Predictions of Industrial and Commercial Electricity Sales in Taiwan Using ARIMA and Artificial Neural Networks Techniques

Yuehjen E. Shao*, Yi-Shan Tsai
Department of Statistics and Information Science, Fu Jen Catholic University, New Taipei City, Taiwan, R.O.C.

* Corresponding author. Email: stat1003@mail.fju.edu.tw
Manuscript submitted April 2, 2017; accepted July 1, 2017.
doi: 10.17706/ijeeee.2018.8.2.74-81

Abstract: Electricity is one of the most important sources of energy on earth. Today, electricity has become a part of our life. Electricity is the key component to modern technology and without it most of the products that we use simply could not work. Without doubt, the economic growth for almost every country in the world is affected by electricity rates. Therefore, the prediction of electricity sales is very important for Taiwanese economy. This study employs the autoregressive integrated moving average (ARIMA), artificial neural networks (ANN) and the integrated ARIMA-ANN approaches for predicting the industrial electricity and commercial electricity sales (IECES) in Taiwan. The forecasting accuracy measure is based on the mean absolute percentage error. The real dataset, from the years 2006 to 2016, for IECES in Taiwan are collected and analyzed. The prediction results show that the ARIMA-ANN model has the most satisfactory forecasting accuracy for predictions of IECES in Taiwan.

Key words: Prediction, electricity sales, ARIMA, artificial neural networks.

1. Introduction

The 'Energy Development Group' (EDG) in Taiwan was initially established to formulate the energy–related regulations in 1968. One of the major functions of the Bureau of Energy is to forecast or evaluate and plan the energy supply and demand. In addition, the Taiwan Power Company (Taipower), a national electrical power utility company, was established in 1946. The scope of Taipower includes generation, distribution and sales of electricity. In 2014, the electricity sales accounted for 98.4% of Taipower's revenue. Taipower's energy sales reached 205,956 Gigawatt hour (GWh) in 2014, 2% up from 2013. The electricity sector has been a powerful engine and provides competitively priced electricity that is critical for Taiwan's productivity.

The economic growth in a country is indeed affected by stable electricity sales with reasonable rates. Accordingly, the prediction of electricity sales is very important for a rich economy. In contrast to the multiple regression technique [1]-[7], this study employs two types of forecasting techniques to predict the industrial electricity (IE) and commercial electricity (CE) sales in Taiwan. The first technique is the single-stage forecasting method, which includes the autoregressive integrated moving average (ARIMA) and the artificial neural networks (ANN). ARIMA modeling is appropriate because the seasonal effects would be involved in forecasting industrial electricity and commercial electricity sales (IECES) in Taiwan. ARIMA is a well-known forecasting technique, and it consists of different types of time series process, such as pure autoregressive (AR), pure moving average (MA) and the combined ARMA. Nevertheless, the linear form of ARIMA models may encounter difficulties in capturing the nonlinear pattern of the real data [8]-[10]. The ANN technique contains fewer a priori assumptions. Therefore, ANN can be served as another alternative modeling scheme for predicting IECES because it allows to model nonlinearity and provides good forecasting characteristics [9]-[12]. Nevertheless, ANN is criticized for its long training process in designing the optimal structure [13]-[16].

The second technique is the two-stage modeling scheme. The fundamental concept an integrated forecasting scheme is to capture various patterns in the data by taking advantage of each individual model's capability. The findings have been reported that the two-stage modeling is superior for improving the performance of single-stage modeling [17]-[23]. Since both ARIMA and ANN are appropriate for predicting the IECES, this study considers an integrated ARIMA-ANN as the two-stage model.

The real dataset of IECES in Taiwan was sampled from January of 2006 to July of 2016. This study uses the first 115 data records to build the forecasting models, and performs a confirmation using the last 12 data records. The structure of this study is described as follows. Section 2 addresses the forecasting methodologies in this study. Section 3 discusses the development and design of the single and two-stage models. Real data of IECES are fitted to obtain the forecasting models, and the predictions of IECES in Taiwan are produced by using the single-stage and the two-stage models. The final section presents the research findings and conclusions inferred from this study.

2. The Methodologies

2.1. Autoregressive Integrated Moving Average

A general SARIMA (p, d, q, P, D, Q) model can be described as the following:

$$\phi_p(B)\Phi_P\left(B^L\right)\left(1-B^d\right)\left(1-B^L\right)^D Z_t = \delta + \theta_q(B)\Theta_Q\left(B^L\right)a_t, \tag{1}$$

where

δ: an unknown constant,

Z_t: the working series values as a function of time t, which are stationary after fitting a suitable transformation from the original time series Y,

d, D: the values of non-seasonal and seasonal transformations, respectively,

a_t: white noise at time t, which are independent and identical (iid) with normal distribution,

p,P,q,Q: the order (parameters) of autoregressive (AR) and moving average (MA) models, respectively,

B: the backward shift operator, defined as: $B^j a_t = a_{t-j}$,

L: the number of months in a year, and *L* = 12 for the monthly data,

$\emptyset_p(B)$: A polynomial function for a non-seasonal AR model, defined as:

$$\phi_p(B) = (1-\phi_1 B - \phi_2 B^2 - \cdots\cdots - \phi_p B^p),$$

$\emptyset_p(B^L)$: A polynomial function for a seasonal AR model, defined as:

$$\Phi_P\left(B^L\right) = \left(1-\Phi_{1,L}B - \Phi_{2,L}B^{2L} - \cdots\cdots - \Phi_{P,L}B^{PL}\right),$$

$\theta_q(B)$: A polynomial function for a non-seasonal MA model, defined as:

$$\theta_q(B) = (1-\theta_1 B - \theta_2 B^2 - \cdots\cdots - \theta_q B^q),$$

$\theta_q(B^L)$: A polynomial function for a seasonal MA model, defined as:

$$\Theta_Q\left(B^L\right) = \left(1-\Theta_{1,L}B - \Theta_{2,L}B^{2L} - \cdots\cdots - \Theta_{Q,L}B^{QL}\right)$$

The original nonstationary time series may be transformed into a stationary working series by taking differencing. The transformation is usually performed using four combinations of d and D; that is, (d, D) = (0, 0), (d, D) = (1, 0), (d, D) = (0, 1) and (d, D) = (1, 1). Once the stationary working series have been obtained, we can apply the sample autocorrelation function (ACF) and sample partial autocorrelation function (PACF) to determine the order of p, P, q, and Q for the seasonal ARIMA models. Usually, after performing a diagnostic checking, Ljung-Box test [24], for the parameters and residuals, one is able to obtain the forecasting models. In this study, the prediction capability of the models is compared using the MAPE criteria. This measurement is described as follows.

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|e_t|}{Y_t}\times 100 \tag{2}$$

where e_t stands for the value of the residual at time t.

2.2. Artificial Neural Networks

ANN is a parallel system comprised of highly interconnected processing elements that are based on neurobiological models. ANN processes information through the interactions of a large number of simple processing elements, the neurons. ANN modeling can be described briefly as follows. The relationship between output (y) and inputs $(x_1, x_2, \dots, x_a)$ in an ANN model can be formed as:

$$y = \alpha_0 + \sum_{j=1}^{b}\alpha_j g(\delta_{0_j} + \sum_{i=1}^{a}\delta_{ij}x_i) + \varepsilon, \tag{3}$$

where α_j $(j = 0, 1, 2, \dots, \mathrm{b})$ and δ_{ij} $(i = 0, 1, 2, \dots, a\,; j = 0, 1, 2, \dots, b)$ are model connection weights; a is the number of input nodes; b is the number of hidden nodes, and is the error term. The transfer function in the hidden layer is often represented by a logistic function,

$$g(z) = \frac{1}{1+\exp(-z)}. \tag{4}$$

Accordingly, the ANN in Equation (3) accomplishes a nonlinear functional mapping from the inputs $(x_1, x_2, \dots, x_a)$ to the output y,

$$y = f(x_1, x_2, \dots, x_a, w) + \varepsilon, \tag{5}$$

where w is a vector of all model parameters, and f is a function determined by the ANN structure and connection weights.

3. Results and Discussion

3.1. ARIMA Modeling Results

This study divides the IECES data into two sets for ARIMA procedure. The first set has 115 records which are used for the model building, and the second set contains 12 records which are used for the model confirmation. Figures 1 and 2 demonstrate the original time plot for the IE and CE sales in Taiwan, respectively. After performing the ARIMA modeling using SAS package, we have the parameter estimates for IE and CE sales which are listed in Tables 1 and 2. Therefore, this study employs the Equations (6) and (7) to predict the IE sales and CE sales in Taiwan.

$$\mathrm{Z}_t = -0.947Z_{t-1} - 0.502Z_{t-2} + a_t - 0.628a_{t-12}, \tag{6}$$

where

$$\mathrm{Z}_t = (y_t - y_{t-1}) - (y_{t-12} - y_{t-13}),$$

$$\mathrm{Z}_t = -0.686Z_{t-1} - 0.465Z_{t-2} + a_t - 0.738a_{t-12}, \tag{7}$$

where

$$\mathrm{Z}_t = (y_t - y_{t-1}) - (y_{t-12} - y_{t-13}),$$

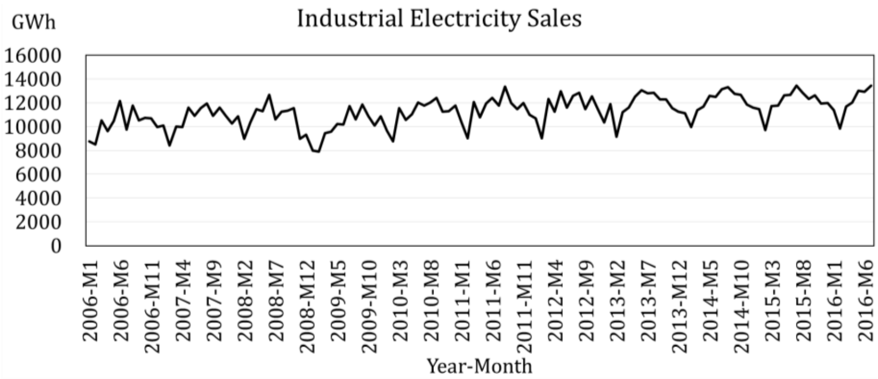

Fig. 1. Time plot for IE sales in Taiwan (unit: GWh).

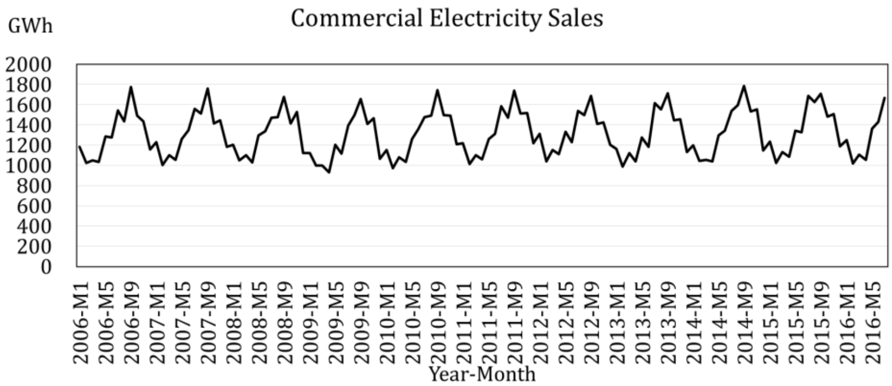

Fig. 2. Time plot for CE sales in Taiwan (unit: GWh).

Table 1. Parameter Estimates for IE Sales

The ARIMA Procedure

Conditional Least Squares Estimation

Parameter	Estimate	Standard error	t value	Approx Pr > |t|	Lag
MA1,1	0.62838	0.08064	7.79	<.0001	12
AR1,1	-0.94744	0.08769	-10.80	<.0001	1
AR1,2	-0.50233	0.08732	-5.75	<.0001	2

Table 2. Parameter Estimates for CE Sales

The ARIMA Procedure

Conditional Least Squares Estimation

Parameter	Estimate	Standard error	t value	Approx Pr > |t|	Lag
MA1,1	0.73818	0.07617	9.69	<.0001	12
AR1,1	-0.68591	0.09043	-7.58	<.0001	1
AR1,2	-0.46463	0.09121	-5.09	<.0001	2

3.2. ANN Modeling Results

Since more than 75% of ANN applications use the back-propagation neural network (BPNN), this study employs the BPNN in developing the ANN forecasting model [25]-[28]. In addition, since there is only one variable (i.e., y_t, or IECES) available, this study uses two designs for ANN modeling. The first design considers the past 2 observations at time t-1 and t-2 (y_{t-1} and y_{t-2}) to predict the observation at current time t (y_t). This design uses two variable, y_{t-1} and y_{t-2}, to serve as the inputs and it uses a single variable, y_t, to serve as the output for ANN modeling. Accordingly, the first design employs 2 input nodes and one output node for the ANN structure. Also, since the seasonal effects may be involved in the IECES predictions, the second ANN design of this study uses past 12 observations to predict the observation at time t. That is, the second design uses 12 variables, y_{t-1}, y_{t-2}, y_{t-3}, ..., and y_{t-12} to serve as the inputs and it uses a single variable, y_t, to serve as the output for ANN modeling. Therefore, the second design employs 12 input nodes and one output node for the ANN structure.

For the ANN developed herein, this study utilizes 2 and 12 input nodes for first and second designs, respectively. The hidden nodes was set to $2n\pm2$, where n is the number of input variables. Thus, the hidden nodes were chosen as 2, 3, 4, 5 and 6 for the first design. The hidden nodes were chosen as 22, 23, 24, 25 and 26 for the second design. Since the learning rate of 0.01 is a very effective setting [22], [28], this study sets the values of learning setting as 0.01 for ANN modeling. After performing the ANN modeling, the first design obtained that the {2-6-1} and {2-2-1} structures provided the best results and a minimum testing MAPE for IE and CE sales, respectively. In this study, {ni-nh-no} represents the number of neurons in the input layer, number of neurons in the hidden layer and number of neurons in the output layer, respectively. For the second design, we have the best structures of {12-25-1} and {12-25-1} for IE and CE sales, respectively. In addition, Tables 3 and 4 lists the corresponding MAPE values for different setting of the ANN topologies for the two designs.

Table 3. The First Design ANN Topology for IECES

ANN topology IE Sales	MAPE	ANN topology CE Sales	MAPE
{2-2-1}	5.97	{2-2-1}	10.67
{2-3-1}	5.97	{2-3-1}	10.71
{2-4-1}	5.97	{2-4-1}	10.72
{2-5-1}	5.97	{2-5-1}	10.74
{2-6-1}	5.96	{2-6-1}	10.73

Table 4. The Second Design ANN Topology for IECES

ANN topology IE Sales	MAPE	ANN topology CE Sales	MAPE
{12-22-1}	2.92	{12-22-1}	2.22
{12-23-1}	2.94	{12-23-1}	2.15
{12-24-1}	2.89	{12-24-1}	2.25
{12-25-1}	2.87	{12-25-1}	2.14
{12-26-1}	2.91	{12-26-1}	2.20

3.3. ARIMA-ANN Modeling Results

Since an SARIMA (2, 1, 0, 0, 1, 1) model was developed for representing process of IE and CE sales, this study uses the past 3 observations at time t-1, t-2 and t-12 (y_{t-1}, y_{t-2} and y_{t-12}) to predict the observation at current time t (y_t). That is, the ARIMA-ANN design employs 3 input nodes and one output node for the ANN structure. The hidden nodes were chosen as 4, 5, 6, 7 and 8 for this design. Table 5 reports the corresponding MAPE values for different setting of the ANN topologies. For the combined ARIMA-ANN design, we have the best structures of {3-4-1} and {3-7-1} for IE and CE sales, respectively.

3.4. The Results and Analysis

The well-known MAPE measure is applied to evaluate the performance of the forecasting models. A low value of MAPE is associated with better forecasting accuracy. This study considers various forecasting models for predictions IECES in Taiwan. Table 6 displays the MAPE for different forecasting models.

Table 5. The ANN Topology for The ARIMA-ANN Model

ANN topology (IE Sales)	MAPE	ANN topology (CE Sales)	MAPE
{3-4-1}	2.21	{3-4-1}	2.72
{3-5-1}	2.37	{3-5-1}	2.66
{3-6-1}	2.30	{3-6-1}	2.64
{3-7-1}	2.32	{3-7-1}	2.63
{3-8-1}	2.37	{3-8-1}	2.67

As shown in Table 4, the ARIMA and ANN (second design) models have minimum MAPEs for IE and CE sales, respectively. However, one can notice that the combined ARIMA-ANN has the best averaged performance for both IE and CE sales in Taiwan. In addition, based on the results in Table Ⅵ, we notice that the MAPE percentage improvements of the proposed ARIMA-ANN model over the ANN-first design model are 62.92% and 75.35%, respectively.

4. Conclusion

This study applies four forecasting modeling schemes to the prediction of IECES Taiwan. The forecasting results are compared and discussed. The results reveal that although the forecasting performances of these four models are all acceptable, the proposed ARIMA-ANN model outperforms the other three forecasting schemes.

The modeling process of this study could provide the guidelines for developing the forecasting models for other energies. In addition, some other forecasting schemes, such as support vector regression (SVR) or extreme learning machine (ELM), can be further investigated to predict the demand or consumption for different types of electricity in the future.

Table 6. MAPE for Various Forecasting Models

Models	IE Sales	CE Sales	Average MAPE
ANN (first design)	5.96	10.67	8.32
ANN (second design)	2.87	2.14	2.51
ARIMA	1.72	5.39	3.56
ARIMA-ANN	2.21	2.63	2.42

References

[1] Sadownik, R., & Barbosa, E. P. (1999). Short-term forecasting of industrial electricity consumption in Brazil. *Journal of Forecasting*, *18(3)*, 215-224.

[2] Harris, J. L., & Liu, L. M. (1993). Dynamic structural analysis and forecasting of residential electricity consumption. *International Journal of Forecasting, 9(4)*, 437-455.

[3] Cebula, R. J., & Herder, N. (2010). Recent evidence on residential electricity consumption determinants: a panel two-stage least squares analysis, 2002-2005. *Research in Business and Economics Journal, 2*, 1-8.

[4] Ranjan, M., & Jain, V. K. (1999). Modelling of electrical energy consumption in Delhi. *Energy, 24(4)*, 351-361.

[5] Holtedahl, P., & Joutz, F. L. (2004). Residential electricity demand in Taiwan. *Energy economics, 26(2)*, 201-224.

[6] Yan, Y. Y. (1998). Climate and residential electricity consumption in Hong Kong. *Energy, 23(1)*, 17-20.

[7] Shao, Y. E., Hou, C. D., & Lin, C. S. (2015). Applying artificial neural networks, multiple regression and their hybrid models for predictions of electricity sales in Taiwan. *ICIC express letters. Part B, Applications: An International Journal of Research and Surveys, 6(2)*, 485-490.

[8] Pai, P. F., & Lin, C. S. (2005). A hybrid ARIMA and support vector machines model in stock price forecasting. *Omega, 33(6)*, 497-505.

[9] Khashei, M., & Bijari, M. (2010). An artificial neural network (p, d, q) model for timeseries forecasting. *Expert Systems with Applications, 37(1)*, 479-489.

[10] Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing, 50*, 159-175.

[11] Muniz, A. M. S., Liu, H., Lyons, K. E., Pahwa, R., Liu, W., Nobre, F. F., & Nadal, J. (2010). Comparison among probabilistic neural network, support vector machine and logistic regression for evaluating the effect of subthalamic stimulation in Parkinson disease on ground reaction force during gait. *Journal of biomechanics, 43(4)*, 720-726.

[12] Caselli, M., Trizio, L., Gennaro, G., & Ielpo, P. (2009). A simple feedforward neural network for the PM10 forecasting: comparison with a radial basis function network and a multivariate linear regression model. *Water, Air, and Soil Pollution, 201(1-4)*, 365-377.

[13] Lee, T. S., Chiu, C. C., Lu, C. J., Chen, & I. F. (2002). Credit scoring using the hybrid neural discriminant technique. *Expert Systems with Applications, 23(3)*, 245-254.

[14] Chou, S. M., Lee, T. S., Shao, Y. E., & Chen, I. F. (2004). Mining the breast cancer pattern using artificial neural networks and multivariate adaptive regression splines. *Expert Systems with Applications, 27(1)*, 133-142.

[15] Dai, W., Shao, Y. E., & Lu, C. J. (2013). Incorporating feature selection method into support vector regression for stock index forecasting. *Neural Computing and Applications, 23(6)*, 1551-1561.

[16] Craven, M. W., & Shavlik, J. W. (1997). Using neural networks for data mining. *Future Generation Computer Systems, 13(2-3)*, 211-229.

[17] Co, H. C., & Boosarawongse, R. (2007). Forecasting Thailand's rice export: Statistical techniques vs. artificial neural networks. *Computers & Industrial Engineering, 53(4)*, 610-627.

[18] Shao, Y. E., Lu, C. J., & Wang, Y. C. (2012). A hybrid ICA-SVM approach for determining the quality variables at fault in a multivariate process. *Mathematical Problems in Engineering, 2012*, 1-12.

[19] Zhang, G. P. (2007). A neural network ensemble method with jittered training data for time series forecasting. *Information Sciences, 177(23)*, 5329-5346.

[20] Lu, C. J., Shao, Y. E., Li, & P. H. (2011). Mixture control chart patterns recognition using independent component analysis and support vector machine. *Neurocomputing, 74(11)*, 1908-1914.

[21] Hou, C. D., & Shao, Y. E. (2015) Integrated use of statistical-based approaches and machine learning techniques for tumors classification uing microarray data. *Discrete Dynamics in Nature and Society*, *2015*, 1-8.

[22] Shao, Y. E., & Hou, C. D. (2013). Change point determination for a multivariate process using a two-stage hybrid scheme. *Applied Soft Computing, 13(3)*, 1520-1527.

[23] Lu, C. J., & Shao, Y. E. (2012). Forecasting computer products sales by integrating ensemble empirical mode decomposition and extreme learning machine. *Mathematical Problems in Engineering, 2012*, 1-12.

[24] Ljung, G. M., & Box, G. E. (1978). On a measure of lack of fit in time series models. *Biometrika, 297-303*.

[25] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation in parallel distributed processing. *MIT Press Cambridge*, MA, USA.

[26] Chiu, C. C., Shao, Y. E., Lee, T. S., & Lee, K. M. (2003). Identification of process disturbance using SPC/EPC and neural networks. *Journal of Intelligent Manufacturing, 14(3)*, 379-388.

[27] Shao, Y. E., & Hsu, B. S. (2009). Determining the contributors for a multivariate SPC chart signal using artificial neural networks and support vector machine. *International Journal of Innovative Computing, Information and Control, 5(12)*, 4899-4906.

[28] Shao, Y. E., Lu, C. J., & Chiu, C. C. (2011). A fault detection system for an autocorrelated process using SPC/EPC/ANN AND SPC/EPC/SVM schemes. *International Journal of Innovative Computing, Information and Control, 7(9)*, 5417-5428.

Yuehjen E. Shao is a professor in the Department of Statistics and Information Science at Fu Jen Catholic University, New Taipei City, Taiwan. He holds a Ph.D. in decision sciences and engineering systems from Rensselaer Polytechnic Institute, New York, USA. His present research interests include statistical process control, data pattern recognition and classification, forecasting methods and data mining.

Yi-Shan Tsai is a graduate student in the Department of Statistics and Information Science at Fu Jen Catholic University, New Taipei City, Taiwan. Her present research interests include data mining and forecasting methods.