Available online at www.sciencedirect.com
SCIENCE DIRECT
ENERGY
Energy 31 (2006) 2129–2141
www.elsevier.com/locate/energy

Comparing linear and nonlinear forecasts for Taiwan's electricity consumption

Hsiao-Tien Pao*

Department of Management Science, National Chiao Tung University, Hsinchu, Taiwan, ROC

Received 28 June 2004

Abstract

This paper uses linear and nonlinear statistical models, including artificial neural network (ANN) methods, to investigate the influence of the four economic factors, which are the national income (NI), population (POP), gross of domestic production (GDP), and consumer price index (CPI) on the electricity consumption in Taiwan and then to develop an economic forecasting model. Both methods agree that POP and NI influence electricity consumption the most, whereas GDP the least. The results of comparing the out-of-sample forecasting capabilities of the two methods indicate the following. (1) If given a large amount of historical data, the forecasts of ARMAX are better than the other linear models. (2) The linear model is weaker on foretelling peaks and bottoms regardless the amount of historical data. (3) The forecasting performance of ANN is higher than the other linear models based on two sets of historical data considered in the paper. This is probably due to the fact that the ANN model is capable of catching sophisticated nonlinear integrating effects through a learning process. To sum up, the ANN method is more appropriate than the linear method for developing a forecasting model of electricity consumption. Moreover, researchers can employ either ANN or linear model to extract the important economic factors of the electricity consumption in Taiwan.

Keywords: Artificial neural networks; Energy forecasting; ARMAX models

1. Introduction

Modeling electrical demand and energy consumption is usually based on historical consumption and the relationship of this consumption to other relevant variables, such as: economic, demographic, climatic, energy price, etc. Multivariate modeling along with cointegration techniques or regression analysis were used in a number of studies on different countries [1–7] to investigate the influence of different determinants on energy consumption. Glasure and Lee [1] presented a bidirectional causality relationship between energy consumption and gross of domestic production (GDP) for South Korea and Singapore using the cointegration and error-correction models (ECM). Shiu and Lam [2] found that there is a unidirectional relationship running from electricity consumption to real GDP for China using ECM model. Soytas and Sari [3] examines

*Tel.: +886 35131578; fax: +886 3 5710102.
E-mail address: htpao@cc.nctu.edu.tw.

0360-5442/$ - see front matter
doi:10.1016/j.energy.2005.08.010

the causal relationship between GDP and energy consumption in the top 10 emerging markets and the G-7 countries. Yang [4] found a bidirectional causality between energy consumption and GDP for Taiwan. Additionally, Yan [5] presented residential electricity consumption models using climate variables for Hong Kong. Rajan and Jain [6] expressed energy consumption patterns for Delhi as functions of weather and population. Moreover, Egelioglu et al. [7] found that the model using the number of customers, the number of tourists, and the electricity prices as regressors has very strong predictive ability for Northern Cyprus.

Recently, some studies have analyzed forecasting performance for energy consumption using different models on different countries [8–10]. Darbellay and Slama [8] compared the predictions of nonlinear artificial neural networks (ANNs) with linear ARIMA models for Czech electric consumption. They found that, for univariate modeling, the forecasting abilities of a linear model and a nonlinear model were not very different. For multivariate modeling, adding the temperature as an external input allows ANNs to integrate more information and thus produce better forecasts. Saab et al. [9] investigated three univariate models, AR, ARIMA, and AR(1)/highpass filter, to forecast electrical energy consumption in Lebanon. They found that AR(1)/highpass filter model yielded the best forecast for this data set. Fatai et al. [10] used three econometric approaches to analyze the pattern of electricity consumption in New Zealand. They found that autoregressive distributed lag approach has the best forecasting performance. Variables affecting demand and energy consumption may vary from one region to another. A model developed for one region may not be appropriate for another region. Electrical consumption models are required for a variety of utility activities. Therefore, a model should be developed in different regions for efficient planning and organization.

Taiwan's energy consumption rises sharply from 52.01 million kiloliters of oil in 1990 to 103.42 million-kiloliters of oil in 2003 because of rapid economy growth and higher living standard. Among the energy forms consumed, petroleum took up 44.2% in 1990 to 39% in 2003; coal 12.8–10.7%; natural gas & liquid natural gas 3–2.3%; and electricity 40–48%. Electricity consumption takes up almost 50% of the total final energy consumption in 2003.

The main object of this study is to analyze and forecast Taiwan's electricity consumption. Variables national income (NI), population (POP), GDP, consumer price index (CPI), climate, and electricity price are concluded to be the most possible factors to affect electricity consumption according to relating literature worldwide [1–10], but the electricity price is not used because it is fixed and rarely changed in Taiwan. Therefore, this paper proposes linear and nonlinear economy models to investigate the influence of NI, POP, GDP, and CPI variables on the electricity consumption and then to forecast the consumption. They are useful for the government authorities to control electrical energy supply. The rest of this paper is organized as follows: Section 2 records the results of data analysis and review. Section 3 proposes three types of linear model and nonlinear ANN models to analyze the relativity between the four economic factors and Taiwan's electricity consumption and to build economic forecasting models. Three statistics for comparison of linear and nonlinear models are also introduced in this section. Section 4 presents the empirical results and comparisons of the models, whereas the last section briefs our findings and presents conclusions.

2. Data analysis

The data applied here are 156 monthly data recorded in January 1990 through December 2002. The values of ELEC, NI, POP, GDP and CPI are collected from the TEDC (Taiwan Economic Data Center) database supervised by the Education Ministry Taiwan. Additionally, the monthly temperature (TEMP) record is collected from 25 representative weather stations, set up by Taiwan Central Weather Bureau.

Table 1 shows the descriptive statistics for variables, and their linear relativities are recorded in Table 2. As we can see, NI has rather high linear relativity with POP (0.99) and CPI (0.97), while POP reacts the same to CPI (0.96). Fig. 1 shows the scatter plot of ELEC to four economic factors. Graphs exhibit logarithmic trend. So, ln (ELEC) is considered dependent variable in this research. The linear relativities that ln (ELEC) has with GDP, POP, NI and CPI are shown in Fig. 2; that of ELEC and ln (ELEC) with TEMP is shown in Fig. 3. Fig. 4 shows the electricity consumed annually from 1990 to 2002. The consumption gradually increases to hit the peak in summertime, and decreases to the bottom in wintertime.

Table 1
Descriptive statistic analysis of variables

Variables	Size	Mean	SD	Minimum	Maximum
ELEC (million)	156	10981.82	2968.50	5545.46	17850.50
LELEC (million)	156	9.27	0.28	8.62	9.79
GDP (million)	156	2629508	381299	1848485	3522562
POP (million)	156	21.379	0.688	20.117	22.453
NI (million)	156	576632	143911	322459	770914
CPI (%)	156	92.42	7.78	75.12	101.97
TEMP (°C)	156	21.91	3.91	14.44	27.96

Table 2
Correlations of the variables

	ELEC	LELEC	GDP	POP	NI	CPI	TEMP
ELEC	1.00	0.990 (<0.0001)	0.63 (<0.0001)	0.92 (<0.0001)	0.89 (<0.0001)	0.86 (<0.0001)	0.45 (<0.0001)
LELEC		1.00	0.65 (<0.0001)	0.93 (<0.0001)	0.91 (<0.0001)	0.90 (<0.0001)	0.45 (<0.0001)
GDP			1.00	0.68 (<0.0001)	0.68 (<0.0001)	0.67 (<0.0001)	0.06 (0.4894)
POP				1.00	**0.99** (<0.0001)	**0.96** (<0.0001)	0.15 (0.0558)
NI					1.00	**0.97** (<0.0001)	0.11 (0.1572)
CPI						1.00	0.16 (0.0478)
TEMP							1.00

p-Values are in the parentheses.

3. Methodology

Electricity demand has always been one of the critical economic issues in Taiwan. To investigate the influence of the economic variables on electricity consumption and then to forecast the consumption, the nonlinear ANN and three types of linear models, multiple log-linear regression (LNREG), response surface regression (RSREG), and regression with ARMA errors model (ARMAX) are proposed. The out-of-sample forecasting capabilities of all models are then compared.

3.1. LNREG, ARMAX and RSREG linear models

This section describes three types of linear model to analyze and forecast Taiwan's electricity consumption. Fig. 2 shows a linear relativity between logarithm function of electricity consumption and individual variable, and Table 2 shows a close linear relativity among the four factors. As a result, the following seven log-linear regression models are employed firstly. The last four models (Models 4–7) are nested, variables GDP, CPI, and POP are added into Model 4 stepwise. The experimental formula (Model 5) is extracted after deleting POP

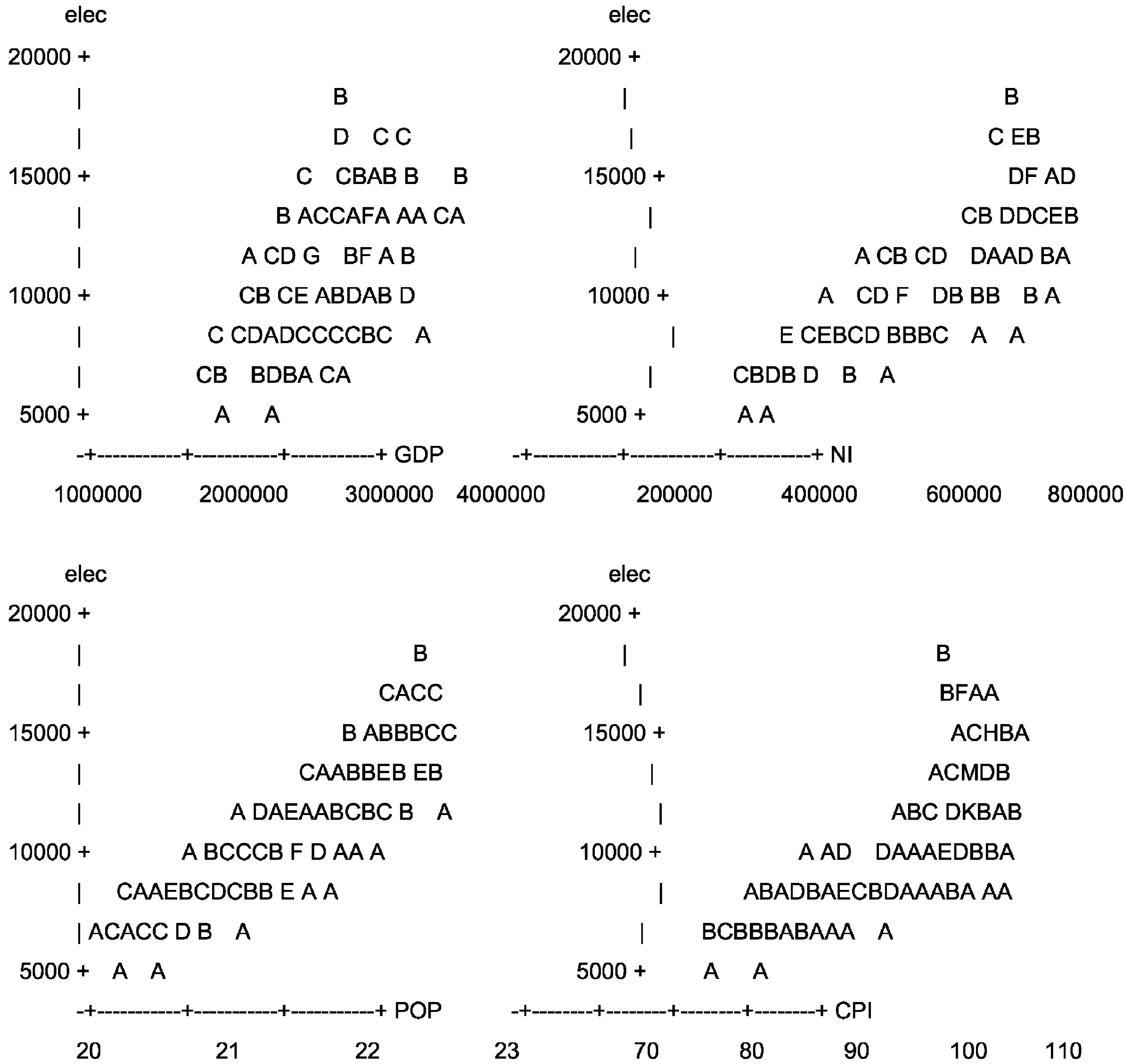

Fig. 1. Scatter plot of ELEC vs. independent variables.

and CPI variables selected by the Model 7 with statistics of Mallows $C(p)$ [11], Adj-R^2, and VIF (Variance Inflation Factor). The procedure is practiced to prevent multicollinearity.

$$\text{Model 1}: \quad \mathrm{LELEC}_t = a_1 + b_1\mathrm{TEMP}_t + c_1\mathrm{GDP}_t + \varepsilon_{1t}, \tag{1}$$

$$\text{Model 2}: \quad \mathrm{LELEC}_t = a_2 + b_2\mathrm{TEMP}_t + c_2\mathrm{CPI}_t + \varepsilon_{2t}, \tag{2}$$

$$\text{Model 3}: \quad \mathrm{LELEC}_t = a_3 + b_3\mathrm{TEMP}_t + c_3\mathrm{POP}_t + \varepsilon_{3t}, \tag{3}$$

$$\text{Model 4}: \quad \mathrm{LELEC}_t = a_4 + b_4\mathrm{TEMP}_t + c_4\mathrm{NI}_t + \varepsilon_{4t}, \tag{4}$$

$$\text{Model 5}: \quad \mathrm{LELEC}_t = a_5 + b_5\mathrm{TEMP}_t + c_5\mathrm{NI}_t + d_5\mathrm{GDP}_t + \varepsilon_{5t}, \tag{5}$$

$$\text{Model 6}: \quad \mathrm{LELEC}_t = a_6 + b_6\mathrm{TEMP}_t + c_6\mathrm{NI}_t + d_6\mathrm{GDP}_t + e_6\mathrm{CPI}_t + \varepsilon_{6t}, \tag{6}$$

$$\text{Model 7}: \quad \mathrm{LELEC}_t = a_7 + b_7\mathrm{TEMP}_t + c_7\mathrm{NI}_t + d_7\mathrm{GDP}_t + e_7\mathrm{CPI}_t + f_7\mathrm{POP}_t + \varepsilon_{7t}, \tag{7}$$

where LELEC = ln (ELEC).

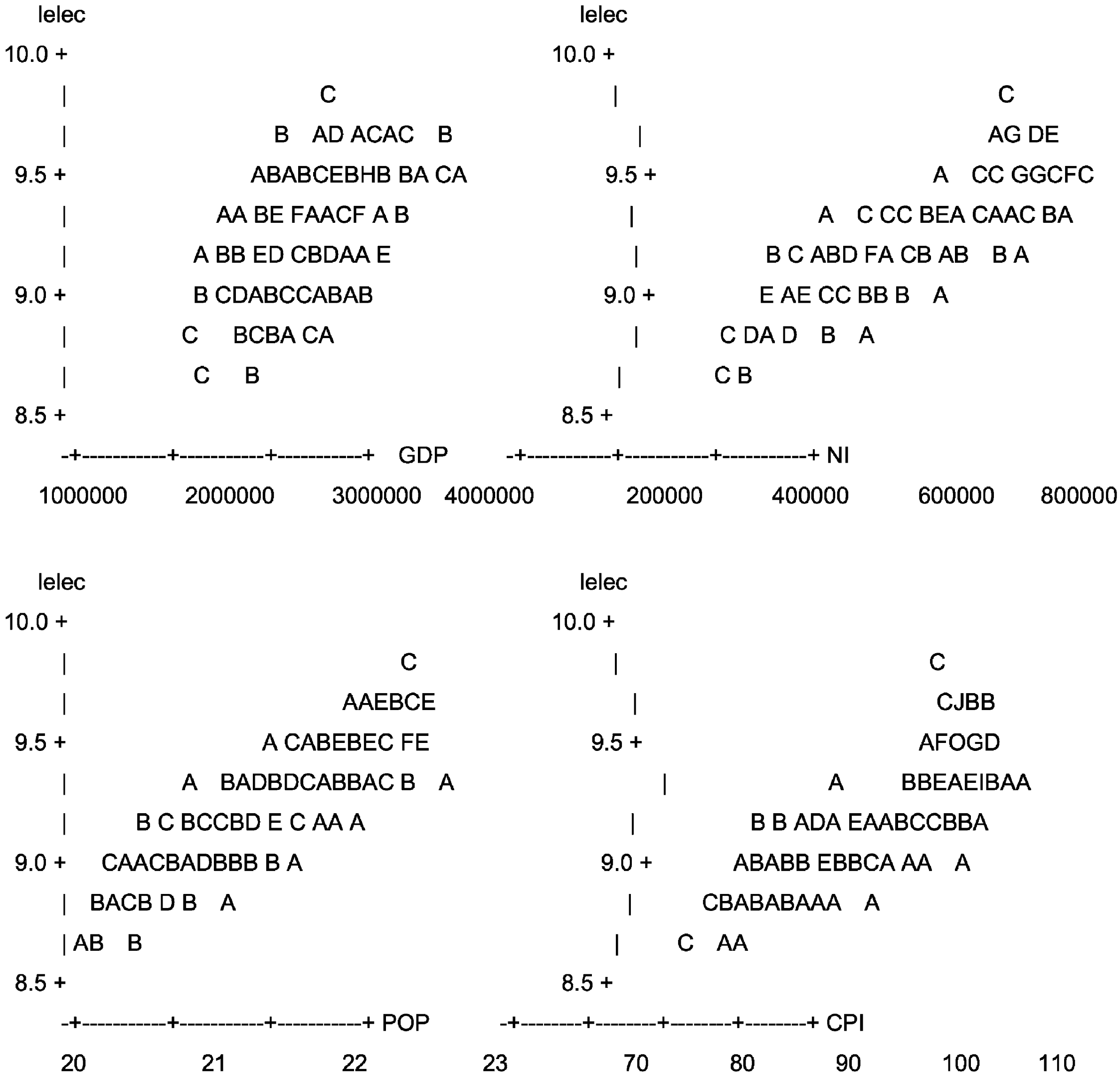

Fig. 2. Scatter plot of ln (ELEC) vs. independent variables.

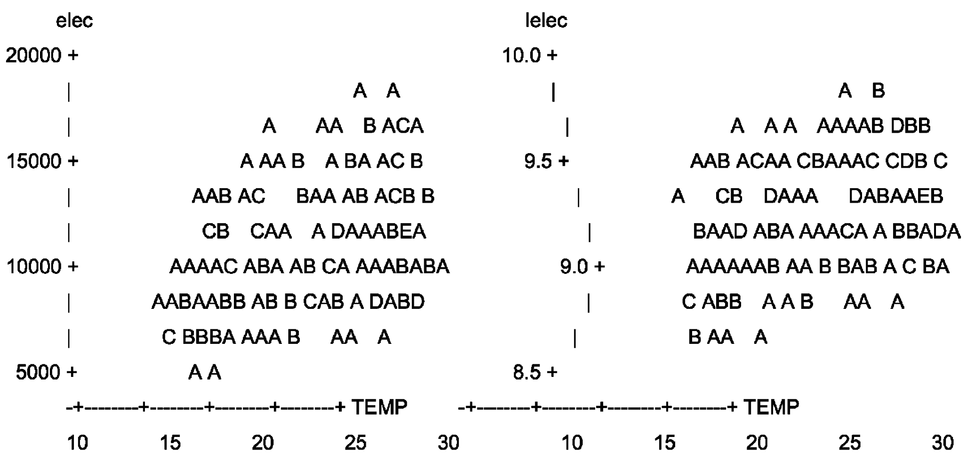

Fig. 3. Scatter plot of (ELEC, ln (ELEC)) vs. temperature variable.

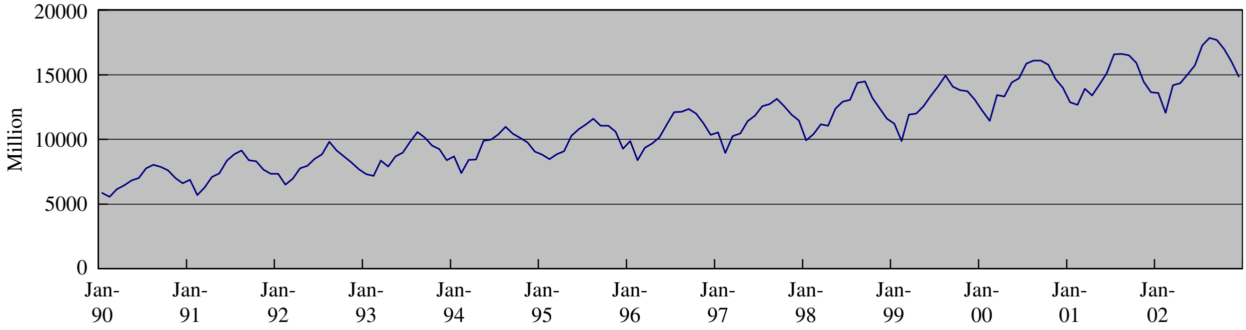

Fig. 4. Electricity consumption in Taiwan.

Among the above models, Model 1 through Model 4 present how the four economic factors affect electricity consumption under controlled the temperature, respectively. And Model 5 is the economic forecasting model of electricity consumption. Compared with Model 4, Model 5 is able to tell how the GDP has an effect on Taiwan's electricity consumption. On the other hand, standardized regression model (SRM) is applied to estimate coefficients values. The standardized estimates may be used to compare the impact of independent variables on the dependent variable. The Durbin–Watson test is applied to test the existence of the first order autoregressive process for each model. The results of normality test, *F*-test, and DW test are presented in Table 3.

The model residuals are amplified by exponential transformation because we need to fit them with the log-linear regression model and then must convert prediction for logarithmic back to untransformed units. So, we build ARMAX and RSREG, two linear models, on the original values to analyze and forecast electricity consumption. When the residuals in the regression model are significantly auto-correlated, the general ARMAX model is used, which is written as

$$\mathrm{ELEC}_t = \mu + \omega_1 \mathrm{TEMP}_t + \omega_2 \mathrm{NI}_t + \omega_3 \mathrm{GDP}_t + \frac{\theta_q(B)}{\phi_p(B)} a_t, \tag{8}$$

where

$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q \quad \text{and}$$
$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p.$$

The roots of $\phi_p(B) = 0$ and $\theta_q(B) = 0$ should all lie outside the unit circle, where B is a backward shift operator: $B^m(Y_t) = Y_{t-m}$. This model shows that the response series is a combination of past values of random shocks a_t and past values of other input series. ELEC_t serves as the response series while TEMP_t, NI_t, GDP_t are the input series.

The response surface regression model (RSREG) is another linear model to be used on the original values to analyze and forecast electricity consumption. It fits the parameters of a complete quadratic response surface and analyzes the fitted surface to determine the factor levels of optimal response, which is written as:

$$\begin{aligned}\mathrm{ELEC}_t = {} & \omega_0 + \omega_1 \mathrm{TEMP}_t + \omega_2 \mathrm{NI}_t + \omega_3 \mathrm{GDP}_t + \omega_4 \mathrm{TEMP}_t \times \mathrm{TEMP}_t + \omega_5 \mathrm{NI}_t \times \mathrm{NI}_t \\ & + \omega_6 \mathrm{GDP}_t \times \mathrm{GDP}_t + \omega_7 \mathrm{TEMP}_t \times \mathrm{NI}_t + \omega_8 \mathrm{TEMP}_t \times \mathrm{GDP}_t + \omega_9 \mathrm{NI}_t \times \mathrm{GDP}_t + \varepsilon_t.\end{aligned} \tag{9}$$

An important limitation of response models is that they provide approximations to a surface, and the model coefficients usually have no practical interpretation. This paper uses the above three models to predict Taiwan's electricity consumption.

3.2. ANN models

The (ANN) consists of an input layer, an output layer and one or more intervening layers also referred to as hidden layers. Each layer consists of multiple neurons that are connected to neurons in adjacent layers. The connection weights and node biases are the model parameters. Fig. 5 is a popular feed forward multilayer

Table 3
Coefficients using regression and standardized regression

		Intercept	TEMP	GDP	CPI	POP	NI	Adj-R^2	Normality (p-val)	F value (p-val)	DW	Size
Model 1	Estimate	7.554*	0.028*	4.07E-7*						96.24	0.225	132
		(63.05)	(7.77)	(10.92)				**0.5925**	0.0060	(<0.0001)		
	Stand.	0	0.434	0.610								
Model 2	Estimate	6.226*	0.022*		0.027*					806.92	1.137	132
		(83.6)	(14.42)		(34.88)			0.9248	0.3100	(<0.0001)		
	Stand.	0	0.349		0.845							
Model 3	Estimate	1.111*	0.023*			0.359*				1294.55	1.325	132
		(6.55)	(18.29)			(44.38)		0.9518	0.0379	(<0.0001)		
	Stand.	0	0.355			0.860						
Model 4	Estimate	7.818*	0.024*				1.57E-6*			1291.61	1.265	132
		(247.62)	(19.69)				(44.33)	**0.9517**	0.0197	(<0.0001)		
	Stand.	0	0.381				0.857					
Model 5	Estimate	**7.732***	**0.024***	**5.44E-8***			**1.47E-6***			**930.59**	**1.373**	132
		(192.66)	(20.45)	(3.30)			(32.30)	**0.9551**	**0.1438**	**(<0.0001)**		
	Stand.	0	**0.381**	**0.082**			**0.803**					
	VIF	0	**1.01**	**1.79**			**1.80**					
Model 6	Estimate	7.449*	0.024*	5.39E-8*	0.005		1.21E-6*			712.53	1.442	132
		(47.83)	(19.99)	(3.30)	(1.87)		(8.30)	0.9560	0.1898	(<0.0001)		
	Stand.	0	0.375	0.081	0.147		0.661					
	VIF	0	1.05	1.79	**18.38**		**18.84**					
Model 7	Estimate	5.247*	0.023*	5.04E-8*	0.004	0.121	7.39E-7*			575.841	1.463	132
		(3.51)	(18.93)	(3.07)	(1.50)	(1.48)	(2.11)	0.9564	0.0823	(<0.0001)		
	Stand.	0	0.367	0.076	0.120	0.290	0.403					
	VIF	0	1.13	1.82	19.36	**114.70**	**109.71**					

t Statistics are in the parentheses.
*Indicates statistical significance at the 0.05 level.

ANN model. The input layer can be represented by a vector $X = (x_1, x_2, \ldots, x_m)'$, the middle layer can be represented by a vector $M = (m_1, m_2, \ldots, m_h)'$, and y is the output.

To use an ANN model for forecasting, forecasters must first build it. The model building process is called the network training or learning. Usually, in applications of ANNs, total available data are split into a training set and a test set. The training set is used to build the network model and then the forecasting ability of the network is evaluated from the test set. During the training process, a weighted sum of the inputs is calculated at tth hidden node:

$$\mathrm{NET}_t = \sum_{i=1}^{m} w_{ti} x_i + b_t, \quad t = 1, 2, \ldots, h. \tag{10}$$

Each hidden node then uses a sigmoid transfer function to generate an output:

$$m_t = [1 + \exp(-\mathrm{NET}_t)]^{-1} = f(\mathrm{NET}_t), \quad t = 1, 2, \ldots, h. \tag{11}$$

It is between 0 and 1. The outputs from each of the hidden nodes, along with the bias input b_o, are then sent to the output node and again calculated a weighted sum,

$$\mathrm{NET}_0 = \sum_{t=1}^{h} v_t m_t + b_0. \tag{12}$$

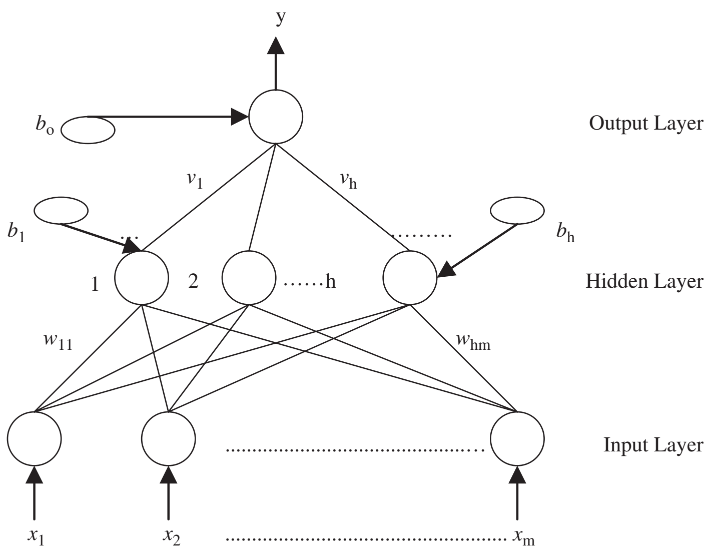

Fig. 5. Neural network model.

The weighted sum becomes the input to the sigmoid transfer function of the output node. The jth resulting output is then scaled to provide the predicted output value:

$$\hat{Y}_j = f(\mathrm{NET}_0) = [1 + \exp(-\mathrm{NET}_0)]^{-1}, \quad j = 1, 2, \ldots, n. \tag{13}$$

At this point, the second phase of the back-propagation algorithm, adjustment of the connection weights, begins. The connection weights of the neural network can be determined by minimizing the objective function of SSE in the training process:

$$\mathrm{SSE} = \sum_{j=1}^{n} (y_j - \hat{Y}_j)^2, \tag{14}$$

where n is the number of the training data.

Assume the relationship of Y and X is monotone, then to calculate the sensitivity S_i of the outputs to each of the ith inputs as a partial derivative of the output with respect to the input [12],

$$S_i = \frac{\partial \hat{Y}}{\partial X_i} = \sum_{t=1}^{h} \frac{\partial \hat{Y}}{\partial \mathrm{NET}_0} \frac{\partial \mathrm{NET}_0}{\partial m_t} \frac{\partial m_t}{\partial \mathrm{NET}_t} \frac{\partial \mathrm{NET}_t}{\partial X_i} = \sum_{t=1}^{h} [f'(\mathrm{NET}_0) v_t f'(\mathrm{NET}_t) w_{ti}]. \tag{15}$$

Assume f' (NET_0) and f' (NET_t) are constants and we ignore them, the relative sensitivity is

$$\hat{S}_i = \sum_{t=1}^{h} v_t w_{ti}. \tag{16}$$

The input variable with higher absolute value of relative sensitivity has the bigger impact on the output variable. The step-by-step process is given below.

1. Training a neural network on all available data.
2. Using these weights, compute the sensitivity for each input variable.

More detailed materials about neural network learning can be found in Bishop [13].

Based on the above models, the following section focuses on the out-of-sample forecasting ability of the linear and nonlinear models. More advanced models assume their residual values behave differently and offer more options for better estimates. Conclusions and policies implications are then drawn as valuable tools for managers in achieving optimal forecasting model.

3.3. Forecasting evaluation methods

For the purpose of evaluating out-of-sample forecasting capability, we examine forecast accuracy by calculating three different evaluation statistics, which are the root mean square error (RMSE), the mean absolute error (MAE), and the mean absolute percentage error (MAPE). Their definitions are in the following:

$$\mathrm{RMSE} = \sqrt{\sum_{i=1}^{n} (P_i - A_i)^2 \Big/ n}, \tag{17}$$

$$\mathrm{MAE} = \sum_{i=1}^{n} |P_i - A_i| \Big/ n, \tag{18}$$

$$\mathrm{MAPE} = \left(\sum_{i=1}^{n} |(P_i - A_i)/A_i| \Big/ n \right) \times 100, \tag{19}$$

where P_i and A_i are ith predicted and actual value, and n is the total number of predictions.

4. Empirical results

In this section, we use LNREG, RSREG, ARMAX linear models, and ANN nonlinear models to investigate the influence of the economic variables, POP, NI, GDP, and CPI on the electricity consumption and then to build a forecasting model in Taiwan. The forecasting performances of the ANN models are compared with those of the linear models using electricity consumption. The period under examination extends from January 1990 to December 2002, with a total of 156 observations for each series. The data set is used in two different ways in an attempt to experiment on the amount of historical data required for generating better forecasts. The first experiment uses the 11-year data set from January 1990 to December 2000 as the training data, and forecast from January 2001 to December 2002. The second experiment uses only the 4-year data set from January 1997 to December 2000 as the training data, and forecasts for the same time horizon. The in-sample training period is used to build models and the out-of-sample testing period is used to evaluate prediction ability.

4.1. Building LNREG, ARMAX and RSREG linear models

Models 1–4 in Section 3.1 investigate the influence of the four economic factors on the electricity consumption, respectively, under the fixed temperatures. Table 3 indicates that POP and NI affect Taiwan's electricity consumption the most reaching 95.18% and 95.17%, respectively, followed by CPI (92.48%), and GDP scores the least (59.25%) by using a 11-year training data set. The normality test and Durbin–Watson test for residuals of each model are recorded in Table 3. Though GDP and CPI are both significant indicators to a country's economy, they do not affect Taiwan's electricity consumption as much as the others. This could result from Taiwan government's enforcement on electricity saving and economizing policies.

Model 4 to Model 7 are nested models. Variables GDP, CPI and POP are added into Model 4 stepwise. Adding CPI and POP to Model 5 increases Adj-R^2 from 0.9551 to 0.9564, but their VIF are 114.70 and 109.71. According as the value of VIF in Table 3 on the Models 4–7, Model 5 has much less multicollinearity than Models 6 and 7. Moreover, from Table 2, we know that NI has rather high linear correlation with POP (0.99) and CPI (0.97). All of them indicate that NI and POP, NI and CPI contribute very redundant information. To avoid tedious linear model, TEMP, NI, and GDP are chosen as explanatory variables to predict electricity consumption (Model 5). Results of Table 3 show that Model 5 has no multicollinearity problem, and both NI and GDP are important significantly. Compared with Model 4, GDP on Model 5 has significant effect on Taiwan's electricity consumption, but it only surpasses Model 4 by 0.34% (= 95.51%−95.17%). The estimated coefficients of standardized regression model show the most crucial factors affecting Taiwan's

Table 4
Coefficients of the linear models

	Incept.	Temp	NI	GDP	Temp^2	NI^2	GDP^2	Temp* NI	Temp* GDP	NI* GDP	R^2
Log-linear regression model											
11-year	7.73*	0.02*	1.5E-6*	5.4E-8*							0.96
	(0.04)	(0.001)	(4.6E-8)	(1.7E-8)							
4-year	7.42*	0.02*	1.8E-6*	9.0E-8*							0.87
	(0.14)	(0.002)	(1.9E-7)	(2.6E-8)							
Response surface regression model											
11-year	4190.0	13.9	−0.001	−0.001	0.119	1.3E-9	2.3E-10	3.0E-3*	0.2E-3	4.1E-9	0.96
	(3672.1)	(191.7)	(0.01)	(0.03)	(3.96)	(5.0E-5)	6.1E-10	(0.00)	(0.00)	(2.7E-9)	
4-year	45013	−220.1	−0.15	0.006	−12.22	9.5E-8	−1.3E-9	0.002*	4.2E-6	2.8E-9	0.90
	(26713)	(537.7)	(0.09)	(0.01)	(9.27)	(7.8E-8)	(1.2E-9	(0.00)	(0.00)	(1.6E-8)	
ARMAX model											
11-year	−1461.2	106.0*	0.02*	3.0E-3*							
	(830.29)	(27.83)	(0.002)	(0.00)							
4-year	−12017*	305.0*	0.02*	0.001*							
	(2018.6)	(24.69)	(0.003)	(0.00)							

Standard errors are in the parentheses.
*Indicates statistical significance at the 0.05 level.

electricity consumption, according to priority, are NI (0.803), TEMP (0.381), and GDP (0.082). The results agree with those of Model 1–Model 4. The Adj-R^2 of Model 5 reaches 95.51%, and passes the normality test for residuals (p-val $= 0.1438$). The Durbin–Watson statistic is $d = 1.373$. To test positive autocorrelation at $\alpha = 0.05$, we use $d_{L,0.05} = 1.64$. Since $d < d_{L,0.05}$. We reject H_0. That is, the error terms are positively autocorrelated. The Adj-R^2 and DW statistics of Model 5 using 4-year training data set are 85.78% and 1.624, and coefficients are in Table 4. The forecast made by the log-linear regression should contain the variance term in that if y is log-normally distributed and $\ln(y)$ has mean m and variance $s2$, $E(y) = \exp(m + s2/2)$. Table 6 presents the forecasting results.

Because of the autocorrelations for residuals in Model 5, we use ARMA with input series, TEMP, NI and GDP to build electricity consumption model on the original value. Table 4 shows the coefficient estimations of Eq. (8) for 11-year and 4-year training data sets. Below is the error term equations for 11-year data set (figures in parentheses are standard errors):

$$\theta_q(B) = (1 + \underset{(0.09)}{0.255^*}B^3)(1 + \underset{(0.09)}{0.212^*}B^4)(1 + \underset{(0.10)}{0.226^*}B^{11}) \quad \text{and}$$

$$\phi_p(B) = 1 - \underset{(0.07)}{0.241^*}B - \underset{(0.07)}{0.712^*}B^{12}. \tag{20}$$

Using 4-year training data set, the error term equations are:

$$\theta_q(B) = 1, \quad \phi_p(B) = 1 - \underset{(0.16)}{0.296^{**}}B^3. \tag{21}$$

The * or ** on the coefficient indicates that it is significant at the 5% or 10% level. Eq. (8) with Eq. (20) or (21) gives the predicted values of electricity consumption.

We also build response surface regression models on the original values of electricity consumption. The estimation values of Eq. (9) are presented in Table 4 for RSREG model using 11-year and 4-year training data sets. Table 6 presents the forecasting results.

4.2. Building ANN models

We use 11-year (1990–2000) and 4-year (1997–2000) data sets as the training samples, while the 24 monthly data recorded 2001 through 2002 are used as testing samples. Because the ANN can catch sophisticated

Table 5
Sensitivities of electricity consumption to five variables

	132 historical data					48 historical data				
	TEMP	GDP	CPI	POP	NI	TEMP	GDP	CPI	POP	NI
Sensitivity	**3.068**	**0.989**	**−2.344**	**7.696**	**15.142**	**2.453**	**0.809**	**−1.831**	**4.098**	**6.816**

nonlinear integrating effects through a learning process, all of the explanatory variables, POP, NI, GDP, CPI and TEMP, are used as the network input values, and electricity consumption is the network output value. More than 60 experiments are performed to determine the best combination of the learning rate, momentum, number of hidden layers, number of hidden nodes, learning rule, and transfer function to utilize.

Throughout the training, the NeuralWare utility, 'SAVEBEST' is used to monitor and save the lowest root mean square (RMS) error from the testing set. The best RMS error result is obtained using a learning rate of 0.2, a momentum of 0.1, and 6 neurons in a single hidden layer that use the generalized delta learning rule and a sigmoid transfer function. The best architecture for 11-year data set contains 5 input layer neurons, 6 hidden neurons, and 1 output layer neuron (5:6:1 architecture). For 4-year data set, the best architecture is 5:5:1. The estimations of connection weights w_{ti} and v_t are obtained, then apply them to formula (16) to calculate the sensitivities $\hat{S}_i$. The absolute sensitivity value of each input variable represents its relative effect with electricity consumption. Crucial factors affecting electricity consumption are, by priority, NI, POP, TEMP, CPI and GDP. Their absolute sensitivities are 15.142, 7.696, 3.068, 2.344 and 0.989 for 11-year data set, 6.816, 4.098, 2.453, 1.831 and 0.809 for 4-year data set (see Table 5). The results agree with that of regression model on which the economy indicators, GDP and CPI, are not the key factors to Taiwan's electricity consumption, but POP and NI are.

Once the eight models to predict monthly electricity consumption are developed, we empirically examine the relative effectiveness of the models in predicting electricity consumption using the data from January 2001 to December 2002.

4.3. Out-of-sample forecasting performance results

We use three statistics, RMSE, MAE, MAPE, and scatter diagram to see the out-of-sample ability of the linear and nonlinear models. Panel A of Table 6 shows that the statistics of the ARMAX model (the optimal linear model) are 931.13%, 764.90% and 4.83% for an 11-year training data set while those of the ANN model decrease to 635.38%, 460.74%, and 3.19%. Panel B of Table 6 shows the statistics of the RSREG model (the optimal linear model) are 1295.43%, 1171.78% and 7.58% for a 4-year training data set while those of the ANN model fall sharply to 709.25%, 598.65% and 4.02%. It is because: (1) the linear model has been tested to make good forecasts for seasonal time series over a short-term period; however, it requires large amount of historical data; (2) the ANN is capable of catching sophisticated nonlinear integrating effect through a learning process. Fig. 6 presents both forecasting values (the ANN and linear models) and actual value of ELEC. It clearly shows that the ANN forecast value is closer to the ELEC actual value, and the linear model is weaker on foretelling peaks and bottoms. This result proves that the ANN is more appropriate to be applied to build an economic forecasting model for Taiwan's electricity consumption regardless of the amount of historical data.

5. Conclusion

By adopting the linear and nonlinear ANN methods, surprisingly, we find that economy indicators, GDP and CPI, have less effect on Taiwan's electricity consumption than POP and NI. Factors NI, TEMP, and GDP are contained in the superior linear economic forecasting model after excluding the multicollinearity. From the ANN model, we obtain the crucial factors, by priority, NI, POP, TEMP, CPI and GDP. By comparing scatter diagram and three statistics, RMSE, MAE, and MAPE for out-of-sample forecasting

Table 6
Comparing forecasting measurement errors

	Panel A 132 historical data			Panel B 48 historical data		
	RMSE	MAE	MAPE	RMSE	MAE	MAPE
LNREG	1508.96	1341.57	8.60%	1542.43	1376.26	8.84%
RSREG	1701.90	1489.72	9.51%	1295.43	1171.78	7.58%
ARMAX	931.13	764.90	4.83%	1566.34	1386.99	8.88%
ANN	**635.38**	**460.74**	**3.19%**	**709.25**	**598.65**	**4.02%**

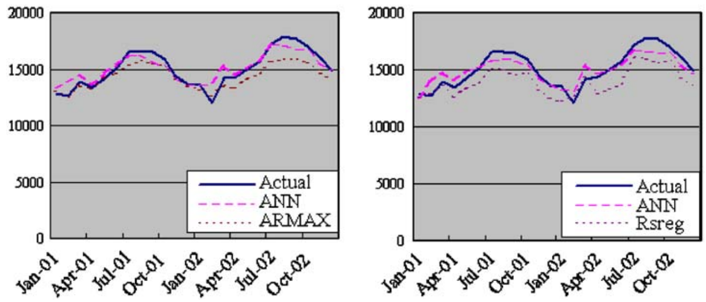

Fig. 6. Actual and forecasted values of Taiwan's electricity consumption ((a) 132 historical data, (b) 48 historical data).

ability of linear and the ANN models, the major findings included the following. (1) If given a large amount of historical data, the forecasts of ARMAX is better than the other linear models. (2) The linear model is weaker on foretelling peaks and bottoms regardless the amount of historical data. (3) The forecasting performance of ANN is higher than the other linear models based on two sets of historical data considered in the paper. This is probably due to the fact that the ANN model is capable of catching sophisticated nonlinear integrating effects through a learning process. To sum up, the important economic variables of the electricity consumption selected by both methods are consistent. Researchers can employ either ANN or linear model to extract them. Moreover, the ANN model is more appropriate between the two to help us build the economy-forecasting model of Taiwan's electricity consumption. Furthermore, it is possible to use the linear and nonlinear hybrid model and univariate time series to forecast energy consumption.

Acknowledgements

The author would like to thank three anonymous reviewers and the Editor of the Journal for comments and helpful suggestions, which have significantly improved the original version of this paper.

References

[1] Glasure YU, Lee AR. Cointegration, error-correction, and the relationship between GDP and electricity: the case of South Korea and Singapore. Resource Electricity Econ 1997;20:17–25.
[2] Shiu A, Lam PL. Electricity consumption and economic growth in China. Energy Policy 2004;32:47–54.
[3] Soytas U, Sari R. Energy consumption and GDP: causality relationship in G-7 countries and emerging markets. Energy Econ 2003;25:33–7.
[4] Yang HY. A note on the causal relationship between electricity and GDP in Taiwan. Energy Econ 2000;22:309–17.
[5] Yan YY. Climate and residential electricity consumption in Hong Kong. Energy 1998;23(1):17–20.

[6] Ranjan M, Jain VK. Modeling of electrical energy consumption in Delhi. Energy 1999;24:351–61.
[7] Egelioglu F, Mohamad AA, Guven H. Economic variables and electricity consumption in Northern Cyprus. Energy 2001;26:355–62.
[8] Darbellay GA, Slama M. Forecasting the short-term demand for electricity: do neural networks stand a better chance. Int J Forecasting 2000;16:71–83.
[9] Saab S, Badr E, Nasr G. Univariate modeling and forecasting of energy consumption: the case of electricity in Lebanon. Energy 2001;26:1–14.
[10] Fatai K, Oxley L, Scrimgeour FG. Modeling and forecasting the demand for electricity in New Zealand: a comparison of alternative approaches. Energy J 2003;24(1):75–102.
[11] Mallows CP. Some comments on C (p). Technometrics 1973;15:661–75.
[12] Hwang JN, Choi JJ, Oh S, Marks RJ. Query based learning applied to partially trained multilayer perceptron. IEEE T-NN 1991;2(1):131–6.
[13] Bishop CM. Neural networks for pattern recognition. Oxford: Oxford University Press; 1995.