

# Research on Prediction of Architectural Engineering Cost based on the Time Series Method

Qian Zhang

Construction Department, State Grid Baoding Electric Power Supply Company, Baoding, China

34063533@qq.com

**Keywords:** Architectural engineering; Project cost; Time series; cost prediction

**Abstract.** Cost plays a very important role in architectural engineering, and the methods to predict the cost rapidly and accurately are needed. On the basis of the valuation mode of quantities bill, this paper proposed the time series method for predicting project cost. According to the theory of test, modeling and prediction, the paper carried out time series analysis, taking steel price for example. It withdrew the steel price tendency item, then conducted modeling using the price residual item, finally established the model of steel price. In line with the model, the steel price of the first quarter in 2014 was forecast. By comparing with the real price, it shows that the error was in 5 percent, which proves the model is correct and the time series method based on the valuation mode of quantities bill can be adopted widely to predict manufacturing cost.

## Introduction

Project cost is one of the very important work of construction project management, which runs through the whole process of project construction. It displays in the various stages of economic file such as feasibility study estimated budget for the design, construction budget, bidding tender and bidding, project completion settlement. Construction cost is not only related to the building size, building structure, building function and so on, but also depend on the building site and the environment, and these factors have increased the difficulty of building project cost [1].

Bill of quantities (BOQ) valuation model is the most commonly used project cost method in domestic and foreign. Although it needs a preparation of complex and computationally intensive, without considering the impact of time and other shortcomings, but it is consistent with the entire process of project control theory, relatively speaking, and can more accurately construction cost calculated [2]. This paper thought that only by adopting a fast and accurate method for improving the efficiency of bill of quantities, taking into account the impact of time on the construction project cost to overcome shortcomings in the bill of quantities in construction engineering cost. Based on this, the author used the fast prediction function and the characteristics analysis of time series, considered the steel price of construction works supplies as the object, and implemented time series analysis, which achieved good forecast values for construction cost quick and accurate prediction and provided a new way of thinking.

## Time series prediction theory

### Time series test.

(1) Stationary test. That is to test whether the time series is a stationary time series, which includes timing chart test and autocorrelation coefficient test. For not obvious trends timing chart, it needs the autocorrelation coefficient test to make further judgments. If the autocorrelation coefficient on the graph shows a triangular symmetry, it is regarded as the non-stationary time series. If the autocorrelation coefficients on the graph gradually decays to zero, it is regarded as stationary time series. If the time series is not smooth, you need first to sequence extraction cycle trend item or items, so the sequence becomes stable, then for subsequent analysis [3].

(2) Pure randomness test. That is judged by  $L_B$  statistics which is defined as follows:

$$L_B = n(n+2) \sum_{k=1}^m \left( \frac{\hat{\rho}_k^2}{n-k} \right) \quad (1)$$

Where  $n$  represents the series observation period,  $m$  represents certain delay period.

$L_B$  statistics approximately obey the chi-square distribution whose freedom degree is  $m$ . When  $L_B$  statistics is above the sub-sites of  $\chi_{1-a}^2(m)$ , or the  $P$  value of the statistics is below  $a$ , the series is defined as the non-white noise series. Otherwise, it is defined as the pure random series.

### Time series modeling.

The time series which satisfies the stationary test and pure randomness tests can establish the time series analysis model.

(1) Model identification. Time series analysis stochastic model has AR, MA and ARMA three categories, which model the specific choice should be judged according to the autocorrelation coefficient and partial correlation coefficient time series.

(2) Parameter estimation. As the special column of ARMA model, the parameter estimation method of AR and MA model are the same to ARMA model, which shown as follows.

For ARMA(p, q) model, it has

$$x_t = \sum_{i=1}^p \varphi_i x_{t-i} - \sum_{j=1}^q \theta_j a_{t-j} + a_t \quad (2)$$

Where  $\varphi$  satisfies the relationship of Eq.(3).

$$R_A = R_{B\varphi} \quad (3)$$

$$\text{Where } \varphi = \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_p \end{bmatrix}, R_A = \begin{bmatrix} R_{q+1} \\ R_{q+2} \\ \vdots \\ R_{q+p} \end{bmatrix}, R_B = \begin{bmatrix} R_q & R_{q-1} & \cdots & R_{q-p+1} \\ R_{q+1} & R_q & \cdots & R_{q-p+2} \\ \vdots & \vdots & \ddots & \vdots \\ R_{q+p-1} & R_{q+p-2} & \cdots & R_q \end{bmatrix}.$$

In which,  $R_k$  represents the autocovariance function of the series,  $R_k = E[x_t x_{t-k}]$ .

In the ARMA(p, q) model, set

$$z_t = x_t - \sum_{i=1}^p \varphi_i x_{t-i} = - \sum_{i=0}^p \varphi_i x_{t-i}, (\varphi_0 = -1) \quad (4)$$

Similarly,

$$z_t = a_t - \sum_{j=1}^q \theta_j a_{t-j} = \sum_{j=0}^q \theta_j a_{t-j}, (\theta_0 = -1) \quad (5)$$

According to the spectral density function formula, we can get:

$$\sum_{k=0}^q R_{z,k} \left( \frac{1}{\eta_j} \right)^k = 0 \quad (6)$$

The  $B$  operator polynomial  $\theta(B)$  of average sliding part satisfies the Eq.(7).

$$\theta(B) = \prod_{j=1}^q (1 - \eta_j B) = 1 - \sum_{j=1}^q \theta_j B^j \quad (7)$$

$\theta_j$  can be obtained through comparing the same power coefficient of  $B$  operator.

(3) Model test. In order to analyze the model meets the statistical law, the need for statistical testing, the test method is to construct statistics. The degree of high order model and the low order model are respectively defined as  $p_h$  and  $p_l$ . Establish the following statistics:

$$F = \frac{(\Omega_{p_l} - \Omega_{p_h})(N - \Omega_{p_h})}{\Omega_{p_h}(p_h - p_l)} \quad (8)$$

Where  $\Omega$  represents the residual sum of squares,  $\Omega = (V^T V)$ .

$$F \sim F(p_h - p_l, N - p_h) \quad (9)$$

According to Eq.(8), if  $F$  is above  $F(p_h - p_l, N - p_h)$ , it indicates that compared with  $p_l$ ,  $p_h$  has significant differences. The low order model is not applicable, also rose order modeling; on the contrary, is considered low-end model is applicable.

#### Time series prediction.

For the ARMA (p, q) model, best predicted values are calculated from the Eq.(10) returns, calculated by Eq.(11) predictive value of 95% confidence interval.

$$\begin{cases} \bar{x}_t(1) = E[x_{t+1}] = \varphi_1 x_t + \varphi_2 x_{t-1} + \cdots + \varphi_p x_{t+1-p} - \theta_1 a_t - \theta_2 a_{t-1} - \cdots - \theta_q a_{t+1-q} \\ \bar{x}_t(2) = E[x_{t+2}] = \varphi_1 \bar{x}_t(1) + \varphi_2 x_{t-1} + \cdots + \varphi_p x_{t+2-p} - \theta_2 a_t - \cdots - \theta_q a_{t+2-q} \\ \bar{x}_t(q) = E[x_{t+q}] = \varphi_1 \bar{x}_t(q-1) + \varphi_2 \bar{x}_t(q-2) + \cdots + \varphi_p \bar{x}_t(q-p) - \theta_q a_t \\ \bar{x}_t(l) = E[x_{t+l}] = \sum_{i=1}^p \varphi_i \bar{x}_t(l-i), (l > q) \end{cases} \quad (10)$$

$$\left( y(t+l) + \bar{x}_t(l) - 1.96\sqrt{\text{var}[e_t(l)]}, y(t+l) + \bar{x}_t(l) + 1.96\sqrt{\text{var}[e_t(l)]} \right) \quad (11)$$

#### Case study

Steel as a building project is important supplies, accurately predict the price of steel is essential for the construction project cost forecasting. The author at market prices of certain steel products as an example, steel prices forecast analysis to provide initial information for the construction cost prediction.

According to the time series can be drawn steel prices, steel prices in the non-stationary time series, we need to extract their trend items. Using the least squares method to extract the steel price trend term time series, and with Matlab fast calculations. After calculated, the trend function can be obtained, which is shown as follows:

$$y(t) = 3.865.833 - 49.3t + 3.7t^2$$

Steel prices residual term time series trend item after extraction, after test, the residuals belongs to stationary time series. Perform time series modeling residuals as described above, draw residual term time series model:

$$x_t = -0.069x_{t-1} + 0.259a_{t-1} + a_t$$

The steel prices time series models can be obtained by adding the trend term function and the residual term time series model:

$$x_t = 3865.833 - 49.3t + 3.7t^2 - 0.069x_{t-1} + 0.259a_{t-1} + a_t$$

Steel prices time series values are shown in Table 1, for the first quarter of 2014, steel prices to predict and analyze the actual price comparison on the market, predict and analyze the relevant data in Table 2.

Table 1 Steel prices time series

| Date    | Price [RMB/t] | Trend term series | Autoregression residuals term series | White noise series | Error series |
|---------|---------------|-------------------|--------------------------------------|--------------------|--------------|
| 2012.01 | 3940          | 3820.23           | 119.77                               | 119.77             | -            |
| 2012.02 | 3820          | 3782.03           | 46.23                                | 15.21              | -15.21       |
| 2012.03 | 3600          | 3751.23           | -148.61                              | -152.55            | 152.55       |
| 2012.04 | 3570          | 3727.83           | -168.27                              | -128.76            | 128.76       |
| 2012.05 | 3645          | 3711.83           | -77.72                               | -44.38             | 44.38        |
| 2012.06 | 3720          | 3703.23           | 12.16                                | 23.65              | -23.65       |
| 2012.07 | 3920          | 3702.03           | 219.12                               | 213.00             | -213.00      |
| 2012.08 | 3630          | 3708.23           | -63.19                               | -118.36            | 118.36       |
| 2012.09 | 3655          | 3721.83           | -72.23                               | -41.48             | 41.48        |
| 2012.10 | 3710          | 3742.83           | -37.44                               | -26.68             | 26.68        |
| 2012.11 | 3900          | 3771.23           | 126.50                               | 133.41             | -133.41      |
| 2012.12 | 3950          | 3807.03           | 151.85                               | 117.30             | -117.30      |
| 2013.01 | 4050          | 3850.23           | 209.63                               | 179.25             | -179.25      |
| 2013.02 | 3820          | 3900.83           | -67.05                               | -113.48            | 113.48       |
| 2013.03 | 3830          | 3958.83           | -134.41                              | -105.02            | 105.02       |
| 2013.04 | 4030          | 4024.23           | -3.12                                | 24.08              | -24.08       |
| 2013.05 | 4160          | 4097.03           | 63.37                                | 57.13              | -57.13       |
| 2013.06 | 4100          | 4177.23           | -72.89                               | -87.69             | 87.69        |
| 2013.07 | 4020          | 4264.83           | -250.16                              | -227.45            | 227.45       |
| 2013.08 | 4350          | 4359.83           | -26.73                               | 32.184             | -32.184      |
| 2013.09 | 4570          | 4462.23           | 107.09                               | 98.75              | -98.75       |
| 2013.10 | 4510          | 4572.03           | -54.60                               | -80.17             | 80.17        |
| 2013.11 | 4730          | 46889.23          | 36.49                                | 57.25              | -57.25       |
| 2013.12 | 4890          | 4813.83           | 78.98                                | 64.15              | -64.15       |

Table 2 Results analysis

| Date    | Actual price [RMB/t] | Prediction price [RMB/t] | Relative error [%] | 95% confidence interval |
|---------|----------------------|--------------------------|--------------------|-------------------------|
| 2014.01 | 4865                 | 4957.00                  | 1.89               | (4957.00 ± 211.67)      |
| 2014.02 | 4960                 | 5067.26                  | 2.16               | (5067.26 ± 219.07)      |
| 2014.03 | 4990                 | 5208.52                  | 4.39               | (5208.82 ± 220.08)      |

As can be seen from steel prices time series analysis and prediction, steel prices remain steady upward trend, and keeps the items consistent with the trend, but the trend term trend somewhat fitting too, in the latter part of the fitting, the trend can make the appropriate entry adjustment. Residuals term forecasting of trends constantly revised, complements the trend term fit, which is a time series and regression analysis mainly reflected favorable. The time series prediction analysis drawn steel prices in the first quarter of 2014 and close to the actual price, prediction relative error does not exceed 5%, the actual price in the forecast 95 percent confidence interval, predicted values are accurate. Time series forecasting steel prices is feasible.

Construction works in a variety of building supplies, cement, gravel, stone, wood, etc., you can get the price predicted value engineering during the specified time by time series forecasting. The predictive value of various supplies for engineering bill pricing mode, a time factor in building construction cost can be obtained.

## Conclusions

This paper considered the defects of BOQ valuation model of project cost prediction and the characteristics of time series prediction, and proposed construction project cost fast forecasting methods based on time series, and summarized the following conclusions: (1) The operation of time series prediction is simple, has faster computing, which can reduce the construction cost calculation time and improve the cost efficiency calculation; (2) The time series prediction considers the effect of time on the cost to overcome the deficiencies in the past for building static calculation of the project cost; (3) Time series prediction steel prices to maintain a prediction error less than 5%, and the actual values are falling into the predictive value of the 95% confidence intervals, prediction is accurate and reliable; (4) Construction cost method using time-series forecasting, to ensure the accuracy of cost forecasting calculations in favor of the construction project cost control and management.

## References

- [1] Yongming Wang, Junzhong Gu, Zili Zhou, Zhijin Wang. Diarrhoea outpatient visits prediction based on time series decomposition and multi-local predictor fusion. *Knowledge-Based Systems*, Vol. 88 (2015), p. 12-23.
- [2] Hon-lun Yip, Hongqin Fan, Yat-hung Chiang. Predicting the maintenance cost of construction equipment: Comparison between general regression neural network and Box-Jenkins time series models. *Automation in Construction*, Vol. 38 (2014), p. 30-38.
- [3] Yangkyun Kim, Amund Bruland. A study on the establishment of Tunnel Contour Quality Index considering construction cost. *Tunnelling and Underground Space Technology*, Vol. 50 (2015), p. 218-225.