

Comparative Analysis of Linear and Quadratic Polynomial Regression Models for Peak Load Forecasting in Distribution Feeders

Titiek Suheta^{1*}, Misbahul Munir¹, Achmad David Julianto¹, Muhammad Iqbal Zakaria¹, Rhesa Ardra Nugraha¹, Faisal Yusana¹, Acok Ahdillah Usman Almah¹, Novendra L Rivaldo Bere¹

¹Department of Electrical Engineering, Institut Teknologi Adhi Tama Surabaya, Surabaya, Indonesia

Article Info

Article history:

Received August 19, 2026

Revised September 16, 2026

Accepted October 20, 2026

Keywords:

peak load
distribution feeder
load forecasting
linear regression
quadratic polynomial regression

ABSTRACT

Peak load forecasting provides essential information for planning and operating electrical distribution systems. This study investigates the forecasting performance of linear and quadratic polynomial regression models using monthly historical load observations from seven distribution feeders over the 2021–2022 period. To account for differences in loading patterns, each feeder was analysed independently. The linear model uses the time index as the predictor, whereas the quadratic polynomial approach is represented by a second-order polynomial regression. Model accuracy was assessed using Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), and the coefficient of determination (R^2). The results reveal variations in forecasting performance among the feeders. In general, quadratic polynomial regression yields slightly lower RMSE values, although the relative performance based on MAPE and R^2 differs across feeders. The differences between the two approaches are relatively small for most feeders, suggesting that greater model complexity does not necessarily provide a substantial improvement in forecasting accuracy. These findings highlight the importance of considering feeder-specific load characteristics and multiple evaluation indicators when selecting a forecasting model. Overall, the proposed approach offers a simple and computationally efficient alternative for feeder-level peak load forecasting based solely on historical load information.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Titiek Suheta

Department of Electrical Engineering, Institut Teknologi Adhi Tama Surabaya, Surabaya, Indonesia

Email: hita@itats.ac.id; Orcid ID: <http://orcid.org/0000-0001-9027-4707>

1. INTRODUCTION

Accurate estimation of future electricity demand plays an important role in the planning and operation of distribution systems. Load forecasting provides useful information for network development, equipment loading evaluation, maintenance planning, and operational decisions, thereby helping utilities allocate system capacity more effectively[1], [2]. Various forecasting approaches have been introduced, including conventional statistical techniques as well as machine learning and deep learning methods[3], [4]. Despite the increasing use of artificial intelligence techniques, their implementation may involve extensive datasets, preprocessing procedures, and considerable computational effort. In contrast, conventional regression approaches offer simpler implementation, easier interpretation, and relatively low computational requirements, making them suitable for practical load forecasting applications[5], [6], [7], [8].

Linear regression describes changes in electrical load through a linear relationship with time, while quadratic polynomial regression can represent more complex variations in load patterns[9], [10]. Previous load forecasting studies have also incorporated additional information, including weather, calendar, socioeconomic, and industrial variables, to improve prediction performance[11]. However, such studies generally focus on improving forecasting performance through different input variables or modelling approaches rather than directly examining whether quadratic polynomial regression provides a meaningful improvement over linear regression when both models are applied to the same historical peak load data at the individual feeder level. This leaves a specific research gap concerning the comparative performance of these two regression approaches under identical input conditions across feeders with different loading characteristics.

To address this research gap, the present study compares linear regression and quadratic polynomial regression for forecasting monthly peak load across seven distribution feeders using historical observations from 2021 to 2022. Separate models are developed for each feeder to preserve their individual loading characteristics. Forecasting accuracy is assessed through Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), and the coefficient of determination (R^2). The study emphasizes a practical feeder-level forecasting approach that relies only on historical load information while maintaining simplicity, interpretability, and low computational requirements for distribution system planning and loading assessment.

The novelty of this study lies in the feeder-level comparison of linear regression and quadratic polynomial regression using identical historical load inputs across seven distribution feeders. Unlike forecasting approaches that incorporate multiple exogenous variables or computationally intensive models, the proposed approach relies solely on historical monthly peak load data while preserving the individual characteristics of each feeder. This comparative framework enables a direct assessment of whether the additional complexity of quadratic polynomial regression provides a meaningful improvement over linear regression. Therefore, the study provides a simple, interpretable, and computationally efficient framework for selecting regression models for feeder-level peak load forecasting.

2. METHOD

2.1 Research Framework

The research was conducted through a sequence of stages designed to compare the forecasting capability of two regression approaches. First, relevant literature on electrical load forecasting and regression-based methods was reviewed to establish the theoretical basis of the study. Historical monthly peak load records from seven distribution feeders covering 2021–2022 were then collected and arranged for further analysis. Using the same dataset, separate linear regression and quadratic polynomial regression models were developed for each feeder. Their forecasting performance was assessed using Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), and the coefficient of determination (R^2). The resulting performance indicators were subsequently compared to evaluate the forecasting capability of both regression approaches.

The regression models were fitted using all available monthly observations for each feeder. Most feeders contained 24 observations covering January 2021 to December 2022, while feeders with unavailable monthly records were modelled using the 22 available observations. No separate training–testing split was applied because the objective of this study was to compare the fitting capability of linear regression and quadratic polynomial regression under identical historical data conditions rather than to perform out-of-sample forecasting. Accordingly, MAPE and RMSE represent in-sample fitting errors calculated between the observed and estimated peak load values, while R^2 represents the goodness of fit of each regression model.

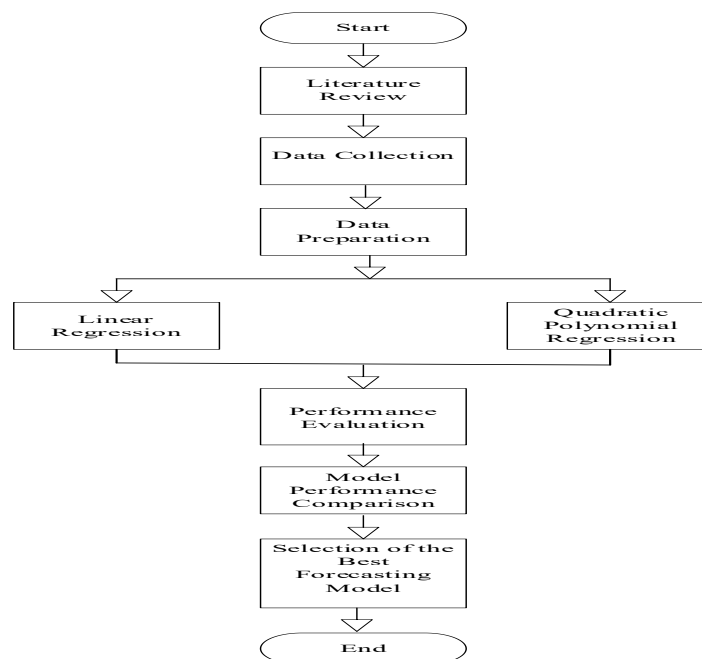


Figure 1. Research framework of the proposed forecasting method.

Figure 1 summarizes the overall procedure used in this study. Following data collection and preparation, the monthly peak load records of each feeder were processed separately using linear regression and quadratic polynomial regression. The fitted values obtained from both models were then evaluated using MAPE, RMSE, and R^2 . The final stage compares the performance of both regression approaches to determine their relative effectiveness in representing feeder-level peak load patterns.

2.2 Data Collection

This study used monthly peak load records from seven distribution feeders covering January 2021 to December 2022. Most feeders contained 24 monthly observations, while feeders with unavailable monthly records contained 22 observations. The data were obtained from distribution system operational records and arranged chronologically according to the observation period. Peak load, expressed as apparent power in MVA, was used as the main variable to represent changes in feeder loading over time. The analysis relied exclusively on historical load observations; therefore, external factors such as weather conditions, calendar variables, and customer characteristics were not incorporated. Using the same available historical information for both regression approaches ensured a consistent basis for comparing their performance. The collected data were subsequently prepared for individual feeder modelling and performance evaluation.

2.3 Data Preparation

Before regression modelling, the historical peak load records were examined to ensure their suitability for analysis. The available observations for each feeder were arranged chronologically, and each monthly value was associated with a corresponding time index. Missing monthly records were not estimated or replaced; therefore, the regression models were developed using only the available observations for each feeder. The same dataset for each feeder was applied to both linear regression and quadratic polynomial regression to ensure consistent comparison conditions. No normalization, feature extraction, or additional exogenous variables were introduced during data preparation. The prepared datasets were subsequently used for model fitting and performance evaluation.

2.4 Linear Regression Model

Linear regression was used to represent the relationship between monthly peak load and the corresponding time index. This approach provides a straightforward means of identifying linear trends in historical load observations and has been applied in various electricity load forecasting studies [12], [13], [14].

The linear regression model is expressed as: $\hat{y} = \beta_0 + \beta_1 x$ (1)

Where:

$\hat{y}$ = predicted monthly peak load

x = time index

β_0 = intercept

β_1 = regression coefficient

The objective function of the Ordinary Least Squares (OLS) method is formulated as:

$$S = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (2)$$

The regression coefficients are calculated as:

$$\beta_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (3)$$

$$\beta_0 = \bar{y} - \beta_1 \bar{x} \quad (4)$$

2.5 Quadratic Polynomial Regression Model

Quadratic polynomial regression was employed to capture load variations that may not be adequately represented by a linear relationship. In this study, a second-order polynomial regression model was used to represent possible nonlinear patterns between the time index and monthly peak load [9], [15]:

$$\hat{y} = \beta_0 + \beta_1 x + \beta_2 x^2 \quad (5)$$

Where:

$\hat{y}$ = predicted monthly peak load

x = time index

β_0 = intercept

β_1 = linear regression coefficient

β_2 = quadratic regression coefficient

2.6 Performance Evaluation

The performance of the linear regression and quadratic polynomial regression models was assessed using three evaluation indicators: Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), and the coefficient of determination (R^2). MAPE measures the relative difference between the observed and fitted values in percentage terms, while RMSE quantifies the magnitude of fitting errors with greater sensitivity to larger deviations. R^2 indicates how well each regression model represents the variation in the observed data. Since no separate testing dataset was used, MAPE and RMSE in this study represent in-sample fitting errors rather than out-of-sample forecasting errors. Better model performance is generally associated with lower MAPE and RMSE values and a higher R^2 value [16], [17].

$$\text{MAPE} = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (6)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (7)$$

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} \quad (8)$$

3. RESULTS

3.1. Peak Load Characteristics

Figure 2 illustrates the monthly peak load patterns of the seven distribution feeders throughout 2021–2022. Differences in load magnitude and temporal variation can be observed among the feeders, indicating that each feeder exhibits its own loading characteristics. These differences provide the basis for developing separate linear regression and quadratic polynomial regression models for each feeder.

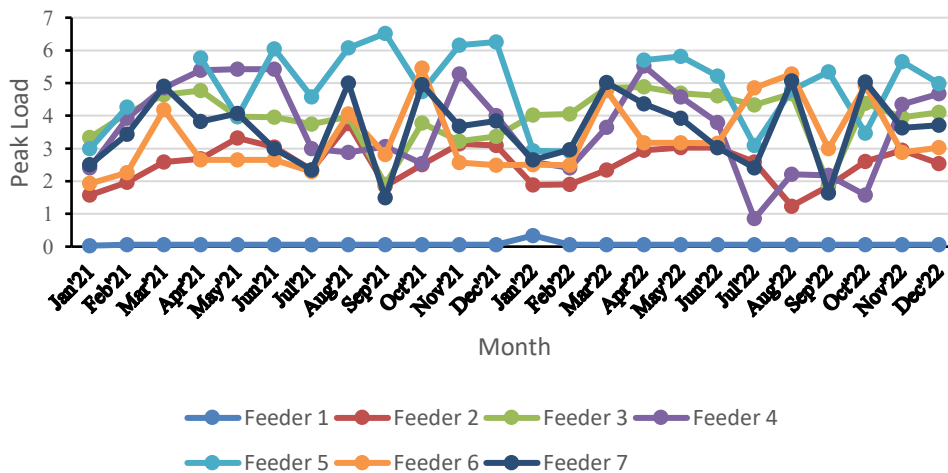


Figure 2. Monthly peak load profiles of seven distribution feeders during 2021–2022

3.2. Linear Regression Results

A separate linear regression model was developed for each feeder with the time index used as the predictor. Table 1 summarizes the resulting regression equations and their corresponding coefficients of determination (R^2).

Table 1. Linear Regression Results for Seven Distribution Feeders

Feeder	Linear Regression Equation	R^2
F1	$\hat{Y} = 0.065145 + 0.000422x$	0.002667
F2	$\hat{Y} = 2.562681 - 0.002450x$	0.000798
F3	$\hat{Y} = 3.847971 + 0.009496x$	0.007164
F4	$\hat{Y} = 4.296087 - 0.055190x$	0.081847
F5	$\hat{Y} = 4.864259 + 0.000767x$	0.000021
F6	$\hat{Y} = 2.618804 + 0.054696x$	0.135102
F7	$\hat{Y} = 3.501377 + 0.008257x$	0.002876

As shown in Table 1, the linear models yielded relatively low R^2 values for all seven feeders. The highest value was obtained for Feeder 6 (0.135102), followed by Feeder 4 (0.081847), whereas considerably smaller values were observed for the other feeders. This suggests that a simple linear trend explains only a limited proportion of the monthly peak load variation, providing a basis for examining whether a quadratic polynomial relationship can offer an improved representation.

3.3. Quadratic Polynomial Regression Results

A second-order quadratic polynomial regression model was developed separately for each feeder to examine nonlinear variations in monthly peak load. The resulting equations and R^2 values are presented in Table 2.

Table 2. Quadratic Polynomial Regression Results for Seven Distribution Feeders

Feeder	Quadratic Polynomial Regression Equation	R^2
F1	$\hat{Y} = -0.000362x^2 + 0.009473x + 0.025924$	0.077612
F2	$\hat{Y} = -0.003111x^2 + 0.075325x + 2.225667$	0.049941
F3	$\hat{Y} = 0.000514x^2 - 0.003350x + 3.903636$	0.007964
F4	$\hat{Y} = 0.000982x^2 - 0.079730x + 4.402436$	0.082834
F5	$\hat{Y} = -0.006380x^2 + 0.163328x + 4.133554$	0.054442
F6	$\hat{Y} = -0.002750x^2 + 0.123379x + 2.321176$	0.148100
F7	$\hat{Y} = -0.000360x^2 + 0.017103x + 3.463442$	0.002709

The results in Table 2 show that the ability of the quadratic polynomial regression model to represent peak load variation differs among feeders. Feeder 6 obtained the highest R^2 value of 0.148100, followed by Feeder 4 at 0.082834 and Feeder 1 at 0.077612. Although these values are generally higher than those obtained from the corresponding linear models, the overall R^2 values remain relatively low. This indicates that the quadratic polynomial regression model provides only a limited improvement in representing the observed monthly load variations.

3.4 Model Performance Comparison

The performance of the two regression approaches was evaluated for each feeder using MAPE, RMSE, and R^2 . MAPE and RMSE were used to quantify the differences between the observed and fitted peak load values, while R^2 was used to assess the goodness of fit of each model. By considering these three indicators together, the performance of linear regression and quadratic polynomial regression can be compared across the seven feeders. Table 3 summarizes the results obtained from both regression approaches.

Table 3. Performance Comparison of Linear and Quadratic Polynomial Regression Models

Feeder	Model	MAPE (%)	RMSE	R^2
F1	Linear	24.4828	0.056455	0.002667
	Quadratic Polynomial	28.0876	0.054292	0.077612
F2	Linear	22.3699	0.599622	0.000798
	Quadratic Polynomial	22.8224	0.584690	0.049941
F3	Linear	17.8254	0.773812	0.007164
	Quadratic Polynomial	17.8344	0.773500	0.007964
F4	Linear	43.8193	1.279491	0.081847
	Quadratic Polynomial	43.8543	1.278803	0.082834
F5	Linear	23.9529	1.167221	0.000021
	Quadratic Polynomial	23.3869	1.135015	0.054442
F6	Linear	24.3407	0.957960	0.135102
	Quadratic Polynomial	24.0692	0.950735	0.148100
F7	Linear	30.3935	1.064127	0.002876
	Quadratic Polynomial	30.3577	1.064035	0.002709

As presented in Table 3, the performance of the two regression models varied among the seven feeders. Quadratic polynomial regression produced slightly lower RMSE values than linear regression for all feeders, although the magnitude of improvement was relatively small. The MAPE results did not show the same pattern, as the better-performing model differed among feeders. Similarly, the R^2 values varied between the two models, with the highest R^2 value of 0.148100 obtained by the quadratic polynomial regression model for Feeder 6.

Overall, the numerical results indicate that the differences between the two regression approaches are relatively limited for most feeders.

4. DISCUSSION

The comparison between linear regression and quadratic polynomial regression indicates that increasing model complexity does not necessarily result in a substantial improvement in model performance. Although quadratic polynomial regression produced slightly lower RMSE values for all seven feeders, the improvements were relatively small. Moreover, the MAPE and R^2 results varied among feeders, indicating that the effectiveness of each regression approach depends on the individual loading characteristics of the feeder. The relatively low R^2 values obtained by both models also suggest that variations in monthly peak load cannot be fully represented by the time index alone.

These findings are consistent with previous load forecasting studies showing that model performance can vary depending on the forecasting approach and the characteristics of the load data [18], [19], [20]. The results of the present study further show that the use of a more complex regression form does not automatically provide a substantial improvement when only historical peak load and time index are considered. Therefore, model selection at the feeder level should consider not only model complexity but also the specific load variation of each feeder and multiple performance indicators.

A limitation of this study is that the model evaluation was conducted using in-sample fitting without a separate testing dataset. Therefore, the reported MAPE and RMSE values describe how well the models fit the available historical observations rather than their out-of-sample forecasting capability. In addition, the models rely solely on historical peak load and time index without incorporating external variables that may influence load variations. Future studies should consider longer observation periods, additional explanatory variables, and out-of-sample validation strategies to evaluate the predictive capability of the models more comprehensively.

5. CONCLUSION

This study evaluated linear and quadratic polynomial regression models for monthly peak load forecasting across seven distribution feeders using historical load data. The results demonstrate that forecasting performance differs among feeders due to variations in their loading characteristics. Overall, the quadratic polynomial regression model tends to yield slightly lower RMSE values, although the relative performance based on MAPE and R^2 differs from one feeder to another. The highest R^2 value was obtained for Feeder 6 using quadratic polynomial regression, with $R^2 = 0.148100$ compared with 0.135102 for linear regression. The small differences observed between the two approaches for most feeders suggest that the use of a more complex regression model does not necessarily provide a substantial improvement in forecasting accuracy. Consequently, selecting an appropriate forecasting model should account for feeder-specific load characteristics and consider MAPE, RMSE, and R^2 collectively rather than using a single performance indicator.

ACKNOWLEDGMENTS

The authors would like to thank the Department of Electrical Engineering, Institut Teknologi Adhi Tama Surabaya, for the academic support provided during this study.

REFERENCES

- [1] K. Fida, U. Abbasi, M. Adnan, and S. Iqbal, "Results in Engineering Review article A comprehensive survey on load forecasting hybrid models : Navigating the Futuristic demand response patterns through experts and intelligent systems," *Results Eng.*, vol. 23, no. June, p. 102773, 2024, doi: 10.1016/j.rineng.2024.102773.
- [2] M. Hasan *et al.*, "Energy Conversion and Management : X A state-of-the-art comparative review of load forecasting methods : Characteristics , perspectives , and applications," *Energy Convers. Manag.*, vol. 26, no. February, p. 100922, 2025, doi: 10.1016/j.ecmx.2025.100922.
- [3] X. Wu, Y. Wang, Y. Bai, Z. Zhu, and A. Xia, "Online short-term load forecasting methods using hybrids of single multiplicative neuron model , particle swarm optimization variants and nonlinear filters," *Energy Reports*, vol. 7, pp. 683–692, 2021, doi: 10.1016/j.egyr.2021.01.030.
- [4] E. O. Jnr, Y. Yevenyo, and S. Relvas, "Hybrid ensemble intelligent model based on wavelet transform , swarm intelligence and artificial neural network for electricity demand forecasting," *Sustain. Cities Soc.*, vol. 66, no. October 2020, p. 102679, 2021, doi: 10.1016/j.scs.2020.102679.
- [5] N. Elamin and M. Fukushige, "Modeling and forecasting hourly electricity demand by SARIMAX with interactions," *Energy*, vol. 165, pp. 257–268, 2018, doi: 10.1016/j.energy.2018.09.157.

- [6] Q. Dong and R. Huang, "Engineering Applications of Artificial Intelligence Short-Term Electricity-Load Forecasting by deep learning : A comprehensive survey," *Eng. Appl. Artif. Intell.*, vol. 154, no. December 2024, p. 110980, 2025, doi: 10.1016/j.engappai.2025.110980.
- [7] A. Jain and S. C. Gupta, "Evaluation of electrical load demand forecasting using various machine learning algorithms," no. June, pp. 1–18, 2024, doi: 10.3389/fenrg.2024.1408119.
- [8] D. Vlah, A. Kastrin, J. Povh, and N. Vuka, "Advanced Engineering Informatics Data-driven engineering design : A systematic review using scientometric approach," vol. 54, no. October, 2022, doi: 10.1016/j.aei.2022.101774.
- [9] A. Setiawan, "Enhancing Medium-Term Load Forecasting Accuracy in Post-Pandemic Tropical Regions : A Comparative Analysis of Polynomial Regression , Split Polynomial Regression , and LSTM Networks," 2025.
- [10] V. Y. Kondaiah, "A review on short-term load forecasting models for micro-grid application.pdf," 2022. doi: DOI: 10.1049/tje2.12151 Revised:
- [11] G. Lee, "A Regression-Based Method for Monthly Electric Load Forecasting in South Korea," *Energies*, vol. 17, p. 2 of 16, 2024, doi: doi.org/10.3390/en17235860.
- [12] B. Dhaval and A. Deshpande, "Short-term load forecasting with using multiple linear regression," *Int. J. Electr. Comput. Eng.*, vol. 10, no. 4, pp. 3911–3917, 2020, doi: 10.11591/ijece.v10i4.pp3911-3917.
- [13] K. Bayu, D. R. Henidar, F. Hermastiandi, G. P. N, and A. Nugraha, "Forecasting Electrical Energy Loads at PT Krakatau Daya Electric Using the Linear Regression Method," *J. Mechatronics Artif. Intell. Homepage*, vol. 1, no. 1, pp. 45–54, 2024, [Online]. Available: <http://ejournal.upi.edu/index.php/jmai/>
- [14] J. Wu, Y. Wang, Y. Tian, K. Burrage, and T. Cao, "Support vector regression with asymmetric loss for optimal electric load forecasting," *Energy*, vol. 223, p. 119969, 2021, doi: 10.1016/j.energy.2021.119969.
- [15] A. Kamooni *et al.*, "Machine learning based energy demand prediction ☆," *Energy Reports*, vol. 9, no. S12, pp. 171–176, 2023, doi: 10.1016/j.egyr.2023.09.151.
- [16] A. Mathew, R. Chikte, S. K. Sadanandan, S. Abdelaziz, and S. Ijaz, "Medium-term feeder load forecasting and boosting peak accuracy prediction using the PWP-XGBoost model," *Electr. Power Syst. Res.*, vol. 237, no. September, p. 111051, 2024, doi: 10.1016/j.epsr.2024.111051.
- [17] K. Misiurek and T. Olkusi, "Review of Methods and Models for Forecasting Electricity Consumption," *Energies*, vol. 18, p. 2 of 27, 2025, doi: <https://doi.org/10.3390/en18154032>.
- [18] Y. Zhou and W. Wang, "MLFGCN : short-term residential load forecasting via graph attention temporal convolution network," vol. 2023, no. September, 2024, doi: 10.3389/fnbot.2024.1461403.
- [19] P. Koukaras, "Short-Term Load Forecasting in the Greek Power Distribution System : A Comparative Study of Gradient Boosting and Deep Learning Models," pp. 1–27, 2025.
- [20] G. Tziolis *et al.*, "Short-term electric net load forecasting for solar-integrated distribution systems based on Bayesian neural networks and statistical post-processing," *Energy*, vol. 271, no. February, p. 127018, 2023, doi: 10.1016/j.energy.2023.127018.