

## Research Article

# Forecasting Primary Energy Requirements of Territories by Autoregressive Integrated Moving Average and Backpropagation Neural Network Models

Ning-Kang Pan and Chunwan Lv 

*School of Mathematics and Big Data, Foshan University, Foshan 528000, China*

Correspondence should be addressed to Chunwan Lv; lvchunwan@163.com

Received 19 May 2019; Revised 10 July 2019; Accepted 14 July 2019; Published 4 August 2019

Academic Editor: Vincenzo Bianco

Copyright © 2019 Ning-Kang Pan and Chunwan Lv. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Forecasting energy data, especially the primary energy requirement, is the key part of policy-making. For those territories of different developing types, seeking a knowledge-based and dependable forecasting model is an essential prerequisite for the prosperous development of policy-making. In this paper, both autoregressive integrated moving average and backpropagation neural network models which have been proved to be very efficient in forecasting are applied to the forecasts of the primary energy consumption of three different developing types of territories. It is shown that the average relative errors between the actual data and simulated value are from 4.5% to 5.9% by the autoregressive integrated moving average and from 0.04% to 0.47% by the backpropagation neural network. Specially, this research shows that the backpropagation neural network model presents a better prediction of primary energy requirement when considering gross domestic product, population, and the particular values as predictors. Furthermore, we indicate that the single-input backpropagation neural network model can still work when the particular values have contributed most to the energy consumption.

## 1. Introduction

Since the emergence of the issue about applying mathematical models to predict energy consumption, relevant researches have been carried out [1] and several related experiments have been implemented [2]. The recent study [3] has combined the nonlinear metabolic grey model (NMGM) and autoregressive integrated moving average (ARIMA) model and has used the linear ARIMA to correct NMGM forecasting residuals, which has improved forecasting accuracy steadily and given useful policy recommendation. The time-series forecasting techniques based on the metabolic grey model, autoregressive integrated moving average model-grey model, and induced ordered weighted geometric averaging operator have been investigated in [4], which has found a way to provide reliable information and has indicated that the results from the time-series and econometric forecasting

technique are consistent. The above two studies have shown that the prediction of relevant models is effective and also indicate the research prospects and directions in this field. In addition, the single-linear, hybrid-linear, and nonlinear forecasting techniques based on grey theory have been presented in [5] to forecast energy demand in both China and India more accurately. In a novel study [6], researchers have found that the NMGM-ARIMA technique can significantly improve forecasting effectiveness and outperform other related forecasting models, which have been a guide to practical applications. In the field of primary energy consumption (PEC), many mathematical models have been successfully applied in prediction. For example, a new hybrid method (HAP) has been proposed for estimating energy demand in Turkey using particle swarm optimization (PSO) and ant colony optimization (ACO), and the research has found that different models have different advantages [7]. Two new

models based on artificial bee colony (ABC) and particle swarm optimization (PSO) techniques have been proposed in [8] to estimate electrical energy demand in Turkey. In addition, the research methods in [9, 10] also have made good predictions. Among these models, the ARIMA model [11] and backpropagation (BP) neural networks [12] are the two most commonly used predictive models.

The forecast of energy consumption, especially the PEC, has always been a significant reference point for a territory [7]. Specifically, on the one hand, it is always important for a territory to make the future energy policies, to achieve a better development and to reduce greenhouse gas emissions [13]. In the recent study [14], particle swarm optimization (PSO) and artificial bee colony (ABC) techniques have been applied in estimating CO<sub>2</sub> emission in Turkey based on socioeconomic indicators. On the other hand, the PEC refers to the overall energy consumption within a geographic territory and represents the total supply of energy available to the territory which supports all the requirements for energy transformation and final consumption in that territory. In general, the PEC of a territory includes both its indigenous energy sources and imported energy commodities consumed within the territory [15].

As for models, both economists and applied mathematicians have built a large number of models and accumulated experiences in forecasting energy [16]. For instance, the macroeconomic analysis of energy [17] and the analytic network process in energy policy planning [18] are two methods which are based on econometrics transforming the forecasting problem into an economic one and are seemingly in accordance with the thought of “big data.” However, applied mathematicians have developed serial linear or nonlinear mathematical models to complete the prediction [19].

Motivated by previous studies which show that the consumption of primary energy could be approximated by using linear or nonlinear econometrics with economic and/or noneconomic indicators, mathematical models [20] and especially the ARIMA model [21] and the BP neural network [22] are common models used for prediction. Therefore, in this paper, we will adopt the ARIMA model and the BP neural network one to make predictions and draw conclusions through making comparisons of models’ correctness.

This paper also focuses on solving some gaps in previous research and filling these gaps. Compared with the previous research concerning the single model [1], this paper adopts two common and effective mathematical models and then makes a comparison. In contrast to the research taking into account multiple models [2], this paper not only explores the accuracy of comparison but also opens up a new way to use neural networks to reveal the possible time of economic transformation.

The novelty and usefulness of this paper can be summed up as follows: (a) Compared with the available studies focusing on only one territory or one mathematical model which leads to the lack of comparison between different models and different types of territories, this study,

however, includes three cities of different developing types from Greater Bay Area to make a comparison and then obtain more reasonable conclusions. It is shown in this paper that both the ARIMA model and the BP neural network model can be applied in different developing types of territories. (b) As for the novelty, this paper concludes that the single-input BP neural network model can be used for judging whether the territory has transformed completely. (c) In terms of practicality, this paper explores that ARIMA can be applied to predict the PEC in addition to manufacturing territories, and the BP neural network model can be applied to predict the PEC in different types of territories.

## 2. Methodology

*2.1. Models and Territories.* It is shown in [23] that both the ARIMA model and the BP neural network model are suitable for time-series data and can better catch on the related data and predict future points in the series. In this paper, we consider three cities including Hong Kong, Shunde, and Zhaoqing from Greater Bay Area in China as involved representative territories. The reason why we choose these territories is that they have varying development models; that is, they are three different industry-dominated cities. Hong Kong has been a service-dominated city since the 2000s or earlier. Shunde has turned to be a manufacturing-dominated city since the beginning of the Chinese economic reform. But Zhaoqing is famous for its agriculture, so it is regarded as an agriculture-dominated one.

The ARIMA is a commonly used time-series model which achieves the object’s characteristics of self-similarity, periodicity, suddenness, and trends [23] and has a better achievement in the short-term subject’s forecast. Therefore, it has been applied in the prediction of the stock price index, the blood glucose concentrations [24], the current blockchain technology [25], the wind generation [26], and so on. The BP neural network is a commonly used time-series and nonlinear prediction model applied in prediction of short-term wind power, indoor temperature, wind speed [27], and hydraulic press machine.

In short, we deliberately consider the agricultural territory, the manufacturing territory, and the service-dominated territory combined with effective ARIMA and BP neural network models.

*2.2. Datasets.* The monthly data are adopted to be the time step of the data in this paper. Four monthly data series adopted in this paper are addressed in detail below.

The first one is the PEC which means the total energy consumed by a geographic territory. Generally, it represents the energy including the transformation parts, the final consumption ones [28], the energy produced locally, and the imported energy sources consumed locally. Therefore, to calculate the PEC, we take Hong Kong as an example; first, we need to sum up the renewable energy produced locally, net import of coal, oil products, and

electricity and then minus the net usage of energy storage needed, and finally, we need to adjust the results by the supply from stock. This type of data comes from 2000 to 2017 for the agricultural territory, comes from 2005 to 2017 for the manufacturing one, and comes from 1979 to 2017 for the service-dominated one. Data are obtained from Statistical Yearbook of Shunde published by Development Planning and Statistics Department of Shunde, Statistical Yearbook of Guangdong Province published by Statistics Department of Guangdong, and Hong Kong Energy Statistics published by Census and Statistics Department of Hong Kong [28].

The second one and the third one are population and GDP. Population is used as a proxy for the demands of humans, and GDP can be regarded as a symbol of the growth situation of the economy. These two types of data are selected from 2000 to 2017 for the agricultural territory, from 2005 to 2017 for the manufacturing one, and from 1979 to 2017 for the service-dominated one. It should be noticed that the way we obtain these data is the same as the one we obtain PEC [29].

Last but not least, additional three different forecasting parameters are employed in this paper, which are the value added from primary products, the manufacturing industry-added value, and the values of total services. These values are of great importance for cities of different developing types. The value added from primary products indicates the growth of products directly from the natural sector (including plantation, forestry, animal husbandry, and fisheries) in this liquidation cycle compared to the previous liquidation cycle. It is shown in [29] how the territory gains from the agricultural products. In addition, the manufacturing industry-added value represents the final result of the production activities of the secondary industry in the form of money during the reporting period, which is the total result of all production activities of the production unit minus the value of the physical products and services consumed or transferred in the production process. Finally, the values of total services include the exports of services and imports of services, where the exports of services refer to the services that Hong Kong have sold to other entities and the imports of services mean the services purchased from the other entities. The data are obtained from Hong Kong Trade in Services Statistics published by Census and Statistics Department of Hong Kong [30].

**2.3. ARIMA Modeling.** Developed by Box and Jenkins, the ARIMA model is a linear regression model and is widely applied in forecasting when it is with time-series data [31]. ARIMA( $p, d, q$ ), an autoregressive integrated moving average model where parameters  $p$ ,  $d$ , and  $q$  are non-negative integers, consists of three parts including the autoregressive (AR) model, differencing (I) model, and moving average (MA) model. Specially,  $p$  represents the order of the AR model,  $d$  means the degree of I, and  $q$  is the order of the MA, respectively. To achieve the prediction through ARIMA( $p, d, q$ ), these three steps must be completed, that is, to smooth the sequence through  $d$ -order difference calculation, to obtain a more suitable  $p$  and  $d$  by

calculating and comparing the autocorrelation coefficient with the partial autocorrelation one, and to carry out the prediction by the selected model. Mathematically, the expression of ARIMA( $p, d, q$ ) is as follows [31]:

$$X_i = \phi_1 X_{i-1} + \phi_2 X_{i-2} + \cdots + \phi_p X_{i-p} + \varepsilon_i + \theta_1 \varepsilon_{i-1} + \theta_2 \varepsilon_{i-2} + \cdots + \theta_q \varepsilon_{i-q}, \quad (1)$$

where  $X_i$  and  $\varepsilon_i$  are the actual value and random error at every year and  $\phi_i$  ( $i = 1, 2, \dots, p$ ) and  $\theta_i$  ( $i = 1, 2, \dots, q$ ) are two parameters of the ARIMA model. To be clear,  $p$  and  $q$  are integers and the orders of the ARIMA model. By extracting the common factor, a new expression can be obtained as follows:

$$\begin{aligned} & (1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p) X_i \\ & = (1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q) \varepsilon_i. \end{aligned} \quad (2)$$

It can be expressed as

$$\phi(L) X_i = \Theta(L) \varepsilon_i, \quad (3)$$

by setting

$$\begin{aligned} \phi(L) &= 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p, \\ \Theta(L) &= 1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q, \end{aligned} \quad (4)$$

where  $\phi(L)$  and  $\Theta(L)$  are the  $p$ -order and  $q$ -order characteristic polynomials of the independent variable “ $L$ .”

To conclude, in ARIMA( $p, d, q$ ),  $d$  is the integral sum order,  $p$  is the autoregressive coefficient, and  $q$  is the moving average coefficient [32].

**2.4. BP Neural Network Modeling.** Suppose that the network has  $R$  nodes [33] and the transfer function of each layer is of sigmoid type [34]. The following notations are used throughout this section:  $\alpha_{1i}$  denotes the output of the  $i$ -th hidden-layer node;  $\alpha_{2k}$  denotes the output of the  $k$ -th output-layer node;  $\omega_{1ij}$  and  $\omega_{2ki}$  represent the weight between node  $i$  and node  $j$  and the weight between node  $k$  and node  $i$ , respectively;  $b_{1i}$  denotes the threshold for hidden-layer node  $i$ ; and  $b_{2k}$  denotes the threshold for output-layer node  $k$  [35].

The BP neural network can be understood in detail below [36].

Firstly, the sample forward propagation is input, which is a way of calculating the sample from the input layer, passing through the hidden layer, and finally to the output layer. The input samples should be known, and then the corresponding output is obtained by the above relevant steps.

In the hidden layer, the output of the  $i$ -th neuron is

$$\alpha_{1i} = f_1 \left( \sum_{j=1}^n \omega_{1ij} p_j + b_{1i} \right), \quad i = 1, 2, \dots, s_1, \quad (5)$$

where  $p_j$  is the input value,  $s_1$  is an integer denoting the number of neurons in the hidden layer, and  $f_1$  is the activation function defined as follows:

$$f_1(p_j) = \frac{1}{1 + e^{p_j}} \quad (6)$$

In the output layer, the output of the  $k$ -th neuron is

$$\alpha_{2k} = f_2 \left( \sum_{i=1}^{s_1} \omega_{2ki} \alpha_{1i} + b_{2k} \right), \quad k = 1, 2, \dots, s_2, \quad (7)$$

where  $s_2$  is an integer denoting the number of neurons in the output layer and  $f_2$  is defined as follows:

$$f_2(\alpha_{1i}) = \frac{1}{1 + e^{\alpha_{1i}}} \quad (8)$$

Secondly, the error backpropagation is output.

Through the above forward propagation calculation, the actual output can be obtained. However, in general, the actual expectations are different. When two values are different or the errors of two values exceed some specified values, corresponding learning corrections should be made for the network.

In particular, the error function  $E$  is defined as follows:

$$E = \frac{1}{2} \sum_{k=1}^{s_2} (t_k - \alpha_{2k})^2, \quad (9)$$

where  $t_k$  is the predicted value.

In the output layer, the weights  $\Delta\omega_{2ki}$  from the  $i$ -th input to the  $k$ -th output are

$$\Delta\omega_{2ki} = -\eta \frac{\partial E}{\partial \omega_{2ki}} = \eta \cdot \delta_{ki} \cdot p_j, \quad 0 < \eta < 1, \quad (10)$$

where  $p_j$  is the input value and  $\delta_{kj}$  can be described as

$$\begin{aligned} \delta_{kj} &= (p_j - \alpha_{2k}) \cdot f_1(1 - f_1), \\ e_k &= \sum_{j=1}^{s_2} \delta_{kj} \omega_{2ki} = t_k - \alpha_{2k}, \end{aligned} \quad (11)$$

where  $e_k$  is the range fault.

After implicating the layer weight correction, the weight ( $\Delta\omega_{ij}$ ) from the  $j$ -th input to the  $i$ -th output can be described as follows:

$$\begin{aligned} \Delta\omega_{ij} &= -\eta \frac{\partial E}{\partial \omega_{lij}} = \eta \cdot \delta_{ij} \cdot p_j, \quad 0 < \eta < 1, \\ \delta_{ij} &= e_i f_1', \\ e_i &= \sum_{k=1}^{s_2} \delta_{kj} \omega_{2ki}, \end{aligned} \quad (12)$$

where  $\eta$  is the learning coefficient.

The error of the output is reversed from the output layer, through the intermediate layer and finally to the input layer, where each layer is corrected once. The threshold can be seen as one of the weights and also contributes to the adjustment of the weights.

Thirdly, the memory training is made to be cyclic. In order to improve the accuracy of the network and reduce the appearance of the output error, it is necessary to carry on the loop memory training for all the samples input to the network. In addition, the number of loops should not be too few so that this sample mode can be effectively remembered by the network.

Last but not least, the end of the study is checked. The output error is checked to see if it meets the required standard whenever the loop memory training is completed. If it meets the requirements, the process is ended. Otherwise, the loop training is made again until it meets the requirements [37].

### 3. Results and Discussion

**3.1. ARIMA Modeling.** Following the above methodology, the PEC data of three different developing territories are used as the training dataset in the ARIMA model. In addition, unit root tests are stochastic trends in time series which are executed for the stationarity in time series. The stationarity in time series is that a shift in time does not cause a change in the shape of the distribution [38]. To obtain a smooth sequence and “ $p$ ” parameter of each model, we conduct three different unit root tests from zero orders to three orders for each model and make a comparison of three types of results. The unit root tests and results are shown in Tables 1–6.

From Tables 1–6, the augmented Dickey–Fuller (ADF) statistic is a statistical method to check whether a time series is stationary or not [37]. And the test critical value means threshold values, comparing the ADF statistic with the following three values at three significant levels. It is observed from the data from Tables 1–6 that the second-order difference is suitable for both the agricultural territory’s model and the model of the manufacturing territory and the zero-order difference is suitable for the one of the service-dominated territory. Therefore, we obtain the correlogram which shows the autocorrelation and the partial correlation with the help of EVIEWS, where the autocorrelation is the correlation of a signal with its delayed copy and the partial correlation means the degree of association between two random variables [39]. Both the autocorrelation and the partial correlation are shown in Figures 1–3. In addition, the AC and PAC in Figures 1–3 represent the autocorrelation coefficient of the sequence and the partial correlation coefficient of the sequence, respectively. Q-Stat represents the output of the statistical test about whether any of a group of autocorrelations of a time series are different from zero which obeys the chi-square distribution, and Prob represents its possibilities [40]. Representatively, it is presented in Figure 3 that the smoothness of the correlation data is good without difference. Furthermore, the autocorrelation function image is decreased, and the partial autocorrelation function image is truncated after the first image. Therefore, we can determine the specific values of the coefficients  $p$ ,  $d$ , and  $q$  based on these facts.

TABLE 1: Unit root tests of the agricultural city.

Sequence	ADF statistic	Test critical value		
		1%	5%	10%
Zero order	-2.275	-3.616	-2.941	-2.609
First order	-4.974	-3.621	-2.943	-2.610
Second order	-5.892	-3.654	-2.957	-2.617

TABLE 2: Unit root tests of the manufacturing city.

Sequence	ADF statistic	Test critical value		
		1%	5%	10%
Zero order	-1.524	-4.122	-3.145	-2.714
First order	-4.067	-4.200	-3.175	-2.729
Second order	-5.275	-4.297	-3.213	-2.748

TABLE 3: Unit root tests of the service-dominated city.

Sequence	ADF statistic	Test critical value		
		1%	5%	10%
Zero order	0.130	-4.004	-3.099	-2.690
First order	-3.905	-4.004	-3.099	-2.690
Second order	-3.373	-4.058	-3.120	-2.701

TABLE 4: Probability of each sequence (the agricultural city).

Sequence	Probability
Zero order	0.956
First order	0.012
Second order	0.032

TABLE 5: Probability of each sequence (the manufacturing city).

Sequence	Probability
Zero order	0.488
First order	0.012
Second order	0.003

TABLE 6: Probability of each sequence (the service-dominated city).

Sequence	Probability
Zero order	0.185
First order	0.000
Second order	0.000

The specific determination method can be shown as follows.

The determination of the coefficients  $p$ ,  $d$ , and  $q$  requires experience, but the theoretical support is still very meaningful. According to some available literatures like [41], the determination of the values of  $p$ ,  $d$ , and  $q$  can be summed up as follows: (a) Since the sequence requires a first-order difference to achieve smoothness, we determine that the

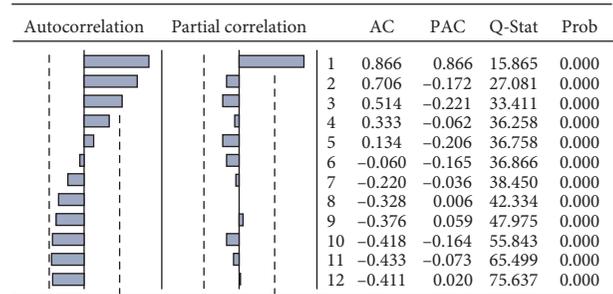


FIGURE 1: Correlogram (primary industry dominated).

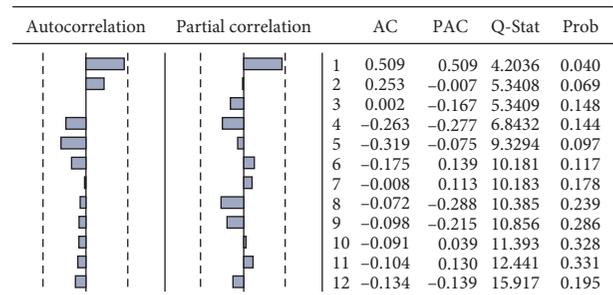


FIGURE 2: Correlogram (manufacturing dominated).

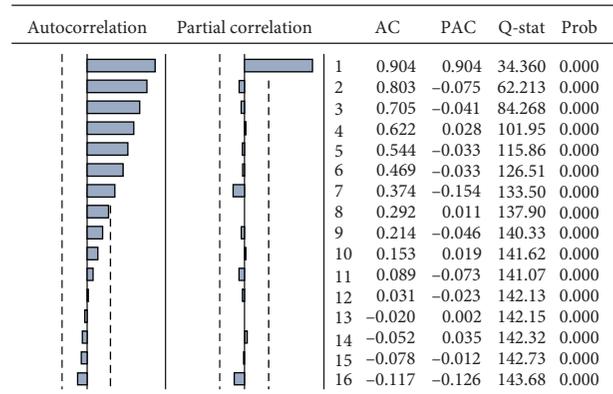


FIGURE 3: Correlogram (service dominated).

value of  $d$  is 1. (b) Next, based on the fact that the autocorrelation images are truncated and the partial correlation images are smeared, we conclude that the sequence is suitable for the AR model. Among them, since none of the partial correlation images has been outside the confidence interval, we determine that the value of  $p$  is 0. (c) We also determine that the value of  $q$  is 1 because the autocorrelation images are truncated. (d) Finally, we find out that the best model is obtained by taking a combination of several values of  $p$  and  $q$  and simulating the prediction. To conclude, as it is shown in the correlogram, ARIMA(0,1,1) can be used for the prediction of the primary energy consumption of the service-dominated territory when both AC and PAC are considered. Similarly, it is seen that the applicability of ARIMA(2, 0, 2) and ARIMA(2, 0, 5) for the agricultural territory and the manufacturing territory can be guaranteed,

TABLE 7: Summaries of the three ARIMA models.

The number of predictions	Stationary $R^2$	$R^2$	The number of outliers
$ARIMA(2, 0, 2)$			
1	0.941	0.941	0
$ARIMA(2, 0, 5)$			
1	0.742	0.721	0
$ARIMA(0, 1, 1)$			
1	0.88	0.955	0

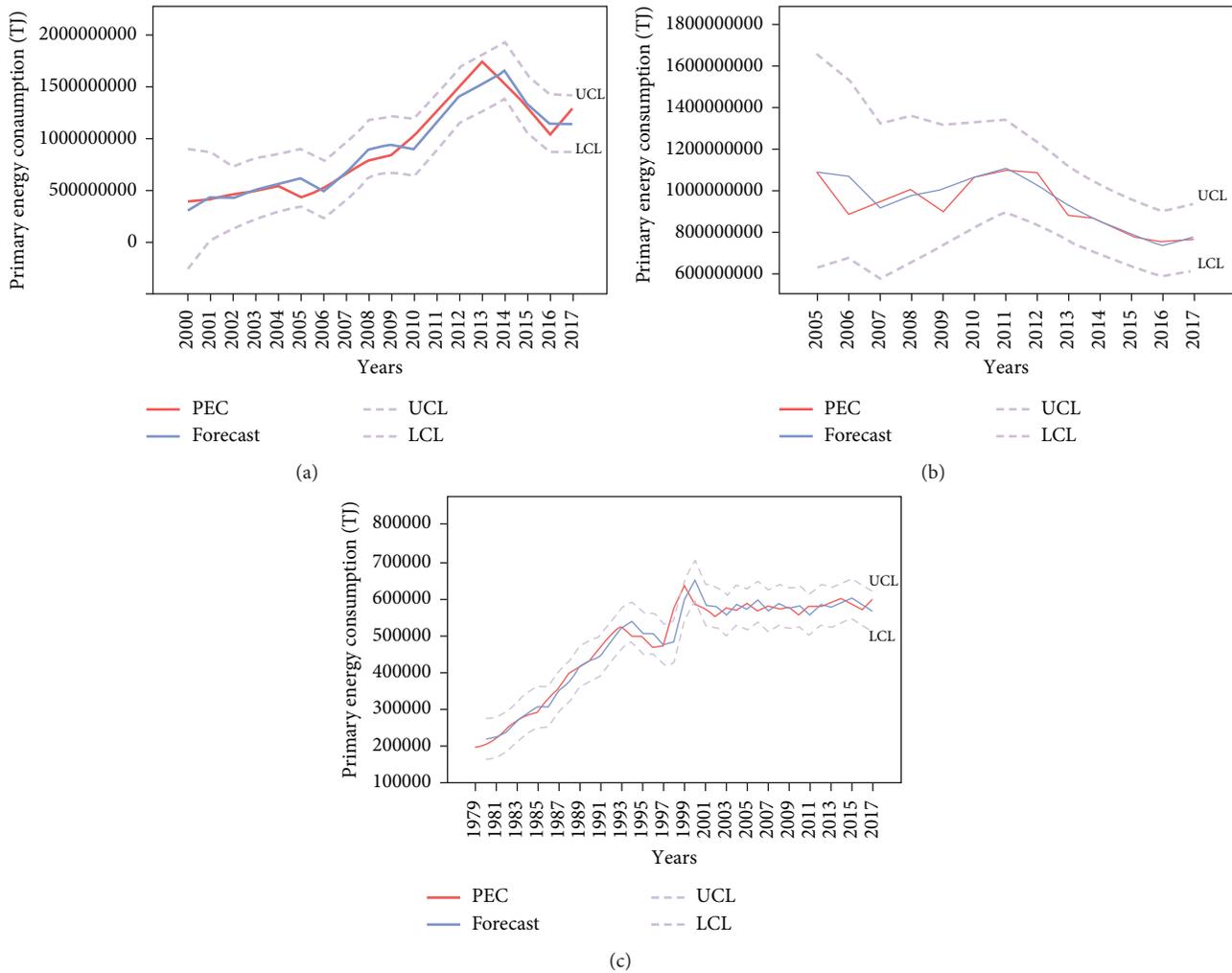


FIGURE 4: Forecasts by ARIMA. (a) ARIMA for the agricultural territory. (b) ARIMA for the manufacturing territory. (c) ARIMA for the service-dominated territory.

respectively. Then, we construct three models, and the summaries of these three models are shown in Table 7. The stationary  $R^2$  in Table 7 represents the measurement which compares the stationary part of the model with a simple mean model. A positive stationary  $R^2$  means the model under consideration is better than the baseline one. In addition, the  $R^2$  in Table 7 means the goodness of fit. The closer the value of  $R^2$  to 1, the better the fit of the regression line to the observations [42].

It is noticed that the  $R^2$  of  $ARIMA(2, 0, 2)$  is 0.941 and the corresponding one of  $ARIMA(0, 1, 1)$  is 0.955, which means that both of them are greater than 0.80. So the models are considered as the fitting ones for the agricultural and service-dominated territories [37]. However, ARIMA is not suitable for the manufacturing territory because the  $R^2$  of  $ARIMA(2, 0, 5)$  is 0.721 which is less than 0.80. Finally, the results of predicting numbers of primary energy consumption of three territories are

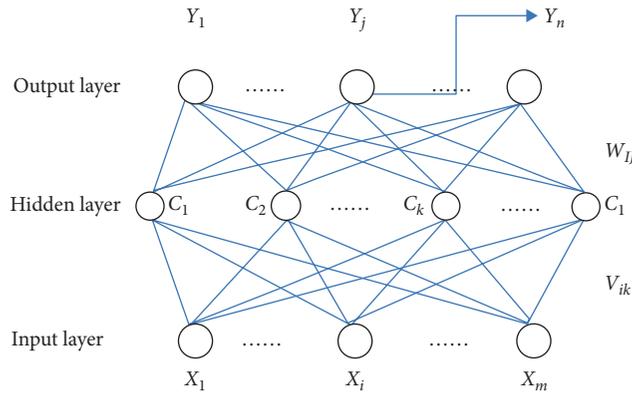


FIGURE 5: Schematic diagram of the BP neural network model.

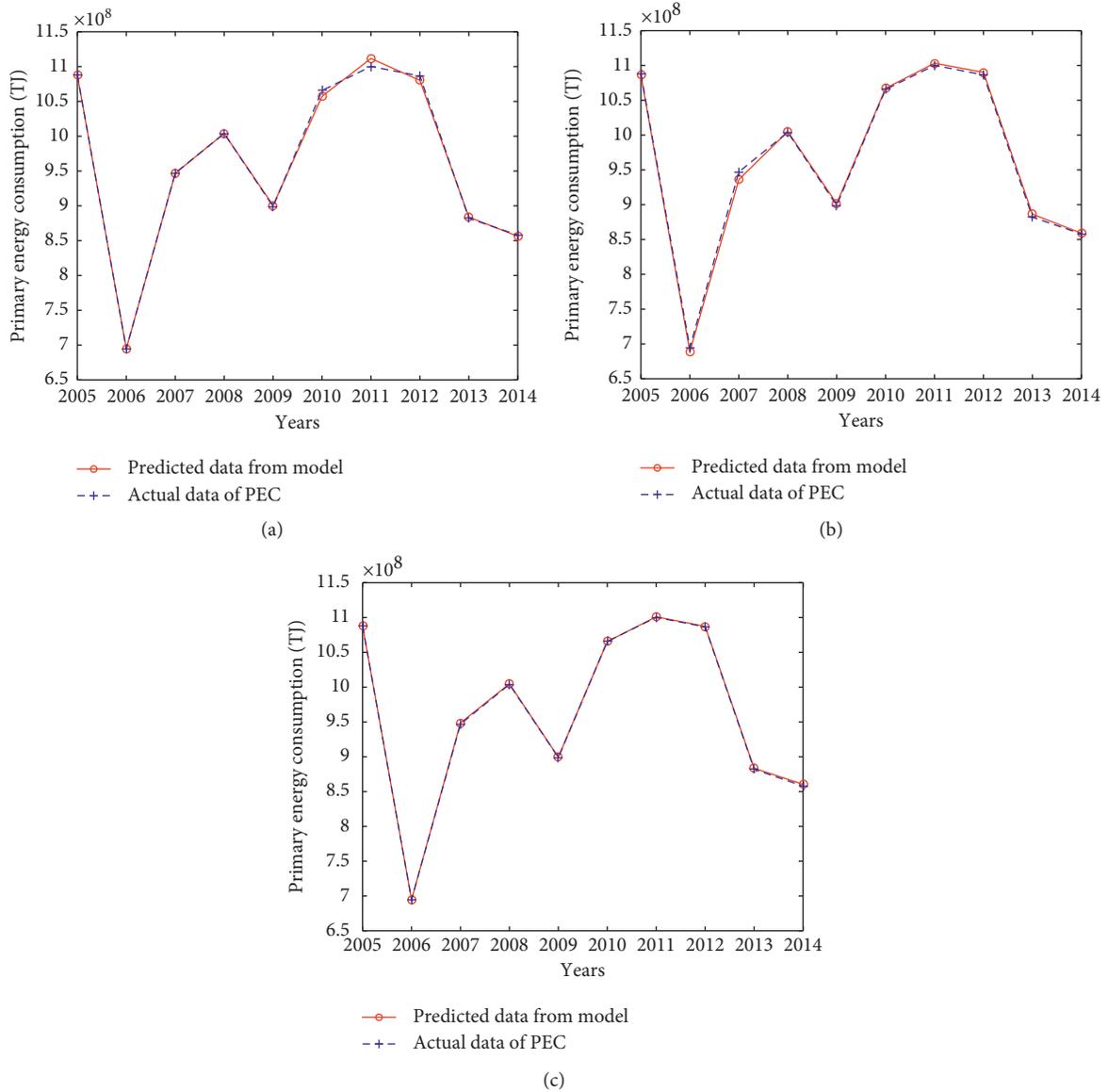


FIGURE 6: Single-input models (the manufacturing-dominated territory) for comparison of predicted data and actual data. (a) GDP. (b) Population. (c) Manufacturing.

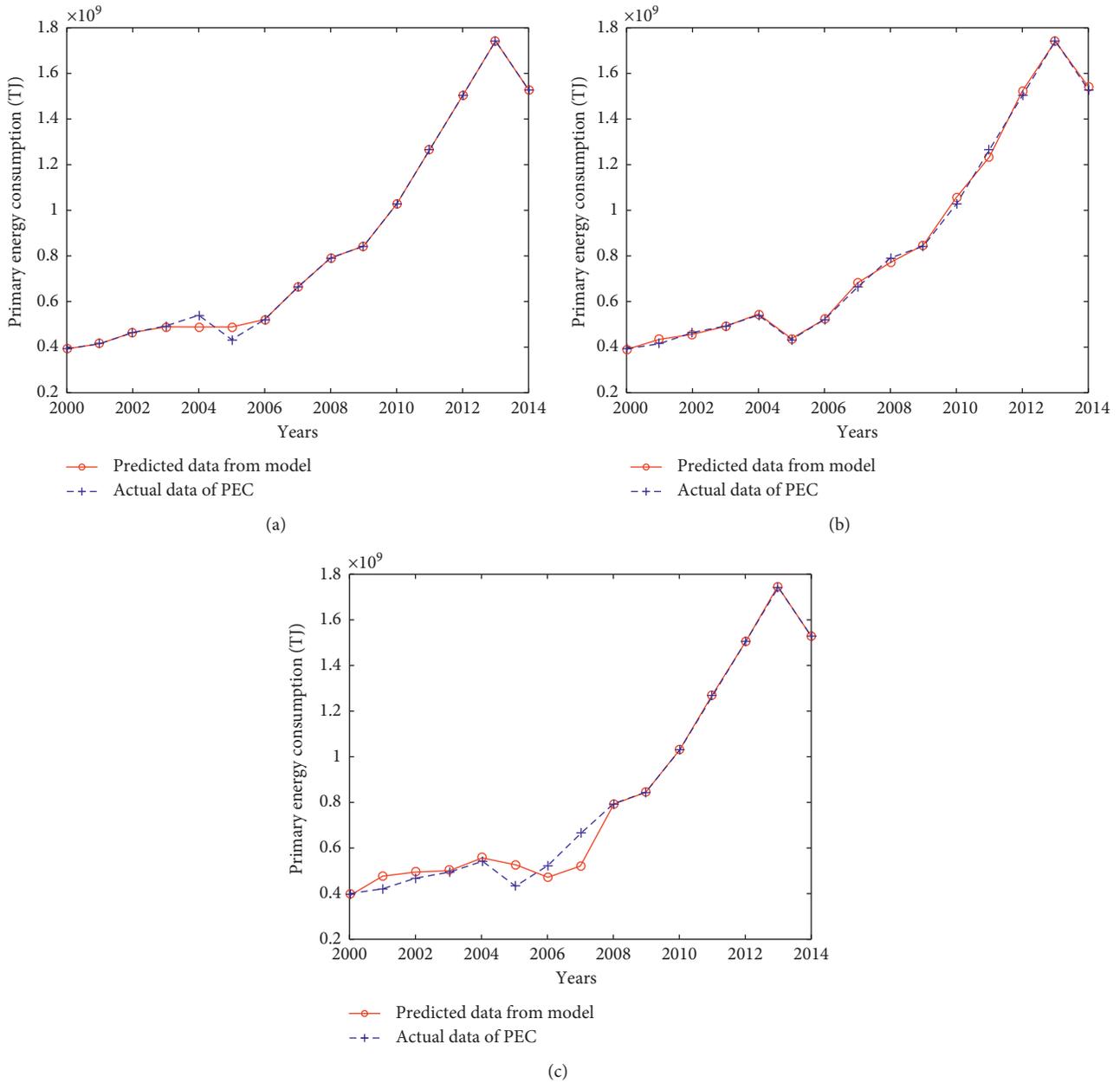


FIGURE 7: Single-input models (the agricultural territory) for comparison of predicted data and actual data. (a) GDP. (b) Population. (c) Agriculture.

shown in Figure 4, where UCL and LCL denote the upper control line and the lower control line, respectively. Their mathematical definitions are as follows:

$$\begin{aligned} \text{UCL} &= \mu + 3\sigma, \\ \text{LCL} &= \mu - 3\sigma, \end{aligned} \quad (13)$$

where  $\mu$  is the average value of the output and  $\sigma$  is the standard deviation of the output.

**3.2. BP Neural Network Modeling.** A three-layer back-propagation neural network with three hidden layers is developed for predicting three territories' primary energy consumption. Based on Kolmogorov's theorem [43], for any continuous function, it can be implemented with a three-layer network, where the input layer has  $m$  units, the hidden layer has  $2m + 1$  nodes, and the output layer has  $n$  units. So the number of hidden nodes is set to be three. And the following results show that our models are feasible under this condition. Two

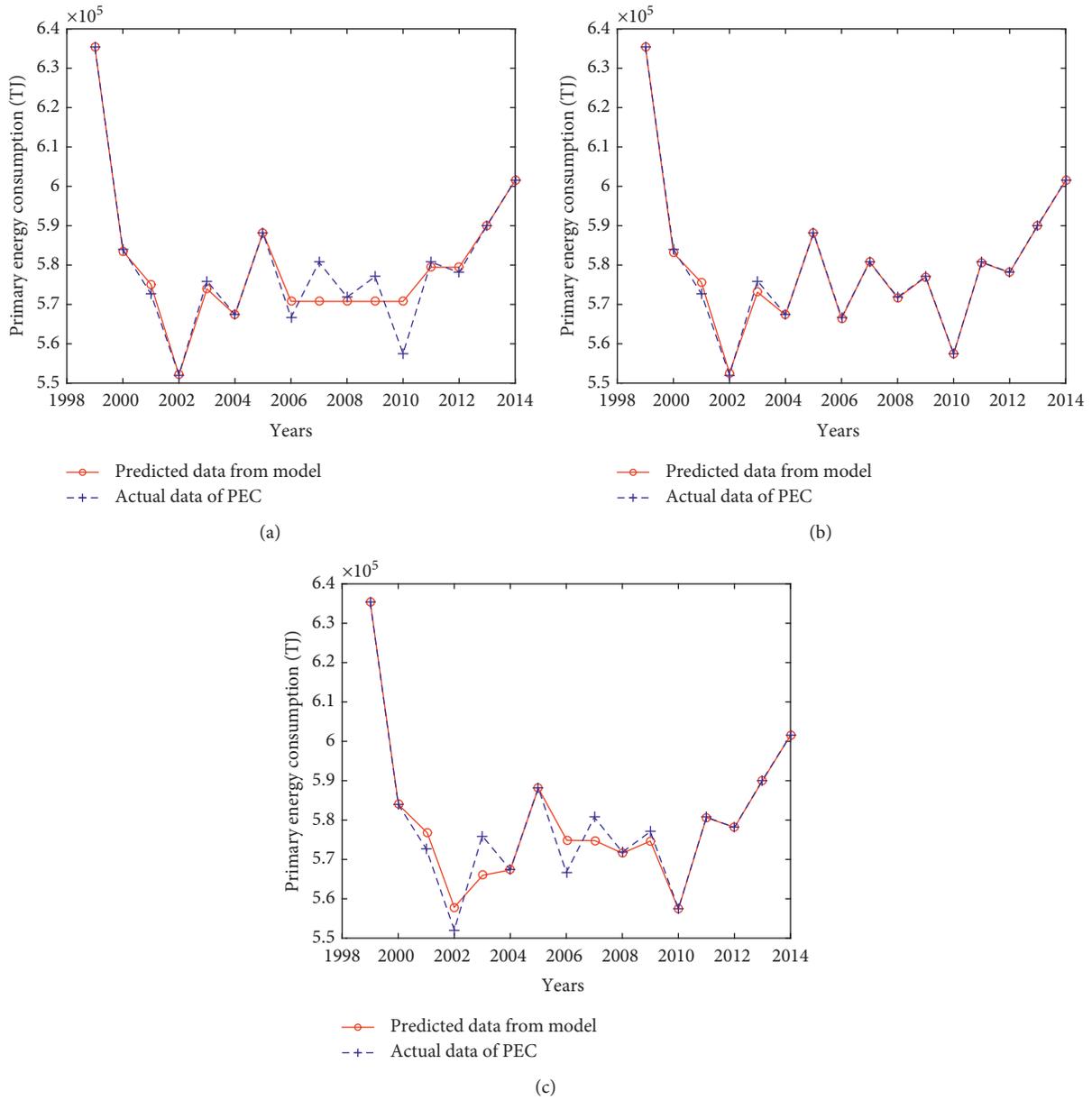


FIGURE 8: Single-input models (the service-dominated territory) for comparison of predicted data and actual data. (a) GDP. (b) Population. (c) Services.

activation functions, i.e., the pure linear function and the logsigmoid transfer function, are considered in the hidden layers. With the help of MATLAB software, we establish two types of models including three inputs and single input. The single-input model means only one among four predictors (population/primary GDPs/manufacturing industry-added value/values of total services) is adopted. The single-input model can be used to observe whether one variable has a significant impact on the PEC and can also be used to determine whether the prediction can be completed with only one variable. In contrast, the three-input model includes three predictors in the model. The schematic diagram is presented in Figure 5.

After debugging repetitions, we set the number of training iterations to be 50000. It should be noticed that

the number of hidden-layer nodes can influence the network structure to some extent. On the one hand, if the number of hidden-layer nodes is too large, an over-matching phenomenon will occur. On the other hand, if the number of hidden-layer nodes is too small, the useful information obtained by the network from the original input may be scanty, which is not enough to discover the characteristics of the data, and the generalized nonlinear learning ability of the model may be weaker. Therefore, according to Kolmogorov's theorem, for any continuous function, it can be implemented with a three-layer network, where the input layer has  $m$  units, the hidden layer has  $2m + 1$  nodes, and the output layer has  $n$  units [44]. So the number of hidden nodes is set to be 3. As mentioned

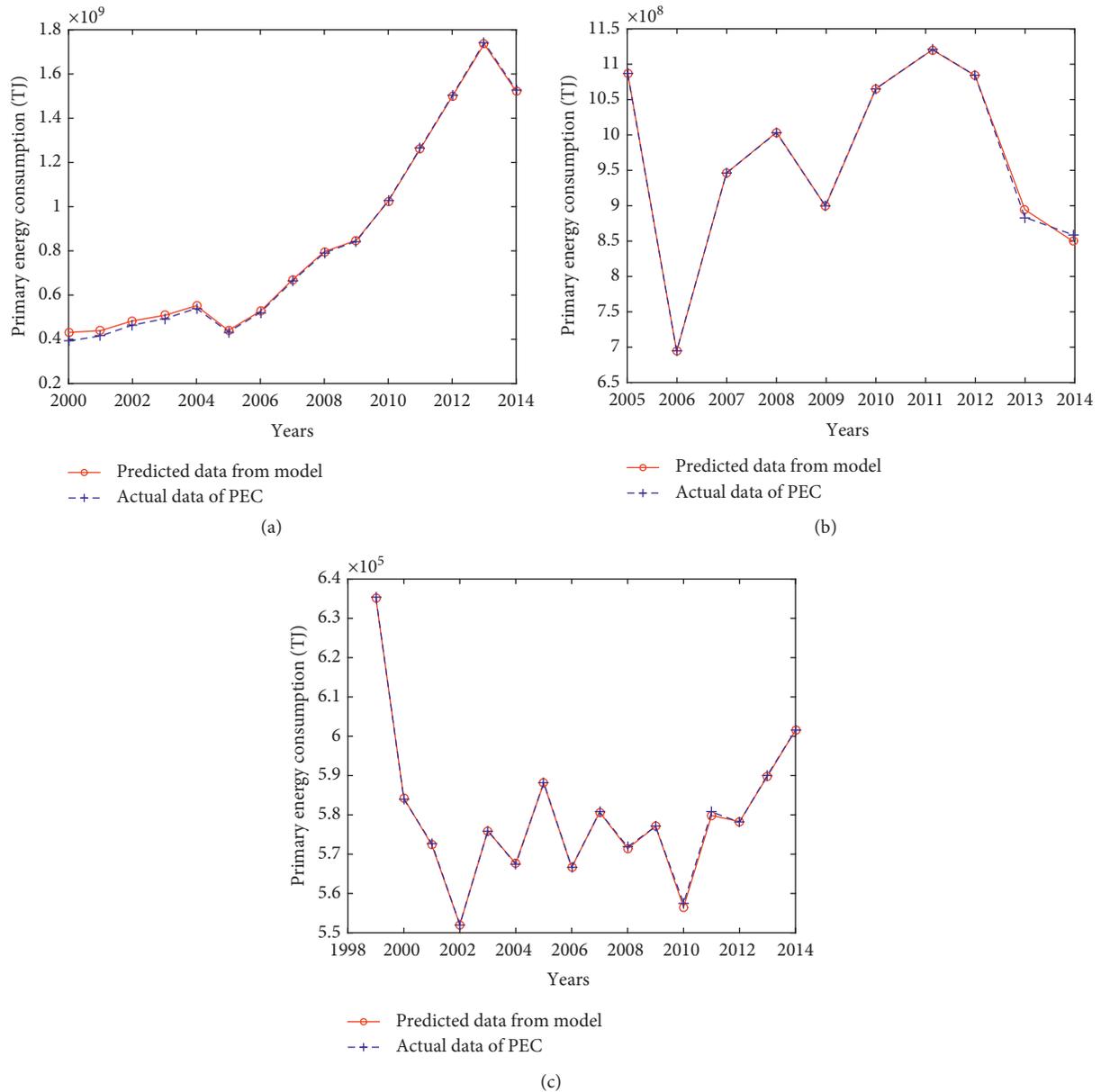


FIGURE 9: Three-input models for comparison of predicted data and actual data. (a) Agricultural city. (b) Manufacturing city. (c) Service-dominated city.

above, we selected the relevant data from 1999 to 2014 as the training sample for the service-dominated territory and for other territories similarly. Different models' predictions are presented in Figures 6–9.

We can see that the predicting data of the three-input BP neural network model are more close to the actual data compared to the three other single-input models. Therefore, we consider the three-input model as the best BP neural network for prediction. Furthermore, we think some values have a significant impact on the PEC of both the agricultural and manufacturing territories, which means that the prediction of the PEC can be conducted with only one value. However, we see from Figure 8 that

the single-input model is not suitable for the service-dominated territory. The regression of the three-input models is presented in Figure 10. As we can see in Figure 10, the regression line fits the observations very well, and the values of  $R^2$  are very close to 1, which indicates that the three models for three different developing types of cities perform well.

**3.3. Comparison of Models.** We examine two forecasting models in three different developing types of cities in this paper. From Figures 4, 9, and 10, we draw an elementary conclusion that the forecasting capability of BP neural

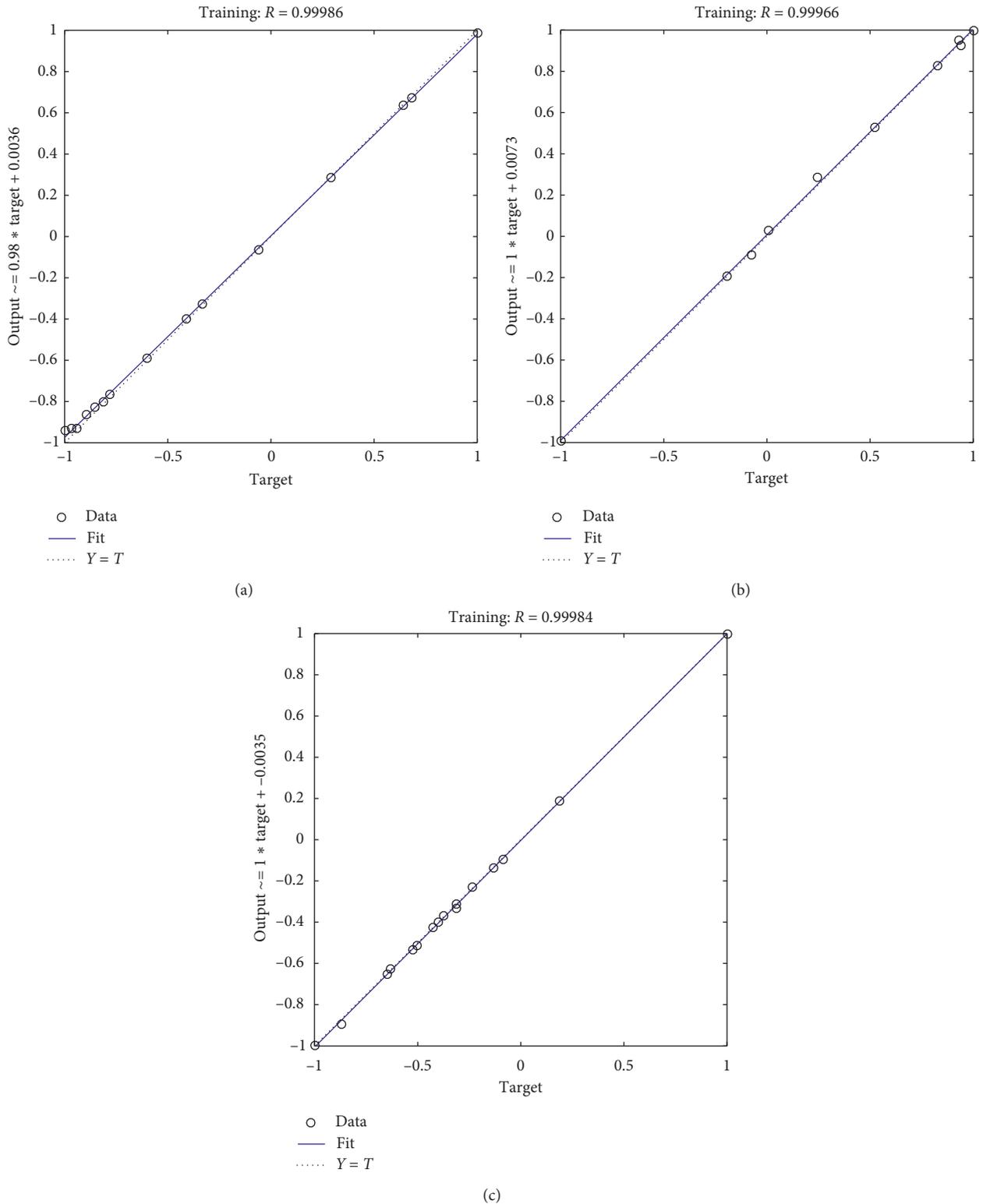


FIGURE 10: Regressions. (a) Agricultural city. (b) Manufacturing city. (c) Service-dominated city.

network models is better than that of ARIMA models on the whole.

In order to verify accurately the correctness, we make a comparison of these two types of models

through calculating the mean absolute percentage error (MAPE) and the root mean square error (RMSE). The mathematical definition of the MAPE and RMSE is as follows:

TABLE 8: Results of the correctness.

Agricultural city	Three-input BP neural network	ARIMA (2, 0, 2)
MAPE	0.004	0.105
RMSE	0.146	0.233
Manufacturing city	Three-input BP neural network	ARIMA (2, 0, 5)
MAPE	0.016	0.042
RMSE	0.155	0.363
Service-dominated city	Three-input BP neural network	ARIMA (0, 1, 1)
MAPE	0.047	0.039
RMSE	0.161	0.281

$$\text{MAPE} = \frac{\sum_{t=1}^n |(A_t - F_t)/A_t|}{n}, \quad (14)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^n (A_t - F_t)^2}{n}},$$

where  $A_t$  is the accurate value of the data,  $F_t$  is the predictive value of the data, and  $n$  is the number of data we employed. The result of the correctness is shown in Table 8.

#### 4. Conclusion

In this paper, the autoregressive integrated moving average (ARIMA) model and the backpropagation (BP) neural network model are applied to forecast the primary energy requirement of three different developing types of territories, and a comparison of the accuracy is made.

Three different models including ARIMA(2, 0, 2), ARIMA(2, 0, 5), and ARIMA(0, 1, 1) are conducted to make predictions. Their accuracies are within the range from 0.721 to 0.941. We believe that the ARIMA model is not suitable for the manufacturing territories since the  $R^2$  of ARIMA(2, 0, 5) is 0.721 which is less than 0.80. Furthermore, we conclude that ARIMA can predict accurately without employing too much data.

As for the BP neural network model, the data concerning primary products, manufacturing industry-added value, and values of total services are employed as the predictor in the model, and the most accurate prediction is made. In addition, we also use only one predictor to build three single-input models, and the corresponding result shows that not only the accuracy of the single-input models is acceptable but also the single-input models can demonstrate whether the territory has been totally supported by certain industry. This conclusion is supported by economics. In the economic literature [45], experts have pointed out that Hong Kong has completely transformed into a service-dominated territory since 2009. And in the literature [46], Shunde was regarded as a manufacturing-dominated territory in 2005. The above results support the conclusions of this paper. As for limitations, we believe that if researchers aim to obtain more accurate predictions, they need to collect multiple types of and a large amount of data, preferably monthly data, which will make the research work cumbersome. As for policy-

makers, this paper presents a judgment to determine whether the territory is fully transformed. When the territory has completely transformed into a specific development model, the corresponding energy policy should be more inclined to the industry.

In the process of this research, we also find that some new technologies can be well applied in predicting energy. In particular, the following papers deserve further research in the future: the literature [8], in which swarm intelligence approaches have achieved excellent predictions and comparable advantages over artificial neural networks (ANNs), and the literature [14], in which swarm intelligence approaches have also been applied in predicting greenhouse gas emissions, which is very important for sustainable development.

#### Data Availability

Previously reported data (four monthly data series) were used to support this study, as addressed in Section 2.2. The prior studies (and datasets) are cited at relevant places as references [18, 28, 30].

#### Conflicts of Interest

The authors declare that they have no conflicts of interest.

#### Acknowledgments

This work was supported by the Natural Science Foundation of China (Grant no. 11801077) and the Natural Science Foundation of Guangdong Province (Grant nos. 2018A1660005 and 2018A030310357).

#### References

- [1] C. R. Rivero, V. Sauchelli, H. D. Patio, J. A. Pucheta, and S. Laboret, "Long-term power consumption demand prediction: a comparison of energy associated and bayesian modeling approach," in *Proceedings of the 2015 Latin America Congress on Computational Intelligence (LA-CCI)*, pp. 1–6, Curitiba, Brazil, October 2015.
- [2] Y. He, J. Y. Wu, Y. Ge, D. Z. Li, and H. G. Yan, "Research on prediction of energy demand in china based on grey theory and system dynamics," in *Proceedings of the 2017 IEEE Conference on Energy Internet and Energy System Integration (EI2)*, pp. 1–6, Beijing, China, November 2017.
- [3] Q. Wang, S. Y. Li, and R. R. Li, "China's dependency on foreign oil will exceed 80% by 2030: developing a novel NMGM-ARIMA to forecast China's foreign oil dependence from two dimensions," *Energy*, vol. 163, pp. 151–167, 2018.
- [4] Q. Wang, S. Li, and R. Li, "Will Trump's coal revival plan work?—comparison of results based on the optimal combined forecasting technique and an extended IPAT forecasting technique," *Energy*, vol. 169, pp. 762–775, 2019.
- [5] Q. Wang, S. Li, and R. Li, "Forecasting energy demand in China and India: using single-linear, hybrid-linear, and non-linear time series forecast techniques," *Energy*, vol. 161, pp. 821–831, 2018.
- [6] Q. Wang, X. Song, and R. Li, "A novel hybridization of nonlinear grey model and linear ARIMA residual correction

- for forecasting U.S. shale oil production,” *Energy*, vol. 165, pp. 1320–1331, 2018.
- [7] M. S. Kiran, E. Özceylan, M. Gündüz, and T. Paksoy, “A novel hybrid approach based on particle swarm optimization and ant colony algorithm to forecast energy demand of Turkey,” *Energy Conversion and Management*, vol. 53, no. 1, pp. 75–83, 2012.
- [8] M. S. Kiran, E. Özceylan, M. Gündüz, and T. Paksoy, “Swarm intelligence approaches to estimate electricity energy demand in Turkey,” *Knowledge-Based Systems*, vol. 36, pp. 93–103, 2012.
- [9] X. Wang, D. Luo, X. Zhao, and Z. Sun, “Estimates of energy consumption in China using a self-adaptive multi-verse optimizer-based support vector machine with rolling cross-validation,” *Energy*, vol. 152, pp. 539–548, 2018.
- [10] D. V. D. Meer, G. R. C. Mouli, G. M. E. Mouli, L. R. Elizondo, and P. Bauer, “Energy management system with PV power forecast to optimally charge EVs at the workplace,” *IEEE Transactions on Industrial Informatics*, vol. 14, no. 1, pp. 311–320, 2018.
- [11] S. L. Lai, M. Liu, K. C. Kuo, and R. Chang, “Energy consumption forecasting in Hong Kong using ARIMA and artificial neural networks models,” *Applied Mechanics and Materials*, vol. 672-674, pp. 2085–2097, 2014.
- [12] C. Wan, J. Zhao, Y. Song, Z. Xu, J. Lin, and Z. Hu, “Photovoltaic and solar power forecasting for smart grid energy management,” *CSEE Journal of Power and Energy Systems*, vol. 1, no. 4, pp. 38–46, 2015.
- [13] T. Ma, P. A. Østergaard, H. Lund, H. Yang, and L. Lu, “An energy system model for Hong Kong in 2020,” *Energy*, vol. 68, pp. 301–310, 2014.
- [14] E. Özceylan, “Forecasting CO<sub>2</sub> emission of Turkey: swarm intelligence approaches,” *International Journal of Global Warming*, vol. 9, no. 3, pp. 337–361, 2016.
- [15] Hong Kong Energy Statistics, *Census and Statistics Department Hong Kong Special Administrative Region*, Hong Kong Energy Statistics, Hong Kong, China, 2017.
- [16] F. Jiang, X. Yang, and S. Li, “Comparison of forecasting India’s energy demand using an MGM, ARIMA model, MGM-ARIMA model, and BP neural network model,” *Sustainability*, vol. 10, no. 2, p. 2225, 2018.
- [17] J. Xia and W. Wu, “Application of the gray neural network combination model in forecasting energy demand of China,” *Computer Science and Application*, vol. 8, no. 2, pp. 181–189, 2018.
- [18] R. P. Hämäläinen and T. O. Seppäläinen, “The analytic network process in energy policy planning,” *Socio-Economic Planning Sciences*, vol. 20, no. 6, pp. 399–405, 1986.
- [19] X. F. Song, G. Liang, C. Z. Li, and W. W. Chen, “Electricity consumption prediction for Xinjiang electric energy replacement,” *Mathematical Problems in Engineering*, vol. 2019, Article ID 3262591, 11 pages, 2019.
- [20] H. R. Cui, R. R. Wu, and T. Zhao, “Dynamic decomposition analysis and forecasting of energy consumption in shanxi Province based on VAR and GM (1, 1) models,” *Mathematical Problems in Engineering*, vol. 2018, Article ID 4234206, 11 pages, 2018.
- [21] S. A. Javed and S. Liu, “Predicting the research output/growth of selected countries: application of even GM (1, 1) and NDGM models,” *Scientometrics*, vol. 115, no. 1, pp. 395–413, 2018.
- [22] J. Y. Liu, H. X. Wang, Y. Y. Sun, C. Q. Fu, and J. Guo, “Real-coded quantum-inspired genetic algorithm-based BP neural network algorithm,” *Mathematical Problems in Engineering*, vol. 2015, Article ID 571295, 10 pages, 2015.
- [23] E. Egrigöglü, U. Yolcu, C. H. Aladag, and C. Kocak, “An arma type fuzzy time series forecasting method based on particle swarm optimization,” *Mathematical Problems in Engineering*, vol. 2013, Article ID 935815, 12 pages, 2013.
- [24] J. Yang, L. Li, Y. Shi, and X. Xie, “An ARIMA model with adaptive orders for predicting blood glucose concentrations and hypoglycemia,” *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 3, pp. 1251–1260, 2019.
- [25] B. Wang, X. Zhu, Q. S. He, and G. D. Gu, “The forecast on the customers of the member point platform built on the blockchain technology by arima and lstm,” in *Proceedings of the 2018 IEEE 3rd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)*, Chengdu, China, April 2018.
- [26] R. S. F. Ferraz, R. S. F. Ferraz, F. C. Cruz, E. F. S. Filho, and A. F. S. Correia, “Multi step forecasting of the wind generation by ARIMA and K-NN,” in *Proceedings of the 2018 Simposio Brasileiro de Sistemas Eletricos (SBSE)*, Niteroi, Brazil, May 2018.
- [27] J. Li, S. W. Yin, G. S. Shi, and L. Wang, “Optimization of indoor thermal comfort parameters with the adaptive network-based fuzzy inference system and particle swarm optimization algorithm,” *Mathematical Problems in Engineering*, vol. 2017, Article ID 3075432, 13 pages, 2017.
- [28] Hong Kong Energy Statistics, *Census and Statistics Department Hong Kong Special Administrative Region*, Hong Kong Energy Statistics, Hong Kong, China, 1979.
- [29] B. S. Liang, B. Gao, and F. Liu, “Research on the dynamic relationship between fiscal expenditure on supporting agriculture and development of the primary industry in Hunan Province,” *Journal of Hunan Agricultural University*, vol. 43, pp. 450–458, 2009, in Chinese.
- [30] Hong Kong Trade in Services Statistics, *Census and Statistics Department Hong Kong Special Administrative Region*, Hong Kong Trade in Services Statistics, Hong Kong, China, 1999.
- [31] A. Jeyasekar, S. V. Kasmir, and R. A. Uthra, “Congestion avoidance algorithm using ARIMA(2,1,1) model-based RTT estimation and RSS in heterogeneous wired-wireless network,” *Journal of Network and Computer Applications*, vol. 93, no. 1, pp. 91–109, 2017.
- [32] A. Rath, S. Samantaray, and P. P. C. Swain, “Flow forecasting of hirakud reservoir with ARIMA model,” in *Proceedings of the 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)*, Chennai, India, August 2017.
- [33] C. Ren, N. An, J. Wang, L. Li, B. Hu, and D. Shang, “Optimal Parameters Selection for BP neural network based on particle swarm optimization: a case Study of wind speed forecasting,” *Knowledge-Based Systems*, vol. 56, pp. 226–239, 2014.
- [34] B. Y. Xu, H. J. Zhang, Z. T. Wang, H. X. Wang, and Y. L. Zhang, “Model and algorithm of BP neural network based on expanded multichain quantum optimization,” *Mathematical Problems in Engineering*, vol. 2015, Article ID 362150, 11 pages, 2015.
- [35] H. Z. Li and S. Guo, “External economies evaluation of wind power engineering project based on analytic hierarchy process and matter-element extension model,” *Mathematical Problems in Engineering*, vol. 2013, Article ID 848901, 11 pages, 2013.
- [36] P. Zhou, M. Yuan, H. Wang, and T. Y. Chai, “Data-driven dynamic modeling for prediction of molten iron silicon content using ELM with self-feedback,” *Mathematical*

- Problems in Engineering*, vol. 2015, Article ID 326160, 11 pages, 2015.
- [37] P. G. Li, J. Zhang, C. B. Li et al., "Dynamic similar sub-series selection method for time series forecasting," *IEEE Access*, vol. 6, pp. 32532–32542, 2018.
  - [38] A. Bhargava, "On the theory of testing for unit roots in observed time series," *The Review of Economic Studies*, vol. 53, no. 3, pp. 369–384, 1986.
  - [39] S. Vergura and M. Carpentieri, "Phase coherence index, HHT and wavelet analysis to extract features from active and passive distribution networks," *Applied Sciences*, vol. 8, no. 1, p. 71, 2018.
  - [40] C. Guarnaccia, J. Quartieri, E. R. Rodrigues, and C. Tepedino, "Acoustical noise analysis and prediction by means of multiple seasonality time series model," *International Journal of Mathematical Models and Method in Applied Sciences*, vol. 8, pp. 384–393, 2014.
  - [41] J. Liu, G. Y. Shi, and K. G. Zhu, "High-precision combined tidal forecasting model," *Algorithms*, vol. 12, no. 3, p. 65, 2019.
  - [42] Z. S. Wu and R. Xue, "A cyclical non-linear inertia-weighted teaching-learning-based optimization algorithm," *Algorithms*, vol. 12, no. 5, p. 94, 2019.
  - [43] A. N. Kolmogorov, "On tables of random numbers," *Theoretical Computer Science*, vol. 207, no. 2, pp. 387–395, 1998.
  - [44] M. Paluch and L. Jackowska-Strumillo, "Hybrid models combining technical and fractal analysis with ANN for short-term prediction of close values on the warsaw stock exchange," *Applied Sciences*, vol. 8, no. 12, p. 2473, 2018.
  - [45] Y. Li, *Returning to 20 Years: Review and Prospect of Hong Kong's Economic Development*, Chinese Finance, Beijing, China, 2017.
  - [46] R. N. Su, "Study on the Path of Division of Labor, cooperation and competition in the international value chain in the development strategy of "Guangzhou-foshan-one-city," *Special Economic Zone*, vol. 3, pp. 152–161, 2017.

