□   4892

Applying of Double Seasonal ARIMA Model for Electrical Power Demand Forecasting at PT. PLN Gresik Indonesia

Ismit Mado[1], Adi Soeprijanto[2], Suhartono[3]
[1,2]Department of Electrical Engineering, Institut Teknologi Sepuluh Nopember, Indonesia
[3]Department of Statistics, Institut Teknologi Sepuluh Nopember, Indonesia
[1]Department of Electrical Engineering, Universitas Borneo Tarakan, Indonesia

ABSTRACT

The prediction of the use of electric power is very important to maintain a balance between the supply and demand of electric power in the power generation system. Due to a fluctuating of electrical power demand in the electricity load center, an accurate forecasting method is required to maintain the efficiency and reliability of power generation system continuously. Such conditions greatly affect the dynamic stability of power generation systems. The objective of this research is to propose Double Seasonal Autoregressive Integrated Moving Average (DSARIMA) to predict electricity load. Half hourly load data for of three years period at PT. PLN Gresik Indonesia power plant unit are used as case study. The parameters of DSARIMA model are estimated by using least squares method. The result shows that the best model to predict these data is subset DSARIMA with order $([1,2,7,16,18,35,46], 1, [1,3,13,21,27,46])(1,1,1)^{48}(0,0,1)^{336}$ with MAPE about 2.06%. Thus, future research could be done by using these predictive results as models of optimal control parameters on the power system side.

Corresponding Author:

Ismit Mado,
Department of Electrical Engineering,
Institut Teknologi Sepuluh Nopember,
Arief Rahman Hakim Road, Keputih Campus, Sukolilo 60111, Surabaya, Indonesia.
Email: ismitmado@gmail.com

1.   INTRODUCTION

Prediction of electrical power demand is an important first step in planning of power plant system operation [1]. Planning of power plant system operation is about building a plan for the preparation of power plant system operation for a certain time period. Based on the issues to be addressed, the power plant system operation plan is divided into several types of time period plan, i.e. annual plan, quarterly plan, monthly plan, weekly plan and daily plan. The operation plan of the power plant system also includes managerial, maintenance, and operations in systems and equipment to ensure good economic value of financing, system reliability, and service quality.

Referring to the problems solving in power system operation, power prediction is classified into three categories, i.e. long-term, medium-term and short-term predictions. Long term electrical power predictionsarerequired for peak load capacity planning and system maintenance schedule [2], medium-term predictions are required for planning and operation of the power plant system [3], and short-term predictions are needed for controlling and scheduling the power plant system [4].

One of the concerns in the power plant system operation is the quality of electrical power. The electric power generated shall be always equal to the electric power consumed by the electric power user. If the power that is distributed is greater than required, then there will be wastage of energy. And if the power produced is smaller than required, it will occur over load which will affect the occurrence of power outages.

In fact, the use of electrical energy tends to change at any time in accordance with the needs of consumers. Therefore, it is necessary to predict the use of electrical power that is able to maintain a balance between the supply and consumption of electrical power in the power plant system.

Fluctuations in loads mean small disturbances in the dynamical stability studies of power generation systems. This problem occurs after the first swing when control equipment such as governor and excitation systems have worked. Dynamic stability studies in plant systems are still being developed [5],[6]. Analysis of small signal stability of power systems with probabilistic uncertainty on renewable energy generation equipment becomes the topic of the current study [7].

The time series prediction model is an accurate choice and growing continuously to this day for power prediction and forecasting [8]-[10]. The study of time series-based forecasting of electrical power consumption evolved into two parts, i.e. prediction models based on statistical mathematical models such as moving average, exponential smoothing, regression, and ARIMA Box-Jenkins; and artificial intelligence-based prediction models such as neural networks, genetic algorithms, simulated annealing, genetic programming, classification, and hybrid. Some time series research based on statistical mathematics as in [8], [11]-[15]. The application of the Box-Jenkins ARIMA model is developed through a seasonal pattern [16]-[22].For models based on artificial intelligence has also become the attention of researchers as in [23]-[27]. The hybrid model has also been developed to obtain the best data in electrical load prediction study as in [28]-[36].

This study proposes a Double Seasonal ARIMA (DSARIMA) method for predicting or forecasting power demand model in PT. PLN Gresik Indonesia based on three years load training and testing data (daily data every half hour). The prediction results are used as reference parameters for optimum control in improving the stability of electrically generated systems in our research. Time series method based on statistical mathematical model is chosen because of its superiority to process data which is not stationary and not linear. Statistical mathematical models are also capable of generating data that is not included in the training process.

2.   FORECASTING METHODS
2.1. ARIMA Model

ARIMA method or commonly referred to as Box-Jenkins method is a model intensively developed by George Box and Gwilyn Jenkins in 1970. This forecasting model still dominates many areas of research to date. Unfortunately, ARIMA model can only be applied for stationary time-series data. If the data is not stationary, then to make the data becomes stationary it is necessary to do the differentiation process [37].The autoregressive (AR) model indicates a connection between a value at the present time (Z_t) with a value in the previous time (Z_{t-k}), plus a random value. While the moving average (MA) model shows the dependence of the current time value (Z_t) with the residual value at the previous time (Z_{t-k}) with $k = 1,2, \dots n$.

The ARIMA model (p, d, q) is a combination of AR(p) and MA(q) models, with dth-order differentiation process when the data pattern is not stationary. General form of the ARIMA model (p, d, q) is as follows:

$$\emptyset_p(B)(1 - B)^d \dot{Z}_t = \theta_q(B)a_t \tag{1}$$

where B is the backshift operator and a_t is the random process values, and

$$\emptyset_p(B) = \left(1 - \emptyset_1 B - \dots - \emptyset_p B^p\right)$$

$$\theta_q(B) = \left(1 - \theta_1 B - \dots - \theta_q B^q\right)$$

Generalization of the ARIMA model for data that has a seasonal pattern is expressed by ARIMA$(p, d, q)(P, D, Q)^s$ and formulated as follows [37]:

$$\emptyset_p(B)\Phi_P(B^s)(1 - B^s)^d(1 - B^s)^D \dot{Z}_t = \theta_q(B)\Theta_Q(B^s)a_t \tag{2}$$

where s is the seasonal period, and

$$\emptyset_p(B) = 1 - \emptyset_1 B - \emptyset_2 B^2 - \dots - \emptyset_p B^p$$

$$\Phi_P(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_P B^{Ps}$$

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

$$\Theta_Q(B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} - \ldots - \Theta_Q Q^{Qs}$$

Short-term power consumption data by consumers has a double seasonal pattern, i.e. daily and weekly season. The ARIMA model with a double seasonal pattern is expressed by ARIMA $(p, d, q)(P_1, D_1, Q_1)^{s_1}(P_2, D_2, Q_2)^{s_2}$ and has the following general form [38]:

$$\emptyset_p(B)\Phi_{P1}(B^{s_1})\Phi_{P2}(B^{s_2})(1-B)^d(1-B^{s_1})^{D_1}(1-B^{s_2})^{D_2}\dot{Z}_t = \theta_q(B)\Theta_{Q1}(B^{s_1})\Theta_{Q2}(B^{s_2})a_t \quad (3)$$

where s_1 and s_2 are different seasonal periods.

2.2. ARIMA Box-Jenkins Model Procedure

The prediction procedure of ARIMA Box-Jenkins model through five stages of iteration, as follows:
i.   Preparation of data, including checking of data stationary.
ii. Identification of ARIMA model through autocorrelation function and partial autocorrelation function.
iii. Estimation of ARIMA model parameters: *p*, *d*, and *q*.
iv. Determination of ARIMA model equations.
v.   Prediction.

2.3. Least Squares Estimation

One method that can be used to estimate ARIMA model parameters is the least squares method [39]. For AR (1), carried out by including non-zero mean parameter, μ. This parameter is further estimated by least squares. Consider the first-order case where $Z_t - \mu = \emptyset(Z_{t-1} - \mu) + a_t$: The equation is a regression model with Y_{t-1} as the predictor variable and Y_t as the response variable. Least squares estimation then processed by minimizing the sum of squares of differences

$$(Z_t - \mu) - \emptyset(Z_{t-1} - \mu) \tag{4}$$

Since only $Z_1, Z_2, \ldots, Z_n$ are observed, then we can only sum from $t = 2$ to $t = n$. Let it

$$S_c(\emptyset, \mu) = \sum_{t=2}^{n}\{(Z_t - \mu) - \emptyset(Z_{t-1} - \mu)\}^2 \tag{5}$$

This equation is called the conditional sum squared function. According to the basic principle of the least squares method, we can estimate $\emptyset$ and μ by the respective values through parameter values $Z_1, Z_2, \ldots, Z_n$. Consider the equation $\partial S_c / \partial \mu = 0$. We have

$$\frac{\partial S_c}{\partial \mu} = \sum_{t=2}^{n} 2[(Z_t - \mu) - \emptyset(Z_{t-1} - \mu)](-1 + \emptyset) = 0 \tag{6}$$

or, simplifying and solving for μ,

$$\mu = \frac{1}{(n-1)(1-\emptyset)}\left[\sum_{t=2}^{n} Z_t - \emptyset\sum_{t=2}^{n} Z_{t-1}\right] \tag{7}$$

Now, for large *n*,

$$\frac{1}{n-1}\sum_{t=2}^{n} Z_t = \frac{1}{n-1}\sum_{t=2}^{n} Z_{t-1} = \bar{Z}$$

Thus, regardless of the value of $\emptyset$, equation (7) reduces to

$$\hat{\mu} = \frac{1}{1-\emptyset}(\bar{Z} - \emptyset\bar{Z}) = \bar{Z} \tag{8}$$

then it can be written, $\hat{u} = \bar{Z}$

Next we reconsider minimizing the equation $S_c(\emptyset, \bar{Z})$. We have

$$\frac{\partial S_C}{\partial \emptyset} = \sum_{t=2}^{n} 2[(Z_t - \bar{Z}) - \emptyset(Z_{t-1} - \bar{Z})](Z_{t-1} - \bar{Z}) \tag{9}$$

Setting this equal to zero and solving for $\emptyset$ yields

$$\widehat{\emptyset} = \frac{\sum_{t=2}^{n}(Z_t - \bar{Z})(Z_{t-1} - \bar{Z})}{\sum_{t=2}^{n}(Z_{t-1} - \bar{Z})^2} \tag{10}$$

Except for one term missing in the denominator $(Z_n - \bar{Z})^2$, this is the same as r_1. The lone missing term is negligible for stationary prosses, and thus the least squares and method-of-moments estimators are nearly identical, especially for large samples.

For the general AR(p) process, the methods used to obtain equations (7) and (8) can easily be extended to yield the same result, namely $\hat{\mu} = \bar{Z}$. To generalize the estimation $\emptyset$'s, we can be considered through a second-order equation. So, we replace μ by $\bar{Z}$ in the conditional sum-of-squares function,

$$S_C(\emptyset_1, \emptyset_2, \bar{Z}) = \sum_{t=3}^{n}[(Z_t - \bar{Z}) - \emptyset_1(Z_{t-1} - \bar{Z}) - \emptyset_2(Z_{t-2} - \bar{Z})]^2 \tag{11}$$

Setting $\partial S_c/\partial \emptyset_1 = 0$, we have

$$-2\sum_{t=3}^{n}(Z_t - \bar{Z}) - \emptyset_1(Z_{t-1} - \bar{Z}) - \emptyset_2(Z_{t-2} - Z) = 0 \tag{12}$$

Which we can rewrite as

$$\sum_{t=3}^{n}(Z_t - \bar{Z})(Z_{t-1} - \bar{Z}) = (\sum_{t=3}^{n}(Z_{t-1} - \bar{Z})^2)\emptyset_1 + (\sum_{t=3}^{n}(Z_{t-1} - \bar{Z})(Z_{t-2} - \bar{Z}))\emptyset_2 \tag{13}$$

The sum of the lagged products by $\sum_{t=3}^{n}(Z_t - \bar{Z})(Z_{t-1} - \bar{Z})$ is very nearly the numerator of r_1. We can divide both sides of the equation by $\sum_{t=3}^{n}(Z_t - \bar{Z})^2$ then, excepts for end effects, which are negligible under the stationary assumption, we obtain

$$r_1 = \emptyset_1 + r_1\emptyset_2 \tag{14}$$

With the same approach for differentiation equations $\partial S_C/\partial \emptyset_2 = 0$, obtained

$$r_2 = r_1\emptyset_1 + \emptyset_2 \tag{15}$$

These two equations are called Yule-Walker equations for the AR(2) model. For MA(1), $Z_t = a_t - \theta a_{t-1}$; if the MA model is invertible then

$$Z_t = -\theta Z_{t-1} - \theta^2 Z_{t-2} - \theta^3 Z_{t-3} - \cdots + a_t \tag{16}$$

Then the least squares can be done by selecting a value θ that minimizes.

$$S_C(\theta) = \sum(a_t)^2 = \sum[Z_t + \theta Z_{t-1} + \theta^2 Z_{t-2} + \theta^3 Z_{t-3} + \cdots]^2 \tag{17}$$

It is obvious that the least squares problem in equation (17) is nonlinear in the parameter. We will not be able to minimize $S_C(\theta)$ by taking a derivative with respect to θ, setting it to zero, and solving. To address these issues, consider evaluating $S_C(\theta)$ for a single given value of θ. Rewrite first-order equation as

$$a_t = Z_t + \theta a_{t-1} \tag{18}$$

Using this equation, $a_1, a_2, \ldots, a_n$ can be calculated recursively if we have the initial value a_0. A common approximation is to set $a_0 = 0$. We can obtain

$$\left. \begin{array}{l} a_1 = Z_1 \\ a_2 = Z_2 + \theta a_1 \\ a_3 = Z_3 + \theta a_2 \\ \quad\vdots \\ a_n = Z_n + \theta a_{n-1} \end{array} \right\} \tag{19}$$

For higher-order moving average models, we can compute $a_t = a_t(\theta_1, \theta_2, \ldots, \theta_q)$ recursively from

$$a_t = Z_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \cdots + \theta_q a_{t-q} \tag{20}$$

With $a_0 = a_{-1} = \cdots = a_{-q} = 0$. The sum of squares is minimized jointly in $\theta_1, \theta_2, \ldots, \theta_q$ using a multivariate numerical method. For a general ARMA model: the conditional sum ofsquares can be defined much as in the MA case, we consider $a_t = a_t(\emptyset, \theta)$ and wish to minimize $S_C(\emptyset, \theta) = \sum a_t^2$, so

$$a_t = Z_t - \emptyset Z_{t-1} + \theta a_{t-1} \tag{21}$$

To obtain a_1, we now have an additional problem, namely Z_0. One approach is to set $Z_0 = 0$ or to $\bar{Z}$ if our model contains a nonzero mean. However, a better approach is to begin the recursion at $t = 2$, thus avoiding Z_0 altogether, and simply minimize, $S_C(\emptyset, \theta) = \sum_{t=2}^{n} a_t^2$. For general ARMA$(p, q)$ model, we compute

$$a_t = Z_t - \emptyset_1 Z_{t-1} - \emptyset_2 Z_{t-2} - \cdots - \emptyset_p Z_{t-p} + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \cdots + \theta_q a_{t-q} \tag{22}$$

With $a_p = a_{p-1} = \cdots = a_{p+1-q} = 0$ and then minimize $S_C(\emptyset_1, \emptyset_2, \ldots, \emptyset_p, \theta_1, \theta_2, \ldots, \theta_q)$ numerically to obtain the conditional least squares estimates of all parameters.

2.4. Measuring Accuracy Level of Prediction Results

Basically, measuring the accuracy of prediction results can be done by various statistical analysis methods; such as the root mean of square error (RMSE), the mean of absolute error (MAE) value and the mean of absolute percentage error (MAPE). In this research, we used MAPE as standard measurement of prediction results accuracy. MAPE is defined as follows [40]:

$$MAPE = \frac{\sum_{t=1}^{n} \left| \frac{Z_t - \hat{Z}_t}{Z_t} \right|}{n} \times 100\% \tag{23}$$

where Z_t and $\hat{Z}_t$ are the actual value and prediction value, while n is the number of prediction values.

2.5. Data Set

This study uses electrical power demand in the load center data (taken every half hour) at power plant unit of PT. PLN Gresik Indonesia on January 1, 2009 - December 31, 2011. Where, data from January 1, 2009 - December 24, 2011 is used for forecasting and data from December 25 - 31, 2011 is used for testing.

3. RESULTS AND DISCUSSION

3.1. Model Fitting and Identification Parameter

Exploration of electrical power consumption data is done through time series plot on January 1, 2009 - December 24, 2011. The data pattern is very fluctuate as shown in Figure 1. This condition may be influenced by the integrated power distribution system in Java-Madura-Bali interconnection system in Indonesia. From the picture it shown that the data has not been stationary.
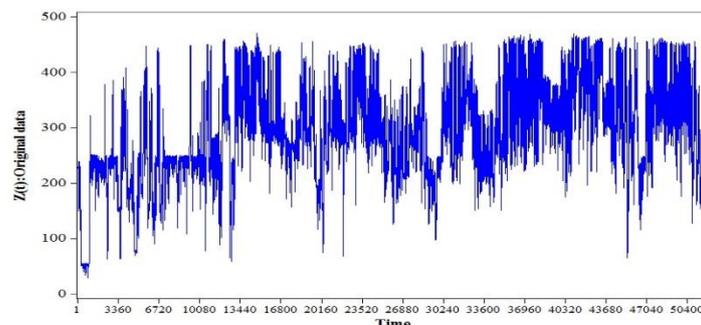

Figure 1. Electrical power data every half hour on January 1, 2009 - December 24, 2011

Figure 2 shows that the ACF coefficient is significantly different from zero and the PACF coefficient is close to zero after the first lag. Based on these two points, it shows that the average data has not been stationary and there is a seasonal patterned trend. Therefore, it is necessary to process 1st level differencing ($d = 1$).
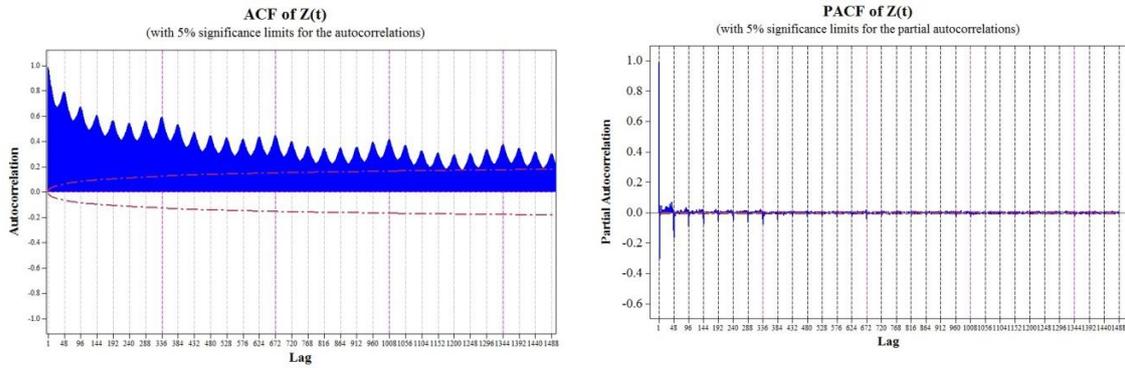

Figure 2. The ACF and PACF of Z_t

The plots of ACF and PACF in Figure 3 have through a differentiation process so that the non-seasonal conditions have been stationary in the mean value. Based on the ACF plot it shows that the autocorrelation values of the stationary data go down to zero after the second lag and the third lag, while for the seasonal data it is still not stationary in the mean value. The ACF plot also shows the existence of another seasonal pattern which is a weekly seasonal pattern on lag 336, 672 and so on.
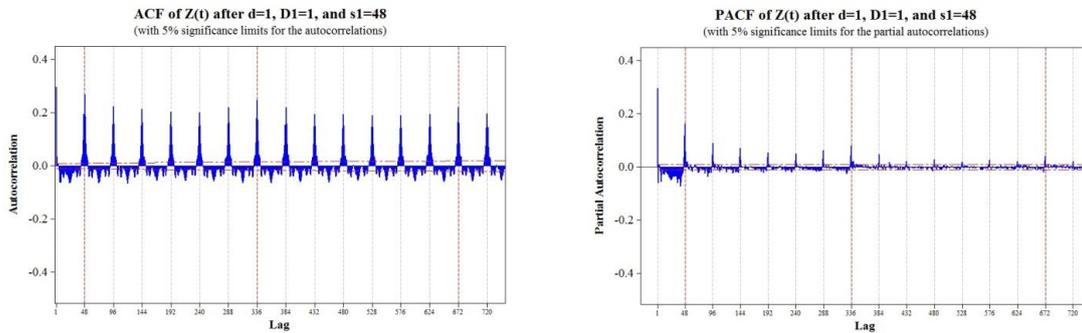

Figure 3. The ACF and PACF of Z_t after $d = 1, D_1 = 1$ and $s_1 = 48$
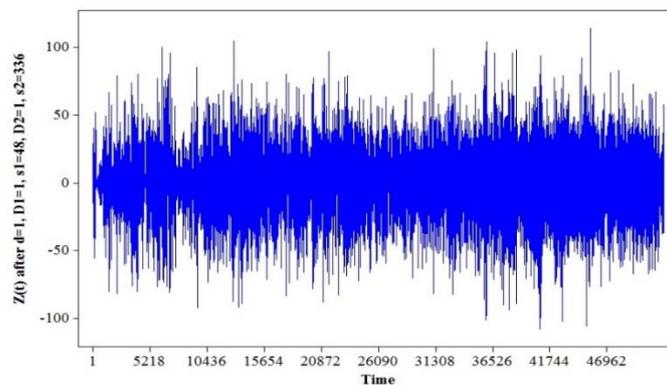

Figure 4: Load demand series after $d = 1, D_1 = 1, s_1 = 48, D_2 = 1,$ and $s_2 = 336$

The ACF plot of the load data pattern in Figure 5 has been stationary in the mean value after going through a second order differentiation process (with $D_2 = 1$). For non-seasonal patterns based on ACF and PACF plots, the load tends to decrease gradually especially after the first lag so that the non-seasonal model assumption is the ARMA(1,1) model. While the daily seasonal pattern ($s_1 = 48$) shows that the plot of the ACF data pattern tends to be interrupted after the lag 48 and the PACF pattern indicates the data pattern tends to decrease gradually, then the assumption of the daily seasonal model is $MA(1)^{48}$model.For the weekly seasonal pattern ($s_2 = 336$) it shows that ACF and PACF plots along lag 336, 672, and so on tend to fall gradually, then the assumption of the weekly seasonal model is $MA(1,1)^{336}$model.
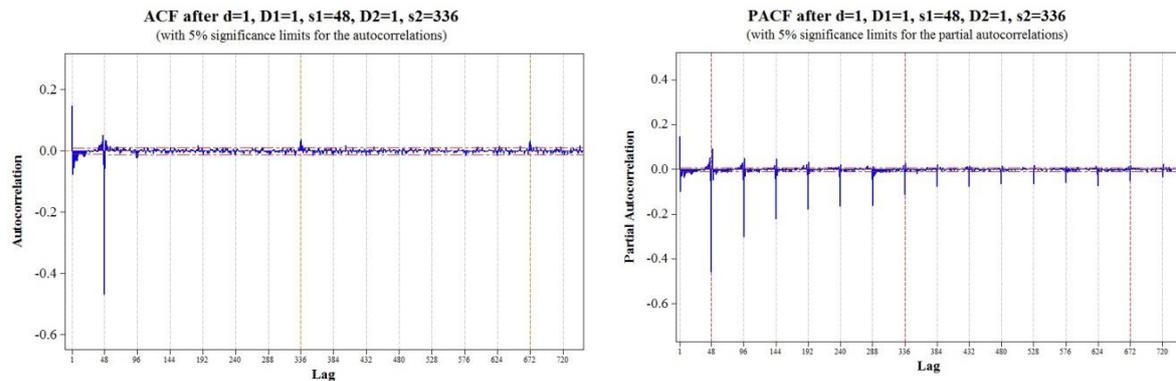

Figure 5. The ACF and PACF of Z_t after $d = 1, D_1 = 1, s_1 = 48, D_2 = 1$ and $s_2 = 336$

Based on the identification of ACF and PACF patterns, the assumption of an appropriate ARIMA model is a double seasonal ARIMA model $(1,1,1)(0,1,1)^{48}(0,0,1)^{336}$.However, if white noise is detected in the test data, it is necessary to add or substitute the order of differentiation process. In this study, Statistical Analysis System (SAS) programming tools are used to analyze load data of double seasonal ARIMA models.

3.2. DSARIMA Model Parameter Estimation

The AR and MA coefficients in the DSARIMA model are estimated by the least squares method. The initial estimate obtained has been used as the initial value of the iterative estimation method. Through the double seasonal ARIMA model $(1,1,1)(0,1,1)^{48}(0,0,1)^{336}$, the initial data of the AR and MA coefficients were obtained as follows:

Based on Table 1, to meet the white noise criterion, p-value must be greater than fault tolerance $\alpha = 5\%$, with alpha significance level less than 0.0001.In addition, the model has an improvement pattern with 3 MA parameters i.e. MA(1,1), MA(2,1) dan MA(3,1), so that these three parameters should be included in the model estimation. While the residual assumption test that includes the assumption of white noise must meet the criteria of independent and normal distribution $(0, \sigma^2)$.

The Ljung-Box test is used to check the residual independence assumption with the following hypothesis:

H_0 : $\rho_1 = \rho_2 = \ldots = \rho_K = 0$
H_1 : at least one $\rho_1$which is not equal to zero for $i = 1, 2, \ldots, K$.

with a fault tolerance of 5% then H_0 is rejected if p-value $< \alpha$, which means the residual does not meet the white noise assumption.

Based on the final result of AR and MA coefficient parameter estimation in Table 2, it can be plotted the residual normal probability to determine whether residual has fulfilled white noise assumption with limit of $\pm 1.96/\sqrt{n} \approx \pm 0.009$. By iterating the addition of AR and MA parameters, the best iteration value has been obtained which has fulfilled white noise assumption, i.e. double seasonal ARIMA $([1,2,7,16,18,35,46], 1, [1,3,13,21,27,46])(1,1,1)^{48}(0,0,1)^{336}$.

Table 1. An Output SAS of model with CLS iterative

The ARINA Procedure Conditional Least Squares Estimation					
Paramater	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag
MA1, 1	-0.35184	0.01899	-18.53	<.0001	1
MA2, 1	0.95734	0.0013007	736.02	<.0001	48
MA3, 1	-0.04526	0.0045103	-10.03	<.0001	336
AR1, 1	-0.14578	0.02006	-7.27	<.0001	1
Variance Estimate				130.7053	
Std Error Estimate				11.43264	
AIC				401946	
SBC				401981	
Number of Residuals				52127	
*AIC and SBC do not include log determinant					

Autocorrelation Check of Residuals									
To Lag	Chi-Square	DF	Pr>ChiSq	Autocorrelations					
6	153.39	2	<.0001	-0.002	-0.019	-0.041	-0.017	-0.023	-0008
12	274.15	8	<. 0001	-0.033	-0.027	-0.014	-0.009	-0.014	-0.007
18	342.13	14	<.0001	-0.009	-0.009	-0.008	-0.017	-0.016	-0.023
24	422.74	20	<.0001	-0.023	-0.018	-0.020	-0.011	-0.013	-0.003
30	473.05	26	<.0001	-0.009	-0.008	-0.017	-0.008	-0.017	-0.014
36	489.03	32	<.0001	-0.011	-0.009	-0.002	0.000	-0.010	0.000
42	497.60	38	<.0001	-0.007	-0.008	0.002	-0.005	-0.004	0.003
48	804.03	44	<.0001	0.001	0.002	0.006	0.018	0.004	0.060

Table 2. An Output SAS of model

Conditional Least Squares Estimation					
Parameter	Estimate	Standard Error	T Value	Approx Pr>|t|	Lag
MA1,1	0.95131	0.01292	73.64	<.0001	1
MA1,2	-0.07269	0.0064826	-11.21	<.0001	3
MA1,3	0.01357	0.0030786	4.41	<.0001	13
MA1,4	0.01053	0.0028620	3.68	0.0002	21
MA1,5	0.01727	0.0025884	6.67	<.0001	27
MA1,6	0.05074	0.0045478	11.16	<.0001	46
MA2,1	0.97155	0.0010890	892.14	<.0001	48
MA3,1	-0.05320	0.0044774	-11.88	<.0001	336
AR1,1	1.13300	0.01402	80.80	<.0001	1
AR1,2	-0.28056	0.0078422	-35.78	<.0001	2
AR1,3	-0.01869	0.0025332	-7.38	<.0001	7
AR1,4	-0.01052	0.0032983	-3.19	0.0014	16
AR1,5	-0.01024	0.0032260	-3.18	0.0015	18
AR1,6	-0.0068893	0.0021609	-3.19	0.0014	35
AR1,7	0.07395	0.0049768	14.86	<.0001	46
AR2,1	0.04051	0.0049489	8.19	<.0001	48
Variance Estimate				127.2317	
Std Error Estimate				11.2797	
AIC				400554	
SBC				400695.7	
Number of Residuals				52127	
*AIC and SBC do not include log determinant					

Autocprrelation Check of Residuals									
To Lag	Chi-Square	DF	Pr>ChiSq	Autocorrelations					
6	.	0	.	-0.000	-0.000	0.000	0.005	-0.004	0.004
12	.	0	.	-0.006	-0.002	0.006	0.003	-0.007	-0.004
18	12.54	2	0.0019	0.004	0.000	-0.002	-0.002	-0.003	-0.002
24	16.38	8	0.0373	-0.004	-0.004	0.001	0.005	-0.001	0.004
30	23.11	14	0.0586	-0.004	-0.006	0.001	0.008	-0.003	-0.002
36	27.18	20	0.1304	-0.003	-0.003	0.001	0.001	-0.003	0.007
42	31.53	26	0.2093	-0.004	-0.004	0.006	-0.002	0.007	-0.011
48	39.26	32	0.1765	0.003	0.003	0.006	-0.006	0.007	-0.001

3.3. Model Testing and Measuring of Forecasting Level Accuracy

For accuracy measurement, testing uses the MAPE procedure. Based on the comparison of prediction load data with actual load data, the average data accuracy of 2,06% is obtained.

Applying of Double Seasonal ARIMA Model for Electrical Power Demand Forecasting ... (Ismit Mado)

3.4. Prediction

Based on the final result of parameter estimation in Table 1 we get the ARIMA coefficient parameters as follows: AR(1,1) = 1.133, AR(1,2) = -0.28056, AR(1,3) = -0.01869, AR(1,4) = -0.01052, AR(1,5) = -0.01024, AR(1,6) = -0.0068893, AR(1,7) = 0.07395, AR(2,1) = 0.04051 MA(1,1) = 0.95131, MA(1,2) = -0.07269, MA(1,3) = 0.01357, MA(1,4) = 0.01053, MA(1,5) = 0.01727, MA(1,6) = 0.05074, MA(2,1) = 0.97155, MA(3,1) = -0.0532.

Through the parameter of this prediction model we get the equation for DSARIMA model $([1,2,7,16,18,35,46], 1, [1,3,13,21,27,46])(1,1,1)^{48}(0,0,1)^{336}$ as follows:

$$(1 - B)(1 - B^{48})(1 - 1.133B + 0.281B^2 + 0.019B^7 + 0.011B^{16} + 0.010B^{18} + 0.007B^{35} - 0.074B^{46})(1 - 0,041B^{48})\dot{Z}_t = (1 - 0.951B + 0.073B^3 - 0.014B^{13} - 0.011B^{21}   -0.017B^{27} - 0.051B^{46}) (1 - 0.972B^{48})(1 + 0.053B^{336})a_t$$
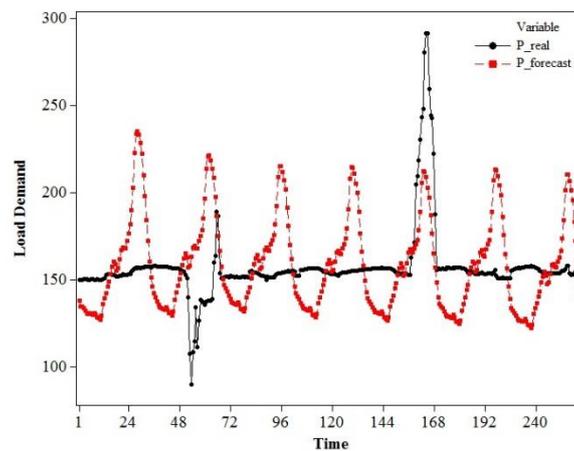

Figure 6. The out-samples of actual data and one-step ahead out-sample forecasts

4.   CONCLUSION

Statistical analysis based on DSARIMA model is suitable with electrical power characteristics with continuous and fluctuating load patterns. The load changes are always unexpected at any time depending on electrical power demand in the load center. With the statistical analysis model, predictions are able to generate data that is not included in the data training process. Through the best model assumption, the model in this study was able to predict with the average accuracy of MAPE of 2.06%. Further research that can be developed is the pattern of electrical power demand on a large-scale area such as Java-Madura-Bali, Indonesia electricity network interconnection.

REFERENCES

[1]   G. J. Tsekouras, *et al.*, "A non-linear multivariable regression model for midterm energy forecasting of power systems," *Electr. Power Syst. Res.*, vol/issue: 77(12), pp. 1560–1568, 2007.
[2]   P. E. McSharry, *et al.*, "Probabilistic forecasts of the magnitude and timing of peak electricity demand," *IEEE Trans. Power Syst.*, vol/issue: 20(2), pp. 1166–1172, 2005.
[3]   E. G. Romera, "Monthly Electric Energy Demand Forecasting Based on Trend Extraction," *IEEE Trans. Power Syst.*, vol/issue: 21(4), pp. 1946–1953, 2006.
[4]   J. W. Taylor and P. E. McSharry, "Short-term load forecasting methods: An evaluation based on European data," *IEEE Trans. Power Syst.*, vol/issue: 22(4), pp. 2213–2219, 2007.
[5]   F. W. Bernal, *et al.*, "Small-signal analysis of power system swing modes as affected by wind turbine-generators," *2016 IEEE Power and Energy Conference at Illinois, PECI 2016*, 2016.
[6]   K. K. Al-Dalwi and A. M. Vural, "Modeling and simulation of Bazyan Gas Power Plant," *2017 4th International Conference on Electrical and Electronics Engineering, ICEEE 2017*, 2017.
[7]   Y. Zhou, *et al.*, "The Stochastic Response Surface Method for Small-signal Stability Study of Power System with Probabilistic Uncertainties in Correlated Photovoltaic and Load," *IEEE Transactions on Power Systems*, vol. 99, pp. 1, 2017.
[8]   H. Hahn, *et al.*, "Electric load forecasting methods: Tools for decision making," *Eur. J. Oper. Res.*, vol/issue:

199(3), pp. 902–907, 2009.

[9] M. M. Nia, *et al.*, "Stochastic approach to a rain attenuation time seris synthesizer for heavy rain regions," *International Journal of Electrical and Computer Engineering (IJECE)*, vol/issue: 6(5), pp. 2379–2386, 2016.

[10] S. Kamley, *et al.*, "Performance forecasting of share market using machine learning techniques: A review," *International Journal of Electrical and Computer Engineering (IJECE)*, vol/issue: 6(6). pp. 3196–3204, 2016.

[11] A. T. Lora, *et al.*, "Time-series prediction: Application to the short-term electric energy demand," *Curr. Top. Artif. Intell.*, pp. 577–586, 2004.

[12] H. M. Al-Hamadi and S. A. Soliman, "Long-term/mid-term electric load forecasting based on short-term correlation and annual growth," *Electr. Power Syst. Res.*, vol/issue: 74(3), pp. 353–361, 2005.

[13] H. M. Al-Hamadi and S. A. Soliman, "Short-term electric load forecasting based on Kalman filtering algorithm with moving window weather and load model," *Electr. Power Syst. Res.*, vol/issue: 68(1), pp. 47–59, 2004.

[14] J. W. Taylor, "Density forecasting for the efficient balancing of the generation and consumption of electricity," *Int. J. Forecast.*, vol/issue: 22(4), pp. 707–724, 2006.

[15] S. A. Soliman and A. M. Al-Kandari, "Electrical Load Forecasting," 2010.

[16] S. N. Hassan, *et al.*, "A comparison of the forecast performance of double seasonal ARIMA and double seasonal ARFIMA models of electricity load demand," *Appl. Math. Sci.*, vol/issue: 6(133–136), pp. 6705–6712, 2012.

[17] N. Mohamed, *et al.*, "Double Seasonal ARIMA Model for Forecasting Load Demand," *Matematika*, vol/issue: 26(2), pp. 217–231, 2010.

[18] N. Mohamed, *et al.*, "Improving Short Term Load Forecasting Using Double Seasonal Arima Model," *World Appl. Sci. J.*, vol/issue: 15(2), pp. 223–231, 2011.

[19] Suhartono, "Time Series Forecasting by using Seasonal Autoregressive Integrated Moving Average: Subset, Multiplicative or Additive Model," *J. Math. Stat.*, vol/issue: 7(1), pp. 20–27, 2011.

[20] J. D. J. Deng and P. Jirutitijaroen, "Short-term load forecasting using time series analysis: A case study for Singapore," *Cybern. Intell. Syst. (CIS), 2010 IEEE Conf.*, vol. 1, pp. 1–6, 2010.

[21] C. Nataraja, *et al.*, "Short Term Load Forecasting Using Time Series Analysis: A Case Study for Karnataka, India," *Int. J. Eng. Sci. Innov. Technol.*, vol/issue: 1(2), pp. 45–53, 2012.

[22] N. Mohamed, *et al.*, "Improving Short Term Load Forecasting Using Double Seasonal Arima Model," *Appl. Sci.*, vol/issue: 15(2), pp. 223–231, 2011.

[23] C. M. Pereira, *et al.*, "Fuzzy modeling to forecast an electric load time series," in *Procedia Computer Science*, vol. 55, pp. 395–404, 2015.

[24] R. Mamlook, *et al.*, "A fuzzy inference model for short-term load forecasting," *Energy Policy*, vol/issue: 37(4), pp. 1239–1248, 2009.

[25] S. Ramos, *et al.*, "Short-term load forecasting based on load profiling," in *2013 IEEE Power & Energy Society General Meeting*, pp. 1–5, 2013.

[26] N. Kandil, *et al.*, "An efficient approach for short term load forecasting using artificial neural networks," *Int. J. Electr. Power Energy Syst.*, vol/issue: 28(8), pp. 525–530, 2006.

[27] M. R. AlRashidi and K. M. EL-Naggar, "Long term electric load forecasting based on particle swarm optimization," *Appl. Energy*, vol/issue: 87(1), pp. 320–326, 2010.

[28] A. A. El Desouky and M. M. Elkateb, "Hybrid adaptive techniques for electric-load forecast using ANN and ARIMA," *Gener. Transm. Distrib. IEE Proceedings*, vol/issue: 147(4), pp. 213–217, 2000.

[29] H. Nie, *et al.*, "Hybrid of ARIMA and SVMs for short-term load forecasting," in *Energy Procedia*, vol/issue: 16(PART C), pp. 1455–1460, 2011.

[30] R. H. Liang and C. C. Cheng, "Combined regression-fuzzy approach for short-term load forecasting," *IEE Proc. - Gener. Transm. Distrib.*, vol/issue: 147(4), pp. 261, 2000.

[31] Suhartono, *et al.*, "Two-level seasonal model based on hybrid ARIMA-ANFIS for forecasting short-term electricity load in Indonesia," *ICSSBE 2012 - Proceedings, 2012 International Conference on Statistics in Science, Business and Engineering: "Empowering Decision Making with Statistical Sciences,"* pp. 634–638, 2012.

[32] D. Niu, *et al.*, "Combination forecasting model for mid-long term load based on least squares support vector machines and a mended particle swarm optimization algorithm," in *Proceedings - 2009 International Joint Conference on Bioinformatics, Systems Biology and Intelligent Computing, IJCBS 2009*, pp. 525–528, 2009.

[33] S. Sunaryo, *et al.*, "Double Seasonal Recurrent Neural Networks for Forecasting Short Term Electricity Load Demand in Indonesia," *Recurr. Neural Networks Temporal Data Process.*, 2011.

[34] E. Srinivas and A. Jain, "A Methodology for Short Term Load Forecasting Using Fuzzy Logic and Similarity," *Syst. Res.*, pp. 1–6, 2009.

[35] H. H. Cevic and M. Cunkas, "Short-term load forecasting using fuzzy logic and ANFIS," *Neural Comput. Appl.*, vol/issue: 26(6), pp. 1355–1367, 2015.

[36] P. Bunnoon, "Electricity Peak Load Demand using De-noising Wavelet Transform integrated with Neural Network Methods," *Int. J. Electr. Comput. Eng.*, vol/issue: 6(1), pp. 12, 2016.

[37] W. Wei, "Time Series Analysis: Univariate and Multivariate Methods," 2nd ed. New York: Pearson, Addison Wesley, 2006.

[38] G. E. P. Box, *et al.*, "Time series analysis: Forecasting and Control," 2008.

[39] J. D. Cryer and K. S. Chan, "Time Series Analysis: With Applications in R," 2008.

[40] N. Mohamecl, *et al.*, "Forecasting short term load demand using double seasonal arima model," *World Appl. Sci. J.*, vol/issue: 13(1), pp. 27–35, 2011.