Jurnal Teknologi

An Overview of the Forecasting Methods Used in Real Estate Housing Price Modelling

Mohan Munusamy[a], Chitrakala Muthuveerappan[b*], Maizan Baba[b], Mat Naim Abdullah @ Mohd Asmoni[b]

[a]*Researcher, Faculty of Geoinformation & Real Estate, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Malaysia*
[b]*Centre of Real Estate Studies, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Malaysia*

*Corresponding author: chitrakala@utm.my and chitrakala_m@yahoo.com

Graphical abstract
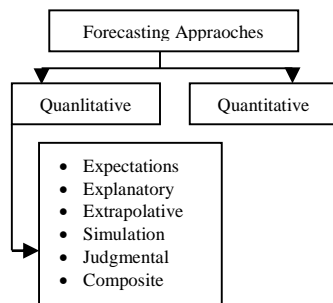

Abstract

Forecasting is very fundamental in real estate where the past transactions become the evidences while decision making for the present and the future. Several techniques and validation approached that were commonly used in housing price index forecasting. Beside the appropriate forecasting method, error calculation is one of the critical constraints in accuracy out of all methods. This paper overview the available methods and the types of error being considered in forecasting techniques. Then the forecasting methods, namely Multiple Regression Analysis (MRA) and Artificial Neural Network which are highly applied in forecasting modelling are compared over its error accuracy.

Keywords: Forecasting methods; forecasting errors; real estate modelling

Abstrak

Ramalan adalah suatu perkara asas dalam hartanah di mana transaksi yang lalu dijadikan bukti semasa membuat keputusan untuk masa kini dan masa hadapan. Beberapa teknik dan pendekatan validasi adalah yang biasa digunakan dalam meramal indeks harga perumahan. Selain ramalan, pengiraan ralat adalah salah satu kekangan ketepatan yang kritikal daripada semua kaedah-kaedah yang lain. Kajian ini mengkaji tentang kaedah-kaedah yang sesuai dan jenis-jenis ralat yang dipertimbangkan dalam teknik ramalan. Teknik ramalan iaitu Analisis Regresi Berganda (MRA) dan Rangkaian Neural Buatan yang kerap digunakan di dalam model ramalan dibandingkan dengan ralat ketepatannya.

Kata kunci: Kaedah ramalan; ralat ramalan; permodelan harta tanah

■1.0 INTRODUCTION

Real estate is a unique investment category influenced by many aspects in every nation's economy either macro or micro economic factors. Investors, banking sectors, economists, developers and even private individuals need accurate information while decision making. Thus, forecasting become vital in the property investment decision-making process for institutional investors, supporting asset allocation, property fund strategy and stock selection for a mixed-asset portfolio (Chaplin, 1999; Mitchell and McNamara, 1997).

Forecasting plays vital role in real estate in the decision making of individual or organization knowing the potential or trend would allow for an investment plan. In general, forecasting is an attempt to predict and explains that forecasting is the prediction of how does the actual value or situation of a subject will be in future rather than what should be (Armstrong 2001). Garrick and Raymond (2001), explain that every forecasting method is based on the assumption that the pattern or behaviour of the parameter and its accuracy which are reflected by the mathematical relation adopted where it tends to represent the actual subject condition. Therefore, generally it uses historical data to examine the potential condition or trend of the unknown output forecasted subject.

Choosing the suitable method of forecasting technique is essential which may influence in generating good forecasted information. Armstrong (2001) suggests convenience, market popularity, structured judgment, statistical criteria, relative track records, and guidelines from prior research while selecting a suitable forecasting method in Real Estate. However, many empirical researches had shown that accuracy varies according to different method and since accuracy is very important, therefore choosing the right technique of forecasting is vital consequently arose questions arise on the selection of appropriate forecasting technique, namely: What are the available approaches and its methods in forecasting?, What are the sources of errors and types of errors considered in forecasting?, and What are the common forecasting tools in real estate?. A detailed knowledge is needed while selecting the method while considering forecasting methods

may assists the ability to forecast the future trends either, demand, price or supply in Real Estate would be an advantage especially for investors and speculators.

Consequently, there have been many different techniques being developed to fulfil these needs. This paper overviews the forecasting techniques, and the common techniques and the error calculation for its accuracy that had been used throughout the years in real estate.

The contents of this article are presented in a manner that initially it explains the approaches in forecasting and the common methods in each approach are summarised. Then sources of errors and types of errors used for accuracy in forecasting are elaborated. Finally, the forecasting methods that are commonly used in real estate and the justification of adoption in terms of error handling are discussed.

■2.0 APPROACHES IN FORECASTING

Forecasting become essential in decision making to be rational substantially increased the forecasting knowledge through high level research activities in recent years. The statistical and stochastic methods have major advantages in designing probability distribution based forecast rather than a deterministic point of forecasting. Moreover, these methods become the base or fundamental design of any forecasting modelling.

Forecasting falls into two categories, which is a quantitative approach and qualitative approach (Seow and Teck, 1996). Quantitative approached organize past information about an activity to forecast the future condition. Qualitative approaches are used when historical data related to the events to be forecasted either limited or unavailable, or when the events are affected by non-quantifiable information (Gurney K 1997). Heather and Leonie (2008) classify Expectation, Extrapolative and Explanatory methods are widely used in mortality forecasting. Actuaries and Demographers have increasingly been using standard statistical forecasting methods. The Society of Actuaries (SOA) a worldwide organization for all actuarial had subdivided forecasting methods, namely Extrapolative, Explanatory, Simulation, Judgmental, and Composite.

2.1 Expectation

Expectation in the form of experts' opinion based on linear models with high and low scenario. Waldron (2005) explains the scenario where expectation approach has been adopted in forecasting. Actuaries also used this approach using simple linear models in the past and moved to extrapolate approaches (Idler & Benyamini, 1997; Heather and Leonie, 2008).

2.2 Explanatory

Explanatory Method is used when the causes and risk factors are known, and is generally limited to short-term forecasting. Regression analysis; Predictive modelling; Artificial Neural Network; and Econometric modelling are the main methods being used under this approach (Gardner, 1985; Makridakis *et al.* 1998; Armstrong J.S., 2001; Montgomery *et al.*, 2008).

2.3 Extrapolative

Extrapolative method assumes that future the trend always be the continuation of the past. This method uses the historical date with a weight where higher weight will be given to the recent year data. Widely used due to its simplicity, accuracy and ease of use in business (REF). Simple moving average, time series, and exponential or complex moving average is the common methods under this approach in practice. Especially in time series, being a stochastic model enables the probabilistic prediction interval calculation is the advantage (Makridakis *et al.* 1998; Shapiro *et al.* 1999; Armstrong J.S., 2001; Fox & Fox 2008).

2.4 Simulation

Simulation Modelling Method provides just-in-time results according to any parametric value changes. Cell based modelling, dynamic simulation, and multi agent simulation fall under this approach. Complex modelling structures usually modelled using simulation to produce analog results to study the complexity impact (Makridakis *et al.*, 1998; Epstein, 1996; Miller & Page 2007; Gilbert, 2008).

2.5 Judgmental

Judgmental Method is based on the experts' consensus mechanism of similar cases. Qualitative method totally relies on the experts' opinion rather than historical data or any statistical methods. Delphi technique, Focus group is some common methods under this approach (Makridakis *et al.*, 1998; Lawrence *et al.*, 2006).

2.6 Composite

Composite approach is more of the similarity parameters of similar case studies on case based reasoning. This approach needs huge data with short histories. Bayesian forecasting method is the common method under this approach (Lapide, 2008; Makridakis *et al.* 1998; West, M., & Harrison, J. 1997).

■3.0 SOURCE OF ERRORS IN MODELLING

Accuracy is regarded as the level of significance for acceptance in forecasting is termed by Armstrong (2001) as an "optimist term for forecasting error". Forecast errors, on the other hand, represent the difference between the forecast value and the actual value. Due to parameter selection, calculation methods, and forecasting models, errors are considered as a variable with significant value in modelling. Alho (1990) informs that there are many sources of error in forecasting demographic rates. Heather and Leonie (2008) and Keilman (1990) mention that sources of error may be classified into four various but related sources: a) wrong model specification for forecasting, b) parameter consideration, c) errors in the dataset, and randomness inherent variation. Garrick and Raymand (2001) had underline four techniques in measuring the accuracy of forecasted information. Table 1 brief about the common error measuring methods available in forecasting.

Table 1 Methods of measuring the forecasting accuracy

Technique	Description	Formula
Coefficient of Determination (R^2)	Measures how well the model is. It is the square of correlation between Y (the model) $\check{Y}$ (the computing)	$Y = \beta0+\beta1+Xi+\beta2X2i+....\beta nXni+\varepsilon$ Where Y = value of the dependent variable $\alpha\beta0$ = intercept or constant $\beta1....n$ = slope of the regression X1i = given value of the independent variable ε = observed error or residual
Root Mean Square Error (RMSE)	Differences between the two data sets where one data set from the computed and the second data set will be the actual.	$\sqrt{(\sum|"X"("f")/"N"|^{"2"})}$
Mean Squared Errors (MSE)	The accuracy measure computes by squaring the individual error for each item in a data set and then finding the average or mean value of the sum of those squares.	$MSE = \frac{1}{n}\sum_{t=1}^{n}e^2 t$ Where N = period E2 = forecasting error
Mean Absolute Percentage Error (MAPE)	The mean or average of the sum of all of the percentage errors for a given data set taken without regard to sign so as to avoid the problem of positive and negative values cancelling each other.	$MAPE = \frac{1}{n}\sum_{t=1}^{n}[PE_t]$ $PE_t = \left(\frac{Y_t - F_t}{Y_t}\right)X100$ Where PE = percentage error, n = time period Y_t = actual observation for time period t F_t = forecast for the same period
Theil's U-statistic	Developed by Henry Theil (1966), it is an accurate measure that emphasizes the importance of large errors (as in MSE) as well as providing a relative basis for comparison with naïve forecasting methods (Garrick and Raymand, 2001). Makridakis *et al*. (1998) have simplified Theil's equation.	**U =** U= $$U = \sqrt{\frac{\sum_{t=1}^{n-1}\left(\frac{F_{t+1}-Y_{t+1}}{Y_t}\right)^2}{\sum_{t=1}^{n-1}\left(\frac{Y_{t+1}-Y_t}{Y_t}\right)^2}}$$ Where U = Theil's statistics, F = Forecast Y = Observation.

Source: Makridakis *et al*. (1998); and Garrick and Raymond (2001)

Considering accuracy and reliability in forecasted results, all types of errors should be taken account while predicting intervals in forecasting. Omitting one or more sources of uncertainty may mislead the level of confidence in forecasting.

■4.0 FORECASTING METHODS IN REAL ESTATE

In Real Estate, large number of forecasting models including housing price are developed using Multiple Regression Analysis (MRA) and Artificial Neural Network (ANN). Partial Least Square (PLS), Genetic Algorithmnm (GA), Fuzzy Logic, and Support Vector Regression (SVR) are also used in a few studies of housing price index modelling. Regression techniques have been long central to the field of economic statistic (Alho 1990). For instance,

Mark and Goldbery (1988) explain that the main traditional method in property mass assessment had always been multiple regression analysis.

The MRA based method has been popular because of their established methodology, long history of application and wide acceptance among both practitioners and academicians. Besides that, multiple regression analysis had also been used in other field such as in forecasting potential demand. Hua (1997) had used multiple regressions to forecast the residential construction demand in Singapore. Barr *et al*. (1996) developed model selection criteria such as firm merging, share prices in an abnormal event in finance using multiple regression with Mean Square Error of Prediction (MSEP) and Mean Square Error (MSE).

Cheng *et al*. (2000) applied MRA to analyse the pricing dynamics of index in Hong Kong. Cegielski (2002) applied Multiple Regression to analyse the price of public stocks of real-estate firms. Korean Public Administration Review (2003) applied MRA to analyse the assessment model value to validate the assessment value over cost approach and transacted price approach. Ibrahim (2006) used multiple regression analysis in spatial data using geographical information system for property mass appraisal processes. Ng *et al*. (2007) had combine linear regression analysis and genetic algorithm technique in forecasting the private housing demand in Hong Kong. Xiaoli Yang and Qiong He (2013) forecasted China's Yunnan province housing price from 1999 to 2009 by applying MRA.

The limitation of traditional linear multiple regression analysis had been recognized for some time (Mark and Goldberg, 1998; Do and Grudnitski, 1992). The methodological problem associated with non-linearity, multicollinearity, function form misspecification, and heteroscedasticity had limited the accuracies of the technique (Larsen and Peterson, 1988; Mark and Goldberg, 1988; Do and Grudnitski, 1992). Therefore, several other techniques had been introduced to overcome the limitation faced by multiple regression analysis method. According to Do and Grudnitski (1992), the most commonly used methods in addressing the problem faced by multiple regression analysis is a method based on the neural network system.

The use of artificial neural network in real estate had dated around 20 years ago (Julio and Esperanza, 2004). Peter Rossini (1997) and Stanley (1998) applied ANN in residential property value forecasting. Stephen *et al*. (2001) used ANN to forecast construction cost of buildings at the pre-design stage and applied MSE to estimate the error. Ian Wilson *et al*. (2002) used ANN in residential price forecasting of national and regional levels in the UK market. Cegielski (2002) used the ANN price of public stocks of real-estate firms and applied MRA to study the significance. Ge and Runeson (2004) modelled Property price index using Neural Network Model for Hong Kong property market. Mohamad (2012) used ANN to forecast the housing price index in Malaysia. Hasan Selim (2004). ANN in real estate application, especially related to pricing index can be found from Cheng *et al*. (2000), Cegielski (2002).

■5.0 DISCUSSION

Since MRA and ANN are highly used in real estate forecasting, for accuracy and reliability, many studies have been carried out to compare ANN with MRA. Majority of the forecasting studies calculates the percentage of error of both the artificial neural network and multiple regression analysis models which then been applied in a market sample in which the sales price is known. Ng and Cripps (2001) compared multiple regression analysis and artificial neural network in predicting housing value. Peter and Ralph (2008) had also used neural network and compare the

accuracies of multiple regression analysis in evaluating of housing price in Nigeria. Mohamad *et al*. (2012) had also compared multiple regression analysis and artificial neural network method in mass appraisal context. Besides that Julio and Esperanza (2004), simulate the hypothesis by applying it in valuing real estate value in Madrid. Hasan (2008) forecasted the housing price in the Turkey market using MRA & ANN and demonstrated that ANN is more significant than MRA. The majority of the real estate analysis using multiple regression or neural network is carried on macro economic parameters. All these macroeconomic parameters on housing price index forecasting using MRA and ANN confirm that ANN shows an average error rate between 5 to 10% inaccurate while multiple regression analysis shows a higher average which is 10 to 15% inaccurate. Hasan (2008) applied MRA and ANN on microeconomic parameters of housing prices is confirm that ANN provides more significant result over MRA.

■6.0 CONCLUSION

From the literatures, there are many techniques that have been used in forecasting the housing price in real estate market. Explanatory and exploratory research approaches are used in housing price forecasting in real estate. Selecting the most appropriate technique is very crucial to ensure that the model being developed should really reflect the actual market condition to reduce the source of errors in modelling. Among the forecasting methods, MRA and ANN are the most adopted methods in housing price forecasting.

All empirical studies carried on to evaluate the performance of MRA and ANN confirms that ANN eliminates the non-linearity issue in modelling and more significant than MRA in housing price index forecasting.

References

[1]   Alho, J. M. 1990. Stochastic Methods in Population Forecasting. *International Journal of Forecasting.* 6(4): 521–530.
[2]   Armstrong, J. S. 2001. *Principles of Forecasting. A Handbook for Researchers and Practioners*. Springer 2001. ISBN: 0-7923-7930-6.
[3]   Barr, G. D. I, Sparks, R., Affleck Graves, J., & Thompson, M. L. 1996. Selection Criteria for Regression Models and Their Application to Abnormal Event Studies in Finance. *South African Statistical Journal*. 30(2): 155–173.
[4]   Chaplin, R. 2000. Predicting Real Estate Rents: Walking Backwards into the Future. *Journal of Property Investment and Finance.* 18: 352–370.
[5]   Cegielski, C., Kitchens, F., Rebman, C., & Lang, T. 2002. What is a Dot-Com Worth? An Exploration of the Value Estimates of Information Technology-based Firms. *Journal of Computer Information Systems*. 43(2): 31–35.
[6]   Cheng, L. T. W., Fung, J. K. W., & Chan, K. C. 2000. Pricing Dynamics of Index Options and Index Futures in Hong Kong Before and During the Asian Financial Crisis. *Journal of Futures Markets.* 20(2): 145–166.
[7]   Epstein, J. M., Axtell, R. 1996. *2050 Project. Growing Artifcial Societies: Social Science from the Bottom Up*. Washington, D.C.: Brookings Institution Press.
[8]   Fox, J., & Fox, J. 2008. *Applied Regression Analysis and Generalized Linear Models*. 2nd Ed. Los Angeles: Sage.
[9]   Gardner, E. S. 1985. Exponential Smoothing: the State of the Art. *Journal of Forecasting*. 4(1): 1–28.
[10] Garrick S., & Raymond. W. 2001. *The Validity of Forecasting*. Pacific Rim Real Estate Society International Conference Christchurch.
[11] Gilbert, G. N. 2008. *Agent-Based Models*. Los Angeles: Sage Publications.
[12] Gurney, K. 1997. *An Introduction to Neural Networks*. London: Routledge. ISBN 1-85728-673-1.
[13] Hasan Selim. 2009. Determinants of House Prices in Trukey: Hedonic Regression Versus Artificial Neural Network. *Journal of Expert Systems with Applications*. 36(2). ISBN: 2843-2852.
[14] Heather B., & Leonie T. 2008. *Mortality Modelling and Forecasting: A Review of Methods*. The Australian Demographic & Social Research Institute Working Papers. Official url : http://adsri.anu.edu.au [Accessed June 2013].

[15] Theil, H., Raj, B., & Koerts, J. 1992. *Henri Theil's Contributions to Economics and Econometrics: Econometric Theory and Methodology*. Publisher: Springer Science & Business Media. ISBN 0-7923-1548-0.

[16] Hua, G. B. 1997. An Evaluation of the Accuracy of the Multiple Regression Approach in Forecasting Sectoral Construction Demand in Singapore. *Construction Management*. 17(2): 231–241.

[17] Julio Gallego, M. E. 2004. *Artificial Intelligence Applied to Real Estate Valuation: An Example for the Appriasal of Madrid*. JULIO GALLEGO MORA –ESPERANZA. 255–265.

[18] Keilman, N. 1990. *Uncertainty in Population Forecasting Analyses, Recommendations*. Swets & Zeitlinger, Amsterdam.

[19] KOREAN PUBLIC ADMINISTRATION REVIEW. 2003. *Alternatives for Improving the Indexes for Building Assessment in Property Tax*. Publisher: The Korean Association for Public Administration. ISSN: 1226-2536. 37(3): 263–289.

[20] Lawrence, M., Goodwin, P., O'Connor, M., & Onkal, D. 2006. Judgemental Forecasting: A Review of Progress Over the Last 25 Years. *International Journal of Forecasting*. 22: 493–518.

[21] Lapide, L. 2008. Thinking About Composite Forecasting. The Journal of Business Forecasting. 27(2).

[22] Makridakis, S. G., Wheelwright, S. C., & Hyndman, R. J. 1998. *Forecasting: Methods and Applications*. 3rd ed. New York: John Wiley & Sons.

[23] Mark, J., & Goldberg, M. 1998. Multiple Regression Analysis and Market Assessment: A Review of the Issues. *Appraisal Journal*. 56(1): 89–109.

[24] Miller, J. H., & Page, S. E. 2007. *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*. Princeton, N.J.: Princeton University Press.

[25] Mitchell, P.and McNamara, P. 1997. Issues in the Development and Application of Property Market Forecasting: The Investor's Perspective. *Journal of Property Finance*. 8: 363–376.

[26] Mohamad, S. M. R., Chitrakala, M., Norhaya, K., & Izran, S. 2012. Forecasting Housing Price Index Using Artificial Neural Network Technique. *Malaysian Journal of Real Estate*. 7(1). ISBN: 1832-8505.

[27] Montgomery, D. C., Jennings, C. L., & Kulahci, M. 2008. *Introduction to Time Series Analysis and Forecasting*. Hoboken, N.J.: 34.: Wiley-Interscience.

[28] Ng, S. T., M. Skitmore, *et al*. 2008. Using Genetic Algorithms and Linear Regression Analysis for Private Housing Demand Forecast. *Building and Environment*. 43(6): 1171–1184.

[29] Nguyen, N. and A. Cripps. 2001. Predicting Housing Value: A Comparison of Multiple Regression Analysis and Artificial Neural Networks. *Journal of Real Estate Research*. 22(3): 313–336.

[30] Peter, O., & Eriki, R. I. U. 2008. Application of Neural Network in Evaluating Price of Housing Units in Nigeria: A Preliminary Investigation. *Journal of Artificial Inteligence*.

[31] Peter Rossini. 1997. *Application of Artificial Neural Networks to the Valuation of Residential Property*. Proceedings from the 3rd annaul pacific-Rim Real Estate Society Conference, New Zealand, Jan. 1997.

[32] Shapiro, A. F., Pfumm, J. S., & DeFilippo, T. A. 1999. The Inner Workings of Neural Networks and Genetic Algorithms. Actuarial Research Clearing House. 1: 415–426.

[33] Seow E. O., Teck. I. C. 1996. Singapore Residential Market: An Expert Judgemental Forecast Incorporating the Analytical Hierarchy Process. *Journal of Property Valuation and Investment*. 14.

[34] Stanley McGreal, A. A., Dylan McBurney, David Patterson. 1998. Neural Networks: The Prediction of Residential Values. *Journal of Property Valuation and Investment*. 16(1): 57–70.

[35] Xiaoli Yang, & Qiong He. 2013. Influence of Modeling Methods for Housing Price Forecasting. *Advanced Materials Research*. 798–799: 885–888.

[36] West, M., & Harrison, J. 1997. *Bayesian Forecasting and Dynamic Models*. 2nd edition. Publisher: Springer, New York.