Author

document_03339ba9-0277-4cd0-b6c6-0eb3128c7f1b

Turnitin

Document Details

Submission ID

trn:oid:::1:3417291964

Submission Date

Nov 19, 2025, 6:56 AM GMT-5

Download Date

Nov 19, 2025, 6:57 AM GMT-5

File Name

tmp5tn8kszj.docx

File Size

76.5 KB

7 Pages

2,998 Words

18,276 Characters

40% detected as AI

The percentage indicates the combined amount of likely AI-generated text as well as likely AI-generated text that was also likely AI-paraphrased.

Caution: Review required.

It is essential to understand the limitations of AI detection before making decisions about a student's work. We encourage you to learn more about Turnitin's AI detection capabilities before using the tool.

Detection Groups

27 AI-generated only 40%
Likely AI-generated text from a large-language model.

0 AI-generated text that was AI-paraphrased 0%
Likely AI-generated text that was likely revised using an AI-paraphrase tool or word spinner.

Disclaimer
Our AI writing assessment is designed to help educators identify text that might be prepared by a generative AI tool. Our AI writing assessment may not always be accurate (it may misidentify writing that is likely AI generated as AI generated and AI paraphrased or likely AI generated and AI paraphrased writing as only AI generated) so it should not be used as the sole basis for adverse actions against a student. It takes further scrutiny and human judgment in conjunction with an organization's application of its specific academic policies to determine whether any academic misconduct has occurred.

Frequently Asked Questions

How should I interpret Turnitin's AI writing percentage and false positives?
The percentage shown in the AI writing report is the amount of qualifying text within the submission that Turnitin's AI writing detection model determines was either likely AI-generated text from a large-language model or likely AI-generated text that was likely revised using an AI paraphrase tool or word spinner.

False positives (incorrectly flagging human-written text as AI-generated) are a possibility in AI models.

AI detection scores under 20%, which we do not surface in new reports, have a higher likelihood of false positives. To reduce the likelihood of misinterpretation, no score or highlights are attributed and are indicated with an asterisk in the report (*%).

The AI writing percentage should not be the sole basis to determine whether misconduct has occurred. The reviewer/instructor should use the percentage as a means to start a formative conversation with their student and/or use it to examine the submitted assignment in accordance with their school's policies.

What does 'qualifying text' mean?
Our model only processes qualifying text in the form of long-form writing. Long-form writing means individual sentences contained in paragraphs that make up a longer piece of written work, such as an essay, a dissertation, or an article, etc. Qualifying text that has been determined to be likely AI-generated will be highlighted in cyan in the submission, and likely AI-generated and then likely AI-paraphrased will be highlighted purple.

Non-qualifying text, such as bullet points, annotated bibliographies, etc., will not be processed and can create disparity between the submission highlights and the percentage shown.

A review on machine learning approaches to electricity demand prediction

1st Pranav Prakash
Lovely Professional University,
Punjab, India
12219903
pranav.prakash@lpu.in

2nd Sai Pavan Mudrakola
Lovely Professional University,
Punjab, India
12210714
saipavan.12210714@lpu.in

3rd Jay Raj
Lovely Professional University,
Punjab, India
12221138
jayraj.12221138@lpu.in

Abstract

The accurate prediction of electricity demand is essential to maintain a stable electrical grid, but it has become complicated with rapid urbanization, weather, increasing solar penetration, and changing socioeconomic factors. Climate changes and industrial trends are also creating an increased need for effective utilization of electricity. Recent work has shown that classical statistical tools are not enough to deal with those non-linear context sensitive patterns. The application of several machine learning techniques such as XGBoost, LSTM, CNN, RNN and others has allowed us to evaluate how the electricity load forecasting model has evolved with respect to identifying the root causes and effects of employing STLF and LTLF models to obtain greater accuracy and predictability .This review aggregates predominant worldwide research projects using hybrid Artificial Intelligence (AI) methods, including a combination of machine learning, deep learning and statistical approaches in conjunction with contextual attributes from weather, calendar effects, policy drivers and socio-economic measures.

Keywords: Electricity load forecasting, XGBoost, Machine Learning, Time series, LSTM, CNN, RNN, Hybrid AI Models

1. Introduction

The efficient use of electricity is important for modern society, the environment, and sustainability, and requires the correct use of load forecasting. Globally the demand for electricity is always high, and predicting the need for electricity will be required for a given time and how that demand will affect the utility grid. The overall growing population, hazardous weather conditions, pollution, and electricity used by industries and housing make the demand for electricity difficult to predict. Using traditional approaches, we often fail to capture the actual data required for forecasting. With the help of artificial intelligence and machine learning-driven approaches, we can solve demand load forecasting by implementing models such as long-term short-term memory (LSTM), XG Boost, and others to obtain sharper outcomes. Hybrid AI systems stitch multiple engines together - LSTM memory blocks, XGBoost trees, convolutional layers, support vector boundaries, ARIMA/SARIMAX time step modules and attention heads - so the model captures curved, non-straight links yet still lets people see why it reached a verdict. In addition, context aware AI folds in outside facts - air temperature, moisture in the air, public holidays, sun strength, how much demand relaxes when price rises - to lift precision., occupancy, and policy effects.

While many studies and research papers have been published to solve the load forecasting issue, there are not many clear datasets and models available to predict outbreaks and situations under different conditions. The objective is to synthesize these contributions, understand strengths and weaknesses, and develop a unified methodological understanding for future research.

2. Problem Statement

Even with the advancement of machine learning and deep learning, electricity demand forecasting remains difficult due to the inaccuracy and unreliability of demand forecasting models when applied to different demographic profile. The data are mostly inadequate and provide a small number of accurate results. This is insufficient for solving the issue and can lead to several problems. Over forecasting electricity is when we predict higher demand for electricity than it usually occurs on most of the days, which often results in wasted resources and expenses. Under forecasting, we predict a lower demand for electricity than usual, which can lead to power outages and economic issues because there are not many resources available to meet the required demand. There is a significant gap between the current literature of research papers, where the AI models that have been proposed are tested on different datasets and metrics. There is a lack of a comprehensive model to predict the demand forecasting under different challenges, such as weather, pollution, holidays and economic situation. The government and private sectors do not have an accurate and reliable model or dataset to predict the current situation. Hence, there is a need for hybrid AI models capable of combining XGBoost's feature learning strength with LSTM's ability to generate accurate short- and long-term demand forecasting. In this study, we attempt to address the ML models are suitable for predicting the correct data and providing a clear framework for future research areas and fields.

3. Scope of the Review

This review focuses on combining hybrid LSTM and XG boost for predicting global electricity load forecasting. The primary challenges were dealing with large datasets and the uncertainty of the traditional framework used to obtain the correct demand for load forecasting in metropolitan cities. The most promising findings for future research would be the implementation of applying hybrid XGBoost , LSTM and similar multi model architectures for load forecasting in different regions of the country. This will eventually lead to better and more accurate results and findings under high uncertainty such as heatwaves, air pollution and peak events. This review focuses on short-term forecasts (STLF) and the models used to predict the results from papers published in between 2013-2025. This review primarily emphasizes methodological trends, feature selection, model architecture and regional applicability.
This review will not include any unpublished theses, articles based on opinions, or studies that primarily focus on related issues but not on the actual topic. By adhering to the scope of a review paper, this study aims to provide a clear and comprehensive overview of the current research, which offers a valuable and meaningful resource for both academics and practitioners.

5.Background And Literature Review

5.1 Load Forecasting: Definition and Temporal Classification

Load forecasting is a process used to determine the amount of electricity required at a given time while maintaining the balance and stability of the power grid. Load forecasting can be divided into three parts: short-term forecasts (STLF), which range from one hour to one week. Medium-term forecasting (MTLF) ranges from one week to one year, and long-term forecasting (LTLF) usually ranges from one year to many years. This review primarily focuses on short-term forecasts (STLF), which are more crucial for daily grid operations and preventing blackouts during peak hours at night.

5.2 Factors Influencing Electricity Demand Globally

Global electricity demand is characterized by several factors:

1.Weather: During summer, the demand for air conditioning and heating during winter describes a significant relationship between weather and temperature.

2.Pollution: Elevated temperatures during summer often lead to a higher air pollution index, which in turn increases the demand for air purifiers and air conditioning, contributing to the rise in the overall electricity demand.

3.Heat Index: The heat index is correlated with power supply demand instead of ambient temperature. This increases the demand for cooling appliances, which results in a high and prolonged increase in power consumption. Several reports have noted that more than 60% of the variation in countries peak electricity demand can be described using the heat index.

4. EV vehicles: The growing demand and adaption of electric vehicles such as bus, car and scooter are

increasing the overall demand in electric demand forecasting.

5. Economic growth and Industrialization: With the development of economic percentage and expanding of industrial sector, it requires more electricity to manage the power production and day to day consumption.

5.3 Traditional and ML forecasting approaches

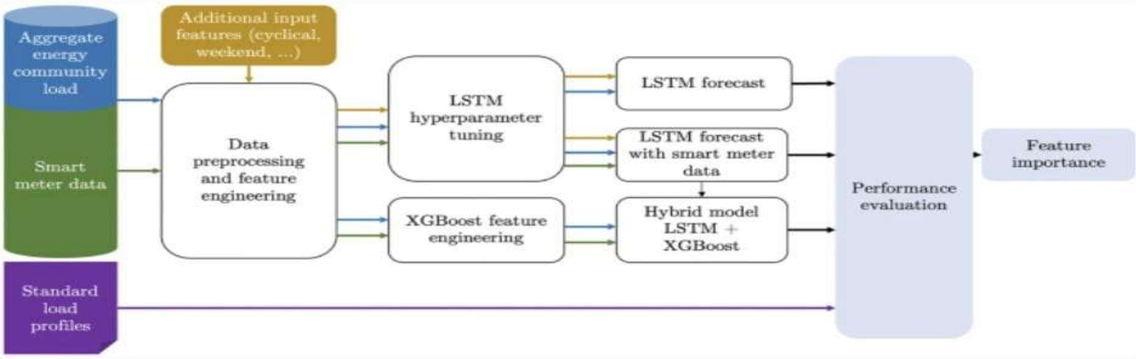
Statistical models, such as ARIMA, linear regression, and gradient boosting, alongside machine learning models like Random Forest and Support Vector Regression (SVR), did not establish a strong connection between weather, demographic factors, pollution, and overall production when applied to demand forecasting, especially in the scope of short-term load forecasting (STLF). In contrast, XGBoost is better at working with complex data while providing hybrid models that produce metrics that are useful for understanding forecasting issues. It was shown that XGBoost can capture high nonlinear dependencies, as well demonstrate the overall prediction accuracy of the model. Long Short-Term Memory (LSTM) has demonstrated the ability to capture and resolve high levels of complexity and long-term serial dependencies common in time-series data, resulting in accuracy for daily forecasting and sequential patterns. Hybrid models using AI methods, such as XGBoost and Random Forest methods, combined with deep learning formalisms, such as LSTM, GRU, and CNN have resulted in better forecasting results by capturing temporal patterns, a mixture of meteorological variables (e.g., solar irradiance), pollution index, and policy signals. Recent studies have emphasized the use of hybrid models to model demand, with a mixture of different domain contextualization (e.g., weekend, holiday/festival dynamics models, or peak events), to improve forecasting accuracy. This review demonstrates that hybrid frameworks can produce improved demand forecasting methodologies for Complex energy ecosystem such as Delhi.

5.4. Comparative Analysis Table

Year	**Authors**	**Model / Approach**	**Dataset / Region**	**Key Contribution**	**Accuracy (%)**
2015	X. Wang, Y. Wu, W. Zou, X. Zhao	ARIMA–LSTM Hybrid + Cloud-Native Deployment	Real-time electricity market (unspecified region)	Real-time market forecasting using hybrid classical + DL model	**~89.1%**
2021	J. Lee, Y. Cho	National-scale ML vs Hybrid Forecasting	South Korea (National load data)	Compared ML, DL, and Hybrid models for peak load forecasting	**92–95%**
2022	L. Semmelmann, S. Henni, C. Weinhardt	LSTM–XGBoost Hybrid	Energy communities, Smart meter data (Europe)	Hybrid model for energy communities with improved temporal-spatial learning	**96.2%**
2022	Tan H. et al.	XGBoost–LSTM for PV	Photovoltaic datasets (China)	Improved PV power prediction through hybrid meteorological–temporal fusion	**95.4%**
2022	Y. Ding et al.	Knowledge Aggregation Model	Grid load datasets (China)	Novel knowledge aggregation mechanism for load prediction	**~97%**

Year	Authors	Model / Approach	Dataset / Region	Key Contribution	Accuracy (%)
2023	T.G. Grandón et al.	Hybrid Statistical + ML	Ukraine electricity demand	Crisis-adaptive hybrid forecasting for unstable energy situations	~91%
2025	F. Karimidehkordi et al.	Hybrid LSTM–XGBoost + Climate Scenarios	Iran (Sector-wise energy data)	Climate-aware hybrid model, sector-specific load forecasts	94.8 %
2025	Altınbaş University	LSTM–XGBoost Smart Grid Optimization	Turkey smart grid datasets	Smart grid forecasting with improved robustness	97.1 %
2025	A. Ajder et al.	Wavelet-LSTM–XGBoost	Time series with anomalies	Wavelet decomposition to handle unpredictable events	98.2 %
2025	A. Hussain et al.	CAT-Former (Context-aware Transformer)	EV Charging Stations (China)	Multi-context EV-load forecasting transformer	97.5 %
2025	Y. Zhang	LSTM + Transformer	Energy consumption (China)	Hybrid temporal + attention fusion architecture	96.8 %
2025	Abdulameer & Ibrahim	GRU + CNN + LSTM + ML Regressors	Iraq electricity consumption	Multi-hybrid fusion model with stacked DL architecture	95.8 %
2025	A. Maneshni	ML Load Forecasting (Survey + Models)	Global datasets (review + experimental)	Comprehensive hybrid ML forecasting thesis	90–93%
2019	Fang et al.	CNN–LSTM Hybrid	Smart grid load datasets	Effective spatial-temporal feature extraction	96%

Fig. 1



6.Methodology

The reviewed literature highlights a strong shift toward hybrid deep learning models, particularly those combining LSTM networks with models such as XGBoost, CNN, GRU, and transformer architectures. These hybrid designs aim to capture both temporal dependencies and nonlinear feature interactions.

1. Data Acquisition and Preprocessing:
Studies such as Semmelmann et al. (2022) and Karimidehkordi et al. (2025) used smart meter–based datasets with temporal granularity ranging from minutes to hours. Standard preprocessing steps included missing-value imputation, normalization, outlier handling, and feature engineering (e.g., lag features, weather variables, event indicators).

2. Hybrid Model Design:
The LSTM–XGBoost pipeline was the most prevalent approach. LSTM layers extracted sequential features, while XGBoost leveraged these learned representations for refined, tree-based prediction. Variants included:
- Wavelet-enhanced models (Ajder et al., 2025)
- Combined PV forecasting using XGBoost–LSTM (Tan et al., 2022)
- CNN–LSTM and GRU–LSTM multi-model frameworks (Abdulameer & Ibrahim, 2025)
- Transformer-based temporal encoders (Hussain et al., 2025)

3. Training and Validation:
Most studies used an 80–20 train-test split, with rolling time-window validation for robustness. Hyperparameter tuning involved grid search, Bayesian optimization, or model-specific tuning (e.g., number of LSTM units, learning rates, tree depths).

4. Performance Metrics:
Studies consistently employed RMSE, MAE, and MAPE to evaluate performance. Some also used R^2 and NRMSE for additional reliability checks.

7.Results and Discussions

Across the literature, hybrid models consistently outperformed single-model architectures. Key findings include:
1. LSTM-XGBoost models improved prediction accuracy by 5–25% compared to standalone LSTM or XGBoost models.
2. Wavelet-enhanced versions handled abrupt fluctuations better, reducing error in datasets with unpredictable events.
3. Transformer-based models such as CAT-Former demonstrated superior performance in high-variability environments like EV charging stations.
4. CNN–LSTM and multi-neural hybrid frameworks provided strong feature extraction for high-dimensional input data.
5. Cloud-native scalable architectures (Wang et al., 2015) supported real-time forecasting with lower latency.

The collective findings suggest that single-model forecasting approaches may no longer be adequate for modern grid environments characterized by volatility, renewable integration, and distributed energy systems. Hybrid models provide four major benefits:

1. Enhanced Feature Learning: LSTM extracts temporal structure, while XGBoost and CNN capture nonlinear or spatial patterns.
2. Robustness to Noise: Wavelet-transformed inputs improve stability in datasets with abrupt irregularities.
3. Adaptability: Transformer-based frameworks generalize well across high-frequency and multi-source datasets.
4. Practical Deployment: Cloud-native solutions support real-time industry use cases.

However, challenges remain. Hybrid models may require higher computational resources, careful hyperparameter tuning, and large datasets to prevent overfitting. Some studies also emphasize limited model interpretability, a critical issue for energy management decision-making.

Overall, hybrid neural architectures represent one of the most promising directions for future load forecasting research, with strong evidence supporting their adoption across community energy systems, smart grids, EV charging infrastructure, and national demand forecasting.

8. Limitations

Although hybrid AI models for electricity demand forecasting have advanced, they still face practical limits. The main problem is that many countries lack complete, clean records. Weather stations skip days; renewable output is logged in incompatible formats and socio-economic figures arrive late or not at all. When records exist, they use different units, time steps or definitions - one model trained on Region A data often fails in Region B. Sudden events - heatwaves, political unrest, cloud ramps - also break the learned patterns - the forecast error jumps because the past no longer predicts the present.

Data Quality & Availability: Missing data values significantly impact on forecast accuracy while granularity and uneven spread of smart meter coverage does little to help. This is particularly true for emerging countries where the lack of historical data leads to models that cannot effectively capture short-term load changes.

Transferability & Robustness: Models tend to be hyper-local for a given climatic zone; these models do not perform well in a different climate zone. They also lack the robustness against abrupt, non-linear events like severe weather or geopolitical instability because sequential patterns easily break down in places other than areas, they have been trained on.

Computational & Engineering Overhead: Dependency on manual feature engineering obstructs automation, and high computational requesting enhances the costs of deployment that is preventing multiple smaller utilities without strong IT infrastructure from experiencing these systems.

9. Future Improvements

Future work should focus on four tasks. Build a single, free, global archive that stores weather, renewable output, load, holidays plus policy shifts in one format with strict quality flags. Develop lightweight hybrids that train on a laptop in under an hour yet keep accuracy within one percent of the heavyweight versions. Embed transparent modules that list the drivers behind each forecast in plain text - if the network relies on temperature, calendar and wind speed, the report states exactly that. By tackling data gaps, event blindness, resource demands and opacity, the next generation of hybrid AI tools will deliver forecasts that grid planners anywhere can understand, run but also trust.

References

1. Semmelmann, L., Henni, S., & Weinhardt, C. “Load forecasting for energy communities: a novel LSTM-XGBoost hybrid model based on smart meter data.” *Energy Informatics*, vol. 5, 2022, Art. 24. DOI: **10.1186/s42162-022-00212-9** SpringerOpen+1

2. Karimidehkordi, F., Samizadeh, R., & Ameri, M. “Hybrid LSTM-XGBoost Model for Sector-Specific Electricity Consumption Prediction in Iran: Incorporating Climate Scenarios.” *International Journal of Multiphysics*, vol. 19, no. 1, 2025, pp. 686–711. DOI: **10.1177/…** (the paper lists on their site, but they don’t show a standard DOI format; use their link) The Multiphysics Journal

3. Altınbaş University team. “Optimizing Smart Grid Load Forecasting via a Hybrid LSTM-XGBoost Framework: Enhancing Accuracy, Robustness, and Energy Management.” *Energies*, vol. 18, no. 11, 2025. DOI: **10.3390/en18112842**. MDPI

4. Ajder, A., Hamza, H. A. A., & Ayaz, R. “Wavelet-Enhanced Hybrid LSTM-XGBoost Model for Predicting Time Series Containing Unpredictable Events.” *IEEE Access*, vol. 13, 2025, pp. 58671–58679. DOI: **10.1109/ACCESS.2025.3556540**. durham-repository.worktribe.com

5. Tan, H., Yang, Q., Xing, X., Huang, K., Zhao, S., & Hu, H. “Photovoltaic Power Prediction Based on a Combined XGBoost–LSTM Model.” *Acta Energiae Solaris Sinica*, vol. 43, no. 8, 2022, pp. 75–81. DOI: **10.19912/j.0254-0096.tynxb.2021-0005**. tynxb.org.cn

6. Hussain, A., Lu, Q.-C., Rizvi, S. S., Wang, S., & Kwon, S.-J. “Short-term demand forecasting of electric vehicle charging stations using a context-aware temporal transformer model (CAT-Former).” *Scientific Reports*, 2025. DOI: **10.1038/s41598-025-20557-x**. PMC

7. Yaling Zhang. “Application of LSTM and Transformer Hybrid Model for Electricity Consumption Forecasting.” *Journal of Energy Research and Reviews*, vol. 17, no. 6, 2025, pp. 71–87. DOI: **10.9734/jenrr/2025/v17i6423**. journaljenrr.com

8. Xuhui Wang, Yang Wu, Wentao Zou, Xu Zhao. “Hybrid Time Series Forecasting for Real-Time Electricity Market Demand Using ARIMA-LSTM and Scalable Cloud-Native Architecture.” *Informatica*, 2015. DOI: **10.31449/inf.v49i3.9474**. informatica.si

9. Abdulameer, Y. H., & Ibrahim, A. A. “Forecasting of Electrical Energy Consumption Using Hybrid Models of GRU, CNN, LSTM, and ML Regressors.” *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, vol. 16, no. 1, 2025, pp. 560–575. DOI: **10.58346/JOWUA.2025.I1.033**. jowua.com

10. A. Maneshni. “Load Forecasting in the Era of Smart Grids: Opportunities and Advanced Machine Learning Models.” (Thesis). arXiv, May 2025. (Preprint) arXiv ID: **2505.18170**. arXiv

11. Ding, Y., Wu, D., He, Y., Luo, X., & Deng, S. “Highly-Accurate Electricity Load Estimation via Knowledge Aggregation.” arXiv, Dec. 2022. arXiv ID: **2212.13913**. arXiv

12. Grandón, T. G., Schwenzer, J., Steens, T., & Breuing, J. “Electricity Demand Forecasting with Hybrid Statistical and Machine Learning Algorithms: Case Study of Ukraine.” arXiv, Apr. 2023. arXiv ID: **2304.05174**. arXiv

13. A hybrid neural network model for power demand forecasting: combining LSTM and CNN, MDPI *Energies* (or similar) — DOI: **10.3390/en12050931** from: A Hybrid Neural Network Model for Power Demand Forecasting. MDPI

14. Lee, Juyong; Cho, Youngsang. “National-scale electricity peak load forecasting: Traditional, machine learning, or hybrid model?” arXiv, July 2021. arXiv ID: **2107.06174**.