

Article

Probabilistic Forecasting of Available Load Supply Capacity for Renewable-Energy-Based Power Systems

Qizhuan Shao ¹, Shuangquan Liu ^{1,*} , Yigong Xie ¹, Xinchun Zhu ¹, Yilin Zhang ², Junzhou Wang ² and Junjie Tang ^{2,*} 

¹ System Operation Department, Yunnan Power Grid Co., Ltd., Kunming 650011, China; shaoqizhuan@yn.csg.cn (Q.S.); xieyigong@yn.csg.cn (Y.X.); zhuxinchun@yn.csg.cn (X.Z.)

² The State Key Laboratory of Power Transmission Equipment & System Security and New Technology, Chongqing University, Chongqing 400044, China; zhangyilin0109@163.com (Y.Z.); wangjunzhou@cqu.edu.cn (J.W.)

* Correspondence: liushuangquan@yn.csg.cn (S.L.); tangjunjie@cqu.edu.cn (J.T.)

Abstract: In order to accurately analyze the load supply capability of power systems with high penetration of renewable energy generation, this paper proposes a probabilistic available load supply capability (ALSC) forecasting method. Firstly, the optimal input features are selected by calculating the maximal information coefficient (MIC) between the input features and the target output. Based on this, a stacking ensemble learning model is applied for the prediction of wind power, photovoltaic power and load power. Secondly, the distributions of the forecasting objects are obtained based on forecasting errors and the error statistics method. Finally, the forecasting distributions of wind power, photovoltaic power and load are set as the parameters of a power system, and then probabilistic ALSC is calculated using Latin hypercube sampling (LHS) and repeated power flow (RPF). In order to simulate a more realistic power system, multiple slack buses are introduced to conduct two types of power imbalance allocations with novel allocation principles during the RPF calculation, which makes the ALSC evaluation results more reasonable and accurate. The results of probabilistic ALSC forecasting can provide a reference for the load power supply capacity of a power system in the future, and they can also provide an early warning for the risk of ALSC threshold overlimit. Case studies carried out on the modified IEEE 39-bus system verify the feasibility and effectiveness of the proposed methods.

Keywords: stacking ensemble learning model; probabilistic forecasting; repeated power flow; multi-slack buses; Latin hypercube sampling



Citation: Shao, Q.; Liu, S.; Xie, Y.; Zhu, X.; Zhang, Y.; Wang, J.; Tang, J. Probabilistic Forecasting of Available Load Supply Capacity for Renewable-Energy-Based Power Systems. *Appl. Sci.* **2023**, *13*, 8860. <https://doi.org/10.3390/app13158860>

Academic Editor: Pierluigi Siano

Received: 30 June 2023

Revised: 24 July 2023

Accepted: 26 July 2023

Published: 31 July 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Renewable energy, such as solar energy and wind energy, is clean and sustainable, and can relieve human dependence on fossil fuels. Therefore, it is necessary to study how renewable energy impacts the power system operation. As power grids continue expanding, uncertainties from load growth and massive integration of renewable energy have increased the randomness in power systems [1], which can deteriorate the security of the power system [2]. Therefore, against the background of large-scale integration of new energy generation into the grid, the security of the power system encounters serious challenges. The available load supply capability (ALSC) of the power system is an important indicator to measure the security of the power system. The power system must maintain sufficient ALSC to meet power demand, support socio-economic development, and overcome challenges such as extreme weather conditions. Consequently, it is critical to calculate accurately the available load supply capability (ALSC) of the power system, since an adequate load supply capability works as the foundation for guaranteeing the security of the power system, and accurate evaluation of ALSC can provide quantitative references for the load supply capability of the power system, and hence assist system operators develop

grid operation plans. The definition of ALSC is the maximum load growing percentage without violating any electrical constraint. ALSC evaluation can quantify the flexibility about the load growth and security margin of the power system. This information is important and can be helpful for reactive power compensation device placement, network reconfiguration, network planning, etc. [3].

There are some studies with regard to the calculation methods of the supply capability of power system. References [4,5] use the continuation power flow (CPF) to calculate the available transfer capability of power systems. References [6,7] evaluate the transfer capability by optimal power flow (OPF). CPF has a fast computing speed but lacks flexibility, and it is difficult to handle constraints and control strategies. OPF can consider the effects of various uncertainties of renewable energy but suffers from high computational complexity [8]. Therefore, in order to remain a good balance between the accuracy and feasibility for the ALSC of renewable-energy-based power systems, this paper adopts the repeated power flow (RPF) method [9,10], which has the advantages of simple calculation procedure and high flexibility.

RPF calculation is based on the multiple normal power flow calculations. In normal power flow calculation, only one slack bus is set up, and all power imbalance of the network is handled by this slack bus. When the power imbalance of the system is small, the power flow calculation model with a single slack bus will not affect the calculation results. However, when the power imbalance becomes large, it is not suitable to use the power flow calculation model with a single slack bus to calculate the power flow. Due to the limitation of the maximum output of generators, it is difficult for a single slack bus to bear all the power imbalance of the system. The continuously increasing load in RPF calculation process will bring a significant power imbalance. Therefore, a single slack bus is not suitable for the RPF calculation. The power flow calculation model considering multiple slack buses can make up for the above deficiencies [11,12]. It sets up multiple slack buses to share the power imbalance of the power system together, which can make the results of the power flow calculation fit the practical situation better. But according to the best knowledge of the authors, in existing studies on the RPF method for solving ALSC problem, some studies do not consider multi-slack buses [3]. In the studies that consider multi-slack buses, power imbalance allocation is only conducted once to share the power imbalance caused by load increment and the allocation principle is the ratio of the current power of the generators [13]. To simulate a more realistic power system operation, this paper proposes an RPF method with multi-slack buses that includes two types of power imbalance allocations and considers the economy of generators in the second power imbalance allocation, which is a novel principle of power imbalance allocation.

With wind power, photovoltaic power and other renewable energy generation integrated to the grid, the uncertainty factors in the power system have greatly increased. Therefore, the deterministic power flow is not suitable for solving the problem of probabilistic ALSC evaluation. Probabilistic power flow (PPF) can be introduced to address this issue. PPF can be generally divided into three methods: the Monte Carlo simulation (MCS) method [14], analytical method [15] and approximation method [16]. The analytical method has a high dependence on complicated theory. The approximate method has the disadvantage of low accuracy on the probability distribution. Therefore, the MCS method, with simple calculation and relatively high accuracy, is selected herein. The traditional random-sample Monte Carlo simulation (RS-MCS) achieves a high computational accuracy when the sample size is large enough, but it takes a long time and suffers from low computational efficiency. Therefore, it is usually used as a benchmark for verifying the accuracy of other methods. Latin hypercube sampling (LHS) can efficiently cover the entire distribution range of random input variables. It is advantageous at high sampling efficiency, good robustness and simple implementation. Thus, it has been widely applied in probabilistic power flow calculation and probabilistic ALSC analysis [3,13].

In the existing studies about probabilistic ALSC evaluation, the probability distributions of uncertainty sources, such as loads, wind power generators and photovoltaic power

generators are artificially preset as normal. This does not conform to the actual situation and cannot reflect the operating conditions of the future power grid. If the forecasting techniques are used to predict the probability distributions of these uncertainty sources in advance, based on which the ALSC analysis is conducted, then the forecasting result of ALSC can be obtained. The predicted ALSC can help power grid operators assess the security risk of future power systems, which brings an important guiding significance. Therefore, this is a meaningful, but neglected, problem.

In recent years, with the improvement in computing power and big data technology [17], artificial intelligence methods have been widely applied in the field of forecasting. Reference [18] proposes a sliding-window-based LightGBM model for the short-term load forecasting. Reference [19] applies XGBoost to predict photovoltaic production. However, the generalization ability and application scope of a single algorithm are quite limited. In order to further improve the forecasting accuracy and stability, some studies adopt the ensemble learning methods, which refer to the method of combining several different machine learning techniques into one forecasting model. Reference [20] combines the decomposition method and stacking-ensemble learning method to forecast wind power generation. Reference [21] proposes a feature-selective ensemble learning model for long-term regional PV generation forecasting. In addition, reference [22] points out that selecting the optimal model input features is also a critical step in the forecasting problem. Therefore, it is necessary to study a method to identify the critical inputs of the model.

Correspondingly, the main contributions of this paper are as follows:

- (1) The maximal information coefficient (MIC) is used to select appropriate model input features, and the stacking ensemble learning model is used to predict the wind power, photovoltaic power and load power. Both the MIC feature selection method and stacking ensemble learning model are beneficial for forecasting accuracy improvement.
- (2) An RPF method considering multi-slack buses is proposed to solve the ALSC evaluation problem. In the RPF process, multi-slack buses are applied for the two power imbalance allocations with novel allocation principles, which can make the RPF calculation match better with the actual situation, and hence obtain more accurate and reasonable ALSC evaluation results.
- (3) LHS is combined with the RPF to solve the probabilistic evaluation problem of ALSC. The probability distributions of uncertainty sources are determined by the forecasting results of renewable energy and load, so the probabilistic ALSC forecasting can be accomplished. Based on the probabilistic forecasting result of ALSC, the risk of ALSC for exceeding the threshold can be calculated, thus providing a reference for the security of power system in the future.

The framework of this paper is as follows. Section 2 introduces the model for wind power forecasting, photovoltaic power forecasting and load power forecasting, including feature selection method, stacking ensemble learning model and probabilistic forecasting method. Section 3 firstly introduces the definition of ALSC and the RPF method, considering the multi-slack buses, and secondly introduces the basic theories of LHS, which are fundamental for probabilistic ALSC analysis and the procedure of probabilistic ALSC forecasting. In Section 4, the performance of the proposed methods is verified through the use of real data sets and the modified IEEE 39-bus system.

2. Wind Power, Photovoltaic Power and Load Power Forecasting Model

2.1. Maximal Information Coefficient

The MIC can measure the degree of correlation between any two random variables. In this study, MIC is used to quantify the correlation between the input features and the true value of the target output. The core idea is that if two variables are correlated, dividing their scatter plot into grids allows calculating their mutual information (MI) based on the approximate probability distributions within each grid. The normalized MI can quantify the correlation between any two variables.

Given two random variables X and Y and their sample pairs $D = \{(x_i, y_i), i = 1, \dots, N\}$, where N is the number of samples, the calculation steps of MIC are as follows [22].

First, the sample space is divided into an $m \times n$ grid G . Second, the empirical joint probability density function $p(x, y)$ and empirical marginal probability density function $p(x)$ and $p(y)$ are estimated. Then, the MI is calculated as

$$MI(X, Y|D, G) = \sum_{x \in X} \sum_{y \in Y} p(x, y) \log_2 \left(\frac{p(x, y)}{p(x)p(y)} \right) \tag{1}$$

where $MI(X, Y|D, G)$ denotes the MI between X and Y with sample set D and grid G . There are various partitioning methods for the same grid size $m \times n$. Among all possible grids, the maximal MI is defined as

$$MI^*(D, m, n) = \max_G MI(X, Y|D, G) \tag{2}$$

where $MI^*(D, m, n)$ denotes the maximal MI.

Normalize the maximum MI into an interval $[0, 1]$, as follows:

$$NMI^*(D, m, n) = \frac{MI^*(D, m, n)}{\log \min\{m, n\}} \tag{3}$$

where $NMI^*(D, m, n)$ represents the normalized maximal MI.

Among all the grids satisfying $m \times n < k(N)$, the $NMI^*(D, m, n)$ can be calculated. Then the MIC is the maximal $NMI^*(D, m, n)$ over all possible grids, as follows:

$$MIC(X, Y) = \max_{m \times n < k(N)} \{NMI^*(D, m, n)\} \tag{4}$$

where $k(N)$ is normally set at $k(N) = N^{0.6}$ [23].

The value of MIC is between 0 and 1. The larger the MIC, the stronger the correlation between the two random variables.

2.2. Stacking Ensemble Learning Model

Stacking is an ensemble learning framework. Ensemble learning combines multiple learners to leverage their complementary strengths. It aggregates the outputs from multiple learners to achieve enhanced learning performance. The stacking ensemble learning model has a two-layer structure, as shown in Figure 1. The first layer uses multiple base learners to make predictions and obtains multiple sets of forecasting results. The second layer model is a meta-learner, which takes the outputs of the base learners as inputs and makes predictions again to obtain the final forecasting power. The base learners and meta-learner refer to different regression learning algorithms. As shown in Table 1, the base learners used in this study include the gradient boosting decision tree (GBDT) [24], XGBoost [25], the light gradient boosting machine (LGBM) [26], ridge regression (RR) [27], support vector regression (SVR) [28], and K-nearest neighbor (KNN) [29]. The meta-learner is linear regression (LR) [30]. The following subsections will briefly introduce the techniques used for the base learners and meta-learner, individually.

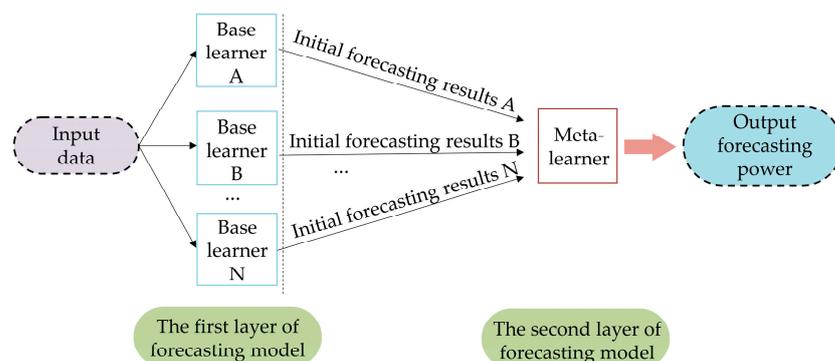


Figure 1. The structure of stacking ensemble learning model used for power prediction.

Table 1. The technique of the first and second layer of forecasting.

The first layer of forecasting (base learners)	GBDT	XGBoost	LGBM	RR	SVR	KNN
The second layer of forecasting (meta-learner)	LR					

Due to the higher forecasting accuracy of the stacking ensemble learning model compared to the single algorithm model, this paper selects the stacking ensemble learning model to predict wind power, photovoltaic power and load power. In addition, it is noteworthy that the base learners and meta-learner adopted in stacking ensemble learning may be different depending on different forecasting objects.

2.2.1. GBDT

GBDT is a boosting algorithm. Specifically, the residuals from prior decision trees are fed into succeeding ones, which are then combined progressively for the final forecasting results. Since GBDT makes good use of multiple weak learners in cascade, it has decent accuracy. But at the same time, due to the cascade relationship between weak learners, GBDT is difficult to train data in parallel. The calculation procedure of GBDT is as follows:

- (1) Given a sample set $\{(x_i, y_i) | x_i \in \mathbb{R}^n, y_i \in \mathbb{R}\}_i^N$, initialize the weak learner as follows:

$$f_0(x) = \operatorname{argmin}_c \sum_{i=1}^N \mathcal{L}(y_i, c) \tag{5}$$

where $\mathcal{L}(\cdot)$ is the loss function.

- (2) Calculate the negative gradient, namely the residual, which is defined as

$$r_i^m = - \left[\frac{\partial \mathcal{L}(y_i, f(x_i))}{\partial f(x_i)} \right]_{f(x)=f_{m-1}(x)}, i = 1, 2, \dots, N; m = 1, 2, \dots, M \tag{6}$$

where m denotes the number of iterations and M denotes the number of decision trees.

- (3) Take the residuals obtained in step 2 as the new sample labels, and take the dataset $\{(x_i, r_i^m) | x_i \in \mathbb{R}^n, r_i^m \in \mathbb{R}\}_i^N$ as the training data for the next decision tree, where \mathbb{R}_j^m is the j th leaf node region of the m th decision tree. Then the value of the j th leaf node region of the m th decision tree is as follows:

$$\mathcal{R}_j^m = \operatorname{argmin}_{\mathcal{R}} \sum_{x_i \in \mathbb{R}_j^m} \mathcal{L}(y_i, f_{m-1}(x_i) + \mathcal{R}), m = 1, 2, \dots, M; j = 1, 2, \dots, J \tag{7}$$

where J is the number of leaf node.

- (4) Update the strong learner as follows:

$$f_m(x) = f_{m-1}(x) + \sum_{j=1}^J \mathcal{R}_j^m I(x_i \in \mathbb{R}_j^m) \tag{8}$$

where $I(\cdot)$ represents the indicator function.

- (5) Obtain the final learner $f(x_i)$ as follows:

$$f(x_i) = f_M(x_i) = f_0(x_i) + \sum_{m=1}^M \sum_{j=1}^J \mathcal{R}_j^m I(x_i \in \mathbb{R}_j^m) \tag{9}$$

2.2.2. XGBoost

XGBoost is a variant of the GBDT. Differently from GBDT, XGBoost performs a second-order Taylor expansion on the objective function, with both first-order and second-order derivatives, thus improving the forecasting accuracy. In addition, XGBoost adds a regularization term in the objective function to control the complexity of the model, which can effectively prevent overfitting. However, XGBoost needs to traverse to select the optimal split point for decision trees. When the data volume is high, it occupies much memory and takes a long time. The objective function of XGBoost is defined as

$$\mathcal{L}(q) = -\frac{1}{2} \sum_{j=1}^T \frac{(\sum_{i \in I_j} g_i)^2}{\sum_{i \in I_j} h_i + \lambda} + \gamma T \tag{10}$$

where T is the number of leaf node; $I_j = \{i | q(\mathbf{x}_i) = j\}$ denotes the instance set of leaf j ; λ and γ are the L1 and L2 regularization parameters, respectively; and g_i and h_i are the first- and second-order gradient statistics on the loss function, respectively.

2.2.3. LGBM

LGBM is another variant of the GBDT, which can solve the problems encountered by GBDT in massive data. First, the histogram algorithm is introduced into the LGBM, which can effectively reduce the memory consumption and improve calculation speed. Second, LGBM uses a leaf-wise growth strategy with depth limitations, which can ensure high efficiency while preventing overfitting. In summary, compared with XGBoost and GBDT, LGBM has faster training speed and lower memory usage; therefore, it can be better applied in industrial practice.

2.2.4. SVR

SVR applies support vector machine (SVM) to regression problems. Its main idea is to find a nonlinear mapping from the input space to the output space, that is, an optimal hyperplane, which has at most ϵ deviation from the target y_i . The above can be described mathematically as follows:

$$f(x) = [w, \Phi(x)] + b \tag{11}$$

with $\Phi : \mathbb{R}^n \rightarrow F, w \in F$

where $\Phi(\cdot)$ represents the nonlinear mapping; F represents a high-dimensional space; w is the weights and b is the bias. Equation (11) can be further transformed into an optimization problem, as shown in Reference [28].

The advantages of SVR lie in that it has good efficiency and robustness, so it is often applied to small datasets. However, when encountering large-scale datasets, SVR will suffer from problems like sensitivity to hyper-parameters, low accuracy, and low speed.

2.2.5. KNN

The main idea of KNN is that each sample can be represented by its K nearest neighbors. When performing regression tasks, KNN selects the K data points closest to each sample from the training data set, takes the average of their label values, and uses this average as the forecasting value for corresponding sample. This paper uses the most commonly used Euclidean distance [31] to measure the distance between each sample in what follows:

$$d = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2} \tag{12}$$

where (x_1, x_2, \dots, x_d) and (y_1, y_2, \dots, y_d) denote the coordinates of any two points in n dimensional space.

KNN is an algorithm with simple principles and is easy to implement and understand. It also has the advantage of being insensitive to outliers. However, since KNN needs to calculate the distances from one sample point to all other sample points, it has large computational overhead and high space complexity, and often suffers from decreased accuracy when dealing with high-dimensional data.

2.2.6. LR and RR

LR is a statistical analysis method to determine the quantitative relationship between two or more variables. The general form of the linear regression equation [32] is

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \epsilon \tag{13}$$

Write (13) in terms of matrices as follows:

$$y = W^T X + \epsilon \tag{14}$$

where $W = [\beta_0, \beta_1, \dots, \beta_k]^T$ is the regression coefficient matrix; $X = [1, x_1, x_2, \dots, x_k]^T$ is the independent variable matrix; y represents the dependent variable; and ε is white noise. The least square method is used to obtain the solution of W as follows:

$$W = (X^T X)^{-1} X^T y \quad (15)$$

Linear regression modeling is simple and has strong interpretability but is only suitable for linear problems. It is sensitive to outliers and multicollinearity and has limited fitting and predicting performance. Therefore, LR is more suitable as a meta-learner to integrate the information learned by base learners, rather than as a base learner to extract the features of original data.

RR is an improved method of least squares estimation. It obtains regression coefficients that are more realistic and reliable at the cost of losing some information and reducing accuracy. Specifically, the ridge regression adds a matrix λI in (15), so the result of ridge estimation is

$$W = (X^T X + \lambda I)^{-1} X^T y \quad (16)$$

Ridge regression prevents overfitting by shrinking the regression coefficients, and is suitable for multicollinearity problems. But there is also a risk of underfitting the model if the parameters are not selected properly.

2.3. Probabilistic Forecasting Model Based on Error Statistics

In power forecasting, given a set of pairs $\{(x_i, t_i)\}$, the measured value of power can be defined as [33]

$$t_i = f(x_i) + \varepsilon(x_i) = \hat{y}(x_i) + \varepsilon(x_i) \quad (17)$$

where t_i is the i th measured value; x_i is the input variables including numerical weather prediction (NWP) data and historical power data; $\hat{y}(x_i) = f(x_i)$ denotes the regression value of stacking ensemble learning model; and $\varepsilon(x_i)$ is the forecasting error which is the noise with a zero mean.

According to (17), the forecasting error $\varepsilon(x_i)$ moves the regression value of the forecasting model away from the true measured value. The $\varepsilon(x_i)$ is assumed to obey a Gaussian distribution, as follows:

$$\varepsilon(x_i) \sim N(0, \sigma_\varepsilon^2(x_i)) \quad (18)$$

where $\sigma_\varepsilon^2(x_i)$ can be estimated from the forecasting errors of test set as follows:

$$\sigma_\varepsilon^2(x_i) = \frac{1}{N-1} \sum_{i=1}^N (t_i - \hat{y}(x_i) - \frac{1}{N} \sum_{i=1}^N t_i - \hat{y}(x_i))^2 \quad (19)$$

Many studies have proven that using a Gaussian distribution to model the forecasting error is reasonable and effective [34,35]. Actually, reference [33] points out that though the actual error distribution may be non-Gaussian, the time series models under an assumption of Gaussian distribution can still work well.

After obtaining the probability distribution of forecasting errors $\varepsilon(x_i)$, superimposing this probability distribution on the point forecasting value can yield the probability distribution of prediction for renewable energy generation and load.

3. Probabilistic Forecasting of ALSC

3.1. Deterministic ALSC Problem

The maximum load growing percentage λ can be used to quantify the ALSC. The λ is obtained by gradually increasing load until encountering an electrical constraint violation. The deterministic ALSC problem can be described mathematically as follows [3]:

Objective:

$$\max \lambda \quad (20)$$

Subject to:

$$\begin{cases} f(x) + \lambda b = 0 \\ U_{i\min} \leq U_i \leq U_{i\max} \\ S_{ij} \leq S_{ij\max} \\ P_{Gi} \leq P_{Gi\max} \end{cases} \quad (21)$$

where $f(x)$ denotes the power flow equation under basic load; b denotes load growing vector; U_i is the voltage amplitude of bus i ; $U_{i\min}$ and $U_{i\max}$ are the minimum and the maximum voltage amplitude limitation of bus i , respectively; S_{ij} is the apparent power of the branch between bus i and bus j ; $S_{ij\max}$ is the maximum apparent power the branch can be transferred; P_{Gi} is the active power of the generator at bus i ; and $P_{Gi\max}$ denotes the maximum power limitation of the generator at bus i .

3.2. The Repeated Power Flow Considering Multi-Slack Buses

RPF is a common method for solving the deterministic ALSC problem, whose basic principle is that, starting from the basic condition, the load is gradually increased according to the given load growing pattern and power flow is calculated repeatedly until any electrical constraint is violated. The maximum load growing percentage λ is defined as

$$\lambda = \frac{S_{Li} - S_{Li0}}{S_{Li0}} \quad (22)$$

where S_{Li} is the increased load of bus i and S_{Li0} is the basic load of bus i .

In the process of RPF, the load will continue to increase. If all the growing load is borne solely by the slack bus, the generator at the slack bus have a high probability of exceeding its maximum output capacity, which does not conform to the actual situation and affects the accurate evaluation of ALSC. Therefore, it is necessary to set multi-slack buses during the RPF calculation. The main task of multi-slack buses is to allocate the power imbalance of the system. In this study, apart from the growing load, the forecasting errors of wind power, photovoltaic power and load power in the basic condition are another source of power imbalance. For two different power imbalance sources, two types of allocations are proposed and performed amongst multiple slack buses herein.

- (1) The first power imbalance allocation: the power imbalance source is the power forecasting errors. More specifically, the power forecasting errors are modeled as having a Gaussian distribution in Section 2.3. When LHS is used to sample from the distribution of random variables, each sampling value is different, which corresponds to the fluctuation of renewable energy generation and load power. This will lead to the power imbalance, which is relatively small. Therefore, the generator power adjustment process is similar to the secondary frequency regulation in the power system. For this kind of power imbalance, power allocation is performed according to the ratio of the remaining capacity of each generator amongst multi-slack buses.
- (2) The second power imbalance allocation: the power imbalance source arises from the growing load power. During the RPF calculation, loads need to be increased continuously to find the maximum load point, which can cause a significant power imbalance. This part of the power imbalance is relatively large and the load change is regular. Therefore, the generator power adjustment process is similar to the tertiary frequency regulation. The tertiary frequency regulation is essentially a problem of economic dispatch. Therefore, the allocation principle of power imbalance is selected according to the ratio of generator economy, which can be represented by the reciprocal of the cost of each generator.

The calculation steps of solving the deterministic ALSC problem by RPF are introduced as follows:

- (1) Obtain the reference value of power system loss P_{loss} by a power flow calculation under the mean value of all random input variables. The active power equation of power flow equations is as follows:

$$P_{Gi} - P_{Di} - U_i \sum_{j=1}^N U_j (G_{ij} \cos \delta_{ij} + B_{ij} \sin \delta_{ij}) = 0 \tag{23}$$

where N is the number of buses; P_{Gi} is the active power of generator at bus i ; P_{Di} is the active power of load at bus i ; U_i is the voltage amplitude of bus i ; δ_{ij} is the phase angle difference between bus i and bus j ; and G_{ij} and B_{ij} are the real part and imaginary parts of the bus admittance matrix, respectively.

- (2) Sample from the forecasting distribution of load, wind power and photovoltaic power at the current moment t and get the sampling points. Given a system with N buses, multi-slack bus set \mathbf{B} , traditional generator bus set \mathbf{T} ($\mathbf{B} \in \mathbf{T}$), and renewable energy generator bus set \mathbf{N} , the first part of the power imbalance ΔP_1 is

$$\Delta P_1 = \sum_{i \in \mathbf{R}} P_{Gi} + \sum_{i \in \mathbf{T}} P_{Gi} - \sum P_{Di} - P_{loss} \tag{24}$$

where $\sum_{i \in \mathbf{R}} P_{Gi}$ represents the power of all traditional generators and $\sum_{i \in \mathbf{T}} P_{Gi}$ represents the power of all renewable energy generators.

- (3) According to the ratio of remaining capacity of each generator, determine the allocation ratio $k_{1,i}$ of the power imbalance ΔP_1 amongst slack buses as

$$k_{1,i} = \frac{P_{r,i}}{\sum P_{r,i}}, i \in \mathbf{B} \tag{25}$$

where $P_{r,i}$ is the remaining power of the generator at bus i .

- (4) Change the active power equation of the slack buses into

$$P_{Gi} - k_{1,i} \Delta P_1 - P_{Di} - U_i \sum_{j=1}^N U_j (G_{ij} \cos \delta_{ij} + B_{ij} \sin \delta_{ij}) = 0, i \in \mathbf{B} \tag{26}$$

then perform the power flow calculation.

- (5) Make an increase of the load, $S_{Li} = S_{Li} + hb_i$ and the second part of power imbalance ΔP_2 is

$$\Delta P_2 = \sum_i^{N_l} hb_i \tag{27}$$

where N_l is the number of loads and b_i is the element of load growing vector b .

- (6) According to the ratio of reciprocal of cost of each generator, determine the allocation ratio $k_{2,i}$ of power imbalance ΔP_2 amongst slack buses as

$$k_{2,i} = \frac{\frac{1}{C_i}}{\sum \frac{1}{C_i}}, i \in \mathbf{B} \tag{28}$$

$$C_i(P_{Gi}) = 0.01P_{Gi}^2 + 0.3P_{Gi} + 0.2 \tag{29}$$

where C_i is the cost of the generator of bus i .

- (7) Change the active power equation of the slack buses into

$$P_{Gi} - k_{1,i} \Delta P_1 - k_{2,i} \Delta P_2 - P_{Di} - U_i \sum_{j=1}^N U_j (G_{ij} \cos \delta_{ij} + B_{ij} \sin \delta_{ij}) = 0, i \in \mathbf{B} \tag{30}$$

and then perform the power flow calculation.

- (8) Check whether some constraints have been violated. If there is any constraint violation, $S_{Li} = S_{Li} - hb_i$ and the generators of slack buses minus the corresponding allocation power, then go to step 9. Otherwise, go to step 5.
- (9) Check whether h is smaller than the convergence condition ϵ . If yes, go to step 10. Otherwise, cut the step h in half, i.e., $h = h/2$, and go to step 5.
- (10) Calculate the ALSC as shown in (22).
- (11) Go to step 2 and calculate the ALSC at the next moment $t + 1$.

In order to intuitively illustrate the two power imbalance allocations amongst the multi-slack buses, a schematic diagram is drawn, as shown in Figure 2.

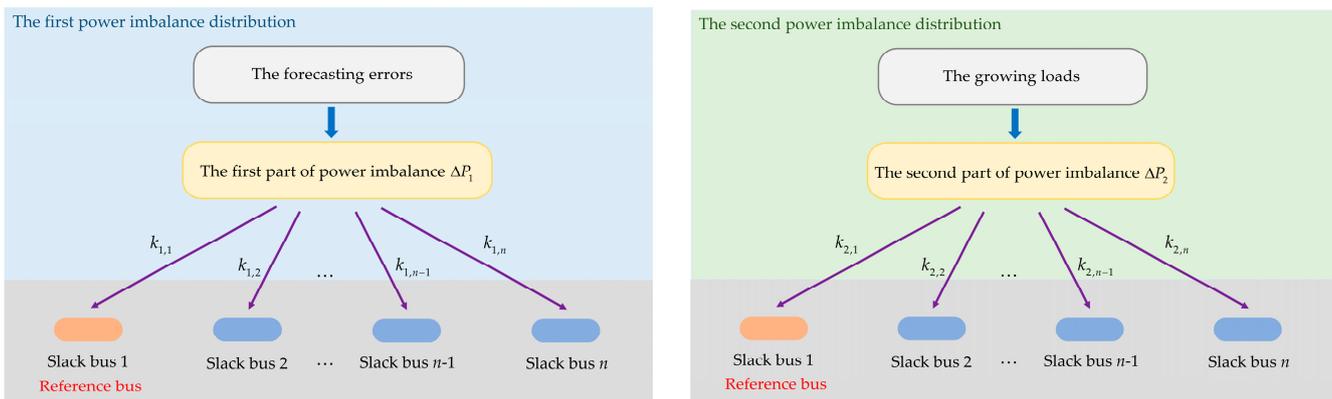


Figure 2. The schematic diagram of two power imbalance allocations.

3.3. The Basic Theories of LHS

The sampling method is crucial for Monte Carlo simulation (MCS). Traditional random sampling (RS) requires numerous samples to characterize random variables accurately but reduces computational efficiency. Evidently, a smaller sample size is expected to decrease calculation time while maintaining accuracy. Latin hypercube sampling (LHS), which is a typical stratified sampling method, can cover more space of random variables than RS with the same sample size.

With the same sample size N , RS and LHS achieve different joint coverage of two independent random variables. The expected values of joint coverage space for RS E_{RS} and LHS E_{LHS} are shown as (31) and (32), respectively [36].

$$E_{RS} = \left(\frac{N - 1}{N + 1}\right)^2 \times 100\% \tag{31}$$

$$E_{LHS} = \left(\frac{N - 1}{N}\right)^2 \times 100\% \tag{32}$$

When $N \geq 2$, E_{LHS} is always larger than E_{RS} . Therefore, LHS always covers more space of random input variables than RS with the same sample size.

3.3.1. The Sampling

The calculation procedure of sampling is as follows. Given M random input variables X_1, X_2, \dots, X_M , the cumulative distribution function (CDF) of X_M is defined as

$$Y_m = F_m(X_m), \quad m = 1, 2, \dots, M \tag{33}$$

First, divide the interval $[0, 1]$ into N nonoverlapping subintervals of equal length. Then choose one value randomly from each subinterval. In the end, the sampling values of random input variables are calculated as

$$x_{mn} = F_m^{-1}\left(\frac{n - rand}{N}\right), \quad m = 1, 2, \dots, M, \quad n = 1, 2, \dots, N \tag{34}$$

where $F_m^{-1}(\cdot)$ denotes the inverse cumulative distribution function and $rand$ represents the random number changing from 0 to 1.

Then organize the N sample values of X_M into a vector $[x_{k1}, x_{k2}, \dots, x_{kN}]$. Once the samples of all random input variables are drawn, an $M \times N$ initial sampling matrix S can be formulated correspondingly.

The above sampling process is shown in the Figure 3.

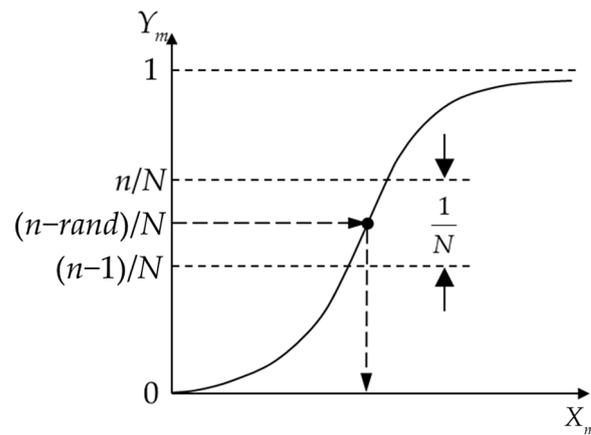


Figure 3. Schematic diagram of sampling process of LHS.

3.3.2. The Permutation

For a probabilistic problem, the random input variables often have certain correlations. However, the correlations between random variables after sampling are uncontrollable, which can negatively impact the results. Therefore, it is necessary to reorder the sampling values to adjust the correlations to the expected level. There are several permutation methods, such as ranked Gram–Schmidt orthogonalization [37], simulated annealing [38], the genetic algorithm [39], and Cholesky decomposition [40]. Among them, Cholesky decomposition is the most popular method due to its efficiency and accuracy.

The detailed computational steps of Cholesky decomposition method for permutation are as follows.

- (1) According to the procedure of sampling, obtain the initial sampling matrix S ;
- (2) Randomly permute 1 to N in every row to obtain an $M \times N$ ranking matrix R ;
- (3) Calculate the Spearman rank correlation coefficient matrix ρ_R of R as shown in (35) and (36)

$$\rho_R = \begin{bmatrix} 1 & \rho_{12} & \dots & \rho_{1M} \\ \rho_{21} & 1 & \dots & \rho_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{M1} & \rho_{M2} & \dots & 1 \end{bmatrix} \tag{35}$$

$$\rho_{ij} = \frac{\text{cov}(R_i, R_j)}{\sigma_i \sigma_j} = 1 - \frac{6 \sum_{n=1}^N (r_{in} - r_{jn})^2}{N(N-1)(N+1)} \tag{36}$$

where ρ_{ij} denotes the Spearman rank correlation coefficient between the i th row and the j th row of R ; $\text{cov}(R_i, R_j)$ denotes the covariance between the i th row and the j th row of R ; σ_i and σ_j are the standard deviations of the i th row and the j th row of R , respectively; r_{in} represents the n th element in the i th row of R ; and r_{jn} represents the n th element in the j th row of R ;

- (4) Perform Cholesky decomposition on ρ_R as shown in (37)

$$\rho_R = DD^T \tag{37}$$

where D is the lower triangular matrix from the decomposition.

- (5) Multiply the inverse of D with R , and then obtain an identity correlation matrix G as follows:

$$G = D^{-1}R \tag{38}$$

- (6) Conduct Cholesky decomposition on the expected Spearman rank correlation coefficient matrix ρ_{exp} of random input variables as follows:

$$\rho_{\text{exp}} = D_{\text{exp}}D_{\text{exp}}^T \tag{39}$$

- (7) Multiply G to D_{exp} , then obtain a matrix G_{new} whose Spearman rank correlation coefficient is approximately equal to the ρ_{exp} as follows:

$$G_{\text{new}} = D_{\text{exp}}G \tag{40}$$

- (8) According to the amplitude order of elements in each row of matrix G , reorder the corresponding rows of R to obtain a new permutation matrix R_{new} ;
- (9) According to the order in each row of R_{new} , reorder the rows of the initial sampling matrix S to obtain the final sampling matrix S_{new} .

3.4. The Procedure of Probabilistic Forecasting of ALS C

The calculation steps of probabilistic ALS C forecasting can be summarized as: First, the stacking ensemble learning model is used to forecast wind power, photovoltaic power and load power. Secondly, the forecasting distribution is obtained based on the point forecasting results. Then, the LHS method is used to sample the forecasting distribution, and the sampling values are input into the RPF considering multi-slack buses to calculate the ALS C, and finally the probabilistic prediction results of the ALS C are obtained. The whole calculation process is shown in Figure 4.

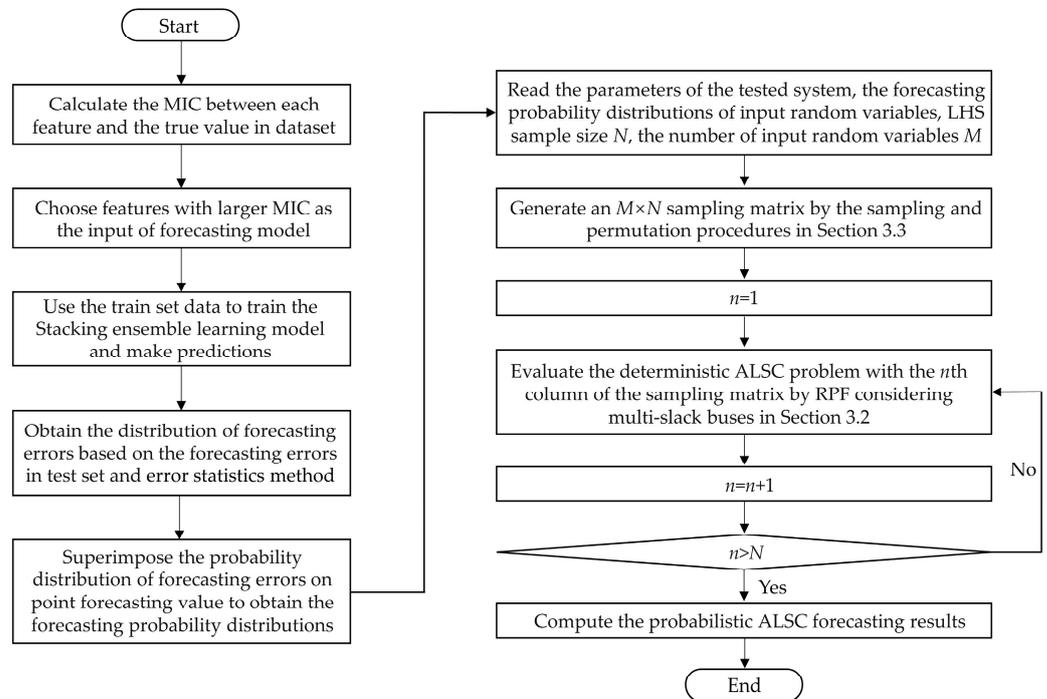


Figure 4. Flowchart of probabilistic ALS C forecasting.

3.5. Probabilistic Indices for ALS C

In order to accurately evaluate ALS C under uncertainties, five probabilistic evaluation indices are defined as follows.

- (1) Expectation of ALS C:

$$E_{\text{ALS C}} = \frac{1}{N_M} \sum_{i=1}^N \lambda_i \tag{41}$$

where N_M is the number of the Monte Carlo simulation and λ_i denotes the ALS C value of the i th Monte Carlo simulation.

- (2) Standard deviation of ALS C:

$$S_{\text{ALS C}} = \sqrt{\frac{1}{N_M} \sum_{i=1}^N (\lambda_i - E_{\text{ALS C}})^2} \tag{42}$$

- (3) Variation coefficient of ALS C:

$$\beta_{\text{ALS C}} = \frac{S_{\text{ALS C}}}{E_{\text{ALS C}}} \tag{43}$$

- (4) Probability of exceeding threshold for ALSC:

$$P_{ET} = \frac{M_{ALSC < \bar{\lambda}}}{N_M} \tag{44}$$

where $M_{ALSC < \bar{\lambda}}$ denotes the number at which ALSC is less than the given threshold $\bar{\lambda}$.

- (5) Probability of ALSC equal to zero:

$$P_{EZ} = \frac{M_{ALSC=0}}{N_M} \tag{45}$$

where $M_{ALSC=0}$ denotes the number at which ALSC is equal to zero.

E_{ALSC} gives a general estimation of the ALSC. S_{ALSC} and β_{ALSC} represent the fluctuation of the ALSC. P_{ET} and P_{EZ} are the risk indices of power system.

4. Case Studies

4.1. Wind Power, Photovoltaic Power and Load Power Forecasting

4.1.1. Data Description

Three real datasets are used to conduct wind power forecasting, photovoltaic power forecasting and load power forecasting [41]. The wind power data comes from a certain wind power station in Inner Mongolia, China. The photovoltaic power prediction data comes from a photovoltaic power station in the same area. The load data comes from a substation connected to a distribution network and transmission network. There are a total of 7788 data in the wind power dataset. The first 5500 data are used as the training set, and the remaining data are used as the test set. There are a total of 8688 data in the photovoltaic power dataset. The first 6500 data are used as the training set, and the remaining data are used as the test set. There are a total of 9432 data in the load power dataset. The first 7000 data are used as the training set, and the remaining data are used as the test set. Feature engineering is performed on the basis of the original dataset, as shown in Table 2. The original features refer to the features contained in the original dataset, and the added features refer to the features constructed from the original features. The number of features after the feature engineering are 14, 44 and 44, respectively, for wind power forecasting, photovoltaic power forecasting and load power forecasting.

Table 2. Feature engineering in different prediction models.

Features	Wind Power Forecasting	Photovoltaic Power Forecasting	Load Power Forecasting
Original features	Forecasting wind speed	Forecasting illuminance	Forecasting temperature
	Forecasting wind direction	Forecasting temperature	Forecasting humidity
	Forecasting humidity	Forecasting air pressure	Forecasting wind speed
	Forecasting temperature	Forecasting humidity	Forecasting precipitation
	Forecasting air pressure	Historical power at same moment.	Historical load at same moment.
Added features	Square of forecasting wind speed	Average power at the same moment in previous three days	Average load at the same moment in previous three days
	Cube of forecasting wind speed	Power at the same moment in previous three days	Load at the same moment in previous three days
	Product of wind speed and humidity	Power of previous 24 time points in one day	Load of previous 24 time points in one day
	Historical wind speed for previous four hours	Difference between predicted and previous meteorological data	Difference between predicted and previous meteorological data

4.1.2. Evaluation Index for Forecasting

Mean absolute error (MAE) [42], root mean square error (RMSE) [43] and mean absolute percentage error (MAPE) [44] are employed for evaluating the forecasting results; the formulas are as follows:

$$MAE = \frac{1}{N} \sum_{t=1}^N |\hat{y}(t) - y(t)| \tag{46}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N [\hat{y}(t) - y(t)]^2} \tag{47}$$

$$MAPE = \frac{1}{N} \sum_{t=1}^N \left| \frac{\hat{y}(t) - y(t)}{y(t)} \right| \tag{48}$$

where N denotes the number of samples; $\hat{y}(t)$ is the forecasting value; and $y(t)$ is the real value.

4.1.3. Model Parameter Settings

In stacking ensemble learning, there are many options of learners. Through the experiments, we can find the best combination of learners and the optimal hyper-parameters for each forecasting problem, as shown in Tables 3–5.

Table 3. Learners and hyper-parameters of wind power forecasting.

Type of Learners	Algorithm	Hyper-Parameters Setting
base learners	LGBM	Boosting tree is GBDT, number of leaves is 20, number of trees is 300, tree depth is 6, learning rate is 0.1, training data ratio is 0.8, feature column sampling rate is 0.9.
	XGBoost	
	RR	The regularization term coefficient is 1.
	SVR	The distance error is 0.15.
	KNN	The number of neighbors is 5.
meta-learner	LR	/

Table 4. Learners and hyper-parameters of photovoltaic power forecasting.

Type of Learners	Algorithm	Hyper-Parameters Setting
base learners	LGBM	Boosting tree is GBDT, number of leaves is 100, number of trees is 100, tree depth is 2, learning rate is 0.1, training data ratio is 0.8, feature column sampling rate is 0.9.
	XGBoost	
	RR	The regularization term coefficient is 1.
	KNN	The number of neighbors is 100.
	LR	/
meta-learner	LR	/

As shown from Tables 3–5, the best forecasting effects of different forecasting problems require different combinations of learners. Through experiments on combinations of learners, it is found that the base learners for wind power forecasting are LGBM, XGB, RR, SVR and KNN. The base learners for photovoltaic power forecasting are LGBM, XGB, RR, KNN and LR. The base learners for load power forecasting are LGBM, XGB and KNN. The meta-learner for all three forecasting objects is LR.

Table 5. Learners and hyper-parameters of load power forecasting.

Type of Learners	Algorithm	Hyper-Parameters Setting
base learners	LGBM	Boosting tree is GBDT, number of leaves is 100, number of trees is 100, tree depth is 7, learning rate is 0.1, training data ratio is 0.8, feature column sampling rate is 0.9.
	XGBoost	
	KNN	The number of Neighbors is 10.
meta-learner	LR	/

4.1.4. The Results of Feature Selection

For wind power forecasting, photovoltaic power forecasting and load power forecasting, MICs between features and true value are calculated respectively, and the features with larger MICs are selected as the input features. The optimal number of input features is determined by an exhaustive search.

It can be seen from Figure 5 that different numbers of input features will lead to different forecasting results. For wind power forecasting, photovoltaic power forecasting and load power forecasting separately select the number of input features that can minimize the RMSE, which are 13, 8 and 33 respectively, as shown by the red dot in Figure 5.

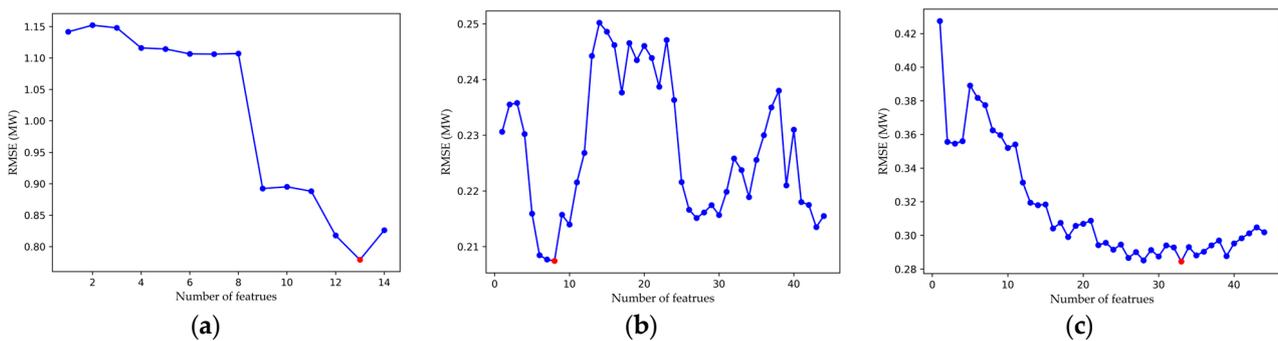


Figure 5. The relationship between the number of input features and forecasting results: (a) Wind power; (b) Photovoltaic power; (c) Load power.

4.1.5. The Results of Forecasting

In order to illustrate the effectiveness of the MIC feature selection method, the results of the forecasting model with feature selection and those of the forecasting model without feature selection are compared, as shown in Table 6. In addition, in order to illustrate the improvement of stacking ensemble learning on the forecasting accuracy of a single model, for each power forecasting problem, this study compares the MAE, RMSE and MAPE of the stacking ensemble learning model with the MAE, RMSE and MAPE of a single algorithm. At the same time, this study is also compared to the forecasting results of a long-short term memory (LSTM) network, a representative method in the forecasting field, to reflect the accuracy of the model. The results of the comparative experiment are as shown in Table 7.

Table 6. Forecasting results comparison of with and without feature selection.

Index	Without Feature Selection			With Feature Selection		
	Wind	Photovoltaic	Load	Wind	Photovoltaic	Load
MAE (MW)	0.650	0.118	0.192	0.597	0.108	0.178
RMSE (MW)	0.826	0.220	0.301	0.779	0.207	0.284
MAPE (%)	35.369	10.394	3.923	27.418	9.977	3.005

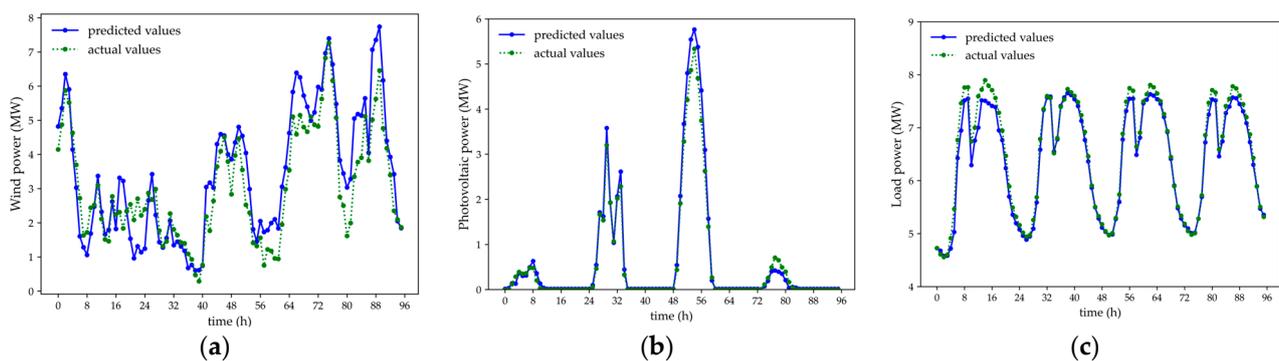
Table 7. Forecasting results comparison of stacking and different algorithms.

Model	MAE (MW)			RMSE (MW)			MAPE (%)		
	Wind	Photovoltaic	Load	Wind	Photovoltaic	Load	Wind	Photovoltaic	Load
LGBM [26]	0.621	0.109	0.183	0.798	0.226	0.297	29.123	10.640	3.079
XGBoost [25]	0.624	0.122	0.185	0.813	0.238	0.289	27.712	11.348	3.120
RR [27]	0.857	0.123	0.254 *	1.055	0.228	0.347 *	44.474	12.986	4.385 *
SVR [28]	0.846	0.155 *	0.216 *	1.060	0.216 *	0.346 *	40.634	11.275 *	3.542 *
KNN [29]	0.756	0.127	0.255	0.977	0.260	0.369	35.309	13.491	4.317
LR [30]	0.835	0.123	0.278	1.040	0.228	0.369	41.702	10.372	4.805
LSTM [45]	0.611 *	0.115 *	0.189 *	0.828 *	0.213 *	0.288 *	28.187 *	11.161 *	3.248 *
Stacking	0.597	0.108	0.178	0.779	0.207	0.284	27.462	9.977	3.005

The results in Table 6 prove the effectiveness of MIC feature selection method, because the MAE, RMSE and MAPE of the forecasting model with feature selection are all lower than those without feature selection. In Table 7, data with * indicate that the corresponding algorithm is not used in the stacking ensemble model of this forecasting problem. The results show that even if the accuracy of some algorithms is not high when predicting independently, the accuracy of the ensemble model can be improved after they join the stacking model. Therefore, they will be also included in the stacking model. For example, the RMSE of the KNN algorithm in photovoltaic and load power forecasting is relatively high, but it is still integrated into the stacking model. At the same time, although the LSTM model can achieve a good forecasting accuracy, its accuracy is still inferior to that of the stacking model.

To further understand the forecasting accuracy of the stacking ensemble learning model, 96 time points (4 days) are selected herein to display the forecasting value of the stacking ensemble models of the three forecasting problems and the actual value, respectively.

As shown in Figure 6, the forecasting model based on the stacking ensemble learning model achieves a good accuracy in wind power forecasting, photovoltaic power forecasting and load power forecasting, which can provide the accurate wind power, photovoltaic power and load data for RPF calculation.

**Figure 6.** The point forecasting results: (a) Wind power; (b) Photovoltaic power; (c) Load power.

It has been mentioned in Section 2.3 that the forecasting errors of the test set can be statistically obtained to get the probability distribution of the forecasting errors. The forecasting errors of wind power, photovoltaic power and load power follow the Gaussian distribution as follows:

$$\varepsilon_{wind} \sim N(0, 0.779) \quad (49)$$

$$\varepsilon_{pv} \sim N(0, 0.207) \quad (50)$$

$$\varepsilon_{load} \sim N(0, 0.284) \quad (51)$$

where ε_{wind} , ε_{pv} and ε_{load} represent the forecasting errors of wind power, photovoltaic power and load power, respectively.

The forecasting probability distribution can be obtained by superimposing the distribution of forecasting errors on the point forecasting value. For example, if the point forecasting value of wind power at the next moment is 5, its forecasting probability distribution is $N(5,0.779)$, correspondingly.

4.2. Probabilistic Forecasting of ALSC

In the ALSC calculation, the probability distributions of uncertainty sources are determined by the forecasting results of wind power, photovoltaic power and load power herein. To avoid confusion, it is necessary to explain how to use the forecasting results in the following subsections. Sections 4.2.2–4.2.4 and 4.2.6 do not involve time series data, so the power probabilistic forecasting results at a certain time point can be represented as the final distribution of renewable energy and load in the modified IEEE 39-bus system. Section 4.2.5 shows the probabilistic forecasting results of ALSC varying along with time. Therefore, the probabilistic power forecasting results over a period of time are investigated in Section 4.2.5.

4.2.1. Modified IEEE 39-Bus System

To evaluate the performance of the proposed methods, a modified IEEE 39-bus system is designed. The base power is 100 MVA. Bus 31, bus 35, bus 36, bus 37, bus 38 and bus 39 are selected as the slack buses, and bus 31 is assigned as the reference bus. A wind farm is connected at bus 34. The power of the wind farm is determined by 100 times of the predicted wind power. The photovoltaic power plant is connected at bus 30. The power of the photovoltaic power plant is determined by 70 times of the predicted photovoltaic power. The active power of loads is determined by integer multiples of the forecasting load power, as shown in Table 8. The reactive power of loads is determined by keeping the power factor of the load constant. In the RPF calculation, it is assumed that loads grow in accordance with the original power factor. There are a total of 23 uncertainty sources. The allowed bus voltage range is 0.94–1.06 p.u. The active power capacity of each branch is specified as in Table 9 and the maximum active power of each generator is set as shown in Table 10. The figure of modified IEEE 39-bus system is as shown in Figure 7.

Table 8. The active power of loads of each bus.

Load of Bus	Multiple of Forecasting Load
Bus 9, 12, 31	1
Bus 1, 18, 26	20
Bus 7, 23, 25, 28	35
Bus 3, 15, 16, 21, 24, 27, 29	45
Bus 4, 8	85
Bus 20	110
Bus 39	185

Correlations always exist among loads, among renewable generation, and between loads and renewable generation in the power system. The wind power data, photovoltaic power data and load power data are used to calculate the correlation amongst each type of the random input variables. The correlation coefficient between the wind power and photovoltaic power is calculated as -0.0419 ; the correlation coefficient between the wind power and loads is calculated as -0.1203 ; the correlation coefficient between the photovoltaic power and photovoltaic power is calculated as 0.3755 ; and the correlation coefficient among loads is set as 0.7 according to [3].

Table 9. The active power capacity of each branch.

Branch	1–2	1–39	2–3	2–25	2–30	3–4	3–18	4–5	4–14	5–6	5–8	6–7	6–11
$S_{ijmax}/p.u.$	3.75	3	6	6	6	3	3	4.5	4.5	9	6	7.5	6
Branch	6–31	7–8	8–9	9–39	10–11	10–13	10–32	12–11	12–13	13–14	14–15	15–16	16–17
$S_{ijmax}/p.u.$	12	4.5	3	3	6	6	10.5	1.5	1.5	6	1.5	6	5.25
Branch	16–19	16–21	16–24	17–18	17–27	19–20	19–33	20–34	21–22	22–23	22–35	23–24	23–36
$S_{ijmax}/p.u.$	7.5	6.75	3	4.5	1.5	4.5	10.5	9	10.5	1.5	12	6	10.5
Branch	25–26	25–37	26–27	26–28	26–29	28–29	29–38						
$S_{ijmax}/p.u.$	1.5	9.75	4.5	3	4.5	6.75	15						

Table 10. The maximum active power of each generator.

Generator	30	31	32	33	34	35	36	37	38	39
P_{Gimax}/MW	1040	969	725	652	508	1030.5	870	846	1297.5	1650

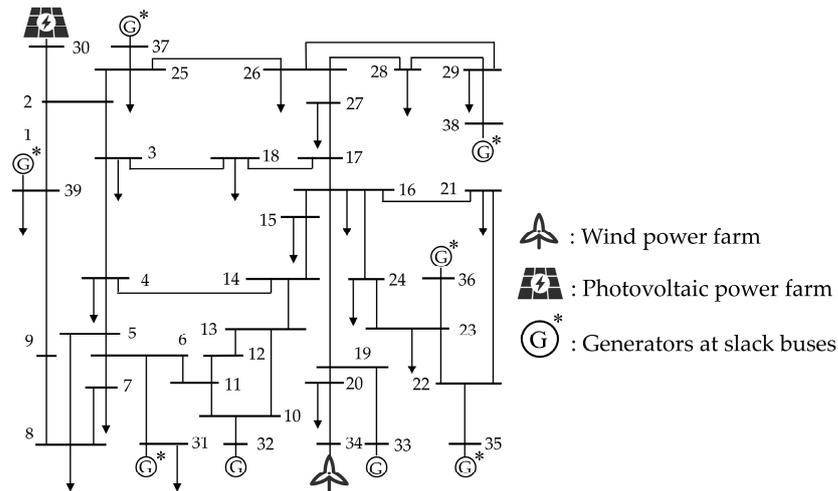


Figure 7. Modified IEEE 39-bus system.

4.2.2. The Effectiveness of LHS-MCS

In order to evaluate the performance of LHS-MCS, the results of the probabilistic indices for ALSC (expectation, standard deviation and variation coefficient) of RS-MCS are taken, with a sample size of 20,000 as the benchmark, the results of which are expressed as $I_{accurate}$. Likewise, the results of probabilistic indices obtained from the test sample size of LHS-MCS and RS-MCS are expressed as $I_{simulate}$. The error of indices is introduced to evaluate the accuracy of the probabilistic methods, as follows:

$$\epsilon_I = \left| \frac{I_{simulate} - I_{accurate}}{I_{accurate}} \right| \times 100\% \tag{52}$$

Due to the random factors in the probabilistic evaluation of ALSC, the error of probabilistic indices of both methods are obtained by the results of 50 simulations with the given sample size. The mean error of 50 simulations, which is expressed as $\epsilon_{I,50}^{mean}$, is taken to demonstrate the accuracy of the tested method. In addition, to explain the robustness of the tested method, the standard deviation of the error of 50 simulations is taken, which can be expressed as $\epsilon_{I,50}^{std}$.

The curves of $\epsilon_{E_{ALSC},50}^{mean}$, $\epsilon_{S_{ALSC},50}^{mean}$, $\epsilon_{\beta_{ALSC},50}^{mean}$ and $\epsilon_{E_{ALSC},50}^{std}$, $\epsilon_{S_{ALSC},50}^{std}$ and $\epsilon_{\beta_{ALSC},50}^{std}$ with different sample sizes are shown in Figures 8 and 9, respectively. It can be seen that the mean and standard deviation of the error of LHS-MCS are smaller than those of RS-MCS

under different sample sizes, demonstrating that LHS-MCS is more accurate and stable than RS-MCS. In addition, the mean and standard deviation of the error of LHS-MCS converge faster, indicating that LHS-MCS needs a smaller sample size to achieve good accuracy. In contrast, RS-MCS requires a large sample size to achieve an acceptable accuracy.

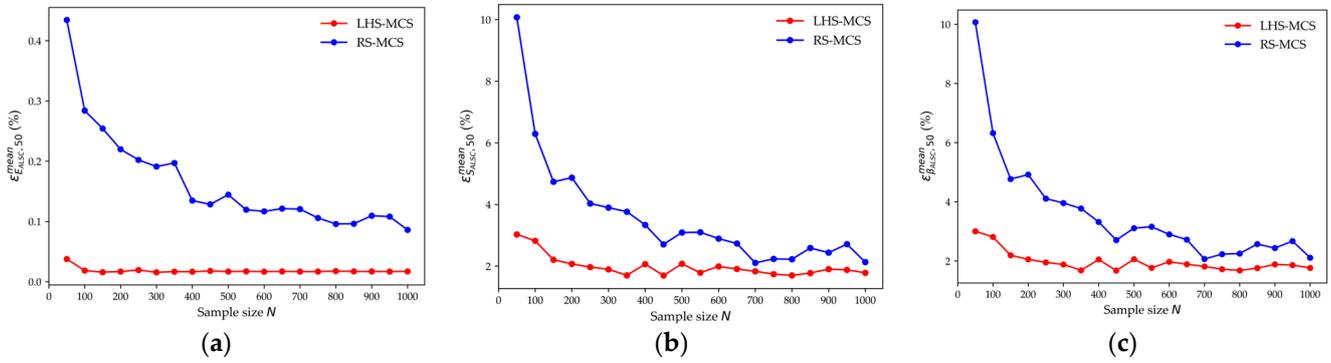


Figure 8. Error comparisons of $\varepsilon_{1,50}^{mean}$ between LHS-MCS and RS-MCS: (a) Error comparison of $\varepsilon_{E_{ALSC},50}^{mean}$; (b) Error comparison of $\varepsilon_{E_{SALSC},50}^{mean}$; (c) Error comparison of $\varepsilon_{\beta_{ALSC},50}^{mean}$.

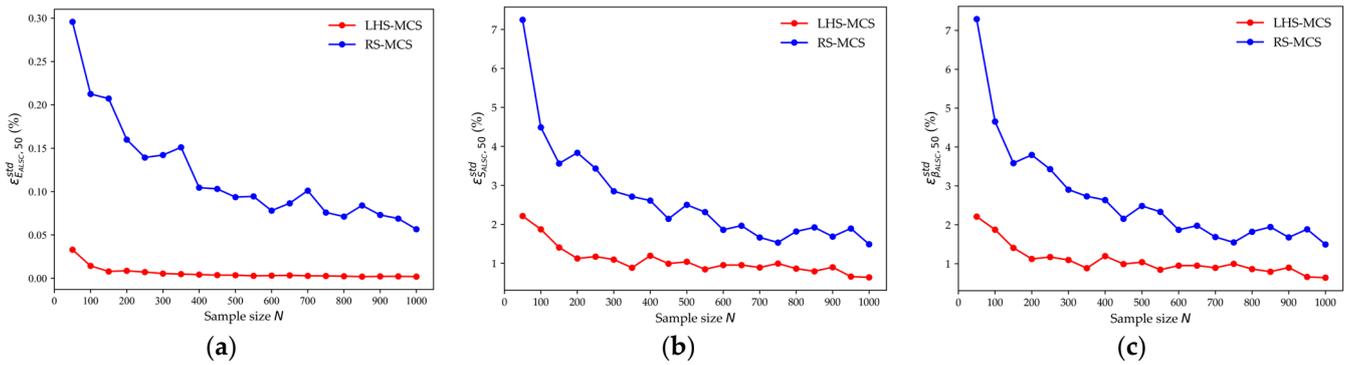


Figure 9. Error comparisons of $\varepsilon_{1,50}^{std}$ between LHS-MCS and RS-MCS: (a) Error comparison of $\varepsilon_{E_{ALSC},50}^{std}$; (b) Error comparison of $\varepsilon_{E_{SALSC},50}^{std}$; (c) Error comparison of $\varepsilon_{\beta_{ALSC},50}^{std}$.

4.2.3. The Impact of Multi-Slack Buses on Probabilistic ALSC

In this study, two types of power imbalance allocations are performed at multi-slack buses. In order to study the impact of multi-slack based on probabilistic ALSC, according to the different number of times and principles of power imbalance allocation, four different methods are considered, which are shown in Table 11. Probabilistic evaluations of ALSC are performed with four different methods, and the results are shown in Table 12. More specifically, the influence of multi-slack buses on probabilistic ALSC is analyzed from two perspectives: (1) the impact of the principle of the first power imbalance allocation; (2) the impact of the principle of the second power imbalance allocation.

First, the influence of the first power imbalance allocation on the results of probabilistic ALSC is discussed. By comparing the results of Method 1 and Method 3 (i.e., the first and the third rows in Table 12) and the results of Method 2 and Method 4 (i.e., the second and the fourth rows in Table 12), it can be found that the probabilistic results of ALSC of Method 3 and Method 4 containing the first power imbalance allocation are different from those of Method 1 and Method 2 without the first power imbalance allocation, demonstrating the first part of power imbalance caused by forecasting errors can affect the probabilistic results of ALSC. The formation mechanism of the first power imbalance is as follows. Due to the forecasting errors, wind power, photovoltaic power and load power are all random variables. Probabilistic ALSC calculation needs samples from these random variables and then performs deterministic ALSC calculation. The sampling values are different each time,

resulting in the first part of power imbalance according to (24). The first power imbalance affects the results of probabilistic ALSC in two ways. First, with different sampling values, generators at each slack bus change their productions to share this power imbalance. Secondly, the first power imbalance allocation determines the starting point of the second power allocation; in other words, it determines the growth direction of generator power in the subsequent power imbalance allocation. Figure 10 shows the change of power of generators at each slack bus under different sampling values of random variables in the process of ALSC probabilistic evaluation (the sample size is 1000, but only 40 of them are shown in this figure). It can be seen that the power of each generator is adjusted with the different sampling values. In addition, since the allocation principle of this power imbalance is formed based on the ratio of remaining capacity, the power of each generator has a similar changing trend.

Table 11. Introduction of four different methods.

Method	Description
Method 1 [13]	Only the second power imbalance allocation is conducted; the allocation principle is the ratio of the current power of the generators
Method 2	Only the second power imbalance allocation is conducted; the allocation principle is the ratio of reciprocal of each generator cost
Method 3	Both power imbalance allocations are conducted; the allocation principle of the second power imbalance allocation is the ratio of the current power of the generators
Method 4	Both power imbalance allocations are conducted; the allocation principle of the second power imbalance allocation is the ratio of reciprocal of each generator cost

Table 12. Probabilistic results of ALSC with different methods.

Method	E_{ALSC}	S_{ALSC}	β_{ALSC}
Method 1	0.3255	0.0375	0.1152
Method 2	0.2282	0.0226	0.0990
Method 3	0.4042	0.0482	0.1193
Method 4	0.2351	0.0366	0.1557

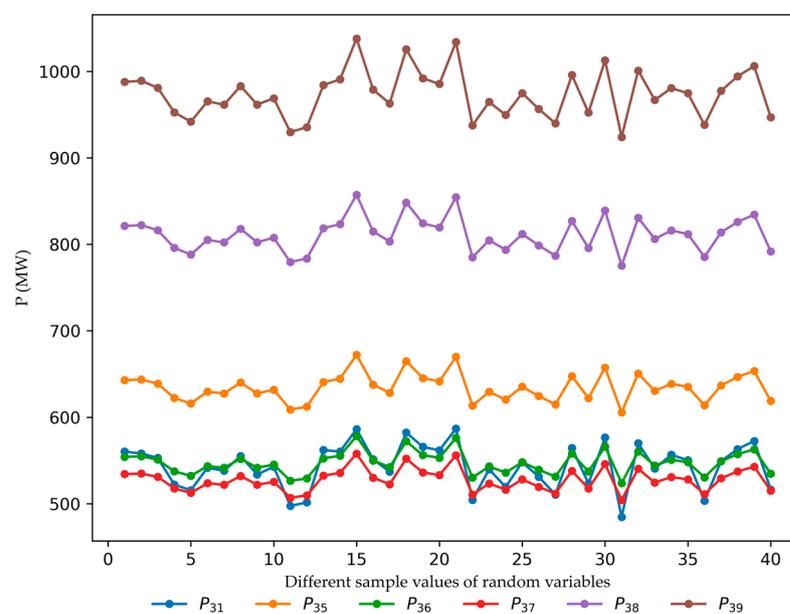
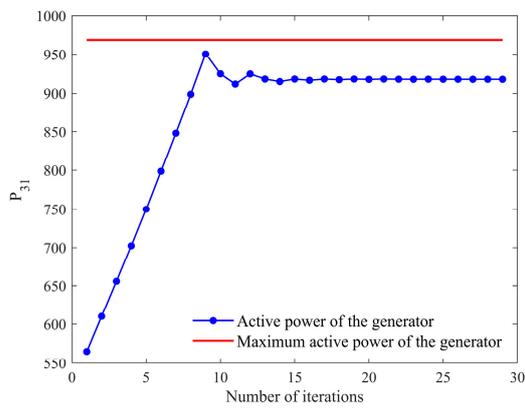


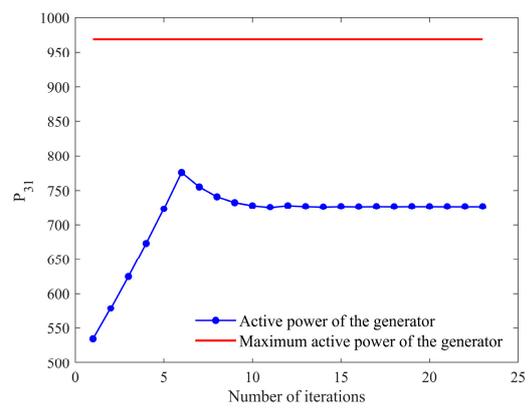
Figure 10. The power change of generators at slack buses with different sampling values of random variables based on modified IEEE 39-bus system.

Furthermore, it can be seen that the S_{ALSC} of Method 3 is larger than that of Method 1 and the S_{ALSC} of Method 4 is larger than that of Method 2, indicating that the consideration of the first power imbalance allocation increases the S_{ALSC} . The reason lies in that the first power imbalance allocation brings a larger difference of ALSC calculation results with different sampling values. To be more specific, the cause of the S_{ALSC} is that the different sampling values of wind power, photovoltaic power and load power lead to different results of deterministic ALSC calculation. When the sampling values of wind power and photovoltaic power are large and the sampling value of load power is small (i.e., the first part of power imbalance $\Delta P_1 > 0$), the first power imbalance allocation will decrease the power of traditional generators at slack buses, while when the sampling values of wind power and photovoltaic power are small and the sampling value of load power is large (i.e., the first part of power imbalance $\Delta P_1 < 0$), the first power imbalance allocation will increase the power of traditional generators at slack buses. If the first power imbalance allocation is not required, the initial power (i.e., the power before the load growth and the second power imbalance allocation) of the traditional generator is constant. With the first power imbalance allocation, different sampling values of wind power and photovoltaic power and load power will cause the initial power of the traditional generators to change, which magnifies the difference between the different sampling values and hence increases the S_{ALSC} .

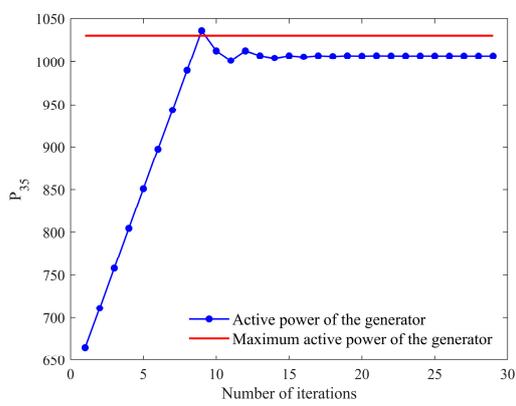
Next, the influence of the principles of the second power imbalance allocation on the results of probabilistic ALSC is discussed. Two different allocation principles of the second power imbalance allocation are considered herein: (1) Method 1 and Method 3 with the ratio of the current power of the generators; (2) Method 2 and Method 4 with the ratio of reciprocal of each generator cost. According to the results in Table 11, it is obvious that the E_{ALSC} calculated by the method with the second allocation principle is smaller. Taking Method 3 and Method 4 as an example, Figure 11 shows the power changes of each generator during the increase of load under different allocation principles. From Figure 11, it can be found that Method 3 makes fuller use of the capacity of each generator, so the E_{ALSC} calculated in Method 3 is greater than that in Method 4. This is because the principle of the second power imbalance allocation in Method 4 works according to the ratio of reciprocal of cost of each generator. The cost of generators is positively correlated with the power, as shown in (29). Therefore, generators with higher capacity, such as P_{38} and P_{39} have higher cost and get a smaller share in power allocation during the whole load increase process. Generators with lower capacity, such as P_{36} and P_{37} , have lower cost and have a larger allocation ratio more quickly reach their output limits. On the contrary, the allocation principle of Method 3 is the ratio of the current power of generators. Therefore, P_{38} and P_{39} , which have higher capacity, have a larger allocation ratio, while P_{36} and P_{37} with lower capacity have a smaller allocation ratio. Since the output limit of the high-capacity generator is correspondingly larger, this allocation principle can make full use of the capacity of each generator and obtain a larger E_{ALSC} . However, it should be emphasized that although the E_{ALSC} of method with the first allocation principle is larger, the second allocation principle is more reasonable, because considering the economics in the face of sustained regular load growth is a common and indispensable practice in actual operation. Therefore, the evaluation results of the load supply capability obtained by the first allocation principle are more optimistic than the actual situation, which amplifies the risk of encountering security issues.



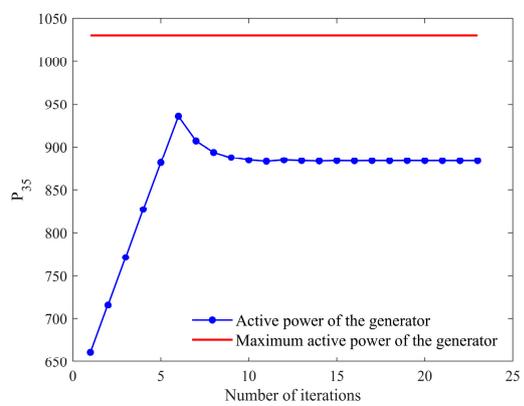
(a)



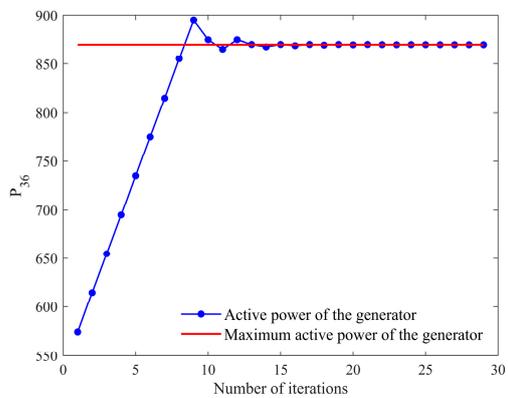
(b)



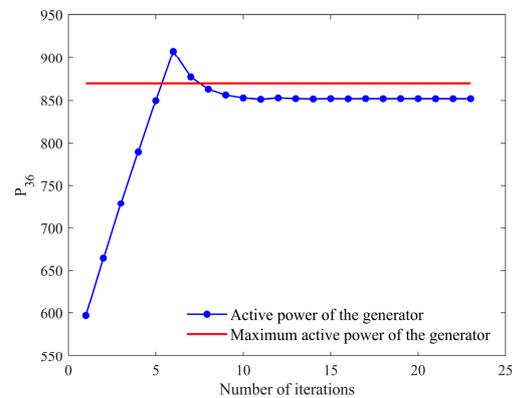
(c)



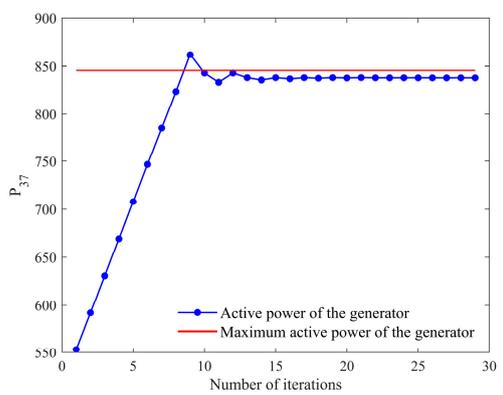
(d)



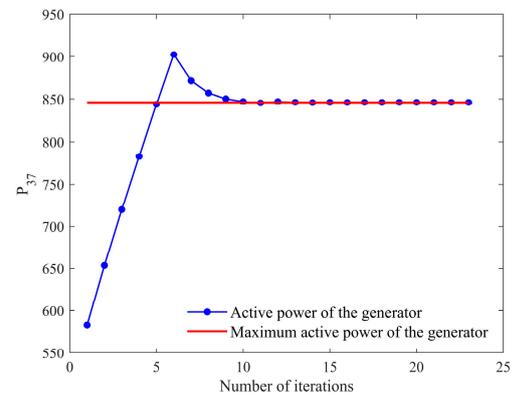
(e)



(f)



(g)



(h)

Figure 11. Cont.

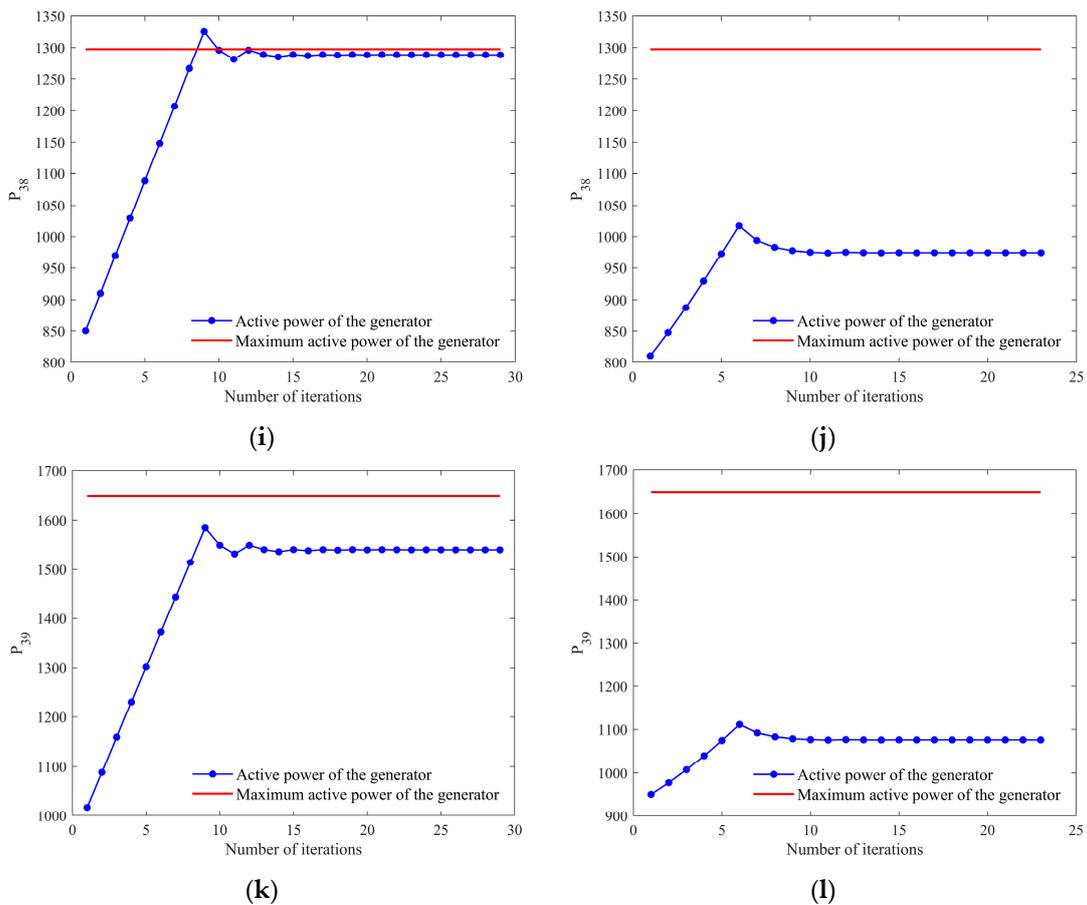


Figure 11. The power variation of generators during load growth based on modified IEEE 39-bus systems: (a) Generator at bus 31 in Method 3; (b) Generator at bus 31 in Method 4; (c) Generator at bus 35 in Method 3; (d) Generator at bus 35 in Method 4; (e) Generator at bus 36 in Method 3; (f) Generator at bus 36 in Method 4; (g) Generator at bus 37 in Method 3; (h) Generator at bus 37 in Method 4; (i) Generator at bus 38 in Method 3; (j) Generator at bus 38 in Method 4; (k) Generator at bus 39 in Method 3; (l) Generator at bus 39 in Method 4.

4.2.4. The Impact of Power Forecasting on Probabilistic ALS

In addition to the fact that two types of power imbalance allocations affect the results of probabilistic ALS, the forecasting accuracy of wind power, photovoltaic power and load power are influential as well. In order to study the impact, eight different forecasting models mentioned in Table 6 (LGBM, XGB, RR, SVR, KNN, LR, LSTM and Stacking ensemble learning model) are used to predict the wind power, photovoltaic power and load power at the same time point, respectively. The probabilistic forecasting results are used to calculate the probabilistic ALS. Correspondingly, the probability density functions (PDFs) of ALS obtained by different forecasting models are shown in Figure 12.

From Figure 12, it can be seen that the PDFs of ALS with different forecasting models are evidently different, thus demonstrating that the different prediction accuracy of forecasting models have significant influence on the results of probabilistic ALS. More specifically, in Figure 12a, the PDF of XGB is most similar to the PDF of Stacking. Compared with the PDF of stacking, the PDFs of LGBM and RR have lower sharpness, and the tail of the PDF of RR is thicker. In Figure 12b, the PDF of LSTM is most similar to the PDF of Stacking. The sharpness of the PDFs of SVR, LR is apparently lower and the PDF of KNN is more right-skewed. In addition, according to Table 6, it can be found that the two models with the highest accuracy closest to the stacking ensemble learning model in the wind power, photovoltaic power and load power forecasting: XGB and LSTM can also obtain the PDFs closest to stacking ensemble learning model.

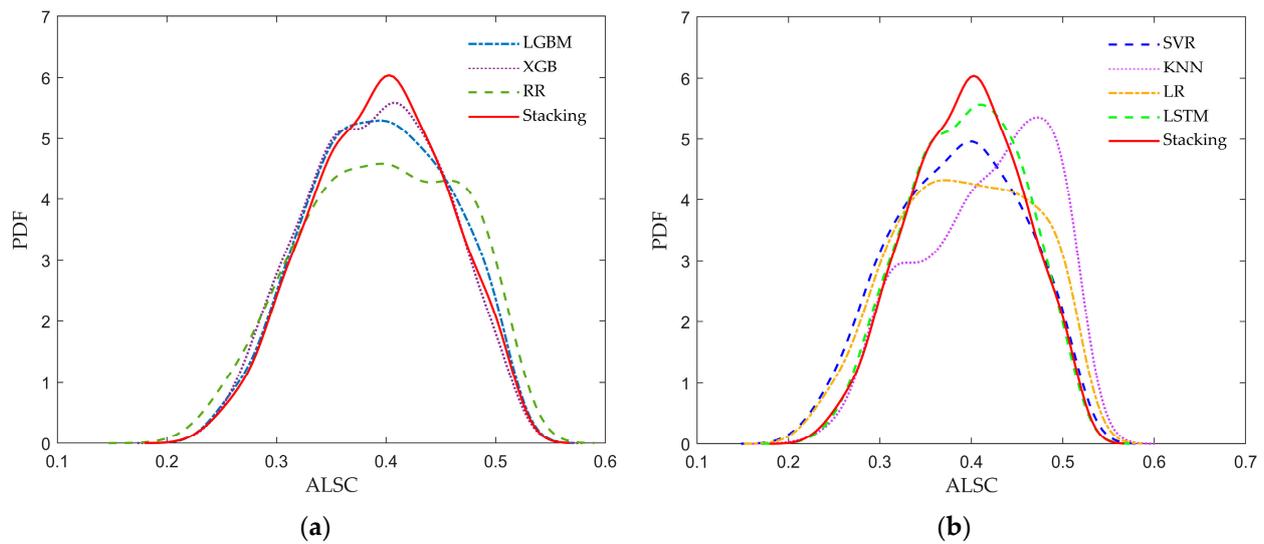


Figure 12. Comparisons on PDF of the ALSC with different forecasting models: (a) Comparisons amongst LGBM, XGB, RR and Stacking; (b) Comparisons amongst SVR, KNN, LR, LSTM and Stacking.

4.2.5. The Probabilistic Forecasting Results of ALSC and Risk Analysis

Using the probabilistic forecasting data of wind power, photovoltaic power and load power can predict the probabilistic ALSC results at future moments. Therefore, we can have an understanding of the load supply capability of the power grid at future moments. More importantly, based on such probabilistic information, we can determine the risk of ALSC threshold overlimit P_{ET} and the probability that ALSC is equal to 0 P_{EZ} , and, based on these indices, an early warning on the operation of power grid in the future can be hence achieved. Twenty-four time points (one day) are selected herein to present the probabilistic forecasting results of ALSC, as shown in Table 13.

Table 13 provides the probabilistic forecasting results of ALSC: expectation of ALSC E_{ALSC} ; standard deviation of ALSC S_{ALSC} ; variation coefficient of β_{ALSC} ; the results of risk: probability of exceeding threshold for ALSC P_{ET} under different thresholds (0.05, 0.10, 0.15, 0.20, 0.25, 0.30); and the probability of ALSC equal to zero P_{EZ} . It can be seen that the fluctuation of renewable energy generation and load over time poses a significant influence on the ALSC. As the forecasting values of wind power P_{wind} , photovoltaic power P_{pv} and load power P_{load} change, the results of probabilistic ALSC evaluation are different for each moment. According to the forecasting results, we can estimate and warn the load supply capacity of the power grid at future moments. For example, from the 1st time point to the 6th time point, the E_{ALSC} is relatively high, the P_{ET} with threshold less than 0.2 equal to 0, and even the P_{ET} with threshold higher than 0.2 are relatively low and the P_{EZ} all equal to 0, illustrating that in this period, the load power supply capacity of the power system is sufficient. But from the 13th to the 18th time point, the E_{ALSC} is low, the P_{ET} almost equal to 1 and the P_{EZ} is high, illustrating that in this period, the load power supply capacity of the power system is seriously inadequate. Therefore, during this period, corresponding measures should be taken to increase the load supply capacity of power system to enhance the security and stability of the power grid. For example, the energy storage facilities such as pumped storage and large-scale battery energy storage systems are influential, and can change the probabilistic distribution of the generation and load power through their charging/discharging characteristics and can also provide power when generation is insufficient, so as to increase ALSC. Specifically, the impact of energy storage facilities on the probabilistic ALSC will be discussed in the next subsection.

Table 13. Probabilistic forecasting results of ALSC.

Time Point	E_{ALSC}	S_{ALSC}	β_{ALSC}	P_{ET} ($\bar{\lambda} = 0.05$)	P_{ET} ($\bar{\lambda} = 0.10$)	P_{ET} ($\bar{\lambda} = 0.15$)	P_{ET} ($\bar{\lambda} = 0.20$)	P_{ET} ($\bar{\lambda} = 0.25$)	P_{ET} ($\bar{\lambda} = 0.30$)	P_{EZ}
1	0.377	0.057	0.151	0	0	0	0	0.002	0.059	0
2	0.400	0.061	0.152	0	0	0	0	0.002	0.025	0
3	0.431	0.066	0.153	0	0	0	0	0	0.008	0
4	0.449	0.068	0.152	0	0	0	0	0	0.003	0
5	0.393	0.059	0.151	0	0	0	0	0	0.037	0
6	0.336	0.049	0.146	0	0	0	0	0.018	0.23	0
7	0.176	0.031	0.179	0	0.006	0.209	0.761	0.992	1	0
8	0.111	0.032	0.289	0.012	0.32	0.872	0.997	1	1	0.012
9	0.010	0.024	2.428	0.859	0.997	1	1	1	1	0.843
10	0.063	0.040	0.633	0.24	0.824	0.993	1	1	1	0.229
11	0.180	0.034	0.187	0	0.007	0.18	0.7	0.977	0.999	0
12	0.130	0.034	0.263	0.006	0.168	0.694	0.977	1	1	0.006
13	0.002	0.012	5.014	0.975	1	1	1	1	1	0.95
14	0.001	0.007	6.060	0.992	1	1	1	1	1	0.961
15	0.001	0.008	5.426	0.989	1	1	1	1	1	0.95
16	0.002	0.011	4.718	0.976	1	1	1	1	1	0.943
17	0.008	0.022	2.798	0.892	0.995	1	1	1	1	0.877
18	0.025	0.035	1.416	0.656	0.975	1	1	1	1	0.649
19	0.067	0.042	0.622	0.228	0.776	0.986	1	1	1	0.221
20	0.108	0.036	0.338	0.034	0.375	0.87	0.998	1	1	0.032
21	0.171	0.035	0.203	0	0.022	0.259	0.796	0.991	1	0
22	0.237	0.037	0.158	0	0	0.006	0.141	0.628	0.935	0
23	0.302	0.045	0.149	0	0	0	0.005	0.096	0.502	0
24	0.342	0.051	0.149	0	0	0	0	0.015	0.21	0

To illustrate the results of probabilistic ALSC forecasting more intuitively, Figure 13 provides the prediction intervals (PIs) of ALSC at different confidence levels (10%, 20%, . . . , 90%). The PI is an interval range constructed around the predicted value of the model given a confidence level, used to indicate the probability that the true target value of a new sample will fall within this interval. The upper and lower bounds of the PIs are obtained through two quantiles. For example, the upper bound of the PI with 90% confidence level is the 0.95 quantile, and the lower bound is the 0.05 quantile. From Figure 13, it can be seen that as the confidence level increases, the PI gradually becomes narrow. In practical applications, the appropriate confidence level should be chosen according to the requirements to guide the operation of power systems.

4.2.6. The Impact of Energy Storage on Probabilistic ALSC

Energy storage systems (ESS) refer to various devices or systems used for keeping the power balance. ESS in power systems can be usually divided into three types according to their installation locations: (1) installed on the generation side, (2) installed on the load side, (3) installed on the grid side. For ESS installed on the generation side, considering the smoothing effect of ESS on the uncertainty of wind turbine and photovoltaic cell outputs [46], it can be assumed that the ESS can reduce the standard deviation of the probability distribution of renewable energy generation. In addition, some energy storage elements such as superconducting magnetic energy storage can discharge instantly and increase power output [47]. Therefore, ESS can also increase the mean of the probability distribution of power output in a short period of time. Both of these effects aim at improving the ALSC. Similarly, for ESS installed on the load side, it can also reduce the expectation and standard deviation of the probability distribution of the load to enhance the ALSC. For ESS installed on the grid side, its charging and discharging characteristics are set to be determined by the power imbalance in the whole system [48]. Specifically, when the power generation is greater than the load demand in the power system, the ESS charges; when

the power generation is less than the load demand, the ESS discharges. The mathematical expression is shown in (53):

$$P_{ESS,i} = \frac{1}{N_{ESS}} (\sum P_{Gi} - \sum P_{Di}) \quad (53)$$

where P_{ESS} is the power of the ESS and N_{ESS} is the number of ESS. As shown in (53), each ESS shares the total power imbalance equally. When the power generation is greater than the load demand, the ESS charges and is treated as a positive load; when the power generation is less than the load demand, the ESS discharges and is treated as a negative load.

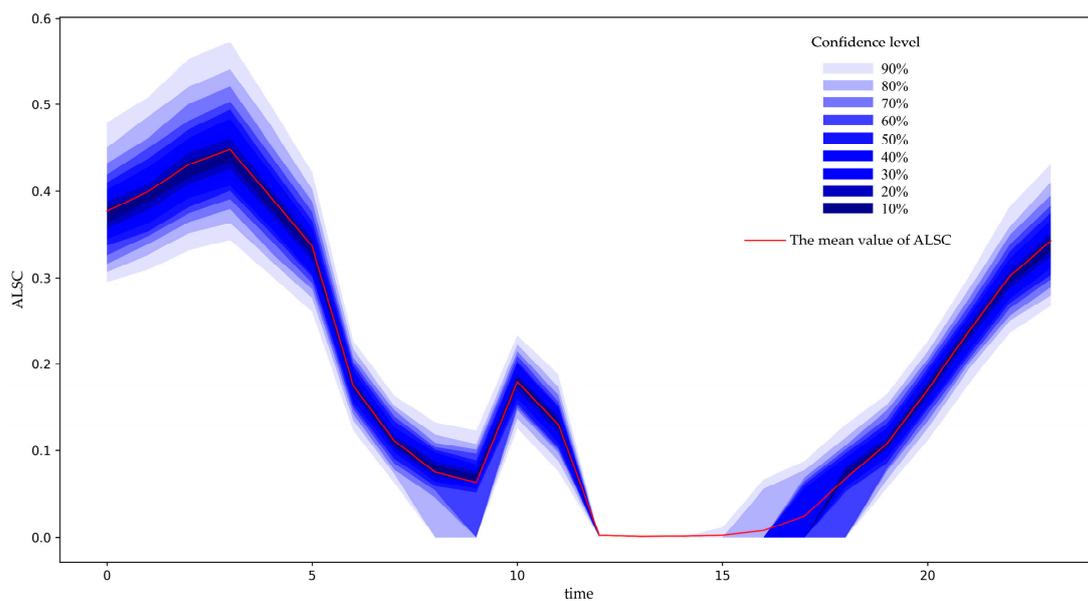


Figure 13. The prediction interval of ALSC at different moments.

In order to study the impact of different installation quantities and locations of the ESS on probabilistic ALSC, take the time 9 in Table 13 as an example; the experimental design is shown in Table 14. All cases are modified based on the base value at time 9. Cases 1~4 only study the case where ESS is installed together with wind turbines. Specifically, in Case 1 and Case 2, ESS only changes the mean of wind power output, checking the impact on ALSC of ESS with the mean changed to different extents. In Case 3, ESS only changes the standard deviation of wind power output, which can study the impact brought by ESS changing different characteristics of the probability distribution of wind power output. In Case 4, ESS changes both the mean and standard deviation, while Case 5 increases the number of ESS, with ESS installed on both wind power generators and photovoltaic power farms. Case 6 and 7 study the case where ESS is installed on the load side. Case 6 focuses on the three loads with close distances, and Case 7 covers the three loads which are far away with one another. It is worth noting that, in order to study the impact of different locations of energy storage on ALSC, the capacities of loads in Case 6 and Case 7 are close to each other, which implies that the energy storage facilities also have similar capacities. Case 8 studies the case where ESS is installed on the grid side.

Table 14. Introduction of eight different cases.

Cases	Description
Case 1	ESS is installed at bus 34, increasing the expectation of wind power to 1.5 times.
Case 2	ESS is installed at bus 34, increasing the expectation of wind power to 2 times.
Case 3	ESS is installed at bus 34, decreasing the standard deviation of wind power to 0.5 times.
Case 4	ESS is installed at bus 34, increasing the expectation of wind power to 2 times and decreasing the standard deviation of wind power to 0.5 times.
Case 5	ESS are installed at bus 30 and 34, increasing the mean of wind power and photovoltaic power to 2 times and decreasing the standard deviations to 0.5 times.
Case 6	ESS are installed at bus 1, 3, 4, decreasing the mean of load to 0.5 times and decreasing the standard deviations to 0.5 times.
Case 7	ESS are installed at bus 1, 15, 28, decreasing the mean of load to 0.5 times and decreasing the standard deviations to 0.5 times.
Case 8	ESS are installed at bus 5, 6, 13, 14 and 19.

Table 15 shows the experimental results on the impact of energy storage on ALSC. Through the results of Case 1 and Case 2, it can be seen that when ESS increases the mean value of wind power output, the expectation of ALSC also becomes larger. Accordingly, the risk index decreases, indicating that ESS plays a role in increasing the load supply capability of the power system. In addition, it can be seen that the more ESS increases the mean value of wind power output, the greater the increase on the expectation of ALSC and the smaller the risk index. This shows that ESS with larger capacity has a stronger regulation capability for integrating new energy generation and can more effectively ensure an adequate load supply capability. Moreover, as the expectation of ALSC rises, the standard deviation of ALSC also increases. This is because the larger mean value of wind power output makes more sample values nonzero, which is also reflected in the P_{EZ} index, thus resulting in a larger standard deviation of ALSC. But in fact, the ALSC in Cases 1 and 2 has a smaller degree of dispersion compared to the original case (Time 9), because they have smaller variation coefficient β_{ALSC} . By comparing Case 3 with the original case and comparing Case 4 with Case 2, it can be found that reducing the standard deviation of wind power output with ESS decreases the degree of dispersion of ALSC; that is, a smaller standard deviation and coefficient of variation. The results of Case 5 indicate that after installing ESS on both the wind power and PV generation side, the level of ALSC will be further improved. The results of Case 6 and Case 7 indicate that installing ESS at three distant loads improves ALSC more significantly. This is probably because the three distant ESSs can cover a larger area in the grid, while the three close ESSs concentrate in one region and impose limited impact on the other buses. Therefore, in practical operation, the most suitable ESS installation location should be selected for improvement of ALSC to the greatest extent. In the end, from the results of Case 8, it can be seen that under reasonable control strategies, ESS installed on the grid side can effectively improve the expectation of ALSC and reduce the degree of dispersion of ALSC. This is because when there is power imbalance in the power system, ESS can be effective at depressing such imbalance very well, thus improving ALSC.

In general, whether ESS is installed on the generation side, load side or grid side, it can always effectively improve ALSC. This is reflected from increasing the expected value of ALSC on one hand, and reducing the degree of dispersion of ALSC on the other hand.

Table 15. Probabilistic results of ALS C under different cases.

Cases	$E_{ALS C}$	$S_{ALS C}$	$\beta_{ALS C}$	P_{ET} ($\bar{\lambda} = 0.05$)	P_{ET} ($\bar{\lambda} = 0.10$)	P_{ET} ($\bar{\lambda} = 0.15$)	P_{ET} ($\bar{\lambda} = 0.20$)	P_{ET} ($\bar{\lambda} = 0.25$)	P_{ET} ($\bar{\lambda} = 0.30$)	P_{EZ}
Time 9	0.010	0.024	2.428	0.859	0.997	1	1	1	1	0.843
Case 1	0.018	0.031	1.735	0.748	0.987	1	1	1	1	0.736
Case 2	0.029	0.037	1.271	0.604	0.959	1	1	1	1	0.597
Case 3	0.010	0.020	2.188	0.865	0.996	1	1	1	1	0.845
Case 4	0.029	0.036	1.253	0.602	0.972	1	1	1	1	0.590
Case 5	0.071	0.038	0.544	0.181	0.773	0.989	1	1	1	0.174
Case 6	0.065	0.043	0.658	0.260	0.775	0.982	1	1	1	0.097
Case 7	0.124	0.032	0.260	0.012	0.177	0.772	0.990	1	1	0.008
Case 8	0.159	0.057	0.360	0.079	0.213	0.213	0.99	1	1	0

5. Conclusions

Firstly, the forecasting models of wind power, photovoltaic power and load are introduced. By adopting the stacking ensemble learning model, the results of multiple base learners are integrated and outputted, which effectively improves the forecasting accuracy. In addition, MIC is used for feature selection to select the optimal set of input variables. Secondly, the forecasting power data is used to predict the probability of ALS C. Specifically, the RPF method considering multi-slack buses is used to calculate the deterministic ALS C problem, and the LHS-MCS method is used to generate probabilistic results. The distributions of wind power, photovoltaic power and load power are determined by the forecasting distributions generated by point forecasting value and error statistics. Consideration of the multi-slack buses makes the RPF calculation more in line with the actual situation, and hence can obtain more accurate and reasonable ALS C evaluation results. Finally, the modified IEEE 39-bused system is used to verify the effectiveness of the proposed method.

The results demonstrate that both the stacking ensemble learning model and the MIC feature selection method can improve the forecasting accuracy. Besides the influence from the prediction of renewable energy and load, the multi-slack buses and the power imbalance allocation principles of multi-slack buses both have an impact on the calculation results of ALS C. Finally, the forecasting results of ALS C probabilistic evaluation can provide a reference for the load power supply capacity of power system in the future and can also provide an early warning for the risk of ALS C threshold overlimit.

There are also some limitations in the research of this paper. For example, the accuracy of ALS C probabilistic forecasting significantly relies on the accuracy of wind power, photovoltaic power and load power forecasting. Therefore, some more advanced algorithms can be adopted in power forecasting to obtain more accurate power predictions.

Author Contributions: Conceptualization, S.L. and J.T.; methodology, Q.S., S.L., Y.X. and Y.Z.; software, Q.S., Y.Z. and J.W.; validation, Q.S., S.L., X.Z. and Y.Z.; formal analysis, S.L., Y.X. and J.W.; investigation, Q.S., S.L. and Y.Z.; resources, Q.S., S.L., Y.X. and X.Z.; data curation, Y.Z. and J.W.; writing—original draft preparation, Q.S. and Y.Z.; writing—review and editing, Q.S., S.L., Y.Z. and J.T.; visualization, Q.S., S.L. and Y.X.; supervision, Q.S., S.L. and J.T.; project administration, Q.S., S.L. and X.Z.; funding acquisition, Q.S., S.L., Y.X. and X.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Science and Technology Program of China Southern Power Grid Co., Ltd., grant number YNKJXM20222174 and the Reserve Talents Program for Middle-aged and Young Leaders of Disciplines in Science and Technology of Yunnan Province, China, grant number 202105AC160014.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available on reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

ALSC	available load supply capability
MIC	maximal information coefficient
LHS	Latin hypercube sampling
RPF	repeated power flow
CPF	continuation power flow
OPF	optimal power flow
PPF	probabilistic power flow
MCS	Monte Carlo simulation
GBDT	gradient boosting decision tree
LGBM	light gradient boosting machine
RR	ridge regression
SVR	support vector regression
SVM	support vector machine
KNN	K-nearest neighbor
LR	linear regression
MAE	mean absolute error
RMSE	root mean square error
MAPE	mean absolute percentage error
PDF	probability density function
PI	prediction interval

References

1. Ali, M.A.; Kamel, S.; Hassan, M.H.; Ahmed, E.M.; Alanazi, M. Optimal Power Flow Solution of Power Systems with Renewable Energy Sources Using White Sharks Algorithm. *Sustainability* **2022**, *14*, 6049. [[CrossRef](#)]
2. Rubino, L.; Rubino, G.; Esemio, R. Linear Programming-Based Power Management for a Multi-Feeder Ultra-Fast DC Charging Station. *Energies* **2023**, *16*, 1213. [[CrossRef](#)]
3. Zhang, S.; Cheng, H.; Zhang, L.; Bazargan, M.; Yao, L. Probabilistic Evaluation of Available Load Supply Capability for Distribution System. *IEEE Trans. Power Syst.* **2013**, *28*, 3215–3225. [[CrossRef](#)]
4. Ejebe, G.C.; Tong, J.; Waight, J.G.; Frame, J.G.; Wang, X.; Tinney, W.F. Tinney Available Transfer Capability Calculations. *IEEE Trans. Power Syst.* **1998**, *13*, 1521–1527. [[CrossRef](#)]
5. Kim, S.-S.; Kim, M.K.; Park, J.-K. Consideration of Multiple Uncertainties for Evaluation of Available Transfer Capability Using Fuzzy Continuation Power Flow. *Int. J. Electr. Power Energy Syst.* **2008**, *30*, 581–593. [[CrossRef](#)]
6. Hakim, L.; Kubokawa, J.; Yuan, Y.; Mitani, T.; Zoka, Y.; Yorino, N.; Niwa, Y.; Shimomura, K.; Takeuchi, A. A Study on the Effect of Generation Shedding to Total Transfer Capability by Means of Transient Stability Constrained Optimal Power Flow. *IEEE Trans. Power Syst.* **2009**, *24*, 347–355. [[CrossRef](#)]
7. Bhaskar, M.A.; Jimoh, A.A. Available Transfer Capability Calculation Using PTDF and Implementation of Optimal Power Flow in Power Markets. In Proceedings of the 2016 IEEE International Conference on Renewable Energy Research and Applications (ICRERA), Birmingham, UK, 20 November 2016; pp. 219–223.
8. Dai, Y.; Xu, W.; Hua, W.; Li, W.; Dong, W.; Zhao, M. Calculation method of thermal stability secure power supply interval considering source-grid-load interaction. *Electr. Power Autom. Equip.* **2023**, *43*, 178–183. [[CrossRef](#)]
9. Ou, Y.; Singh, C. Singh Assessment of Available Transfer Capability and Margins. *IEEE Trans. Power Syst.* **2002**, *17*, 463–468. [[CrossRef](#)]
10. Gravener, M.H.; Nwankpa, C. Nwankpa Available Transfer Capability and First Order Sensitivity. *IEEE Trans. Power Syst.* **1999**, *14*, 512–518. [[CrossRef](#)]
11. Lu, L.; Hou, Y.; Wu, Y.; Xiong, X. An approach for calculating Lerner index to evaluate market power of each generator. *Relay* **2005**, *33*, 9–11.
12. Wang, H.; Guan, Q.; Chen, X.; Wang, B.; Ding, H. Energy Flow Analysis of Combined Natural Gas and Electric System Using the Multi-Balanced Buses Power Flow Calculation. In Proceedings of the 2018 International Conference on Robotics, Control and Automation Engineering, Beijing, China, 26–28 December 2018; ACM: Beijing China, 2018; pp. 5–9.
13. Ou, Z.; Lou, Y.; Wang, J.; Li, Y.; Yang, K.; Peng, S.; Tang, J. The Effect of Power Flow Entropy on Available Load Supply Capacity under Stochastic Scenarios with Different Control Coefficients of UPFC. *Sustainability* **2023**, *15*, 6997. [[CrossRef](#)]
14. Xiang, M.; Yu, J.; Yang, Z.; Yang, Y.; Yu, H.; He, H. Probabilistic Power Flow with Topology Changes Based on Deep Neural Network. *Int. J. Electr. Power Energy Syst.* **2020**, *117*, 105650. [[CrossRef](#)]
15. Wei, P.; Liu, J.K.; Zhou, Q.; Wang, D.J. A Probabilistic Power Flow Algorithm Based on Semi-Variable and Series Expansion. In Proceedings of the 2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA), Beijing, China, 10 March 2017; pp. 563–567.

16. Xiao, H.; Pei, W.; Wu, L.; Ma, L.; Ma, T.; Hua, W. A Novel Deep Learning Based Probabilistic Power Flow Method for Multi-Microgrids Distribution System with Incomplete Network Information. *Appl. Energy* **2023**, *335*, 120716. [[CrossRef](#)]
17. Zheng, H.; Yuan, J.; Chen, L. Short-Term Load Forecasting Using EMD-LSTM Neural Networks with a Xgboost Algorithm for Feature Importance Evaluation. *Energies* **2017**, *10*, 1168. [[CrossRef](#)]
18. Park, S.; Jung, S.; Jung, S.; Rho, S.; Hwang, E. Sliding Window-Based LightGBM Model for Electric Load Forecasting Using Anomaly Repair. *J. Supercomput.* **2021**, *77*, 12857–12878. [[CrossRef](#)]
19. Cordeiro-Costas, M.; Villanueva, D.; Eguía-Oller, P.; Granada-Álvarez, E. Machine Learning and Deep Learning Models Applied to Photovoltaic Production Forecasting. *Appl. Sci.* **2022**, *12*, 8769. [[CrossRef](#)]
20. da Silva, R.G.; Ribeiro, M.H.D.M.; Moreno, S.R.; Mariani, V.C.; Coelho, L. dos S. A Novel Decomposition-Ensemble Learning Framework for Multi-Step Ahead Wind Energy Forecasting. *Energy* **2021**, *216*, 119174. [[CrossRef](#)]
21. Eom, H.; Son, Y.; Choi, S. Choi Feature-Selective Ensemble Learning-Based Long-Term Regional PV Generation Forecasting. *IEEE Access* **2020**, *8*, 54620–54630. [[CrossRef](#)]
22. Wang, Y.; Xu, H.; Song, M.; Zhang, F.; Li, Y.; Zhou, S.; Zhang, L. A Convolutional Transformer-Based Truncated Gaussian Density Network with Data Denoising for Wind Speed Forecasting. *Appl. Energy* **2023**, *333*, 120601. [[CrossRef](#)]
23. Reshef, D.N.; Reshef, Y.A.; Finucane, H.K.; Grossman, S.R.; McVean, G.; Turnbaugh, P.J.; Lander, E.S.; Mitzenmacher, M.; Sabeti, P.C. Detecting Novel Associations in Large Data Sets. *Science* **2011**, *334*, 1518–1524. [[CrossRef](#)]
24. Liu, S.; Cui, Y.; Ma, Y.; Liu, P. Short-Term Load Forecasting Based on GBDT Combinatorial Optimization. In Proceedings of the 2018 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2), Beijing, China, 20 October 2018; pp. 1–5.
25. Wang, Y.; Sun, S.; Chen, X.; Zeng, X.; Kong, Y.; Chen, J.; Guo, Y.; Wang, T. Short-Term Load Forecasting of Industrial Customers Based on SVM and XGBoost. *Int. J. Electr. Power Energy Syst.* **2021**, *129*, 106830. [[CrossRef](#)]
26. Massaoudi, M.; Refaat, S.S.; Chihi, I.; Trabelsi, M.; Oueslati, F.S.; Abu-Rub, H. A Novel Stacked Generalization Ensemble-Based Hybrid LGBM-XGB-MLP Model for Short-Term Load Forecasting. *Energy* **2021**, *214*, 118874. [[CrossRef](#)]
27. Hoerl, A.E.; Kennard, R.W. Ridge Regression: Applications to Nonorthogonal Problems. *Technometrics* **1970**, *12*, 69–82. [[CrossRef](#)]
28. Wei, G.; Yu, X.; Long, X. Novel Approach for Identifying Z-Axis Drift of RLG Based on GA-SVR Model. *J. Syst. Eng. Electron.* **2014**, *25*, 115–121. [[CrossRef](#)]
29. Kramer, O. (Ed.) K-Nearest Neighbors. In *Dimensionality Reduction with Unsupervised Nearest Neighbors*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 13–23, ISBN 978-3-642-38652-7.
30. Su, X.; Yan, X.; Tsai, C.-L. Linear Regression. *WIREs Comput. Stat.* **2012**, *4*, 275–294. [[CrossRef](#)]
31. McRoberts, R.E. Estimating Forest Attribute Parameters for Small Areas Using Nearest Neighbors Techniques. *For. Ecol. Manag.* **2012**, *272*, 3–12. [[CrossRef](#)]
32. Huang, W.; Song, Q.; Huang, Y. Two-Stage Short-Term Power Load Forecasting Based on SSA-VMD and Feature Selection. *Appl. Sci.* **2023**, *13*, 6845. [[CrossRef](#)]
33. Wan, C.; Xu, Z.; Pinson, P.; Dong, Z.Y.; Wong, K.P. Probabilistic Forecasting of Wind Power Generation Using Extreme Learning Machine. *IEEE Trans. Power Syst.* **2014**, *29*, 1033–1044. [[CrossRef](#)]
34. Cao, Z.; Wan, C.; Zhang, Z.; Li, F.; Song, Y. Hybrid Ensemble Deep Learning for Deterministic and Probabilistic Low-Voltage Load Forecasting. *IEEE Trans. Power Syst.* **2020**, *35*, 1881–1897. [[CrossRef](#)]
35. Liu, W.; Xu, Y. Randomised Learning-Based Hybrid Ensemble Model for Probabilistic Forecasting of PV Power Generation. *IET Gener. Transm. Distrib.* **2020**, *14*, 5909–5917. [[CrossRef](#)]
36. Iman, R.L. Uncertainty and Sensitivity Analysis for Computer Modeling Applications. In Proceedings of the Winter Annual Meeting of ASME, Anaheim, CA, USA, 8–12 November 1992.
37. Owen, A.B. Controlling Correlations in Latin Hypercube Samples. *J. Am. Stat. Assoc.* **1994**, *89*, 1517–1522. [[CrossRef](#)]
38. Vořechovský, M.; Novák, D. Correlation Control in Small-Sample Monte Carlo Type Simulations I: A Simulated Annealing Approach. *Probabilistic Eng. Mech.* **2009**, *24*, 452–462. [[CrossRef](#)]
39. Liefvendahl, M.; Stocki, R. A Study on Algorithms for Optimization of Latin Hypercubes. *J. Stat. Plan. Inference* **2006**, *136*, 3231–3247. [[CrossRef](#)]
40. Iman, R.L.; Conover, W.J. A Distribution-Free Approach to Inducing Rank Correlation among Input Variables. *Commun. Stat. Simul. Comput.* **1982**, *11*, 311–334. [[CrossRef](#)]
41. Yang, C.; Du, X.; Xu, D.; Tang, J.; Lin, X.; Xie, K.; Li, W. Optimal Bidding Strategy of Renewable-Based Virtual Power Plant in the Day-Ahead Market. *Int. J. Electr. Power Energy Syst.* **2023**, *144*, 108557. [[CrossRef](#)]
42. Liu, H.; Duan, Z.; Chen, C. Wind Speed Big Data Forecasting Using Time-Variant Multi-Resolution Ensemble Model with Clustering Auto-Encoder. *Appl. Energy* **2020**, *280*, 115975. [[CrossRef](#)]
43. Liu, H. Wind Speed Forecasting Using a New Multi-Factor Fusion and Multi-Resolution Ensemble Model with Real-Time Decomposition and Adaptive Error Correction. *Energy Convers. Manag.* **2020**, *112*, 112995. [[CrossRef](#)]
44. Liu, H.; Duan, Z.; Chen, C.; Wu, H. A Novel Two-Stage Deep Learning Wind Speed Forecasting Method with Adaptive Multiple Error Corrections and Bivariate Dirichlet Process Mixture Model. *Energy Convers. Manag.* **2019**, *199*, 111975. [[CrossRef](#)]
45. Cordeiro-Costas, M.; Villanueva, D.; Eguía-Oller, P.; Martínez-Comesaña, M.; Ramos, S. Load Forecasting with Machine Learning and Deep Learning Methods. *Appl. Sci.* **2023**, *13*, 7399. [[CrossRef](#)]
46. Wan, C.; Qian, W.; Zhao, C.; Song, Y.; Yang, G. Probabilistic Forecasting Based Sizing and Control of Hybrid Energy Storage for Wind Power Smoothing. *IEEE Trans. Sustain. Energy* **2021**, *12*, 1841–1852. [[CrossRef](#)]

47. Ould Amrouche, S.; Rekioua, D.; Rekioua, T.; Bacha, S. Overview of Energy Storage in Renewable Energy Systems. *Int. J. Hydrogen Energy* **2016**, *41*, 20914–20927. [[CrossRef](#)]
48. Almehezia, A.A.; Al-Masri, H.M.; Ehsani, M. Integration of Renewable Energy Sources by Load Shifting and Utilizing Value Storage. *IEEE Trans. Smart Grid* **2019**, *10*, 4974–4984. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.