

Research on the Improved Combinatorial Prediction Model of Steel Price Based on Time Series

Qingyun ZHANG, Dingxia LIU, Xiaojiang WANG, Zhiqing YE, Hanxi JIANG, Wei WEI*

Abstract: Accurately predicting the price change of steel (main building materials) is an effective means to control and manage the cost of construction projects. It is one of the ways for construction enterprises to reasonably allocate building materials, save resources, reduce carbon emissions and reduce environmental pollution. Based on the monthly historical price data of 100 steel rebar (16 mm) from November 2010 to February 2019, the separation and retrieval process of the four components in the time series are improved. The improved multiplicative and additive models were used to make separate predictions, and the reasonable weight is given to combine the multiplication and addition model by the reciprocal of variance method. Finally, an improved prediction model of steel bar price combination with higher prediction accuracy is obtained. The prediction results show that the improved multiplication model and addition model have higher prediction accuracy, their MAPE are 2.62% and 2.36% respectively. Moreover, the prediction accuracy of the combined model is even higher, its MAPE is 2.29%. The prediction accuracy of the improved composite model is higher than that of the individual models. The improved combined prediction model of reinforcement price based on time series method can provide some reference and help for cost control and management in construction engineering, further reduce resource waste and construction non-point source pollution.

Keywords: additive model; combined prediction model; multiplication model; steel prices time series

1 INTRODUCTION

Steel is one of the main materials commonly used in construction engineering [1]. Due to the huge consumption of steel in construction engineering, in general civil construction engineering, the cost of steel can account for about 30% or more of the material cost, and the fluctuation of its price has a great impact on the project cost [1]. The ability to accurately predict the steel price is of great help to the construction and construction units to reasonably control the cost of the project and even achieve the objectives of quality and construction period. It also plays a significant role in the rational development and allocation of iron ore and related resources, and reducing carbon emissions. Therefore, the research and construction of appropriate prediction model is a scientific and accurate method to realize steel price prediction.

Time series is a set of random variables sorted by time, which is generally the result of observation of a certain potential process according to the given sampling rate in equal intervals [2, 3]. It is a simple and effective method for short-term prediction of time series to establish the corresponding model based on the sequence itself, and monthly steel price data and its essence is a set of time series. At present, steel price prediction models based on time series mainly include the ARIMA model [4-6], exponential smoothing prediction model [7], multiplication or addition model [8], etc.

There is little research on steel price prediction by using multiplication model and addition model in the research of various prediction models and methods based on time series. Guo [8] used the multiplication model to predict the steel bar price in the short term. But the traditional methods were used in the modeling process, and the steel bar price data only included the historical price of steel bar in Baotou city from 2010 to 2012. Therefore, its prediction accuracy and applicability need to be further improved. The multiplication and addition models are mostly used to predict building energy consumption [9], the number of tourists [10, 11], the cigarette sales [12] and GDP data [13].

In this paper, the multiplicative model, the additive model and their combination model are used to predict the steel bar price. Compared with the traditional processing methods [8], the modeling process is improved. According to the characteristics of the time series diagram of steel bar price (Fig. 1) and the analysis of the factors affecting the change of steel bar price, a piecewise function is proposed to fit the long-term trend value T . The separation process of seasonal index S and random item I in time series is also improved. The seasonal product model ARMA (p, q) (P, Q) n was used to predict and fit S and I in the short term. Finally, the reciprocal variance method is applied to give reasonable weight, and the multiplicative and additive models are combined to obtain an improved combination prediction model with higher prediction accuracy. It is conducive to the scientific and reasonable prediction of steel prices by construction enterprises, the management and control of enterprise costs, and then reduce the waste of resources and environmental pollution in the construction process.

2 MODEL INTRODUCTION AND SELECTION DESCRIPTION

The constituent elements of time series mainly include long-term trend T , seasonal variations S , cyclic variation C and irregular variation I , and its time series diagram may show equal width advance or radial pattern with time [14]. The basic models in time series include multiplication model, addition model and hybrid model (both multiplication and addition). Due to the different data types and varying trends, the multiplication model and the addition model have different selection criteria. When the seasonal variation of time series shown in the time series graph is roughly equal, or when the time series graph advances with the same width over time, the addition model should be adopted. When the seasonal variation of time series shown in the time series graph is roughly proportional to the long-term trend of time series, the multiplication model should be applied [14].

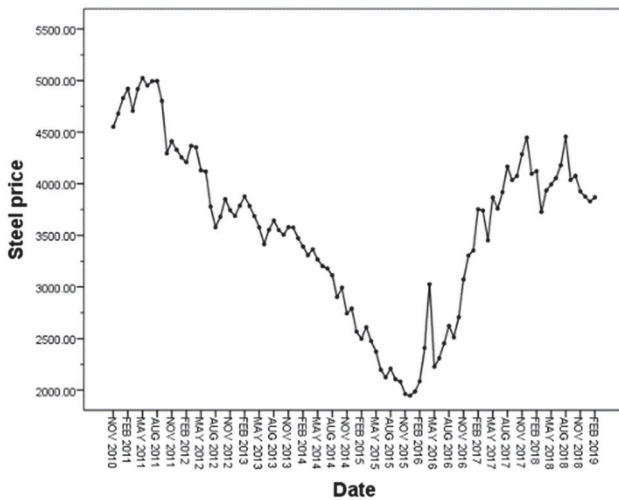


Figure 1 2010.11~2019.2 price change trend of rebar steel (16 mm) in China

The time series diagram of steel bar price used for modeling in this paper is shown in Fig. 1. As can be seen from Fig. 1, the long-term trend of the time series diagram is not only roughly equal to or equal width with the passage of time, but also in a proportional progression stage. Combined with the selection characteristics of multiplication model and addition model, the multiplication model, addition model and multiplication-addition combination model are used to predict the steel price respectively.

3 MODEL BUILDING

3.1 Introduction to the Improved Model

In order to improve the prediction accuracy, the monthly price data of rebar (16 mm) for a total of 100 issues from November 2010 to February 2019 were collected and sorted out from "China metallurgical price information center". Under the influence of market supply and demand, production costs, economic cycles, steel stock price and futures option price changes, there are certain rules for the change of steel bar price. In general, there is usually a long-term trend value (T), and a seasonal periodic value (S), and there may be periodic cyclical fluctuation (C) in a long period of time (usually several years). Moreover, it is also affected by some sudden factors (I) [15].

The basic formulas of multiplication model, addition model and multiplication-addition combination model established in this paper are as follows:

① In the multiplication model, it is assumed that the influences of four factors on the development of the phenomena are interactive, based on the absolute quantity of long-term trend components, and the rest are expressed in terms of ratios, then there is the following expression:

$$Y_m = T \times S \times C \times I \tag{1}$$

② In the addition model, assuming that the influences of the four factors are independent of each other, and each component is expressed as an absolute quantity, then the expression can be described as:

$$Y_a = T + S + C + I \tag{2}$$

③ By applying the reciprocal of variance method and giving reasonable weight to the improved multiplication and addition model, and the improved multiplication-addition combination model is obtained as follows:

$$Y = \alpha_1 Y_m + \alpha_2 Y_a \tag{3}$$

wherein, T is the long-term trend value, S is the seasonal factor, C is the periodic variation factor, I is the irregular factor, α_1 and α_2 represent the weights assigned to the multiplication model and the addition model in the combination model respectively.

3.2 The Data Processing

(1) Smooth processing of data

In the process of time series decomposition, the original time series is firstly smoothed, and the moving weighted average method is used to smoothen the original data.

Assuming the original time series is $y_t, t = 1, 2, \dots, N$.

The number of terms in the moving average method is $n, n < N$. Then, the formula of moving weighted average method can be described as follows:

When n is odd:

$$M_{t+\frac{n-1}{2}} = \frac{W_1 y_t + W_2 y_{t+1} + \dots + W_n y_{t+n-1}}{W_1 + W_2 + \dots + W_n}, \tag{4}$$

$$1 \leq t \leq N - n + 1$$

When n is even:

$$M_{t+\frac{n}{2}} = \frac{W_1 y_t + W_2 y_{t+1} + \dots + W_n y_{t+n-1}}{W_1 + W_2 + \dots + W_n}, \tag{5}$$

$$1 \leq t \leq N - n + 1$$

wherein, M is the average of moving weights, y_t is actual steel price data of the period t , W_1, W_2, \dots, W_n are the respective weights of item 1 through item n .

The determination process of moving item number N and the weight W is as follows:

① The number of moving terms n

Theoretically, when the number of moving average item n is equal to the seasonal periodic item in the original sequence, the sequence weighted by n terms will no longer contain the seasonal component, and the irregular component can be greatly eliminated [16]. In order to determine the seasonal periodic term n in the original time series, SPSS22.0 is used to conduct a difference processing on the original time series. Then, according to the time series diagram after difference, the optimal number of moving items is determined by trying to make weighted movement of terms $n = 3, 4$ and 6 .

② The weight W

Since only the prediction research based on the change law of the data is carried out, the method that the weight value of the observation value closer to the prediction period is larger is selected to complete the determination of the weight W . The weight values under various moving items are shown in Tab. 1.

Table 1 Table of weights of moving weighting method

The average number of moving items	The weight					
	W_1	W_2	W_3	W_4	W_5	W_6
$n = 3$	0.2	0.3	0.5	—	—	—
$n = 4$	0.1	0.2	0.3	0.4	—	—
$n = 6$	0.1	0.1	0.15	0.15	0.25	0.25

After determining the number of moving items n and the weight W , the original time series is processed by moving weighted average with $n = 3, 4$ and 6 . The result is shown in Fig. 2.

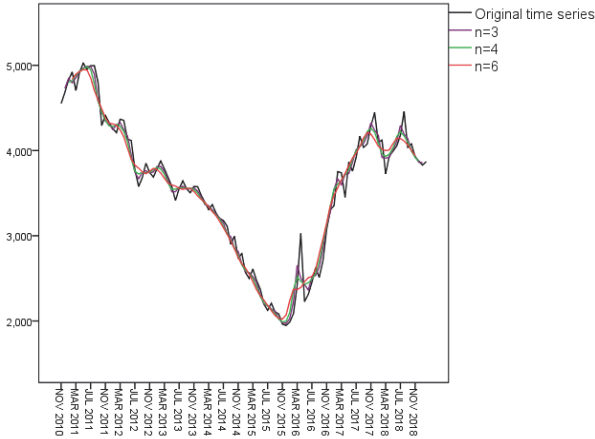


Figure 2 Weighted moving average graph of original time series $n = 3, 4, 6$

As can be seen from Fig. 2, when the average number of moving items is $n = 6$, the obtained time series curve is smoother, which shows that the separation effect of seasonal factors and irregular factors in the original sequence is better, and it also has a better performance on the long-term variation trend of the growth and decline of the original sequence. Therefore, the smooth curve obtained when $n = 6$ is used to fit the long-term trend of steel price fluctuation.

(2) The segmentation of time series

There is a certain functional relationship between the long-term trend value and the increase in time. It can be seen from Fig. 2 that the long-term trend value of steel bar price does not increase or decrease constantly, but will change accordingly with the influence of market environment and changes in national policies in a certain period of time. Thus, according to the six moving weighted average curves of the original time series, and through the observation data of the inflection point, the long-term trend of steel prices change can be divided into four stages (as shown in Fig. 3). Four linear regression equations are established. When solving the long-term trend value of the prediction results, the weighted average results of four linear regression equations are obtained. The four linear regression equations fitted by the smooth curve can be described as follows:

$$T_{t1} = -52.832X_t + 5080.5, R^2 = 0.86 \tag{6}$$

$$T_{t2} = -55.664X_t + 4010.5, R^2 = 0.96 \tag{7}$$

$$T_{t3} = 105.65X_t + 1886.9, R^2 = 0.97 \tag{8}$$

$$T_{t4} = -13.823X_t + 4183.6, R^2 = 0.51 \tag{9}$$

The six items moving weighted average method being used to process the original series, the first three items of the time series are missing when solving the regression Eq. (6), and the last three items of the time series are missing when solving the regression Eq. (9). Therefore, the data of the original time series are used to supplement their own missing items.

According to the principle that the closer the forecast period is, the higher the weight is given, the weights of the calculation results of the above four equations are 0.1, 0.1, 0.4 and 0.4 respectively, and the long-term trend value in the next six periods can be obtained. The calculation formula can be expressed as:

$$T_t = 0.1T_{t1} + 0.1T_{t2} + 0.4T_{t3} + 0.4T_{t4} \tag{10}$$

The calculation results are shown in Tab. 2.

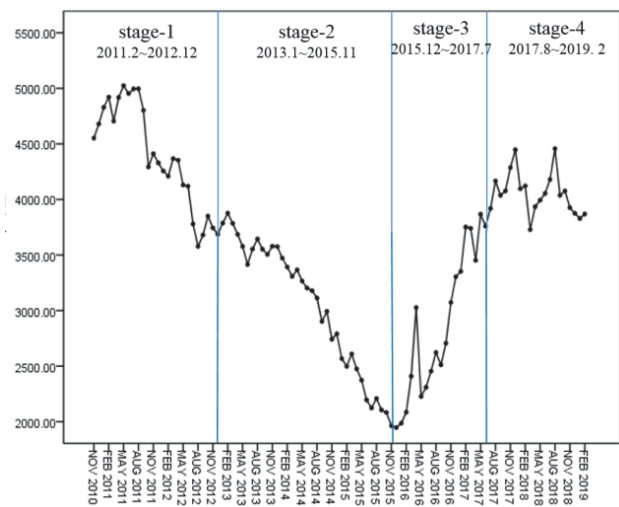


Figure 3 Original time series segment diagram

Table 2 Long-term trend value prediction results

The number of future issues X_t	1	2	3	4	5
Trend values T	3786.99	3812.87	3838.75	3864.63	3890.51

Table 3 Long-term trend value separation table

Year	Month	The real value Y_t / yuan	Long-term trend value T_t / yuan	$Y_t / T_t = S \times C \times I$
2010	11	4552.00	5027.67	0.9054
	12	4680.00	4974.84	0.9407
2011	1	4829.00	4922.00	0.9811
	2	4921.00	4869.17	1.0106
	3	4705.00	4816.34	0.9769
	4	4918.00	4763.51	1.0324
	5	5024.00	4710.68	1.0665
	6	4953.00	4657.84	1.0634
.....				
2019	1	3829.00	3907.14	0.9800
	2	3869.00	3893.32	0.9938

3.3 Improved Model Construction and Validation

(1) The multiplication model

The modeling steps and test results of the improved multiplication model can be summarized as follows:

① The trend value in the original time series is eliminated to obtain $S \times C \times I$. According to Eq. (1),

through the observed value dividing by the long-term trend value with 6 weighted moving averages, $S \times C \times I$ can be obtained, that is $Y_i/T_i = S \times C \times I$. The calculation results are shown in Tab. 3.

② Eliminating the "CI", getting the monthly seasonal index "S", and adjusting it.

In this paper, trend elimination method is used to calculate the seasonal index S_i of each month.

$$\frac{Y_i}{T_i} = \frac{\text{actual value}}{\text{trend value}}$$
 is used to get "SCI";

(2) Averaging the "SCI" coefficient of the same month in each year, so that we can eliminate the influence of cyclic irregular change "CI" and get the seasonal index \bar{S}_i , $i = 1, 2, \dots, 12$;

(3) Using monthly average seasonal index \bar{S}_i divided by the average monthly seasonal index \bar{S} , and S_i is obtained, that is $S_i = \bar{S}_i / \bar{S}$, $i=1, 2, \dots, 12$;

(4) Adjusting S_i , so that $\sum_{i=1}^{12} S_i = 1$.

According to step (2), the final result of S_i is shown in Tab. 4.

Table 4 The calculation table of S_i

month	\bar{S}_i	\bar{S}	$S_i = \bar{S}_i / \bar{S}$	Adjusting S_i
1	0.98779	1.00096	0.9868	0.9864
2	1.00586	1.00096	1.0049	1.0045
3	1.01152	1.00096	1.0105	1.0101
4	1.03850	1.00096	1.0375	1.0371
5	0.99826	1.00096	0.9973	0.9969
6	0.98360	1.00096	0.9827	0.9823
7	0.99280	1.00096	0.9918	0.9914
8	1.01687	1.00096	1.0159	1.0155
9	0.98227	1.00096	0.9813	0.9809
10	0.99313	1.00096	0.9922	0.9918
11	0.99480	1.00096	0.9938	0.9934
12	1.01103	1.00096	1.0101	1.0097

③ Calculating the relative number "CI" of cyclic change and irregular change, and testing the stationaries of random item I_t , the fitting prediction model of I_t can be obtained.

According to formula $CI_t = \frac{I_t}{T_i S_i}$, that is

$$CI_t = \frac{\text{observations}}{\text{Long-term trend value} \times \text{Seasonal index}}$$

get the period-random term CI_t . Since the short-term prediction is carried out in this paper, the influence of cycle factor C can be ignored, the time series of random item I_t can be obtained (as shown in Fig. 4).

The necessary premise for fitting the prediction model of I_t is that the random item I_t is a stationary sequence. Therefore, I_t sequence should be tested for stationarity. At present, there are mainly two kinds of testing methods for the stationarity of time series: subjective testing method (such as autocorrelation function graph [17], partial autocorrelation function graph [18], etc.) and objective testing method (such as unit root ADF test [19], etc.).

SPSS22.0 is used to make diagram of autocorrelation function (ACF) and diagram of partial autocorrelation function (PACF) for the time series of I_t , as shown in Fig. 5 and Fig. 6. At the same time, Eviews10.0 is used to

perform ADF unit root test for I_t sequence including constant term and trend term, and the results are shown in Tab. 5.

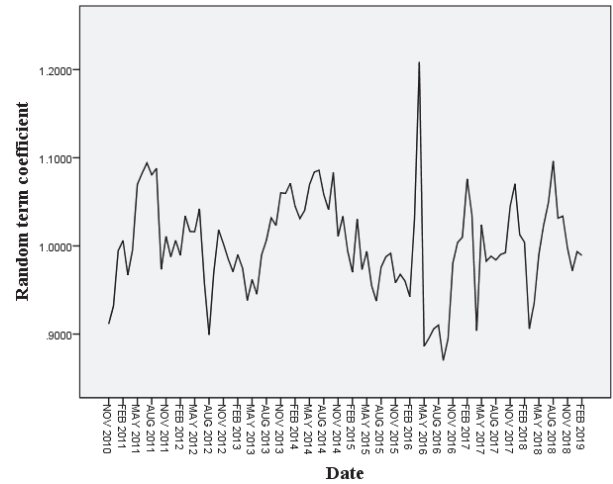


Figure 4 Time series diagram of random item I_t in the multiplication model

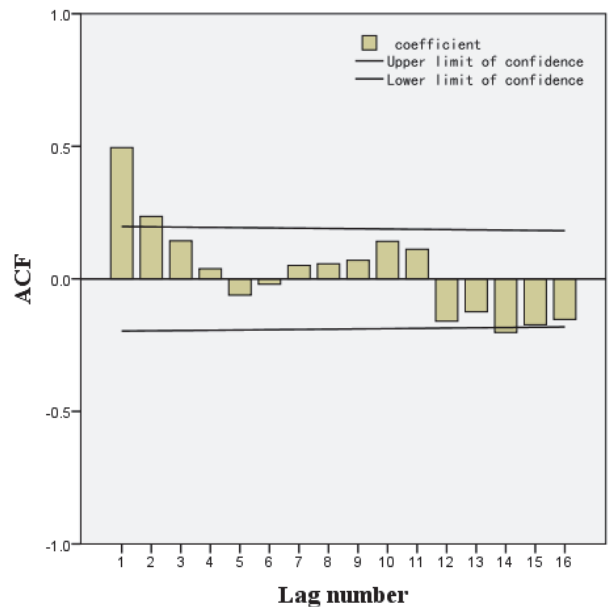


Figure 5 Graph of the autocorrelation function of the random term I_t

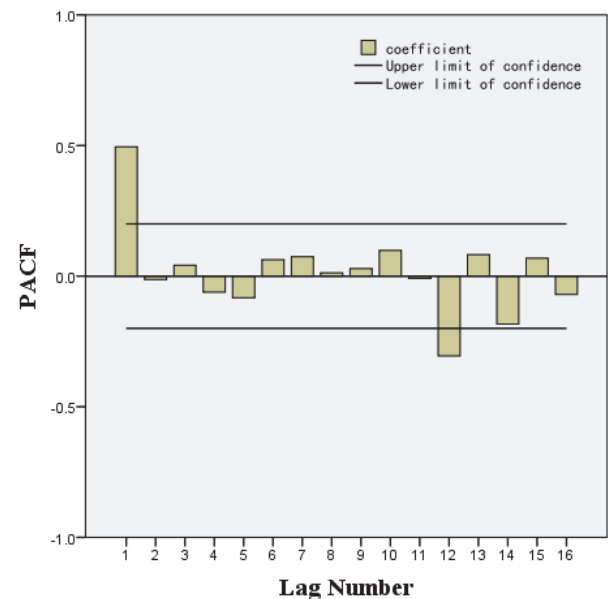


Figure 6 The random term I_t partial autocorrelation function graph

Table 5 ADF test results of random item I_t sequence

		<i>t</i> -Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-5.868186	
Test critical values:	1% level	-4.053392	
	5% level	-3.455842	
	10% level	-3.153710	

Table 6 ARMA model selection of random item I_t

model	Steady <i>R</i> squared	The value of BIC	Ljung-box statistics	The value of <i>P</i>
ARMA (1, 2) (2, 1) ₁₂	0.411	-5.907	13.716	0.319
ARMA (2, 2) (2, 1) ₁₂	0.410	-5.793	13.526	0.260
ARMA (1, 2) (1, 1) ₁₂	0.407	-5.901	12.995	0.448
ARMA (0, 2) (2, 0) ₁₂	0.404	-5.896	13.753	0.391
ARMA (0, 2) (1, 0) ₁₂	0.400	-5.945	13.127	0.517
ARMA (2, 2) (1, 1) ₁₂	0.400	-5.832	12.940	0.373

As can be seen from Fig. 5 and Fig. 6, the autocorrelation function and partial correlation function of the random item I_t are mostly within the confidence limit, and the significance of Box-Ljung statistics is 0. At this time, it can be roughly judged that I_t is a stationary series. In addition, it can be seen from Tab. 5 that the test value of ADF is -5.868186, which is even less than the critical value of -4.053392 when significance level $\alpha = 0.01$. Therefore, the null hypothesis is rejected and it is believed that there is no unit root in this sequence. It shows that after the treatment of $I_t = \frac{Y_t}{T_t S_i}$, the trend term and period term

in the observed value are completely separated, that is, the time series about the random item I_t at this time is determined to be a stationary series. However, at this time,

the seasonal index represented by the formula $I_t = \frac{Y_t}{T_t S_i}$ is

invariable, that is, it is repeated every year, which is not quite consistent with the actual situation. Therefore, it is believed that there are still some seasonal factors in the random item I_t obtained at this time that cannot be completely separated, so the traditional probability distribution model is not adopted for fitting and predicting the random item I_t . It can be judged from Fig. 5 and Fig. 6 that the autocorrelation function and partial correlation function of the random item I_t are trailing. Therefore, ARMA seasonal product model is used to fit and predict the random term. The basic form of ARMA seasonal product model is ARMA (p, q) (p, q)_{*n*}. For the selection of parameters p, q, P and Q in the ARMA seasonal product model, the correlation function diagram of the random item I_t (Fig. 5 and Fig. 6) is firstly determined. For those that cannot be determined directly, the software SPSS22.0 is used to carry out sequential fitting of values, which is determined according to the basic principle that the larger the coefficient R_2 value and the smaller the standardized BIC value are in the model fitting results, the better the model fitting will be. Finally, the model is determined as ARMA (1, 2) (2, 1)₁₂, and Q test was conducted on the residual error of the model. The results showed that Ljung-Box $Q = 13.799, P = 0.315 (P > 0.05)$, and the difference was not statistically significant. In other words, the residual

sequence of ARMA (1, 2) (2, 1)₁₂ was white noise, and the selected model was suitable for prediction. The comparison and selection process of ARMA model is shown in Tab. 6, the final fitting and prediction results of random item I_t are shown in Fig. 7.

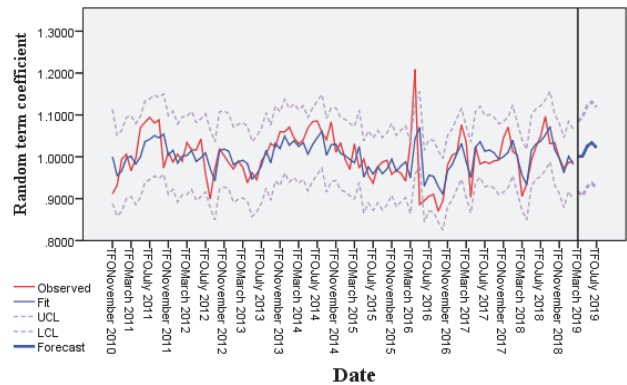


Figure 7 Random item I_t prediction model fitting results

④ The random term coefficient series I_{in} (which is the random term coefficient for the n th month of the n th year, $n = 2010, 2011, \dots, 2019$) fitted by the ARMA model in step (3) is substituted back into the "SCI" sequence in step (1). Then the stationarity of S_i sequence is tested and the model equation of season index S_i of each month is fitted.

In practice, the seasonal index S is not constant every year (each cycle), but will change with the year. Therefore,

in this paper, the formula $S_i = \frac{Y_t}{T_t I_{in}}$ is used to obtain the

time series of the seasonal index S_i of each month, and the optimal fitting equation is obtained according to the time series of S_i . In the step (1), indirectly using the calculation

results of the formula $\frac{Y_t}{T_t} = S \times C \times I$ to fit the model

equation of the months seasonal index S_i , because the results of $\frac{Y_t}{T_t}$ contain not only the season's variation, but

also contain random item I_t . Therefore, only the first use ARMA (1, 2) (2, 1)₁₂ model to fit out random sequence I_{in} , and then the seasonal index sequence S_i (Fig. 8) can be completely separated.

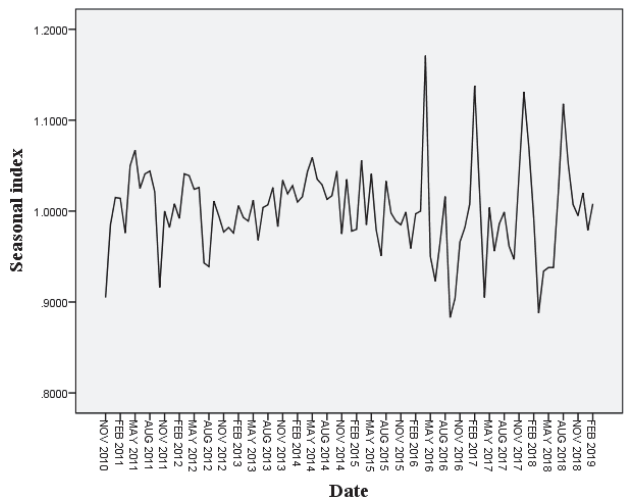


Figure 8 Time series diagram of seasonal index S_i

Before solving the fitting equation of seasonal exponential sequence S_t , the stationarity of the sequence should also be tested. SPSS22.0 was used to make diagram of autocorrelation function (ACF) and diagram of partial autocorrelation function (PACF) for S_t time series, as shown in Fig. 9 and Fig. 10. At the same time, Eviews10.0 is applied to perform ADF unit root test for S_t sequence including constant term and trend term, and the results are shown in Tab. 7.

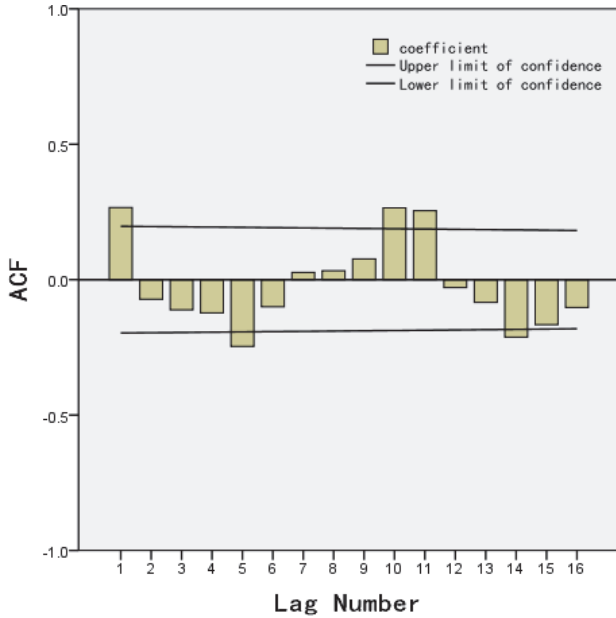


Figure 9 Autocorrelation function graph of S_t sequence

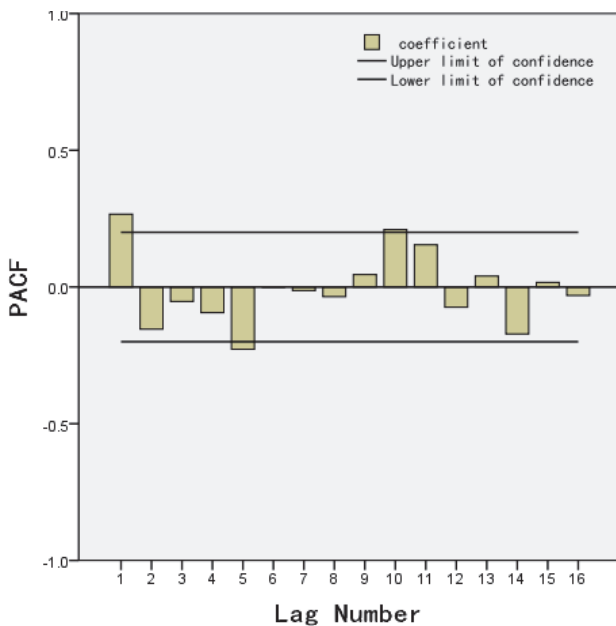


Figure 10 Partial autocorrelation function graph of S_t sequence

It can be seen from Fig. 9 and Fig. 10 that the autocorrelation function and partial correlation function of seasonal index S_t sequence are mostly within the confidence limit, and the significance of box-Ljung statistics is 0. At this time, it can be roughly judged that S_t sequence is a stationary sequence. In addition, it can be seen from Tab. 7 that the test value of ADF is -7.671418 , even less than the critical value of -4.053392 when the significance level α is 0.01. Therefore, the original

hypothesis is rejected and it is considered that there is no unit root in the sequence. In other words, it is determined about the seasonal index S_t is a stationary series at this time.

Table 7 ADF test results of seasonal exponential S_t sequence

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-7.671418	0.0000
Test critical values:	1% level	-4.053392	
	5% level	-3.455842	
	10% level	-3.153710	

The season product model of ARMA (p, q) $(p, q)_n$ is used to fit and predict S_t series. According to the information from Fig. 9 and Fig. 10, SPSS22.0 is used for value comparison, ultimately ARMA $(1, 2)$ $(1, 1)_{12}$ is determined as the optimal seasonal product model of S_t sequence. Further Q test is carried out on the residual error of the model. The results show that Ljung-box $Q = 20.662$, $P = 0.08$ ($P > 0.05$), the difference is not statistically significant, that is, the residual sequence of ARMA $(1, 2)$ $(1, 1)_{12}$ was white noise, and the selected model was suitable for prediction. The final fitting and prediction result of seasonal index S_t sequence is shown in Fig. 11.

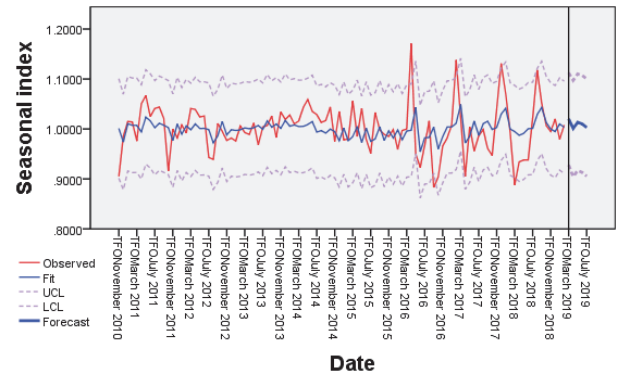


Figure 11 The fitting and prediction results of S_t series of seasonal index

⑤ The obtained T_t, S_{in} and I_{in} are combined and multiplied to obtain the final multiplication model, and the model is

$$Y_m = T_t \times S_{in} \times I_{in} \tag{11}$$

According to the calculation and prediction results in step (1) to step (4), the multiplication model (11) is used to predict the price of rebar (16 mm) for a total of 5 phases from march 2019 to July 2019, and test the prediction accuracy of the model. The prediction results are shown in Tab. 8.

As can be seen from Tab. 8, the minimum relative error between the predicted price of steel reinforcement and the real price in the fifth period is 0.24%, the maximum is -7.75% , and $MAPE = 2.62\%$, and the accuracy of the model is good.

(2) Additive model

The modeling steps of the addition model are similar to those of the multiplication model, and the steps and results are described as follows:

- ① the trend value in the original time series is eliminated to get the expression " $S + C + I$ ".
- ② the "CI" is eliminated to obtain the monthly seasonal variation "S" and adjust it.

In this paper, the trend elimination method is used to calculate the seasonal index S_i of each month.

1) The formula " $Y_t - T_t = \text{actual value} - \text{trend value}$ " is used to obtain " $S + C + I$ ";

2) The average value of " $S + C + I$ " in the same month of each year is calculated to eliminate the influence of the cyclic and irregular variation value " $C + I$ " and get the monthly average seasonal variation $S_i, i=1, 2, \dots, 12$;

3) The monthly average seasonal variation \bar{S}_i is used to subtract the average monthly seasonal variation \bar{S} and obtain S_i , that is $S_i = \bar{S}_i - \bar{S}, i = 1, 2, \dots, 12$;

4) S_i is adjusted so that $\sum_{i=1}^{12} S_i = 0$.

③ The sum of cyclic variation and irregular variation values " $C+I$ " is calculated and the stationarity of the random term I_t is tested, so the fitting model of I_t is obtained;

④ The random term coefficient I_{in} fitted by the ARMA model in step (3) is substituted back into the " $S + C + I$ " sequence in step (1) and the model equation of monthly seasonal variation S_i is fitted;

⑤ By using the obtained T_t, S_{in} and I_t being combined and added to obtain the final addition model, which can be described as follows:

$$Y_a = T_t + S_{in} + I_{in} \tag{12}$$

According to the calculation and prediction results in the previous steps (1) ~ (4), the additive model (12) is used to predict the price of rebar (16 mm) for a total of 5 phases from march 2019 to July 2019 to test the prediction accuracy of the model. The prediction results are shown in Tab. 8.

As can be seen from Tab. 8, the minimum relative error between the predicted price of steel reinforcement in the fifth period and the real price is -1.32%, the maximum is -3.75%, $MAPE = 2.36\%$, and the prediction accuracy is better.

(3) Combined model

The purpose of combining the models is to give full play to the advantages of the two models and further improve the prediction accuracy. The reciprocal variance method (error square sum method) is used to determine the weights of the multiplication model and the addition model in the combined model. The inverse variance method is based on the sum of error squares of each model, the smaller the sum of error squares of the model is, the higher the prediction accuracy of the model is, and the greater the weight should be given to it in the combined model, otherwise the smaller the weight is. The formula for the reciprocal of variance is described as follows:

$$\omega_j = \frac{e_j^{-1}}{\sum_{j=1}^m e_j^{-1}}, \sum_{j=1}^m e_j^{-1} = 1, j = 1, 2, \dots, m \tag{13}$$

wherein, ω_j is the weight of the j -th model.

$$\sum_{j=1}^m \omega_j = 1, j = 1, 2, \dots, m, e_j = \sum_{t=1}^n (x_t - x_{jt})^2 \tag{14}$$

wherein, e_j is the error squared sum of the j -th model; x_t is the observation at time t ; x_{jt} is the predicted value of model j at time t .

According to the formula of reciprocal variance above, the weight of the multiplication model is 0.3169, and that of the addition model is 0.6831. Therefore, the improved combination model can be expressed as:

$$Y = 0.3169Y_m + 0.6831Y_a \tag{15}$$

The prediction comparison results of the combined model can be calculated from Eqs. (13), (14) and (15), as shown in Tab. 8. It can be seen from Tab. 10 that the improved multiplication and addition prediction model has higher prediction accuracy, while the forecast effect and precision of combination forecast model is further improved, which shows that the improved combination forecast model can be used for the short-term forecasting of steel bar market price, and the prediction accuracy is higher than that of a single time series model.

Table 8 Prediction results of multiplication model, addition model, and combination model

Forecast time	The actual price	Combined model Y			Multiplication model Y_m			Additive model Y_a		
		Predictive value	The relative error %	MAPE	Predictive value	The relative error %	MAPE	Predictive value	The relative error %	MAPE
2019.03	3953.00	3878.962	-1.87%	2.29%	3863.49	-2.26%	2.62%	3886.14	-1.69%	2.36%
2019.04	4145.00	3937.113	-5.02%		3823.93	-7.75%		3989.62	-3.75%	
2019.05	4069.00	3949.596	-2.93%		3991.73	-1.90%		3930.05	-3.41%	
2019.06	4002.00	3969.125	-0.82%		4040.29	0.96%		3936.11	-1.65%	
2019.07	3978.00	3945.073	-0.83%		3987.63	0.24%		3925.33	-1.32%	

4 CONCLUSIONS

In this paper, the following new methods and viewpoints are obtained from the research of the improved combined prediction model of steel price based on time series:

(1) Compared to single multiplication or addition prediction models, combined prediction models can better reflect and predict the changing trends of steel bar prices in time series with multiple changing trends.

(2) Based on the characteristics of long-term trend values, the long-term trend values of steel bar prices are segmented and weighted according to the principle of "the closer the time, the greater the weight". The results have shown that for time series data with high variability, fitting long-term trend values using segmented functions has better predictive performance than using a single curve function. In addition, this article should have more quantitative judgment basis and methods when determining the segmentation cutoff points and weights, and further improvements will be made in the future.

(3) In reality, the seasonal term S has the characteristics of volatility and wide amplitude. In this paper, the traditional method firstly is used to obtain the seasonal term S in a cycle, fit the prediction of the random term I_t is fitted, then the prediction of seasonal term S_{in} with varying over time is fitted by retrogression method. This improvement can greatly improve the prediction accuracy of the time series prediction model. In addition, the selection and improvement of the seasonal term S_{in} , and random term I_t is also a direction for future research.

Acknowledgements

The authors would like to thank the Sichuan Youth Software Innovation Project [Grant No.: 2021012] for its support.

5 REFERENCES

- [1] Zieliński, A., Wersta, R., & Sroka, M. (2021). Analysis of the precipitation process of secondary phases after long-term ageing of the S304H steel. *Bulletin of the Polish Academy of Sciences. Technical Sciences*, 69(5). <https://doi.org/10.24425/bpasts.2021.137520>
- [2] Heydari, M., Ghadim, H. B., Rashidi, M., & Noori, M. (2020). Application of holt-winters time series models for predicting climatic parameters (case study: Robot Garah-Bil Station, Iran). *Polish Journal of Environmental Studies*, 29(1), 617-627. <https://doi.org/10.15244/pjoes/100496>
- [3] Adedotun, A. F. (2022). Hybrid Neural Network Prediction for Time Series Analysis of COVID-19 Cases in Nigeria. *Journal of Intelligent Management Decision*, 1(1), 46-55. <https://doi.org/10.56578/jimd010106>
- [4] AAdli, K. (2020). Forecasting steel price susing ARIMAX model: A case study of Turkey. *The International Journal of Business Management and Technology*, 4(5), 2581.
- [5] Yang, C., Wang, D. M., & Wang, H. L. (2019). Analysis and forecast of steel composite price index based on ARIMA model. *Industrial & Science Tribune*, 18(17), 43.
- [6] Zhou, Y. B. & Wang, M. (2018). A preliminary analysis of the factors affecting the price of steel. *Science and Technology & Innovation*, 20(68).
- [7] Zhang, Q. (2013). Price prediction of building materials based on cubic exponential smoothing method. *Railway Engineering Technology and Economy*, 28(1), 48.
- [8] Zhu, R. (2019). *Analysis of price fluctuation of construction materials*. M. A. thesis, North China University of Water Resources and Electric Power, China.
- [9] Deb, C., Zhang, F., Yang, J., Lee, S. E., & Shah, K. W. (2017). A review on time series forecasting techniques for building energy consumption. *Renewable and Sustainable Energy Reviews*, 74, 902-924. <https://doi.org/10.1016/j.rser.2017.02.085>
- [10] Kong, C. L. (2018). Forecast of tourist Numbers in Hainan based on multiplication season model. *Statics and Management*, 2(62).
- [11] Ivanovski, Z., Milenkovski, A., & Narasanov, Z. (2018). Timeseries forecasting using a moving average model for extrapolation of number of tourist. *UTMS Journal of Economics*, 9(2), 121-132.
- [12] Qian, M. J. & Yzern, L. (2020). Method on monthly railway passenger traffic forecast based on season alde composition and SARIMA-GARCH model. *Journal of the China Railway Society*, 42(6), 25-34.
- [13] Luo, Z. D. (2016). The quarterly adjust model and forecast analysis of China' GDP. *Statistic & Decision*, 20(92).
- [14] Gonzalez-Mulé, E., Kim, M. M., & Ryu, J. W. (2021). A meta-analytic test of multiplicative and additive models of jobdemands, resources, and stress. *Journal of Applied Psychology*, 106(9), 1391. <https://doi.org/10.1037/apl0000840>
- [15] Chawsheen, T. A. & Broom, M. (2017). Seasonal time-series modeling and forecasting of monthly mean temperature for decision making in the Kurdistan Region of Iraq. *Journal of Statistical Theory and Practice*, 11, 604-633. <https://doi.org/10.1080/15598608.2017.1292484>
- [16] Lu, L. (2007). *Research on decomposition of timeseries addition model*. M. A. thesis, Southwest Jiaotong University, China.
- [17] Aich, U. & Banerjee, S. (2017). Characterizing topography of EDM generated surface by timeseries and autocorrelation function. *Tribology International*, 111, 73-90. <https://doi.org/10.1016/j.triboint.2017.02.016>
- [18] Yeh, H. F. & Hsu, H. L. (2019). Stochastic model for drought forecasting in the Southern Taiwan Basin. *Water*, 11(10), 2041. <https://doi.org/10.3390/w11102041>
- [19] Aylar, E., Smeekes, S., & Westerlund, J. (2019). Lag truncation and the local asymptotic distribution of the ADF test for aunitroot. *Statistical Papers*, 60, 2109-2118. <https://doi.org/10.1007/s00362-017-0911-y>

Contact information:

Qingyun ZHANG

403 Geological Team of SiChuan Bureau of Geology and Mineral Exploration and Development,
Leshan 614200, China
E-mail: Zhangqy0101@163.com

Dingxia LIU

403 Geological Team of SiChuan Bureau of Geology and Mineral Exploration and Development,
Leshan 614200, China
E-mail: Liudx1201@126.com

Xiaojiang WANG

403 Geological Team of SiChuan Bureau of Geology and Mineral Exploration and Development,
Leshan 614200, China
E-mail: 18505956@qq.com

Zhiqing YE

403 Geological Team of SiChuan Bureau of Geology and Mineral Exploration and Development,
Leshan 614200, China
E-mail: 275154645@qq.com

Hanxi JIANG

Engineering Cost Center of Liangshan Transportation Construction,
Xichang 615000, China
College of Architecture and Urban-Rural Planning,
Sichuan Agricultural University,
Chengdu 611830, China
E-mail: 2219151450@qq.com

Wei WEI

(Corresponding author)
College of Architecture and Urban-Rural Planning,
Sichuan Agricultural University,
Chengdu 611830, China
E-mail: jpwflz@sicau.edu.cn