planning is how to deal with uncertainty as the uncertainty of the source side leads to complex and changeable operating scenarios of REN. This brings challenges to the planning of DRE. In the literature, there are three major planning methods: (a) deterministic, (b) robust, and (c) probabilistic planning.

In terms of deterministic planning, daily typical load curves in spring, summer, autumn and winter play important roles in power grid planning. The operating scenarios are divided into typical-day scenarios and extreme scenarios [5], and multi-stage planning is carried out with the scenarios. In [6], multi-stage planning is divided into different time stages, and each stage includes different scenarios. A multi-objective planning model is presented in [7], which analyzes uncertainty using scenarios, and deterministic power flow (DPF) is carried out for each scenario. In essence, the above studies are

*Correspondence: fuxueqian@cau.edu.cn
[1] College of Information and Electrical Engineering, China Agricultural University, 100083 Beijing, China
Full list of author information is available at the end of the article

scenario planning methods based on knowledge-driven schemes, and the effectiveness depends on whether the selected typical load curves can represent the load profile over the whole planning period. The fewer the number of operating scenarios, the more accurate the deterministic method can be. As new energy sources and demand response loads are widely integrated into REN, the uncertain disturbances on both power supply and load are becoming stronger, making the operation of energy systems increasingly complicated. The traditional typical-day method of spring, summer, autumn and winter is no longer suitable as it cannot cope with the complexity of the operating scenarios in REN. Since the effective refined operation scenarios cannot be passed to the planning decision-makers, it is difficult to ensure the accuracy and safety of REN uncertainty planning. Reference [8] proposes a distribution-free scenario generation method based on generative adversarial networks (GAN), which can be deliberately modified according to statistical characteristics for power system planning and operation. The conditional variational automatic encoder method is used to simulate the refined random scenarios of renewable energy, while the simulation of complex operation scenarios based on data-driven theory has become the core of energy network uncertainty planning [9].

In terms of robust planning, interval power flow and interval optimization are the keys to its success. The idea of interval flow is to get the solution scheme of the interval range of the state variables based on mathematical planning with uncertain injection power as a feasible region [10]. To avoid nonlinear planning, a linear approximation is a feasible scheme to realize interval optimization [11]. It has benn proved that the interior point method can solve the nonlinear planning models [12]. In essence, the above studies use a planning method based on extreme condition scenarios, and the effectiveness depends on whether the selected extreme condition scenarios can represent the boundaries of the state variables over the whole planning period. Robust planning can be used for dispatching, and uncertainty planning is suitable for power grid planning. In addition, robust planning can ensure operational safety, while uncertainty planning can balance the needs between economy and safety.

In terms of probabilistic planning, it is well known that the classical way of solving the uncertainty problem is to use probability, and probabilistic power flow (PPF) may be one of the key issues. Both the Gaussian process emulator and Bayesian inference are the basic theories of PPF calculation [13]. When planning the best placement of electric vehicle (EV) charging stations and wind turbines, uncertainty planning should be involved in the planning framework [14]. An uncertainty planning model considering different operational constraints in multi-periods is presented in [15], and the planning model solves uncertain dual dynamic integer planning. Few studies exist on how to apply PPF to uncertainty planning, while researchers focus on how to enhance the calculation accuracy and efficiency of PPF. As PPF is largely absent in probabilistic planning, the framework combining PPF and probabilistic planning is a problem worthy of discussion.

In terms of probabilistic planning, it is easy to estimate the probability distribution function (PDF) of renewable energy generation when there is randomness, so it is suitable to use the probability theory random method [16]. However, it is difficult to determine the probability distribution and membership relationship of variables in many cases. Therefore, the interval method is more suitable for dealing with uncertainty. Compared with probability theory, the interval method used in robust programming finds it easier to describe various uncertain factors. Because it requires less data that are the upper and lower boundaries. The disadvantage is that interval expansion will occur [17]. The wider the uncertainty set is, the more conservative the optimal solution is in the economy.

University of California Berkeley considers statistical learning, also known as statistical machine learning (SML), to involve probability theory, statistics, approximation theory, convex analysis, algorithm complexity theory, machine learning and other disciplines [18]. The development of DRE planning theory has been accompanied by the development of SML algorithms in recent decades, showing a strong trend of development [19]. However, there is no systematic method for SML to deal with uncertainty planning.

REN with a high proportion of DRE generation and EV loads is a large-scale dynamic system with high dimension, is nonlinear, and has uncertain and complex characteristics. Most traditional scenario analysis methods are based on probabilistic models [9], and the model capacity is small. Digital features can only capture local data features, which cannot fully characterize the complex high-dimensional and large data features of renewable energy output. Thus, it is insufficient for solving the new energy uncertainty planning problem. Probability theory is model-oriented and relies on rigorous mathematical derivation, emphasizing model interpretation. However, it has limitations in the face of complex high-dimensional data. Machine learning (ML) is algorithm-oriented and pays attention to the prediction results, and has good controllability and scalability. Deep learning (DL) is a branch of ML and is based on a neural network, and can effectively solve high-dimensional problems with a strong ability for autonomous learning, pattern recognition and optimization. However, ML has poor interpretability. A general knowledge of probabilistic theory combined with ML, SML emphasizes the intelligibility, precision and uncertainty of models, and this is consistent with the engineering requirements of uncertain DRE planning. New energy planning needs to provide decision-makers with explanations, and SML calculation results can be understood and trusted [20, 21] (achieving ML understanding-to-human understanding).

The contributions of this paper can be summarized as follows.

(1) Uncertainty analysis of renewable energy is conducted via statistical machine learning.
(2) Novel techniques in uncertainty modeling are illustrated by principle demonstration.
(3) Renewable energy allocation algorithms are analyzed in an uncertain environment.
(4) Future development trend of uncertainty theory is analyzed for renewable energy integration.

2 Literature review

China actively supports regional integrated energy development at the policy level and has promulgated a series of policies to promote the development of power energy storage, the energy Internet and other technologies to push the development of smart energy management. With the current research background, renewable energy generation, integrated energy and EV have become important factors in regional planning. At present, many actual REN projects have been built in China. For example, the integrated energy system (IES) of Qiantang New District Power Supply Company, shown in Fig. 1, has built distributed clean energy services and EV services. Its power generation mainly covers PV and WP, and the load side mainly includes an intelligent energy storage system. Thus, the described research object directly targets practical systems.

Uncertainty in source-network-load leads to complex and changeable operating scenarios of REN, which makes the planning problem uncertain. In this paper, the probability model is used to describe uncertainty. Based

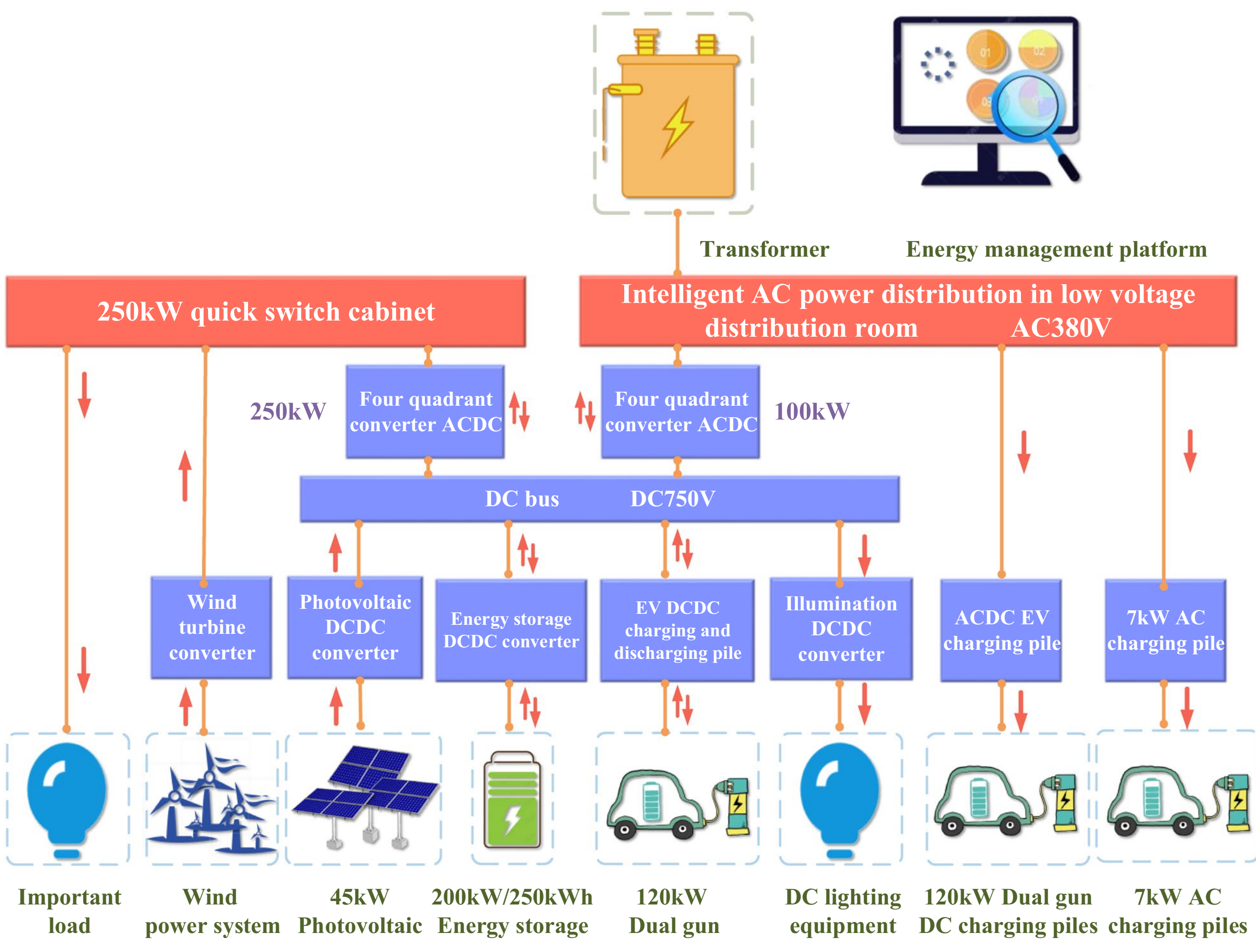

Fig. 1 An actual project with distributed renewable energy in China

on the SML tool, the objective function and constraint conditions of the probabilistic planning model are established. Section 2.1 introduces and evaluates the common methods of distributed renewable energy and electric vehicle uncertainty modeling, including probabilistic and ML modeling. In Sect. 2.2, the characteristics of uncertainty planning methods of distributed renewable energy in REN and IES are analyzed, including models and solving algorithms. The planning model includes chance-constrained, two-stage, multi-stage, distributed robust planning, multi-objective and bi-level models. The solving algorithm is composed of the traditional, intelligent and hybrid algorithms. Probability theory and ML theory belong in the SML category.

2.1 Uncertainty modeling

Figure 2 depicts a high proportion of intermittent renewable energy, load side EV grid connection and air conditioning regulation in modern REN. On the source side, PV and WP generation are the typical uncertain power sources, which are affected by changing weather conditions. On the load side, the uncertainty of EV charging time is an important source of uncertainty. In addition to renewable energy and EV loads, there are a large number of other uncertainties in modern REN, such as air conditioning loads, and the uncertainties of source and load spread through transmission lines, leading to power flow uncertainty. Therefore, source-network-load uncertainty makes power grid operation conditions complex and changeable, and the power flow presents strong uncertainty, which needs to be described by probability.

2.1.1 Distributed new energy modeling

Figure 3 shows the actual research project named "panoramic viewable", a renewable energy control sub-station of Qilinshan wind farm in Shangyi, Hebei Province. As seen, Fig. 3a shows the operational status of all PV inverters, and Fig. 3b, c show the voltage distribution map. In Fig. 3c, the red indicates a high voltage area and the blue indicates a low voltage area. Figure 3d and e show the active power distribution, while the three-dimensional pipeline is used to represent the active power and grid loss of all feeders. The thicker pipelines with brighter color indicate that the active power flow and active power loss are greater. It can be seen that the system voltage and power flow are greatly affected by DRE in the projects. Therefore, the problem studied in this paper has great practical significance.

With DRE generation applied widely, the output characteristics of renewable energy shown in Fig. 4 fluctuate greatly and have strong randomness, which together lead to the uncertainty and complexity of energy system operation. IES planning characteristics have undergone fundamental changes, such as diversification and

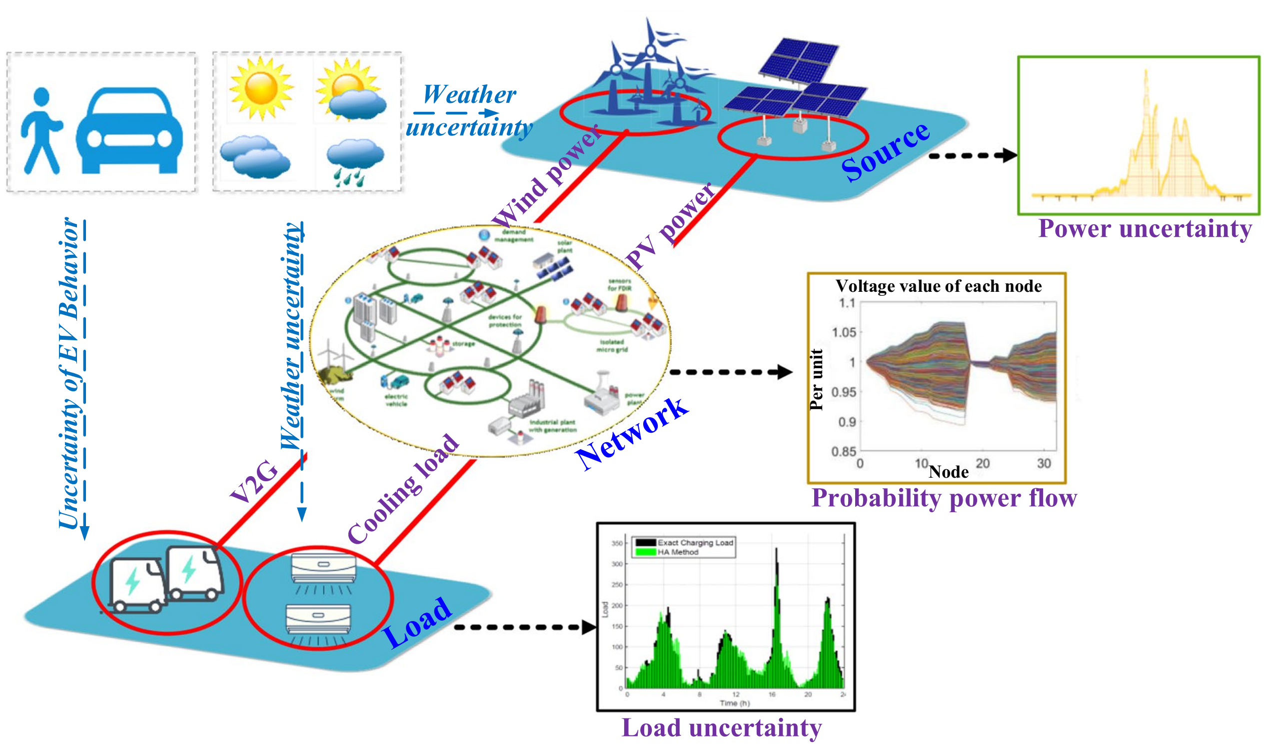

Fig. 2 Source-network-load uncertainty

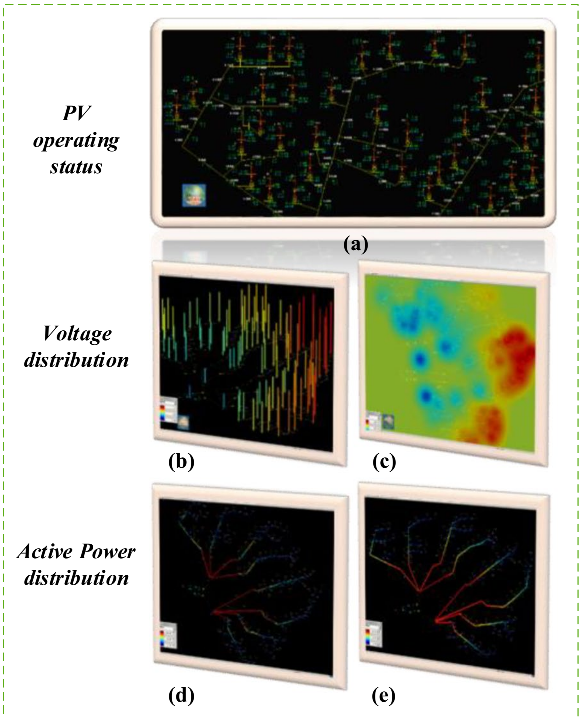

Fig. 3 The great impact of PV on voltage and power flow

differentiation. Therefore, it is necessary to carry out renewable energy modeling accurately.

In scenario generation, random distribution is introduced to simulate the probability distribution of renewable energy. By sampling the continuous probability distribution of random variables, discrete sample sets and probability values are obtained, thereby turning the uncertainty problem into a certain problem. Reference [22] generates a set of dynamic-wind speed scenarios that consider wind speed fluctuations to describe the WP uncertainty by using a multivariate normal distribution and inverse transform sampling, while a scenario generation method sampled from non-Gaussian and interdependent conditional distributions is presented in [23]. The quantile prediction method of solar and WP generators is applied in [24]. In addition, a large amount of empirical research has shown that the probability distribution of scenario prediction errors doesn't follow any parameter density. The maximum entropy theory is introduced to analyze the uncertainty of WP output and EV loads, and a refined seasonal analysis of the complex operating scenarios of REN is conducted in [25]. In [26], a linear planning optimization model directly generates quantiles with different proportions, effectively generating non-parametric probability forecasts for WP generation, whereas a limited Boltzmann machine is used to capture the time characteristics of wind speed in [27]. This uses divergence and Gibbs sampling to fit the

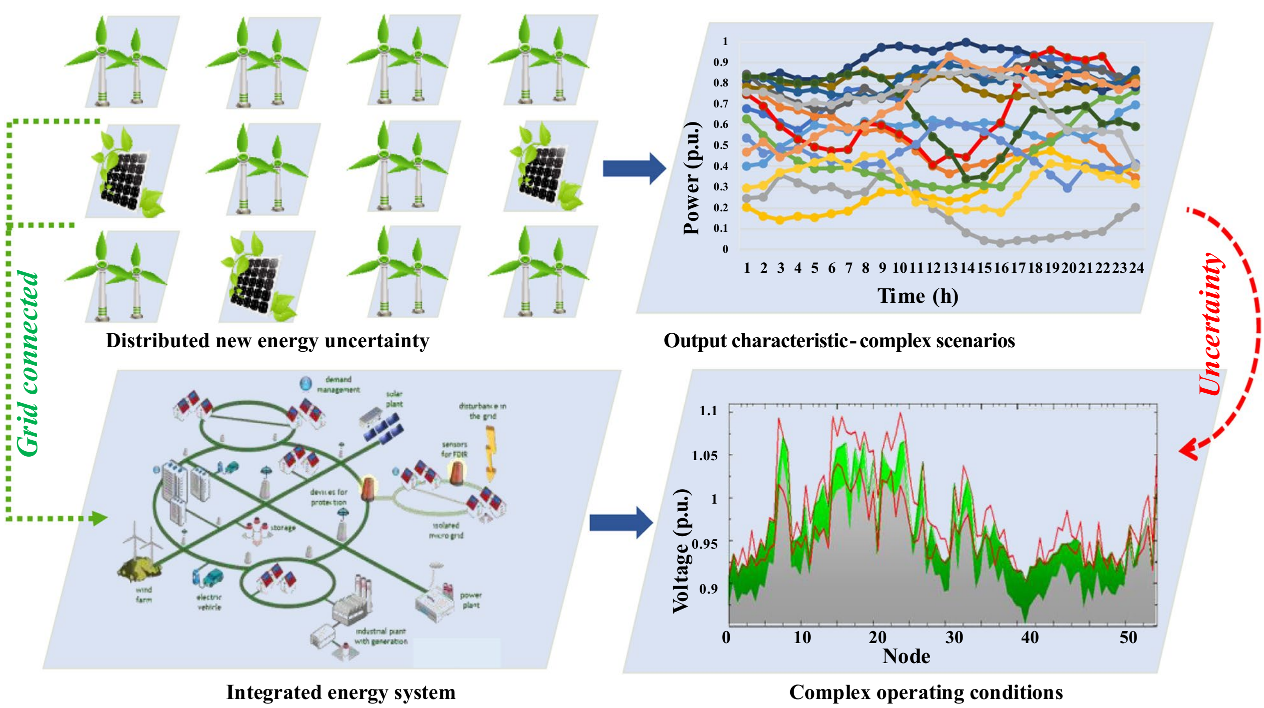

Fig. 4 Distributed renewable energy uncertainty modeling

probability distribution of WP scenarios. Monte Carlo sampling (MCS) is combined with quantile estimation technology to estimate the dimensions of the hybrid renewable energy systems [28], while in [29], MCS is presented to represent the uncertainty of PV generation. In many cases, Latin hypercube sampling (LHS) is considered to be an improvement on rough MCS. In [30], LHS is used to improve sampling efficiency and deal with the relevance problem of DRE generation access nodes, while the methods of MCS and LHS are analyzed to generate WP scenarios in [31].

At present, many scholars have verified that the copula function in statistical science is effective in the correlation modeling between random variables. In [32], scenario generation is generated based on the multivariate estimated distribution of WP in multiple regions. Pair copula and vine copula were first introduced to the WP scenario generation to discuss their temporal coupling [33]. The presented model can easily obtain accurate and sufficient WP scenarios with time and space correlation. In [34], a combination of nuclear density estimation and regular vine copula is also used to describe the spatial correlation between WP plants, whereas in [35], Brownian motion is combined with copula to generate correlated dynamic PV scenarios. In [36], a dynamic factor model is presented to generate scenarios similar to the statistical characteristics of the actual scenarios. For some DRE power plants that do not have enough historical data, multiple scenarios and a knowledge-based scenario generation algorithm are presented, which capture the potential distribution of target wind speed with high precision [37].

The time series method is another emerging scenario generation method. The common methods used have been the auto-regressive moving average model and Markov random process. In [38], random load and generated renewable energy scenarios are considered on multiple operating time scales. A double-layer WP time series model is presented that considers daily weather changes and intraday WP fluctuations [39]. The Markov chain and time series are used to simulate the daily typical weather and intraday WP scenarios of each typical weather state, and then are integrated into a complete WP time series. An inhomogeneous Markov chain is also presented to depict the characteristics of diurnal and seasonal changes in wind speed [40]. Both models presented in the above two references have good ergodicity and consistency. In [41], an auto-regressive moving average (ARMA) model is presented to grasp renewable energy spatiotemporal features. Based on this, [42] improves ARMA that can grasp the time correlation of WP scenarios. Generalized auto-regressive scoring models are presented to generate renewable energy multivariate time series scenarios in [43], while [44] presents state space models to generate renewable energy WP scenarios.

Recently, the concept of machine learning has gradually become the direction pursued by the random scenario generation of the IES. In [45], the radial basis function (RBF) and tabu search-based metaheuristic algorithm are combined to effectively capture the important features of the WP scenarios, while artificial neural networks are used to generate representative multi-dimensional DRE scenarios based on historical time series values [46]. Reference [47] adopts stacked autoencoders to improve the precision of WP scenario generation, whereas for reducing the prediction error in scenario generation, the concept of a generative confrontation network is presented in [48]. This was first used to generate continuous sequences in the field of artificial intelligence. In [49], an unsupervised scenario generation method of GANs is presented to capture the dynamic characteristics of DRE generation. However, the method has some problems such as vanishing gradients, difficulty in training, and lack of diversity in generated scenario samples. Therefore, Wasserstein distance is used instead of JS distance on the basis of GANs, and Wasserstein-GAN is used to train DRE scenarios. This effectively solves the above problems [50]. The scenario generation ability of GANs is further improved by adding Lipschitz constraints on the discriminator network [51], whereas in [52], a Bayesian formula is introduced into a genetic neural network to realize semi-supervised learning. Reference [53] applies this method to DRE scenario generation. Even intentionally mixing PV and WP scenarios, this method can identify and generate different scenarios at the same time, so as to better represent the DRE generation process. Sequence generation is used for the network to capture temporal correlations. It adopts a long and short-term memory structure and uses GAN combined with reinforcement learning [54]. In the work of [55], GAN based on two interconnected deep neural networks is used to generate real WP and PV power distributions with completely diverse behaviors. It also explains how to use labeled data in the training process to generate the required scenarios based on different interest conditions.

Research on accurate renewable energy uncertainty modeling has been carried out to a considerable extent, while scenario analysis methods provide an effective way to deal with DRE uncertainty modeling. Table 1 shows the characteristics of scenario analysis methods based on probability models, time series methods and AI algorithms.

2.1.2 Electric vehicle load modeling

More active devices and adjustable loads will be introduced into future intelligent energy systems, and the

Table 1 The induction and evaluation of new energy uncertainty modeling methods

Method	References	Characteristic	Evaluation
Parameter probability prediction	[22–24]	The premise is that the known data conform to a certain probability model; hypothesis testing is required	The assumptions of the probability model and the determination of the relevant parameters affect the precision of the generated scenarios.
Nonparametric probability prediction	[26, 27, 56]	Don't need global parameter assumption; simple calculation and wide applicability	Unable to process massive sample data. And some data suitable for parameter estimation may lose some scenario characteristics.
MCS sampling	[28, 29, 57]	Close to the actual sample	The accuracy is low when the sample is small, and the efficiency is low when the sample is large.
LHS sampling	[30, 31]	Suitable for uniform sampling in multi-dimensional space and small samples; high sampling efficiency	The correlation coefficient matrix (CCM) of random variables is required as input parameters, but CCM is difficult to extract.
Copula and its improvement	[32–34, 36, 37, 58]	The correlation characteristics of weather sensitive factors are captured	It is difficult to process high-dimensional and complex scenario samples.
Auto regressive moving Average	[38, 40, 59]	Use the mean variance normalization method to preprocess the data, which is simple	Prone to over-fitting and pattern recognition errors; insufficient data diversity.
Markov stochastic process	[41, 44]	The daily cycle and seasonal change characteristics of scenarios are considered; the scope of application is wide	Due to the lack of memory, only short-term autocorrelation characteristics are retained.
Radial basis function	[45]	The results of fitting the actual scenarios are of high accuracy	When the sample size is large, there are many hidden layers and the network structure is complex, resulting in low computational efficiency.
Artificial neural networks	[46]	The training process is stable and the convergence rate is fast	The neural network design is complex; the interpretation ability is weak; machine learning feature selection will have a great impact on the fitting effect.
GANs and its improvements	[48, 55, 60, 61, 62, 63]	High accuracy in processing high-dimensional samples; no need to manually label data; capture correlation features	GANs have the problem of mode collapse and weak interpretability.
Variational automatic encoder	[47]	The mathematical proof is clear; the interpretability is strong; the long-term and short-term characteristics can be captured	It is almost impossible to characterize the correlation of scenario historical data.

EV charging loads are the most typical uncertain loads. The load-side uncertainty caused by the characteristics of EV and the grouping characteristics of drivers' social networks, increases the operating pressure on the power grid. As shown in Fig. 5, grid connection of EV brings uncertainty to REN, with not only the uncertain grid connection time but also the random grid node. Among them, electric taxis are more uncertain and more difficult to predict than private cars and buses. Moreover, during the peak charging period of EV, the charging loads increase sharply, and thus the local peak loads increase significantly. During the low-peak charging period, many EVs discharge power in reverse to the grid, which is filled with excess power load. The power in both periods may be higher than the equipment capacity of REN under traditional planning, resulting in local blocking and paralysis of the power grid in severe cases.

The EV charging load uncertainty models are based on the premise of realizing the charging stations' scientific planning in REN, and probability statistics are one of the key methods. Probability is the mathematics of uncertainty. It uses the actual historical data to simulate the PDF of the initial EV charging time and state of charge (SOC), and then models the charging loads by random sampling methods. In [64], MCS is used to fit the charging load distribution of each EV. The model assumes that the EV relevant parameters obey the above distribution, e.g., leaving home in the morning to stop charging and returning home at night to start charging. Reference [65] takes load samples by MCS and estimates the total average weight by a discrete probability formula, while [66] first selects whether the EV is V2G through roulette, then generates random numbers, and finally uses MCS to produce massive random scenarios. Reference [67] discretizes EV charging duration and charging-start time, and obtains joint uncertainty by Cartesian product. In addition to the commonly used probability distribution function, the spherical simplex unscented transformation (UT) is used to approximate the probability distribution in [68]. The point set σ can accurately represent the statistical information required by random variables. In [69], the charging pull-out time of the electric taxi is affected by the initial charging time and duration. Considering the disadvantages of MCS with more samples and long calculation time, LHS is presented. LHS is used to simulate the EV charging loads, and the EV relevant parameters obey the general probability distribution [70]. In [71], a combination of LHS and K-means is used to model the uncertain scenarios of EV charging loads, and the EV relevant parameters obey a truncated Gaussian distribution.

The above references, except [67], assume that the EV parameters are independent of each other, and the

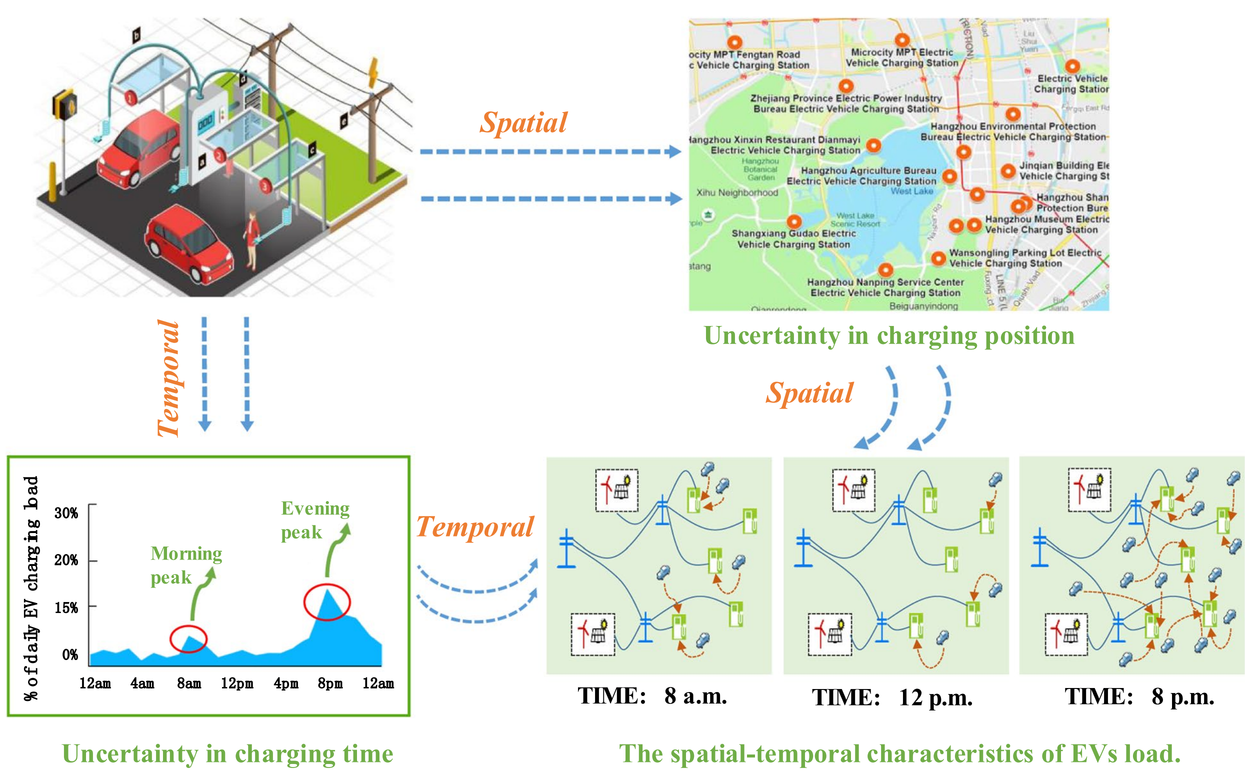

Fig. 5 Spatial and temporal uncertainty of EV load

relationships between parameters are not considered. The copula is a function to describe the dependency between variables, and can create a multivariate distribution of data. Multivariate t-copula [72–74], Gaussian copula [75] and Archimedes copula [76] are used to describe the dependence between the EV starting time, arrival time and vehicle distance, to model the driving mode more accurately. In addition, a 3-dim kernel density estimation (KDE) is considered to model the relationship between different variables [74]. In [77], the 3-dim KDE is also used to describe the uncertainty of the EV charging mode. Traveling is a kind of activity arrangement in time and space, which connects two or more places, for home or non-home travel purposes. First, the time and space probabilities are obtained respectively, and then the total probability is calculated to obtain the travel probability model of the EV. In [78–81], some prediction models of EV charging behavior are presented based on the travel chain. These methods can simulate the travel behavior of users after constructing the travel chain according to the different purposes of driving, so as to analyze the driving law of users more accurately. In the above probability models, the EV relevant parameters are fitted to synthesize the relevant probability distribution from the actual data. However, the accuracy of the fitting results remains to be discussed.

Considering the temporal and spatial uncertainty of EV charging loads, advanced machine learning has been used for their uncertainty modeling. Reference [82] uses a non-homogeneous Markov chain to simulate and generate the EV usage pattern, and a fast-forward selection method to reduce scenarios. The EV state is divided into normal charging, fast charging, driving and parking in [83], while in public charging stations, the continuous-time Markov chain method is used to depict the uncertainty of EV charging behaviors in [84]. The uncertainty of passengers, charging stations, and public transport usually leads to the uncertainty of electric taxi loads. Therefore, the operational behavior model of electric plug-in taxis (PET) is presented based on environmental uncertainty [85]. In [86], regarding the movements of electric taxis as random walks, the Markov process is used to simulate the distribution of charging demand in static space, while in [87], the random forest (RF) method based on a regression tree is used to predict the driving characteristics of each EV, so as to obtain the travel mode data set of the EV. From the measured charging information and big data mining technology, the EV charging behavior model is presented based on RF and principal component analysis (PCA), which captures the EV with different charging characteristics based on a data-driven model [88]. The gradient boosting model (GBM) is used to simulate the SOC state at the end of a plug-in EV (PEV) daily schedule [89], and in [90], a generalized regression neural network (GRNN) is used to realize power prediction. This can effectively describe the uncertainty on EV charging loads. DL has also received extensive attention in solving uncertainty problems. A back propagation (BP) neural network is presented to forecast EV charging loads in [91], while a convolutional neural network (CNN) is used to reduce the impact of uncertainty on EV charging demand accuracy in [92]. In [93], a method combining a CNN and a deep belief network (DBN) is presented to describe the uncertainty of EV charging demand. With the wide application of computer technology and social networks, multi-agent technology has been developed. Based on multi-agent technology, a spatiotemporal EV charging demand simulation model considering passenger travel demand is constructed [94], where the travel demand of passengers, taxi decision-making process and taxi queuing process are managed by a multi-agent. Reference [95] uses multi-agent and multi-step Q-learning methods to describe the uncertainty of electric taxi charging loads on time and space scales. As the above data-driven methods only need to collect real data without the need to obey various distributions, the correlation between the starting time, arrival time, vehicle distance and other parameters of EV can be retained to ensure accuracy.

As shown in Table 2, uncertainty modeling methods for EV charging loads can be divided into two types: probability statistics theory and machine learning. There are six methods in probability statistics: UT, MCS, LHS, copula, 3-dim KDE and the travel chain model. These methods approximate PDF through the actual historical data of EV parameters, and thus the planning results are greatly influenced by the accuracy of the PDF. Moreover, the relationship between the EV parameters would be ignored if only MCS or LHS were used. Therefore, MCS or LHS needs to be used together with methods that describe correlation, such as copula and 3-dim KDE. Machine learning methods mainly include the Markov chain model, RBF, GBM, RF, DL and multi-agent technology. These methods consider the temporal and EV spatial characteristics, and do not require the parameters of EVs to obey various distributions. However, the accuracy of the results depends on the adjustment of parameters and the extraction of features. In addition, from the perspective of the types of EVs, electric taxis have characteristics such as weak regularity of driving routes and diverse selection of charging stations, which are more uncertain than traditional private electric cars. Thus, the uncertainty modeling of electric taxis is more difficult.

Table 2 The classification of EV loads uncertainty modeling methods

References	Types of EVs	Modeling methods											
		Probability statistics						Machine learning					
		MCS	LHS	UT	Copula	3-Dim KDE	Travel chain	Markov chain	RBF	GBM	RF	DL	Multi-agent technology
[64–67]	Private	√											
[68]	Private			√									
[69]	Taxi	√											
[71]	Private		√										
[70]	Taxi		√										
[72–76]	Private				√								
[77]	Private					√							
[78–81]	Private						√						
[82–84]	Private							√					
[85, 86]	Taxi							√					
[87, 88]	Private										√		
[89]	Private									√			
[90]	Private								√				
[91–93]	Private											√	
[94, 95]	Taxi												√

2.1.3 Summary

The uncertainty of renewable energy power generation is very significant, and is the main source of uncertainty in REN planning. EV policies are being promoted all over the world, and it has become a trend for EVs to replace fuel vehicles. Consequently, EV loads have become the leading factor of load uncertainty. With the background of low-carbon goals, renewable energy and EV loads have the most prominent impact on power systems. Uncertainty modeling theory has been analyzed in detail for renewable and EV loads in the previous sections. On the load side, air conditioning load is another important uncertainty factor. Because of the variations in weather and electricity price, the air conditioning load is always in a disordered operation state. The air conditioning load is sensitive to outdoor temperature and is related to residents' needs and habits. When the outdoor temperature changes, the air conditioning load will also change, and thus, the variation in outdoor temperature leads to an uncertainty of air conditioning load. Moreover, the behavior of residents is random and disordered, as the indoor temperature set point can randomly vary from 16 to 30 °C, so the air conditioning load is uncertain even in the same weather conditions [96]. The demand side response under different electricity price policies will increase the difficulty of analyzing the randomness of residents' behavior. In addition, network parameter uncertainty is also a source of uncertainty in REN, such as the parameter uncertainty of resistance, reactance or capacitance [97].

2.2 Uncertainty planning modeling

As shown in Fig. 6, the uncertainty of DRE makes IES operating scenarios complex and changeable, and the calculation burden of planning solutions caused by uncertainty increases greatly. The number of power flow calculations is the number of solutions traversed in the solution process and the number of operating scenarios. This multiplication makes the number of calculations increase explosively, while more decision variables and stronger uncertainty make the complexity of the solution process greater, in company with the increased grid connection capacity of renewable energy.

2.2.1 Distributed renewable energy planning in REN

Compared with traditional power supply, DRE power supplies have lower pollution emission, higher power supply efficiency, and are more flexible and convenient. On the basis of the known load prediction results and REN operating conditions, determining the installation location and configuration capacity of DRE supply can optimize the economy and reliability of REN in the whole planning period, thus improving power quality, economy and reliability.

To get the optimal configuration scheme of DRE, plant propagation and water cycle algorithms are presented in [98], while the power factors of DRE generators are

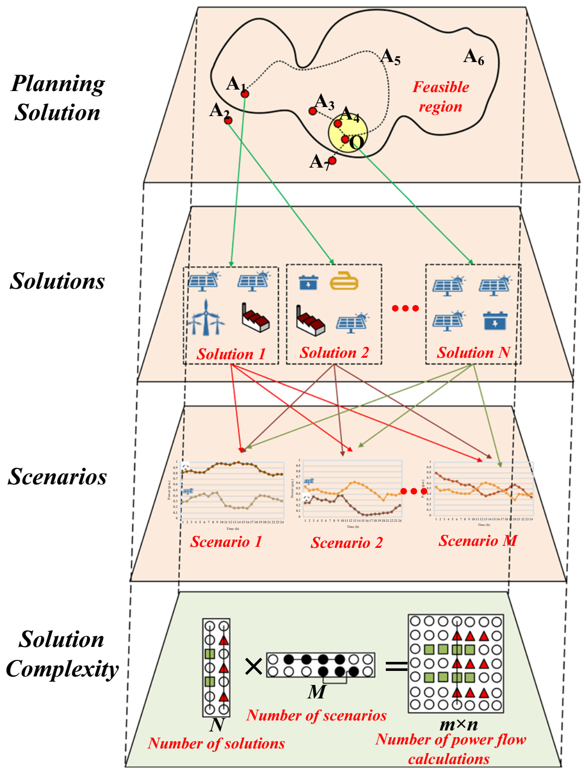

Fig. 6 The complexity and difficulty of uncertainty planning

also optimized to reduce power losses. A methodology is presented to find the optimal size, location and technology of distributed new energy generator units considering economic, technical, and environmental issues simultaneously, using multi-objective particle swarm optimization and fuzzy decision-making techniques to find the optimal solution [99]. In [100], a fuzzy multi criteria decision making approach is used to sort the load points and locations/feeder sections, and particle swarm optimization (PSO) is used to evaluate the optimal size and position of distributed new energy generators. To get the best configuration of DRE, reference [101] overcomes the shortcomings of the previous algorithms and presents the multi-leader PSO solution algorithm, while a distributional robust optimization model is presented, also known as a data-driven model in [102]. In addition, whale optimization algorithms are used to calculate the optimal layout and size of a multi-objective DRE generator set, which improves voltage distribution and minimizes power losses and operating costs [103].

To solve the large-scale and complex DRE supply optimization configuration problem with multiple constraints, many optimization models have been presented from different perspectives. In [104, 105], bi-layer planning models are presented to allocate more DRE, which generate power system operating scenarios in different periods by using the multi-state models of renewable energy and loads. In [106], a bi-level planning method is also presented to achieve optimal distribution of hybrid distribution transformers. This integrates distribution and control issues. The method achieves an optimal distribution in a dynamic environment, reducing investment costs and improving voltage curves. Given the DRE power output intermittence and the load uncertainty, a bi-level planning approach is presented to achieve an integrated optimal decision at the upper level and comprehensively considers the benefits of DRE [107]. A multi-objective distributed generation planning model is established with the objective functions of minimizing both the annual total cost and the risk [108], while a risk-based multi-objective optimal allocation model is presented to optimize the placement and configuration of distributed new energy generators, to provide a reliable and cost-effective system [109]. The collaborative planning of PV and energy storage is realized in [110].

For the problem of optimal long-term distribution of DRE, uncertainty two-stage planning models are established [111, 112]. The formerly determined investment variables of DRE, before the scenarios, are known in the first stage, while the maintenance problem is solved, one that depended on scenario variables after the scenarios become known in the second stage. The latter also puts forward the solution to the problem of the scenario generation process, develops and launchs a general framework, as open-source DRE optimization of distribution. A two-stage stochastic programming model is presented that considers the uncertainty of renewable distributed generator sets, demand and energy prices in [113], while in [114], a two-stage stochastic programming model is developed to optimally determine the layout and size of PV modules in a distribution system, where the uncertainties of solar irradiance and load are modeled in finite scenarios to minimize installation and network operating costs while meeting the necessary operating constraints. This model can effectively reduce the cost of system operation. A novel multi-stage planning method is presented in [115], one which takes into account operational changeability and randomness introduced by emission price and demand growth of intermittent power generation sources through probabilistic and uncertainty methods.

For these models, various optimization algorithms are proposed. A comprehensive solution algorithm for optimal position setting and output power of DRE supply is presented, one which is suitable for all possible load changes of distributed generator sets, and uses a lightning search algorithm to find the optimal allocation of DRE generator units [116]. In [117], a multi-objective symbiont search method is presented to get the best placement of DRE, whereas a linearized load flow method is used to find the optimal allocation of DRE generator units in [118]. Reference [119] combines PSO with voltage and reactive power control of the DRE supply, and adjusts the voltage amplitude within the allowable range to get the optimal capacity and location of DRE. In [120], a chance-constrained uncertainty planning model is introduced in the optimal placement of PV, and uses the combination of PSO and SVM algorithms to quickly obtain the optimal scheme. A sustainable unit planning method is presented for a distribution system, which tries to optimize a multi-objective index and uses Tabu Search [121]. An improved AI optimization algorithm-based ecosystem and a hybrid grey wolf metaheuristic algorithm are presented in [122] to get the best configuration of DRE to minimize the power loss of REN, while an evolutionary algorithm-based approach is used to solve optimization problems that can reduce the impact of the environmental system and the total cost [123].

2.2.2 Planning of DRE in the IES

With the correlation and coupling of ever increasing renewable energy and different types of loads, the uncertainty of energy and loads in the IES becomes increasingly prominent. This poses a great challenge to the planning and regulation of the IES. An adequately coordinated operation method of IES can ensure the supply and demand balance, and further promote the safe, economic, flexible and efficient operation of the energy Internet. However, it will affect the operation of the IES.

In practical application, uncertainty planning methods of DRE and loads have been applied. A column generation and sharing algorithm is presented to solve the computational burden of large capacity and multi-stage uncertainty mixed-integer planning model in [124], while a new congestion control method is developed to enable flexible DRE to participate in solving transmission system operator congestion problems, and to use the sequential least squares method for solving the optimization problem [125]. In [126], k-means clustering is used to extract typical scenarios, and considers changes in sunshine, wind speed, and power demand to adjust the scale of the DRE. An iterative bi-layer planning model is presented to deal with the nonlinear and nonconvex state space of the systems [127], while a combination of the hybrid big bang-big crunch solution algorithm and PSO is presented to obtain higher precision in the optimization performance of a high-dimensional function [128]. In [129], distributed algorithms for multiple/single storages are used to realize the coordinated allocation of DRE without the need of a central coordinator. The modified Quasi-oppositional chaotic Selfish-herd optimization algorithm in [130], with the advantages of both chaotic linear search and quasi-opposition based learning supported, is a faster search process than normal uncertain search. In [131], PSO is improved based on the map-reduce, which reduces the particle search range of a single evolutionary algorithm with a faster calculation speed. Multi Agent-Hierarchical Task Network (HTN) adopts a heuristic search strategy based on depth-first in [132], and compared with the classic HTN algorithm, it has higher quality and efficiency, better convergence and adaptability in solving complex problems. Based on the PSO algorithm, an iterative three-stage planning model is used to get the optimal price of non-utility DRE contributions in different operation conditions [133]. To optimize the solution, the multi-objective PSO algorithm is used in [134], and proves that the comprehensive benefit of multi-objective optimization is better than that of single economic optimization system. In [135], the augmented epsilon constraint method is used to find the Pareto optimality region to analyse a complex mixed integer linear programming model. The developed multi-objective optimization method in [136] adopts the strength Pareto evolutionary algorithm 2 technique to promote and distribute the benefits of distributed new energy, and is expected to promote the update of current power industry regulatory proposals. An iterative constraint-based search method is presented to optimize the microgrid DRE configuration [137], while an improved teaching–learning optimization algorithm is used to enhance the performance of the algorithm in global search [138]. Limited by the non-analytical mapping between uncertainty planning strategies, an iterative bi-layer optimization algorithm is presented to deal with the problem in [139], while [140] uses a genetic algorithm on a mixed-integer linear planning model and gives the approximate optimal solution with accurate power flow. The harmony search algorithm and the firefly algorithm are combined in [141], and the solution results are of higher quality, have better convergence characteristics, and higher computational efficiency. An improved PSO is presented to overcome an inherent tendency [142].

Reference [143] presents a multi-temporal scale operation planning model covering the upper, middle and lower levels, where the upper model optimizes the whole distributed IES of the previous day, the middle model optimizes each community in the daytime, and the lower model sets the power consumption of every community real-time adjustment. A multi-energy micro-grid optimal design mixed-integer linear planning model is established to achieve optimal scheduling in [144], while an uncertain two-stage planning model is proposed in [145], where the first stage gets the virtual power plant configuration decision and the second stage deals with renewable energy power generator and storage. The integrated system tool is presented in [146], where a customer adoption model (CAM) optimizes dynamic model-in-the-loop, and realizes distributed energy scheduling optimization in power grid transactions. Based on the mixed-integer linear planning model, the role of DRE in the dynamic microgrid is used to improve the recovery ability of REN under extreme climate events in [147], while a mixed-integer nonlinear model of an island microgrid is established, which uses PDF to describe the uncertainty of DRE [148].

2.2.3 Summary

From different perspectives, many have used a variety of algorithms to model and solve the DRE planning problem in IES. Various classical mathematical, machine learning and intelligent algorithms have been adopted or improved to simulate and constrain a variety of uncertain factors of distributed resources, so as to solve the modeling, planning and configuration of distributed resources. Table 3 summarizes the common uncertainty planning models of renewable energy, while Table 4 compares the characteristics of the specific solution algorithms from the literature. The advantages and disadvantages of the traditional algorithms and intelligent algorithms are analyzed as follows:

(1) Common traditional solution algorithms include the interior point method [149], and algorithms such as the sequential least squares [125], iterative bi-layer optimization [127, 139], augmented epsilon constraint [135] and constraint-based iterative search algorithms [137]. Such algorithms generally require the first or second order gradient of the objective function to the control variable. The structure and parameters are fixed in the process of the solution, and the computational complexity and convergence can be proved by means of the analysis. In theory, a global optimal solution can be obtained quickly. However, these algorithms require the objective function to be derivable and continuous. This limits their application in large-scale nonlinear and uncertain planning problems.

(2) The intelligent solution algorithms are generally heuristic. The algorithms are independent of derivatives, without assumptions and approximations. They have a certain randomness and can jump out of the local optimum. In addition, their inherent parallelism helps to deal with complex planning problems in a discontinuous, non-smooth and highly nonlinear solution space, showing good robustness, convergence characteristics and optimization ability. The main disadvantage is that the optimal solution relies on continuous parameter adjustment experiments and experience, resulting in high computational complexity and slow convergence rate. In addition, finding the global optimal solution cannot be guaranteed.

3 Advanced technology

The authors have conducted a series of studies in the uncertainty planning and inquiry field. In 2015, the support vector machine (SVM), a popular machine learning theory, was used to calculate power flow approximately and quickly [120], while in 2016, an analytic statistics method was presented in [151] to estimate and reduce the power loss in REN uncertainty planning. In 2017, the combination of information theory and probability theory was studied to analyze and solve the uncertainty problem of a power grid [152], whereas in 2018, a new data-driven approach was used to analyze how uncertain power flows affect power grid reliability [153]. In the 2020s, the theories of unsupervised dimensionality reduction, probability inequality and statistical estimation have been combined to deal with the uncertainty planning problem in distribution systems [96], and it was found that the above theories belong to the theoretical system of SML.

This section presents the primary advanced SML technologies for uncertainty planning. In Sect. 3.1, a Bayesian generative adversarial network (Bayesian GAN) and semi-supervised algorithm are presented for renewable energy output scenario generation. This is the core of REN refined planning decision calculation. Section 3.2 presents information entropy theory for the uncertainty measurement of REN probabilistic power flow calculation results. This is an indispensable part of the constraint condition set in the uncertainty planning model. In Sect. 3.3, the economy of typical scenarios and the small probability of extreme scenarios for renewable energy uncertainty planning modeling are presented. The Bayesian GAN and semi-supervised algorithm in Sect. 3.1, information entropy in Sect. 3.2 and small probability calculation in Sect. 3.3 all belong to the SML technologies.

3.1 Uncertainty modeling of new energy output

3.1.1 Scenario generation

The key to random scenario simulation is to learn the probability distribution. To analyze the characteristics of various DRE sources, SML is applied to study the uncertainty research of DRE. GAN combined with the Bayesian formula can use historical data to train and generate renewable energy operating scenarios in unsupervised and semi-supervised learning methods.

Table 3 DRE uncertainty planning model characteristics

Model	References	Objective	Constraint	Characteristic
Chance constrained	[108]	Annual total costs; risk in REN	Chance constraints of power, voltage; correlations; DRE investment	The probability that the constraint holds is not less than a certain confidence level, which can be set artificially
	[120]	Power losses; voltage quality	Chance constraints of power, voltage; PV capacity and number	
Two-stage	[111, 112]	First: DRE investment costs Second: operation and maintenance costs	Power; DRE size; generation; investment; expected energy not served (EENS)	The first stage: before the realization of random variables, independent of decision variables; The second stage: the objective function is improved based on decision variables.
	[113]	First: DRE investments costs Second: Operation costs	Power; reliability; the operation of DRE unit	
	[114]	First: installation costs Second: power losses	Power; PV placement and size	
Multi-stage	[115]	Different costs per stage	Power; generation; investment; reliability; EENS; DRE penetration;	Two-stage expansion; complex calculations
Distribute robust	[102]	DRE capacity meeting the risk level in the case of the worst probability distribution	Power; DRE output; the risk level δ	The model combines uncertainty planning and robust optimization, and considers the optimization under the worst distribution.
Multi-objective	[99]	DRE costs and profit; technical issues; pollutant gas emissions	Power; critical transmitted power; DRE capacity; DRE cost and investment	More than two conflicting targets, and generally turn them into single target through weighting method to solve.
	[100]	Restore the critical loads; EENS	Power; DRE capacity and location	
	[103]	Power losses; operating costs; voltage quality	Power; DRE capacity	
	[109]	Operational risk; costs and profit	Power; DRE power; reverse power flow	
	[116]	Reliable power supply; costs and profit	Power; DRE power; load shed	
	[150]	Power losses; voltage stability	Power; DRE size and capacity	
Bi-level	[104, 105]	Upper: Annual total costs Lower: Power losses	Upper: Power; capacities of devices Lower: Power; switching number of capacitor bank; regulation ranges of devices	The upper and lower planning have some independent decision variables to optimize their respective objective functions, and they interact with each other. Generally, the lower level makes decisions on the basis of the upper level.
	[107]	Upper: Annual total costs; voltage quality Lower: Annual DRE costs and generation; power losses; voltage quality	Power; DRE capacity	

Table 4 The induction and evaluation of DRE uncertainty planning solution algorithms

Classification	Algorithm	References	Characteristic
Traditional algorithm	Sequential least squares	[125]	Minimize the sum of error squares
	Iterative bi-layer optimization algorithm	[127, 139]	Nonlinear and non-convex function and constraints
	Augmented epsilon constraint algorithm	[135]	The most used algorithm for multi-objective optimization
	Constraint-based iterative search algorithm	[137]	Based on maximum reliability and minimum cost, the optimal solution result is moderate
Intelligent algorithm	Improved PSO algorithm based on map-reduce	[131]	Reduce the particle search scope of a single evolutionary algorithm
	Multi-objective PSO algorithm	[134]	Use random selection and adaptive grid method
	Strength Pareto evolutionary algorithm 2	[136]	Use a set of chromosome number chain solutions. Higher fitness value
	Improve teaching optimization algorithm	[138]	Enhances the performance of the solution algorithm in global search
	Improved PSO algorithm	[142]	Overcome the inherent trend of local traps in particle swarm optimization
Hybrid algorithm	Column generation and sharing algorithm	[124]	Reduce the computational burden of the long-term planning uncertainty model
	Hybrid big bang-big collision algorithm	[128]	Higher precision in the optimization performance of the high-dimensional function
	Algorithm based on consensus and gradient strategy	[129]	It's proved that the distributed energy coordination problem can be modified into a convex equivalence problem
	Quasi-opposite chaos selfish herd optimization algorithm	[130]	Combine the chaotic linear search and quasi-oppositional learning to have a faster solution
	Genetic algorithm	[140]	Higher precision of optimal solution
	Harmony search algorithm and firefly algorithm combination	[141]	High quality, good convergence characteristic and less iterative process

The structure of the Bayesian GAN Is shown in Fig. 7. Bayesian formula probabilistic reasoning is introduced to set the weight parameters of the discriminator network (DN) and generator network (GN) iteratively sampled from the conditional posterior distribution. The update is as follows:

$$p\left\{\theta_d\middle|z_j^{noise},X,\theta_g\right\}\propto\prod_{i=1}^{N_d}DN(x_i;\theta_d)\times\left(\prod_{i=1}^{m}\left(1-DN\left(GN\left(z_j^{noise};\theta_g\right);\theta_d\right)\right)\right)\times p\{\theta_d|\sigma_d\}\tag{1}$$

$$p\left\{\theta_g\middle|z_j^{noise},\theta_d\right\}\propto\left(\prod_{i=1}^{N_g}\left(DN\left(GN\left(z_j^{noise};\theta_g\right);\theta_d\right)\right)\right)\times p\{\theta_g|\sigma_g\}\tag{2}$$

where x_i is the training sample, X is the new test sample, θ_d and θ_g are the weight parameters of DN and GN, respectively. σ_d and σ_g are the respective hyper parameters of the weight parameters of DN and GN, $p\{\theta_d|\sigma_d\}$ and $p\{\theta_g|\sigma_g\}$ are the respective prior distributions of the weight parameters of DN and GN, N_d and N_g are the numbers of input samples of DN and GN, respectively.

Combining the Bayesian formula for the edge processing of noise z_j^{noise}, a simple Monte Carlo method can be used to marginalize:

$$\begin{aligned}p\{\theta_g|\theta_d\}&=\int p\left\{\theta_g\middle|z_j^{noise},\theta_d\right\}dz\\&=\int p\left\{\theta_g\middle|z_j^{noise},\theta_d\right\}p\left\{z_j^{noise}|\theta_d\right\}dz\\&\approx\frac{1}{J}\sum_{j=1}^{J}p\left\{\theta_g\middle|z_j^{noise},\theta_d\right\}\end{aligned}\tag{3}$$

The same can be obtained:

$$p\{\theta_d|\theta_g\}\approx\frac{1}{J}\sum_{j=1}^{J}p\left\{\theta_d\middle|z_j^{noise},X,\theta_g\right\}\tag{4}$$

It is worth noting that when (3) and (4) are regarded as functions of noise z_j^{noise}, the distribution of $p\left\{\theta_g\middle|z_j^{noise},\theta_d\right\}$ and $p\left\{\theta_d\middle|z_j^{noise},X,\theta_g\right\}$ should be broad,

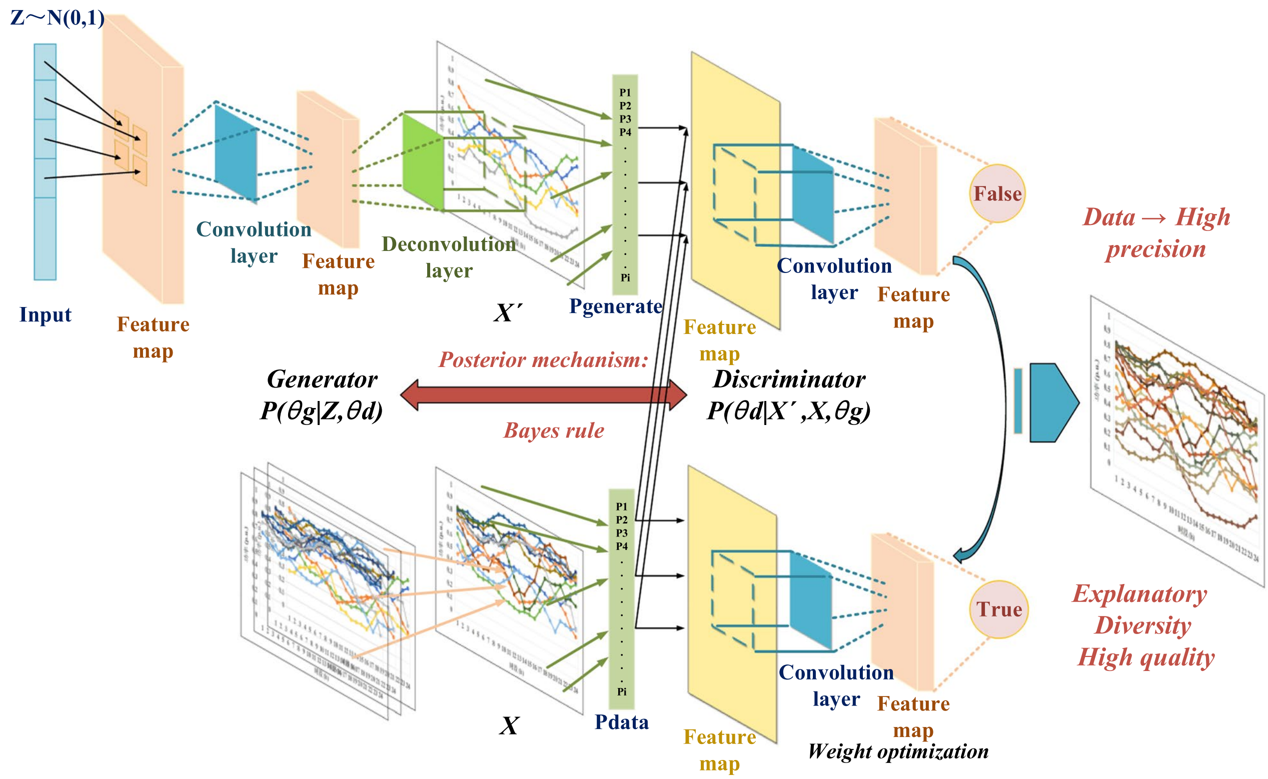

Fig. 7 New energy scenario generation structure framework based on Bayesian GAN

because z_j^{noise} is used to generate the candidate sample data. Therefore, each of the simple Monte Carlo sums usually makes a reasonable contribution to the total marginal posterior estimate. Through iterative sampling, the candidate sample data can be obtained from the approximate posterior of θ_g and θ_d in the extreme case.

Figure 8 depicts the comparison between the data generated by Bayesian GAN and GAN, and the original data from 3×3 identical WP stations, where the abscissa represents the time sampling point, and the ordinate represents the power generation of each station. The difference between the real scenarios and the generated scenarios of multiple sites can be found in the figure. In general, the quality accuracy of renewable energy output scenarios generated by Bayesian GAN is higher than that of GAN. At the same time, this technology improves traditional GAN with the expressiveness of Bayesian formula probabilistic reasoning. The mathematical proof is clear and the interpretability is good, which prove the effectiveness of SML in realizing the uncertainty modeling of renewable energy output.

3.1.2 Typical scenario generation

For the IES uncertainty planning problem, PPF is calculated by using all the data of the large complex high-dimensional scenarios. The more decision variables and calculations there are, the more complex renewable energy planning is, and the more difficult to solve. In addition, the amount of information passed to the IES planning in most scenarios is not completely used by decision-makers, and the value density is low. Therefore, it is meaningful to effectively extract and use the key information of the large complex high-dimensional scenarios.

Typical scenario generation is to generate scenarios with similar statistical, correlation and shape characteristics to the original scenarios, and the Wasserstein distance metric is adopted here. Suppose $p_a(x)$ is a continuous PDF of one-dimensional random variable x, then Z_t is used to calculate the corresponding probability p_d^t as:

$$\int_{-\infty}^{Z_t} p_a(x)^{1/(1+h)} dx = \frac{2t-1}{2T} \int_{-\infty}^{\infty} p_a(x)^{1/(1+h)} dx \quad (5)$$

$$p_d^t = \int_{\frac{Z_t + Z_{t-1}}{2}}^{\frac{Z_t + Z_{t+1}}{2}} p_a(x) dx, \quad t = 1, \ldots, T \quad (6)$$

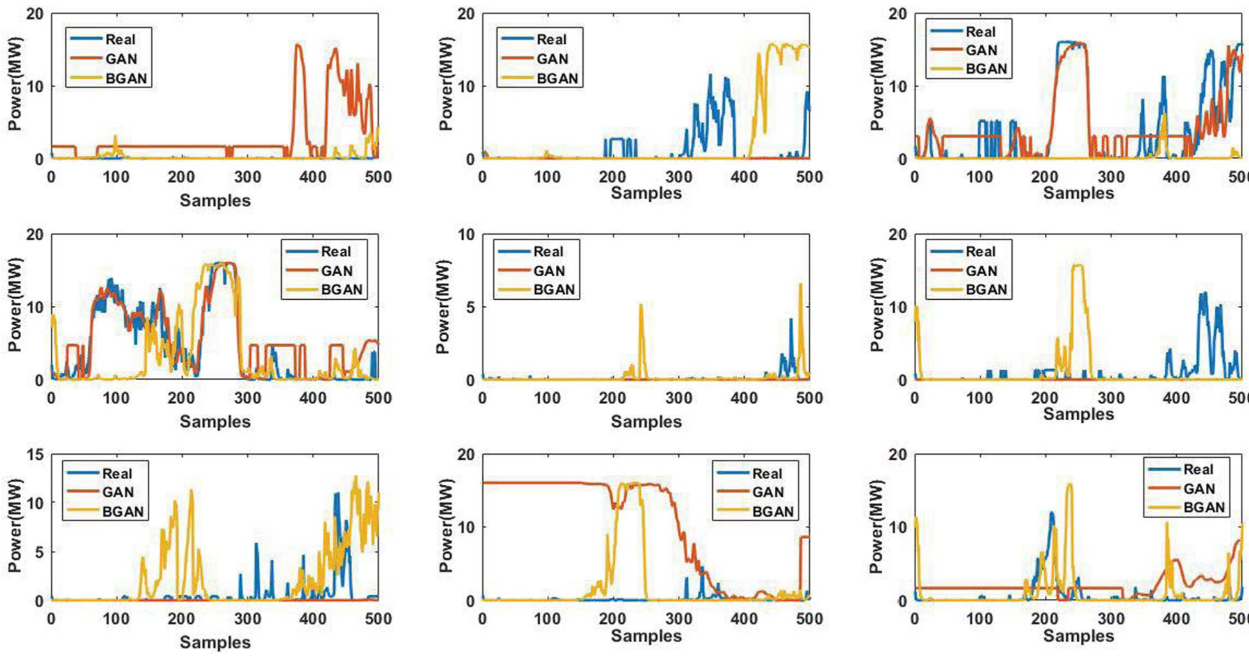

Fig. 8 WP simulation results using Bayesian GAN and GAN

In the optimal quantile theory presented in [154], the total of p_d^t calculated is not equal to 1. For IES, the energy variable always changes in an interval with upper and lower limits, so the following two different variable boundary approximation formulas are more consistent with the discretization of IES variables [155]:

$$p_d^0 = \int_{-\infty}^{\frac{Z_0+Z_1}{2}} p_a(x)dx \tag{7}$$

$$p_d^t = \int_{\frac{Z_t+Z_{t+1}}{2}}^{\infty} p_a(x)dx \tag{8}$$

As shown in Figs. 9 and 10, discretization of continuous variables is achieved through the Wasserstein distance. The PDF distributions represent the actual scenarios of CHP and GAS. Comparing PDF distribution shapes and discrete scenarios, one can see that the Wasserstein distance metric method is effective and accurate. Figure 11 achieves clustering by unsupervised learning k-means (Euclidean distance). They all refine complex operating scenarios into typical scenarios in engineering. The number of typical scenarios is small, and are used for planning modeling. The calculation

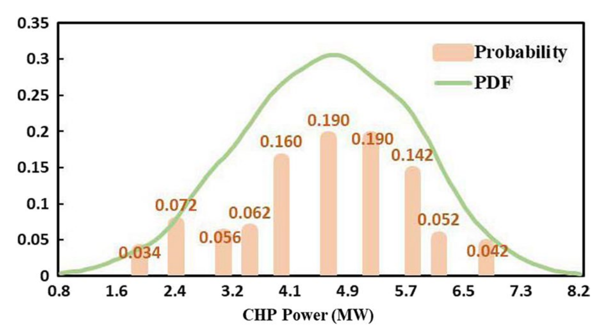

Fig. 9 Typical scenario set for the CHP based on the Wasserstein distance

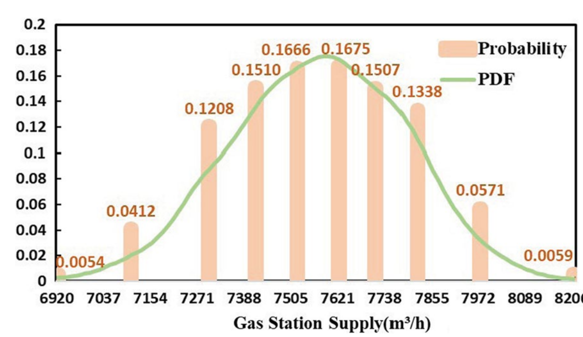

Fig. 10 Typical scenario set for the Gas based on the Wasserstein distance

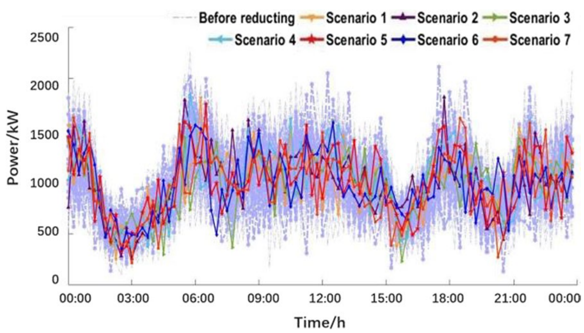

Fig. 11 Typical scenario set for the WP based on the K-means

efficiency is high, while the accuracy of the planning results is close to the planning accuracy of complex operating scenarios, which meets the engineering needs.

3.1.3 Semi-supervised learning algorithm

As shown in Fig. 12 and listed in Table 5, the conclusion is that semi-supervised learning is more suitable than unsupervised learning for the typical operating scenario for generation of REN with renewable energy. Because power grid regulatory personnel require knowledge of

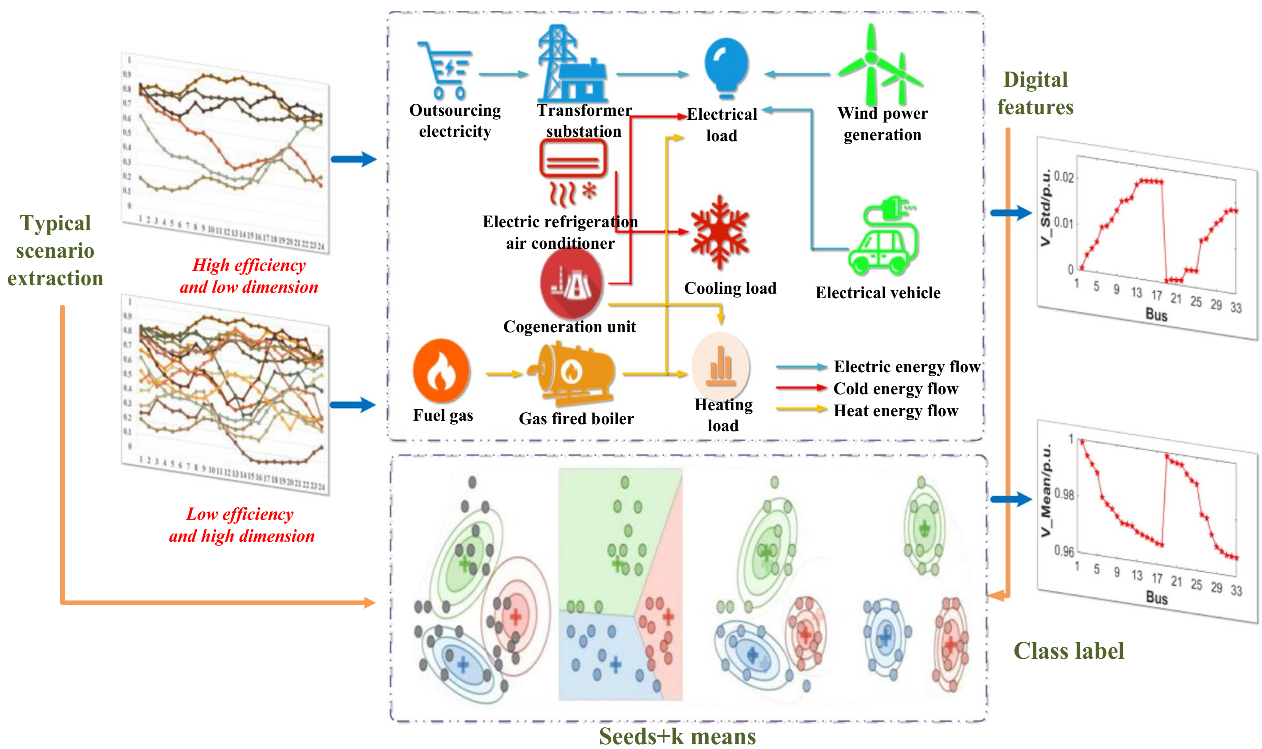

Fig. 12 Typical scenario extraction based on semi-supervised learning

Table 5 Typical scenario generation methods comparison

Method	Characteristic
Wasserstein distance	(1) High accuracy.
	(2) Able to generate extreme scenarios.
	(3) It can only handle a single continuous variable, that is, a weather correlation.
K-means	This method can handle multiple continuous variables.
Semi-supervised learning	(1) Able to generate multivariate typical scenarios.
	(2) Use the data features of a small number of samples as markers to retain the probability features of all state variables for PPF in massive scenarios.
	(3) Ensure the calculation accuracy of uncertainty planning.

the power flow of the grid, it needs to calculate the network power flow concerned with renewable energy power generation. The typical scenario extraction of renewable energy power generation based on unsupervised learning is not as good as the semi-supervised learning method considering renewable energy output x and grid state variable y.

3.2 Uncertainty modeling of power flow

As shown in Fig. 13, PPF can reflect the influence of system operation caused by random changes in various factors in the power system. It can comprehensively consider the uncertain situations of variable changes such as power system network topology, component parameters, node load values, generator output, etc. At the same time, it can also analyze the randomness of PV, WP, EV charging loads and thermal loads caused by temperature change, wind speed fluctuation, and solar radiation change and travel behavior. The above can contribute to discovering the frail segments in REN and provide valuable information to the planning and dispatching departments to help them make decisions.

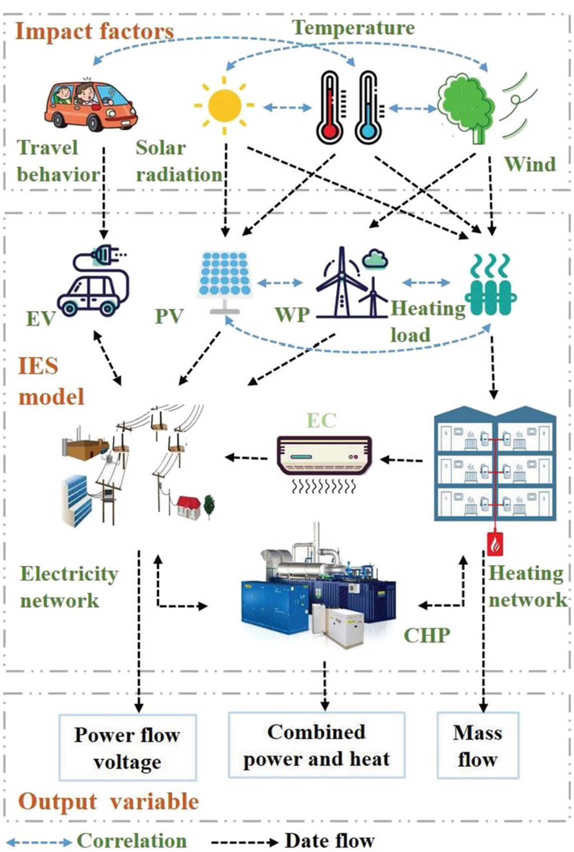

Fig. 13 PPF calculation process: the influence of PV, WP, EV load and thermal load on IES operation

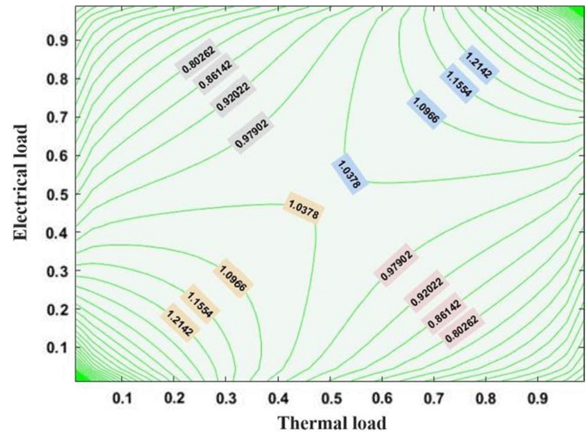

Fig. 14 Correlation between electrical load and thermal load

The uncertainty of power load and thermal load in networks will affect each other's energy networks through thermoelectric coupling. Figure 14 shows the bivariate copula function distribution of electrical load and thermal load. The larger the PDF value, the stronger the correlation.

Information entropy is perceived as the uncertainty of a random event or a measure of the amount of information. Greater information entropy indicates greater uncertainty, and greater uncertainty indicates greater complexity. The entropy $H(x)$ can be expressed as:

$$H(x) = -\sum_{k=1}^{M} p(x_k) \log_2 p(x_k) \tag{9}$$

where x_k is the k_{th} random variable.

As shown in Figs. 15, 16 and 17, information entropy is used to depict the strength of uncertainty. This paper sets three different scenarios with information entropy of [3.676, 7.015, 6.31]. In Fig. 15, CHP output characteristics are affected by uncertain variation of thermal load, and the entropies are in the area of 3.6. In Figs. 16 and 17, the uncertainties of the IES heating network and electricity network are jointly affected by PV, WP, EV charging loads

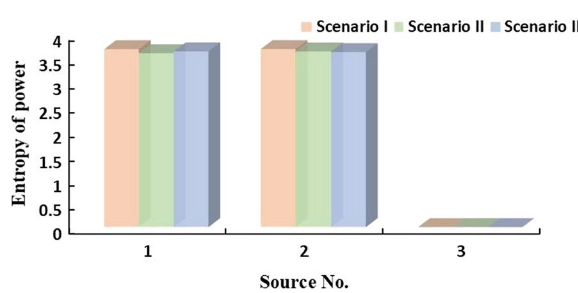

Fig. 15 CHP power outputs under different uncertainty

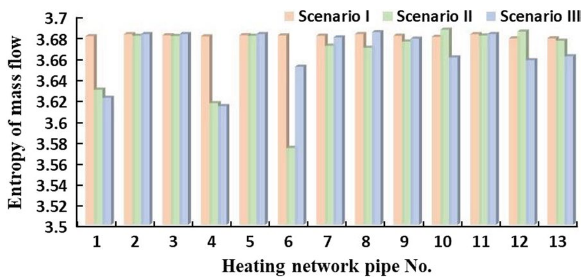

Fig. 16 Mass flows under different uncertainty

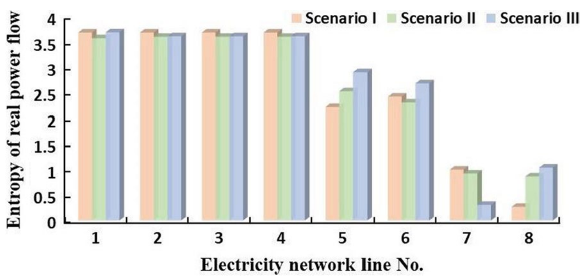

Fig. 17 Power flows under different uncertainty

and thermal loads. However, it can be seen that when the information entropy increases, that is, when the influencing factors increase, the entropies of scenarios II and III do not become larger. Therefore, it can be concluded that when more types and higher proportions of uncertain energies access the grid, the uncertainty of the IES overall operation will not increase.

3.3 Uncertainty modeling of renewable energy planning

A large amount of DRE integration will bring high-dimensional characteristics and uncertainty to REN. As shown in Table 6, the following conclusions are verified by practical engineering experience: (a) the calculation efficacy of probability theory is good when there are few random variables; (b) in the high-dimensional random variable modeling, the calculation efficacy is not good; and (c) with respect to the modeling of correlation, the method of the correlation coefficient is invalid. It can be concluded that dimensionality disaster is the difficulty of scientific data modeling. Modeling high-dimensional correlation variables is the key to break through the accuracy issue of distributed power cluster uncertainty planning modeling, which needs to be modeled by SML technologies. The key SML technologies considered in this paper are shown in Fig. 18 and listed in Table 7.

The decision variables of uncertainty planning are the location and capacity of DRE. High dimensional correlation modeling can connect deterministic decisions with uncertain objectives and constraints, and then the uncertain planning model can be established. Distribution network operating scenarios can be divided into typical scenarios and extreme scenarios. The planning results need to make typical scenarios economical and maintain high reliability in extreme scenarios. The key scientific problem of typical scenario extraction of an REN is to realize the classification of operation data. Similar scenarios are treated as one category, with an obvious effect in improving the efficiency of planning and arriving at solutions. Small probability events in extreme scenarios are

Table 6 Analysis of random variable dimensions and simulation results in practical engineering

Method	Reference and Journal Title	Renewable energy dimension and data sources	Analysis of simulation results
Probability Theory	[156] Electrical Power and Energy Systems	4 wind farms in China	The PDF of WP scenarios calculated by the RVM-copula method are very similar to the empirical copula, and the correlation of small-scale WP scenarios is more accurately simulated.
Probability Theory	[32] Journal of Modern Power Systems and Clean Energy	26 wind farms in East China	The R-vine copula model is introduced to deal with the high-dimensional characteristics and correlation of WP scenarios, which is more accurate and flexible than the Gaussian copula calculation results. When the scenario dimension increases, the accuracy of this method decreases due to the limitation of computational ability.
SML	[55] IEEE Transactions on Power Systems	24 wind farms, 32 solar power plants located in the Washington from NREL	The marginal distribution of generated scenario by the model-free GAN network is closer to the actual distribution than the Gaussian copula. When the power fluctuation of renewable energy is large and the spatial correlation is enhanced, the calculation accuracy of traditional probability theory is greatly reduced.

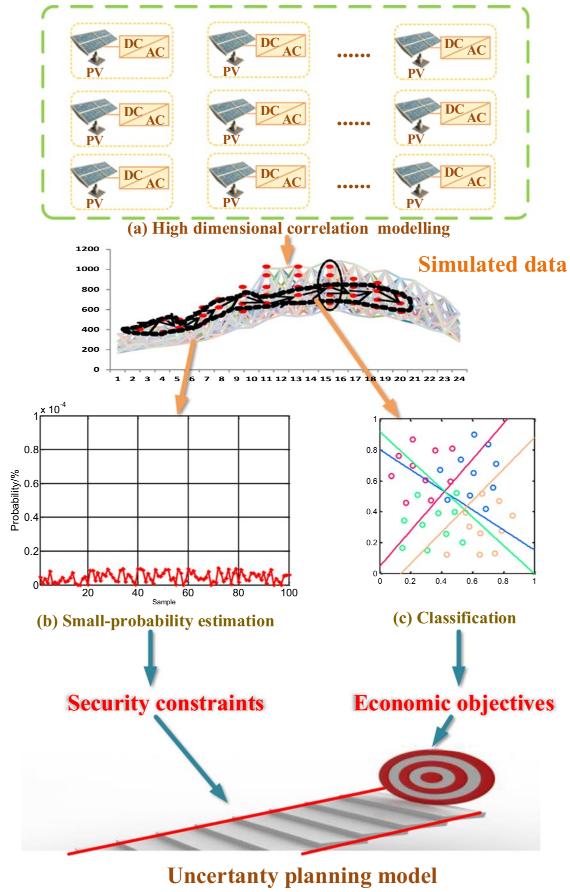

Fig. 18 The key SML technologies of uncertainty planning

the key factors affecting the reliability of REN planning. Therefore, how to use SML to estimate small probability events and realize classification is key and constitutes the main research area of DRE uncertainty planning.

4 Discussion

The topic of this paper is the optimal placement method of DRE. A deterministic decision-making scheme should be given considering the uncertainty of DRE. The difficulty in this field is how to model the uncertainty so that the deterministic planning results can meet the safety and economic requirements in uncertain environments.

For uncertainty modeling, Monte Carlo and machine learning are the two most popular and effective methods. When the search string "Monte Carlo uncertainty renewable energy" was entered on December 9, 2021, there were 856 publications listed by Web of Science, while there were 302 publications corresponding to the search string "machine learning uncertainty renewable energy". It can be concluded that both Monte Carlo and machine learning have received significant attention in the area of renewable energy.

However, there are areas that need to be further improved, as follows:

(a) The Monte Carlo method is based on the probability distribution function in DRE uncertainty simulation. The capacity of the probability model is small, and all information depends on several digital characteristics, such as expectation, variance and correlation coefficient. The complex relationship of the time–space correlation of multiple renewable energies cannot be fully described via some digital characteristics. Engineering practice shows that the probability modeling of a single DRE is accurate, while multiple DRE models based on probability correlation are not. Thus, probabilistic models of uncertainty need to be further improved to ensure the simulation performance of multiple DREs.

(b) There are two general uses for machine learning in the uncertainty modeling of DREs. The first is DRE generation prediction, and the second is stochastic scenario simulation of DREs. However, it is documented that a prediction method based on machine learning performs poorly when renewable energy generation fluctuates violently. Concerning stochastic scenario simulation of DREs, machine learning has weak explanatory power for the results. Therefore, machine learning models need to be further improved to ensure DRE model reliability and interpretability.

Below is a discussion on the development prospects of SML in future planning technology. It can solve large-scale planning problems subject to high renewable energy penetration. High renewable energy penetration is one of the most prominent features of the new generation REN, in which the number of operation scenarios is very large, and extreme scenarios are difficult to accurately predict. It is thus difficult to find a deterministic solution, one which can enhance the overall optimization under an ocean of operational scenarios and meet with a high probability strict constraint in each scenario.

Reference [164] considers that the probabilistic framework of machine intelligence has an attractive advantage in dealing with uncertainty modeling. It can be concluded that SML is an effective tool to deal with uncertainty. SML combines the advantages of machine learning and probability theory and is one of the effective tools to deal with uncertainty planning. On the one hand, the powerful data learning ability of machine learning can be used to deal with high-dimensional, complex, and uncertainty modeling, which greatly improves calculation efficiency and maximizes the use of existing electricity datasets. On the other hand, rigorous probability theory is used to ensure the interpretability and precision of the solution results of the REN planning model. The planning results can be quantified as the probability level so that planners can make decisions to meet the engineering

Table 7 Key SML techniques

Engineering problems	Scientific problems	Difficulty description	Key SML techniques	Role in planning
Large-scale renewable energy grid-connection	(a) High dimensional correlation modeling	(1) A large amount of renewable energy bring dimension disaster to uncertain modeling.	Singular value decomposition && principal component analysis [157]	Uncertainty modeling of decision variables, which includes capacity or location.
		(2) The high dimension reduces the accuracy of probability modeling.		
			Convolutional neural network[158]	
		(3) The correlation coefficient can only grasp the overall characteristics, and the correlation simulation is not accurate.		
Extreme operation scenarios	(b) Small probability estimation	Small probability event is difficult to estimate accurately, but it affects the electric network reliability.	Response surface methodology [153]	Uncertainty constraint modeling, which includes voltage amplitude and static voltage index.
			First-order reliability method [159]	
			Second-order reliability method [160]	
			Analytical method [161]	
			Central moment method [162]	
Typical operation scenarios	(c) Classification	Typical scenario extraction is usually based on clustering algorithm, but the correctness of unsupervised learning is difficult to verify.	The nearest neighbor approach && nonnegative matrix factorization [96]	Uncertainty objective modeling, which includes network loss and return on investment.
			Wasserstein distance metric [155]	
			k-means clustering algorithm [163]	

requirements. Based on the above analysis, the future development trend of DRE and grid connection technology can be explored. This is an information solution. SML and other artificial intelligence technologies need to be further developed to reduce the impact of DRE generation uncertainty.

5 Conclusions

Given the uncertainty planning problem of renewable energy, a robust planning method has been proposed in Chinese academia. In theory, robust planning is based on interval mathematics, and the idea of interval optimization is to use closed sets to express uncertainty. Solving the maximum and minimum state variables based on mathematical programming is essentially a planning method based on extreme condition scenarios, and its effectiveness depends on whether the selected extreme condition scenarios can represent the boundary of state variables throughout the planning period. Although the planning results are economically conservative, they can ensure the operational safety of distribution networks. The uncertainty optimization method, which is based on probability and statistics in theory, has been proposed by international academia. Probabilistic planning represents operational and security constraints at a specific expected confidence level so that decision-makers can see the relationship between risk and possible planning results. However, probability theory cannot fully describe the complex relationship of renewable energy spatial–temporal correlations, so it may lead to potential safety hazards hidden in the planning results. Because of the high proportion of renewable energy, it is increasingly difficult to grasp the features of power grid operation, and economy and security cannot be guaranteed by the traditional model-driven methods. Based on robust programming and probabilistic programming, distributional robust optimization has gradually attracted extensive attention by finding the decision results under the worst probability distribution of new energy uncertain parameter. Combining the advantages of both methods, the planning results show good performance in terms of economy and robustness. However, most of the distributional robust optimization processes need approximate transformation, and this reduces the accuracy of the results to some degree. SML is an effective way to solve the problem of uncertainty planning for DRE. In essence, it is a way to reduce the scope of DRE uncertainty, so that uncertainty planning can become closer to deterministic planning. This paper puts forward complex operating scenarios based on adversarial networks, extreme scenarios based on deep learning and representative scenarios based on semi-supervised learning. These can guarantee the efficiency and accuracy of uncertainty planning of distributed generation. Extreme scenarios guarantee the security of planning results, and representative scenarios guarantee the economy of planning results.

Acknowledgements
The author would like to thank the referees and the editor of this journal for valuable comments.

Authors' information
X.Q. Fu (1985-) received his B.S. and M.S. degrees from North China Electric Power University in 2008 and 2011, respectively. He received his Ph.D. degree from South China University of Technology in 2015. From 2011 to 2015, he was an electrical engineer with Guangzhou Power Supply Co. Ltd. From 2015 to 2017, he was a Post-Doctoral Researcher with Tsinghua University. He is currently an Associate Professor at China Agricultural University. His current research interests include statistical machine learning, Agricultural Energy Internet, and PV system integration. He is an associate Editor-in-Chief of "Information Processing in Agriculture", an editorial member of "Journal of Solar Energy Research Updates", and an associate editor of "Protection and Control of Modern Power Systems".
X.P. Wu (1997-), Major in the planning of integrated energy system based on the statistical machine learning.
C.Y. Zhang (2000-), Major in the statistical machine learning.
S.Q. Fan (1997-), Major in the planning of distribution network based on the statistical machine learning.
N. Liu (1981-), PHD and Professor, Major in the Smart Power Distribution and Consumption and Multi-Energy System Integration.

Author contributions
XF: conceptualization, methodology, writing—review & editing, supervision. XW, CZ: methodology, formal analysis, investigation, writing—original draft. SF, NL: investigation, writing—original draft. All authors read and approved the final manuscript.

Funding
This study is supported by the State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources under Grant No.LAPS21016, the National Natural Science Foundation of China under Grant 52007193, and the 2115 Talent Development Program of China Agricultural University.

Availability of data and materials
The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Declarations

Competing interests
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.

Author details
[1]College of Information and Electrical Engineering, China Agricultural University, 100083 Beijing, China. [2]Beijing Engineering and Technology Research Center for Internet of Things in Agriculture, Beijing, People's Republic of China. [3]State Key Laboratory of Alternate Electrical Power System With Renewable Energy Sources, North China Electric Power University, 102206 Beijing, China.

Received: 26 May 2022 Accepted: 10 October 2022
Published online: 19 October 2022

References

1. Wang, H., Liu, Y., Zhou, B., Li, C., & Barakhtenko, E. (2020). Taxonomy research of artificial intelligence for deterministic solar power forecasting. *Energy Conversion and Management, 214*, 112909.
2. Murty, V. V. S. N., & Kumar, A. (2020). Multi-objective energy management in microgrids with hybrid energy sources and battery energy storage systems. *Protection and Control of Modern Power Systems, 5*(1), 1–20.
3. Liu, Y., Lin, J., Wu, Q. H., & Zhou, X. (2017). Frequency control of DFIG-based wind power penetrated power systems using switching angle controller and AGC. *IEEE Transactions on Power Systems, 32*(2), 1553–1567.
4. Zheng, J. H., Chen, J., Wu, Q. H., & Jing, Z. (2015). Multi-objective optimization and decision making for power dispatch of a large-scale integrated energy system with distributed DHCs embedded. *Applied Energy, 154*, 369–379.
5. Shen, X., Shahidehpour, M., Han, Y., Zhu, S., & Zheng, J. (2017). Expansion planning of active distribution networks with centralized and distributed energy storage systems. *IEEE Transactions on Sustainable Energy, 8*(1), 126–134.
6. Shen, X., Shahidehpour, M., Zhu, S., Han, Y., & Zheng, J. (2018). Multi-stage planning of active distribution networks considering the co-optimization of operation strategies. *IEEE Transactions on Smart Grid, 9*(2), 1425–1433.
7. Maghouli, P., Hosseini, S. H., Buygi, M. O., & Shahidehpour, M. (2011). A scenario-based multi-objective model for multi-stage transmission expansion planning. *IEEE Transactions on Power Systems, 26*(1), 470–478.
8. Dong, W., Chen, X., & Yang, Q. (2022). Data-driven scenario generation of renewable energy production based on controllable generative adversarial networks with interpretability. *Applied Energy, 308*, 118387.
9. Wang, S., Chen, H., Li, X., & Shu, X. (2018). Conditional variational automatic encoder method for stochastic scenario generation of wind power and photovoltaic system. *Power System Technology, 42*(6), 1860–1867.
10. Wang, H., Ruan, J., Wang, G., Zhou, B., Liu, Y., Fu, X., & Peng, J. (2018). Deep learning-based interval state estimation of ac smart grids against sparse cyber attacks. *IEEE Transactions on Industrial Informatics, 14*(11), 4766–4778.
11. Zhang, C., Chen, H., Liang, Z., Guo, M., Hua, D., & Ngan, H. (2018). Reactive power optimization under interval uncertainty by the linear approximation method and its modified method. *IEEE Transactions on Smart Grid, 9*(5), 4587–4600.
12. Zhang, C., Chen, H., Shi, K., Qiu, M., Hua, D., & Ngan, H. (2018). An interval power flow analysis through optimizing-scenarios method. *IEEE Transactions on Smart Grid, 9*(5), 5217–5226.
13. Xu, Y., Hu, Z., Mili, L., Korkali, M., & Chen, X. (2020). Probabilistic power flow based on a Gaussian process emulator. *IEEE Transactions on Power Systems, 35*(4), 3278–3281.
14. Amer, A., Azab, A., Azzouz, M. A., & Awad, A. (2021). A stochastic program for siting and sizing fast charging stations and small wind turbines in urban areas. *IEEE Transactions on Sustainable Energy, 12*(2), 1217–1228.
15. Shahidehpour, M., Ding, T., Ming, Q., Huang, C., & Du, P. (2021). Multi-period active distribution network planning using multi-stage stochastic programming and nested decomposition by SDDIP. *IEEE Transactions on Power Systems, 36*(3), 2281–2292.
16. Yao, Z., & Wang, Z. (2020). Two-level collaborative optimal allocation method of integrated energy system considering wind and solar uncertainty. *Power System Technology, 44*(12), 4521–4531.
17. Liu, W., Li, Z., Yang, Y., Yi, F., & Wang, Y. (2020). Collaborative optimal configuration for integrated energy system considering uncertainties of demand response. *Automation of Electric Power Systems, 44*(10), 41–49.
18. https://www.stat.berkeley.edu/~statlearning/index.html.
19. Zhao, C., Wang, J., Watson, J. P., & Guan, Y. (2013). Multi-stage robust unit commitment considering wind and demand response uncertainties. *IEEE Transactions on Power Systems, 28*(3), 2708–2717.
20. Fu, X., Wu, X., & Liu, N. (2021). Statistical machine learning model for uncertainty planning of distributed renewable energy sources in distribution networks. *Frontiers in Energy Research, 9*, 809254.
21. Fu, X., Guo, Q., & Sun, H. (2022). Statistical machine learning model for capacitor planning considering uncertainties in photovoltaic power. *Protection and Control of Modern Power Systems, 7*(1), 1–13.
22. Xu, J., Yi, X., Sun, Y., Lan, T., & Sun, H. (2017). Stochastic optimal scheduling based on scenario analysis for wind farms. *IEEE Transactions on Sustainable Energy, 8*(4), 1548–1559.
23. Wang, Z., Shen, C., & Liu, F. (2017). A conditional model of wind power forecast errors and its application in scenario generation. *Applied Energy, 212*, 771–785.
24. Quan, H., Srinivasan, D., Khambadkone, A. M., & Khosravi, A. (2015). A computational framework for uncertainty integration in stochastic unit commitment with intermittent renewable energy sources. *Applied Energy, 152*, 71–82.
25. Chen, Z., Gao, Z., Chen, J., Wu, X., Fu, X., & Chen, X. (2021). Research on cooperative planning of an integrated energy system considering uncertainty. *Power System Protection and Control, 49*(8), 32–40.
26. Wan, C., Lin, J., Wang, J., Song, Y., & Dong, Z. Y. (2017). Direct quantile regression for nonparametric probabilistic forecasting of wind power generation. *IEEE Transactions on Power Systems, 32*(4), 2767–2778.
27. Mahdi, K., Wang, J., & Mohammad, M. (2019). Interval deep generative neural network for wind speed forecasting. *IEEE Transactions on Smart Grid, 10*(4), 3974–3989.
28. Chang, K.-H. (2016). A quantile-based simulation optimization model for sizing hybrid renewable energy systems. *Simulation Modelling Practice & Theory, 66*, 94–103.
29. Mokryani, G., Majumdar, A., & Pal, B. C. (2016). A probabilistic method for the operation of three-phase unbalanced active distribution networks. *IET Renewable Power Generation, 10*(7), 944–954.
30. Li, Q., Wang, X., & Rong, S. (2018). Probabilistic load flow method based on modified Latin hypercube-important sampling. *Energies, 11*(11), 317.
31. Diaz, G., Casielles, P. G., & Coto, J. (2014). Simulation of spatially correlated wind power in small geographic areas-sampling methods and evaluation. *International Journal of Electrical Power & Energy Systems, 63*, 513–522.
32. Wang, Z., Wang, W., Liu, C., & Wang, B. (2020). Forecasted scenarios of regional wind farms based on regular vine copulas. *Journal of Modern Power Systems and Clean Energy, 8*(1), 77–85.
33. Becker, R. (2018). Generation of time-coupled wind power infeed scenarios using pair-copula construction. *IEEE Transactions on Sustainable Energy, 9*(3), 1298–1306.
34. Wang, Z., Wang, W., Liu, C., Wang, Z., & Hou, Y. (2018). Probabilistic forecast for multiple wind farms based on regular vine copulas. *IEEE Transactions on Power Systems, 33*(1), 578–589.
35. Fu, X., Jia, Q., & Wu, X. (2021). Fast probability power flow calculation of distribution networks considering dynamic correlation and high-dimensional uncertainty. In *Renewable energy microgeneration systems* (pp. 17–48).
36. Lee, D., & Baldick, R. (2017). Load and wind power scenario generation through the generalized dynamic factor model. *IEEE Transactions on Power Systems, 32*(1), 400–410.
37. Hu, J., Li, H., & Liu, Z. (2021). A novel scenario generation framework based on the knowledge of existing wind power plants. *IEEE Transactions on Sustainable Energy, 12*(2), 1229–1241.
38. Wu, H., Krad, I., Florita, A., Hodge, B. M., Ibanez, E., Zhang, J., & Ela, E. (2017). Stochastic multi-timescale power system operations with variable wind generation. *IEEE Transactions on Power Systems, 32*(5), 3325–3337.
39. Li, D., Yan, W., Li, W., & Ren, Z. (2016). A two-tier wind power time series model considering day-to-day weather transition and intraday wind power fluctuations. *IEEE Transactions on Power Systems, 31*(6), 4330–4339.
40. Xie, K., Liao, Q., Tai, H., & Hu, B. (2017). Non-homogeneous Markov wind speed time series model considering daily and seasonal variation characteristics. *IEEE Transactions on Sustainable Energy, 8*(3), 1281–1290.
41. Morales, J. M., Mínguez, R., & Conejo, A. J. (2010). A methodology to generate statistically dependent wind speed scenarios. *Applied Energy, 87*(3), 843–855.
42. Yunus, K. J., Thiringer, T., & Chen, P. (2016). ARIMA-based frequency-decomposed modeling of wind speed time series. *IEEE Transactions on Power Systems, 31*(4), 2546–2556.
43. Hoeltgebaum, H., Fernandes, C., & Street, A. (2018). Generating joint scenarios for renewable generation: The case for non-Gaussian models with time-varying parameters. *IEEE Transactions on Power Systems, 33*(6), 7011–7019.

44. Diaz, G., Gomez-Aleixandre, J., & Coto, J. (2016). Wind power scenario generation through state-space specifications for uncertainty analysis of wind power plants. *Applied Energy, 162*, 21–30.
45. Karamichailidou, D. A., Kaloutsa, V., & Alexandridis, A. (2021). Wind turbine power curve modeling using radial basis function neural networks and Tabu search. *Renewable Energy, 163*, 2137–2152.
46. Vagropoulos, S. I., Kardakos, E. G., Simoglou, C. K., Bakirtzis, A. G., & Catalao, J. P. S. (2016). ANN-based scenario generation methodology for stochastic variables of electric power systems. *Electric Power Systems Research, 134*, 9–18.
47. Khodayar, M., Kaynak, O., & Khodayar, M. E. (2017). Rough deep neural architecture for short-term wind speed forecasting. *IEEE Transactions on Industrial Informatics, 13*(6), 2770–2779.
48. Yu, L., Zhang, W., Wang, J., & Yong, Y. (2017). SeqGAN: Sequence generative adversarial nets with policy gradient. In *2017 31st AAAI conference on artificial intelligence*.
49. Chen, Y., Wang, X., & Zhang, B. (2018). An unsupervised deep learning approach for scenario forecasts. *Power Systems Computation Conference (PSCC), 2018*, 1–7.
50. Liu, Y., Zhou, Y., Liu, X., Dong, F., Wang, C., & Wang, Z. (2019). Wasserstein GAN-based small-sample augmentation for new-generation artificial intelligence: A case study of cancer-staging data in biology. *Engineering, 5*(1), 156–163.
51. Jiang, C., Mao, Y., Chai, Y., Yu, M., & Tao, S. (2018). Scenario generation for wind power using improved generative adversarial networks. *IEEE Access, 6*, 62193–62203.
52. Saatchi, Y., & Wilson, A. G. (2017). Bayesian GANs. In *2017 conference and workshop on neural information processing systems* (p. 30).
53. Chen, Y., Li, P., & Zhang, B. (2018). Bayesian renewables scenario generation via deep generative networks. In *2018 52nd annual conference on information sciences and systems (CISS)* (pp. 1–6).
54. Liang, J., & Tang, W. (2020). Sequence generative adversarial networks for wind power scenario generation. *IEEE Journal on Selected Areas in Communications, 38*(1), 110–118.
55. Chen, Y., Wang, Y., Kirschen, D. S., & Zhang, B. (2018). Model-free renewable scenario generation using generative adversarial networks. *IEEE Transactions on Power Systems, 33*(3), 3265–3275.
56. Golestaneh, F., Pinson, P., & Gooi, H. B. (2016). Very short-term nonparametric probabilistic forecasting of renewable energy generation-with application to solar energy. *IEEE Transactions on Power Systems, 31*(5), 3850–3863.
57. Carpinelli, G., Caramia, P., & Varilone, P. (2015). Multi-linear Monte Carlo simulation method for probabilistic load flow of distribution systems with wind and photovoltaic generation systems. *Renewable Energy, 76*, 283–295.
58. Qiu, Y., Li, Q., Pan, Y., Yang, H., & Chen, W. (2019). A scenario generation method based on the mixture vine copula and its application in the power system with wind/hydrogen production. *International Journal of Hydrogen Energy, 44*(11), 5162–5170.
59. Ailliot, P., Bessac, J., Monbet, V., & Pene, F. (2015). Non-homogeneous hidden Markov-switching models for wind time series. *Journal of Statistical Planning and Inference, 160*, 75–88.
60. Yang, M., Lin, Y., Zhu, S., Han, X., & Wang, H. (2015). Multi-dimensional scenario forecast for generation of multiple wind farms. *Journal of Modern Power Systems and Clean Energy, 3*(3), 361–370.
61. Samadianfard, S., Hashemi, S., Kargar, K., Izadyar, M., Mostafaeipour, A., Mosavi, A. H., Nabipour, N., & Shamshirband, S. (2020). Wind speed prediction using a hybrid model of the multi-layer perceptron and whale optimization algorithm. *Energy Reports, 6*, 1147–1159.
62. Wasilewski, J., & Baczynski, D. (2017). Short-term electric energy production forecasting at wind power plants in Pareto-optimality context. *Renewable & Sustainable Energy Reviews, 69*, 177–187.
63. Vassallo, D., Krishnamurthy, R., Sherman, T., & Fernando, H. (2020). Analysis of random forest modeling strategies for multi-step wind speed forecasting. *Energies, 13*(20), 5488.
64. Lu, X., Liu, Z., Ma, L., Wang, L., Zhou, K., & Feng, N. (2020). A robust optimization approach for optimal load dispatch of community energy hub. *Applied Energy, 259*, 114195.
65. Abdi, H., Moradi, M., & Lumbreras, S. (2021). Metaheuristics and transmission expansion planning: A comparative case study. *Energies, 14*(12), 3618.
66. Rasouli, B., Salehpour, M., Wang, J., & Kim, S. J. (2019). Optimal day-ahead scheduling of a smart micro-grid via a probabilistic model for considering the uncertainty of electric vehicles' load. *Applied Sciences, 9*(22), 4872.
67. Zhang, J., Wang, S., Zhang, C., Luo, F., Dong, Z., & Li, Y. (2021). Planning of electric vehicle charging stations and distribution system with highly renewable penetrations. *IET Electrical Systems in Transportation, 11*(3), 256–268.
68. Karimiarpanahi, S., Jooshaki, M., Moeinaghtaie, M., Fotuhifiruzabad, M., & Lehtonen, M. (2020). Considering forecasting errors in flexibility-oriented distribution network expansion planning using the spherical simplex unscented transformation. *IET Generation Transmission & Distribution, 14*(24), 5970–5983.
69. Zheng, Y., Shao, Z., Zhang, Y., & Jian, L. (2020). A systematic methodology for mid-and-long term electric vehicle charging load forecasting: The case study of Shenzhen, China. *Sustainable Cities and Society, 56*, 102084.
70. Chen, W., Zhang, L., & Pei, X. (2021). Probability evaluation of excess voltage in a distribution network with uneven charging electric vehicle load. *Journal of Electrical Engineering & Technology, 17*(1), 15–23.
71. Alilou, M., Tousi, B., & Shayeghi, H. (2020). Home energy management in a residential smart micro grid under stochastic penetration of solar panels and electric vehicles. *Solar Energy, 212*, 6–18.
72. Vatandoust, B., Ahmadian, A., & Golkar, M. A. (2016). Stochastic copula-based multivariate modeling of plug-in hybrid electric vehicles load demand in residential distribution network. *Smart Grids Conference (SGC), 2016*, 1–7.
73. Tehrani, N. H., & Wang, P. (2015). Probabilistic estimation of plug-in electric vehicles charging load profile. *Electric Power Systems Research, 124*, 133–143.
74. Sarabi, S., Davigny, A., Courtecuisse, V., Riffonneau, Y., & Robyns, B. (2016). Potential of vehicle-to-grid ancillary services considering the uncertainties in plug-in electric vehicle availability and service/localization limitations in distribution grids. *Applied Energy, 171*, 523–540.
75. Wang, L., Sharkh, S., & Chipperfield, A. (2018). Optimal decentralized coordination of electric vehicles and renewable generators in a distribution network using a search. *International Journal of Electrical Power & Energy Systems, 98*, 474–487.
76. Hakimi, S. M., Hasankhani, A., Shafie-Khah, M., & Catalo, J. (2021). Stochastic planning of a multi-microgrid considering integration of renewable energy resources and real-time electricity market. *Applied Energy, 298*, 117215.
77. Seddig, K., Jochem, P., & Fichtner, W. (2019). Two-stage stochastic optimization for cost-minimal charging of electric vehicles at public charging stations with photovoltaics. *Applied Energy, 242*, 769–781.
78. Tao, S., Liao, K., Xiao, X., Wen, J., Yang, Y., & Jian, Z. (2016). Charging demand for electric vehicle based on stochastic analysis of trip chain. *IET Generation Transmission & Distribution, 10*(11), 2689–2698.
79. Xiang, Y., Jiang, Z., Gu, C., Teng, F., Wei, X., & Wang, Y. (2019). Electric vehicle charging in smart grid: A spatial-temporal simulation method. *Energy, 189*, 116221.
80. Xue, P., Xiang, Y., Gou, J., Xu, W., & Liu, J. (2021). Impact of large-scale mobile electric vehicle charging in smart grids: A reliability perspective. *Frontiers in Energy Research, 9*, 688034.
81. Lin, H., Fu, K., Wang, Y., Sun, Q., & Wennersten, R. (2019). Characteristics of electric vehicle charging demand at multiple types of location-application of an agent-based trip chain model. *Energy, 188*, 116122.
82. Wang, Z., Jochem, P., & Fichtner, W. (2020). A scenario-based stochastic optimization model for charging scheduling of electric vehicles under uncertainties of vehicle availability and charging demand. *Journal of Cleaner Production, 254*, 119886.
83. Ehsan, A., & Yang, Q. (2020). Active distribution system reinforcement planning with EV charging stations—part I: Uncertainty modeling and problem formulation. *IEEE Transactions on Sustainable Energy, 11*(2), 970–978.
84. Fan, V. H., Dong, Z., & Meng, K. (2020). Integrated distribution expansion planning considering stochastic renewable energy resources and electric vehicles. *Applied Energy, 278*, 115720.
85. You, Y., Zhu, J., Huang, Y., & Jing, Z. (2021). Optimal decision-making method for a plug-in electric taxi in uncertain environment. *IEEE Access, 9*, 62467–62477.

86. Qian, X., Xue, J., Sobolevsky, S., Yang, C., & Ukkusuri, S. V. (2019). Stationary spatial charging demand distribution for commercial electric vehicles in urban area. *IEEE Intelligent Transportation Systems Conference (ITSC), 2019*, 220–225.
87. Lakshminarayanan, V., Chemudupati, V., Pramanick, S. K., & Rajashekara, K. (2019). Real-time optimal energy management controller for electric vehicle integration in workplace microgrid. *IEEE Transactions on Transportation Electrification, 5*(1), 174–185.
88. Liu, Y., Hong, Y., & Hu, C. (2019). Research on big data mining technology of electric vehicle charging behaviour. *Elektronika Ir Elektrotechnika, 25*(6), 55–61.
89. Deb, S., Goswami, A. K., Harsh, P., Sahoo, J. P., & Shekhawat, A. S. (2020). Charging coordination of plug-in electric vehicle for congestion management in distribution system integrated with renewable energy sources. *IEEE Transactions on Industry Applications, 56*(5), 5452–5462.
90. Hao, Y., Dong, L., Liang, J., Liao, X., Wang, L., & Shi, L. (2020). Power forecasting-based coordination dispatch of PV power generation and electric vehicles charging in microgrid. *Renewable Energy, 155*, 1191–1210.
91. Tan, B., & Chen, H. (2020). Multi-objective energy management of multiple microgrids under random electric vehicle charging. *Energy, 208*, 118360.
92. Zhang, X., Chan, K., Li, H., Wang, H., & Wang, G. (2021). Deep-learning-based probabilistic forecasting of electric vehicle charging load with a novel queuing model. *IEEE Transactions on Cybernetics, 51*(6), 3157–3170.
93. Zhong, J., & Xiong, X. (2021). An orderly EV charging scheduling method based on deep learning in cloud-edge collaborative environment. *Advances in Civil Engineering*, 6690610.
94. Pan, A., Zhao, T., Yu, H., & Zhang, Y. (2019). Deploying public charging stations for electric taxis: A charging demand simulation embedded approach. *IEEE Access, 7*, 17412–17424.
95. Jiang, C., Jing, Z., Ji, T., & Wu, Q. H. (2018). Optimal location of PEVCSs using MAS and ER approach. *IET Generation Transmission & Distribution, 12*(20), 4377–4387.
96. Fu, X., Guo, Q., & Sun, H. (2020). Statistical machine learning model for stochastic optimal planning of distribution networks considering a dynamic correlation and dimension reduction. *IEEE Transactions on Smart Grid, 11*(4), 2904–2917.
97. Rakpenthai, C., Uatrongjjit, S., & Premrudeepreechacharn, S. (2012). State estimation of power system considering network parameter uncertainty based on parametric interval linear systems. *IEEE Transactions on Power Systems, 27*(1), 305–313.
98. Muhammad, M. A., Mokhlis, H., Naidu, K., Amin, A., Franco, J. F., & Othman, M. (2019). Distribution network planning enhancement via network reconfiguration and DG integration using dataset approach and water cycle algorithm. *Journal of Modern Power Systems and Clean Energy, 8*(1), 86–93.
99. Ameli, A., Farrokhifard, M. R., Davari-nejad, E., Oraee, H., & Haghifam, M. R. (2015). Profit-based DG planning considering environmental and operational issues: A multiobjective approach. *IEEE Systems Journal, 11*(4), 1959–1970.
100. Reddy, G. H., Chakrapani, P., Goswami, A. K., & Choudhury, N. B. D. (2017). Optimal distributed generation placement in distribution system to improve reliability and critical loads pick up after natural disasters. *Engineering science and technology, an international journal, 20*(3), 825–832.
101. Karunarathne, E., Pasupuleti, J., Ekanayake, J., & Almeida, D. (2020). Optimal placement and sizing of DGs in distribution networks using MLPSO algorithm. *Energies, 13*(23), 6185.
102. Chen, X., Wu, W., Zhang, B., & Lin, C. (2016). Data-driven DG capacity assessment method for active distribution networks. *IEEE Transactions on Power Systems, 32*(5), 3946–3957.
103. Prakash, D. B., & Lakshminarayana, C. (2018). Multiple DG placements in radial distribution system for multi objectives using Whale Optimization Algorithm. *Alexandria Engineering Journal, 57*(4), 2797–2806.
104. Zhang, L., Shen, C., Chen, Y., Huang, S., & Tang, W. (2018). Coordinated allocation of distributed generation, capacitor banks and soft open points in active distribution networks considering dispatching results. *Applied Energy, 231*, 1122–1131.
105. Zhang, L., Shen, C., Chen, Y., Huang, S., & Tang, W. (2017). Coordinated optimal allocation of DGs, capacitor banks and SOPs in active distribution network considering dispatching results through bi-level programming. *Energy Procedia, 142*, 2065–2071.
106. Gao, R., Kou, P., Liang, D., & Yang, C. (2021). Optimal allocation of hybrid distribution transformers considering dynamical control. *International Journal of Electrical Power & Energy Systems, 131*, 107052.
107. Liu, S., Zhou, C., Guo, H., Shi, Q., Song, T. E., Schomer, I., & Liu, Y. (2021). Operational optimization of a building-level integrated energy system considering additional potential benefits of energy storage. *Protection and Control of Modern Power Systems, 6*(1), 1–10.
108. Zhang, S., Cheng, H., Li, K., Tai, N., Wang, D., & Li, F. (2018). Multi-objective distributed generation planning in distribution network considering correlations among uncertainties. *Applied Energy, 226*, 743–755.
109. Xiong, X., Wu, W., Li, N., Yang, L., Zhang, J., & Wei, Z. (2019). Risk-based multi-objective optimization of distributed generation based on GPSO-BFA algorithm. *IEEE Access, 7*, 30563–30572.
110. Chen, J., Xiao, Y., Mo, R., & Tian, Y. (2021). Optimized allocation of microgrid energy storage capacity considering photovoltaic correction. *Power System Control and Protection, 49*, 59–66.
111. Tanaka, I., & Ohmori, H. (2017). New scenario-based stochastic programming problem for long-term allocation of renewable distributed generations. In *6th International conference on operations research and enterprise systems, ICORES* (pp. 96–107).
112. Tanaka, I., Yuge, H., & Ohmori, H. (2017). Formulation and evaluation of long-term allocation problem for renewable distributed generations. *IET Renewable Power Generation, 11*(12), 1584–1596.
113. De Lima, T. D., Tabares, A., Arias, N. B., & Franco, J. F. (2019). A stochastic programming model for the planning of distribution systems considering renewable distributed generation and CO2 emissions. *IEEE PES Innovative Smart Grid Technologies Conference-Latin America (ISGT Latin America), 2019*, 1–6.
114. Bazrafshan, M., Yalamanchili, L., Gatsis, N., & Gomez, J. (2019). Stochastic planning of distributed PV generation. *Energies, 12*(3), 459.
115. Santos, S. F., Fitiwi, D. Z., Bizuayehu, A. W., Shafie-khah, M., Asensio, M., Contreras, J., Pereira Cabrita, C. M., & Catalão, J. P. (2017). Novel multi-stage stochastic DG investment planning with recourse. *IEEE Transactions on Sustainable Energy, 8*(1), 164–178.
116. Manna, D., & Goswami, S. K. (2020). Optimum placement of distributed generation considering economics as well as operational issues. *International Transactions on Electrical Energy Systems, 30*(3), 12246.
117. Saha, S., & Mukherjee, V. (2021). A novel multi-objective modified symbiotic organisms search algorithm for optimal allocation of distributed generation in radial distribution system. *Neural Computing and Applications, 33*(6), 1751–1771.
118. Da Rosa, W., Gerez, C., & Belati, E. (2018). Optimal distributed generation allocating using particle swarm optimization and linearized ac load flow. *IEEE Latin America Transactions, 16*(10), 2665–2670.
119. Lee, D., Son, S., & Kim, I. (2021). Optimal allocation of large-capacity distributed generation with the volt/var control capability using particle swarm optimization. *Energies, 14*(11), 3112.
120. Fu, X., Chen, H., Cai, R., & Yang, P. (2015). Optimal allocation and adaptive VAR control of PV-DG in distribution networks. *Applied Energy, 137*, 173–182.
121. Khanbabapour, S., & Golshan, M. H. (2016). Synchronous DG planning for simultaneous improvement of technical, overcurrent, and timely anti-islanding protection indices of the network to preserve protection coordination. *IEEE Transactions on Power Delivery, 32*(1), 474–483.
122. Eid, A., Kamel, S., Korashy, A., & Khurshaid, T. (2020). An enhanced artificial ecosystem-based optimization for optimal allocation of multiple distributed generations. *IEEE Access, 8*, 178493–178513.
123. Fatemi, S. S., & Samet, H. (2019). Allocation of renewables, switches, and relays considering relays actual operation time. *IEEE Systems Journal, 14*(1), 950–959.
124. Flores-Quiroz, A., & Strunz, K. (2021). A distributed computing framework for multi-stage stochastic planning of renewable power systems with energy storage as flexibility option. *Applied Energy, 291*, 116736.
125. Graditi, G., Ciavarella, R., Di Somma, M., & Valenti, M. (2019). A control strategy for participation of DSO flexible resources in TSO ancillary services provision. In *2019 international conference on clean electrical power (ICCEP)* (pp. 586–592).
126. Guerrero, R. C., & Pedrasa, M. A. A. (2019). An integrated clustering-milp-simulation approach for distributed energy resource sizing. In *2019 IEEE*

international conference on environment and electrical engineering and 2019 IEEE industrial and commercial power systems Europe (EEEIC/I&CPS Europe) (pp. 1–6).
127. Ahmadi, S., & Abdi, S. (2016). Application of the Hybrid Big Bang-Big Crunch algorithm for optimal sizing of a stand-alone hybrid PV/wind/battery system. *Solar Energy, 134*, 366–374.
128. Varasteh, F., Nazar, M. S., Heidari, A., Shafie-khah, M., & Catalão, J. P. (2019). Distributed energy resource and network expansion planning of a CCHP based active microgrid considering demand response programs. *Energy, 172*, 79–105.
129. Wu, D., Yang, T., Stoorvogel, A. A., & Stoustrup, J. (2016). Distributed optimal coordination for distributed energy resources in power systems. *IEEE Transactions on Automation Science and Engineering, 14*(2), 414–424.
130. Barik, A. K., Das, D. C., Latif, A., Hussain, S. S., & Ustun, T. S. (2021). Optimal voltage–frequency regulation in distributed sustainable energy-based hybrid microgrids with integrated resource planning. *Energies, 14*(10), 2735.
131. Maity, D., Ghosh, A., Banerjee, S., & Chanda, C. K. (2018). Microgrid: Planning of optimal placing of distributed energy resources by loss and fuel cost reduction using map reduce optimization algorithm approach. In *2018 national power engineering conference (NPEC)* (pp. 1–5).
132. Zhang, J., Wang, G., Song, Y., Zhao, F., & Wang, S. (2019). Multiagent task planning based on distributed resource scheduling under command and control structure. *Mathematical Problems in Engineering, 2019*.
133. Zakernezhad, H., Nazar, M. S., Shafie-khah, M., & Catalão, J. P. (2021). Multi-level optimization framework for resilient distribution system expansion planning with distributed energy resources. *Energy, 214*, 118807.
134. Sun, K., Wang, J., Wang, L., Chen, X., Pan, H., & Zheng, J. H. (2018). Multi-objective capacity planning for distributed energy interconnection system in distributed networks. In *The 11th IET international conference on advances in power system control, operation and management (APSCOM 2018)* (pp. 1–6).
135. Maroufmashat, A., Sattari, S., Roshandel, R., Fowler, M., & Elkamel, A. (2016). Multi-objective optimization for design and operation of distributed energy systems through the multi-energy hub network approach. *Industrial & Engineering Chemistry Research, 55*(33), 8950–8966.
136. Ney, R. C., Canha, L. N., Adeyanju, O. M., & Arend, G. (2019). Multi-objective optimal planning of distributed energy resources using SPEA2 algorithms considering multi-agent participation. In *2019 54th international universities power engineering conference (UPEC)* (pp. 1–6).
137. Quan, Z., Jian, W., Liqun, R., Yuxu, J., Dan, Z., Jiayao, Q., & Tianyu, J. (2020). Optimal planning method for distributed wind/solar/battery intergraded microgrid. *Journal of Physics: Conference Series, 1659*, 012024.
138. Kumawat, M., Gupta, N., Jain, N., & Bansal, R. C. (2018). Optimal planning of distributed energy resources in harmonics polluted distribution system. *Swarm and Evolutionary Computation, 39*, 99–113.
139. Wang, J., Zhong, H., Xia, Q., & Kang, C. (2017). Optimal planning strategy for distributed energy resources considering structural transmission cost allocation. *IEEE Transactions on Smart Grid, 9*(5), 5236–5248.
140. Morvaj, B., Evins, R., & Carmeliet, J. (2016). Optimization framework for distributed energy systems with integrated electrical grid constraints. *Applied Energy, 171*, 296–313.
141. Jeddi, B., Vahidinasab, V., Ramezanpour, P., Aghaei, J., Shafie-khah, M., & Catalão, J. P. (2019). Robust optimization framework for dynamic distributed energy resources planning in distribution networks. *International Journal of Electrical Power & Energy Systems, 110*, 419–433.
142. Kanwar, N., Gupta, N., Niazi, K. R., Swarnkar, A., & Bansal, R. C. (2017). Simultaneous allocation of distributed energy resource using improved particle swarm optimization. *Applied Energy, 185*, 1684–1693.
143. Li, P., Wang, Z., Wang, J., Guo, T., & Yin, Y. (2021). A multi-time-space scale optimal operation strategy for a distributed integrated energy system. *Applied Energy, 289*, 116698.
144. Mashayekh, S., Stadler, M., Cardoso, G., & Heleno, M. (2017). A mixed integer linear programming approach for optimal DER portfolio, sizing, and placement in multi-energy microgrids. *Applied Energy, 187*, 154–168.
145. Ichoua, S. (2014). A stochastic approach for the integration of distributed energy resources. *IEEE International Technology Management Conference, 2014*, 1–4.
146. Ghatikar, G., Mashayekh, S., Stadler, M., Yin, R., & Liu, Z. (2016). Distributed energy systems integration and demand optimization for autonomous operations and electric grid transactions. *Applied Energy, 167*, 432–448.
147. Gilani, M. A., Kazemi, A., & Ghasemi, M. (2020). Distribution system resilience enhancement by microgrid formation considering distributed energy resources. *Energy, 191*, 116442.
148. Thang, V. V. (2021). Optimal sizing of distributed energy resources and battery energy storage system in planning of islanded micro-grids based on life cycle cost. *Energy Systems, 12*(3), 637–656.
149. Kardoš, J., Kourounis, D., & Schenk, O. (2019). Two-level parallel augmented schur complement interior-point algorithms for the solution of security constrained optimal power flow problems. *IEEE Transactions on Power Systems, 35*(2), 1340–1350.
150. Ali, A., Keerio, M. U., & Laghari, J. A. (2020). Optimal site and size of distributed generation allocation in radial distribution network using multi-objective optimization. *Journal of Modern Power Systems and Clean Energy, 9*(2), 404–415.
151. Fu, X., Chen, H., Cai, R., & Xuan, P. (2016). Improved LSF method for loss estimation and its application in DG allocation. *IET Generation, Transmission & Distribution, 10*(10), 2512–2519.
152. Fu, X., Sun, H., Guo, Q., Pan, Z., Xiong, W., & Wang, L. (2017). Uncertainty analysis of an integrated energy system based on information theory. *Energy, 122*, 649–662.
153. Fu, X., Li, G., Zhang, X., & Qiao, Z. (2018). Failure probability estimation of the gas supply using a data-driven model in an integrated energy system. *Applied Energy, 232*, 704–714.
154. Li, J., Wei, H., & Mo, D. (2012). Asymptotically optimal scenario analysis and wait-and-see model for optimal power flow with wind power. *Zhongguo Dianji Gongcheng Xuebao (Proceedings of the Chinese Society of Electrical Engineering), 32*(22), 15–23.
155. Fu, X., Guo, Q., Sun, H., Pan, Z., Xiong, W., & Wang, L. (2017). Typical scenario set generation algorithm for an integrated energy system based on the Wasserstein distance metric. *Energy, 135*, 153–170.
156. Su, C., Liu, C., Jiang, S., & Wang, Y. (2021). Probabilistic power flow for multiple wind farms based on RVM and holomorphic embedding method. *International Journal of Electrical Power & Energy Systems, 130*, 106843.
157. Fu, X. (2019). Recommended air conditioner temperature based on probabilistic power flow considering high-dimensional stochastic variables. *IEEE Access, 7*, 133951–133961.
158. Hu, T., Wu, W., Guo, Q., Sun, H., Shi, L., & Shen, X. (2019). Very short-term spatial and temporal wind power forecasting: A deep learning approach. *CSEE Journal of Power and Energy Systems, 6*(2), 434–443.
159. Fu, X., Guo, Q., Sun, H., Zhang, X., & Wang, L. (2017). Estimation of the failure probability of an integrated energy system based on the first order reliability method. *Energy, 134*, 1068–1078.
160. Fu, X., Li, G., & Wang, H. (2018). Use of a second-order reliability method to estimate the failure probability of an integrated energy system. *Energy, 161*, 425–434.
161. Fu, X., Zhang, X., Qiao, Z., & Li, G. (2019). Estimating the failure probability in an integrated energy system considering correlations among failure patterns. *Energy, 178*, 656–666.
162. Fu, X., & Zhang, X. (2018). Failure probability estimation of gas supply using the central moment method in an integrated energy system. *Applied Energy, 219*, 1–10.
163. Li, J., Wen, J., Cheng, S., & Wei, H. (2013). A scene generation method considering copula correlation relationship of multi-wind farms power. *Proceedings of the CSEE, 33*(16), 30–36.
164. Ghahramani, Z. (2015). Probabilistic machine learning and artificial intelligence. *Nature, 521*(7553), 452–459.