Journal of Mechatronics, Electrical Power, and Vehicular Technology 13 (2022) 214-221

Journal of Mechatronics, Electrical Power, and Vehicular Technology

CrossMark

e-ISSN: 2088-6985
p-ISSN: 2087-3379

mev.lipi.go.id

Long-term forecasting for growth of electricity load based on customer sectors

Sujito [a, *], Ridho Riski Hadi [b], Langlang Gumilar [a], Abdullah Iskandar Syah [b], Moh. Zainul Falah [b], Tran Huy Duy [c]

[a] *Intelligent Power and Advance Energy System, Jurusan Teknik Elektro, Universitas Negeri Malang*
Jl. Semarang No. 5 Sumbersari, Lowokwaru, Malang, East Java, 65145, Indonesia
[b] *Electrical Engineering, Electrical Engineering Department, Universitas Negeri Malang*
Jl. Semarang No. 5 Sumbersari, Lowokwaru, Malang, East Java, 65145, Indonesia
[c] *Electrical Engineering Department, Dalat University, Vietnam*
1 Đường Phù Đổng Thiên Vương, Phường 8, Thành phố Đà Lạt, Lâm Đồng, Vietnam

Received 25 May 2022; 1st revision 18 October 2022; 2nd revision 2 November 2022; 3rd revision 24 November 2022; 4th revision 14 December 2022; Accepted 16 December 2022; Published online 29 December 2022

Abstract

The availability of electrical energy is an important issue. Along with the growth of the human population, electrical energy also increases. This study addresses problems in the operation of the electric power system. One of the problems that occur is the power imbalance due to scale growth between demand and generation. Alternative countermeasures that can be done are to prepare for the possibility that will occur in the future or what we are familiar with forecasting. Forecasting using the multiple linear regression method with this research variable assumes the household sector, business, industry, and public sectors, and is considered by the influence of population, gross regional domestic product, and District Minimum Wage. In forecasting, it is necessary to evaluate the accuracy using mean absolute percentage error (MAPE). MAPE evaluation results show a value of 0.142 % in the household sector, 0.085 % in the business sector, 1.983 % in the industrial sector, and 0.131 % in the total customer sector.

Keywords: district minimum wage; gross regional domestic product; long-term forecasting; mean absolute percentage error; multiple linear regression;

I. Introduction

Electricity is one of many energies that are needed by society to support daily activities [1][2]. Along with the increase in the total electrical loads, the electrical energy required was increased [3][4][5]. Generally, an electric power system is divided into several customer sectors, including a housing sector, a commercial sector, an industrial sector, and a public or general sector [6][7]. One problem in the operation of the electric power system is a power imbalance between the power required and the power generated. The imbalance results in disruption of frequency stability and voltage drop in the system [8]. According to [9], which was explained in the Electricity Supply Business Plan, it was said that the one target of the State Electricity Company or *Perusahaan Listrik Negara* ("PLN") is to be able to provide power capacity and electrical energy every year. To keep an electricity demand fulfilled, it is necessary to connect a supply of electrical energy according to demand and load forecasting that will take place in the future [1][10].

Forecasting is a process of predicting a possible structured way that can take place in the future based on data from the past and the current periods to minimize errors [11][12]. The main factors that influence the forecasting of electricity load growth are macroeconomic problems such as economic growth, population, gross regional domestic product (GRDP), etc. [13][14]. According to [15][16], predicting expenses in the future is usually done by analyzing a graph of expenses in the past to future

* Corresponding Author. Tel: +62-851-32014085
E-mail address: sujito.ft@um.ac.id

doi: https://dx.doi.org/10.14203/j.mev.2022.v13.214-221
2088-6985 / 2087-3379 ©2022 National Research and Innovation Agency

How to Cite: Sujito *et al.*, "Long-term forecasting for growth of electricity load based on customer sectors," *Journal of Mechatronics, Electrical Power, and Vehicular Technology*, vol. 13, no. 2, pp. 214-221, Dec. 2022.

periods, including short-term, medium-term, or long-term. Short-term, mid-term, and long-term models are important to account for the complexity of load data and produce reliable forecasts [15]. Forecasting methods are divided into three types, namely, monitoring, causal, and time series methods [17]. The three types are described in the form of Table 1.

Based on Table 1, this study uses a causal method with two techniques in determining forecasting results [18], namely simple regression and multiple regression, where both are required in forecasting for each explanatory variable (population, GRDP, and district minimum wage).

This research activity was done at West Kotawaringin Regency because West Kotawaringin Regency is the shaded area of PT PLN (Persero) ULP Pangkalan Bun. According to [19], in 2020, West Kotawaringin Regency was the district with the 4th largest population of 14 regencies or cities in Central Kalimantan Province, with an average population growth rate in the last three years of 2.71 % annually. In addition, according to historical data from PT PLN (Persero) ULP Pangkalan Bun 2021, the growth rate in the number of customers has increased by 0.52 % every month starting from January 2018 to December 2020. Therefore, the availability of electrical energy must be continuously monitored and maintained. Therefore, PT PLN (Persero) ULP Pangkalan Bun needs to maintain a continuity of distribution of electrical energy. One of them by forecasting the growth of the electrical energy load in the future.

In this research, long-term forecasting was done in each sector, such as the industrial, household, business, and general sectors. This is to support the maintenance of the availability of electrical energy and planning for the addition of electrical energy in the future by PT PLN [20]. In addition, the data used is a type of linear interpolation because the data always increases and is stable every year, so this research can only be done at the end of each year. The factors that influence a forecasting growth of an electrical load used are the total population, GRDP, and the minimum wage in the West Kotawaringin area.

This research uses a linear regression forecasting method. Simple linear regression was used to predict the total population, GRDP, and minimum wage in the West Kotawaringin Regency area. Forecasting results from these variables with multiple linear regression were used to predict an amount of electricity load growth in industrial, household, business, and general sectors. Calculation of the magnitude of the impact of each independent variable on the dependent variable was done using Pearson's product moment analysis (r) and coefficient of determination (r^2). Then, the accuracy of measured forecasting is determined by mean absolute percent error (MAPE).

Table 1.
Types of forecasting

Forecasting methods	Description
Monitoring	Tracking signals
Causal	Linear regression, multiple regression, ARIMA
Time series	Simple time series, advanced time series

II. Materials and Methods

A. Correlation theory

Correlation is one of the existing methods in statistical analysis as a search for continuity between two variables with quantitative properties, i.e., to see the size of the impact given by an independent variable (independent) to a dependent variable (dependent). Two variables are said to be correlated if a change in one variable will be accompanied by a change in another variable in an organized manner in one direction (positive correlation) or contradictory (negative correlation) [21][22][23]. This research was done using a Pearson product-moment correlation technique and coefficient of determination.

1) Pearson product-moment

According to [24][25], in principle, Pearson product-moment correlation is used to determine a correlation between two variables (bivariate model) which has an interval or ratio scale. Pearson product-moment correlation can be interpreted to find whether there is a relationship between the x variable and the y variable. It is useful as an explanation of how much one variable contributes to another variable. The positive (+) and negative (-) signs represent a type of relationship between variables, and the value ranges from -1.0 to 1.0 as a strengthening statement of the relationship. Mathematically, the Pearson product-moment is described in equation (1):

$$r = \frac{n\sum XY - \sum X \sum Y}{\sqrt{\{n\sum X^2 - (\sum X)^2\}\{n\sum Y^2 - (\sum Y)^2\}}} \quad (1)$$

where r is the Pearson product-moment, n is the customers point (X, Y), X is the independent variable, and Y is the dependent variable.

The results of correlation calculation using a Pearson product-moment method can represent how strong the relationship between variables was given. Table 2 describes a representation strength of continuity between variables calculated using a Pearson product-moment method [26]. The correlation rate is used as a result of decisions from variables that influence each other.

Table 2.
Correlation representation using Pearson product moment method

Value of r	Correlation rate
0.00 – 0.199	Very Low
0.20 – 0.399	Low
0.40 – 0.599	Medium
0.60 – 0.799	Strong
0.80 – 1.000	Very strong

2) Coefficient of determination

The coefficient of determination (equation (2)) is symbolized by r^2 and is usually represented using a percentage (%). The coefficient of determination is a value used to measure the contribution of the independent variable (X) to the variation (increase/decrease) of the dependent variable (Y). Another explanation is that variable Y can be described by variable X with a magnitude of r^2 then the rest was described by other variables. Another variation of y (the rest) was caused by many different factors that influence Y as well and has been included in disturbance error [27][28]. The range coefficient of determination is 0 to 1, with 0 representing no continuity between the independent variables to the dependent variable and 1 representing a perfect relationship between these variables [28][29].

$$r^2 = \frac{((n)(\sum XY) - (\sum X)(\sum Y))^2}{((n(\sum X^2) - (\sum X)^2)(n(\sum X^2) - (\sum Y)^2))} \quad (2)$$

where r^2 is the coefficient of determination.

B. Linear regression

Regression is one of many statistical methods that serve to see the pattern of the relationship between the response variable and predictor variable [30].

1) Simple linear regression

According to [31][32], a simple linear regression is a linear regression that only considers one independent variable (X) and one dependent variable (Y). In linear regression, the variable (Y) can be expressed as a response variable, or in other terms, an output variable and not independent. While the variable (X) can be said to be a predictor variable (used to estimate Y value), it can be expressed as an explanatory variable, input regressor, and independent.

The simple linear regression equation model can be explained through equation (3)

$$Y = a + bX \quad (3)$$

where a is the constant and b is the coefficient.

The value of a and b needs to be calculated using equation (4) and equation (5):

$$a = \frac{(\sum Y)(\sum X^2) - (\sum X)(\sum XY)}{(n)(\sum X^2) - (\sum X)^2} \quad (4)$$

$$b = \frac{n \sum XY - (\sum X)(\sum Y)}{n(\sum X^2) - (\sum X)^2} \quad (5)$$

2) Multiple linear regression

According to [33][34], multiple linear regression is a continuation of simple linear regression analysis. Multiple linear regression uses one dependent variable (Y) as the predicted target and several independent variables (X) as the variable used to predict a target.

The coefficient on linear regression is an estimated value of the parameter in the regression model for the real condition (true condition). The coefficient for the linear regression model is an average value that has a chance of appearing on the Y variable if a value of X_1, X_2, and X_n was given. Mathematically, multiple linear regression analysis is explained in equation (6).

$$Y = a + b_1X_1 + b_2X_2 + \dots \dots \dots b_kX_k \quad (6)$$

where b_k is the coefficient X_k, X_1 is the 1st independent variable, b_1 is the coefficient X_1, X_2 is the 2nd independent variable, b_2 is the coefficient X_2, and X_3 is the 3rd independent variable.

3) Forecasting accuracy

Forecasting accuracy is a measure of forecast error based on the magnitude of the difference between forecast results and actual demand. The measurement of forecasting accuracy aims to determine the performance or accuracy of forecasting results that have been done using certain methods and techniques. In this research, the mean absolute percentage error (MAPE) method was used to assess forecasting accuracy. According to [35][36], MAPE can be calculated by the following mathematical equation (7)

$$MAPE = \frac{\sum_{t=1}^{n} \frac{|At - Ft|}{At}}{n} \times 100 \quad (7)$$

where t is the period on data, F_t is the forecasting data in period t, A_t is the actual data in period t, and n is the total forecasting data.

MAPE results can be grouped based on the level of forecasting accuracy [37] as shown in Table 3. Parameter results are evaluated using MAPE and compared with the range value from MAPE to determine whether the accuracy level is high or low.

4) Systematic research

Figure 1 shows the flow of research from

Table 3.
MAPE parameters in forecasting

Value of MAPE	Prediction accuracy
MAPE ≤ 10%	High
10% < MAPE ≤ 20%	Good
20% < MAPE ≤ 50%	Reasonable
MAPE > 50%	Low

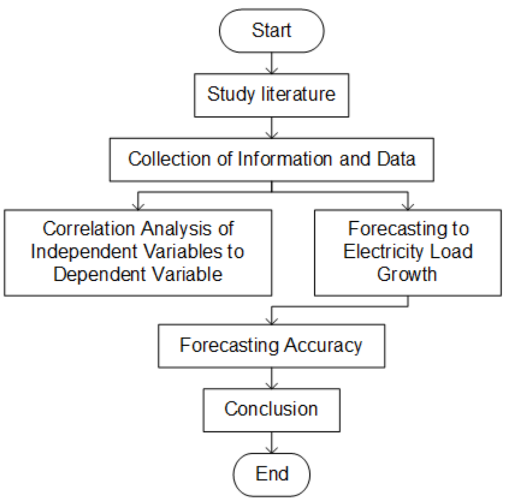

Figure 1. Research completion flow

beginning to end. It shows research data that has been collected and analyzed with two different output lines. In the first analysis, an output produced was to identify the correlation between independent variables and the dependent variable using a Pearson product-moment method and coefficient of determination. While the other analysis, a results output, is the forecasting of electrical energy needs and continued by measuring the accuracy of the forecast.

III. Results and Discussions

A. Effect of correlation of forecasting variables

In this study, the data used consisted of two types of variables, namely the independent variable and the dependent variable. Independent variables are macroeconomic factors that can affect the dependent variable, such as population, GRDP, and city minimum wage. The dependent variable is a variable that was influenced by macroeconomic factors; here is electricity data which is divided into four sectors, namely the household, business, industrial, and public sectors.

Based on the calculations performed using equation (1) and equation (2), the results of the correlation between the independent and the dependent variables are obtained. The correlation results are shown by the value of the Pearson product moment (r) and the coefficient of determination (r^2) listed in Table 4. Thus, we can see how much influence the independent variable has on the dependent variable based on the r and r^2 values obtained.

The results of the correlation calculation obtained for each independent variable and the dependent variable were corrected with the r value in Table 2. These results indicate that the correlation value obtained is very strong and positive. This signifies that there is a very close relationship between the independent variable and the dependent variable. This means that if there is a change in the independent variable, there will also be a significant change in the dependent variable.

Table 4.
Correlation between independent variable and dependent variable

Variable	Pearson product moment (r)	Coefficient of determination (r^2)
X1 to Y1	0.998	0.995
X2 to Y1	0.984	0.969
X3 to Y1	0.996	0.991
X1, X2, X3 to Y1	0.993	0.999
X1 to Y2	0.978	0.957
X2 to Y2	0.972	0.944
X3 to Y2	0.975	0.950
X1, X2, X3 to Y2	0.975	0.994
X1 to Y3	0.976	0.953
X2 to Y3	0.928	0.861
X3 to Y3	0.978	0.957
X1, X2, X3 to Y3	0.960	0.961
X1 to Y4	0.999	0.997
X2 to Y4	0.974	0.948
X3 to Y4	0.998	0.996
X1, X2, X3 to Y4	0.990	0.997
X1 to Ytot	0.998	0.996
X2 to Ytot	0.984	0.968
X3 to Ytot	0.996	0.992
X1, X2, X3 to Ytot	0.992	0.999

B. Forecasting based on variables that affect load demand

Based on calculations made using equations (3) to (6), the forecasting results for the growth of electricity loads for each sector are obtained. These sectors are the household, business, industrial, and public sectors. The calculation takes into account macro factors that can affect the growth of electricity expenses, such as population, GRDP, and city minimum wage. The forecasting results are shown in Table 5. Thus, we can estimate the electricity load growth for each sector based on influencing macro factors.

In general, each year, the energy growth in each sector has increased (see Figure 2). This is because every year, it is influenced by population growth, as

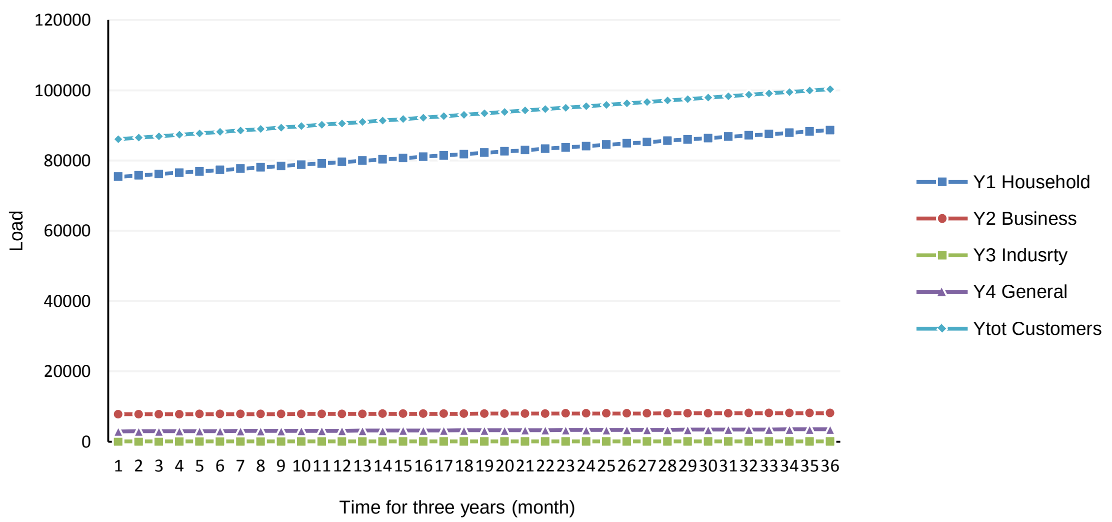

Figure 2. Load demand growth

Table 5.
The result of forecasting the growth of electricity load for each sector

Year	Month	Customer				
		Y1 Household	Y2 Business	Y3 Industry	Y4 General	Ytot Total customers
2021	January	75377	7770	32	2926	86104
	February	75755	7780	32	2943	86510
	March	76134	7790	32	2960	86916
	April	76513	7800	32	2977	87322
	May	76892	7810	33	2993	87729
	June	77271	7820	33	3010	88135
	July	77650	7830	33	3027	88541
	August	78029	7840	34	3044	88947
	September	78408	7850	34	3061	89353
	October	78787	7861	34	3077	89759
	November	79166	7871	34	3094	90165
	December	79545	7881	35	3111	90571
2022	January	79924	7891	35	3128	90978
	February	80303	7901	35	3145	91384
	March	80682	7911	36	3162	91790
	April	81061	7921	36	3178	92196
	May	81440	7931	36	3195	92602
	June	81818	7941	36	3212	93008
	July	82197	7951	37	3229	93414
	August	82576	7961	37	3246	93820
	September	82955	7972	37	3262	94227
	October	83334	7982	38	3279	94633
	November	83713	7992	38	3296	95039
	December	84092	8002	38	3313	95445
2023	January	84471	8012	38	3330	95851
	February	84850	8022	39	3347	96257
	March	85229	8032	39	3363	96663
	April	85608	8042	39	3380	97069
	May	85987	8052	40	3397	97476
	June	86366	8062	40	3414	97882
	July	86745	8072	40	3431	98288
	August	87124	8083	40	3447	98694
	September	87503	8093	41	3464	99100
	October	87881	8103	41	3481	99506
	November	88260	8113	41	3498	99912
	December	88639	8123	42	3515	100318

it is known through Table 4 that the correlation between the two factors is strong.

Based on Table 5, it can be seen that the number of electricity customers in each sector has always increased from January 2021 to December 2023. This increase is due to population growth and other sectors which have a high correlation.

To measure the accuracy of forecasting the growth of this electrical load, equation (7), which is called the mean absolute percentage error (MAPE), can be used. The results of the calculation of forecasting accuracy according to this method can be seen in Table 6. From these results, it can be seen that forecasting the growth of electricity loads in each sector has fairly high accuracy.

In this case, the results of calculating the accuracy of the load forecasting using the linear regression method show that the mean absolute percentage error (MAPE) is quite small. This shows that load forecasting using the linear regression method is quite accurate in contrast to the results of the previous study [14], which showed that the linear regression method for forecasting urban demand loads has a larger MAPE value compared to the K-nearest neighbour (KNN) method they use, which in their research method KNN is more appropriate to the cases they use compared to the linear regression method.

In the same case study, the auto-regressive (AR) method was applied to the same dataset to forecast each sector in January and December to ensure the training results. The results of the auto-regressive (AR) test are listed in Table 7.

In Table 6, the MAPE value for forecasting the independent variables is 0.005 % for the population and 0.764 % for GRDP. Meanwhile, the MAPE value for forecasting the dependent variable is shown in Table 6, which is 0.142 % for household sector electricity costs, 0.085 % for the business sector, 1.983 % for the industrial sector, and 0.131 % for

Table 6.
Forecasting accuracy results with MAPE

Year	Month	Error (%)				
		Y1	Y2	Y3	Y4	Ytot
		Household	Business	Industry	General	Total customers
2021	January	0.169	0.084	3.884	0.549	0.171
	February	0.134	0.051	4.970	0.118	0.123
	March	0.411	0.005	3.166	0.198	0.347
	April	0.234	0.013	2.175	0.271	0.209
	May	0.045	0.006	1.184	0.357	0.051
	June	0.187	0.052	0.192	0.358	0.145
	July	0.268	0.085	3.401	0.276	0.216
	August	0.326	0.050	2.451	0.030	0.278
	September	0.013	0.096	1.501	0.115	0.024
	October	0.343	0.088	0.550	0.118	0.310
	November	0.381	0.134	4.765	0.081	0.343
	December	0.169	0.008	2.704	0.316	0.134
2022	January	0.019	0.170	1.647	0.367	0.021
	February	0.195	0.120	0.590	0.379	0.171
	March	0.277	0.102	0.466	0.195	0.257
	April	0.202	0.269	1.523	0.091	0.204
	May	0.285	0.317	2.580	0.335	0.288
	June	0.037	0.102	3.637	0.151	0.029
	July	0.030	0.150	0.667	0.482	0.028
	August	0.224	0.120	1.684	0.848	0.212
	September	0.138	0.037	2.700	0.676	0.140
	October	0.037	0.007	3.716	0.206	0.039
	November	0.117	0.063	2.749	0.229	0.114
	December	0.076	0.080	1.805	0.228	0.065
2023	January	0.050	0.069	0.490	0.153	0.045
	February	0.115	0.072	0.826	0.478	0.109
	March	0.126	0.049	1.381	0.618	0.127
	April	0.104	0.039	0.110	0.436	0.102
	May	0.003	0.081	2.212	0.219	0.003
	June	0.040	0.058	0.985	0.040	0.032
	July	0.076	0.074	2.990	0.138	0.056
	August	0.014	0.077	1.802	0.031	0.019
	September	0.103	0.066	0.614	0.313	0.106
	October	0.103	0.034	0.574	0.487	0.104
	November	0.027	0.096	1.762	0.002	0.015
	December	0.053	0.132	2.951	0.242	0.065
MAPE		0.142	0.085	1.983	0.281	0.131

Table 7.
Forecasting with auto-regressive (AR) method

Month (2023)	Customer per sector				Total customer
	House	Business	Industry	General	
January	84470896	8012142	38183383	3329843	95851117
December	8863886	8123124	41202297	3514834	100318076

customers as a whole. In addition, in other tests using different methods, as forecasting comparisons are shown in Table 7, the forecasting results prove the same as the calculation of equations (3) to (6). This proves that the cases used are in accordance with the linear regression method, where the data used is time series.

Based on these results, forecasting the growth of electricity loads in all sectors can be used with a high accuracy value because the MAPE value is below 10 %, according to Table 3. With the results of forecasting in 2023, it can be used as a comparison in increasing power capacity that is able to support the energy demand needs of consumers from various sectors.

IV. Conclusion

The population, GRDP, and minimum wage have a very strong (positive) influence on the growth of the electricity load in each customer sector. The results of forecasting the growth of the electrical load showed a steady increase. The results of the measurement of the accuracy of the electricity load growth forecast conducted by MAPE in the household sector are 0.142 %, business is 0.085 %, the industry is 1.983 %, and the number of customers is 0.131 %. MAPE value < 10 %, so the accuracy of forecasting the growth of electricity load in all sectors is high. In further research, the use of other training methods can be used as a comparison for the results of forecasting the electrical loads in Kalimantan.

Acknowledgments

Thanks are conveyed to the Institute for Research and Community Service (LP2M) State University of Malang for funding this research activity through Basic Research Grants – Thesis Research in 2022.

Declarations

Author contribution

Sujito: Formal Analysis, Validation, Data Curation. R.R. Hadi, L. Gumilar, T.H. Duy: Conceptualization, Formal Analysis, Resources. A.I. Syah, M.Z. Falah: Writing, Formal Analysis, Software Operations, Visualization, Funding Acquisition.

Funding statement

Thanks are conveyed to the Institute for Research and Community Service (LP2M) State University of Malang for funding this research activity through Basic Research Grants – Thesis Research in 2022 No. 19.5.1259/UN32.20.1/LT/2022.

Competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Additional information

Reprints and permission: information is available at https://mev.lipi.go.id/.

Publisher's Note: National Research and Innovation Agency (BRIN) remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

[1] M. Subekti, I. A. Rahardjo, and D. Rosyanti, "Forecasting electrical energy demand of PT. PLN (Persero) UP3 Sukabumi using analytical, econometrics, and trends methods," *IOP Conference Series: Materials Science and Engineering*, vol. 1098, no. 4, p. 042032, 2021.

[2] K. B. Lindberg, P. Seljom, H. Madsen, D. Fischer, and M. Korpås, "Long-term electricity load forecasting: Current and future trends," *Utilities Policy*, vol. 58, pp. 102–119, 2019.

[3] A. Setiawan, "Losses sharing regulation as a solution for spike losses at PLN's Distribution system caused by energy over production from independent power producers," in *2020 International Conference on Technology and Policy in Energy and Electric Power (ICT-PEP)*, Bandung, Indonesia, Sep., pp. 219–224, 2020.

[4] G. Zhang and J. Guo, "A novel method for hourly electricity demand forecasting," *IEEE Transactions on Power Systems*, vol. 35, no. 2, pp. 1351–1363, 2019.

[5] S. Aslam, H. Herodotou, S. M. Mohsin, N. Javaid, N. Ashraf, and S. Aslam, "A survey on deep learning methods for power load and renewable energy forecasting in smart microgrids," *Renewable and Sustainable Energy Reviews*, vol. 144, p. 110992, 2021.

[6] Z. A. Khan *et al.*, "Efficient short-term electricity load forecasting for effective energy management," *Sustainable Energy Technologies and Assessments*, vol. 53, p. 102337,, 2022.

[7] L. Peng, L. Wang, D. Xia, and Q. Gao, "Effective energy consumption forecasting using empirical wavelet transform and long short-term memory," energy, vol. 238, p. 121756,, 2022.

[8] H. Zhou, Y. Zhang, L. Yang, Q. Liu, K. Yan, and Y. Du, "Short-term photovoltaic power forecasting based on long short term memory neural network and attention mechanism," *IEEE Access*, vol. 7, pp. 78063–78074, 2019.

[9] J. Fahmi, J. Windarta, and A. Y. Wardaya, "Studi awal penerapan distributed generation untuk optimalisasi PLTS atap on grid pada pelanggan PLN sistem jawa bali untuk memenuhi target EBT nasional," *Jurnal Energi Baru dan Terbarukan*, vol. 2, no. 1, pp. 1–13, 2021.

[10] L. Wen, K. Zhou, and S. Yang, "Load demand forecasting of residential buildings using a deep learning model," *Electric Power Systems Research*, vol. 179, p. 106073, 2020.

[11] Y. Yu, J. Cao, and J. Zhu, "An LSTM short-term solar irradiance forecasting under complicated weather conditions," *IEEE Access*, vol. 7, pp. 145651–145666, 2019.

[12] M. Pavlicko, M. Vojteková, and O. Blažeková, "Forecasting of electrical energy consumption in Slovakia," *Mathematics*, vol. 10, no. 4, p. 577, 2022.

[13] B. Nepal, M. Yamaha, A. Yokoe, and T. Yamaji, "Electricity load forecasting using clustering and ARIMA model for energy management in buildings," *Japan Architectural Review*, vol. 3, no. 1, pp. 62–76, 2020.

[14] N. J. Johannesen, M. Kolhe, and M. Goodwin, "Relative evaluation of regression tools for urban area electrical energy demand forecasting," *Journal of cleaner production*, vol. 218, pp. 555–564, 2019.

[15] H. Riahi-Madvar, M. Dehghani, R. Memarzadeh, and B. Gharabaghi, "Short to long-term forecasting of river flows by heuristic optimization algorithms hybridized with ANFIS," *Water Resources Management*, vol. 35, no. 4, pp. 1149–1166, 2021.

[16] A. Sherstinsky, "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network," *Physica D: Nonlinear Phenomena*, vol. 404, p. 132306, 2020.

[17] P. Malhan and M. Mittal, "A novel ensemble model for long-term forecasting of wind and hydro power generation," *Energy Conversion and Management*, vol. 251, p. 114983, 2022.

[18] N. Talkhi, N. A. Fatemi, Z. Ataei, and M. J. Nooghabi, "Modeling and forecasting number of confirmed and death caused COVID-19 in IRAN: A comparison of time series forecasting methods," *Biomedical Signal Processing and Control*, vol. 66, p. 102494, 2021.

[19] R. Sihite, "Analisis pengaruh pendapatan perkapita, jumlah konsumsi dan pertumbuhan ekonomi di Kabupaten/Kota Provinsi Kalimantan Tengah," *JEPP: Jurnal Ekonomi Pembangunan Dan Pariwisata*, vol. 2, no. 1, pp. 46–57, 2022

[20] 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, 2018.

[21] J. Deng, Y. Deng, and K. H. Cheong, "Combining conflicting evidence based on Pearson correlation coefficient and weighted graph," *International Journal of Intelligent Systems*, vol. 36, no. 12, pp. 7443–7460, 2021.

[22] H. Zhu, X. You, and S. Liu, "Multiple ant colony optimization based on pearson correlation coefficient," *IEEE Access*, vol. 7, pp. 61628–61638, 2019.

[23] P. Schober, C. Boer, and L. A. Schwarte, "Correlation coefficients: appropriate use and interpretation," *Anesthesia & Analgesia*, vol. 126, no. 5, pp. 1763–1768, 2018.

[24] F. Zinzendoff Okwonu, B. Laro Asaju, and F. Irimisose Arunaye, "Breakdown analysis of Pearson correlation coefficient and robust correlation methods," *IOP Conference Series: Materials Science and Engineering*, vol. 917, no. 1, 2020.

[25] A. Ali *et al.*, "A k-Nearest neighbours based ensemble via optimal model selection for regression," *IEEE Access*, vol. 8, pp. 132095–132105, 2020.

[26] M. Ausloos, A. Eskandary, P. Kaur, and G. Dhesi, "Evidence for Gross Domestic Product growth time delay dependence over Foreign Direct Investment. A time-lag dependent correlation study," *Physica A: Statistical Mechanics and Its Applications*, vol. 527, p. 121181, 2019.

[27] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Computer Science*, vol. 7, pp. 1–24, 2021.

[28] S. Nakagawa, P. C. D. Johnson, and H. Schielzeth, "The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded," *Journal of the Royal Society Interface*, vol. 14, no. 134, 2017.

[29] A. M. da Silva Filho, G. F. Zebende, A. P. N. de Castro, and E. F. Guedes, "Statistical test for multiple detrended cross-correlation coefficient," *Physica A: Statistical Mechanics and its Applications*, vol. 562, p. 125285, 2021.

[30] T. M. Hope, "Linear regression," in *Machine Learning*, Elsevier, pp. 67–81, 2020.

[31] M. Sharifzadeh, A. Sikinioti-Lock, and N. Shah, "Machine-learning methods for integrated renewable power generation: A comparative study of artificial neural networks, support vector regression, and Gaussian Process Regression," *Renewable and Sustainable Energy Reviews*, vol. 108, pp. 513–538, 2019.

[32] H.-Y. Kim, "Statistical notes for clinical researchers: simple linear regression 3 – residual analysis," *Restorative Dentistry & Endodontics*, vol. 44, no. 1, pp. 1–8, 2019.

[33] T. A. Trunfio, A. Scala, A. D. Vecchia, A. Marra, and A. Borrelli, "Multiple regression model to predict length of hospital stay for patients undergoing femur fracture surgery at 'San Giovanni Di Dio e Ruggi d'Aragona' University Hospital," in *European Medical and Biological Engineering Conference*, pp. 840–847, 2020.

[34] M. Korkmaz, "A study over the general formula of regression sum of squares in multiple linear regression," *Numerical Methods for Partial Differential Equations*, vol. 37, no. 1, pp. 406–421, 2021.

[35] N. Kusuma, M. Roestam, and L. Pasca, "The analysis of forecasting demand method of linear exponential smoothing," *International Journal of Educational Administration, Management, and Leadership*, pp. 7–18, 2020.

[36] S. Prayudani, A. Hizriadi, Y. Y. Lase, and Y. Fatmi, "Analysis accuracy of forecasting measurement technique on random K-nearest neighbor (RKNN) using MAPE and MSE," in *Journal of Physics: Conference Series*, vol. 1361, no. 1, p. 012089, 2019.

[37] E. Vivas, H. Allende-Cid, and R. Salas, "A systematic review of statistical and machine learning methods for electrical power forecasting with reported mape score," *Entropy*, vol. 22, no. 12, p. 1412, 2020.