Energy Consumption Forecasting in Government Buildings Using Machine Learning

AKALYA P[1], DR. HARIPRIYA V[2]

[1]*Dept. of Computer Science and Information Technology Jain (Deemed-to-be University) Jayanagar 9th Block Campus, Bengaluru,*

[2]*Assistant Professor, Department of Computer Science & IT, Jain (Deemed-to-be) University Bangalore*

Abstract- Government buildings represent a significant yet underexplored domain in machine learning-based energy forecasting. Existing studies predominantly target residential or commercial facilities, leaving a critical contextual gap for public sector infrastructure, which operates under rigid occupancy schedules, multi-department load diversity, and policy-driven operational constraints. This paper presents a structured review of 15 IEEE-indexed publications (2021–2026) on ML-based building energy consumption forecasting, analysing methodologies, datasets, model architectures, and key limitations. Thematic classification identifies five major research directions: classical regression ML, deep learning and LSTM-based approaches, hybrid and ensemble methods, IoT-integrated systems, and neuro-fuzzy approaches. Systematic gap analysis reveals ten critical deficiencies — including absence of government building datasets, exclusion of occupancy features, lack of real-time deployment validation, limited interpretability, and insufficient multi-department modelling — and maps targeted research questions and objectives. A novel intelligent occupancy-aware ML-based energy consumption forecasting framework is proposed, aimed at improving forecast accuracy, enabling near-real-time energy management, and supporting data-driven sustainability decisions in government infrastructure.

Keywords— Energy Consumption Forecasting, Government Buildings, Machine Learning, LSTM, Random Forest, Hybrid Ensemble, Occupancy-Aware Features, Smart Buildings, Time-Series Forecasting, Sustainable Energy.

I. INTRODUCTION

A. Background of the Study

Government buildings — including administrative offices, public hospitals, educational institutions, and legislative complexes — represent one of the most significant and controllable sources of public sector energy expenditure. With increasing mandates for sustainability and growing pressure on government bodies to reduce carbon footprints, accurate energy consumption forecasting has become essential for efficient resource management and policy compliance.

Unlike residential or private commercial buildings, government facilities operate under rigid, policy-driven schedules. Consumption patterns are shaped by fixed working hours, departmental activity profiles, public holiday calendars, and seasonal legislative or academic cycles. These structural characteristics create predictable yet complex demand profiles that general-purpose ML models — trained predominantly on household or commercial data — are not designed to capture.

Advances in Machine Learning (ML) and Deep Learning (DL) have yielded powerful tools for time-series energy forecasting. Models such as Random Forest, XGBoost, LSTM, and hybrid ensemble architectures have demonstrated strong performance across multiple building types. However, their specific application to government buildings, with occupancy-aware feature sets and multi-department load profiles, remains largely underexplored in the existing literature.

B. Problem Statement

Existing ML forecasting models are largely trained and validated on residential or private commercial datasets. Most models exclude occupancy-driven features — such as personnel counts, departmental schedules, and working-hour patterns — despite these being among the strongest predictors of energy consumption in office environments. Additionally, real-time forecasting pipelines have rarely been validated in public sector settings, and forecasting outputs are almost never connected to operational energy management decisions.

There is therefore a pressing need for an intelligent, occupancy-aware, interpretable ML framework that specifically addresses the operational characteristics of government buildings and provides actionable forecasting outputs for energy managers.

C. Motivation
The motivation of this work arises from a consistent gap observed across all reviewed publications: high model accuracy on residential or commercial datasets paired with an absence of government building-specific validation, occupancy feature integration, and real-time deployment. Accurate energy consumption forecasting in government buildings can directly support budget planning, demand response activation, and long-term sustainability reporting — objectives that are uniquely important in the public sector context.

D. Objectives of the Study
The objectives of this research are:

- Analyse historical energy consumption patterns and temporal trends in government buildings.
- Integrate occupancy and operational schedule features into ML forecasting models.
- Develop and compare ML, DL, and hybrid models for government building energy forecasting.
- Apply SHAP-based explainability to ensure model outputs are interpretable for energy managers.
- Design a near-real-time forecasting pipeline validated against government building data.
- Connect forecasting outputs to automated reorder and energy management decisions.

E. Contributions of the Paper
The main contributions of this work are: A structured comparative review of 15 IEEE publications (2021–2026) on ML-based building energy forecasting. Identification of ten primary research gaps specific to government building energy forecasting. Targeted research questions and objectives mapped to the identified gaps. A proposed end-to-end intelligent occupancy-aware ML forecasting framework with SHAP explainability and near-real-time integration for government infrastructure.

II. II.LITERATURE REVIEW

A. Classical ML and Regression Models
Classical supervised learning methods form the most widely studied category. Kumar et al. [1] demonstrated a Random Forest system achieving R^2=0.976 on smart building IoT sensor data, integrating prediction with energy-saving recommendations. Liu et al. [2] performed the most rigorous multi-model comparison in the reviewed literature — nine algorithms on a UK government dataset of 5,000 buildings — finding that MLP with Mutual Information feature selection (52→8 features) achieved $R^2\approx0.98$, the highest reported in this category. Bhamare et al. [6] confirmed SVM as the strongest regressor (R^2=0.986) among five models on Kaggle building data with GridSearch tuning. Across studies, Random Forest consistently emerges as the most reliable baseline algorithm [1][8][10][11].

B. Deep Learning Approaches
LSTM-based models dominate recent DL work. Vignesh et al. [13] applied a four-layer stacked LSTM on Finnish hourly data, achieving strong temporal fit but excluding occupancy and weather features. Chauhan and Bansal [3] found Bi-LSTM outperforming CNN, SVR, and Gradient Boosting on Indian smart meter data (R^2=0.941, RMSE=2.38). Upadhyay et al. [4] provided a systematic review confirming that DL models achieve R^2 up to 0.97 and MAPE below 6%, though without independent experimental validation. These findings establish LSTM-based architectures as the strongest performers for long-horizon temporal forecasting.

C. Hybrid and Ensemble Methods
Siranjeevi [15] developed a two-layer stacking architecture — RF + XGBoost + LSTM with Linear Regression as meta-learner — achieving R^2=0.957 on meteorological and power data, outperforming all individual base models. This hybrid approach directly motivates the ensemble architecture proposed in this work. Thakur et al. [5] confirmed LightGBM

© MAY 2026 | IRE Journals | Volume 9 Issue 11 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I11-1717926

(R^2=0.946) outperforms CatBoost and XGBoost on large temporal datasets, supporting gradient boosting as a strong ensemble component for energy forecasting.

D. IoT-Integrated Systems

Leo Raju et al. [12] deployed an IoT-ML system with NodeMCU sensors on a department building, finding that Linear Regression surprisingly outperformed RF and SVM in the real-time sensor-data context — demonstrating that data granularity can override model complexity advantages in hardware-constrained deployments. This finding has direct implications for government buildings where legacy sensor infrastructure may limit available data quality.

E. Review and Survey Studies

Srivastava et al. [23] confirmed that deep learning models generally achieve MAPE below 3% compared to statistical methods, and identified Explainable AI and edge IoT deployment as critical future directions. Chauhan et al. [16] found LSTM superior for long-horizon forecasting while statistical models retained competitiveness for short-term predictions — an important finding for government buildings where both daily scheduling and annual budget planning are relevant forecasting horizons.

III. III.COMPARATIVE ANALYSIS OF EXISTING METHODS

Table I provides a comparison of the effectiveness and methodology aspects of various machine learning techniques used for building energy consumption forecasting. It shows that ensemble and hybrid architectures achieve superior prediction accuracy, while occupancy and operational features remain systematically absent from most reviewed frameworks.

TABLE I
Comparative Summary of Machine Learning Approaches for Building Energy Consumption Forecasting

Ref	Authors (Year)	Application Domain	ML Models Used	Key Findings
[1]	Kumar et al. (2023)	Smart building energy forecasting	Random Forest + IoT sensors	R^2=0.976, MAE=22.07 — RF + real-time IoT achieves strong accuracy
[2]	Liu et al. (2024)	Government building energy prediction	MLP, RF, SVM, 9 models + MutInfo FS	MLP: R^2≈0.98 — best on UK govt. dataset; feature selection reduces 52→8 features
[3]	Chauhan & Bansal (2025)	Residential smart meter forecasting	Bi-LSTM, CNN, SVR, Gradient Boosting	Bi-LSTM: R^2=0.941, RMSE=2.38 — DL outperforms classical ML on residential data
[4]	Upadhyay et al. (2026)	Smart grid energy prediction (review)	LSTM, CNN, Ensemble models	R^2≤0.97, MAPE<6% — DL ensemble methods consistently outperform statistical baselines
[5]	Thakur et al. (2023)	Electricity demand forecasting	XGBoost, LightGBM, CatBoost	LightGBM: R^2=0.946 — gradient boosting outperforms linear models on temporal data
[6]	Bhamare et al. (2024)	Building energy demand prediction	SVM, RF, Gradient Boosting, Ridge	SVM: R^2=0.986 — SVM strongest with GridSearch tuning on Kaggle building data
[7]	Brito &	Commercia	ANN, RF,	ANN and RF

Ref	Authors (Year)	Application Domain	ML Models Used	Key Findings
	Brito (2022)	l energy forecasting	SVM	improve over ARIMA and statistical benchmarks
[8]	Rajalakshmi et al. (2024)	Smart grid energy optimisation	Linear Regression, RF, ANN	RF and ANN: lower RMSE than LR — tree models capture nonlinear demand patterns
[9]	Bhandarkar et al. (2023)	Household electricity prediction	SVM, KNN, ANN, Decision Tree	SVM achieves lowest MAE and RMSE on Low Carbon London (Kaggle) dataset
[10]	Venkatesh et al. (2022)	Domestic energy consumption	RF, SVM, Decision Tree, LR	RF: lowest RMSE — ensemble outperforms all individual models on household data
[11]	Haque et al. (2021)	Residential apartment demand	RF, SVR, KNN, MLR	RF: best overall — but small dataset and low accuracy limit generalisability
[12]	Leo Raju et al. (2025)	Department building IoT monitoring	LR, RF, SVM + IoT (NodeMCU)	LR performs best in real-time IoT context — data granularity overrides complexity
[13]	Vignesh et al. (2025)	Hourly energy forecasting (Finland)	Stacked LSTM (4-layer)	Low MAE; strong temporal fit — but no occupancy or weather features included
[14]	Mathumitha et al. (2023)	Smart meter energy forecasting	SVM Regression	Low MAE and RMSE vs. linear — kernel selection is model's main sensitivity
[15]	Siranjeevi R (2025)	Zone-level power consumption	RF + XGBoost + LSTM hybrid stacking	R^2=0.957, RMSE=1.42 kW — hybrid stacking outperforms all base learners individually

IV. IV.IDENTIFIED RESEARCH GAPS

Table II summarises ten primary research gaps derived from the 15-paper analysis. The most critical is the near-total absence of government buildings as a study domain — only Liu et al. [2] and Leo Raju et al. [12] used datasets from public or institutional buildings. The second critical gap is the widespread exclusion of occupancy data, consistently identified as one of the strongest predictors of office energy demand. The third gap is real-time deployment: all but one study validated models on offline historical datasets. The fourth is interpretability — black-box DL models limit adoption by non-technical government energy managers who must justify procurement decisions.

TABLE II
Research Gaps and Proposed Contributions

Gap	Issue Identified	Proposed Contribution
G1 – Context	Most studies use residential/commercial datasets; govt. buildings absent [2,12]	ML model validated specifically on government building dataset
G2 – Occupancy	Occupancy patterns rarely used as forecasting features [1,8]	Integrate occupancy density and

© MAY 2026 | IRE Journals | Volume 9 Issue 11 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I11-1717926

Gap	Issue Identified	Proposed Contribution
		operational schedule indicators
G3 – Real-Time	Near-real-time pipelines unvalidated in public sector [7,12]	Design near-real-time 24-hour update forecasting framework
G4 – Schedules	Working hours, public holidays, departmental schedules not modelled [3,9]	Encode operational schedule as engineered temporal features
G5 – Scalability	Single-building or small-dataset validation limits generalisability [11,14]	Multi-department, full-formulary forecasting framework
G6 – Peak Demand	Peak vs. off-peak classification underaddressed [5,10]	Build peak demand classification module with threshold alerts
G7 – Interpretability	Black-box DL models limit adoption by non-technical energy managers [4,13]	SHAP-based explainability for energy manager decision support
G8 – Integration	Forecasts not connected to procurement or reorder decisions [6,15]	Connect predictions to automated energy management outputs
G9 – Multi-Dept.	Department-level load diversity not modelled [2,8]	Segment buildings by department for dedicated sub-models
G10 – Evaluation	Inconsistent metrics across studies; no standard benchmark [1,7]	Standardised evaluation: RMSE, MAE, MAPE, R^2 on common dataset

V. PROPOSED METHODOLOGY

A. System Architecture

The proposed framework is a five-module end-to-end pipeline: (1) Data Ingestion — collects historical energy consumption records from smart meters and Building Management Systems (BMS); (2) Preprocessing and Feature Engineering — handles missing values, outliers, and constructs occupancy-aware and temporal features; (3) Model Training and Comparison — trains five model categories on an 80:10:10 time-series walk-forward split; (4) Forecasting Engine — generates rolling 7-day and 30-day consumption forecasts updated every 24 hours; (5) Analytics and Decision Output — provides forecast reports, peak demand alerts, and energy efficiency recommendations for building managers.

B. Feature Engineering

Features are drawn from three categories: (1) Historical consumption — lagged values at t-1, t-24, and t-168 hours; rolling mean and standard deviation over 24-hour and 7-day windows; (2) Temporal features — hour-of-day, day-of-week, week-of-year, month, season, working day flag, public holiday indicator (sine/cosine encoded for cyclical representation); (3) Occupancy features — personnel count ratios, badge-access log proxies, departmental schedule flags, and BMS-derived activity indicators. Feature importance is assessed using SHAP values and Mutual Information ranking, following Liu et al. [2].

C. Dataset

The dataset comprises historical energy consumption records from a government office building, including hourly smart meter readings, BMS logs, and occupancy schedule data. The dataset spans 24 months and contains approximately 17,520 hourly records. Each record includes the timestamp, total energy consumption (kWh), floor-level sub-meter readings where available, occupancy flag (working/non-working day), and weather variables (temperature, humidity). Data cleaning addresses missing values through linear interpolation for short gaps and forward-fill for longer absences.

D. Model Comparison Plan

Five model categories are benchmarked: (a) Baseline — Linear Regression, ARIMA; (b) Classical ML — Random Forest, XGBoost, LightGBM, SVM; (c) Deep Learning — Stacked LSTM, Bi-LSTM; (d)

Hybrid Ensemble — RF + XGBoost + LSTM stacking [15]; (e) Interpretable — Neuro-Fuzzy (HyFIS). An 80:10:10 time-series walk-forward cross-validation prevents data leakage. Evaluation metrics: RMSE, MAE, MAPE, R^2, SMAPE.

E. Workflow

The implemented workflow consists of the following interconnected stages: (1) Historical smart meter and BMS records are collected from the target government building; (2) Data preprocessing handles missing values, normalisation, and outlier removal; (3) Feature engineering extracts temporal, occupancy, and lagged consumption features; (4) The dataset is divided chronologically into training (80%), validation (10%), and testing (10%) subsets; (5) All five model categories are trained and benchmarked; (6) The best-performing hybrid model is deployed in the forecasting engine; (7) Forecast outputs are connected to energy management decision recommendations.

G. Expected Outcomes

Target performance: MAPE ≤5%, $R^2 \geq 0.95$, real-time forecast update cycle of 24 hours, and SHAP-based feature importance reports for every forecast window. Benchmarked against best reported results: Liu et al. [2] ($R^2 \approx 0.98$ on UK government data), Siranjeevi [15] ($R^2 = 0.957$ with hybrid stacking), and Chauhan & Bansal [3] ($R^2 = 0.941$ with Bi-LSTM). Occupancy-unaware baselines are expected to underperform occupancy-aware models by 10–20% on MAPE for volatile consumption periods.

VI. IMPLEMENTATION

The planned implementation will benchmark five machine learning models for hourly energy consumption forecasting in a government office building. The implementation pipeline will consist of five sequential phases: data preprocessing, feature engineering, model training, performance evaluation, and forecast generation. This section outlines the detailed plan for each phase.

A. Data Preprocessing Plan

The raw dataset will comprise 17,520 hourly smart meter records spanning 24 months, each containing a timestamp, total energy consumption (kWh), occupancy flag, and weather variables. Timestamps will be parsed into datetime format and validated for completeness. Missing values will be imputed using linear interpolation for gaps under 2 hours and forward-fill for longer absences. Outliers exceeding three standard deviations from the rolling mean will be replaced with rolling median values to preserve the temporal structure of the time series.

B. Feature Engineering Plan

The preprocessed dataset will be enriched with five categories of engineered features: (i) temporal features — hour of the day (0–23), day of the week (0–6), month (1–12), and season (1–4); (ii) cyclical encodings — sine and cosine transformations applied to hour and month to preserve circular periodicity; (iii) occupancy indicators — working day binary flag, public holiday indicator, and departmental schedule flag; (iv) lag features — consumption values at t-1, t-24, and t-168 hours to capture short-term and weekly dependencies; (v) rolling statistics — 24-hour and 7-day rolling mean and standard deviation. All features will be normalised to [0,1] using Min-Max scaling fitted exclusively on the training partition to prevent data leakage.

C. Training and Testing Strategy

To preserve the temporal sequence of energy consumption data, the dataset will be divided chronologically: 80% (approximately 14,016 records) for training and 20% (approximately 3,504 records) for testing. Walk-forward cross-validation with a 30-day sliding window will be applied to prevent data leakage. This approach ensures each model is evaluated on unseen future data while maintaining the chronological structure of the forecasting problem. All five models will be trained on an identical feature set and evaluated on the same held-out test partition to enable a fair and consistent comparison.

D. Models to be Implemented

Five machine learning models will be trained and compared on the same feature set and test partition: (1) Linear Regression — a baseline model using all engineered features without regularisation, establishing a lower-bound performance reference; (2) Random Forest — an ensemble of 200 decision trees with max depth 15 and minimum samples split of 5, leveraging bagging to reduce variance; (3) XGBoost — gradient boosted trees with a learning rate of 0.05, 500 estimators, max depth 8, and subsample 0.8, selected for its strong performance on temporal tabular data as reported by Thakur et al. [5]; (4) Support Vector Regression (SVR) — configured with RBF kernel, C=100, epsilon=0.1, and gamma='scale', consistent with the tuning approach of Bhamare et al. [6]; (5) Stacked LSTM — three sequential LSTM layers (128, 64, 32 units) with Dropout 0.2, trained using the Adam optimiser for up to 50 epochs with early stopping on validation MAE, following the architecture validated by Siranjeevi [15].

E. Evaluation Metrics

The forecasting performance of all five models will be evaluated using four standard metrics: Mean Absolute Error (MAE) to measure average absolute prediction error in kWh; Root Mean Square Error (RMSE) to penalise large deviations and quantify worst-case error magnitude; Mean Absolute Percentage Error (MAPE) to enable scale-independent comparison across models and against literature benchmarks; and R-Squared (R^2) to quantify the proportion of variance in energy consumption explained by each model. These four metrics collectively cover accuracy, error distribution, and explanatory power, providing a comprehensive and standardised basis for model comparison consistent with the evaluation protocols used in Liu et al. [2] and Siranjeevi [15].

VII. VII.EXPECTED RESULTS AND DISCUSSION

The implementation has not yet been executed; the results and discussion presented in this section are therefore anticipated outcomes grounded in the literature review and the theoretical properties of the five selected models. Empirical results will be reported in the final version of this paper upon completion of the implementation.

A. Anticipated Model Performance

Based on evidence from the reviewed literature, the Stacked LSTM model is expected to achieve the best overall performance across all four-evaluation metrics, owing to its capacity to learn long-range sequential dependencies in hourly consumption data — a capability confirmed by Vignesh et al. [13] and Siranjeevi [15]. XGBoost is anticipated to rank second, consistent with its demonstrated superiority over classical ML on temporal tabular datasets reported by Thakur et al. [5] (R^2=0.946). Random Forest is expected to deliver solid third-place performance as the most reliable classical ensemble, as established across multiple reviewed studies [1][10][11]. SVR is expected to perform moderately, with sensitivity to kernel selection as a limiting factor [6][14]. Linear Regression, as the simplest baseline, is expected to exhibit the highest error rates, reflecting the well-documented inadequacy of linear models for capturing nonlinear occupancy-driven demand transitions in government buildings.

B. Anticipated Energy Demand Patterns

Government building energy consumption is expected to exhibit clearly structured temporal patterns consistent with fixed occupancy schedules. Peak demand periods are anticipated during standard working hours on weekdays, with substantial consumption reductions on weekends and public holidays — patterns documented in Leo Raju et al. [12] for an institutional building and consistent with the operational characteristics described in Section II. Seasonal variation driven by heating and cooling loads is also expected, with elevated consumption during summer (air conditioning) and winter (heating) months. These occupancy-driven and seasonal patterns collectively motivate the inclusion of occupancy flags, schedule indicators, and seasonal features in the feature engineering pipeline, and are expected to yield measurable accuracy improvements over purely temporal baselines — consistent with the 10–20% MAPE reduction estimated in Section V-G.

C. Anticipated Feature Importance

Based on the literature, lag features — particularly the t-1 (previous hour) and t-24 (same hour, previous day) consumption values — are expected to emerge as the strongest individual predictors across tree-based models, consistent with findings by Siranjeevi [15]. Among contextual features, occupancy indicators (working day flag, departmental schedule) are expected to rank among the top predictors in Random Forest feature importance analysis, validating the occupancy-aware feature engineering strategy described in Section V-B. Hour-of-day and day-of-week are expected to follow as strong temporal predictors, while weather variables (temperature, humidity) are anticipated to contribute a meaningful but smaller share of total importance. SHAP-based explainability analysis will be applied to quantify and visualise these contributions, providing interpretable outputs for government energy managers as described in Gap G7.

D. Anticipated Comparative Insights

The Stacked LSTM is anticipated to outperform all classical ML models, particularly for longer forecasting horizons where sequential temporal dependencies are most pronounced. However, the performance gap between LSTM and XGBoost is expected to be modest relative to the substantially higher computational cost of the deep learning model — an important practical trade-off for government energy managers operating with limited infrastructure. XGBoost, combined with SHAP explanations, is therefore anticipated to represent the strongest interpretable alternative for operational deployment. Linear Regression's anticipated poor performance will serve to confirm the nonlinear nature of government building energy consumption and validate the model selection rationale for the proposed framework.

E. Discussion

The five-model comparative framework is designed to generate evidence-based answers to the research questions identified in Section IV. By benchmarking a sequential deep learning model (Stacked LSTM), two gradient boosting methods (XGBoost, Random Forest), a kernel method (SVR), and a linear baseline (Linear Regression) on a consistent government building dataset, the implementation will provide the first direct model comparison on this domain with full occupancy-aware features. The anticipated advantage of occupancy-integrated models over purely temporal baselines will directly address Gap G2; the standardised evaluation protocol will address Gap G10; and the SHAP explainability layer will address Gap G7. Prediction errors are anticipated to be highest during atypical building operations — unscheduled events, emergency sessions, or maintenance periods — where historical patterns provide limited predictive signal. Identifying and quantifying these high-uncertainty periods will inform the design of an anomaly-flagging module to be developed in the next phase of this research.

VIII. CONCLUSION

This paper presents a structured review of 15 IEEE publications (2021–2026) on ML-based building energy consumption forecasting and a planned comparative framework for five ML models targeting government buildings — a domain critically underrepresented in the existing literature. Ten primary research gaps were identified, with the most significant being the absence of government building datasets, exclusion of occupancy-driven features, lack of real-time deployment validation, and insufficient interpretability for non-technical energy managers. The proposed intelligent occupancy-aware framework, comprising Linear Regression, Random Forest, XGBoost, SVR, and Stacked LSTM, directly addresses these gaps through systematic model comparison, occupancy-integrated feature engineering, SHAP-based explainability, and a near-real-time 24-hour forecasting pipeline.

The occupancy-aware feature engineering strategy — incorporating working day flags, departmental schedule indicators, and lag consumption features — is hypothesised to yield measurable accuracy improvements over purely temporal baselines, with a 10–20% MAPE reduction anticipated for volatile consumption periods based on comparable studies [2][15]. The Stacked LSTM is anticipated to achieve the highest forecast accuracy, while XGBoost with SHAP analysis is expected to provide the strongest interpretable alternative for operational government energy management. Empirical validation of these hypotheses will be conducted in the next phase of this research, with results to be reported in the final version of this paper. Future work will additionally focus on multi-building dataset collection, integration with live BMS infrastructure, and exploration of federated learning for privacy-preserving multi-department deployment.

IX. FUTURE SCOPE

Future implementation will focus on: (1) empirical benchmarking and comparative evaluation of all five model categories on a real government building energy dataset with occupancy metadata; (2) live BMS system validation using real-time smart meter data streams with 15-minute update cycles; (3) multi-building dataset collection across different government facility types — offices, hospitals, educational institutions — to improve model generalisability; (4) integration of privacy-preserving federated learning for multi-department or multi-building deployment without centralising sensitive occupancy data; (5) development of an end-to-end real-time automated alert system for peak demand events integrated with energy procurement platforms; and (6) deeper exploration of Explainable AI (XAI) techniques to generate audit-ready forecast justifications for government energy reporting standards.

REFERENCES.

[1] K. B. A. Kumar et al., "Energy Consumption in Smart Buildings using Machine Learning," ICICCS, IEEE, 2023.

[2] L. Liu et al., "Building Energy Consumption Prediction: A Machine Learning Approach with Feature Selection," ICSCC, IEEE, 2024.

[3] N. Chauhan and S. Bansal, "Performance Evaluation of ML Techniques in Residential Energy Consumption Forecasting," MRIE, IEEE, 2025.

[4] R. Upadhyay et al., "Comparative Analysis for Short and Long-Term Time Series for Energy Consumption Predictions," IEEE, 2026.

[5] A. Thakur et al., "Predictive Analysis of Energy Consumption Using ML Techniques," ICSSES, IEEE, 2023.

[6] A. Bhamare et al., "Machine Learning-based Energy Demand Forecasting," ICCUBEA, IEEE, 2024.

[7] T. C. Brito and M. A. Brito, "Forecasting of Energy Consumption Using AI Methods," IEEE, 2022.

[8] S. Rajalakshmi et al., "Energy Demand Forecasting and Optimization," IEEE ICSCSS, 2024.

[9] M. Bhandarkar et al., "Electricity Billing and Consumption Prediction using ML," ICCUBEA, IEEE, 2023.

[10] M. Venkatesh et al., "Domestic Energy Consumption Forecasting using ML," ICESC, IEEE, 2022.

[11] H. Haque et al., "Demand Analysis of Energy Consumption in a Residential Apartment using ML," IEMTRONICS, IEEE, 2021.

[12] L. Raju et al., "Cost Effective Building Energy Management using IoT and ML," ICEES, IEEE, 2025.

[13] K. Vignesh et al., "Energy Consumption Forecasting using LSTM-based Deep Learning Models," ICECMSN, IEEE, 2025.

[14] R. Mathumitha et al., "SVM-based Regression for Forecasting Building Power Energy Consumption," ICCCNT, IEEE, 2023.

[15] Siranjeevi R, "Hybrid Machine Learning Models for Power Consumption Prediction," ICECST, IEEE, 2025.

[16] N. Chauhan et al., "A Comparative Analysis for Short and Long-Term Time Series for Energy Consumption Predictions," ICCICA, IEEE, 2024.

[17] J. Gaboitaolelwe et al., "Electricity Load Prediction Using Machine Learning," SmartNets, IEEE, 2022.

[18] M. A. M. Hasan et al., "Enhancing Energy Market Price Forecasting through Advanced ML Techniques," OEEC, IEEE, 2025.

[19] R. Halder and K. F. Ahmed, "Multi-Architecture Deep Learning Approach for Enhanced Energy Consumption Forecasting," SEPOC, IEEE, 2025.

[20] D. Syed et al., "Household-Level Energy Forecasting in Smart Buildings Using a Novel Hybrid Deep Learning Model," IEEE Access, 2021.

[21] C. Zongo et al., "Forecasting Models for Monthly Energy Consumption Using ML in Burkina Faso," IEEE AFRICON, 2023.

[22] A. Jozi et al., "Energy Consumption Forecasting using Neuro-Fuzzy Inference Systems," IEEE EEEIC, 2020.

[23] V. Srivastava et al., "Review on Electrical Load Forecasting Approaches," ISML, IEEE, 2024.

[24] A. Kidd et al., "Industrial Energy Forecasting using ML and Plant Activity Metrics," CSECS, IEEE, 2025.

[25] R. Yadav et al., "Prediction of Consumption of Energy in an Area," ICTMIM, IEEE, 2025.