Theme:

Title:

Forecasting Construction Industry-level Total Factor Productivity Growth using Neural Network Modeling

Author(s):   Mao Zhi, Goh Bee Hua, Wang Shouqing and George Ofori

Institution(s):   National University of Singapore

E-mail(s):   brep9098@nus.edu.sg

Abstract: *Total Factor Productivity (TFP) is widely recognised as a better indicator than Labour Productivity and Multi-Factor Productivity to represent industry-level productivity performance. Productivity is the key determinant of a nation's standard of living and an industry's competitiveness. As such, the ability to predict trends in TFP growth in the construction industry is very important. The factors influencing TFP growth in the construction industry are complicatedly interrelated. This fact made the conventional regression method highly inadaptable to such complex multi-attribute nonlinear mappings.*

As an AI information-processing tool, the artificial neural network (ANN) system has been proven to be a powerful approach to solving complex nonlinear mappings with higher accuracy than regression methods. However, so far, there has been little application of ANNs in predicting TFP growth in the construction field. This study will for the first time, apply the concepts of ANNs to develop a model to forecast the TFP growth in the case of the construction industry of Singapore.

Macro-level information processing models are useful in monitoring and predicting the performance of the construction industry as a whole. With the need to manage construction performance information at all three levels, namely, industry, firm and site, this study looks specifically at developing an 'intelligent' model for forecasting industry-level productivity.

Meanwhile, using the same set of data, a model developed by the Multiple Linear Regression method will serve as a benchmark to judge the performance of the ANN model. The ANN model, compared with the traditional regression model, would be expected to have better forecasting ability for TFP growth in the construction industry, in terms of accuracy.

Keywords: *Total factor productivity, growth accounting, artificial neural networks, multiple regression, forecasting*

Introduction

Recognized as a better indicator than labour productivity and Multi-factor productivity to evaluate the efficiency in use of resources in the construction industry, total-factor productivity (TFP) has aroused great interest in recent years and become a key current consideration in Singapore's economy in general.

Theoretically, TFP is a relevant measure for technological change by measuring the real growth in production value, which cannot be explained by changes in the input of labour, capital and intermediate input (Pedersen, 1994). However in reality, due to lack of data, most of the works in TFP measurement are limited on the basis of two factor inputs, usually capital and labour. Thus the term "multifactor-

productivity" (MFP) is prevailingly used instead of TFP. Using translog value-added production function and Tornqvist index, the major statistical agencies (Department Statistics of Singapore, US Bureau of Labour Statistics, Australian Bureau of Statistics) have published MFP indices at both national and industry level. In Singapore, by using a Tornqvist index, Tan (2000) estimated the value-added two-factor (capital and labour) productivity of the Singapore construction industry between 1980 and 1996.

The increasingly role of materials input on productivity force researchers to reconstruct their productivity measurement model to incorporate all factor inputs, labour, capital and intermediate inputs. Chau and Walker developed (1988; 1993) two indirect approaches to measure TFP from construction cost and price data. By defining TFP as the ratio of quantity of output to quantity of aggregate input calculated by an input aggregator function, assuming Hicks Neutral Technical change and constant to return, the TFP change be represented by Chau as the sum of weighted change in the price of factor inputs: labour, capital and materials, minus the change in the price of output. However, the fact of existence of bias in technology change made the Chau's neutral technology change assumption restrictive and vulnerable. A more appropriate and solid measurement of TFP that is able to release the restriction on Neutral Technology Change while incorporating all the factor inputs is that of Jorgenson et.al (1987).

Meanwhile, the government of Singapore has set a target of attainment of MFP growth rate of 2 per cent per annum. As such, the ability to predict trends in TFP growth in the construction industry is very important for monitor and control purpose. Conventionally, regression techniques have been used to develop forecast model of productivity growth in construction field. Dacy (1965) established a simple linear regression model for estimating labour productivity indexes with three independent variables.

Owing to the possibility that factors influencing construction industry-level TFP growth are highly interactive, this implied that the traditional regression models may be unadaptable to such complex multi-attribute non-linear function approximation problem. As an AI information-processing tool, the artificial neural network (ANN) system has been proven to be a powerful approach to solving complex nonlinear mappings with higher accuracy than regression methods. Their applications in construction, as summarized by Moselhi et al. (1991), includes predicting project cash flow and costs, risk analysis, decision making, resources optimisation, prediction of tendering outcomes, classification and selection. So far, there has been little application of ANNs in predicting TFP growth in the construction field. Their applicability can be judged in terms of forecasting accuracy compared with the conventional regression methods.

Research Objectives

The main objective of this paper is to develop a relatively accurate forecasting model for TFP growth of construction industry of Singapore. To achieve this goal, four sub-objectives are involved and listed bellow:

1. to compute industry-level TFP growth for construction in Singapore by using Jorgenson's TFP measurement,
2. to identify and select factors that are significantly related to TFP growth for construction industry of Singapore,
3. to apply the ANN technique to develop an forecasting model by using the selected significant factors.
4. to apply regression method to develop another forecasting model to serve as an benchmark to judge the performance of the ANN model.

Data and Methodology

The study period is from the 2nd quarter 1984 to the 4th quarter 1988, totalling 59 quarterly records.

To achieve the first goal, Jorgenson's translog production function based TFP measurement is used to compute the real quarterly TFP growth for construction industry in Singapore during the 2nd quarter 1984 to the 4th quarter 1988. Jorgenson's approach is presented bellow:

It assumes that for each industry, there exists a transcendental logarithmic (translog) production function, giving output as a function of intermediate input, capital input, labour input, and time.

$$Z = F(X, K, L, T) \hspace{4cm} \text{equation [1]}$$

Where T is time, Z is quantity of output, X, K, L are quantity of the intermediate, capital and labour inputs.

Under the condition of constant returns to scale and producer equilibrium, the average growth rate of TFP can be expressed as the growth rate of output less the sum of weighted average of growth rate of intermediate, capital and labour inputs:

$$\bar{V_T} = [\ln Z(T) - \ln Z(T-1)] - \bar{V_X}[\ln X(T) - \ln X(T-1)] -$$
$$\bar{V_K}[\ln K(T) - \ln K(T-1)] - \bar{V_L}[\ln L(T) - \ln L(T-1)] \quad \text{equation [2]}$$
where

$$\bar{V_X} = \frac{1}{2}[V_X(T) + V_X(T-1)]$$

$$\bar{V_K} = \frac{1}{2}[V_K(T) + V_K(T-1)]$$

$$\bar{V_L} = \frac{1}{2}[V_L(T) + V_L(T-1)]$$

$$\bar{V_T} = \frac{1}{2}[V_T(T) + V_T(T-1)]$$

$$V_X = \frac{p_X X}{qZ} \qquad V_K = \frac{p_K K}{qZ} \qquad V_L = \frac{p_L L}{qZ}$$

In the above equations, $\bar{V_X}, \bar{V_K}, \bar{V_L}$ represents the respective shares of intermediate, capital and labour input averaged over time T and T-1, while q, p_x, p_k, p_L denote the prices of the output and intermediate, capital and labour inputs.

Each of the input can be expressed as a translog function of its individual inputs. Under constant return to scales and producer equilibrium, the growth rate of each input can be expressed as a weighted average of growth rates of its individual input, with weights given by average value shares:

$$\ln X(T) - \ln X(T-1) = \sum \bar{V_{Xi}}[\ln X_i(T) - \ln X_i(T-1)] \hspace{2cm} \text{equation [3]}$$

$$\ln K(T) - \ln K(T-1) = \sum \bar{V_{Kj}}[\ln K_j(T) - \ln K_j(T-1)] \hspace{2cm} \text{equation [4]}$$

$$\ln L(T) - \ln L(T-1) = \sum \bar{V_{Li}}[\ln L_i(T) - \ln L_i(T-1)] \hspace{2cm} \text{equation [5]}$$

To compute the TFP growth using the above equations, the data need to be compiled are translog index of real output growth, translog indices of real intermediate, capital and labour growth and value share of intermediate, capital and labour input.

Translog index of output growth

In Singapore, there is no direct record on construction industry-level output. The alternative way is to use coefficients of value-added and output published on *Singapore I-O Table*. The value of gross output can be

acquired by dividing the value-added by the coefficients. As *the Singapore I-O table* is published only every 5 years and it assumes that the economic structure during the benchmark years remain stable, the coefficients between benchmark years can be developed by interpolation. Meanwhile the output should be valued in the producer's prices. Adjustment to nominal gross output is needed to exclude all indirect taxes and intra-industry purchase. To get the real output growth, the nominal output are deflated by using tender price index as a deflator.

Translog index of intermediate input growth

From equation 3, translog index of intermediate input growth is aggregate growth of its individual components. The primary sources for the individual components are Table 2 COMMODITY ANALYSIS OF PURCHASES FROM DOMESTIC PRODUCTION, and Table 3 COMMODITY ANALYSIS OF RETAINED IMPORTS, on *Singapore I-O Table*. The Singapore I-O table allows two types of intermediate input to be distinguished: (1) materials, purchased; (2) non-industrial services. To get the real value of each component, for materials, building materials price index are used as deflators; for non-industrial services, GDP deflators of each industry are used as deflators.

Translog index of capital growth

Theoretically, quantity of capital input should be measured in the terms of capital services provided. In reality, such data is seldom recorded in Singapore as well as in other countries. The alternative method assumes that the services provided by a durable good are proportional to initial investment in this good. Our method mainly follows Willie Tan's (2000) approach, which use value of net fixed assets, published on the *Corporate Sector Performance and structure, 1992 and 1997*, as an approximate to capital stock. For simplicity purpose, two types of fixed asset are distinguished: machinery and equipment; land and building. As data on net fixed assets are available in aggregate level, a survey was conducted to estimate the proportion of each category in the composition of fixed asset. Results from the 20 responding construction firms indicates that land and building constitute about 24.5% (mean value) of all net fixed assets. Thus, 24.5% of net fixed assets is attributed to land and buildings and Property Price Index (all types of properties) is used to deflate the real property proportion of net fixed assets; and the rest 73.5% is attributed to machinery and equipment, and Domestic Supply Price Index for Machinery and Transport Equipment is used as the deflator.

Translog index of labour growth

Labour input is measured by quarterly hours worked, which is computed as the product of number of workers employed, weekly hours worked, and weeks worked per quarter. Workers are cross-classified by age, sex, education. Totally 72 categories of workers are identified (6 age group, 2 sex group and 6 education levels).

To measure quarterly hours worked, the first step is to obtain employment matrices of construction workforce, cross-classified by age, sex, education. The main data source is *profile of the labour force of Singapore 1983-1994*. Second, weekly hours worked by each category of workers are from *TREND* data on *average weekly hours worked by industry annually*. Third, quarterly hours worked for each group of workers are assumed to be 12 weeks for all groups of workers (divide the 48 weeks for annual hours worked by 4).

Value share of intermediate, capital and labour input

Under the condition for producer's equilibrium, the sum of the value intermediate, capital and labour is equal to value of output. Value share of intermediate, capital and labour input is derived in benchmark year are from *Singapore Input-Output Table*, and between benchmark years using interpolation.

To achieve the second objective, that is to identify and select factors that are significantly influence the TFP growth for construction industry of Singapore, three stages are involved. First, a comprehensive

literature search of potential factors that theoretically identified as influencing TFP growth in the construction industry are conducted. Second, a list of statistical indicators or proxies is abstracted, from statistics resources in Singapore to create the time-series dataset. Third, STEPWISE procedure, available on linear regression in SPSS, is applied to select the statistically significant indicators. The significance level is set on 10%.

To develop a forecasting model using ANN, three stages involved. First is to design the architecture of the neural network. Second is to train and develop the model by using the 55 dataset, starting from 2nd Q 1984 to 4th Q 1997. Third is to test and evaluate the forecasting capability of the developed model by using the 4 dataset on the 4 quarters of 1988.

To achieve the last objective, first a MR model is developed by using the same dataset. Second, the predication errors for both models are computed to compare their forecasting capability in terms of accuracy.

TFP growth rate of construction industry in Singapore

Using Jorgenson's TFP measure, the real quarterly TFP growth trend for construction industry in Singapore during the 2nd quarter 1984 to the 4th quarter 1988 is depicted in Fig.1.

Figure 1. *TFP growth for construction industry of Singapore during the 1984 2Q - 1998 4Q*
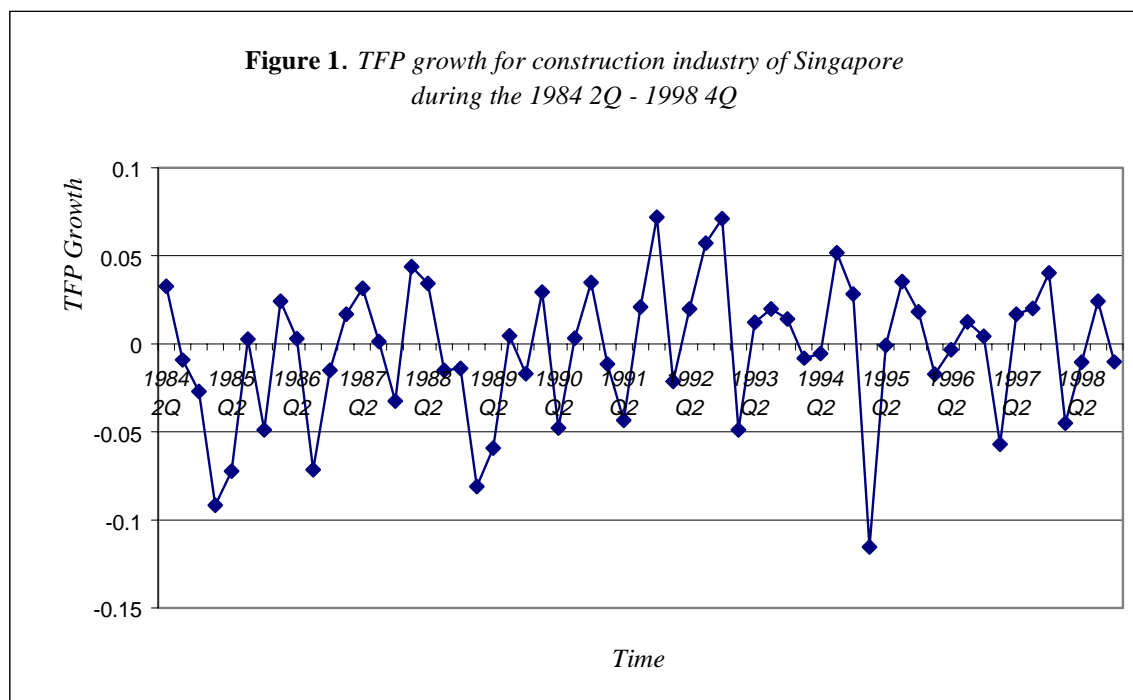
Selection of significant factors

A comprehensive literature search of theoretical factors influencing on TFP growth in construction industry has been conducted in the case of the construction industry of Singapore. Theoretical factors then have been translated into statistical indicators in order to gather the data. Table 1 lists the statistical indicators for each theoretical factor and the data sources for each indicator.

Table 1 Theoretical factors, statistical indicators and data sources

Theoretical factors	Statistical indicators	Data sources
Composition of output	Index of composition of output	Labour productivity by type of work from *CIDB annual report*; Value share of each type of work from progress payment certified by construction industry from TREND
	Ration of new construction to repair and maintenance works	Dataset recorded is not long enough
Technology progress	R&D in construction per million contract awarded	*CIDB annual report*; *National survey of R & D expenditure & manpower*; *National Survey Of R & D In Singapore.*
	Domestic R&D per Million GDP	*National survey of R & D expenditure & manpower*; *National Survey Of R & D In Singapore.*
	The average age of real fixed capital stock employed in construction industry	N/A
	Rate of International Technology transfer	N/A
Labour quality	Index of labour quality, by age-sex-education	See Translog index of labour growth above
	foreign worker's proportion in total labour force	*CIDB construction manpower report 1987*; *construction industry manpower survey* by CIDB
	Certified worker's proportion in total labour force	*Construction industry manpower survey* by CIDB; *CIDB annual report*; *CIDB economics report*
Materials quality	Pre-cast level	Precast level *from CIDB annual report*; *TREND*: total consumption of cement for precast concrete; *CIDB economics report*: TOTAL CEMENT DEMAND; *TREND*:
	Buildable score	Dataset recorded is not long enough
Economies of scale	Output per capita	Data on output see translog index of output growth above; capita is number of labour employed from construction industry. Data is from *TREND: employed person aged 15 and over by construction industry, annual* .
Government regulations	Investment allowance granted	CIDB annual report
	Efficiency indicators of the CIDB- Total granted to CIDB(%)/contract awarded	CIDB annual report
Cyclical factors	Energy price (oil)	*TREND*: Domestic Price Index-petroleum product refined
	Inflation rate	Inflation rate is computed as growth rate of consumer price index. *TREND*: Consumer price index
Industrial relation polices	Percentage of contract that establish of labour management committees	N/A
	Percentage of contract that allow for incentive wage payment	N/A

	Percentage of contract that allow for subcontracting	N/A
	Percentage of union	Memberships of employee's trade Unions by sector at year end, from *Singapore yearbook of labour statistics* and *Singapore yearbook of manpower statistics*
Irregular factors	Number of industrial accidents	Industrial accidents by number and industry, from *Singapore yearbook of labour statistics* and *Singapore yearbook of manpower statistics*
Other factors	Such as social and institutional change	N/A

A total of 14 indicators have been quantified as theoretical factors influencing TFP growth in of construction industry in Singapore. The STEPWISE procedure selected 7 indicators that are significantly related to TFP growth for construction industry of Singapore. Those met the significance 10 % are:

1. Output per capita (OPC)
2. Foreign worker's proportion (FWP)
3. Precast level (PL)
4. R&D per million contract awarded (RDPCA)
5. Percentage union (PU)
6. Accident numbers (AN)
7. Investment allowance granted (IAG)

Theory of ANN

In essence, ANN is an information technology that mimics the human brain and nervous system to learn from experience or generalize from previous examples to generate new ones by abstracting essential characteristics from inputs in the pattern of variable interconnection weights among the processing elements. The most import advantage of ANNs over conventional mathematical and statistical models is their adaptivity**.**

Among the various ANN architectures and paradigms, the back-propagation (BP) neural network is one of the simplest and the most practicable networks. Backpropagation is a learning algorithm that relies on backward-error propagation. BP works extremely well in diagnosis, classification, decision-making, planning and scheduling. It has the ability to acquire arbitrarily complex nonlinear mappings with great reliability. Properly trained backpropagation networks tend to give reasonable answers when presented with inputs they have never seen.

Develop the ANN model

There are three steps in the process to develop the ANN model:
1. design the architecture for the ANN network
2. train the network
3. test the model

Architecture of the ANN network

A typical three-layered backpropagation feedforward neural network is chosen for the ANN model. It consists an input layer with 7 nodes, a hidden layer with 5 nodes and an output layer with one node. Fig. 2

displays the network structure created for the ANN model. The size of 5 nodes in the hidden layer is found to be optimum for the ANN model in terms of accuracy.
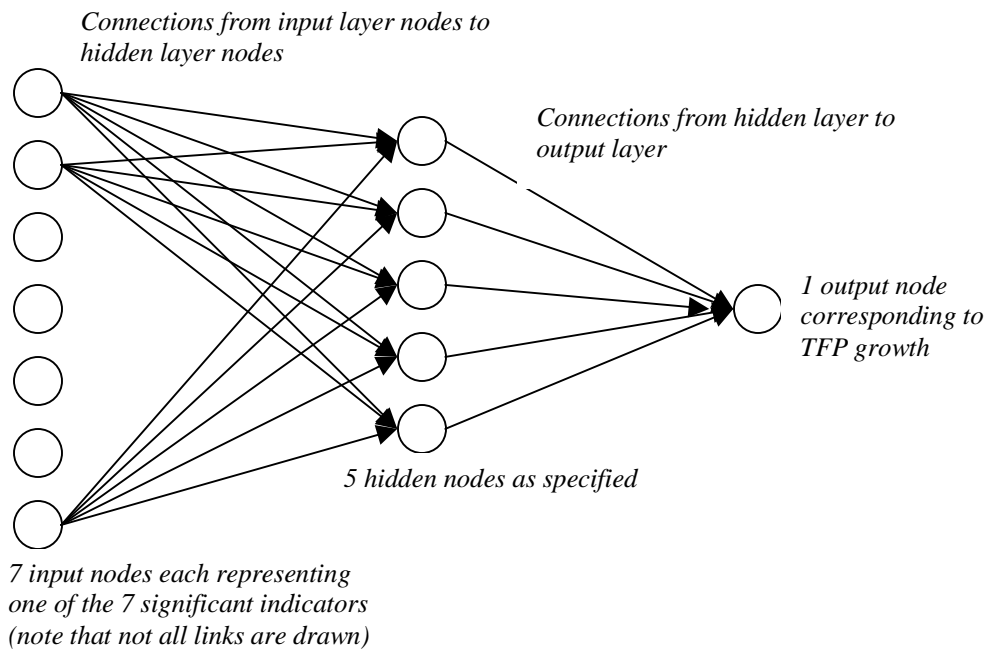

Figure 2. three-layered back-propagation feedfoward neural network adopted for the ANN model

There are many variations of the backpropagation training algorithms, available on MATLAB package. In this case, as the total number of points in the training set is not much bigger than the number of the parameters in the network, there is a high chance of overfitting. The network tends to memorize the training example without the ability to generalize the new situations. To deal with this problem, the Bayesian Regulation has been used as the training function. The entire 55 training set with each set of 7 inputs and one output (TFP growth) was applied to the network one at a time to conduct batch learning. The training stops with the message " Maximum MU reached". The network is regarded as fully converged.

Testing the ANN model

After the network is trained, the network performance is evaluated on the test dataset. The test set is the 4 quarters of 1988. The forecast results of the ANN model for the 4 quarters of 1988 are shown in Table 2.

Table 2 Results of forecast by the ANN model

Yr/Qtr	Actual	Forecast	Error
1998 1Q	-0.04497	-0.0455	0.00053
1998 2Q	-0.01045	-0.0087	-0.00175
1998 3Q	0.02417	0.0292	-0.00503
1998 4Q	-0.01027	-0.0104	0.00013

Forecasting results of the MR model

The 55 training dataset used in the ANN model was applied to develop the MR model The 7 significant factors identified during the ealier stage of the study are adopted as independent variables in the regression analysis to estimate the dependent variables, TFP growth. The MR model is developed using the REG procedure in the SPSS package with the forced entry option selected. The results of the REG procedure are presented in Table 3. The analysis indicates that all the variable entered were statistically significant at the 10% level with the R-square value of 0.907.

Table 3 Estimated TFP growth function for construction industry of Singapore

Independent variables	Coefficients	t-value	Sig.t
(Constant)	-7.986E-03	-4.276	.000
OPC	.424	17.853	.000
FWP	-.111	-2.974	.005
PL	-7.635E-02	-2.477	.017
RDPCA	1.079E-02	1.732	.090
PU	1.535E-02	2.075	.043
AN	-3.795E-02	-2.062	.045
IAR	9.914E-03	1.964	.056

R-square=0.907; adjusted R-square=0.893
Final Parameter: Number of residuals=54; Standard error=1.3261850E-02.
Analysis of Variance:

DF
Sum of squares

Residuals
47
8.266E-03

By summarizing the results from Table 3, the MR model for TFP growth for construction industry of Singapore is represented as:

$$TFP\ growth = -0.007986 + 0.424(\ln(OPC)\ growth) - 0.111(\ln(FWP)\ growth) - 0.07635(\ln(PL)\ growth)$$
$$+ 0.01079(\ln(RDPCA)\ growth) + 0.01535(\ln(PU)\ growth) - 0.03795(\ln(AN)\ growth)$$
$$+ 0.009914(\ln(IAR\ growth))$$

The forecasts generated by the MR model are shown in Table 4.

Table 4. Results of the forecast obtained by the MR model

Yr/Qtr	Actual	Forecast	Error
1998 1Q	-0.04497	-4.51E-02	9.5338E-05
1998 2Q	-0.01045	-7.58E-03	-0.00286548
1998 3Q	0.02417	2.68E-02	-0.00262879
1998 4Q	-0.01027	-9.17E-03	-0.00109651

Comparative study of the ANN and MR models

The forecasting results of the ANN and the MR model are compared in terms of relative measures of the forecasting accuracy. The measures used in this study are:

1. Percentage error (PE)

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

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)

$$MAPE = \sum | PE_i | / n$$

The results of the comparative study are shown in Table 5. As shown in Table 5, the MPE and MAPE of the ANN model were found to be 1.63% and 10.00% respectively. While the MPE of and MAPE the MR model were 6.75% and 12.30%. As the limit error for an accurate forecast model is less than 10%, results of the MAPE values indicates that the ANN model is able to produce more accurate forecasts for TFP growth of construction industry of Singapore than the MR model.

Table 5 Results of the comparative study

Measures of accuracy	Forecast by ANN model %	Forecast by MR model %
PE for 1998 1Q	-1.18	-0.21
PE for 1998 2Q	16.75	27.42
PE for 1998 3Q	-20.81	-10.88
PE for 1998 4Q	-1.27	10.68
MPE	-1.63	6.75
MAPE	10.00	12.30

Summary of findings

Firstly, the actual TFP growth in construction industry of Singapore has been computed using Jorgenson's method and the growth trend was presented in Figure 2.

Secondly, among the 14 statistical indicators identified, 7 has been found to be significantly related to TFP growth in the construction industry of Singapore. They are: output per capita; foreign worker's proportion; pre-cast level; R&D per million contract awarded; percentage union; accident numbers; investment allowance granted. Those which did not meet the 10% significance level were: certified worker's proportion; development granted to CIDB to the size of construction industry; energy price; inflation; R&D per GDP, output composition and labour quality index. Among the 7 selected indicators, the one with controversial signs of coefficient is pre-cast level. According to the literature review, pre-cast level is expected to have a positive relationship with TFP growth but the reverse was obtained. This departure from the theory may be justifiable as currently, the prefabrication although is labour saving but its cost is more expensive than casting on site. Another possibility is that when computing pre-cast level, the quality data may not be reliable.

Thirdly, MPE and MAPE derived from the ANN and MR models indicate that the ANN model is capable of producing more accurate forecasts (less than 10%) than the MR model.

Conclusion

Macro-level information processing models are useful in monitoring and predicting the performance of the construction industry as a whole. With the need to manage construction performance information at all

three levels, namely, industry, firm and site, this study looks specifically at developing an 'intelligent' model for forecasting industry-level productivity. The major contributions of this study are highlighted bellow:

- Using total factor productivity concept and Jorgenson's translog production function to measure industry-level productivity performance;
- Applying an AI information-processing tool, the artificial neural network (ANN) system to predict TFP growth in the construction field

The outstanding forecast performance of the ANN model over MR model is mainly attributed to the characteristics of the neural network model, which are designed to capture the non-linear relationship between the input and output variables and are highly fault tolerant. This study continues to prove that the ANNs can be used to predict accurate outcomes, particularly in complex non-linear case and its application can be extended to other construction management areas where exists complex non-linear relationship between input and output variables. Finally, the methodological framework set up in this study can in general be applicable to other counties as well.

References

Chau K.W.and Walker A. (1988), "the measurement of total factor prodcutvity of the HONG Kong Construction Industry", *Construction Management and Economics*, E. & F.N Spon, 6, 209-224.

Chau K.W. (1993), "Estimating industry-level productivity trends in the building industry from building cost and price data", *Construction Management and Economics*, E. & F.N Spon, 11, 370-383.

Construction Industry Development Board (CIDB) (1996), Census of the Construction Industry 1996. CIDB, Singapore.

Dacy, D.C. (1965), " Productivity Price Trends in Construction Since 1947", *Review of Economics and Statistics*, 47
Department of Statistics Singapore (1992), Singapore's Corporate Sector: Size, Composition and Financial Structure, DOS, Singapore

Department of Statistics Singapore (1998), Singapore's Corporate Sector 1986-1995, DOS, Singapore.

Department of Statistics Singapore (2000), Singapore's Corporate Sector 1996-1997, DOS, Singapore.

Department of Statistics Singapore (1978) (1983) (1985) (1988) (1990) (1995), Singapore Input-Output Tables, DOS, Singapore

Department of Statistics Singapore (1997), Occasional paper series: Multi-factor productivity growth in Singapore: concepts methodology and trends, Singapore: Department of Statistics.

Jorgeson et al..(1987), *productivity and US economic growth*, Cambridge, MA, Harvard University Press.

Moselhi, O., Hegazy, T., and Fazio, P.(1991), "Neural network as tools in construction", *Journal of Construction Engineering and Management*, 117(4), 606-623.

Tan Willie (2000), "Total factor productivity in Singapore construction", *Engineering, Construction and Architecture Management*, 2000, 7(2) 154-158.