S1-1

BIM-BASED TIME SERIES COST MODEL FOR BUILDING PROJECTS: FOCUSING ON MATERIAL PRICES

Sungjoo Hwang[1], Moonseo Park[2], Hyun-Soo Lee[3] and Hyunsoo Kim[4]

1 PhD Student, Seoul National University, Seoul, Korea
2 Associate Professor, Seoul National University, Seoul, Korea
3 Professor, Seoul National University, Seoul, Korea
4 PhD Student, Seoul National University, Seoul, Korea
Correspond to nkkt14@snu.ac.kr

ABSTRACT: As large-scale building projects have recently increased for the residential, commercial and office facilities, construction costs for these projects have become a matter of great concern, due to their significant construction cost implications, as well as unpredictable market conditions and fluctuations in the rate of inflation during the projects' long-term construction periods. In particular, recent volatile fluctuations of construction material prices fueled such problems as cost forecasting. This research develops a time series model using the Box-Jenkins approach and material price time series data in Korea in order to forecast trends in the unit prices of required materials. Building information modeling (BIM) approaches are also used to analyze injection times of construction resources and to conduct quantity take-off so that total material prices can be forecast. To determine an optimal time series model for forecasting price trends, comparative analysis of predictability of tentative autoregressive integrated moving average (ARIMA) models is conducted. The proposed BIM-based time series forecasting model can help to deal with sudden changes in economic conditions by estimating material prices that correspond to resource injection times.

Keywords: Material Prices, Time series model, ARIMA model, BIM, Cost estimating

1. INTRODUCTION

Large-scale construction projects have recently increased for residential, commercial and office facilities worldwide. Examples of this trend are numerous high-rise buildings over 100 stories being constructed or planned as urban landmarks. In these circumstances, construction costs for large-scale building projects have been a matter of much concern due to their significant cost implications and frequent design changes during long-term construction periods. Many factors may influence the final project costs over the long time span from project start-up to completion of construction [1]. These volatile factors, such as prices of resources injected into the overall project phase, can lead to overestimations or underestimations of total project costs because construction resource prices vary due to changes in demand, market conditions, and macroeconomic conditions [2]. Recent volatile fluctuations of construction material prices fueled such problems as cost forecasting. For example, structural steel prices tripled from 2001 to 2008 in Korea. As material costs are approximately one-fourth of the total project costs, the prediction of material prices can improve the accuracy of cost estimates.

This research, therefore, develops time series forecasting model to help understand past material price trends and accurately predict construction material unit price trends. Also, this research suggests building information model (BIM)-based time series cost forecasting model framework to better reflect economic fluctuations and design changes by estimating time periodic material costs as well as by analyzing changes of resources inputs.

2. PRELIMINARY RESEARCH

2.1 Literature Review

To address cost escalation factors for construction projects and accurately predict costs, many studies have focused on construction cost estimation. Trost and Oberlender [3] presented a mathematical model for evaluating the accuracy of early estimates using factor analysis and multivariate regression analysis. Shaheen et al. [4] proposed an alternative approach to cost range estimating by Monte Carlo simulation. Touran [5] proposed a probabilistic model for the calculation of project cost contingency by considering the expected cost changes. Sonmez [6] developed an integrated approach for conceptual cost estimating including the advantages of parametric and probabilistic estimating techniques.

Although the above-mentioned models are useful for addressing cost escalation factors and preliminary estimate in the early design phase, some restrictions exist for time-varying variables and for reflecting different time lags between variables. In reality, much of the time-related data are dependent or has an autocorrelation [7]. To solve these problems, Issa [8] applied artificial neural networks to help contractors track past prices and conditions affecting these prices and then predict prices. Ashuri and Lu [9] developed a time series model that determined trends based on past values and corresponding errors and then provided a more accurate prediction of a construction cost index. Akintoye et al. [10] identified leading indicators of construction cost oscillations in the United Kingdom using time series approach. Lu and Abourizk [7] studied the Box-Jenkins approach that can be applied to the capital planning of infrastructure systems. Ng et al. [11] outlined the procedures for integrating the regression analysis and time series models (ARIMA) to forecast tender price index for Hong Kong construction projects. Fellows [12] provided reliable forecasts of tender prices, building costs, and the effects of inflation on building projects using a time series model.

The time series models presented above provide a systematic and time-related approach to forecast trends. That is, it is possible to make useful projections based on historical patterns [13]. However, such research predicting total costs or total prices for construction projects have limitations in that they include detailed information required by construction managers, even though it is useful for owners and bidders. Therefore, this research applies a time series model to predict the material unit cost in order to estimate total construction costs at a more detailed level. After this, this BIM-based time series model can calculate stocked-time periodic future material prices by conducting monthly quantity take-off and analyzing material injection time.

2.2 Time Series Analysis

Time series data sets are a sequence of data points, measured typically at successive times and spaced at uniform time intervals. A time series model determines trends based on past values and corresponding errors. Since it only requires the historical data of the forecast variable itself, it is widely used to develop predictive models [7]. Among these models, the Box-Jenkins approach, a procedure suggested by Box and Jenkins (1994) for carrying out stochastic time series modeling, is useful for univariate time series forecasting. The Box-Jenkins approach is also called the autoregressive integrated moving average (ARIMA) model, which considers the underlying trends, cyclic and seasonal elements and takes into account the particular repetitive continuing patterns exhibited by past data [11]. The ARIMA model analyzes seasonal and trend factors, estimates appropriate weighting parameters, tests the model, and repeats the cycle as appropriate [7]. The AR model estimates the stochastic process underlying a time series where the time series values exhibit a non-zero auto correlation, while the MA model estimates the process where the current time series value is related to the random errors from previous time period [11]. The procedure of Box-Jenkins approach has three stages: (a) tentative model identification, (b) parameter estimation, and (c) diagnostic checking. In the tentative model identification stage, stationary time series data is judged and non-stationary data is transformed into stationary data. After this, tentative ARIMA models are determined by analyzing auto-correlation functions (ACF) plot and partial auto-correlation functions (PACF) plot. In the parameter estimation stage, fitted ARIMA models are identified based on goodness-of-fit statistics and model parameters are estimated. The best suitable model is finally determined by the residual test and over-fitting examination in the diagnostic checking stage. Figure 1 shows the modeling process of the ARIMA model.

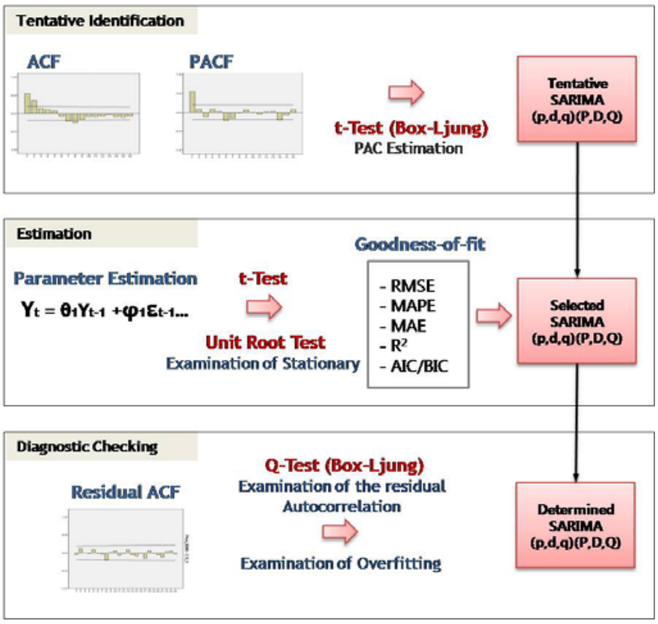

Figure 1. Procedures of the ARIMA Model

2.3 Building Information Modeling (BIM)

BIM is object-based parametric modeling that represents objects by parameters and rules that involve the geometric information as well as non-geometric properties and features [14]. By using object-based properties information, BIM can be used in overall project stages that involve design process, construction management, facility management, etc. In this research, properties within each object are used for monthly quantity take-off and analyzing material injection times. For these purposes, this research utilizes the following properties: geometric information and unit quantity of each material in each object yielded from the 3D model, material stocked times extracted from the scheduler, and the predicted values of unit material prices estimated by the time series forecasting model.

3. TIME SERIES MODEL FOR MATERIAL PRICE FORECASTING

3.1 Data Sources

This research develops a time series model to extract BIM properties of the predicted values of unit material prices prior to proposing the BIM-based construction material cost forecasting model. By analyzing past price time series data sets using a univariate ARIMA model, time variant predicted values of unit materials can be estimated. For time series modeling, structural steel is selected among various construction materials due to its wide range of fluctuations in the past 10 years. This research utilizes approximately 120 monthly time series data sets of high-tensile rebar (SD400: D51, 15.9kg/m, Seoul) from December 2000 to August 2010. Among these, the last 15 data sets from June 2009 to August 2010 are used for out-of-sample forecasting to check the validity of the selected model. Figure 2 represents the time series data of the high-tensile rebar in Seoul, Korea.

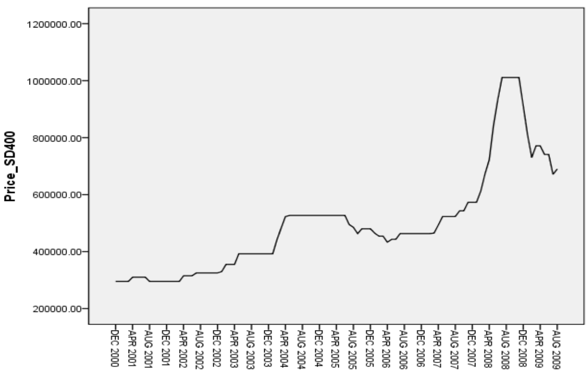

Figure 2. Row Time Series Data of High-Tensile Rebar

3.2 Fitting Time Series Models

The ARIMA model is applied to analyze the stationary time series data. Therefore, non-stationary original data sets should be transformed into stationary data sets by applying regular differencing or logarithmic transformation prior to determining a tentative ARIMA model. Because data sets of high-tensile rebar in this research show non-stationary averages and non-stationary variances, this model applies first or second order differencing and the logarithmic transformation and then obtains stationary data sets as shown in Figure 3.

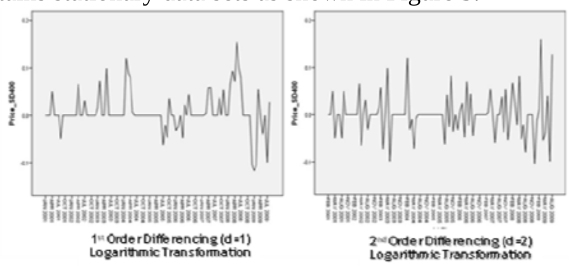

Figure 3. Transformed Data Sets

To determine tentative ARIMA(p,d,q) models after the process above, the order of the AR model (p), the order of the MA model (q), and the order of differencing (d) are determined by analyzing auto-correlation functions (ACF) plots and partial auto-correlation functions (PACF) plots. The possible values of p and q are set as 0 to 2 based on principle of parsimony as well as considering the efficiency of the modeling process.

In the tentative model identification stage, suitable models are selected by conducting test of goodness-of-fit. According to the results of this goodness-of-fit test and the parameter estimation, two suitable models are selected: ARIMA(1,1,0) and ARIMA(1,2,0). Table 1 shows the goodness-of-fit statistics of five tentative models.

Table 1. Test of Goodness-of-fit

Models	Goodness-fit-statistics			
	RMSE	MAPE	Normalized R^2	Normalized BIC
ARIMA(1,1,0)	24236.2	2.527	0.194	20.281
ARIMA(2,1,0)	24024.8	2.501	0.208	20.323
ARIMA(0,1,1)	25659.1	2.606	0.143	20.395
ARIMA(1,1,1)	24259.6	2.510	0.205	20.327
ARIMA(1,2,0)	26434.5	2.675	0.142	20.455

In Table 1, the smaller values of root mean square error (RMSE), mean absolute percentage error (MAPE) means the higher goodness-of-fit for predicting values while the larger values of normalized R^2 means a higher goodness-of-fit. Normalized Bayesian information criterion (BIC) rules are used to select tentative parameters for p and q. Table 2 represents the estimated parameters of the two selected models.

Table 2. Results of Parameter Estimation

Models	Constant			AR(1)		
	Value	SE	Prob.	Value	SE	Prob.
ARIMA(1,1,0)	0.008	0.007	0.211	0.437	0.089	0.000
ARIMA(1,2,0)	-0.001	0.003	0.980	-0.389	0.096	0.000

In the diagnostic checking stage, the residual autocorrelation function is used to examine the residuals series of the tentative models. Residuals mean the differences between the actual (observed) and the fitted values. The closer to zero and the more random are the residual, the better is the fit [15]. The Box-Ljung chi-square test examines the autocorrelation of the residuals. When probability is larger than the level of significance (generally 0.05), the null hypothesis that correlation coefficients of white noises are all zero is adopted. Table 3 represents the results of the residual test.

Table 3. Rresults of the Residual Test.

Models	Examination of the residual Autocorrelation: Ljung-Box Q (18)		
	Q-Statistics	Degree of freedom	Prob.
ARIMA(1,1,0)	14.904	17	0.602
ARIMA(1,2,0)	23.978	17	0.120

According to the results of the residual test, selected ARIMA(1,1,0) and ARIMA(1,2,0) are fitted to high-tensile rebar unit price forecasting models.

3.3 Forecasting

This research conducts comparative analysis of the prediction capability between two selected models using the out-of-sample forecast. In this stage, the last 15 data sets from June 2009 to August 2010 are utilized to analyze the predictive power of the models as well as check the validity of the selected models. From the results of out-of-sample forecasting shown in Figure 4, the ARIMA(1,1,0) model reflects the trends of original data more than the ARIMA(1,2,0) model, although the former has larger variance of errors than the latter model.

Figure 4. Comparative Analysis of Predictive Power (Out-of-sample forecast)

Therefore, this research finally selects the ARIMA(1,1,0) model and then forecasts material prices using this model for 12 months from September 2010. The results of the predictions in Figure 5 show that material prices can steadily increase, thus material cost escalation should be considered in construction cost estimations as well as in overall project management phases.

4. BIM-BASED COST FORECASTING MODEL

4.1 Exporting Properties

This research proposes the BIM-based time series construction material cost forecasting model. This model utilizes the time series unit material price forecasting model mentioned above and uses object properties information related to quantities and activity times of materials extracted from 3D modeling and construction scheduling.

Figure 5 represents the process of exporting properties from the BIM 3D model, time series model, and quantity take-off system connected with scheduling module. Each object separately has a standard and unique ID classified according to located and installed area as well as includes various properties encompassing material types and sizes of installed materials (length, area, and volume).

This information is imported to the quantity take-off module connected to the scheduling module. Information of activity start times and material stocked times extracted from this model are added to objects as new properties. With respect to predicted values of material unit prices, these prices' properties are extracted from the time series model and then added to objects corresponding to activity time periods when each material is installed. Finally, each object includes properties about material types, material quantity information, material stocked time and future value of material prices corresponding to the materials' stocked time.

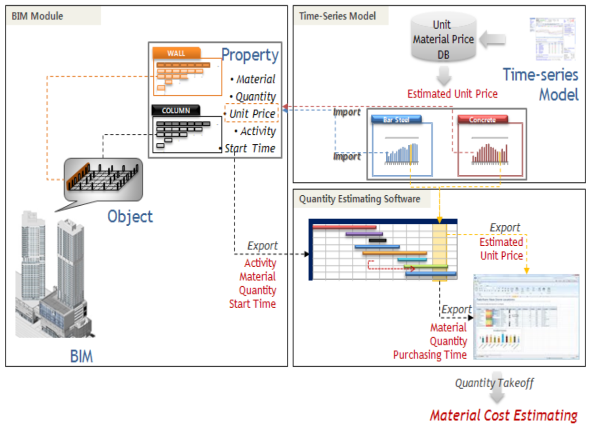

Figure 5. Process of Exporting Properties

4.2 Model Framework

Figure 6 represents the BIM-based material cost forecasting model framework after applying trends of material price fluctuations. This model is constructed based on the extracted properties mentioned above and the database. By using extracted properties including material types, quantity information, material stocked times, and future values of material prices, the quantities and times of materials inputs are classified according to corresponding activity start time, so monthly (as well as weekly or quarterly) material quantity inputs can be yielded from the quantity take-off system and the 5D simulator.

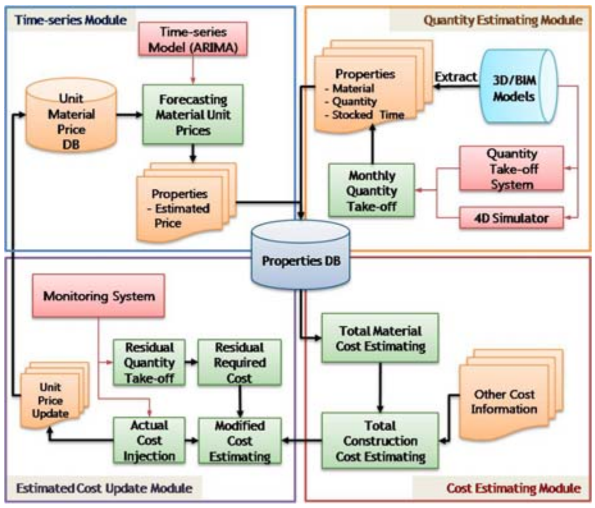

Figure 6. Model Framework

In the cost estimating module, each monthly quantity is multiplied by forecasted material prices of equivalent periods so that total material cost can be yielded. Total construction cost can be also estimated by applying other cost information such as labor costs, subcontracts costs, and equipment costs.

On the other hand, new time series data sets will be updated as the construction progresses. By using this updated material prices data, more accurate estimated values of material prices can be forecast to modify budgeted costs for residual quantities and work schedules. Therefore, the cost update module can more accurately estimate modified total construction costs by combining budgeted costs for residual works and actual costs for work performed.

Figure 7 shows the example of a properties database of the BIM-based cost model. In a properties database, each object has a unique ID and includes the materials' names, quantity information, stocked time, and unit price of materials, so total material prices can be automatically calculated in this module.

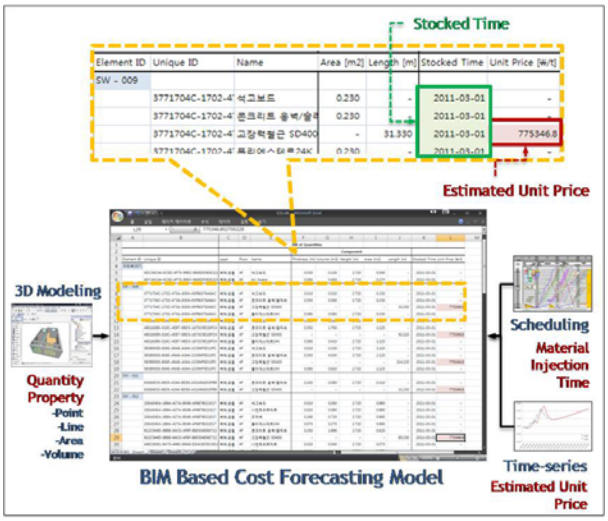

Figure 7. An Example of the Model

5. CONCLUSIONS

This research developed a time series model with the Box-Jenkins methodologies and material price time series data in Korea in order to forecast trends of unit prices of required materials. BIM approaches are also used to analyze injection times of construction resources and to conduct quantity take-off so that total material prices can be forecast. To determine an optimal time series model for forecasting prices trends, particularly focusing on rebar, comparative analysis of the predictability of tentative ARIMA models was conducted.

In recent volatile economic conditions, good prior knowledge of future material prices can help to increase the accuracy of cost estimates. Therefore, the proposed BIM-based time series forecasting model can help to deal with sudden changes in economic conditions and design changes by estimating future material prices corresponding to resource injection times.

Although this research tried to forecast trends of injected construction resources, this model can produce large forecast errors if discontinuities occur within the projection time periods. Also, it can make trends reliable only in the short run as well as involve complicated and several iterative procedures. In future research, several attempts should be made to reduce prediction errors by building multivariate time series models that consider various influence factors and by applying an intervention model to deal with sudden changes of market conditions. Auto-selected procedures and algorithms for determining suitable time series models can also help to simplify complicated and iterative forecasting procedures of the present model.

ACKNOWLEDGEMENTS

This research was supported by a grant (code # 09 R&D A01) from Super-Tall Building R&D Project funded by the Ministry of Land, Transport and Maritime Affairs of Korean government.

REFERENCES

[1] Shane, J. S., Molenaar, K. R., Anderson, S., and Schexnayder, C., "Construction Project Cost Escalation Factors", *Journal of Management in Engineering*, Vol. 25(4), pp.221-229, 2009.

[2] Williams, T. P., "Predicting Changes in Construction Cost Indexes Using Neural Networks", *The Journal of Construction Engineering and Management*, Vol. 120(2), pp.306-320, 1994.

[3] Trost, S. and Oberlender, G., "Predicting Accuracy of Early Cost Estimates Using Factor Analysis and Multivariate Regression", *The Journal of Construction Engineering and Management*, Vol. 29(2), pp.198-204, 2003.

[4] Shaheen, A. A., Fayek, A. R., and Abourizk, S. M., "Fuzzy Numbers in Cost Range Estimating", *The Journal of Construction Engineering and Management*, Vol. 133(4), pp.325-334, 2007.

[5] Touran, A., "Probabilistic Model for Cost Contingency", *The Journal of Construction Engineering and Management*, Vol. 129(3), pp.280-284, 2003.

[6] Sonmez, R., "Parametric Range Estimating of Building Costs Using Regression Models and Bootstrap", *The Journal of Construction Engineering and Management*, Vol. 134(12), pp.1011-1016, 2008.

[7] Lu, T. and Abourizk, S. M., "Automated Box-Jenkins Forecasting Modeling", *Automation in Construction*, Vol. 18(5), pp.547-558, 2009.

[8] Issa, R. A., "Application of Artificial Neural Networks to Predicting Construction Material Prices", *Proceedings of the Eighth International Conference on Computing in Civil and Building Engineering*, ASCE, pp.1129-1132, 2000.

[9] Ashuri, B. and Lu, J., "Time Series Analysis of ENR Construction Cost Index", *The Journal of Construction Engineering and Management*, Vol. 136(11), pp.1227-1237, 2010.

[10] Akintoye, A., Bowen, P., and Hardcastle, C., "Macro-economic leading Indicators of Construction

Contract Prices", *Construction Management and Economics*, Vol. 16(2), pp.159-175, 1998.
[11] Ng, S. T., Cheung, S. O., Skitmore, M., and Wong, T. C., "An Integrated Regression Analysis and Time Series Model for Construction Tender Price Index Forecasting", *Construction Management and Economics*, Vol. 22(5), pp.483-493, 2004.
[12] Fellows, R. F., "Escalation Management: Forecasting the Effects of Inflation on Building Projects", *Construction Management and Economics*, Vol. 9(2), pp.187-204, 1991.
[13] Wong, J., Chan , A., and Chiang , Y., "Time Series Forecasts of Construction Labour Market in Hong Kong: The Box-Jenkins Approach", *Construction Management and Economics*, Vol. 23(9), pp.979-991, 2005.
[14] Eastman, C. Teicholz, P., Sacks, R., and Liston, K., *BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers, and Contractors*, John Whiley & Sons Inc., USA, pp.25-64, 2008.
[15] Khosrowshahi, F. and Alani, A. M., "A model for Smoothing Time-series Data in Construction", *Construction Management and Economics*, Vol. 21(5), pp.483-494, 2003.