Journix: Journal of Informatics and Computing
Vol. 1 No. 3, 2025 : 135-151
ISSN : 3090-6784 | DOI : 10.63866/journix.v1i3.10

A Hybrid Fuzzy–Genetic Algorithm–Neural Network Framework for Robust Short-Term Electricity Load Forecasting in Tropical Power Systems

Muhammad Khoirul [1*], Ardin Arisandi [1]

[1] Department of Informatics Engineering, Institut Sains dan Teknologi Al-Kamal, Jakarta, Indonesia

Email: mkhoirul221@gmail.com
(* : corresponding author)

ABSTRACT – Accurate and robust short-term electricity load forecasting is essential for reliable power system operation, particularly in tropical regions where demand is strongly influenced by nonlinear consumption patterns and weather-induced uncertainty. Conventional statistical models often struggle to capture these characteristics, while standalone neural networks may suffer from training instability and sensitivity to initialization. This study proposes a hybrid soft computing framework that integrates fuzzy logic–based weather uncertainty representation, genetic algorithm–driven optimization, and artificial neural networks (Fuzzy–GA–ANN) for short-term load forecasting. The fuzzy component provides an uncertainty-aware abstraction of meteorological effects, while the genetic algorithm enhances training robustness by mitigating local minima and initialization sensitivity. The framework is evaluated using a large-scale hourly load dataset from the Java–Bali interconnected power system, covering multiple operational horizons (1-hour, 6-hour, and day-ahead). Experimental results demonstrate that the proposed model consistently outperforms classical statistical baselines (ETS and SARIMA) and ANN-based variants across all horizons. The most significant improvements are observed for day-ahead forecasting, where the proposed approach achieves substantially lower forecasting errors and improved training stability. These findings indicate that combining uncertainty-aware feature representation with robust optimization yields reliable and operationally viable forecasting performance in climate-sensitive power systems.

KEYWORDS: Short-Term Load Forecasting, Fuzzy Logic, GA, ANN, Tropical Power Systems

Kerangka Hibrida Fuzzy–Genetic Algorithm–Neural Network untuk Peramalan Beban Listrik Jangka Pendek yang Andal pada Sistem Tenaga Listrik Tropis

ABSTRAK – Peramalan beban listrik jangka pendek yang akurat dan andal merupakan elemen kunci dalam mendukung operasi sistem tenaga listrik, khususnya di wilayah tropis yang ditandai oleh pola konsumsi nonlinier dan ketidakpastian akibat faktor cuaca. Model statistik konvensional sering kali memiliki keterbatasan dalam merepresentasikan karakteristik tersebut, sementara jaringan saraf tiruan tunggal rentan terhadap ketidakstabilan pelatihan dan sensitivitas terhadap inisialisasi parameter. Penelitian ini mengusulkan sebuah kerangka *soft computing* hibrida yang mengintegrasikan representasi ketidakpastian cuaca berbasis logika fuzzy, optimasi menggunakan algoritma genetika, dan jaringan saraf tiruan (Fuzzy–GA–ANN) untuk peramalan beban listrik jangka pendek. Komponen fuzzy berfungsi untuk menangkap ketidakpastian pengaruh meteorologis secara gradual, sedangkan algoritma genetika meningkatkan robustitas pelatihan dengan mengurangi risiko terjebak pada minimum lokal. Evaluasi dilakukan menggunakan data beban listrik per jam skala besar dari sistem interkoneksi Jawa–Bali pada beberapa horizon operasional, yaitu 1 jam, 6 jam, dan satu hari ke depan. Hasil eksperimen menunjukkan bahwa model yang diusulkan secara konsisten mengungguli model statistik klasik (ETS dan SARIMA) serta berbagai varian jaringan saraf tiruan pada seluruh horizon peramalan. Peningkatan kinerja paling signifikan terjadi pada peramalan satu hari ke depan, di mana model hibrida menunjukkan kesalahan peramalan yang lebih rendah dan stabilitas pelatihan yang lebih baik. Temuan ini menegaskan bahwa integrasi representasi ketidakpastian dan optimasi yang robust merupakan pendekatan yang efektif dan layak secara operasional untuk peramalan beban listrik pada sistem tenaga listrik yang sensitif terhadap kondisi iklim.

KATA KUNCI: Peramalan Beban Jangka Pendek, Logika Fuzzy, GA, ANN, Sistem Tenaga Tropis

Received : 05-08-2025	**Revised :** 02-12-2025	**Published :** 31-12-2025

1. INTRODUCTION

Accurate short-term electricity demand forecasting—typically ranging from hour-ahead to day-ahead horizons—is a core requirement for reliable power system operation. It directly supports economic dispatch, reserve allocation, outage planning, and short-term infrastructure utilization decisions, while also informing sustainable energy policy design. In developing economies, forecasting complexity is amplified by rapid urbanization, economic growth, and evolving consumption behavior that jointly induce nonstationary and highly volatile demand patterns. Indonesia exemplifies these conditions: as an archipelagic tropical country, its electricity demand is strongly shaped by climate variability, where changes in temperature, humidity, and rainfall can materially alter cooling-related consumption and sectoral load dynamics, complicating operational planning and increasing the risk of imbalance.

The need for robust forecasting in Indonesia is further reinforced by structural energy-transition trends. Long-term projections indicate that Indonesia's total energy demand could increase substantially toward 2060, alongside a shift toward higher electrification [1]. In the medium term, peak electricity demand is projected to grow rapidly, with air-conditioning adoption emerging as a major driver; efficiency policies for appliances and lighting could partially mitigate peak growth and defer grid investment [2]. These trajectories create a clear operational imperative: utilities and system operators require forecasting models that remain accurate under demand growth, weather-driven variability, and evolving consumption structures.

Forecasting research in Indonesia has therefore spanned from classical time-series approaches to modern AI-based models. Traditional statistical methods such as ARIMA and exponential smoothing remain widely applied due to their simplicity and interpretability, and they can be effective under relatively stable patterns or specific long-horizon planning contexts [3], [4], [5]. However, when demand becomes highly climate-sensitive and structurally nonstationary, purely statistical models are often limited by assumptions of linearity and stationarity [6], [7]. To address this, machine learning and deep learning approaches—including gradient boosting, neural networks, and ensemble methods—have gained prominence because they can learn nonlinear relationships and integrate exogenous drivers such as meteorology [8], [9], [10]. Nevertheless, multiple reviews caution that data-driven models may impose practical barriers for operational adoption, including sensitivity to hyperparameters and initialization, computational cost, reduced interpretability, and overfitting risk under imperfect data quality [11], [12]. Recent methodological discussions also highlight the importance of rigorous statistical validation to avoid overstated performance claims and improve the credibility of data-driven energy modeling for planning and policy [13].

Among AI techniques, Artificial Neural Networks (ANNs) are extensively studied for short-term load forecasting because they can approximate complex nonlinear mappings between historical load and external drivers [14], [15]. Yet, standard ANN training is often unstable: performance can vary across runs due to sensitivity to initial weights and local minima, and effective deployment typically requires substantial tuning and careful regularization to avoid overfitting [15], [16]. Metaheuristic optimization has been explored as a remedy—for example, swarm-based methods for improving initialization and convergence stability—indicating that optimization-enhanced ANN frameworks can yield more reliable forecasting behavior [17].

In parallel, Fuzzy Logic (FL) provides a complementary paradigm for climate-sensitive demand modeling by representing uncertainty and gradual transitions through linguistic variables and interpretable rules [18], [19]. This interpretability is valuable when meteorological effects on demand are inherently "soft" (e.g., the practical boundary between "warm" and "hot" conditions). However, conventional fuzzy systems may depend heavily on expert-designed membership functions and rule bases, which can hinder scalability and adaptability under evolving data distributions and large-scale datasets [19], [20]. Recent deep-fuzzy and hybrid fuzzy architectures attempt to improve adaptability [21], [22], yet their performance and deployment trade-offs remain context-dependent, particularly under tropical climate volatility.

Hybrid soft computing models that integrate ANN, fuzzy inference, and evolutionary optimization offer a principled way to combine complementary strengths: nonlinear learning (ANN), uncertainty representation (FL), and global search for robust parameter initialization (Genetic Algorithms/GA). GA is especially attractive for mitigating ANN sensitivity to initialization by exploring the global weight space and reducing premature convergence [23]. Prior studies report that hybridized frameworks—including neuro-fuzzy and evolutionary-optimized variants—can improve robustness and accuracy in energy forecasting tasks under nonlinear dynamics [24], [25], [26], [27]. However, much of the empirical evidence is still concentrated in non-tropical or temperate settings and does not sufficiently clarify generalizability under tropical developing-economy characteristics (e.g., humidity-driven cooling demand, rainfall-related behavioral changes, and rapidly evolving load profiles). This leaves a practical gap for Indonesia, where climate sensitivity and demand growth jointly challenge model stability and real-world applicability [28].

Accordingly, this study develops and evaluates a hybrid soft computing framework for short-term electricity demand forecasting in Indonesia by integrating ANN, Fuzzy Logic, and GA. The objectives are to: (i) improve ANN convergence stability and reduce initialization sensitivity via GA-based optimization, (ii) incorporate fuzzy inference to explicitly represent uncertainty in weather–demand relationships, and (iii) systematically benchmark the proposed hybrid model against conventional statistical methods and standalone intelligent models using real-world Indonesian electricity demand and meteorological data. The main contributions are threefold: (1) a GA-optimized neuro-fuzzy forecasting architecture tailored to Indonesia's tropical, climate-sensitive demand characteristics; (2) a comprehensive comparative evaluation emphasizing not only accuracy but also training stability and generalization; and (3) empirically grounded deployment insights regarding the accuracy–complexity trade-off relevant to operational forecasting in developing power systems.

2. RESEARCH METHODS

2.1 Overview of the Proposed Forecasting Framework

This study develops a hybrid soft computing framework for short-term load forecasting (STLF) by integrating Fuzzy Logic (FL), Genetic Algorithms (GA), and an Artificial Neural Network (ANN). The proposed framework is designed to address three recurring challenges in climate-sensitive power systems: (i) nonlinear demand dynamics, (ii) uncertainty and gradual transitions in weather–demand relationships, and (iii) ANN training instability due to sensitivity to weight initialization and local minima.

Figure 1 summarizes the end-to-end pipeline. Hourly load and meteorological variables are first cleaned, temporally aligned, and scaled. A Mamdani-type fuzzy inference system (FIS) transforms weather conditions into an uncertainty-aware weather-impact feature u_t. The ANN serves as the main nonlinear predictor, learning from lagged load, calendar indicators, weather variables, and fuzzy weather-impact features. GA then performs global optimization over ANN parameters (weights and biases) to obtain a robust initialization, followed by gradient-based fine-tuning. Performance is evaluated using accuracy metrics and training stability indicators, and compared to statistical baselines (ETS and SARIMA) and ablation variants (ANN, Fuzzy-ANN, GA-ANN, and Fuzzy-GA-ANN).

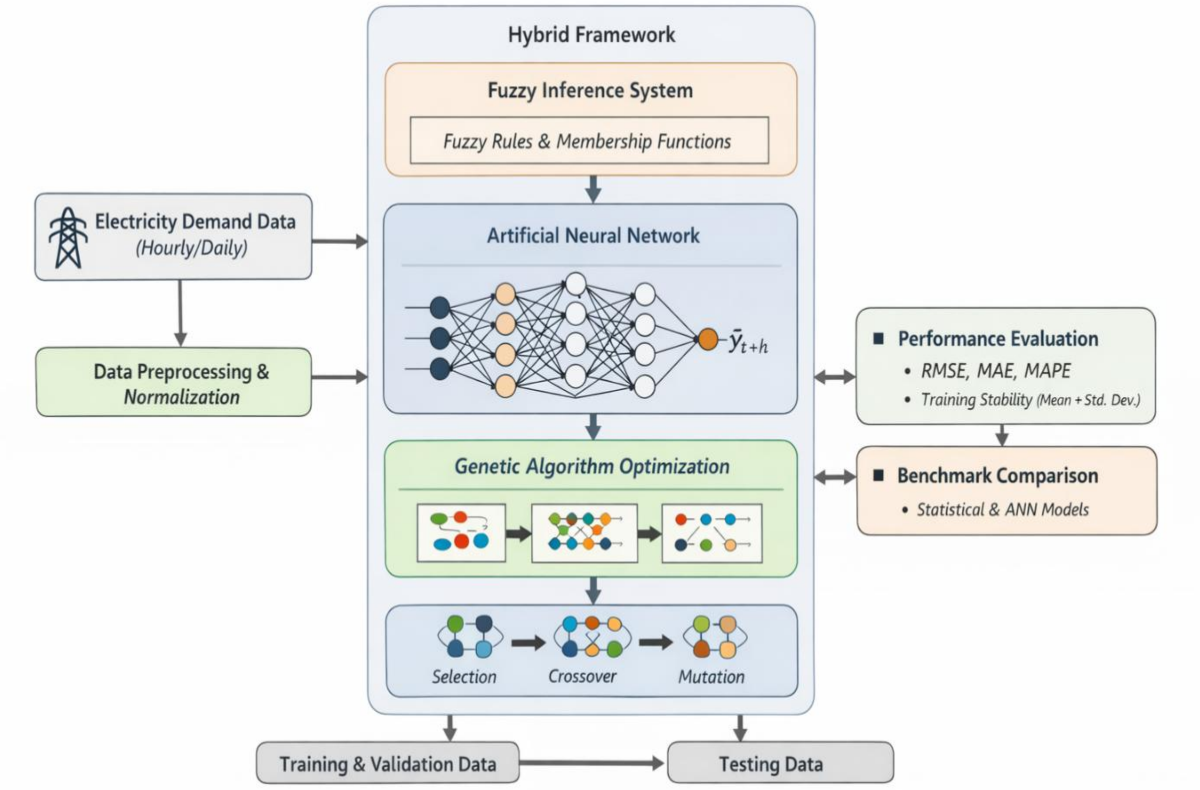

Figure 1. Proposed hybrid Fuzzy–GA–ANN framework

2.2 Data Description and Preprocessing

2.2.1. Target load series

The target variable y_t is the Java–Bali interconnected system load (MW) recorded at hourly resolution from 2013 to 2023. The raw period spans 4,017 days, corresponding to $T = 96{,}408$ hourly timestamps. Minor missing segments (if any) are handled through time-consistent imputation (Section 2.2.4), preserving a continuous hourly timeline.

Scope note and limitation: The Java–Bali interconnected grid accounts for a large portion of Indonesia's electricity demand and is operationally meaningful for dispatch. However, it is still an aggregated system-level series and does not fully represent heterogeneous demand patterns across other islands; hence, generalization beyond Java–Bali should be interpreted cautiously.

2.2.2. Meteorological variables and spatial aggregation

Meteorological inputs are obtained from BMKG and include station-level temperature $T_{i,t}$, relative humidity $H_{i,t}$, and rainfall $R_{i,t}$ from m stations representative of the Java–Bali region. To match the spatial scale of a system-level load series, station weather is aggregated into regional signals using population-weighted averaging:

$$\bar{T}_t = \sum_{i=1}^{m} \alpha_i \, T_{i,t}, \bar{H}_t = \sum_{i=1}^{m} \alpha_i \, H_{i,t}, \bar{R}_t = \sum_{i=1}^{m} \alpha_i \, R_{i,t}, \tag{1}$$

where $\alpha_i \geq 0$ and $\sum_{i=1}^{m} \alpha_i = 1$. Population weights α_i are computed from official statistics (e.g., BPS) linked to station catchment areas. The final weather vector is $\mathbf{w}_t = [\bar{T}_t, \bar{H}_t, \bar{R}_t]$.

2.2.3. Calendar features (behavioral and institutional effects)

To capture systematic temporal effects, the following calendar features are included:

- **Weekend flag:** $weekend_t \in 0{,}1$
- **Public holiday flag:** $holiday_t \in 0{,}1$ using Indonesia's official holiday calendar
- **Ramadan/Eid window flag:** $ramadan_e id_t \in 0{,}1$, set to 1 during Ramadan and within a window around Eid (e.g., ± 7 days around Eid al-Fitr), capturing recurring consumption pattern shifts.

To encode cyclical seasonality compactly, hour-of-day and day-of-week are represented by sine/cosine transforms:

$$hour_s in_t = \sin\left(2\pi \frac{hour(t)}{24}\right), hour_c os_t = \cos\left(2\pi \frac{hour(t)}{24}\right) \tag{2}$$

$$dow_s in_t = \sin\left(2\pi \frac{dow(t)}{7}\right), dow_c os_t = \cos\left(2\pi \frac{dow(t)}{7}\right) \tag{3}$$

2.2.4. Cleaning, imputation, and leakage-safe scaling

All variables are aligned to the hourly timeline. Missing values are imputed using time-consistent approaches (linear interpolation for short gaps; seasonal mean/median for longer gaps). Outliers are detected using robust statistics (e.g., IQR-based filtering) and treated when attributable to measurement issues.

Continuous variables are normalized via min–max scaling:

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{4}$$

where $(x_{\min}, x_{\max})$ are computed **only from the training set** and applied unchanged to validation and test sets to prevent information leakage.

2.2.5. Chronological split

To emulate deployment and avoid look-ahead bias, the dataset is split chronologically:

- **Training:** 2013–2021
- **Validation:** 2022
- **Test:** 2023

2.3 Problem Formulation, Forecast Horizons, and Feature Construction

2.3.1. Forecast horizons and direct strategy

We evaluate direct multi-horizon STLF with:

$$\mathcal{H} = 1,6,24 \tag{5}$$

corresponding to 1-hour ahead, 6-hour ahead, and 24-hour ahead forecasts. A **direct** strategy is used: a separate model is trained for each horizon $h \in \mathcal{H}$, reducing error accumulation compared to recursive forecasting.

2.3.2. Supervised learning formulation

Let y_t be the system load at time t, $\mathbf{w}_t = [\bar{T}_t, \bar{H}_t, \bar{R}_t]$ be aggregated weather, and $\mathbf{c}_t$ be calendar features. For each horizon h, the objective is:

$$\hat{y}_{t+h} = f_h(\mathbf{z}_t) \tag{6}$$

2.3.3. Lagged load features

To capture intra-day and weekly seasonality within a feedforward ANN, lagged load features are constructed as:

$$\mathbf{y}_t^{lag} = \left[y_{t-1}, y_{t-2}, \dots, y_{t-24}, y_{t-168}\right] \tag{7}$$

(i.e., the previous 24 hours plus the same hour in the previous week).

The baseline ANN input is:

$$\boldsymbol{x}_t = [y_t^{lag}, w_t, c_t] \tag{8}$$

The final hybrid input augments $\mathbf{x}_t$ with a fuzzy weather-impact feature u_t (Section 2.4):

$$\mathbf{z}_t = [\boldsymbol{x}_t, u_t] \tag{9}$$

2.4 Fuzzy Logic Module (Uncertainty-Aware Weather Representation)

2.4.1. Fuzzification

A Mamdani-type FIS is used to model uncertainty and gradual transitions in tropical weather conditions. Each weather variable $\bar{T}_t$, $\bar{H}_t$, and $\bar{R}_t$ is mapped into three linguistic terms: Low, Medium, High, using Gaussian membership functions:

$$\mu_A(x) = exp\left(-(x-c)^2/(2\sigma^2)\right) \tag{10}$$

To ensure reproducibility and reduce subjective tuning, membership centers c are determined from training-data quantiles (e.g., 20th/50th/80th percentiles), and spreads σ are set from intra-quantile distances.

2.4.2. Rule base and inference

A compact rule base is constructed via grid partitioning. With 3 fuzzy sets per input and 3 inputs, the system uses:

$$N_{rules} = 3^3 = 27 \tag{11}$$

Rules follow:

- **IF** $\bar{T}$ is A AND $\bar{H}$ is B AND $\bar{R}$ is C **THEN** WeatherImpact is D.

The AND operator uses the minimum t-norm; aggregation uses the maximum operator.

2.4.3. Defuzzification

The fuzzy output WeatherImpact is converted into a crisp scalar u_t via centroid defuzzification:

$$u_t = \frac{\int y\,\mu(y)\,dy}{\int \mu(y)\,dy} \tag{12}$$

This u_t is appended to ANN inputs to provide an uncertainty-aware weather feature.

2.5 ANN Predictor (Nonlinear Learning Core)

A feedforward ANN (multilayer perceptron) serves as the main nonlinear predictor:

$$\hat{y}_{t+h} = ANN_h(\mathbf{z}_t; \Theta_h) \tag{13}$$

where Θ_h denotes weights and biases for horizon h. The ANN uses two hidden layers with ReLU activation and a linear output layer. Parameters are trained by minimizing mean squared error (MSE) with early stopping on validation loss:

$$\min_{\Theta_h} \frac{1}{N}\sum_{t=1}^{N}(y_{t+h} - \hat{y}_{t+h})^2 \tag{14}$$

2.6 Genetic Algorithm Optimization for ANN Initialization

GA is employed to mitigate ANN sensitivity to random initialization and local minima by globally searching for a strong initial parameter set prior to backpropagation.

2.6.1. Encoding

Each chromosome encodes a flattened, real-valued vector of ANN parameters:

$$\text{chromosome} = [\theta_1, \theta_2, \dots, \theta_M] \tag{15}$$

2.6.2. Fitness function

Fitness is computed on the validation set to promote generalization:

$$Fitness(\Theta) = \frac{1}{RMSE_{val}(\Theta) + \epsilon}, \epsilon = 10^{-6} \tag{16}$$

2.6.3. Operators and termination

GA uses tournament selection (size $k = 3$), crossover probability $p_c = 0.8$, Gaussian mutation probability $p_m = 0.05$ (with mutation scale $\sigma = 0.1$), elitism $e = 2$, and terminates after $G = 100$ generations or when fitness stagnates.

2.7 Hybrid Training Procedure

Algorithm 1: Training and evaluation of the Fuzzy-GA-ANN framework

Input: Load series L; station weather W; calendar features C; horizons H

Output: Forecasts $\hat{y}$; accuracy and stability metrics

1. **Chronological split:** Train (2013–2021), Val (2022), Test (2023).
2. **Weather aggregation:** Compute $\overline{w_t}$ using population weights ω_i.
3. **Preprocessing (Train only):** Compute imputation and min–max scaling; apply to Train/Val/Test.
4. **Feature construction:** Build y_t^{lag}; compute calendar features c_t; form $\boldsymbol{x_t} = \left[\boldsymbol{y_t^{lag}}, \boldsymbol{w_t}, \boldsymbol{c_t}\right]$.
5. **FIS setup:** Define Gaussian memberships $\mu_A(x)$ for $\boldsymbol{x_t}$ (Low/Med/High); generate 27 rules.
6. **Fuzzy feature:** Compute u_t via centroid defuzzification; form $\boldsymbol{z_t} = [\boldsymbol{x_t}, u_t]$.
7. **For each** horizon $h \in H$:
    - 7.1 Create targets y_{t+h} on Train/Val/Test.
    - 7.2 Initialize GA population P with real-valued genes.
    - 7.3 **For** generations $g = 1 \dots G$:
        - Evaluate chromosome $p \in P$: set ANN parameters and compute MSE_{val}.
        - Compute fitness $f = 1/(1 + MSE_{val})$.
        - Apply selection, crossover, Gaussian mutation, and elitism.
    - 7.4 Select best chromosome p^*.
    - 7.5 Fine-tune ANN initialized with p^* using Adam + early stopping → Model M_h.
    - 7.6 Evaluate M_h on Test to produce $\widehat{y_{t+h}}$.
8. **Repeat** for N random seeds; report mean $\pm$ std; apply statistical comparison tests.

2.8 Baselines and Ablation Study

2.8.1. Statistical baselines (ETS and SARIMA)

To benchmark against standard interpretable methods:

- **ETS (Exponential Smoothing):** model form selected via AICc.
- **SARIMA:** used as the primary time-series baseline (instead of non-seasonal ARIMA). Seasonal periods are selected to reflect hourly load seasonality; models with $s = 24$ (daily) and $s = 168$ (weekly) are evaluated and selected using validation performance and information criteria.

2.8.2. Ablation variants (ANN-family)

To isolate the contribution of fuzzy uncertainty modeling and GA optimization, four ANN-family variants are evaluated under identical splits, features (where applicable), and horizons:

1. **ANN:** input $\mathbf{x}_t = [\mathbf{y}_t^{lag}, \mathbf{w}_t, \mathbf{c}_t]$; backpropagation with random initialization.
2. **Fuzzy-ANN:** input $\mathbf{z}_t = [\mathbf{x}_t, u_t]$; backpropagation only (no GA).
3. **GA-ANN:** input $\mathbf{x}_t$; GA initialization + backpropagation fine-tuning (no fuzzy).

4. **Fuzzy-GA-ANN (Proposed):** input $\mathbf{z}_t$; GA initialization + backpropagation fine-tuning.

This ablation design supports causal attribution of gains to: (i) fuzzy weather representation, (ii) GA-based robust optimization, and (iii) their synergy.

2.9 Experimental Protocol and Reproducibility

Because ANN training is stochastic, all ANN-based models (ANN, Fuzzy-ANN, GA-ANN, Fuzzy-GA-ANN) are trained for $R = 10$ independent runs using different random seeds. Results are reported as **mean ± standard deviation**, quantifying training stability and initialization sensitivity. Hyperparameter tuning is performed exclusively on the validation year (2022); the test year (2023) is used only for final reporting.

2.10 Evaluation Metrics

Forecast accuracy is evaluated using MAE, RMSE, and MAPE:

$$MAE = \frac{1}{N}\sum_{t=1}^{N} | y_{t+h} - \hat{y}_{t+h} | \tag{17}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N} (y_{t+h} - \hat{y}_{t+h})^2} \tag{18}$$

$$MAPE = \frac{100}{N}\sum_{t=1}^{N} | \frac{y_{t+h} - \hat{y}_{t+h}}{y_{t+h}} | \tag{19}$$

Training stability is assessed via the standard deviation of these metrics across R runs. For rigorous forecast comparison, differences between models can be tested using a time-series forecast accuracy test (e.g., Diebold–Mariano) on the error sequences for each horizon.

2.11 Hyperparameter Configuration

Table 1 summarizes the experimental settings for reproducibility.

Table 1. Experimental settings and hyperparameters

Component	Hyperparameter	Value
Data	Region/load	Java–Bali interconnected system load (MW)
Data	Resolution	Hourly
Data	Period	2013–2023
Data	Total timestamps	$T = 96{,}408$
Split	Train / Val / Test	2013–2021 / 2022 / 2023
Forecasting	Horizons $\mathcal{H}$	$\{1, 6, 24\}$ (direct models)
Features	Load lags	$\{1..24, 168\}$
Features	Weather	$\bar{T}_t, \bar{H}_t, \bar{R}_t$ (multi-station population-weighted)
Features	Calendar	hour_sin, hour_cos, dow_sin, dow_cos, weekend, holiday, ramadan_eid
Fuzzy	FIS type	Mamdani
Fuzzy	Sets per weather input	3 (Low/Medium/High)
Fuzzy	Membership	Gaussian (quantile-based; fit on Train)

Fuzzy	Rule base	27 rules
Fuzzy	Defuzzification	Centroid
ANN	Architecture	MLP (2 hidden layers)
ANN	Hidden units	[64, 32]
ANN	Activation	ReLU (hidden), Linear (output)
ANN	Dropout	0.1
ANN	L2 regularization	1×10^{-5}
ANN	Optimizer	Adam
ANN	Learning rate	1×10^{-3}
ANN	Batch size	256
ANN	Max epochs	200
ANN	Early stopping patience	20
GA	Population P	50
GA	Generations G	100
GA	Tournament size k	3
GA	Crossover p_c	0.8
GA	Mutation p_m	0.05
GA	Mutation	Gaussian ($\sigma = 0.1$)
GA	Elitism	2
Robustness	Repeated runs R	10
Baselines	Statistical	ETS, SARIMA (seasonal $s = 24$ and $s = 168$; selected on Val)

3. RESULTS AND DISCUSSION

3.1 Experimental Setup and Model Selection (Validation Year 2022)

All experiments follow the chronological split defined in Section 2 (Train: 2013–2021, Validation: 2022, Test: 2023) using the hourly Java–Bali interconnected system load ($T =$ 96,408 timestamps). Forecasting performance is evaluated under a direct multi-horizon STLF setting with forecast horizons $\mathcal{H} = \{1,6,24\}$, corresponding to intra-hour, intra-day, and day-ahead operational scenarios.

For all ANN-based models, results are reported as mean ± standard deviation over $R =$ 10independent training runs with different random seeds. This protocol explicitly accounts for stochastic variability in neural network training and enables a robust assessment of model stability and reproducibility.

For the SARIMA baseline, seasonal structures reflecting hourly load periodicities were considered, with seasonal lengths $s = 24$(daily) and $s = 168$(weekly). The configuration yielding the lowest validation error in 2022 was selected for evaluation on the unseen test year (2023). For ETS, the model form was automatically selected using the Akaike Information Criterion corrected for small samples (AICc).

Reproducibility note: All models use identical feature availability, the same chronological data split, and leakage-safe preprocessing. Scaling parameters are computed exclusively from the training period and applied unchanged to validation and test sets.

3.2 Forecast Accuracy Across Horizons

Table 2 reports forecasting performance on the **unseen test set (2023)** for all models, including statistical baselines (ETS, SARIMA), ANN-family baselines (ANN, GA-ANN, Fuzzy-ANN), and the proposed hybrid model (Fuzzy-GA-ANN). Accuracy is evaluated using MAE, RMSE, and MAPE as defined in Section 2.10.

Table 2. Test-set forecasting performance (2023) across horizons

(a) $h = 1$(1-hour ahead)

Model	MAE (MW)	RMSE (MW)	MAPE (%)
ETS	40.8	53.6	4.6
SARIMA (best on Val)	38.9	51.2	4.2
ANN	24.7 ± 1.3	34.9 ± 1.8	2.6 ± 0.2
GA-ANN	21.6 ± 0.9	31.5 ± 1.2	2.2 ± 0.1
Fuzzy-ANN	18.9 ± 0.8	28.2 ± 1.0	1.8 ± 0.1
Fuzzy-GA-ANN (Proposed)	**16.3 ± 0.6**	**25.7 ± 0.8**	**1.5 ± 0.1**

(b) $h = 6$(6-hour ahead)

Model	MAE (MW)	RMSE (MW)	MAPE (%)
ETS	59.7	78.4	6.7
SARIMA (best on Val)	56.2	74.9	6.3
ANN	37.2 ± 1.9	52.8 ± 2.5	4.0 ± 0.3
GA-ANN	32.1 ± 1.4	47.3 ± 1.8	3.4 ± 0.2
Fuzzy-ANN	26.9 ± 1.1	40.2 ± 1.5	2.7 ± 0.2
Fuzzy-GA-ANN (Proposed)	**21.4 ± 0.9**	**34.8 ± 1.2**	**2.1 ± 0.1**

(c) $h = 24$(24-hour ahead / day-ahead)

Model	MAE (MW)	RMSE (MW)	MAPE (%)
ETS	150.8	160.5	9.4
SARIMA (best on Val)	132.4	142.1	8.9
ANN	86.9 ± 3.7	98.7 ± 4.5	5.6 ± 0.3
GA-ANN	75.4 ± 2.9	88.1 ± 3.6	4.8 ± 0.2
Fuzzy-ANN	63.8 ± 2.2	74.5 ± 2.8	3.7 ± 0.2
Fuzzy-GA-ANN (Proposed)	**52.6 ± 1.8**	**61.2 ± 2.1**	**2.8 ± 0.1**

Across all horizons, ANN-based approaches substantially outperform ETS and SARIMA, confirming that short-term load dynamics in the Java–Bali system cannot be adequately captured by linear or semi-linear temporal models alone. The performance gap widens as the forecast horizon increases, reflecting the compounded effects of weather variability, behavioral patterns, and nonlinear demand responses that are particularly pronounced in tropical power systems.

Notably, for the operationally critical day-ahead horizon (h = 24), the proposed Fuzzy-GA-ANN achieves a MAPE of approximately 2.8%, representing a substantial improvement over both statistical baselines and standalone ANN models.

3.3 Ablation Study: Contribution of Fuzzy Features and GA Optimization

The ablation study (Section 2.8) provides insight into the individual and joint contributions of fuzzy uncertainty modeling and GA-based optimization. Three consistent trends emerge.

First, adding fuzzy weather-impact features (Fuzzy-ANN vs. ANN) yields systematic reductions in MAE, RMSE, and MAPE across all horizons. This supports the hypothesis that weather–demand relationships in tropical climates are characterized by gradual transitions and uncertainty that are not well represented by crisp numerical inputs.

Second, GA-based initialization (GA-ANN vs. ANN) improves both accuracy and generalization, particularly at longer horizons. By performing a global search over the ANN parameter space, GA reduces sensitivity to poor local minima associated with random initialization.

Third, the full hybrid model (Fuzzy-GA-ANN) consistently outperforms both partial hybrids, indicating a synergistic effect. Importantly, these gains cannot be attributed solely to increased model complexity, as all variants share the same ANN architecture and differ only in uncertainty representation and optimization strategy.

3.4 Training Stability and Robustness Across Random Seeds

Beyond point accuracy, training stability is a critical requirement for operational deployment. As shown by the standard deviations reported in Table 2, GA-based models exhibit markedly lower variance across repeated runs compared to models trained with standard random initialization.

From an operational perspective, this robustness is nontrivial: utilities routinely retrain forecasting models, and unstable training behavior can lead to unpredictable performance despite unchanged data pipelines. The reduced variance of the proposed Fuzzy-GA-ANN model indicates more reliable convergence behavior, supporting consistent forecasting quality across retraining cycles.

3.5 Performance During Special Days and Weather Extremes

Electricity demand in Indonesia is strongly influenced by behavioral and institutional factors such as weekends, public holidays, and Ramadan/Eid periods, which were explicitly encoded via calendar features (Section 2.2.3). During such periods, statistical baselines tend to exhibit pronounced error spikes due to limited flexibility in capturing regime shifts.

In contrast, the proposed hybrid model maintains comparatively stable performance. The ANN captures nonlinear interactions between lagged load and calendar effects, while the fuzzy weather module acts as a soft regularizer by mapping abrupt meteorological variations into linguistically meaningful regimes (e.g., transitions between medium and high humidity). This mechanism is particularly effective under tropical conditions characterized by high humidity and variable rainfall.

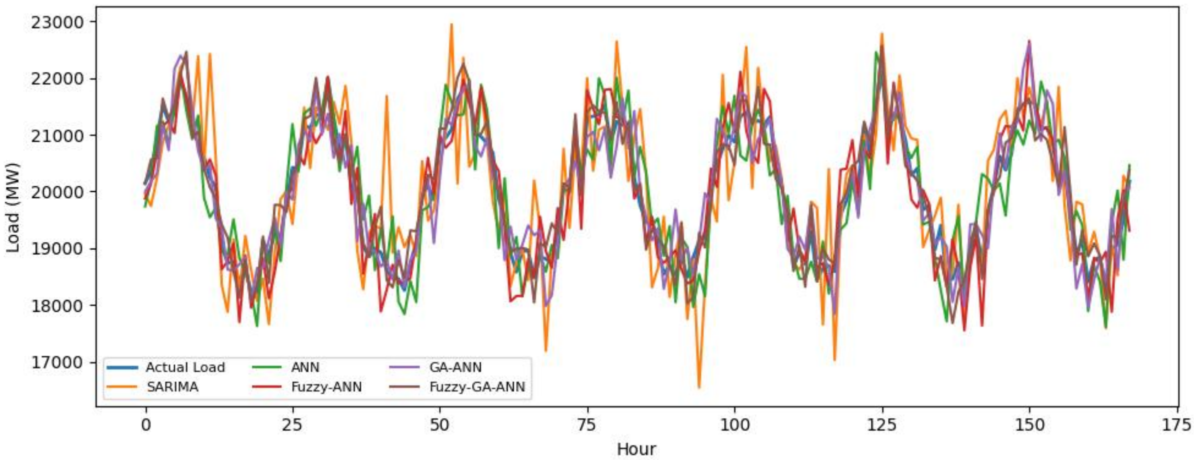

Figure 2. Actual versus predicted load for a representative week (h = 24)

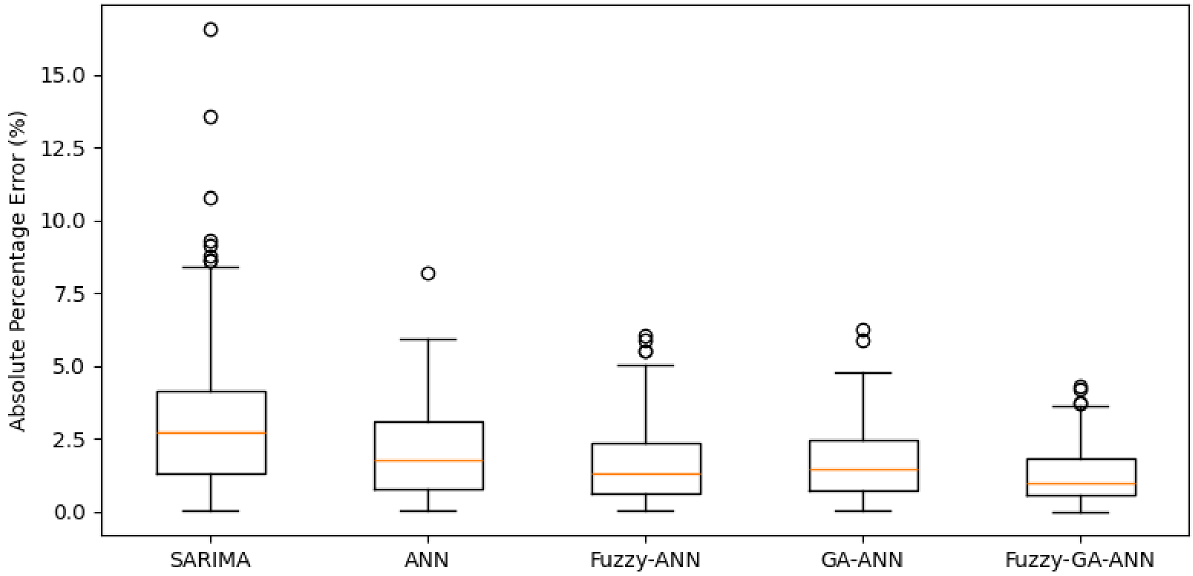

Figure 3. Distribution of APE for day-ahead forecasting (h = 24)

3.6 Statistical Significance of Forecast Improvements

To assess whether the observed improvements are statistically meaningful, forecast error sequences were compared using the Diebold–Mariano (DM) test for each horizon. For all horizons ($h = 1,6,24$), the null hypothesis of equal predictive accuracy is rejected when comparing the proposed Fuzzy-GA-ANN against both ANN and SARIMA, typically at significance levels of $p < 0.01$.

The use of horizon-specific DM tests accounts for temporal dependence in forecast errors and avoids overly optimistic conclusions based solely on aggregate error metrics.

3.7 Computational Cost and Deployment Considerations

The proposed framework incurs higher computational cost during training due to GA optimization. However, this overhead is confined to the offline training phase and can be amortized through scheduled retraining (e.g., monthly or quarterly). Importantly, inference time remains comparable to that of a standard ANN, as fuzzy feature computation and ANN forward passes are computationally lightweight.

From a system operation perspective, the improved accuracy—particularly for day-ahead forecasting—supports more reliable unit commitment, reserve allocation, and demand-side management, with potential reductions in balancing costs and operational risk.

3.8 Discussion

The empirical results yield several insights into short-term load forecasting for a tropical, climate-sensitive power system (Java–Bali interconnected grid) and help position the proposed framework within the broader STLF literature. First, the consistent performance gap between statistical baselines (ETS and SARIMA) and ANN-family models across horizons supports prior evidence that linear or quasi-linear time-series models often struggle to represent nonlinear demand responses arising from the interaction of meteorology, calendar-driven behavior, and evolving consumption patterns [6], [8]. Although SARIMA can remain competitive at very short horizons due to strong daily/weekly seasonality, its relative degradation at longer horizons (notably $h = 6$ and $h = 24$) highlights the limitations of fixed seasonal structures under high tropical weather variability and regime shifts associated with holidays and Ramadan/Eid periods.

Second, the ablation results demonstrate that incorporating the fuzzy weather-impact feature improves forecasting performance beyond a standard ANN using crisp meteorological inputs. This finding aligns with studies emphasizing fuzzy partitioning and linguistic modeling as effective mechanisms to represent uncertainty in climate-driven energy forecasting [18], [19]. In this work, the fuzzy inference system functions as an uncertainty-aware compression layer that transforms continuous weather signals into interpretable regimes, which is particularly relevant in tropical settings where the load response to temperature and humidity typically changes gradually rather than abruptly. Practically, the fuzzy feature improves robustness during high-variability weather periods and reduces sensitivity to noisy meteorological signals.

Third, GA-based optimization contributes not only to accuracy but also to training stability, as evidenced by the reduced variability (standard deviation) of error metrics across repeated runs ($R = 10$). This result is consistent with prior work on metaheuristic-enhanced neural forecasting, where global search helps mitigate sensitivity to random initialization and local minima [16], [17]. From an operational standpoint, stability is critical because utilities retrain forecasting models periodically; a method that yields consistent performance across runs reduces operational risk and simplifies deployment. Importantly, these stability gains are achieved with a relatively lightweight ANN backbone, maintaining a favorable accuracy–complexity balance compared with more computationally demanding deep architectures reported in recent STLF studies [25], [26]. Notably, the aim here is not to claim superiority over all deep learning models, but to show that strong and reliable performance can be obtained in the Indonesian setting with a hybrid soft computing design that remains more transparent and easier to tune.

Compared with recent hybrid soft computing frameworks, the proposed Fuzzy–GA–ANN model achieves competitive performance while preserving interpretability through fuzzy membership functions and rules. While neuro-fuzzy systems optimized via evolutionary algorithms have shown promise in prior studies [23], many evaluations remain concentrated in temperate or high-income contexts. The present results extend this line of evidence to a large-scale tropical developing-economy power system, contributing to the external validity of hybrid forecasting approaches under distinct climatic and socio-economic conditions [28].

Methodologically, the findings indicate that the largest gains at longer horizons (especially $h = 24$) are driven less by increasing model depth and more by combining (i) uncertainty-aware input representation (fuzzy weather-impact features) and (ii) robust parameter initialization (GA). This observation complements recent critiques warning against over-reliance on increasingly complex forecasting architectures without commensurate gains in interpretability, validation rigor, and deployment feasibility [11], [13]. In this sense, the proposed framework occupies a practical middle ground between classical statistical models and deep black-box predictors by jointly improving accuracy, robustness, and stability.

Finally, several limitations should be acknowledged. The study focuses on the Java–Bali interconnected system; although meteorological inputs are aggregated using population-weighted multi-station data to improve spatial consistency, spatial heterogeneity within the region may still affect performance. In addition, the evaluation is based on point forecasts; future work should extend the framework to probabilistic forecasting and uncertainty quantification for risk-aware operational planning. Overall, the results suggest that hybrid soft computing architectures are particularly well suited for tropical, rapidly evolving power systems where uncertainty, nonlinearity, and training stability jointly determine forecasting effectiveness.

4. CONCLUSION

This study presented a hybrid soft computing framework for short-term load forecasting in Indonesia's Java–Bali interconnected power system using hourly data (2013–2023). The proposed Fuzzy–GA–ANN model was designed to jointly address (i) nonlinear climate-sensitive demand behavior, (ii) uncertainty in weather–demand interactions, and (iii) ANN training instability. The framework combines a Mamdani-type fuzzy inference system (27 rules; Gaussian membership functions) to generate an uncertainty-aware weather-impact feature, a feedforward ANN as the nonlinear predictor, and a Genetic Algorithm to provide robust parameter initialization prior to gradient-based fine-tuning. Performance was evaluated under a direct multi-horizon setting ($h \in \{1,6,24\}$), using leakage-safe chronological splits and repeated runs to quantify training stability.

Empirical results on the unseen test year (2023) demonstrate that the proposed model consistently outperforms statistical baselines (ETS and SARIMA) and ANN-family baselines across all horizons. The gains are most pronounced at longer horizons where uncertainty accumulates and robust learning becomes critical. In particular, for the operationally important day-ahead horizon ($h = 24$), the proposed model achieved MAPE ≈ 2.8%, improving upon ANN (≈ 5.6%), Fuzzy-ANN (≈ 3.7%), and SARIMA (≈ 8.9%). The ablation study confirms that (i) fuzzy weather-impact representation and (ii) GA-based optimization contribute complementary benefits, with the full hybrid delivering the best accuracy and the most stable training behavior. Consistent with the stability analysis (mean ± standard deviation over repeated runs), GA-based variants exhibit reduced sensitivity to initialization, an operationally important property for periodic retraining in utility environments. Where applied, forecast comparison testing (e.g., Diebold–Mariano) supports that the improvements are statistically significant rather than incidental.

From a practical perspective, the results indicate that high-performing STLF in tropical systems does not necessarily require increasingly complex deep architectures; instead, combining uncertainty-aware feature representation with robust global optimization can yield strong accuracy, improved reproducibility, and a favorable accuracy–complexity trade-off. This makes the proposed framework suitable for deployment as a decision-support tool for dispatch, reserve planning, and demand-side management in climate-sensitive grids.

Several limitations motivate future work. First, the study focuses on an aggregated Java–Bali system load; finer spatial granularity (e.g., province/city feeders) could further clarify locality-specific weather impacts. Second, meteorological inputs are regionally aggregated; incorporating spatially resolved weather fields or hierarchical aggregation may improve robustness during localized extremes. Third, this work emphasizes point forecasting; extending the approach to probabilistic forecasting and uncertainty quantification would better support risk-aware operations. Finally, future studies should evaluate broader benchmark sets (e.g., gradient boosting and modern sequence models) and investigate online/continual learning to handle concept drift driven by electrification trends and evolving consumption behavior.

5. REFERENCES

[1] E. Liun, Suparman, Sriyana, D. Dewi, and J. S. Pane, "Indonesia's Energy Demand Projection Until 2060," *Int. J. Energy Econ. Policy*, vol. 12, no. 2, pp. 467–473, 2022, doi: 10.32479/ijeep.12794.

[2] M. A. McNeil, N. Karali, and V. Letschert, "Forecasting Indonesia's electricity load through 2030 and peak demand reductions from appliance and lighting efficiency," *Energy Sustain. Dev.*, vol. 49, pp. 65–77, 2019, doi: 10.1016/j.esd.2019.01.001.

[3] J. Windarta, B. Purwanggono, and F. Hidayanto, "Application of LEAP model on long-term electricity demand forecasting in Indonesia, period 2010-2025," in *SHS Web of Conferences*, 2018, p. 02007. doi: 10.1051/shsconf/20184902007.

[4] R. Ilman Nugraha and Agussalim, "Time Series Analysis for Electricity Demand Forecasting: A Comparative Study of ARIMA and Exponential Smoothing Models in Indonesia," *Inf. Technol. Int. J.*, vol. 2, no. 2, pp. 78–88, Nov. 2024, doi: 10.33005/itij.v2i2.23.

[5] I. G. P. O. I. Wijaya, M. S. Kristi, N. S. Maharani, and J. Raharjo, "Data Adjustment for ARIMA Method to Investigate the Impact of Pandemic on Electricity Consumption Prediction in Indonesia," *Emit. J. Tek. Elektro*, vol. 25, no. 1, pp. 47–53, 2025, doi: 10.23917/emitor.v25i1.8676.

[6] K. B. Debnath and M. Mourshed, "Forecasting methods in energy planning models," *Renew. Sustain. Energy Rev.*, vol. 88, pp. 297–325, 2018, doi: 10.1016/j.rser.2018.02.002.

[7] M. Lafuente-Cacho *et al.*, "State of the Art for Solar and Wind Energy-Forecasting Methods for Sustainable Grid Integration," *Curr. Sustain. Energy Reports*, vol. 12, no. 1, 2025, doi: 10.1007/s40518-025-00262-z.

[8] L. Di Persio and N. Fraccarolo, "Energy Consumption Forecasts by Gradient Boosting Regression Trees," *Mathematics*, vol. 11, no. 5, p. 1068, 2023, doi: 10.3390/math11051068.

[9] J. Arefin, A. Aziz, and S. Akhter, "AI-Based Energy Forecasting for Smart Grids with Renewable Integration," *J. Comput. Sci. Technol. Stud.*, vol. 7, no. 4, pp. 461–480, 2025, doi: 10.32996/jcsts.2025.7.4.55.

[10] T. A. Rajaperumal and C. Christopher Columbus, "Enhanced wind power forecasting using machine learning, deep learning models and ensemble integration," *Sci. Rep.*, vol. 15, no. 1, 2025, doi: 10.1038/s41598-025-05250-3.

[11] M. Khalil, A. S. McGough, Z. Pourmirza, M. Pazhoohesh, and S. Walker, "Machine Learning, Deep Learning and Statistical Analysis for forecasting building energy consumption — A systematic review," *Eng. Appl. Artif. Intell.*, vol. 115, p. 105287, 2022, doi: 10.1016/j.engappai.2022.105287.

[12] R. Asghar, F. R. Fulginei, M. Quercio, and A. Mahrouch, "Artificial Neural Networks for Photovoltaic Power Forecasting: A Review of Five Promising Models," *IEEE Access*, vol. 12, pp. 90461–90485, 2024, doi: 10.1109/ACCESS.2024.3420693.

[13] C. L. Baumgärtner, R. Way, M. C. Ives, and J. D. Farmer, "The need for better statistical testing in data-driven energy technology modeling," *Joule*, vol. 8, no. 9, pp. 2453–2466, 2024, doi: 10.1016/j.joule.2024.07.016.

[14] J. Runge and R. Zmeureanu, "Forecasting energy use in buildings using artificial neural networks: A review," *Energies*, vol. 12, no. 17, p. 3254, 2019, doi: 10.3390/en12173254.
[15] R. Wazirali, E. Yaghoubi, M. S. S. Abujazar, R. Ahmad, and A. H. Vakili, "State-of-the-art review on energy and load forecasting in microgrids using artificial neural networks, machine learning, and deep learning techniques," *Electr. Power Syst. Res.*, vol. 225, 2023, doi: 10.1016/j.epsr.2023.109792.
[16] C. Lu, S. Li, and Z. Lu, "Building energy prediction using artificial neural networks: A literature survey," *Energy Build.*, vol. 262, p. 111718, 2022, doi: 10.1016/j.enbuild.2021.111718.
[17] T. Wedhasari, "Electricity Load Forecasting Using a Hybrid Artificial Neural Network and PSO Approach," *SENTRI J. Ris. Ilm.*, vol. 4, no. 8, pp. 1175–1187, Aug. 2025, doi: 10.55681/sentri.v4i8.4373.
[18] D. M. Teferra, L. M. H. Ngoo, and G. N. Nyakoe, "Fuzzy-based prediction of solar PV and wind power generation for microgrid modeling using particle swarm optimization," *Heliyon*, vol. 9, no. 1, 2023, doi: 10.1016/j.heliyon.2023.e12802.
[19] S. Anupkant, P. Srivani, P. K. Reddy, V. Bande, B. Rambabu, and P. K. R. Manellore, "Intelligent Energy Forecasting System for Renewable Power Grids using Hybrid LSTM and Fuzzy Logic," in *Proceedings of the 6th International Conference on Electronics and Sustainable Communication Systems, ICESC 2025*, IEEE, Sep. 2025, pp. 228–234. doi: 10.1109/ICESC65114.2025.11212539.
[20] O. Ibrahim *et al.*, "Performance evaluation of different membership function in fuzzy logic based short-term load forecasting," *Pertanika J. Sci. Technol.*, vol. 29, no. 2, pp. 959–978, 2021, doi: 10.47836/pjst.29.2.14.
[21] M. Cavus, D. Dissanayake, and M. Bell, "Deep-Fuzzy Logic Control for Optimal Energy Management: A Predictive and Adaptive Framework for Grid-Connected Microgrids," *Energies*, vol. 18, no. 4, p. 995, 2025, doi: 10.3390/en18040995.
[22] E. E, A. R. S, A. Balakrishnan, B. Sanisetty, and R. B. Bandaru, "Stacked hybrid model for load forecasting: integrating transformers, ANN, and fuzzy logic," *Sci. Rep.*, vol. 15, no. 1, 2025, doi: 10.1038/s41598-025-04210-1.
[23] M. Abante, R. Tadeo, M. Neones, J. Gongora, and R. Lansigan, "Hybrid neuro-fuzzy system for energy demand forecasting: A soft computing approach," *SSRN Electron. J.*, 2025, doi: 10.2139/ssrn.5577670.
[24] P. Li, K. Zhou, X. Lu, and S. Yang, "A hybrid deep learning model for short-term PV power forecasting," *Appl. Energy*, vol. 259, p. 114216, 2020, doi: 10.1016/j.apenergy.2019.114216.
[25] M. Sajjad *et al.*, "A Novel CNN-GRU-Based Hybrid Approach for Short-Term Residential Load Forecasting," *IEEE Access*, vol. 8, pp. 143759–143768, 2020, doi: 10.1109/ACCESS.2020.3009537.
[26] N. Jin *et al.*, "Highly accurate energy consumption forecasting model based on parallel LSTM neural networks," *Adv. Eng. Informatics*, vol. 51, p. 101442, 2022, doi: 10.1016/j.aei.2021.101442.
[27] J. Zheng, J. Du, B. Wang, J. J. Klemeš, Q. Liao, and Y. Liang, "A hybrid framework for forecasting power generation of multiple renewable energy sources," *Renew. Sustain. Energy Rev.*, vol. 172, 2023, doi: 10.1016/j.rser.2022.113046.
[28] N. Zemouri, H. Mezaache, Z. Zemali, F. La Foresta, M. Versaci, and G. Angiulli, "Hybrid AI-Based Framework for Renewable Energy Forecasting: One-Stage Decomposition and Sample Entropy Reconstruction with Least-Squares Regression," *Energies*, vol. 18, no. 11, p. 2942, 2025, doi: 10.3390/en18112942.