



Article Multi-Objective Interval Prediction of Load Based on the Conditional Copula Function

Gang Zhang ¹,*, Zhixuan Li¹, Jinwang Hou¹, Kaoshe Zhang ¹, Fuchao Liu² and Xin Zhang ³

- ¹ Institute of Water Resources and Hydro-Electric Engineering, Xi'an University of Technology, Xi'an 710048, China; 15594818128@163.com (Z.L.); hjw20190124@163.com (J.H.); zhangks@263.net (K.Z.)
- ² State Grid Gansu Electric Power Company, Gansu Electric Power Research Institute, Lanzhou 730050, China; liufc8127@sina.com
- ³ Shaanxi Electric Power Trading Center Co., Ltd.; Xi'an 710048, China; zhangxin8459@126.com
- * Correspondence: zhanggang3463003@xaut.edu.cn

Received: 24 January 2019; Accepted: 1 March 2019; Published: 6 March 2019



Featured Application: As the load characteristics of power systems tend to be complex, the difficulty of accurate and reliable load point prediction is constantly increasing, and it is more and more difficult to only use the point prediction results to ensure the safe and stable operation of the power system. By using the multi-objective interval prediction algorithm based on the conditional copula function, the power system decision-makers can better understand the fluctuation range of future load changes, so that they can clearly recognize the possible uncertainties and facing risk factors when they are performing tasks such as real-time scheduling and system security analysis of the future load, and then make more reasonable, safe, and economical scheduling decisions in a timely manner, and more in line with the future development needs of the power market.

Abstract: Compared with the point prediction, the interval prediction of the load could more effectively guarantee the safe operation of the power system. In view of the problem that the correlation between adjacent load data is not fully utilized so that the prediction accuracy is reduced, this paper proposes the conditional copula function interval prediction method, which could make full use of the correlation relationship between adjacent load data so as to obtain the interval prediction result. At the same time, there are the different prediction results of the method under different parameters, and the evaluation results of the two accuracy evaluation indicators containing PICP (prediction interval coverage probability) and the PIAW (prediction interval average width) are inconsistent, the above result that the optimal parameters and prediction results cannot be obtained, therefore, the NSGA-II (Non-dominated Sorting Genetic Algorithm-II) multi-objective optimization algorithm is proposed to seek out the optimal solution set, and by evaluating the solution set, obtain the optimal prediction model parameters and the corresponding prediction results. Finally, the proposed method is applied to the three regions of Shaanxi Province, China to conduct ultra-short-term load prediction, and compare it with the commonly used load interval prediction method such as Gaussian process regression (GPR) algorithm, artificial neural network (ANN), extreme learning machine (ELM) and others, and the results show that the proposed method always has better prediction accuracy when applying it to different regions.

Keywords: load prediction; interval prediction; conditional copula function; multi-objective optimization algorithm

1. Introduction

With the rapid development of China's economy, the power industry, which is transitioning from a monopoly business model to a competitive relationship, has put forward higher requirements for all departments of the power system [1–3]. The power sector conducts a comprehensive and detailed study of the data related to load prediction, formulates an efficient and economical power generation plan, and rationally arranges the output of the unit so that it can continuously provide users with safe and reliable power, meet the needs of each user, ensure the safety and stable operation of the power system, reduce power generation costs, and improve economic efficiency [4,5]. Power system load prediction involves studying the change law of historical load data and finding out the relationship between load and various influencing factors, so as to predict future load. Load prediction can be generally be divided into the following four categories according to the length of the forecast: long-term, medium-term, short-term, and ultra-short-term. Long-term load prediction is mainly used to predict the load situation in the next several years, generally used for grid planning and reconstruction work [6,7]; medium-term load prediction refers to prediction the load in the next few months to one year, mainly for reservoirs' operation scheduling, unit maintenance, and usage planning for fuel [8–10]; short-term load prediction is mainly for the next day to one week of load forecast, often used for optimizing the combination of water and thermal power and control of economic flow [11-13]; ultra-short-term load prediction is a hot spot of research in recent years [14,15], and mainly refers to load prediction for the next few minutes. Usually used for power quality control, safety monitoring for online operation, prevention and emergency control, its prediction accuracy directly affects economic dispatch, online safety monitoring, automatic generation control (AGC) frequency modulation, and preventive control emergency situations [16].

Since the early 1960s, load prediction has begun to develop toward exploration, research, and application. At present, most researchers use some methods related to artificial neural network (ANN) and support vector machine (SVM) [17–20]. Some scholars have proposed selecting a similar daily load to the input load, and then applying wavelet decomposition to decompose the load into low-frequency components and high-frequency components. Finally, a single neural network is used to predict the future load of these two components. Some scholars have used asymmetric quadratic loss functions to support vector regression to accurately predict the load, which effectively improves the accuracy of the power load prediction model. In the literature [21], a new hybrid intelligent algorithm based on Wavelet Transform (WT) and Fuzzy Adaptive Resonance Theory Maps (F-ARTMAP) is proposed. The prediction model has been proven by extensive comparison of predictions. In the literature [22], the empirical mode decomposition method is used to divide the time series into two parts, describing the energy consumption values of the trend and the local oscillation, respectively; these are then used for training the support vector regression model. In the literature [23], for problems applying the SVM to short-term load prediction, such as data preprocessing, kernel function construction and selection and parameter optimization, it does analysis, summarizes the existing solutions, and proposes the key issues to be solved in the next step. Some scholars have studied the electric load prediction problem in local areas [24] and proposed a new multi-model variable structure load prediction method based on adaptive clustering partitioning and support vector regression. Experiments show that this prediction method has higher precision and stronger robustness than the traditional neural network prediction method.

At present, all the traditional load prediction methods can only obtain the determined point prediction result, that is, only a certain value is given; the disadvantage is that the possible fluctuation range of the prediction result cannot be determined. In fact, the power system contains various uncertain factors, which makes decision-making risky. Therefore, the uncertainty of the power system must be fully considered in the decision-making, and the interval prediction is more in line with the objective requirements [25]. In addition, as the power system load characteristics tend to be complex, the difficulty of accurate and reliable load point prediction is increasing, and it is more difficult to only use the point prediction results to ensure the safe and stable operation of the power system [26].

Compared to deterministic point prediction, the result of probabilistic interval prediction is not a simple deterministic value, but an interval. Based on this result, power system decision-makers can better understand the fluctuation range of future load changes, so that they can clearly recognize the potential uncertainties and risk factors of the future load when performing real-time scheduling and system safety analysis, and then make more reasonable, safe, and economical scheduling decisions [27], and also be more in line with the future development needs of the power market [28]. At present, many scholars have explored and studied the load interval prediction and formed a variety of prediction methods [26–31]. They can be summarized as follows: the upper and lower bound estimation method based on ANN [29], the mixed construction interval method [26–30], and the probabilistic prediction method [28–31]. The upper and lower bound estimation method based on ANN directly adopts the multi-input and double-output structure of feedforward neural network and makes the objective function optimal to obtain the prediction interval. In the literature [29], this method is applied to load interval prediction, and the particle swarm optimization algorithm is used to optimize the parameters in the ANN to improve the quality of the prediction interval. The hybrid construction interval method is usually based on a certain point prediction method, and then the interval results are constructed by error analysis. The authors of Reference [26] proposed a method based on the point prediction method of the extreme learning machine to construct the proportional coefficient so as to predict the short-term load interval. It has a wide application range and strong robustness; the literature [30] gives a prediction method based on multi-layer wavelet ANN, which constructs a load prediction interval by adding a hybrid Kalman filter. However, this kind of method is to perform interval prediction based on the traditional point prediction result, and the effect of point prediction will directly affect the effect of interval prediction. The probabilistic prediction method is mostly based on Bayesian theory [31]. This method mainly proceeds by constructing a distribution model of the predictive quantity, which can directly obtain the expected value of the predicted quantity, and its distribution, so that interval prediction results can be more conveniently obtained under any confidence level. The literature [28] also describes a brand-new machine learning algorithm based on Bayesian theory and statistical learning theory. Compared with the traditional interval prediction method, it has the advantages of strong generalization ability of machine learning algorithms. However, in practical applications, using different kernel functions would obtain different prediction results; when conducting an accuracy evaluation on the prediction results with different kernel functions, the evaluation results of the evaluation indicators often prove to be contradictory, resulting in an inability to obtain the optimal kernel function and prediction results.

The copula function can describe the correlation relationship between multiple variables more comprehensively and accurately; due to this excellent characteristic, some scholars have tried to apply it in the field of wind power prediction in recent years [32,33]. This method has a fast calculation speed and relatively good prediction effect, but there are also some problems, which are manifested in the following aspects: (1) In the research, relevant scholars did not consider the value boundary problem of conditional copula function prediction model parameters (interval division number *K*, condition number *t*); (2) different parameter values (*K*, *t*) will result in interval prediction outcome with different precision, and the values of inappropriate parameters will reduce the prediction accuracy; (3) interval accuracy evaluation index containing the PICP and the PIAW generally contradict each other (when the effect of PICP is preferable, PIAW has poorer evaluation results and vice versa), which means selecting the parameter is difficult in the application of the current method.

Aiming at the above problems, this paper proposes a conditional copula multi-objective prediction algorithm based on discrete form. This method takes advantage of the copula function and makes full use of the correlation existing in the adjacent load sequences to obtain the prediction interval to be predicted. A rolling prediction method is adopted to obtain prediction results for the entire period. At the same time, we further improve the prediction method of the conditional copula function in the process of research, and the value boundary problem of the model parameters was found and fully considered. Based on this, the NSGA-II (Non-dominated Sorting Genetic Algorithm-II)

multi-objective optimization algorithm was used to obtain the optimal solution of *K* and *t*. After that, the optimal parameters and corresponding prediction results are obtained by evaluating the solution set, which improves the prediction accuracy and practicability of the method. Compared with the current common load prediction methods, the conditional copula multi-target load interval prediction method proposed in this paper has the following advantages:

- (1) The conditional copula multi-objective interval prediction algorithm does not need to process a large amount of original data, and the calculation takes a short time.
- (2) Compared with the traditional machine learning algorithm represented by ANN, the prediction algorithm proposed in this paper has fewer calculation model parameters, does not require long-term machine learning and is convenient to apply.
- (3) The prediction algorithm proposed in this paper does not depend on the accuracy of point prediction and can directly obtain the probability distribution of the load to be predicted, thus, on the whole, improving the prediction accuracy of the load.
- (4) Compared with the probabilistic prediction method represented by the literature [28], there is no problem of selecting the kernel function, so that it can be applied better in practice and has the advantages of wide application range and strong expansibility.
- (5) Due to the excellent characteristics of the copula function, the prediction method proposed in this paper can make full use of the correlation between load sequences in adjacent time periods, and, compared with the current commonly used load prediction methods, it has better prediction accuracy and effect.

2. Materials and Methods

2.1. Conditional Copula Function

The copula function, also known as the connection function, is a type of function that connects multiple random variable joint distribution functions with their respective marginal distribution functions. It can describe the correlation between multiple variables more comprehensively and accurately. Due to this characteristic, it has a very wide application in risk prediction in the fields of finance and insurance.

Using the copula function for load interval prediction mainly involves the calculation of conditional copula function construction and confidence intervals. The copula function in its general form (Gaussian copula function, t-copula function, and others) always reveals differences from the actual situation. At the same time, the use of the continuous copula function form in calculations is relatively difficult because it requires obtaining the inverse function in the calculation of the confidence interval. Therefore, this paper uses the empirical conditional copula function, which could more accurately describe the correlation between random variables and uses the conditional copula function in the discrete form [34–36] to calculate the confidence interval at a given confidence level more conveniently.

The construction method of the discrete condition copula function is given as follows. It is assumed that N independent samples for t + 1 variables are known, as shown in Equation (1), where each row represents one sample, and each column represents the value of the marginal distribution function of one variable under different samples:

$$[X_1, \cdots, X_t, X_{t+1}] = \begin{bmatrix} F_1(X_1^1) & \cdots & F_t(X_t^1) & F_{t+1}(X_{t+1}^1) \\ F_1(X_1^2) & \cdots & F_t(X_t^2) & F_{t+1}(X_{t+1}^2) \\ \cdots & \cdots & \cdots & \cdots \\ F_1(X_1^N) & \cdots & F_t(X_t^N) & F_{t+1}(X_{t+1}^N) \end{bmatrix}.$$
(1)

The copula function is connected to the edge distribution function between t + 1 variables. We obtain the value of the edge distribution function of the corresponding Equation (1) by the literature [32,33]. As shown in Equation (2):

 F_1, \dots, F_t, F_{t+1} in Equation (2) are the edge distribution function of the variable 1, ..., variable *t*, variable *t* + 1, respectively. The value of each edge distribution function in the matrix is between 0 and 1. We divide the interval [0, 1] into *K* subintervals equally. Each interval ranges is as in Equation (3):

$$\delta = 1/K, S_1 \in [0, \delta], S_j = [(j-1)\delta, j\delta], j = 2, 3, \cdots K.$$
(3)

It can be seen that there are *K* subintervals for each variable, and t + 1 variables will together form K^{t+1} subspaces. Any line in Equation (2) must fall within a certain subspace.

The condition copula function refers to the probability distribution function of the corresponding t + 1th variable, with the former t variable values of a certain sample F being known. That is, the $F_{t+1}(x_{t+1}^F)$ probability distribution function in Equation (4) is the conditional copula function that needs to be obtained:

$$F_1(x_1^F), F_2(x_2^F), \cdots, F_t(x_t^F).$$
 (4)

In order to approximately describe a continuous distribution function using a set of discrete probability distributions, we adopted a scenario approach [32,33]. The specific operation is as follows: in N samples of Equation (2), the samples in the former t variables that fall in the same subspace as Equation (4) are removed to form a conditional matrix. Suppose there are N_1 samples in the matrix, all samples in the matrix are then classified using K-means clustering.

The method consists of two steps. First, the samples for which all of the elements in the t + 1th column fall in the same subinterval are classified into the same class, and the subinterval definition is the same as in the previous definition. In this way, we can classify N_1 samples into J classes, and the numbers of samples included in each class are recorded, respectively, as M_1, M_2, \dots, M_J ($M_1 + M_2 + \dots + M_J = N_1$).

Then, the mean F_{t+1}^{j} ($j = 1, 2, \dots J$) is calculated using Equation (5) to represent the *t*+1th element in each class, and the probability of occurrence of each class is given:

$$F_{t+1}^{j} = \frac{1}{M_{j}} \sum_{i=1}^{M_{j}} F_{t+1}\left(x_{t+1}^{i}\right), p^{j} = M_{j}/N_{1}.$$
(5)

Therefore, $[F_{t+1}^{j}, p^{j}]_{j=1,2,...j}$ is the discrete probability distribution function of $F_{t+1}(x_{t+1}^{F})$ based on Equation (4), which is the discrete conditional copula function.

2.2. Multi-Objective Prediction Method Based on the Conditional Copula Function

The load prediction model proposed in this paper includes two main components. One component is the conditional copula interval ultra-short-term prediction method based on the discrete form, which uses load historical data after giving a set of prediction model parameters (K, t), to establish the probability distribution function of the load in the period to be predicted by calculating the marginal distribution function such that the upper and lower bounds of the prediction interval under the given confidence level are obtained; the prediction of the entire prediction period is then completed using rolling prediction, and the interval prediction effect is measured by two evaluation indices containing PICP and PIAW. The second component is the NSGA-II multi-objective optimization method, which

uses PICP and PIAW as the objective functions within the value range of the copula interval prediction model parameters *K* and *t* and finds the non-inferior solution set for predicting the parameters of *K* and *t*. It is assumed that each evaluation index is equally important such that each index is endowed with 50% weights, and a set of optimal prediction model parameters is obtained by weighting calculations. The copula interval prediction is performed with the recently obtained parameter, and the results are compared with the prediction results of the currently used interval prediction methods.

To better understand the detailed process of this prediction method, we give the specific implementation steps of the prediction method, as shown in Figure 1:

Step 1: Using the historical load sequence, in the process of establishing conditional copula distribution function in discrete form, we first determine the value range of K (interval division number) and t (condition number) that is applicable to the current historical data.

Step 2: We use two indicators, PIAW and PICP, to evaluate the accuracy of the conditional copula interval prediction results under different *K* and *t* values.

Step 3: Taking PICP and PIAW as objective functions, we use the NSGA-II multi-objective optimization method to find the non-inferior solution set of the conditional copula prediction model parameters within the range of *K*, *t*.

Step 4: We obtain a set of optimal parameters for the prediction model using the average weight method.

Step 5: We use the optimal group of parameters of *K*, *t* for conditional copula interval prediction and compare the prediction results with those of GPR, ANN and ELM (three commonly used interval prediction algorithm results) to verify the effectiveness of the proposed method.



Figure 1. Process diagram of the multi-objective prediction method based on the conditional copula function.

2.2.1. Load Interval Ultra-Short-Term Prediction

A relatively strong correlation exists between load sequences in short time intervals. Therefore, we can take the continuous load sequences collected from certain regions as historical data and use the copula function to describe this correlation between load sequences. We assume that the continuous load sequence of a certain region is $[y_1, y_2, \dots, y_{n-1}, y_n]$, and we need to predict the interval of the load y_{n+1} at the next moment. In using the discrete conditional copula function to make a load prediction, assuming that the condition number is t, the (t + 1) variables and (n - t) samples can be transformed from the historical data:

$$\begin{bmatrix} y_1 & \cdots & y_t & y_{t+1} \\ y_2 & \cdots & y_{t+1} & y_{t+2} \\ \cdots & \cdots & \cdots & \cdots \\ y_{n-t+1} & \cdots & y_{n-1} & y_n \end{bmatrix}.$$
 (6)

Each column in the matrix is a load sequence at certain time intervals, which is viewed as $[X_1, X_2, \dots, X_t, X_{t+1}]$. Thus, we establish the conditional copula distribution function of load to be predicted. In other words, we use the marginal distribution function value $F_1(x_1^F), F_2(x_2^F), \dots, F_t(x_t^F)$ of the load $[y_{n-t+1}, \dots, y_{n-1}, y_n]$ as a known condition and obtain the probability distribution function of y_{n+1} .

We assume that the conditional copula function obtained is $[F_{t+1}^j, p^j]_{j=1,2,\cdots,J'}$, where p^1 is the largest and p^J is the smallest. As all values of the marginal distribution functions fall into the interval [0, 1], we divide the interval into K subintervals, which are denoted as S_1, S_2, \cdots, S_K in turn. Thus p^j is accumulated from j = 1 until the sum is greater than or equal to the preset confidence level β . The value of β needs to be preset by humans according to actual needs and has nothing to do with the other factors. In the process of grid dispatching and decision-making, too small a β value will make the interval coverage rate low and the scheduling risk large, while too big a β value will make the interval width large and the application effect not good. As a compromise, we set $\beta = 0.9$ to get a better prediction.

We assume that when accumulating to j = q, $\sum_{j=1}^{j=q} p^{(j)} \ge \beta$, the subinterval corresponding to

$$F_{t+1}^1, F_{t+1}^2, \cdots F_{t+1}^q$$
 is given as follows:

$$F_{t+1}^{1} \in \left(S_{1}^{l}, S_{1}^{u}\right], F_{t+1}^{2} \in \left(S_{2}^{l}, S_{2}^{u}\right] \cdots F_{t+1}^{q} \in \left(S_{q}^{l}, S_{q}^{u}\right],$$
(7)

where *l* (lower) refers to the lower bound of the interval, and *u* (upper) refers to the upper bound of the interval. Under the confidence level β , the prediction interval of the load at the (*t* + 1)th moment is the union of Equation (5), namely:

$$\left(S_1^l, S_1^u\right] \cup \left(S_2^l, S_2^u\right] \cup \cdots \left(S_q^l, S_q^u\right] = \left(S^l, S^u\right].$$
(8)

Using the inverse operation of the marginal distribution function, we obtain the prediction interval $F_{t+1}^{-1}(S^l) \le W_{t+1} \le F_{t+1}^{-1}(S^u)$ of the load at the (t + 1)th moment, and the load interval prediction at (t + 1)th moment is realized.

2.2.2. Interval Prediction Evaluation Index

We adopt PICP and PIAW as the evaluation indices to validate the prediction accuracy of the proposed algorithm.

(1) PICP

Prediction interval coverage probability is the most important index to measure the quality of the predicted interval. It defines the probability that the actual observation value falls within the prediction interval which is enveloped by the upper and lower bounds. A larger PICP value indicates that more observation values are covered by the constructed prediction interval, and vice versa. Under ideal conditions, PICP = 100% and all observation values are within the prediction interval. The expected observation generally falls within the constructed prediction interval at a probability *k* above the preset confidence level, that is:

$$k = \mathbf{P}(V_u \in \left[\underline{V_u}, \overline{V_u}\right]),\tag{9}$$

where P() is the probability; \overline{V}_u and \underline{V}_u are the upper and lower bounds of the prediction interval. V_u is the corresponding observational value and β is the preset confidence level. According to Bernoulli's law of large numbers, interval coverage can be expressed directly by the frequency of the predicted interval coverage observation value, which converges to *k* according to a certain probability.

$$PICP = \frac{1}{U} \sum_{u=1}^{U} A_u \times 100\% \rightarrow P(V_u \in [\underline{V_u}, \overline{V_u}]),$$
(10)

where *U* is the total amount of load to be predicted, $u = 1, 2 \dots U$. A_u is a Boolean quantity that is defined as follows:

$$A_{u} = \left\{ \begin{array}{cc} 1, & V_{u} \in \left[\underline{V}_{u}, \overline{V}_{u} \right] \\ 0, & V_{u} \notin \left[\underline{V}_{u}, \overline{V}_{u} \right] \end{array} \right\}.$$
(11)

When the actual value of load to be predicted falls into the prediction interval, the value of A_u is 1. Otherwise, A_u is 0.

To construct an effective prediction interval, the PICP should be as high as possible compared to the rated confidence level β ; if the PICP is much smaller than β , the prediction interval is treated as an invalid interval and needs to be reconstructed.

(2) PIAW

PIAW is an important reference for evaluating the quality of the predicted interval. If the interval is sufficiently wide, it can easily satisfy the ideal coverage of PICP = 100%. However, such an interval is too conservative to give effective information about the uncertainty of the value to be predicted—this makes the interval prediction altogether impractical. It is necessary to measure the interval width in order to reasonably evaluate the prediction interval. The PIAW index is defined as follows:

$$PIAW = \frac{1}{U} \sum_{u=1}^{U} \left(\overline{V_u} - \underline{V_u} \right).$$
(12)

The index measures the PIAW of each point to be predicted. The prediction interval width is positively correlated with the observed value; the interval width can be evaluated very accurately by PIAW. Under extreme conditions, if the upper and lower bounds of the prediction interval are the same for N points to be predicted, the width of the predicted interval is zero. The interval prediction is reduced to a point prediction and the evaluation index of the prediction interval loses its practical significance. Here, we strictly distinguish point prediction from interval prediction and ensure that the prediction interval satisfies $PIAW \neq 0$.

2.2.3. NSGA-II Multi-Objective Optimization Algorithm

The NSGA (Non-dominated Sorting Genetic Algorithm) is derived from the Genetic Algorithm (GA) but is different from the traditional GA. The NSGA algorithm is based on a genetic algorithm, and its core technology lies in two aspects: one is to rank individuals in the population by using the non-dominated sorting principle; the other is to calculate the virtual crowding distance of each individual by the Niched Genetic Algorithms (NGR). Based on the above two technical principles, the NSGA algorithm evaluates and selects the population in the "selection" segment, thus achieving multi-objective optimization, and exhibits strong advantages such as strong searchability and robustness in the field of multi-objective optimization. However, with the increase in the complexity of problems, the NSGA algorithm also exposes the disadvantages of high computational complexity, the easy coverage of excellent individuals in the iterative process, and the need to manually specify shared parameters [37].

In view of the shortcomings of the NSGA algorithm, in 2002 Deb and other scholars proposed NSGA-II algorithm, which is superior to the NSGA algorithm: it uses a fast non-dominated sorting algorithm, and the computational complexity is much lower than that of NSGA; the comparison operator of congestion degree and congestion degree is adopted to replace the shared radius: share Q that needs to be specified and serves as the winning criterion in the peer comparison after quick sorting. It extends the distribution range of the solution set in the Pareto frontier and distributes evenly, maintaining the diversity of the population. The elite strategy is introduced to expand the sampling space, prevent the loss of the best individual, and improve the computing speed and robustness of the algorithm [38].

In this paper, the two evaluation indexes consisting of PIAW and PICP are used as the adaptive function, *K* and *t* of copula interval prediction under the confidence level β are used as decision variables, according to the definition of *K* and *t* [32,33], their value should be a positive integer. But the total number of subspaces of Equation (2) is K^{t+1} , the subspace is too much. There will be no sample in the former N samples who fall within the same subinterval as Equation (4), so fail to construct the conditional matrix. Therefore, we first assume that *t* = 1, *K* gradually increases from 2, take obtain the condition matrix as the principle, and determine the value range of *K* when *t* = 1; then assume that *t* = 2, *K* gradually increases from 2, and determine the value range of *K* when *t* = 2. After such an exhaustive operation, the value range of all parameters (*K*, *t*) is obtained. And then NSGA-II is used for multi-objective optimization to obtain a non-inferior solution set of the conditional copula prediction model. In order to better understand the specific process of this method, the specific implementation steps of the optimization method are given below, and the flowchart is shown in Figure 2:

Step 1: Monte Carlo random sampling is carried out for two values of *K* and *t*.

Step 2: Coding. Adopt the real number coding method.

Step 3: Set the population size "Pop Size" and generate the initial population, PO.

Step 4: Conduct the fast-non-dominated sorting and virtual crowding distance calculation for the contemporary population, which is one of the core steps of this algorithm. The fast-non-dominated sorting is based on the two adaptability function values including PIAW and PICP, while the virtual crowding distance is obtained from the distance information of the individual vectors in the variable space.

Step 5: Conduct a genetic operation, which includes selection, crossover, and mutation. The sub-population can be obtained in this step.

Step 6: Conduct an elitist strategy. The parent population and sub-population are combined, selection operation based on the fast-non-dominated sorting and the virtual crowding distance is carried out, and we can thus obtain the next parent generation.

Step 7: The iterations number plus 1; return to step 4 until the iterations number reaches the given maximum.





Figure 2. The specific process of the NSGA-II (Non-dominated Sorting Genetic Algorithm-II) algorithm.

3. Results

3.1. Example Data Source

To verify the effectiveness of the proposed method in this paper, we apply the multi-objective ultra-short-term prediction method of the conditional copula function to three regions in Shaanxi Province, China, which are recorded as region 1, region 2, and region 3. The load data of each region are derived from the local power grid. According to China's power industry standard (DL/T 1711-2017) and the actual application requirements of ultra-short-term load prediction, the resolution of the load data is determined to be 15 min.

We divide the historical load data collected in one area into three parts, which is called respectively training data, validation data and test data, while the NSGA-II algorithm is only used during the validation data.

(1) training data: We use most of the historical load data of a region as training data, and use these data to construct the conditional copula function. At this time, the prediction model parameters (K, t) are not fixed, if the selected parameters are not suitable, the prediction effect will be poor.

(2) validation data: In the ultra-short-term load prediction, due to the full consideration of the principle of "near load with large correlation and far load with small correlation," namely, the future change trend of load is more dependent on the development law of recent load in the historical period, and the correlation with the development trend of long-term loads is weak. Therefore, we use the load data of the day before the test period to find the optimal parameters of the prediction model, and use

the NSGA-II multi-objective optimization algorithm to obtain the optimal solution set of the prediction model parameter values (K, t), finally we obtain a group of optimal parameters (K, t) by evaluating the solution set.

(3) test data: In order to verify the actual accuracy of the optimal parameter prediction model, we use the test data to test to ensure the unbiasedness of the prediction model, and compare the results with the three commonly used load interval prediction methods.

Region 1 is located in Yan'an City, Shaanxi Province, China. We use the load data from the Yan'an Power Grid Corporation from 1 January 2016 to 31 January 2017 as the training data, and we take the load data of 1 February 2017 as the validation data. Finally, the test is performed using the (96 points) load data of 2 February 2017 to verify the actual accuracy of the prediction model with the optimal parameters.

Region 2 is located in Hanzhong City, Shaanxi Province, China. We use the load data from the Hanzhong Power Grid Corporation from 1 March 2016 to 31 March 2017 as the training data, and we take the load data of 1 May 2017 as the validation data. Finally, a test is performed using the load data (96 points) of 2 May 2017 to verify the actual accuracy of the prediction model.

Region 3 is located in Shangluo City, Shaanxi Province, China. We use the load data from the Shangluo Power Grid Corporation from April 2016 to August 2016 as the training data, and we take the load data of 1 September 2016 as the validation data. Finally, the test is performed using the load data (96 points) of 2 September 2016 to verify the actual accuracy of the prediction model.

The computer CPU model we use is Core i7-7700K with 4 core 8 thread, clock speed is 4.2–4.5 GHz; computer memory is 16 GB, solid state disk is SAMSUNG 960 EVO 500G; the operating system is Windows 10 64-bit; the proposed load ultra-short-term prediction model was built on MATLAB R2016b. In the validation period, the execution time is 17.815 s; while in the testing period, conduct the load prediction of a point (15 min), the execution time is 0.841 s.

3.2. Algorithm Application Process

In establishing the discrete conditional copula function, t and K must be determined. The confidence level β is a free parameter that can reflect the prediction accuracy, and it is unified as β = 0.9 in this paper. The values of parameter K and t are limited by the sample number of the historical load sequence. After rigorous calculation, we obtain the following range: in region 1: when t = 1, the range of K is from 2 to 1087; when t = 2, the range of K is from 2 to 13; and when t = 3, the range of K is from 2 to 10. In region 2: when t = 1, the range of K is from 2 to 10; and when t = 3, the range of K is from 2 to 5. In region 3: when t = 1, the range of K is from 2 to 651; when t = 2, the range of K is from 2 to 7; and when t = 3, the range of K is from 2 to 4.

In the value range of model parameters *K* and *t*, the two evaluation indices of PICP and PIAW are taken as the objective function for calculation using the NSGA-II multi-objective optimization method, with programming in MATLAB according to the algorithm optimization principle in Section 3.3 to find the Pareto optimal solution of the conditional copula prediction model parameters. The obtained Pareto optimal solution set is shown in Figure 3.



Figure 3. Pareto optimal solution set: P: point, P1: point 1, (a) region 1, (b) region 2, (c) region 3.

It can be seen that the Pareto optimal solution sets of the three regions each have 12 best points, the PICP is greater than or equal to 0.7, and the PIAW are between 200 and 400 MW, which explains why the points in the solution set have better prediction performance. However, in Figure 3a,b, point 3 can be regarded as the sudden-change point of the entire fitting Pareto optimal solution set curve, which can also be said to be an inflection point, and in Figure 3c, the inflection point of the entire curve becomes point 4. In the figure, we mark the general Pareto optimal solution with a red dot and the inflection point with a blue circle.

In the Pareto optimal solution set, the larger the PICP point, the larger the PIAW. The PIAW and PICP values are contradictory to each other; the solution sets cannot be compared with each other to seek the better one, so all solution sets are considered optimal solutions. In order to obtain the optimal parameter setting, we first use the extreme value method to standardize the PIAW and PICP values of each group of parameters. According to the definition in Section 3.2, the smaller the PIAW value, the better the prediction performance, so we use Equation (13) to standardize the index PIAW. Meanwhile, the larger the PICP value, the better the prediction performance, so we adopt Equation (14) to standardize the index PICP. x_i is the value to be normalized and x_j is the normalized value, $MAX = \max(x_i)(1 \le i \le 12)$, $MIN = \min(x_i)(1 \le i \le 12)$.

$$x_j = \frac{MAX - x_i}{MAX - MIN} \tag{13}$$

$$x_j = \frac{x_i - MIN}{MAX - MIN} \tag{14}$$

Equations (13) and (14) are used to standardize the PIAW and PICP value of all points, and we consider that the two precision indexes are equally important for evaluating the effect of interval prediction. Therefore, the weights of PICP and PIAW can all be considered as 50%. Finally, we conduct a weight calculation for the two evaluation indexes of all the points in the Pareto optimal solutions set, with results shown in Tables 1–3.

| Point | PICP | PIAW | Standardized PICP | Standardized PIAW | Weight Calculation |
|--------------------------|-------|---------|----------------------|----------------------|-----------------------|
| P1 ($t = 1, K = 25$) | 1.000 | 340.810 | 1 | 0.000 | 0.500 |
| P2 $(t = 1, K = 52)$ | 0.990 | 270.125 | 0.967 | 0.556 | 0.762 |
| P3 ($t = 1, K = 155$) | 0.979 | 219.224 | 0.930 | 0.957 | 0.944 |
| P4 $(t = 1, K = 199)$ | 0.969 | 218.039 | 0.897 | 0.966 | 0.932 |
| P5 ($t = 1, K = 251$) | 0.958 | 218.009 | 0.861 | 0.966 | 0.914 |
| P6 ($t = 1, K = 255$) | 0.948 | 217.979 | 0.828 | 0.967 | 0.898 |
| P7 $(t = 1, K = 317)$ | 0.938 | 217.910 | 0.795 | 0.967 | 0.881 |
| P8 ($t = 1, K = 375$) | 0.844 | 217.315 | 0.483 | 0.972 | 0.728 |
| P9 $(t = 1, K = 358)$ | 0.813 | 216.661 | 0.381 | 0.977 | 0.679 |
| P10 ($t = 1, K = 421$) | 0.760 | 216.597 | 0.205 | 0.978 | 0.592 |
| P11 ($t = 1, K = 457$) | 0.740 | 215.418 | 0.139 | 0.987 | 0.563 |
| P12 ($t = 1, K = 514$) | 0.700 | 213.751 | 0 | 1.000 | 0.500 |

Table 1. Weight calculation results of the points in the Pareto optimal solutions set in region 1.

Table 2. Weight calculation results of the points in the Pareto optimal solutions set in region 2.

| Point | PICP | PIAW | Standardized PICP | Standardized PIAW | Weight Calculation |
|-------------------------|-------|---------|----------------------|----------------------|-----------------------|
| P1 ($t = 1, K = 28$) | 1.000 | 355.244 | 1 | 0.000 | 0.500 |
| P2 $(t = 1, K = 99)$ | 0.990 | 266.483 | 0.967 | 0.657 | 0.812 |
| P3 ($t = 1, K = 180$) | 0.979 | 226.465 | 0.930 | 0.953 | 0.942 |
| P4 ($t = 1, K = 188$) | 0.969 | 224.155 | 0.897 | 0.971 | 0.934 |
| P5 $(t = 1, K = 288)$ | 0.958 | 223.975 | 0.860 | 0.972 | 0.916 |
| P6 ($t = 1, K = 347$) | 0.948 | 223.138 | 0.827 | 0.978 | 0.903 |
| P7 $(t = 1, K = 349)$ | 0.938 | 222.474 | 0.793 | 0.983 | 0.889 |

| Point | PICP | PIAW | Standardized PICP | Standardized PIAW | Weight Calculation |
|--------------------------|-------|---------|----------------------|----------------------|-----------------------|
| P8 ($t = 1, K = 377$) | 0.844 | 222.148 | 0.480 | 0.985 | 0.734 |
| P9 ($t = 1, K = 394$) | 0.813 | 221.481 | 0.377 | 0.990 | 0.686 |
| P10 $(t = 1, K = 418)$ | 0.760 | 221.149 | 0.200 | 0.993 | 0.599 |
| P11 $(t = 1, K = 456)$ | 0.740 | 220.967 | 0.133 | 0.994 | 0.567 |
| P12 ($t = 1, K = 544$) | 0.700 | 220.171 | 0 | 1.000 | 0.500 |

Table 2. Cont.

Table 3. Weight calculation results of the points in the Pareto optimal solutions set in region 3.

| Point | PICP | PIAW | Standardized PICP | Standardized PIAW | Weight Calculation |
|--------------------------|-------|---------|----------------------|----------------------|-----------------------|
| P1 ($t = 1, K = 48$) | 1.000 | 377.187 | 1 | 0.000 | 0.500 |
| P2 $(t = 1, K = 69)$ | 0.990 | 298.588 | 0.967 | 0.609 | 0.788 |
| P3 ($t = 1, K = 119$) | 0.979 | 288.157 | 0.930 | 0.690 | 0.810 |
| P4 ($t = 1, K = 134$) | 0.969 | 251.691 | 0.897 | 0.972 | 0.935 |
| P5 ($t = 1, K = 245$) | 0.958 | 251.563 | 0.861 | 0.973 | 0.917 |
| P6 $(t = 1, K = 287)$ | 0.948 | 250.422 | 0.828 | 0.982 | 0.905 |
| P7 ($t = 1, K = 348$) | 0.938 | 250.154 | 0.795 | 0.984 | 0.889 |
| P8 $(t = 1, K = 395)$ | 0.844 | 249.565 | 0.483 | 0.989 | 0.736 |
| P9 ($t = 1, K = 457$) | 0.813 | 249.194 | 0.381 | 0.991 | 0.686 |
| P10 ($t = 1, K = 478$) | 0.760 | 248.568 | 0.205 | 0.996 | 0.601 |
| P11 ($t = 1, K = 515$) | 0.740 | 248.421 | 0.139 | 0.997 | 0.568 |
| P12 ($t = 1, K = 556$) | 0.700 | 248.084 | 0 | 1.000 | 0.500 |

It can be seen from Tables 1–3 that, after normalizing the two interval prediction indexes of each point, the PIAW and PICP values of the Pareto optimal solution in each region are between 0 and 1. The PIAW value of the point 1 to point 12 gradually increases, and the PICP value gradually decreases, indicating that the PICP of the 1st point is optimal, but the PIAW is the worst; otherwise, the PICP of the 12th point is the worst, but the PIAW is the best. After assigning 50% weights to each of the two indicators and adding them up, we can see that the weight of the 3rd point in Table 1 is 0.944, which is the maximum in 12 points. This indicate that the interval prediction result is optimal in this parameter setting. In Figure 3a, it can be seen that the point is located at the inflection point of the approximate curve, and the specific prediction model parameters are set to K = 155, t = 1. The weight of point 3 in Table 2 is 0.942, which is the maximum in 12 points, which explains that the interval prediction result is optimal in this parameter setting. In Figure 3b, it can be seen that the point is also located at the inflection point of the approximate curve, and the specific prediction model parameters are set to K = 180, t = 1. The weight of point 4 in Table 3 is the maximum in 12 points. In Figure 3c, it can be seen that the point is still at the inflection point of the approximate curve, and its specific prediction model parameters are set to K = 134, t = 1.

3.3. Results Comparison for Different Algorithms

To validate the effectiveness of the proposed method, ELM [26], GPR algorithm [28], and ANN [29] are introduced to compare the results of interval prediction with those of the proposed method.

Figures 4–6 show the prediction results of applying the four methods in region 1, region 2, and region 3, respectively. In the figure, the black x mark represents the actual load, the red line is the upper limit of the predicted interval, the purple line is the lower limit of the predicted interval, and the red and purple lines form 96 little intervals, which are the results of the interval prediction. If the black x mark is between the red line and the purple line, the point prediction can be considered successful. If the black x mark is beyond the red line, then the measured value is considered beyond the upper limit of the prediction. If the black x mark is lower than the purple line, then the measured value is lower than the lower limit of the prediction. When the measured value exceeds the upper limit or lower limit, the point prediction is considered to be a failure.

It can be observed from Figure 4a that the measured values of two points (34, 73) exceed the upper limit of the prediction, so they do not meet the prediction requirements. In Figure 4b, the measured values of three points (27, 34, 73) exceed the upper limit of the prediction, so they do not meet the prediction requirements. In Figure 4c, the measured values of four points (33, 34,37, 73) exceed the upper limit of the prediction, so they do not meet the prediction requirements. In Figure 4d, the measured values of five points (24, 33, 34, 37, 73) exceed the upper limit of the prediction, so they do not meet the prediction requirements. In Figure 4, we see that the load values of points 33–34 and 72–73 undergo a sudden change, but the four prediction methods cannot predict the dramatic change of this load in advance, so all the interval prediction upper bounds are a bit low, resulting in failed prediction with each method at points 34 and 73. However, since the method can make full use of the correlation between adjacent load sequences when the load is abruptly changed at other periods such as points 26–27 and 32–33, a certain interval prediction margin is retained, thereby reducing the number of points of failed prediction. From the prediction results of the four methods, when using the method of this paper to predict region 1, only the upper bounds of the two prediction intervals do not meet the requirements, while more than three appear in (b), (c), (d), which indicates that the PICP obtained by this method is higher. At the same time, the PIAW of the prediction results obtained by this method is the lowest among all methods, which shows that the interval prediction algorithm provided in this paper has the highest accuracy.



Figure 4. Cont.



Figure 4. Prediction results of applying four methods in region 1: (**a**) Prediction results based on the optimal parameter *K* and *t*; (**b**) prediction results of artificial neural network (ANN); (**c**) prediction results based on extreme learning machine (ELM); (**d**) prediction results of Gaussian process regression (GPR) algorithm.

In Figure 5a, only two points (34, 74) in the predicted segment do not meet the prediction requirements. In Figure 5b, only three points (27, 34, 71) in the predicted segment do not meet the prediction requirements. In Figure 5c, only four points (2, 34, 35, 71) in the predicted segment do not meet the prediction requirements. In Figure 5d, only five points (34, 35, 39, 71, 74) in the predicted segment do not meet the prediction requirements. It can be observed that the method of this paper is still the most reliable. Similarly, the main reason for the failure of the 34th point prediction may be a sudden load change during this period. In Figure 5, we find that at the 90–96 point, all the lower bounds of the interval predictions given by the four methods are low, which makes the actual power values close to the lower bound of the interval prediction, so the prediction of this period (86–90 points), such that the predicted lower bound and increasing the success rate. However, by comparing the predicted lower bound and increasing the success rate. However, by comparing the prediction results of the four methods, we can also see that the method proposed in this paper has the largest PICP and the smallest PIAW, and a better interval prediction effect.





Figure 5. Prediction results of applying four methods in region 2: (**a**) Prediction results based on the optimal parameter *K* and *t*; (**b**) prediction results of ANN; (**c**) prediction results based on ELM; (**d**) prediction results of GPR algorithm.

In region 3, the accuracy of the method proposed in this paper is still the highest, and three points (32, 34, 89) in Figure 6a do not meet the requirements. Four points (31, 32, 34, 35) in Figure 6c do not meet the requirements. Five points (31, 32, 34, 35, 36) in Figure 6d do not meet the requirements. It can be seen that the overall prediction effect of each method in region 3 is inferior to region 1 and region 2, which may be caused by the relatively bigger fluctuation of load in the forecast period of region 3, and the overall prediction effect of each method in points 30 to 40 is poor; we speculate that this may be related to the fact that the load before the 30th point is relatively stable, so that each model believes that this "smooth" process will continue, but in fact, the load has increased by about 500 MW in one hour. The excessive increase of the load makes all the upper bounds of the interval prediction bound, causing prediction failure. However, the method in this paper can make full use of the correlations between loads and finds the "trend" of such a continuous decline is actually weakened. The prediction successful at points 35 and 36.



Figure 6. Cont.



Figure 6. Prediction results of applying four methods in region 3: (a) Prediction results based on the optimal parameter K and t; (b) prediction results of ANN algorithm; (c) prediction results based on ELM; (d) prediction results of GPR algorithm.

4. Discussion

After obtaining the prediction results of each method from the three regions, for further analysis and discussion, we use Equations (10) and (12) to calculate the prediction accuracy indicators of each method separately. The results are shown in Table 4.

| Method | PICP(R1) | PIAW(R1) | PICP(R2) | PIAW(R2) | PICP(R3) | PIAW(R3) |
|--------|----------|----------|----------|----------|----------|----------|
| M1 | 0.979 | 219.238 | 0.979 | 228.249 | 0.969 | 253.741 |
| M2 | 0.969 | 240.773 | 0.969 | 243.602 | 0.958 | 263.314 |
| M3 | 0.958 | 235.523 | 0.958 | 237.006 | 0.958 | 260.336 |
| M4 | 0.948 | 220.913 | 0.948 | 231.248 | 0.948 | 258.471 |

 Table 4. Accuracy comparison between the four methods.

M1: the method proposed by this paper, M2: ANN, M3: ELM, M4: GPR, R1: region 1, R2: region 2, R3: region 3.

It can be observed from Table 4 that in the three regions, the PICP values obtained by each method are all greater than 0.90, indicating that the prediction effects of these methods are excellent. However, the PICP of the interval prediction method (M1) proposed in this paper is the largest, and at the same time, the corresponding PIAW is smallest. This observation illustrates that the proposed method in this paper is superior to the commonly used interval prediction methods of which these three are representative.

The comparison shows that the prediction interval coverage of M2 is better than that of M3 and M4, but its interval width is too large. The interval average width of M4 is narrower, but the success rate of interval prediction is not high. In general, each of these three methods has advantages and disadvantages, and it is impossible to choose an optimal method. However, the optimal parameters selected in this paper are based on the same importance of PICP and PIAW, but at times it is necessary to emphasize security in the scheduling process, and thus the requirements for PICP are higher in the two evaluation indicators. Applying the method in this paper endows the PICP with a higher weight, but because of the need to consider high security, the prediction process tends to be conservative, resulting in an increase in the interval width. In contrast, at times we require a narrower prediction interval in the scheduling process. In this case, the requirements for PIAW are higher. The smaller the

PIAW, the easier it is to meet the requirements such that the PIAW is endowed with a higher weight when applying this method, but this increases the probability of prediction failure and decreases PICP. It can be observed that the selection of the optimal parameters for the prediction model actually changes with different situations, and the actual scheduling must be considered.

5. Conclusions

Due to the strong randomness and volatility of the load, accurate value prediction is difficult. After analyzing the problems existing in the current load prediction method, this paper proposed a multi-objective interval prediction method based on the conditional copula function in discrete form, which helps power system staff to make more reasonable, safe, and economical scheduling decisions. Specifically, we did the following research:

- (1) At present, some load interval prediction methods need to rely on point prediction or require appropriate kernel functions to be selected in advance, which reduces the prediction accuracy. With the excellent characteristics of the copula function, we could take full advantage of the correlation relationship between load sequences in adjacent time periods by establishing the discrete condition copula function of the point to be predicted, to obtain the prediction interval of a point to be predicted, and then, the rolling prediction method is adopted to obtain the prediction result of the whole prediction period, which improves the interval prediction accuracy; the method has the advantages of being simple and quick, and easy to expand.
- (2) Based on the problem of contradiction between multiple evaluation indicators, when the values of the conditional copula prediction model parameters (*K*, *t*) change, we use the NSGA-II multi-objective algorithm to seek the optimal model parameter within its value range such that the non-inferior solution set of the conditional copula prediction model parameters is obtained. Then we use the average weight method to get a group of the optimal set of conditional copula prediction model parameters.
- (3) To validate the effectiveness and applicability of the load interval prediction method proposed in this paper, the method was applied to three regions in China. We compared the prediction results obtained using the proposed method with those of three more mature methods (ELM [26], GPR algorithm [28], and ANN [29]). The comparative analysis showed that the method proposed in this paper has the smallest PIAW, the largest PICP, and the best effect.
- (4) As is mentioned in the discussion, we believe that when PICP and PIAW are endowed with their weights, we must combine the actual requirements of scheduling. The prediction results obtained using this method are not exactly the same in different practical situations. At the same time, the method proposed in this paper can provide a new idea for dispatching departments, because it can be more effective and safer to conduct power system dispatching after combining with point value prediction. Meanwhile, the construction process of a conditional copula function can also be extended to other fields such as wind speed, PV, and so on.

However, there are still some shortcomings in our work, in that if the correlation relationship between load sequences is weak, the prediction results using this method are usually bad. On the other hand, since the ultra-short-term load prediction resolution is relatively high (within 15 min), the influence of weather factors on the prediction results is small, and the numerical weather prediction model is very complicated, so we did not consider the weather factors and prediction day types features in the study. In future work, if this method could be applied to medium- and long-term load prediction, we could further study how to better combine meteorological factors, day type features, and copula interval prediction methods.

Author Contributions: G.Z. and Z.L. conceived and designed the experiments; G.Z., Z.L. and J.H. performed the experiments; Z.L., J.H. and K.Z. analyzed the data; J.H. and F.L. contributed reagents/materials/analysis tools; G.Z., Z.L. and X.Z. wrote the paper.

Funding: This work was supported by the National Key Research and Development Program of China (2016YFC0401409), the National Natural Science Foundation of China under Grant No.51679188 and the Key research and development plan of Shaanxi Province (2018ZDCXL-GY-10-04).

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

References

- 1. Guo, H.C. Research on Development of Demand Side Management System Based on Load Analysis in Smart Grid. Ph.D. Dissertation, North China Electric Power University, Beijing, China, 2014.
- 2. Li, P. Research on Risk Element Transfer Model and Decision Support System of Smart Grid Operation Management. Ph.D. Dissertation, North China Electric Power University, Beijing, China, 2014.
- 3. Chen, X.L. Load Forecasting and Pricing Mechanism of Interactive Smart Grid. Master's Dissertation, Shanghai Jiaotong University, Shanghai, China, 2013.
- 4. Carcedo, J.M.; García, J.P. Integrating long-term economic scenarios into peak load forecasting: An application to Spain. *Energy* **2017**, *140*, 682–695. [CrossRef]
- 5. Matijaš, M.; Vukićević, M.; Krajcar, S. Supplier Short Term Load Forecasting Using Support Vector Regression and Exogenous Input. *J. Electr. Eng.* **2011**, *62*, 280–285. [CrossRef]
- 6. Singh, N.; Singh, A.; Tripathy, M. Selection of Hidden Layer Neurons and Best Training Method for FFNN in Application of Long Term Load Forecasting. *J. Electr. Eng.* **2012**, *63*, 153–161. [CrossRef]
- Li, H.Z.; Guo, S.; Li, C.J.; Sun, J.Q. A hybrid annual power load forecasting model based on generalized regression neural network with fruit fly optimization algorithm. *Knowl. Based Syst.* 2013, 37, 378–387. [CrossRef]
- 8. Mandal, P.; Senjyu, T.; Urasaki, N.; Funabashi, T. Electricity Price and Load Short-Term Forecasting Using Artificial Neural Networks. *Int. J. Emerg. Electr. Power Syst.* **2011**, 7. [CrossRef]
- Xiao, B.; Wang, H.; Mu, G. Spatial Load Forecasting Based on Load Forecasting Reliability. *Appl. Mech. Mater.* 2014, 3546, 1075–1080. [CrossRef]
- 10. Zhao, H.; Guo, S. An optimized grey model for annual power load forecasting. *Energy* **2016**, *107*, 272–286. [CrossRef]
- 11. Ioannis, P. Application of hybrid computational intelligence models in short-term bus load forecasting. *Expert Syst. Appl.* **2016**, *54*, 105–120.
- 12. Yalcinoz, T.; Eminoglu, U. Short term and medium term power distribution load forecasting by neural networks. *Energy Convers. Manag.* **2004**, *46*, 1393–1405. [CrossRef]
- Kandil, N.; Wamkeue, R.; Saad, M.; Georges, S. An efficient approach for short term load forecasting using artificial neural networks. In Proceedings of the 2006 IEEE International Symposium on Industrial Electronics, Montreal, QC, Canada, 9–13 July 2006. [CrossRef]
- 14. Siwek, K.; Osowski, S.; Szupiluk, R. Ensemble Neural Network Approach for Accurate Load Forecasting in a Power System. *Int. J. Appl. Math. Comput. Sci.* **2009**, *19*, 303–315. [CrossRef]
- 15. Radziukynas, V.; Klementavičius, A. Short-Term Forecasting of Loads and Wind Power for Latvian Power System: Accuracy and Capacity of the Developed Tools. *Latvian J. Phys. Tech. Sci.* **2016**, *53*, 3–13. [CrossRef]
- 16. Shi, J.G. Design Technology of CAN Fieldbus System; National Defense Industry Press: Beijing, China, 2004.
- Hassan, S.; Khosravi, A.; Jaafar, J. Improving load forecasting accuracy through combination of best forecasts. In Proceedings of the IEEE International Conference on Power System Technology, Auckland, New Zealand, 23–26 October 2012.
- Singh, A.K.; Khatoon, S.; Muazzam, M.; Chaturvedi, D.K. Load forecasting techniques and methodologies: A review. In Proceedings of the International Conference on Power, Control and Embedded Systems (ICPCES), Allahabad, India, 17–19 December 2012.
- Chen, Y.; Luh, P.; Guan, C.; Zhao, Y.; Michel, L.D.; Coolbeth, M.A.; Friedland, P.B.; Rourke, S. Short-Term load forecasting: Similar day-based wavelet neural networks. *IEEE Trans. Power Syst.* 2010, 25, 322–330. [CrossRef]
- 20. Stockman, M.; Eiramli, R.S.; Awad, M.; Jabr, R. An asymmetrical and quadratic support vector regression loss function for beirut short term load forecast. In Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (SMC), Seoul, Korea, 14–17 October 2012.

- Haque, A.U.; Mandal, P.; Meng, J.; Srivastava, A.K.; Tseng, T.L.; Senjyu, T. A novel hybrid approach based on wavelet transform and fuzzy ARTMAP network for predicting wind farm power production. In Proceedings of the IEEE Industry Applications Society Annual Meeting (IAS), Las Vegas, NV, USA, 7–11 October 2012.
- 22. Ghelardoni, L.; Ghio, A.; Anguita, D. Energy load forecasting using empirical mode decomposition and Support Vector Regression. *IEEE Trans. Smart Grid* **2013**, *4*, 549–556. [CrossRef]
- 23. Wang, B.; Leng, B.X.; Zhang, X.H.; Shan, C.H.; Cong, Z. Application of Support Vector Machine in Short-term Load Forecasting. *J. Power Syst. Autom.* **2011**, *23*, 115–121.
- 24. Gu, Y.D.; Hu, F.F. Partition Support Vector Regression and Its Application in Load Forecasting. J. Power Syst. Autom. 2013, 25, 20–24.
- 25. Kang, C.Q.; Xia, Q.; Zhang, B. Review of power system load forecasting and its development. *Autom. Electr. Power Syst.* **2004**, *28*, 1–11.
- 26. Li, Z.Y.; Ding, J.Y.; Wu, D.; Wen, F. An ensemble model of the extreme learning machine for load interval prediction. *J. North China Electr. Power Univ.* **2014**, *41*, 78–88.
- 27. Yang, W.J.; Kang, C.Q.; Xia, Q.; Liu, R.; Tang, T.; Wang, P.; Zhang, L. Short term probabilistic load forecasting based on statistics of probability distribution of forecasting errors. *Autom. Electr. Power Syst.* **2006**, *30*, 47–52.
- 28. Zong, W.; Wei, Z.; Sun, G.; Li, H.; Cheung, K.W.; Sun, Y. Short-term Load Interval Forecasting Based on Improved Gauss Process Regression Model. *J. Power Syst. Autom.* **2017**, *29*, 22–28.
- 29. Quan, H.; Srinivasan, D.; Khosravi, A. Probabilistic uncertainty handling using neural network-based prediction intervals for electrical load forecasting. *Energy* **2014**, *73*, 916–925. [CrossRef]
- 30. Che, G.; Luh, P.B.; Michel, L.D.; Chi, Z. Hybrid Kalman filters for very short-term load forecasting and prediction interval estimation. *IEEE Trans. Power Syst.* **2013**, *28*, 3806–3817.
- 31. Charytoniuk, W.; Niebrzydowski, J. Confidence interval construction for load forecast. *Electr. Power Syst. Res.* **1998**, *48*, 97–103. [CrossRef]
- 32. Li, J.H.; Wen, J.Y.; Cheng, S.J.; Wei, H. A scene generation method considering copula correlation relationship of multi-wind farms power. *Proc. CSEE* **2013**, *33*, 30–36.
- 33. Lan, F.; Nong, Z.G.; Li, J.H. Research on spatial and temporal correlation of wind power sequence. *Proc. CSU-EPSA* **2016**, *28*, 24–31.
- 34. Yu, Z.P. The properties of conditional copula. J. Jiamusi Univ. 2011, 29, 763–765.
- 35. Wang, L. A Study on Financial Contagion Effect among Stock Markets Based on the Conditional Copula Model. Master's Dissertation, Xiamen University, Xiamen, China, 2009.
- 36. Cao, Y.D. The Simulation of Garch Model by Multiple Conditions Copula via Vine Struction. Master's Dissertation, Lanzhou University, Lanzhou, China, 2014.
- 37. Xu, L. Research and Application of Multi-objective Optimization Based on Genetic Algorithms. Master's Dissertation, Central South University, Changsha, China, 2007.
- 38. Chen, X.Q.; Hou, Z.X.; Guo, L.M.; Luo, W.C. Improved multi-objective genetic algorithm based on NSGA-II. *J. Comput. Appl.* **2006**, *10*, 2453–2456.



© 2019 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 (http://creativecommons.org/licenses/by/4.0/).