Research Article

The Combination Forecasting of Electricity Price Based on Price Spikes Processing: A Case Study in South Australia

Jianzhou Wang,[1] **Ling Xiao,**[2] **and Jun Shi**[3]

[1] *School of Statistics, Dongbei University of Finance and Economics, Dalian 116025, China*
[2] *School of Mathematics and Statistics, Lanzhou University, Lanzhou 730000, China*
[3] *School of Information Science and Engineering, Lanzhou University, Lanzhou 730000, China*

Correspondence should be addressed to Ling Xiao; xiaoling5231@163.com

Electricity price forecasting holds very important position in the electricity market. Inaccurate price forecasting may cause energy waste and management chaos in the electricity market. However, electricity price forecasting has always been regarded as one of the largest challenges in the electricity market because it shows high volatility, which makes electricity price forecasting difficult. This paper proposes the use of artificial intelligence optimization combination forecasting models based on preprocessing data, called "chaos particles optimization (CPSO) weight-determined combination models." These models allow for the weight of the combined model to take values of $[-1, 1]$. In the proposed models, the density-based spatial clustering of applications with noise (DBSCAN) algorithm is used to identify outliers, and the outliers are replaced by a new data-produced linear interpolation function. The proposed CPSO weight-determined combination models are then used to forecast the projected future electricity price. In this case study, the electricity price data of South Australia are simulated. The results indicate that, while the weight of the combined model takes values of $[-1, 1]$, the proposed combination model can always provide adaptive, reliable, and comparatively accurate forecast results in comparison to traditional combination models.

1. Introduction

Electricity has been regarded as a unique commodity in the energy market under the reformation of the energy economy because it is hard to store and it is closely related to peoples' livelihood. Electricity price plays an important role in the power market because it balances market demand and market supply [1]. The future price of electricity is the indicator of the market demand for electricity generators as well as for the performance of the profitability of the electricity market. In addition, the nonstorable nature of electricity requires maintaining a constant balance between demand and supply. Therefore, an accurate electricity price forecasting method plays an important role in the electricity industry because not only can it help the generators and decision-makers to generate more profit, but it can also help to reduce waste of the electricity. However, the price of electricity is affected by many factors, including the demand of the market, electricity net transport, weather, and many other uncertain factors [2].

Thus, the price is highly volatile, and the abnormal behaviors of jump points amass abnormal price volatility. The abnormal price points are also called price spikes, and these spikes cause the forecasting of the price of electricity to be a challenge.

The common models used to forecast the price of electricity are classified as time series models and artificial intelligence based models [3]. The time series models are some of the most widely used techniques, including the ARIMA and GARCH models, which are regarded as efficient ways to forecast the price of electricity [4]. Time series models are specifically developed for forecasting the price of electricity [2]. Artificial intelligence based models, including neural network and data-mining models, are regarded as the most popular techniques applied in electricity price forecasting [5]. A fuzzy inference system is also applied in electricity price forecasting, and the desired results have been obtained [6]. Francesco Serinaldi has introduced the application of generalized additive models for location, scale, and shape (GAMLSS), which is a dynamical distribution model that has

proved to be a flexible model in electricity price forecasting [5].

However, the models listed above are all just based on the original data without preprocessing any outliers. Electricity prices can change suddenly and can make large jumps to extreme levels; this phenomenon may be caused by unexpected increases in demand, an unexpected supply shortage, or failures in the transmission net [7]. The extreme price values are also called "price spikes," which negatively affect forecasting accuracies because many forecasting models are sensitive to extreme observations [5]. Thus, the detection of price spikes is a very important issue for practical situations. In this paper, the density-based spatial clustering of applications with noise (DBSCAN) algorithm is used to identify and remove the electricity price spikes. The linear interpolation function is used to produce new points that replace the price spikes for the development of the preprocessed data. Furthermore, the original data and the preprocessed data are used to establish the forecasting models, and then the forecasting results are compared in this study. The simulation results show that the forecasts built using the preprocessed data perform better than when the original data were used.

It is widely accepted that combination forecasts can achieve desirable results [8, 9]. In several research studies, combination forecasts have generally been shown to have outstanding performance in comparison to single forecasts [10]. Because the main idea of a combination forecast is assigning weights to the combined forecasts according to the contribution of the combined forecasts, it is an important task to select a reasonable combination method. Traditional combination methods have restrictions on weight: assigning negative values is not allowed and the sum of the weights must be 1. In this way, traditional combination methods have low forecasting accuracies. The CPSO weight-determined combination method allows weights of combined models to take values of $[-1, 1]$, which overcomes the shortage of traditional combination models caused by the no negative constraint on weight. Furthermore, the CPSO weight-determined combination method has inherited several advantages of an artificial intelligence algorithm to flexibly search for the best solutions. Several combination models are compared, and their forecasting results are analyzed carefully in this paper. The results indicate that CPSO weight-determined combination models outperform traditional combination models.

This paper is organized as follows: Section 1 gives a brief introduction of this paper; Section 2 proposes theories of individual forecasts and combination methods; Section 3 illustrates the electricity price forecasting of South Australia (SA) and contains the analysis of the forecasting results and a comparison of the forecasting models. Lastly, Section 4 concludes with the contributions of this paper.

2. Materials and Methods

Three individual forecasting models including the autoregressive integrated moving average (ARIMA) model, the back propagation neural network (BPNN) model, and Kalman filtering (KF) model are chosen for this study. Each of these models has its own advantages in electricity price forecasting. The ARIMA model can analyze the linear relationship of variables and implement various exponential smoothing models [11]. The BPNN model is an efficient way to analyze nonlinear data [12], and the KF model provides an efficient computational solution using the least square method with only a minor computational cost and with easy adaptation to any alteration of the observations [13]. However, the ARIAM model is not always efficient when applied to any data set [11]; the global solution determined using the BPNN model depends on the training of its nonlinear functions, which may cause inaccurate prediction results [14]; and the KF models cannot always produce desirable forecasting results [15]. As a widely accepted way to improve forecasting accuracies, combination methods are proposed in this paper to combine the individual forecasts.

2.1. Autoregressive Integrated Moving Average (ARIMA) Model. The ARIMA model was developed by Box and Jenkins, also called "Box-Jenkins," and has been widely used in many areas in the last several decades [11]. This model can capture the linear relationship between different variables in the real world [16]. An ARIMA (p, d, q) model can be expressed as follows:

$$\varphi(B)\left(1 - B^d\right)Y_t = \left(1 - \theta\left(B^r\right)\right)\varepsilon_t, \tag{1}$$

where Y_i $(i = 1, 2, \ldots, t)$ is the time series value and $\varphi(B)$ and $\theta(B)$ are the coefficients related to the lag operator or back operator B. In addition, $d = 1, 2, \ldots, p$, $r = 1, 2, \ldots, q$, p, and q are inter numbers often referred to as autoregressive and moving average polynomials, respectively. Finally, ε_i $(i = 1, 2, \ldots, t)$ are the random errors at time t, which are assumed to be independently and identically distributed with mean = 0 and a variance of σ^2. In an ARIMA model, the future time series value is regarded to be a linear function of several past observations and random errors [11]. The form of (1) can be written as (2), and then it is easy to obtain the future series according to the several past values:

$$Y_t = \frac{\left(1 - \theta\left(B^r\right)\right)\varepsilon_t}{\varphi(B)\left(1 - B^d\right)}. \tag{2}$$

To obtain accurate forecasting results, three phases must be performed [16]:

(1) identify the structure of the model;

(2) estimate the parameters of the model;

(3) establish the model and forecast the future series.

2.2. Back Propagation Neural Network (BPNN) Model. An artificial neural network system is a nonlinear model composed of three layers, including an input layer, one or more hidden layers, and one output layer. There is a large amount of nodes to connect the three or more layers using different weights. The BPNN model is a type of multilayer feed-forward neural network with a wide variety of applications,

which is one of the artificial intelligence types of models. Input vectors and the corresponding reference vectors are used to train a network until it can approximate a function and associate input vectors with specific output vectors [12]. In this paper, the Hecht-Nelson method is used to determine the number of hidden layer nodes [13], which means that when the number of input layer nodes is n, the number of hidden layer nodes is chosen as $2n + 1$. The structure of BPNN is shown in Figure 1 and consists of three layers: the input layer, the hidden layer, and the output layer.

Before training the net, it is necessary to process data. To ensure that the input vectors are compatible when there are significant differences in their magnitudes, we determine the input vector by normalizing each input value as follows [14, 17]:

$$P = \{P_i\} = \frac{x_i - x_{i\,min}}{x_{i\,max} - x_{i\,min}}, \quad i = 1, 2, \ldots, n, \quad (3)$$

where $x_{i\,max}$ and $x_{i\,min}$ are the maximum and the minimum values of the input data, respectively. x_i is the real value of each vector. Subsequent steps for training a BP net model for application in electricity price forecasting are necessary. Figure 1 shows the topology of the BP network.

The net is initialized by giving weights ω_{ij}, v_{jt}, and threshold value θ_i randomly. The input set is then preprocessed by (3), and the output of hidden layer is calculated by the hidden layer function:

$$y_i = f_1 \left(\sum_i w_{ij} x_i - \theta_j \right) \quad (4)$$

$$(i = 1, \ldots, n; \; j = 1, \ldots, 2n + 1),$$

where y_i is the output of the hidden layer node j and f_1 represents the activation function of a node, which is usually a sigmoid function.

The outputs of the output layer data are calculated with the following form:

$$\text{output}_1 = f_2 \left(\sum_i v_{ij} y_j - r_1 \right) \quad (j = 1, \ldots, 2n + 1), \quad (5)$$

where r_1 represents the bias of the neuron, output_1 represents the output data of the network, and f_2 stands for the activation function of the output layer node [14].

2.3. Kalman Filtering (KF) Model.

Kalman filters (KF) were proposed by Kalman in 1960. These filters are dynamic behavioral procedure systems for optimal time series estimation. Observations are recursively combined with recent forecasting values with weights that minimize the corresponding biases [15]. The discrete procedure used is stated by the following state equation (6) and output equation (7):

$$X(k + 1) = AX(k) + V_1(k), \quad (6)$$

$$Y(k) = CX(k) + V_2(k), \quad (7)$$

where k stands for the time, X is the actual state at time k, and A and C are the coefficient matrices that must be
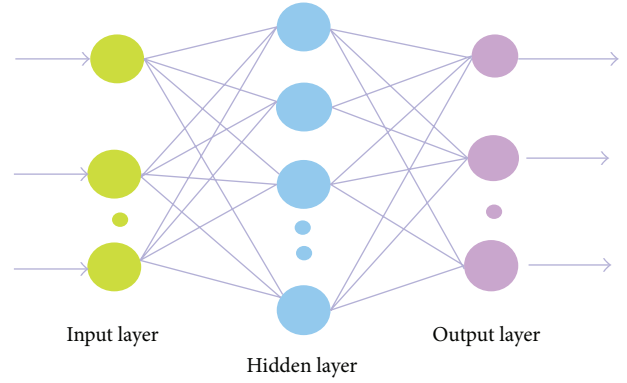

FIGURE 1: The topology of the BP neural network.

determined before the model is applied in the forecasting system. Y represents the output variable vector. V_1 and V_2 are the process series and measurement noise series, respectively, and they are assumed to be white Gaussian noise [18].

The KF model gives a dynamic estimation of the state X based on the observation value Y along with time k. The present estimate of $X(k)$ is based on the previous value $X(k-1)$, $k = 1, 2, \ldots$, and the relationship can be given as

$$X_{k/k-1} = AX(k - 1), \quad (8)$$

$$Y_{k/k-1} = AY(k - 1) + V_1(k). \quad (9)$$

As the new observation value $Y(k)$ is known, the estimate of state X is updated to be

$$X_k = X_{k/k-1} + K_k \left(Y_k - C_k X_{k/k-1} \right), \quad (10)$$

where

$$K_k = P_{k/k-1} \cdot C_k^T \left(C_k \cdot P_{k/k-1} \cdot C_k^T + V_2(k) \right)^{-1}. \quad (11)$$

The final estimate of P is

$$P_k = \left(I - K_k \cdot C_k \right) P_{k/k-1}. \quad (12)$$

Equations (8)–(12) provide a detailed updating procedure for the Kalman filter algorithm where K is the forecasting gain and P is the forecasting covariance. The ARIMA $(1, 0, 0)$ model is used to estimate C in this paper, and the observation equation can be denoted as [15]:

$$y_k = x_{0,k} + x_{1,k} \cdot d_k \quad (13)$$

and the observation matrix $C = [1, d_k]^T$.

2.4. Traditional Combination Method.

In general, the forecasting output of a combination forecasting model based on K component forecasts that produces forecasts $\widehat{p}_t^{(1)}, \ldots, \widehat{p}_t^{(K)}$ of p_t is given by the following form:

$$\widehat{p}_t = f \left(\widehat{p}_t^{(1)}, \ldots, \widehat{p}_t^{(k)}; w \right), \quad (14)$$

where f is a function and w is a weight parameter vector [8]. In this study, two methods are chosen to determine the parameter vector w.

Let $\widehat{x}_{it}$ denote the forecasting output of the ith individual forecasting model at time point t for time series x_t. The combination forecasting output at time t can then be denoted in the following form:

$$\widehat{x}_{ct} = \sum_{i=1}^{n} w_i \widehat{x}_{it}, \quad t = 1, 2, \ldots, \tag{15}$$

where $\widehat{x}_{ct}$ is the combination model's forecasting output, n is the number of the component forecasting models, and w_i is the weight of the ith individual forecasting model. Then, the constraints for w_i must be considered. In conventional research studies, the constraints must meet the following requirement to ensure that the combined model is reasonable [19]:

$$\sum_{i=1}^{n} w_i = 1, \quad 0 \leq w_i \leq 1. \tag{16}$$

To estimate the performance of the combined models, the key issue is to adjust the weight of the component forecasts. There are many methods for obtaining reasonable weight values, and the most commonly used method is the weighted average (WA) method. Let $e_{it} = \widehat{x}_{it} - x_t$ denote the residual of the ith individual forecasting model at time t. The residual of the combined model is denoted as follows:

$$e_{ct} = \widehat{x}_{ct} - x_t = \sum_{i=1}^{n} w_i e_{it}. \tag{17}$$

Therefore, the WA minimizes the mean absolute percentage error (MAPE) of the output of the combination forecasts to obtain the optimal weights as follows:

$$\min \quad M = \sum_{t=1}^{T} \sum_{j=1}^{m} \sum_{i=1}^{m} w_i w_j \frac{e_{it}}{x_t} \frac{e_{jt}}{x_t}$$

$$\text{subject to} \quad \begin{cases} \sum_{i=1}^{m} w_i = 1 \\ w_i \geq 0, \quad i = 1, 2, \ldots, m. \end{cases} \tag{18}$$

2.5. The Proposed CPSO Weight-Determined Combination Method.
Particle swarm optimization (PSO), developed by Kennedy and Eberhart, is a new intelligence algorithm based on a swarm intelligence algorithm inspired by the cooperation and communication of a swarm of birds or fish looking for food [20–22].

2.5.1. CPSO Algorithm.
Chaos is a common phenomenon that consists of unstable dynamic behavior that is sensitive to initial conditions. However, it contains an inherent regularity and randomness whereby the ergodicity and regularity of chaotic variables can be used to optimize a search. Recently, the chaos series replaced the random series for the development of the chaos optimization algorithm (COA), which has

achieved good results [23]. A logistic equation is a chaotic system, and the chaotic sequence is expressed as follows:

$$C_{n+1} = \mu C_n (1 - z_n), \quad n = 0, 1, 2, \ldots, \tag{19}$$

where $C_0 \in [0, 1]$, μ is the control parameter. C_n is the chaotic sequence at point n.

The chaos particle swarm optimization (CPSO) algorithm is developed based on the COA, which replaces random series with chaos series to avoid premature convergence [24]. The CPSO algorithm integrates the fast computational time of the PSO algorithm and the strong ability of the COA to examine values beyond the local extrema. In addition, CPSO can avoid the shortcomings of the PSO algorithm because it easily arrives at the local extrema by maintaining the diversity of the swarm [25]. The speed of the particles is $V_i = (v_{i1}, v_{i2}, \ldots, v_{in})$, which varies using the following form:

$$\begin{aligned} v_{id}&(t + 1) \\ &= w v_{id}(t) + c_1 \text{rand1}()\left(p_{id} - x_{id}(t)\right) \\ &\quad + c_2 \text{rand2}()\left(p_{gd} - x_{id}(t)\right) \quad i = 1, 2, \ldots, N. \end{aligned} \tag{20}$$

The position of each particle is changing according to the following equation:

$$x_{id}(t + 1) = x_{id}(t) + v_{id}(t + 1), \quad i = 1, 2, \ldots, N, \tag{21}$$

where t is the time, c_1 and c_2 are acceleration coefficients, w is the inertia factor, rand1() and rand2() are two independent random numbers uniformly distributed in the range of [0, 1], and N indicates the number of particles. Many studies have demonstrated that better results can be achieved when $c_1 = c_2 = 1.5$ and when w changes from 0.4 to 0.9 [26]. Table 1 shows the advantage of using CPSO compared to using basic PSO. Generally, the logistic chaos series performs better than the random series under Sphere, Rosenbrock, Rastrigrin and Schaffer f6 test functions.

2.5.2. Selecting Weights of Combination Models Using the CPSO Algorithm.
In this paper, the CPSO algorithm is used to determine the weights of the combination models. The combination model is shown in (1), and the objective function is shown as (5). However, to obtain better results, the constraints must be changed to the following:

$$\text{subject to} \quad \begin{cases} \sum_{i=1}^{m} w_i \approx 1 \\ \\ -1 \leq w_i \leq 1, \quad i = 1, 2, \ldots, m \end{cases} \quad w_i \geq 0, \ i = 1, 2, \ldots, m. \tag{22}$$

Weights are regarded as particles and are optimized by

TABLE 1: The search success ratio of functions (%) [26].

Success ratio	Logistic	Random
Sphere	100.0	100.0
Rosenbrock	94.2	97.1
Rastrigrin	92.4	90.3
Schaffer f6	96.7	90.4

the CPSO algorithm; the best particles have optimal weights. The optimal processing is expressed using the following steps [27].

Step 1. Initialize the parameters containing the particle population N, acceleration coefficients c_1 and c_2, inertia factor w, and the iteration times.

Step 2. Let the particles fly with the speed and location of particles designated under the control of (7) and (8), in which the random series is replaced by the chaos series.

Step 3. Denote the best particle as the ith particle and map $p_{g,i}$ into the interval $[0, 1]$ with the form $C_i = (p_{g,i} - a_i)/(b_i - a_i)$, $(i = 1, 2, \ldots, n)$ to generate the chaos series $C_t^{(m)}$, $m = 1, 2, \ldots$, and then return the chaos series $C_t^{(m)}$ to the original solution space using the form $P_{g,t}^{(m)} = a_t + (b_t - a_t)C_t^{(m)}$, $m = 1, 2, \ldots$, to obtain $P_g^m = (p_{g,1}^{(m)}, p_{g,2}^{(m)}, \ldots, p_{g,N}^{(m)})$.

Step 4. Calculate the fitness value of the current best particle that is the MAPE in this study, and then find the best individual P_*.

Step 5. Use the best individual P_* to replace a current particle.

Step 6. Check whether the optimal solution or maximum number of iterations has been achieved. If not, return to Step 2, or break.

3. A Case Study

The electricity price data used in this study is from South Australia (SA) (see Figure 2), which is located in the south of Australia [28].

3.1. Data Description. The available data were collected every 30 minutes from 1 March to 21 March 2009, and the collection time began at 00:30 and ended at 12:00 each day. The collected data are grouped into two sets: the "training set" was used to build the forecasting models, and the other set was the test set. The collected data are shown in Figure 3, which shows the high volatility of the data. In other words, there are many price spikes in the data.

3.2. Data Processing. Inaccurate forecasting results will arise if the training set used to build the forecasting model includes price spikes. Therefore, in this paper, several steps are taken to revise this model. The density-based spatial clustering of applications with noise (DBSCAN) algorithm is used
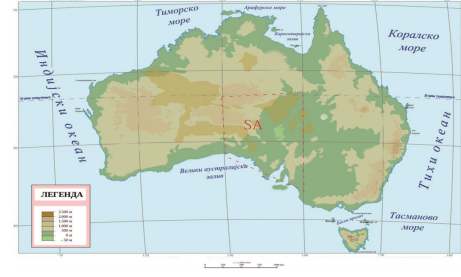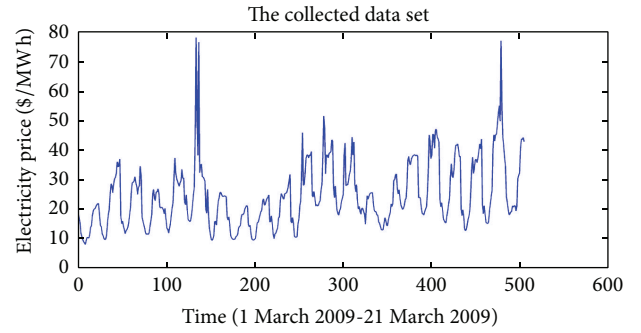

FIGURE 2: Study area.

FIGURE 3: The collected data.

in this study to identify the extreme data, and the linear interpolation function is used to generate normal values to replace the price spikes in the training set.

3.2.1. Density-Based Spatial Clustering of Applications with Noise (DBSCAN) Algorithm. The DBSCAN algorithm was proposed in 1996 and is a simple and efficient clustering algorithm based on density clustering [29]. The DBSCAN algorithm regards the cluster as a region and assumes that the objects in the region are dense. DBSCAN is the best choice for identifying price spikes, for it is able to discover clusters of arbitrary shapes [30]. The DBSCAN algorithm includes two parameters, ε and *Minpts*. ε is the radius of a neighborhood, *Minpts* is the number of points, and D is the given data. The following concepts are introduced for understanding and implementing the process of the algorithm. Figure 4 shows a brief scan of the DBSCAN algorithm.

Definition 1 (core object). The ε-neighborhood of a point p, containing at least *Minpts* points, denoted by $N_{Eps}(p)$, namely, which should satisfy $|N_{Eps}(p)| \geq Minpts$.

Definition 2 (directly density-reachable). Two conditions should be satisfied where an object q in the data is directly density-reachable from an object $p \cdot q \in N_{Eps}(p)$, $|N_{Eps}(p)| \geq Minpts$.

Definition 3 (density-connected). Objects p and q are density-connected in the data set D if for the given ε and *Minpts*; p and q are all directly density-reachable from the same object o.

Table 2: The forecasting results of three individual models.

| Time	Actual price value ($/MWh)	ARIMA		BP		Kalman	
		Forecasting value ($/MWh)	MAPE (%)	Forecasting value ($/MWh)	MAPE (%)	Forecasting value ($/MWh)	MAPE (%)
00:30	41.91	22.04	47.41	17.77	57.61	24.58	41.34
01:00	37.92	19.25	49.24	18.43	51.40	20.04	47.15
01:30	34.93	20.99	39.90	20.61	40.99	23.86	31.70
02:00	24.75	18.54	25.07	18.45	25.45	21.34	13.79
02:30	22.86	15.42	32.53	10.50	54.07	16.54	27.66
03:00	20.49	15.33	25.17	17.59	14.16	17.24	15.87
03:30	18.55	15.33	17.35	17.59	5.18	17.24	7.07
04:00	18.82	15.82	15.92	17.82	5.31	16.95	9.92
04:30	19.59	18.44	5.86	18.32	6.46	19.09	2.57
05:00	19.57	20.31	3.80	19.27	1.52	20.55	5.00
05:30	21.61	22.51	4.15	19.77	8.50	22.70	5.03
06:00	21.14	24.63	16.50	26.05	23.24	24.41	15.47
06:30	21.63	20.51	5.16	34.76	60.72	30.35	40.33
07:00	21.63	47.05	117.54	26.16	20.94	36.19	67.31
07:30	19.85	25.52	28.54	26.72	34.60	28.81	45.12
08:00	21.66	27.48	26.88	32.36	49.38	30.00	38.50
08:30	30.75	30.15	1.94	34.72	12.91	32.48	5.61
09:00	31.95	29.32	8.23	30.23	5.39	38.02	19.00
09:30	32.57	27.89	14.37	28.75	11.74	37.98	16.62
10:00	40.81	28.76	29.52	28.44	30.31	40.39	1.03
10:30	44.21	28.91	34.61	32.05	27.51	38.72	12.42
11:00	43.97	26.24	40.32	32.49	26.10	39.26	10.71
11:30	44.66	31.46	29.55	35.30	20.97	36.95	17.26
12:00	43.25	40.80	5.66	37.51	13.27	38.74	10.42
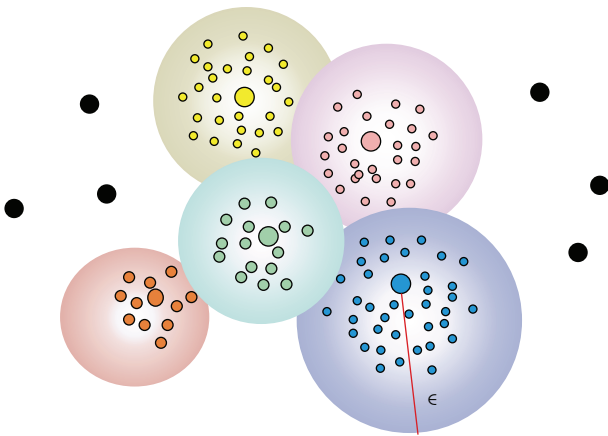
Figure 4: A sketch of DBSCAN algorithm.

Definition 4 (cluster). Clusters are the nonempty subs of the data set D that are density-connected from all core points.

Definition 5 (noise). Suppose that $C_1, C_2, \ldots, C_r$ are all clusters of data set D and noise is the set of points in D not in any cluster. Thus, noise is treated as a set of points that satisfy the form $\{p \in D \mid p \notin C_i, \ i = 1, 2, \ldots, k\}$.

3.2.2. Interpolation. In Section 2.2, the DBSCAN algorithm was used to identify the anomaly data. However, elimination of singular points can cause missing data, which is troublesome when the preprocessed data are used to build the forecasting models. In this study, a linear interpolation function is used to generate data to fill the holes left by the eliminated data. The *interp1* function in the toolbox of MATLAB 2013a is applied in interpolation.

The preprocessed data are shown in Figure 5.

3.3. Statistic Measurements of Forecasting Performance. The forecasting performance of the individual forecasts and combination models is evaluated by three common criteria: the mean absolute percentage error (MAPE), the root mean square error (RMSE), and the mean absolute percentage error
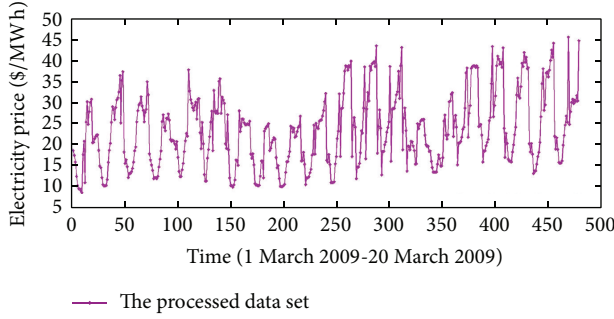
FIGURE 5: The preprocessed data.

(MAE). The accuracy of the forecasting results increase as these three error values decrease [31]:

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{x_i - \widehat{x}_i}{x_i} \times 100\%\right|,$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(x_t - \widehat{x}_t)^2}, \qquad (23)$$

$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n}|x_t - \widehat{x}_t|,$$

where x_i is the actual value at time i, $\widehat{x}_i$ is the forecasting value at time, and $n = 24$ in this study.

3.4. Simulation and Individual Models Performance.

The electricity price data of SA are preprocessed and divided into two sets: one set contains the electricity price data from 1 March 2009 to 20 March 2009, which is used to build the forecasting model, and the remaining data are included in the other set, which is the out-of-sample data. The detailed process of the model building is shown in Figure 6, and three individual models are chosen for the simulation. The forecasting result is shown in Figure 7, which suggests the forecasting model is only accurate in the local area. Table 2 provides a more detailed description of Figure 7 in which the forecasting results of the three individual models are listed. Three criteria are chosen to evaluate performance of the three individual forecasts, and the results are shown in Table 3. The data in Table 3 indicates that the KF model is the best model of the three forecasting models because it produces the minimum amount of error indicators in MAPE, RMSE, and MAE. The ARIMA model performs undesirably upon comparison with other two models.

3.5. Simulations and the Evaluation of the Combination Model's Performance.

The proposed combination method combines three individual forecasts. As a comparison, the traditional combination method is adopted to combine individual forecasts as well. Finally, the forecasting results are compared in this section. The simulation process can be seen in Figure 8.

TABLE 3: Comparison of three individual forecasting models.

	ARIMA	BP	Kalman
MAPE (%)	26.05	25.32	21.12
RMSE ($/MWh)	10.62	9.88	7.81
MAE ($/MWh)	7.93	7.73	6.04

Step 1. Combination simulation: CPSO weight-determined combination method combines the three individual forecasts. As a comparison, the three individual models are also combined using the traditional combination method.

Step 2. Analysis of the results: the performance of the combination models based on two combination methods is shown in Tables 4 and 5, respectively. A-B represents the combination model of ARIMA and BP; B-K represents the combination models of BP and Kalman; A-K represents the combination model of ARIMA and KF; A-B-K represents the combination model of ARIMA, BP, and KF; T means the traditional combination method; and C means the CPSO weight-determined combination method. Table 4 shows the forecasting results of the CPSO weight-determined combination models in the forecast period from 00:30 to 12:00 on 21 March and Table 5 shows the forecasting results of the traditional models in the same forecasting period, where the forecasting performance in every step can be observed easily. In comparing the forecasting results shown in Tables 4 and 5, a piece of information is revealed: while the forecasting points are moving, the proposed CPSO weight-determined combination models always perform better than the traditional combination models. For the CPSO weight-determined combination models, the A-B, B-K, A-K, and A-B-K models have a minimum MAPE value of 0.00% at time points 04:30, 05:00, 08:30, and 05:30, respectively. However, for the traditional combination method based A-B, A-K, B-K, and A-B-K models, the MAPE values were 6.33%, 3.37%, 3.73%, and 2.89%, respectively. This study illustrates the outstanding performance of the CPSO weight-determined combination models. The average evaluation criteria listed in Table 6 suggest that the A-B model based on the traditional combination method, which produced MAPE, RMSE, and MAE values of 25.19%, 9.74, and 7.72, respectively, is higher than the corresponding values of A-B model based on the CPSO weight-determined combination method. The same analyses were conducted using the B-K, A-K, and A-B-K models, and these results indicate that the combination models based on the CPSO weight-determined combination method perform better than the traditional combination method overall because CPSO algorithm searches the weights of combination models by the way of artificial intelligence.

Table 6 lists the comparison of the individual and combination models. On the whole, the combination models outperform the best combined individual models because these combination models generate smaller values of MAPE, RMSE, and MAE. The MAPE values of the A-B-K model are 21.12% (T) and 20.79% (C), the RMSE values of the A-B-K model are 8.19 (T) and 7.93 (C), and the MAE values are 6.32 (T) and 6.12 (C), which are all smaller than the evaluation
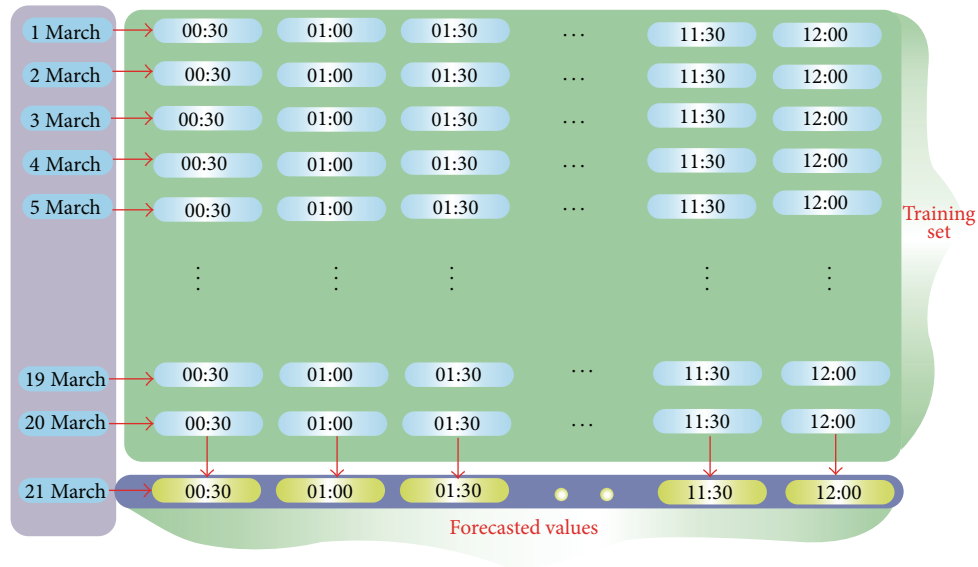
FIGURE 6: Simulation process description.

TABLE 4: CPSO weight-determined combination models forecasting results.

Time	Actual price value ($/MWh)	A-B (C)		B-K (C)		A-K (C)		A-B-K (C)	
		Forecasting value ($/MWh)	MAPE (%)	Forecasting value ($/MWh)	MAPE (%)	Forecasting value ($/MWh)	MAPE (%)	Forecasting value ($/MWh)	MAPE (%)
00:30	41.91	19.75	52.88	23.41	44.15	23.69	43.46	23.51	43.91
01:00	37.92	19.83	47.70	19.08	49.67	19.12	49.57	19.11	49.60
01:30	34.93	22.08	36.78	22.72	34.95	22.83	34.63	22.85	34.59
02:00	24.75	19.72	20.31	20.32	17.89	20.42	17.49	20.45	17.38
02:30	22.86	12.10	47.07	15.75	31.10	16.02	29.94	15.79	30.95
03:00	20.49	18.38	10.29	16.42	19.87	16.36	20.14	16.51	19.44
03:30	18.55	18.38	0.90	16.42	11.49	16.36	11.79	16.51	11.01
04:00	18.82	18.67	0.77	16.14	14.23	16.06	14.65	16.19	13.96
04:30	19.59	19.59	0.00	18.18	7.20	18.18	7.21	18.20	7.08
05:00	19.57	20.77	6.13	19.57	0.00	19.59	0.10	19.57	0.00
05:30	21.61	21.61	0.00	21.62	0.03	21.71	0.49	21.61	0.00
06:00	21.14	27.57	30.42	23.25	9.96	23.11	9.33	23.23	9.88
06:30	21.63	34.57	59.81	28.90	33.62	28.62	32.31	29.40	35.91
07:00	21.63	31.70	46.55	34.46	59.33	34.89	61.29	33.87	56.61
07:30	19.85	28.33	42.70	27.44	38.22	27.48	38.43	27.58	38.97
08:00	21.66	33.69	55.53	28.57	31.90	28.39	31.06	28.69	32.45
08:30	30.75	36.26	17.93	30.93	0.59	30.75	0.00	31.04	0.94
09:00	31.95	32.13	0.55	36.21	13.32	36.51	14.28	36.61	14.59
09:30	32.57	30.55	6.19	36.17	11.05	36.55	12.21	36.64	12.50
10:00	40.81	30.44	25.42	38.46	5.75	38.97	4.50	39.01	4.42
10:30	44.21	33.67	23.84	36.87	16.59	37.12	16.03	37.34	15.55
11:00	43.97	33.58	23.63	37.39	14.97	37.64	14.39	38.01	13.55
11:30	44.66	37.01	17.12	35.19	21.21	35.19	21.20	35.45	20.63
12:00	43.25	40.66	5.99	36.89	14.70	36.87	14.75	36.77	14.99

TABLE 5: Traditional combination models forecasting results.

Time	Actual price value ($/MWh)	A-B (T)		B-K (T)		A-K (T)		A-B-K (T)	
		Forecasting value ($/MWh)	MAPE (%)	Forecasting value ($/MWh)	MAPE (%)	Forecasting value ($/MWh)	MAPE (%)	Forecasting value ($/MWh)	MAPE (%)
00:30	41.91	18.84	55.05	22.88	45.41	23.95	42.87	22.29	46.81
01:00	37.92	18.64	50.86	19.64	48.21	19.84	47.67	19.49	48.61
01:30	34.93	20.71	40.72	23.05	34.02	23.14	33.75	22.67	35.11
02:00	24.75	18.47	25.36	20.62	16.70	20.64	16.61	20.26	18.13
02:30	22.86	11.73	48.69	15.03	34.25	16.26	28.87	14.58	36.21
03:00	20.49	17.03	16.91	17.33	15.43	16.76	18.19	17.23	15.91
03:30	18.55	17.03	8.22	17.33	6.59	16.76	9.64	17.23	7.12
04:00	18.82	17.32	7.97	17.17	8.78	16.67	11.44	17.15	8.87
04:30	19.59	18.35	6.33	18.90	3.54	18.93	3.38	18.81	3.99
05:00	19.57	19.53	0.20	20.23	3.37	20.49	4.70	20.14	2.89
05:30	21.61	20.46	5.34	21.97	1.65	22.65	4.82	21.77	0.75
06:00	21.14	25.70	21.55	24.82	17.41	24.47	15.73	24.94	17.96
06:30	21.63	31.20	44.23	31.45	45.41	27.89	28.94	31.11	43.84
07:00	21.63	31.38	45.09	33.68	55.72	38.91	79.87	33.73	55.96
07:30	19.85	26.42	33.10	28.29	42.51	27.99	41.00	27.95	40.81
08:00	21.66	31.14	43.77	30.59	41.23	29.37	35.60	30.58	41.18
08:30	30.75	33.58	9.20	33.04	7.45	31.90	3.73	33.03	7.43
09:00	31.95	30.00	6.10	36.07	12.90	35.85	12.19	35.04	9.68
09:30	32.57	28.54	12.39	35.67	9.53	35.46	8.87	34.47	5.82
10:00	40.81	28.52	30.12	37.40	8.35	37.48	8.15	35.93	11.96
10:30	44.21	31.27	29.28	37.05	16.19	36.27	17.97	36.02	18.52
11:00	43.97	30.93	29.66	37.57	14.56	36.01	18.11	36.33	17.37
11:30	44.66	34.34	23.11	36.54	18.19	35.58	20.34	36.09	19.19
12:00	43.25	38.33	11.37	38.43	11.14	39.26	9.24	38.48	11.02
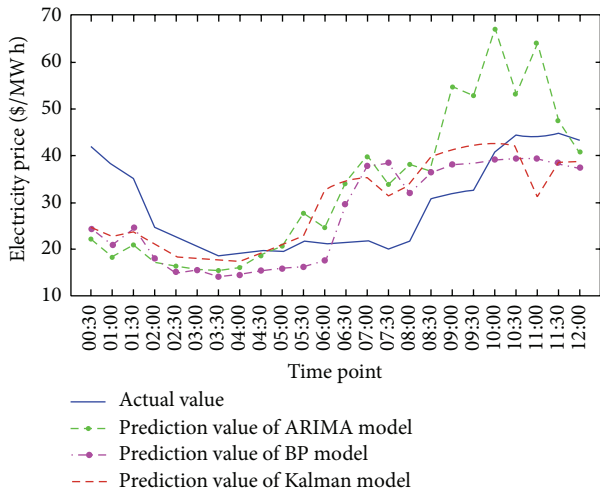
FIGURE 7: Comparison of the three individual forecasting results.

TABLE 6: Comparison of the performance of different forecasting models.

Models	MAPE (%)	RMSE ($/MWh)	MAE ($/MWh)
ARIMA	26.05	10.62	7.93
BP	25.32	9.88	7.73
Kalman	21.12	7.81	6.04
A-B (T)	25.19	9.74	7.72
A-B (C)	24.10	9.32	7.15
B-K (T)	21.12	8.07	6.30
B-K (C)	20.91	8.00	6.18
A-K (T)	21.12	8.19	6.32
A-K (C)	20.80	7.97	6.13
A-B-K (T)	21.12	8.19	6.32
A-B-K (C)	20.79	7.93	6.12

criteria of A-B-K, A-B, B-K, and A-K models. Therefore, A-B-K models perform better than A-B, B-K, and A-K models no matter if these other models are combined by the traditional combination method or the CPSO weight-determined combination method.

Table 7 shows the detailed weight assignment of the combination models and compares the two different combination methods. The comparison results indicate that the CPSO weight-determined combination method assigns

The compared results

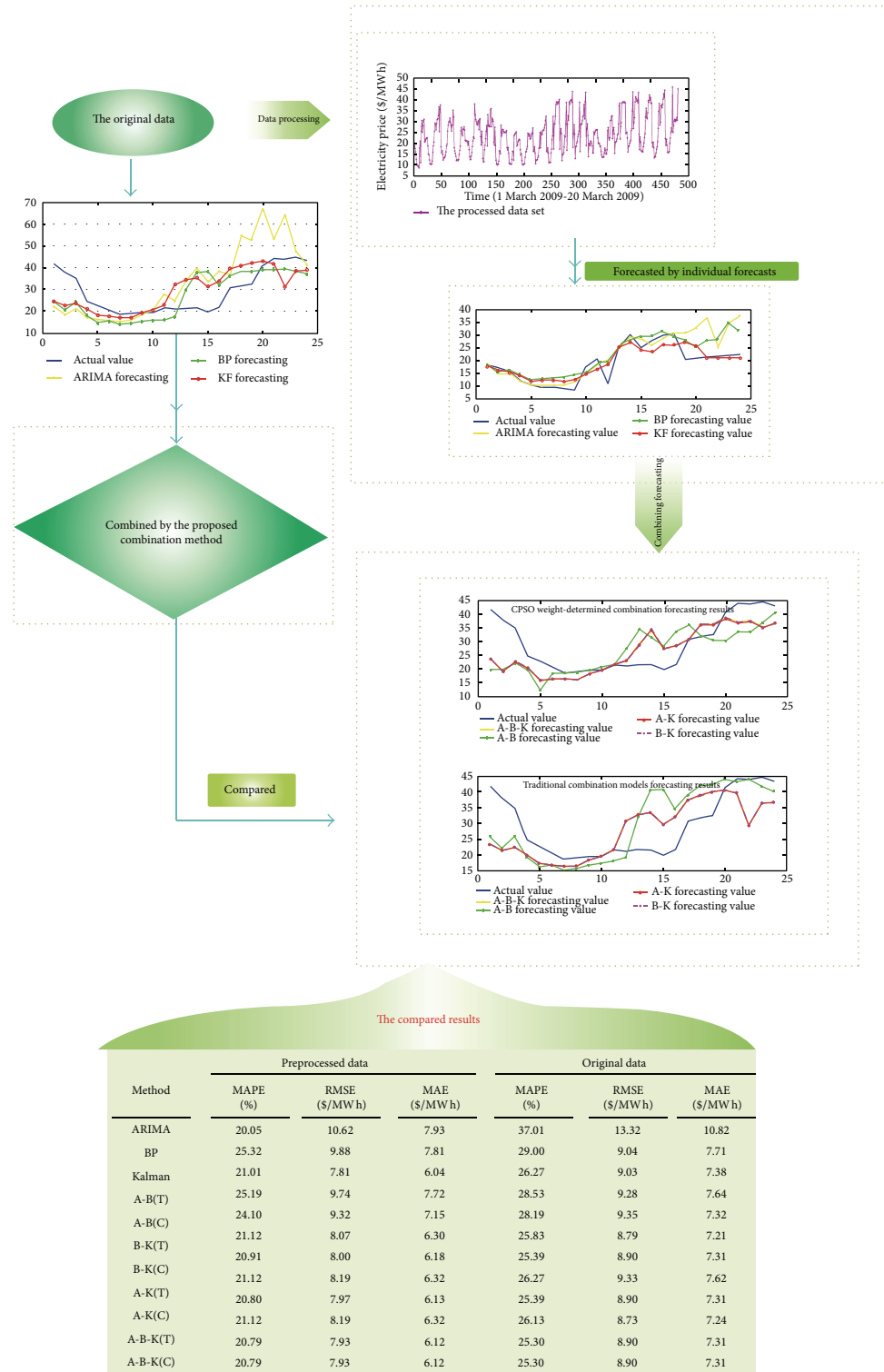| Method	Preprocessed data			Original data		
	MAPE (%)	RMSE ($/MW h)	MAE ($/MW h)	MAPE (%)	RMSE ($/MW h)	MAE ($/MW h)
ARIMA	20.05	10.62	7.93	37.01	13.32	10.82
BP	25.32	9.88	7.81	29.00	9.04	7.71
Kalman	21.01	7.81	6.04	26.27	9.03	7.38
A-B(T)	25.19	9.74	7.72	28.53	9.28	7.64
A-B(C)	24.10	9.32	7.15	28.19	9.35	7.32
B-K(T)	21.12	8.07	6.30	25.83	8.79	7.21
B-K(C)	20.91	8.00	6.18	25.39	8.90	7.31
A-K(T)	21.12	8.19	6.32	26.27	9.33	7.62
A-K(C)	20.80	7.97	6.13	25.39	8.90	7.31
A-B-K(T)	21.12	8.19	6.32	26.13	8.73	7.24
A-B-K(C)	20.79	7.93	6.12	25.30	8.90	7.31

FIGURE 8: The results of the simulation process.

weights to the individual models flexibly without the constraint of the sum of weights being 1 and with no negative assignments, which makes it easy to obtain more accurate forecasting results. The results in Table 7 suggest that, when the combination models perform undesirably, the traditional combination method would assign the weight of the best combined model a value of 1 and the others a value of 0. However, the CPSO weight-determined combination method would adjust the assignment of the weights without the constraint of the sum of weights being 1 with no negative

TABLE 7: Weights of the individual models in the combination models.

Weight models	Traditional combination method				CPSO weight-determined combination method			
	ARIMA	BP	Kalman	MAPE (%)	ARIMA	BP	Kalman	MAPE (%)
A-B	0.2500	0.7500	—	25.19	0.1798	0.8883	—	24.10
B-K	—	0.0000	1.0000	21.12	—	0	0.9523	14.70
A-K	0.0000	—	1.0000	21.12	−0.0498	1.0000	—	20.80
A-B-K	0.0000	0	1.0000	21.12	−0.0506	0.0024	1.0000	20.79

TABLE 8: Comparison of forecasting models built using preprocessed data and original data.

Method	Preprocessed data			Original data		
	MAPE (%)	RMSE ($/MWh)	MAE ($/MWh)	MAPE (%)	RMSE ($/MWh)	MAE ($/MWh)
ARIMA	20.05	10.62	7.93	37.01	13.32	10.82
BP	25.32	9.88	7.81	29.00	9.04	7.71
Kalman	21.01	7.81	6.04	26.27	9.03	7.38
A-B (T)	25.19	9.74	7.72	28.53	9.28	7.64
A-B (C)	24.10	9.32	7.15	28.19	9.35	7.32
B-K (T)	21.12	8.07	6.30	25.83	8.79	7.21
B-K (C)	20.91	8.00	6.18	25.39	8.90	7.31
A-K (T)	21.12	8.19	6.32	26.27	9.33	7.62
A-K (C)	20.80	7.97	6.13	25.39	8.90	7.31
A-B-K (T)	21.12	8.19	6.32	26.13	8.73	7.24
A-B-K (C)	20.79	7.93	6.12	25.30	8.90	7.31

assignments to find the best combination. The MAPE values indicate that the performance of the B-K model based on the traditional combination method is undesirable. Therefore, it is reasonable to improve the combination rule using the CPSO weight-determined combination method.

3.6. Comparison of the Forecasting Results of the Preprocessed Data and the Original Data. It is well known that the preprocessing of anomaly data is very useful in the building of forecasting models. In this section, the forecasting results of the forecasting models built by the preprocessed data (PF) and the forecasting models built by the original data (ODF) are compared in Table 8.

Table 8 shows that the PF based models have the best performance compared to the ODF models for both individual models and combination models. For the ARIMA models, the PF based model has a MAPE value of 20.05%, which is lower than the MAPE value of the ODF based model at 37.01%. The BP and KF models based on preprocessed data are superior to the models based on the original data. Although the combination models based on the original data (CODF) have good performance in comparing individual (IODF) models, the CODF models are inferior to the combination models based on the preprocessed data (CPF).

4. Conclusion

As electricity forecasting becomes increasingly important, this paper proposes the use of the CPSO weight-determined combination method based on DBSCAN algorithm preprocessed data to improve electricity price forecasting.

The traditional combination method has a no-negative weight constraint, and its sum must be 1, which limits the range of weights and thus causes lower forecasting accuracies. In addition, original data-based forecasts do not consider the fact that electricity price spikes are sensitive to outliers, which leads to inaccurate forecasting. To overcome the limitations of original data based traditional combination models, the proposed CPSO weight-determined combination method allows the weights of combined models to take values of $[-1, 1]$, and the original data are preprocessed by the DBSCAN algorithm. The electricity price data of South Australia have been simulated by the proposed, preprocessed, and data based CPSO weight-determined combination models to forecast the electricity price for the period from 00:30 to 12:00 on 21 March 2009, in SA. The preprocessed data-based CPSO weight-determined combination models were observed to have optimal performance in comparison to the traditional combination method. This conclusion was based on the original data in terms of three metrics: MAPE, RMSE, and MAE.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

References

[1] J. Janczura, S. Trückb, R. Weron, and R. C. Wolff, "Identifying spikes and seasonal components in electricity spotprice data: a guide to robust modeling," *Energy Economics*, vol. 38, pp. 96–110, 2013.

[2] A. Cifter, "Forecasting electricity price volatility with the Markov-switching GARCH model: evidence from the Nordic electric power market," *Electric Power Systems Research*, vol. 102, pp. 61–67, 2013.

[3] S. K. Aggarwal, L. M. Saini, and A. Kumar, "Electricity price forecasting in deregulated markets: a review and evaluation," *International Journal of Electrical Power and Energy Systems*, vol. 31, no. 1, pp. 13–22, 2009.

[4] E. Hickey, D. G. Loomis, and H. Mohammadi, "Forecasting hourly electricity prices using ARMAX-GARCH models: an application to MISO hubs," *Energy Economics*, vol. 34, no. 1, pp. 307–315, 2012.

[5] N. Amjady and F. Keynia, "A new prediction strategy for price spike forecasting of day-ahead electricity markets," *Applied Soft Computing Journal*, vol. 11, no. 6, pp. 4246–4256, 2011.

[6] A. I. Arciniegas and I. E. Arciniegas Rueda, "Forecasting short-term power prices in the Ontario Electricity Market (OEM) with a fuzzy logic based inference system," *Utilities Policy*, vol. 16, no. 1, pp. 39–48, 2008.

[7] T. M. Christensen, A. S. Hurn, and K. A. Lindsay, "Forecasting spikes in electricity prices," *International Journal of Forecasting*, vol. 28, no. 2, pp. 400–411, 2012.

[8] S. Bordignon, D. W. Bunn, F. Lisi, and F. Nan, "Combining day-ahead forecasts for British electricity prices," *Energy Economics*, vol. 35, pp. 88–103, 2013.

[9] J. M. Bates and C. W. J. Granger, "The combination of forecasts," *Operational Research Quarterly*, vol. 20, no. 4, pp. 451–468, 1969.

[10] R. R. Andrawis, A. F. Atiya, and H. El-Shishiny, "Combination of long term and short term forecasts, with application to tourism demand forecasting," *International Journal of Forecasting*, vol. 27, no. 3, pp. 870–886, 2011.

[11] M. Khashei, M. Bijari, and G. A. Raissi Ardali, "Hybridization of autoregressive integrated moving average (ARIMA) with probabilistic neural networks (PNNs)," *Computers and Industrial Engineering*, vol. 63, no. 1, pp. 37–45, 2012.

[12] X. WeiLi, S. J. Cho, and S. T. Kim, "Combined use of BP neural network and computational integral imaging reconstruction for optical multiple-image security," *Optics Communications*, vol. 315, pp. 147–158, 2014.

[13] R. Hecht-Nielsen, "Kolmogorov's mapping neural network existence theorem," in *Proceedings of the IEEE 1st International Conference on Neural Networks*, pp. 11–13, 1987.

[14] C. Ren, N. An, J. Wang, L. Li, B. Hu, and D. Shang, "Optimal parameters selection for BP neural network based on particleswarm optimization: a case study of wind speed forecasting," *Knowledge-Based Systems*, vol. 56, pp. 226–239, 2014.

[15] P. Louka, G. Galanis, N. Siebert et al., "Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 96, no. 12, pp. 2348–2362, 2008.

[16] Y. Wang, J. Wang, G. Zhao, and Y. Dong, "Application of residual modification approach in seasonal ARIMA for electricity demand forecasting: a case study of China," *Energy Policy*, vol. 48, pp. 284–294, 2012.

[17] Z. H. Guo, J. Wu, H. Y. Lu, and J. Z. Wang, "A case study on a hybrid wind speed forecasting method using BP neural network," *Knowledge-Based Systems*, vol. 24, no. 7, pp. 1048–1056, 2011.

[18] M. Chaabene and M. Ben Ammar, "Neuro-fuzzy dynamic model with Kalman filter to forecast irradiance and temperature for solar energy systems," *Renewable Energy*, vol. 33, no. 7, pp. 1435–1443, 2008.

[19] H. Chen, *The Validity of the Theory and Its Application of Combination Forecast Methods*, Science Press, Beijing, China, 2008.

[20] S. Alam, G. Dobbie, Y. S. Koh, P. Riddlea, and S. U. Rehman, "Research on particle swarm optimization based clustering: a systematic review of literature and techniques," *Swarm and Evolutionary Computation*, 2014.

[21] W. Zhang, D. Ma, J. Wei, and H. Liang, "A parameter selection strategy for particle swarm optimization basedon particle positions," *Expert Systems with Applications*, vol. 41, no. 7, pp. 3576–3584, 2014.

[22] S. J. Nanda, G. Panda, and B. Majhi, "Improved identification of Hammerstein plants using new CPSO and IPSO algorithms," *Expert Systems with Applications*, vol. 37, no. 10, pp. 6818–6831, 2010.

[23] X. Yuan, J. Zhao, Y. Yang, and Y. Wang, "Hybrid parallel chaos optimization algorithm with harmony searchalgorithm," *Applied Soft Computing*, vol. 17, pp. 12–22, 2014.

[24] X. Y. Meng, S. Huang, and Y. Li, "Analysis of chaotic PSO and application in ship design," *Computer Engineeringand Applications*, vol. 46, no. 17, pp. 224–228, 2010.

[25] H. Jiang, C. K. Kwong, Z. Chen, and Y. C. Ysim, "Chaos particle swarm optimization and T-S fuzzy modeling approaches to constrained predictive control," *Expert Systems with Applications*, vol. 39, no. 1, pp. 194–201, 2012.

[26] J. M. Liu and Y. L. Gao, "Chaos particle swarm optimization algorithm," *Computer Application*, vol. 28, no. 2, pp. 322–325, 2008.

[27] Y. Gao and S. L. Xie, "Chaos particle swarm optimization algorithm," in *Computer Science and Technology*, 2004.

[28] J. Forrest and K. Dunn, "Cultural diversity, racialisation and the experience of racism in rural Australia: the South Australian case," *Journal of Rural Studies*, vol. 30, pp. 1–9, 2013.

[29] Z. Francis, C. Villagrasa, and I. Clairand, "Simulation of DNA damage clustering after proton irradiation using an adapted DBSCAN algorithm," *Computer Methods and Programs in Biomedicine*, vol. 101, no. 3, pp. 265–270, 2011.

[30] H. Jiang, J. Li, S. Yi, X. Wang, and X. Hu, "A new hybrid method based on partitioning-based DBSCAN and ant clustering," *Expert Systems with Applications*, vol. 38, no. 8, pp. 9373–9381, 2011.

[31] W. Zhao, J. Wang, and H. Lu, "Combining forecasts of electricity consumption in China with time-varying weights updated by a high-order Markov chain model," *Omega*, vol. 45, pp. 80–91, 2014.