

Article

Parallel Multi-Model Energy Demand Forecasting with Cloud Redundancy: Leveraging Trend Correction, Feature Selection, and Machine Learning

Kamran Hassanpouri Baesmat *, Zeinab Farrokhi, Grzegorz Chmaj and Emma E. Regentova 

Department of ECE, University of Nevada, Las Vegas, NV 89154, USA; zeinab.farrokhi@unlv.edu (Z.F.); grzegorz.chmaj@unlv.edu (G.C.); emma.regentova@unlv.edu (E.E.R.)

* Correspondence: kamran@unlv.edu

Abstract: In this work, we present a novel approach for predicting short-term electrical energy consumption. Most energy consumption methods work well for their case study datasets. The proposed method utilizes a cloud computing platform that allows for integrating information from different sources, such as weather data and historical energy consumption, while employing machine learning techniques to achieve higher accuracy in forecasting. We collected detailed weather data from the “Weather Underground Company” website, known for its accurate records. Then, we studied past energy consumption data provided by PJM (focusing on DEO&K, which serves Cincinnati and northern Kentucky) and identified features that significantly impact energy consumption. We also introduced a processing step to ensure accurate predictions for holidays. Our goal is to predict the next 24 h of load consumption. We developed a hybrid, generalizable forecasting methodology with deviation correction. The methodology is characterized by fault tolerance due to distributed cloud deployment and an introduced voting mechanism. The proposed approach improved the accuracy of LSTM, SARIMAX, and SARIMAX + SVM, with MAPE values of 5.17%, 4.21%, and 2.21% reduced to 1.65%, 1.00%, and 0.88%, respectively, using our CM-LSTM-DC, CM-SARIMAX-DC, and CM-SARIMAX + SVM-DC models.



Academic Editors: Luigi Grossi and Sonia Leva

Received: 23 April 2025

Revised: 21 May 2025

Accepted: 27 May 2025

Published: 31 May 2025

Citation: Hassanpouri Baesmat, K.; Farrokhi, Z.; Chmaj, G.; Regentova, E.E. Parallel Multi-Model Energy Demand Forecasting with Cloud Redundancy: Leveraging Trend Correction, Feature Selection, and Machine Learning. *Forecasting* **2025**, *7*, 25. <https://doi.org/10.3390/forecast7020025>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: IoT; cloud computing; machine learning; feature selection; electrical load prediction; power systems; statistical methods; deep learning; artificial intelligence; deviation correction

1. Introduction

Modern energy management systems rely on accurate energy consumption forecasting to maintain power grid reliability and efficiency. Electrical load forecasting plays a crucial role in ensuring effective energy distribution and grid stability, allowing utilities to predict demand fluctuations. Reliable predictions lead to better resource management, cost savings, and enhanced power system dependability. Two major aspects affect the efficiency of forecasting: model and data.

Machine learning and hybrid models that combine different methods to improve prediction accuracy have gained popularity due to their high accuracy. The authors in [1] used a machine learning approach based on household usage data to improve time series forecasting and demonstrated that the hybrid model outperforms both individual and ensemble models in terms of accuracy. In [2], the authors introduced a new method combining statistical methods with a land consumption model for future energy demand. In [3], the researchers used neural networks and the bee colony method for better short-term

load forecasting. The authors in [4] used hybrid regression techniques for their predictions. A 99.2% prediction accuracy for peak electricity demand was attained with a hybrid model based on faster k-medoids clustering, support vector machine, and artificial neural network in [5].

In [6], the authors showed that deep learning methods in hybrid models efficiently handle nonlinear patterns. The authors in [7] tested four ML models on two different sites and across four forecasting horizons. Khan and Byun [8] created an ensemble method with optimized feature selection that outperformed single models. The authors in [9] used a clustering–classification hybrid method for short-term consumption forecasting. The researchers in [10] reviewed hybrid methods (such as statistical and machine learning techniques) in building energy prediction, and Chou and Ngo [11] applied a combination of metaheuristic sliding window methods based on ML systems with time-series models for identifying building energy consumption patterns. Models like SVM, ANN, DNN, and optimization algorithms were demonstrated in [12–14]. It was also shown that combining ML methods with ensemble weather forecasts could produce more accurate estimates of future energy consumption [15].

The quality and contents of input data play a crucial role in the performance of energy forecasting models. It was explained in [16] how the selection of weather stations affects electric load forecasting. Jain et al. [17] showed that hourly and location-specific sensor data help predict consumption in multi-family buildings. In [18], the authors emphasized the crucial role of data preprocessing and feature selection in ensuring high-quality inputs for artificial neural network models, thereby improving the accuracy and reliability of short-term load forecasting. Researchers in [19] found that the real weather values are better than averages. Ahmad et al. [20] and Shin and Woo [21] used time series data to show how data type affects prediction results. In [22], the researchers discussed the missing effect of human behavior in many datasets, which makes some models less accurate. Authors of [23] found that LSTM has problems predicting cooling power when behavior is not included. Demolli et al. [24] and Persson et al. [14] used long-term and multi-site data to improve solar and wind predictions. Herrera et al. [25] showed that energy commodity prices are useful for long-term forecasts.

Some papers warned about biased or overfitting problems. Wei et al. [26] concluded that ML models might cause overfitting in long-term predictions. Beyca et al. [27] demonstrated that, in stable time series, statistical methods can work better than ML.

Both energy providers and grid managers have identified smart meter information as a key factor in improving local forecast accuracy [28]. Aside from analyzing error-causing factors like weather changes, holidays, and weekends, it is necessary to analyze factors that could lead to errors in predicting energy consumption [29]. Jain et al. [17] highlighted that the time and location of sensors matters in support vector regression (SVR).

As much as data quality, feature selection significantly affects the forecasting performance, and the proper selection of input features affects model performance [30,31]. In [32], a hybrid system whose forecasting performance was largely improved by selecting useful features was discussed. The authors of [8] used genetic algorithms for feature selection and improved results. Ahmad et al. [7] focused on climate-sensitive and non-sensitive conditions by choosing suitable features. Deb et al. [33] also combined statistical knowledge with engineering ideas for better building energy forecasting. Dong et al. [12] earlier proved that SVR works well in tropical climates when features are well selected.

Smart grid technologies continue to improve rapidly in the energy field. Moreover, supervised learning has shown promise in reducing noise (random, misleading data points) from systems with imbalanced data in the network formulation [34].

For model implementation, cloud computing provides access to more computing power and real-time processing. It offers scalable resources, enhanced flexibility, and cost-efficiency, enabling businesses to access and manage data and applications remotely without the need for extensive on-premise infrastructure.

Hou et al. [35] used a cloud-edge system to handle abnormal data and improve prediction. Zhang et al. [36] created a hybrid method with edge and cloud computing and used ELM models for fast forecasting. Raju and Laxmi [37] used IoT and ML together for real-time load prediction. Geberslassie and Bitzer [38] showed how cloud computing supports future smart grid systems. Hou et al. [39] added a privacy-safe edge-fog-cloud system for load forecasting. Pelluri et al. [40] and Devaraj et al. [41] explained some of the problems of using cloud in forecasting, like resource allocation and privacy. But, overall, the future is moving toward more real-time, scalable, and secure cloud-based models.

The current methodologies for forecasting energy consumption show both their potential and challenges. These studies also demonstrate how cloud computing in electrical load forecasting is diverse and evolving. In addition to addressing challenges such as privacy and real-time processing, cloud solutions have the potential to improve load forecasting model accuracy, efficiency, and scalability. In the above review, we highlighted the transformative impact of parallel approaches that combine ML and cloud computing, underscoring the scalability and resilience these methods bring to energy consumption forecasting.

Our research introduces an innovative and robust energy consumption forecasting model that leverages a parallel approach, employing three distinct machine learning methods alongside weather and dynamic correction strategies. This approach includes an advanced “2-out-of-3” system utilizing three independent cloud platforms, ensuring operational reliability and robustness. Each method operates independently across cloud services such as AWS, Google Cloud, and Microsoft Azure, allowing for fault tolerance in cases where one or two models or cloud infrastructures encounter technical issues. This configuration enhances system resilience and maintains optimal performance, even with substantial datasets, dynamically allocating resources to meet computational demands. The innovative parallelized prediction approach—integrating ML with cloud computing resources—adds a significant layer of reliability and robustness to energy forecasting. Our proposed system, outlined in a multiphase structure (as seen in Figure 1), begins by collecting accredited data, followed by feature selection to identify the most relevant predictors. Subsequent noise reduction techniques address holiday-related variabilities, preparing the data for forecasting. Finally, we run three ML methods across three cloud platforms, achieving greater reliability and robustness due to the unique “2-out-of-3” system—a novel contribution of this research that underpins the system’s high fault tolerance and operational resilience.

The key contributions of our paper can be summarized as follows:

- (a) Hybrid forecasting framework with deviation correction:

We develop a two-stage forecasting approach using historical load data (1 April 2022–30 March 2025). The framework partitions the dataset using the first 80% for training and the remaining 20% for testing.

The proposed model first generates a 24 h load forecast (for 31 March 2025) and then refines the initial prediction by analyzing the deviation trend present within the test data. For the first 5 h of 1 April 2025, the forecast is computed as the sum of the initial predictions and the forecasted deviation, thereby reducing systematic errors in volatile periods.

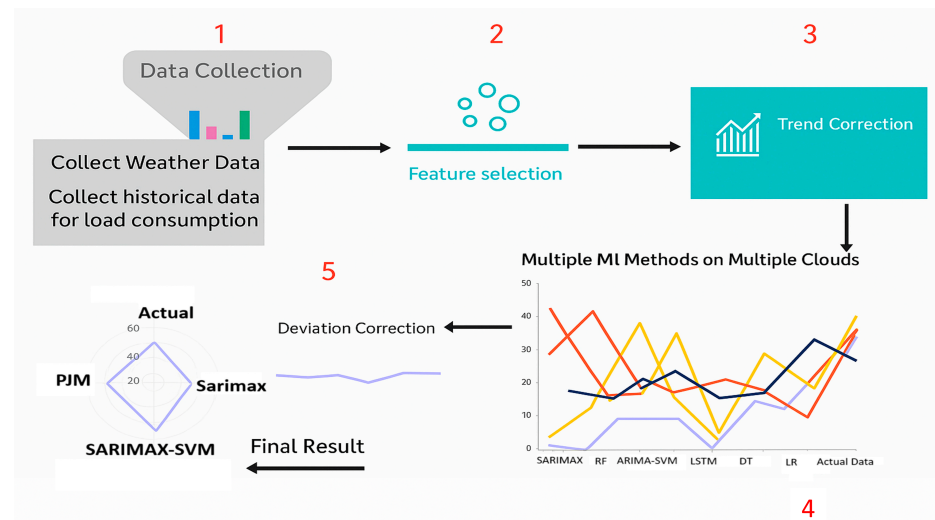


Figure 1. Steps of the proposed load consumption forecasting system.

(b) Fault Tolerance via Distributed Cloud Deployment and Voting:

To enhance the reliability of the forecasting system, we deploy three distinct forecasting models across three independent cloud platforms. We use a “2-out-of-3” voting scheme to determine the final prediction at each time point, ensuring that even if one platform or model produces an outlier result, the consensus from the other two can override it. This approach significantly improves the fault tolerance of the system as the overall prediction reliability remains high unless multiple components fail concurrently.

(c) Generalizable and Adaptable Methodology

By integrating robust feature selection, machine learning, statistical methods, and cloud computing, our framework does not rely on region-specific assumptions. This design facilitates its application to various types of electrical load data and across different geographic regions.

Collectively, these contributions demonstrate the novelty of our approach in simultaneously addressing accuracy through deviation correction and ensuring operational reliability with a fault-tolerant cloud-based ensemble system. This comprehensive methodology paves the way for future enhancements using diverse energy market data.

The rest of the paper is organized as follows: Section 2 explains the methodology of prediction. Then, we show results in Section 3 and present a discussion in Section 4, followed by the conclusion in Section 5.

2. Methodology

2.1. Procedures

The methodology consists of implementing five steps. The procedures of each step are shown in Figure 1.

First, we gathered weather information. To do this, we checked a variety of sources such as the National Weather Service, Weather Channel, AccuWeather Co., Weather Underground Co., and wunderground.com. We selected Wunderground data (1 April 2022–1 April 2025) as the primary source of information based on our consideration of the available sources. Our project significantly benefited from the database’s comprehensive weather information and historical data, and this source provided us with reliable and accurate information to work with. However, the proposed model is built to be flexible for different regions and types of electrical loads. It combines machine learning, statistical methods, and strong feature selection and deviation correction to capture the key patterns of energy use

without being limited to a particular area. Its modular design allows it to work with various datasets and operating conditions, ensuring reliable forecasts even when the data quality and structure differ. This broadly applicable method maintains consistent performance and sets the stage for future improvements with diverse energy market data.

2.2. 1-Collecting Historical Data for Load Consumption

The PJM (PJM is a regional transmission organization (RTO) that coordinates the movement of wholesale electricity in all or parts of 13 states and the District of Columbia) website represents data from the electrical grid in 13 of the United States (Figure 2) and provided us with resources to track and analyze power usage. They are known for providing reliable and official datasets that are essential for studying energy consumption patterns across a wide region.

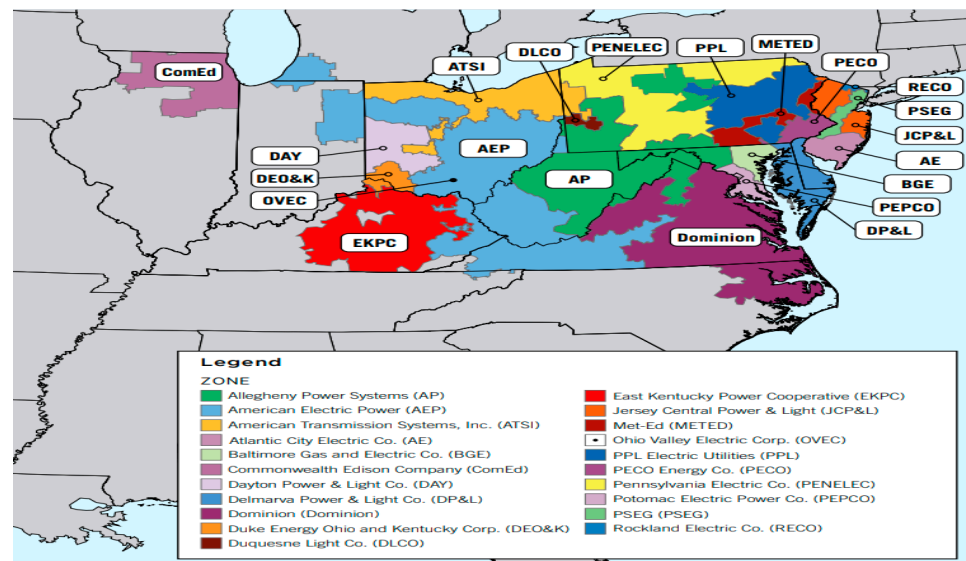


Figure 2. PJM’s dataset from several companies in the United States—courtesy of PJM [42].

As mentioned, the analysis focused on DEO&K Co (Figure 3). According to existing data, the area is the smallest, and the weather data are accurate due to its size.

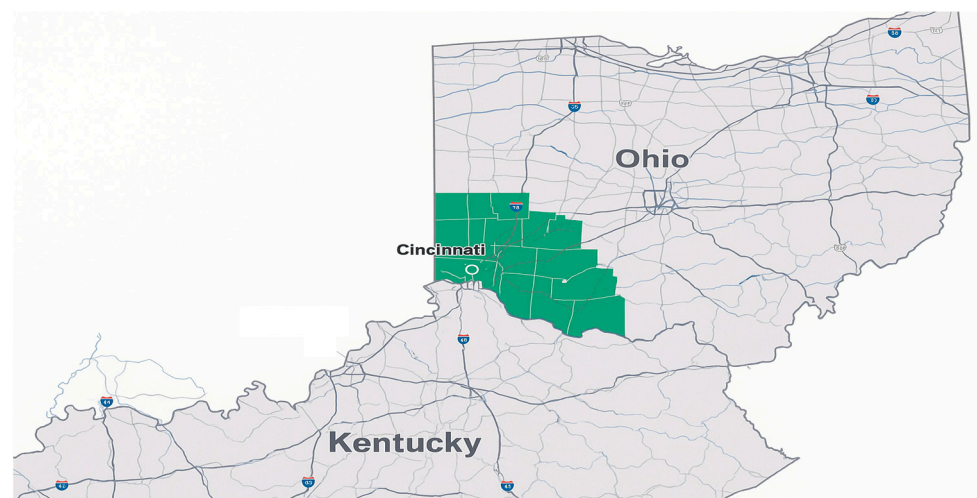


Figure 3. Duke Energy Ohio and Kentucky (DEO&K) Co serving area—courtesy of PJM [42].

In order to access these data directly, we used the “Data Miner 2” platform. Data Miner 2 provides us with a comprehensive feed of operational summaries from previous periods

and a wealth of information to analyze electrical energy consumption. Figures 4 and 5 show the collected data for load consumption from 1 April 2022 to 1 April 2025.

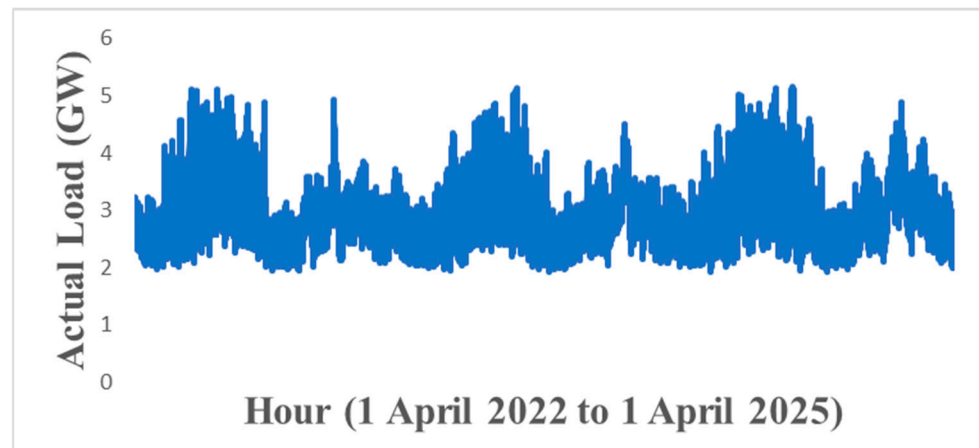


Figure 4. Consumption data from 1 April 2022 to 1 April 2025.

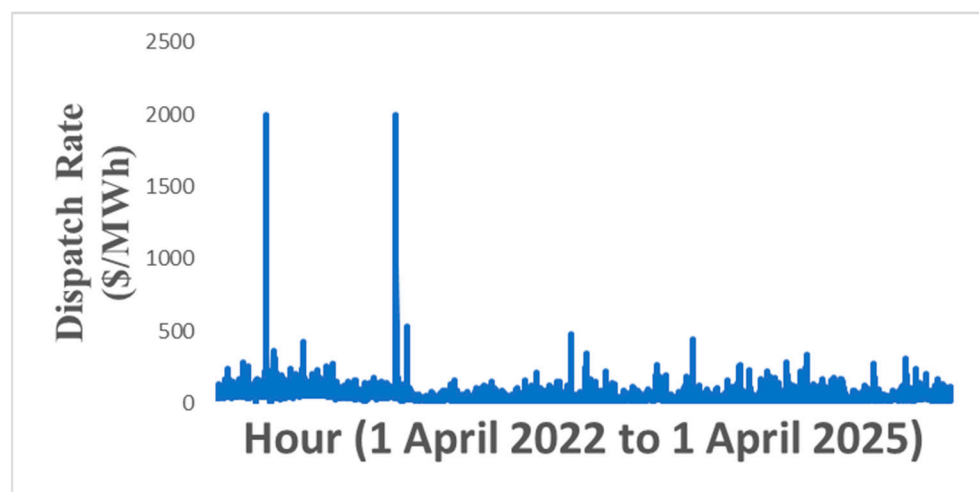


Figure 5. Dispatch rate data from 1 April 2022 to 1 April 2025.

Duke Energy Ohio and Kentucky (DEO&K) is a major utility company serving over a million customers across southwestern Ohio and northern Kentucky. As a subsidiary of Duke Energy, it operates and maintains the electric grid, ensuring reliable power delivery through a mix of traditional generation assets and investments in modern infrastructure.

2.3. 2-Feature Selection

Correlation Matrix (CM)

A correlation matrix is a matrix of correlation coefficients between two variables (in this case, load consumption and weather features). The values range from -1 to 1 . Correlation matrices represent the following values:

- $+1$: means that the two variables increase proportionally as one increases.
- -1 : indicates that a perfect negative correlation exists, meaning that as one variable increases, the other variable decreases proportionally.
- A correlation of 0 : there is no linear relationship between the variables. It is still possible; however, that another nonlinear relationship exists.

The best feature can be defined as the smallest values for “actual load—feature”. The results for the calculated correlation matrix can be seen in Table 1.

Table 1. Correlation by CM.

Features	Correlation Value	Features	Correlation Value
Dispatch rate	0.33921	Condition	−0.029156
Wind Speed	0.11236	Pressure	−0.091274
Wind Gust	0.11188	Dew Point	−0.10462
Humidity	0.052206	Temperature	−0.1339
Precipitation	0.0035134	Actual Load	1

As follows from Table 1, the most important predictors are wind speed, wind gusts, and temperature. We also incorporate the energy price (Figure 5, “Dispatch Rate”) and local population growth (<https://news.duke-energy.com/>, access on 30 March 2025). The Cincinnati area grew at an average annual growth rate of 0.56% from 2000 to 2024, which we apply to the DEO and K territory’s forecast (<https://fred.stlouisfed.org/series/CTIPOP>, access on 21 May 2025).

2.4. 3-Trend Correction

Adjusting the “actual load” data for holidays and weekends is crucial for creating accurate energy consumption forecasts because these days often exhibit unique usage patterns distinct from regular weekdays. During holidays and weekends, typical daily routines shift significantly—people may stay at home longer, leading to increased residential energy use, while many businesses close or reduce operations, decreasing commercial energy demand. Ignoring these variations could skew data analysis, leading to inaccurate models that perceive these anomalies as regular consumption trends, thus mistaking significant deviations for ‘noise’ in the data.

To accurately reflect energy consumption patterns, we employ a systematic approach to adjust for these discrepancies. Firstly, we identify public holidays and weekends within the dataset. We then calculate the average energy usage on these days and compare it to the average on typical weekdays to understand the variance in consumption. By calculating a holiday coefficient, a metric that quantifies the difference in energy use, we can adjust daily load values accordingly. In this case, the holiday coefficient is 0.937; this indicates that energy use on holidays and weekends is 93.7% of that on a regular day.

Applying this coefficient systematically across the dataset enables us to adjust the actual load values for holidays and weekends. This adjustment ensures that our energy consumption data more accurately reflects the true patterns, taking into account the lower or shifted energy use on these non-standard days. By doing so, we mitigate the risk of misinterpreting the seasonal or occasional drops in energy usage as errors or outliers, thereby enhancing the overall accuracy and reliability of our energy forecasting models. This nuanced approach is not only methodologically robust but also essential for sectors like utility planning and energy policy making, where predicting precise demand is critical.

It is important to mention that using a holiday factor helps reduce the variation caused by holidays. This adjustment is only applied during model training to lessen the impact of irregular consumption on holidays and weekends. When making predictions, if the target date is a holiday, weekend, or special day, we reverse the adjustment by dividing the forecasted value by the holiday coefficient. This approach helps us have smoother trends during training, while still being able to accurately predict holiday consumption during the prediction phase.

2.5. Implementation of Multiple Machine Learning Methods on Multiple Clouds

To forecast consumption for the next 24 h, we use data from 1 April 2022 to 1 April 2025. The consumption for the next 24 h (31 March) can be seen at Table 2 with different methods by using heart AI [43–45], namely the following machine learning methods: Support Vector Machine (SVM), Seasonal Autoregressive Integrated Moving Average with exogenous regressors (SARIMAX), SARIMAX + SVM, Long Short-Term Memory (LSTM), Linear Regression (LR), and forecasting values by Decision Tree (DT).

Table 2. Results of the first prediction for the next 24 h by different methods in gigawatt.

Points	Actual Data	SVM	SARIMAX	SARIMAX-SVM	LSTM	DT	LR
1	2.7142	2.6015	2.6639	2.6643	2.5890	2.5957	2.6134
2	2.4560	2.3966	2.4690	2.4344	2.5842	2.4083	2.3944
3	2.3348	2.3481	2.4151	2.3484	2.4777	2.3426	2.3371
4	2.2407	2.3904	2.3956	2.2992	2.2989	2.5218	2.3509
5	2.1184	2.3425	2.3332	2.2162	2.1163	2.5564	2.3206
6	2.0390	2.3999	2.2708	2.1503	2.0707	2.2839	2.3288
7	2.0019	2.3812	2.2343	2.1183	2.1638	2.2340	2.3469
8	1.9885	2.2186	2.2119	2.1052	2.2506	2.2557	2.2873
9	2.0104	2.2542	2.1978	2.1095	2.2660	2.1821	2.2571
10	2.1826	2.3480	2.3251	2.2581	2.3503	2.2845	2.3648
11	2.4312	2.5364	2.5075	2.4677	2.4726	2.3892	2.5374
12	2.5925	2.5118	2.6105	2.5944	2.5176	2.5190	2.5796
13	2.6633	2.5209	2.6243	2.6300	2.5002	2.4803	2.5489
14	2.7386	2.5162	2.6607	2.6812	2.5322	2.4634	2.5464
15	2.8005	2.5569	2.7059	2.7328	2.5907	2.6012	2.5706
16	2.8049	2.5523	2.7000	2.7302	2.6029	2.6616	2.5567
17	2.8078	2.6193	2.7314	2.7517	2.6543	2.7387	2.6047
18	2.8121	2.8028	2.7914	2.7927	2.7578	2.9074	2.7348
19	2.8116	3.0136	2.8453	2.8277	2.9006	2.9053	2.9066
20	2.8131	2.8910	2.8782	2.8489	2.9880	2.7959	2.9388
21	2.8432	2.8151	2.9128	2.8805	2.9828	2.7885	2.8989
22	2.8704	2.7953	2.9080	2.8846	2.8775	2.7140	2.8317
23	2.9160	2.7785	2.9365	2.9175	2.8068	2.7983	2.8109
24	2.9439	2.7573	2.8954	2.9234	2.8334	2.7480	2.7514

In order to evaluate the methods, Deviation, MSE (Mean Squared Error), MAE (Mean Absolute Error), MAPE (Mean Absolute Percentage Error), and RMSE (Root Mean Squared Error) are calculated using Equation (1) through (4), where A stands for actual and F for forecasted loads. The results are shown in Figure 6.

$$\text{MSE : } \text{MSE} = \frac{1}{N} \sum_{i=1}^N (A_i - F_i)^2 \quad (1)$$

$$\text{MAE : } \text{MAE} = \frac{1}{N} \sum_{i=1}^N |F_i - A_i| \quad (2)$$

$$\text{MAPE : } \text{MAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{A_i - F_i}{A_i} \right| \times 100\% \quad (3)$$

$$\text{RMSE : } \text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (F_i - A_i)^2} \quad (4)$$

Thus, SARIMAX and SARIMAX + SVM yield the top two results. Third, we select LSTM forecasts. The first cloud represents SARIMAX, the second cloud represents SARIMAX + SVM, and the third cloud represents LSTM. All results from these methods should always be close to one another. One of the results should not be considered until it is aligned with the other two if it deviates significantly. This way, we are always able to forecast accurately.

In our analysis, we implemented a threshold-based approach to evaluate the closeness of predictions from LSTM, SARIMAX, and SARIMAX + SVM models, establishing a 5% threshold relative to actual load values to ensure robustness in our forecasting accuracy. By

calculating the absolute differences between each pair of forecasted results and comparing these to the threshold, we determined whether the deviations fell within acceptable limits. This method enabled us to apply a systematic and quantifiable measure to assess the prediction alignment, requiring at least two out of the three models to agree within this predefined margin. This approach ensures that our forecasting system remains reliable by incorporating a consensus mechanism, which enhances the validity of the predictions by mitigating the impact of outlier forecasts and refining the overall prediction accuracy.

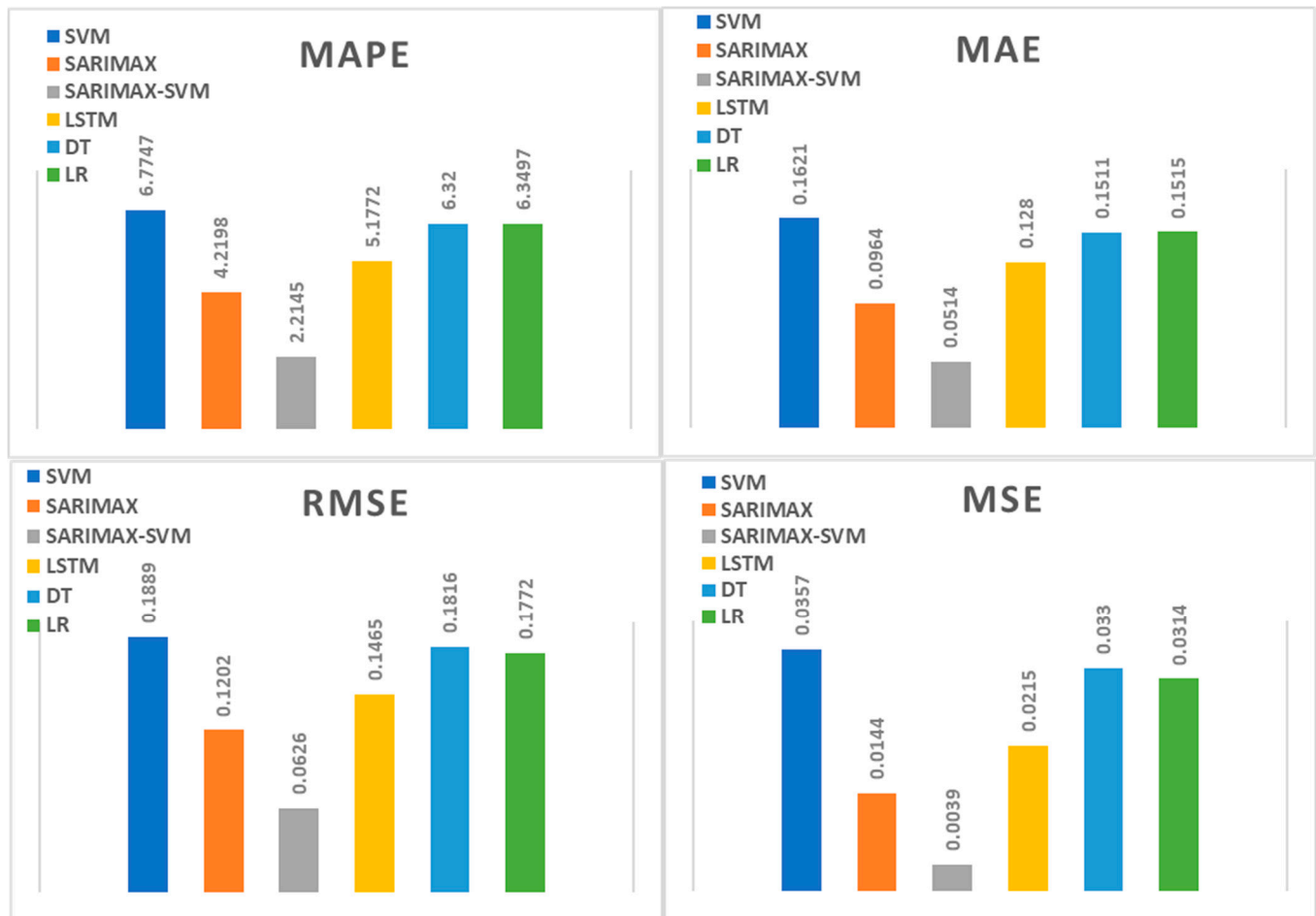


Figure 6. MAE, MSE, MAPE, and RMSE errors produced by the methods under analysis.

2.6. Deviation Correction (DC)

In this step, we calculate how much each outcome deviates from the actual data. Then, by understanding the pattern of these differences, we can accurately predict the future deviation. We now have a deviation for the first 24 h and can predict the next 5 h of deviation, which means the first 5 h of 1 April 2025 (Figure 7).

Now, we have the predicted values for the next 5 h as well as the forecasted deviation for the last 5 h (Table 3).

Table 3. Predicted consumption and deviation for the last 5 h—values given in gigawatt (GW).

Predicted Consumption			Predicted Deviation		
LSTM	SARIMAX	SARIMAX + SVM	LSTM	SARIMAX	SARIMAX + SVM
2.9623	2.9698	2.9696	0.0118	0.01	0.0114
2.868	2.852	2.8763	0.0123	0.0415	0.0184

Table 3. Cont.

Predicted Consumption			Predicted Deviation		
LSTM	SARIMAX	SARIMAX + SVM	LSTM	SARIMAX	SARIMAX + SVM
2.6871	2.6976	2.6977	0.0141	0.024	0.0274
2.5187	2.5452	2.5309	0.0150	0.0152	0.0361
2.4111	2.4233	2.4239	0.0154	0.0334	0.0389

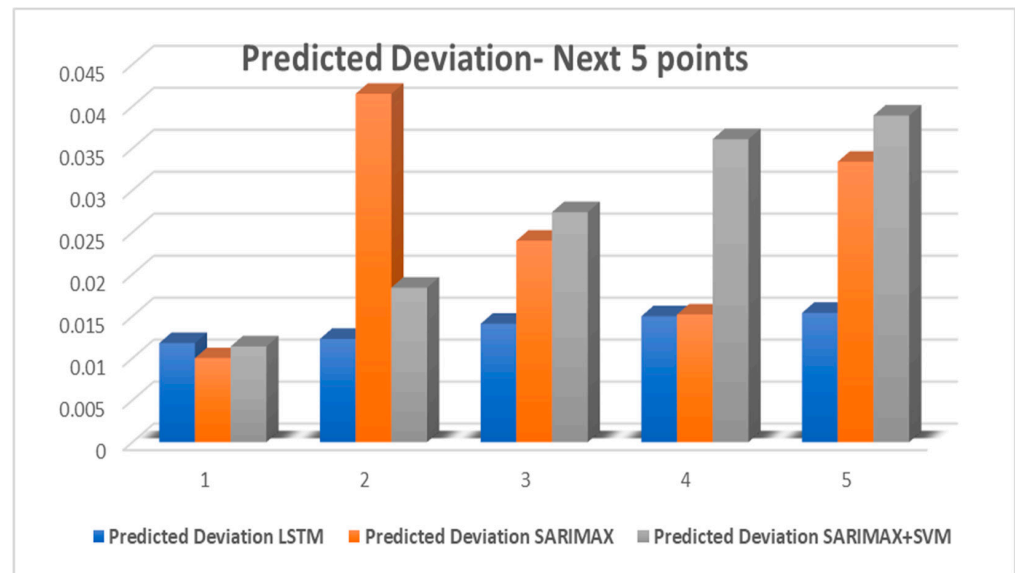


Figure 7. Forecasted value deviations for the next 5 h using selected methods.

So, after collecting historical data for load consumption, feature selection, and trend correction, the next steps are explained in parameters of the proposed model in clearer way.

Description 1: Parameters of the Proposed Model

So, after collecting historical data for load consumption, feature selection, and trend correction, the next steps are explained in the parameters of the proposed model more clearly.

Input:

- Dataset D from 1 April 2022, to 30 March 2025;
- Forecast horizon: next 24 h (31 March 2025).

Output:

- Hourly load forecast for 31 March 2025, with refined predictions for the first 5 h of 1 April 2025.

Data Partitioning

The training set consists of the first 80% of D.

The test set consists of the remaining 20% of D.

24 h Forecasting

- 2.1. Train the forecasting models on the training set.
- 2.2. Generate the forecast for the next 29 h: 24 h for 31 March 2025 (A), and the first 5 h of 1 April 2025 (B).

Deviation Analysis and Correction

- 3.1. Analyze the deviation trend within the test set.
- 3.2. Forecast the deviation for the first 5 h of 1 April 2025 (C).
- 3.3. For the first 5 h, compute the corrected forecast:

$$\text{Final Forecast} = B + C$$

Fault Tolerance via 2-out-of-3 Voting on Three Cloud Platforms

- 4.1. Deploy three different models on three cloud platforms.
- 4.2. Determine the final prediction by majority voting (averaging the consensus) if at least two out of the three platforms produce consistent predictions.

For instance, if all three systems work, reliability follows R^3 .

If exactly two systems function correctly while one fails, the reliability contribution is proportional to $3R^2(1-R)$.

The system fails only if two or more components (cloud or forecasting method) fail, which is not considered in the formula.

- R represents reliability.

3. Results

Table 3 presents the predicted consumption values alongside the forecasted deviations. To derive the final forecast for the last five targeted hours, we sum these two metrics. The outcomes of this computation are documented in Table 4, representing the finalized forecasted values.

In our analysis, Figure 6 illustrates the errors in the forecast before applying DC. Figure 8 shows the final prediction.

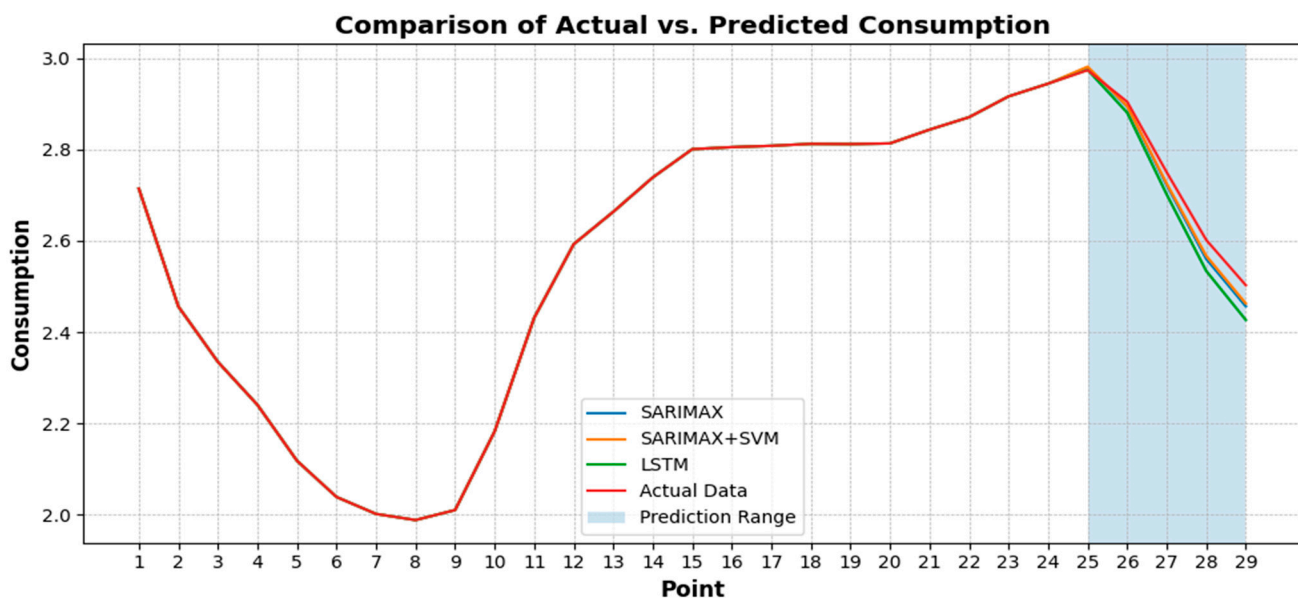


Figure 8. Final prediction for the next 5 h.

Table 4. Predicted consumption and deviation for the last 5 h—value given in gigawatt (GW).

Actual Real Data	Predicted Consumption by Proposed Methods			Predicted Consumption Without DC					
	LSTM-DC	SARIMAX-DC	SARIMAX-SVM-DC	LSTM	SARIMAX	SARIMAX SVM	SVM	LR	DT
2.9734	2.9741	2.9798	2.981	2.9623	2.9698	2.9696	2.9108	2.9201	2.9267
2.9042	2.8803	2.8935	2.8947	2.868	2.852	2.8763	2.8002	2.8065	2.7971

Table 4. Cont.

Actual Real Data	Predicted Consumption by Proposed Methods			Predicted Consumption Without DC					
	LSTM-DC	SARIMAX-DC	SARIMAX-SVM-DC	LSTM	SARIMAX	SARIMAX SVM	SVM	LR	DT
2.7497	2.7012	2.7216	2.7251	2.6871	2.6976	2.6977	2.5907	2.5951	2.6522
2.6016	2.5337	2.5604	2.567	2.5187	2.5452	2.5309	2.4679	2.4688	2.5314
2.5026	2.4265	2.4567	2.4628	2.4111	2.4233	2.4239	2.3556	2.3538	2.4221

4. Discussion

The errors depicted in Figure 6 represent discrepancies likely caused by assumptions, external factors not accounted for in the initial model, or inherent data variability and other factors. By comparing these results with those in Table 5, Figures 8 and 9, which showcase the errors after applying deviation correction and final prediction, we observe a marked improvement in forecast accuracy.

Table 5. MAE, MAPE, RMSE, and MAPE produced by the methods under analysis.

Metric	CM-LSTM-DC	CM-SARIMAX-DC	CM-SARIMAX-SVM-DC
MAE	0.0434	0.0264	0.0232
MSE	0.0027	0.0009	0.0007
RMSE	0.0516	0.0308	0.0266
MAPE	1.65%	1.00%	0.88%

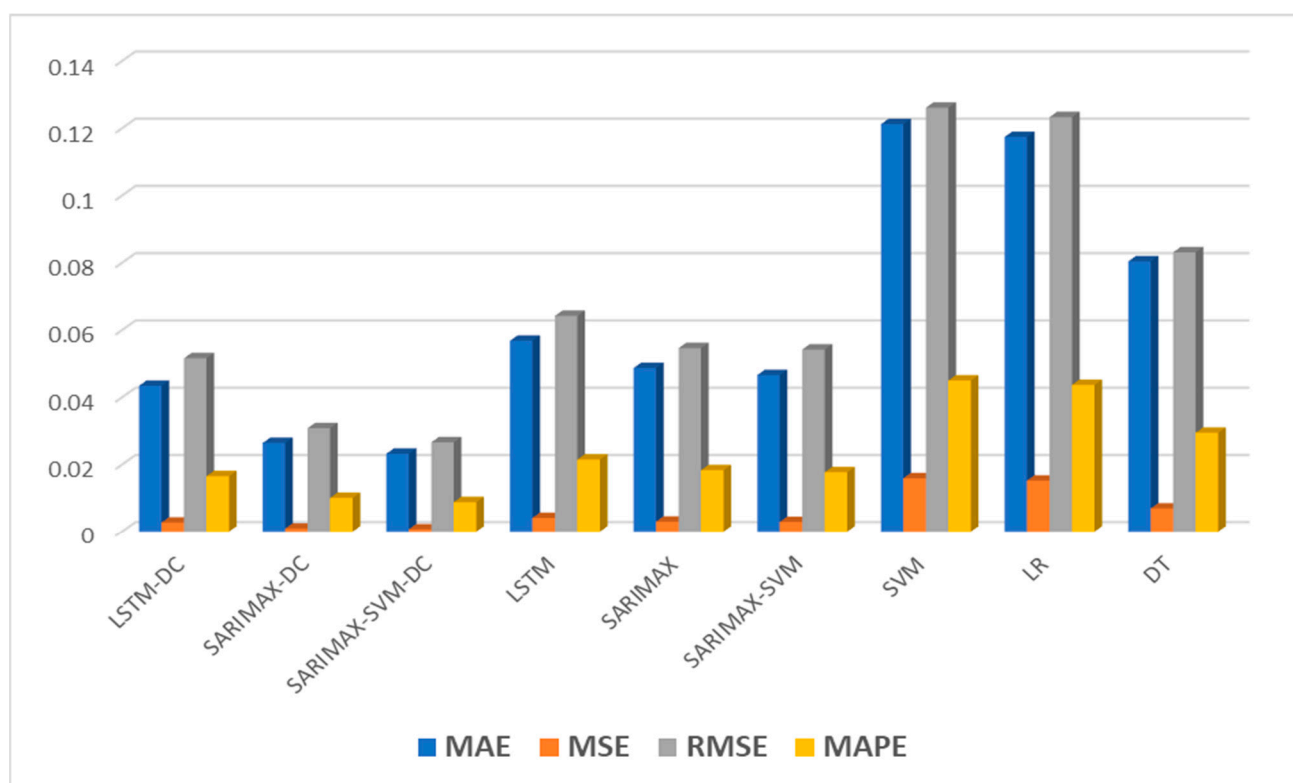


Figure 9. MAE, MSE, MAPE, and RMSE errors produced by the methods under analysis after/before DC.

By implementing the proposed technique, we have seen a marked improvement in the accuracy of our forecasting models across all metrics. As detailed in Table 4, the initial forecasts from models such as LSTM, SARIMAX, and SARIMAX + SVM show

certain deviations from actual consumption values. However, after applying the proposed method, there is a significant reduction in both MAE and MAPE, indicating a closer alignment with the actual data, as shown in Table 5. The RMSE and MSE metrics also show a substantial improvement, confirming the effectiveness of the proposed method in refining our predictions. This enhancement is further evidenced by the visual analysis in Figure 6, where the post-proposed system predictions align more closely with the true consumption patterns.

By comparing the forecasted values with recorded data in Table 5, the SARIMAX model shows a notable increase in forecast accuracy, as evidenced by the progression of predicted values compared to the relatively stable initial results. This suggests that the proposed method enhances the model's ability to better capture and adjust to underlying trends, possibly by improving the handling of seasonal patterns or anomalies previously not as accurately modeled, with a clear shift indicating a reduction in systematic underestimations. For the SARIMAX + SVM model, the results indicate an improved model response to data variance, starting with higher predictions and tapering to values that more closely match actual measurements. Meanwhile, even the already highly accurate LSTM model shows subtle but important enhancements; slight increases in forecast values reflect the fine-tuning provided by the proposed method, which, although minor, significantly boosts the model's precision and alignment with real consumption data. These nuanced adjustments, while numerically slight, signify meaningful improvements in the models' abilities to predict energy consumption with reduced error margins, demonstrating the effectiveness of the proposed technique in refining forecast accuracy and providing a more reliable, adaptable tool for energy management and planning.

Overall, the proposed system not only enhances the accuracy but also bolsters the reliability of forecasting by integrating advanced analytical techniques that effectively adapt to varying data patterns and seasonal fluctuations. This integration enables the system to handle anomalies and outliers with greater precision, ensuring that forecasts remain robust under diverse conditions. Additionally, by leveraging real-time data processing capabilities, the system continuously updates its predictive models to reflect the latest trends, further enhancing the forecast reliability. The incorporation of machine learning algorithms also allows the system to learn from past discrepancies, automatically adjusting its parameters for future predictions, thereby reducing the likelihood of significant forecasting errors and improving the overall decision making in energy management. This holistic approach not only optimizes the accuracy and reliability of forecasts but also enhances the operational efficiency of energy systems, providing a valuable tool for both immediate and long-term planning.

The proposed 2-out-of-3 cloud-based voting system works as follows (R is the reliability of each cloud, such as AWS):

R^3 : All three systems work correctly.

$(3R^2)(1-R)$: Exactly two systems work and one fails (and the two correct results "outvote" the incorrect one).

The system only fails if two or more components fail, which is excluded in this formula. The structure could be seen in Figure 10.

The proposed method aims to predict the grid demand. Consumer behavior in using renewable energy is affected by personal motivation, making it difficult to study with numbers alone. Instead, forecasts focus on the results of consumer actions. The study in [46] looks at the future of renewable energy in Poland by reviewing consumer opinions, legal rules, and energy balance.

Another study [47] shows that events like war can change energy management conditions and increase the focus on energy efficiency. War can raise natural gas prices, which in

turn may push up electricity prices in some markets. Unfortunately, unexpected events like war also affect oil prices, the economy, industry, and even physical and mental health. Energy demand, prices, and environmental concerns may accelerate the shift to cleaner and more sustainable sources, such as wind and solar energy.

Another interesting area for future work is energy forecasting at a local level, such as for a specific region, building, or location. For example, the authors in [48] describe using machine learning in smart buildings to boost the energy efficiency by analyzing energy use, occupancy patterns, and environmental conditions. They tested prediction models using algorithms like LSTM, RF, and the gradient boosting regressor, and applied these models to real-life case studies in educational buildings. The results show that customizing predictive models to fit each building's unique energy consumption characteristics is crucial.

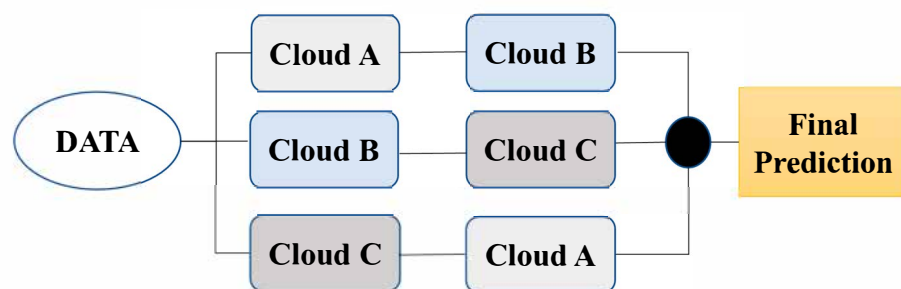


Figure 10. The proposed 2-out-of-3 cloud based voting system for energy prediction.

5. Conclusions

In this paper, we introduced a novel parallel method for short-term energy forecasting in electrical networks, leveraging the DEO&K Co data from the PJM website. The system was rigorously applied to real hourly energy consumption data spanning from 1 April 2022 to 1 April 2025. We provided forecasts for the five subsequent intervals of the case study and demonstrated increased forecasting accuracy through the integration of deviation correction, trend correction, and redundant machine learning across three diverse clouds. This approach not only enhanced reliability through a ‘2-out-of-3’ cloud-based system, but also significantly improved the precision of our machine learning models. Specifically, the LSTM model’s Mean Absolute Percentage Error (MAPE) decreased from 5.17% to 1.65%, reflecting an absolute improvement of 3.52% (a 68.09% relative reduction). The SARIMAX model showed even greater gains, with MAPE dropping from 4.21% to 1.00%—an absolute improvement of 3.21% (76.25% relative reduction). Similarly, the SARIMAX + SVM model achieved a MAPE reduction from 2.21% to 0.88%, marking an absolute improvement of 1.33% (60.18% relative reduction). These results underscore the potential of integrating cloud-based infrastructure into real-time grid management, offering substantial advancements over traditional forecasting methods. Future work will aim to further refine these techniques and expand their application to additional datasets and forecasting scenarios.

Author Contributions: Conceptualization, K.H.B. and G.C.; methodology, K.H.B.; software, K.H.B. and Z.F.; validation, K.H.B. and Z.F.; formal analysis, K.H.B. and Z.F.; investigation, K.H.B. and Z.F.; resources, E.E.R.; data curation, K.H.B. and Z.F.; writing—original draft preparation, K.H.B.; writing—review and editing, E.E.R.; visualization, K.H.B. and E.E.R.; supervision, E.E.R.; project administration, E.E.R.; funding acquisition, E.E.R. and G.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: <https://www.pjm.com/>; access on 21 May 2025.

Conflicts of Interest: The authors declare no conflict of interest.

Nomenclature

Metric	Definition
MSE	Mean square error
MAE	Mean absolute error
MAPE	Mean absolute percentage error
RMSE	Root mean square error
AE	Average error
DC	Deviation correction
ML	Machine learning
SARIMAX	Seasonal autoregressive integrated moving average with exogenous regressors
DT	Decision tree
LSTM	Long short-term memory
SVM	Support vector machine
LR	Linear regression
ANN	Artificial neural networks
DEO&K Co	Duke Energy Ohio and Kentucky
TC	Trend correction
AWS	Amazon Web Services
CM	Correlation matrix
PJM	PJM is a regional transmission organization (RTO) that coordinates the movement of wholesale electricity in all or parts of 13 states and the District of Columbia

References

1. Chou, J.-S.; Tran, D.-S. Forecasting Energy Consumption Time Series Using Machine Learning Techniques Based on Usage Patterns of Residential Householders. *Energy* **2018**, *165*, 709–726. [\[CrossRef\]](#)
2. Baesmat, K.H.; Shiri, A. A New Combined Method for Future Energy Forecasting in Electrical Networks. *Int. Trans. Electr. Energy Syst.* **2019**, *29*, e2749. [\[CrossRef\]](#)
3. Baesmat, K.H.; Masoudipour, I.; Samet, H. Improving the Performance of Short-Term Load Forecast Using a Hybrid Artificial Neural Network and Artificial Bee Colony Algorithm. *IEEE Can. J. Electr. Comput. Eng.* **2021**, *44*, 275–282. [\[CrossRef\]](#)
4. Singh, U.; Rizwan, M.; Alaraj, M.; Alsaidan, I. A Machine Learning-Based Gradient Boosting Regression Approach for Wind Power Production Forecasting: A Step towards Smart Grid Environments. *Energies* **2021**, *14*, 5196. [\[CrossRef\]](#)
5. Haq, E.U.; Lyu, X.; Jia, Y.; Hua, M.; Ahmad, F. Forecasting Household Electric Appliances Consumption and Peak Demand Based on Hybrid Machine Learning Approach. *Energy Rep.* **2020**, *6* (Suppl. 9), 1099–1105. [\[CrossRef\]](#)
6. Baesmat, K.H.; Shokoochi, F.; Farrokhi, Z. SP-RF-ARIMA: A sparse random forest and ARIMA hybrid model for electric load forecasting. *Glob. Energy Interconnect.* in press. **2025**. [\[CrossRef\]](#)
7. Ahmad, T.; Chen, H.; Zhang, D.; Zhang, H. Smart Energy Forecasting Strategy with Four Machine Learning Models for Climate-Sensitive and Non-Climate-Sensitive Conditions. *Energy* **2020**, *198*, 117283. [\[CrossRef\]](#)
8. Khan, P.W.; Byun, Y.-C. Genetic Algorithm Based Optimized Feature Engineering and Hybrid Machine Learning for Effective Energy Consumption Prediction. *IEEE Access* **2020**, *8*, 196274–196286. [\[CrossRef\]](#)
9. Torabi, M.; Hashemi, S.; Saybani, M.; Shamshirband, S.; Mosavi, A. A Hybrid Clustering and Classification Technique for Forecasting Short-Term Energy Consumption. *Environ. Prog. Sustain. Energy* **2018**, *38*, 66–76. [\[CrossRef\]](#)
10. Deb, C.; Zhang, F.; Yang, J.; Lee, S.; Shah, K. A Review on Time Series Forecasting Techniques for Building Energy Consumption. *Renew. Sustain. Energy Rev.* **2017**, *74*, 902–924. [\[CrossRef\]](#)
11. Chou, J.-S.; Ngo, N.-T. Time Series Analytics Using Sliding Window Metaheuristic Optimization-Based Machine Learning System for Identifying Building Energy Consumption Patterns. *Appl. Energy* **2016**, *177*, 751–770. [\[CrossRef\]](#)
12. Dong, B.; Cao, C.; Lee, S.E. Applying Support Vector Machines to Predict Building Energy Consumption in Tropical Region. *Energy Build.* **2005**, *37*, 545–553. [\[CrossRef\]](#)
13. Mawson, V.J.; Hughes, B.R. Deep Learning Techniques for Energy Forecasting and Condition Monitoring in the Manufacturing Sector. *Energy Build.* **2020**, *217*, 109966. [\[CrossRef\]](#)
14. Persson, C.; Bacher, P.; Shiga, T.; Madsen, H. Multi-Site Solar Power Forecasting Using Gradient Boosted Regression Trees. *Sol. Energy* **2017**, *150*, 423–436. [\[CrossRef\]](#)

15. Bogner, K.; Pappenberger, F.; Zappa, M. Machine Learning Techniques for Predicting the Energy Consumption/Production and Its Uncertainties Driven by Meteorological Observations and Forecasts. *Sustainability* **2019**, *11*, 3328. [\[CrossRef\]](#)
16. Hong, T.; Wang, P.; White, L. Weather station selection for electric load forecasting. *Int. J. Forecast.* **2015**, *31*, 286–295. [\[CrossRef\]](#)
17. Jain, R.; Smith, K.; Culligan, P.; Taylor, J. Forecasting Energy Consumption of Multi-Family Residential Buildings Using Support Vector Regression: Investigating the Impact of Temporal and Spatial Monitoring Granularity on Performance Accuracy. *Appl. Energy* **2014**, *123*, 168–178. [\[CrossRef\]](#)
18. Ma, J.; Oppong, A.; Acheampong, K.N.; Abruquah, L.A. Forecasting Renewable Energy Consumption under Zero Assumptions. *Sustainability* **2018**, *10*, 576. [\[CrossRef\]](#)
19. Abdel-Aal, R.E.; Al-Garni, A.Z.; Al-Nassar, Y.N. Modelling and Forecasting Monthly Electric Energy Consumption in Eastern Saudi Arabia Using Abductive Networks. *Energy* **1997**, *22*, 911–921. [\[CrossRef\]](#)
20. Ahmad, T.; Chen, H.; Huang, R.; Guo, Y.; Wang, J.; Shair, J.; Akram, H.M.A.; Mohsan, S.A.H.; Kazim, M. Supervised Based Machine Learning Models for Short-, Medium- and Long-Term Energy Prediction in Distinct Building Environments. *Energy* **2018**, *158*, 17–32. [\[CrossRef\]](#)
21. Shin, S.; Woo, H. Energy Consumption Forecasting in Korea Using Machine Learning Algorithms. *Energies* **2022**, *15*, 4880. [\[CrossRef\]](#)
22. Bourdeau, M.; Zhai, X.-Q.; Nefzaoui, E.; Guo, X.; Chatellier, P. Modeling and Forecasting Building Energy Consumption: A Review of Data-Driven Techniques. *Sustain. Cities Soc.* **2019**, *48*, 101533. [\[CrossRef\]](#)
23. Lin, X.; Yu, H.; Wang, M.; Li, C.; Wang, Z.; Tang, Y. Electricity Consumption Forecast of High-Rise Office Buildings Based on the Long Short-Term Memory Method. *Energies* **2021**, *14*, 4785. [\[CrossRef\]](#)
24. Demolli, H.; Dokuz, A.S.; Ecemis, A.; Gokcek, M. Wind Power Forecasting Based on Daily Wind Speed Data Using Machine Learning Algorithms. *Energy Convers. Manag.* **2019**, *198*, 111823. [\[CrossRef\]](#)
25. Herrera, G.P.; Constantino, M.; Miranda Tabak, B.; Pistori, H.; Su, J.-J. Long-Term Forecast of Energy Commodities Price Using Machine Learning. *Energy* **2019**, *179*, 214–221. [\[CrossRef\]](#)
26. Wei, N.; Li, C.; Peng, X.; Zeng, F.; Lu, X. Conventional Models and Artificial Intelligence-Based Models for Energy Consumption Forecasting: A Review. *J. Pet. Sci. Eng.* **2019**, *181*, 106187. [\[CrossRef\]](#)
27. Beyca, O.; Ervural, B.; Tatoğlu, E.; Ozuyar, P.; Zaim, S. Using Machine Learning Tools for Forecasting Natural Gas Consumption in the Province of Istanbul. *Energy Econ.* **2019**, *20*, 937–949. [\[CrossRef\]](#)
28. Baesmat, K.H. Impedance Analysis of Adaptive Distance Relays Using Machine Learning. In *ITNG 2024: 21st International Conference on Information Technology-New Generations*; Latifi, S., Ed.; Advances in Intelligent Systems and Computing; Springer: Cham, Switzerland, 2024; Volume 1456. [\[CrossRef\]](#)
29. Khan, P.W.; Kim, Y.; Byun, Y.-C.; Lee, S.-J. Influencing Factors Evaluation of Machine Learning-Based Energy Consumption Prediction. *Energies* **2021**, *14*, 7167. [\[CrossRef\]](#)
30. Divina, F.; García Torres, M.; Gómez Vela, F.A.; Vázquez Noguera, J.L. A Comparative Study of Time Series Forecasting Methods for Short Term Electric Energy Consumption Prediction in Smart Buildings. *Energies* **2019**, *12*, 1934. [\[CrossRef\]](#)
31. Nizami, S.J.; Al-Garni, A.Z. Forecasting Electric Energy Consumption Using Neural Networks. *Energy Policy* **1995**, *23*, 1097–1104. [\[CrossRef\]](#)
32. Khan, P.W.; Byun, Y.-C.; Lee, S.-J.; Park, N. Machine Learning Based Hybrid System for Imputation and Efficient Energy Demand Forecasting. *Energies* **2020**, *13*, 2681. [\[CrossRef\]](#)
33. Hafeez, G.; Alimgeer, K.; Khan, I. Electric Load Forecasting Based on Deep Learning and Optimized by Heuristic Algorithm in Smart Grid. *Appl. Energy* **2020**, *269*, 114915. [\[CrossRef\]](#)
34. Ahmad, T.; Chen, H. A Review on Machine Learning Forecasting Growth Trends and Their Real-Time Applications in Different Energy Systems. *Sustain. Cities Soc.* **2020**, *54*, 102010. [\[CrossRef\]](#)
35. Hou, W.; Wen, H.; Zhang, N.; Lei, W.; Lin, H. Dynamic Load Combined Prediction Framework with Collaborative Cloud-Edge for Microgrid. In *Proceedings of the IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, New York, NY, USA, 2–5 May 2022; pp. 1–2. [\[CrossRef\]](#)
36. Zhang, X.; Zeng, Z.; Wang, P.; Song, J.; Kong, Z. A Hybrid Edge-Cloud Computing Method for Short-Term Electric Load Forecasting Based on Smart Metering Terminal. In *Proceedings of the 2020 IEEE 4th Conference on Energy Internet and Energy System Integration (EI2 2020)*, Wuhan, China, 30 October–1 November 2020; pp. 3101–3105. [\[CrossRef\]](#)
37. Raju, M.; Laxmi, A. IOT Based Online Load Forecasting Using Machine Learning Algorithms. *Procedia Comput. Sci.* **2020**, *171*, 551–560. [\[CrossRef\]](#)
38. Geberslassie, M.; Bitzer, B. Cloud Computing for Renewable Power Systems. *Renew. Energy Power Qual.* **2012**, *10*, 1250–1254. [\[CrossRef\]](#)
39. Hou, S.; Li, H.; Yang, C.; Wang, L. A New Privacy-Preserving Framework Based on Edge-Fog-Cloud Continuum for Load Forecasting. In *Proceedings of the 2020 IEEE Wireless Communications and Networking Conference (WCNC)*, Seoul, Republic of Korea, 25–28 May 2020; pp. 1–8. [\[CrossRef\]](#)

40. Pelluri, S.; Srinivas, P.; Sneha, S. A Compendium on Load Forecasting Approaches and Models. *Int. J. Res. Eng. Technol.* **2014**, *3*, 27–36. [[CrossRef](#)]
41. Devaraj, J.; Madurai Elavarasan, R.; Shafiullah, G.; Jamal, T.; Khan, I. A Holistic Review on Energy Forecasting Using Big Data and Deep Learning Models. *Int. J. Energy Res.* **2021**, *45*, 13489–13530. [[CrossRef](#)]
42. PJM Website. Available online: <https://www.pjm.com/> (accessed on 21 May 2025).
43. Baesmat, K.H.; Latifi, S. A New Hybrid Method for Electrical Load Forecasting Based on Deviation Correction and MRMRMS. In Proceedings of the 30th International Conference on Systems Engineering, ICSEng 2023, Las Vegas, NV, USA, 22–24 August 2023; Volume 761. [[CrossRef](#)]
44. Zeinab, F.; Kalhor, A.; Tale Masouleh, M. Implementation and Evaluation of Object Identification Techniques on Nao Robot Platform. *Int. J. Mechatron. Electr. Comput. Technol.* **2018**, *8*, 3947–3958. Available online: https://aeuso.org/includes/files/articles/Vol8_Iss29_3947-3958_Implementation_and_Evaluation_of_Ob.pdf (accessed on 21 May 2025).
45. Farrokhi, Z.; Baesmat, K.; Regentova, E. Enhancing Urban Intelligence Energy Management: Innovative Load Forecasting Techniques for Electrical Networks. *J. Power Energy Eng.* **2024**, *12*, 72–88. [[CrossRef](#)]
46. Zielińska-Sitkiewicz, M.; Chrzanowska, M.; Furmańczyk, K.; Paczutkowski, K. Analysis of Electricity Consumption in Poland Using Prediction Models and Neural Networks. *Energies* **2021**, *14*, 6619. [[CrossRef](#)]
47. Gajdzik, B.; Wolniak, R.; Nagaj, R.; Žuromskaitė-Nagaj, B.; Grebski, W.W. The Influence of the Global Energy Crisis on Energy Efficiency: A Comprehensive Analysis. *Energies* **2024**, *17*, 947. [[CrossRef](#)]
48. Elhaby, K.; Baina, A.; Bellafkih, M.; Deifalla, A.F. Machine Learning Algorithms for Predicting Energy Consumption in Educational Buildings. *Int. J. Energy Res.* **2024**, *1*, 6812425. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.