Jurnal Teknologi dan Manajemen Industri Terapan (JTMIT) Vol. 5, No. 3, September 2026 pp. 1790 - 1797
P-ISSN: 2829-0232 E-ISSN: 2829-0038

Machine Learning-Based Daily Feeder Load Forecasting for Operation and Maintenance Strategy in 20 kV Distribution Systems

Fika Ardilla Rista[1*], Herry Nugraha[2], Widya Nita Suliyanti[3]
[1,2,3]Department of Electrical Engineering, Faculty of Engineering, Institut Teknologi PLN, Jakarta, Indonesia
Email: fikaardilla5@gmail.com

ABSTRACT

Accurate daily feeder load forecasting is essential for supporting operation and maintenance strategies in electrical distribution systems, where nonlinear and dynamic load variations may limit the performance of conventional forecasting methods. This study develops a machine learning-based framework for daily load forecasting of six 20 kV distribution feeders (PLR16–PLR21) supplied by Transformer 1 at Palur Substation, Indonesia. We used daily operational load data from January 2020 to April 2026, incorporating customer-growth information as an additional forecasting feature. We used Random Forest as the primary model and Holt–Winters Exponential Smoothing as a benchmark. We evaluated forecasting performance using MAE, RMSE, MAPE, and sMAPE. Random Forest outperformed Holt–Winters, achieving MAE of 292.45 A, RMSE of 417.85 A, and MAPE of 12.98%, compared with 720.30 A, 853.37 A, and 26.79%, respectively. The selected model was further applied to forecast feeder loading through 2028 and identify priority feeders for operation and maintenance planning. The proposed framework provides a practical, data-driven approach to feeder load forecasting and operational decision support in 20 kV distribution systems.

Keywords: *Daily Load Forecasting; Distribution Feeders; Random Forest; Holt-Winters; Power Distribution System.*

Introduction

Accurate electrical load forecasting is essential for power system operation, planning, asset management, and maintenance. In distribution networks, daily feeder load forecasting is particularly important because demand varies with customer consumption, operational conditions, and network growth. Reliable forecasts therefore support load assessment, operational prioritization, and maintenance planning [1] – [4]. The increasing complexity of load behavior has encouraged the use of machine learning alongside conventional forecasting methods. Holt–Winters Exponential Smoothing is widely used because of its simplicity and ability to model trend and seasonal patterns [5] – [8]. but its performance may decrease under nonlinear and abrupt load variations[2], [9]. Random Forest (RF), in contrast, can model nonlinear relationships and has shown effective performance in short-term load forecasting [10]–[14]. Nevertheless, feeder-level daily load forecasting using long-term operational data in Indonesian 20 kV distribution systems, particularly comparative evaluations of RF and classical statistical methods combined with priority feeder identification, remains relatively underexplored [15] – [17]. This study uses more than six years of daily operational data from six 20 kV feeders (PLR16–PLR21) supplied by Transformer 1 at Palur Substation, covering January 2020 to April 2026. Daily RST load is used as the forecasting target, with customer growth as an explanatory feature. RF is the primary model, with Holt–Winters as the benchmark, evaluated using MAE, RMSE, MAPE, and sMAPE. The best-performing model is then used to forecast future feeder loading and identify priority feeders for operation and maintenance planning [18] – [20]. The main contributions are: (1) development of a daily feeder load forecasting framework using more than six years of real operational data from six Indonesian 20 kV feeders; (2) comparative evaluation of RF and Holt–Winters for feeder-level forecasting; and (3) integration of future load projection and priority feeder identification to support operational monitoring, maintenance, and planning [21] – [23].

Research Methods

Study Area and Dataset

This study utilized daily three-phase current data from six 20 kV distribution feeders (PLR16–PLR21) supplied by Transformer 1 at Palur Substation, covering January 2020 to April 2026. For each feeder, the daily load was represented by the RST load, calculated as:

$$RST_t = R_t + S_t + T_t$$

where R_t, S_t, and T_t denote the phase currents at time t. The RST load represents the feeder loading conditions and actual operating characteristics of the distribution system [1], [2].

The dataset captures daily fluctuations, seasonal variations, and long-term load growth, with customer-growth information included as an additional forecasting feature to improve prediction and support priority feeder identification for operation and maintenance planning [24], [25].

Data Preprocessing and Partitioning

Before model development, the data were sorted chronologically and checked for duplicates, missing values, temporal continuity, and abnormal measurements. Missing values were interpolated, while abnormal measurements were inspected and corrected when necessary. The dataset was then chronologically divided into training data (January 2020–December 2023) and testing data (January 2024–April 2026) to preserve temporal characteristics, prevent data leakage, and objectively evaluate model generalization under real operating conditions [2].

Feature Engineering

Feature engineering transformed the time-series data into a supervised learning format for Random Forest forecasting. We generated lag, rolling-statistical, calendar, and customer-growth features to capture temporal dependencies, load trends, and recurring patterns. Table I presents the selected features.

Table 1. Input Features Used for Random Forest Forecasting

Lag-1	Lag-7	Lag-14	Rolling Mean-7	Rolling Mean-14	Day of Week	Day of Month	Month
Previous day load	One-week earlier load	Two-week earlier load	7-day moving average	14-day moving average	Calendar feature	Calendar feature	Calendar feature

Lag features capture short-term dependencies, while rolling statistics represent local load trends. Calendar features capture weekly and monthly patterns [10], [13]. Customer-growth information was included to represent long-term demand evolution and improve forecasting capability.

Random Forest Forecasting Model

We used Random Forest (RF) as the primary forecasting model. RF combines multiple decision trees using bootstrap aggregation (bagging) to improve prediction and reduce overfitting [10].For regression, the final prediction is obtained by averaging tree outputs:

$$\hat{y} = \frac{1}{N}\sum_{i=1}^{N} T_i(x)$$

where $\hat{y}$ represents the predicted load, N denotes the number of decision trees, and $T_i(x)$ represents the prediction generated by the $i-th$ tree. RF was selected for its ability to model nonlinear relationships under varying operating conditions [10], [11].

Holt–Winters Forecasting Model

Holt–Winters Exponential Smoothing was used as the benchmark because of its ability to represent trend and seasonal behavior with a simple forecasting structure [5] – [8]. The additive formulation:

$$l_t = \alpha y_t + (1-\alpha)(l_{t-1} + b_{t-1})$$
$$b_t = \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1}$$
$$\widehat{y_{t+h}} = l_t + hb_t$$

where l_t is the level, b_t is the trend, α and β are smoothing coefficients, and h is the forecasting horizon.

Experimental Setup and Hyperparameters

We implemented the RF model using Scikit-Learn in Python. The selected configuration used 100 estimators, unlimited tree depth, a minimum split of 2, a minimum leaf size of 1, bootstrap enabled, and random state 42. We applied this configuration consistently for model evaluation and future load forecasting to maintain reproducibility.

Table 2. Random Forest Hyperparameters

n_estimators	max_depth	min_samples_split	min_samples_leaf	bootstrap	random_state
100	None	2	1	True	42

Performance Evaluation Metrics

RF and Holt–Winters were evaluated using MAE, RMSE, MAPE, and sMAPE:

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y}_t|$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}$$

$$MAPE = \frac{100}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right|$$

$$sMAPE = \frac{100}{n}\sum_{t=1}^{n}\frac{|y_t - \hat{y}_t|}{(|y_t| + |\hat{y}_t|)/2}$$

where y_t and $\hat{y}_t$ denote actual and predicted load values. These metrics provide complementary measures of forecasting accuracy and are widely used in load forecasting studies [2], [26].

Research Workflow

The overall workflow adopted in this study is illustrated in Fig. 1.

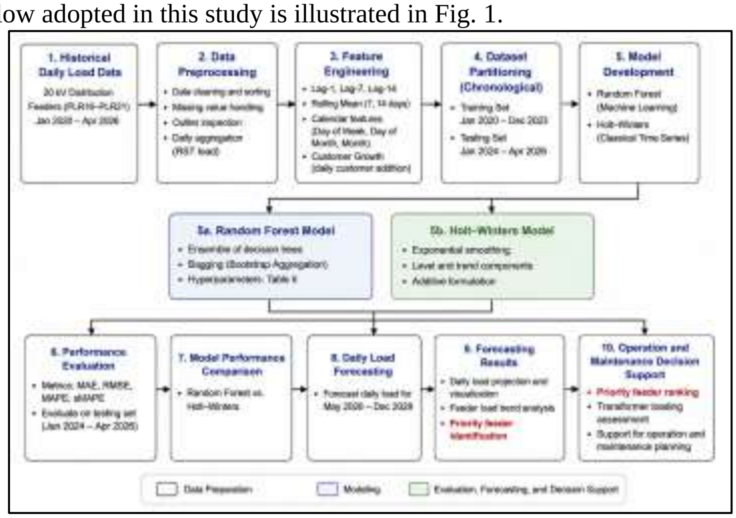

Figure 1. Overall research workflow adopted in this study

The process comprises historical data collection and preprocessing, feature engineering, model development, model evaluation, performance comparison, and future load projection. We evaluated Random Forest and Holt–Winters using the same test data, then applied the best-performing model for future feeder load forecasting and operational analysis.

Results And Discussion

Historical Load Characteristics
Daily Load Behavior

Fig. 2 shows the historical daily RST load of Transformer 1 from January 2020 to April 2026. The load exhibits substantial daily fluctuations and an overall increasing trend, reflecting variations in customer consumption and operating conditions [1], [2]. Daily load ranges from 383 A to 4363 A, indicating considerable variability and dynamic temporal behavior. These characteristics support the use of forecasting methods capable of capturing both nonlinear short-term fluctuations and long-term load evolution [10], [12], [13].

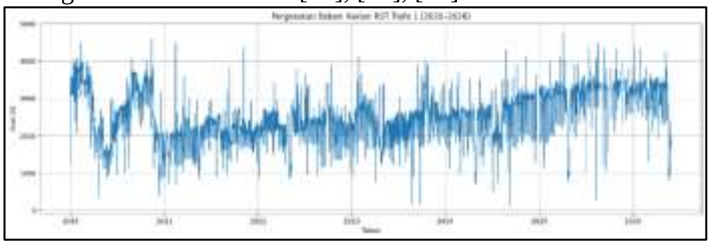

Figure 2. Historical daily RST load of Transformer 1 (2020–2026)

Statistical Characteristics

Table 3. Descriptive Statistics Of Daily Rst Load

Minimum (A)	Maximum (A)	Mean (A)	Median (A)	Standard Deviation (A)
383	4363	2505	2468	613

Table III summarizes the historical load characteristics: a minimum of 383 A, a maximum of 4363 A, a mean of 2505 A, a median of 2468 A, and a standard deviation of 613 A. The wide range and relatively high standard deviation indicate substantial load variability, supporting machine learning-based forecasting methods [2], [10].

Feeder Load Contribution Analysis

Table 4. Average Feeder Load Contribution

Feeder	Average Load (A)	Contribution (%)
PLR16	814.90	30.62
PLR19	508.17	19.09
PLR20	501.21	18.83
PLR18	356.96	13.41
PLR17	273.17	10.26
PLR21	206.92	7.77

Table IV shows the average contribution of each feeder. PLR16 is the dominant feeder with 30.62%, followed by PLR19 (19.09%) and PLR20 (18.83%). Together, these feeders account for 68.54% of total feeder demand. Their relatively high contributions make them important for load assessment, priority identification, and operation and maintenance planning.

Forecasting Performance Evaluation
Random Forest Prediction Results

Fig. 3 shows that Random Forest closely follows the actual load profile and captures most short-term fluctuations, with minor deviations during extreme loading conditions, demonstrating its ability to model nonlinear and temporal feeder load characteristics for future load projection and operational planning.

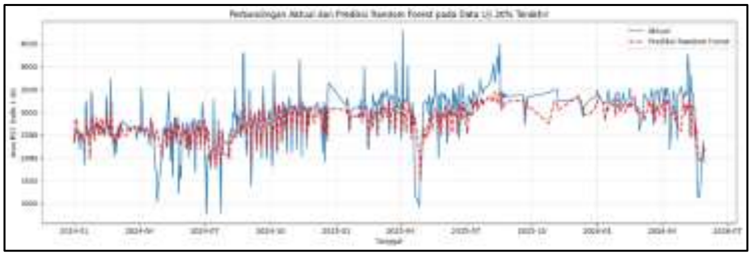

Figure 3. Actual versus predicted load using Random Forest

Holt–Winters Prediction Results

Fig. 4 shows that Holt–Winters produces smoother forecasts that capture the general trend but have limited ability to represent abrupt load variations. Consequently, larger errors occur during periods of significant load fluctuations.

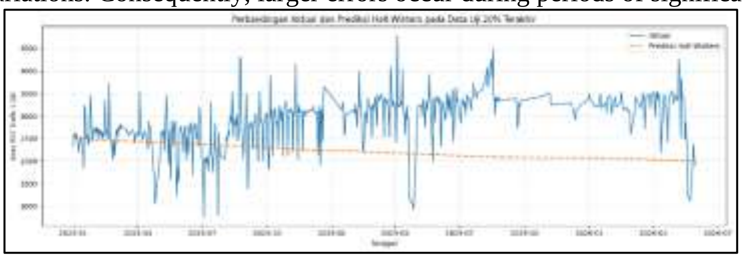

Figure 4. Actual versus predicted load using Holt–Winters

Quantitative Performance Comparison

Table 5. Forecasting Performance Comparison

Model	MAE (A)	RMSE (A)	MAPE (%)
Random Forest	292.45	417.85	12.98
Holt–Winters	720.30	853.37	26.79

Table V shows that Random Forest outperformed Holt–Winters across all metrics, achieving an MAE of 292.45 A, RMSE of 417.85 A, and MAPE of 12.98%, compared with 720.30 A, 853.37 A, and 26.79%, respectively. The superior RF performance is attributed to its ensemble structure, which uses bootstrap aggregation to capture nonlinear relationships and reduce model variance [10], [11]. Holt–Winters relies primarily on smoothing and trend representation, limiting its ability to capture highly dynamic feeder behavior [5] – [8]. The results are also consistent with previous studies reporting the effectiveness of Random Forest for short-term load forecasting[13], [14].Therefore, RF was selected for future feeder load projection and priority feeder identification.

Comparison with Previous Studies

Table 6. Comparison With Previous Studies

Study	Method	Forecasting Level
Ahmad et al. [1]	ML-based Forecasting	Smart Grid
Bielecki and Dudek [8]	Random Forest	STLF
Magalhães et al. [9]	Optimized Random Forest	Load Forecasting
This Study	Random Forest	20 kV Feeders

Table VI compares the proposed approach with previous studies. The results are consistent with previous findings on the effectiveness of Random Forest for short-term load forecasting [13], [14], while this study applies RF specifically to 20 kV distribution feeders.

Feature Importance Analysis

To better understand the Random Forest model's behavior, we performed a feature importance analysis. Fig. 5 presents the relative importance of the forecasting features.

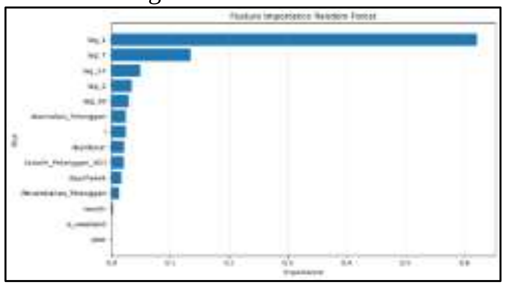

Figure 5. Random Forest feature importance ranking

Fig. 5 shows that Lag-1 and Lag-7 are the most influential forecasting features. Lag-1 indicates that the previous day’s load provides the strongest information for predicting the next daily load, while Lag-7 reflects weekly temporal dependency. Thus, short-term temporal features are more influential than calendar and customer-growth variables in the present framework, consistent with previous findings [13], [14]. Calendar and customer-growth features provide supplementary information for recurring patterns and long-term demand evolution. Overall, RF provides superior forecasting accuracy while preserving realistic feeder-level load behavior.

Forecasted Load Trend Analysis (2026–2028)

The selected RF model was applied to forecast daily feeder loads from May 2026 through December 2028. Fig. 6 shows that the forecast trajectories generally follow historical patterns without unrealistic discontinuities, indicating that the model captures the temporal structure of feeder loading. PLR16 remains the highest-loaded feeder, followed by PLR19 and PLR20, while PLR21 remains the lowest. Table VII summarizes the projected feeder contributions.

Table 7. Forecasted Feeder Load Contribution (2026–2028)

Feeder	Average Forecast Load (A)	Contribution (%)
PLR16	760.00	**27.00**
PLR19	555.00	**19.72**
PLR18	420.00	**14.92**
PLR20	410.00	**14.56**
PLR21	360.00	**12.79**
PLR17	310.00	**11.01**
Total	**2,815.00**	**100.00**

PLR16 has the highest projected loading at approximately 760 A (27.00%), followed by PLR19 at 555 A (19.72%), warranting greater attention in future monitoring and maintenance planning. The forecast indicates gradual load growth across feeders and provides a baseline based on historical load patterns and customer growth. Because weather, maintenance, and network reconfiguration were not explicitly modeled due to data limitations, treat the projections as a reference for operational planning, with future studies incorporating scenario-based analysis to assess forecast robustness.

Figure 6. Historical and forecasted daily RST load profiles of feeders PLR16–PLR21 from 2020 to 2028

Operational Implications

The forecasting results inform load assessment, resource allocation, and operational decision-making [1], [2]. Based on historical and forecasted loading contributions, PLR16, PLR19, and PLR20 require greater attention in operational monitoring and maintenance planning, with forecast information supporting inspection, preventive maintenance, and load management based on projected demand. The study is limited to historical load and customer-growth data from a single distribution area and excludes weather, economic, and social variables. Future research should incorporate these variables and customer-growth scenarios through scenario-based forecasting, and explore hybrid machine learning and deep learning approaches to improve forecasting accuracy and robustness.

Conclusion

This study developed a machine learning-based framework for daily feeder load forecasting in six 20 kV distribution feeders (PLR16–PLR21) supplied by Transformer 1 at Palur Substation, using daily operational data from January 2020 to April 2026. We used Random Forest as the primary model, with Holt–Winters Exponential Smoothing as the benchmark.

Random Forest outperformed Holt–Winters, achieving an MAE of 292.45 A, RMSE of 417.85 A, and MAPE of 12.98%. The forecasting results indicated gradual feeder load growth during 2026–2028, with PLR16 remaining the dominant contributor, followed by PLR19 and PLR20. The resulting forecasts support priority feeder identification, load monitoring, and operation and maintenance planning.

This study contributes a practical framework for feeder-level daily load forecasting in Indonesian 20 kV distribution systems by integrating model comparison, future load projection, and priority feeder identification using more than six years of real operational data. Future research may incorporate weather and economic variables, customer-growth scenarios, and hybrid forecasting architectures to improve prediction accuracy and robustness.

ACKNOWLEDGMENT

The authors would like to thank all parties who provided operational data, technical assistance, and valuable support throughout the completion of this research.

References

[1] A. Ahmad, N. Javaid, A. Mateen, M. Awais, and Z. A. Khan, "Short-Term Load Forecasting in Smart Grids: An Intelligent Modular Approach," *Energies 2019, Vol. 12, Page 164*, vol. 12, no. 1, p. 164, Jan. 2019, doi: 10.3390/EN12010164.

[2] T. Hong and S. Fan, "Probabilistic electric load forecasting: A tutorial review," *Int. J. Forecast.*, vol. 32, no. 3, pp. 914–938, Jul. 2016, doi: 10.1016/J.IJFORECAST.2015.11.011.

[3] J. W. Taylor and P. E. McSharry, "Short-term load forecasting methods: An evaluation based on European data," *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 2213–2219, Nov. 2007, doi: 10.1109/TPWRS.2007.907583.

[4] S. Fan and R. J. Hyndman, "Short-term load forecasting based on a semi-parametric additive model," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 134–141, Feb. 2012, doi: 10.1109/TPWRS.2011.2162082.

[5] P. R. Winters, "Forecasting Sales by Exponentially Weighted Moving Averages," *https://doi.org/10.1287/mnsc.6.3.324*, vol. 6, no. 3, pp. 324–342, Apr. 1960, doi: 10.1287/MNSC.6.3.324.

[6] C. C. Holt, "Forecasting seasonals and trends by exponentially weighted moving averages," *Int. J. Forecast.*, vol. 20, no. 1, pp. 5–10, Jan. 2004, doi: 10.1016/J.IJFORECAST.2003.09.015.

[7] E. S. Gardner, "Exponential smoothing: The state of the art," *J. Forecast.*, vol. 4, no. 1, pp. 1–28, Jan. 1985, doi: 10.1002/FOR.3980040103;REQUESTEDJOURNAL:JOURNAL:1099131X;WGROUP:STRING:PUBLICATION.

[8] E. S. Gardner, "Exponential smoothing: The state of the art—Part II," *Int. J. Forecast.*, vol. 22, no. 4, pp. 637–666, Oct. 2006, doi: 10.1016/J.IJFORECAST.2006.03.005.

[9] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice (3rd ed)*. Melbourne, Australia: OTexts, 2021. Accessed: Jun. 11, 2026. [Online]. Available: https://otexts.com/fpp3/

[10] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324/METRICS.

[11] G. Biau and E. Scornet, "A random forest guided tour," *TEST 2016 25:2*, vol. 25, no. 2, pp. 197–227, Apr. 2016, doi: 10.1007/S11749-016-0481-7.

[12] A. Lahouar and J. Ben Hadj Slama, "Day-ahead load forecast using random forest and expert input selection," *Energy Convers. Manag.*, vol. 103, pp. 1040–1051, Oct. 2015, doi: 10.1016/J.ENCONMAN.2015.07.041.

[13] A. Bielecki and G. Dudek, "A Comprehensive Study of Random Forest for Short-Term Load Forecasting," *Energies 2022, Vol. 15, Page 7547*, vol. 15, no. 20, p. 7547, Oct. 2022, doi: 10.3390/EN15207547.

[14] B. Magalhães, P. Bento, J. Pombo, M. do R. Calado, and S. Mariano, "Short-Term Load Forecasting Based on Optimized Random Forest and Optimal Feature Selection," *Energies 2024, Vol. 17, Page 1926*, vol. 17, no. 8, p. 1926, Apr. 2024, doi: 10.3390/EN17081926.

[15] A. Asnil, K. Krismadinata, F. Eliza, I. Husnaini, and R. Maulana, "Aplikasi IoT untuk Kendali Beban Listrik," *JTEIN: Jurnal Teknik Elektro Indonesia*, vol. 1, no. 2, pp. 207–211, Nov. 2020, doi: 10.24036/JTEIN.V1I2.78.

[16] S. Nurdiyanti and O. Candra, "Sistem Monitoring Daya Listrik 3 Fasa Berbasis IoT (Internet of Things)," *JTEIN: Jurnal Teknik Elektro Indonesia*, vol. 4, no. 2, pp. 924–933, Oct. 2023, doi: 10.24036/JTEIN.V4I2.532.

[17] S. Seabold and J. Perktold, "Statsmodels: Econometric and Statistical Modeling with Python," *Proceedings of the 9th Python in Science Conference*, pp. 92–96, 2010, doi: 10.25080/MAJORA-92BF1922-011.

[18] PedregosaFabian *et al.*, "Scikit-learn: Machine Learning in Python," *The Journal of Machine Learning Research*, Nov. 2011, doi: 10.5555/1953048.2078195.

[19] "IEEE Guide for Loading Mineral-Oil-Immersed Transformers and Step-Voltage Regulators," Dec. 2011, doi: 10.1109/IEEESTD.2012.6166928.

[20] R. Billinton and R. N. Allan, "Reliability Evaluation of Power Systems," *Reliability Evaluation of Power Systems*, 1996, doi: 10.1007/978-1-4899-1860-4.

[21] A. J. Wood, B. F. Wollenberg, and G. B. Sheblé, *Power generation, operation, and control.* Hoboken, NJ, USA: John Wiley & Sons, 2013. Accessed: Nov. 02, 2023. [Online]. Available: https://www.wiley.com/en-gb/Power+Generation%2C+Operation%2C+and+Control%2C+3rd+Edition-p-9781118733912

[22] H. L. Willis, "Spatial Electric Load Forecasting," *Spatial Electric Load Forecasting*, Aug. 2002, doi: 10.1201/9780203910764.

[23] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "The M4 Competition: 100,000 time series and 61 forecasting methods," *Int. J. Forecast.*, vol. 36, no. 1, pp. 54–74, Jan. 2020, doi: 10.1016/J.IJFORECAST.2019.04.014.

[24] C. J. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," *Clim. Res.*, vol. 30, no. 1, pp. 79–82, 2005, [Online]. Available: http://www.jstor.org/stable/24869236

[25] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *Int. J. Forecast.*, vol. 22, no. 4, pp. 679–688, Oct. 2006, doi: 10.1016/j.ijforecast.2006.03.001.

[26] A. S. Afrah *et al.*, "Comparative Study of Machine Learning and Holt-Winters Exponential Smoothing Models for Prediction of CPI's Seasonal Data," *2024 2nd International Conference on Software Engineering and Information Technology, ICoSEIT 2024*, pp. 144–148, 2024, doi: 10.1109/ICOSEIT60086.2024.10497509.