Forecasting low-cost housing demand in an urban area in Malaysia using artificial neural networks: Batu Pahat, Johor

N. Y. Zainun[1], I. A. Rahman[1] & M. Eftekhari[2]
[1]*Universiti Tun Hussein Onn Malaysia, Johor, Malaysia*
[2]*Loughborough University, Leicestershire, UK*

Abstract

Over the past decade, the rate of growth of housing construction in Malaysia has been dramatic. The level of the urbanization process in the various states in Peninsular Malaysia is considered to be important in planning for low-cost housing needs. The aim of this study is to develop a Neural Networks model to forecast low-cost housing demand in Batu Pahat, Johor, one of the states in Peninsular Malaysia. The time series data was analyzed using Principal Component Analysis to determine the significant indicators which will be the input in Neural Networks model. The feed forward network with the most commonly used training algorithm, back propagation networks is used to develop the model. The results show that the best Neural Network model is 2-25-1 with 0.7 learning rate and 0.4 momentum rate. Neural Networks can forecast low-cost housing demand in Batu Pahat very well with 0% of MAPE value.

Keywords: forecasting, low-cost housing, artificial neural networks.

1 Introduction

In each five year National Plan, Malaysia's government has focused on various housing programmes to ensure that all Malaysians, particularly the low income groups, have access to adequate and affordable shelter and related facilities [1].

During the Ninth Plan period, the development of the housing sector continues to focus on the provision of adequate, affordable and quality houses for all Malaysians [2]. The housing is divided into four main categories; low cost, low medium cost, medium cost and high cost housing. In Malaysia low cost

WIT Transactions on Ecology and the Environment, Vol 142, © 2010 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/SW100051

housing is defined at a ceiling price RM25,000 per unit or less. Low cost housing can be sold to households with a monthly income between RM500 to RM750 while low medium is defined at a ceiling price of RM25,001 to RM60,000 and can be sold to households with monthly incomes between RM750 to RM1,500 [3]. On the other hand, the construction cost alone ranges from a low of RM12,000 per unit to a high of RM43,000, with average cost RM23,000 per unit for terrace house [1]. Therefore, To ensure an adequate supply of low cost houses for the low income group, any mixed-development projects undertaken by private developers, continued to be guided by the 30% low cost housing policy requirement [2]. Construction each category of housing should build fairly especially in such area which located level of people with the different incomes. By develop low cost and low medium cost housing it can reduce housing growth illegally on the government's land and also prevent the public creating other new squatters.

In the year 2003, Selangor, Johor, Perak, Federal Territory Kuala Lumpur and Penang dominated housing existing stock and together contributed 68.9% (2,133,128 units) of total existing housing stock in Malaysia [4]. All these states experience a high migration of people because of many vacancies offered in industry also well maintain economy flow. Residential Property Stock Report in that year reported that housing stock in the fourth quarter was increased by 1.3% to 3,237,599 units over third quarter.

Due to the increment of the demand for low cost houses it is very significant and vital; the selection of the best method on forecasting of demand is also becoming an important factor. All this while, the number of unit of low cost houses have been built by practice the requirement imposed by the government which is 30% of the total development. Obviously, by following this requirement, the numbers of low cost houses to be built do not reflect the actual demand of low cost housing. Henceforth, developing a model as an alternative way to forecast the number of units of low cost houses is therefore timely and imperative for a developing nation.

2 Objective

The aim of this paper is to develop a model to forecast low-cost housing demand in the district of Batu Pahat, Johor, using Artificial Neural Networks. The actual and forecasted data will be compared and validated using Mean Absolute Percentage Error (MAPE).

3 Methodology

The methodologies of this study are including finding out the significant indicator using Principal Component Analysis (PCA) adapted from SPSS and a Neural Network (NN) model development adopted from NeuroShell2. PCA is used to derive new indicators; that is the significant indicators from the nine selected indicators. The indicators are: (1) population growth; (2) birth rate; (3) mortality baby rate; (4) inflation rate; (5) income rate; (6) housing stock;

(7) GDP rate; (8) unemployment rate; and (9) poverty rate. The new indicators in terms of Principal Component (PC) will be the inputs in the NN model development. The dependent indicator is the monthly time series data on low cost housing demand starting from January 2000 to December 2002.

In NN model development, a series of trial and error process are done to find the suitable number of neurons in the hidden layer, learning rate, momentum rate and screening the result using the best NN model.

4 Significant indicators

The determinant correlation matrix for Batu Pahat is |R| = 7.47× 10-17, that is very close to zero. The hypothesis assumes the population matrix is equal to the identify matrix, that is all indicators are uncorrelated when the data are multivariate normal. For Batu Pahat, there are 9 indicators, p = 9 and 36 data, N =36 will be analyzed. The value for test statistic is 1157.313 with the critical point for chi-square p(p-1)/2 = 9(9 – 1)/2 = 36. Degree of freedom, α = 0.001 with the critical point from chi-square table is 67.923. As a result, the hypothesis will be rejected at 0.001 significant level because 1157.313 > 67.923. Therefore, PCA can be performed.

PC1 gave the largest eigenvalue with 6.748 consist 74.97% of total variation while PC2 showed 1.332 eigenvalue with 14.8% of total variation. PC3 is 0.904 eigenvalue covered 10.04% of variation. For PC4 to PC9, the eigenvalues is 0.017 with only 0.07% of variation.

From the scree plot in Figure 1, eigenvalues for PC4 to PC9 are close to zero. While the eigenvalues for principle component three is less than one. Since the eigenvalues for principle component four to nine is close to zero and PC3 less than one, they can be ignored. Therefore, there is two PC that will be as the input to develop the neural network model.

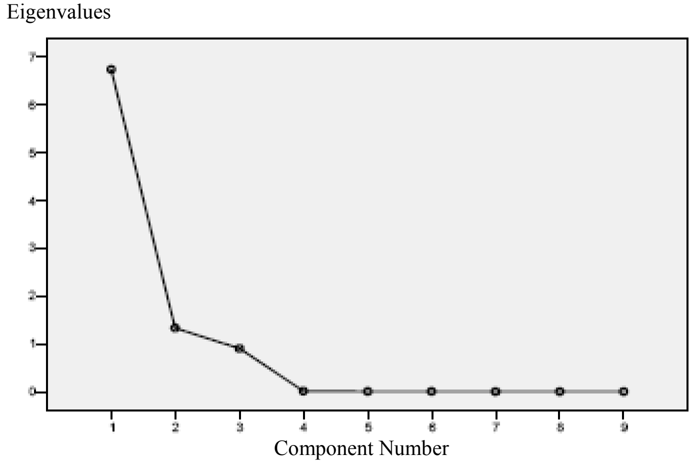

Figure 1: Scree plot for Batu Pahat.

Table 1: Component score coefficient matrix for Batu Pahat.

Indicators	Component	
	1	2
Population growth	0.683	0.381
Birth rate	-0.975	0.025
Mortality baby rate	0.567	-0.806
Unemployment rate	0.730	0.553
Inflation rate	0.990	-8.271×10^{-2}
Gross Domestic Product	-0.914	0.393
Poverty rate	0.989	3.989×10^{-2}
Income rate	-0.975	3.905×10^{-3}
Housing stock	0.855	0.261

Table 1 shows the component score coefficient matrix in Batu Pahat for the nine indicators. The number of component is to be equal to the number of eigenvalue of R which is equal to 1 [5]. The most significant indicators are evaluated using component score coefficient matrix nearest to 1. For PC 1, the most significant indicator is inflation rate and for PC 2 the most significant indicator is unemployment rate.

5 Model development

The learning and momentum rate is determined by means trial and error. From the previous researchers [6] and [7], the value of learning and momentum rate can be use as shows in Table 2.

Table 2: Determination of learning and momentum rate.

	Phase 1	Phase 2	Phase 3	Phase 4
Learning rate	0.9	0.7	0.5	0.4
Momentum rate	0.1	0.4	0.5	0.6

From thirty six data, thirty three data had been set as a training data while three data as a testing data. The evaluation for testing data is using linear correlation coefficient, r. The highest r from the training and testing will be selected. During training and testing, a series of trial and error process by varying the number of neurons in hidden layer. The range of neurons is between 2 to 40 and the network can approximate a target function of complexity if it has an enough numbers of hidden nods. The learning process is continues either the error reach 0.001 or 40,000 cycles is achieve. The successful trained networks would be trained again with different number of epoch. The final set of weights and biases would be obtained when one of the two criteria is met. In this study, two PC as the input layer and 1 output that is the actual housing demand. The graphs below show the network performance of testing for each phase.

The best linear correlation shows in phase two with r =0.8255 at neuron number twenty five. So the best NN model for Batu Pahat is 2 -25 -1 with learning and momentum rate 0.7 and 0.4 respectively.

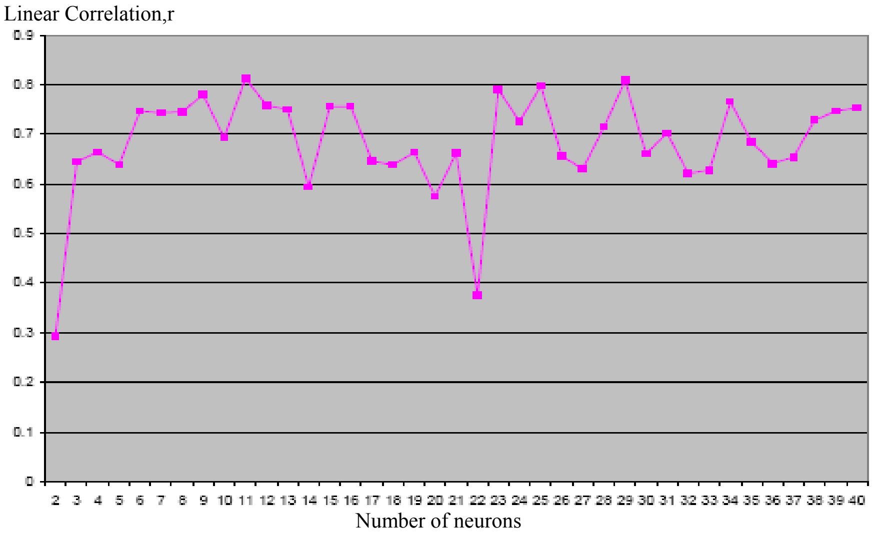

Figure 2: Network performance of testing with different number of neurons for phase 1.

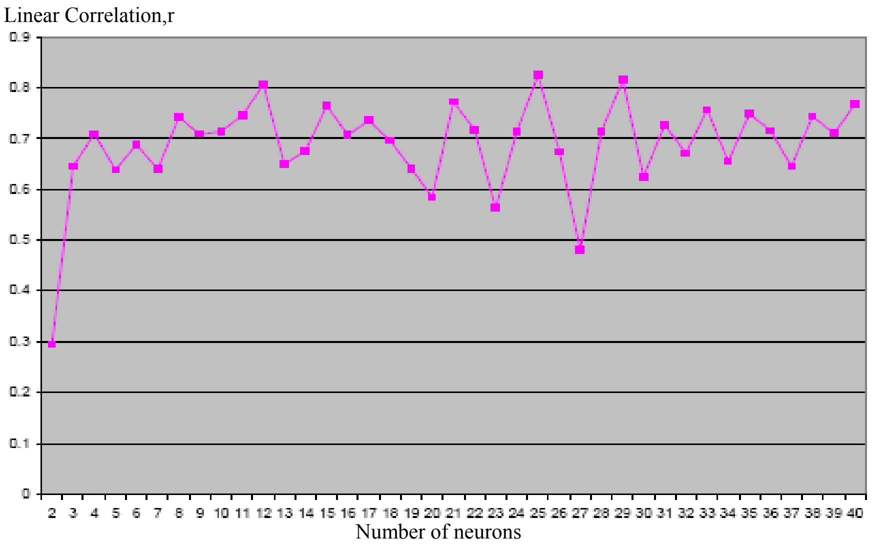

Figure 3: Network performance of testing with different number of neurons for phase 2.

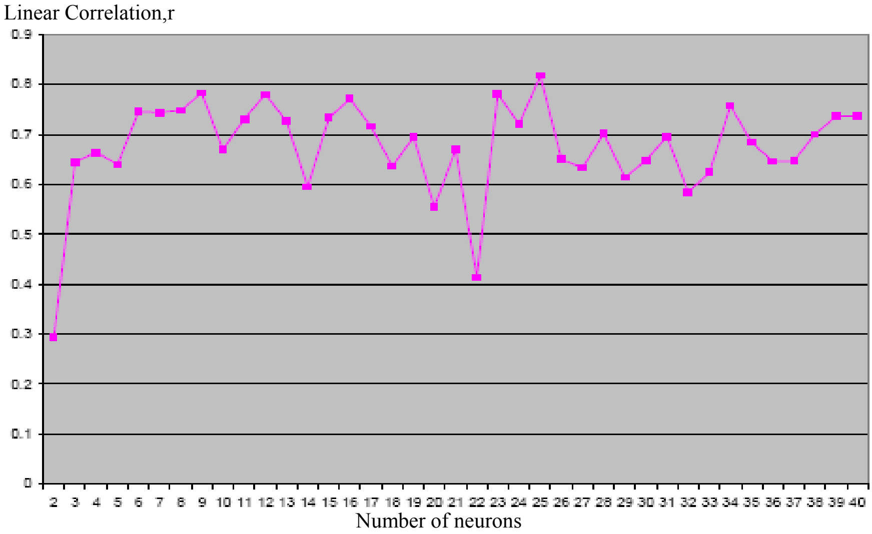

Figure 4: Network performance of testing with different number of neurons for phase 3.

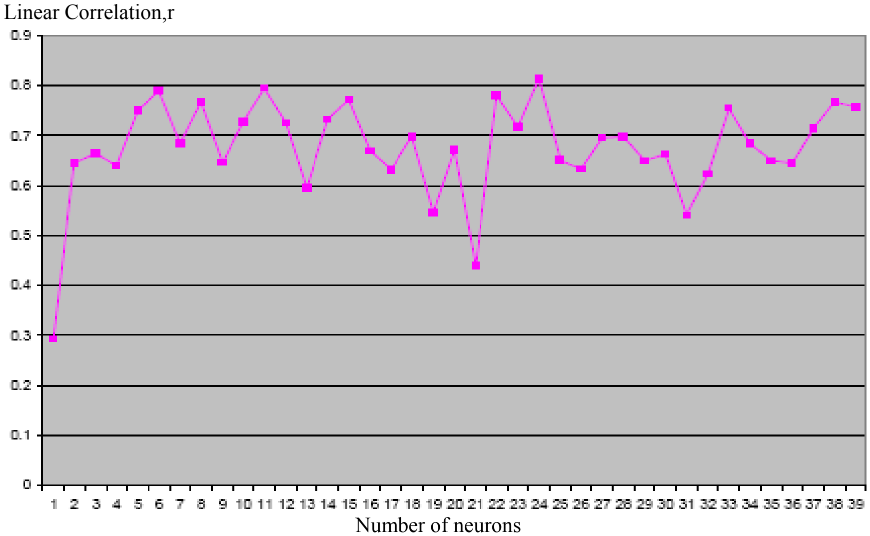

Figure 5: Network performance of testing with different number of neurons for phase 4.

The forecasted demand for low cost housing demand for Batu Pahat in October, November and December 2002 are 9, 10 and 11. MAPE is calculated to evaluate the forecasting performance. According to Sobri Harun [8], the forecasting result is very good if MAPE value is less than 10% while it is good if MAPE value is less 20%. The MAPE value of 19.7% shows that NN can forecast low cost housing demand in Batu Pahat quite good.

Table 3: Actual and forecasted data for 3 month ahead.

Time series	Actual data	Forecasted data	PE (%)
October 2002	12	9	25.0
November 2002	15	10	33.3
December 2002	11	11	0.0
MAPE			19.7

6 Conclusion

The results shows that NN capable to forecast low cost housing demand in Batu Pahat. By developing this model, it is hoped that there will be no more under or over construction of low-cost houses in Malaysia. It is also hoped that it can be helpful to the related agencies such as developer or any other relevant government agencies in making their development planning for low cost housing demand in urban area in Malaysia towards the future as there is no model have been created yet. Furthermore, there are a lot of advantages through the better planning of low cost housing construction, such as savings in expenditure, in time, manpower and also less paper.

References

[1] Ministry of Housing and Local Government Malaysia (1999). “Housing in the New Millennium – Malaysian Perspective.” http: www.kpkt.gov.my/jpn/artikel3.htm.
[2] Government of Malaysia (2006). The Ninth Malaysia Plan, 2006-2010. Kuala Lumpur: Percetakan Nasional Malaysia Berhad.
[3] Ong, Han Ching and Lenard, D (2002). “Partnership Between Stakeholders in the Provision of an Access to Affordable Housing in Malaysia.” FIG XXII International Congress. USA, Washington D.C.
[4] Government of Malaysia (2001). The Eight Malaysia Plan, 2001-2005. Kuala Lumpur: Percetakan Nasional Malaysia Berhad.
[5] Johnson, D.E (1998). “Applied Multivariate Methods for Data Analyst”. United States of America: Brooks/Cole Publishing Company.
[6] Sobri Harun, Nor Irwan Ahmat Nor, and Amir Hashim Mohd. Kassim (2001). “Rainfall-Runoff Modeling Using Artificial Neural Network”. *Malaysian Journal of Civil Engineering*, 13 (1). pp. 37-50.

[7] Khairulzan Yahya and Muhd Zaimi Abd. Majid (2002). "Comparative Study on Forecasting Demand on Low-cost House in Urban Areas Using Artificial Neural Networks and ARIMA Model." First International Conference on Construction in the 21st Century (CITC2002), USA, p. 687-694.
[8] Sobri Harun (1999) "Forecasting and Simulation of Net Inflows for Reservoir Operation and Management", PhD thesis., Universiti Teknologi Malaysia, Skudai, Malaysia.