Probabilistic Revenue Analysis of Microgrid Considering Source-Load and Forecast Uncertainties

YANG YANG, YUANFAN JI, GUANGCHAO GENG, AND QUANYUAN JIANG

1College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China
2Polytechnic Institute, Zhejiang University, Hangzhou 310027, China

Corresponding author: Guangchao Geng (ggc@zju.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 51977188.

ABSTRACT Due to the randomness of load and renewable energy generation (REG), microgrids face multiple uncertainties. These uncertainties lead to the uncertainty of microgrid operation and bring more challenges to the economic evaluation of microgrids. In this paper, an economic evaluation method for determining microgrid revenue distribution is proposed. Considering the dual uncertainties of source-load and forecast, and temporal autocorrelation of time series, the probabilistic model of uncertainties is established by multivariate kernel density estimation (KDE). Then the random scenarios including forecasting values are generated and used in optimal dispatch calculation for the detailed production simulation. The probabilistic revenue is derived with a method based on Monte Carlo method. Finally, a case study is carried out based on the real data of an industrial park. The results demonstrate the necessity and effectiveness of the probabilistic revenue analysis proposed in this paper. This method can reveal the actual values of each component of a microgrid (e.g., device or algorithm) in specific scenes and provides more insights into investment decisions.

INDEX TERMS Autocorrelation, cost-benefit analysis, microgrid, Monte Carlo method, production simulation, uncertainty.

NOMENCLATURE

BOS Balance of system.
CUC Component use costs.
DC Demand charges.
DER Distributed energy resource.
DG Distributed generation.
EC Energy charges.
ESS Energy storage system.
KDE Kernel density estimation.
LCOE Levelized cost of energy.
MAPE Mean absolute percentage error.
PCS Power conversion system.
PDF Probability density function.
PV Photovoltaic.
REG Renewable energy generation.
ROI Return on investment.
RMSE Root-mean-square error.
SOC State of charge.
TC Total costs.

I. INTRODUCTION

Microgrids are the systems that integrate distributed energy resources (DERs), energy storage systems (ESSs), and flexible loads [1]. In recent years, microgrids have been increasingly used to improve local reliability, reduce costs, and promote renewable energy consumption [2]. However, due to the high price of DERs and ESSs, economic viability has become an essential issue in microgrids’ large-scale applications. Cost-benefit analysis is an essential means to provide decision support for the planning, reconstruction, and operation of microgrids.

Most conventional planning tools in power systems are based on deterministic cost-benefit analysis. That is, evaluation and decision-making are based entirely on historical data. QuESt is an open-source software for energy storage simulation and analysis which can estimate the maximum revenue from participating in energy arbitrage or providing ancillary...
services [3]. HOMER software performs a one-year hourly simulation of microgrids to help users design a microgrid according to the costs [4]. References [5], [6] use load duration curve to perform production simulation and economic analysis. Life-cycle-cost analysis is often used to compare the economic feasibility of projects [7]. In [8], an optimal design model of microgrids is proposed along with three economic evaluation indicators, which are combined to comprehensively evaluate the microgrids’ economic performances and investment risk. Deterministic cost-benefit analyses rely upon some deterministic assumptions about uncertain issues, such as load demand and economic parameters. If the costs and benefits are sensitive to some influencing factors, the results will have fortuities and limitations.

In power system studies, uncertainty has always been a topic of interest. With the development of smart grids, renewable energy usage, and load diversification, the influence of uncertainty has become increasingly prominent. The common methods of dealing with power system uncertainties can be classified into GUM (guide to the expression of uncertainty in measurement) approaches, probabilistic approaches, and non-probabilistic approaches. Some distribution models, such as normal, Gamma and Weibull distribution, are often used to fit the distribution of samples. In [9], it is believed that the probability distributions of load and wind power in power systems obey the normal distribution centered on the day-ahead forecast, and the standard deviation is related to the average forecast accuracy. When there is not enough prior system knowledge, kernel density estimation (KDE) is more suitable for uncertainty quantification [10]. In [11], a data-driven temporal-dependency Haar expansions approach is used to quantify the household energy demand. In [12], information gap decision theory (IGDT) is employed to model the load uncertainty. There is no unified conclusion on which approach to adopt to model uncertainties, which depends on the situation. However, many existing methods are too extensive to take data’s detail characteristics into account. For example, the temporal autocorrelation and nonstationarity of time series cannot be considered. In this study, the modeling method based on multivariate KDE can solve the above difficulties.

Probabilistic analysis can incorporate uncertainty factors and give the statistical distribution of results, which is widely used in power systems [13]–[18]. In recent years, probabilistic analysis is applied to cost-benefit analysis. In some studies, the costs and benefits are modeled as Gaussian variables or uniform variables [19]–[21]. The parameters of variable distribution (e.g., mean, standard deviation, and correlation coefficient) are estimated according to historical data. A recognized defect of this method is the accuracy of the statistical modeling of the cost and benefit. When the probability density function (PDF) of input variables is simple, the results of probabilistic analyses can be obtained by system theoretic methods [21]. For cases of increasing complexity, Monte Carlo simulation is the more commonly used solution technique [21]–[24]. In [22], considering the uncertainties of unit capital costs and power demand, the cost uncertainty for different generation portfolios is qualified. Reference [23] considers the uncertainties of inputs and the endogeneity between inputs when calculating the levelized energy cost of a nuclear and gas power project. Under the uncertainty of energy prices, the optimal dispatch calculation is carried out to determine the economic value of the energy hub [24]. The details of power system operation, such as unit commitment and power reserve, are not considered in most of the above studies. Besides, when solving optimal dispatch, all studies assume perfect foresight of the future data (e.g., knowing the real load for the next day in advance), which is inconsistent with the facts.

The demand and supply of microgrids are more uncertain compared with the bulk power system. On the other hand, microgrids have more operational flexibility. Probabilistic cost-benefit analysis is more challenging and meaningful to microgrids. In this paper, with the temporal autocorrelation considered, the dual uncertainties of source-load and forecast are modeled by multivariate KDE. Since the uncertainty of forecasting errors has been quantified, the random scenarios with forecasting values can be generated and used in optimal dispatch calculation for the detailed production simulation. When using Monte Carlo simulation to calculate the revenue distribution, sampling from the scenario set reduces the computational time to some extent.

Compared with existing state-of-the-art, the main contributions of this study are as follows:

i) The potential information in historical data is fully utilized by modeling the dual uncertainties of source-load and forecast using multivariate KDE. The generated random scenarios with forecasting values are close to reality, so that we can simulate microgrid operation in detail and get more accurate cost-benefit analysis results.

ii) The proposed economic evaluation method can reveal the actual values of each component of a microgrid (e.g., device or algorithm) from the perspective of probability, which provides more insights into economic evaluation.

This paper is organized as follows: In Section II, the uncertainties involved in this paper are quantified by probabilistic models. Section III describes the basic issues and steps of probabilistic revenue analysis. Section IV presents the results of a use case. Section V concludes the paper.

II. PROBABILISTIC MODELING OF UNCERTAINTIES

For a microgrid with a given configuration and energy management strategy, the uncertainties of economic dispatch and operating revenue stem from the uncertainties of load and renewable energy generation (REG) such as PV and wind power. The historical load and REG reflect the basic situation of a microgrid’s load demand and renewable energy resources. However, historical data is only a possible scenario
that cannot reappear. Only using historical data to analyze the operating revenue of microgrids will ignore many possible situations.

Economic dispatch of microgrids based on the forecasting results of load and REG is helpful to optimize the use of dispatchable distributed generation (DG) units and ESSs, which has been widely adopted in many studies and practical projects [25]–[28]. For a determined load demand and renewable energy generation power, the accuracy of forecasting will affect the final dispatching result. Considering this, the forecast of load and REG is also the reason for the uncertain microgrid revenue.

Therefore, the uncertainties include two aspects: source-load uncertainty and forecast uncertainty. Reasonable modeling of theses uncertainties is the basis of probabilistic revenue analysis. In this section, the uncertainties are quantified by probabilistic models.

A. SOURCE-LOAD UNCERTAINTY

In microgrids, the daily load profiles are mainly affected by the production and living behaviors, presenting several typical patterns. The daily REG profiles are mainly affected by weather conditions, which also have several typical patterns. Different patterns of load/REG have different characteristics (e.g., forecasting error distributions). Therefore, classifying load/REG according to profile similarity can simplify the subsequent analysis.

Although it is impossible to know the load/REG pattern of every specific day in advance, each load/REG pattern’s probability can be obtained according to historical data. The K-means clustering algorithm is used to identify each historical day’s load/REG pattern [29], and then the proportion of each load/REG pattern can be calculated. The number of clusters (i.e., the number of patterns) is optimized by the Calinski-Harabasz index [30]. These days with the same load pattern and REG pattern are said to be of the same day type. Considering the independence of load and REG, we can get the occurrence probability of each day type.

For example, historical data of a season are selected for analysis, including m kinds of load patterns and n kinds of REG patterns. After obtaining the proportions of each load/REG pattern in historical days, we can calculate the occurrence probabilities of $m \times n$ kinds of day types:

$$p_k = \alpha_{Li} \times \alpha_{REGj}$$

(1)

where $p_k$ is the probability that a day of the season belongs to the $k$th day type; $\alpha_{Li}$ and $\alpha_{REGj}$ are the proportions of the $i$th load pattern and the $j$th REG pattern; $i \in 1 : m, j \in 1 : n, k \in 1 : m \times n$.

Even in the same load/REG pattern, the load/REG at each time step will show randomness affected by some random factors. To describe this randomness, KDE is used to establish the ambiguous distribution of load/REG in the same pattern at each time step. KDE is a nonparametric estimation method, which has a wide range of applications. It studies the data distribution characteristics from the data sample itself, which do not rely on any data distribution assumptions. Because the kernel function and bandwidth can be chosen flexibly, the distribution of load/REG at each time step can be well fitted.

Let $x_1, x_2, \ldots, x_n$ be an independent and identically distributed sample drawn from some distribution with an unknown density $f$. Its kernel density estimator is [31]:

$$\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^{n} K_h(x - x_i) = \frac{1}{nh} \sum_{i=1}^{n} K(\frac{x - x_i}{h})$$

(2)

where $K$ is the kernel function, and the normal kernel is often used due to its convenient mathematical properties; $h$ is a smoothing parameter called the bandwidth.

Since the load/REG sequence is continuous and usually will not change drastically, it has strong autocorrelation [32]–[34]. Autocorrelation is the correlation of values of a time series at different time steps. The correlation strength can be described by Pearson correlation coefficient (also called Pearson’s r). Given paired data $(x_1, y_1), \ldots, (x_n, y_n)$ consisting of n pairs, Pearson correlation coefficient is defined as:

$$r_{xy} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \sum_{i=1}^{n} (y_i - \bar{y})^2}}$$

(3)

where $r_{xy}$ is Pearson correlation coefficient between variables $x$ and $y$; $\bar{x}$ and $\bar{y}$ are the sample mean.

In order to simplify the model, only the correlation between the load/REG at two consecutive time steps is considered. Therefore, it is necessary to introduce the multivariate KDE theory which is used to establish joint PDF.

Let $x_1, x_2, \ldots, x_n$ be samples of d-variate random vectors drawn from a common distribution described by the density function $f$. The kernel density estimate is defined to be [31]:

$$\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^{n} K_h(x - x_i)$$

(4)

where $H$ is the bandwidth $d \times d$ matrix which is symmetric and positive definite; $K$ is the kernel function which is a symmetric multivariate density.

Use the standard multivariate normal kernel throughout:

$$K_H(x) = (2\pi)^{-d/2} |H|^{-\frac{1}{2}} e^{-\frac{1}{2} x^T H^{-1} x}.$$  

(5)

Bandwidth $h$ or bandwidth matrix $H$ has a great influence on the estimation results. Bandwidth or bandwidth matrix can be selected based on minimizing the mean integrated squared error (MISE) or directly using some rules of thumb [35].

B. FORECAST UNCERTAINTY

The defects of forecasting models, the influence of unexpected events, the quality of input data, and other problems will cause forecasting errors. Compared with the bulk power system, the load and distributed REG in microgrids are more uncertain and more difficult to forecast. Therefore, it is critical to consider the influence of forecast in microgrid production simulation. We first need to model the uncertainty...
of forecasting errors so that the forecasting results of each scenario’s load/REG can be generated. In the same pattern and time step, the numerical results show that the actual values and forecasting errors of load/REG are weakly correlated. Under such conditions, we can assume that they are independent.

Some studies point out that the forecasting errors of short-term load, PV, and wind power are approximately Gaussian distribution [18]. In [36], the study examines the distribution of errors from operational forecasting systems in two different Independent System Operator (ISO) regions for both wind power and load forecasts; The conclusion is that the distribution of forecasting error is leptokurtic, and the bimodal distribution provides a better fit for it. The forecasting error is affected by many factors, such as the forecasting algorithm and application scenario, and its distribution model is inconclusive. In the microgrid used in the case study, it is difficult to describe the forecasting error of load/REG at each time step with a unified distribution model. Fig. 1 is the frequency distribution histogram of load forecasting errors in a pattern at a time step and the Gaussian fitting. It can be seen that the fitting effect with Gaussian distribution is not good. At some time periods, the forecasting errors may even show multi-peak distribution.

Since KDE does not need to assume the distribution in advance, it is used to fit the distribution of load/REG forecasting errors. Besides, it is also found that the forecasting errors have a strong autocorrelation [37]. For example, if the load/REG is over-forecast for one hour, the error tends to persist in the next several hours. The most important reason is that forecasting errors mainly stem from the inability to accurately predict weather conditions and production plans, and the influence of these factors is continuous. Therefore, the joint distribution of forecasting errors at two consecutive time steps is also modeled by multivariate KDE.

III. PROBABILISTIC REVENUE ANALYSIS

The main steps of probabilistic revenue analysis of a microgrid are as follows: First, a series of random scenarios of load and REG are generated by sampling in time step order. Then, according to the energy management strategy, microgrid production simulation is performed in each generated scenario. The operation of each unit is simulated, and the microgrid’s operating cost is calculated. Finally, the revenue distribution of the microgrid is obtained by Monte Carlo simulation.

A. RANDOM SCENARIO GENERATION

When using Monte Carlo simulation to calculate the monthly revenue distribution of a microgrid, if the load and REG must be regenerated every time and the production simulation must be performed repeatedly, it will lead to high computational efforts. In fact, although there are countless scenarios, there are only slight differences between many scenarios, which have little impact on the operation and revenue of microgrids. We only need to generate enough random scenarios for a day type so that the revenue in these scenarios can represent the revenue in this day type. Enough here means that the generated scenarios obey the characteristics of the original data. The characteristics include probability distribution and autocorrelation, which can be tested by statistical methods [38]. Based on experience, the number is usually set to hundreds.

If there are M day types and set the number of each day type’s random scenarios to N, there are $M \times N$ scenarios. Different scenarios include different actual and forecasting values of load and REG. Based on the probabilistic models in Section II, the steps to generate random scenarios for each day type are as follows:

1. Randomly sample the PDF of the actual load/REG ($f(P^1)$) and the PDF of its forecasting error ($g(\Delta P^1)$) at the first time step, getting the simulated actual value ($P^1$) and forecasting error ($\Delta P^1$).
2. Given the simulated load/REG ($P^{t-1}$) and the simulated value of forecasting error ($\Delta P^{t-1}$) at time step $t - 1$ ($t \geq 2$), the conditional PDF of the actual value at time step $t$ is calculated by the joint PDF of the actual values at time step $t$ and $t - 1$ ($f(P^t, P^{t-1})$); The conditional PDF of the forecasting error at time step $t$ is calculated by the joint PDF of the forecasting error at time step $t$ and $t - 1$ ($g(\Delta P^t, \Delta P^{t-1})$):

$$f(P^t | P^{t-1}) = \frac{f(P^t, P^{t-1})}{f(P^{t-1})}$$

$$g(\Delta P^t | \Delta P^{t-1}) = \frac{g(\Delta P^t, \Delta P^{t-1})}{g(\Delta P^{t-1})}$$

Random sampling is carried out on the above distribution to obtain the simulated actual value and forecasting error of load/REG at time step $t$. Sampling is carried out in time step order until the last time step.

3. The simulated actual values ($P = [P^1, P^2, \ldots, P^{nT}]$) and forecasting errors ($\Delta P = [\Delta P^1, \Delta P^2, \ldots, \Delta P^{nT}]$) of the whole day are obtained. Then the day-ahead forecasting result is:

$$P' = P + \Delta P$$

4. Step 1-3 is repeated $N$ times to obtain the $N$ random scenarios of the day type.
B. MICROGRID PRODUCTION SIMULATION

Under the uncertainties of load demand and REG, a double-layer energy management model is used to simulate the operation and cost of a microgrid. This energy management strategy is based on the double-layer coordinated control framework in [25], and some improvements are made for demand charge management. Demand charges are fees charged by utilities based upon the peak power demand during a month. To ensure stable operation and minimize the operating costs of microgrids, the strategy is divided into two steps, including day-ahead scheduling and real-time dispatching.

In day-ahead scheduling, the whole day is divided into several time steps. Considering the technical and economic constraints, the scheme for the next day is made based on the load and REG forecasting results. In order to minimize a microgrid’s operating cost, the day-ahead scheduling is formulated as a mixed-integer linear programming (MILP), and it is solved by CPLEX [39].

The objective function is given as follows:

\[
\min f(x, u) = f_G(x, u) + f_B(x, u) + f_{grid}(x, u)
\]

\[
f_G(x, u) = \sum_{t \in \theta_T} \sum_{i \in \theta_G} (c^G_{Gi} P^i_{Gi} + c^m_G M^i_{Gi})
\]

\[
f_B(x, u) = \sum_{t \in \theta_T} \sum_{i \in \theta_B} \left[ c_{Bi} (P^i_{Bi+} + P^i_{Bi-}) 
+ c_{Bi}^c (M^i_{Bi+} + M^i_{Bi-}) \right]
\]

\[
f_{grid}(x, u) = \sum_{t \in \theta_T} \left( c^f_{grid+} P^f_{grid+} + c^f_{grid-} P^f_{grid-} \right) + c^d_{dc} D_{grid}
\]

where \(f_G\) is the cost of dispatchable DG units, including fuel costs and startup costs; \(f_B\) is the wear cost of ESSs; \(f_{grid}\) is the utility bill charged by public grids, which includes two parts: electricity consumption charges and demand charges; \(P^i_{Gi}\) is the output power of dispatchable DG units; \(P^i_{Bi+}\) and \(P^i_{Bi-}\) are the discharge and charge power of ESSs; \(P^f_{grid+}\) and \(P^f_{grid-}\) are the power bought from and sold to public grids, respectively; \(M^i_{Gi}\) is the sign of change in dispatchable DG units’ operation state; \(M^i_{Bi+}\) and \(M^i_{Bi-}\) are the sign of change in ESSs’ discharging and charging states; \(D_{grid}\) is the maximum demand variable; \(c^d_{dc}\) is demand charge rate converted to a day; \(c^f_{Gi}, c^f_{Gi}^+\), and \(c^f_{grid+}\) are the corresponding cost coefficients; \(\theta_T, \theta_G\) and \(\theta_B\) represent the sets of all time steps, dispatchable DG units and ESSs.

In real-time dispatching, the dispatching plan for the next time step is calculated based on the real-time data. The dispatching plan should compensate for the forecasting errors while following the day-ahead scheme so that the dispatching can follow the economic operation scheme without affecting microgrids’ stable operation. The real-time dispatching is formulated as a quadratic programming (QP) and solved by IPOPT [40], aiming at minimizing the deviation between the dispatching plan and the day-ahead scheme.

The objective function is given as follows:

\[
f(x) = \sum_{i \in \theta_G} \mu_G (P^i_{Gi} - \hat{P}^i_{Gi})^2 + \sum_{i \in \theta_B} \mu_B (P^i_{Bi} - \hat{P}^i_{Bi})^2 
+ \mu_{grid} (P_{grid} - \hat{P}_{grid})^2
\]

where \(P^i_{Gi}, P^i_{Bi}\) and \(P_{grid}\) are the output power of dispatchable DG units, ESSs, and public grids; \(\hat{P}^i_{Gi}, \hat{P}^i_{Bi}\) and \(\hat{P}_{grid}\) are the output power at the current time step in the day-ahead scheme; \(\mu_G, \mu_B\) and \(\mu_{grid}\) are the deviation penalty factors set artificially, which affect the sensitivity of each variable deviating from the day-ahead scheme.

In [25], the exchange power between microgrids and public grids has a fixed upper limit value limited by the transformer’s capacity. In contrast, the upper limit value in this model’s day-ahead scheduling stage is set as an optimization variable to reduce demand charges. In real-time dispatching stage, the optimized demand limit needs to be followed and it can be automatically enlarged when the problem fails to be solved.

In this strategy, ESSs need to satisfy the state of charge (SOC) constraints and maximum charge/discharge power constraints. The setting of cycle costs in \(f_B(x, u)\) avoids frequent changes of ESSs’ charging and discharging states. There are also constraints such as the power balance equation and units’ technical constraints to be considered in the model. It should be noted that the proposed probabilistic revenue analysis applies to not only this energy management strategy but also any other forecast-based dispatch strategy.

C. REVENUE ESTIMATION USING MONTE CARLO SIMULATION

The microgrid operation and costs under each random scenario are obtained by substituting the generated scenarios into the production simulation program. Since demand charges are accounted monthly, the costs and benefits of microgrids are calculated on a monthly basis. Monthly total costs (TC) include energy charges (EC), demand charges (DC), and component use costs (CUC):

\[
TC = EC + DC + CUC
\]

\[
EC = \sum_{d=1}^{D} (C_{ec,d} - B_{fit,d})
\]

\[
DC = c_{dc} \times P_{max_{grid+}}
\]

\[
CUC = \sum_{d=1}^{D} (C_{B,d} + C_{REG,d} + C_{G,d})
\]

where \(D\) is the number of days in the month; \(C_{ec,d}\) is the electricity consumption charge of day \(d\); \(B_{fit,d}\) is the income from feed-in tariff (FIT) of day \(d\); \(c_{dc}\) is the monthly demand charge rate; \(P_{max_{grid+}}\) is the maximum amount of power purchased from public grids during the month; \(C_{B,d}, C_{REG,d}\) and \(C_{G,d}\) are life cycle costs of ESSs, REG, and dispatchable DG units in day \(d\), which are calculated based on levelized cost of energy (LCOE).
The monthly net revenue of a microgrid investment is defined as:

$$ R = TC_b - TC_a $$  \hspace{1cm} (18)$$

where $TC_b$ and $TC_a$ are the monthly total cost before and after the investment, respectively.

The Monte Carlo method is a repeated solution of a given problem by random sampling the inputs, which is used to calculate the monthly revenue distribution of microgrids. The process is shown in Fig. 2. At each iteration of simulations, the day type of each day in the month is determined according to each day type’s occurrence probability. Thus, the number of days for each day type in a month is determined, which obeys the multinomial distribution. Then, the corresponding days are randomly selected from the corresponding random scenario set, and the microgrid’s monthly revenue is obtained. Through a large number of iterations, the revenue tends to converge in distribution.

### IV. CASE STUDY

#### A. BASIC DATA

The typical microgrid studied in the case study is located in an industrial park in Shanghai, China. The microgrid is connected with the public grid through a 1600-kV A transformer, which contains PV generators of 150 kWp and is equipped with a Li-ion ESS of 150 kW/500 kWh. The peak load demand in summer is about 650 kW. The schematic of the microgrid is shown in Fig. 3. The analyses are based on the historical and forecasting data in the summer of 2019 (from July to September) with an interval of 15 minutes.

The time-of-use electricity prices in summer are shown in Table 1. In addition, the monthly demand charge rate is 42 CNY/kW (1 CNY = 0.155 USD).

The basic technical and economic parameters of the ESS are shown in Table 2. The investment cost of the ESS is divided into two parts: energy cost and power cost, which are proportional to rated energy and rated power, respectively. Energy cost mainly relates to the cost of battery packs. For commercial Li-ion ESS, the power cost mainly includes power conversion system (PCS) cost and balance of system (BOS) cost, accounting for a large part of the system cost [41].

#### B. CHARACTERISTICS OF HISTORICAL DATA

The historical daily load are clustered into three patterns by K-means clustering: “high level”, “medium level” and “low level”, which are mainly determined by the daily production schedule. The daily PV power are clustered into two patterns: “high level” and “low level”, which are determined by daily
solar resources. The historical data and typical profiles of three load patterns and two PV power patterns are shown in Fig. 4. Three load patterns and two PV power patterns constitute six different day types. The proportion of each load pattern and PV power pattern, and the occurrence probability of each day type are shown in Table 3.

Whether actual values should be considered in the forecasting error modeling depends on the correlation between them. According to historical data, after distinguishing the pattern and time step, Pearson correlation coefficients between load forecasting errors and actual values in more than 90% of time steps are less than 0.5. For PV, the ratio is 70%. In general, Pearson correlation coefficients whose magnitude are less than 0.5 indicate that variables have a low correlation or little correlation. Therefore, we assume that the forecasting errors and the actual values are independent to simplify the model.

In order to validate the autocorrelation of load/PV power and its forecasting errors, the correlation coefficient is calculated. Table 4 shows the Pearson correlation coefficient between the load/PV power at two consecutive time steps and between their forecasting errors. In general, two random variables can be considered highly correlated if the absolute value of Pearson’s $r$ is greater than 0.7. The results show that the load and the PV power and their forecasting errors all have significant autocorrelation. We divide the whole day into 96 time steps, and the data of each time step is regarded as a random variable. Therefore, the $96 \times 96$ correlation coefficient matrix can be obtained as shown in Fig. 5 in the form of heat maps. Most of the elements near the main diagonal are close to 1, which indicates a strong correlation between the actual values/forecasting errors of adjacent time steps, and the closer the time steps are, the stronger the correlation is. It proves that it is necessary to consider the autocorrelation of both actual values and forecasting errors when establishing the probabilistic models.

The generated scenarios not only retain the inherent characteristics of load/PV power and its forecasting errors but also embodies the randomness, which increases the reliability of production simulation results. By using the probabilistic models in this paper, hundreds scenarios and the corresponding day-ahead forecasting results are generated. For example, the simulated “high level” load scenarios are shown in Fig. 6.

### C. ESS VALUATION

Probabilistic revenue analysis can be used to evaluate the values of each component of a microgrid in specific scenes. In this part, taking the ESS as an example, the value of the ESS is quantified by calculating the difference between microgrid monthly total cost with and without ESS. In this case, the ESS mainly contributes to the microgrid in the following two ways:

#### TABLE 3. Probability of load/PV patterns and day types.

| Load       | PV          | High level (64.5%) | Low level (35.5%) |
|------------|-------------|--------------------|-------------------|
| High level | Day type 1: 18.7% | Day type 4: 10.3%  |
| Medium     | Day type 2: 29.2% | Day type 5: 16.0%  |
| Low level  | Day type 3: 16.6% | Day type 6: 9.2%   |

#### TABLE 4. Autocorrelation of actual values and forecasting errors.

|                  | Pearson correlation coefficient |            |
|------------------|--------------------------------|------------|
|                  | actual values                  | Forecasting Errors |
| Load             | 0.945                          | 0.856      |
| PV power         | 0.914                          | 0.850      |

#### FIGURE 4. Historical load/PV data and pattern division. a) Daily load profiles. b) Daily PV power profiles.

#### FIGURE 5. Heatmaps of correlation coefficient matrix. a) Actual load. b) Actual PV power. c) Load forecasting error. d) PV power forecasting error.
- Energy arbitrage: charging at the off-peak period of TOU or when there is PV surplus, and discharging at the peak period;
- Reducing demand charges: discharging at demand peak to reduce peak power demand.

Due to the lack of actual data for a whole year, July is selected as the target for analysis. In summer, the peak load is higher, and the difference between peak and valley electricity prices is larger, which is beneficial for ESS to realize its value in demand charge management and energy arbitrage. It should be noted that the proposed probabilistic revenue analysis method is universal. Although the analysis results of the whole year may be more convincing, the results of a single season are enough to demonstrate the necessity and effectiveness of the method.

Fig. 7 visually presents the calculation results. Through probabilistic revenue analysis, we can get some conclusions that are hard to find in deterministic economic analysis. Considering the uncertainties of source-load and forecast, DC savings and net revenue fluctuate considerably, while the fluctuation of EC savings is slight. It means the uncertainties of load and PV power have a significant influence on demand charge management, but little influence on energy arbitrage in this case.

Several revenue statistics indicators are shown in Table 5. Mean revenue is the average of 10000 (enough to make the revenue converge in distribution) simulations. This value can be approximately regarded as the mathematical expectation of revenue, so it has great reference value for the long-term benefit that ESS can gain under the uncertainties of load and REG. The probability that the actual revenue is higher than conservative revenue is 90%. It means that the revenue is less likely to be lower than 7300. Optimistic revenue is the revenue in the 90th percentile, which is the optimistic estimate of the revenue. Compared with the maximum revenue with extremely low possibility, this value is more suitable as the maximum achievable revenue. These statistical indicators quantify the value of ESS under the dual uncertainties of source-load and forecasting, which provide more insights for users when they evaluate ESS.

D. ESS SIZING BASED ON PROBABILISTIC REVENUE ANALYSIS

For ESS with fixed rated energy, the higher the rated power, the better it can play the role of demand reduction and energy arbitrage. However, with the increasing cost of ESS, the benefits it brings cannot increase indefinitely, which will lead to the decline of its economic efficiency. To evaluate the rationality of the current ESS’s rated power configuration, four ESSs with the same rated energy and different rated power are selected for comparison. The initial investment costs and probabilistic revenue analysis results of four ESS configurations are shown in Table 6.

According to the analysis results, mean DC savings gradually increase with the increase of ESS rated power. In contrast, the extra DC savings brought by the increase of rated power from 200 kW to 250 kW significantly decline. This is because with the increase of demand reduction degree, the required ESS capacity becomes larger and larger, and the demand reduction is limited by the rated energy of ESS. The difference of EC savings in the four cases is slight, indicating that the rated power of ESS has little influence.
on energy arbitrage. With the rated power of ESS increasing from 200 kW to 250 kW, the mean net revenue hardly changes and has a downward trend. At this time, it is not cost-effective to increase the rated power of ESS, which is consistent with the previous judgment.

The probabilistic analysis provides more comprehensive results and various analysis perspectives. When comparing the economic efficiency of different investments under uncertainty, the investment costs, the expected value and fluctuation degree of revenue usually need to be considered at the same time. Therefore, annual return on investment (ROI) and utility function in the field of investment are introduced, which are defined as:

\[
ROI = \frac{R_y}{C_i} \times 100\% \tag{19}
\]

\[
U = \mathbb{E}(ROI) - \frac{1}{2} A \sigma^2 \tag{20}
\]

where \(R_y\) is annual net revenue; \(C_i\) is the initial investment cost; \(U\) is the utility function; \(\mathbb{E}(ROI)\) and \(\sigma^2\) are the expected value and variance of \(ROI\); \(A\) is the coefficient of risk aversion, which represents the investor’s degree of risk aversion.

In this paper, the utility function is used as the basis for selecting the optimal rated power of ESS. For risk averters, the value of \(A\) is positive. This means that under the same expected value of revenue, risk averters prefer schemes with smaller variance. For ordinary investors, the value of \(A\) is usually between 2 and 10, and the middle value of 6 is taken here. To highlight the difference between probabilistic revenue analysis and conventional deterministic method, the actual situations in July 2019 are simulated and taken as deterministic analysis results. The two types of results are shown in Table 7.

The best choice is case 3 based on probabilistic revenue analysis. However, based on the deterministic analysis, the best choice is case 2. The reason for the difference is that the ESS selected by probabilistic revenue analysis is the most suitable choice for thousands of random scenarios, while in deterministic analysis the ESS is selected only based on the unique historical situation. In case 2, the revenue calculated according to the historical situation is 10750 CNY, which is close to the optimistic revenue in probabilistic analysis. Therefore, in this case, using the results of deterministic economic analysis to guide the ESS sizing may lead to a high expectation of the revenue and ultimately fail to get the expected revenue. On the contrary, if the results of deterministic economic analysis are close to the conservative revenue, the results cannot fully reflect the benefits that this investment can bring. Overall, it is effective and necessary to carry out probabilistic revenue analysis considering source-load and forecast uncertainties in microgrids.

### E. IMPACT OF FORECAST ON MICROGRID REVENUE

To illustrate the impact of forecast on revenue analysis, we process the original load forecasting data of the industrial park (MAPE = 21.8%), and assume that there are three different load forecasting algorithms, of which the MAPE are 21.8%, 10%, and 0, respectively. Different load forecasting data will only affect the modeling of forecasting errors, and will not affect the actual load in production simulation.

Through the probabilistic revenue analysis, the distribution of monthly revenue obtained by ESS are shown in Fig. 8. As the load forecasts become more and more accurate, the revenue gradually moves to the right of coordinates and becomes more concentrated. That is to say, reducing the forecasting error can not only improve the revenue but also reduce the uncertainty of the revenue. The mean and standard deviation of the three cases’ revenue in Table 8 also reflect this phenomenon.
TABLE 8. Revenue distribution parameters.

| Case | Mean revenue (CNY) | Standard deviation of revenue (CNY) |
|------|-------------------|------------------------------------|
| MAPE=21.8%| 8990 | 1260 |
| MAPE=10%| 10300 | 580 |
| MAPE=0 | 10750 | 170 |

The case “MAPE = 0” can be regarded as the result of other probabilistic analysis methods without considering forecasts. It can be seen that forecasts have a great impact on the results of microgrid economic analysis. Therefore, it is unreasonable to ignore the impact of forecasts and use historical data for production simulation when analyzing the costs and benefits of microgrids. The above explains why this paper models the uncertainty of forecast and considers forecasts when analyzing the revenue.

Besides, when evaluating a forecasting algorithm, we usually use some evaluation indexes, such as MAPE, RMSE. However, the numerical value of these indexes may not reflect the value brought by the algorithm. Compared with these evaluation indexes, microgrid owners may be more concerned about the actual benefits brought by the improvement of forecasting accuracy. In this regard, the proposed probabilistic revenue analysis method also provides a perspective to evaluate the forecasting algorithm under the uncertainties of load and REG.

V. CONCLUSION

Aiming at the problem that a microgrid’s revenue cannot be accurately described under the randomness of load and REG, this study establishes the probabilistic models considering the dual uncertainties of source-load and forecast to generate random scenarios, and proposes the probabilistic revenue analysis method. Through the simulation and numerical analysis, the following conclusions can be drawn:

1) The actual values and forecasting errors of load/REG in microgrids have strong temporal autocorrelation. Preserving the autocorrelation when probabilistic modeling can increase the fidelity of the generated scenarios, which contributes to better estimation of microgrid revenue;

2) The uncertainties of source-load and forecast cannot be ignored, since they make a microgrid’s operating revenue fluctuate within a certain range. Conventional deterministic cost-benefit analyses have fortuities and limitations, so it is necessary to introduce probabilistic revenue analysis;

3) The proposed probabilistic revenue analysis method quantifies the revenue from microgrid operation and assists the sizing of ESS. This method can reveal the actual values of each component of a microgrid in specific scenes and be used to guide investment decisions.

REFERENCES

[1] N. Hatzigiourgiou, H. Asano, R. Iravani, and C. Marnay, “Microgrids,” IEEE Power Energy Mag., vol. 5, no. 4, pp. 78–94, Jul./Aug. 2007.

[2] M. Stadler, G. Cardoso, S. Mashayekh, T. Forget, N. DeForest, A. Agarwal, and A. Schönbein, “Value streams in microgrids: A literature review,” Appl. Energy, vol. 162, pp. 980–989, Jan. 2016.

[3] R. Concepcion. (May 2021). SNL-Quest. [Online]. Available: https://github.com/snl-quest/snliquest

[4] U. Sureshkumar, P. S. Manoharan, and A. P. S. Ramalakshmi, “Economic cost analysis of hybrid renewable energy system using HOMER,” in Proc. IEEE Int. Conf. Adv. Eng., Sci. Manage. (ICAESM), Mar. 2012, pp. 94–99.

[5] Z. Yamaye and H. Hakimimashadi, “Production simulation for power system studies,” IEEE Trans. Power App. Syst., vol. PAS-104, no. 12, pp. 3376–3381, Dec. 1985.

[6] A. Poulin, M. Dostie, M. Fournier, and S. Sansregret, “Load duration curve: A tool for techno-economic analysis of energy solutions,” Energy Buildings, vol. 40, no. 4, pp. 29–35, Jan. 2008.

[7] D.-R. Thiam, “Renewable decentralized in developing countries: Appraisal from microgrids project in Senegal,” Renew. Energy, vol. 35, no. 8, pp. 1615–1623, Aug. 2010.

[8] M. Mao, P. Jin, L. Chang, and H. Xu, “Economic analysis and optimal design on microgrids with SS-PVs for industries,” IEEE Trans. Sustain. Energy, vol. 5, no. 4, pp. 1328–1336, Oct. 2014.

[9] W. Yang, C. Kang, Q. Xia, R. Liu, T. Tang, P. Wang, and L. Zhang, “Short-term probabilistic load forecasting based on statistics of probability distribution of forecasting errors,” Autom. Electr. Power Syst., vol. 30, no. 19, pp. 47–52, 2006.

[10] Q. Hou, N. Zhang, E. Du, M. Miao, F. Peng, and C. Kang, “Probabilistic duck curve in high PV penetration power system: Concept, modeling, and empirical analysis in China,” Appl. Energy, vol. 242, no. 1, pp. 205–215, May 2019.

[11] H. Shi, Q. Ma, N. Smith, and F. Li, “Data-driven uncertainty quantification and characterization for household energy demand across multiple time-scales,” IEEE Trans. Smart Grid, vol. 10, no. 3, pp. 3092–3102, May 2019.

[12] S. Nojavan, M. Majidi, and K. Zare, “Performance improvement of a battery/PV/fuel cell/grid hybrid system by considering load uncertainty modeling using IGIDT,” Energy Convers. Manage., vol. 147, pp. 29–39, Sep. 2017.

[13] H. Khorramdel, J. Aghaei, B. Khorramdel, and P. Siano, “Optimal battery sizing in microgrids using probabilistic unit commitment,” IEEE Trans. Ind. Informat., vol. 12, no. 2, pp. 834–843, Apr. 2016.

[14] R. Fernández-Blanco, Y. Dvorkin, and M. A. Ortega-Vazquez, “Probabilistic security-constrained unit commitment with generation and transmission contingencies,” IEEE Trans. Power Syst., vol. 32, no. 1, pp. 228–239, Jan. 2017.

[15] B. Rügen and D. Huang, “A subsystem based probabilistic approach for the assessment of transmission capital project reliability impacts,” Electr. Power Syst. Res., vol. 150, pp. 76–83, Sep. 2017.

[16] Y. Ding, C. Shao, J. Yan, Y. Song, C. Zhang, and C. Guo, “Economical flexibility options for integrating fluctuating wind energy in power systems: The case of China,” Appl. Energy, vol. 228, pp. 426–436, Oct. 2018.

[17] A. Hadian, M.-R. Haghigham, J. Zohrevand, and E. Akhavan-Rezai, “Probabilistic approach for renewable DG placement in distribution systems with uncertain and time varying loads,” in Proc. IEEE Power Energy Soc. Gen. Meeting (PES), Jul. 2009, pp. 1–8.

[18] X. Wen, W. Yan, L. Guo, Y. Li, J. Yu, X. Zhao, and H. Zhang, “A probability assessment method for energy-saving benefit of daily generation scheduling considering uncertainties of wind power and load,” Power Syst. Technol., vol. 38, no. 4, pp. 143–150, 2014.

[19] G. T. Heydt, “A probabilistic cost/benefit analysis of transmission and distribution asset expansion projects,” IEEE Trans. Power Syst., vol. 32, no. 5, pp. 4151–4152, Sep. 2017.

[20] J. Eng., vol. 2019, no. 16, pp. 680–685, Mar. 2019.

[21] G. T. Heydt, “The probabilistic evaluation of net present value of electric power distribution systems based on the Kaldor–Hicks compensation principle,” IEEE Trans. Power Syst., vol. 33, no. 4, pp. 4488–4495, Jul. 2018.

[22] P. Vithayasrichareon and I. F. MacGill, “A Monte Carlo based decision-support tool for assessing generation portfolios in future carbon constrained electricity industries,” Energy Policy, vol. 41, pp. 374–392, Feb. 2012.

[23] T. Geissmann and O. Ponta, “A probabilistic approach to the computation of the levelized cost of electricity,” Energy, vol. 124, pp. 372–381, Apr. 2017.
YANG YANG received the B.S. degree in electrical engineering from the College of Electrical Engineering, Zhejiang University, Hangzhou, China, in 2019, where he is currently pursuing the M.S. degree.

His research interests include micro-grid energy management and economic evaluation.

QUAN YUAN received the B.S. degree in electrical engineering from the College of Electrical Engineering, Zhejiang University, Hangzhou, China, in 1997, 2000, and 2003, respectively.

From 2012 to 2013, he was a Visiting Student at the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA, USA. From 2014 to 2017, he was a Postdoctoral Fellow at the College of Control Science and Engineering, Zhejiang University, and the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada. From 2017 to 2019, he was a Research Assistant Professor at the College of Electrical Engineering, Zhejiang University, where he is currently an Associate Professor. His research interests include power non-intrusive sensing technology, data analytics in power systems, and power system stability and control.

GUAN CHAO received the B.S. and Ph.D. degrees in electrical engineering from the College of Electrical Engineering, Zhejiang University, Hangzhou, China, in 2009 and 2014, respectively.

From 2012 to 2013, he was a Visiting Student at the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA, USA. From 2014 to 2017, he was a Postdoctoral Fellow at the College of Control Science and Engineering, Zhejiang University, and the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada. From 2017 to 2019, he was a Research Assistant Professor at the College of Electrical Engineering, Zhejiang University, where he is currently an Associate Professor. His research interests include power non-intrusive sensing technology, data analytics in power systems, and power system stability and control.