Electricity Consumption Forecasting Using Time Series Analysis

Ms. Mahi Mistry[1], Mr. Ashish Modi[2]
[1]Student, Department of MSc. IT, Nagindas Khandwala College, Mumbai, Maharashtra, India
mistrymahi3@gmail.com

[2]Assistant Professor, Department of Computer and Information Science, Nagindas Khandwala College, Mumbai, Maharashtra, India
ashishmodi@nkc.ac.in

ABSTRACT: Forecasting electricity consumption is essential for effective energy management and long-term sustainability. In this study, household electricity demand is forecasted by leveraging historical consumption patterns along with external factors such as temperature, weekends, and holidays. Multiple forecasting models, including ARIMA, SARIMA, and LSTM, were evaluated across both short-term and long-term horizons to explore their strengths and limitations in different scenarios. The dataset used for this research was sourced from the UCI Machine Learning Repository and contains detailed household power measurements such as active power, voltage, and sub-metering for specific appliances. After extensive preprocessing and resampling to ensure data consistency, the models were trained and their performance was compared using evaluation metrics like MAE, RMSE, and MAPE. The results indicate that statistical models such as ARIMA and SARIMA are well suited for short-term forecasts, while deep learning models like LSTM provide better accuracy for long-term predictions. Incorporating external factors through the SARIMAX model further enhanced forecasting accuracy, particularly in capturing spikes in consumption during extreme weather or holiday periods. The findings emphasize the importance of selecting the appropriate forecasting model based on the prediction horizon and integrating contextual data to improve reliability. Additionally, this research highlights the value of hybrid modeling approaches and paves the way for smarter, data-driven strategies that can support efficient energy use and grid stability in increasingly complex consumption environments.

KEYWORDS: Electricity Forecasting, Time Series Analysis, ARIMA, SARIMA, LSTM, SARIMAX, Smart Grid, Energy Demand, Household Consumption, Exogenous Variables, Forecast Accuracy.

I. INTRODUCTION

Electricity consumption patterns are increasingly complex due to factors like urbanization, technological advancements, and changing lifestyles. Accurate forecasting of household energy demand is crucial for efficient grid management, cost reduction, and the integration of renewable energy sources. While traditional statistical models like ARIMA and SARIMA have been effective for short-term predictions, they often fall short in capturing long-term dependencies and non-linear patterns. Recent advancements in deep learning, particularly Long Short-Term Memory (LSTM) networks, offer promising solutions by modeling intricate temporal relationships in data. However, the performance of these models can be significantly influenced by external variables such as temperature, weekends, and holidays, which affect household energy usage.

This study utilizes the "Individual Household Electric Power Consumption" dataset from the UCI Machine Learning Repository, comprising over 2 million measurements collected at a one-minute sampling rate over four years. The dataset includes various features like global active power, voltage, and sub-metering values for specific appliances. By integrating this rich dataset with advanced forecasting models, we aim to enhance the accuracy of electricity consumption predictions. The primary objectives of this research are to evaluate the performance of ARIMA, SARIMA, and LSTM models across different forecasting horizons and to assess the impact of incorporating external variables on forecasting accuracy. Through this approach, we seek to contribute valuable insights for energy providers, policymakers, and consumers in managing electricity demand more effectively.

II. RESEARCH OBJECTIVE

- Evaluate the differences in model performance for short-term vs. long-term electricity consumption forecasting.

- Analyze the impact of external variables (e.g., temperature, weekends, holidays) on electricity usage trends.

III. LITERATURE REVIEW

Zhao & Wang (2017) Promotion strategy is directly tied to competitiveness when using deep learning for sales forecasting. Their convolutional neural network (CNN) model showed that integrating data on price changes, promotional activities, and consumer preferences leads to more accurate sales predictions.

Alizamir et. al (2022) Personal and psychological factors strongly influence consumer buying behaviour when analysed through hybrid statistical and machine learning methods. Their large-scale empirical study of online shopping behaviour revealed that price sensitivity, demographic traits, and perceptions of delivery time significantly shape consumer decisions.

Schultz & Schultz et. al (2020) Personal and psychological factors are associated with the study of the impact of sales promotion on consumer buying behaviour because psychological perception determines the behavioural economics of the consumers regarding their purchasing behaviour about products of the apparel industry.

Darmawan et al., 2018; Gorji & Siami, (2020) Promotions and price discounts are benchmarks for a company's success in attracting consumer interest. This condition is because consumers are interested in buying the product. Therefore, the right marketing strategy is one way to achieve company goals, namely by knowing the needs and desires of consumers to create the correct product so that it can achieve the goal of increasing sales and winning the competition.

Carreón et. al (2019) Social and demographic factors are crucial in shaping consumer buying behaviour because their study using machine learning models showed that demographics have a stronger influence on purchase decisions than simple advertisement exposure. This highlights the importance of social identity and segmentation in the apparel industry.

IV. METHODOLOGY

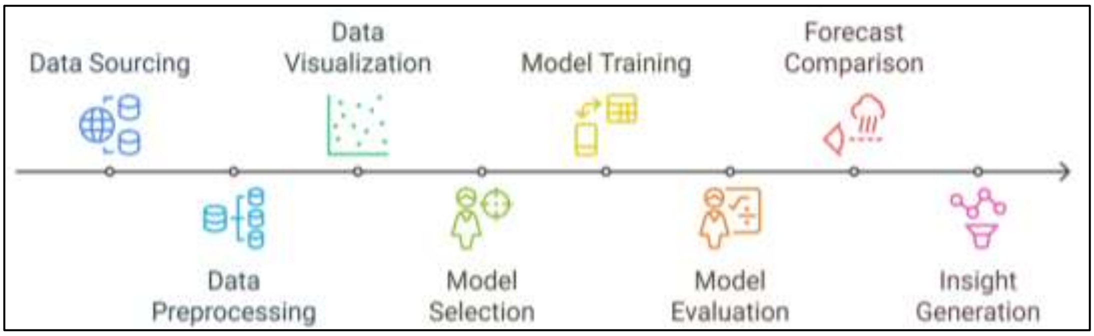

Figure 1: Methodology

The dataset used in this study was sourced from the UCI Machine Learning Repository and contains detailed household electricity consumption records collected over time. Data preprocessing involved handling missing values, converting timestamps into a standard datetime format, and resampling at both daily and hourly levels to better explore usage patterns and variability. Visualizations such as line plots and heatmaps were used to identify trends, seasonality, and anomalies, helping prepare the data for more accurate modeling. These steps ensured that the data was clean, structured, and suitable for in-depth analysis and forecasting. For forecasting, several models were applied, including ARIMA, SARIMA, SARIMAX, and LSTM, each chosen to address specific challenges. ARIMA and SARIMA focused on capturing trends and seasonal patterns, while SARIMAX extended this by including external variables like temperature, weekends, and holidays to enhance model accuracy. LSTM, a deep learning model, was used to capture more complex and long-term dependencies in the sequential data. The models were trained using historical data and evaluated using metrics such as MAE, RMSE, and MAPE to measure prediction accuracy. Comparisons were made across different forecast horizons, such as one day, one week, and one month, to assess the performance under various scenarios. The insights from these models provide a strong basis for improving energy management and planning, helping stakeholders make informed decisions. The study also highlighted how advanced analytics and data-driven approaches can lead to smarter forecasting and better resource allocation.

V. RESULT

The visual comparison of the actual electricity consumption against the forecasts from ARIMA, SARIMA, and LSTM models clearly shows how each model performs over time. The actual usage line fluctuates significantly, with sharp peaks and drops that reflect real-life variations in household consumption. The ARIMA forecast appears flat and struggles to track these changes, while the SARIMA model follows the general seasonal trends but misses abrupt increases. The LSTM forecast, on the other hand, closely mirrors the ups and downs of the actual data, responding better to sudden changes and providing a smoother yet more accurate representation of demand over longer periods. In terms of results, the error metrics further emphasize the differences between the models. The ARIMA model produces the highest errors, indicating poor short-term accuracy and an inability to adapt to sudden changes. SARIMA performs better, thanks to its consideration of seasonality, but still falls short in capturing sharp fluctuations. The LSTM model stands out with much lower errors, suggesting its ability to understand complex patterns and dependencies in the data. These findings highlight that while statistical models may work for simple, short-term forecasts, deep learning methods like LSTM are far superior for longer, more complicated prediction tasks.

Figure 2: Electricity Forecast

Model	Forecast Horizon	MAE	RMSE	MAPE (%)
ARIMA	Short-term (1-7 days)	1272.07	1337.35	74.19
SARIMA	Short-term (1-7 days)	917.76	999.98	52.06
LSTM	Short-term (1 months)	489.58	598.31	27.95

Table 1: Model Performance Summary

The SARIMAX model, which incorporates external factors such as temperature, weekends, and holidays, proved to be an effective forecasting approach for predicting household electricity consumption over a 7-day horizon. With a mean absolute error of 365.53 and a root mean square error of 0.1592, the model showed a significant improvement in accuracy compared to models that only rely on historical data. The forecast closely followed the general trend of actual consumption, while the confidence interval accounted for potential variability due to changes in weather and user behavior. These results highlight how important contextual information is for demand forecasting, as it helps anticipate fluctuations more accurately and enhances the reliability of short-term predictions.

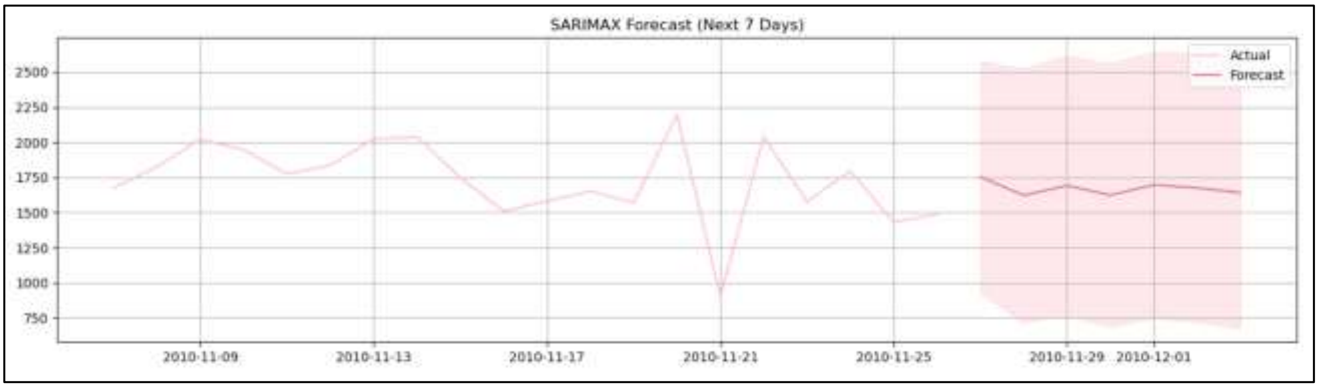

Figure 3: SARIMAX Forecast

The visualization further illustrates the benefits of including external variables, as the forecast line aligns well with actual consumption despite some uncertainty. The shaded area represents potential variability caused by unforeseen factors such as extreme weather or changes in usage patterns, but overall, the model captures the consumption trend effectively. The performance outcomes confirm that adding exogenous variables strengthens the forecasting process by reducing errors, especially when predicting spikes and drops during weekends or holidays. At the same time, the results indicate that electricity demand remains complex and influenced by many unpredictable factors, emphasizing the need for richer datasets and hybrid modeling techniques to further improve forecasting accuracy and support efficient energy management.

VI. CONCLUSION

Accurate forecasting of electricity consumption is critical for ensuring efficient energy management and supporting sustainable development. This study demonstrated that combining historical consumption data with external factors such as temperature, weekends, and holidays can significantly enhance forecasting performance. By applying multiple models, including ARIMA, SARIMA, LSTM, and SARIMAX, the research provided a comparative understanding of how different techniques perform across short-term and long-term forecasting scenarios. The results clearly showed that while statistical models like ARIMA and SARIMA are effective for capturing short-term trends and fluctuations, deep learning models such as LSTM are better equipped to handle complex, long-term patterns. Furthermore, the inclusion of external variables through SARIMAX improved forecast accuracy.

Overall, this study emphasizes that no single model is universally best; instead, the choice of forecasting method should be aligned with the specific requirements and timeframes of energy planning. The findings support the development of hybrid approaches that combine the interpretability of traditional models with the predictive power of deep learning. These insights can help utilities, policymakers, and households alike make informed decisions and adopt smarter energy management practices in today's data-driven environment.

REFERENCE

1. Jain PK, Quamer W, Pamula R. **Electricity Consumption Forecasting Using Time Series Analysis**. In: *ICACDS 2018* (Dehradun, India, April 2018). Singapore: Springer; 2018. p. 327–335.
2. Karthika D, Karthikeyan K. **Forecasting Electricity Consumption in India by Using Univariate Time-Series Analysis**. *Advances and Applications in Mathematical Sciences*. 2022;22(1):71–78.
3. Bandyopadhyay A, Sarkar BD, Hossain ME, Rej S, Mallick MA. **Modelling and Forecasting India's Electricity Consumption Using Artificial Neural Networks**. 2024; 00:1–13.
4. Niharika, Kaur H, Singh J. **Electricity Consumption Forecasting System of India: A Review**. *Advances in Mathematics: Scientific Journal*. 2022;11(3):249–273.
5. Research on the Forecasting of Electricity Demand in India State—**Electricity Demand Prediction for Maharashtra**. (2024). In: *Highlights in Business, Economics and Management*, EMFRM; 45:780–792.
6. Reddy SK, Gopakumar G. **PM-Gati Shakti: Advancing India's Energy Future through Demand Forecasting – A Case Study**. arXiv Preprint. 2023 Jul 30.

7. Barbar M, Mallapragada D, Alsup M, Stoner R. **Scenarios of Future Indian Electricity Demand Accounting for Space Cooling and Electric Vehicle Adoption**. arXiv Preprint. 2021 Jun 14.
8. Gupta S, Chitkara P. **Day Ahead Price Forecasting Models in Thin Electricity Market (India)**. arXiv Preprint. 2020 Nov 3.
9. Roy K, Ishmam A, Taher KA. **Demand Forecasting in Smart Grid Using Long Short-Term Memory**. arXiv Preprint. 2021 Jul 28.
10. Neslin SA. A market response model for coupon promotions. Marketing science. 1990(2):125–145.