Journal of Mechatronics and Artificial Intelligence

Homepage: http://ejournal.upi.edu/index.php/jmai/

Forecasting Electrical Energy Loads at PT Krakatau Daya Electric Using the Linear Regression Method

Krisna Bayu[1], Dhea Rahmalia Henidar[2], Fahmi Hermastiandi[3], Galih Prasetya N[4], Adi Nugraha[5,*]

[1,2,3,4,5] *Electrical Engineering, Sultan Ageng Tirtayasa University, Indonesia*
**Correspondence author email: adi.n@untirta.ac.id*

ARTICLE INFO

Article History:

Submitted/Received 25 May 2024
First Revised 30 May 2024
Accepted 31 May 2024
Publication Date 02 Jun 2024

Keywords:
Electrical energy,
Forecasting,
Linear regression

ABSTRACT

The importance of the role of electrical energy at this time cannot be denied and it is difficult to imagine how life would be without electricity, not only as a source of light at night in Cilegon City because it is rich in resources, especially in the industrial sector. Therefore, the existence of a guaranteed power supply is very important. PT Krakatau Daya Listrik, as the main provider and distributor of electrical energy in the KIEC Area (Krakatau Industrial Estate Cilegon), indirectly becomes the backbone for the economy of the people in the trading area of PT Krakatau Daya Listrik. The method used in making predictions is the linear regression method which is a method to test how accurate the relationship between x and y is. In addition, to do forecasting or similarity testing, use Google Colab. The results of the two show a correlation coefficient of 0.4 which is enough to have a relationship between x and y, the more years the more power or electrical energy is needed. This is very relevant considering that electrical energy has become a necessity, so this forecast can help electricity service providers meet consumer needs.

1. Introduction

The important role of electrical energy today cannot be denied and it is difficult to imagine what life would be like without electricity, not only as a source of light at night in our cities. Cilegon City is rich in resources, especially in the industrial sector. Responsibility for managing energy resources, including electrical energy, lies with local governments. The development of an area in the industrial, commercial ,and public service sectors is influenced by the availability of electrical energy, and even has an impact on the economy and community welfare [1,2]. Therefore, the existence of a guaranteed electricity supply is very important. PT Krakatau Daya Electric, as the main provider and distributor of electrical energy in the KIEC (Krakatau Industrial Estate Cilegon) area, indirectly becomes the backbone of the economy of the community in the PT Krakatau Daya Electric buying and selling area.

There are various types of electricity customers distributed to the community, such as households, offices, and industry. In planning the development and expansion of electricity systems, predicting electricity needs is the first step [3]. The next step is to determine the required capacity, type of electrical energy generation, cost allocation, and predict income from sales of the electricity produced. If the prediction is too high, the electricity supplier may experience losses because the costs incurred are too high. On the other hand, if the prediction is too low, the electricity supply will be less, disrupting economic conditions in the area. Therefore, predicting electrical energy needs is very important in companies operating in the electrical energy industry [4].

With advances in technology, problems often occur in terms of providing electrical energy, especially in terms of electrical energy needs. The demand for electrical energy continues to increase along with the increase in the number of consumers, business growth, and the industrial sector [5]. Therefore, it is necessary to forecast electrical energy for the future. Current use of electrical energy is very high and is not only limited to lighting purposes but is also important in supporting economic activities. In the future, demand for electrical energy is expected to continue to increase along with population growth, increased investment, and technological developments, including in the industrial world. Increasing technology in the industrial sector also means increasing the amount of production and increasing connected power to improve production services. However, currently, there is an imbalance between the increase in electrical energy demand and the increase in electrical energy supply, where the available power capacity remains the same, while demand continues to increase and burdens the existing energy supply [6].

Forecasting has an important role in predicting the electrical power requirements needed to meet the load and energy requirements in electricity distribution. Apart from technical considerations, economic factors are also crucial aspects that must be taken into account. If the estimated power requirements are inaccurate, it can result in a lack of available power capacity to meet the existing load. Internally, forecasting aims to accommodate the addition of new loads, while externally, forecasting aims to ensure the system remains stable when new loads are added. Apart from that, the central government also encourages local governments to increase the use of alternative energy in order to achieve national energy supply security. By considering geographical conditions, electricity load predictions can help electricity providers carry out accurate planning and approach a high level of accuracy for maintenance purposes [7].

In recent years, the use of linear regression methods for forecasting electrical energy loads has been supported by advances in computing technology that enable larger and faster data processing [8]. In addition, with developments in the fields of data analytics and artificial intelligence, electrical energy load forecasting models using linear regression methods can be improved through the use of more sophisticated techniques and integration with other factors that have the potential to influence electrical energy loads [8]. Simple Linear Regression is a statistical method used to test the extent of the causal relationship between causal factor variables (X) and effect variables (Y) [9]. The causal factor variable is generally denoted by X or also called a predictor, while the effect variable is symbolized by Y or also called a response. Simple Linear Regression, or abbreviated as SLR (Simple Linear Regression), is also a statistical method used in production to forecast or predict quality or quantity characteristics [10]. Therefore, from this it was compiled.

2. Methods

Research methodology is a systematic approach used to design, conduct, and analyze research. This method helps collect relevant data, analyze information, and make conclusions based on existing evidence. The following flow diagram illustrates the process of forecasting electrical energy loads at PT Krakatau Daya Electric using the linear regression method. Forecasting electrical energy loads is very important in planning and managing efficient electrical energy supplies. In this flow diagram, the steps required to forecast electrical energy loads in a systematic and structured manner will be explained.

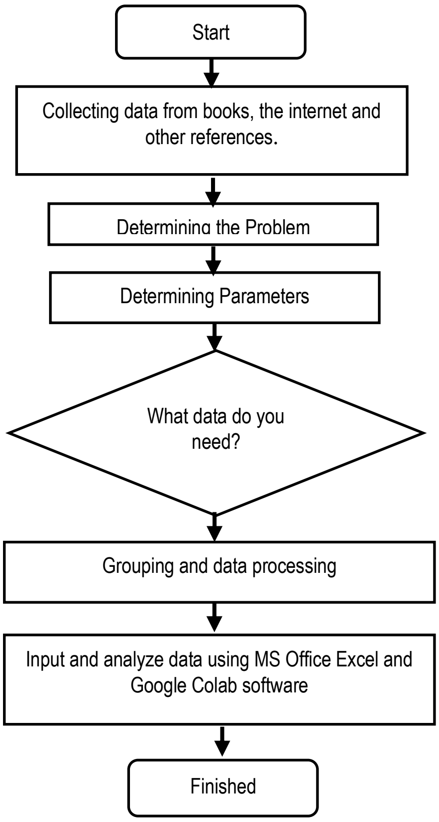

Figure 1: Research flow chart.

Table 1: Power connection data.

Years	Industry	Business	Social	Household	Total
2018	1,497,347.80	19,434.60	2,065.90	1,627.60	1,520,475.90
2019	1,547,422.10	19,574.30	2,616.90	1,318.70	1,570,932.00
2020	1,556,608.20	23,059.50	2,616.90	1,545.50	1,583,830.10
2021	1,591,930.90	19,545.90	2,616.80	1,259.40	1,615,353.00
2022	1,542,514.40	19,715.00	2,376.30	1,287.40	1,565,893.10

Table 2: Linear regression variables.

Years	Variable x: Year Period	Variable y: Connection Data	x^2	x.y	y^2
2018	1	1,520,475.90	1	1520475.9	2,311,846,942,480.81
2019	2	1,570,932.00	4	3141864	2,467,827,348,624.00
2020	3	1,583,830.10	9	4751490.3	2,508,517,785,666.01
2021	4	1,615,353.00	16	6461412	2,609,365,314,609.00
2022	5	1,565,893.10	25	7829465.5	2,452,021,200,627.61

2.1. Linear Regression Method

The equation used in this research are:

$$a = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2} \tag{1}$$

$$b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2} \quad (2)$$

$$y = a + b\ (x) \quad (3)$$

$$R^2 = \frac{n(\sum xy) - \sum x \sum y}{\sqrt{[n(\sum x^2) - (\sum x)^2][n(\sum y^2) - (\sum y)^2]}} \quad (4)$$

3. Results and Discussion

3.1. Results

To find out the value of constant (a) and constant (b) it can use equations (1) and (2). So,

$$a = \frac{(7856484)(55) - (15)(23704708)}{5(55) - 15^2} = 1530720$$

$$b = \frac{5(23704708) - (15)(7856484)}{5(55) - 15^2} = 13525.6$$

The coefficient a produces a value of 1,530,720 and b obtains 13,525.6. Both of these coefficients are positive, indicating that there is a direct relationship between the predictor variable (the cause variable) and the effect variable (the variable being measured). Specifically, a positive coefficient suggests that as the value of the predictor variable increases, the value of the effect variable also tends to increase. Conversely, if the coefficients a and b were negative, it would indicate an inverse relationship between the predictor variable and the effect variable. In such a case, an increase in the predictor variable would be associated with a decrease in the effect variable, and vice versa. Thus, the negative values of coefficients a and b would demonstrate a tendency for a negative influence, where higher values of the cause variable would correspond to lower values of the consequence variable in the linear relationship being analyzed. This interpretation highlights the nature of the relationship between the variables studied, emphasizing the direction and strength of the influence that the predictor variables exert on the effect variable within the context of the linear regression model.

So, to predict the load forecast for the next 5 years can use equation (3), so

y = 1530720 + 13525.6 (13) = 1706553

y = 1530720 + 13525.6 (25) = 1868860

y = 1530720 + 13525.6 (37) = 2031167

y = 1530720 + 13525.6 (49) = 2193474

y = 1530720 + 13525.6 (61) = 2355782

The x value contains the contents of the year which is the sum of 12 months plus 1 and produces a value of 13 for forecasting in 2023, as well as for the next year, namely 2024, which contains x in the form of 25 which is obtained from 13 plus 12. Likewise, 37 is produced from 25 plus 12 for forecasting for 2025. In 2026 there is an x value of 49 which was obtained from the previous result, namely 37, and when added 12. The forecast for the last year, namely 2027, is the same as the previous year which was added by 12 from the 2026 result, so that 49 plus 12 produces a value of 61. It could say it is a multiple of 12 or 1 year.

To ensure that the forecasting results are accurate, the correlation coefficient is calculated which will give results in the range 0 – 1, so that to calculate the correlation coefficient can use equation (4).

$$r = \frac{5(23{,}704{,}708) - (15)(7{,}856{,}484)}{\sqrt{[5(55) - (15)^2][5(61{,}724{,}342{,}413{,}552.80) - (7{,}856{,}484)^2]}} = 0{,}4$$

A result of 0.4 for the correlation coefficient indicates that there is a relationship between x (year) and y (connection data). Specifically, a correlation coefficient of 0.4 suggests that there is a moderate positive relationship between these two variables. This means that while the relationship is not extremely strong, there is still a discernible and positive association between the progression of years and the increase in connection data. In practical terms, this indicates that as the years pass, the amount of connection data tends to rise. This moderate positive relationship implies that as time progresses, there is a growing power demand, highlighting that the need for electricity is becoming increasingly significant and unavoidable. Therefore, the data suggests that with each passing year, more power is required to meet the rising demand, underlining the interconnected nature of temporal progression and electricity consumption. This relationship, while moderate, still signifies that changes in one variable (year) can influence changes in the other variable (connection data), demonstrating the dynamic interaction between these two factors over time.

3.2. Discussion

To confirm the results try using one of the latest trends, namely Google Colab. To carry out forecasting using Google Colab, to carry out data analysis and modeling. Can also import relevant datasets, prepare data, and build linear regression forecasting models. As seen in the listing code below.

```
import numpy as np
from sklearn.linier_model import LinierRegression
import matplotlib.pyplot as plt
# Masukkan data total
X = np.array([1, 2, 3, 4, 5]).reshape((-1, 1))
y = np.array([1520475.90, 1570932, 1583830.10, 1615353, 1565893.10])
# Create a linear regression model
model = LinierRegression().fit(X, y)
# Predict value
x_pred = np.array([10]).reshape((-1, 1))
y_pred = model.predict(x_pred)
a = model.Intercept_
b = model.coef_
# Calculating r-squared
r_sq = model.score(X, y)
# Displays results
print('Koefisien regresi: ', model.coef_) #m/b
print('Intercept: ', model.Intercept_) #c/a
print('Persamaan Regresi : {} + {}x'.format(a,b))
print('Hasil prediksi: ', y_pred)
print('R-squared: ', r_sq)
plt.figure(figsize = (10,8))
plt.plot(X, y, 'b.')
plt.plot(X,model.Intercept_ + model.coef_*X, 'r')
plt.xlabel('x')
plt.ylabel('y')
```

The results of the experiment on Google Colab can be seen that the result of the regression coefficient is 13525.54, with an Intercept of 1530720.12, as well as the Regression Equation with a value of 1530720.12 + [13525.54]x, the R-squared result is 0.388. The result of the regression coefficient is [13525.54]. This indicates that every one-unit increase in the independent variable (x) will be followed by an increase of 13525.54 in the dependent variable (y). If the independent variable is time, then every increase of one unit of time will be followed by an increase of 13525.54 in the dependent variable.

Intercept (Intercept) is 1530720.12. This is the value of y when the independent variable (x) is zero. In this context, if the independent variable is time, then when time is zero, the value of the dependent variable (y) will be 1530720.12. The Regression Equation is 1530720.12 + [13525.54]x. This equation describes the relationship between the independent variable (x) and the dependent variable (y) in linear regression. In this equation, 1530720.12 is the Intercept (c/a), and [13525.54] is the coefficient (m/b).

The prediction result is [1665975.6]. This is the prediction result for the value of the dependent variable (y) when the value of the independent variable (x) is 10, based on the trained linear regression model. R-squared is 0.388. This is the R-squared value that indicates how well the linear regression model fits the data. The R-squared value ranges from 0 to 1, where closer to 1 indicates that the model is better at explaining variations in the data. In this context, the R-squared value of 0.388 shows that the linear regression model can explain around 38.84% of the data variation.

The R-squared of 0.388 indicates that around 38.84% of the data variability in the dependent variable (y) can be explained by the independent variable (x) used in the linear regression model. This means that most of the data variations still cannot be explained by this linear regression model, and other factors outside the independent variable (x) influence data variations in the dependent variable (y). In a simple interpretation, the R-squared value of 0.388 shows that the linear regression model is only able to explain around 38.84% of the variation in the data, while the remaining 61.16% is a variation that cannot be explained by the model. So, the results of the regression equation are the same as those calculated manually using Microsoft Excel, and also the R-squared result is 0.388. R-squared is a correlation coefficient that produces a value of 0.4 if rounded. Based on the results of this data, from 2023 to 2027, power requirements will increase over time, which can be seen in Table 3 below.

Table 3: Connected power prediction results.

Years	Variable x: Year Period	Variable y: Connection Data	x^2	x.y	y^2
2018	1	1,520,475.90	1	1520475.9	2,311,846,942,480.81
2019	2	1,570,932.00	4	3141864	2,467,827,348,624.00
2020	3	1,583,830.10	9	4751490.3	2,508,517,785,666.01
2021	4	1,615,353.00	16	6461412	2,609,365,314,609.00
2022	5	1,565,893.10	25	7829465.5	2,452,021,200,627.61

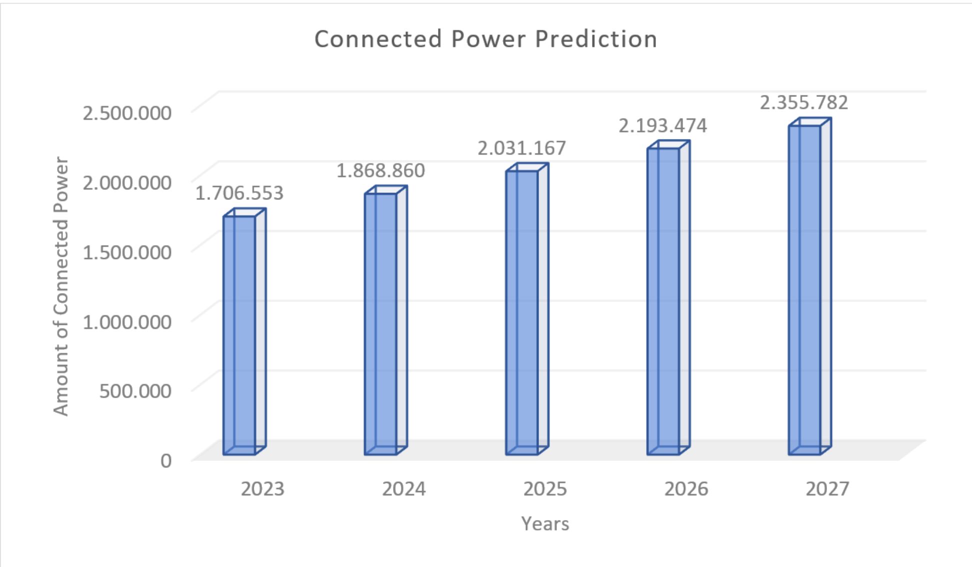

Figure 2: Connected power growth graph.

As seen in Figure 2, which shows the connected power growth graph, it is clear that as time goes on, the power requirements increase. This indicates that the use of electricity has become a necessity, especially in industry. The industry relies on machines and electrical equipment to run production operations. With an adequate and stable electricity supply, this equipment can function properly and provide optimal output. A strong and reliable power supply enables the industry to increase operational power and productivity. KIEC areas, such as manufacturing, mining, and petrochemical, use electrical equipment that requires large amounts of power. For example, heavy machinery, heating systems, material processing, control systems, and others. An adequate electricity supply is required for this equipment to operate smoothly and efficiently.

4. Conclusion

Based on the results of research regarding Electrical Energy Load Forecasting using the Linear Regression Method at PT Krakatau Daya Electric, it was concluded that:

1. The prediction of electrical energy load requirements for five years, until 2027 at PT Krakatau Daya Electric is 1,706,553 kVA in 2023, while in 2024 it will have connected power worth 1,868,860 kVA. In 2025 it is predicted to be 2,031,167 kVA. In the following year, 2026, 2,193,474 kVA is predicted. In the last year, 2027 was predicted to have a power of 2,355,782 kVA.
2. After processing the data, some results always increase over time. This can also be seen when calculating the correlation coefficient obtained at 0.4. So, the relationship between positive x and y is quite related to each other.
3. By obtaining results each year, PT Krakatau Daya Electric is expected to be able to produce electrical energy according to customer needs.

Acknowledgments

I would like to express my sincere gratitude to my supervisor, Agung Baktiarto, for his continued support, guidance, and feedback throughout this research. His expertise and encouragement were instrumental in completing this study. I would also like to thank my colleagues, Adi Nugraha, Dhea Rahmalia Henidar, Fahmi Hermastiandi, and Galih Prasetya N, for their invaluable discussions and collaboration, which greatly enriched this research. Their contributions and collaboration were critical in developing and refining the ideas presented in this paper. Special thanks go to the Commercial Division, Electricity Sales Department for their assistance in data collection and analysis. Their expertise and dedication were critical in ensuring the accuracy and reliability of the results of this study. Finally, I am extremely grateful to my family and friends for their never-ending support and encouragement. Their patience and understanding allowed me to focus on my research and achieve my goals.

References

[1] P. H. Hutabarat, M. F. Zambak, And S. Suwarno, “Prediksi Kebutuhan Energi Listrik Wilayah Pln Kota Parapat Simalungun Sampai Tahun 2024,” Journal Of Electrical And System Control Engineering, Vol. 5, No. 2, Pp. 52–58, Feb. 2022, Doi: 10.31289/Jesce.V5i2.5757.

[2] I. Sholeh, U. Situmeang, P. Studi Teknik Elektro, F. Teknik, And U. Lancang Kuning Pekanbaru Jl Yos Sudarso Rumbai, “Studi Peramalan Beban Pada Gardu Induk Teluk Lembu Uip3bs Upt Pekanbaru PT Pln (Persero),” 2021.

[3] J. T. Putra, N. Fadhilah, And M. Arrofiq, “Peramalan Beban Pada Gardu Induk Mantingan Dalam Penentuan Kapasitas Transformator Dengan Metode Regresi Linear,” 2021.

[4] P. Mangera, “Perkiraan Kebutuhan Energi Listrik Jangka Panjang Pada PT Pln (Persero) Wilayah Papua Dan Papua Barat Area Merauke Dengan Menggunakan Metode Regresi Linier,” Mustek Anim Ha, Vol. 7, No. 3, Pp. 247–256, Dec. 2018, Doi: 10.35724/Mustek.V7i3.1736.

[5] N. Rahmadani, M. Musaruddin, M. N. A. Nur, H. T. Mokui, and A. N. Aliansyah, “Analisis Prakiraan Kebutuhan Energi Listrik di Kabupaten Kolaka Utara menggunakan Metode Dkl 3.2, Regresi Linear dan Software Leap,” J. Fokus Elektroda Energi List. Telekomun. Komputer, Elektron. dan Kendali), vol. 8, no. 2, pp. 101–109, 2023.

[6] H. Riandi, “Analysis Forecasting Electrical Loads In Rsup. Dr. M. Djamil Padang Until 2029,” No. 01, 2020.

[7] J. Wu, Y. G. Wang, Y. C. Tian, K. Burrage, And T. Cao, “Support Vector Regression With Asymmetric Loss For Optimal Electric Load Forecasting,” Energy, Vol. 223, May 2021, Doi: 10.1016/J.Energy.2021.119969.

[8] O. D. P. Yuan, A. N. Afandi, and H. Putranto, “Tekno Jurnal Teknologi Elektro Dan Kejuruan Studi Prakiraan Beban Listrik Menggunakan Metode Artificial Neural Network,” 2018. [Online]. Available: Http://Journal2.Um.Ac.Id/Index.Php/Tekno.

[9] Rizqulloh, F. R., Prasetyono, S., & Cahyadi, W. (2020). Analisis Perbandingan Peramalan Beban Listrik Jangka Pendek Antara Metode Backpropagation Neural Network Dengan Metode Regresi Linear. Jurnal Arus Elektro Indonesia, 6(3), 69-77.

[10] E. Triyanto, H. Sismoro, And A. D. Laksito, “Implementasi Algoritma Regresi Linear Berganda Untuk Memprediksi Produksi Padi Di Kabupaten Bantul,” Rabit : Jurnal Teknologi Dan Sistem Informasi Univrab, Vol. 4, No. 2, Pp. 66–75, Jul. 2019, Doi: 10.36341/Rabit.V4i2.666.