Journal of Social Research

P-ISSN: 2827-9832
E-ISSN: 2828-335x
http://ijsr.internationaljournallabs.com/index.php/ijsr

Forecasting Household Electricity Consumption by Seasonal Autoregressive Integrated Moving Average (SARIMA) Method

Riana Wijayanti, Esther Ria Matulessy*, Nurhaida
Universitas Papua, Manokwari, Indonesia
Email: rianawijayanti20@gmail.com, e.matulessy@unipa.ac.id*, n.nurhaida@unipa.ac.id

ABSTRACT

Electricity is a vital energy source for households, with consumption patterns influenced by seasonal trends and external factors such as the COVID-19 pandemic. This study aims to forecast monthly household electricity consumption at *ULP Manokwari Kota* using the *Seasonal Autoregressive Integrated Moving Average* (*SARIMA*) method. The research utilizes secondary data from January 2013 to December 2022, which exhibits an upward trend and seasonal fluctuations. The *Box-Jenkins* methodology is employed, involving stationarity checks, model identification, parameter estimation, diagnostic testing, and forecasting. The dataset, spanning January 2013 to December 2022, demonstrates both an upward trend and a seasonal pattern. The forecasting process follows the *Box-Jenkins* approach: checking stationarity, identifying the model, estimating parameters, diagnosing the model, and performing forecasting. The optimal model for predicting electricity demand in the residential sector at *ULP Manokwari Kota* is SARIMA (1,1,0) *((0,0,1))^12*, with parameters indicating significant autoregressive and seasonal effects. Using this model, monthly electricity demand from January to December 2023 is forecasted. The lowest demand is projected for February 2023 (7,274,147 kWh), while the highest is in December 2023 (7,481,067 kWh). This research provides valuable insights for PT. PLN (Persero) *ULP Manokwari Kota* in planning electricity supply and ensuring system reliability, particularly in addressing seasonal demand variations. The study contributes to the literature by applying *SARIMA* to household electricity forecasting in a region with isolated power systems, highlighting its utility for energy management and policy formulation.

Keywords: *electricity, forecasting, SARIMA.*

INTRODUCTION

Electricity is a fundamental energy source that society relies on. Both state-owned and private offices, as well as industries, are highly dependent on electricity for their activities, making it nearly impossible to perform most tasks without it (Arnaz, 2018; Burke & Kurniawati, 2018; Koepke et al., 2023; Obeng-Darko, 2019). This is because, in general, tools or equipment—such as lighting, computers, printers, room temperature controllers, and information and communication tools (internet)—are powered by electricity. Similarly, many household appliances, including televisions, washing machines, irons, refrigerators, fans, and cooking utensils, also depend on electricity. This illustrates that society now considers electrical energy a basic necessity.

PT. PLN (Persero) is an Indonesian State-Owned Enterprise (BUMN) responsible for serving the public interest by providing electricity. In line with its corporate mandate, PLN manages the supply of electricity, which includes generating, distributing, planning, and developing electricity infrastructure. Based on the amount of electricity consumed, PLN divides its customers into five categories: households, businesses, industry, public service, and social service .

According to , there are 12 regencies and 1 municipality in West Papua Province with isolated electricity systems, consisting of seven 20 kV systems with loads exceeding 2 MW, namely Sorong, Fakfak, Manokwari, Kaimana, Teminabuan, and Bintuni. An isolated power system also supplies rural electricity to 56 districts, with a peak load of less than 2 MW. PLN's electrical energy sales over the last five years (2015–2019) averaged 405 GWh per year, while from 2011 to 2020, the average annual growth in electrical energy sales (in GWh) was 9.6%.

Based on electricity sales data at ULP Manokwari Kota from 2013 to 2022 obtained from PT. PLN (Persero) ULP Manokwari Kota, it is evident that household electricity usage fluctuates over certain periods and repeats annually. The highest electricity usage occurred in January 2021, amounting to 8,173,444 kWh. This increase was due to the rise in COVID-19 cases in 2021, which caused many people's activities to shift to the home. This situation is a concern for PLN, as it must continue to maintain the quality and reliability of electricity supply in the household sector . Therefore, it is necessary to forecast the need for electrical energy, especially in the household sector, for the coming years to ensure that supply meets demand. To meet the demand for electrical energy, an electricity development plan is needed as a guideline for future implementation . In other words, forecasting electricity demand greatly assists the government, particularly PT. PLN (Persero) ULP Manokwari Kota, in setting strategies for electricity supply.

In this article, the forecasting method used to predict household electricity usage in kWh is SARIMA (Seasonal Autoregressive Integrated Moving Average). SARIMA is an ARIMA model that has been modified to account for seasonal factors. Previous research on electricity forecasting using SARIMA has been conducted by several researchers. For example, Desvina forecasted household electricity consumption in Pekanbaru using SARIMA (0,1,1)(0,1,1)^12, where the model was deemed feasible because the residuals met the required assumptions. Sim also used SARIMA (0,1,1)(0,1,1)^12 to forecast energy consumption in Malaysia, achieving a MAPE of 8.4%, which is considered very good. Sosa , in research titled "Forecasting Electric Power Consumption Using the ARIMA Method Based on kWh of Energy Sold," obtained ARIMA (1,1,0)(0,1,1)^12 with a MAPE of 7.966%, rated as excellent.

Accurate forecasting of electricity consumption is critical for ensuring stable supply and efficient resource allocation. Previous studies have explored various time series models to predict electricity demand, with the Seasonal Autoregressive Integrated Moving Average (SARIMA) method emerging as a robust tool for capturing seasonal patterns and trends. For instance, Desvina et al. (2018) successfully applied SARIMA to forecast household electricity usage in Pekanbaru, while Sim et al. (2019) demonstrated its effectiveness in predicting energy consumption in Malaysia. Similarly, Sosa et al. (2020) utilized ARIMA models to forecast power consumption, highlighting the method's adaptability to different contexts. These studies underscore the versatility of SARIMA in addressing seasonal fluctuations, yet gaps remain in its application to regions with unique consumption patterns, such as isolated power systems.

Despite advancements in forecasting techniques, there is limited research on household electricity consumption in isolated regions like Manokwari Kota, where infrastructure and demand dynamics differ significantly from urban centers. Existing studies often focus on largescale grids or densely populated areas, leaving a gap in understanding how seasonal models perform in smaller, geographically constrained systems. This gap is particularly pressing given the region's reliance on localized power grids and the potential for disruptions due to external shocks, such as the COVID-19 pandemic, which caused unprecedented spikes in household electricity use. Addressing this gap is essential for improving energy planning and ensuring the reliability of supply in underserved areas.

The urgency of this research is underscored by the growing need for precise energy forecasts to support sustainable development and infrastructure planning. In Manokwari Kota, where electricity demand exhibits pronounced seasonal variations, inaccurate predictions could lead to either shortages or wasteful overproduction, both of which have economic and social consequences. The pandemic further highlighted the vulnerability of energy systems to sudden demand shifts, emphasizing the need for resilient forecasting tools. By developing a reliable model tailored to this region, stakeholders can optimize resource allocation, reduce costs, and enhance service delivery, ultimately contributing to energy security and community well-being.

This study introduces novelty by applying the SARIMA model to household electricity consumption in an isolated power system, a context rarely explored in previous research. Unlike broader regional studies, this work focuses on granular, monthly data from a specific locality, offering insights into how seasonal patterns manifest in smaller-scale grids. Additionally, the research incorporates recent data that reflect the impact of the pandemic, providing a contemporary perspective on demand fluctuations. The methodological rigor, including thorough stationarity checks and diagnostic testing, ensures the model's robustness, while the focus on practical applicability distinguishes it from purely theoretical approaches.

The primary objective of this research is to forecast household electricity consumption in Manokwari Kota using the SARIMA model, identifying seasonal trends and validating the model's accuracy. By achieving this, the study aims to equip PT. PLN (Persero) ULP Manokwari Kota with actionable insights for energy planning and infrastructure development. The benefits extend beyond immediate utility management, offering a template for similar regions facing seasonal demand challenges. Furthermore, the findings contribute to the broader discourse on energy forecasting, demonstrating the adaptability of SARIMA in diverse settings. Ultimately, this research supports sustainable energy practices and enhances the resilience of isolated power systems.

Considering that the time series plot of electricity usage in the household sector contains seasonality and trends, and that there has been no previous research related to forecasting the amount of kWh usage in the household sector using SARIMA, this research aims to apply the model to obtain the best fit and predict the electricity needs of households in Manokwari Kota.

METHOD

The Box-Jenkins methodology is employed, involving stationarity checks, model identification, parameter estimation, diagnostic testing, and forecasting. The data which spans from January 2013 to December 2022 has an upward trend and a seasonal pattern. Our research data is secondary data obtained from PT. PLN (Persero) ULP Manokwari Kota. The data is a report on results of the household sector's monthly electricity usage at ULP Manokwari Kota for 10 years, from January 2013 to December 2022. The research method is Seasonal Autoregressive Integrated Moving Average (SARIMA) following Box-Jenkis methodology. Data processing was carried out using the Microsoft Office Excel to input data. Meanwhile, the forecasting process uses the R Studio version 1.4.1106. The R Studio is used to simplify calculations. Several RStudio packages used for the forecasting process, i.e., library(readxl), library(astsa), library(FitAR), library(tseries), library(urca), library(forecast), library(sarima), library(nortest), and library(stats) (R Team, 2020).

The forecasting stages are as follows:

1. Checking stationarity

   Stationarity of series can be checked either visually or theoretically. Time series plot shows these two properties, stationarity in mean and stationarity in variance. A series is said to be mean stationary if the curve lies around a horizontal line. In other words, the trend is flat. Furthermore a variance stationarity series can be shown as a plot where the curve fluctuates uniformly around the trend. Theoretically, non-stationarity in the variance can be overcome by carrying out Box-Cox transformation. In the Box-Cox transformation process, the first step is to estimate a parametera λ. Let $Z_t, t > 0$ be a time series . The general equation for the Box-Cox transformation is as follows:

$$(Z_t) = \{\frac{{Z_t}^{\lambda} - 1}{\lambda} \quad ; \ \lambda \neq 0 \ \ ln\, ln\, Z_t \quad ; \ \lambda = 0 \tag{1}$$

   When λ equals 1, the series is stationary in variance. Meanwhile, non-stationarity in mean can be treated by applying a differentiation process on original data. Given our data is called **data**, R-code for the differencing process is '**diff(data)**'. Theoretically, a differencing process in the first order is a difference between the t-th and $(t-1)$-th data. Notating the first differenced series by ΔZ_t, the differencing process in the first order is written as:

$$\Delta Z_t = Z_t - Z_{t-1} \tag{2}$$

   Checking stationarity either in variance and in mean can be theoretically done by conducting a statistic test, Augmented Dickey-Fuller (ADF) test. The R-code of the test are **'adf.test(data) and ur.df(data difference, lags=1, type="trend") '.** The test is used to test the following hypothesis(Enders, 2008):

   H_0 : $\delta = 0$ (Contains a unit root or is not stationary)

   H_1 : $\delta \neq 0$ (Does not contain unit roots or stationary)

   and the test statistics:

$$\tau = \frac{\hat{\delta}}{SE(\hat{\delta})} \tag{3}$$

   where $SE(\hat{\delta})$ is standard error of the least squares estimate of $\hat{\delta}$ and $\hat{\delta} = -(1 - \sum_{i=2}^{p} \quad \beta_i)$. The rejection criteria is to reject H_0 if $p\ value < \alpha$ or $|\tau| > |M|$, where τ is the statistical test and *M* is the Mackinnon's critical value. It needs to mention that we use $\alpha = 0.05$ in all statistical tests.

2. Model Identification

   In this step, a non-seasonal ARIMA and a seasonal ARMA model on differentiated series are identified using ACF and PACF plots. Theoretically, the ACF and PACF plot patterns of non-seasonal orders ($p\ and\ q$) are explained in **Table 1**.

Table 1. Characteristics of theoretical ACF and PACF for stationary processes (Wei, 2006)

Process	ACF	PACF
AR $(\boldsymbol{p})$	Tails off as an exponential decay or a damped sine wave	Cut of after lag p
MA $(\boldsymbol{q})$	Cutt of after lag q	Tails off as an exponential decay or a damped sine wave
ARMA $(\boldsymbol{p}, \boldsymbol{q})$	Tails off after lag $(q-p)$	Tails off after lag $(p-q)$

The seasonal part of an AR or MA model will be seen in seasonal lags (S) of the PACF and ACF. However, before determining seasonal orders, we need to first check the stationarity of the seasonal lags using the ADF test. The R code used to check stationarity in seasonal lags is **adf.test(data difference, k=S)** and **ur.df(data difference, type="trend", lags =12)**, if it reaches stationary then the value of D=0, but if it is not stationary then the differential is carried out on the seasonal lags. The next step is then to proceed to determining some tentative models.

3. Parameter Estimation

The most commonly used parameter estimation is the Conditional Least Square (CLS) model. And The next step is to test the significance of the parameters using a hypothesis. Based on the SARIMA model that has been obtained, then significancy of the parameters are carried out under the following hypotheses:

H_0 : $\emptyset_p = 0$ or $\theta_q = 0$ (parameters are not significant)

H_1 : $\emptyset_p \neq 0$ or $\emptyset_q \neq 0$ (parameters are significant)

Test statistic:

$$t = \frac{\hat{\emptyset}_p}{SE(\hat{\emptyset}_p)} \text{ or } t = \frac{\hat{\theta}_q}{SE(\hat{\theta}_q)} \quad (4)$$

Rejection Criteria:

reject H_0 if $|t| > t_{\frac{\alpha}{2};n-p}$ where n is the number of data and p is the number of parameters, or a similar way is to compare the p-value and the significance level α, namely reject H_0 if p-value $< \alpha$. After carrying out significance test, hopefully we are able to obtained a model that meets the parameter significance test. Given p, d, q, P, D, Q , and S are the SARIMA orders obtained in the previous step, estimating parameters in R done by **'sarima(datadifference, p,d,q,P,D,Q,S, no.constant=TRUE)'.**

4. Model Diagnostic

Diagnostic of models is carried out with the aim of checking white noise and normality on residuals. white noise residuals can be seen visually on its ACF/PACF plot which aims to see whether there is a serial correlation in the residuals of the observed model. A normal data quantile vs theoretical quantile plot (Q-Q lot), on the other hand, can give a visual plot of normality on the residuals. The residual is assumed to be normal if the series spreads around the diagonal line. Moreover, these two properties can be theoretically checked by Ljung-Box test and Kolmogorov-Smirnov test.

Under Ljung-Box test, a white noise assumption is tested by the following hypothesis:

H_0 : $\rho_{a_1} = \rho_{a_2} = \cdots = \rho_{a_k} = 0$ (residual meets white noise)
H_1 : $a_{i;i=1,\dots,k} \neq 0$ (residual does not meet the white noise)
Test statistics:
$Q = n(n+2)\sum_{k=1}^{n} \quad (n-k)^{-1}\hat{\rho}_{a_k}{}^2$
Rejection criteria:
reject H_0 if $Q > \chi^2{}_{(K-p-q)}$ or the $p-value < \alpha$, where n is the number of data, and $\hat{\rho}_{a_k}$ is an estimated autocorrelation of the kth lag of residuals.
Rcode for white noise assumption used is '**res<-residuals(model$fit)**; **Box.test(res,12,"Ljung-Box"), Box.test(res,24,"Ljung-Box"); Box.test(res,36,"Ljung-Box"), Box.test(res,48,"Ljung-Box")'.**

The next step is then carried out a residual normal distribution test. After visually see the normal Q-Q plot of residuals, the investigation can be explained further through the Kolmogrov-Smirnov test with hypothesis:
H_0 : $F(a_t) = F_0(a_t)$ (residuals are normally distributed)
H_1 : $F(a_t) \neq F_0(a_t)$ (residuals are not normally distributed)
Test statistic:

$$D = xsup|S(a_t) - F_0(a_t)| \tag{5}$$

Rejection criteria:
Reject H_0 if the value $D > D_{(1-\alpha);n}\ or\ p-value < \alpha$ where n is the number of data.
R.code for checking normality in residuals : **'lillie.test(res)'**

5. Forecasting

Based on historical data up to time t, namely $Z_{1.}Z_2, \dots Z_{t-1}, Z_t$, we will predict the value of Z_{t+l} which will occur *l* units of time in the future. For example, given time *t* is called the start of the forecast (forecast origin) and *l* is the forecast period (lead time), the forecast value is denoted as Z_{t+l}. The minimum mean square error estimates $\hat{Z}_t(l)$ and Z_{t+l} in the initial forecast *t* are given the following conditional expectations:

$$\hat{Z}_t(l) = E(Z_1, Z_2, \dots Z_t) \tag{6}$$

As in all statistical endeavors, apart from predicting or predicting the unknown Z_{t+l}, an accurate prediction for deterministic models with a white noise stochastic component $\{Z_t\}$ i.e.

$$\hat{Z}_t(l) = \mu_{t+l} \tag{7}$$

and

$$Var\big(a_t(l)\big) = Var(Z_{t+l}) \tag{8}$$

If the stochastic component is assumed to be normally distributed, then the forecast error

$$a_t(l) = Z_{t+l} - \hat{Z}_t(l) \tag{9}$$

So for a given confidence level $(1-\alpha)$, one can use a standard normal percentile, $Z_{1-\frac{\alpha}{2}}$ to set

$$P\left[-z_{1-\frac{\alpha}{2}} < \frac{Z_{t+l}-\hat{Z}_t(l)}{\sqrt{Var\big(a_t(l)\big)}} = \frac{Z_{t+l}-\hat{Z}_t(l)}{SE\big(a_t(l)\big)} < z_{1-\frac{\alpha}{2}}\right] = 1-\alpha. \tag{10}$$

or equivalent to $P\left[\hat{Z}_t(l) - z_{1-\frac{\alpha}{2}}\sqrt{Var\big(a_t(l)\big)} < Z_{t+l} < \hat{Z}_t(l) + z_{1-\frac{\alpha}{2}}\sqrt{Var\big(a_t(l)\big)}\right] = 1 - \alpha.$

Thus $(1-\alpha)100\%$ future observation interval Z_{t+l} will have a prediction bound as

$$\hat{Z}_t(l) \pm z_{1-\frac{\alpha}{2}}\sqrt{Var\big(a_t(l)\big)} \tag{11}$$

where $Z_{1-\frac{\alpha}{2}}$ is chosen to get the desired level of confidence. For example, if the process is Gaussian, then choosing $Z_{1-\frac{\alpha}{2}} = 2$ will produce a prediction of approximately 95% of the intervals for $\hat{Z}_t(l)$ (Shumway & Stoffer, 2016). In obtaining the forecast, we used the following series of R.code : **'sarima.for(datadifferenced,p,d,q,P,D,Q,S, no.constant = TRUE)'.**

R.code of prediction limits of 95% confidence interval: '**Upper=forecast$pred + 2*forecast$se ;Lower=forecast$pred - 2*forecast$se**', and

R.code of prediction limits 80% confidence interval: '**Upper= forecast$pred + 1*forecast$se ;Lower=forecast$pred - 1*forecast$se**'.

RESULTS AND DISCUSSION

Checking Stasionary

The following plot shows a data pattern on the number of kWh usage in the household sector at PT. PLN Persero ULP Manokwari Kota.

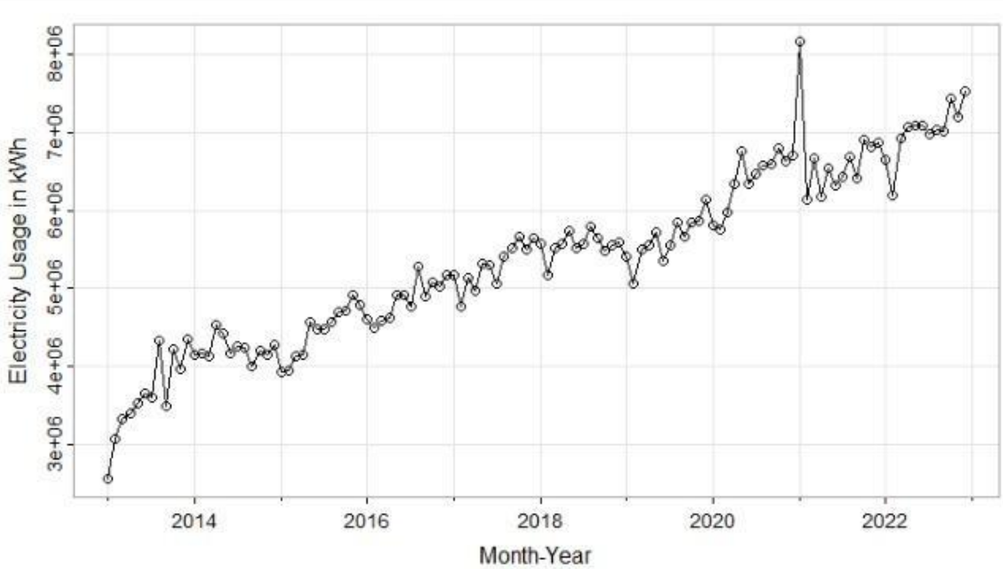


Figure 1. Time Series Plot Total Electricity Usage kWh in Household Sector at PT. PLN Persero ULP Manokwari Kota

Figure 1 shows that visually the data has a seasonal pattern as seen from the data plot which increases and decreases over each annual period, and there is also an upward trend or the data continues to experience an increase. Seasonal patterns can be detected from repeating patterns and the data will show ups and downs within a fixed period of time. The data plot shows a seasonal effect, which is denoted by a pattern of strong decreases and increases that repeat over time. **Figure 1** shows that the amount of electricity used by the household sector from 2013 to 2022 experienced decreases and increases over a certain period of time and repeated every year. This is marked by a decrease every January and February from 2017 to 2022, also marked by an increase at the end of each year where the increase occurs in the period October to December from 2013 to 2022. In January 2021 electricity usage increased sharply of 8.173.444 kWh due to the large number of community activities that have shifted to work from home suspected due to the increasing cases of the Covid-19 pandemic.

The next step is to identify the stationarity of the data, both in the mean and in the variance. If the data is not yet stationary, the data will be stationary first.

1. Stationary in variance

At this stage, the first thing to do is to estimate the lambda parameter (λ), with the lambda value obtained being 0.999959. Because the value of λ is close to 1, it means that the data is stationary in variance so it does not need to be transformed.

2. Stationary in mean

The next step after stationary data in the variance is to have the data stationary relative to the mean, because in **Figure 1** it shows that the data is still not stationary with respect to the mean and does not fluctuate around a line parallel to the time axis, so differentiation needs to be carried out so that the data can be used for forecasting. The following is a time series plot of electricity consumption (kWh) after the first differentiation.

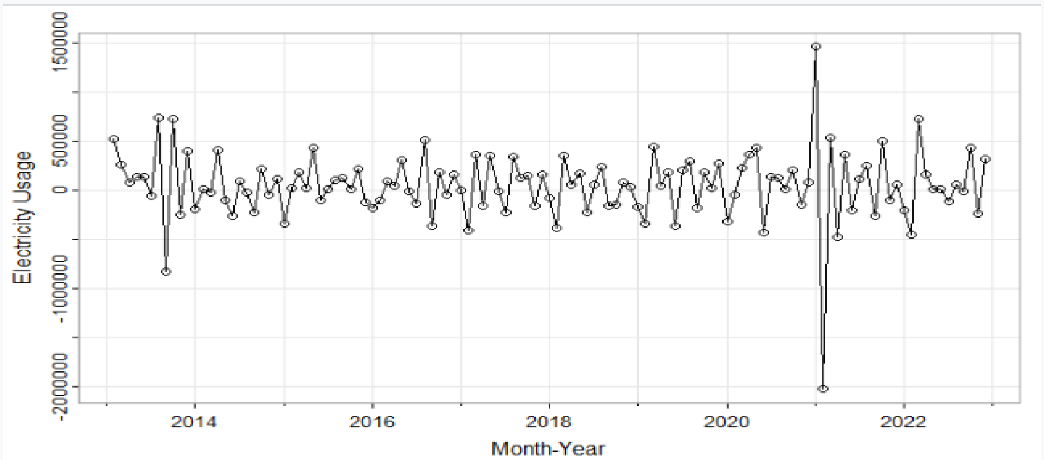
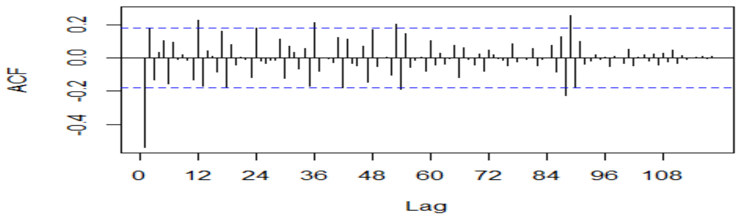


Figure 2. Time series plot of differencing data

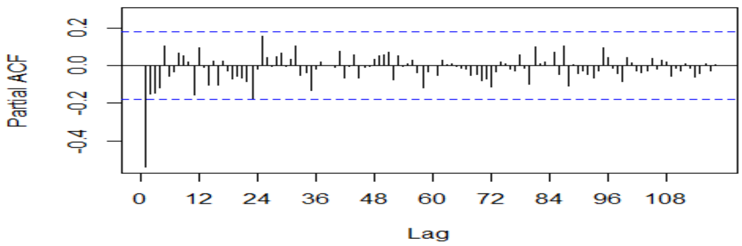
After going through the differencing process, the results of the time series plot in **Figure 2** show that the data is stationary with respect to the mean and there is no longer any trend in it. Using ADF test result, we obtained statistical test of $\tau = -6.0559$ which is less than the critical value of -3.99, -3.43, and -3.13 at 1%, 5%, and 10% significance level, respectively. Also, p-value $= 0.01$ meaning p-value $< \alpha$. So, H_0 is rejected, it means that the series has no unit root and is statistically stationary.

Identify the Tentative Model

Analyzing differenced series on its ACF and PACF plots is shown in **Figure 3**.



(a)



(b)

Figure 3. Plot of ACF and PACF of Differenced Series

Based on **Figure 3**, with $d = 1$, identification of non-seasonal orders is explained in **Table 2**.

Table 2. Identification of Non-Seasonal Orders (p and q) in Figure 3

Identification	Non-seasonal order
From the ACF plot in Figure 3. It can be seen that a cuts off occurs after the 1st lag. Then, the PACF plot is considered tails off	MA(1)
PACF plot in Figure 3. It can be seen that a cuts off occurs after the 1st lag. Then, the ACF Plot is considered tails off	AR(1)
Combined ACF and PACF identification results	ARMA(1,1)

Based on **Figure 3**, which is an ACF plot of differencing data, it indicates that the data has seasonal characteristics because there is a significant ACF value at lag 12 so that the S value = 12. Then again the seasonality of 12 is tested for stationarity by carrying out the ADF test on lag multiples of 12. The results of the ADF test on lag multiples of 12 have a p-value of 0.03923 which is smaller than $\alpha = 0.05$, which means the seasonal lag has stationary, and no differentiation is needed, the $D = 0$. In **Figure 3** it can be seen that the ACF plot is truncated after 24 lags, so that the order $\mathrm{MA}(1)^{12}$ atauiMA$(2)^{12}$ is obtained for seasonal models.

Based on **Table 2**, identifying the non-seasonal order and the order for the seasonal model, the Tentative model SARIMA $(0,1,1)(0,0,1)^{12}$, SARIMA $(0,1,1)(0,0,2)^{12}$, SARIMA $(1,1,0)(0,0,2)^{12}$, SARIMA $(1,1,0)(0,0,1)^{12}$, SARIMA $(1,1,1)(0,0,2)^{12}$, and SARIMA$(1,1,1)(0,0,1)^{12}$.

Parameter Estimation and Significance Test

Summary Significant tests for several tentative models can be seen in **Table 3.**

Table 3. Model Parameter Estimation

Model	Type	Parameter Estimated Value	Nilai $\lvert t \rvert$	*p-value*	Significance	AIC
SARIMA $(0,1,1)(0,0,1)^{12}$	MA 1	-0,5245	-7,4715	0,0000	Significant	28.1014
	SMA 12	0,1684	1,9446	0,0541	Not Significant	
SARIMA $(0,1,1)(0,0,2)^{12}$	MA 1	-0,5606	-8,0546	0,0000	Significant	28.0614
	SMA 12	0,0781	0,7354	0,4635	Not Significant	
	SMA 24	0,3190	2,5459	0,0122	Significant	
SARIMA $(1,1,0)(0,0,1)^{12}$	AR 1	-0,5204	-6,5873	0,0000	Significant	28.0911
	SMA 12	0,1813	2,0815	0,0395	Significant	
SARIMA $(1,1,0)(0,0,2)^{12}$	AR 1	-0,5396	-6,9003	0,0000	Significant	28.0602
	SMA 12	0,0999	0,9416	0,3483	Not Significant	
	SMA 24	0,2887	2,3414	0,0209	Significant	
SARIMA $(1,1,1)(0,0,1)^{12}$	AR 1	-0,3103	-1,8264	0,0703	Not Significant	28.0909
	MA 1	-0,2931	-1,6402	0,1036	Not Significant	
	SMA 12	0,1645	1,8995	0,0599	Not Significant	
SARIMA $(1,1,1)(0,0,2)^{12}$	AR 1	-0,3044	-1,9948	0,0484	Significant	28.0472
	MA 1	-0,3506	-2,2474	0,0265	Significant	
	SMA 12	0,0633	0,5994	0,5500	Not Significant	
	SMA 24	0,3322	2,6219	0,0099	Significant	

Table 3 shows that from the results of estimating model parameters, only one model with significant parameters (p-value < 0.05). Because only one model satisfies the significant parameter test, the AIC criteria are not needed to determine the best model. Based on **TABLE 3**, the model that satisfies the significant parameter test or rejects H_0 is the SARIMA $(1,1,0)(0,0,1)^{12}$ model, so it is suitable for use in the next stage, namely checking the residual assumptions.

Diagnostic Model

Diagnostic model using graphics from the data can be seen in **Figure 4**.

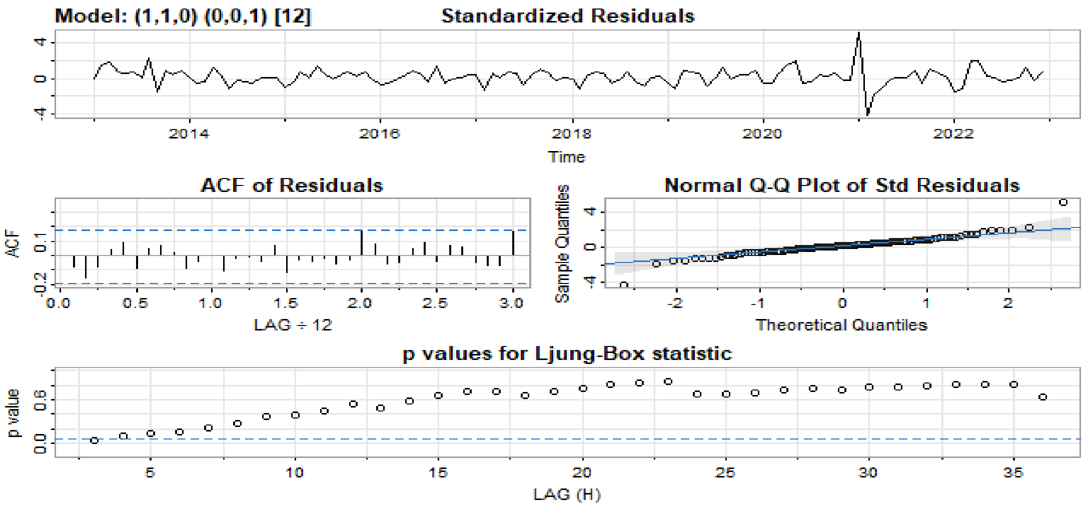


Figure 4. Diagnostic Plots SARIMA $(1, 1, 0)(0, 0, 1)^{12}$

It can be seen from the model diagnostic results in **Figure 4**, the SARIMA $(1, 1, 0)(0, 0, 1)^{12}$ model is a good model, from the ACF plot it can be seen that the residual is already a white noise model, indicated by the absence of lag (≥ 1) outside the significance band, also the p-value of the Ljung-Box statistic is also above the 5% significance level line, which indicates that the null hypothesis of residuals not containing serial correlation is accepted. This can be further explained by residual testing using the Ljung-Box Test.

a. White noise assumption

A summary of the results of the Ljung - Box test on lines 12, 24, 36, and 48 of the Rstudio output can be seen in **Table 4**:

Table 4. *p-value* of Ljung-Box Test

	lag	*p-value*	Independency of Residuals
SARIMA $(\mathbf{1}, \mathbf{1}, \mathbf{0})(\mathbf{0}, \mathbf{0}, \mathbf{1})^{\mathbf{12}}$	12	0,7175	independence
	24	0,773	independence
	36	0,7341	independence
	48	0,7954	independence

Based on **Table 4**, the SARIMA model $(1, 1, 0)(0, 0, 1)^{12}$ fulfills the assumption of residual independence at lags 12, 24, 36, and 48. It can be concluded that it failed to reject H_0 because the p-value of Ljung-Box is more than α (α=0.05).

b. Normal Distribution Assumption

The next stage is after the assumption of white noise is met, then we do the residual normal distribution test. Checking the normality of the residuals visually is shown through the residual probability plot shown in **figure 4**, especially in the QQ-norm Standard residuals plot, it can be seen that most of the residuals spread around the diagonal line and follow the direction of the diagonal line, but there are two points that spread far from the diagonal line and causes the residuals to not have a normal distribution. This can be further explained through the Kolmogrov-Smirnov test. A summary of the results of the Kolmogrov-Smirnov test can be seen in **table 5**.

Table 5. Kolmogrov-Smirnov test

Kolmogrov-Smirnov test	
D	0,10802
p-value	0,001541

Based on **table 5**, the SARIMA model $(1, 1, 0)(0, 0, 1)^{12}$ does not meet the residual normality assumption because it rejects H_0 or $p - value < \alpha$ $(\alpha = 0.05)$, it can be concluded that the residuals are not normally distributed. This can be ignored and continued at the next stage, because the residual normality test is not as important as the white noise test (Rosadi, 2014).

Forecasting

The following are the results of forecasting the electrical energy needs of the household sector for the 12 future periods shown in **Table 3**, using the SARIMA model $(1,1,0)(0,0,1)^{12}$ with a value of $\emptyset_1$ =-0,5204, $\Theta_1 = 0{,}1813$. The form of the SARIMA $(1,1,0)(0,0,1)^{12}$ model is expanded as follows:

$$\emptyset_p(B)\Phi_P(B)^S(1-B)^d(1-B^S)^D\dot{Z}_t = \theta_q(B)\Theta_Q(B^S)a_t$$
$$\emptyset_1(B)(\Phi_0)(B)^{12}(1-B)^1(1-B^{12})^0\dot{Z}_t = \theta_0(B)\Theta_1(B^{12})a_t$$
$$\emptyset_1(B)(1-B)^1\dot{Z}_t = \Theta_1(B^{12})a_t$$
$$(1-\emptyset_1 B)(1-B)\dot{Z}_t = (1-\Theta_1 B^{12})a_t$$
$$(1-B-\emptyset_1 B+\emptyset_1 B^2)\dot{Z}_t = (1-\Theta_1 B^{12})a_t$$
$$\dot{Z}_t - B\dot{Z}_t - \emptyset_1 B\dot{Z}_t + \emptyset_1 B^2\dot{Z}_t = a_t - \Theta_1 B^{12}a_t$$
$$\dot{Z}_t - \dot{Z}_{t-1} - \emptyset_1\dot{Z}_{t-1} + \emptyset_1\dot{Z}_{t-2} = a_t - \Theta_1 a_{t-12}$$
$$\dot{Z}_t = \dot{Z}_{t-1} + \emptyset_1\dot{Z}_{t-1} - \emptyset_1\dot{Z}_{t-2} + a_t - \Theta_1 a_{t-12}$$
$$\dot{Z}_t = (1+\emptyset_1)\dot{Z}_{t-1} - \emptyset_1\dot{Z}_{t-2} + a_t + \Theta_1 a_{t-12}$$
$$\dot{Z}_t = (1+0{,}5204)\dot{Z}_{t-1} + 0{,}5204\dot{Z}_{t-2} + a_t + 0{,}1813a_{t-12}$$

With $a_t \sim N(0,\sigma_a), {\sigma_a}^2 = 8.766 \times 10^{10}$ and $\dot{Z}_t = Z_t - Z_{t-1}$.

Table 6. Forecasts. The demand for Electrical Energy in the ULP Manokwari Kota Household Sector January 2023- December 2023

Month	Prediction Results	Lo 80	Hi 80	Lo 95	Hi 95
January 2023	7.293.715	6.996.044	7.591.386	6.698.373	7.889.057
February 2023	7.274.147	6.944.539	7.603.756	6.614.930	7.933.364
March 2023	7.382.270	7.023.556	7.740.983	6.664.843	8.099.697
april 2023	7.424.495	7.038.867	7.810.123	6.653.239	8.195.751
may 2023	7.417.191	7.006.408	7.827.974	6.595.625	8.238.757
June 2023	7.422.706	6.988.222	7.857.189	6.553.738	8.291.673
July 2023	7.400.869	6.943.912	7.857.826	6.486.955	8.314.783
August 2023	7.402.529	6.924.153	7.880.905	6.445.777	8.359.280
september 2023	7.408.057	6.909.181	7.906.933	6.410.306	8.405.808
October 2023	7.466.918	6.948.352	7.985.484	6.429.786	8.504.049
November 2023	7.428.624	6.891.088	7.966.159	6.353.553	8.503.694
december 2023	7.481.067	6.925.209	8.036.924	6.369.351	8.592.782

where,

Hi 80 means Upper bound with 80% confidence interval;
Lo 80 means Lower bound with 80% confidence interval;
Hi 95 means Upper bound with 95% confidence interval;
Lo 95 means Lower bound with 95% confidence interval.

Based on **Table 6**, The lowest demand occurred in February 2023 and the highest is in December 2023, i.e., 7.274.147 kWh and 7.481.067 kWh, respectively. The plot of forecasting the energy needs of the electricity for the months January 2023 to December 2023 can be seen in **Figure 5.**

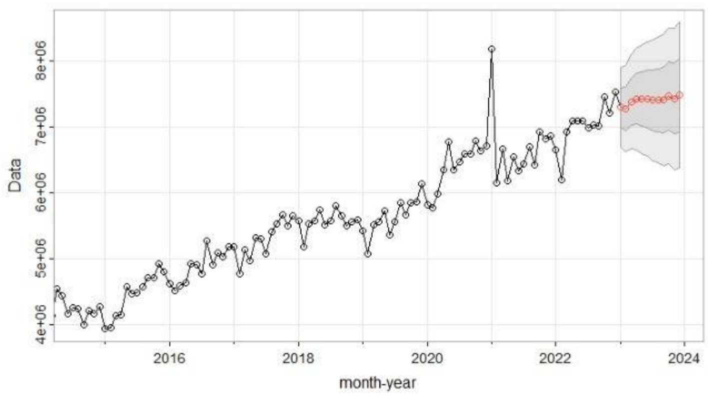


Figure 5. Forecasting Results of Electrical Energy Needs for the ULP Manokawari City Household Sector, January 2023 – December 2023

Figure 5 shows that the results of forecasting the electricity demand for the household sector at PT. PLN (Persero) Manokwari Kota ULP for January 2023 to December 2023 is in the upper and lower bound intervals, with the 80% confidence interval in light gray and the 95% confidence interval in dark grey.

CONCLUSION

Based on the application of the *Seasonal Autoregressive Integrated Moving Average* (*SARIMA*) method, the study identified SARIMA (1,1,0)(0,0,1)1212 as the most effective model for forecasting household electricity demand in ULP Manokwari Kota, accurately capturing both seasonal and trend components. The model forecasts that electricity demand in 2023 will fluctuate, with the lowest consumption anticipated in February at 7,274,147 kWh and the highest in December at 7,481,067 kWh, reflecting the influence of seasonal factors over the 12-month period. For future research, it is recommended to explore the integration of exogenous variables—such as weather patterns, economic indicators, or policy changes—into the forecasting model to further enhance prediction accuracy and provide deeper insights into the drivers of electricity demand in isolated power systems.

REFERENCES

Arnaz, M. Z. (2018). The aggregate and residential electricity demand in Indonesia. *Jurnal Ekonomi Pembangunan*, *26*(1), 1–20. https://doi.org/10.14203/jep.26.1.2018.1-20

Burke, P. J., & Kurniawati, S. (2018). Electricity subsidy reform in Indonesia: Demand-side effects on electricity use. *Energy Policy*, *116*, 410–421. https://doi.org/10.1016/j.enpol.2018.02.018

Desvina, A. P., A'Zima, R., Susanti, E., Yendra, R., & Rahmadeni. (2018). Prediksi jumlah pemakaian kWh sektor rumah tangga di Pekanbaru menggunakan model SARIMA (0,1,1)((0,1,1)^2). *Jurnal Statistika*, *6*(2), 697–704.

Enders, W. (2008). *Applied econometric time series* (3rd ed.). John Wiley & Sons.

Gardapapua. (2021, February 17). *Covid-19 belum pergi*. Gardapapua. https://gardapapua.com

Hantje, P. (2018). *Dasar teknik listrik*. CV. Budi Utomo.

Kementerian Energi dan Sumber Daya Mineral Republik Indonesia. (2021). *Rencana usaha penyediaan tenaga listrik (RUPTL) PT PLN (Persero) 2021–2030*.

Koepke, M., Monstadt, J., & Pilo', F. (2023). Urban electricity governance and the (re)production of heterogeneous electricity constellations in Dar es Salaam. *Energy, Sustainability and Society*, *13*(1), Article 10. https://doi.org/10.1186/s13705-023-00401-8

Obeng-Darko, N. A. (2019). Why Ghana will not achieve its renewable energy target for electricity: Policy, legal and regulatory implications. *Energy Policy*, *128*, 75–83. https://doi.org/10.1016/j.enpol.2018.12.059

Perusahaan. (2019). *Laporan keberlanjutan*. Jakarta.

R Core Team. (2020). *RStudio: Integrated development for R*. RStudio, PBC.

Rosadi, D. (2014). *Analisis runtun waktu dan aplikasinya dengan R*. Gadjah Mada University Press.

Shumway, R. H., & Stoffer, D. S. (2016). *Time series analysis and its applications with R examples* (4th ed.). Springer.

Sim, S. E., Tay, K. G., Huong, A., & King, T. W. (2019). Forecasting electricity consumption using SARIMA. *International Journal of Electrical and Computer Engineering*, *6*(5), 103–114.

Sosa, G. R., Falah, M. Z., L, D. F., Wibawa, A. P., Handayani, A. N., & Hammad, J. A. H. (2020). Forecasting electrical power consumption using ARIMA method based on kWh of sold energy. *Science in Information Technology Letters*, *1*(1), 9–15.

Syafriwel, S. (2015). Analisis ramalan kebutuhan energi listrik di regional Sumatera Utara tahun 2015–2019 dengan metode gabungan. *Jurnal Teknovasi*, *2*(2), 53–62.

Wei, W. S. (2006). *Time series analysis: Univariate and multivariate methods* (2nd ed.). Pearson Education.