Regional Electricity Load Prediction Using Hybrid Machine Learning Models Under Limited Data Conditions: A Case Study of Uganda

Kwesiga Kato Steven
(2025/HD05/26350U, 2500726350)

Makerere University, College of Computing and Informatics Science
Master's in Computer Science (MCS)
MCS 7103: MACHINE LEARNING :- Project Report
Academic Year: 2025/2026

Abstract

This report presents a comprehensive analysis of regional electricity load forecasting in Uganda, a context characterized by limited historical data. Utilizing daily power-sales data (Jan 2021–Nov 2025) from the UEDCL Kabalagala district, this study systematically develops and evaluates a spectrum of forecasting models, including classical SARIMAX, machine learning (Linear Regression, Random Forest, XGBoost), deep learning (LSTM), and a hybrid weighted ensemble.

The investigation reveals that a Linear Regression model, when empowered by a methodically engineered feature set, achieves the highest predictive accuracy (RMSE: 1609.02), decisively outperforming more architecturally complex models like LSTM and XGBoost. This outcome underscores a critical principle for data-scarce environments: disciplined, domain-aware feature engineering is a more potent strategy for achieving high-accuracy forecasts than reliance on complex model architectures alone. The findings provide a robust, interpretable, and immediately practical forecasting solution for utility providers like UEDCL.

Contents

1 Introduction

Accurate electricity load forecasting is fundamental to the operational efficiency, grid stability, and strategic planning of power utilities. In developing regions such as Uganda, this task is challenging due to limited long-term, granular historical consumption data. These constraints hinder the straightforward deployment of sophisticated forecasting models typically effective in data-rich contexts.

A potentially valuable source of information – weather data(for Kabalagala) – could not be obtained for this study. Variables such as temperature, humidity, and solar irradiance are known to influence electricity demand, particularly for cooling, heating, and lighting loads. Their absence limits the models' ability to account for external environmental factors that typically improve forecast accuracy.

This project investigates how statistical, machine learning, and deep learning approaches can be adapted to deliver accurate short-term load forecasts under real-world data scarcity.

Research Question: How can hybrid machine learning models and strategic feature engineering be leveraged to accurately forecast short-term electricity demand in Uganda given limited data conditions?

Objectives:

1. Develop and evaluate diverse short-term forecasting models based on available customer purchase data.

2. Compare the performance of SARIMAX, feature-driven ML models (Linear Regression, Random Forest, XGBoost), LSTM, and a hybrid ensemble.

3. Identify the most performant and practical approach suitable for Ugandan power sector deployment.

4. Analyze the trade-offs between model complexity and feature engineering under constrained data conditions.

2 Data and Methodology

2.1 Dataset Description

Daily electricity consumption data for the UEDCL Kabalagala district was used, covering January 1, 2021 to November 13, 2025. The dataset includes:

- `datetime`
- `consumption_kWh`

2.2 Data Preprocessing and EDA

Preprocessing steps:

- Removed formatting artifacts and converted consumption values to floats.
- Parsed timestamps and established a daily time-series index.
- Enforced date continuity using:
  - 3-day forward fill for short gaps.
  - 7-day rolling median for remaining missing values.

EDA observations:

- Long-term upward consumption trend.
- Strong weekly and annual seasonality.
- Consumption peaks on weekends.

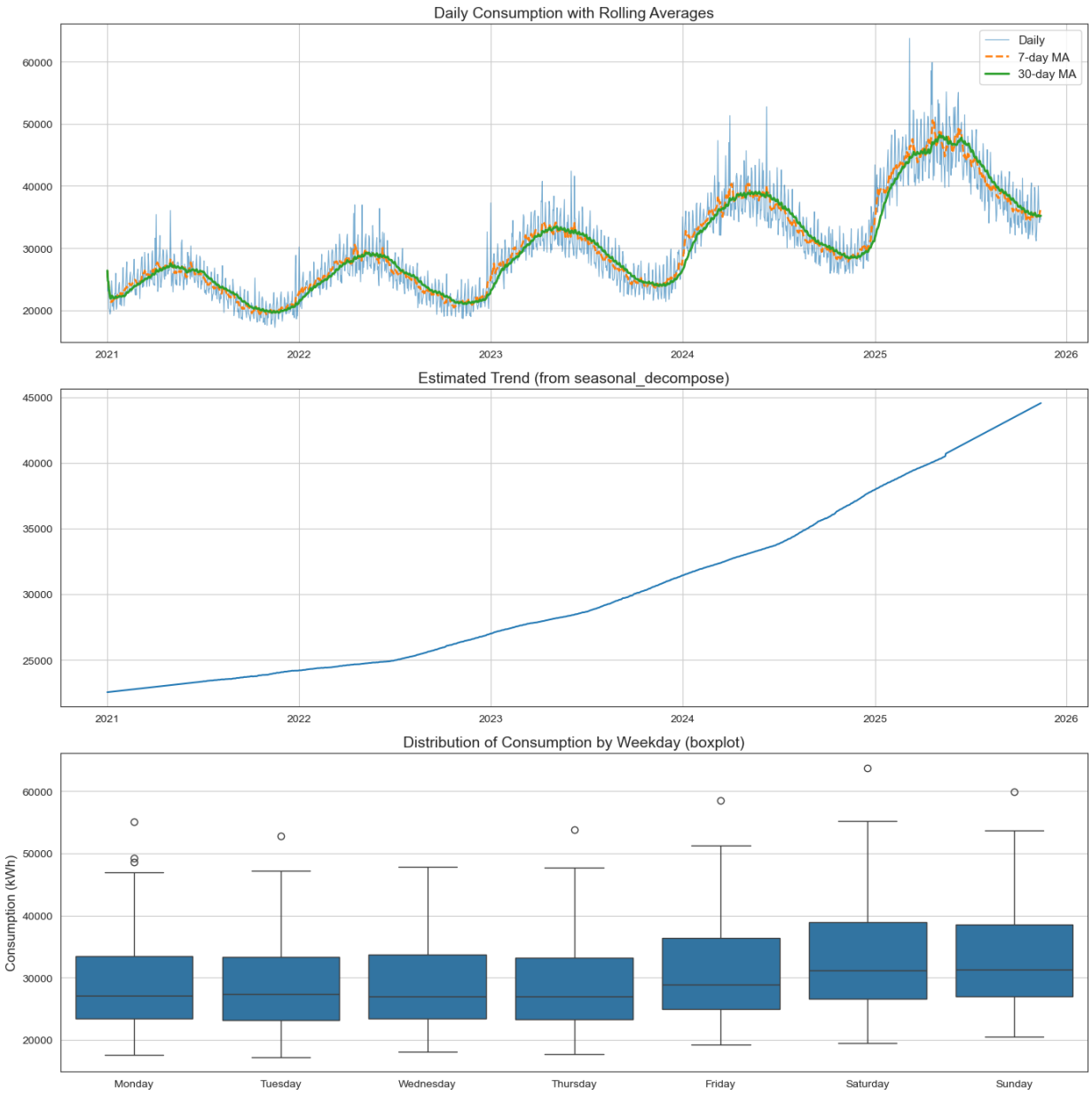

Figure 1: Energy consumption trend overview

2.3 Feature Engineering

To prepare the raw time-series data for supervised learning, a focused set of features was created. These features help the model capture key time-based patterns and external factors that influence electricity demand. They include:

Lag Features: 1, 2, 3, 7, 14, and 30 days. **Rolling Statistics:** 7, 14, 30-day rolling mean and standard deviation. **Calendar Features:** day of week, month, weekend flag, and Ugandan public holidays.

2.4 Modeling Approach

Temporal Split: All data except the last 90 days used for training; final 90 days used for testing.

Scaling: `StandardScaler` applied to Linear Regression and LSTM.

Models Implemented:

- Baselines: Naive, Linear Regression, Random Forest (`n_estimators=200`)
- SARIMAX $(2, 1, 1)$ with seasonal $(1, 0, 1, 7)$
- XGBoost tuned via Optuna with:
  - `learning_rate = 0.206`
  - `max_depth = 9`
  - `subsample = 0.766`
- LSTM: 14-step sequence; LSTM(64), Dropout(0.19), Dense(128), Adam($lr = 0.005$)
- Hybrid Ensemble: Weighted by inverse RMSE

3 Performance Metrics

Models were evaluated using standard regression metrics that capture different aspects of predictive accuracy:

3.1 Root Mean Squared Error (RMSE)

Measures the square root of the average squared differences between predicted and actual values. It penalizes large errors more, making it sensitive to outliers, and is in the same units as consumption (kWh).

3.2 Mean Absolute Error (MAE)

Average absolute difference between predictions and actual values. Treats all errors equally and gives a clear measure of typical prediction error. Comparing MAE and RMSE indicates the impact of outliers.

3.3 Accuracy (%)

Defined as $100 - \text{MAPE}$. Provides a relative, intuitive measure of model performance, useful for non-technical stakeholders.

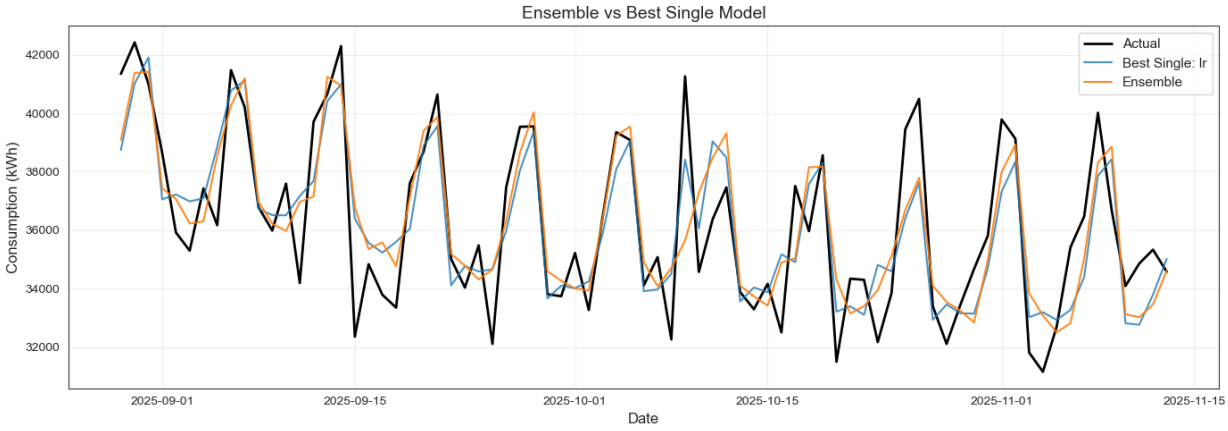

Figure 2: Weighted ensemble predictions compared to the best single model. The ensemble combined all available models inversely weighted by RMSE, producing a slightly more robust forecast. Linear Regression dominated due to careful feature engineering, which effectively captured temporal patterns despite its simpler architecture. More complex models like LSTM underperformed due to limited training data.

Note: The ensemble leveraged complementary strengths of multiple models, but in this dataset, the Linear Regression model remained the strongest contributor because it generalized well using engineered lag, rolling statistics, and calendar features. Complex models failed to outperform due to insufficient data for training effectively.

4 Results and Discussion

4.1 Model Performance (90-Day Test Set)

Model	RMSE	MAE	Accuracy (%)
Linear Regression	**1609.02**	**1348.53**	**96.27**
Hybrid Ensemble	1689.45	1368.54	96.20
SARIMA	1879.07	1517.17	95.78
XGBoost	1954.85	1518.77	95.80
Random Forest	1944.94	1574.79	95.60
LSTM	2457.21	1962.29	94.78
Naive	3254.62	2490.75	93.10

4.2 Key Insights

Linear Regression Dominance: Feature engineering enabled the simplest model to outperform all advanced algorithms.

XGBoost: Important features included lag_7, roll_std_7, lag_1, and roll_mean_14.

LSTM: Underperformed due to insufficient training data for its parameter scale.

Hybrid Ensemble: Strong second-best performer.

Residual Diagnostics: Linear Regression residuals were approximately normal with no significant autocorrelation.

4.3 Justification for Model Selection

The choice of models spans from simple statistical approaches to advanced deep learning to assess the trade-off between feature engineering and model complexity in a data-constrained context.

Baselines: Linear Regression and Random Forest were included to benchmark interpretability and non-linear capabilities, respectively. Linear Regression tests whether engineered temporal features suffice for accurate linear prediction, while Random Forest captures complex feature interactions with minimal tuning.

Classical Time-Series: SARIMAX models trend, seasonality, and autocorrelation explicitly, providing probabilistic forecasts for risk assessment and serving as a statistical benchmark.

Gradient Boosting: XGBoost evaluates whether non-linear interactions in engineered features improve predictive accuracy and provides feature importance for validation.

Deep Learning: LSTM assesses the potential to autonomously learn long-range temporal dependencies beyond manual feature engineering.

Meta-Model: A Weighted Ensemble aggregates predictions inversely by RMSE, leveraging complementary model strengths for a more robust forecast.

5 Conclusion

This study demonstrates that feature engineering outweighs model complexity in data-limited forecasting contexts. A feature-rich Linear Regression model delivers superior accuracy, interpretability, and operational practicality for UEDCL. It stands as the recommended approach for short-term load forecasting and resource planning.