Research Article

Adaptive Combination Forecasting Model Based on Area Correlation Degree with Application to China's Energy Consumption

Zhou Cheng[1] and Chen XiYang[2]

[1] *School of Logistics and Engineering Management, Hubei University of Economics, Hubei, Wuhan 430205, China*
[2] *School of Energy & Power Engineering, Huazhong University of Science & Technology, Hubei, Wuhan 430074, China*

Correspondence should be addressed to Chen XiYang; chenxiyang@mail.hust.edu.cn

To accurately forecast energy consumption plays a vital part in rational energy planning formulation for a country. This study applies individual models (BP, GM (1, 1), triple exponential smoothing model, and polynomial trend extrapolation model) and combination forecasting models to predict China's energy consumption. Since area correlation degree (ACD) can comprehensively evaluate both the correlation and fitting error of forecasting model, it is more effective to evaluate the performance of forecasting model. Firstly, the forecasting model's performances rank in line with ACD. Then ACD is firstly proposed to choose individual models for combination and determine combination weight in this paper. Forecast results show that combination models usually have more accurate forecasting performance than individual models. The new method based on ACD shows its superiority in determining combination weights, compared with some other combination weight assignment methods such as: entropy weight method, reciprocal of mean absolute percentage error weight method, and optimal method of absolute percentage error minimization. By using combination forecasting model based on ACD, China's energy consumption will be up to 5.7988 billion tons of standard coal in 2018.

1. Introduction

Energy is a substantial foundation for the consistently high growth rate of China's economy. Since the policy of Reforming and Opening was initialized, China has become the second largest energy consumer of the world by 2005. As the most flourishing developing country that has a significant influence on current global energy field, China's energy consumption has become a global concern [1]. Its primary energy consumption has been more than doubled in less than a decade, according to the National Energy Administration of China. The GDP of China maintained a high growth rate of 10.3% in 2010, reaching RMB (Chinese currency) 39.80 trillion Yuan. Meanwhile, its energy consumption is up to 3.25 billion tons of standard coal, which increased by 5.9% compared to 2009 [2, 3]. It is well known that the consistently high growth rate of China's economy is supported by huge energy consumption. With tremendously increasing energy consumption, the problems of greenhouse gas emissions and energy shortage have become more serious.

In order to steer its energy-economy-environment system onto a more sustainable pathway, China is now changing its economic development model from energy-intensive economy to low-carbon economy. During the process of economic development model transition, one of challenges that China continues to face in the energy field is the consistent imbalance between energy supply and demand. Most of all, the imbalance will not be solved in the foreseeable future. It will be more serious in the coming years. To accurately forecast China's energy consumption can be profoundly instructive to the economic sustainable development of China and even the world. Energy consumption forecasting is a vital part of energy policy for a country all over the world.

For a rapidly developing country, the main factors that influence China's energy consumption are sustaining high

economic growth rate, increasing human population, large-scale industrialization, motorization process, energy structure optimization, reducing greenhouse gas, and improving energy efficiency [4]. Among these influential factors, the economic growth rate is identified as a major determining factor, and the primary energy structure is identified as a major restricted factor. It is very difficult to accurately forecast China's energy consumption for these reasons. In view of this, many scholars at home and aboard try to create advanced energy consumption forecasting models to help China make better government policies. Various studies of energy consumption forecasting models can be broadly classified into two overall groups: single forecasting model and combination forecasting model.

A substantial body of the research had been conducted on the single forecasting model in the energy consumption field during the past decade. Christodoulakis et al. [5] forecasted the future demand for energy in Greece based on scenario analysis. Murat and Ceylan [6] adopted an artificial neural network approach to forecast the transport energy demand, by use of socioeconomic and transport related indicators in social economy and transport. Crompton and Wu [7] applied a Bayesian vector autoregressive method to forecast China's energy consumption. Chen [8] proposed a collaborative fuzzy-neural approach to forecast the long-term load, which aggregated the fuzzy different experts' annual energy consumption forecasts. Geem [9] developed an artificial neural network model by using various independent variables (such as GDP, population, oil price, number of vehicle registrations, and passenger transport amount) as input parameters to forecast South Korea's transport energy demand. Forouzanfar et al. [10] introduced a new multilevel genetic programming approach for forecasting transport energy demand in Iran, and its result had smaller error compared with the result obtained using neural network or fuzzy linear regression approach. Kumar and Jain [11] applied Grey-Markov model to forecast crude-petroleum consumption, Grey model with rolling mechanism to forecast coal and electricity (in utilities) consumption, and singular spectrum analysis to predict natural gas consumption in India. Lu et al. [12] adopted GM $(1, 1)$ by modifying the value of the development coefficient to reflect the influence of economic growth on the forecasting results. Hodge et al. [13] used ARIMA models to forecast future wind power output from historical data and applied a systematic approach to determine the best values for the assortment of variables associated with the models, such as training period length and model orders. Kankal et al. [14] forecasted energy consumption in Turkey based on socioeconomic and demographic variables (gross domestic product (GDP), population, import and export amounts, and employment) using artificial neural network and regression analyses. Lee and Tong [15] developed an improved grey forecasting model for Chinese energy consumption, which combined residual modification with genetic programming sign estimation. Chenchen et al. [16] established a new class of weakening buffer operators with exponential type to forecast energy demand by using the exponential average, weighted average, and geometrical average weakening buffer operators. Pao [17] found ANN method more appropriate than the linear method to forecast the electricity consumption in Taiwan. Dong et al. [18] adopted support vector machines, a new neural network algorithm, to forecast building energy consumption in the tropical region. Chedid et al.[19] built an econometric model to predict the future production trend of crude oil and natural gas based on the behavior of them. Some energy or related agencies practiced energy elasticity coefficients method [20], input-output method [21], cointegration theory, and error correction models [22] to forecast energy consumption, and so on.

But for all of these single forecasting models mentioned above, there are some limitations due to the hypothetical conditions, the applying scope, and especially not using lots of useful information effectively. Therefore, combination method was put forward to make good use of source information from different perspectives. Since combination forecasting model properly links the different forecasting methods and makes the best use of the information given by each of them, it can overcome the limitations of individual models and gather the information more effectively.

A linear combination forecast model was first established by Bates and Granger [23] in 1969 which considered as a successful alternative for using an individual forecasting method. Xue et al. [24] made use of neural network, grey forecast, and time series to construct a combination forecasting model to predict Chinese energy consumption. Under departures from elliptically symmetric distributions, Elliott and Timmermann [25] proved that the optimal combination weights under asymmetric loss can be much more different than under mean square error (MSE) loss. Guohao et al. [26] developed a combination forecasting model by use of grey system and multiple regression method to predict China's coal demand, and the variable combination weight was the type of function that changed with forecast time. Chai et al. [27] found that the outcome of combining forecasting will be more precise based on the Bayesian error correction approach. In order to improve the forecasting accuracy of energy consumption of Hebei Province in China, Song et al. [28] proposed a combination forecasting model based on introduction orderly weighted average (IOWA) operator. Xu and Wang [29] used polynomial curve and moving average combination projection model to estimate the future natural gas consumption in China from 2009 to 2015 and employed ordinary least squares to estimate the combination weights. Yu et al. [2] proposed a combination forecasting model that optimized the coefficients of the three forms of the model (linear, exponential, and quadratic model) based on PSO-GA. Lemke and Gabrys [30] developed a ranking-based combination of methods to improve forecasting performance over simple model selection approaches. Zhou and Wang [31] proposed a combination forecasting model based on multi-criteria, in which the combination weights are obtained by using genetic algorithms. Sánchez [32] proposed an adaptive forecast combination (AEC) procedure based on a two-step procedure: firstly, different combination procedures are used; secondly, the AEC is used to combine the combinations from the first step. In view of the individual forecasting model with the smaller variance of mean absolute percentage error (MAPE) has higher stability, its combination weight should

be larger. Meanwhile information entropy is a measure of the uncertainty associated with a random variable; an assessment method of the combination weights is adopted on the basis of entropy approach by Zhao [33]. The literature in the field of nonlinear forecasting combination is very sparse due to the lack of evidence of success such as what stated in [34]. Only the linear combination method is discussed in this paper.

Although a large body of literature applied various methods to construct combination forecasting models, there are some limitations of choosing the individual models for combination and determining the combination weight. The weights of individual forecasting model to be combined are based on the relative past forecasting performance, and no single method can guarantee more accurate results. The commonly used criterions in evaluating out-of-sample forecast capability include simple average, root mean square error, mean absolute error, and mean absolute percentage forecast error. Creating an efficient method of distributing combination weights is not easy because of the difficulty in choosing the estimation criterion.

The method for optimally selecting models is vital for the combination forecasting performance. Therefore, ACD is proposed to be the selection criterion in constructing the combination model in this paper. Also the combination weights are determined based on the sequence relative area correlation degree. The ACD of individual model is obtained by finding the area of the curves between actual and forecasted values, which is a comprehensive performance indicator of correlation and matching error between actual and forecasted values.

In this paper, based on China's energy consumption data (1979~2010), a combination forecasting model is developed to forecast the energy consumption of China during 2011~2018. Most of all, ACD is proposed as the selection criterion of individual model and the method of determining the combination weights.

The rest of this paper is organized as follows. In Section 2, four single forecasting methods are briefly introduced; the individual model selection method based on ACD is proposed; a combination forecasting model is developed, in which the combination weights are also determined based on the ACD of individual forecasting model. Section 3 presents the empirical analysis using the methods mentioned in Section 2, all of which are adopted to forecast China's energy consumption. Section 4 concludes the paper and suggestions on China's energy policy.

2. Methodology

Combination forecasting method is under the guideline of linking the different forecasting methods and making good use of the information given by each one. The performance of combination forecasting model is mainly associated with the performance consistency of each individual forecasting model and the assignment combination weights. Meanwhile the forecasting error becomes a random variable because of the great uncertainty of future. In order to develop a

highly forecasting performance combination model to predict China's energy consumption, individual model choice and combination weights assignment are emphasized in this paper. In this section, the theories of four individual forecasting models are discussed, individual model selection based on ACD is proposed, also the combination weights assignment based on ACD is adopted, and a new combination forecasting modeling method is developed.

2.1. Individual Forecasting Model. Although some individual models adopted to forecast time series are considered simple, they are popular and surprisingly powerful. BP neural network, GM (1, 1), exponential smoothing model, and polynomial trend extrapolation model are presented as follows since they are used widely and efficiently in the field of energy consumption forecasting.

2.1.1. BP Neural Network Model. For the ability of learning, storing, and recalling information based on a given data set, neural network model has been successfully employed in the field of energy consumption forecasting. Many of multivariate nonlinear projects can be solved by neural network. BP neural network is a kind of multilayer forward propagation neural network based on the data stream forward propagation and inverse error propagation algorithm. BP consists of input layer, hidden layer, and output layer, with the different layer neurons by connective weight and threshold. More detailed introduction about BP neural network can be found in [35]. The total available data are split into a training set and a test set in the process of BP forecasting. The training set is to calibrate the neural network and the test set is to evaluate BP's forecasting performance. The modeling steps of BP neural network forecasting model are briefly described as below [36].

Step 1. Normalize the input and output data.

Step 2. Initialize the value for every connective weight of network, and set up the maximum permitted error, learning rate, and momentum term.

Step 3. Prepare the training input data and expect output data for training network.

Step 4. Perform feed-forward computation during the process of network training.
Input of input layer is χ_i; input net_j of hidden layer is as formula (1); output out_j of hidden layer is as formula (2):

$$\text{net}_j = \sum_{j=1}^{k} \omega_{ji} \chi_i + \theta_j, \tag{1}$$

$$\text{out}_j = \frac{1}{1 + e^{-\text{net}_j}}, \tag{2}$$

where ω_{ji} is connecting weight between node j and node i and θ_j is the threshold of node j.

Step 5. The computation of error back-forward propagation is as follows:

$$\omega_{ij}(k+1) = \omega_{ij}(k) + \alpha \cdot \sigma_j \cdot \text{out}_j, \tag{3}$$

where σ_j is the error between input layer and output layer and α ($0 < \alpha < 1$) is the learning rate.

Step 6. In order to meet the need of the maximum permitted error during the process of network training, fore-forward and back-forward propagation computation are computed recursively until convergence.

Step 7. Evaluate the network performance by use of the test data. Also the BP neural network forecasting model is established.

2.1.2. GM (1, 1) Model. Grey system theory is proposed by Professor Deng in 1982. It is spread into the field of prediction of known information to uncertain areas. GM (1, 1) is applied widely in grey system forecasting. The modeling steps of GM (1, 1) forecasting model are described as follows [37].

Step 1. Assume the actual data series $x^{(0)}$ with n samples is expressed as

$$x^{(0)} = \left[x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n)\right], \quad n \geq 4, \tag{4}$$

where the actual data is assumed to be positive, which is the need of Grey system forecasting theory.

Step 2. Generate a new data series $x^{(1)}$ by accumulated operation (AGO)

$$x^{(1)}(0) = \left[x^{(1)}(1), x^{(1)}(2), \ldots, x^{(1)}(n)\right], \tag{5}$$

where

$$x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), \quad i = 1, 2, \ldots, k. \tag{6}$$

Step 3. Establish GM (1, N) matrix equation.

Since both data series $x^{(1)}$ and the solution of a first order difference equation meet the requirements of the exponential increase law, $x^{(1)}$ should conform to a first order difference equation as $dx^{(1)}/dt + ax^{(1)} = u$. By use of dispersed method solving, GM (1, N) matrix equation can be established: $Y_n = BA$, where

$$Y_n = \begin{pmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{pmatrix}, \qquad A = \begin{pmatrix} a \\ u \end{pmatrix},$$

$$B = \begin{pmatrix} -\dfrac{1}{2}\left[x^{(1)}(1) + x^{(1)}(2)\right] & 1 \\ -\dfrac{1}{2}\left[x^{(1)}(2) + x^{(1)}(3)\right] & 1 \\ \vdots & \vdots \\ -\dfrac{1}{2}\left[x^{(1)}(n-1) + x^{(1)}(n)\right] & 1 \end{pmatrix}. \tag{7}$$

Then the only unknown parameter is matrix A.

Step 4. Solve parameters matrix A based on least squares method. Let $x^{(1)}(1) = x^{(0)}(1)$; then

$$\hat{x}^{(1)}(k+1) = \left[x^{(0)}(1) - \frac{\hat{u}}{\hat{a}}\right] e^{-\hat{a}k} + \frac{\hat{u}}{\hat{a}}, \tag{8}$$

where $k = 0, 1, 2, \ldots$.

Step 5. By taking inverse AGO on sequence $\hat{x}^{(1)}$, the actual data forecasting $\hat{x}^{(0)}$ is obtained. Also, GM (1, 1) forecasting model is established:

$$\hat{x}^{(0)}(k+1) = \left(1 - e^{\hat{a}}\right)\left(x^{(0)}(1) - \frac{\hat{u}}{\hat{a}}\right) e^{-\hat{a}k}, \tag{9}$$

where $k = 0, 1, 2, \ldots$.

2.1.3. Triple Exponential Smoothing Model. Triple exponential smoothing model (TESM) has been around since 1950s and is still one of the most popular forecasting methods. It produces an exponentially weighted moving average in which each smoothing calculation or forecasting is dependent on the previous data. The modeling steps of TESM are as follows [38].

Step 1. Implement three-time smoothing calculation in line with formula (10) as follows:

$$\begin{aligned} S_t^{(1)} &= \partial y_t + (1 - \partial) S_{t-1}^{(1)}, \\ S_t^{(2)} &= \partial S_t^{(1)} + (1 - \partial) S_{t-1}^{(2)}, \\ S_t^{(3)} &= \partial S_t^{(2)} + (1 - \partial) S_{t-1}^{(3)}, \end{aligned} \tag{10}$$

where ∂ is the smoothing factor ($0 \leq \partial \leq 1$), when ∂ closer to 1 means more weigh to the recent observations and hence more rapidly changing forecast; y_t is the actual data set; and $S_T^{(n)}$ is the estimated value for the n order exponential smoothing.

Step 2. Construct the triple exponential smoothing model as

$$\hat{y}_{t+m} = a_t + b_t m + c_t m^2, \tag{11}$$

where $m = 1, 2, 3, 4, \ldots,$

$$a_t = 3S_t^{(1)} - 3S_t^{(2)} + S_t^{(3)},$$

$$b_t = \frac{\partial}{2(1-\partial)^2}$$
$$\times \left[(6 - 5\partial) S_t^{(1)} - 2 (5 - 4\partial) S_t^{(2)} + (4 - 3\partial) S_t^{(3)} \right], \quad (12)$$

$$c_t = \frac{\partial^2}{2(1-\partial)^2} \left[S_t^{(1)} - 2S_t^{(2)} + S_t^{(3)} \right].$$

Let $m = 1$ and the triple exponential smoothing model be as

$$\widehat{y}_{t+1} = \frac{3 - 3\partial + \partial^2}{(1-\partial)^2} S_t^{(1)} - \frac{3 - \partial}{(1-\partial)^2} S_t^{(2)} + \frac{1}{(1-\partial)^2} S_t^{(3)}. \quad (13)$$

The estimating value of each moment can be calculated by assigning $t = 0, 1, 2, \ldots$.

2.1.4. Polynomial Trend Extrapolation Model. Polynomial trend extrapolation model (PTEM) is a very simple forecasting method that adopts previous data trends to project future data. Polynomial trend extrapolation forecasting model is to identify the overall past trend and to fit proper polynomial $\Phi(x)$. The object function for $\Phi(x)$ can be expressed as

$$\min I = \sum_{i=0}^{m} \left(\sum_{k=0}^{n} a_k x_i^k - y_i \right)^2, \quad (14)$$

where $\Phi(x) = \sum_{k=0}^{n} a_k x_i^k$; (x_i, y_i) is an actual data point; $i = 0, 1, 2, 3, \ldots, m$; and a_k is the coefficient of $\Phi(x)$; $n \le m$.

The necessary condition for the extreme values of polynomial is $\partial I / \partial a_j = 0$.

Therefore

$$\sum_{k=0}^{n} \left(\sum_{i=0}^{m} x_i^{j+k} \right) a_k = \sum_{i=0}^{m} x_i^j y_i, \quad (15)$$

where $j = 0, 1, \ldots, n$.

The matrix representation of formula (16) is as follows:

$$\begin{bmatrix} m+1 & \sum_{i=0}^{m} x_i & \cdots & \sum_{i=0}^{m} x_{i_i}^n \\ \sum_{i=0}^{m} x_i & \sum_{i=0}^{m} x_i^2 & \cdots & \sum_{i=0}^{m} x_{i_i}^{n+1} \\ \vdots & \vdots & & \vdots \\ \sum_{i=0}^{m} x_{i_i}^n & \sum_{i=0}^{m} x_{i_i}^{n+1} & \cdots & \sum_{i=0}^{m} x_{i_i}^{2n} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \sum_{i=0}^{m} y_i \\ \sum_{i=0}^{m} x_i y_i \\ \vdots \\ \sum_{i=0}^{m} x_i^n y_i \end{bmatrix}. \quad (16)$$

Therefore the coefficients of $\Phi(x)$ can be obtained through formula (16) based on least squares method.

2.2. Combination Forecasting Model. Combination forecasting model is a reliable method of improving forecast accuracy by making good use of the different strength of various models. If there are m models to forecast an actual data set $y(t)$, the method of combination forecasting model constructed by these individual models is described as

$$\widehat{y}(t) = \sum_{i=1}^{m} w_i y_i'(t), \quad (17)$$

where w_i is the weight of the i model and $\sum_{i=1}^{m} w_i = 1$; $y_i'(t)$ is the forecasting value of the i model; and $\widehat{y}(t)$ is the forecasting value of combination forecasting model.

In line with formula (17), we can infer that the forecasting performance of combination model is typically influenced by the individual forecasting model used and the combination weights assignment. This paper proposes the selection of individual forecasting model and the combination weights assignment based on ACD to improve the performance of combination forecasting model.

2.2.1. Selection of Forecasting Method Based on ACD. The ability of evaluating and selecting a best-fitting model among the different methods that is appropriate for a particular time series is considered to be one of the most important aspects in the field of combination forecasting. MAPE is the most widely adopted one to describe the accuracy of forecasting models in statistics. The expression of MAPE is as follows:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - \widehat{y}_i}{y_i} \right|, \quad (18)$$

where y_i is the actual value; $\widehat{y}$ is the forecast value; and n is the number of data series.

Although MAPE is very simple, a very high MAPE might distort a comparison among models when calculating the MAPE for a few data. In other words, MAPE cannot always be effective to represent the correlation between forecasting model and actual time series' development. This paper proposes a new area correlation degree criterion for the model selection, which is easy to identify a candidate model for the selection of combination forecasting. The principle of ACD criterion is the better high performance, the larger ACD. Since it comprehensively evaluates both the correlation and fitting error of the forecasting and actual values', ACD model selection criteria can be introduced as an indicator for the performance of forecasting model. ACD is obtained by way of calculating the area of curves between actual and forecasted value.

First of all, given X_0 is the actual data, X_j is the forecasting data. Then X_0 and X_j can consist of two curves (actual curve and forecasting curve) as in Figure 1. When the area is zero, the forecasting model is completely fitting to the actual data and the ACD is 1. The smaller the area of the curves between the actual and forecasted value is, the larger the ACD becomes. At the same time, the forecasting model has more high performance.

The process of calculating the ACD is as follows [39].

Step 1. Given X_0 is the actual data, X_j is the forecasting data:

$$X_o = \left(x_o\left(1\right), x_o\left(2\right), \ldots, x_o\left(n\right)\right);$$
$$X_j = \left(x_j\left(1\right), x_j\left(2\right), \ldots, x_j\left(n\right)\right). \tag{19}$$

Step 2. Define $f(i)$ as a sign function as follows:

(1) $f(i) > 0$, when $x_o(i) - x_j(i) > 0$;

(2) $f(i) = 0$, when $x_o(i) - x_j(i) = 0$;

(3) $f(i) < 0$, when $x_o(i) - x_j(i) < 0$.

Step 3. Calculate the area of curves between actual and forecasted curves. Firstly, we calculate S_{ji}, which is the area of time span from number (i) to number ($i+1$).

(1) If $f(i) \times f(i + 1) \geq 0$, then

$$S_{ji} = \frac{1}{2} \left| \left[\left(x_o\left(i\right) - x_j\left(i\right)\right) + \left(x_o\left(i+1\right) - x_j\left(i+1\right)\right) \right] \times 1 \right|. \tag{20}$$

(2) If $f(i) \times f(i + 1) < 0$, then

$$S_{ji} = \frac{1}{2} \left| \left(x_o\left(i\right) - x_j\left(i\right)\right) \times \left(x - i\right) \right| + \frac{1}{2} \left| \left(x_o\left(i+1\right) - x_j\left(i+1\right)\right) \times \left((i+1) - x\right) \right|, \tag{21}$$

where

$$x = i + \frac{\left| \left(x_j\left(i\right) - x_o\left(i\right)\right) \right|}{\left| \left(x_j\left(i+1\right) - x_o\left(i+1\right)\right) \right| + \left| \left(x_j\left(i\right) - x_o\left(i\right)\right) \right|}. \tag{22}$$

Step 4. Define γ_i as the ACD of actual data curve and forecasting data curve:

$$\gamma_i = \frac{1}{1 + \left(4S_{ji} / \sum_{j=1}^{M} \sum_{i=1}^{N-1} S_{ji}\right)}, \tag{23}$$

where M is the number of individual models for combination model selecting candidates and N is the number of data series for calculation.

In view of removing the effect of base number (various energy consumption at each time moment), ACD is calculated based on the area of the curves between x-axis and forecasted percentage error curve in this paper.

2.2.2. Combination Weights Assignment Based on ACD. The assessment of the individual weights has become a hot point for research institutes. Generally, it can be objectively obtained by solving a mathematical optimal object function with some constraints by regression method. There are some drawbacks during the assignment of weights; for example, the objective function is not easy to resolve and the weights are not the optimal ones for the nonpositive constraint condition of weights. Most of all, it cannot grasp the key principle of combination forecasting that makes full use of information from different aspects.
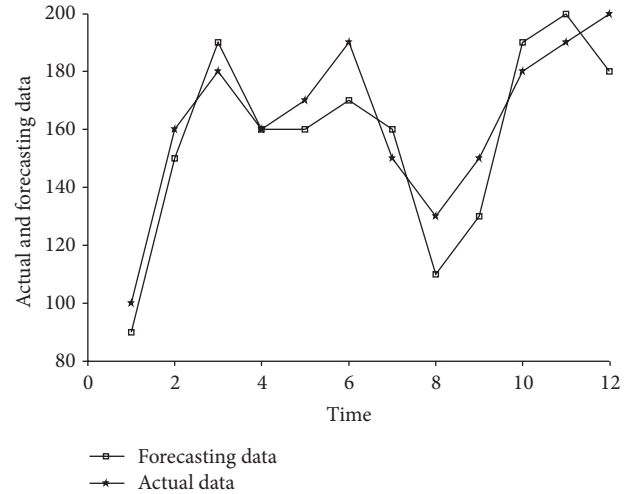

FIGURE 1: Schematic diagram of actual and forecasting curve.

As mentioned above, the changing trend of area between actual and forecasted can be an indicator for the stability of individual forecasting model. Since the stability of individual forecasting model is higher when the changing trend of the curve area is smaller, its weight of combination forecasting should be larger. For the individual forecasting model, the better the performance is, the larger the ACD is. For these reasons, an assessment method of the combination weights is adopted on the basis of ACD. The combination weights assignment based on ACD is formed as follows.

Step 1. Calculate ACD of each individual forecasting model, given that γ_j is the ACD of number j individual forecasting model.

Step 2. Obtain the variance coefficient D_j:

$$D_j = 1 - \gamma_j, \tag{24}$$

where γ_j is the ACD of number j individual forecasting model, $j = 1, 2, \ldots, k$.

Step 3. Assess the individual weights l_j:

$$l_j = \frac{1}{k - 1} \left[1 - \left(\frac{D_j}{\sum_{j=1}^{k} D_j} \right) \right], \tag{25}$$

where l_j is the weight of number j individual forecasting model, $j = 1, 2, \ldots, k, \sum_{j=1}^{k} l_j = 1$.

3. Result and Discussion

3.1. Data Source. Actual data on "Total Energy Consumption in China 1979~2010" is from various issues of the China Statistical Yearbook and China Energy Statistical Yearbook. Meanwhile, the unit of energy is ten thousand tons of standard coal. Figure 2 shows the annual energy consumption of China during the period of 1979~2010.
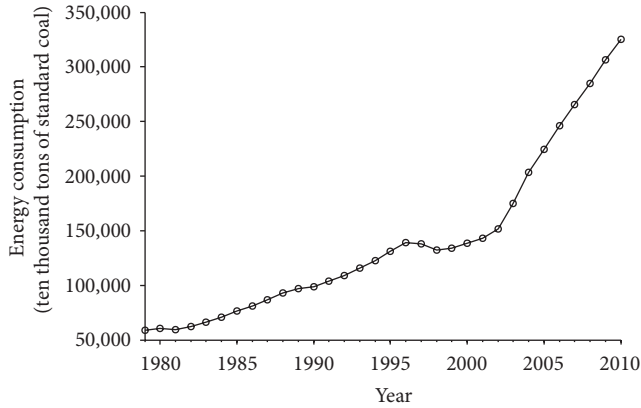
Figure 2: Annual energy consumption of China during the period of 1979~2010.

Figure 3: Forecast of China's energy consumption by BP.

Due to strong economic growth, China's demand for energy is surging rapidly. From 1979 to 2010, the energy consumption of China has increased from 0.58588 billion tons of standard coal to 3.25 billion tons of standard coal. During the early period of 1979–1996, with the energy consumption growth slowing down, the energy-intensive of China economy continued to decline. From 1997 to 1999, the energy consumption of China declined because of 1997 Asian Financial Crisis. Since 2000 to 2003, its energy consumption had entered a new round of increasing process since China had gradually recovered from the influence of 1997 Asian Financial Crisis. Especially during the period of 2003~2010, because of conducting large-scale infrastructure construction and the high economy growth, China's energy consumption showed a rapid growth.

3.2. Forecast Result. According to the above energy consumption trend and its annual data from 1979 to 2010 shown in Figure 2, this section contains the forecasting results obtained from the forecasting models described above. All the computations required for the present study have been carried out on MATLAB 2010.

3.2.1. Result Using BP. Under the restriction of a limited training set, we use BP neural networks in this paper with a 4-11-1 structure (three layers: one input layer with 4 neurons, one hidden layer with 11 neurons, and one output layer with 1 neuron, with a learning rate of 0.01) to construct the forecasting model. BP neural network was trained with the annual data from 1979 to 2004, and the remaining data from 2005 to 2010 was used for evaluating its forecasting performance. Forecasting result by BP is shown in Figure 3.

3.2.2. Result Using GM (1, 1). Using the energy consumption data of 1995~2010 for the initial data sequence to create the GM (1, 1) forecasting model, where $\hat{u}$ and $\hat{a}$ are obtained as 94084 and -0.07833, respectively, the forecasting model that based on GM (1, 1) is formed as $\hat{x}^{(0)}(t+1) = 100376.5e^{0.07833t}$, where $t = 1$ in 1996. China's energy consumption forecasting result based on GM (1, 1) is as shown in Figure 4.
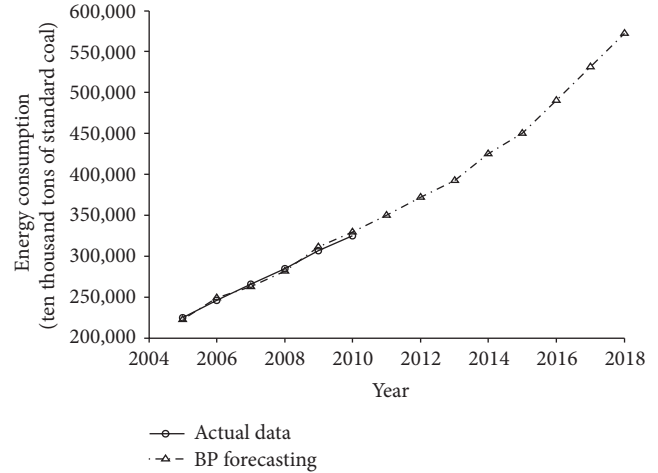
3.2.3. Result Using TESM. The energy consumption data of 1979~2010 is adopted for the initial data sequence to construct TESM. In light of the fact that the smoothing factor ∂ closer to 1 means more weigh to the recent observations and hence more rapidly changing forecast, ∂ is assumed to be 0.27 by forecasting simulation tests. The triple exponential smoothing mode is formed as $x_{2010+t} = 328350.5 + 24286t + 612.67t^2$, where $t = 1$ in 2011. In light of the forecasting process based on TESM mentioned above in the description of TESM process, China's energy consumption forecasting result based on TESM is as shown in Figure 5.

3.2.4. Result Using PTEM. Using the energy consumption data of 1979~2010 for the initial data sequence to construct PTEM as $x(t) = 18.607t^3 - 631.31t^2 + 10070t + 36885$, where $t = 0$ when in 1979, China's energy consumption forecasting result based on PTEM is as shown in Figure 6.

3.2.5. Result Using Combination Forecasting Model. Appropriate individual forecasting models should be applied to develop the combination forecasting model. Since combination forecasting model is usually constructed by 3~5 types of individual model, this paper will select 3 types of individual methods from BP, GM (1, 1), exponential smoothing model, and polynomial trend extrapolation model. Area correlation degree of each individual forecasting model is calculated based on the forecast result as reported in Table 1.

To remove the effect of base number (energy consumption at each time) during the process of calculating the correlation areas, ACD is calculated in this paper based on the area of the curves between x-axis and each individual forecasting model's percentage error curve during the period of 2005~2010. The ACD results of each individual forecasting model are reported in Table 2.

Because the better the performance, the larger the ACD, according to this study, TESM, BP, GM, and PTEM are listed in terms of the forecasting performance as shown in Table 2. Therefore, TESM, BP, and GM are chosen for developing

TABLE 1: Forecasting results of China's energy consumption based on individual model in 2005~2010.

| Year	Actual value	BP forecasting	BP APE[a] (%)	GM forecasting	GM APE (%)	TESM forecasting	TESM APE (%)	PTEM forecasting	PTEM APE (%)
2005	224682	222677.2	0.8923	219684.8	2.2241	222220	1.09577	214791.6	4.402
2006	246270	249307.2	1.2333	237583.9	3.5271	251050	1.94096	232358.8	5.6487
2007	265583	262882.2	1.0169	256941.3	3.2539	275170	3.60979	251789.4	5.1937
2008	285000	281967.7	1.064	277875.8	2.4997	294190	3.22456	273195	4.1421
2009	306600	311303	1.5339	300516	1.9843	311730	1.67319	296687.2	3.2331
2010	325000	329039	1.2428	325000.9	0.000265	331770	2.08308	322377.7	0.8069
MAPE (%)		1.1639		2.2482		2.2712		3.9044	

[a]APE: absolute percentage error.
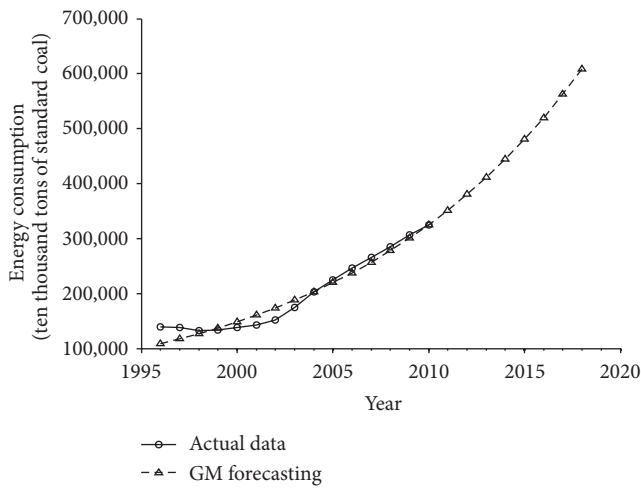
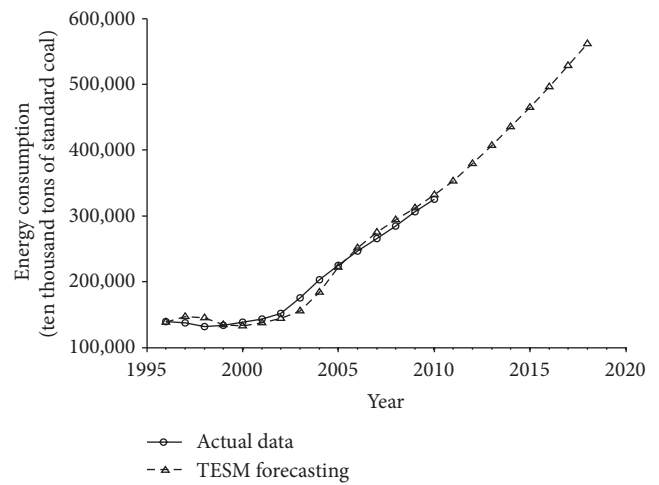FIGURE 4: Forecast of China's energy consumption by GM (1, 1).

FIGURE 5: Forecast of China's energy consumption by TESM.

TABLE 2: The ACD results of each individual forecasting model.

Number	Method	ACD
1	BP	0.7431433
2	GM	0.496139088
3	TESM	0.518059155
4	PTEM	0.369205433

FIGURE 6: Forecast of China's energy consumption by PTEM.

the combination forecasting model. Also the combination weights assignment based on ACD are reported in Table 3.

For the purpose of evaluating the performance of combination weight assignment method based on ACD, it also adopts some other combination weight assignment methods such as entropy weight method, reciprocal of MAPE weight method, and optimal method of APE minimization method. Table 4 shows the forecasting results of combination models based on different weights assignment methods (the combination forecasting model is developed by BP, GM, and TESM).
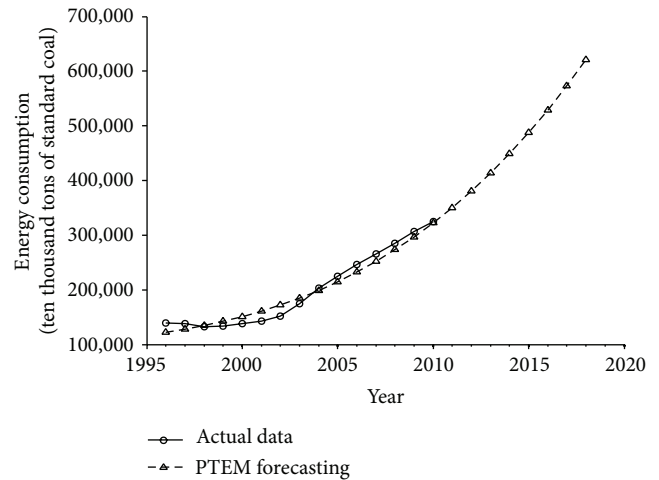
The forecasting results of China's energy consumption of 2011~2018 by combination forecasting model are shown in Table 5, which is developed by using of BP, GM, and TESM based on ACD.

3.3. Discussion. Facing great energy challenges and an energy dilemma, it is essential for China to forecast energy consumption for its sustainable development. Moreover, combination forecasting can improve the fitting effect and increase the forecasting accuracy, which makes good use of the information source from each individual forecasting model. Besides that, the relative stability of forecasting model is improved. For the reasons above, it is useful to project with a relatively simple system based on combination forecasting model in this paper. Thus, the specific purpose of this study was to determine a suitable combination forecasting model for short-term forecasting of China's energy consumption with limited data. In view of the fact that the area of the curves between actual and forecasted value is a more comprehensive performance indicator for forecasting correlation and matching error, ACD is proposed as the individual model selection criterion to develop combination model and the determination method of the combination weights. It establishes a procedure for increasing the efficiency of the forecasting combination method based on ACD. Consequently, four kinds of individual forecasting model, model selection method for the purpose of constructing combination forecasting model, and combination weight assignment method are emphatically researched in this paper. In light of all the forecasting results of China's energy consumption based on this study, we can infer the following.

As can be seen from the individual model forecasting results reported in Figures 3 to 6, China's energy consumption is also on its rapid growing process during the coming years. During the periods of 1997~1999 and 2000~2003, the simulation values based on three types of individual models (TESM, GM, and PTEM) have significant deviation from actual value. The reasons for huge simulation errors are the Asian Financial Crisis at 1997 and the fact that China's energy consumption has entered into a new round of increasing process since 2003. It proves that sudden change factors that influence China's energy consumption will be shown in forecasting results. It does not lie in forecasting model itself, but in the distortion of the behavioral data series with the disturbance of the distracter.

Table 1 summarizes the simulation results of China's energy consumption during the period of 2005~2010 based on BP, TESM, GM, and PTEM. These four individual forecasting models rank as BP, GM, TESM, and PTEM by using MAPE to be an indicator of forecasting performance. All the individual forecasting models have highly accurate forecasts (MAPE results are less than 4%).

Most of all, this study proposed ACD as an indicator of forecasting performance. Table 2 shows that these four individual forecasting models rank as BP, TESM, GM, and PTEM in terms of the forecasting correlation and fitting error based on ACD. The performance of forecasting based on ACD is different from that based on MAPE. Since MAPE might distort a comparison among models when calculating a few data, it is not always effective to represent the correlation between forecasting model and actual data series' development trend. Therefore, the MAPE of forecasting result based on GM, TESM, and PTEM in the period of 1997~2010 is calculated as 5.43%, 4.65%, and 5.79%. The

Table 3: Combination weight of combination forecasting model by TESM, BP, and GM.

Individual model	BP	GM	TESM
Combination weight	0.39665032	0.29726492	0.30608475

performance of forecasting models rank based on MAPE by using relative more data is in accordance with its ranking based on ACD by using more data. We can infer that ACD is more effective to represent the correlation and fitting error between forecasting model and actual time series' changing trend when calculating a few data. Hence ACD can be an indicator of the forecasting model's performance.

By using ACD to choose appropriate candidate individual models for constructing combination model and determining the combination weight, a new combination forecasting method is proposed in this paper. Table 3 reports the combination weight assignment based on ACD. It also reveals that the larger ACD, the higher the forecasting performance. Furthermore, Table 4 lists the combination forecasting results based on different weights assignment methods. It concludes that combination forecasting model can improve forecasting accuracy by making good use of source information from each individual forecasting model. In comparison of forecasting performance, only the combination forecasting model based on optimal method is better than the combination forecasting model based on ACD. Taking into account the time cost of calculation and the fact that it does not grasp the key principle of combination forecasting, combination assignment based on optimal method does not make full use of information from different aspects. As a result, ACD can be used to combine the selected forecasts; in other words, it is feasible to determine the combination weights based on ACD.

From the above analysis and comparison, we adopt BP, TESM, and GM to develop a combination forecasting model based on ACD for the purpose of forecasting China's energy consumption of 2011~2018. From the forecasting results in Table 5, we can infer that China is still on a rapid energy consumption path. The total energy consumption of China in 2018 will increase to 5.798812 billion tons of standard coal. With the rapid growth of energy consumption, China should adopt energy-saving policies to reduce greenhouse gas emissions intensity, develop circular economy, and actively promote low-carbon technologies to respond to climate change.

4. Conclusions

A new combination forecasting model based on ACD is developed to accurately predict China's energy consumption by using the data of "Total Energy Consumption in China 1979~2010". Since ACD can be an indicator of the correlation and fitting error of the forecasted and actual values, it is proposed to choose appropriate candidate individual models for constructing combination model and to determine the combination weight based on ACD in this paper. Based on the forecasting results, it can provide more accurate and reliable forecasting results by using combined forecasting

TABLE 4: Forecasting results of China's energy consumption by combination models based on different weights assignment methods in 2005~2010.

Year	Actual	ACD	APE (%)	Entropy weight[b]	APE (%)	Reciprocal of MAPE weight method[c]	APE (%)	Optimal method[d]	APE (%)
2005	224682	221647.7	1.35	221535.5	1.4	221798.7	1.28	221584.1	1.38
2006	246270	246355.7	0.03	246367	0.04	246757.6	0.2	246270	0
2007	265583	264877.3	0.27	266339.7	0.29	264469	0.42	265361.7	0.08
2008	285000	284492.4	0.18	285979.8	0.34	284009.4	0.35	285000	0
2009	306600	308227.1	0.53	308087.3	0.49	308660	0.67	308093.8	0.49
2010	325000	328674.5	1.13	328950.7	1.22	328698.6	1.14	328743.6	1.15
MAPE (%)		0.58			0.63		0.68		0.51

[b]Entropy weight: weight assigned based on information entropy as adopted in [33].
[c]Reciprocal of MAPE method: weight assigned based on the reciprocal of each individual forecasting model's MAPE.
[d]Optimal method: the fundamental principle used to obtain the weight coefficient of combination forecasting is to make the APE minimization.

TABLE 5: Forecasting results of China's energy consumption by combination model (BP, GM, and TESM) based on ACD in 2011~2018 (unit: ten thousand tons of standard coal).

Year	BP	GM	TESM	Combination forecasting based on ACD
2011	350628.9	351497.7	353249.2	351689.2
2012	372956.5	380137.5	379373.2	377055.2
2013	394915.3	411111.0	406722.5	403343.7
2014	425057.9	444608.1	435297.2	434003.6
2015	450432.6	480834.5	465097.3	463958.6
2016	490235.2	520012.6	496122.6	500889.0
2017	531573.5	562383.0	528373.3	539752.6
2018	572568.4	608205.7	561849.4	579881.2

method based on ACD. Therefore, ACD is more effective to evaluate the forecasting performance than MAPE with fewer data. All of these prove that it is feasible for us to develop a combination forecasting model based on ACD. The combination forecasting model proposed in this paper can be reliably and accurately used for forecasting energy consumption of China. It also can be widely applied to most of other countries' energy consumption forecasting.

Moreover, in light of China's energy consumption forecasting results of 2011~2018, we can conclude that the current status of China's energy policies still faces challenges. The policy implication of China in this study is summarized as follows.

(1) Insisting on the implementation of environmental protection and resource conservation as well as energy saving as a fundamental national policy: to be the largest developing country and energy consumer in the world, with the largest population, complex climate, insufficient energy resources, and fragile eco-environment, China should pay more attention to the global climate change.

(2) Upgrading the traditional industries: China has cautiously scheduled to restructure and rejuvenate high-energy-consuming industries.

(3) Realizing the energy conservation by promoting energy-saving technology and developing a circular economy, especially improving energy utilization efficiency.

(4) Developing the low-carbon energy by accelerating the development of hydropower, geothermal, solar power, wind power, biomass energy, natural gas, and other clean resources: most of all, we should pay more attention to develop and utilize nonfossil energy. Meanwhile, since China energy consumption structure mainly depends on coal, the clean coal technology is still essential.

(5) Joining forces with other countries to respond to energy dilemma and climate change problems.

Conflict of Interests

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

References

[1] U. K. Rout, A. Voß, A. Singh, U. Fahl, M. Blesl, and B. P. Ó Gallachóir, "Energy and emissions forecast of China over a long-time horizon," *Energy*, vol. 36, no. 1, pp. 1–11, 2011.

[2] S. Yu, K. Zhu, and X. Zhang, "Energy demand projection of China using a path-coefficient analysis and PSO-GA approach,"

Energy Conversion and Management, vol. 53, no. 1, pp. 142–153, 2012.

[3] China statistics yearbook, 2010, http://www.stats.gov.cn/tjsj/ndsj/2010/indexch.htm.

[4] H. T. Pao, "Forecasting energy consumption in Taiwan using hybrid nonlinear models," *Energy*, vol. 34, no. 10, pp. 1438–1446, 2009.

[5] N. M. Christodoulakis, S. C. Kalyvitis, D. P. Lalas, and S. Pesmajoglou, "Forecasting energy consumption and energy related CO_2 emissions in Greece: an evaluation of the consequences of the community support framework II and natural gas penetration," *Energy Economics*, vol. 22, no. 4, pp. 395–422, 2000.

[6] Y. S. Murat and H. Ceylan, "Use of artificial neural networks for transport energy demand modeling," *Energy Policy*, vol. 34, no. 17, pp. 3165–3172, 2006.

[7] P. Crompton and Y. Wu, "Energy consumption in China: past trends and future directions," *Energy Economics*, vol. 27, no. 1, pp. 195–208, 2005.

[8] T. Chen, "A collaborative fuzzy-neural approach for long-term load forecasting in Taiwan," *Computers and Industrial Engineering*, vol. 63, no. 3, pp. 663–670, 2012.

[9] Z. W. Geem, "Transport energy demand modeling of South Korea using artificial neural network," *Energy Policy*, vol. 39, no. 8, pp. 4644–4650, 2011.

[10] M. Forouzanfar, A. Doustmohammadi, S. Hasanzadeh, and H. Shakouri G, "Transport energy demand forecast using multilevel genetic programming," *Applied Energy*, vol. 91, no. 1, pp. 496–503, 2012.

[11] U. Kumar and V. K. Jain, "Time series models (Grey-Markov, Grey Model with rolling mechanism and singular spectrum analysis) to forecast energy consumption in India," *Energy*, vol. 35, no. 4, pp. 1709–1716, 2010.

[12] I. J. Lu, C. Lewis, and S. J. Lin, "The forecast of motor vehicle, energy demand and CO_2 emission from Taiwan's road transportation sector," *Energy Policy*, vol. 37, no. 8, pp. 2952–2961, 2009.

[13] B. Hodge, A. Zeiler, D. Brooks, G. Blau, J. Pekny, and G. Reklatis, "Improved wind power forecasting with ARIMA models," *Computer Aided Chemical Engineering*, vol. 29, pp. 1789–1793, 2011.

[14] M. Kankal, A. Akpinar, M. I. Kömürcü, and T. Ş. Özşahin, "Modeling and forecasting of Turkey's energy consumption using socio-economic and demographic variables," *Applied Energy*, vol. 88, no. 5, pp. 1927–1939, 2011.

[15] Y.-S. Lee and L.-I. Tong, "Forecasting energy consumption using a grey model improved by incorporating genetic programming," *Energy Conversion and Management*, vol. 52, no. 1, pp. 147–152, 2011.

[16] L. Chenchen, G. Yan, and S. Zhuyan, "Energy demand forecast based on the new weakening buffer operators with exponential type," *Energy Procedia*, vol. 5, pp. 2522–2526, 2011.

[17] H. Pao, "Comparing linear and nonlinear forecasts for Taiwan's electricity consumption," *Energy*, vol. 31, no. 12, pp. 2129–2141, 2006.

[18] B. Dong, C. Cao, and S. E. Lee, "Applying support vector machines to predict building energy consumption in tropical region," *Energy and Buildings*, vol. 37, no. 5, pp. 545–553, 2005.

[19] R. Chedid, M. Kobrosly, and R. Ghajar, "A supply model for crude oil and natural gas in the Middle East," *Energy Policy*, vol. 35, no. 4, pp. 2096–2109, 2007.

[20] D. Shi, "The impact of changes in the industrial structure on energy consumption," *Quantitative & Technical Economics*, no. 12, pp. 50–52, 1999 (Chinese).

[21] B. Y. Yuan and Y. C. Yao, "The input-output analysis of coal resources based on the sustainable development," *Energy Technology and Management*, vol. 2, pp. 78–80, 2004 (Chinese).

[22] G. Hondroyiannis, "Estimating residential demand for electricity in Greece," *Energy Economics*, vol. 26, no. 3, pp. 319–334, 2004.

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

[24] Y. Xue, Z. Cao, and L. Xu, "The application of combination forecasting model in energy consumption system," *Energy Procedia*, vol. 5, pp. 2599–2603, 2011.

[25] G. Elliott and A. Timmermann, "Optimal forecast combinations under general loss functions and forecast error distributions," *Journal of Econometrics*, vol. 122, no. 1, pp. 47–79, 2004.

[26] Z. Guohao, G. Shufen, S. Jing, and W. Yongguang, "An investigation of coal demand in China based on the variable weight combination forecasting model," *Journal of Resources and Ecology*, vol. 2, no. 2, pp. 126–131, 2011.

[27] J. Chai, J. Guo, and H. Lu, "Forecasting energy demand of China using bayesian combination model," *China Population Resources and Environment*, vol. 18, no. 4, pp. 50–55, 2008.

[28] H. Song, R. Zhang, Y. Zhang, F. Xia, and Q. Z. Miao, "Energy consumption combination forecast of Hebei province based on the IOWA operator," *Energy Procedia*, vol. 5, pp. 2224–2229, 2011.

[29] G. Xu and W. Wang, "Forecasting China's natural gas consumption based on a combination model," *Journal of Natural Gas Chemistry*, vol. 19, no. 5, pp. 493–496, 2010.

[30] C. Lemke and B. Gabrys, "Meta-learning for time series forecasting and forecast combination," *Neurocomputing*, vol. 73, no. 10, pp. 2006–2016, 2010.

[31] S.-Q. Zhou and J.-Q. Wang, "Combined forecasting approach based on multi-criteria optimization," *Systems Engineering and Electronics*, vol. 31, no. 7, pp. 1651–1654, 2009 (Chinese).

[32] I. Sánchez, "Adaptive combination of forecasts with application to wind energy," *International Journal of Forecasting*, vol. 24, no. 4, pp. 679–693, 2008.

[33] H.-Q. Zhao, "The combination forecasting model based on the entropy approach and application in power peak load forecasting," *College Mathematics*, vol. 27, no. 3, pp. 157–160, 2011 (Chinese).

[34] A. Timmermann, "Forecast combinations," in *Handbook of Economic Forecasting*, C. Elliott, A. Granger, and A. Timmermann, Eds., chapter 4, pp. 135–196, Elsevier, Amsterdam, The Netherlands, 2006.

[35] C. M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, Oxford, Miss, USA, 1996.

[36] X. Y. Chen and W. Wu, "Remote diagnosis system of hydrogenerator sets based on wavelet-neural network," *International Journal of Power and Energy Systems*, vol. 27, no. 3, pp. 305–310, 2007.

[37] A. I. Dounis, P. Tiropanis, D. Tseles, G. Nikolaou, and G. P. Sycros, "A comparison of grey model and fuzzy predictive model for time series," *International Journal of Computational Intelligence*, vol. 2, no. 3, pp. 176–181, 2005.

[38] G. L. Li, Z. H. Cai, X. J. Kang, Z. D. Wu, and Y. Z. Wang, "ESPSA: a prediction-based algorithm for streaming time series segmentation," *Expert Systems with Applications*, vol. 41, no. 14, pp. 6098–6105, 2014.

[39] C. Sun, J. Li, H. Zheng, J. Yang, and X. Fang, "A new method of faulty insulation diagnosis in power transformer based on degree of area incidence analysis," *Power System Technology*, vol. 26, no. 7, pp. 24–29, 2002 (Chinese).