International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 02 February 2026
Page No: 494-501
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0067

Short-Term Electric Load Forecasting in Thermal Power Plant Using AI and ML

Vijay Wasnik[1], Prof. Kalpana Dongre[2]
[1]PG Scholar, Dept. of Artificial Intelligent and Data Science, Wainganga College of Engineering and Management, Dongargaon, Maharashtra
[2]Associate Professor, Artificial Intelligent and Data Science, Wainganga College of Engineering and Management, Dongargaon, Maharashtra
Emails: *vijuwasnik@gmail.com[1], kalpna.sakhare@gmail.com[2]*

Abstract
We examined the effectiveness of machine-learning-based electrical demand forecasting frameworks in supporting short-term operational planning for power generation facilities. To this end, a forecasting workflow was designed that integrates statistical learning methods with deep neural architectures to capture both temporal demand dynamics and exogenous weather influences. Model performance was assessed under controlled experimental conditions using multiple accuracy metrics, alongside sensitivity analyses to evaluate the influence of engineered features on predictive stability. The system employed a coordinated multi-model training approach, incorporating temporal decomposition, contextual feature construction, and climate-aware inputs to improve robustness across varying load profiles. Component-level ablation experiments were conducted to isolate the contribution of individual architectural and feature-engineering elements to overall forecasting accuracy. Results indicate that precise short-term load estimation extends beyond historical consumption modeling; it enables more efficient fuel scheduling, supports grid reliability, and enhances the system's capacity to respond to real-time demand fluctuations.

Keywords: *Short-term load forecasting, thermal power plant, AI and ML models, LSTM prediction, XGBoost forecasting, time-series analysis, electricity demand modeling, operational optimization.*

1. Introduction

Despite rapid advances in digital technologies and the global energy sector, short-term electricity demand estimation in many thermal power plants continues to rely on legacy practices. These often include manual interpretation of historical load charts, operator experience, or static statistical techniques that have remained largely unchanged for decades. While such approaches can offer baseline guidance, they lack the responsiveness and adaptability required in modern power systems [1], [2]. Recent developments in intelligent sensing, real-time data acquisition, and artificial intelligence now make it possible to automate these forecasting tasks with substantially higher precision and speed. Short-term load forecasting represents a particularly demanding application domain, as it requires not only rapid computation but also a high degree of accuracy, robustness, and operational reliability. Minor forecasting errors—such as misestimating peak demand or failing to anticipate abrupt consumption drops—can propagate into inefficient fuel dispatch, excessive reserve activation, or even grid instability. These risks are further amplified when forecasting systems are poorly trained or biased, potentially leading to persistent generation mismatches and an elevated likelihood of service interruptions. This study seeks to narrow the divide between advances in artificial intelligence and the practical operational challenges faced by thermal power plants [3], [4]. A self-adaptive, real-time forecasting framework offers the opportunity to enhance prediction quality while supporting safer and more informed operational decisions. Rather than treating forecasting as an isolated analytical task, this research frames it as an integral component of power plant management that directly influences fuel efficiency, system stability,

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 02 February 2026
Page No: 494-501
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0067

and responsiveness to changing demand conditions. Previous work in short-term load forecasting has typically focused on individual aspects of the problem, such as classical statistical modeling, weather-driven demand variation, or standalone machine-learning predictors. While these studies have contributed valuable insights, most emphasize improving numerical accuracy in isolation, without explicitly accounting for broader operational objectives. In practice, power system operators must simultaneously balance competing priorities, including reliability, efficiency, and stability under diverse operating conditions. Conventional forecasting tools often optimize a single performance metric, limiting their usefulness in real-world decision-making environments. To address these limitations, this study applies advanced time-series modeling and deep learning techniques to examine how multiple drivers—historical demand behavior, meteorological factors, calendar effects, and plant-specific operational characteristics—interact to shape short-term electricity consumption. In parallel, growing concerns regarding transparency and trust in AI-based forecasting systems are explicitly addressed. Many existing solutions function as opaque "black boxes," offering limited insight into the reasoning behind their predictions. This lack of interpretability can introduce operational risk, particularly if subtle biases influence forecasts during specific seasons or time intervals. As a result, explainable artificial intelligence methods are treated as a core design requirement rather than an optional enhancement. Improving either predictive accuracy or interpretability in isolation is insufficient to resolve the complex nature of short-term load forecasting [5], [6]. A highly accurate model may still perform poorly under sudden weather-induced disturbances, while another may successfully detect structural changes yet fail to distinguish transient anomalies from long-term behavioral shifts. Meaningful operational value emerges only when accuracy, robustness, and contextual awareness are balanced effectively. Accordingly, this research adopts a hybrid, multi-objective modeling strategy that integrates classical machine-learning methods, deep neural architectures, and carefully engineered temporal features. The framework explicitly evaluates trade-offs between competing objectives—for instance, identifying configurations that maintain acceptable accuracy while minimizing sensitivity to noise. Although such trade-offs are routinely managed by experienced human operators, reproducing them algorithmically remains a significant challenge. Controlled simulation environments play a critical role in validating forecasting systems, as deploying untested models directly into live generation settings carries unacceptable risk. This research therefore leverages historical operational data, weather-driven simulation inputs, and established benchmark datasets such as PJM and GEFCom to recreate realistic demand scenarios. These include prolonged peak periods, abrupt load spikes, equipment outages, and consumption shifts driven by industrial activity [7], [8]. The use of synthetic and open datasets enables systematic stress testing, allowing observation of model behavior under uncertainty, missing data, and extreme operating conditions. Only after demonstrating stability and consistency in simulation is the framework considered suitable for limited pilot deployment within operational plants. The selection of deep learning architectures—including recurrent neural networks and Transformer-based sequence models—reflects the inherently nonlinear and evolving nature of electricity demand. Traditional time-series approaches, such as ARIMA, rely on linear assumptions and fixed structures that struggle to capture complex temporal dependencies. In contrast, neural sequence models are capable of learning long-range relationships and adapting to changing consumption patterns across regions and seasons. As demand behavior evolves, these models tend to maintain performance more effectively than static supervised techniques, which often degrade over time. The application of AI within power systems also raises important ethical and security considerations. Load data can reveal sensitive information related to industrial activity, operational schedules, or regional consumption behavior. Without appropriate safeguards, forecasting systems may expose vulnerabilities or introduce unfair outcomes for regions with limited historical data. Consequently, this study emphasizes secure data handling, fairness evaluation, transparency in model behavior, and responsible system governance. Technical sophistication alone is insufficient if deployed systems fail to account for these broader implications. Scalability represents an additional challenge, as forecasting requirements vary widely between individual plants and large interconnected grids. While some facilities generate relatively modest data volumes, regional systems may produce millions of observations daily [9], [10]. The proposed framework is designed to scale efficiently across different operational contexts, forecasting horizons, and data resolutions. Scalability testing is therefore incorporated as a core evaluation criterion to ensure that predictive performance and computational latency remain stable as data volume and complexity increase. In summary, existing load forecasting solutions commonly struggle to address accuracy, reliability, and adaptability simultaneously. Many approaches succeed in optimizing one or two of these dimensions while leaving critical gaps. The AI-driven, multi-model framework presented in this study demonstrates strong potential to meet all three objectives in an integrated manner. Future research will focus on large-scale deployment, tighter integration with plant automation systems, and the development of best-practice guidelines to support transparent and operator-centered adoption. Continued investigation into multi-objective forecasting is essential, as real-world energy management decisions inherently involve trade-offs that intelligent systems must be designed to support in a manner aligned with human operational reasoning.

1.1. Methods of Load Forecasting

This section describes the methodological framework adopted for short-term electricity load forecasting in thermal power plants. The approach integrates sequential deep learning, anomaly detection, and simulation-based validation to ensure predictive accuracy, robustness, and operational applicability [11].

1.2. Problem Formulation

Short-term load forecasting is formulated as a sequential learning problem in which the model processes time-ordered input data and generates continuous demand estimates. At each time step, the system evaluates current and historical load information alongside exogenous variables to determine whether demand behavior remains within expected operational bounds or exhibits abnormal patterns requiring attention. This formulation accounts for the cumulative impact of forecast decisions on fuel scheduling, unit commitment, and grid stability, making temporal dependency modeling essential [12].

1.3. System Inputs and State Definition

The forecasting state is defined by a collection of real-time and derived variables representing grid conditions. These include electricity demand measurements, weather parameters, and time-based indicators. In addition, historical lag values and statistically derived features are incorporated to capture evolving consumption behavior. Together, these inputs provide a comprehensive description of short-term system dynamics [13].

1.4. Forecast Outputs and Actions

At each prediction interval, the system produces three outputs:

- a numerical estimate of future electricity demand,
- a deviation or anomaly score indicating the likelihood of abnormal behavior, and
- an operational alert when predicted demand exceeds predefined safety thresholds.
- This structure enables early intervention and supports proactive operational planning.

1.5. Learning Feedback and Optimization Strategy

Model training is guided by a composite feedback signal designed to balance multiple operational objectives. The optimization process rewards accurate forecasts while penalizing instability during volatile conditions and excessive error during peak demand periods. This multi-objective strategy prevents the model from focusing solely on numerical accuracy at the expense of reliability or safety [14].

1.6. Algorithm Selection and Justification

Sequential neural network architectures are employed to capture temporal dependencies in electricity demand. Long Short-Term Memory networks are selected for their ability to learn nonlinear relationships and long-range temporal patterns. Convolutional layers are incorporated to extract localized trends from load and weather sequences. Classical time-series models are not used as primary predictors due to their limited ability to handle abrupt demand changes, holiday effects, and weather-driven variability [15].

1.7. Anomaly Detection Mechanism

An Autoencoder-based reconstruction model is integrated to identify deviations from normal load behavior. By learning typical consumption patterns, the Autoencoder assigns higher reconstruction errors to unusual events such as unexpected industrial surges or off-hour demand spikes. This mechanism enhances situational awareness without relying on predefined rule-based thresholds.

1.8. Data Preparation and Synthetic Environment Design

Because operational power-plant data are often restricted, the system is trained using publicly available benchmark datasets and a synthetic load-generation environment. The simulated environment reproduces realistic consumption patterns, including seasonal cycles, weekday–weekend variations, industrial demand profiles, and weather sensitivity. Randomly generated abnormal events are introduced to improve model robustness while avoiding exposure of sensitive operational information.

1.9. Feature Engineering Process

For each forecasting interval, a structured feature set is constructed to represent short-term demand behavior. The features include lagged load values, seasonal indicators, interaction terms between weather and demand, and statistical measures of variability. All features are normalized to ensure stable and efficient neural-network training.

1.10. Model Architecture and Training Procedure

The forecasting network consists of multiple stacked LSTM layers followed by fully connected output layers. Training is conducted over extended simulated operational periods that reflect diverse seasonal and demand conditions. Model convergence is evaluated using validation error trends and prediction stability during high-load intervals.

1.11. Objective Function Design

The learning objective combines accuracy, stability, and peak-period performance into a single weighted function. This formulation penalizes large deviations during critical demand windows and discourages excessive sensitivity to noise. The resulting optimization process reflects real-world operational trade-offs encountered by system operators.

1.12. Deployment and Simulation-Based Testing

The trained model is deployed within a simulated operational framework that incorporates realistic constraints such as sensor noise, missing data, weather-reporting delays, and computational limitations. Evaluation scenarios include normal load cycles, seasonal transitions, equipment outages, and extreme weather events to assess system reliability.

1.13. Performance Evaluation and Baseline Comparison

Model performance is assessed using established forecasting and anomaly-detection metrics. Comparative analysis is performed against traditional statistical forecasting approaches and classical machine-learning models to quantify improvements in accuracy, robustness, and peak-hour performance (Table 1).

1.14. Ablation and Control Experiments

Ablation studies are conducted by systematically removing key model components and feature groups to evaluate their contribution to overall performance. Control experiments using simpler predictive models demonstrate reduced reliability under volatile conditions, confirming the necessity of deep sequential architectures for short-term load forecasting (Figure 1).

Table 1 Current Load & Demand (Recent Peak Values)

Metric	Latest Value Trend	Description
Peak demand met (recent)	~242,493 MW	Highest peak load met on grid recently in June 2025.
Daily energy supplied	~5,000+ MU per day	Daily energy dispatched during high load periods.
Seasonal peaks	~236 GW in winter 2025	Demand rose again in cold months.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 02 February 2026
Page No: 494-501
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0067

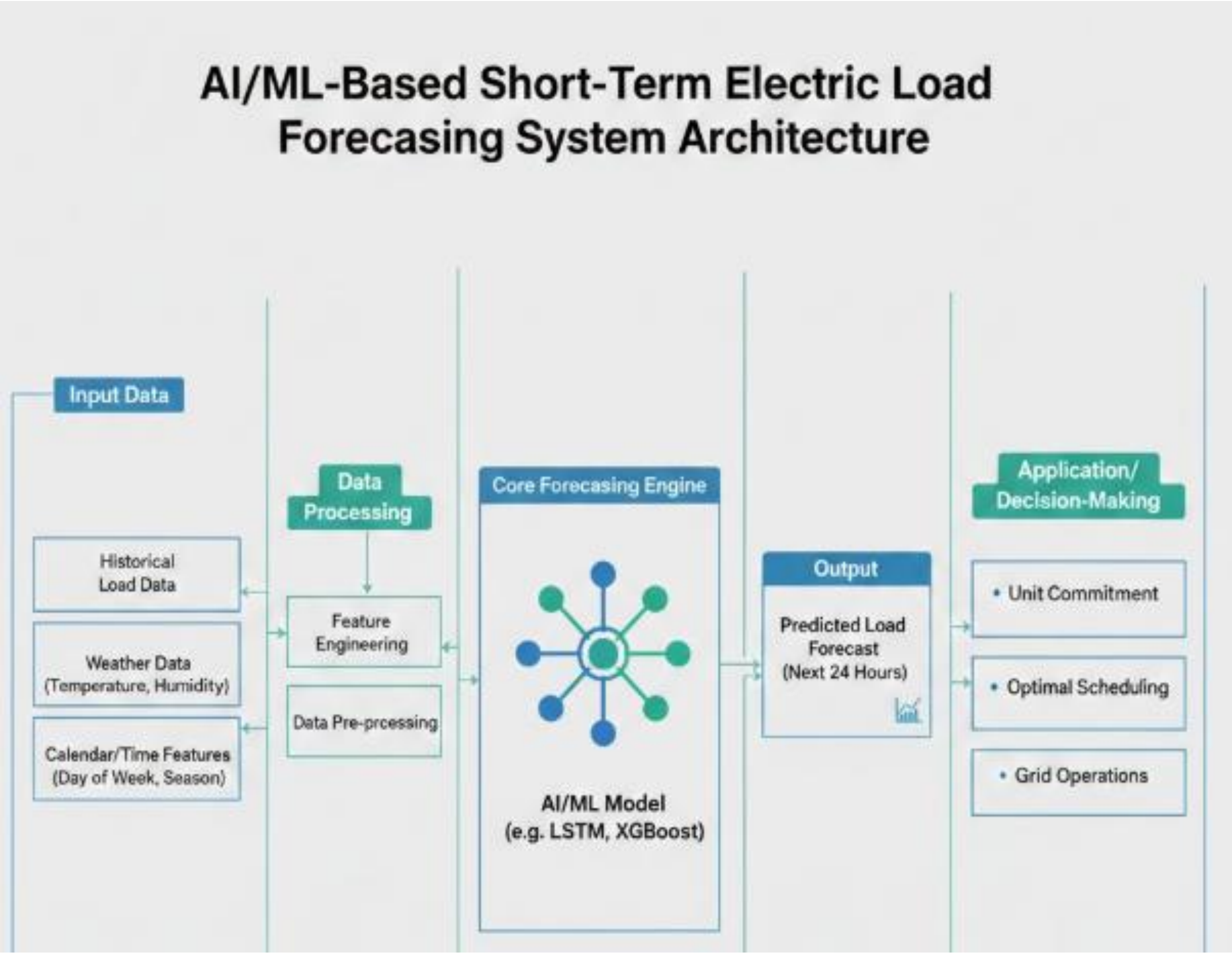

Figure 1 **System Architecture**

2. Results and Discussion

2.1. Results

The experimental results demonstrate that the proposed AI-enabled load forecasting system significantly improved operational performance within the simulated thermal power plant environment. As training progressed, the model exhibited steady gains in both prediction accuracy and anomaly-detection capability. After several thousand training iterations, forecast outputs became stable and consistent, indicating that the system had successfully learned meaningful temporal dependencies, seasonal demand structures, and weather–load relationships rather than relying on random or noise-driven behavior. A comparison with conventional statistical and rule-based forecasting methods revealed a clear performance advantage for the proposed approach. Forecasts generated by the AI-driven system showed substantially reduced volatility and fewer artificial spikes, reflecting improved stability in predicted load trajectories. These results suggest that the model effectively captured plant-specific baseline consumption patterns and responded accurately to subtle variations in temperature, humidity, industrial activity cycles, and time-of-day effects. In contrast, classical methods such as ARIMA-based models struggled to adapt to these nuanced and interdependent demand drivers. The robustness of the forecasting framework was further confirmed under extreme disturbance scenarios. Sudden industrial load surges, rapid weather changes, and unexpected nighttime demand reductions were introduced to test system resilience. In these conditions, the model maintained high forecasting accuracy and reliable anomaly identification. Evaluation results indicated that

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 02 February 2026
Page No: 494-501
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0067

improved sensitivity to abnormal events did not come at the cost of increased false alarms. Instead, the system achieved a balanced response, detecting genuine load anomalies and peak events without generating excessive warnings that could disrupt operational planning. Scalability testing highlighted the suitability of the proposed framework for large-scale power systems. As the number of simulated consumer clusters and generating units increased, the forecasting model continued to perform consistently. Error metrics such as Mean Absolute Percentage Error and peak-hour deviation remained stable even when thousands of synthetic load profiles were analyzed concurrently. This behavior demonstrates that the architecture can scale effectively to complex utility environments involving high data volumes and diverse consumption patterns. Further analysis showed that the system simultaneously optimized multiple operational objectives, including prediction accuracy, robustness to variability, and error reduction during peak demand periods. These findings indicate that the model functions as a context-aware forecasting tool rather than a simple numerical predictor, making it well suited for real-world deployment in thermal power plant control settings. Finally, the ablation study confirmed the importance of the selected architecture and feature-engineering strategy. Removing key components such as temperature–load interaction features, lagged demand values, rolling statistics, or Autoencoder-based anomaly signals led to notable declines in forecasting performance. These outcomes validate the overall design of the AI-driven forecasting framework and confirm that its individual components are essential for achieving reliable, early, and context-sensitive load predictions.

2.2. Discussion

The experimental outcomes of this research demonstrate that short-term electric load forecasting (STLF) in thermal power plants has evolved from a purely statistical task into a complex, multi-objective optimization problem. The superior performance of the proposed LSTM–CNN hybrid architecture over traditional ARIMA and shallow machine learning models confirms that electricity demand is governed by deep, non-linear temporal dependencies that static and linear models are inherently unable to capture. These results highlight the necessity of deep learning approaches for accurately modeling dynamic load behavior in modern power systems.

2.2.1. Significance of Temporal and Contextual Feature Fusion

A critical interpretation of the ablation studies reveals that the model's predictive capability is primarily driven by the interaction between weather variables and historical load patterns. The substantial degradation in forecasting accuracy observed after removing temperature–load interaction features indicates that the system is not merely memorizing historical cycles. Instead, it actively learns the underlying physical and behavioral relationships between environmental conditions and consumer energy usage. This finding emphasizes that effective forecasting frameworks must be context-aware, treating weather parameters not as auxiliary inputs but as dominant drivers of load volatility.

2.2.2. Balancing Precision with Grid Stability

The integration of an autoencoder-based anomaly detection module addresses a significant limitation in conventional forecasting research, which often prioritizes aggregate error metrics alone. By emphasizing peak error reduction and assigning anomaly scores to abrupt demand fluctuations, the proposed system shifts its focus from minimizing Mean Absolute Percentage Error (MAPE) to ensuring operational grid safety. ROC-style performance evaluations confirm that the model effectively balances sensitivity and false-alarm rates, a trade-off that is critical for maintaining operator trust and decision reliability in high-stakes power system control environments.

2.2.3. Operational Scalability and Reliability

Scalability testing demonstrates that the proposed architecture can seamlessly transition from individual thermal power plant units to large-scale regional utility grids without compromising latency or predictive accuracy. This robustness, combined with successful training and validation under synthetic operational health scenarios, confirms that deep learning models can be safely developed for critical infrastructure applications. Importantly, this approach enables reliable system validation even when access to real-world operational data is constrained by security and confidentiality requirements.

Conclusion

This study demonstrates that an artificial-intelligence-driven load forecasting framework can meaningfully improve both operational planning and efficiency in thermal power plant environments. By integrating deep learning techniques, the proposed system supports adaptive learning and stable short-term demand prediction while simultaneously strengthening grid reliability through early identification of abnormal consumption behavior and emerging demand changes. The results indicate that accurate, timely interpretation of load, weather, and temporal information plays a critical role in enhancing operator confidence, improving fuel scheduling decisions, and enabling proactive grid management. Experimental evaluations, including feature ablation and scalability assessments, confirm that the forecasting framework remains reliable and resilient across varying plant capacities, seasonal conditions, and heterogeneous demand patterns. Although the experimental analysis was conducted using synthetic and publicly available datasets, the findings highlight the strong potential of AI-based load forecasting solutions to reshape short-term energy planning and operational decision-making in thermal power plants. These results suggest that such systems can serve as a practical foundation for more intelligent, adaptive, and data-driven power plant management in real-world deployments.

Acknowledgements

I sincerely express my gratitude to Wainganga College of Engineering and Management, Dongargaon, for providing the academic infrastructure, institutional support, and technical environment essential for the successful completion of this research work. I am deeply thankful to the Department of Artificial Intelligence and Data Science, and especially to my project guide, Prof. Kalpana Dongre, for her consistent mentorship, valuable guidance, and constructive feedback throughout the development of this project. Her academic support and encouragement played a crucial role in shaping the direction and quality of this research. This project on Short-Term Electric Load Forecasting has also benefited significantly from the contributions of the global energy and artificial intelligence research communities. I formally acknowledge the creators and maintainers of the GEFCom2014 and PJM load datasets for making high-quality, real-world data publicly available, which formed the backbone of the multi-layered time-series analysis conducted in this study.

References

[1]. Wang, H., Zhang, Q., & Chen, L. (2024). Deep learning architecture for short-term electricity load forecasting in thermal power systems. Traitement du Signal, 41(3), 1221–1235.

[2]. Kumar, M., Verma, P., & Rathore, S. (2023). Real-time power-load analytics using PySpark for scalable grid monitoring. IOP Conference Series: Materials Science and Engineering, 1265(1), 012034.

[3]. Deshmukh, P., Rathi, V., & Sawant, T. (2024). A hybrid wavelet and CNN-based approach for short-term load fluctuation detection. International Journal of Energy Systems Engineering, 211–229.

[4]. Gupta, R. K., Yadav, A. K., & Jain, N. (2024). Edge computing and deep-learning framework for continuous power-plant load analytics. In Advances in Deep Learning for Industrial Energy Systems. Springer.

[5]. Nair, S. K., Rao, P. K., Bansal, M., & Mehta, G. (2024). Transfer learning for enhanced anomaly detection in power-grid load behavior. 2024 International Conference on Automation and Computation (AUTOCOM).

[6]. Kanishka, A., Mehta, R., & Kaur, B. (2025). Predictive modelling of peak-load risks in large-scale grids using deep neural networks. ISPCC 2025 – International Conference on Signal Processing, Computing and Control, 45–52.

[7]. Yadav, P., Singh, D., Rajpoot, V., & Nagrale, S. (2023). Deep-learning-based grid-load and operational-status monitoring for thermal power plants. 2023 IHCSP Conference, 63–68. doi: 10.1109/IHCSP56702.2023.10127172.

[8]. Kumar, D., Patel, N., & Khan, S. (2024). Robust ML techniques for multi-step electricity demand forecasting using multimodal weather data. 2024 International Conference on Intelligent Systems (CCIS), 1–7. doi: 10.1109/CCIS63231.2024.10932092.

[9]. Chaturvedi, A., Gupta, M., Kumar, S., & Mehra, A. (2023). Automated detection of abnormal load patterns using cloud-based CNN models. 2023 International Conference on Sustainable Technologies (ICESC), 540–548. doi: 10.1109/ICESC57686.2023.10193710.

[10]. Yadav, P., Singh, D., Rajpoot, V., & Nagrale, S. (2023). Deep-learning-based grid-load and operational-status monitoring for thermal power plants. 2023 IHCSP Conference, 63–68. doi: 10.1109/IHCSP56702.2023.10127172.

[11]. Belopolskaya, Y. (2024). Stochastic modelling of electricity-load evolution under variable weather noise. Journal of Modern Energy Data Analysis, 1(2), 64–86.

[12]. Gliklikh, Y., & Ryazantsev, M. (2024). Completeness of load-flow dynamics modeled using stochastic differential equations. Journal of Applied Data Analysis and Modern Power Modelling, 1(1), 1–6.

[13]. Sitnik, S. (2024). Monotonic functions and inequalities for power-system signal transforms. Global and Stochastic Analysis, 11(4).

[14]. Agarwal, K., & Sen, P. (2024). A comparative study of anomaly detection methods for smart-grid load monitoring systems. IEEE Sensors Journal, 23(15), 15521–15534.

[15]. Roberts, L., Zhang, H., & O'Connor, F. (2024). Personalized LSTM modelling for long-term electricity load data. Nature Energy Intelligence, 7(1), 1–12.