



Article A Multi-Stage Planning Method for Distribution Networks Based on ARIMA with Error Gradient Sampling for Source-Load Prediction

Sheng Yan * and Minqiang Hu

School of Electrical Engineering, Southeast University, Nanjing 210096, China * Correspondence: yansgcc@foxmail.com; Tel.: +86-138-1141-1788

Abstract: As the scale of distributed renewable energy represented by wind power and photovoltaic continues to expand and load demand gradually changes, the future evolution of the smart distribution network will be directly driven by both distributed generation and user demand. The smart distribution network contains a wide range of flexible resources, and its flexibility and uncertainty will bring great challenges to grid data acquisition and control feedback. To adapt to the precise control and feedback of smart distribution network access equipment under the high proportion of new energy access and to ensure the safe operation of the system, it is urgent to accelerate the study of the evolution of the future distribution grid based on the existing distribution grid. Hence, a multi-stage planning method for distribution networks based on source-load prediction is proposed in this paper. Firstly, a distribution network source-load prediction method based on the autoregressive integrated moving average model (ARIMA) and error gradient sampling is proposed, using ARIMA to predict the scale of source-load development and error gradient sampling based on the generation of source-load scenarios with error intervals. K-means is further used for scenario reduction, to explore multiple operating scenarios of China's distribution network source-load, and the unit's output forecast interval and load demand from 2021 to 2030 for typical regions are derived using rolling forecasts by combining the unit's output, end-demand and clean energy share over the years. Secondly, the planning model of distribution grid evolution in different stages is constructed to analyze the future evolution form of the distribution grid considering the distribution network's load cross-section, respectively, and to provide a development path reference for the future construction of distribution grid form in China.

Keywords: distribution network; source–load prediction; multi-stage planning method; ARIMA; error gradient sampling; data acquisition; precise control; sensory feedback

1. Introduction

With the broad access to new energy sources such as wind power and photovoltaic, and the development of customer-side electricity load, the increase in its controllability has put higher requirements on the operation mode and structure form of the distribution network. The distribution network is no longer a passive network that simply receives and distributes electricity but a primary support platform that carries a large amount of distributed energy access and supports electricity interaction between users in production and sales. Due to the significant uncertainty and volatility of distributed clean energy sources such as wind power and photovoltaic, it is essential to consider not only the planning constraints of component and equipment investments but also the operation of the distribution network under different distributed renewable energy output scenarios in the distribution network planning [1,2]. Not only the traditional investment in transformers and distribution lines should be considered, but also the allocation and utilization of flexible regulation resources such as energy storage and user-side controllable loads to enhance



Citation: Yan, S.; Hu, M. A Multi-Stage Planning Method for Distribution Networks Based on ARIMA with Error Gradient Sampling for Source–Load Prediction. *Sensors* 2022, 22, 8403. https:// doi.org/10.3390/s22218403

Academic Editor: Juan M. Corchado

Received: 30 September 2022 Accepted: 26 October 2022 Published: 1 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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/). the flexibility of the distribution network's operation, realize the consumption of largescale distributed renewable power, and help the whole power system achieve its energy transformation goals [3].

The distribution network is the most active component of the grid in the current technological innovation, and its participation in grid operation is a key part of the evolution of the traditional grid to the smart grid. The widespread access to distributed energy sources such as distributed generation, energy storage, and customer-side controllable resources makes the distribution system no longer a sourceless and passive receiving network in the traditional sense. To achieve effective management and overall integration of distributed energy, energy management concepts and institutions with different coverage and distinct hierarchical characteristics such as microgrids, community microgrids [4], and active distribution networks [5] have emerged. With the coupling characteristics and complementary potential of different energy forms, the active interaction between users and the grid is also gradually emerging. Most studies of transmission and distribution networks mostly consider network planning problems under probabilistic scenarios, focusing on the construction of network planning models, and the uncertainty of source and load is mostly quantified in a probabilistic manner and transformed into deterministic scenarios. Planning studies for distribution networks have focused on the consideration of the flexibility of distributed resources in the network, starting from a single aspect such as the flexibility of source–load resources [6], demand response uncertainty [7], the controllability/uncontrollability of the grid [8,9], and the proactivity/passivity of system control [10–12], to plan the expansion of the grid or equipment of the distribution network. However, the intersection of many factors will undoubtedly further complicate the analysis and planning of future transmission and distribution systems. Therefore, in order to cope with the randomness and diversity of regional growth of source–loads in distribution networks, some studies have started to partition the distribution networks and plan the networks considering multiple factors on the basis of partitioning. A GIS-based dynamic planning method for distribution networks was proposed in [13], and the problem scale was simplified by the established spatio-temporal load distribution model under clustered partitioning. The cost of network planning and construction, operation and maintenance costs can be effectively reduced. The overall framework system and planning process of territorial spatial planning was analyzed in [14], a distribution grid planning process adapted to territorial spatial planning was proposed, and the spatial layout of power facilities in distribution grids was studied. The above studies mostly focus on the spatial layout of the distribution network and the cross-influence of multiple types of factors for the medium and long-term line and equipment planning of the distribution network, but with the large-scale new energy access to the distribution network and the increasing flexibility of the load, it is necessary to take into account the changing trend of the source and load and the scale of development, and realize the multi-stage planning of the distribution network based on the dual-drive digital model. Distribution network planning that has been considered data-driven or data-mode dual-driven mostly considers a single power structure or load type by mining short-term and medium-term data on source and loads as data input for the generation of future sourced load scenarios [15–17]. The problem of load data mining in the field of distribution network planning was addressed in [18], a model for air conditioning load mining and demand-side response evaluation was established, and a distribution network expansion planning model was further proposed to optimize the distribution network planning scheme on the basis of fully exploiting the potential of customer demand response. In [19], a data-driven distribution network planning and investment decision method based on deep learning was proposed to construct a distribution network investment return model and migration learning to solve the problem that deep learning is difficult to use effectively in the case of small samples. Although the above studies consider the development trend prediction of source-load big data but failed to analyze the evolution of the long-term scale source–load development scale trend.

In addition to models and forecasting methods, smart grid optimization techniques can also provide effective assurance of planning results in distribution network planning. The optimal configuration of systems and equipment for smart grids can also effectively help users interact with the grid for accurate data and reliable information, enhancing efficient interaction between the grid and users [20]. At this stage, optimization algorithms for distribution network planning mainly contain particle swarm, heuristic, Benders, and other optimization algorithms [21,22], Using such methods, a two-level optimization search of the planning model is performed to find the optimal planning solution. In [23], a distribution network zoning planning technique based on an improved K-means clustering algorithm for overseas less developed cities was proposed, which can effectively solve the problems of chaotic distribution network planning and uneven distribution of load points in overseas less developed cities. In summary, the exact solution of the planning model can now be achieved by selecting a suitable optimization algorithm for optimal planning of the power grid. In this paper, the MINLP algorithm in GAMS is used to optimally solve the proposed planning model, and the results show that the optimization algorithm can also achieve the exact optimization and solution of the model.

In summary, the uncertainty of source–load is the premise of all factors. The uncertainty of source-load development is quantitatively analyzed to predict its future scale trend. Based on this, the phased planning of the transmission network and distribution network is carried out, and the morphological evolution route of the power grid is finally obtained. There is not enough research in this area. Therefore, a multi-stage planning route for future distribution networks based on the prediction of future distribution network source-load morphology is investigated in this paper. Firstly, the morphological characteristics of distributed generation sources and multiple loads in different stages of the distribution network are analyzed as the basis for source–load forecasting in the distribution network. Secondly, ARIMA forecasting and error gradient sampling methods are used to make multi-year rolling forecasts of source-load data in typical areas of Jiangsu Province to form the source–load variation trends and load deviation intervals from 2020 to 2030. Then, multi-stage distribution network planning models are constructed. Finally, based on the minimum operation scenario of source-load prediction, the changing trend of source-load under different time sections is considered, and the morphological changes in the distribution network under different time nodes are analyzed and elaborated based on the source–load prediction data of Jiangsu province as input, considering the access of electric vehicles, hot and cold equipment, etc., and the morphological evolution route of the future distribution network is derived.

2. Morphological Evolution Mechanism Based on the Source–Load Prediction of Distribution Network

With the continuous development and construction of new power systems, there will be a large number of new energy sources connected to the distribution network in the future, and the uncertainty of new energy sources will have a greater impact on the power grid. At the same time, the scale and type of customer-side load will gradually expand from the original single fixed load to multiple flexible loads. The source-load situation of the distribution network will develop from "source follows load" to "load follows source." Therefore, the morphological evolution of the distribution network will be mainly reflected in the increase, decrease, and expansion of grid power and customer-side equipment. In order to cope with the impact of the uncertainty of the extensive access of new energy sources on the distribution network and the interaction of multiple flexible loads on the customer side, it is necessary to analyze and predict the future source-load trend more accurately, so as to reduce the impact of source-load fluctuations on the distribution network morphology in a long-term scale. Based on the results of different types of load forecasting in typical areas, it is also necessary to consider the actual situation in different areas of the distribution network, including the development of the urban residential population, the investment and expansion of factories and shopping malls, and the growth of diversified loads and other time section nodes. It is necessary to consider the above factors, construct the planning model of the power grid, and form the optimal evolution route of the distribution network with the objective of minimizing the total planning cost during the planning period of the distribution network. Therefore, the framework of morphological evolution analysis of the distribution grid at different stages is shown in Figure 1.



Network at Different Stages

Figure 1. Analysis framework of multi-stage morphological evolutionary route of distribution network.

3. Prediction of Source–Load Morphology in Distribution Networks Based on ARIMA and Error Gradient Sampling

The planning elements of the future distribution network will be richer, and various types of sources and loads will have greater variability in physical properties and dynamic characteristics. In particular, access to distributed generation sources, energy storage, electric vehicles, and other devices often has a strong degree of freedom, which leads to a significant increase in the differences and uncertainties of individual user characteristics, which not only makes it difficult to carry out uniform metrics, but also greatly increases the difficulty of analysis and prediction. In this regard, a specific analysis of various source and load planning elements is needed to condense the common features of a large number of source–load responses, while distinguishing the typical individual differences between different sources and loads.

In the analysis of the morphological evolution of the distribution network, it is necessary to take the trend of the evolution of the unsupported source–load morphology of the distribution network as the premise, while there are more types of source–loads in China's distribution network. In the prediction and analysis of the source–load morphology, it is necessary to effectively grasp the trend of the source–load development and dig into the trend and periodicity of the source–load evolution on the basis of the current stage of the source–load development state. The ARIMA prediction model can exploit the trend of the series to be predicted on the basis of the existing time series data. Compared with the general prediction methods, its model is simple and can effectively take into account the smoothness of time series data to accurately predict the data development trend. The source–load growth rate is introduced to dynamically correct the error interval range of adjacent moments in the error gradient sampling process, and then multiple load scenarios are derived. Based on the load scenarios, the ARIMA model is used to make rolling forecasts of power sources, which eventually form multi-class source–load prediction scenarios, combined with K-means clustering to reduce the scenarios, and then derive the evolution curve of the source–load development trend.

3.1. The Source–Load Prediction Model Based on ARIMA

ARIMA is a time series forecasting analysis method that combines the autoregressive model and moving average model to form ARMIA(p, d, q) model. The input data, i.e., source–load historical data, are transformed into smooth data by difference, and then the dependent variable is regressed only on its own lagged value and the present and lagged values of the random error term to derive the final source–load forecasting results, where *p* is the number of autoregressive terms, *q* is the number of moving average terms, and *d* is the number of differences (order) made to make it a smooth series.

The smoothness of the source–load data from the input history is first tested using the difference method, which is the process of converting non-smooth data to smooth by removing its non-constant trend. Firstly, the source–load data of the input history is smoothed using the difference method defining the source–load series of the *t*th day of the history as x_t and the difference process as the difference between the time series at *t* and t - 1 moments, as shown in Equation (1).

$$\Delta x_t = x_t - x_{t-1} \tag{1}$$

where Δx_t is the first-order difference of the time series at moment *t*. The second-order difference is performed on the basis of the first-order difference to observe whether the data are smooth or not.

$$\Delta(\Delta x_t) = \Delta x_{t+1} - \Delta x_t \tag{2}$$

where $\Delta(\Delta x_t)$ is the second-order difference of the time series at moment *t*.

The difference method is used to determine the difference order d, when the time series tends to be smooth. On this basis, the ARIMA model is constructed, and the orders p and q are determined at the same time. The autoregressive order p and the moving average order q need to be determined in the modeling process of the ARIMA model. p and q are usually selected using the autocorrelation function (*ACF*) and the partial autocorrelation function (*PACF*). *ACF* is used to describe the linear correlation between the predicted and historical values of the time series, and *PACF* is used to describe the linear correlations given intermediate observations. The formulas for *ACF* and *PACF* are as follows.

It is known that the input sequence $X = \{x_1, x_2, x_t, ..., x_n\}$, whose mean is denoted by u as shown in Equation (3), and variance is denoted by σ^2 as shown in Equation (4).

$$u = \frac{1}{n} \sum_{i=1}^{n} x_i \tag{3}$$

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - u)^2$$
(4)

For two different series *X* and *Y* of equal length, the covariance can be used to portray their correlation, and the specific covariance formula is as follows.

$$cov(X,Y) = \frac{1}{n} \sum_{i=1}^{n} (x_i - u_x)(y_i - u_y)$$
(5)

where u_x and u_y denote the mean values of series *X* and *Y*, respectively. The larger the value of the covariance *cov* (*X*, *Y*), the stronger the correlation between the sequences *X* and *Y* (positive correlation when greater than 0, negative correlation when less than 0). Similarly, for the sequence *X*, the corresponding serial self-covariance is calculated based on the number of lags *k* of the sequence, i.e.,

$$c_k = \frac{1}{n-k} \sum_{i=k+1}^n (x_i - u)(x_{i-k} - u)$$
(6)

According to c_k , the autocorrelation coefficient *ACF* can be obtained as:

$$ACF(k) = \frac{c_k}{c_0} = \frac{n}{n-k} \frac{\sum_{i=k+1}^n (x_i - u)(x_{i-k} - u)}{\sum_{i=1}^n (x_i - u)(x_t - u)}$$
(7)

The partial correlation coefficient *PACF* of the series can be obtained from the c_k and Toeplitz matrix as:

$$PACF(k) = (R^{-1}r)_k^{(k)}$$
 (8)

where *R* is the Toeplitz matrix, which can be expressed by Equation (10), and r is the autocorrelation vector, which can be expressed by Equation (11).

$$r_i = \frac{c_i}{c_0}, \ i \in [0, k]$$
 (9)

$$R = \begin{pmatrix} 1 & r_1 & \cdots & r_{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ r_{k-1} & r_{k-3} & \cdots & r_1 \\ r_{k-1} & r_{k-1} & \cdots & 1 \end{pmatrix}$$
(10)

$$r = \begin{pmatrix} r_2 \\ \vdots \\ r_{k-1} \\ r_k \end{pmatrix}$$
(11)

The *ACF* and *PACF* of the model are calculated by choosing the appropriate autocorrelation order p and moving average order q. When the *ACF* converges quickly to 0 after being greater than a certain constant, this constant is called the autoregressive order p. Similarly, the *PACF* converges quickly after being greater than a certain constant, which is the moving average order q; then the model is fitted.

There is subjectivity in selecting p and q using *ACF* and *PACF*, so the Akaike information criterion (AIC) and Bayesian information criterion (BIC) are introduced to choose p and q more objectively, whose *AIC* can be expressed by Equation (12) and *BIC* can be described by Equation (13).

$$AIC = -2\ln(L) + 2K \tag{12}$$

$$BIC = -2\ln(L) + K\ln(n) \tag{13}$$

where *L* denotes the great likelihood function of the model, *K* denotes the number of model parameters, and *AIC* and *BIC* can balance the prediction error and model complexity, and according to the information criterion function, determine the model's order.

After determining the difference order d, the autoregressive order p and the moving average order q, the parameters of the ARMIA(p, d, q) model are formed, based on which the input smoothed posterior time series are predicted as shown in the following equation.

$$x_t = \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \dots + \alpha_p x_{t-p} + u_t$$
(14)

where x_t is the predicted output of the autoregressive model at moment t, x_t , $i \in [t - 1, t - p]$ is the historical input sequence, and u_t is the random perturbation. When the random perturbation term is assumed to be white noise, Equation (14) can be rewritten in the following form.

$$x_t = \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \dots + \alpha_p x_{t-p} + \varepsilon_t \tag{15}$$

where ε_t is the variance of white noise. α_p is the weighting factor of the historical input series. If u_t is not white noise, a moving average of order q is used to represent it. That is:

$$u_t = \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \dots + \beta_q \varepsilon_{t-q}$$
(16)

where ε_t , $i \in [t - 1, t - q]$ denotes the white noise series. β_q is the weighting factor of the white noise sequence. In particular, when $x_t = u_t$, i.e., the current value of the time series is not related to the historical values and depends only on the linear combination of historical white noise, the moving average term is obtained as:

$$x_t = \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \dots + \beta_q \varepsilon_{t-q}$$
(17)

The ARIMA(p, q) is obtained by combining the autoregressive model and the moving average model based on the difference method, as shown in the following equation.

$$x_{t} = \sum_{i=1}^{p} \alpha_{i} x_{t-i} + u_{t} + \sum_{j=1}^{q} \beta_{j} \varepsilon_{t-j}$$
(18)

3.2. Source–Load Scene Generation Based on Error Gradient Sampling Method

In order to generate specific source–load scenarios for the next 10 years, this paper will use the random sampling method to construct source–load scenarios for the next 10 years. However, in the random sampling process the source–load value at the future t + 1 moment will be constrained by the source–load value at the t moment, i.e., after determining the source–load scenario at the t moment, the source–load distribution at the t + 1 moment is not in satisfying the initial predicted distribution, to solve this problem this paper introduces the concept of source–load growth rate, whose specific expression is shown in Equations (19) and (20).

$$r_P^t = \frac{P_{t+1} - P_t}{P_t}$$
(19)

$$r_{P,\min} \le r_P^\iota \le r_{P,\max} \tag{20}$$

where P_{t+1} and P_t are the source–load values at moments t + 1 and t, respectively, and $r_{P,\min}$ and $r_{P,\max}$ are the minimum and maximum growth rates of historical source–loads, respectively. The predicted distribution at the moment t + 1 is dynamically corrected in real time after randomly generating the source–load scenario at moment t, and the specific correction formula is as follows.

$$P_{\min}^{t+1} = \begin{cases} P_{\min,\min}^{t+1} & \text{if } P_{\max}^{t+1} - P_{\min}^{t+1} > r_{P,\min} \\ P_{\max}^{t+1} - r_{P,\min} & \text{else} \end{cases}$$
(21)

$$P_{\max}^{t+1} = \begin{cases} P_{\min,\max}^{t+1} & \text{if } P_{\max}^{t+1} - P_{\max}^{t+1} > r_{P,\max} \\ P_{\max}^{t+1} + r_{P,\max} & \text{else} \end{cases}$$
(22)

8 of 28

where $P_{t+1,\max}$, $P_{t+1,\min}$ and $P_{t+1,\max}$ are the initial predicted maximum, minimum and trend baseline values of the source–load at time t + 1, respectively.

4. Multi-Stage Morphological Evolution Planning Model of Distribution Network

According to the above source–load prediction results, the distribution network can access a variety of flexible adjustment resources at different time points and adjust the operation mode of the network by optimizing its own scheduling or adding flexible adjustment resources. In this chapter, the multi-stage planning model of the distribution network is constructed to analyze the morphological evolution route of distribution network at different time points.

4.1. Objective Function

As the proportion of new energy sources such as wind and solar energy is gradually increasing, and demand-side resources are more widely involved in power grid operation and dispatching, when analyzing the evolution of distribution network, it is necessary to consider the development trend of customer-side resources and wind and solar resources. Based on this, the planning model of distribution network is constructed. The planning model takes the minimum planning cost of the distribution network system as the objective function, and considers the distribution network loss, system equipment investment cost and equipment operation cost, as shown in Equation (23).

$$\min F_1 = C_S + C_M + C_e E_{eloss} + C_Q \tag{23}$$

$$C_{\rm S} = \sum_{j=1}^{N} \frac{C_j Q_j}{(1+i)^n} (1-t_{\rm r})$$
(24)

$$C_{\rm M} = \sum_{n=1}^{L_{\rm P}} \frac{C_{\rm MP}}{\left(1+i\right)^n} (1-t_{\rm r}) \alpha_n \tag{25}$$

$$C_{\rm MP} = r_{\rm M} \sum_{j=1}^{N} C_j Q_j \tag{26}$$

$$C_{\rm Q} = \sum_{n=1}^{L_{\rm P}} \frac{C_{\rm QP}}{\left(1-i\right)^n} (1-t_{\rm r})$$
(27)

$$C_{\rm QP} = C_{\rm r} P_{\rm r} \tag{28}$$

where C_e denotes the time coefficient of power grid loss cost, E_{eloss} denotes the power loss of distribution network, C_s denotes the initial operation cost of power distribution system, C_M denotes the total cost of operation and maintenance of distribution system equipment, C_{MP} is the operation and maintenance cost of the equipment in the first year; nis the total number of equipment that the system needs to build; C_j is the initial investment cost per unit capacity of equipment j; Q_j is the capacity of equipment j; t_r is the tax rate; i is the interest rate; L_p is the equipment planning year; r_M is the maintenance rate of the equipment; α_n is the development degree of the system in the nth year; C_Q is the cost of abandoned wind and light, C_{QP} is the loss of abandoned wind and light in the first year of the system; C_r is the penalty coefficient of abandoned wind and light, and P_r is the amount of abandoned wind and light.

4.2. Constraint Condition

4.2.1. Equipment Capacity Constraint

The capacity constraint of the above involved devices is shown in Equation (29).

$$Q_{\text{EX},k}^{\min} \leq Q_{\text{EX},k} \leq Q_{\text{EX},k}^{\max}$$

$$P_{j}^{\min} \leq P_{j} \leq P_{j}^{\max}$$

$$P_{u,j}^{\min} \leq P_{u,j} \leq P_{u,j}^{\max}$$

$$P_{d,j}^{\min} \leq P_{d,j} \leq P_{d,j}^{\max}$$
(29)

where $Q_{\text{EX},k}$ denotes the actual reserve capacity of the resource k, $Q_{\text{EX},k}^{\min}$ and $Q_{\text{EX},k}^{\max}$ denote the minimum and maximum reserve capacity of the resource k, P_j denotes the actual output of the device j, P_{j}^{\min} and P_j^{\max} denote the minimum and maximum output of the device j, $P_{u,j}$ denotes the actual upward climbing rate of the device j, $P_{d,j}^{\min}$ and $P_{u,j}^{\max}$ denote the minimum and maximum output of the device device j, $P_{d,j}$ denotes the actual upward climbing rate of the device j, $P_{d,j}$ denotes the actual downward climbing rate of the device j, $P_{d,j}^{\min}$ and $P_{d,j}^{\max}$ denote the minimum and maximum downward climbing rate of the device j.

4.2.2. Power Grid Operation Constraint

The distribution network system model mainly includes power balance constraint, nodal voltage constraint, three-phase power flow constraint, line power constraint, generator output constraint, and generator climbing constraint.

V

$$\sum_{j \in \Omega_l^i} P_{ij,t} = \sum_{g \in \Omega_G^i} P_{g,t} + P_{i,t}^{WT} + P_{i,t}^{PV} - P_{i,t}^{\text{indl}} - P_{i,t}^{\text{ecol}} - P_{i,t}^{\text{resl}} + P_{i,t}^{EV,in} - P_{i,t}^{EV,out}$$
(30)

$$\sum_{j\in\Omega_l^i} Q_{ij,t} = \sum_{g\in\Omega_G^i} Q_{g,t} - Q_{i,t}^{EV} - Q_{i,t}^{indl} - Q_{i,t}^{ecol} - Q_{i,t}^{resl}$$
(31)

$$V_i^{\min} \le V_{i,t} \le V_i^{\max}$$
 (32)

$$\sum_{\in \Omega_i^l} P_{ij,t} - V_{i,t} \sum_{j \in \Omega_i^l} V_{j,t} \left(G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij} \right) = 0$$
(33)

$$\sum_{\in \Omega_l^i} Q_{ij,t} - V_{i,t} \sum_{j \in \Omega_l^i} V_{j,t} \left(G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij} \right) = 0$$
(34)

$$-P_{ij}^{\max} \le P_{ij,t} \le P_{ij}^{\max} \tag{35}$$

$$P_g^{\min} \le P_{g,t} \le P_g^{\max} \tag{36}$$

$$Q_g^{\min} \le Q_{g,t} \le Q_g^{\max} \tag{37}$$

$$P_{g,t} - P_{g,t-1} \le RU_g \tag{38}$$

$$P_{g,t-1} - P_{g,t} \le RD_g \tag{39}$$

where $P_{ij,t}$ denotes the transmission power between node *I* and node *j* at time *t*, $P_{g,t}$ denotes the active power of the generator at time *t*, $P_{i,t}^{WT}$ denotes the active power of wind power at node *i* at time *t*, $P_{i,t}^{PV}$ denotes the active power of photovoltaic at node *i* at time *t*, $P_{i,t}^{indl}$ denotes the industrial load power at node *i*, $P_{i,t}^{ecol}$ denotes the commercial load power at node *i*, $P_{i,t}^{resl}$ denotes the residential load power at node *i*, $P_{i,t}^{EV,in}$ and $P_{i,t}^{EV,out}$ denote the charging power and the discharge power of the electric vehicle at node *i* at time *t*, $Q_{ij,t}$ denotes the reactive power between node *i* and node *j* at time *t*, $Q_{g,t}$ denotes the reactive power of the generator at time *t*, $Q_{i,t}^{EV}$ denotes the reactive power of electric vehicle of node *i* at time *t*, $Q_{i,t}^{indl}$ denotes the industrial load reactive power at node *i*, $Q_{i,t}^{ecol}$ denotes the commercial load reactive power at node *i*, $Q_{i,t}^{resl}$ denotes the residential load reactive power at node *i*, $V_{i,t}$ denotes the voltage of node *i* at time *t*, G_{ij} denotes the conductance between node *i* and node *j* at time *t*, B_{ij} denotes the susceptance between node *i* and node *j*, θ_{ij} denotes the voltage phase angle between node *i* and node *j*, P_{ij}^{max} denotes the maximum transmission power, P_g^{min} , P_g^{max} denote the upper and lower limits of the generator active power, respectively, Q_g^{min} , Q_g^{max} denote the upper and lower limits of the generator reactive power, respectively, RU_g , RD_g denote the upper and lower climbing limits of the generator.

5. Morphological Evolution Planning Analysis of Typical Distribution Network Based on Source–Load Prediction

In this paper, the source–load prediction model proposed in Section 4.1 is used to predict the trend of the nodes of the typical example system. The example system is shown in Figure 2.



Figure 2. IEEE14 node distribution system.

In this example, the load demand of each node is equivalently converted from the predicted load of typical areas. Among them, nodes 1, node 5, node 6, node 7, and node 10 are mainly industrial loads, and the total load of nodes 1 and 7 is higher among all nodes. The total load of nodes 2, 8, and 9 is lower, the economic growth of the region represented by node 2 is slower, and the commercial and industrial loads account for a more average share. Nodes 8 and node 9 are relatively small and dominated by commercial loads due to their geographical location; load distribution and the total amount of remaining nodes are more uniform.

Nodes 1, 4, and 7 also have a higher demand for electric vehicles, which is consistent with their larger industrial loads. Due to the impact of the epidemic, the load growth trend of each node is small before 2024, the economy is on track after 2024, and the population of residents in various regions is gradually rising, driving the load growth of the commercial complex. With the morphological evolution of the distribution network, some nodes will be connected to new energy, and urban users will also carry out some demand response behavior of adjustable load.

In order to verify the effectiveness of the method proposed in this paper, two scenarios are set up for comparative analysis:

S1: Single-stage planning scenario of distribution network considering the source–load prediction;

S2: Multi-stage planning scenario of distribution network considering the source–load prediction.

The effectiveness of the proposed method is verified by comparing the results of the two scenarios.

5.1. Prediction Analysis of the Coming Source–Load Morphology in the Distribution Network 5.1.1. Prediction Analysis of the Coming Distributed Generations Morphology in the Distribution Network

As the ARIMA model is not capable of multi-step forecasting, a rolling forecast is used for multi-step forecasting, and a schematic diagram of the rolling forecast is shown in Figure 3.



Figure 3. Rolling forecast diagram.

The annual load, gas turbine output, wind power output, and photovoltaic output data of a region are used to predict the annual load, gas turbine output, wind power output, and photovoltaic output of the region in the next 10 years. The data of load, gas turbine output, wind power output, and photovoltaic output over the years are shown in Figure 4.





In Figure 4, blue is the historical output curve of gas turbines. It can be found that the output of gas turbines shows a downward trend from 2013 to 2019. The red is the wind power output curve, and the wind power output is in a high growth state. The yellow is the photovoltaic output curve, and the photovoltaic output is small, but the overall trend is still growing.

Firstly, the ARIMA model is established for the historical load data, and the ADF unit root is used to test the stability of the original historical load, first-order difference, and second-order difference sequence. The results are shown in Figure 5.



Figure 5. Load data stability detection.

It can be seen from Figure 5 that the *ADF* unit root of the second-order difference sequence is -5.78 much smaller than the *ACF* unit root of the first-order difference and the original sequence, indicating that the stability of the second-order difference is stronger, and it can be seen that the values fluctuate around 0, so the difference coefficient d of ARIMA is 2.

ACF and PACF are used to select *p* and *q* for the ARIMA model. The results of ACF and PACF for the second-order differential load sequences are shown in Figure 6.





It can be seen from Figure 6 that both *ACF* and *PACF* are first-order truncations, that is the value of the sequence becomes very small after the first point, so p and q should be taken as 1. However, due to the subjectivity of using *ACF* and *PACF* to select p and q, p and q are corrected by Equations (12) and (13).

The prediction process of gas turbine output, wind power output, and photovoltaic output in the next 10 years is similar to the appeal process. The relevant process is no



longer given specifically, and the modeling results and prediction results are given directly. The modeling results are ARIMA (1,2,3), ARIMA (1,2,0), ARIMA (0,2,0). The prediction results are shown in Figures 7–9, respectively.

Figure 7. Gas turbine output prediction results. (a) Gas turbine prediction error band. (b) Gas turbine prediction scene generation. (c) Gas turbine output scenarios for the next 10 years after reduction.(d) Actual selected prediction curve of gas turbine.



Figure 8. Wind power output forecast chart. (a) Wind power output prediction error band. (b) Wind power output prediction scenario generation. (c) Wind power output scenarios in the next 10 years after reduction. (d) Actual selected wind power output scenarios.



Figure 9. Photovoltaic output prediction results. (a) Photovoltaic output prediction error band. (b) Photovoltaic output prediction scenario generation. (c) Photovoltaic output scenarios for the next 10 years after reduction. (d) Actual selected photovoltaic output scenarios.

On the basis of the historical gas turbine output curve, combined with the characteristics of the distributed generation at the present stage [24], the prediction trend and error interval of the gas turbine in the next 10 years are formed. Due to the large-scale access to new energy in the future, the utilization rate of primary energy is gradually reduced, and the output of the traditional unit also shows a gradual downward trend. Because the future energy development trend cannot be accurately obtained, the rolling prediction is carried out by using the three times variance of the data of adjacent years, so the error interval will be generated, as shown in Figure 7a. The proposed error gradient sampling method is used to generate the scene of the gas turbine output trend in the error interval, and the K-means clustering is used to reduce the generated scene, as shown in Figure 7b,c. On this basis, the final output curve of the selected gas turbine in the next 10 years is shown in Figure 7d, and the overall output curve shows a downward trend.

There are abundant wind power resources in some areas of China. With the development of new power systems and the breakthrough of wind power technology, the output of wind power in China has generally shown an upward trend in the past 10 years. Based on the analysis of the development trend of wind power [25], ARIMA combined with the principle of three times variance is used to predict the output of wind power in the next 10 years, and the trend and error interval of wind power output in the next 10 years is obtained, as shown in Figure 8. On this basis, the error gradient sampling method is used to generate multiple wind power output trend scenarios in the error interval and K-means clustering is used to reduce them, as shown in Figure 8b,c. The final selected wind power output scenario is shown in Figure 8d. It can be seen that the wind power output will show a rapid upward trend in the next 10 years, but the update in technology can drive the development of wind power resources. Therefore, the growth trend of wind power output will occasionally slow down in the next 10 years.

From the historical data of photovoltaic, it can be found that its historical output curve is similar to that of wind power. ARIMA and three times variance principle are used to predict the photovoltaic output, and the output curve and error interval of photovoltaic in the next 10 years are obtained, as shown in Figure 9a. The error gradient sampling method is used to generate the photovoltaic output scene in the error interval, and the K-means clustering is used to reduce the scene. Finally, the typical photovoltaic output scene is selected, as shown in Figure 9b,d. It can be seen from the diagram that photovoltaic output is similar to wind power output, and it is also growing rapidly in the next 10 years, which is in line with the characteristics of China's future new power system construction. In the future, wind power and photovoltaics will be connected to the power grid on a large scale, gradually occupying the main body of power generation in China.

5.1.2. Prediction Analysis of the Coming Multi-Type Load Morphology in the Distribution Network

The network search method is used to search and traverse between $p \in [0, 5]$ and $q \in [0, 5]$. Finally, the minimum AIC is 72.227 when p = 0, q = 2, so the ARIMA (0, 2, 2) model is finally established. The model is trained, and the load data for the next 10 years are predicted. The results are shown in Figure 10.





From Figure 10, it can be seen that the load in the region will show an increasing trend in the next 10 years, and its value fluctuates within the error range that satisfies the Gaussian distribution. After processing by this method, the final load scenario of node 1 in the next 10 years is shown in Figure 11.



Then, K-means clustering is used to reduce the generated scenarios, and the typical load scenarios in the next 10 years are obtained. The clustering center k is 4, and the typical load scenarios in the next 10 years after reductions are obtained as shown in Figure 12. According to the obtained clustering results, a typical scene is randomly selected for analysis, and the typical scene is selected as shown in Figure 13.



Figure 12. Load scenarios in the next 10 years after reduction.



Figure 13. Actual selected typical load scenarios.

Using the same method, the load scenarios of other nodes in the next 10 years are obtained, and the results are shown in Figure 14.



Figure 14. Load prediction results of each node.

It can be seen from Figure 14 that the load of each node is on the rise, with nodes 1, 4, and 7 showing a large increase and the other nodes showing a slow growth trend in load. At the same time, in order to study the composition of each node load, the proportion of industrial heating load, industrial refrigeration load, other industrial load, commercial heating load, commercial refrigeration load, other commercial load, residential heating load, residential refrigeration load and other residential load of each node from 2010 to 2020 is analyzed, and the proportion distribution of each node load in the next 10 years is analyzed by statistical method. For the load with obvious change trend, the mean and variance of its growth rate are calculated, and the growth rate at time t + 1 is randomly selected by using the 3 σ principle. For the load with no obvious change trend, the mean and variance of its proportion are counted, and the proportion at t + 1 is randomly selected by using the 3 σ principle. The specific formula is as follows:

$$P_{o}^{t+1} = P_{o}^{t} + \operatorname{random}(r_{P,\text{mean}}, r_{P,\text{std}})P_{o}^{t}$$

$$\tag{40}$$

$$P_{\rm no}^{t+1} = {\rm random}(P_{\rm no,mean}, P_{\rm no,std})$$
(41)

where p_o^{t+1} , p_o^t denote the load values with obvious proportion changes at tome t + 1 and t, p_{no}^{t+1} denotes the load value whose proportion change is not obvious at time t + 1, $r_{P,\text{mean}}$ and $r_{P,\text{std}}$ denote the mean and variance of the obvious load growth rate of the proportion change, $P_{\text{no,mean}}$ and $P_{\text{no,std}}$ denote the mean and variance of the load with no obvious change in the proportion, and random (*) is a random function. The proportion of each type of load at node 1 from 2021 to 2030 is obtained through the appeal method, as shown in Figures 15 and 16.



Figure 15. Node 1 load proportion prediction.



Figure 16. Node 1 load prediction by type.

As shown in Figure 15, the load of node 1 in the next 10 years is mainly industrial load, accounting for about 50% of the total load demand. Among them, the proportion of other industrial loads is about 0.35, and the proportion of commercial load and residential load is about 0.25. The proportion of load in node 1 in 10 years is basically unchanged, and the proportion of industrial load has declined. On the contrary, the proportion of commercial load has increased. This is mainly due to the regional load is always the main part of the industrial load, large industrial users occupy a major position. It can be seen from Figure 16 that although the change in proportion of load is small, the annual load demand continues to increase. The total load of node 1 in 2030 is about 27.09% higher than that in 2021, which is mainly due to China 's future economic development and population. From 2021 to 2030, the industrial load will increase by about 19.06%, the commercial load will increase by about 53.89%, and the residential load will increase by about 24.12%. It can be seen that the commercial load will have a greater growth trend in the next 10 years, mainly because the development of the city will eventually be biased towards comfort and user satisfaction. The industrial load will gradually shift to areas with less cost or population, and the commercial load in areas with more population will gradually increase. The industrial, commercial and residential loads only account for a small proportion of the hot and cold loads, and the loads are still dominated by other loads.

5.2. Analysis of Multi-Stage Configuration Results

5.2.1. Multi-Stage Planning Results of Distribution Network

Aiming at the source–load demand from 2021 to 2030 predicted in Section 3.2, based on the IEEE 14-node distribution system, the evolution route of the distribution network in the next ten years is analyzed under the predicted typical source–load scenarios. The node source–load data of the 14-node distribution system is shown in Figure 17.





On the basis of Section 5.3, consider the time cross-section of source–load development, as shown in Figure 18. Combined with the source–load prediction results and development trend in Section 5.2, it can be seen that in 2025, the load demand shows a large increase, compared with the previous year increased by about 25%, and by 2030, the total load demand increased by 40%. Therefore, considering that there is a time cross-section from 2021 to 2030, a multi-stage evolution route of the distribution network is formed. As shown in Tables 1 and 2. The morphological evolution of distribution network is shown in Figures 19–22.



Figure 18. Load growth trend from 2020 to 2030.



Figure 19. Morphological evolution of distribution network in 2021.



Figure 20. Morphological evolution of distribution network in 2024.



Figure 21. Morphological evolution of distribution network in 2025.



Figure 22. Morphological evolution of distribution network in 2028.

| Resource Type | Node | Time/(Year) | Capacity/(MW) |
|--------------------|------|-------------|---------------|
| Electric vehicle | 1 | 2021 | 0.955 |
| | 3 | 2022 | 0.225 |
| | 4 | 2021 | 0.759 |
| | 5 | 2021 | 0.368 |
| | 6 | 2021 | 0.439 |
| | 7 | 2021 | 1.44 |
| | 10 | 2021 | 0.574 |
| | 11 | 2021 | 0.345 |
| | 12 | 2021 | 0.409 |
| | 13 | 2021 | 0.324 |
| | 14 | 2021 | 0.249 |
| Industrial heating | 7 | 2023 | 0.176 |
| Industrial | 1 | 2021 | 0.724 |
| refrigeration | 7 | 2023 | 0.169 |
| Commercial heating | 1 | 2021 | 0.234 |
| | 4 | 2021 | 0.234 |
| | 5 | 2023 | 0.155 |
| | 10 | 2021 | 0.169 |
| | 12 | 2021 | 0.183 |

 Table 1. Resource allocation results of distribution network from 2020 to 2025.

| Resource Type | Node | Time/(Year) | Capacity/(MW) |
|----------------------|------|-------------|---------------|
| | 1 | 2021 | 0.365 |
| | 4 | 2021 | 0.46 |
| Commorcial | 5 | 2021 | 0.392 |
| | 7 | 2021 | 0.381 |
| reingeration | 10 | 2021 | 0.344 |
| | 12 | 2021 | 0.226 |
| | 13 | 2023 | 0.174 |
| | 1 | 2021 | 0.339 |
| Providential brating | 4 | 2021 | 0.285 |
| Residential heating | 11 | 2021 | 0.343 |
| | 12 | 2021 | 0.268 |
| | 1 | 2021 | 0.921 |
| | 3 | 2021 | 0.166 |
| | 4 | 2021 | 0.56 |
| | 5 | 2021 | 0.445 |
| Decidential | 6 | 2021 | 0.326 |
| refrigeration | 7 | 2021 | 0.716 |
| refrigeration | 10 | 2021 | 0.346 |
| | 11 | 2021 | 0.533 |
| | 12 | 2021 | 0.674 |
| | 13 | 2021 | 0.311 |
| | 14 | 2021 | 0.25 |

Table 1. Cont.

Table 2. Optimal allocation results of distribution network resources from 2025 to 2030.

| Resource Type | Node | Time/(Year) | Capacity/(MW) |
|----------------------|------|-------------|---------------|
| | 1 | 2025 | 2.129 |
| | 2 | 2026 | 0.24 |
| | 3 | 2025 | 0.536 |
| | 4 | 2025 | 1.856 |
| | 5 | 2025 | 0.843 |
| | 6 | 2025 | 0.876 |
| | 7 | 2025 | 2.571 |
| Electric vehicle | 8 | 2029 | 0.206 |
| | 9 | 2028 | 0.211 |
| | 10 | 2025 | 1.23 |
| | 11 | 2025 | 0.756 |
| | 12 | 2025 | 0.986 |
| | 13 | 2025 | 0.682 |
| | 14 | 2025 | 0.565 |
| Industrial heating | 7 | 2025 | 0.22 |
| Industrial | 1 | 2025 | 0.866 |
| refrigeration | 7 | 2025 | 0.209 |
| Commercial heating | 1 | 2025 | 0.279 |
| | 4 | 2025 | 0.306 |
| | 5 | 2025 | 0.206 |
| | 7 | 2027 | 0.179 |
| | 10 | 2025 | 0.211 |
| | 12 | 2025 | 0.22 |

| Resource Type | Node | Time/(Year) | Capacity/(MW) |
|---------------------------|------|-------------|---------------|
| Commercial | 1 | 2025 | 0.426 |
| | 4 | 2025 | 0.591 |
| | 5 | 2025 | 0.508 |
| | 7 | 2025 | 0.474 |
| | 10 | 2025 | 0.41 |
| Temperation | 11 | 2029 | 0.154 |
| | 12 | 2025 | 0.26 |
| | 13 | 2025 | 0.222 |
| | 14 | 2027 | 0.179 |
| | 1 | 2025 | 0.4 |
| | 4 | 2025 | 0.39 |
| Posidontial beating | 5 | 2030 | 0.161 |
| Residential heating | 11 | 2025 | 0.423 |
| | 12 | 2025 | 0.322 |
| | 14 | 2028 | 0.166 |
| | 1 | 2025 | 1.073 |
| | 3 | 2025 | 0.209 |
| | 4 | 2025 | 0.755 |
| Residential refrigeration | 5 | 2025 | 0.558 |
| | 6 | 2025 | 0.366 |
| | 7 | 2025 | 0.86 |
| | 10 | 2025 | 0.429 |
| | 11 | 2025 | 0.651 |
| | 12 | 2025 | 0.801 |
| | 13 | 2025 | 0.392 |
| | 14 | 2025 | 0.324 |

Table 2. Cont.

As a result of the double carbon target, China has started to reduce carbon emissions, and with cars as the main greenhouse gas emitters, it is an inevitable trend to study the use of clean energy electric vehicles. Therefore, in the future evolution of the distribution network, electric vehicles will participate in power grid planning as the main load. By comparing Figures 19–22, considering the load growth cross-section in 2025, it can be found that the increase in electric vehicles, industrial and commercial refrigeration, and heating equipment shows a certain timeliness, mostly concentrated in 2021 and 2023. This is because before 2025, the load change trend is small, and the electric vehicles, industrial, commercial and residential refrigeration and heating equipment configured in 2021 can effectively meet the overall operation of the distribution network. The power grid can rely on its own dispatching operation to ensure the balance of load supply and demand, and there is no need to configure new equipment. In 2025, as the load shows massive growth and the scale of new energy is expanding before 2025 and new energy is only considered for local consumption, but after 2025, the installed capacity of new energy increases massively, and the original local balance has been unable to meet the consumption of new energy such as wind power and photovoltaic. On the basis of adding electric vehicles, and industrial, commercial, and residential refrigeration and heating equipment, new energy gradually forms cross-regional return with external power grids to promote the consumption of new energy. It can be seen from Table 2 that in the distribution network planning from 2025 to 2030, the configuration scheme of the electric vehicle charging station of node 8 and node 9 is added, and the residential heating equipment is added at nodes 5 and 14.

5.2.2. Comparative Analysis of Different Scene Configuration Results

Based on the above-proposed distribution network source–load prediction results, the distribution network is planned in a single phase without considering the increase or decrease in source–load in a specific year. The development of the distribution network is always accompanied by changes in energy, the alternation of new and old towns, and the addition of emerging entities (such as factories, load aggregators, residential communities, etc.). In view of the changes in node load demand and the trend of total node load under the access of different entities, considering the configuration of electric vehicles, industrial, commercial, and residential heating equipment and refrigeration equipment, the results of resource optimization configuration of the distribution network from 2021 to 2030 are shown in Table 3.

Through the comparison of planning costs in the S1 and S2 scenarios, it can be found that the total planning cost in the S1 scenario is about CNY 121,900.69, and the planning cost in the S2 scenario is CNY 117,289.49. Compared with the S1 scenario, the S2 planning cost considering the multi-stage planning scenario is reduced by about 3%, indicating that multi-stage planning can better improve the economy of planning compared with single-stage planning. From Table 3, it can be found that when the single-stage planning is carried out without considering the change in the load cross-section, compared with the configuration results of S2, most of the equipment in the S1 scenario is mainly configured in the first year of planning, and the configuration capacity is large. This is mainly because in the single-stage planning, the time cross-section change is not considered, and the cost of the entire planning cycle is minimized. The majority of the capacity allocation in the first year of planning allows for optimal system planning costs. Electric vehicle charging stations are planned at each node, and mostly in the first year. According to the load prediction results of 5.1, it can be seen that the electric vehicle load of each node of the system shows an increasing trend year by year, while the electric vehicle load of node 1 and node 7 accounts for a large proportion of the electric vehicle load of the whole system. Calculating the proportion of electric vehicles to the total node load for all nodes shows that there is less variability in the proportion of electric vehicles at each node, although node 1 and node 7 have a much larger electric vehicles allocation capacity than the other nodes, this is also a result of the total load at node 1 and node 7 having a larger demand relative to the other nodes, and therefore nodes 1 and 7 have a much larger electric vehicles charging station allocation capacity than the other nodes.

The configuration of industrial equipment is mostly concentrated on node 1 and node 7, which is mainly due to the large proportion of industrial load of these two nodes, and the relative proportion of industrial heat load demand of node 7 is relatively large. However, in the early stage of planning, the thermal load demand of nodes can be met through the dispatching instructions of power grid, so the industrial heating equipment will not be configured until 2023. For industrial refrigeration, the cooling load demand of node 1 is high, and the refrigeration equipment configuration has been carried out in 2021. When configuring equipment for commercial loads, nodes 3, 8, and 9 are considered to have a small geographical area and a smaller overall share of the load required relative to the other nodes, while nodes 4 and 5 have a larger demand, so these two nodes are allocated a large proportion of the capacity.

Comparing the configuration results of the S1 and S2 scenarios, namely Tables 2 and 3, it can be found that compared with the single-stage planning results, when considering 2025 as the time cross-section, the distribution network does not need to plan most of the capacity in the first year, which can effectively reduce the cost of the equipment configuration. Before the second stage configuration, the system can balance the source–load through its own optimal scheduling. When the line reaches the maximum capacity or the load demand increases significantly, the second configuration is performed. Compared with the overall configuration of the single-stage configuration in the first year, the equipment configuration considering the time cross-section of the distribution network can effectively reduce the planning cost of the distribution network and flexibly adjust the planning scheme of the power grid.

| Resource Type | Node | Time/(Year) | Capacity/(MW) |
|---------------------------|------|-------------|---------------|
| | 1 | 2021 | 2.129 |
| | 2 | 2026 | 1.288 |
| | 3 | 2022 | 0.536 |
| | 4 | 2021 | 1.856 |
| | 5 | 2021 | 0.843 |
| | 6 | 2021 | 0.876 |
| | 7 | 2021 | 2.571 |
| Electric vehicle | 8 | 2029 | 0.206 |
| | 9 | 2028 | 0.211 |
| | 10 | 2021 | 1.23 |
| | 11 | 2021 | 0.756 |
| | 12 | 2021 | 0.986 |
| | 13 | 2021 | 0.682 |
| | 14 | 2021 | 0.565 |
| Industrial heating | 7 | 2023 | 0.22 |
| Industrial | 1 | 2020 | 0.866 |
| refrigeration | 1 7 | 2021 | 0.000 |
| Temperation | 7 | 2023 | 0.209 |
| | 1 | 2021 | 0.279 |
| | 4 | 2021 | 0.306 |
| Commercial heating | 5 | 2023 | 0.206 |
| Commercial nearing | 7 | 2027 | 0.179 |
| | 10 | 2021 | 0.211 |
| | 12 | 2021 | 0.22 |
| | 1 | 2021 | 0.426 |
| | 4 | 2021 | 0.591 |
| | 5 | 2021 | 0.508 |
| Communial | 7 | 2021 | 0.474 |
| Commercial | 10 | 2021 | 0.41 |
| refrigeration | 11 | 2029 | 0.154 |
| | 12 | 2021 | 0.26 |
| | 13 | 2023 | 0.222 |
| | 14 | 2027 | 0.179 |
| | 1 | 2021 | 0.4 |
| | 4 | 2021 | 0.39 |
| Desidential besting | 5 | 2030 | 0.161 |
| Residential heating | 11 | 2021 | 0.423 |
| | 12 | 2021 | 0.322 |
| | 14 | 2028 | 0.166 |
| Residential refrigeration | 1 | 2021 | 1.073 |
| | 3 | 2021 | 0.209 |
| | 4 | 2021 | 0.755 |
| | 5 | 2021 | 0.558 |
| | 6 | 2021 | 0.366 |
| | 7 | 2021 | 0.86 |
| | 10 | 2021 | 0.429 |
| | 11 | 2021 | 0.651 |
| | 12 | 2021 | 0.801 |
| | 13 | 2021 | 0.392 |
| | 14 | 2021 | 0.324 |
| | | | |

Table 3. Load configuration results from 2021 to 2030.

5.3. Analysis of Distribution Network Dispatch Results during No New Equipment

From the above analysis, it can be found that the distribution network equipment is only configured in a certain year, and in the unconfigured year, the power grid needs to reasonably plan the power output through its own dispatching operation to achieve the balance of supply and demand of each node in the network. The power output curve of



the interaction between the distribution network and the external power grid in each stage year is shown in Figure 23.

Figure 23. Interactive power between distribution network and main network from 2021 to 2024 years.

It can be seen from the Figure 23 that after the capacity configuration of the distribution network is completed in 2021 because the load is far less than the configuration capacity, the new energy unit has additional output to help the main network to consume; after that, due to the continuous growth of the load, the power output in the distribution network is not enough to meet the load demand. At this time, it is necessary to purchase electricity from the main network to provide the node load, so the interactive power shows an increase. In 2025 form Figure 24, due to the new addition and expansion of node equipment in the distribution network, the pressure on new energy output is alleviated, and the interactive power is reduced compared with 2024, but it still shows an overall growth trend.



Figure 24. Interactive power between distribution network and main network from 2025 to 2030 years.

Between 2020 and 2024, the power output of the distribution network is shown in Figure 25.

It can be seen from the above that 2024 is a time cross-section of load change. Before 2024, the load growth trend is relatively flat. According to Table 1, compared with the single-stage planning evolution, the equipment configuration capacity is smaller. Therefore, the power output is dominated by the gas turbines at node 1 and node 2, with a smaller output from the new energy units, and the output of each unit can meet the load demand at each node of the distribution network.



Figure 25. Power output of each power supply in the distribution network from 2020 to 2024.

It can be seen from Figure 26 that the load will increase greatly in 2025, which is the result of the economic recovery after the domestic epidemic and the construction of new towns. Therefore, most of the equipment expansion or new equipment is carried out in 2025 to meet the demand of power grid load. At the same time, it can be seen from Figure 26 that traditional units are gradually replaced by new energy units, and new energy units gradually become the main power supply on the supply side. The power supply in each year is combined by the units to create an overall output to meet the load variations of the distribution network.



Figure 26. Power output of each power supply in the distribution network from 2025 to 2030.

6. Conclusions

In this paper, ARIMA and error gradient sampling methods are used to predict the future power structure and load demand of the distribution network, and then the multistage planning of a typical distribution network is carried out based on the prediction results and planning model of the coming source–load morphology. The conclusions obtained are shown below.

1. The ARIMA prediction model can be used to effectively explore the trend of power supply and load and take into account the smoothness of the prediction sequence. The error gradient sampling method can effectively divide the error interval formed

by the prediction. The combination of ARIMA and error gradient sampling method can effectively predict the trend of the power structure and load demand scale;

- 2. The proposed multi-stage planning model of a typical distribution network based on source–load prediction can realize the capacity planning of different types of equipment in different stages of the distribution network on the basis of reducing the planning cost. Compared with single-stage planning, it can effectively reduce the cost of equipment investment and adjust the planning scheme of the distribution network more flexibly;
- 3. This paper mainly considers the influence of changes in source–load development on the rigorous route of distribution network planning. With the development and innovation of new energy and load technology in the future and the change in electricity demand and resources of multi-type users, the future distribution network planning also needs to combine the factors of scientific and technological development and the driving conditions of resources and electricity demand, so as to form a comprehensive and flexible planning of distribution network.

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

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

Abbreviations

| Abbreviated Acronyms | Noun Full Name |
|----------------------|--|
| ARIMA | Autoregressive Integrated Moving Average model |
| ACF | Autocorrelation function |
| PACF | Partial autocorrelation function |
| AIC | Akaike information criterion |
| BIC | Bayesian information criterion |
| ADF | Automatic direction finder |
| | |

References

- 1. Kang, C.; Yao, L. Key scientific issues and theoretical research framework for power systems with high proportion of renewable energy. *Autom. Power Syst.* 2017, 41, 2–11.
- Xing, H.; Cheng, H.; Zhang, S.; Zhang, Y. Review of Active Distribution Network Planning. *Power Syst. Technol.* 2015, 39, 2705–2711. [CrossRef]
- 3. Shao, C.; Wang, X.; Wang, X.; Wang, B. Probe into Analysis and Planning of Multi-energy Systems. Proc. CSEE 2016, 36, 3817–3829.

4. Zhang, H.; Cheng, H.; Zeng, P.; Zhang, J.; Lu, J.; Li, C. Overview of Transmission Network Expansion Planning Based on Stochastic Optimization. *Power Syst. Technol.* **2017**, *41*, 3121–3129.

- Liu, L.; Cheng, H.S.; Wu, Y.W.; Chen, H.Y.; Zhang, N.; Wang, Z.D. Research Progress and Prospects of Transmission Expansion Planning Method for High Proportion of Renewable Energy. *Autom. Power Syst.* 2021, 45, 176–183.
- Xu, J.; Wang, T.; Wang, J.; Liu, S.; Ma, Y. Distribution Network Planning Method Considering Flexibility Resource Attribute of Microgrids. *Electr. Power Constr.* 2022, 43, 84–92.
- Qi, X.; Wang, P.; Zhang, F. Fuzzy Credibility Constraint Planning of Distribution Network Considering the Uncertainty of Demand Response. South. Power Syst. Technol. 2022, 1–12.
- 8. Peng, J.; Wang, J.; Qi, F.; Liao, Y. Multi-stage expansion planning of a distribution network with double-carbon policy. *Power Syst. Prot. Control.* **2022**, *50*, 153–161.
- 9. Ma, X. Planning Strategy of Distribution Network in Urban Center Based on Grid Concept. Mod. Ind. Econ. Inf. 2021, 11, 198–200.

- Wu, H.; Sun, L.; Xiang, S.; Yuan, Y. Expansion Planning of Active Distribution Network Considering High-dimensional Temporal Correlation Between Renewable Energy and Load. *Autom. Power Syst.* 2022, 46, 40–51.
- 11. Xu, Z. Coordinated Dispatch and Planning of Multi-Microgrid and Active Distribution Network within Market Environment. Ph.D. Thesis, Southeast University, Jiangsu, China, 2021. [CrossRef]
- 12. Zhao, B.; Wang, C.; Zhou, J.; Zhao, J.; Yang, Y.; Yu, J. Present and Future Development Trend of Active Distribution Network. *Autom. Power Syst.* **2014**, *38*, 125–135.
- 13. Li, Z.; Liu, Y.; Ren, P.; Xiang, Y.; Liu, Q.; Qi, H. Multi-stage Dynamic Planning Method for Distribution Network Considering Clustering and Partitioning of Geographic Information. *Autom. Power Syst.* **2022**, *46*, 38–45.
- 14. Zhang, Q.; Zhu, C.; Gao, Q. Research on Spatial Layout Planning Strategy of Power Distribution Facilities Based on Three-levels and Two-categories System. *Power Capacit. React. Power Compens.* **2022**, *43*, 77–82.
- 15. Yuan, Y.; Wang, H.; Pang, D.; Wang, D. Research on the application of big data in distribution network planning. *Electron. Compon. Inf. Technol.* **2022**, *6*, 40–42.
- Hu, B.; Zhao, S.; Pang, W. Construction of planning model for AC/DC hybrid active distribution network based on big data technology. *Power Syst. Big Data* 2020, 23, 38–46.
- 17. Guo, W.; Xu, J.; Wang, X.; Ding, L.; Zhu, J. Application of Internet Big Data Technology in Intelligent Planning of Distribution Network. *Electr. Power Equip. Manag.* **2020**, *9*, 200–202.
- Su, N.; Li, X.; Liang, H.; Gong, X.; Lv, F.; Xi, Y.; Qi, B. Distribution Network Expansion Planning Method Based on Air-Conditioning Load Data Mining. *Electr. Power Constr.* 2021, 42, 100–107.
- 19. Kong, X.; Sun, Y.; Yao, J.; Liu, M.; Cui, K. A Data-driven Investment Decision Method for Distribution Network Planning. *Power* Supply Consum. 2021, 38, 3–11.
- 20. Recioui, A.; Bentarzi, H. Optimizing Smart Grid Operation and Control; IGI Global: Hershey, PA, USA, 2021.
- Zhong, L.; Wang, Z.; Yu, T.; Liu, Y.; Peng, B.; Zeng, J. Planning Method for AC/DC Hybrid Distribution Network Based on Explicit Expression of Reliability. *Autom. Electr. Power Syst.* 2022, 46, 52–61.
- Yang, D.; Tu, Z.; Xu, J.; Ding, J.; Yang, J.; Xiao, Y. Bi-level planning of active distribution network considering low carbon emissions and demand side management. *Eng. J. Wuhan Univ.* 2022, 55, 709–715.
- 23. Yao, J.; Tang, B.; Liu, Z. The distribution network planning of underdeveloped overseas cities based on improve K-means clustering algorithm. *Electron. Meas. Technol.* **2021**, *44*, 54–60.
- Dong, X.; Hua, Z.; Sahng, L. Morphological Characteristics and Technology Prospect of New Distribution System. *High Volt. Eng.* 2021, 47, 3021–3035.
- Xu, D.; Bai, X.; Sun, J. Characteristic Analysis of Wind Energy Resources in Ulanqab Wind Power Base. Northwest Hydropower 2022, 1, 87–89.