Nigerian Journal of Technology (NIJOTECH)
Vol. 43, No. 4, December, 2024, pp.743 - 753
www.nijotech.com

Print ISSN: 0331-8443
Electronic ISSN: 2467-8821
https://doi.org/10.4314/njt.v43i4.14

MEDIUM TERM LOAD FORECASTING OF 33 kV LINE LOADING: A CASE STUDY OF OTA 132/33 kV SUB-STATION OGUN STATE, NIGERIA

AUTHORS:
O. O. Olusanya[1], A. O. Onokwai[2], and R. O. Oyebisi[3],

AFFILIATIONS:
[1]Department of Computer Engineering, Bells University of Technology, Ota, Ogun State. Nigeria
[2]Department of Mechanical Engineering, Pan Atlantic University, Lagos. Nigeria
[3]Department of Electrical/Electronics and Telecommunication Engineering, Bells University of Technology, Ota, Ogun State. Nigeria.

***CORRESPONDING AUTHOR:**
Email: ooolusanya@bellsuniversity.edu.ng

ARTICLE HISTORY:
Received: 27 April, 2024.
Revised: 17 September, 2024.
Accepted: 20 September, 2024.
Published: 31 December, 2024.

KEYWORDS:
ARIMA: Auto Regressive Integrated Moving Average, MAE: Mean Absolute Error, MAPE: Mean Absolute Percentage Error, RMSE: Root Mean Square Error, and R-squared.

ARTICLE INCLUDES:
Peer review

DATA AVAILABILITY:
On request from author(s)

EDITORS:
Chidozie Charles Nnaji

FUNDING:
None

Abstract

Peak load forecasting plays a pivotal role in the efficient operation and planning of power systems, influencing decision-making processes for resource allocation and infrastructure development. The Ota 132/33 kV substation in Nigeria is facing increasing demand due to rapid industrialization and urbanization. This has strained the substation's infrastructure, leading to issues like transformer overloading, voltage fluctuations, and power outages. The area's trade center status and numerous industries further stress transformers, increasing wear and tear and potentially jeopardizing their reliability. This study delves into the realm of 33 kV feeders at the OTA Transmission Substation, aiming to unravel the intricacies of peak load patterns and provide a forecast for the next five years. Leveraging historical data spanning from 2018 to 2022, sourced from OTA Transmission Substation was used to forecast from 2023 to 2027. The research employs the Auto-Regressive Integrated Moving Average (ARIMA) model to discern trends and project future peak loads. Performance metrics, including Mean Absolute Error, Mean Absolute Percentage Error, Root Mean Square Error, and R-squared, are meticulously evaluated to assess the robustness of the forecasting model. The findings shed light on the unique characteristics of each feeder, with Sumo, Amje, and Idiroko having a better predictive accuracy performance with minimal errors, while Sango, FSM, and Estate show a moderate level of predictive accuracy probably due to the presence of little nuance in their data set. Where the Sango 33 kV feeder displayed an upward trend of 18.91MW in 2023 and 19.34 MW in 2027. Sumo 33 kV feeder exhibits a decline trend from 3.21 MW (2023) to 2.23 MW in 2027. FSM 33 kV feeder shows a fluctuation pattern while Amje 33 kV feeder indicates a highly stable trend of 22.29 MW all over. Idiroko 33 kV feeder shows a steadily increasing trend of 16.09 MW (2024) to 16.25 MW in 2027. The Estate 33 kV feeder on the other hand depicts a relatively stable pattern. This study not only contributes to the localized understanding of peak load dynamics but also serves as a template for similar investigations in other power distribution networks and unveils other alternative data science-based models for future researchers.

HOW TO CITE:
O. O. Olusanya[1], A. O. Onokwai[2], R. O. Oyebisi[3] "Medium Term Load Forecasting of 33 kV Line Loading: A Case Study of Ota 132/33 kV Sub-Station Ogun State, Nigeria", *Nigerian Journal of Technology*, 2024; 43(4), pp. 743 – 753; *https://doi.org/10.4314/njt.v43i4.14*

1.0 INTRODUCTION

Electricity is a fundamental driver of economic growth and social development in modern societies. Reliable and sustainable power supply is essential for meeting the energy demands of industries, households, and other critical sectors of the economy. As such, accurate forecasting of future electricity load is imperative for efficient resource allocation, infrastructure planning, and policy formulation in the power sector [1], [25]. The practice of load forecasting has evolved significantly over the years, driven by advancements in data analytics, technological innovation, and the growing complexity of energy systems.

In the context of Nigeria's electricity industry, which has grappled with issues of chronic underinvestment, infrastructure deficits, and an increasing population [2], accurate load forecasting has taken on unprecedented importance [3]. Historical efforts at forecasting in the Nigerian electricity sector have often faced difficulties stemming from inconsistent data quality [4], policy uncertainties [5], and volatile socioeconomic conditions. Contemporary issues, such as inadequate generation capacity and *distribution* challenges, further compound the complexities of accurate load prediction [2]. With the advent of artificial intelligence, actionable insights could be inferred from Nigeria's electricity big data for future policy direction. The emergence of data science has brought about a paradigm shift in energy data analysis. With the advent of big data, machine learning, and artificial intelligence techniques, the energy sector now possesses the tools to extract valuable insights from vast and diverse datasets. Data science offers the capability to uncover hidden patterns [6], forecast demand with greater accuracy [7], optimize resource allocation [8], and enhance decision-making processes within the energy domain [9].

In the Nigerian context, where data availability and quality have often been stumbling blocks, leveraging data science techniques becomes crucial for informed policymaking and operational efficiency in the electricity industry. Among the various methods available for time series forecasting, the Autoregressive Integrated Moving Average (ARIMA) model stands out as particularly relevant [10 and is recommended for this study. ARIMA is well-suited to capture the temporal dependencies and seasonality inherent in load data [11]. Its ability to accommodate non-stationary data through differencing, coupled with its simplicity and interpretability, makes it an attractive choice.

Furthermore, ARIMA has a proven track record in load forecasting across diverse settings, including electricity markets with dynamics similar to those of the Nigerian context. Its adaptability to noisy data and its capacity to provide both short-term and long-term forecasts align closely with the multifaceted needs of the Nigerian electricity industry. ARIMA, a powerful time series forecasting method, and data visualization play pivotal roles in deciphering load growth trends, identifying seasonal variations, pinpointing yearly trends, and uncovering factors that influence load growth within the context of an electricity substation or grid [12].

This study aims to contribute significantly to the understanding of load forecasting in the Nigerian electricity sector. By employing the ARIMA model in conjunction with contemporary data science techniques, it seeks to improve the accuracy and reliability of load forecasts. The study's findings are expected to inform policy decisions, aid in infrastructure planning, and facilitate efficient resource allocation in Nigeria's electricity industry. Moreover, it underscores the potential of data science to revolutionize energy data analysis, offering a valuable template for similar studies in other emerging economies facing energy supply challenges. The Ota 132/33 kV substation is a crucial component of Nigeria's power infrastructure, which plays a pivotal role in ensuring a reliable supply of electricity to both industrial and residential areas within its service territory. The primary concern lies in the growing demand for electrical power in this region, driven by the rapid industrialization and urbanization of Ota and its neighboring areas. Over the years, the substation has faced an escalating load demand, and this surge in electricity consumption has strained its existing infrastructure [13].

The situation is exacerbated by the fact that this area serves as a trade center for Ogun State and hosts numerous industries, making the need for a reliable power supply even more critical. This places enormous stress on the transformers, increasing their wear and tear and potentially jeopardizing their reliability [13]. It is therefore imperative to have a reliable load forecast system that relies on historical data to predict the short or long-term future electricity needs of the station. Several studies have attempted the forecast of electricity load consumption in existing literature, employing diverse methodological approaches that have richly contributed to knowledge, offering experimental results that avail critical stakeholders with informed decisions for future development in the electricity supply and management industry. [14], examines how temperature affects the short-term forecasting of energy demand in a Nepalese urban region, highlighting the vital role that precise load forecasting has in power system quality and reliability. An extensive analysis of power demand was carried out, taking into account factors like temperature, load fluctuation during the week and on weekends, and the impact of load lags on demand. The study looks at the relationship between load demand, temperature, and demand from the previous day throughout a full year of data collection.

In terms of methodology, the study uses feed-forward neural networks (FF-ANNs), a type of artificial neural network (ANN), in addition to traditional time series models to forecast short-term load. The results of the study reveal that FF-ANNs outperform the conventional time series model in short-term load forecasting, as indicated by the mean absolute percentage error (MAPE). Specifically, FF-ANNs demonstrate a 0.34% improvement on weekdays and an impressive 8.04% enhancement on weekends compared to the conventional model.

[15] came up with a hybrid forecasting scheme for electricity demand, using time series methodology. The methodology leverages adaptive Fourier decomposition as its initial step to extract fluctuation characteristics from the data. The empirical findings of this study demonstrate the remarkable efficacy of the proposed hybrid forecasting scheme. This effectiveness can be attributed to the impact of data pretreatment and optimization through the use of the sine cosine optimization algorithm. The study's results conclusively indicate that the hybrid modeling approach yields promising prediction results while maintaining an acceptable level of computational complexity. [16] proposes a load-forecasting model using deep learning techniques, specifically long short-term memory (LSTM) networks, to improve electric load forecasting accuracy. The model accounts for the periodicity of electric load data by incorporating multiple input time lags a feature not typically addressed by standard LSTM models. The study develops an autoregressive model and autocorrelation function to improve performance in LSTM models. It explores LSTM and GRU variations and compares them against various data mining techniques. The models capture complex time series data characteristics, resulting in more accurate predictions.

The work of [17], examines the effectiveness of the Auto-Regressive Integrated Moving Average (ARIMA) method in forecasting seasonal time series data, particularly for small-scale agricultural load. It suggests further research to improve the method's accuracy and reliability in this context. [18] asserted that COVID-19 pandemic has significantly impacted electricity demand and load forecasting, with a study analyzing data from five years until November 2020. The study introduces a rolling stochastic Regressive Integrated Moving Average with Exogenous (ARIMAX) model to mitigate the pandemic's impact on forecasting models, demonstrating superior performance compared to benchmark models and reducing forecast error by up to 23.7%. [10] focused on the robustness of auto-regressive integrated moving average (ARIMA) models in electrical load forecasting. It uses a simulation-based approach to simulate noise levels and re-identify the model. The results show a weak response to random disturbances and a specific noise threshold that significantly deteriorates the model's forecasting ability. The study emphasizes the importance of data processing in data mining and learning processes.

[19], investigated short-term electricity demand forecasting in deregulated markets using linear regression-based models, spline function-based models, and traditional time series models. It focuses on estimating the yearly cycle within the deterministic component. The research uses data from the Nordic electricity market from 2013 to 2016. The study finds the proposed component-wise estimation method effective, with vector autoregressive modeling and spline function-based regression showing superior performance. A new method for electric load forecasting that uses support vector regression to automatically select lags was suggested by [20]. This method uses gradient descent optimization to fine-tune the widths of an anisotropic Gaussian kernel. The research evaluated its effectiveness on four electricity demand forecasting datasets. The method showed superior predictive accuracy and the ability to identify relevant lags and seasonal patterns. It outperformed existing strategies and state-of-the-art models for automatic model selection, making it a valuable tool for improving electric load forecast accuracy. The research conducted by [21], focused on the use of R programming to analyze and forecast electricity supply and demand in Texas, enabling utility operators to make informed decisions for peak shaving strategies.

[22] presents a forecasting methodology for improving day-ahead electric power load forecasts, utilizing a systematic preprocessing pipeline and gated recurrent units (GRU) model. The methodology uses multivariate time-discrete power data from the ENTSO-E repository and fine-tunes the GRU model to generate precise multi-step forecasts. The methodology outperforms the autoregressive integrated moving average with exogenous variables (ARIMAX) statistical model and the actual day-ahead forecasts generated by the ENTSO-E platform. In [23], the K-means clustering algorithm is used to forecast peak electricity load for university buildings, enhancing energy management and conservation. This hybrid approach improves forecasting performance by gaining insights into electricity consumption patterns. It allows for early peak load prediction, providing

building management with lead time for load reduction strategies, and can be integrated into demand response programs.

2.0 METHODOLOGY

2.1 Description of the Case Study

The Ota 132/33 kV substation is situated in the western part of Ogun State, Nigeria, and is an integral part of the TCN's subsidiary network. It commenced operations in December 2001, initially equipped with a single 40 MVA transformer [13]. This substation is supplied by the Ikeja West 330 kV transmission network. The lone 40 MVA transformer within the station serves two 33 kV outgoing feeders, namely Sango and Idiroko 33 kV Breakers, with their coverage areas presented in Table 1. The Sango 33 kV feeder provides power to the Ota Industrial Estate and the Abule Iroko area. On the other hand, the Idiroko 33 kV breaker supplies electricity to the Idiroko community and the Federated Steel Mill (F.S.M). Ota, being a thriving industrial hub and the central trading point for Ogun State, has experienced rapid and exponential growth [13]. This surge in demand necessitated the urgent installation of an additional 60 MVA transformer, referred to as T2, which became operational in April 2002.

Alakuko and the Abule Egba Axis of Lagos State are currently served by the Amje, the first 60 MVA Transformer feeder, which was previously supplied power by the Ogba 132/33 kV transformer substation in Lagos [24]. Furthermore, the Federal Steel Mill (FSM) in Ota receives the majority of the power from this transformer's second feeder. The Mobitra (Mobile Transformer) 40 MVA was later installed in 2012 (Estate & Idiroko Rd Feeders). T4 was commissioned in 2019, but was having little problems and was taken away for maintenance and later installed in 2020. Hence, could not be captured in this research work.

2.2 Acquisition of Five-Year Monthly Electricity Load Consumption Data for Ota Substation

The Data Acquisition phase of this study involved the collection and compilation of historical electricity load consumption data, which serves as the foundation for the load forecasting analysis. The dataset encompasses a comprehensive set of attributes and columns, each of which is crucial for understanding and predicting load patterns. The primary data is acquired from the OTA transmission station with data types including time-series data, which provides a chronological record of load consumption over time, and numerical data, representing the load consumption values in Megawatts (MW). The dataset comprises columns, with each column corresponding to a specific location or substation within the study area. The columns are labeled as follows: "SUMO," "FSM," "AMJE," "SANGO," "IDIROKO," and "ESTATE," denoting different substation locations or load measurement points.

2.3 Hardware Requirements

The data processing and machine learning implementation of the conceptual model is achieved on a computer laptop with the following configurations:

i. HP Computer laptop
ii. AMD Athlon Silver 3050U 2.30 GHz processor
iii. 4.00 GB RAM
iv. 64-bit Operating System, x64-based processor
v. Windows 10 Pro operating system

Table 1: Distribution of the OTA 132/33 kV coverage Areas

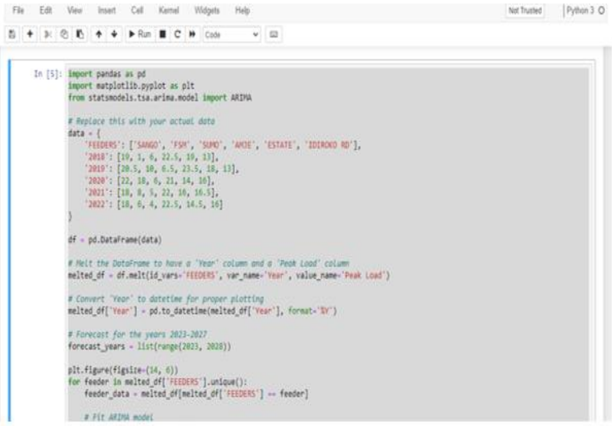
S/N	Transformers	Feeders	Coverage Areas
1	T1 (40 MVA) 132/33 kV	Sango	Sango, Fowobi, Indomie Animasahun, Borehole, Cocacola, Ewupe.
		Tower Alloy	Dedicated
2	T2 (60 MVA) 132/33 kV	FSM	Dedicated
		Amje	Toll Gate, AIT, Jankara Market, Ijaye.
		Sumo	Dedicated
3	MOBITRA (40 MVA) 132/33 kV	Estate	The Bells, Industrial Estate, Iyanayesi, Canan Land.
		Idiroko Rd	Idiroko Rd, Onibuku, Iju and Part of Atan

2.4 Software Requirements

The integrated development environment (IDE) is installed and used to run the Python code for ARIMA implementation, alongside the trend analysis. Other supporting Microsoft Office user packages are used for data cleaning and preprocessing. The following software requirements are used: Anaconda navigator, Jupiter, Microsoft word, and Microsoft Excel.

2.5 Python Implementation

The Python implementation code leverages several essential libraries to conduct the trend analysis and forecast of 33 kV feeders' peak loads for the next five years (2023-2027), based on the historical data of the last five years of 2018 to 2022. The code, as observed from the code snippet in Figure 1 screenshot, relies on the Pandas library for data manipulation, where a *DataFrame* is constructed to organize feeder data across multiple years. *Matplotlib* is utilized for data visualization, enabling the creation of insightful plots. The critical functionality comes from the *Statsmodels* library, particularly the ARIMA model. The ARIMA model instantiated using the ARIMA class, is employed to fit the historical peak load data for each feeder. The *'order'* parameter, set to *(1, 1, 1)*, defines the autoregressive, differencing, and moving average components of the ARIMA model. The code iterates through each feeder, fits an ARIMA model, forecasts peak loads for the next five years, and plots both observed and forecasted data on a single graph. The resulting visualization provides a comprehensive view of historical trends and predicted future peak loads for each feeder, aiding in decision-making for power distribution planning. The *'melt'* function is employed to reshape the *DataFrame*, ensuring compatibility with the ARIMA model and proper plotting. The code seamlessly integrates these libraries, functions, and objects to conduct a robust trend analysis and forecast for the specified feeders, offering a valuable tool for power system planning and management.



```
In [5]: import pandas as pd
import matplotlib.pyplot as plt
from statsmodels.tsa.arima.model import ARIMA

# Replace this with your actual data
data = {
    'FEEDERS': ['SANGO', 'FSM', 'SUMO', 'AMJE', 'ESTATE', 'IDIROKO RD'],
    '2018': [19, 1, 6, 22.5, 19, 13],
    '2019': [20.5, 10, 6.5, 23.5, 18, 13],
    '2020': [22, 18, 6, 21, 14, 16],
    '2021': [18, 8, 5, 22, 16, 16.5],
    '2022': [18, 6, 4, 22.5, 14.5, 16]
}

df = pd.DataFrame(data)

# Melt the DataFrame to have a 'Year' column and a 'Peak Load' column
melted_df = df.melt(id_vars='FEEDERS', var_name='Year', value_name='Peak Load')

# Convert 'Year' to datetime for proper plotting
melted_df['Year'] = pd.to_datetime(melted_df['Year'], format='%Y')

# Forecast for the years 2023-2027
forecast_years = list(range(2023, 2028))

plt.figure(figsize=(14, 6))
for feeder in melted_df['FEEDERS'].unique():
    feeder_data = melted_df[melted_df['FEEDERS'] == feeder]

    # Fit ARIMA model
```

Figure 1: Screenshot of the Python code snippet for ARIMA implementation

2.6 Data Pre-processing

This research uses trend analysis to understand long-term electricity load consumption patterns across five locations of an OTA transmission station. The Autoregressive Integrated Moving Average (ARIMA) model is used, which efficiently handles temporal patterns like trends, seasonality, and cyclic behavior, based on real-world utility requirements. This model comprises three main components: Auto-Regressive (AR), Integrated (I), and Moving Average (MA).

The AR component represents the linear relationship between the current value and previous values in a time series, capturing trends over time [12]. The I component accounts for differencing to make the data stationary, ensuring that the model can handle non-constant variance [12]. The MA component captures the relationship between the current value and past forecast errors [12], helping to account for any random noise in the data. In this study, the ARIMA model was employed to perform a comprehensive trend analysis of the load consumption data for each location.

2.7 The Training Phase

The ARIMA model's Auto-Regressive and Moving Average components during the training phase, with historical Ota which is divided into training and testing phases.

A. Auto-Regressive (AR) Component

The AR component can be expressed mathematically as:

$$X_t = C + \emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \cdots + \emptyset_p X_{t-p} + \epsilon_t \quad (1)$$

Where, X_t is the electricity load consumption at time *t*; C is a constant; $\emptyset_1$, $\emptyset_2$ …$\emptyset_p$ are autoregressive coefficients; p is the order of the autoregressive component; ϵ_t represents white noise, which is the residual error at time *t*.

B. Integrated (I) Component

The integrated component, *d*, represents the order of differencing required to make the data stationary. It can be expressed as:

$$Y_t = (1 - B)^d X_t \quad (2)$$

Where, Y_t is the differenced series; B is the backward shift operator; *d* is the order of differencing.

C. Moving Average (MA) Component

The MA component can be expressed as:

$$X_t = \mu + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q} \quad (3)$$

Where, μ is the mean of the series; ϵ_t represents the white noise error at time *t*; $\theta_1, \theta_2, \ldots, \theta_q$ are the moving average coefficients; *q* is the order of the moving average component.

2.8 The Testing Phase

In the testing phase of the ARIMA the model's performance is assessed to evaluate its ability to accurately forecast future energy requirements. During this phase, historical data not used in training the ARIMA model is employed to simulate real-world conditions. The model's predictions are then compared to the actual values in the test set to measure its accuracy and effectiveness in capturing temporal patterns within the data.

2.9 Model Evaluation

To ensure the effectiveness of the ARIMA model in predicting load consumption, the study focused on several crucial evaluation metrics:

i. Mean Absolute Error (MAE): MAE quantifies the average magnitude of errors between predicted and actual values [22] [26]. The mean absolute error (MAE) is the average of all absolute errors. The formula is

$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n}|x_t - x| \quad (4)$$

Where, n is the number of errors, Σ is the summation symbol (which means 'add them all up'), $|x_{t} - x|$ is the absolute errors.

ii. Mean Absolute Percentage Error (MAPE): MAPE calculates the percentage difference between the predicted and actual values [27].

$$M = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \quad (5)$$

Where, n is the number of fitted points; A_t is the actual value; F_t is the forecast value; Σ is summation notation (the absolute value is summed for every forecasted point in time).

Percentage errors are calculated in terms of absolute errors, without regards to sign. This avoids the problem of positive and negative errors canceling each other out. It offers insights into the relative error of the forecasts, making it a valuable metric for assessing the model's accuracy in percentage terms.

iii. Root Mean Square Error (RMSE): RMSE provides an assessment of the model's prediction error in the same units as the data [27].

Root mean square error can be expresses as:

$$RMSE = \sqrt{\frac{\Sigma(y_i - \hat{y}_i)^2}{N-P}} \quad (6)$$

Where, y_i is the actual value of the ith observation; $\hat{y}_i$ is the predicted value for the ith observation; P is the number of the parameter estimated, including the constant; N is the number of observations.

iv. R-squared (R^2): R^2 measures the proportion of the variance in the dependent variable (electricity load) that is predictable from the independent variable (time). A higher R^2 indicates that the model explains a larger portion of the load's variance, reflecting its goodness of fit [3].

v. Residual Analysis: Examining the residuals (the differences between actual and predicted values) is crucial. Residual plots, autocorrelation plots of residuals, and tests for normality can reveal any systematic errors, bias, or patterns in the model's predictions.

3.0 RESULTS AND DISCUSSION

3.1 Primary Data Acquisition Result

The historical peak load data for each of the feeders, between 2018 and 2022 is acquired for this study at the OTA transmission station. The aggregates of the entire peak load per year are then presented in Table 2. Each row corresponds to a specific feeder, while columns represent individual years. The values in the table denote the recorded peak loads in megawatts (MW) for each respective feeder in the corresponding year. The feeders include SANGO, FSM, SUMO, AMJE, ESTATE, and IDIROKO RD. The aggregates in Table 2 are then used to train the ARIMA model, to have a yearly forecast approach for the next five years. The trend analysis of the data, reveals interesting trends and patterns. For instance, the SANGO feeder experienced a peak load of 19 MW in 2018, which gradually increased to 22 MW in 2020, and then slightly decreased to 18 MW in 2021 and 2022. FSM, on the other hand, exhibits a substantial increase from 1.0 MW in 2018 to 18 MW in 2020, followed by a decrease to 6 MW in 2022. SUMO maintains a relatively stable load of around 6 MW, while AMJE starts at 22.5 MW, fluctuates, and eventually stabilizes around 22 MW. ESTATE and IDIROKO RD show variations in their peak loads over the years. The dataset is crucial for understanding the power consumption patterns of each feeder by the ARIMA, enabling pattern recognition in electricity distribution before its forecast for the next five years.

The dataset in Table 3 provides a detailed overview of the minimum and maximum load values recorded by 33 kV feeders at the OTA Transmission Station during the period from 2018 to 2022. Focusing on the 2018-2022 readings, the observed trends for each feeder are distinct and reveal critical insights into their operational behavior. In January 2019, the SUMO feeder exhibited the lowest recorded load at 2MW, indicative of a period of reduced energy demand. Conversely, FSM experienced its highest load in August 2020, reaching 18 MW, reflecting a significant surge in power consumption during that summer month. AMJE, in June 2019, recorded its maximum load at 23.5 MW, possibly influenced by increased energy demand during that particular month. SANGO, with a peak load of 22 MW in February 2020, and IDIROKO, registering 19.5 MW in July 2021, both captured substantial energy consumption events.

Table 2: Aggregate of the Annual Peak Load for OTA T/S between 2018 and 2022

FEEDERS	**2018** (MW)	**2019** (MW)	**2020** (MW)	**2021** (MW)	**2022** (MW)
SANGO	19	20.5	22	18	18
FSM	1.0	10	18	8	6
SUMO	6	6.5	6	5	4
AMJE	22.5	23.5	21	22	22.5
ESTATE	19	18	14	16	14.5
IDIROKO RD	13	13	16	16.5	16

Finally, ESTATE noted its lowest load in April 2020 at 8 MW, highlighting a period of reduced electricity demand. Examining the temporal aspects, the dates associated with minimum and maximum loads provide additional context. For instance, SUMO's minimum load occurred in October 2022, indicating a recent period of decreased demand. FSM's minimum load was recorded in February 2021, possibly reflecting a rainy season with lower energy usage. AMJE's maximum load in June 2019 aligns with the mid-year timeframe when energy demand often rises. SANGO's maximum load in February 2020 and IDIROKO's in July 2021 coincided with months known for increased power consumption, likely due to seasonal factors.

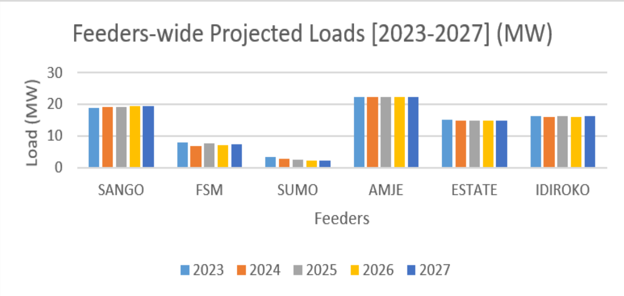


Figure 2: Projected annual peak loads for the five feeders

Table 3: OTA T/S 33 kV Feeders' Minimum and Maximum Load from 2018 to 2022

FEEDERS	MIN LOADS (MW)	DATES	MAX LOADS (MW)	DATES
SUMO	2	OCT 2022	6.5	JAN 2019
FSM	1	FEB 2021	18	AUG 2020
AMJE	14	APR 2021	23.5	JUNE 2019
SANGO	12.5	OCT 2022	22	FEB 2020
IDIROKO	11.5	AUG 2019	19.5	MAY 2018
ESTATE	8	APR 2020	19	OCT 2018

3.2 The Result of ARIMA's Forecasted Peak Loads

The data in Table 2, which is an aggregate of the historical load consumption history of the feeders is used to train the ARIMA machine learning model. During the training, the model does pattern recognition from the data to aid its time series forecasting ability. The experimental result includes the forecasted peak load to be used by the five feeders in the next five years. The ARIMA-based forecasts provide valuable insights into the expected trends in peak loads for each feeder from 2023 to 2027. The forecasted values highlight specific patterns for each feeder, including upward trends, fluctuations, stability, and declines, offering crucial information for power distribution planning and infrastructure development over the projected period. The aggregate forecasted loads in MW are presented in Table 4, while Figure 2 depicts the projected loads in megawatts. The trends of the forecast, including the fluctuations, are presented in the trend plot produced by ARIMA. For the SANGO feeder, the forecasted loads show a steady increase over the five-year period, ranging from approximately 18.91 in 2023 to 19.34 in 2027. This indicates a consistent upward trend in the predicted peak loads for SANGO. In the case of the FSM feeder, the forecasted loads exhibit a slight fluctuation, with values ranging from 6.83 in 2024 to 8.02 in 2023, and this suggests a variable pattern in the predicted peak loads for FSM over the forecasted period. For the SUMO feeder, the forecasted loads show a decrease from 3.21 in 2023 to 2.23 in 2027, indicating a declining trend in the predicted peak loads for SUMO over the five-year period. The AMJE feeder demonstrates a high level of stability in the forecasted loads as observed with values consistently around 22.29 over the forecasted year.

Table 4: Result of ARIMA Anticipated Peak Loads for 2023 to 2027

YEAR	SANGO	FSM	SUMO	AMJE	ESTATE	IDIROKO
2023	18.91028845947822	8.022317917005878	3.208779707265579	22.270411544411846	15.118645482394331	16.25186348321162
2024	19.202603579088986	6.825302129431928	2.7478983897114504	22.294397442400307	14.764479509889457	16.089282909637216
2025	19.296472866635526	7.533819214531323	2.479437645186882	22.291891552891016	14.967234625137143	16.194230409634955
2026	19.326616511467265	7.114445915457997	2.323060789355183	22.292153351812857	14.851160129136439	16.126485675355646
2027	19.336296347625513	7.362674176532284	2.2319721607993976	22.29212600077634	14.91761117085422	16.170215626261317

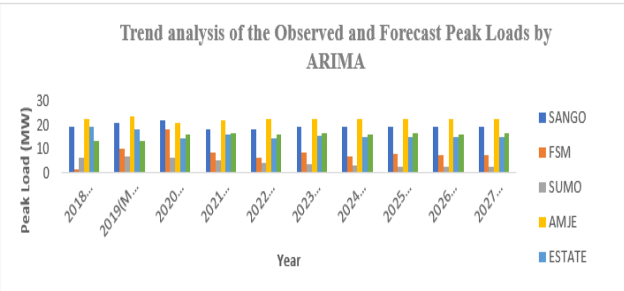


Figure 3: Trend analysis of the observed and forecast peak loads by ARIMA

SUMO on the other hand maintains a consistent load of around 6MW, while AMJE initially peaks at 23.5 MW in 2019 and stabilizes around 22 MW thereafter. IDIROKO demonstrates variations, reaching a peak of 16.5 MW in 2021. Looking ahead to the forecasted values for 2023-2027, the data suggests a continuation of established trends with slight variations. Notably, SANGO's forecast remains within the historical range, and FSM experiences a gradual decline. SUMO maintains its stable load, while AMJE and ESTATE show marginal fluctuations. IDIROKO continues to exhibit variability. The forecasted values align closely with historical patterns, reflecting the robustness of the ARIMA model in capturing and projecting feeder-specific load trends.

By the year 2026 and 2027, the Mobitra transformer should be upgraded to 60 MVA as the total load on it shall be reaching 31 MW by 2026. The transformer is expected to carry 32 MW by its rating 40 MVA 132/33 kV i.e. (MVA rating x Power Factor). Also, T1 40 MVA should be upgraded to 60 MVA too following the addition of Tower Alloy 33 kV feeders as of June 2023. As Sango 33 kV feeder shall be carrying 19.3MW by 2027. The load on Tower Alloy presently is 3 to 4 MW with the rate of development and Industrial growth in Ota Metropolis. Following the trend analysis chart of Figure 3, it could be deduced that never should Sango and Amje be on the same transformer. Both show the highest projected loads between 2023 to 2027. The comprehensive examination of each feeder's trend contributes to a holistic understanding of the dynamic energy landscape, supporting informed decision-making for sustainable power infrastructure.

To maintain power system stability and efficiency during peak periods, several strategies can be employed. These include encouraging consumers to shift electricity usage to off-peak periods, implementing demand-side management programs, deploying large-scale battery systems to balance load, and developing microgrids for localized power generation. When combined, these techniques ensure a stable and efficient power system. This study has proposed recommendations for infrastructures like transformers, and other switch gears needed to improve electricity supply and optimization, risk mitigation strategies, and policy implications for energy planning, infrastructure development, and sustainability initiatives.

3.3 Statistical Summary of the Projected Peak Loads

The statistical summary is presented in Table 5, giving valuable insights into the projected peak loads for the six feeders throughout 2023 to 2027. The table presented the necessary metrics of min, max, mean, median, and interquartile range. Considering the SANGO feeder, the mean projected load is approximately 19.21 with a relatively low standard deviation of 0.18, indicating a consistent and stable forecast. The minimum and maximum values of 18.91 and 19.34, respectively, further reinforce the narrow range of expected peak loads. Quartiles (Q1, Q2, and Q3) illustrate the distribution of the projected values, with Q2 (median) aligning closely with the mean, reinforcing the symmetry in the distribution. In the case of the FSM feeder, the mean load is approximately 7.37, and the standard deviation is 0.45, suggesting a slightly higher variability compared to SANGO.

The range between the minimum (6.83) and maximum (8.02) projected values is wider, indicating a broader spectrum of potential peak loads. Quartiles provide a detailed breakdown of the load distribution, with the median (Q2) aligning with the mean. For the SUMO feeder, the mean projected load is around 2.60, with a standard deviation of 0.39. The minimum and maximum values are 2.23 and 3.21, respectively. The quartiles again showcase the distribution characteristics, highlighting the variability in the projected peak loads. The AMJE feeder exhibits a mean load of approximately 22.29 with a remarkably low standard deviation of 0.01. The narrow range between the minimum (22.27) and maximum (22.29) indicates a high level of precision in the forecasts. Quartiles, especially the median, align closely with the mean, emphasizing the stability in the projections.

For the ESTATE feeder, the mean load is approximately 14.92, and the standard deviation is 0.13. The distribution, as indicated by quartiles, is well-balanced, with a small range between the minimum (14.76) and maximum (15.12) values. The IDIROKO feeder presents a mean load of approximately 16.17, with a standard deviation of 0.06. The quartiles provide insight into the distribution, and the narrow range between the minimum (16.09) and maximum (16.25) values reinforces the consistency in the forecasts.

Table 5: Statistical summary of the projected peak loads

FEEDER	MEAN	STD	MIN	MAX	Q1	Q2	Q3
SANGO	19.214456	0.178052	18.910288	19.336296	19.202604	19.296473	19.326617
FSM	7.371712	0.451220	6.825302	8.022318	7.114446	7.362674	7.533819
SUMO	2.598230	0.393429	2.231972	3.208780	2.323061	2.479438	2.747898
AMJE	22.288196	0.009994	22.270412	22.294397	22.291892	22.292126	22.292153
ESTATE	14.923826	0.132805	14.764480	15.118645	14.851160	14.917611	14.967235
IDIROKO	16.166416	0.062517	16.089283	16.251863	16.126486	16.170216	16.194230

Table 6 presents the percentage improvement based on the average of the present (the observed peak load between the years 2018 to 2022) and predicted (the predicted peak load between the years 2023 to 2027) state of the network. It was observed that Sango, FSM, SUMO, AMJE, ESTATE, and Iddiroko feeders will have 1.46, 14.28, 52.75, 0.05, 8.44, and -8.49 % by 2027 respectively based on the developed model. Table 7 shows the Paired Samples Correlations table ran on SPSS to test the validity of the obtained results. This table provides information on the relationship between the observed and predicted values. This analysis is crucial for assessing how well the predicted values align with the observed ones. The correlation coefficient of 0.989 indicates an extremely strong positive linear relationship between the observed and predicted values. A correlation that is close to 1 suggests that the predicted values are very closely aligned with the observed values. The high correlation coefficient of 0.989 and its statistical significance (p = .000) provide strong evidence that the predicted values are highly valid when compared to the observed values. The minimal bias, low standard error, and the narrow confidence interval that lies near 1.000 all support the conclusion that the predicted values accurately reflect the observed values. This high level of agreement validates the predictive model or method used.

Table 6: Percentage improvement based on the present and future state of the network

FEEDERS	AVERAGE OBSERVED PEAK LOAD (MW)	AVERAGE PREDICTED PEAK LOAD (MW)	% IMPROVEMENT
SANGO	19.5	19.214456	1.464328205
FSM	8.6	7.371712	14.2824186
SUMO	5.5	2.59823	52.75945455
AMJE	22.3	22.288196	0.052932735
ESTATE	16.3	14.923826	8.442785276
IDIROKORD	14.9	16.166416	-8.499436242

Table 7: Paired Samples Correlations

					Bootstrap for Correlation[a]			
							95% Confidence Interval	
		N	Correlation	Sig.	Bias	Std. Error	Lower	Upper
Pair 1	Observed & Predicted	6	.989	.000	-.012	.113	.897	1.000

Table 8: Result of the performance evaluation of the developed ARIMA model

FEEDER	MAE	MAPE	RMSE	R2
SANGO	5.317042848558577	0.2781886522308178	8.687632826629676	-30.4479017209723
FSM	5.5014789943322295	0.7400799645707229	6.763646043342719	-0.4738050193178356
SUMO	1.5754202003120663	0.2672249364063426	2.7265395294350214	-8.292522256964686
AMJE	5.237530376272458	0.23374177757318731	10.1129635380905	-153.95762351931504
ESTATE	5.220834836414644	0.2959693029520959	8.740384559601674	-19.317638896203018
IDIROKO	3.2383749749015633	0.23985614384694826	5.966890803352399	-13.591715516037473

3.4 Result of the Performance Evaluation of the Developed ARIMA Model

The performance metrics of the ARIMA model are computed and presented in Table 8. The ARIMA model results for the six feeders reveal varying levels of accuracy and predictive performance. For SANGO, the model displays a Mean Absolute Error (MAE) of 5.32 MW, indicating an average absolute difference between the observed and predicted values. However, the Mean Absolute Percentage Error (MAPE) of 0.28% suggests a relatively low percentage of error in the predictions. The Root Mean Square Error (RMSE) of 8.69 MW indicates the square root of the average squared differences between observed and predicted values. Surprisingly, the R-squared (R2) value of -30.45 indicates a poor fit of the model to the data, possibly suggesting the presence of unaccounted complexities or trends.

Similarly, for FSM, the model shows an MAE of 5.50 MW, a MAPE of 0.74%, and an RMSE of 6.76 MW. These metrics collectively indicate moderate predictive accuracy. The negative R-squared value of -0.47, though not ideal, might suggest that the model struggles to capture the variability in the data. In the case of SUMO, the model exhibits a lower MAE of 1.58 MW, a MAPE of 0.27%, and a relatively lower RMSE of 2.73 MW, implying a better predictive performance with minimal errors. The negative R-squared value of -8.29, however, indicates a less favorable fit. For AMJE, the model's MAE of 5.24 MW and MAPE of 0.23% show a reasonable accuracy, but the higher RMSE of 10.11 MW suggests larger errors in predicting peak loads. The R-squared value of -153.96 indicates a significant deviation from the actual data, questioning the model's appropriateness for capturing the underlying patterns. In the case of ESTATE, the model displays an MAE of 5.22 MW, a MAPE of 0.30%, and an RMSE of 8.74 MW, suggesting a moderate level of predictive accuracy. The negative R-squared value of -19.32 indicates a suboptimal fit, pointing towards the limitations of the model in capturing the nuances of the data. The model demonstrates an MAE of 3.24 MW, a MAPE of

0.24%, and an RMSE of 5.97 MW, indicating a relatively better performance compared to other feeders, with the IDIROKO RD forecast. The R-squared value of -13.59 suggests a suboptimal fit, but the model still provides reasonably accurate predictions.

Summarily, the lower the MAE the better the model predicts. However, the relationship between MAE values and how well a model performs depends on the data, MAE cannot be compared across different models and data sets. Hence, it can be developed further by calculating the MAPE to make it easier to compare the model performance and interpret error value. Therefore, for MAPE analysis, a MAPE value of <10% indicates high accurate forecasting, that is: <10: Highly Accurate Forecasting; 10 – 20: Good Forecasting; 20 – 50: Reasonable Forecasting; >50: Inaccurate Forecasting.

For Root Mean Square Error, the lower the RMSE, the better the model and its prediction. A higher RMSE value of 20 and above indicate that there is a large deviation from the residual to the ground truth. While in the case of R^2, a higher R-Square value generally mean a better fit.

4.0 CONCLUSION

In conclusion, this study embarked on a comprehensive analysis of the 33 kV feeders at the OTA Transmission Substation, unraveling key insights into the annual peak load patterns from 2018 to 2022. The meticulous process began with the acquisition of data reflecting the dynamic load behaviors of six critical feeders: SANGO, FSM, SUMO, AMJE, ESTATE, and IDIROKO RD. These historical load profiles formed the foundation for a robust pattern recognition and forecasting endeavor using the Auto Regressive Integrated Moving Average (ARIMA) model. The ARIMA algorithm, applied to each feeder, demonstrated varying degrees of success in predicting the annual peak loads for the subsequent five years (2023-2027).

The detailed analysis encompassed a rich array of statistical performance metrics, including Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), and R-squared (R2) values. The obtained results unveiled nuanced patterns, with certain feeders exhibiting notable accuracy in predictions while others faced challenges in capturing the complexities of the underlying data. The conclusion drawn from the comprehensive performance metrics evaluation emphasizes the need for careful consideration of model selection and parameter tuning to enhance predictive accuracy. Moving forward, this study lays the groundwork for continued research and refinement of forecasting methodologies, offering valuable insights for power system planning and resource allocation at the OTA Transmission Substation.

REFERENCES

[1] Alizadeh, R., Soltanisehat, L., Lund, P. and Zamanisabzi, H., “Improving renewable energy policy planning and decision-making through a hybrid MCDM method”. *Energy Policy,* Vol 137, 2020. p. 111174

[2] Chanchangi, Y. N. et al., “Nigeria's energy review: Focusing on solar energy potential and penetration”. *Environment, Development and Sustainability*, 2022. pp. 1-42.

[3] Gomes, I. S. F., Perez, Y. and Suomalainen, E., “Coupling small batteries and PV generation: A review”. *Renewable and Sustainable Energy Reviews*, Vol 126, 2020. p. 109835.

[4] Okafor, C. C. et al., “Biomass utilization for energy production in Nigeria: A review”. *Cleaner Energy Systems*, 2022. p. 100043.

[5] Edomah, N., Ndulue, G. and Lemaire, X., “A review of stakeholders and interventions in Nigeria's electricity sector”. *Heliyon* , 7(9). 2021.

[6] Olaleye, T. et al.,.” Predictive Analytics and Software Defect Severity: A Systematic Review and Future Directions”. *Scientific Programming*, 2023. pp. 1-18.

[7] Tyralis, H., Karakatsanis, G., Tzouka, K. and Mamassis, N., “ Exploratory data analysis of the electrical energy demand in the time domain in Greece”. *Energy*, Vol 134, 2017. pp. 902-918.

[8] Jiang, Z. and Liu, Z., “Policies and exploitative and exploratory innovations of the wind power industry in China: The role of technological path dependence” . *Technological Forecasting and Social Change*, Vol 77 , 2022. p. 121519.

[9] Shamim, G. and Rihan, M., “Exploratory Data Analytics and PCA-Based Dimensionality Reduction for Improvement in Smart Meter Data Clustering”. *Journal of Research*, 2023. pp. 1-10.

[10] Chodakowska, E., Nazarko, J. and Nazarko, Ł., “Arima models in electrical load forecasting and their robustness to noise”. *Energies*, 14(23), 2021. p. 7952.

[11] Dat, N. Q., Anh, N. T. N., Anh, N. N. and Solanki, V. K., “Hybrid online model based multi seasonal decompose for short-term electricity load forecasting using ARIMA and

https://doi.org/10.4314/njt.v43i4.14

online RNN". *Journal of Intelligent & Fuzzy Systems*, 41(5), 2021. pp. 5639-5652.

[12] Gupta, A. and Kumar, A., "Mid term daily load forecasting using ARIMA, wavelet-ARIMA and machine learning", *IEEE*, 2020. pp. 1-5.

[13] Ade-Ikuesan, O. O., Osifeko, M. O., Okakwu, I..K., Folaranmi, K. S., and Alao, P. O "Prediction of electricity consumption demand pattern for 2018 in Ogun State, Nigeria". *Journal of Applied Sciences and Environmental Management*, 22(6), 2018. pp. 883-886.

[14] Rajbhandari, Y., Marahatta, A. and Ghimire, B., "Impact study of temperature on the time series electricity demand of urban nepal for short-term load forecasting". *Applied System Innovation*, 4(3), 2021. p. 43.

[15] Li, R., Jiang, P., Yang, H. and Li, C., "A novel hybrid forecasting scheme for electricity demand time series" . *Sustainable Cities and Society*, Vol 55, 2020. p. 102036.

[16] Bouktif, S., Fiaz, A., Ouni, A. and Serhani, M. A., "Single and Multi-Sequence Deep Learning Models for Short and Medium Term Electric Load Forecasting". *Energies*, 12(1), 2019. p. 149.

[17] Noureen, S., Atique, S., Roy, V. and Bayne, S., "*Analysis and application of seasonal ARIMA model in Energy Demand Forecasting: A case study of small scale agricultural load.* Dallas" , s.n., 2019. pp. 521-524.

[18] Alasali, F., Nusair, K., Alhmoud, L. and Zarour, E., "Impact of the COVID-19 Pandemic on Electricity Demand and Load Forecasting" . *Sustainability*, 13(3). 2021.

[19] Shah, I., Iftikhar, H., Ali, S. and Wang, D., "Short-Term Electricity Demand Forecasting Using Components Estimation Technique" . *Energies*, 12(13), 2019. p. 2532.

[20] Maldonado, S., González, A. and Crone, S., "Automatic time series analysis for electric load forecasting via support vector regression". *Applied Soft Computing*, Vol 83, 2019. p. 105616.

[21] Ngabesong, R. and McLauchlan, L., "Implementing "R" Programming for Time Series Analysis and Forecasting of Electricity Demand for Texas, USA. Lafayette", *IEEE*, 2019.pp. 1-4.

[22] Aseeri, A. O. "Effective RNN-Based Forecasting Methodology Design for Improvi-ng Short-Term Power Load Forecasts: Application to Large-Scale Power-Grid Time Series", *Journal of Computational Science*, Vol 68. 2023. p. 101984.

[23] Nepal, B., Yamaha, M., Yokoe, A., and Yamaji, T., "Electricity load forecasting using clustering and ARIMA model for energy management in buildings". *Japan Architectural Review*, 3(1), 2019. pp. 62-76.

[24] Ade-Ikuesan, O. O., Oyedeji, A. O. and Osifeko, M. O. "Linear regression long-term energy demand forecast modelling in Ogun State, Nigeria". *Journal of Applied Sciences and Environmental Management*, 23(4), 2019. pp. 753-757.

[25] Obi, P. I, Okonkwo. I. I. and. Ogba, C. O., "Power Supply Enhancement in Onitsha Distribution Network using Distribution Generations". Nigerian Journal of Technology (NIJOTECH); 41(2), 2022. pp.318-329 .http://dx.doi.org/10.4314/njt.v41i2.14.

[26] Ashigwuike, E. C., Aluya, A. R. A., Emechebe, J. E. C. and Benson, S. A. " Medium Term Electrical Load Forecast of Abuja Municipal Area Council using Artificial Neural Network Method". *Nigerian Journal of Technology (NIJOTECH)* 39(3), 2020, pp. 860 – 870. http://dx.doi.org/10.4314/njt.v39i3.28

[27] Tian, L. W., Xu, H. and Tong, H., "Short-term power load forecasting model based on t-SNE dimension reduction visualization analysis, VMD and LSSVM improved with chaotic sparrow search algorithm optimization". *Journal of Electrical Engineering & Technology*, 17(5), 2022, pp. 2675-2691.