A variable asymptote logistic (VAL) model to forecast electricity consumption

Zaid Mohamed and Pat S. Bodger*

Department of Electrical and Computer Engineering, University of Canterbury, PO Box 4800, Christchurch, New Zealand

Fax: 64 33642761 E-mail: p.bodger@elec.canterbury.ac.nz

E-mail: zmo10@student.canterbury.ac.nz

*Corresponding author

Abstract: The logistic model has been very effective in forecasting many technological forecasting patterns. However, it has the characteristic of underestimating the forecasts in many situations. This is mainly due to the constraints imposed by the constant saturation level of the logistic growth curve. This paper proposes a variable asymptote logistic (VAL) model for forecasting electricity consumption in New Zealand. The saturation level of electricity consumption is found by the respective degree of correlation with the population of the country and the price of electricity. This is used as a variable asymptote in this logistic model and hence the name variable asymptote logistic (VAL) model. The VAL model gave lower forecasting errors than the logistic model and gave future forecasts that are higher than the logistic model. The software package MATLAB is used in all stages of the simulation.

Keywords: forecasting electricity; forecasting accuracy; logistic model.

Reference to this paper should be made as follows: Mohamed, Z. and Bodger, P.S. (2005) 'A variable asymptote logistic (VAL) model to forecast electricity consumption', *Int. J. Computer Applications in Technology*, Vol. 22, Nos. 2/3, pp.65–72.

Biographical notes: Zaid Mohamed received the BE (Hons) degree from the University of Adelaide, Australia in 2000. From 2000 to 2002, he worked as an engineer at the State Electric Company, Maldives. Currently he is a PhD student at the Department of Electrical and Computer Engineering, University of Canterbury, New Zealand. His research interests include electricity forecasting, planning and management of power systems.

Pat S. Bodger received the BE (Hons) in 1972 and PhD degree in 1977 from the University of Canterbury, Christchurch, New Zealand. From 1977 to 1981 he was with the New Zealand Electricity working on the areas of design and construction, power station maintenance and commercial engineering. He joined the Department of Electrical Engineering at the University of Canterbury in 1982. Currently, he is Professor of Electric Power Engineering, Director of the Electric Power Engineering Centre, and the head of the Department of Electrical and Computer Engineering, University of Canterbury, New Zealand. His research interests include power system simulation, energy forecasting, management and modelling, power applications engineering, high voltage and electro biotechnology.

1 INTRODUCTION

The logistic model has been frequently used in technological forecasting either in its simplest form or in a modified form (Carrillo and Gonzalez, 2002; Bass, 1969; Jain and Rao, 1994; Bhargava et al., 1991; Frank, 2004;

Bewley and Fiebig, 1988; Meyer and Ausubel, 1999; Giovanis and Skiadas, 1999; Mar-Molineo, 1980; Oliver, 1981; Harvey, 1984). The logistic model has also been successfully applied to electricity forecasting (Mohamed and Bodger, 2003; Bodger and Tay, 1987; Skiadas et al., 1993; Sharp and Price, 1990; Young, 1993;

Tingyan, 1990). The logistic model (Mohamed and Bodger, 2003; Bodger and Tay, 1987) proposed for New Zealand electricity consumption was very successful in explaining the past electricity consumption in New Zealand. However, it was found that it underestimated actual consumption values that have occurred since the forecasts were made and were generally lower than other available national forecasts. The main reason behind the low predictions by the logistic model is due to the constraints imposed by the saturation level that was obtained by a Fibonacci search technique (Mohamed and Bodger, 2003; Bodger and Tay, 1987).

Logistic models are based on S-shape growth curves that ultimately reach a saturation level. Such growth curve models may be based on the normal distribution (Stapleton, 1978) and the Weibull function (Sharif and Islam, 1980). In such models, the estimated saturation level is used as one of the fitted parameters of the growth functions. Leach (1981) demonstrated that the prediction power of the logistic growth function can be improved by taking into account the systematic changes in the saturation level. The idea of varying saturation level can also be found in (Maharajan and Peterson, 1978; Maharajan et al., 1979; Sharif and Ramanathan, 1981; Sharif and Ramanathan, 1984). It was shown that the forecasting ability of the logistic function for the Greek electric system is improved using the correlations of some socioeconomic variables to describe the saturation level (Skiadas et al., 1993).

In this paper, a variable asymptote logistic (VAL) model is proposed that improves the forecasts made by the logistic model for New Zealand electricity consumption. The saturation levels of the logistic curve are estimated using the Fibonacci search technique (Mohamed and Bodger, 2003; Bodger and Tay, 1987). The saturation levels are correlated with population, price of electricity and gross domestic product (GDP). The VAL model for forecasting electricity consumption in New Zealand is proposed using the selected variables based on the analysis. Multiple linear regression is used in studying the correlation between the explaining variables and electricity consumption. Since ARIMA techniques are well known for predicting economic and demographic variables (Canarella and Pollard, 1989; Ilkka Virtanen and Yli-Olli, 1987; Pflaumer, 1992; Binroth et al., 1979), they are used in predicting future values of population, price of electricity and GDP.

2 VARIABLE ASYMPTOTE LOGISTIC (VAL) APPROACH

2.1 Proposed model

The logistic model is described by (Mohamed and Bodger, 2003; Bodger and Tay, 1987):

$$f = \frac{F}{1 + \exp(a_0 + a_1 t)} \tag{1}$$

where, F is the saturation level, f is the electricity consumption, a_0 and a_1 are constants and t is the time in years.

In the VAL model, the saturation level F is estimated using the explaining variables $(X_1, ..., X_n)$ and used as a variable asymptote F(X) in equation (1). Thus, the proposed VAL model takes the form:

$$f = \frac{F(X_i)}{1 + \exp(a_0 + a_1 t)}$$

and,

$$F(X_i) = c_0 + \sum_{i=1}^{n} (c_i \times X_i)$$
 (2)

where, $F(X_i)$ is the saturation level expressed as a function of n variables, and c_0 and c_i are the parameters obtained from the explaining variables.

A modified logistic function based on the sigmoid logistic function has been proposed initially for the Greek electric system (Skiadas et al., 1993). However, the sigmoid logistic function is different from the logistic model in terms of its description and the method used to obtain the saturation levels.

The main stages involved in the VAL model are:

- 1 estimate the saturation level *F* using the Fibonacci search technique
- 2 correlate *F* with the explaining variables considered and reestimate *F* using these variables
- 3 if the reestimated F does not adequately explain the saturation level, reconsider the explaining variables and repeat step 2
- 4 obtain forecasts of the explaining variables using ARIMA technique
- 5 estimate future saturation levels *F* using the forecasts of the explaining variables
- 6 use the variable F values in equation (2) to forecast electricity consumption.

2.2 Fibonacci search technique

Initial values of the saturation level are obtained by the Fibonacci search technique (Mohamed and Bodger, 2003; Bodger and Tay, 1987). The Fibonacci search technique is developed using the software MATLAB. The program calculates the saturation level of a given data set with the flexibility of adjusting which data points are to be used in the calculation. The Fibonacci search technique is applied to domestic, non-domestic and total electricity consumption data series 1943-1980, 1943-1981, ..., 1943–1999 of New Zealand and the saturation level F is estimated for each case. The corresponding saturation levels are shown in Figure 1.

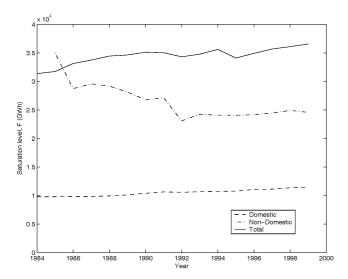


Figure 1 Varying saturation levels of electricity consumption in New Zealand

The domestic saturation level is relatively constant with a slight increase over the data period. The non-domestic saturation level is consistently decreasing until the early 1990s. The decrease of the saturation level in the initial years suggests possible immaturity in the non-domestic data. However, the near constant value beyond 1992 suggests possible maturity in the latter years. The total consumption saturation level is relatively constant with increasing value after 1984. The restrictions on electricity brought by the draught of 1992 are reflected in the variation of the saturation level of the non-domestic consumption for that year, with an overall low saturation beyond 1994. This is not very significant in the domestic or the total consumption. In all cases, deregulation of electricity in New Zealand in 1987 appears to have had little or no influence in the patterns.

2.3 Correlation and reestimation of F

In order to use population, price of electricity and GDP in estimating the saturation levels, it is expected that the chosen variables effectively explain the respective saturation levels. Table 1 shows the correlation between the saturation levels and the explaining variables considered.

Table 1 Correlation between the saturation levels and the explaining variables

	Explaining variables						
Saturation	Population	Price	GDP				
Domestic	0.95	0.81	0.88				
Non-domestic	- 0.49	- 0.61	- 0.51				
Total	0.68	0.67	0.61				

The correlations are high in the domestic sector. The correlation in the total consumption are also acceptable

while those for the non-domestic sector are negative and much lower in magnitude overall. The negative correlation in the non-domestic sector is the result of an overall saturation levels in the non-domestic decrease in consumption as opposed to the domestic and consumptions. This indicates immaturity non-domestic sector and that the growth the non-domestic sector has not gone beyond the early stages of development for the Fibonacci search technique to be applied effectively. In addition to these, the low correlations in the non-domestic sector have excluded the VAL model from being applied to the non-domestic sector.

Statistical tests such as the Durbin-Watson statistic, coefficient of determination and *F*-test are used in evaluating the applicability of the estimated saturation levels using the explaining variables (Makridakis et al., 1998). Initially, the saturation levels are estimated using combinations of all these three variables. The process of reestimation is continued with not only the various combinations of variables, but also with the various saturation level datasets. It was found that all the combinations of the variables are generally acceptable for the saturation datasets of the range from 1980–1999 to 1985–1999. The best results are found for the dataset from 1984–1999. Table 2 summarises the statistical test results and the various combinations of explaining variables used for the saturation dataset 1984–1999.

Table 2 Statistical test results of the various models for the data set from 1984–1999

	Domestic		Total		99%		
Varaibles in the model	DW	r²	F	DW	r²	F	critical F
Population, price and GDP	1.0	0.97	117.1	1.7	0.85	21.8	5.0
Population and price	0.9	0.96	147.4	1.7	0.84	35.0	5.6
Population and GDP	1.0	0.97	187.5	0.7	0.60	9.6	5.6
Price and GDP	8.0	0.88	45.7	1.6	0.84	34.6	5.6

The combination of population and GDP in model 3 for the domestic sector is the best. However, the corresponding results for the total consumption are not satisfactory. The DW, r^2 and F-statistics values are much lower than all the other combinations of explaining variables chosen. Model 4 performs the worst for both the domestic and total consumption. Models 1 and 2 are the most comparable in terms of the statistical tests. In model 1, the addition of a third explaining variable has not improved the DW and r^2 statistical results significantly while the F-statistics have decreased. Overall it is model 2 that performs most consistently. In all the models, the DW values of the domestic sector are much closer to 1 (rather than being closer to 2) indicating some serial correlation of the residuals. The corresponding saturation value estimates for each of the models are:

Model 1:

$$F_{D1}(X) = -6.08 \times 10^{2} + 3.6 \times 10^{-3} X_{1}$$

$$+18.5 X_{2} - 2.33 \times 10^{2} X_{3}$$

$$F_{T1}(X) = 2.95 \times 10^{4} - 1.16 \times 10^{-3} X_{1}$$

$$+1.05 \times 10^{3} X_{2} + 2.54 \times 10^{-2} X_{3}$$
(3)

Model 2:

$$F_{D2}(X) = 9.46 \times 10^{2} + 2.6 \times 10^{-3} X_{1} + 42.1 X_{2}$$

$$F_{T2}(X) = 2.79 \times 10^{4} - 5.28 \times 10^{-4} X_{1} + 1.03 \times 10^{3} X_{2}$$
(4)

Model 3:

$$F_{D3}(X) = -931 + 3.8 \times 10^{-3} X_1 - 2.49 \times 10^{-2} X_3$$

$$F_{T3}(X) = 1.11 \times 10^4 + 8.0 \times 10^{-3} X_1 - 6.42 \times 10^{-2} X_3$$
(5)

Model 4:

$$F_{D4}(X) = 5.74 \times 10^{3} + 182X_{2} + 4.16 \times 10^{-2} X_{3}$$

$$F_{T4}(X) = 2.67 \times 10^{4} + 977X_{2} - 3.8 \times 10^{-3} X_{3}$$
(6)

where, X_1 is population, X_2 is price of electricity, X_3 is GDP, and F(X) denotes the saturation function with respect to the subscript D for domestic sector and T for the total consumption.

In order to propose the VAL model using the logistic model, accurate forecasts of the explaining variables are needed. ARIMA techniques are often used in these types of forecasting (Ilkka Virtanen and Yli-Olli, 1987; Pflaumer, 1992; Binroth et al., 1979; Makridakis et al., 1998). Therefore, ARIMA methods are used in forecasting population, price of electricity, and GDP. The forecasts of these variables are employed in forecasting the saturation levels for electricity.

2.4 VAL method and forecasting

In the logistic model, the asymptote F calculated for a particular data set is a constant for the forecast to be made. If a forecast is to be made from the year 2000 onwards, the Fibonacci search technique is used to obtain a constant F, which is then used in equation (1), as the time t is increased to obtain the forecasts. In the VAL model, the asymptote F for each of the forecast years is different from one another. The forecasts of the explaining variables are used to calculate the saturation values of electricity for the future years using equations (3)–(6).

2.5 Choice of model in terms of forecasting accuracy

Forecasting accuracy is measured over the period from 1991 to 1999. In making the forecast for 1999, all the data up to 1998 is used. The forecasting accuracy is measured by

calculating the mean absolute percentage error (MAPE) of the first year ahead forecast. Similarly, the MAPE values of two years ahead through to nine years ahead forecasts for the domestic sector are calculated and are shown in Figure 2.

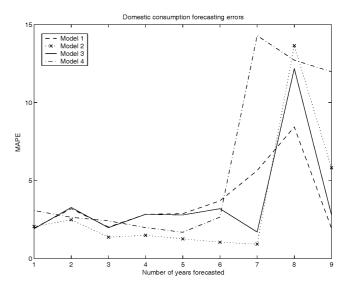


Figure 2 Domestic sector forecasting accuracy for the four models

The best forecast is for model 2 giving the lowest MAPE values from first year through to seven years ahead forecasts. Initially, models 1 and 3 behaved very similarly but model 1 performed better for the longer forecasting periods. Model 4 performed better than models 1 and 3 initially, but gave much higher errors from six years ahead MAPE onwards.

The MAPE values of the first year through to nine years ahead forecasts of the total electricity consumption are shown in Figure 3.

The forecasting errors are closer together especially for the short range forecasts. However, model 2 again gave the lowest MAPE values from one through to six years ahead errors. It can also be seen that model 1 gave the highest MAPE values in general, indicating the worst forecasts.

Beyond year seven, the forecasts of all the models are generally weak giving rise to high MAPE values, especially at year eight. There are two possible reasons for this. Firstly, in the VAL method, the asymptotes are estimated for the years from 1984. Thus, in the eight years ahead forecast the saturation levels are initially calculated from 1984 to 1991. This means that as the forecasted period increases, the number of saturation values calculated decreases. The decrease in the number of data points generally increases the error in the estimate of the coefficients in the regression analysis (Farnum and Stanton, 1989). This will lead to an increase in forecasting errors. Secondly, for both the domestic and total consumption data, the decrease in the year 1992 due to the electricity restrictions brought by drought in that year resulted in an overall increase in the forecasting error.

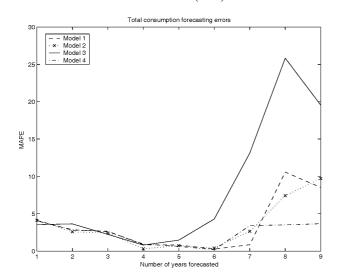


Figure 3 Total consumption forecasting accuracy of the four models

In terms of forecasting accuracy, model 2 is the best model for forecasting electricity consumption. This model uses the explaining variables population and price of electricity only.

3 VAL MODEL FORECASTS

Analyses in Sections 2.3 and 2.5 show that the best VAL model uses population and price of electricity as the explaining variables. The estimated domestic and total consumption saturations using the VAL model are shown in Figures 4 and 5, respectively. There is a good correlation between the actual saturation values obtained by the Fibonacci search technique and those estimated by the VAL model.

The forecasting equation for the VAL model for the domestic sector is:

$$f_D = \frac{F_{D2}(X)}{1 + \exp(-200 + 0.104t)}$$

$$= \frac{9.46 \times 10^2 + 2.6 \times 10^{-3} X_1 + 42.1 X_2}{1 + \exp(-200 + 0.104t)} \,. \tag{7}$$

While that for total consumption is:

$$f_T = \frac{F_{T2}(X)}{1 + \exp(-185 + 0.094t)}$$

$$= \frac{2.79 \times 10^4 - 5.28 \times 10^{-4} X_1 + 1.03 \times 10^3 X_2}{1 + \exp(-185 + 0.094t)}.$$
 (8)

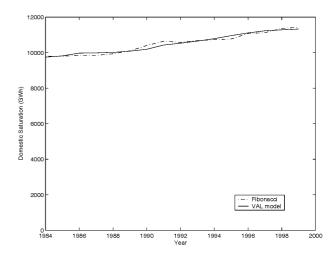


Figure 4 Forecast of the domestic consumption saturation level using price and population

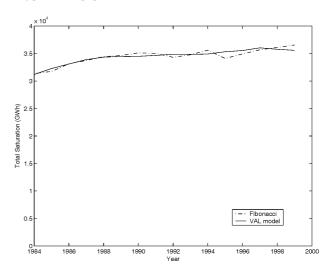


Figure 5 Forecast of the total consumption saturation using price and population

4 IMPROVEMENTS OVER THE LOGISTIC MODEL

The initiative to developing the VAL model is to increase forecasting accuracy. This should be reflected in a reduction in forecasting errors. If the VAL model does not reduce the forecasting error, then the choice of this model over the logistic model could not be supported.

The quality of forecasting of these models can be obtained visually by observing how well the models have forecasted actual data. Figures 6 and 7 show the forecasts of the VAL model and the logistic model from 1995 to 1999. In making these forecasts, it is assumed that no data beyond 1994 is known. The forecasts by the VAL model are much better than the forecasts by the logistic model for the five-year ahead forecasts.

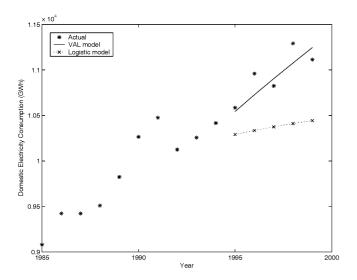


Figure 6 Comparison of domestic forecasts for VAL and logistic models

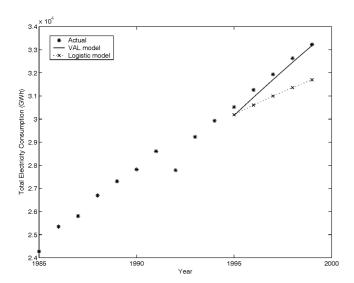


Figure 7 Comparison of total forecasts for the VAL and logistic models

The MAPE values for the domestic consumption are 1.3 for the VAL model while there is a much higher value of 5.3 for the logistic model. Similarly, the MAPE value for the total electricity consumption is 0.7 for the VAL model while that for the logistic is 2.9. These values clearly suggest that the VAL model has performed much better.

The forecasting errors, using MAPE, are calculated not only for the fifth year ahead forecasts but from first year ahead to nine years ahead forecasts for the VAL and logistic models. The resulting MAPE values are shown in Figure 8.

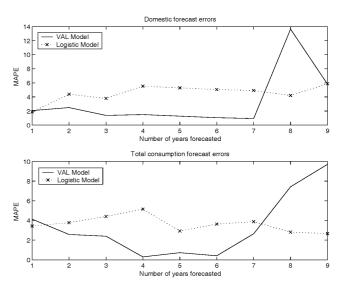


Figure 8 One to nine years ahead MAPE values for the domestic and total electricity consumption

The MAPE values are much lower for the VAL model from first year ahead through to seven years ahead forecasts for both the domestic and the total electricity consumptions. At year eight, the VAL model has higher MAPE values than the logistic model, again reflecting the drought year of 1992 which could not have been predicted by the change in population and price of electricity in the VAL model as compared to the logistic model. Overall, the VAL model has given rise to more accurate forecasts than the logistic model. Forecasts from the year 2000 for the VAL model along with the logistic model for the domestic and total consumption are shown in Figures 9 and 10, respectively.

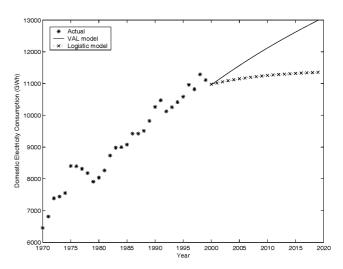


Figure 9 Domestic electricity consumption forecasts of the VAL and logistic model

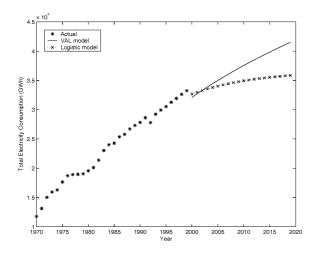


Figure 10 Total electricity consumption forecasts of the VAL and logistic models

The VAL models forecasts appear more consistent with historical data. For the total consumption, the forecasts from the VAL model for the first two years are just lower than those for the logistic model. This is not very significant when the whole set of forecasts are compared. Historically, the logistic model forecasts generally underestimate the actual consumption data.

5 CONCLUSIONS

This paper has proposed a variable asymptote logistic (VAL) model for forecasting electricity consumption in New Zealand. The model has made use of some demographic and economic variables to estimate the saturation levels. The explaining variables were forecasted using ARIMA techniques. A variable saturation level was then used in forecasting electricity consumption. It was found that in the VAL model, the saturation levels were best explained by the combination of population and price of electricity. The addition of the extra variable GDP degraded the model. This indicates that GDP may not be a useful variable in describing forecasting electricity consumption, an observation that is contrary to many forecasting practices using econometrics. The model performed better than the original logistic model in terms of MAPE and gave higher predicted consumption values.

REFERENCES

Bass, F.M. (1969) 'A new product growth for model consumer durables', *Management Science*, Vol. 15, No. 5, pp.215–227.

Bewley, R. and Fiebig, D.G. (1988) 'A flexible logistic growth model with applications in telecommunications', *International Journal of Forecasting*, Vol. 4, pp.177–192.

Bhargava, S., Bhargava, R. and Jain, A. (1991) 'Requirement of dimensional consistency in model equations: diffusion models incorporating price and their applications', *Technological Forecasting and Social Change*, Vol. 41, pp.177–188.

Binroth, W., Burshtein, I., Harboush, R.K. and Hartz, J.R. (1979) 'A comparison of commodity price forecasting by Box-Jenkins and regression-based techniques', *Technological Forecasting and Social Change*, Vol. 14, pp.169–180.

Bodger, P.S. and Tay, H.S. (1987) 'Logistic and energy substitution models for electricity forecasting: a comparison using New Zealand consumption data', *Technological Forecasting and Social Change*, Vol. 31, pp.27–48.

Canarella, G. and Pollard, S.K. (1989) 'Unanticipated monetary growth, output, and the price level in Latin America: an empirical investigation', *Journal of Development Economics*, Vol. 30, No. 2, pp.345–358.

Carrillo, M. and Gonzalez, J.M. (2002) 'A new approach to modeling sigmoid curves', *Technological Forecasting and Social Change*, Vol. 69, pp.233–241.

Farnum, N.R. and Stanton L.W. (1989) *Quantitative Forecasting Methods*, PWS-KENT Publishing Company, USA.

Frank, L.D. (2004) 'An analysis of the effect of the economic situation on modeling and forecasting the diffusion of wireless communications in Finland', *Technological Forecasting and Social Change*, Vol. 71, No. 4, pp.391–403.

Giovanis, A.N. and Skiadas, C.H. (1999) 'A stochastic logistic innovation diffusion model studying the electricity consumption in Greece and the United States', *Technological Forecasting and Social Change*, Vol. 61, pp.235–246.

Harvey, A.C. (1984) 'Time series forecasting based on the logistic curve', *Journal of the Operational Research Society*, Vol. 35, No. 7, pp.641–646.

Ilkka Virtanen, I. and Yli-Olli, P. (1987) 'Forecasting stock market prices in a thin security market', *Omega*, Vol. 15, No. 2, pp.145–155.

Jain, D.C. and Rao, R.C. (1994) 'New concepts and applications in growth phenomena', *Journal of Applied Statistics*, Vol. 21, No. 3, pp.161–190.

Leach, D. (1981) 'Re-evaluation of the logistic curve for human populations', *Journal of the Royal Statistics Society A*, Vol. 144, pp.94–103.

Maharajan, V. and Peterson, R.A. (1978) 'Innovation diffusion in a dynamic-potential adopter population', *Management Science*, Vol. 24, pp.1589–1597.

Maharajan, V., Peterson, R.A., Jain, A.K. and Malhotra, N. (1979) 'A new product growth model with a dynamic market potential. Long range planning', *Technological Forecasting and Social Change*, Vol. 12, pp.51–58.

Makridakis, S., Wheelwright, S.C. and Hyndman, R.J. (1998) Forecasting: Methods and Applications, John Wiley and Sons, Inc., USA.

Mar-Molineo (1980) 'Tractors in Spain: a logistic analysis', Journal of Operational Research Society, Vol. 31, pp.141–152.

Meyer, P.S. and Ausubel, J.H. (1999) 'Carrying capacity: a model with logistically varying limits', *Technological Forecasting and Social Change*, Vol. 61, pp.209–214.

Mohamed, Z. and Bodger, P.S. (2003) 'Analysis of the logistic model for predicting New Zealand electricity consumption', Proceedings of the Electricity Engineer's Association (EEA) New Zealand 2003 Conference, Christchurch, New Zealand, Published in CD-ROM.

Oliver, F.R. (1981) 'Tractors in Spain: a further logistic analysis', *Journal of the Operational Research Society*, Vol. 32, pp.499–502.

Pflaumer, P. (1992) 'Forecasting US population totals with the Box-Jenkins approach', *International Journal of Forecasting*, Vol. 8, pp.329–338.

Sharif, M.N. and Islam, M.N. (1980) 'The Weibull distribution as a general model for forecasting technological change', *Technological Forecasting and Social Change*, Vol. 18, pp.247–256.

- Sharif, M.N. and Ramanathan, K. (1981) 'Binomial innovation diffusion models with dynamic potential adopter population', *Technological Forecasting and Social Change*, Vol. 20, pp.63–87.
- Sharif, M.N. and Ramanathan, K. (1984) 'Temporal models for innovation diffusion', *IEEE Transactions on Engineering Management*, Vol. 31, pp.78–86.
- Sharp, J.A. and Price, H.R. (1990) 'Experience curve models in the electricity supply industry', *International Journal of Forecasting*, Vol. 6, pp.531–540.
- Skiadas, C.H., Papayannakis, L.L. and Mourelatos, A.G. (1993) 'An attempt to improve forecasting ability of growth functions: the Greek electric system', Technological Forecasting and Social Change, Vol. 44, pp.391–404.
- Stapleton, E. (1978) 'The normal distribution as a model for technological substitution', *Technological Forecasting and Social Change*, Vol. 8, pp.325–334.
- Tingyan, X. (1990) 'A combined growth model for trend forecasts', *Technological Forecasting and Social Change*, Vol. 38, pp.175–186.
- Young, P. (1993) 'Technological growth curves: a competition of forecasting models', *Technological Forecasting and Social Change*, Vol. 44, pp.375–389.