Advanced Time Series Forecasting Models for Electricity Consumption Prediction in Abu Dhabi

Jaber Jemai
CIS division
Higher Colleges of Technoogy
Abu Dhabi, UAE
jjemai@hct.ac.ae

Alia Al Hammadi
CIS division
Higher Colleges of Technoogy
Abu Dhabi, UAE
H00492493@hct.ac.ae

Alya Al Zaabi
CIS division
Higher Colleges of Technoogy
Abu Dhabi, UAE
H00503082@hct.ac.ae

Rawdha Al Zaabi
CIS division
Higher Colleges of Technoogy
Abu Dhabi, UAE
H00492127@hct.ac.a

https://doi.org/10.63962/CRZW4432

Abstract - Faced with the constant rise in energy demand and the urgent need to combat climate change, many governments have been driven to wisely consume available energy and explore new resources through innovative and sustainable solutions. Accurate forecasting of electricity demand is critical for operational planning and grid optimization. In this study, we aim to design and develop a comprehensive set of time series forecasting models to predict power consumption for TAQA, the main energy distributor in Abu Dhabi. Using historical load data, we will construct and compare four families of models: baseline econometric models (ARMA and ARIMA), a gradient-boosted tree ensemble (XGBoost), and a deep learning architecture (LSTM). In addition to forecasting performance, the models will undergo rigorous validation using a suite of time-series–specific diagnostic tests, including assessments of heteroscedasticity, autocorrelation, stationarity, and residual distribution properties. These diagnostics ensure the stability and statistical reliability of the models developed. The forecast accuracy of all models will be evaluated across multiple horizons using standard error metrics.

Keywords—Electricity Consumption Forecasting, Time Series Analysis, LSTM, XGBoost.

I. INTRODUCTION

Abu Dhabi is a fast-growing city that faces increasing energy demand driven by rapid urbanization, rising population, and industrial growth. This leads to higher costs of energy production and supply. The technical report of the Department of Energy (DoE) shows that the electricity demand in Abu Dhabi is continuously increasing to reach 13 GigaWatts (GW) in 2023, with a 9.1% increase from 2022 to 2023. TAQA Distribution, as the primary electricity and water distribution company in Abu Dhabi, plays a significant role in addressing these challenges. Such figures reveal a need for reliable and robust solutions to support TAQA in optimizing its electricity distribution grid.

Predicting electricity demand is one of the primary challenges faced by energy supply companies. If future demands are accurately approximated, then energy production factories and distribution grids will operate efficiently, peak demands will be known ahead of time, electricity shortages will be avoided, and customer satisfaction will be improved. To achieve this, TAQA aims to develop and deploy energy consumption forecasting solutions utilizing advanced econometric and AI-enabled models to support Abu Dhabi and the United Arab Emirates' vision for sustainability and its Net Zero 2050 goals. In this paper, we will use historical energy consumption data from TAQA to create short and long-term forecast. The dataset will include historical electricity usage of Abu Dhabi, Al Ain, and Al Dhafra regions. The goal of this paper is to test and compare advanced econometric and AI-enabled models to make the most accurate predictions. Similar studies have been conducted in various countries, including the United States of America [1], China [2], Brazil [3], and Mexico [4].

This project adds significant value to TAQA and other stakeholders by providing a practical solution for short and long-term energy demand forecasting. Through the use of machine learning techniques such as XGBoost and LSTM, the project aims to identify the most accurate predictive model. By analyzing historical energy consumption data alongside external factors like temperature and time patterns, the solution helps reduce uncertainty in demand management. For TAQA, this translates into better resource allocation, cost reduction, and more efficient energy planning. Ultimately, the project resolves the pressing problem of inefficiencies and rising costs in energy management by offering a scalable, intelligent solution.

II. RESEARCH METHODOLOGY

This study follows a structured methodological framework (Fig. 1) consisting of data acquisition, preprocessing, model development, diagnostic validation, and performance evaluation. The methodology ensures systematic comparison between all models to identify the most reliable forecasting approach for TAQA's electricity consumption.



Fig. 1. Research Methodology

III. LITERATURE REVIEW

The application of Artificial Intelligence (AI) and Machine Learning (ML) techniques have significantly transformed electricity consumption forecasting, addressing the limitations of traditional statistical methods that struggle to capture nonlinear and dynamic energy patterns [1]. Early

models, such as ARIMA and regression analysis, provided foundational approaches, but proved inadequate when dealing with complex temporal dependencies and exogenous influences [3]. Consequently, AI-based models have emerged as more powerful alternatives capable of enhancing prediction accuracy. Machine learning approaches, including Logistic Regression (LR), Random Forest (RF), Support Vector Machine (SVM), and ensemble algorithms such as XGBoost and CatBoost have been effectively applied in diverse contexts. For instance, studies in the USA used Energy Information Administration (EIA) and Department of Energy (DOE) [5] datasets demonstrated that ML-based forecasting outperformed traditional methods in modeling nonlinear consumption behaviors, with LR and RF achieving notable accuracy gains. Similarly, CatBoost combined with hybrid optimization algorithms such as Particle Swarm Optimization (PSO) and Ant Lion Optimization (ALO) proved effective in long-term national-scale forecasting for sustainable energy planning in Mexico [6].

Deep learning techniques, particularly LSTM networks and their variants, have become dominant in short-term and medium-term load forecasting due to their ability to model sequential dependencies [7]. Comparative analyses show that Bidirectional LSTM (Bi-LSTM) models achieve superior performance relative to conventional LSTM, ARIMA, and SARIMA methods, particularly for nonlinear and seasonal datasets [8]. Recent advancements, such as the Improved LSTM (ILSTM) [9] integrated with IoT-based data acquisition systems, achieved outstanding results in residential energy forecasting within the United Arab Emirates, with MAE and RMSE of 0.185 and 0.276, respectively.

Hybrid AI architecture has further improved forecasting accuracy by integrating deep learning with optimization algorithms. The Optimized Deep Neural Network (ODNN) enhanced through a hybrid Genetic Algorithm–Particle Swarm Optimization (GA-PSO) achieved higher accuracy and lower error than conventional deep networks [10]. Similarly, Least Squares Support Vector Machine (LSSVM) tuned via the Gooseneck Barnacle Optimizer (GBO) [11] reached exceptional forecasting accuracy, outperforming several other bio-inspired optimizers such as PSO and Dragonfly Algorithm. Combining machine learning and fuzzy logic have demonstrated strong potential in managing uncertainty within microgrids. A study integrating BiLSTM forecasting with Mamdani fuzzy logic [12] achieved 93.2% classification accuracy for microgrid stability assessment. Hybrid optimization methods such as Echo State Networks (ESN) optimized by Differential Evolution (DE) [4] achieved very low forecasting errors using datasets from China and Taiwan. Comprehensive surveys of AI applications in energy systems report that hybrid models integrating deep learning with swarm intelligence or evolutionary optimization achieve the highest accuracy and reliability [1,13]. As reported in [2,14], these models not only enhance prediction accuracy but also support intelligent control and optimization in energy management, as seen in applications of AI to hybrid renewable systems and building-level forecasting.

IV. TAQA DATASET AND DATA PREPROCESSING

Time series data plays a central role in energy forecasting, as it captures changes in consumption over time with strong temporal dependencies. In the context of electrical grids and regional demand analysis, TAQA will provide a dataset reporting electricity consumption in the Emirate of Abu Dhabi. The dataset spans energy usage between 2019 and 2025 on hourly basis. Geographically, data will be covering Abu Dhabi City, Al Ain, and Al Dhafra regions at sub-station granularity level. Additionally, data on humidity, temperature, and precipitation will be collected from the National Center of Meteorology in the UAE.

Effective preprocessing is essential because real-world energy data often contains frequent irregularities, such as missing values, noise, and peaks, caused by unpredictable events. To clean TAQA's dataset, we expect to handle gaps in the time series when timestamps are missing. To reconstruct the timeline at uniform intervals, typical missing values imputation techniques will be used. It is also important to separate and process legitimate and erroneous extreme spikes. Furthermore, feature engineering of raw timestamps will be undertaken to extract time-based information such as hour-of-day, day-of-week, or season. These temporal attributes allow models to recognize recurring behaviors such as lag features, rolling windows, moving averages, and seasonal growth. This makes an early understanding of the critical characteristics of the time series. Properly addressing these and similar issues will transform raw observations into informative time-aware variables capable of revealing trends, recurring patterns, and long-term dependencies.

V. FORECASTING MODELS SELECTION

The selected models include classical econometric approaches, along with advanced machine learning and deep learning models. Together, these models provide a comprehensive forecasting framework combining traditional statistical principles with modern artificial intelligence methods.

A. Baseline forecasting models

Primarily, time-series forecasting relied on econometric models such as Autoregressive (AR), Moving Average (MA), Autoregressive Moving Average (ARMA), and Autoregressive Integrated Moving Average (ARIMA). These models are grounded in statistical theory and remain widely used due to their interpretability and effectiveness in modeling linear temporal patterns. In the context of energy forecasting, econometric models help establish a baseline understanding of consumption trends by capturing dependencies within historical data. They are especially useful when consumption follows stable, predictable cycles. The ARMA model combines two components: an autoregressive term that expresses current values as a function of previous values, and a moving average term that

incorporates past forecast errors. If energy consumption data exhibits non-stationarity due to seasonal effects, temperature variation, and long-term trends, then ARIMA models can be used.

B. XGBoost models

XGBoost is an advanced ensemble model based on gradient boosting principles by building trees sequentially, with each tree correcting the errors of the previous one. It incorporates strong regularization terms to prevent overfitting and includes features such as parallel computation, missing-value handling, and optimized memory usage. This makes XGBoost efficient, scalable, and highly accurate for forecasting tasks involving tabular datasets with many engineered features. XGBoost performs well in structured data environments because it can model high-level feature interactions, such as relationships between temperature, calendar effects, and regional consumption trends.

C. XGBoost models

Long short-term memory networks are a type of recurrent neural network designed specifically to model sequential and time-dependent data. LSTMs can automatically learn temporal dependencies by processing sequences such as the past 24 or 48 hours of consumption data. This makes LSTMs highly effective for modeling cyclic patterns and long-range temporal behavior. LSTMs excel in environments where energy consumption is influenced by strong daily and weekly cycles.

VI. CONCLUSION

This paper presented a data-driven methodological framework for forecasting TAQA's hourly electricity demand using representative statistical, machine learning, and deep learning models. Through a focused literature review and systematic model selection, ARMA, ARIMA, XGBoost, and LSTM were identified as suitable approaches for capturing both linear and nonlinear load dynamics. The proposed methodology emphasizes rigorous preprocessing and diagnostic validation, ensuring that each model meets key time-series reliability criteria. Overall, the framework provides a solid foundation for developing accurate and scalable forecasting solutions in energy analytics. Future work will apply this methodology to full empirical evaluation to TAQA dataset.

REFERENCES

[1.] R. Klyuev, I. Morgoev, A. Morgoeva, O. Gavrina, N. Martyushev, E. Efremenkov, and Q. Mengxu. Methods of forecasting electric energy consumption: A literature review. Energies, 15(23):8919, 2022.

[2] G. Hafeez, Kh. S. Alimgeer, Z. Wadud, Z. Shafiq, M. Ali Khan, I. Khan, F. Khan, and A. Derhab. A novel accurate and fast converging deep learning-based model for electrical energy consumption forecasting in a smart grid. Energies, 13(9):2244, 2020.

[3] J. de Assis Cabral, L. Legey, and M. de Freitas Cabral. Electricity consumption forecasting in brazil: A spatial econometrics approach. Energy, 126:124–131, 2017. \

[4] L. Wang, H. Hu, Xue-Yi Ai, and H. Liu. Effective electricity energy consumption forecasting using echo state network improved by differential evolution algorithm. Energy, 153:801–815, 2018.

[5] S. Hossain, M. Hasanuzzaman, M. Hossain, M. Amjad, M. Shovon, M. Hossain, and M. Rahman. Forecasting energy consumption trends with machine learning models for improved accuracy and resource management in the usa. Journal of Business and Management Studies, 7(1):200–217, 2025.

[6] H. Li and A. Kayae. Predicting energy consumption in mexico: Integrating environmental, economic, and energy data with machine learning techniques for sustainable development. Energy, 324:135992, 2025.

[7] M. W. Hasan. Design of an iot model for forecasting energy consumption of residential buildings based on improved long short-term memory (lstm). Measurement: Energy, 5:100033, 2025.

[8] H. Alizadegan, B. Rashidi Malki, A. Radmehr, H. Karimi, and M. Ilani. Comparative study of long short-term memory (lstm), bidirectional lstm, and traditional machine learning approaches for energy consumption prediction. Energy Exploration & Exploitation, 43(1):281–301, 2025.

[9] E. Uzun, M. Karabinaoˇglu, and M. Ayaz. Lstm-based estimation of energy consumption in energy-intensive facilities. Engineering Research Express, 7(1):015420, 2025.

[10] E. Hosseini, B. Saeedpour, M. Banaei, and R. Ebrahimy. Optimized deep neural network architectures for energy consumption and pv production forecasting. Energy Strategy Reviews, 59:101704, 2025.

[11] M. Ahmed, M. Sulaiman, M. Hassan, M. Rahaman, and M. Amin. Daily allocation of energy consumption forecasting of a power distribution company using optimized least squares support vector machine. Results in Control and Optimization, 18:100518, 2025.

[12] Y. Kholiavka and Y. Parfenenko. Intelligent energy consumption forecasting and microgrid state assessmen