The application of combination forecasting methods in energy consumption based on RS

Shubao Xing

School of Management, Xi'an University of Science and Technology, Xi'an 710054, China

Abstract

Energy demand is a complex nonlinear system, whose developmental changes have dual trends of increase and fluctuation. Scientifically predicting energy consumption is of great significance to optimize the allocation of energy. The combination can take full advantages of known information to improve prediction accuracy .Via establishing the relative data model of prediction model and prediction object, the data value of discrete attribute is used to establish the knowledge expression system and decision table. Based on RS theory, the weights of the single models in combination forecasting are calculated, which can be used to construct the combination forecasting model and predict the energy demand in China. The results indicate that the average error of combination forecast model is 1.96,which is lower than 3.15, 2.09 and 2.45 of linear regression model, GM(1,1) and BP neural network.

Keywords

Rough Set; Energy Consumption; Combination Forecasting; GM（1,1）Model

1. Introduction

Energy is an important material guarantee for economic development and social progress, as well as a vital strategic material related to the economic lifeline and security of a country, which appears to be the significant factor in the modernization drive. Scientific analysis of China's energy demand is of great significance to the formulation of the correct strategy planning of energy security [1-3]. In recent years, the continuous growth of energy demand has brought the state of imbalance between supply and demand, which has become the focus of the industry and academia. Scholars at home and abroad including related institutions have carried out extensive research on energy forecasting. Many scholars have conducted research on the prediction of energy demand, and some prediction methods are frequently used: the trend extrapolation, ARMA model, BP neural network prediction, grey system, exponential smoothing and moving average. Different forecasting methods have their advantages and disadvantages. They are not mutually exclusive, but are interrelated and complementary [4-5].

Since Bates and Granger first put forward the theory, the research and application of combination forecasting methods have developed rapidly. This paper, based on the fundamental idea of combination forecasting, uses some appropriate methods and integrates the results of various single models to complement on each other and to improve the prediction accuracy and increase the reliability of the effect. On one hand, one of the most important facets of combination forecasting is the determination of the weight coefficient of each individual prediction model, for which influence the final effect; On the other hand, it reflects the importance degree of each forecasting model in the combination forecasting method, which is one of the core contents of rough set theory [6].

Through the analysis of the historical data of energy consumption in our country to deal with, this paper compares the fitting degree with various forecasting models and historical data by analysis of time changing series chart. We finally pick three models of unary regression, GM (1, 1) and BP neural network prediction. The combination forecasting model based on rough set weight allocation is used to predict the energy demand in China, with accuracy comparison, we obtain a practical method of energy demand forecasting.

2. Single prediction model

2.1 unary regression model

According to China statistical yearbook of energy consumption data (see table 1 column 2), historical statistics of energy consumption in China have been increasing over time, therefore, we choose unary regression analysis to forecast. Using SPSS statistical software to establish prediction model:

$$Y(t) = 147159.8476 + 120697.1107t \tag{1}$$

$Y(t)$—Energy demand value

t—time variable, corresponding values in 2001~2015, 1~15 respectively.

The variance analysis showed that the tail probability of the significance test of unary regression model was less than 0.0001, which decides the ratio $R^2 = 0.9774$, $F = 606.40 > F_{0.01}$, The fitting tests showed its significance and higher fitting precision for the average relative error was 3.15%, from which we can see in Table1.

2.2 GM(1, 1)Model

The prediction of grey series G (1,1) Model is a realistic and dynamic analysis and prediction, if the raw data sequence is given, that is $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\}$, then continuous data with different lengths are selected as subsequences from the sequence $X^{(0)}$. GM (1, 1) model is established for subsequence, ensure any sequence of sub data as

$$X_i^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(m)\} \tag{2}$$

Generate a cumulative sequence of data sequences, we can obtain

$$X_i^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(m)\} \tag{3}$$

$$x^{(1)}(t) = \sum_{k=1}^{t} x^{(0)}(k), \quad t = 1, 2, \cdots, m$$

Structure accumulation matrix B and constant term vector C, that is

$$B = \begin{bmatrix} -\frac{1}{2}[x^{(1)}(1) + x^{(1)}(2)] & 1 \\ -\frac{1}{2}[x^{(1)}(2) + x^{(1)}(3)] & 1 \\ \vdots & \vdots \\ -\frac{1}{2}[x^{(1)}(m-1) + x^{(1)}(m)] & 1 \end{bmatrix} \tag{4}$$

$$Y_m = [x^{(0)}(2), x^{(0)}(3), \cdots, x^{(0)}(m)]^T \tag{5}$$

Solve grey parameters $\hat{a}$ by least square method

$$\hat{a} = \begin{bmatrix} a \\ u \end{bmatrix} = (B^T B)^{-1} B^T Y_m \tag{6}$$

Establish GM(1, 1)Model

$$\hat{x}^{(0)}(t+1) = [x^{(0)}(1) - \frac{u}{a}]e^{-at} + \frac{u}{a} \tag{7}$$

take the derivative reduction of $\hat{X}^{(1)}$

$$\hat{x}^{(0)}(t+1) = -a[x^{(0)}(1) - \frac{u}{a}]e^{-at} \tag{8}$$

In accordance with statistics of China's energy consumption in 2005~2015, we utilize GM (1, 1) model to build its grey prediction model

$$\hat{x}(t+1) = 284036.41e^{0.05t} - 278682.45 \tag{9}$$

Meanwhile, test variance ratio $C = S_2 / S_1$ (S_2 is quadratic mean deviation of residuals, S_1 is quadratic mean deviation of primitive series) ,so $C = 9421.8695 / 58778.9197 = 0.1603 < 0.35$, while small error probability p=1.000>0.95. We can draw a conclusion that the established GM (1, 1) model has a good fitting accuracy and the average relative error (MAPE) is 2.09% (See Table1), which can be used to predict and analyze energy consumption.

2.3 BP Neural network model

According to the historical data of China's energy consumption in the past 2000-2015, this paper selects data from the first three years to predict the fourth year's, so 3 neurons are selected in the input layer and 1 neurons in the output layer , then search the hidden layer nodes at about 2-3 times of the input layer nodes. It is found that when the input layer node is 3 and the hidden layer node is 8, the fitting precision of BP neural network is higher. Training BP neural network and modeling on the basis of the following parameters:(1)The number of nodes in the input layer is 3; (2)The number of hidden layer nodes is 8;(3)The output node number is 1;(4)The learning rate is 0.5; (5)Training times is 3000; (6)Network training error is 0.0001; The initial weights and thresholds of the network are automatically selected by the network.

The weight between the input layer and the hidden layer can be calculated by MATLAB programming

$$v = \begin{bmatrix} 0.1037 & 0.2098 & -1.1530 & 1.4120 & 1.3421 & -1.2260 & 1.0125 & -1.8497 \\ 0.2999 & -0.3011 & -0.5730 & -0.2951 & 0.3271 & -0.6432 & 1.1975 & -0.5144 \\ -1.2593 & -1.3400 & 0.2699 & 1.2667 & 0.9159 & 0.3379 & 0.4927 & 0.0389 \end{bmatrix}$$

The threshold between the input layer and the hidden layer is

$$st1 = [0.6561 \quad 0.6671 \quad 0.6679 \quad 0.5539 \quad 0.7815 \quad 0.1228 \quad 0.9727 \quad -0.0937]$$

The weight between the hidden layer and the output layer is

$$w = [-0.5985 \quad -1.2048 \quad -1.0936 \quad 1.0825 \quad 0.9346 \quad -0.7860 \quad 1.5565 \quad -1.4325]$$

The threshold between the input layer and the hidden layer is

$$st2 = 0.4731$$

Under the condition that the training times and the overall errors meet the requirements, fitting value and relative error of BP neural network can be seen in table1, and the absolute relative error is 2.15%. Consequently, it can be used in the prediction and analysis of energy consumption.

3. Combination forecasting model based on Rough Set

3.1 Determination of weight coefficient in combination forecasting method

The combination forecasting method is described in mathematical language as follows: Suppose that there are m forecasting models to predict the same forecast object, the combination forecasting model composed of these m single forecasting models is the following

$$\hat{y}_t = \sum_{i=1}^{m} k_i \hat{y}_t^{(i)} \tag{10}$$

$\hat{y}_t$ is the predictive value of t time combination forecasting model; $\hat{y}_t^{(i)}$ is the predictive value of i prediction model at t moment $(i = 1, 2, \cdots, m)$; k_i is the weight of the i prediction model $(i = 1, 2, \cdots, m)$,

$$\sum_{i=1}^{m} k_i = 1, \quad \text{and} \sum_{i=1}^{m} k_i \geq 0$$

3.2 Determination method of weight coefficient based on Rough Set Theory

In 1982, Z.Pawlak, a scholar of Poland, put forward the theory of rough set. Rough set is a mathematical tool for dealing with incompleteness and uncertainty, which can effectively deal with inaccuracy, inconsistent, and other incomplete information, from which we can find out the connotative knowledge and reveal potential laws[7]. Based on rough set, the combination forecasting method transforms the weight coefficient determination issues into the important evaluation issues of attribute in rough set.

The set composed of various forecasting methods is considered as the conditional attribute of decision table, while the observed values of forecast objects are regarded as decision attributes. Then the importance to decision set (prediction index) of various attributes (forecasting methods) is calculated. Finally, the weights of various forecasting methods are determined according to the importance degree. This method fully analyzes the importance of each prediction method from data analysis, and eliminates the subjectivity of combination forecasting method.

3.2.1 Build relational data model

In order to ascertain the weight coefficient in the combination forecasting model, we need to establish relational data model first. Considering each prediction method as condition attribute, then conditional attribute set $C=\{\hat{y}^{(1)},\hat{y}^{(2)},\cdots,\hat{y}^{(m)}\}$; Regarding the index y of prediction object as decision attribute, then decision attribute set D={y}; The predictive value of each prediction method and the historical data of the predicted object in the t period are regarded as a message of some research object u_t, here defines $u_t=\{\hat{y}_t^{(1)},\hat{y}_t^{(2)},\cdots,\hat{y}_t^{(m)},y_t\}$, discourse domain $U=\{u_1,u_2,\cdots,u_n\}$, also called sample set. Then, the attribute value of the research object u_t is $y_i(u_t)=\hat{y}_t^{(i)}(i=1,2,\cdots,m;t=1,2,\cdots,n)$, $y(u_t)=y_t\,(t=1,2,\cdots,n)$. The two-dimensional information table formed of $u_t(t=1,2,\cdots,n)$ is the relational data model about the combination forecasting method. Each row in the table describes an object, while each column describes an attribute of an object.

3.2.2 Characterization of attribute values and Building knowledge expression system

In order to analyze the dependency between knowledge and the importance of attributes from each single model, it is necessary to classify the domain and establish knowledge system on the domain. However, the basis of classification is the characterization of attribute value, namely divide attribute values into several eigenvalues according to different feature, then replace them by eigenvalues. The knowledge expression system can be established after characterization.

The knowledge expression system composed of domain, condition attribute and decision attribute can be expressed as $S=(U,A,C,D)$, among them, $A=C\cup D$, $C\cap D=\varnothing$. Usually, the knowledge expression system with conditional attributes and decision attributes is called decision table, and each row in a decision table represents a decision rule. In the domain, the objects are divided into different decision classes in the light of the different decision rules.

3.2.3 Calculation of the weight coefficients of each model[8-10]

(1)Calculate the dependence of decision attribute set D on condition attribute set C

$$\gamma_C(D)=\frac{|POS_c(D)|}{|U|} \tag{11}$$

$|U|$——cardinality of set U, the number of contained elements in a set of finite sets; $|POS_c(D)|$—— the positive field of decision attribute D about conditional attribute C

(2)Delete the NO I prediction model, and calculate the dependence of decision attribute set D on condition attribute set C-{ci}

$$\gamma_{C-\{c_i\}}(D)=\frac{|POS_{C-\{c_i\}}(D)|}{|U|} \qquad (i=1,2,\cdots,m) \tag{12}$$

(3)Calculate the importance of the i prediction model in all prediction models

$$\mu^D_{C-\{c_i\}}(c_i)=\gamma_C(D)-\gamma_{C-\{c_i\}}(D) \qquad (i=1,2,\cdots,m) \tag{13}$$

(4)Calculate the weight coefficient of the i prediction model

$$k_i = \frac{\mu_{C-\{c_i\}}^D(c_i)}{\sum_{i=1}^{m}\mu_{C-\{c_i\}}^D(c_i)} \quad (i=1,2,\cdots,m) \tag{14}$$

3.2.4 Establishing the combination forecasting model

After calculating the weight coefficients of each single model k_i, the weighted combination of each single model can be achieved by using (10), thus we can obtain the combination forecasting model.

3.3 Combination forecasting model and analysis of energy consumption

3.3.1 Combination forecasting model of energy consumption based on Rough Set

According to the prediction results based on the single model(See Table1), the data in Table 1 is divided into 3 grades by fuzzy cluster analysis, we obtain discrete values instead of attribute values, and establish a knowledge expression system for combination forecasting.

Table 1 Data table of relationship between single model fitting value and energy consumption

Domain U	Condition attribute			Decision attribute
	Model1	Model 2	Model 3	Energy consumption y
u_1	1	1	1	1
u_2	1	2	2	1
u_3	1	2	2	2
u_4	2	2	2	2
u_5	2	2	3	2
u_6	2	2	3	2
u_7	2	3	3	2
u_8	2	3	3	3
u_9	3	3	3	3
u_{10}	3	3	3	3
u_{11}	3	3	3	3

The dependence of energy consumption on 3 forecasting models is calculated by formula (12).

According to the formula (13), a forecasting model is deleted by sequence, we can calculate the dependence of energy consumption y on the remaining 2 forecasting models $\gamma_{c-\{c_i\}}(D)(i=1,2,3)$, then the importance of various forecasting models is calculated by the formula (14) and the formula (15) $\mu_{C-\{c_i\}}^D(c_i)(i=1,2,3)$ as well as the weight coefficient $k_i(i=1,2,3)$. We can see the results in Table2.

According to the formula (13), a forecasting model is deleted by sequence, we can calculate the dependence of energy consumption y on the remaining 2 forecasting models $\gamma_{c-\{c_i\}}(D)(i=1,2,3)$, then the importance of various forecasting models is calculated by the formula (14) and the formula $\mu_{C-\{c_i\}}^D(c_i)(i=1,2,3)$ as well as the weight coefficient $k_i(i=1,2,3)$. We can see the results in Table2.

Table 2 Importance and single model weight

parameter	Model 1	Model 2	Model 3
$\gamma_{C-\{c_i\}}(D)$	0.4545	0.4545	0.6363
$\mu_{C-\{c_i\}}^D(c_i)$	0.1818	0.1818	0
k_i	0.5	0.5	0

From the weight coefficient of each model in Table 4, we can establish the combination forecasting model of energy consumption in China based on Rough Set Theory

$$\hat{y}_t = 0.5\hat{y}_t^{(1)} + 0.5\hat{y}_t^{(2)} \tag{15}$$

We can obtain the predictive value of energy consumption in 2001-2015 by means of above-mentioned combination forecasting model, and calculate the relative error between 3 prediction models of unary regression, GM (1,1), BP neural network and combination forecasting model, see Table3.

Table 3 Relative error of fitting value and relative error of combined model

Year	Actual value	Fitting value of combined model	Relative error of combined model(%)
2005	261369	256009.2	2.05
2006	286467	284293.55	0.76
2007	311442	301551.05	3.18
2008	320611	319129.69	0.46
2009	336126	337044.38	0.27
2010	360648	355310.76	1.48
2011	387043	373945.17	3.38
2012	402138	392964.72	2.28
2013	416913	412387.31	1.09
2014	425806	432231.67	1.51
2015	430000	452517.41	5.24
MARE	—	—	1.96

It can be seen from table 3 that the predicted average error of the combination forecasting model is 1.96, while the average error of unary regression model, GM (1, 1) model and BP neural network is 3.15, 2.09 and 2.45. Consequently, the prediction accuracy of combination forecasting model is higher than that of 3 models.

3.3.2 Prediction results based on rough set combination model

The results of China's energy consumption in 2016~2025 based on based on rough set combination model as shown in Table 4.

Table 4 Combined forecasting value of energy consumption in China during 2016~2025

Year	Unitary regression	GM Model	BP Model	Combined forecasting value
2016	478328.62	468201.46	430240	473265.04
2017	499026.74	489965.36	432780	494496.05
2018	519724.85	512740.93	433460	516232.89
2019	540422.96	536575.21	434700	538499.08
2020	561121.07	561517.40	435440	561319.23
2021	581819.18	587619.00	435760	584719.09
2022	602517.29	614933.91	436120	608725.60
2023	623215.40	643518.53	436340	633366.97
2024	643913.51	673431.87	436460	658672.69
2025	664611.62	704735.71	436570	684673.67

The results shows that China's energy consumption reached 4.7326504 billion tons of standard coal in 2016. Besides, it is estimated that 5.6131923 billion tons of standard coal will be reached in 2020 and 684673.67 billion tons in 2025. The annual increasing rate of energy consumption is 4.46%.

4. Conclusion

Using the advantages of rough set theory in dealing with uncertainty problems to determine the weight of each prediction method, the model of energy consumption combination forecasting is established. Through the combination forecasting, it proves that the way has a better prediction accuracy and is higher than the prediction results of the selected models such as one-order-regress model, the GM (1, 1), each forecast model of BP neural network, therefore, it provides a practical new method for energy demand prediction.

Acknowledgements

National Natural Science Foundation of China（71273206）

Characteristic discipline construction project in Shaanxi

References

[1] ZHANG Jinsuo, ZOU Shaohui.Study on mechanism of coal mining option[J].Journal of XIAN University of Science and Technology,2006, 26（1）: 121-124,136

[2] ZHANG Jinsuo,FENGXue, ZOU Shaohui.China′s Coal Demand Forecast Model basedon the trend Combination[J].Commercial Research2014(6):51-56

[3] YANG Wenqing.Combination forecasting model of China’s coal demand and its application [J].Journal of XIAN University of Science and Technology,2014, 34（5）: 553-557

[4] WANG Liming, ZHAO Qinghua. Gray predict and analysis of Heilongjiang province agriculture machinery total power[J]. Journal of Agriculture Mechanization Research, 2003（2）: 48-49.

[5] ZdzisławPawlak, AndrzejSkowron.Rudiments of rough sets[J].Information Sciences,2007, 177(1): 3-27

[6] Marco Aiolfi, AllanTimmermann.Persistence in forecasting performance and conditional combination strategies[J].Journal of Econometrics,2006,135(1-2): 31-53

[7] LU Qi, GU Peiliang, QIU Shiming. The construction and application of combination forecasting model in Chinese energy consumption system[J]].System Engineering Theory & Practice,2003,23(3):24-30.

[8] LIU Shuan, DU Hongtao, WANG Xiao-ling. Development of the rough set theory and its application[J]. System Engineering Theory & Practice,2001,10:77-81.

[9] Miao Duoqian. A new method of discretization of contuous attributes in rough sets[J].ACTA Automatic SINCA, 2001, 27(3):296-302.

[10] ZHAO Xianqiao, CAO Xinyu, LAN Zequan. Study of the method for determining weighting coefficient of coalash slagging fuzzy combination forecast based on rough set theory[J]. Journal of China Coal Society, 2004, 29(2): 222-225.