Application of the Forecasting Technique by hybrid model for Forecasting the Electricity Demands of Rajabhat Universities

Thanakon Sutthison

Program of Applied Statistics, Faculty of Science, Ubon Ratchathani Rajabhat University, Ubon Ratchathani, 34000
* Corresponding author. E-mail address: Thanakon.sutthison@gmail.com

Received: 4 January 2018; Accepted: 5 April 2018

Abstract

Electricity consumption at the education institutions studied is expected to increase which means that a more efficient and relevant model for managing electricity consumption is needed. To achieve this, statistical forecasting technique was applied. An accurate forecast of consumption would be great help for executives and for those concerning to establish a power-saving policy in the universities. The objective of this research was to forecast electricity consumption by using hybrid models (mixing the Box-Jenkins modeling and a support vector regression model). Three Rajabhat Universities: Nakorn Ratchatsima, Ubon Ratchathani and Loei, were included in the modeling. Time series data on the energy consumption on a monthly basis, covering the period from January, 2006 to September, 2017, were available on the website of the Energy Ministry. Data for the period of January, 2006 to December, 2016 were used for our research purposes, and R-language based forecast model that we developed was used for analysis. The most suitable model was selected by using the Mean Absolute Percent Error (MAPE). The most suitable model obtained from the data series from January to September, 2017 was used to measure the accuracy of the forecast produced. Our results indicated that the proposed model was more efficient, with greater accuracy, in forecasting the energy consumption than the conventional model currently used in the universities participating in the study. For these universities, the MAPE was 7.65428 for Nakorn Ratchatsima Rajabhat University, 6.35679 for Ubon Ratchathani and 4.13581 for Loei Rajabhat Universities.

Keywords: Electricity Consumption, Hybrid models, Box-Jenkins Model, Support Vector Regression Model

Introduction

At present, time series forecasting is very important since it allows us to know what will happen in advance. It also helps in decision making about upcoming events efficiently. Statistical forecasting methods are used to assist in decision making on various tasks. For example, Zhang et al. (2016) accurately forecasted building energy consumption by using weighted support vector. Chen et al. (2017) forecasted the short-term power consumption of office by using support vector regression which was suitable for forecasting the short-term power consumption. Theeraviriya (2017) compared the forecasting methods for electric energy demand in Nakhon Phanom province and found that the decomposition method was suitable for forecasting electric energy demand in Nakhon Phanom province.

From the literature review, time series analysis was divided into two groups. (1) Linear approach: the approach widely used was Box-Jenkins method using ARIMA and SARIMA models. However, it cannot be used to analyze non-linear time series data (Junsagoon, 2015). (2) Non-linear approach: the approaches mostly used were Artificial Neural Network and Support vector regression model (SVR) which were highly flexible and can be used in a variety of highly complex data. They can also predict both short and long term future, but they cannot analyze the seasonal data and trends in linear time series. Since both linear and nonlinear approaches contained different disadvantages, they were mixed together called the Hybrid Models

(Zhang, 2003) in order to adjust the forecasting method and to solve the problems of the linear and non-linear forecasting approaches. This makes the forecast more accurate. Several research studies tried to develop hybrid models to forecast time series data and it was found that hybrid models provided higher accuracy than single forecasting models.

Based on the aforementioned reasons, to increase the accuracy of the forecast in this paper, a total of 5 hybrid models mixing linear and non-linear approaches together were proposed, and the most suitable model was chosen for future forecast. Mean Absolute Percentage Error was employed to examine the proposed models. The researcher applied the time series data with electricity consumption of Nakhon Ratchasima Rajabhat University, Ubon Ratchathani Rajabhat University and Loei Rajabhat University. The research results were expected to be beneficial for the establishment of guidelines and measures for saving energy in the related universities.

Methods and materials

1. Data

The time series data on electricity consumption on a monthly basis of three Rajabhat Universities (kWh / hour) covering the period from January, 2006 to September, 2017, from the website of the Energy Policy and Planning Office, Ministry of Energy (2017): 153 values were collected. The data were divided into 2 sets. The first set was the data from January 2005 to December 2016: 144 values which were used for the study with 6 time series forecasting models. The R-Language program was used for data analysis since it was a downloadable and installable application which was widely used. The appropriate forecasting methods were selected by comparing the forecasting errors of MAE, MAPE as they had been applied in various forecasting methods. For example, Chen et al. (2017) used MAE, MAPE to compare the accuracy in forecasting. The forecasting method with the lowest MAE, MAPE was considered the most accurate forecasting method. The second set was the data from January to September 2017: 9 values which were used to compare the accuracy of the forecast of electricity consumption of each university.

2. Concepts and theories

2.1 Box -Jenkins Method: It is a highly accurate forecasting method because the time series are considered how they are correlated. To create a suitable forecasting model, data forecasting is based on two main types: Autoregressive (AR) which indicates that forecast value at any time is based on previous observation value; and Moving average (MA) which determines that forecast value at any time depends on previous error. The important condition of the use of the Box -Jenkins Method is that the forecasted time series must be stationary time series. However, in general the collected time series are non-stationary time series as it moves according to trends and seasons. SARIMA (p,d,q) $(P,D,Q)_s$ is commonly used (Box et al., 2008) as shown in Equation (1) - (6).

$$\phi_p(B)\Phi_P(B^S)(1-B)^d(1-B^S)^D Z_t = \delta + \theta_q(B)\Theta_Q(B^S)\varepsilon_t \quad (1)$$

$$\delta = \mu\phi_p(B)\Phi_P(B^S) \quad (2)$$

$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p \quad (3)$$

$$\Phi_P(B^s) = 1 - \Phi_s B - \Phi_{2s} B^{2s} - \ldots - \Phi_{Ps} B^{Ps} \quad (4)$$

$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q \quad (5)$$

$$\Theta_Q(B^s) = 1 - \Theta_s B - \Theta_{2s} B^{2s} - \ldots - \Theta_{Ps} B^{Qs} \quad (6)$$

Z_t is time series at time t and $\varepsilon_t \sim NID(0,\sigma^2)$ and t is the time ranging from 1 to n. S is the time unit in 1 season, d and D are the orders of differences and seasonal differences. B is backward operator. $B^s Z_t = Z_{t-s}$ and $BZ_t = Z_{t-1}$, $\phi_p(B)$ is non-seasonal autoregressive operator of order p. $\Phi_P(B^s)$ is seasonal autoregressive operator of order P. $\theta_q(B)$ is non-seasonal moving average operator of order q. $\Theta_Q(B^s)$ is seasonal moving average operator of order Q. Setting up SARIMA requires time series analysis to determine SARIMA (p,d,q)(P,D,Q)$_s$. The researcher used R-language program with auto.arima function that is used to specify appropriate ARIMA (p, d, q) (P,D,Q)$_s$ to write command. The process of setting up model included 1) considering if it is stationary time series or not; 2) determining the forecasting model, 3) estimating the parameters and checking the suitability of the model and 4) forecasting the electricity consumption.

2.2 Support vector regression model : SVR

Support vector regression model is a popular forecast method because it is highly accurate and works well with nonlinear time series data. Its principle is reducing structural risk to the minimum value by low dimension dataset on the input space into high dimension dataset on the feature space. The function for modifying the data format is called Kernel function that input data is in the form of a vector $\{(x_1,y_1), (x_1,y_1),\ldots,(x_n,y_n)\}$, when $x_i \in R^n$, x_i represents the input vector and y_i represents the real value of the output vector. Thus, the function of SVR can be written as Equation (7).

$$y = f(x) = \sum_{i=1}^{N} w_i \phi_i(x) + b \quad (7)$$

$\phi(x)$ is feature space (transformed through nonlinear functions from the input space). In order to create the optimal plane, appropriate w and b parameters are needed (w is the weight of the vector and b is the error value) calculated from the minimum error in the equation (8).

$$R(C) = \frac{1}{2}\|w\|^2 + C.\frac{1}{n}\sum_{i=1}^{N}\left|y_i - f(x)\right|_\varepsilon \quad (8)$$

Using SVR technique predicts output from input vector, epsilon tube is created by using various loss functions such as Quadratic, Laplacian, Huber, and ε -insensitive. ε -insensitive loss function proposed by Vapnik (1988) is commonly and widely used in research. For example, Yashlan and Bican (2017) applied ε -insensitive loss function to forecast the electricity load of China. ε -insensitive loss function can be expressed as Equation (9).

$$\left|y_i - f(x)\right|_\varepsilon = \begin{cases} 0 & \left|y_i - f(x)\right| \le \varepsilon \\ \left|y_i - f(x)\right| - \varepsilon & otherwise \end{cases} \quad (9)$$

ξ_i and ξ_i^* are Slack variables and stay away from ε -tube as shown in Figure 1.

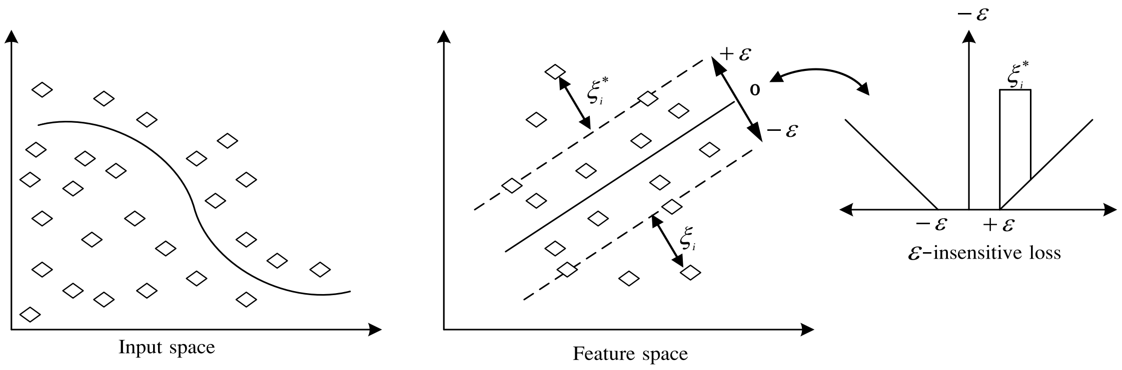


Figure 1 Support vector regression model using ε – insensitive loss function

Adding ξ_i and ξ_i^* is the measurement of all deviations of the data set training outside the edge of the plane or the error of the data from the upper and lower planes. Thus, adding ξ_i and ξ_i^* makes Equation 8 converted to the constrained as in Equations 10 and 11.

$$\text{Minimize} = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*) \tag{10}$$

$$\text{Subject to} \begin{cases} y_i - w\phi_i - b \le \xi_i + \varepsilon \\ -y_i + w\phi_i + b \le \xi_i^* + \varepsilon \\ \xi_i, \xi_i^* \ge 0 \end{cases} \tag{11}$$

Solving Equation 10 with the constrained of Equation 11 can be done by converting to Dual Problem with the Lagrange multipliers as shown in Equations 12 and 13.

$$\begin{aligned} \textit{Maximize}\ (\alpha_i, \alpha_i^*) = &\sum_{i=1}^{N} y_i(\alpha_i + \alpha_j^*) - \varepsilon\sum_{i=1}^{N}(\alpha_i + \alpha_i^*) \\ &- \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}(\alpha_i + \alpha_i^*)(\alpha_j + \alpha_j^*)k(x_i, x_j) \end{aligned} \tag{12}$$

$$\textit{Subject to} \begin{cases} \sum_{i=1}^{N}(\alpha_i + \alpha_i^*) = 0 \\ 0 \le \alpha_i \le C, \quad i = 1,\dots, N \\ 0 \le \alpha_i^* \le C, \ i = 1,\dots, N \end{cases} \tag{13}$$

When α_i, α_i^* are Lagrange multipliers. C is constant. N the number of support vector which the input is support vector: $\alpha_i, \alpha_i^* > 0$. The input is support vector which is not support vector: $\alpha_i, \alpha_i^* = 0$. After calculating the values of α_i and α_i^* from the first data set, SVR equation can be created to predict the output from the input vector as shown in Equation 14.

$$f(x) = \sum_{i=1}^{N}(\alpha_i + \alpha_i^*)k(x, x_i) + b \tag{14}$$

$K(x_i, x_j)$ is Kernel function used to solve nonlinear regression starting from transforming the first data set in input space (nonlinear data) to feature space. Linear regression is analyzed in feature space. Kernel's value is equal to the inner product of the two vectors: x_i and x_j in feature space $\phi(x_i)$ and $\phi(x_j)$ which is $K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$. Kernel functions commonly used in the SVR are Linear, Polynomial, Gaussian kernel and radial basis Function (Chen et al., 2017, Sujjaviriyasup, 2017, Yaslan and Bican, 2017). Based on the constraint conditions of the dual problem with the transformation of the data set training using Kernel function, new weight vector equation is obtained as shown in Equation 15.

$$W = \sum_{i=1}^{N} (\alpha_i + \alpha_i^*) \phi(x_i) \tag{15}$$

When replacing w in Equaion 7, a new regression function is obtained as in Equation 8. For the error of b, KKT (Karush - Kuhn - Tucker) is used to adjust the value in the middle of the edge of the plane as shown in Ewuaiton 16 when x_r and x_s are the vectors on the upper and lower planes, respectively.

$$b = -\frac{1}{2} \sum_{i=1}^{N} (\alpha_i + \alpha_i^*)(K(x_i, x_r) + K(x_i, x_s)) \tag{16}$$

2.3 Hybrid models

SARIMA and SVR are the forecasting methods suitable for different time series forecasting. SARIMA is suitable for analyzing linear time series while SVR is a model that works well with nonlinear time series data. So if only one of the forecasting methods is selected, it will make the forecast inaccurate. Zhang (2003), therefore, proposed a hybrid model that linear and non-linear forecasting values are combined. The forecasting value is more accurate as shown in Equation 17.

$$y_t = L_t + N_t + \varepsilon_t \tag{17}$$

y_t represents the observed time series data at time t. L_t represents the linear data of the parameter term at time t. N_t is the nonlinear data of the parameter term at time t. ε_t represents the error at time t calculated from SARIMA as in Equation 18.

$$\varepsilon_t = y_t - \hat{L}_t \tag{18}$$

$\hat{L}_t$ is the forecasting value of a linear function obtained from SARIMA at time t. When the rest of Equation 18 is obtained, it is used for forecasting the data using SVR method. In this study, five models were presented as follows.

Model 1: SARIMASVR1

$$\varepsilon_t = f(\varepsilon_{t-1}, \varepsilon_{t-12}) + e_t \tag{19}$$

When f is a nonlinear function, derived from SVR and e_t is the random error, the total forecasting value of Model 1 is:

$$\hat{y}_t = \hat{L}_t + \hat{N}_t \tag{20}$$

$\hat{N}_t$ is forecasting value derived from Support Vector Regression.

Model 2: SARIMASVR2

$$\hat{y}_t = f(y_{t-1}, y_{t-12}, \varepsilon_t) + e_t \tag{21}$$

Model 3: SARIMASVR3

$$\hat{y}_t = f(y_{t-1}, y_{t-12}, \hat{L}_t) + e_t \tag{22}$$

Model 4: SARIMASVR4

$$\hat{y}_t = f(y_{t-1}, y_{t-12}, \hat{L}_{t-1}, \hat{L}_{t-12}) + e_t \tag{23}$$

Model 5: SARIMASVR5

$$\hat{y}_t = f(y_{t-1}, y_{t-12}, \hat{L}_{t-1}, \hat{L}_{t-12}, \varepsilon_t) + e_t \tag{24}$$

2.4 Measuring Forecast Errors

(1) Mean Absolute Error: MAE

$$MAE = \sum_{t=1}^{n} \left| y_t - \hat{y}_t \right| / n \tag{25}$$

(2) Mean Absolute Percent Error: MAPE

$$MAPE = 100 \sum_{t=1}^{n} \left| 1 - \hat{y}_t / y_t \right| / n \tag{26}$$

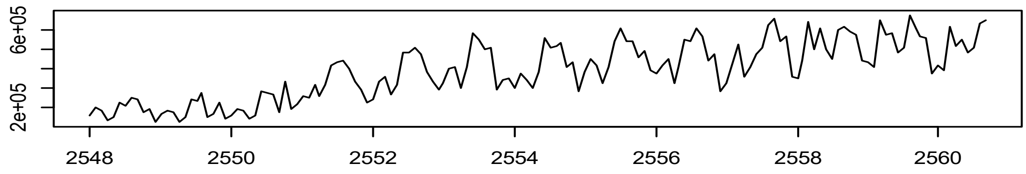
The less MAE and MAPE, the more accurate the forecast is.

2.5 Suitable forecasting method

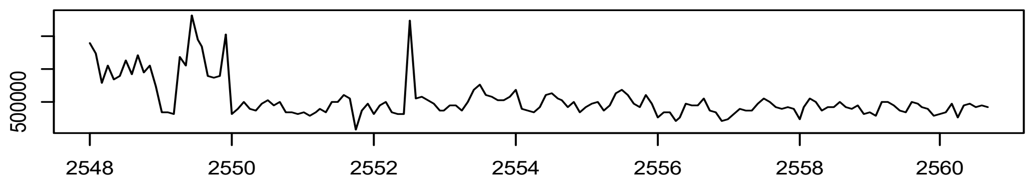
Based on a comparison of the errors of the three forecasting methods, that with the lowest MAE and MAPE was considered the most accurate forecasting method. The forecasting method then was used to forecast the energy consumption of the three Rajabhat Universities and it was compared with the second set of data. The accuracy of the forecast was considered from MAPE. When the appropriate model was obtained, the researcher made 15-month forecasts, as shown in Table 8.

Results

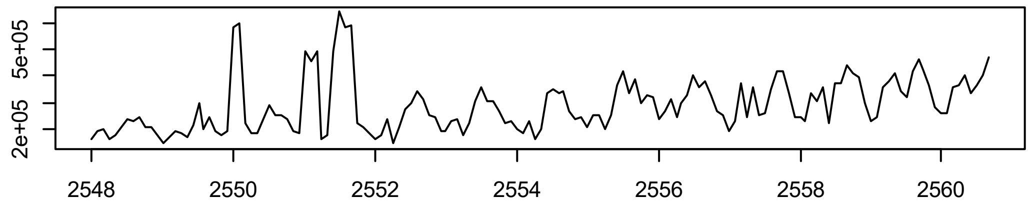
1. The results of the study on the analysis of time series of electricity consumption of Rajabhat Universities from January 2005 to September 2017, as shown in Fig. 2 revealed that the three sets of time series were seasonal movements. That is, the amount of electricity consumption increased with time.



Nakhon Ratchasima Rajabhat University



Ubon Ratchathani Rajabhat University



Loei Rajabhat University

Figure 2 Time series movement of electricity consumption of three Rajabhat Universities from January 2005 to September 2017

2. Results of statistical data analysis

The results of the analysis of basic statistics of the energy consumption of the three Rajabhat Universities found that the average electricity consumption of Nakhon Ratchasima Rajabhat University per month was 391,151.82 kilowatt-hours (kWh). The maximum electricity consumption was 404,918.48 kilowatt-hours (kWh) while the lowest was 130,190.62 kilowatt-hours (kWh). The average electricity consumption of Ubon Ratchathani Rajabhat University per month was 526,615.22 kilowatt-hours (kWh). The maximum electricity consumption was 1,836,456.23 kilowatt-hours (kWh) while the lowest was 78,224.60 kilowatt-hours (kWh). The average electricity consumption of Loei Rajabhat University per month was 289,277.04 kilowatt-hours (kWh). The maximum electricity consumption was 642,140.04 kilowatt-hours (kWh) while the lowest was 144,008.13 kilowatt-hours (kWh) as presented in Table 1.

Table 1 Basic statistics of the energy consumption of the three Rajabhat Universities

Universities	Data	Number	Mean(kWh)	Max. (kWh)	Median(kWh)	Min. (kWh)	Std. (kWh)
Nakhon Ratchasima Rajabhat University	All sample	153	391,151.82	671,632.28	404,918.48	130,190.62	145,475.84
	Training	144	382,578.41	671,632.28	385,365.49	130,190.62	144,127.21
	Testing	9	528,326.48	647,157.19	512,625.64	395,155.44	91,093.04
Ubon Ratchathani Rajabhat University	All sample	153	526,615.22	1,836,456.23	440,475.20	78,224.60	296,960.73
	Training	144	534,904.03	1,836,456.23	446,591.80	78,224.60	303,745.73
	Testing	9	393,994.39	458,360.06	423,577.60	259,001.60	73,011.38
Loei Rajabhat University	All sample	153	289,277.04	642,140.04	254,976.00	144,008.13	102,239.47
	Training	144	284,974.62	642,140.04	251,580.94	144,008.13	102,649.27
	Testing	9	358,115.82	471,501.20	364,193.20	257,427.20	67,781.73

3. Results of the forecasts using Box -Jenkins Method

According to the data on energy consumption of the three Rajabhat Universities, it was found that the data were unstable before analysis. Therefore, the researcher converted the data by logarithm before constructing SARIMA. SARIMA requires orders (p, d, q) and $(P, D, Q)_s$. R-language program with auto.arima function was used to specify appropriate orders (p, d, q) and $(P, D, Q)_s$. The results of the use of auto.arima function to identify the pattern of SARIMA were shown in Table 2.

Table 2 Estimation of SARIMA of the three Rajabhat Universities

Universities		Estimate	Std. Error	z-value	p-value
Nakhon Ratchasima Rajabhat University $SARIMA(0,1,1)(2,0,0)_{12}$	MA(1)	-0.7063	0.0644	-10.974	<0.0001
	SAR(1)	0.5338	0.0800	6.6741	<0.0001
	SAR(2)	0.2684	0.0846	3.1710	0.0015
Ubon Ratchathani Rajabhat University $SARIMA(1,1,1)(1,0,1)_{12}$	AR(1)	0.3225	0.1206	2.968	0.0030
	MA(1)	-0.8605	0.0765	-14.546	<0.0001
	SAR(1)	0.8875	0.0492	7.021	<0.0001
	SMA(1)	-0.7430	0.1255	-4.030	<0.0001
Loei Rajabhat University $SARIMA(1,1,2)(1,0,1)_{12}$	AR(1)	0.3217	0.1646	1.9549	0.0046
	MA(1)	-0.6601	0.1691	-3.9045	0.0001
	MA(2)	-0.2885	0.1511	-1.9094	0.0056
	SAR(1)	0.7697	0.1443	5.3328	<0.0001
	SMA(1)	-0.4495	0.2104	-2.1367	0.0326

When the appropriate model was chosen, the researcher examined the model according to four basic assumptions of SARIMA: 1) the mean error value is zero; 2) The error value is normal distribution; 3) the errors are dependent; and 4) the errors are constant. The results showed that the model passed all the basic assumption, so it was the appropriate model. The results were shown in Table 3 and Table 4, respectively.

Table 3 Examining the basic assumptions of the SARIMA of the three Rajabhat Universities

Universities	The mean error value is 0		The error value is normal distribution		The errors are dependent	
	t-value	p-value	KS -value	p-value	DW-value	p-value
Nakhon Ratchasima Rajabhat University $SARIMA(0,1,1)(2,0,0)_{12}$	0.4237	0.6724	0.0728	0.4307	1.8805	0.2351
Ubon Ratchathani Rajabhat University $SARIMA(1,1,1)(1,0,1)_{12}$	-0.5835	0.5605	1.020	0.0800	2.0139	0.5334
Loei Rajabhat niversity $SARIMA(1,1,2)(1,0,1)_{12}$	1.3811	0.1694	0.1133	0.0501	2.0365	0.5872

Table 4 Ljung – Box Q-statistics used to test Autocorrelation of the rest values of the three Rajabhat Universities

Universities	Ljung – Box Q-statistics		
	Q-statistics	Degree of freedom	p-value
Nakhon Ratchasima Rajabhat University	17.471	16	0.3558
Ubon Ratchathani Rajabhat University	16.994	16	0.3860
Loei Rajabhat University	23.691	16	0.0965

4. Forecasting results using hybrid models of Box –Jenkins Method and support vector regression method

The researcher proposed 5 hybrid models of Box –Jenkins Method and support vector regression: SARIMASVR1, SARIMASVR2, SARIMASVR3, SARIMASVR4 and SARIMASVR5. The details were as described in 2.3. In the creation of the hybrid models, the researcher used R-language program, which contains svm () function in the e1071 package published by Meyer et al. (2015). Before applying SVR to forecast the time series, the researcher adjusted the data set into input vector and target ($D=\{(x_i, y_i)\}_{i=1}^{n}$) according to the models proposed in 2.3. Since the time series used in the study is a monthly period and the number of periods in one season equals 12 calendar periods, the researcher chose to adjust the forecasting value using the past observation values of 12 time periods and used the 3 past observations obtained from SARIMA forecast, including error value, forecasting value and actual value of each model which were different based on the number of vector inputs, respectively. In creating SVR, Kernel function and appropriate parameters must be defined. In this research, the researcher applied a Radial Basis Function (RBF) as it was the most suitable Kernel and it has been applied in a variety of energy forecasting methods. Singchai and Keeratiwintakorn (2014) applied a Radial Basis Function (RBF) to forecast electricity demand for Thailand Demand Side Management Center. The parameters of Cost, ε and γ must be defined by the e1 0 7 1 package in using Radial Basis Function (RBF) as shown in Table 5. After the appropriate parameters of Cost, ε and γ were obtained, they were used for forecasting to compare the models and the appropriate models for predicting the electricity consumption of the three Rajabhat Universities were chosen.

Table 5 The suitability of the parameters of five hybrid models of the three Rajabhat Universities

Models	Universities								
	Nakhon Ratchasima Rajabhat University			Ubon Ratchathani Rajabhat University			Loei Rajabhat University		
	Cost	ε	γ	Cost	ε	γ	Cost	ε	γ
SARIMASVR1	1	0.1	0.0833	1	0.1	0.0833	1	0.1	0.0833
SARIMASVR2	1	0.1	0.07692	1	0.1	0.0769	1	0.1	0.0769
SARIMASVR3	1	0.1	0.07692	1	0.1	0.0769	1	0.1	0.0769
SARIMASVR4	1	0.1	0.04167	1	0.1	0.04167	1	0.1	0.04167
SARIMASVR5	1	0.1	0.0400	1	0.1	0.0400	1	0.1	0.0400

5. Results of forecasting error used to select the appropriate model for forecasting

After the results from the 6 models constructed from the first data set consisting of a single forecasting model, namely SARIMA and 5 hybrid models: SARIMASVR, SARIMASVR2, SARIMASVR3 , SARIMASVR4 and SARIMASVR5 were obtained, in order to obtain the optimal model for forecasting the energy consumption of the three Rajabhat Universities, the researcher compared the forecasting errors by considering the lowest MAPE and MAE values. MAPE was firstly considered. If the values were the same, MAE would be considered. The results of the comparison of the forecast errors were shown in Table 6.

Table 6 The results of the comparison of the forecast errors of 6 models

Models	Universities					
	Nakhon Ratchasima Rajabhat University		Ubon Ratchathani Rajabhat University		Loei Rajabhat University	
	MAPE	MAE	MAPE	MAE	MAPE	MAE
SARIMA	0.87340	0.11090	1.80840	0.23520	1.29954	0.15470
SARIMASVR1	0.95055	0.12018	0.89012	0.11660	1.34992	0.16924
SARIMASVR2	0.29606	0.03795	0.58281	0.03323	0.37309	0.04746
SARIMASVR3	0.60972	0.07756	1.23357	0.06959	0.79602	0.10099
SARIMASVR4	0.63063	0.08017	1.24956	0.07059	0.80804	0.10275
SARIMASVR5	0.31238	0.03995	0.54767	0.03133	0.40036	0.05100

Table 6 showed that SARIMASVR2 was an appropriate model for forecasting the electricity consumption of Nakhon Ratchasima Rajabhat University and Loei Rajabhat University while SARIMASVR5 was a suitable model for forecasting the electricity consumption of Ubon Ratchathani Rajabhat University. The researcher used these models to forecast the electricity consumption for 9 months in advance in order to compare with the second set of data to see the accuracy of the model again. The results were shown in Table 7. After the actual values were compared with the forecasting values, the second data set was mixed with the first data set for forecasting the electricity consumption for 15 months in advance as shown in Table 8 and Figures 3-5.

Table 7 Comparison of actual and forecasting values of energy consumption for 9 months

Months	Universities					
	Nakhon Ratchasima Rajabhat University		Ubon Ratchathani Rajabhat University		Loei Rajabhat University	
	Actual value (kWh)	Forecasting value (kWh)	Actual value (kWh)	Forecasting value (kWh)	Actual value (kWh)	Forecasting value (kWh)
Jan 17	415,651.64	465,508.50	320,596.00	336,848.90	262,297.20	267,827.70
Feb 17	395,155.44	440,082.90	326,490.60	317,228.40	257,427.20	264,020.10
Mar 17	614,988.83	575,863.60	457,340.01	405,883.70	357,396.00	321,096.40
Apr 17	512,625.64	528,793.50	259,001.60	310,177.60	365,756.00	365,044.10
May 17	549,796.76	542,256.10	437,546.40	402,170.90	400,902.81	385,419.70
Jun 17	479,962.24	529,777.30	458,360.06	478,933.60	338,027.60	374,578.00
Jul 17	507,208.20	579,618.20	423,577.60	425,334.10	364,193.20	360,676.30
Aug 17	632,392.40	663,753.00	450,682.00	438,719.70	405,541.21	399,412.40
Sep 17	647,157.19	679,660.60	412,355.20	423,290.90	471,501.20	447,692.30
MAPE	7.65428		6.35679		4.13581	

Table 8 Forecast of electricity consumption for 15 months in advance

Months	Universities		
	Nakhon Ratchasima Rajabhat University	Ubon Ratchathani Rajabhat University	Loei Rajabhat University
Oct 17	592,307.0	413,678.6	429,293.2
Nov 17	551,571.6	422,284.3	360,721.1
Dec 17	455,823.2	312,907.7	335,929.6
Jan 18	448,553.0	343,008.4	248,983.4
Feb 18	405,974.4	271,892.0	217,371.9
Mar 18	624,349.2	476,863.0	290,231.6
Apr 18	596,331.0	512,721.4	346,213.2
May 18	569,810.8	439,920.3	415,641.0
Jun18	483,645.7	335,558.8	321,645.3
Jul 18	527,028.2	312,981.9	370,827.5
Aug 18	641,537.7	483,779.7	313,026.8
Sep 18	584,112.9	436,415.4	465,137.6
Oct18	552,780.8	422,855.2	412,758.5
Nov 18	532,352.8	374,736.9	370,187.2
Dec 18	395,883.5	282,578.3	306,256.0

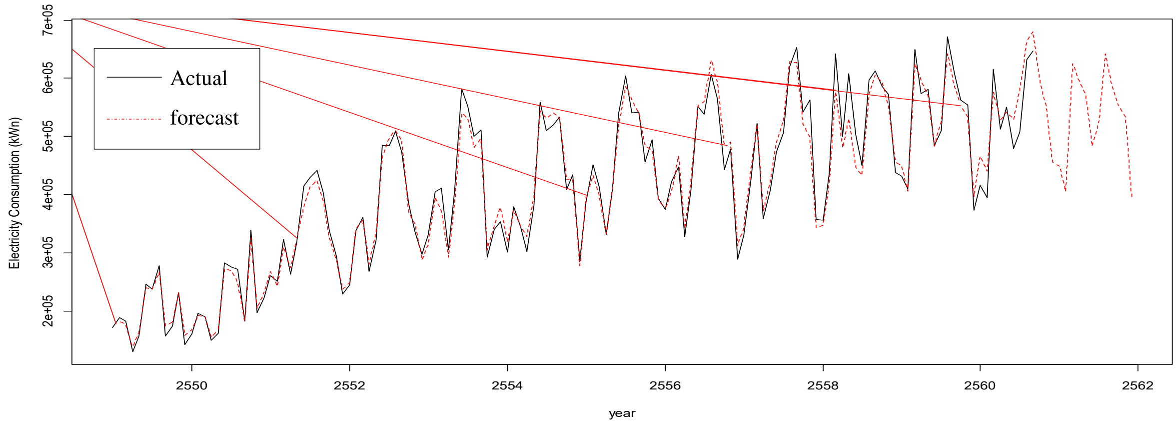


Figure 3 Comparison of actual and forecasting energy consumption values of Nakhon Ratchasima Rajabhat

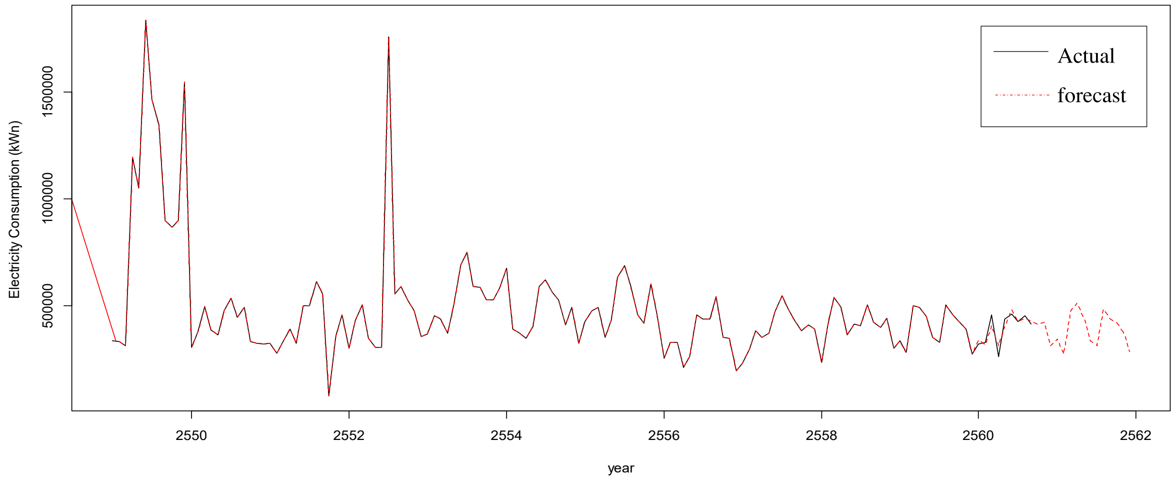


Figure 4 Comparison of actual and forecasting energy consumption values of Ubon Ratchathani Rajabhat University

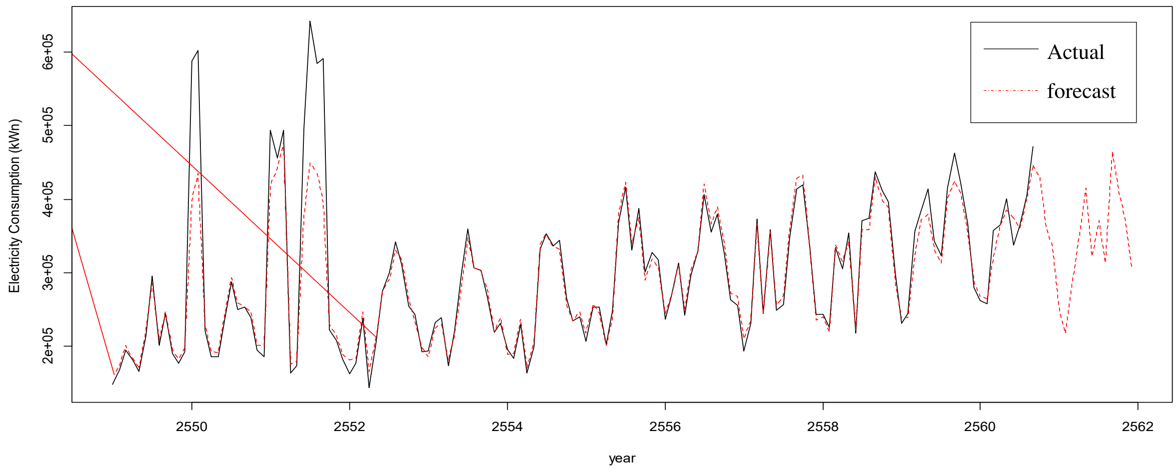


Figure 5 Comparison of actual and forecasting energy consumption values of Loei Rajabhat University

Conclusion, discussion and recommendations

In this research, the researcher investigated the energy consumption of three Rajabhat Universities by using 6 hybrid models of Box –Jenkins Method and support vector regression consisting of a single forecasting model, namely SARIMA and 5 hybrid models: SARIMASVR, SARIMASVR2, SARIMASVR3, SARIMASVR4 and SARIMASVR5. The time series data on electricity consumption on a monthly basis of three Rajabhat Universities (kWh / hour) covering the period from January, 2006 to September, 2017, from the website of the Energy Policy and Planning Office, Ministry of Energy (2017): 153 values were collected. The data were divided into 2 sets. The first set was the data from January 2005 to December 2016: 144 values which were used for the study with 6 time series forecasting models by the R-Language program. The lowest values of MAE, MAPE were considered. The second set was the data from January to September 2017: 9 values which were used to compare the accuracy of the forecasting values. After that electricity consumption of each university was forecasted for 15 months in advance.

The results indicated that from the first set of data, SARIMASVR2 was a suitable model for forecasting the electricity consumption of Nakhon Ratchasima Rajabhat University and Loei Rajabhat University while SARIMASVR5 was a suitable model for forecasting the electricity consumption 0Ubon Ratchathani Rajabhat University. The researcher used these models to forecast the electricity consumption for 9 months in advance in order to compare with the second set of data to see the accuracy of the model again. According to the use of SARIMASVR2 to forecast the electricity consumption of Nakhon Ratchasima Rajabhat University and Loei Rajabhat University, MAPE values were 7.65428 and 4.13581, respectively. In addition, MAPE value from the use of SARIMASVR5 to forecast the electricity consumption of Ubon Ratchathani Rajabhat University was 6.35679. These 3 forecasting models were used to forecast the electricity consumption for 15 months in advance to see the trend of electricity consumption all three Rajabhat universities. Table 8 and Figures 3 -5 showed that the energy consumption of the three Rajabhat Universities was likely to increase and still fluctuated with seasonal factors. That is, in March and June, it was the period with the highest amount of power consumption and it gradually decreased in May to July, and it gradually increased again in August and gradually decreased slightly in September to December which was the lowest amount of power consumption period. The results of this study were consistent with the study conducted by Theeraviriya (2017), which compared the forecasting method for electric energy demand in Nakhon Phanom province and found that the demand for electricity in Nakhon Phanom province was likely to increase with seasonal fluctuations. That is, from March to June, the demand for electricity was high. Singchai and Keeratiwintakorn (2014) forecasted electricity demand for Thailand Demand Side Management Center and found that electricity demand in Thailand was still increasing and based on seasonal fluctuations, especially in houses that the use of electricity was uncertain. It may depend on many external factors such as climate and number of residents, respectively.

Forecasting by the hybrid models of Box -Jenkins Method and support vector regression method was a highly accurate forecasting method since the forecasting values were adjusted and the errors were calculated by SARIMA. However, for the use of forecasting values, various aspects should be taken into consideration because the amount of electricity consumption of each university may not depend on the time factor alone. Other factors may include temperature, number of students, and number of staff or construction of buildings.

Recommendations for further research

1. Since the amount of electricity consumption does not depend on the time factor alone, other factors such as temperature, student number, the number of staff in the office should be considered in the forecast. This will make the forecast more accurate.

2. Before forecasting, data should be filtered as the forecast is more accurate than that of the non-filtered data. Data filtering method, such as EMD (Empirical mode decomposition) is becoming more popular.

Acknowledgement

I would like to express my sincere gratitude to, Assistant Professor Paothai. Wonglao Ph.D., and Associate Professor Jumlong Wongprasert, Ph.D, the lecturers in Applied Statistics Program, Faculty of Science, Ubon Ratchathani Rajabhat University who gave me advice and inspiration in conducting this research. I also would like to thank to the student trainees in Applied Statistics Program who helped record the data in this research.

References

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2008). *Time Series Analysis: Forecasting and Control* (4th ed.). New Jersey: John Wiley and Sons.

Junsagoon, S. (2015). Time Series Forecasting : Linear Approach, Non - Linear Approach and Hybrid Models. *Eau Heritage Journal Science and Technology*, *9*(2), 50-63.

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., Chang, C.C., & Lin, C.C. (2015). *Misc Functions of the Department of Statistics, Probability Theory Group.* Retrieved from http://cran.r-project.org/web/packages/e1071/index.html,5 December

Ministry of Energy. Government Energy Reduction Program. (2017). Retrieved form http://www.e-report.energy.go.th/data/index.php.

Singchai, P., & Keeratiwintakorn, P. (2014). Electricity Demand Forecast for Thailand Demand Side Management Center. *Information Technology Journal*, *10*(2), 32-42.

Sujjaviriyasup, T. (2017). Hybrid Model of Support Vector Machine and Genetic Algorithm for Forecasting the Annual Peak Electricity Demand of Thailand. *Journal of Kmutnb*, *27*(3), 451-463.

Theeraviriya, C. (2017). A Comparison of the Forecasting Method for Electric Energy Demand in Nakhonphanom Province. *Naresuan University Journal: Science and Technology*, *25*(4), 124-137.

Vapnik, V. (1998). *The Nature of Statistical Learning Theory.* Retrieved from https://www.springer.com/gp/book/9780387987804.

Yashlan, Y., & Bican, B. (2017). Empirical mode decomposition based denoising method with support vector regression for time series prediction: A case study for electricity load forecasting. *Measurement*, *103*, 52-61. https://doi.org/10.1016/j.measurement.2017.02.007.

Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputiong*, *50*, 159-175. https://doi.org/10.1016/S0925-2312 (01)00702-0

Zhang, F., Deb, C., Lee, S. E., Yang, J., & Shah, K. W. (2016). Time series forecasting for building energy consumption using weighted Support Vector Regression with differential evolution optimization technique. *Energy and Building*, *126*, 94-103. https://doi.org/10.1016/j.enbuild.2016.05.028