



## Forecasting electricity consumption in New Zealand using economic and demographic variables

Zaid Mohamed, Pat Bodger\*

*Department of Electrical and Computer Engineering, University of Canterbury, P.O. Box 4800, Christchurch, New Zealand*

Received 17 October 2003

---

### Abstract

The influence of selected economic and demographic variables on the annual electricity consumption in New Zealand has been investigated. The study uses gross domestic product, average price of electricity and population of New Zealand during the period 1965–1999. Models are developed using multiple linear regression analysis. It was found that the electricity consumption correlated effectively with all variables. Forecasts made using these models were compared with some available national forecasts. The forecasts are also compared with the forecasts of the previously developed Logistic model.

© 2004 Elsevier Ltd. All rights reserved.

---

### 1. Introduction

A number of electricity forecasting models have been developed using economic, social, geographic and demographic factors [1–7]. Egelioglu et al. [1] studied the influence of economic variables on the annual electricity consumption in Northern Cyprus by using multiple linear regression analysis. It was found that the number of customers, the price of electricity and the number of tourists correlate with annual electricity consumption. Harris and Liu [2] found that price plays a major role in explaining conservation behavior by electricity consumers. Yan [3] proposed residential electricity consumption using climatic variables for Hong Kong. Rajan and Jain [4] expressed energy consumption patterns for Delhi as functions of weather and population. Fung and Tummala [5] concluded that it was reasonable to use electricity price, gross domestic product (GDP), deflated domestic exports and population to forecast electricity consumption in Hong Kong. Liu et al. [6] used GDP, real electricity price and population in forecasting

---

\* Corresponding author. Tel.: +64 3 3642070; fax: +64 3 3642761.

E-mail address: [bodger@elec.canterbury.ac.nz](mailto:bodger@elec.canterbury.ac.nz) (P. Bodger).

electricity consumption of Singapore. Lakhani and Bumb [7] used residential price of electricity, per capita income and the estimated long run elasticity of demand in forecasting demand for electricity in Maryland. These findings suggest that a multiple linear regression model using GDP, price and population would provide an appropriate forecasting model for electricity consumption. Makridakis and Wheelwright [8] state that either a simple model that may not completely duplicate reality can be constructed, or a complex model that is more accurate can be built, but this requires a large amount of effort and resources to be developed and manipulated. Even the most sophisticated model would have some part of reality that could not be explained, as the number of factors in real life phenomena is infinite [8].

In attempting to forecast electricity consumption in New Zealand, a Logistic model based on the logistic growth curve has been proposed initially [9]. The Logistic model is a time series extrapolation model. The model involves obtaining the saturation level of the electricity consumption using a Fibonacci search technique [9]. The estimated saturation level is used as a constant variable to fit the best logistic growth curve and forecasts are obtained by extrapolating the fitted curve [9]. A major finding was that the Logistic model very effectively described the historical electricity consumption while producing lower forecasts in general. This is mainly due to the constraints imposed by the saturation level of the logistic growth curve [9]. This opened the opportunity to search for other models that may give rise to higher forecasts.

This paper proposes a multiple linear regression model using GDP, price of electricity and population as selected variables deemed most relevant for New Zealand electricity consumption. These models make use of the strength of the relationship between each of these independent variables and the electricity consumption.

## 2. Proposed model

The proposed multiple linear regression model to forecast electricity consumption using GDP, price of electricity and population is

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + u \quad (1)$$

where

$Y$  is the electricity consumption (GWh),

$X_1$  is GDP (\$NZ millions),

$X_2$  is electricity price (cents/kWh),

$X_3$  is population, and

$u$  is the error (disturbance term or white noise).

Each of the independent variables  $X_1$ ,  $X_2$  and  $X_3$  are themselves obtained from simple linear regression applied to data sets of these variables over time ( $t$ )

$$X_1 = c_{01} + c_{11}t, \quad X_2 = c_{02} + c_{12}t, \quad X_3 = c_{03} + c_{13}t \quad (2)$$

where  $c_{01}$ ,  $c_{11}$ ,  $c_{02}$ ,  $c_{12}$ ,  $c_{03}$  and  $c_{13}$  are the constants of the respective simple linear regressions.

The electricity consumption data is obtained from the *Energy Planning Report* [10] published by the Ministry of Energy, New Zealand and the *Energy Data File* [11] published by the Ministry of Economic Development (MED), New Zealand. GDP data for New Zealand are obtained from Statistics New Zealand [12,13]. Electricity price data are obtained from the MED, New Zealand [10, 14]. The GDP data are in constant 1991/1992 \$NZ. Attempt to use constant electricity prices instead of the current electricity prices degraded the performance of the models. Therefore, current electricity price data in \$NZ are used. The total electricity consumption data is divided into Domestic and Non-Domestic sectors for the purpose of this study. Fig. 1 shows how each of these variables varied over the years 1965–1999.

Various statistical tests are used to validate the models. They include the adjusted coefficient of determination  $r^2$  to determine how well the model explains the actual consumption data, an  $F$ -test for overall significance of the model and a  $t$ -test for testing the strength of each of the individual coefficients of the model [15]. The residual plots of the error,  $u$ , is also studied to determine the appropriateness of the model.

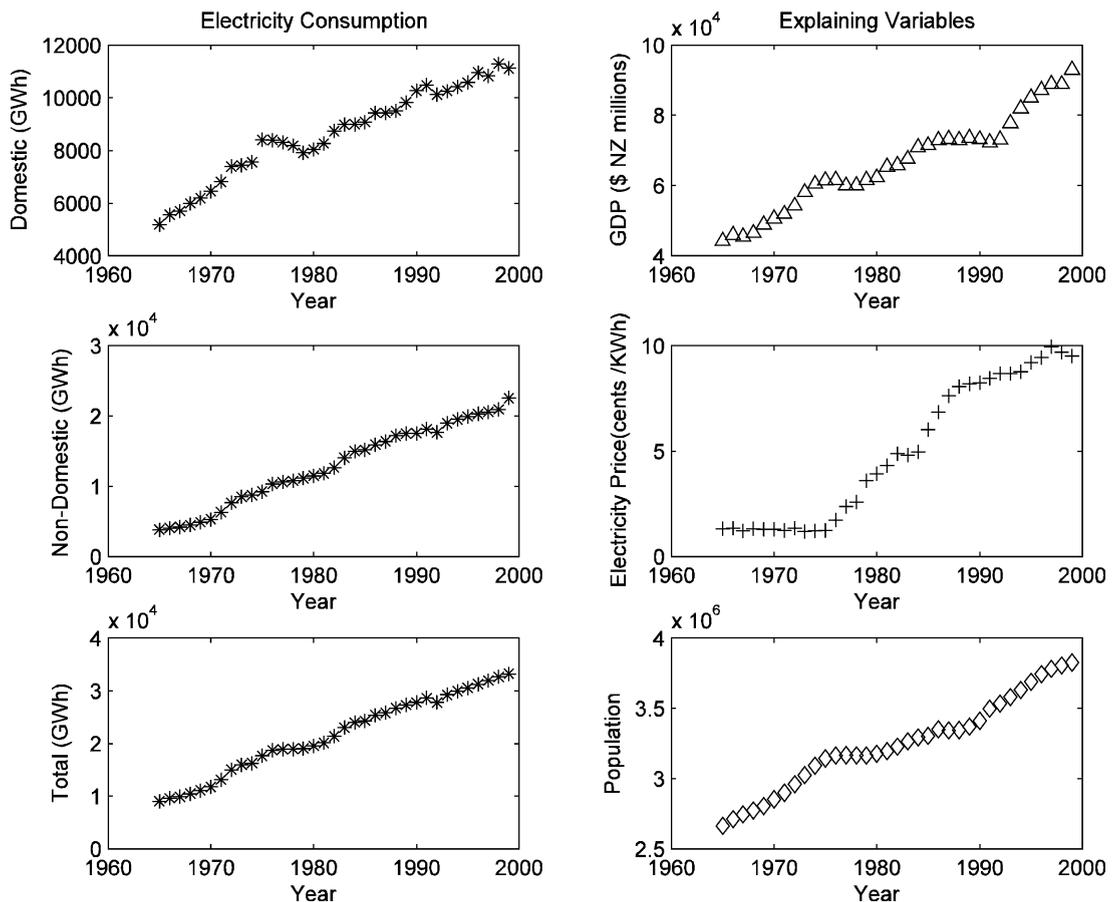


Fig. 1. Historical records of electricity consumption and explaining variables used.

Table 1  
Correlation matrix for variables used in multiple linear regression

	Domestic	Non-domestic	Total	GDP	Price	Population
Domestic	1			0.971	0.921	0.983
Non-domestic		1		0.981	0.963	0.980
Total			1	0.982	0.956	0.984
GDP				1	0.928	0.982
Price					1	0.922
Population						1

The dependency of each of these variables on one another is shown by the respective correlation coefficients. Table 1 shows the correlation matrix for the variables used in the multiple linear regression analysis for modeling New Zealand data from 1965 to 1999.

All the independent variables are highly correlated to the dependent variables (Domestic, Non-Domestic and Total consumption) and therefore are significant in their use in the forecasting model. The correlation coefficient for GDP vs. population is 0.982. It may appear that multicollinearity exists in this situation. Makridakis et al. [15] state that multicollinearity is first and foremost a computational problem. That is, if perfect multicollinearity exists in a regression problem then it is not simply possible to carry out a least squares solution. The program developed for the multiple linear regression model performed effectively for the given data. Attempting to discard population or GDP did not produce a better model. Therefore, it was decided to include these three variables in the model.

The coefficients  $a$ ,  $b_1$ ,  $b_2$  and  $b_3$ , Eq. (1), for each of the Domestic and Non-Domestic sectors and Total consumption are obtained by multiple linear regression using 35 years of data from 1965 to 1999 for each of the variables. The resulting models are:

For Domestic:

$$Y_1 = -5.81 \times 10^3 + 1.5 \times 10^{-3}X_1 - 93.2X_2 + 4.7 \times 10^{-3}X_3 \quad (3)$$

For Non-Domestic:

$$Y_2 = -2.98 \times 10^4 + 2.29 \times 10^{-1}X_1 + 78.7X_2 + 8.2 \times 10^{-2}X_3 \quad (4)$$

For Total:

$$Y_3 = -3.68 \times 10^4 + 2.21 \times 10^{-1}X_1 - 2.10X_2 + 1.34 \times 10^{-2}X_3 \quad (5)$$

The independent variables GDP ( $X_1$ ), price ( $X_2$ ) and population ( $X_3$ ) are estimated using simple linear regression. The resulting equations for the forecasts of  $X_1$ ,  $X_2$  and  $X_3$  are

$$X_1 = -2.49 \times 10^6 + 1.29 \times 10^3t \quad (6)$$

$$X_2 = -267 - 0.130t \quad (7)$$

$$X_3 = -5.91 \times 10^7 + 3.15t \quad (8)$$

where  $t$  is the time in years from 1965.

Table 2 shows the validity test results for the multiple linear regression models. The adjusted coefficient of determination, calculated  $F$  and  $t$  values along with the 99% critical values are given.

Table 2  
Validity test results for multiple linear regression models

Model	Adjusted $r^2$	F-test		t-Test			
		99% critical $F$	$F$	99% critical $t$	$t_1$	$t_2$	$t_3$
Domestic	0.89	4.46	258	2.74	4.04	3.19	30.0
Non-Domestic	0.96	4.46	1077	2.74	17.3	21.6	21.1
Total	0.96	4.46	1024	2.74	10.2	17.8	30.4

For the  $F$  statistics the number of degrees of freedom of the numerator is 3 and the number of degrees of freedom for the denominator is 32. For the  $t$  statistics the number of degrees of freedom is 32.

The Domestic consumption model developed is good with adjusted  $r^2$  of 0.89, but better models may exist as the adjusted  $r^2$  is less than 0.9. During the early 1970s domestic electricity consumption grew rapidly mainly due to the conversion to electric space heating, the near universal use of electric water heating, and the widespread use of appliances such as washing machines and television sets [10]. However, during the late 1970s the electricity consumption dropped noticeably due to a downturn in the economy combined with high electricity prices [10]. Coal and natural gas attracted some of these

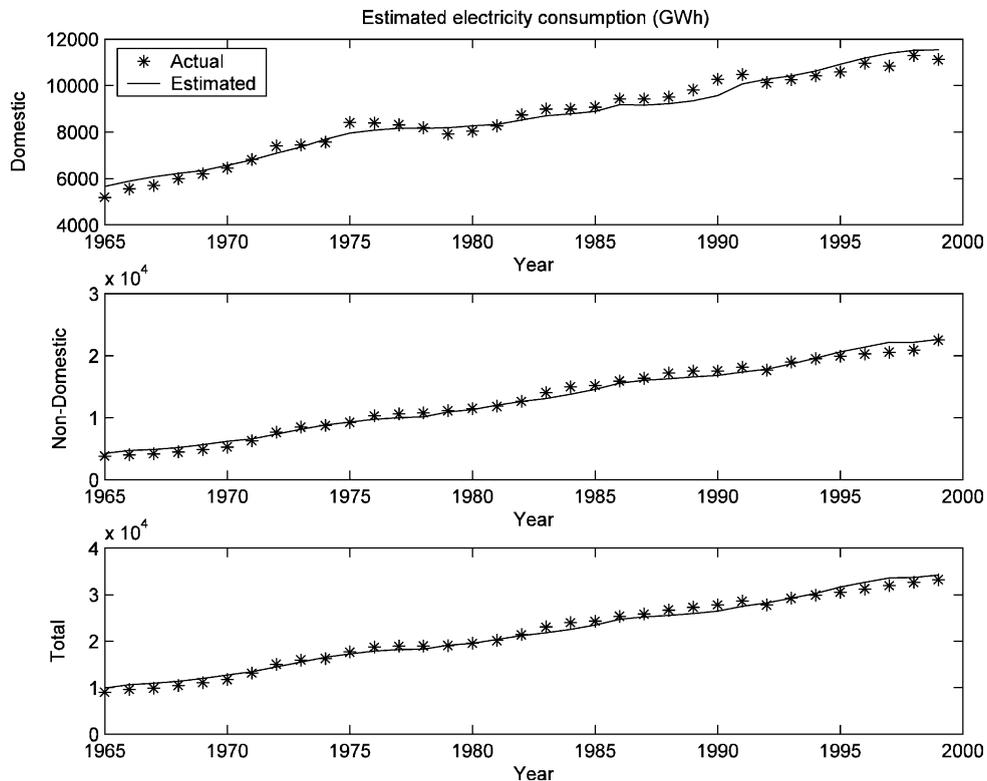


Fig. 2. Actual and estimated electricity consumption for New Zealand.

demand. The increase in domestic electricity consumption in the early 1970s (especially the gap between 1974 and 1975) and the decrease in domestic electricity consumption between 1975 and 1979 may have resulted in the low adjusted  $r^2$  in the Domestic model compared to the other two models.

The adjusted coefficient of determination for the Domestic and Non-Domestic sectors and Total consumption are high implying that even in the worst case of the Domestic sector 89% of the variance in consumption is explained by the combination of GDP, price and population data. Therefore, each of these consumption models coupled with a good forecast of electricity price, GDP and population should produce good forecasts of sector and total electricity consumption. The critical value of  $F$  for each of the sectors is much smaller than the calculated  $F$  value. Therefore, it can be concluded that the multiple linear regression Eqs. (3)–(5) are significant even at the 99% confidence level. Similarly, the  $t$ -test results  $t_1$ ,  $t_2$  and  $t_3$ , for the coefficients of  $X_1$ ,  $X_2$  and  $X_3$ , are higher than the 99% critical value of  $t$ . This means that each of those coefficients  $b_1$ ,  $b_2$  and  $b_3$  for Eqs. (3)–(5) are significant in their use in the models.

Fig. 2 shows the actual electricity consumption along with the estimated values using the models developed. As can be seen, appropriate fit of the historical data is provided by these models.

The residuals produced by these models are also well behaved. Fig. 3 shows the residuals produced by the total consumption model for the three independent variables and the fitted values. The plots are scattered in a horizontal band with no value too far from the band and no patterns such as curvature or increasing spread. Thus, the residuals are deemed to behave randomly and no other explaining variable is required. It seems that there is a relation between residual and price of electricity but this may be due

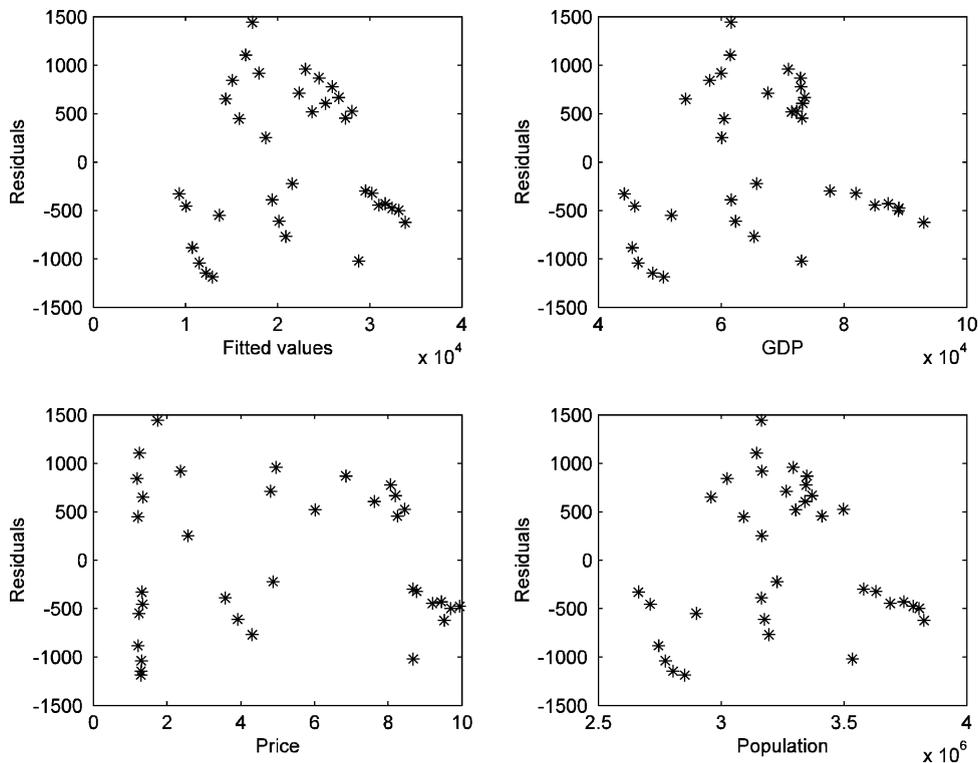


Fig. 3. Plots of residuals against fitted values, GDP, price and population for total consumption.

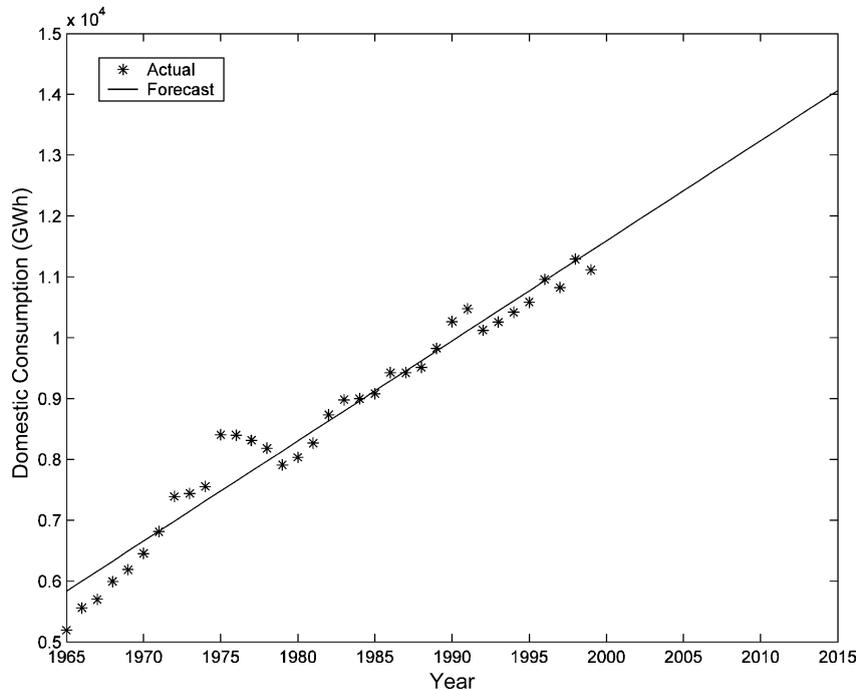


Fig. 4. Forecast for Domestic consumption ( $SSR = 1.05 \times 10^5$ ).

to relatively constant electricity prices between 1965 and 1975. In addition, the Ljung-Box statistic [15] on the residuals with 20 degrees of freedom gave the values of 19.24, 19.64 and 22.81 for the Domestic and Non-Domestic sectors and Total electricity consumption, respectively. These values are much less than the critical  $\chi^2$  value of 31.41 (probability 0.05). These results further suggest that the autocorrelations of the residuals are not significant.

Figs. 4–6 show the forecasts of electricity consumption obtained by the models for the Domestic and Non-Domestic sectors and Total consumption, respectively. The sums of the squared residuals (SSR) of each of the fitted models are also indicated on these figures.

### 3. Comparison of the model

In New Zealand, national electricity forecasts are published by the MED [16] and Sinclair Knight Merz and Centre for Advanced Engineering (CAE) [17]. The MED forecasts are made by the Energy Modelling and Statistics Unit of the Ministry of Commerce (Now Ministry of Economic Development), using its Supply and Demand Energy Model (SADEM). SADEM is a descriptive market equilibrium model focusing on the entire energy sector. The model determines equilibrium in the energy market by projecting demands for a given set of prices and comparing this with the modeled cost of supplying this level of demand [16]. These demands are re-estimated if the prices implied by modeling the level of supply are not consistent with the prices used to determine the initial demand. The process of re-estimation is continued until equilibrium is achieved with demand and supply in balance at market

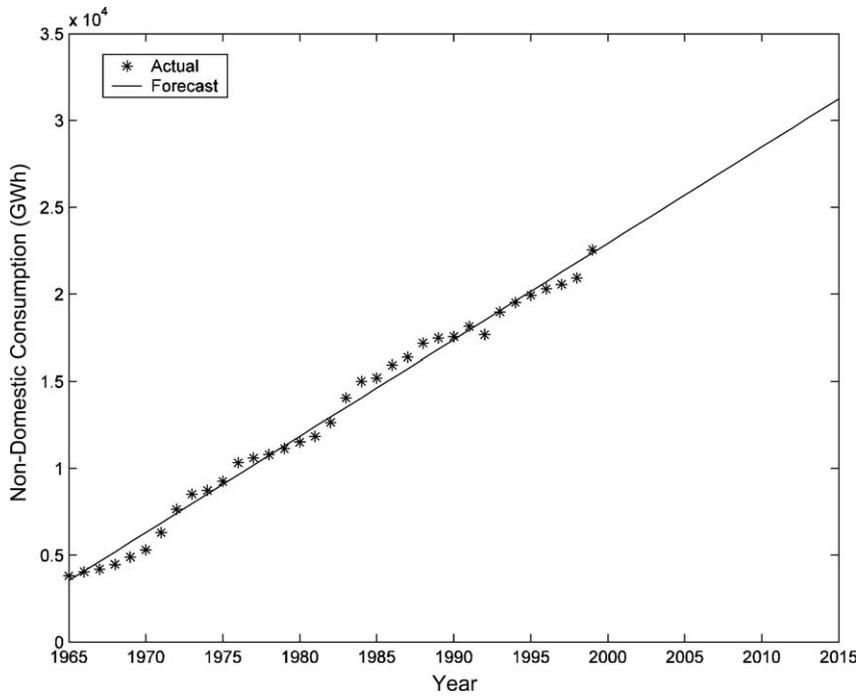


Fig. 5. Forecast for Non-Domestic consumption ( $SSR=3.0 \times 10^5$ ).

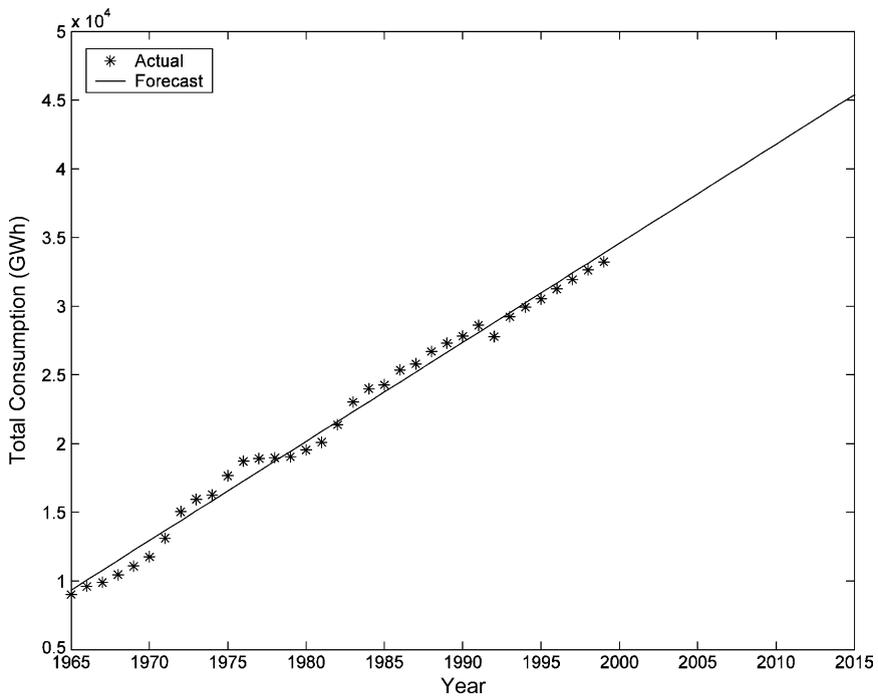


Fig. 6. Forecast for Total consumption ( $SSR=5.30 \times 10^5$ ).

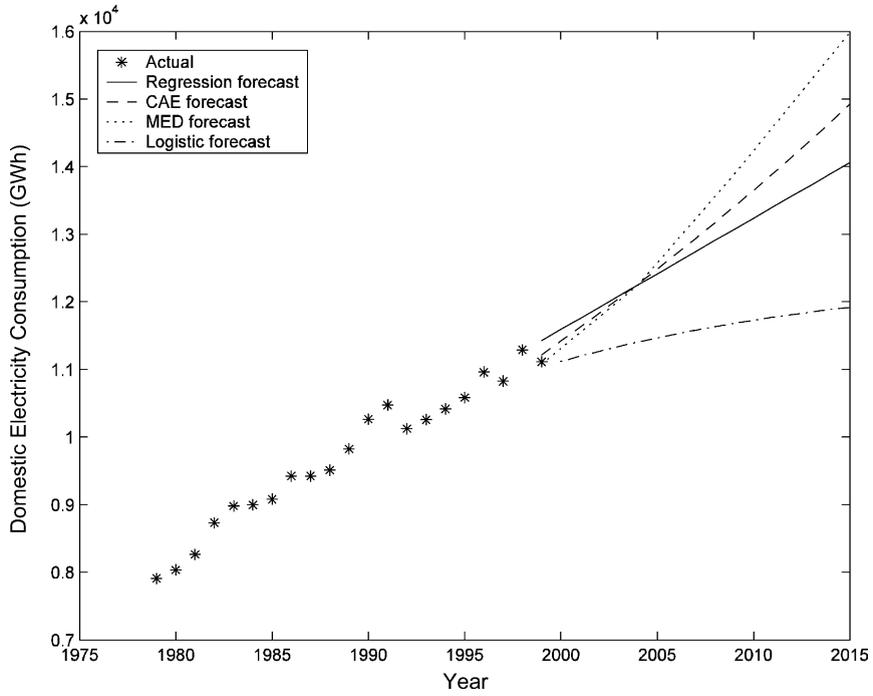


Fig. 7. Comparison of forecasts for the Domestic electricity consumption.

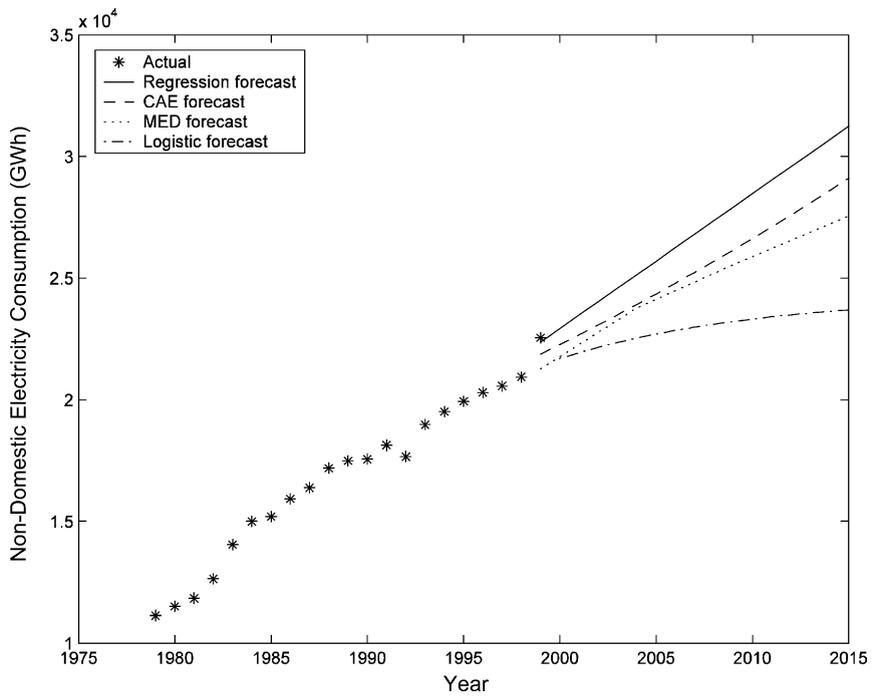


Fig. 8. Comparison of forecasts for the Non-Domestic electricity consumption.

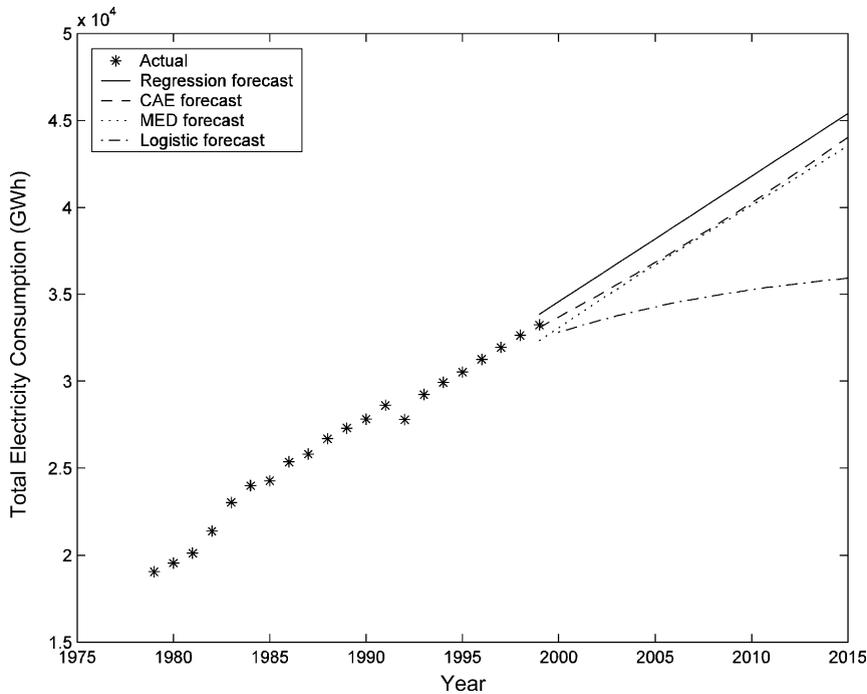


Fig. 9. Comparison of forecasts for the Total electricity consumption.

clearing prices. An interpolation technique is used for updates of the prices. Thus, the forecasts are also interpolated between these years. The CAE forecasts are modeled using an annual load growth of 1.8%. Their study has used 1.8% as the baseline estimate, with 1.3 and 2.3% growth used for sensitivity analysis. This paper uses the 1.8% baseline estimate for comparison purposes.

The forecasts of the MED model, CAE model and Logistic model are compared with the forecasts obtained by the multiple linear regression models. Figs. 7–9 show the corresponding forecasts for the Domestic, Non-Domestic and Total consumption, respectively.

The forecasts obtained by the regression model for the initial years are very similar to CAE and MED forecasts for the Domestic sector. In the latter years, the regression forecasts are lower than CAE and MED forecasts. For the Non-Domestic sector and Total consumption, the regression model forecasts are a little higher on average than the CAE and MED models. In general, the forecasts obtained by these three models are very comparable. It can also be seen that the forecasts obtained by the Logistic model are lower than all the other models for the Domestic and Non-Domestic sectors and Total consumption.

#### 4. Conclusions

Electricity consumption forecasting models based on economic factors for Domestic and Non-Domestic sectors and Total consumption for New Zealand using multiple linear regression have been proposed. The models performed effectively in the statistical tests conducted, implying their significance in forecasting electricity consumption using the explaining variables considered. Comparisons of these

models have been made with the national forecasts available in New Zealand. The comparison revealed that the forecasts made by the regression models are very comparable with the national forecasts.

The accuracy of the forecasts made by these models depends strongly on the accuracy of forecasts made for the explaining variables. In this paper, simple regression had been used to model these variables. Currently other methods of forecasting these variables are being studied.

## References

- [1] Egelioglu F, Mohamad AA, Guven H. Economic variables and electricity consumption in Northern Cyprus. *Energy* 2001; 26:355–62.
- [2] Harris JL, Liu L. Dynamic structural analysis and forecasting of residential electricity consumption. *Int J Forecast* 1993;9: 437–55.
- [3] Yan YY. Climate and residential electricity consumption in Hong Kong. *Energy* 1998;23(1):17–20.
- [4] Rajan M, Jain VK. Modelling of electrical energy consumption in Delhi. *Energy* 1999;24:351–61.
- [5] Fung YH, Tummala VMR. Forecasting of electricity consumption: a comparative analysis of regression and artificial neural network models. *IEE Second International Conference on Advances in Power System Control, Operation and Management, Hong Kong; 1993.* p. 782–7.
- [6] Liu XQ, Ang BW, Goh TN. In: Forecasting of electricity consumption: a comparison between an econometric model and a neural network model. In: *IEEE International Conference on Neural Networks*, vol. 2; 1991. p. 1254–9.
- [7] Lakhani HG, Bumb B. Forecasting demand for electricity in Maryland: an econometric approach. *Technol Forecast Soc Change* 1978;11:237–61.
- [8] Makridakis S, Wheelwright SC. *Forecasting methods for management*, 5th ed. New York: Wiley; 1989.
- [9] Mohamed Z, Bodger PS. Analysis of the Logistic model for predicting New Zealand electricity consumption. *Proceedings of the Electricity Engineer's Association (EEA) New Zealand Conference, Christchurch, New Zealand, Published in CD-ROM, 20–21 June; 2003.*
- [10] Ministry of Energy. *Electricity Forecasting and Planning: a Background Report to the 1984 Energy Plan*. Private Bag, Wellington, New Zealand, issues 1982–1984.
- [11] Ministry of Economic Development. *New Zealand Energy Data File*, Wellington: Crown Copyright; 2002.
- [12] Department of Statistics. *The New Zealand Official Year Book*, Wellington: NZ Govt. Printer, New Zealand; 2000 and 2002.
- [13] Statistics New Zealand. P.O. Box 2922, Wellington, New Zealand, Web page: <http://www.stats.govt.nz>
- [14] Ministry of Economic Development. *Annual Statistics in Relation to Electric Power Development and Operation year ended 1958–1974 condensed statistics for the combined developments and operations of the New Zealand Electricity Department and Distributing Electrical Supply Authorities*, Wellington: A.R. Shearer, Government Printer; 1958–1974.
- [15] Makridakis S, Wheelwright SC, Hyndman RJ. *Forecasting methods and applications*. New York: Wiley; 1998.
- [16] Ministry of Economic Development, Modelling and Statistics Unit. *New Zealand Energy Outlook to 2020*, Wellington: Crown Copyright; 2000.
- [17] Sinclair Knight Merz, CAE (Centre for Advanced Engineering, University of Canterbury, NZ). *Electricity Supply and Demand to 2015*, 5th ed. Christchurch: CAE University of Canterbury Campus; 2000.