www.czasopisma.pan.pl PAN www.journals.pan.pl

GOSPODARKA SUROWCAMI MINERALNYMI – MINERAL RESOURCES MANAGEMENT

2021 Volume 37 Issue 1 Pages 21–38

DOI: 10.24425/gsm.2021.136293

MIN REN[1], JIANYONG DAI[2], WANCHENG ZHU[3], FENG DAI[4]

Combined modelling for iron ore demand forecasting with intelligent optimization algorithms

Introduction

As one of the basic raw materials used for national economic development, a safe and stable iron ore supply can essentially guarantee sustainable development. The National Plan for Mineral Resources (2016–2020) issued by the Ministry of Land and Resources notes that China is currently in the middle stage of industrialization and that energy demand growth will slow, but the total energy demand will remain at a high level. However, the reserve-production ratios of most mineral products are still low in China. Resource security issues are becoming increasingly prevalent. Therefore, it is important to effectively forecast the iron ore demand, and accurate forecasts could aid future development strategies.

Many types of energy demand forecasting models exist in the literature. The models that have been applied mainly include the traditional statistical and econometric techniques of

✉ Corresponding Author: Min Ren; e-mail: renmin@stumail.neu.edu.cn

[1] Northeastern University, Shenyang, China; e-mail: renmin@stumail.neu.edu.cn

[2] University of South China, Hengyang, China; e-mail: 79417049@qq.com

[3] Northeastern University, Shenyang, China; e-mail: zhuwancheng@mail.neu.edu.cn

[4] Northeastern University, Shenyang, China; ORCID iD: 0000-0002-0113-9763; e-mail: typhoondf@stumail.neu.edu.cn

exponential smoothing (Mi et al. 2018), linear regression (D'Amico et al. 2020), autoregressive integrated moving average (Wang et al. 2020), and generalized autoregressive conditional heteroscedasticity (Bikcora et al. 2018) models. If these models are used separately for forecasting, their neglect of the nonlinear characteristics of the energy demand will result in large errors. To solve this problem, machine learning methods have been produced, such as the artificial neural network (Al-Fattah 2020) and support vector machine (Kazemzadeh et al. 2020). The extensive use of these machine learning methods has contributed to improvements in forecasting performance. However, such improvements do not suggest that nonlinear models can completely replace linear models in forecasting the energy demand. Davies and Petruccelli concluded that the conditional mean and conditional median forecasts of nonlinear time series models exhibit poor forecasting performance compared to those of linear models (Davies and Petruccelli 1988). In other words, one forecasting model cannot accurately encompass all the information in a time series, but different types of models can play complementary roles to improve the forecasting accuracy. Thus, combined forecasting models have been developed.

The concept of combining forecasts was first proposed by Bates and Granger (Bates and Granger 1969). They found that a combined forecasting model displayed better forecasting performance than individual models. Combined forecasting models have been applied in various fields, such as energy demand forecasting, tourism demand forecasting, price prediction, electricity consumption forecasting and wind speed forecasting. For example, Liu et al. proposed a combined forecasting model based on the grey forecasting method and a back-propagating neural network to predict energy consumption (Liu et al. 2016). Wang et al. established a combined forecasting model of the tourism demand using an artificial neural network and a clustering algorithm (Wang et al. 2018). Zhang et al. presented a hybrid model based on the autoregressive moving average, a kernel-based extreme learning machine and wavelet transform to forecast electricity prices (Zhang et al. 2017). Zhou et al. proposed a cloud model based hybrid method for combining forecast (Zhou et al. 2019). Song et al. constructed a combined wind speed forecasting model that effectively improved the forecasting accuracy (Song et al. 2018). The applications of combined models have yielded a common result: combined forecasting models generally outperform individual model. Notably, it is not widely used in iron ore demand forecasting.

In this study, the Holt-Winters (HW) non-seasonal exponential smoothing, autoregressive integrated moving average (ARIMA), support vector machine (SVM) and extreme learning machine (ELM) models are combined to capture various relationships and characteristics from time series data to accurately forecast the iron ore demand. As a classical statistical prediction tool, the ARIMA model has been one of the most popular models for time series prediction. Moreover, the exponential smoothing (Holt-Winters) model has the advantages of a low computational demand and easy operation. However, the two methods can only assess linear features of time series data. To overcome this shortcoming, the SVM and ELM algorithms are incorporated into the combined model. The SVM is a neural network algorithm by which the nonlinear features of the iron ore demand can be obtained.

The method adopts the policy of structural risk minimization and is especially suitable for solving problems based on a small sample size. ELM is a single hidden layer feed-forward neural network, which can obtain the unique optimal solution by setting the number of hidden layer neurons. And it has the characteristics of fast learning rate and good generalization performance. The advantages and disadvantages of the four individual models are complementary, which may lead to high prediction accuracy.

In a combined forecasting model, the weighted coefficients of the individual models have a considerable impact on the resulting prediction. Therefore, the theoretical method used to determine the weighted coefficients is particularly important in combined forecasting. Since the 1970s, a series of new intelligent optimization algorithms have been proposed. The optimal solution can be obtained by simple information dissemination and evolution methods. The advantages of intelligent optimization algorithms lie in high parallelism, self-organization, self-learning and self-adaptation and providing a new way to solve complex problems. The related researches on applying the intelligent optimization algorithm to construct combined forecasting models are shown in Table 1. Genetic algorithm (GA), particle swarm optimization algorithm (PSO) and simulated annealing (SA) algorithm are the most commonly used intelligent optimization algorithms in industrial and scientific research fields. Therefore, this paper uses the GA, PSO and SA algorithms to determine the weighted coefficients of the HW, ARIMA, SVM and ELM models and obtain the best combined prediction model of the iron ore demand.

Table 1. Applications of intelligent optimization algorithms in combined forecasting

Tabela 1. Zastosowania inteligentnych algorytmów optymalizacyjnych w połączonym prognozowaniu

Authors	Combined forecasting models	Main results
Zhang et al. (2019)	The non-dominated sorting genetic algorithm III combined system with three objective functions, was proposed and successfully employed to solve the predicament of electricity load forecasting.	Both the stability and the accuracy of the proposed combined system are superior to the compared model which was shown in the experiment results.
Al-Hnaity and Abbod (2016)	A hybrid model combining BPNN, SVM and SVR is presented. The weight of the proposed model is determined by GA.	The numerical results of this model are better than those of all single models, traditional simple average combination models and traditional time series model.
Wang et al. (2010)	This paper presented a new combined model for electric load forecasting, and the adaptive particle swarm optimization was employed to optimize the weight coefficients in the combined forecasting model.	The proposed combined model has been compared with the individual models and the other combined model reported in the literature and its results are promising.
Wang et al. (2012)	The combination of ESM, ARIMA and BPNN includes the advantages of all three models. GA determines the weight of the proposed hybrid model.	Numerical results show that the model outperformed all traditional models, including ESM, ARIMA, BPNN, the equal weight hybrid model and the random walk model.

The structure of this paper is as follows. The first section describes the individual forecasting models, including the HW, ARIMA, SVM and ELM models. The intelligent optimization algorithms of determining the weight are introduced in the second section. In the third section, the experimental results of individual models and combined forecasting models are analyzed, and a comparative study is conducted. Finally, the conclusions are presented.

1. Individual forecasting models used in the combined model

1.1. Holt-Winters (HW) non-seasonal exponential smoothing model

The HW method was proposed by Winters in 1960 (Winters 1960). The Holt-Winters (HW) non-seasonal exponential smoothing model is suitable for time series with linear trends and no seasonal variations. Therefore, this paper applies the Holt-Winters (HW) non-seasonal exponential smoothing model to forecast the iron ore demand. The calculation formula of smooth sequence $\hat{y}_t$ of time series y_t is as follows:

$$\hat{y}_{t+k} = a_t + b_t k \tag{1}$$

where a_t and b_t represent the intercept and slope respectively. These two parameters are defined by the following two recursive formulas:

$$a_t = \alpha y_t + (1 - \alpha)\,(a_{t-1} + b_{t-1}) \tag{2}$$

$$b_t = \beta(a_t - a_{t-1}) + (1 - \beta)\,b_{t-1} \tag{3}$$

where α and β are the damping factors, they are between 0 and 1, $k > 0$. The forecast formula is as follows:

$$\hat{y}_{T+k} = a_T + b_T k \tag{4}$$

The HW model can effectively predict future trends based on historical data, and it meets the requirements of iron ore demand prediction with linear trends. Therefore, this approach can be reliably used to establish a combined model.

1.2. Autoregressive integrated moving average model (ARIMA)

The ARIMA model is a well-known time series forecasting method. It was proposed by Box and Jenkins in the early 1970s (Box and Jenkins 1976). The ARIMA model is an extension of the ARMA model, which includes an autoregressive model (AR) and a moving average model (MA). The ARIMA model can be expressed as ARIMA(p, d, q), where p and q are the orders of the AR and MA models, respectively, and d is the number of differences. The basic concept of this approach is to approximately describe a random sequence with a mathematical model. The random sequence is a data sequence of a predicted object. After the model is identified, it can be used to forecast the trend values at future times based on the past and current values of a time series. This can be written as:

$$\phi(B)\,\nabla^d x_t = \theta(B)\,e_t \tag{5}$$

$$\nabla^d = (1 - B)^d \tag{6}$$

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p \tag{7}$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q \tag{8}$$

where x_t is the historical value at the period t, e_t is the estimated residuals. B is a backward shift operator defined by $Bx_t = x_{t-1}$. $\phi_1, \phi_2, \ldots, \phi_q$ are the autoregressive coefficients. $\theta_1, \theta_2, \ldots, \theta_q$ are the moving average coefficients.

The ARIMA model is simple and easy to understand. It only requires endogenous variables to identify the linear characteristics of stationary time series. Moreover, the model has been successfully applied to univariate time series.

1.3. Support vector machine (SVM)

Support vector machine is a neural network classification technique based on statistic learning theory and the structural risk minimization principle (Vapnik 1995). In the support vector regression model, sample data is divided into a training sample and a test sample. Its fundamental idea is to map training data from input space into high-dimensional feature space. The best fitting effect is obtained in the space of the optimal decision function model, and the test sample is used to validate the analytical model results. Given a training data set of T{(x_1, y_1), ..., (x_i, y_i)} in which $x_i \in R^n$ ($i = 1, \ldots, l$) is the input vector, $y_i \in R$ ($i = 1, \ldots, l$) is the output value and l is the total number of sample data. The key idea of the model is to map the input space into a higher dimensional and possibly infinite-dimensional feature space via nonlinearly mapping $\phi(x)$. The aim is to determine $f(x)$ based on the train data set to approximate the unknown function $g(x)$. The form of the approximation function $f(x)$ has the following format:

$$f(x) = \omega^T \phi(x) + b \tag{9}$$

where $\phi(x)$ is the high dimensional feature space, it is nonlinearly mapped from the input space. The coefficients ω and b are estimated by solving the dual problem.

The method has notable advantages in dealing with small sample sizes, nonlinearity and high-dimensional problems. As an effective method, it can be used to effectively forecast the iron ore demand.

1.4. Extreme learning machine (ELM)

Extreme Learning Machine (ELM) is a learning algorithm (Huang et al. 2006), which is suitable for single hidden layer feed-forward neural network. The main characteristic is that only the hidden layer nodes are required before network training. And the weights of the input node and hidden layer node and hidden node threshold can be assigned randomly. Then the output weight of the hidden layer can be obtained by an analytic operation. Its network training process is completed at a time without complicated iterative operation. Given input samples, the output of ELM having L hidden nodes is modelled by:

$$\sum_{L}^{i=1} \beta_i f(\omega_i x_j + b_i) = y_j, \quad j = 1, 2, ..., N_s \tag{10}$$

where $\{x_j, y_j\}$ represents the training samples, N_s is the number of samples, $\omega_i = [\omega_{i1}, \omega_{i2}, ..., \omega_{in}]^T$ is the weight vector connecting the i-th hidden node and the input nodes, $\beta_i = [\beta_{i1}, \beta_{i2}, ..., \beta_{im}]^T$ is the weight vector connecting the i-th hidden node and the output nodes, and b_i is the threshold of the i-th hidden node.

The ELM algorithm has fast learning speed and generalization performance, which can be applied to the prediction of iron ore demand and play an active role in combined forecasting.

2. Combined forecasting model

2.1. Theory of combined forecasting

The basic theory of combined forecasting involves the linear combination of individual models. To obtain the linear and nonlinear features of time series data, the HW, ARIMA, SVM and ELM models are chosen as the individual forecasting models. Then, the weighted coefficients of the individual forecasting models are determined by three intelligent optimization algorithms.

In this approach, i (i = 1, 2, ..., m; m = 4) represents the number of individual models; t(t = 1, 2, ..., n) is the sample interval; f_{it} is the forecasted value of an individual model; y_t is the actual value; and the weighted coefficients of individual models are w_i, $\sum_{i=1}^{m} w_i = 1$. Thus, the forecasted value using the combined model is as follows.

$$\hat{y}_{combined(t)} = \sum_{i=1}^{m} w_i f_{it} \tag{11}$$

Additionally, the prediction error of the combined model can be calculated as follows.

$$e_t = y_t - \hat{y}_t = \sum_{i=1}^{m} w_i y_t - \sum_{i=1}^{m} w_i f_{it} = \sum_{i=1}^{m} w_i (y_t - f_{it}) = \sum_{i=1}^{m} w_i e_{it} \tag{12}$$

Therefore, the forecasted value of the combined model of the iron ore demand can be expressed as follows:

$$\hat{y}_{combined(t)} = w_1 \hat{Y}_{HW(t)} + w_2 \hat{Y}_{ARIMA(t)} + w_3 \hat{Y}_{SVM(t)} + w_4 \hat{Y}_{ELM(t)} \tag{13}$$

where $\hat{Y}_{HW(t)}$, $\hat{Y}_{ARIMA(t)}$, $\hat{Y}_{SVM(t)}$ and $\hat{Y}_{ELM(t)}$ are the forecasted values of the HW, ARIMA, SVM and ELM models, respectively, in period t. Additionally, w_1, w_2, w_3 and w_4 are the weighted coefficients of the HW, ARIMA, SVM and ELM models, respectively.

2.2. The combined methodology

The weight coefficient plays an important role in the establishment of the combined model. Different theoretical methods will exhibit different forecasting accuracies. To obtain a relatively optimal combined forecasting model, this paper selects the particle swarm optimization algorithm, simulated annealing algorithm and genetic algorithms to allocate the weights. The optimal prediction model of the iron ore demand is identified by comparison and analysis.

2.2.1. Use of the intelligent optimization algorithm

Intelligent optimization algorithms have global optimization performance and versatility, and they are suitable for parallel processing. In general, these algorithms have a rigorous theoretical basis rather than rely on expert experience. They can minimize the sum of absolute errors to get the optimum weights of combination forecasting. Therefore, it is a wise decision to use them in combined forecasting of iron ore demand.

The genetic algorithm (GA) is a randomized search method that evolves from the evolution rule of biological world (Holland 1975). It is characterized by the direct manipulation of structural objects without restrictions on derivative and function continuity. GA has implicit parallelism and global optimization performance. Adopting the probability method, it can get and guide the search space of optimization voluntarily, adjust search direction by adaptation to itself, and has no definite regular. GA is one of the key technologies of intelligent computing.

The simulated annealing (SA) algorithm is a kind of stochastic optimization algorithm based on the Monte Carlo iterative solution method. Its starting point is based on the similarity between the annealing process of solid matter in physics and general combinatorial optimization problems. With probabilistic jumping property, the SA algorithm randomly searches for the global optimal solution of the objective function in the solution space. The SA algorithm is a general optimization algorithm, which has probabilistic global optimization performance.

The particle swarm optimization (PSO) algorithm is a kind of intelligent searching optimization algorithm, originating from bird's searching for food, in which some particles were initiated, and then iterated in them until an optimization solution was obtained (Eberhart and Kennedy 1995). The advantages of the PSO algorithm are simple, efficient and easy to implement. Therefore, it is an effective tool for optimizing the weights of combined forecasting.

2.2.2. Evaluation criteria

In this paper, the root mean square error (RMSE), mean absolute error (MAE) and average relative percentage error (MAPE) are selected as the evaluation index system for the forecasting models of the iron ore demand.

$$MAE = \frac{1}{N}\sum_{t=1}^{N}\left|y_t - \hat{y}_t\right| \tag{14}$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^{N}\left(y_t - \hat{y}_t\right)^2}{N}} \tag{15}$$

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{(y_t - \hat{y}_t)}{y_t}\right| \tag{16}$$

where y_t is the actual value of iron ore demand; $\hat{y}_t$ is the forecasted value; and t is the number of samples.

The MAE, RMSE and MAPE are metrics used to evaluate the predictive effect, and low values of MAE, RMSE and MAPE indicate high forecasting accuracies.

An accuracy improvement (AI) criterion is used to compare the combined models and other models. The AI criterion is designed to compare the two predictive models and is defined as follows:

$$AI = \frac{S - S_{combined}}{S} \cdot 100\% \tag{17}$$

where S is the sum of the absolute errors of a specified model and $S_{combined}$ is the sum of the absolute errors of combined model.

The predictive effect is assessed based on an AI value greater than or less than zero. If $AI > 0$, the combined model is considered the most effective model; if $AI < 0$, the combined model will not be able to overcome the drawbacks of the specific model.

3. Experimental examples and results

3.1. Data description

There are 40 observations about China's iron ore demand between 1980 and 2019 in Figure 1. As Figure 1 shows, China's iron ore demand showed a growth trend from 1980 to 2014, the total demand decreased from 2015 to 2019. A turn occurred in 2001. The iron ore demand grew at a snail's pace in the first twenty years. Then China's iron ore demand stepped into the golden age of booming from 2002. This data was not directly obtained. Some of the data is from the National Bureau of Statistics, and the remainder is from the Chinese Mining Yearbook.

The iron ore demand showed a slow growth trend from 1980 to 1990; the reason was that China was in the early stage of reform and opening up. At that time, the main consumption

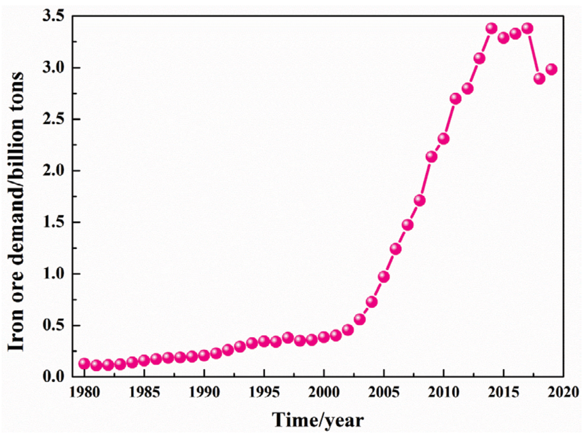

Fig. 1. Chinese annual iron ore demand

Rys. 1. Roczne zapotrzebowanie Chin na rudę żelaza fields of iron ore, including industry and construction, were in their infancy stage. However, the growth of iron ore demand accelerated significantly since 1991. Although the demand for iron ore only reached 0.4016 billion tons by the end of 2001, it has tripled compared with 1980, the average annual rate of increase was up to 5.6%. At this point, the industry and construction are under development, the political strategy of reform and opening up has achieved remarkable results.

From 2002 to 2014, the average annual rate of increase of iron ore demand had reached 18.20%. This is the result of the promotion of national economic strength and the accelerated development of industry and construction. This reduction trend occurred in 2015, and the demand for iron ore decreased by 92.1898 million tons compared with 2014. Although the iron ore demand from 2015 to 2017 showed a very slow growth trend, the reduction of total demand occurred again in 2018, 0.4872 billion tons less than 2017. Then the iron ore demand again started to slowly increase, iron ore demand grew to 2.9817 billion tons from 2018 to 2019 at a rate similar to that of 2002–2003. The occurrence of the reduction trend is related to the adjustment of the industrial structure. At present, China's tertiary industries have surpassed its secondary industries, and the proportion of the mining industry centered on raw materials decreases.

The proper selection of the training set and testing set is an important forecasting consideration. If the training set contains a vast majority of samples and the test set is relatively small, the evaluation results of each model may not be stable and accurate; if the test set contains more samples, the difference between the training set and the data set will be large, and the fidelity of the evaluation results will be reduced. There is no perfect solution to this problem. The common practice is to use about 2/3-4/5 samples for training and the remaining samples for testing (Zhou 2016). In this paper, 3/4 of the data is used to train models, and the remaining 1/4 of the data is used as test sets to determine the model performance. Based on these considerations, the period from 1980 to 2009 is selected as the training period, and 2010 to 2019 as the testing period.

3.2. Models construction

To obtain the optimal iron ore demand prediction model, those individual and combined models, including the HW, ARIMA, SVM, ELM models and PSO, GA and SA combined models are adopted for comparison.

The HW and ARIMA belong to linear models. They only require endogenous variables to identify the linear characteristics of time series. The HW model is used to forecast the iron ore demand, its smoothing coefficient is in the range of 0.01–0.3 according to the principle of minimum sum of squares of the prediction error. The natural logarithm of the iron ore demand data is stationary after one order difference. The logarithm operation does not change the nature and correlation of data, but it compresses the scale of variables, makes the data more stable, and weakens the collinearity and heteroscedasticity of the ARIMA model.

Under the assumption of stationary data, ARIMA (1, 1, 1) is identified based on the comprehensive assessment of the t test and the information criteria. After the residual sequence correlation test and a heteroscedasticity test, the prediction results of ARIMA (1, 1, 1) can be obtained.

ELM and SVM are designed to pick up nonlinear patterns from the iron ore demand time series. Experiments and error analysis show that ELM performs better when four neurons are used as the hidden layer and sigmoid activation function is chosen. The radial basis function (RBF) is used as the kernel function of SVM. Through a series of experiments to establish SVM, the best fitting parameters are obtained as $c = 800$ and $g = 0.05$.

SA, PSO and GA algorithms are popular intelligent optimization algorithms for assigning optimal weights to combinatorial models. For GA, its population size is set as 100, and the number of generations and stall generations are 200. The max iterations of SA are specified as 200. The PSO algorithm performs better when the parameters of c_1 and c_2 are 1.49445, the number of maximum iteration is 200, the number of particles is 100, and ω varies at different data points from 0.4 to 0.9.

3.3. Results analysis

3.3.1. Forecasting results

Figure 2 shows the comparison between each of the model for forecasting the 10 final values from 2010 to 2019. It shows that the HW, ARIMA, SVM and ELM models describe the characteristics of the iron ore demand time series in an approximate manner. As the Figure 2 shows, HW and ARIMA models had a good forecast effect on the iron ore demand from 2010 to 2014, but a relatively poor forecast effect for 2015–2019. The reason may be that the data from 2010 to 2014 tends to be straight lines, while the fluctuation of data from 2015 to 2019 is more obvious. SVM and ELM models have more advantages in the forecasting of iron ore demand from 2015 to 2019 than HW and ARIMA models, which reflect the fluctuation of real data in an approximate manner. This is bound to be inseparable from the fact that HW and ARIMA models belong to linear models, the SVM and ELM are defined as nonlinear models. Usually, linear and nonlinear factors should be considered simultaneously. Only in this way can the prediction results be closer to the true values. It can be observed that when GA, PSO and SA are used to construct the combined forecast models, the forecasts obviously improve. Not only the forecast results of 2010–2014 are pretty close to the true values, but the forecast results from 2015 to 2019 are also satisfactory.

3.3.2. Error analysis

Overall, the individual forecasting models and combined forecasting models all exhibit greater errors from 2015 to 2019 than those in other periods in Figure 2. The most likely

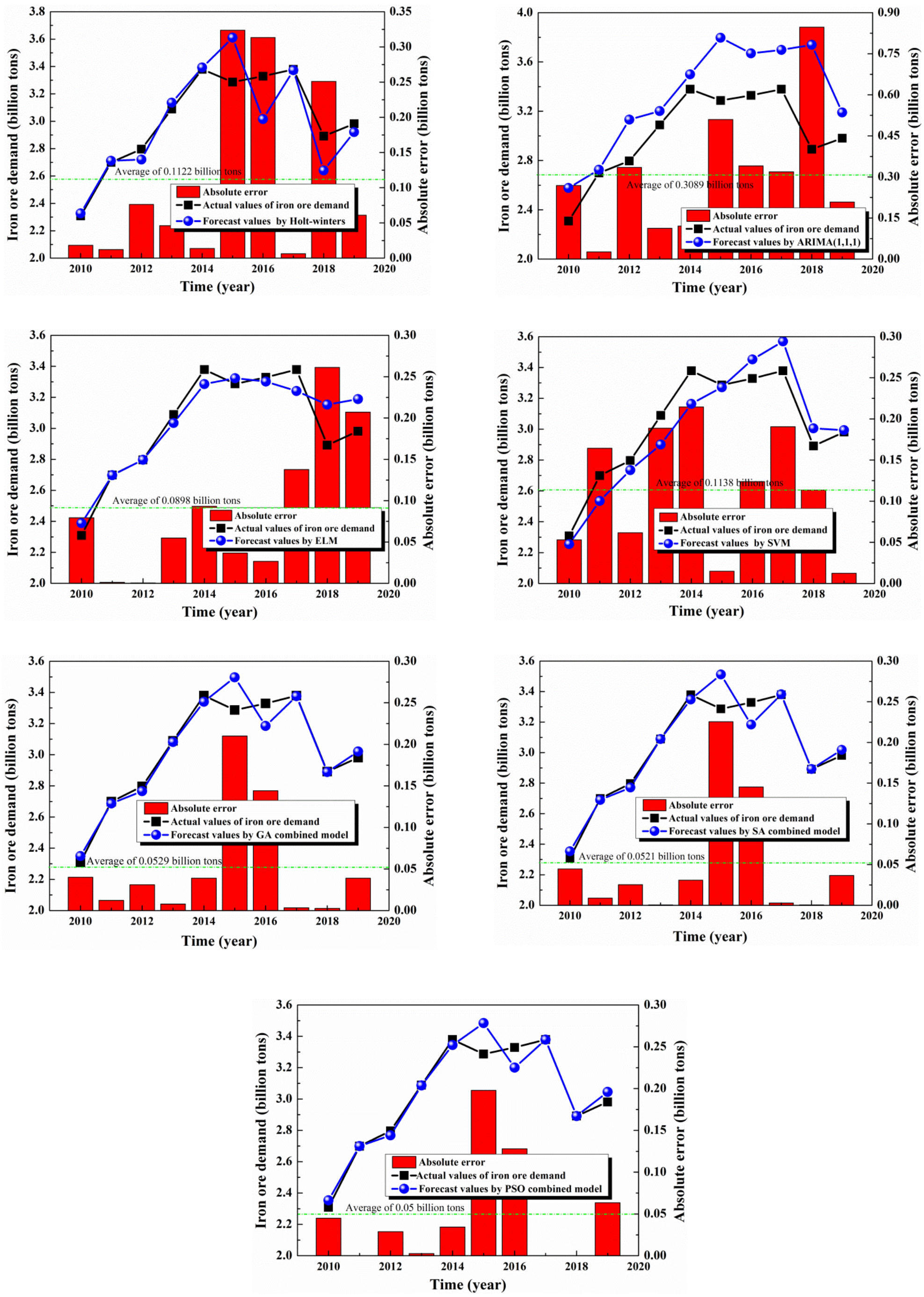

Fig. 2. The actual values and the forecast values by individual models and combined models

Rys. 2. Wartości rzeczywiste i prognozowane według modeli indywidualnych i modeli połączonych reason for this result is the change of the development trend. The iron ore demand showed a growth trend from 1980 to 2014, but it decreased from 2015 to 2019. However, forecasting models must further study the long-term statistical regular and useful information of the iron ore demand from time series data. It is easy to ignore these values with large deviations from the general trend. Under that condition, the HW, ARIMA, ELM and SVM models are combined using weights between 0 and 1. Obviously, the average of forecasting error of combined model from 2015 to 2019 is reduced relative to that of the individual models. This finding suggests that these combined forecasting models encompass the advantages of low error model. The combined model, to a certain extent, improves the predictive effect for abnormal data. Thus, the rationale for applying four individual forecasting models is demonstrated, and the superiority of the combined forecasting model is shown.

By observing the actual and forecasted value of iron ore demand in Figure 2, we find that the forecasted value of the four individual models are less likely to all be simultaneously larger or smaller than the actual value. That is, within the same sample, the signs of the errors are usually inconsistent. For the forecast results of iron ore demand from 2010 to 2015, the forecasted value of the HW and ELM models are close to the actual value. Although the forecast results of the ARIMA and SVM models are not satisfactory, the forecasted value of the ARIMA model in this period are greater than the actual value, while the forecasted value of the SVM model are lower than the actual value. For the forecast results of 2016–2019, although the forecasted value of the ARIMA and SVM models are greater than the actual value, the forecasted value of the HW and ELM models are generally less than the actual value. Therefore, combined forecasting with the four complementary individual models can achieve the best predictions.

3.3.3. Evaluation results

Table 2 shows the evaluation results of the generalization performance of the HW, ARIMA, SVM, ELM forecasting models and GA, PSO, SA combined models for iron ore demand, including MAE, RMSE, MAPE and AI evaluation criteria. From the experimental results given in Table 2, we can find the effectiveness of the combination models. The combination models present the lower error value for the MAE, RMSE, and MAPE criteria than the individual methods. The performance of the individual methods (HW, ARIMA, SVM, and ELM) is inferior to those of the combination models. This is due to the fact that the combination models integrate the linear and nonlinear information to forecast, while the individual model only utilizes the linear or nonlinear information of the time series.

As can be seen from the evaluation results in Table 2, the robustness and generalization of the PSO combined model are better than other models. The results of MAE, RMSE and MAPE support the view that the PSO combined model has the best prediction effect in forecasting iron ore demand. Observing the results of AI criterion, the accuracy of PSO combined model compared to the HW, ARIMA, SVM, ELM, GA and SA combined models improved 55.44%, 83.81%, 56.05%, 44.35%, 5.48% and 4.03%, respectively.

Table 2. The evaluation results for the forecasting models

Tabela 2. Współczynniki korelacji zapotrzebowania na rudę żelaza oraz czynniki wpływające

Evaluation index	Individual forecasting models				Combined forecasting models		
	HW	ARIMA	SVM	ELM	GA	PSO	SA
MAE	0.1122	0.3089	0.1138	0.0898	0.0529	0.0500	0.0521
RMSE	0.1668	0.3810	0.1342	0.1225	0.0841	0.0797	0.0878
MAPE	0.0360	0.1033	0.0372	0.0300	0.0169	0.0161	0.0167
AI	0.5544	0.8381	0.5605	0.4435	0.0548	–	0.0403

These findings indicate that the combined forecasting model can effectively predict the iron ore demand and that the forecasting accuracy reaches a higher level based on these intelligent optimization algorithms. In addition, the evaluation results clearly show how excellent forecasting performance can be achieved by the PSO combined forecasting model.

3.3.4. Comparison of forecasting accuracy

The forecasting accuracy of the iron ore demand of 11 models from three previous studies was compared with that of the PSO combined model in Table 3. The proposed PSO

Table 3. Accuracy assessment of different forecasting models for iron ore demand

Tabela 3. Prognozowane wyniki poszczególnych modeli i modeli połączonych

Forecasting models	Evaluation index MAPE (%)	References
Panel model	7.84	Jia and Xu 2014
Grey model (GM)	12.84	
Co-integration model	14.11	
ARIMA model	16.13	
GM(1,1)	16.88	Ma et al. 2013
PSO-GM(1,1)	14.45	
Rolling GM(1,1)	6.70	
PSO-rolling GM(1,1)	2.31	
SARIMA (seasonal autoregressive integrated moving average)	6.33	Wang et al. 2020
NARNN (non-linear autoregressive neural network)	6.32	
EMD (empirical mode decomposition)-NARNN-ARIMA	3.69	
PSO combined model (the model of this paper)	1.61	–

combined model outperformed the other 11 models, suggesting that it has certain advantages in forecasting the iron ore demand. The PSO combined model performed better than the 10 individual models in Table 3 because it combined nonlinear and linear models to pick up nonlinear and linear information from iron ore demand time series for forecasting. This decreases the differences between the actual and forecast values. As seen, the EMD-NARNN-ARIMA provided better forecasts than its constituent models, and this proves the effectiveness of the combination model again. From Table 3, the difference of MAPE between the PSO combined model and the EMD-NARNN-ARIMA combined model which shows that the PSO combined model is more suitable to forecast the iron ore demand. It also proves that its constituent models (HW, ARIMA, SVM, and ELM) chosen for iron ore demand forecasting are reasonable. These results again demonstrate that the PSO combined forecasting model performs better in iron ore demand forecasting.

Conclusions

With the vigorous development of China's industry, the stable supply of energy resources is very important. One of the most important sources of energy are iron ore resources. From the industrialization initial stage to the industrialization middle period, China's iron ore demand will inevitably experience some changes. Therefore, the accurate forecast of iron ore demand is of great significance to the industrialization development in China and even the world.

Iron ore demand is greatly influenced by national politics and the economy, complex characteristics such as uncertainty, nonlinearity and dynamism make forecasting more difficult. And an individual model cannot always accurately forecast the complex iron ore demand time series. The use of combined models can also reduce the risk of choosing an inappropriate model. The intelligent optimization algorithm can find the optimal weights of combined forecasting. In this study, the paper uses three quite mature intelligent optimization algorithms in engineering applications to determine the optimum weighted coefficients of the HW, ARIMA, SVM and ELM to forecast the iron ore demand. The HW and ARIMA models are selected because they can capture the linear features of time series data, and the SVM and ELM can obtain the nonlinear features of the iron ore demand. Therefore, based on intelligent optimization algorithms, the combination of the four models can reduce information loss and improve forecasting accuracy.

Combined forecasting models of iron ore demand are superior to individual models. This conclusion can be obtained from the MAE, MAPE and RMSE values. Notably, the evaluation results of the PSO combined model are better than the other models, and the forecasting accuracy improved by 55.44%, 83.81%, 56.05%, 44.35%, 5.48% and 4.03% when compared with the HW, ARIMA, SVM, ELM, GA and SA combined model, respectively. This improvement is closely related to the main objective of PSO algorithm. Therefore, there are theoretical and practical foundations to prove that the PSO combined model is the optimal forecasting model for iron ore demand. Moreover, the PSO combined model can also be applied to predict the demand for coal and other types of energy because of its good generalization capability and strong robustness. It is a promising predictive tool.

REFERENCES

Al-Fattah, S.M. 2020. A new artificial intelligence GANNATS model predicts gasoline demand of Saudi Arabia. *Journal of Petroleum Science and Engineering* 194.

Al-Hnaity, B. and Abbod, M. 2016. Predicting Financial Time Series Data Using Hybrid Model. *Intelligent Systems and Applications* 650, pp. 19–41.

Bates, J.M. and Granger, C.W.J. 1969. The combination of forecasts. *Journal of the Operational Research Society* 20(4), pp. 451–468.

Bikcora et al. 2018 – Bikcora, C., Verheijen, L. and Weiland, S. 2018. Density forecasting of daily electricity demand with ARMA-GARCH, CAViaR, and CARE econometric models. *Sustainable Energy Grids and Networks* 13, pp. 148–156.

Box, G.E.P. and Jenkins, G.M. 1976. *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco.

Davies, N.J.P. and Petruccelli, J.D. 1988. An Automatic Procedure for Identification, Estimation and Forecasting Univariate Self Exiting Threshold Autoregressive Models. *Journal of the Royal Statistical Society* 37(2), pp. 199–204.

D'Amico et al. 2020 – D'Amico, A., Ciulla, G., Tupenaite, L. and Kaklauskas, A. 2020. Multiple criteria assessment of methods for forecasting building thermal energy demand. *Energy and Buildings* 224.

Eberhart, R. and Kennedy, J. 1995. A new optimizer using particle swarm theory. [In:] *MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, pp. 39–43.

Holland, J.M. 1975. *Adaptation in Natural and Artificial Systems*. The University of Michigan Press, Ann Arbor.

Huang et al. 2006 – Huang, G.B., Zhu, Q.Y. and Siew, C.K. 2006. Extreme learning machine: theory and applications. *Neurocomputing* 70, pp. 489–501.

Jia, L.W. and Xu, D.Y. 2014. Analysis and Prediction of the Demand for Iron Ore: Using Panel, Grey, Co-Integration and ARIMA Models. *Resources Science* 36(7), pp. 1382–1391.

Kazemzadeh et al. 2020 – Kazemzadeh, M.R., Amjadian, A. and Amraee, T. 2020. A hybrid data mining driven algorithm for long term electric peak load and energy demand forecasting. *Energy* 204.

Liu et al. 2016 – Liu, X.L., Moreno, B. and Garcia, A.S. 2016. A grey neural network and input-output combined forecasting model. *Primary energy consumption forecasts in Spanish economic sectors. Energy* 115, pp. 1042–1054.

Ma et al. 2013 – Ma, W.M., Zhu, X.X. and Wang, M.M. 2013. Forecasting iron ore import and consumption of China using grey model optimized by particle swarm optimization algorithm. *Resources Policy* 38, pp. 613–620.

Mi et al. 2018 – Mi, J., Fan, L., Duan, X. and Qiu, Y. 2018. Short-Term Power Load Forecasting Method Based on Improved Exponential Smoothing Grey Model. *Mathematical Problems in Engineering* 2018, pp. 1–11.

National Bureau of Statistics of China. Output of Industrial Products. [Online] https://data.stats.gov.cn/easyquery.htm?cn=C01&zb=A0E0H&sj=2019 [Accessed: 2020-12-30].

National Bureau of Statistics of China, 2018. Chinese Mining Yearbook. Beijing: China Statistics Press.

Song et al. 2018 – Song, J.J., Wang, J.Z. and Lu, H.Y. 2018. A novel combined model based on advanced optimization algorithm for short-term wind speed forecasting. *Applied Energy* 215, pp. 643–658.

Vapnik, V.N. 1995. *The Nature of Statistical Learning Theory*. New York: Springer.

Wang et al. 2018 – Wang, J., Luo, Y.Y., Tang, T.Y. and Peng, G. 2018. Modeling a combined forecast algorithm based on sequence patterns and near characteristics: An application for tourism demand forecasting. *Chaos, Solitons and Fractals* 108, pp. 136–147.

Wang et al. 2012 – Wang, J.J., Wang, J.Z., Zhang, Z.G. and Guo, S.P. 2012. Stock index forecasting based on a hybrid model. *Omega-International Journal of Management Science* 40, pp. 758–766.

Wang et al. 2010 – Wang, J.Z., Zhu, S.L., Zhang, W.Y. and Lu, H.Y. 2010. Combined modeling for electric load forecasting with adaptive particle swarm optimization. *Energy* 35, pp. 1671–1678.

Wang et al. 2020 – Wang, Z.X., Zhao, Y.F. and He, L.Y. 2020. Forecasting the monthly iron ore import of China using a model combining empirical mode decomposition, non-linear autoregressive neural network, and autoregressive integrated moving average. *Applied Soft Computing* 94.

Winters, P.R. 1960. Forecasting sales by exponentially weighted moving averages. *Management Science* 6(3), pp. 324–42.

Zhang et al. 2019 – Zhang, S.H., Wang, J.Y. and Guo, Z.H. 2019. Research on combined model based on multi--objective optimization and application in time series forecast. *Soft Computing* 23, pp. 11493–11521.

Zhang et al. 2017 – Zhang, Y., Li, C. and Li, L. 2017. Electricity price forecasting by a hybrid model, combining wavelet transform, ARMA and kernel-based extreme learning machine methods. *Applied Energy* 190, pp. 291–305.

Zhou et al. 2019 – Zhou, Z., Si, G.Q., Zheng, K., Xu, X., Qu, K. and Zhang, Y.B. 2019. CMBCF: A Cloud Model Based Hybrid Method for Combining Forecast. *Applied Soft Computing* 85.

Zhou, Z.H. 2016. *Machine Learning*. Beijing: Tsinghua University Press, 425 pp. (*in Chinese*).

COMBINED MODELLING FOR IRON ORE DEMAND FORECASTING WITH INTELLIGENT OPTIMIZATION ALGORITHMS

Keywords

iron ore demand, combined model, intelligent optimization algorithm, forecasting accuracy

Abstract

The stable supply of iron ore resources is not only related to energy security, but also to a country's sustainable development. The accurate forecast of iron ore demand is of great significance to the industrialization development of a country and even the world. Researchers have not yet reached a consensus about the methods of forecasting iron ore demand. Combining different algorithms and making full use of the advantages of each algorithm is an effective way to develop a prediction model with high accuracy, reliability and generalization performance. The traditional statistical and econometric techniques of the Holt–Winters (HW) non-seasonal exponential smoothing model and autoregressive integrated moving average (ARIMA) model can capture linear processes in data time series. The machine learning methods of support vector machine (SVM) and extreme learning machine (ELM) have the ability to obtain nonlinear features from data of iron ore demand. The advantages of the HW, ARIMA, SVM, and ELM methods are combined in various degrees by intelligent optimization algorithms, including the genetic algorithm (GA), particle swarm optimization (PSO) algorithm and simulated annealing (SA) algorithm. Then the combined forecast models are constructed. The contrastive results clearly show that how a high forecasting accuracy and an excellent robustness could be achieved by the particle swarm optimization algorithm combined model, it is more suitable for predicting data pertaining to the iron ore demand.

MODELOWANIE DO PROGNOZOWANIA POPYTU NA RUDĘ ŻELAZA POŁĄCZONE Z INTELIGENTNYMI ALGORYTMAMI OPTYMALIZACJI

Słowa kluczowe

zapotrzebowanie na rudę żelaza, model połączony,
inteligentny algorytm optymalizacji, dokładność prognozowania

Streszczenie

Stabilne dostawy zasobów rudy żelaza związane są nie tylko z bezpieczeństwem energetycznym, ale także ze zrównoważonym rozwojem kraju. Dokładna prognoza zapotrzebowania na rudę żelaza ma ogromne znaczenie dla rozwoju industrializacji kraju, a nawet świata. Naukowcy nie osiągnęli jeszcze konsensusu co do metod prognozowania popytu na rudę żelaza. Łączenie różnych algorytmów i pełne wykorzystanie zalet każdego algorytmu to skuteczny sposób na opracowanie modelu predykcyjnego o wysokiej dokładności i niezawodności. W tej publikacji, model Holta-Wintersa (HW) do wygładzania szeregów czasowych, w których występują wahania przypadkowe, jak również autoregresyjny zintegrowany model średniej ruchomej (ARIMA), a także maszyna wektorów nośnych (SVM) i maszyna do ekstremalnego uczenia się (ELM), zostały połączone w celu uchwycenia różnych relacji i charakterystyk na podstawie danych szeregów czasowych, aby dokładnie przewidzieć zapotrzebowanie na rudę żelaza. Zalety czterech algorytmów są w różnym stopniu łączone przez inteligentne algorytmy optymalizacji, w tym algorytm genetyczny, algorytm optymalizacji roju cząstek oraz algorytm symulowanego wyżarzania. Następnie konstruowane są połączone modele. Kontrastowe wyniki wyraźnie pokazują, w jaki sposób można osiągnąć wysoką dokładność prognozowania i doskonałą solidność za pomocą połączonego modelu algorytmu genetycznego. Model taki jest bardziej odpowiedni do przewidywania danych dotyczących zapotrzebowania na rudę żelaza. Opierając się na prognozowanych wynikach połączonego modelu algorytmu genetycznego, możemy stwierdzić, że oczekuje się, iż krajowy popyt na rudę żelaza będzie w przyszłości wykazywał tendencję rozwojową w postaci trwałego, ale powolnego wzrostu.