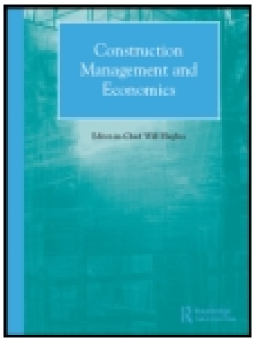

Construction Management and Economics

ISSN: 0144-6193 (Print) 1466-433X (Online) Journal homepage: www.tandfonline.com/journals/rcme20

Residential construction demand forecasting using economic indicators: a comparative study of artificial neural networks and multiple regression

Goh Bee Hua

Construction Management and Economics (1996) 14, 25–34

Residential construction demand forecasting using economic indicators: a comparative study of artificial neural networks and multiple regression

GOH BEE HUA

School of Building and Estate Management, National University of Singapore

Received 11 November 1994; revised 28 February 1995

In recent years, demand for residential construction has been growing rapidly in Singapore. This paper proposes the use of economic indicators to predict demand for residential construction in Singapore. At the same time, two forecasting techniques are applied, namely, Artificial Neural Networks (ANN) and Multiple Regression (MR), the former being a state-of-the-art technique while the latter a conventional one. A comparative study is carried out to determine whether the use of economic indicators with the application of the ANN technique can produce better predictions than with the MR method. A total of 12 economic indicators are identified as significantly related to demand for residential construction. Quarterly data from these 12 indicators are used to develop the ANN model. In order to assess the forecasting performance of this state-of-the-art technique, the same set of data is used to develop a conventional MR model. A comparison is made between the two models, in terms of their forecasting accuracy, by using a relative measure known as the Mean Absolute Percentage Error (MAPE). The forecasting error of the ANN model is found to be about one fifth of that derived from the MR model. The low MAPE values (less than 10%) obtained for both models also indicate that economic indicators may be used as reliable inputs for the modelling of residential construction demand in Singapore.

Keywords: Economic indicators, artificial neural networks, multiple regression, forecasting, demand.

Introduction

In land-scarce Singapore, residential properties have always been of great interest to the government and people. In recent years, the demand for such properties has continually increased, leading to a major surge in construction activity in this sector. The detailed figures of the percentage share of the residential sector in the total value of construction contracts awarded per year are shown in Table 1.

With the share of the residential sector growing, it becomes more important to look into alternative methods of predicting levels of demand for residential construction activity in Singapore.

The construction industry has always been closely related to the national economy. The significant role of the construction industry in the national economy has been highlighted by Turin (1969) and its importance further elaborated by Hillebrandt (1984). In the case of Singapore, the close relationship between its economy and its construction industry has been established chronologically by Ofori (1988). The importance of the construction industry's role in the economy of Singapore, in terms of its contribution to economic growth, is clear. At the same time, the performance of the economy also greatly affects that of the construction industry as 'the construction sector is greatly dependent on changes in the economy' since 'construction output is a response to the demand for buildings, and this is a derived demand for other products and services' (Briscoe, 1988, p. 6). Consequently, studies have shown that building and business cycles are closely related, with one having influence over the other. Bon (1989) relates building cycles to business or economic cycles and discusses how economic fluctuations affect fluctuations in building activity. Again, in Singapore, a close link between the construction cycle and the general business cycle was found to exist by Tan (1989). Therefore economic indicators, which are essentially measures of performance

Address for correspondence: Flat 951, William Goodenough House, Mecklenburgh Square, London WC1N 2AN, UK.

Table 1 Value of contracts awarded per year (in S$million)

Year	1986	1987	1988	1989	1990	1991	1992	1993
All sectors	3897.9	4040.3	3490.6	5500.7	8034.2	7873.7	12832.2	10586.7
Residential sector	1492.4	1638.8	1585.9	1737.1	2695.3	2684.9	5261.7	4638.2
% share of residential sector	38	40	45	32	34	34	41	44

Source: Ministry of National Development, Singapore, 1993.

of the economy, may be proposed as reliable inputs for the modelling of residential construction demand in Singapore.

Artificial Intelligence (AI) forecasting techniques such as neural networks have been receiving much attention lately. They have been cited to have the ability to learn like humans, by accumulating knowledge through repetitive learning activities. Their application in the prediction of economic indicators and financial indices has been demonstrated (White, 1988; Varfis and Versino, 1990; Windsor and Harker, 1990). However, their ability to forecast demand for construction remains to be seen in this study. In order to gauge the success of applying such AI techniques, a comparison needs to be made. Conventional regression techniques have often been used to establish forecasts in the construction industry. Akintoye and Skitmore (1994) produced regression models to predict UK private sector construction demand. Tang *et al.* (1990) also applied regression analysis to project values of total construction activities into the future in their study of demand for the Thai construction industry. In relation to construction tender prices, Flanagan and Norman (1983) used simple linear and curvilinear regressions to obtain extrapolated forecasts, while McCaffer *et al.* (1983) developed a regression model to predict future price movements. This conventional method can therefore serve as a benchmark against which to judge the performance of the neural networks technique in predicting demand for residential construction in Singapore.

Objectives of study

This paper purports to establish whether the use of economic indicators with the application of the ANN technique can produce better predictions of demand for residential construction in Singapore than with the conventional MR method. The following is a list of objectives set out for the study:

1. to identify and select economic indicators that are significantly related to demand for residential construction in Singapore;
2. to apply the ANN technique to develop a forecasting model using values of the selected indicators as the training set of historical data;
3. to apply the MR method to develop another forecasting model, which serves as a benchmark against which to judge the performance of the ANN model, using the same set of data;
4. to carry out a comparative study to evaluate the forecasting performance of the two models.

Data and methodology

The first stage of the study involves a comprehensive literature search of a list of economic indicators often associated theoretically with demand for residential construction. Past quarterly data of these theoretical indicators are abstracted from statistics yearbooks, in the form of national level statistics, to create the time-series dataset. The data collected dates from the third quarter of 1975 to the fourth quarter of 1993, which in total comprises 74 quarterly records within the span of 19 years. These data are entered into a mainframe computer for statistical analysis.

The second stage is carried out using the STEPWISE procedure, available on the SAS statistical package, to select the statistically significant indicators. This statistical variable selection method adds and deletes variables one at a time until a 10% significance level has been met. The significant indicators identified through this process serve as the input and independent variables for the ANN and MR models respectively. It essentially means that the indicators used are the same for both methods.

The third stage comprises the training and testing of the ANN model with the selected indicators. When developing the ANN model the intention is to utilize more input data for training owing to the short span of quarterly data available. Therefore the first 71 quarterly records, commencing from the third quarter of 1975 to the first quarter of 1993, are used for training and developing the model, while the remaining three serve as an independent dataset for testing and evaluating the forecasts. This implies that the predictive capability of the ANN model is tested on a dataset it has not seen during training. The MR model is also developed and tested during this stage using the same sets of data.

The final stage involves the computation of the prediction errors for both models. This allows a comparison of their forecasting accuracy.

Indicators of demand for residential construction

Some indicators that influence the level of demand for residential construction have been identified by several writers (Hillebrandt, 1974; Lange and Mills, 1979; Shutt, 1982; Stone, 1983; Briscoe, 1988; Ofori, 1990; Tang *et al.*, 1990; Raftery, 1991). The indicators may be economic or social and their individual significance varies from one country to another. The following is a list of such indicators.

1. National income per capita.
2. General demand for construction.
3. Size of population.
4. Rate of household formation.
5. Interest rate.
6. Property price.
7. Levels of supply of residential property.
8. Disposable income.
9. Economic growth.
10. Level of unemployment/employment.
11. Existing housing stock.
12. Rate of inflation.
13. Construction cost.
14. Mortgage credit availability/supply.
15. Household personal savings.

They form the list of theoretical indicators for the purpose of this study. Modelling can be based on the relationship between these indicators and demand for residential construction over time.

Theory of ANN

ANNs form a class of pattern recognition and classification devices that model, to varying degrees of exactness depending on their specific application, the workings of the central nervous system. An ANN is made up of a complex network of artificial neurons or processing elements (PEs). Like a neuron, the PE performs three basic functions. Firstly, it receives inputs from other PEs through weighted links; secondly, it processes these inputs; and finally, it outputs the results to other PEs. The processing stage involves the computation of a weighted sum of the inputs and then passing the sum through a mathematical transfer function which acts as a non-linear threshold. The relationship between the input signals and the PE output is given in the following equation.

$$X_{pj} = \sum_{i=1}^{n} W_{pji} Y_{pi}$$

where X_{pj} is the net input to PE(j) following presentation of pattern p; Y_{pi} is the value of input signal from PE(i) of input pattern p; W_{pji} is the weight from PE(i) to PE(j) following presentation of pattern p; and n is the number of PE inputs to PE(j).

The output from PE(j) is a function of the transfer function as follows:

$$a_j = f(t_j)$$

where a_j is the output signal from PE(j) and $f(t_j)$ is the transfer function of PE(j).

Figure 1 shows an example of an artificial neuron performing the three basic functions.

Backpropagation is one of the most widely used training algorithms. It is a supervised learning procedure adopting the error-correction rule. It is capable of learning internal representations involving the presentation of a set of pairs of input and output patterns. Using only internal computation, it can be applied to multilayered neural networks with hidden units, which are neither input nor output units. The application of the generalized delta rule (Rumelhart *et al.*, 1986) to backpropagation involves two stages.

During the first stage, the input is presented and propagated forward to produce the output for each unit. This output is then compared with the desired output to produce the error signal. The second stage involves a backward pass through the network, during which the error signal is passed to each unit in the network. This

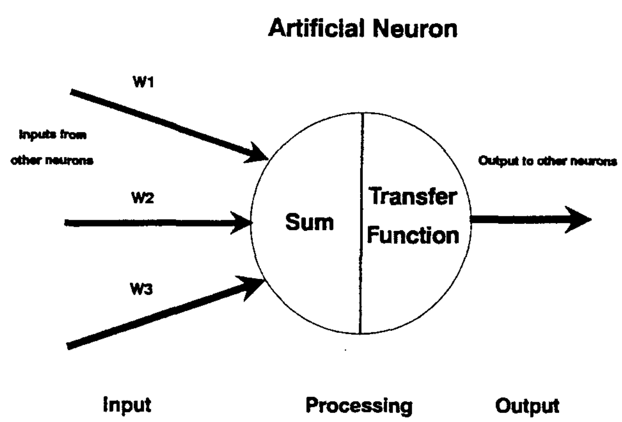

Figure 1 Illustration of an artificial neuron performing its three basic functions

backward pass allows recursive computation of the error signal and appropriate weight changes to be made. These stages are repeated with each set of training data ultimately to find an optimum set of weights so as to minimize the mean sum of squared error between the desired and actual outputs of a multi-layered neural network.

In the training process, the network learns the relationship between a given set of inputs and outputs. After training, if the network is given a set of inputs, it will 'correctly' give the output. Such a network is known as a trained network.

Selection of significant indicators

A list of theoretical indicators has been identified during the first stage of this study. However, owing to the differences in terminology between theory and statistics, the process of data abstraction from statistics yearbooks requires the identification of a comparable list of national level statistics. Table 2 lists, respectively, the equivalent national level statistics for each theoretical indicator.

The national level statistics for Gross Fixed Capital Formation (GFCF) for residential buildings are used to represent demand for residential construction in Singapore. This choice is made on the basis that they best define and measure this type of demand, both theoretically and statistically.

The STEPWISE procedure selected 12 indicators that are significantly related to GFCF for residential buildings. They met the 10% significance level as specified in the procedure. The list of indicators is as follows:

1. Per capita GDP (PCAPGDP).
2. GFCF (construction and works) (GFCFCON).
3. Real GDP (REALGDP).
4. Building material price index (BLGMATPI).
5. Money supply (M2).
6. Money supply (savings and others) (MSAV).
7. CPF withdrawals (home ownership) (CPFWITHD).
8. Prime lending rate (PRIMELR).
9. Consumer price index (CPI).
10. Property price index (residential) (PPIRES).
11. Labour force (LABFOR).
12. Unemployment rate (UNEMPRT).

The ones that did not fall within the significance level are: number of marriages registered, number of planning approvals for residential projects and total resident population.

Table 2 Theoretical and statistical classification of indicators

Theoretical indicators	National level statistics
Demand for residential construction	Gross Fixed Capital Formation (GFCF) (residential buildings)
National income per capita	Per capita Gross Domestic Product (GDP)
General demand for construction	GFCF (construction and works)
Size of population	Total resident population
Rate of household formation	Number of marriages registered
Interest rate	Prime lending rate
Property price	Property price index (residential)
Levels of supply of residential property	Number of planning approvals (residential projects)
Disposable income	Not available[1]
Economic growth	Real GDP
Level of unemployment/employment	Unemployment rate, labour force
Existing housing stock	Not available
Rate of inflation	Consumer price index
Construction cost	Building material price index
Mortgage credit availability/supply	Money supply (M2), Central Provident Fund (CPF) withdrawals (home-ownership)[2]
Household personal savings	Money supply (savings and others)

[1] There are no statistics available in the statistics yearbooks that represent disposable income. The category that relates closest to income tax matters is 'Taxpayers, income assessed and assessed tax'. However, it was not used because it does not measure personal disposable income directly.

[2] In Singapore, the government recognizes that residential property is a good investment and hence has allowed members contributing to the CPF to withdraw their accumulated compulsory savings to purchase such properties. The schemes for public and private residential properties started in 1968 and 1981 respectively. The CPF home ownership schemes serve as important sources of finance for these purchasers.

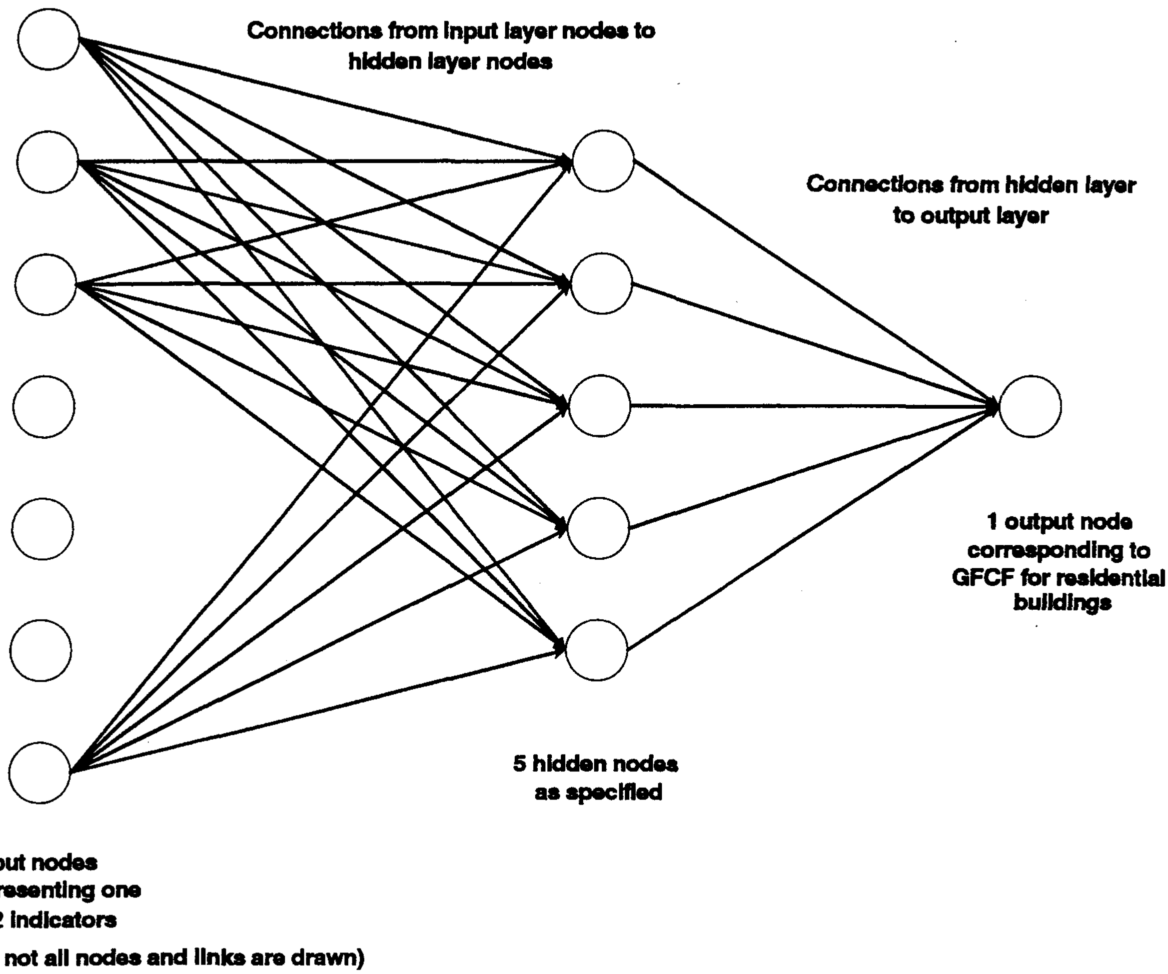

Figure 2 Network architecture adopted for the ANN model

Architecture of the ANN model

A typical three-layered backpropagation neural network architecture is chosen for the ANN model. It consists of an input layer with 12 nodes, a hidden layer with five nodes and an output layer with one node. Figure 2 depicts the network architecture adopted for the ANN model. Although there are no rules governing the design of ANN architecture, a basic understanding of the concepts of the different layers within the network helps to justify the choice of specifications.

The input layer presents data to the network. The size of the input layer, or number of nodes, is determined by the number of data sources. For example, there are 12 indicators used as input variables in the study and so the size of the input layer is 12 nodes. Hidden layers act as layers of abstraction, pulling features from inputs. Increasing the number of sequential hidden layers augments the processing power of the neural network but significantly complicates training and intensifies 'black-box' effects, which results in errors being more difficult to trace. Adding hidden layers will increase both the time and the number of training examples necessary to train the network properly. A rule of thumb is to start with one hidden layer and add more as required. The size of the hidden layer can be ascertained by applying another rule of thumb, which is if there is only one hidden layer, a suitable initial size is 75% of the size of the input layer (Bailey and Thompson, 1990). The general concept is that too few hidden nodes renders the mapping process insufficient, while too many incurs higher training time or allows the network to memorize the training examples rather than extract the general features that will allow it to handle cases it has not seen during training. However, although in this case the application of the rule of thumb works out the size of the hidden layer to be nine nodes (75% of 12 input nodes), experiments with fewer nodes reveal the optimum number to be five. A hidden layer with five nodes is found to be capable of achieving a balance between the objectives of accuracy and generalization. Besides, less time is required to train a network with fewer hidden nodes and interconnections. Lastly, the output layer comprises one node which corresponds with the output variable, GFCF for residential buildings.

The MR model

The 12 indicators used in the ANN model are adopted as independent variables in the regression analysis to estimate the dependent variable, GFCF for residential buildings (GFCFRES). The demand function is expressed as follows:

GFCFRES = f(PCAPGDP, GFCFCON, REALGDP, BLGMATPI, M2, MSAV, CPFWITHD, PRIMELR, CPI, PPIRES, LABFOR, UNEMPRT)

Results and discussion

Forecasting results of the ANN model

The ANN model was built using the backpropagation training algorithm. In the training process, the network was first randomly assigned weights in its input links. Then each set of inputs (that is, the 12 indicators) and output (GFCF for residential buildings) was presented to the network one at a time. In this case, 71 sets were used in the training stage. After each presentation of an input/output pair, the network changed by a small amount the weights of the input links of each node, in such a way that the output may be closer to the one presented. By repeating this process many times, the network was eventually trained and ready to be tested.

In order for the model to achieve internal validity, it has to be trained more than once, each time using different randomly selected initial weights. In this study, the ANN model was trained and tested twice. After each training, it was tested to generate a forecast and therefore two independent forecasts were produced. The purpose of training the network more than once is to test whether the ANN model was able to produce consistent forecasts. Large deviations in independent forecasts often indicate that the network is either insufficiently trained or overtrained. They can be distinguished by examining the average errors. A network is considered insufficiently trained if it produces an uneven plot of its average error values. More data will be required for training or a suitable resampling method has to be adopted if there is limited data available. An overtrained network also has a poor ability to forecast. The process of overtraining would have caused the network to memorize the training data, leading to a failure to generalize on new data. The size of deviation between the two independent forecasts will be examined later in the study. In the meantime, results of the two forecasts for the last three quarters are shown in Table 3.

Table 3 Results of forecast obtained by the ANN model

Qtr/Yr	Actual	Forecast 1	Error 1	Forecast 2	Error 2
2Q93	4819	4897	−78	4936	−117
3Q93	5113	5206	−93	5125	−12
4Q93	5407	5450	−43	5354	+53

Note: The figures represent GFCF for residential buildings (in S$million) obtained by interpolating annual time-series data. To derive the actual and forecasted quarterly values, the figures have to be divided by 4.

Forecasting results of the MR model

The statistical analysis was first carried out using the REG procedure. This time, the forced entry option was adopted where all the variables specified in the demand function were entered in a single step. The stepwise option was not utilized again because the indicators were pre-selected to meet the 10% significance level. There was no need to repeat this selection process.

The results of the REG procedure revealed that all the variables entered were statistically significant at the 5% significance level, and the R-square value obtained was very high (0.99907). However, these initial results seemed too optimistic. The presence of autocorrelated errors was therefore suspected. When time-series data are used for fitting regression models, often the residual or error term is not independent across time. Consequently, when residuals are strongly autocorrelated, the significance levels for the regression coefficients are wrong and the R-square value does not accurately summarize the explanatory power of the independent variables. The Durbin–Watson statistic was used to test for the presence of first-order autocorrelation and the derived value of 0.8872 indicated positively correlated residuals. Occasionally, higher order autoregressive processes for the errors are entertained, 'but this seems to be the exception rather than the rule in current applied work' (Granger and Newbold, 1987, p. 192). Based on this, it was assumed that the problem of autocorrelation of errors was confined to first-order processes only. In order to remedy this problem, an alternative analysis methodology provided by the AUTOREG procedure was adopted. This procedure is able to estimate parameters in regression models when the data are time-series and the error term is an autoregressive process. The results of the analysis are shown in Table 4.

By summarizing the results from Table 4, the derived

Table 4 Estimated demand function for residential construction (dependent variable: GFCFRES)

Independent variables	Coefficients	SEE	t-value	Sig. t
BLGMATPI	−42.02	7.72	−5.44	0.0000
CPFWITHD	−0.79	0.29	−2.68	0.0098
CPI	−54.29	13.83	−3.93	0.0003
GFCFCON	0.643	0.025	25.24	0.0000
LABFOR	−4.17	1.52	−2.74	0.0084
M2	−0.023	0.002	−9.72	0.0000
MSAV	−0.092	0.026	−3.53	0.0008
PCAPGDP	0.557	0.074	7.53	0.0000
PPIRES	−5.55	2.39	−2.32	0.0241
PRIMELR	41.15	31.52	3.04	0.0836
REALGDP	0.365	0.132	2.77	0.0078
UNEMPRT	51.54	15.78	3.27	0.0742
(Constant)	5966.18	467.28	12.77	0.0000

R-square = 0.9685; adjusted R-square = 0.9571; Durbin–Watson value = 1.8766.
Final Parameter: Number of residuals = 71; Standard error = 41.519267.
Analysis of Variance:

	DF	Adj. sum of squares	Residual variance
Residuals	57	88677.374	1723.8495

estimated demand function for residential construction in Singapore is expressed as:

$$\begin{aligned} GFCFRES = {} & 5966.18 - 42.02(BLGMATPI) \\ & - 0.79(CPFWITHD) - 54.29(CPI) \\ & + 0.643(GFCFCON) - 4.17(LABFOR) \\ & - 0.023(M2) - 0.092(MSAV) \\ & + 0.557(PCAPGDP) - 5.55(PPIRES) \\ & + 41.15(PRIMELR) + 0.365(REALGDP) \\ & + 51.54(UNEMPRT) \end{aligned}$$

Another problem often associated with multiple regression analysis is the existence of multicollinearity. It is defined as a high degree of multiple correlation among several independent variables. It occurs because too many variables have been put into the model, and a number of the variables may measure similar phenomena. It often results in coefficient estimates that are not statistically significant or have incorrect signs or magnitude (Myers, 1986). From Table 4, it is evident that the parameter estimates for all the independent variables meet the 10% significance level. Therefore, at this stage, only the signs of the coefficients had to be verified.

Among all the variables, the ones with positive signs of coefficient were GFCF for construction and works; per capita GDP; prime lending rate; real GDP; and unemployment rate. Those with negative signs were: building material price index; CPF withdrawals for home-ownership; consumer price index; labour force; money supply; money supply from savings and others; and property price index for residential type.

According to economic theory, prime lending rate is expected to have an inverse relationship with demand for residential construction but the reverse was obtained from the analysis. This departure from theory may be justifiable in the case of Singapore as it was seen that real GDP for construction grew even with soaring interest rates, most evidently between 1979 and 1982. In Singapore, the public sector accounts for easily over 50% of the annual construction value, and funds are made available through the budget process. Among the various types of buildings, residential construction had as high as a 57% share of total public sector building construction demand in 1990 and 1991 (Construction Industry Development Board, 1991). This was already a drop from 61% in 1988 and 58% in 1989 (Construction Industry Development Board, 1990). The large share of residential construction in the public sector is attributable to the ongoing public housing scheme implemented by the government which has to date housed about 90% of the total population. This may also help to account for the inverse relationship obtained for money supply. At the same time, in the private sector, there are some individuals who view residential investment as a viable hedge against inflation especially in a land-scarce country. They would purchase residential properties even when interest rates were high. This was witnessed in the early 1980s when prime lending rates reached a peak of 13% and demand for residential construction was still increasing.

The analysis also established a positive relationship between demand for residential construction and the rate of unemployment. This, again, is contrary to theoretical belief. Strictly speaking, the Singapore economy has not faced any serious unemployment situation after the early 1960s. By the early 1980s, when many economists came to believe that full employment meant an unemployment rate close to 6%, unemployment in Singapore dropped to a low of 2.6% in 1982 and stood at 3.3% in 1988 (Toh and Low, 1990). In 1986 the rate rose from 4% in the first quarter to 6.5% in the fourth quarter. It was a period of general economic recession in Singapore. However, the government immediately intervened and injected about S$1.3 billion, which constituted about 90% of the total value of building contracts awarded in the residential sector (Construction Industry Development Board, 1988), into public sector housing in the hope of stimulating demand for construction in the private sector and also revitalizing the general economy. The effect of this counter-cyclical spending by the government during a period of rising unemployment may have contributed to the positive relationship found.

The inverse relationship obtained between CPF withdrawals for home ownership and demand for residential construction was also unexpected. Since the withdrawal scheme was set up by the government to encourage home ownership, an increase in withdrawal should imply a greater demand for residential construction. Ever since the scheme came into effect in the late 1960s, it has been very popular and so despite the cyclical effects of the economy, withdrawals have generally been increasing. But since withdrawals can be made to purchase existing and new properties, its persistent rising trend may, in the long run, have gone against declining ones for new residential construction during periods of construction slump. This may have led to the overall inverse relationship.

Although initially, some of the indicators may seem to have opposing signs, indicating the presence of multicollinearity, they were later justified when related to past economic situations in Singapore. Therefore, it was assumed that no serious problem of multicollinearity was present in the regression model. Besides, 'whilst estimators of the regression coefficients can be quite imprecise in the presence of serious multicollinearity, good forecasts may still result from the overall fitted model, provided historical correlation patterns among the independent variables continue to hold in the future' (Newbold and Bos, 1990, p. 29). Considerations for multicollinearity only become critical 'when it is important to ascertain the structure of the relationship of the response to the various independent variables' (Freund and Littell, 1986, p. 76). The forecasts generated by the MR model are shown in Table 5.

Table 5 Results of the forecast obtained by the MR model

Qtr/Yr	Actual	Forecast	Error
2Q93	4819	5205	−386
3Q93	5113	5470	−357
4Q93	5407	5730	−323

Note: The figures represent GFCF for residential buildings (in S$million) obtained by interpolating annual time-series data. To derive the actual and forecasted quarterly values, the figures have to be divided by 4.

Results of comparative study

The forecasting results of the ANN and the MR models were compared using relative measures of forecasting accuracy dealing with percentage errors. The forecasted figures from Tables 3 and 5 were used to calculate the values for these measures. The measures used in the comparative study are:

1. Percentage error (PE)

$$PE_t = \left(\frac{X_t - F_t}{X_t}\right)100\%$$

where X_t = actual value at period t; and
F_t = forecast value at period t.

2. Mean percentage error (MPE)

$$MPE = \sum_{i=1}^{n} PE_i/n$$

3. Mean absolute percentage error (MAPE)

$$MAPE = \sum_{i=1}^{n} | PE_i |/n$$

The results of the comparative study are given in Table 6.

As shown in Table 6, the MAPEs of the two forecasts generated by the ANN model were found to be 1.21% and 1.41%. Their difference was only 0.2%. For the MR model, the MAPE was found to be 6.99%.

Summary of findings

Firstly, among the 15 theoretical indicators identified in the literature search, 12 were found to be significantly related to demand for residential construction in Singapore. Those that did not meet the 10% significance level were: number of marriages registered, number of planning approvals for residential projects, and total resident population. They were the national level statistics for

Table 6 Results of the comparative study

Measures of accuracy	Forecast 1 by ANN model (%)	Forecast 2 by ANN model (%)	Forecast by MR model (%)
PE for 2Q93	−1.62	−2.43	−8.01
PE for 3Q93	−1.82	−0.23	−6.98
PE for 4Q93	−0.80	0.98	−5.97
MPE	−1.41	−0.56	−6.99
MAPE	1.41	1.21	6.99

rate of household formation, level of supply of residential properties, and size of population respectively. Although these indicators have been known theoretically to influence demand for residential construction, their significances were disputed by the statistical analysis. However, this does not necessarily reduce their theoretical importance. A major limitation of any statistical analysis is that the outcome is largely dependent on both the quality and the quantity of data used. This is particularly true when dealing with archival as opposed to empirical data, as there is less control over the recording and updating of data. The competence of the personnel maintaining the data banks has to be relied upon. As more data become available, the statistical significances of these indicators may perhaps improve.

Secondly, the low MAPE values (those less than 10%) obtained for both models imply that economic indicators may be used as reliable inputs for the modelling of residential construction demand. This finding provides further justification for the conclusions drawn by Bon (1989) and Tan (1989) that a close link exists between building and economic cycles.

Thirdly, the small deviation of 0.2% MAPE derived from the two forecasts of the ANN model indicates that the network has achieved internal validity and is capable of producing consistent forecasts.

Finally, the forecasting accuracy of both models can be reflected by their MAPE values. The MAPEs for the ANN and the MR models were found to be 1.31% (average) and 6.99% respectively. Although both models were able to produce accurate forecasts (less than 10% error) the ANN model did outstandingly better. This finding allows external validity to be achieved by the ANN model. Therefore, it is established that the ANN technique can produce more accurate forecasts of demand for residential construction in Singapore than the MR method.

General discussion

In relative terms, the ANN model generated about a fifth of the forecasting error of the MR model. This is largely due to the nature of neural network models. They are designed to capture the non-linear relationship between the input and output variables automatically, without having to specify the non-linear terms to fit the data. Since the demand for construction, in general, is heavily influenced by changes in the economy which often lead to pronounced fluctuations, the adoption of a non-linear representation for demand should well be a more realistic and accurate one. However, it does not imply that regression methods do not have their merits. They are essentially causal methods which allow the relationship between the dependent and the independent variables to be analysed and explained. The absence of this explanatory capability in a neural network model has traditionally termed it a 'black box'. Besides, non-linear terms can be manually specified in regression models. But such judgmental interventions would require the forecaster to have an intimate knowledge of past trends of the variables in order to be objective and precise in their specification. Otherwise they could be regarded as bias and intuitive, undermining the validity of any model specified.

Moreover, it would be appropriate at this juncture for the construction industry to advance further into the area of AI. The use of AI techniques has been lacking in studies relating to macroeconomic aspects of the industry, which primarily focus on broad aggregates and analyse relationships of great complexity. General economists have made a headstart in applying AI techniques to macroeconomic analyses and modelling as they realized that the AI field can have far-reaching consequences for their own activities (Pau, 1986). The field of AI has been described by Moss and Rae (1992, p. 1) as one 'concerned with determining courses of action in problem spaces which are too large or complex for the application of constrained-optimization algorithms'. They explained that 'since all neoclassical economic models describe decision-making as a process of constrained optimization, it would appear that AI techniques offer the possibility of extending our representations of economic behaviour to take complexity into account'. One of the many successful applications in this area is a study undertaken by Varfis and Versino (1990) who used artificial neural networks to carry out univariate economic time-series forecasting, and satisfactory performance was obtained when compared with the conventional Box and Jenkins approach.

Neural network techniques have already been applied to construction management. Applications in the sequencing, simulation and optimization of construction processes have been made (Flood, 1989, 1990; Flood and Kartam, 1993). Gaarslev (1991) also used the technique to examine relationships between variables in the construction bidding process. In relation to the forecasting of construction bids, McKim (1993) discovered that a neural network can outperform traditional statistical methods in both bivariate and multivariate systems. Further adoption of this state-of-the-art technique in the construction industry will have to rely on more successes achieved in a wider range of applications.

Acknowledgements

The author wishes to thank Professor John Andrews and two anonymous referees for their helpful comments on an earlier version of the article.

References

Akintoye, A. and Skitmore, M. (1994) Models of UK private sector quarterly construction demand. *Construction Management and Economics*, 12, 3–13.

Bailey, D. and Thompson, D. (1990) How to develop neural-network applications. *AI Expert*, 5(6), 38–47.

Bon, R. (1989) *Building as an Economic Process*. Prentice-Hall, Englewood Cliffs, NJ.

Briscoe, G. (1988) *The Economics of the Construction Industry*. Mitchell, London.

Construction Industry Development Board (1988) *Sectoral Analysis*. CIDB Construction Economic Report, Singapore, 7–10.

Construction Industry Development Board (1990) *Sectoral Analysis*. CIDB Construction Economic Report, Singapore, 23–8.

Construction Industry Development Board (1991) *Update of Construction Demand in 1991*. CIDB Construction Economic Report, Singapore, 2nd quarter, 3–9.

Flanagan, R. and Norman, G. (1983) The accuracy and monitoring of quantity surveyors' price forecasting for building work. *Construction Management and Economics*, **1**, 157–80.

Flood, I. (1989) 'A neural network approach to the sequencing of construction tasks, in *Proceedings of the 6th International Symposium on Automation and Robotics in Construction*, San Francisco, CA, 204–11.

Flood, I. (1990) Simulating the construction process using neural networks, in *Proceedings of the 7th International Symposium on Automation and Robotics in Construction*, Bristol, UK, pp. 374–82.

Flood, I. and Kartam, N. (1993) The use of artificial neural networks in construction, in *Proceedings of the Conference on Organisation and Management of Construction – The Way Forward*, CIB W-65, Trinidad, West Indies, September, 81–9.

Freund, R.J. and Littell, R.C. (1986) *SAS System for Regression*. SAS Institute Inc, USA.

Gaarslev, A. (1991) Neural networks – technique tested on classical bidding strategic data set, in *Proceedings of the Symposium on Management, Quality and Economics in Buildings*, CIB W-65, Lisbon, Portugal, 1489–97.

Granger, C.W.J. and Newbold, P. (1987) *Forecasting Economic Time Series*. 2nd edition, Academic Press, New York.

Hillbrandt, P. (1974) *Economic Theory and the Construction Industry*. Macmillan, London.

Hillebrandt, P. (1984) *Analysis of the British Construction Industry*. Macmillan, London.

Lange, J.E. and Mills, D.Q. (1979) *The Construction Industry*. D.C. Heath & Co., Canada.

McCaffer, R., McCaffrey, M.J. and Thorpe, A. (1983) The disparity between construction cost and tender price movements. *Construction Papers*, 2, 17–28.

McKim, R.A. (1993) A neural network application to construction management, in *Proceedings of the Conference on Organisation and Management of Construction – The Way Forward*, CIB W-65, Trinidad, West Indies, September, 103–27.

Moss, S. and Rae, J. (1992) *Artificial Intelligence and Economic Analysis: Prospects and Problems*. Edward Elgar Publishing, USA.

Myers, R.H. (1986) *Classical and Modern Regression with Applications*. PWS & Kent Publishing, Boston, USA.

Newbold, P. and Bos, T. (1990) *Introductory Business Forecasting*. South Western Publishing, USA.

Ofori, G. (1988) Construction industry and economic growth in Singapore. *Construction Management and Economics*, **6**, 57–70.

Ofori, G. (1990) *The Construction Industry: Aspects of its Economics and Management*, Singapore University Press, NUS.

Pau, L.F. (1986) *Artificial Intelligence in Economics and Management*. Elsevier Science, Amsterdam.

Raftery, J. (1991) *Principles of Building Economics*. BSP Professional Books.

Rumelhart, D.E., Hinton, G.E. and Williams, R.J. (1986) Learning internal representation by error propagation, in *Parallel Distributed Processing*, Rumelhart, D.E. and McClelland (eds), vol 1, MIT Press, Cambridge, MA, pp. 318–62.

Shutt, R. (1982) *Economics for the Construction Industry*. Longman Group, New York.

Stone, P.A. (1983) *Building Economy: Design, Production and Organisation – A Synoptic View*. 3rd edition, Pergamon Press.

Tan, W. (1989) Subsector fluctuations in construction. *Construction Management and Economics*, 7, 41–51.

Tang, J.C.S., Karasudhi, P. and Tachopiyagoon, P. (1990) Thai construction industry: demand and projection. *Construction Management and Economics*, **8**, 249–57.

Toh, M.H. and Low, L. (1990) *An Economic Framework of Singapore*. McGraw-Hill, Singapore.

Turin, D.A. (1969) Construction Industry, based on the *Proceedings of the International Symposium on Industrial Development* held in Athens in Nov–Dec 1967, New York, Monograph no. 2.

Varfis, A. and Versino, C. (1990) Univariate economic time series forecasting by connectionist methods, in *Proceedings of the International Neural Network Conference* organized by the INNS, IEEE Neural Networks Council, Paris, France, 9–13 July, **1**, 342–5.

White, H. (1988) Economic prediction using neural networks: the case of IBM daily stock returns, in *Proceedings of the IEEE International Conference on Neural Networks*, San Diego, California, 24–27 July, **2**, 451–8.

Windsor, C.G. and Harker, A.H. (1990) Multi-variate financial index prediction – a neural network study, in *Proceedings of the International Neural Network Conference* organized by the INNS, IEEE Neural Networks Council, Paris, France, 9–13 July, **1**, 357–60.