Data-Driven Intelligence: Machine Learning Models for Short-Term Load Forecasting of a Substation Distribution Transformer

Shilpa Aralasurali Subramanya[1,2†] and **A. N. Nagashree**[1], Non-members

ABSTRACT

With the increasing complexity of modern power distribution networks and the rising demand for reliable electricity, accurate load forecasting at the distribution level has become essential for efficient energy management. This work presents a performance-focused analysis of machine learning based short-term load forecasting (STLF) of a substation distribution transformer (SDT). Two STLF methods are proposed: (i) Direct load forecasting method: directly forecasts the SDT load using a dedicated SDT load model; (ii) Feeder aggregation method: forecasts individual feeder load using feeder specific load model and aggregates them with optimized weights to estimate the SDT load. Both methods perform real and reactive power forecasting with a half-hour-ahead prediction horizon and five-minute resolution. The models are trained and tested on data collected from an actual SDT and its connected feeders, ensuring real-world relevance and reliability. The proposed STLF methods offers practical insights for utility operators and system planners by supporting better informed decision-making in short-term load management.

Keywords: Distribution feeder, Genetic algorithm optimization, Model ensembling, Neural network model, Substation distribution transformer, Tree-based model

1. INTRODUCTION

Distribution-level loads are often dominated by large consumers such as industries, commercial buildings, schools etc. These kind of large, diverse and dynamic loads of consumers at the distribution-level often results in a more unpredictable load profile [1]. Electricity demand at the distribution level is also influenced by numerous external variables such as energy prices, weather conditions, demographics and calendar-related factors [2], hence load forecasting at the distribution level is more challenging [3]. The accurate STLF at the distribution level has become further complicated [4] due to the rise in bidirectional power flow, storage, dynamic demand and the growing unpredictability of load behaviors [5].

STLF at the distribution substation level can be carried out using both classical statistical (CS) methods and machine learning (ML) methods[6], based on the required application, availability of data, complexity of the system and desired accuracy. CS methods are easier to implement and interpret, requiring less computational resource and training data. However, they may not be able to capture complex, nonlinear patterns in highly variable load data. Some of the CS methods used for STLF at distribution level include simple linear regression, polynomial regression[7], time series models like ARIMA[8][9] etc. In contrast, ML methods are capable of modeling complex relationships and nonlinearities, making them more suitable for real-world distribution systems. But ML methods require larger datasets for training and require more computational resources for implementation. The increasing digitalization of the grid and the widespread adoption of smart devices have led to access to larger datasets, making ML methods more viable for load forecasting. Some of the ML methods used are neural networks [10]–[14], support vector regression [15][16], regression tree [17] etc.

ML methods can be further extended as hybrid machine learning approaches (HMLA) and ensemble approaches (EA) to address complex forecasting tasks. The HMLA integrates different learning techniques within a single model to improve forecasting accuracy [18][19]. In contrast, EA focus on aggregating the outputs of multiple models to enhance predictive performance [20][21]. The EA uses weights to aggregate the outputs of multiple models. Traditional EA methods often rely on fixed weights or simple averaging method while aggregating the model predictions which may not always yield optimal results. The metaheuristics EA algorithms such as genetic algorithms (GA) [22]–[24], particle swarm optimization, ant colony optimization etc. can dynamically optimize the aggregation process. These algorithms explore a wide range of possible weight configurations and thus help to minimize forecasting errors and improve the quality of the aggregated prediction [25].

The STLF at the distribution level helps in carrying out effective system balancing, voltage regulation, demand-side management, reduce CO_2 emission and other critical

Manuscript received on February 16, 2026; revised on April 17, 2026; accepted on May 25, 2026. This paper was recommended by Associate Editor Komsan Hongesombut.

[1]The authors are with Research Scholar, BMS College of Engineering, Bengaluru, India.

[2]The author is Assistant Professor, K S School of Engineering and Management, Bengaluru, India.

[†]Corresponding author: shilpa.june@gmail.com

©2026 Author(s). This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 License. To view a copy of this license visit: https://creativecommons.org/licenses/by-nc-nd/4.0/.

Digital Object Identifier: 10.37936/ecti-eec.2026243.263694

functions that ensure the stability and efficiency of the distribution grid [26][27]. The STLF of the transformer provides valuable information for optimizing substation operating schemes and planning equipment maintain activities[28]. It also helps to ensure that the transformers are neither overloaded nor underutilized, thereby optimizing transformer's lifespan[29] and the overall system efficiency.

Forecasting of real power (P) helps in transformer loading assessment, peak demand management and energy consumption planning, thereby preventing overloading and improving asset utilization. Forecasting of reactive power (Q) assists in operational decisions such as capacitor bank switching, on-load tap changer control and feeder voltage profile management. It also helps in maintaining acceptable power factor levels and reducing reactive power losses in the distribution network[30]. Since transformer loading and voltage profile depend on both P and Q, accurate forecasting of both P and Q is necessary for reliable and efficient distribution system operation.

This paper focuses on STLF of a SDT and its connected feeders. The SDT is situated in Bengaluru, a city with a tropical climate and relatively stable temperatures. However, in recent years, widespread adoption of cooling systems and Electrica vehicles (EVs), increased commercialization, rising summer temperatures due to global warming have significantly raised the load on the city's electrical distribution grid. Hence in this paper, STLF has been primarily conducted with a focus on summer load patterns.

The main contribution of this work is the development and evaluation of STLF for SDT using two different modeling methods: Direct Load Forecasting Method (DLFM) and Feeder Aggregation Method (FAM). In the FAM, a GA-based weighted aggregation technique is introduced to improve SDT load estimation accuracy by optimally representing the contribution of individual feeders. GA-based weighted aggregation enables dynamic adoption to the distribution network condition. In addition, the work presents forecasting of both P and Q at feeder and SDT levels at five-minute intervals to capture detailed variations in load behavior.

2. DATA COLLECTION

2.1 Load Data Collection

Load data used in this study is from a 12.5 MVA, 66/11 kV SDT, serving a mix of residential, commercial and industrial customers. The SDT is connected to five number of 11 kV distribution feeders that supply power to the different customer segments. The data was collected from the Supervisory Control and Data Acquisition (SCADA) system during the summer months of March and April 2024. The system recorded load data at one-minute intervals for both the SDT and its associated feeders. This one-minute data was averaged and converted into five-minute time step data for further analysis, subsequently used to identify the average daily load profiles for both the SDT (Fig. 1(i)) and all feeders (Fig. 1(ii) to Fig. 1(vi)).

The SDT weekday (WD) and weekend/holiday (WE) load profiles of both P and Q are shown in Fig. 1(i). The daily load range are indicated using slashed vertical lines and the mean daily load are indicated using thick line. SDT daily WD load profile exhibits low demand during 00:00 to 06:00 hours indicating minimal residential, commercial and industrial activities. Subsequently the load demand increases in the morning called as morning peak hours from 08:00 to 11:00. Following this period, the demand begins to stabilize until 18:00. Around 18:00 load demand increases and then gradually decreases by midnight. A similar load pattern can be observed in the WE curve too. In the weekdays the typical peak load was around 10MW while in the weekend it was around 8MW.

The feeders connected to the SDT comprise F1 (commercial), F2 & F3 (industrial) and F4 & F5 (mixed). Among these, F5 contributes the largest share of the total SDT load. Fig. 1(ii) to Fig. 1(vi) shows the WD and WE load profiles of P and Q for each feeder. Feeders serving commercial/mixed loads, such as F1 and F4, show similar demand patterns on both WD and WE due to ongoing certain commercial activities, such as retail and essential services, during weekends. But for the feeders serving industrial/mixed loads such as F2, F3 and F5, the WD and WE load curves differ due to the closure/reduced operation of industrial facilities and other loads on weekends.

2.2 Weather Data Collection

Two summer months weather data, including ambient temperature (T) and relative humidity (H), have been collected at five-minute intervals using two number of Arduino IoT setups. Each setup consists of Arduino Uno boards, DHT11 humidity sensors, DS18B20 temperature sensors, Wi-Fi connectivity for data transmission, power supplies, breadboards, jumper wires and USB-B type cables. The data processing is managed through the Arduino Integrated Development Environment (IDE).

To enhance the accuracy of the readings, the Arduino IoT setups were strategically placed in the north-east and south-west directions. The average of the data from these two locations provides a more reliable estimate of outside ambient data. The sensors were installed at a height of 6 feet above ground, adhering to standard meteorological practices. The daily average T and H data collected over two months, along with their influence on SDT load are shown are shown in Fig. 2. The top plot shows the variation of P and Q, while the bottom plot shows the variation of T and H over a typical working day.

3. METHODOLOGY

3.1 Model Development

The model development includes the following steps: Feature engineering, Feature relevance analysis, ML algorithms used and Evaluation metrics used. Each of

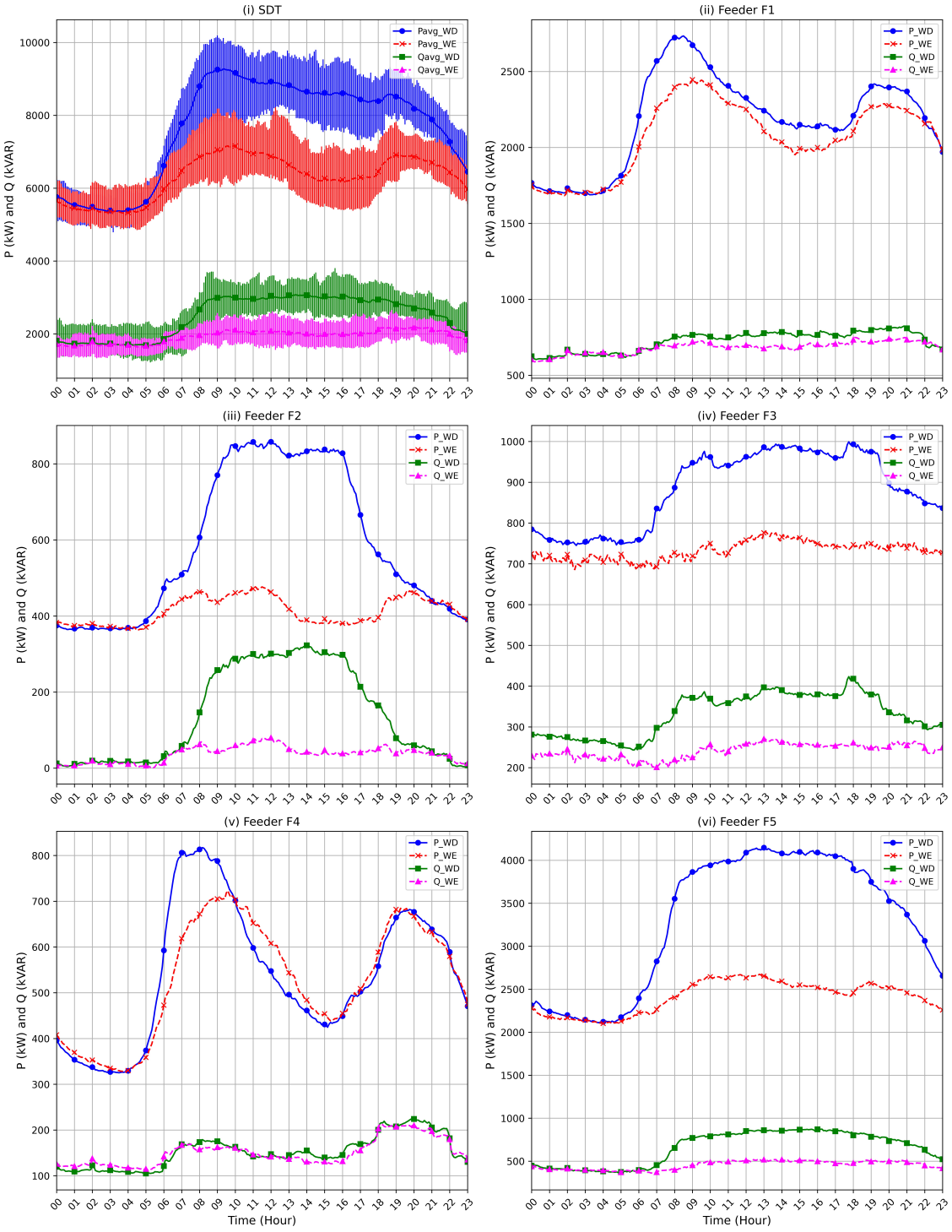

***Fig. 1**: Average daily load profiles: (i) Substation distribution transformer (ii) F1-Commerical feeder (iii) F2-Industrical feeder (iv) F3-Industrical feeder (v) F4-Mixed-type feeder (vi) F5- Mixed-type feeder.*

these sections is described in detail below.

3.1.1 Feature Engineering

Historical Short-Term Average (Psa/Qsa): This feature captures the historical average of P/Q values from the previous four working days (for WD) or two weekend/holidays (for WE) at the same time step. By averaging over multiple days at the same time step, it highlights relevant recent trends and cycles, making it useful for mitigating periodic patterns and noise in the data. For instance, to predict P/Q at 2:00 PM on a WD, the average P/Q values from 2:00 PM on the previous four working days are calculated. For WE predictions, the same time on the last two weekend days is used.

Historical Rolling Average (Pra/Sra): This feature represents the rolling average of P/Q values computed

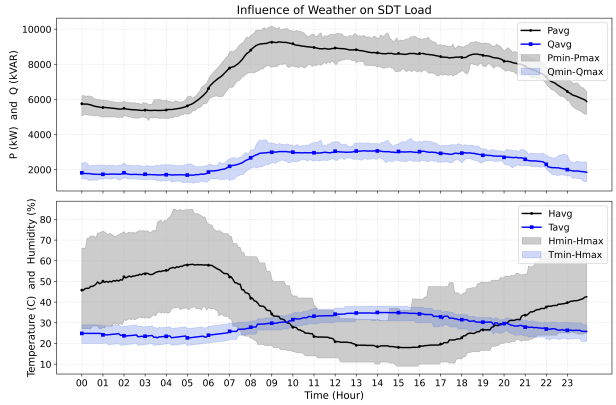

***Fig. 2**: Influence of weather on substation distribution transformer load.*

over the last 30 minutes leading up to the prediction time. It further averages this short-term rolling average over the previous four days (for WD) or two days (for WE). By averaging over multiple days at this timeframe of 30 minutes, it provides a smoothed short-term trend while reducing the impact of fluctuations. For instance, to predict P/Q at 2:00 PM on a WD, average P/Q values from 1:35 PM to 2:00 PM are considered and averaged over the last four working days. For WE predictions, the same timeframe on the last two weekend days is considered.

Lagged demand (Plag/Qlag): This refers to the lagged observation obtained from averaging P/Q values computed over the previous 30 minutes from the same day. For instance, for predicting P/Q at 2:00 on any WD/WE, the average P/Q values from 1:00 to 1:30 PM on the same day is calculated.

Additionally, the factors such as calendar inputs (time of day, day of the week and holidays), consumer behavior (commercial, industrial and residential types), T and H have an impact on load variations, as shown in Fig. 1 and Fig. 2 respectively.

3.1.2 Feature Relevance Analysis

Feature relevance analysis was carried out using both correlation matrix and feature importance analysis. Correlation matrix was used to understand the relationship between exogenous variables and load demand, while feature importance analysis evaluates the contribution of input features to the prediction model.

The corelation of P and Q with weather are shown in Fig. 3. The positive correlation between P and T, especially for SDT and feeders-F1, F3, F5 indicates that as T increases, P also increases. The negative correlation between P and H, indicates that when H increases the P reduces. Similar patterns were observed with Q-demand showing positive correlations with T and negative with H. These patterns highlight the influence of weather on load behaviour.

The feature importance analysis shown in Fig. 4 indicates that both P and Q forecasting are primarily influenced by historical short-term averages. For P forecasting, historical averages dominate, followed by lagged inputs and weather variables, reflecting both periodic consumption patterns and environmental influence. The Q forecasting exhibits relatively higher dependence on lagged inputs, indicating greater sensitivity to short-term system variations.

***Table 1**: Input data for ML models.*

Input data	Input data utilized for prediction of P	Input data utilized for prediction of Q	
Historical short-term average	Psa	Qsa	P in kW Q in kVAR
Lagged input	Plag	Qlag	
Historical rolling average	Pra	Qra	
Weather factors	T H	T H	T in 0C H in %
Calendar inputs	TOD WD WE	TOD WD WE	HH:MM 1 for WD 0 for WE

Table 1 shows a summary of the input data utilized in the development of the ML models. Historical values capture general patterns and trends, while lagged values capture specific temporal dependencies and short-term dynamics. The calendar input is considered by taking WD as 1 and WE as 0 respectively.

3.1.3 ML algorithms used

Given the diverse characteristics of the feeders and the SDT, a set of ML models have been developed for STLF task. This includes three types of tree-based models: Light Gradient Boosting Machine (LGBM), XGBoost (XGB) and Random Forest (RF), as well as two neural network models: Feedforward Neural Network (FNN) and Long Short-Term Memory (LSTM) networks.

3.1.4 Evaluation metrics used

To gauge the performance and predictive accuracy of the models, a range of evaluation metrics were employed. These metrics include: (i) Mean Absolute Error (MAE): which measures the average magnitude of errors in a set of predictions without considering their direction. (ii) Mean Absolute Percentage Error (MAPE): which is useful for understanding error in percentage terms. (iv) Coefficient of Determination (R^2): which indicates the proportion of variance explained by the model, providing insights into the model's goodness of fit. The evaluation metrics are expressed as below:

$$MAE = \frac{1}{nn}\sum_{i=1}^{nn}\left|y_i - \hat{y_i}\right| \quad (1)$$

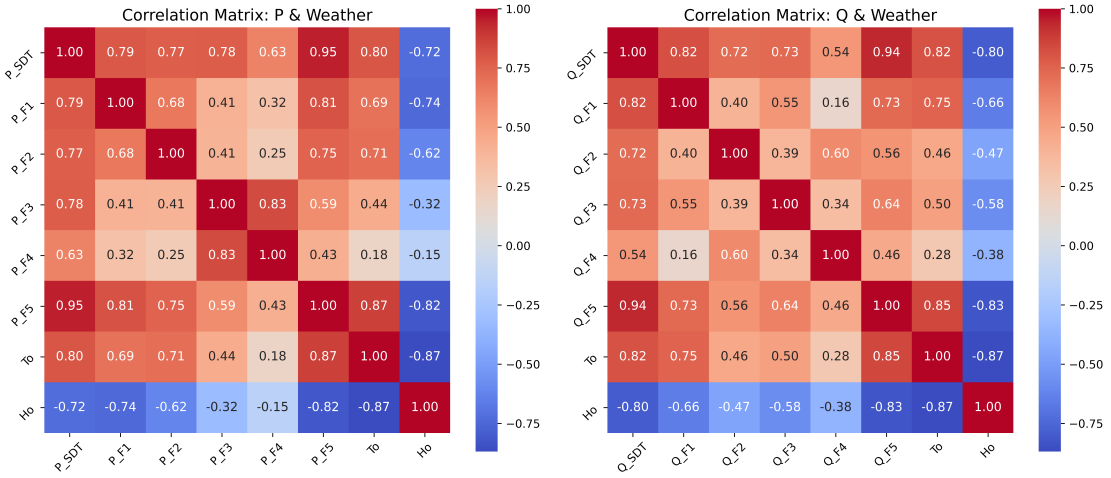

***Fig. 3**: Correlation of real and reactive power with weather variables.*

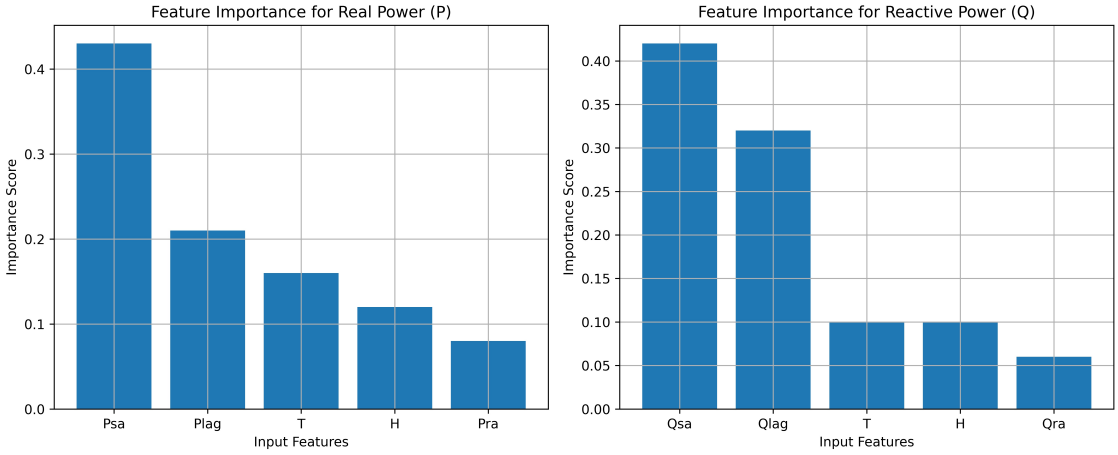

***Fig. 4**: Feature importance analysis for real and reactive power forecasting.*

$$MAPE = \frac{100\%}{nn} \int_{i=1}^{nn} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \tag{2}$$

$$\bar{y} = \frac{1}{nn} \int_{i=1}^{nn} y_i \tag{3}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{nn} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{nn} (y_i - \bar{y})^2} \tag{4}$$

Where:

y_i=Actual/observed value;

$\hat{y}_i$=Predicted value; nn=Number of observations;

$\bar{y}$=Mean of observed values

Table 2 shows the performance evaluation criterion used to assess the developed ML models and Table 3 shows the MAE performance criteria set for the work.

3.2 Short Term Load Forecasting

The collected dataset was cleaned by removal of irrelevant, duplicate and incorrect data. Abnormal data were addressed using horizontal and vertical processing

***Table 2**: Performance evaluation criterion for feeders and SDT.*

Evaluation metrics	Ideal values	Set criterion
MAE (kW/kVAR)	≈0	Based on data scale (P≤ 5% of mean load Q≤ 10% of mean load)
MAPE (%)	0% to 15%	P≤5%, Q≤10%
R2 (Ratio between 0 and 1)	≥0.9 (excellent)	≥0.9

***Table 3**: MAE performance criteria for feeders and SDT.*

MAE	F1	F2	F3	F4	F5	SDT
P (kW)	≤110	≤30	≤45	≤25	≤165	≤375
Q (kVAR)	≤70	≤15	≤30	≤15	≤70	≤225

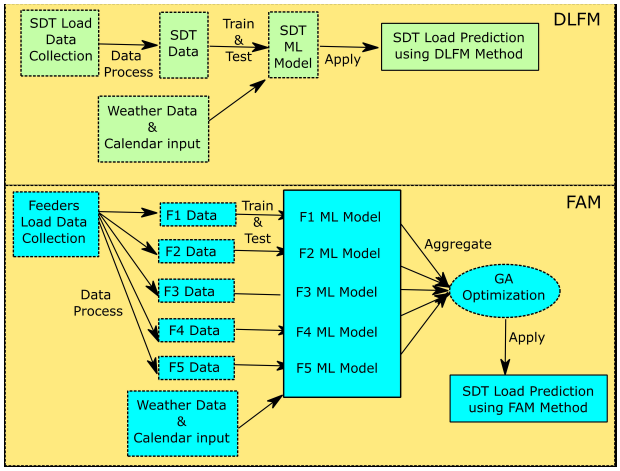

***Fig. 5**: The overall framework for short term load forecasting of a substation distribution transformer.*

techniques by employing Python libraries Pandas and NumPy. Missing values were managed using the Python Pandas interpolate (linear) method.

The dataset considered consists of 17,280 samples, with 11,808 corresponding to WD datasets and 5,472 to WE datasets. Each ML model was trained, tested and evaluated: (i) Trained using 80% of the dataset and tested on the remaining 20%. (ii) Evaluated on an unseen dataset consisting of 288 samples from April 30, 2024. The overall framework for STLF of a SDT is shown in Fig. 5 and the description of each STLF method is provided thereafter.

3.2.1 Direct load forecasting method (DLFM)

The DLFM focuses exclusively on forecasting the load of the SDT by modelling its behaviour independently, without considering the complexities of individual feeders. This approach develops a data-driven load model for the SDT using the historical, lagged and exogenous input variables described in Table 1 to directly predict the SDT's load for short-term periods.

3.2.2 Feeder aggregation method (FAM))

The FAM is an ensemble-based approach that begins with forecasting the load of individual feeders using feeder specific load model (FSLM) for the feeders. The FSLM refers to the load model for either a single feeder or a group of feeders that exhibit similar load behaviour or profiles. When feeders show comparable load characteristics, they can be grouped together, allowing a single model to represent their load behaviour. In the current work, the feeders were of different types, necessitating the development of separate load models for each individual feeder. Once individual feeder forecasts are obtained, the next step is to aggregate these forecasts to estimate the SDT load. Instead of using a simple summation, which may propagate individual feeder errors, an optimized weighted aggregation is employed. This approach minimizes overall forecasting error at the SDT level while preserving feeder-level

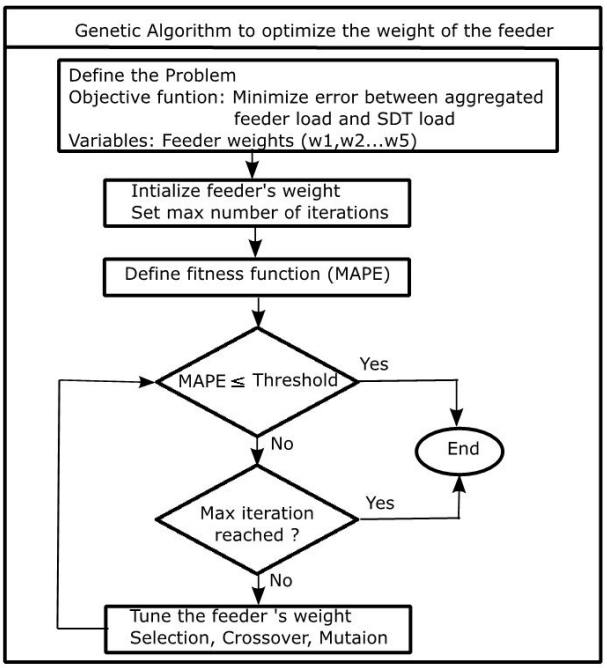

***Fig. 6**: Feeder load aggregation using GA.*

forecast accuracy.

The step-by-step procedure to determine optimal feeder weights using GA is shown in Fig. 6. The GA begins by initializing a population of individuals, where each individual represents a candidate solution comprising randomly assigned feeder weights. A fitness function formulated using MAPE evaluates the accuracy of each solution by comparing the weighted sum of feeder loads with the actual SDT load.

In the proposed GA-based aggregation framework, each candidate represents a set of feeder-specific weightages, where every feeder is assigned with its weight based on its load characteristics and type. Since the feeders connected to the SDT exhibit heterogeneous behaviour, the GA optimizes these weight values rather than treating all feeders equally. The objective is to determine a weight for each feeder such that, when the weighted feeder forecasts are aggregated, the resulting combined prediction closely matches the actual SDT load. Thus, the GA searches for the optimal feeder-weight vector that minimizes the SDT-level forecasting error.

The optimization process proceeds through iterative generations, applying three primary genetic operators: Selection, Crossover and Mutation. Selection identifies the fittest individuals based on their fitness values, ensuring that better-performing solutions have a higher probability of contributing to the next generation. Crossover combines weight vectors from two selected parents to generate offspring, thereby exploring new regions of the search space through recombination. Mutation introduces small random perturbations to individual weights, maintaining population diversity and preventing premature convergence. The process continues until either the maximum number of generations is reached

or the fitness value falls below a specified threshold, ensuring convergence toward an optimal or near-optimal set of feeder weights.

Mathematical formulation:

The aggregation is expressed as:

$$P_{SDT}(t) = \sum_{i=1}^{n} w_i P_{Fi}(t) \tag{5}$$

$$Q_{SDT}(t) = \sum_{i=1}^{n} w_i^* P_{Fi}(t) \tag{6}$$

where:

$P_{Fi}(t)$, $Q_{Fi}(t)$ = predicted real and reactive power of feeder i at time t

$P_{SDT}(t)$, $Q_{SDT}(t)$ = actual SDT real power and reactive power at time t

w_i, w_i^* = optimized weights for real and reactive power aggregation

n = number of feeders

Objective function:

The goal is to minimize the difference between the aggregated SDT forecast and the actual SDT load. To achieve this, the GA employs a fitness function based on the MAPE expressed as:

$$f(w) = MAPE\left(\sum_{i=1}^{n} P_{Fi}(t),\ P_{SDT}(t)\right) \tag{7}$$

$$f(w^*) = MAPE\left(\sum_{i=1}^{n} Q_{Fi}(t),\ Q_{SDT}(t)\right) \tag{8}$$

Constraints:

$$\sum_{i=1}^{n} w_i = 1 \qquad w_i \in [Lower\ bound,\ Upper\ bound] \tag{9}$$

The lower bound prevents any feeder from being completely ignored in the aggregation process, thereby ensuring that all feeders contribute meaningfully to the SDT-level forecast. Conversely, the upper bound of 0.7 restricts any single feeder from dominating the aggregated prediction, which promotes balanced representation across feeders and mitigates the risk of overfitting to the load profile of one particular feeder.

GA configuration:

- Population size: 100
- Max generations: 50 per run
- GA runs: 10
- Crossover probability: 0.5
- Mutation probability: 0.3
- Weight bounds: [Lower bound Upper bound] = [0.05, 0.70]
- Termination: Max generations or fitness threshold
  - $f(w)$: P threshold: MAPE $\leq$ 5%
  - $f(w^*)$: Q threshold: MAPE $\leq$ 10%

4. RESULT

The forecasting was conducted for 30^{th} April 2024, which consists of a total of 288 data points. Data from April 30th were not included in the training or testing phases of the ML model development and were entirely new to the models. The model performed half-hour ahead load forecasting, with predictions made at five-minute intervals. The evaluation metrics for SDT and feeder load forecasting using different ML models are given in Table 4.

For the DLFM (SDT load prediction: given under the column "SDT" in Table 4), LGBM model achieved the lowest MAE and MAPE values and the highest R^2 among all the considered models. The MAE and MAPE values were within the acceptable limits and the R^2 value exceeded 0.9, indicating good prediction accuracy. Although RF, XGB, LSTM and FNN produced reasonable results, their performance was comparatively inferior in terms of error metrics and goodness of fit. Therefore, LGBM was selected as the most suitable model for SDT load forecasting under the DLFM framework.

Similarly, for the FAM (feeder-level load prediction using FSLM: given under the columns "F1, F2, F3, F4 and F5" in Table 4), LGBM performed better compared to the other models for both active and reactive power forecasting across all feeders. The LGBM model consistently produced lower MAE and MAPE values while maintaining high R^2 values. Hence, LGBM was selected for feeder-level forecasting in the FAM framework.

After multiple GA runs, the best weights obtained for P and Q aggregation are:

1. w^P = [0.12, 0.08, 0.08, 0.07, 0.65]
2. w^Q = [0.44, 0.06, 0.06, 0.05, 0.39]

The weight distribution reflects the relative contribution of each feeder to the SDT load as well as the load characteristics of the connected consumers. Feeder F5, being the largest contributor to SDT load, received the highest weight for P (0.65) and a significant weight for Q (0.39). Feeder F1, being a commercial feeder has a moderate weight in P (0.12) but has the highest Q weight (0.44). Commercial buildings consume high active power due to continuous HVAC, lighting, equipment usage etc. Their reactive power demand is also substantial because of inductive loads like HVAC systems, lighting and other motor-driven equipment commonly found in commercial buildings. Feeders F2 and F3, which represent industrial loads and contribute smaller shares to the SDT demand, were assigned lower weights for both active and reactive power (0.08 for P and 0.06 for Q). The smaller weights are attributed not only to their lower contribution to the SDT load but also to the presence of power factor correction in industrial loads and their relatively stable P demand. Mixed Feeder F4, being the smallest load contributor under SDT has a moderate P weight (0.08) and minimal Q weight (0.05).

The forecasted P and Q loads for the SDT in DLFM and FAM using the chosen LGBM models along with T and H are shown in the Fig. 7. On the prediction day,

***Table 4**: Evaluation metrics for load forecasting of SDT and feeders.*

ML Model	Eval metric	P						Q					
		SDT	F1	F2	F3	F4	F5	SDT	F1	F2	F3	F4	F5
LGBM	MAE	259	108	21	25	20	66	153	40	14	17	14	32
	MAPE	1.95	4.46	5.41	3.26	5.97	1.94	3.30	4.90	2.11	7.05	7.26	4.97
	R^2	0.97	0.85	0.96	0.83	0.87	0.99	0.96	0.80	0.92	0.88	0.85	0.96
RF	MAE	266	115	21	26	22	71	166	42	15	19	14	35
	MAPE	2.03	4.71	5.48	3.32	6.41	2.07	3.45	5.25	4.66	8.11	8.58	5.40
	R^2	0.97	0.85	0.96	0.83	0.85	0.98	0.96	0.75	0.92	0.83	0.83	0.95
XGB	MAE	269	113	31	25	25	88	169	47	16	20	18	43
	MAPE	2.09	4.71	5.54	3.44	6.10	2.02	3.55	5.86	5.23	7.89	7.98	5.36
	R^2	0.97	0.84	0.96	0.82	0.86	0.98	0.96	0.77	0.92	0.85	0.88	0.95
LSTM	MAE	291	110	22	27	20	70	191	40	15	19	18.60	34
	MAPE	2.35	4.51	6.72	3.48	8.09	2.56	3.99	4.91	3.02	8.88	11.98	5.35
	R^2	0.96	0.85	0.93	0.82	0.86	0.96	0.95	0.79	0.91	0.85	0.78	0.94
FNN	MAE	313	112	40	25	31	163	213	52	14	20	21.86	49
	MAPE	2.70	4.59	9.66	3.2	11.32	5.48	4.59	6.38	3.2	9.60	14.43	8.04
	R^2	0.95	0.83	0.84	0.83	0.75	0.75	0.94	0.67	0.92	0.82	0.70	0.82

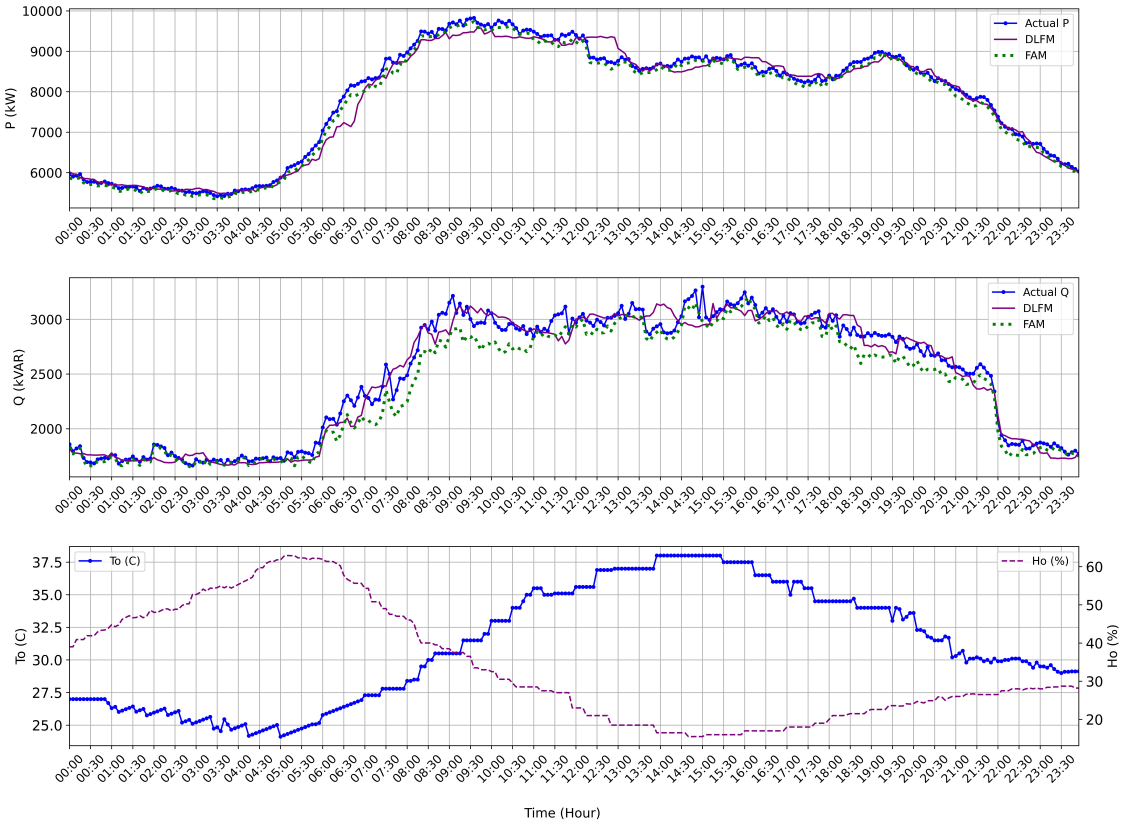

***Fig. 7**: Load on Substation distribution transformer: Feeder aggregation method (using Feeder aggregated load model) with temperature and humidity.*

the average T was approximately 32.0˚C, with a peak of 38.0˚C at 15:30 and a low of 24.18∘C at 04:15. H levels averaged around 25.5%, reaching a maximum of 61.78% at 04:55 and a minimum of 15.5% at 14:40. The evaluation

***Table 5**: Evaluation metric for DLFM and FAM on the forecasting day.*

Evaluation metric	DLFM		FAM	
	P	Q	P	Q
MAE (kW)	259	153	213	159
MAPE (%)	1.95	3.30	1.42	3.85
R^2	0.97	0.96	0.99	0.94

metrics for DLFM and FAM on the forecasting day, i.e., April 30th, are given in Table 5. The evaluation metric indicates that DLFM is more effective for predicting Q, while FAM is superior for predicting P.

DLFM is suited for predicting Q as Q tends to be more stable at the SDT level, eliminating the need to go further down at the feeder level. The SDT-level data itself effectively captured the Q demand. FAM is more effective for predicting P because each feeder's load fluctuates due to varying consumer behaviors and the aggregation of individual feeder loads offered a detailed granular view of each feeder's P contribution to the SDT. Thus, developing individual feeder-specific models and their aggregation through optimization techniques yielded more accurate P demand prediction.

5. CONCLUSION

In this paper, two machine-learning based methods for STLF, namely DLFM and FAM are presented for predicting active and reactive power of an SDT. Among the models evaluated, the LGBM resulted in most accurate performance in predicting both real and reactive power.

DLFM exhibited better performance in forecasting reactive power (MAE of 153KVAR, MAPE of 3.3% and R^2 of 0.96) compared to FAM (MAE of 159KVAR, MAPE of 3.85% and R^2 of 0.94). The FAM was more effective for forecasting real power (MAE of 213kW, MAPE of 1.42% and R^2 of 0.99) compared to DLFM (MAE of 259kW, MAPE of 1.95% and R^2 of 0.97). Overall results suggest that FAM provides more accurate and detailed forecasting, particularly when feeder-level data and substantial computational resources are available. However, in scenarios with limited resource availability or where feeder characteristics are relatively uniform, DLFM remains a viable alternative, offering satisfactory performance with reduced modeling complexity. The proposed forecasting technique may be applicable to any SDT based on the availability of the required historical and operational data.

This work can be extended as a part of future work, to develop a framework for substation-level forecasting under dynamic operating conditions (such as substation overload management, feeder outages, feeder maintenance, load transfer etc.).

FAM can provide a flexible forecasting framework as it is a feeder-level modelling. Also, SDT forecasting can be dynamically adapted by excluding the affected feeder and recalculating the aggregation weights according to the modified network configuration.

The proposed framework can be extended to multiple transformers and a topology aware load forecasting model can be developed that automatically updates aggregation weights in response to network reconfiguration events leading to robustness, scalability and practical applicability of substation-level load forecasting under real-time operating conditions.

ACKNOWLEDGEMENT

The authors are grateful to the Management, BMS Educational Trust, Principal, Vice-Principal, BMS College of Engineering for their continuous support.

DECLARATION

No potential conflict of interest was reported by the authors.

REFERENCES

[1] M. Grabner, Y. Wang, Q. Wen, B. Blažič, and V. Štruc, "A Global Modeling Framework for Load Forecasting in Distribution Networks," *IEEE Trans. Smart Grid*, vol. 14, no. 6, pp. 4927–4941, Nov. 2023, doi: 10.1109/TSG.2023.3264525.

[2] T. Chen, J. M. Lehr, O. Lavrova, and M. Martínez-Ramón, "Distribution feeder-level day-ahead peak load forecasting methods and comparative study," *IET Generation, Transmission & Distribution*, vol. 12, no. 13, pp. 3270–3278, Jul. 2018, doi: 10.1049/iet-gtd.2017.1745.

[3] X. Sun *et al.*, "An Efficient Approach to Short-Term Load Forecasting at the Distribution Level," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 2526–2537, Jul. 2016, doi: 10.1109/TPWRS.2015.2489679.

[4] M. G. Pinheiro, S. C. Madeira, and A. P. Francisco, "Short-term electricity load forecasting—A systematic approach from system level to secondary substations," *Appl. Energy*, vol. 332, p. 120493, Feb. 2023, doi: 10.1016/j.apenergy.2022.120493.

[5] J. O'Donnell and W. Su, "Attention-Focused Machine Learning Method to Provide the Stochastic Load Forecasts Needed by Electric Utilities for the Evolving Electrical Distribution System," *Energies (Basel).*, vol. 16, no. 15, p. 5661, Jul. 2023, doi: 10.3390/en16155661.

[6] T. Matijašević, T. Antić, and T. Capuder, "A systematic review of machine learning applications in the operation of smart distribution systems," *Energy Reports*, vol. 8, pp. 12379–12407, Nov. 2022, doi: 10.1016/j.egyr.2022.09.068.

[7] V. Veeramsetty, A. Mohnot, G. Singal, and S. R. Salkuti, "Short Term Active Power Load Prediction on A 33/11 kV Substation Using Regression Models," *Energies (Basel).*, vol. 14, no. 11, p. 2981, May 2021, doi: 10.3390/en14112981.

[8] O. Rivera-Caballero, A. Cogley, M. Rios, J. Gonzalez, and C. Boya-Lara, "Hierarchical Forecasting of Load Demand With Smart Meter Data for Distribution Networks," in *2022 IEEE 40th Central America and Panama Convention (CONCAPAN)*, IEEE, Nov. 2022, pp. 1–5. doi: 10.1109/CONCAPAN48024.2022.9997671.

[9] C. Tarmanini, N. Sarma, C. Gezegin, and O. Ozgonenel, "Short term load forecasting based on ARIMA and ANN approaches," *Energy Reports*, vol. 9, pp. 550–557, May 2023, doi: 10.1016/j.egyr.2023.01.060.

[10] C. Chupong, N. Junhuathon, S. Dangeam, and B. Plangklang, "Comparison of Deep Learning and Incremental Learning Model for Net Load Forecasting," *ECTI Transactions on Electrical Engineering, Electronics, and Communications*, vol. 22, no. 1, Feb. 2024, doi: 10.37936/ecti-eec.2024221.249273.

[11] 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, p. 127018, May 2023, doi: 10.1016/j.energy.2023.127018.

[12] V. Veeramsetty, D. R. Chandra, F. Grimaccia, and M. Mussetta, "Short Term Electric Power Load Forecasting Using Principal Component Analysis and Recurrent Neural Networks," *Forecasting*, vol. 4, no. 1, pp. 149–164, Jan. 2022, doi: 10.3390/forecast4010008.

[13] Y. K. Ahranjani, M. Beiraghi, and R. Ghanizadeh, "Short time load forecasting for Urmia city using the novel CNN-LTSM deep learning structure," *Electrical Engineering*, vol. 107, no. 1, pp. 1253–1264, Jan. 2025, doi: 10.1007/s00202-024-02361-4.

[14] S. AralasuraliSubramanya and A. N. Nagashree, "BUILDING COOLING: A DATA-DRIVEN APPROACH TO ASSESSING THE IMPACT OF WATER-COOLED HVAC SYSTEM DEMAND RESPONSE," *Journal of Green Building*, vol. 21, no. 1, pp. 247–272, 2026, doi: 10.3992/jgb.21.1.247.

[15] S. W. Hnin and C. Jeenanunta, "A particle swarm optimised support vector regression for short-term load forecasting," *International Journal of Energy Technology and Policy*, vol. 16, no. 4, p. 399, 2020, doi: 10.1504/IJETP.2020.107965.

[16] H. Jiang, Y. Zhang, E. Muljadi, J. J. Zhang, and D. W. Gao, "A Short-Term and High-Resolution Distribution System Load Forecasting Approach Using Support Vector Regression With Hybrid Parameters Optimization," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 3341–3350, Jul. 2018, doi: 10.1109/TSG.2016.2628061.

[17] T. Xia, Y. Zhou, S. Zhan, H. Lin, T. Zhang, and Y. Lan, "Research on short-term load forecasting of power system based on gradient lifting tree," *International Journal of Power and Energy Conversion*, vol. 13, no. 3/4, p. 235, 2022, doi: 10.1504/IJPEC.2022.130951.

[18] L. Wu, C. Kong, X. Hao, and W. Chen, "A Short-Term Load Forecasting Method Based on GRU-CNN Hybrid Neural Network Model," *Math. Probl. Eng.*, vol. 2020, pp. 1–10, Mar. 2020, doi: 10.1155/2020/1428104.

[19] L. Zhang and D. Jánošík, "Enhanced short-term load forecasting with hybrid machine learning models: CatBoost and XGBoost approaches," *Expert Syst. Appl.*, vol. 241, p. 122686, May 2024, doi: 10.1016/j.eswa.2023.122686.

[20] H. Wu and D. Levinson, "The ensemble approach to forecasting: A review and synthesis," *Transp. Res. Part C Emerg. Technol.*, vol. 132, p. 103357, Nov. 2021, doi: 10.1016/j.trc.2021.103357.

[21] M. H. D. M. Ribeiro, R. G. da Silva, G. T. Ribeiro, V. C. Mariani, and L. dos S. Coelho, "Cooperative ensemble learning model improves electric short-term load forecasting," *Chaos Solitons Fractals*, vol. 166, p. 112982, Jan. 2023, doi: 10.1016/j.chaos.2022.112982.

[22] R. Sikora and O. Al-Laymoun, "A Modified Stacking Ensemble Machine Learning Algorithm Using Genetic Algorithms," in *Artificial Intelligence*, IGI Global, pp. 395–405. doi: 10.4018/978-1-5225-1759-7.ch016.

[23] P. Chauhan and A. Swami, "Breast Cancer Prediction Using Genetic Algorithm Based Ensemble Approach," in *2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, IEEE, Jul. 2018, pp. 1–8. doi: 10.1109/ICCCNT.2018.8493927.

[24] J. Abimbola, D. Kostrzewa, and P. Kasprowski, "Genetic Algorithm-Based Optimization ofWeighted Voting in Ensemble Models for Time Signature Detection," in *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, New York, NY, USA: ACM, Jul. 2024, pp. 73–74. doi: 10.1145/3638530.3664083.

[25] S. Boriratrit, C. Srithapon, P. Fuangfoo, and R. Chatthaworn, "Metaheuristic Extreme Learning Machine for Improving Performance of Electric Energy Demand Forecasting," *Computers*, vol. 11, no. 5, p. 66, Apr. 2022, doi: 10.3390/computers11050066.

[26] N. Nigar, M. K. Shahzad, H. M. Faisal, S. A. Ajagbe, and M. O. Adigun, "Improving Demand-Side Energy Management With Energy Advisor Using Machine Learning," *Journal of Electrical and Computer Engineering*, vol. 2024, no. 1, Jan. 2024, doi: 10.1155/2024/6339681.

[27] N. Mounir, H. Ouadi, and I. Jrhilifa, "Short-term electric load forecasting using an EMD-BI-LSTM approach for smart grid energy management system," *Energy Build.*, vol. 288, p. 113022, Jun. 2023, doi: 10.1016/j.enbuild.2023.113022.

[28] H. Liu, Y. Wang, C. Wei, J. Li, and Y. Lin, "Two-Stage Short-Term Load Forecasting for Power Transformers Under Different Substation Operat-

ing Conditions," *IEEE Access*, vol. 7, pp. 161424–161436, 2019, doi: 10.1109/ACCESS.2019.2951422.

[29] S. M. A. Al-Alawi, S. M. Islam, J. Hurley, and S. Dureya, "An intelligent transformer load estimation model using artificial neural networks," *Australian Journal of Electrical and Electronics Engineering*, vol. 3, no. 1, pp. 27–36, Jan. 2006, doi: 10.1080/1448837X.2006.11464142.

[30] G. G. Tolun, Ö. C. Tolun, and K. Zor, "Advanced machine learning algorithms for reactive power forecasting in electric distribution systems," *e-Prime - Advances in Electrical Engineering, Electronics and Energy*, vol. 13, p. 101019, Sep. 2025, doi: 10.1016/j.prime.2025.101019.

Shilpa Aralasurali Subramanya Research Scholar, BMS College of Engineering, Bengaluru, India. M.Tech. in Power Electronics Research Interests: AI & ML, Power Electronics, Demand Side Management, Renewable Energy, Energy Conservation and Energy Auditing. Google Scholar: pdlwoiwAAAAJ

A. N. Nagashree Professor, BMS College of Engineering, Bengaluru, India Ph.D in High Voltage Engineering/ Dielectrics (Solid & Liquid) Research Interests: Insulation Diagnostics and High Voltages, Renewable Energy, Energy Conservation and Energy Auditing.