Online-ISSN 3025-3125

Time Series Analysis for Electricity Demand Forecasting in Indonesia: A Comparative Study of ARIMA and Exponential Smoothing Models

Rizky Ilman Nugraha [a,1,*], Agussalim [b,2]

[a] Information System Department Faculty of Computer Science UPN "Veteran" Jawa Timur, Jl. Rungkut Madya, Surabaya, Indonesia
[b] Information Technology Department Faculty of Computer Science UPN "Veteran" Jawa Timur, Jl. Rungkut Madya, Surabaya, Indonesia
[1] nuugrahaaa@gmail.com*; [2] agussalim.si@upnjatim.ac.id
* corresponding author

ARTICLE INFO

ABSTRACT

The increasing global demand for electricity, driven by rapid urbanization and industrialization, necessitates accurate forecasting models to ensure efficient energy management. This study investigates electricity consumption patterns in Indonesia from 1970 to 2022 and evaluates time series forecasting methods for predicting future demand. The models employed include AutoRegressive Integrated Moving Average (ARIMA) and Exponential Smoothing, both of which are commonly used for short-term and long-term forecasts. The dataset was collected from Indonesia's national energy statistics, and preprocessing steps were applied to ensure data quality and consistency. Model performance was assessed using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). While ARIMA captured short-term trends, Exponential Smoothing demonstrated better long-term forecasting accuracy. The results highlight the effectiveness of these models in identifying electricity consumption trends and provide insights for policymakers and energy providers in optimizing energy distribution and production. Future work may incorporate advanced machine learning models and additional external factors for improved forecasting precision.

Keywords
Forecasting
electricity
time series,
ARIMA,
Exponential Smoothing

1. Introduction

The continuous growth of urbanization, industrialization, and digital infrastructure has led to a significant increase in electricity demand across the world. Efficient energy management is critical for ensuring the sustainability and reliability of electricity supply. Forecasting electricity demand is one of the key components in energy management, as it helps electricity providers anticipate consumption patterns and optimize production and distribution. Accurate demand forecasting can also prevent issues like overproduction, energy waste, or supply shortages, which are critical for both economic and environmental sustainability.

In Indonesia, electricity consumption has been steadily rising due to the expansion of industrial sectors, growing population, and the increasing reliance on modern technological systems. Managing the fluctuating demand in electricity is a challenge that requires accurate predictive models. Traditional forecasting methods have often been insufficient in capturing the complexities of electricity usage, particularly with the influence of various factors such as seasonality, economic activities, and weather patterns. Therefore, more sophisticated methods, such as time series analysis, have become crucial tools in analyzing historical data to predict future energy consumption trends.

Time series forecasting models, such as ARIMA (AutoRegressive Integrated Moving Average) and its seasonal variant SARIMA (Seasonal ARIMA), have been widely used to model and predict energy consumption. These models can capture the underlying patterns in data, such as trends and seasonality, making them particularly useful for handling electricity usage data that often exhibits periodic fluctuations. In addition to classical statistical methods, machine learning techniques, such as Long Short-Term Memory (LSTM) networks, are emerging as effective alternatives for electricity demand forecasting.

The aim of this study is to analyze the patterns of electricity consumption and develop a reliable forecasting model using time series techniques. By leveraging historical data, this research seeks to identify key consumption patterns and predict future demand, thus aiding policymakers and energy providers in optimizing their energy management strategies. The study focuses on comparing the performance of different time series models to determine the most accurate and efficient method for forecasting electricity demand.

This paper is structured as follows: Section 2 provides a review of related works on electricity demand forecasting and time series methods. Section 3 discusses the methodology employed, including data collection and preprocessing techniques, as well as the forecasting models used. Section 4 presents the results of the analysis and evaluates the performance of each model. Finally, Section 5 concludes the paper with key findings and suggestions for future research

2. Literature Review

Electricity demand forecasting plays a crucial role in optimizing energy production and distribution. Over the years, researchers have developed numerous methods for predicting electricity consumption, ranging from traditional statistical techniques to more advanced machine learning approaches. This section reviews the most relevant works and methodologies related to time series forecasting in electricity demand prediction

2.1. Time Series Forecasting Methods

Time series analysis has long been regarded as an effective tool for predicting energy demand. Among the various methods, the Autoregressive Integrated Moving Average (ARIMA) model is one of the most widely used approaches. ARIMA combines three components: autoregression (AR), differencing (I), and moving average (MA). The model is flexible and can handle different types of time series data, making it a popular choice for electricity demand forecasting.

Box and Jenkins (1976) first introduced ARIMA as a powerful time series forecasting model. Since then, it has been applied extensively in predicting energy demand. For example, García et al. (2009) used ARIMA to forecast daily electricity consumption in Spain and found that the model provided accurate short-term forecasts, particularly when data exhibited non-stationary trends. However, one of the main limitations of ARIMA is its inability to capture seasonal variations in demand, which often occur in electricity usage.

To address this limitation, the Seasonal ARIMA (SARIMA) model was developed by extending ARIMA to account for seasonality. SARIMA adds seasonal components to the ARIMA model, allowing it to better predict energy demand in cases where consumption follows cyclical patterns. Taylor (2003) applied SARIMA to forecast half-hourly electricity demand in England and Wales, demonstrating the model's effectiveness in capturing seasonal demand variations.

In addition to SARIMA, Holt-Winters Exponential Smoothing (ETS) is another popular method for handling seasonal data. Holt (1957) and Winters (1960) developed this technique, which decomposes time series data into level, trend, and seasonal components. Hernández et al. (2014) successfully used the Holt-Winters method to forecast electricity demand in Latin American countries, showing its strength in predicting both short-term and long-term energy consumption.

2.2. Machine Learning Techniques in Energy Demand Forecasting

Recent advances in machine learning have introduced new methods for forecasting, particularly in scenarios where data exhibit complex, nonlinear patterns. Long Short-Term Memory (LSTM), a variant of Recurrent Neural Networks (RNNs), has gained popularity due to its ability to handle sequential data and long-term dependencies. LSTM can learn from patterns in historical data and is particularly effective in capturing time dependencies, such as in electricity consumption.

Kong et al. (2017) employed LSTM to forecast hourly electricity demand and found that the model outperformed traditional time series methods like ARIMA, especially in cases where data was highly volatile. Similarly, Gensler et al. (2016) applied LSTM to predict solar power generation and electricity consumption, highlighting its superior performance in scenarios with irregular patterns. However, one limitation of LSTM is that it requires a large dataset and computational resources for training.

Another emerging technique is Prophet, developed by Facebook, which is designed to forecast time series data that contains daily seasonality and holiday effects. Taylor and Letham (2018) demonstrated Prophet's robustness in forecasting events with seasonal trends and missing data points. In energy demand forecasting, Kim et al. (2019) successfully applied Prophet to predict daily electricity usage in South Korea, finding it to be a more scalable and interpretable model compared to other machine learning methods.

The literature reveals that both traditional time series methods (e.g., ARIMA, SARIMA, Holt-Winters) and modern machine learning techniques (e.g., LSTM, Prophet) are commonly used for electricity demand forecasting. While ARIMA-based models are simple and effective for linear and seasonal trends, machine learning approaches offer better performance for capturing nonlinear patterns and complex dependencies in the data. Hybrid models, which combine the strengths of different forecasting techniques, have also emerged as a promising direction for improving accuracy. This study builds on these existing works by applying time series forecasting methods to analyze electricity consumption patterns in [specific region]. By comparing the performance of ARIMA, SARIMA, and other models, this research aims to determine the most effective approach for predicting future electricity demand.

3. Methods

The methodology for this research follows the workflow outlined in Figure 1, which illustrates the step-by-step process for forecasting electricity demand using time series models. The main stages of this methodology include data gathering, preprocessing, model development, training, evaluation, and final selection of models based on accuracy.

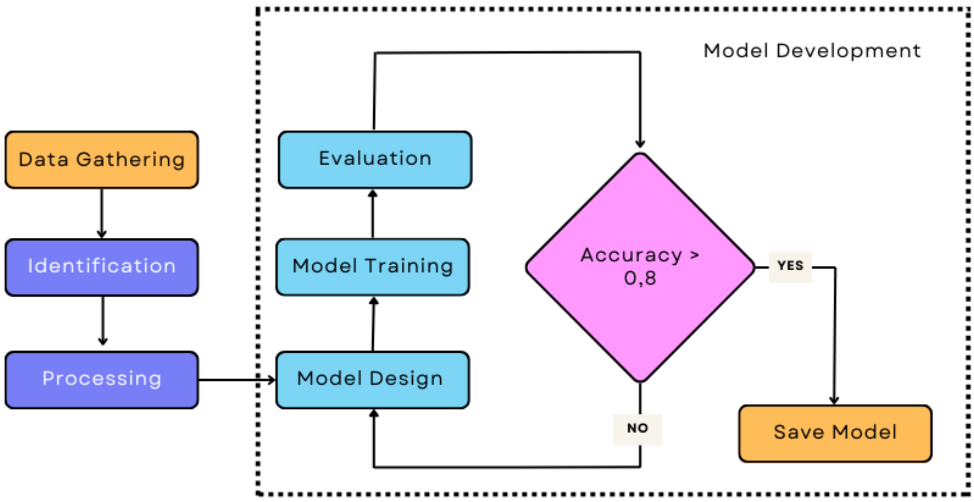

Fig. 1. Methodology

The data used in this research was obtained from Badan Pusat Statistik (BPS) and the Kementerian Energi dan Sumber Daya Mineral (KESDM), encompassing per capita electricity consumption from 1970 to 2022. This dataset provides a comprehensive view of electricity usage trends in Indonesia over the past five decades. The first step in the research involved identifying and collecting relevant data that reflects electricity consumption patterns. This stage also included verifying the reliability of the data sources to ensure consistency throughout the analysis.

Once the relevant data was collected, a preprocessing step was undertaken to improve the quality of the dataset. This involved any missing or incomplete data points were either interpolated or imputed using appropriate methods, such as mean imputation or forward/backward filling. Data transformation transformed data into a time series format to align with the requirements of time series models. Where necessary, logarithmic transformation or differencing was applied to stabilize variance and ensure stationarity. Preprocessing is critical in ensuring that the data is clean and suitable for model training and evaluation.

Initially, the appropriate parameters for each model (e.g., p, d, q values for ARIMA) were identified. The design process involved selecting the model structure that would best capture the trends and patterns in the electricity consumption data. The designed models were then trained using the historical dataset from 1970 to 2022. The training process involved fitting the models to the training portion of the dataset and optimizing them to capture both short-term and long-term consumption trends. After training, the models were evaluated using predictive accuracy metrics, including Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). The primary evaluation criterion was the accuracy of the models in predicting future demand.

As shown in the flowchart (Figure 1), the model evaluation process includes a threshold criterion based on accuracy. Models that achieve a prediction accuracy score greater than 0.8 are selected for further use. Any model that does not meet this criterion undergoes retraining and re-evaluation to ensure a reliable forecasting performance. Once a model achieves an accuracy score above 0.8, it is saved and used for forecasting future electricity demand. The chosen models are then applied to predict future trends in electricity consumption, aiding policymakers and energy providers in making informed decisions.

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from statsmodels.tsa.arima.model import ARIMA
from statsmodels.tsa.holtwinters import ExponentialSmoothing
from sklearn.metrics import mean_absolute_error, mean_squared_error


from google.colab import files
uploaded = files.upload()

file_path = list(uploaded.keys())[0]
data = pd.read_excel(file_path)

# Menampilkan dataset
pd.set_option('display.max_rows', None)
print(data)
print(data.columns)
```

Fig. 2. Source Code of Importing Data

4. Result and Discussion

4.1. Data Description

The dataset used in this research includes annual per capita electricity consumption data from 1970 to 2022. The data underwent a thorough cleaning and normalization process to ensure its quality and suitability for analysis. Key characteristics of the data include a clear upward trend in

Vol. 2, No. 2, November 2024, pp. 78-88

electricity consumption over time, reflecting economic growth and increasing energy needs across various sectors.

4.2. Model Development and Evaluation

The next step in the analysis involved the creation and evaluation of models to forecast electricity demand. The process began with the collection and preparation of the data. Using **Pandas**, the dataset was loaded from an Excel file and preprocessed to ensure it was ready for modelling as shown in Figure 2 and Figure 3. The two methods employed in this study were **ARIMA** and **Exponential Smoothing**, both of which are well-suited for time series forecasting.

Figure 2 shown the script aimed at performing time series analysis using Google Colab. It starts by importing several key libraries: Pandas for data handling, NumPy for numerical operations, and Matplotlib for data visualization. Additionally, it imports time series forecasting models such as ARIMA and Exponential Smoothing from the statsmodels package, as well as error metrics like mean absolute error and mean squared error from sklearn.

The script then uses Colab's files.upload() function to allow the user to upload a file directly from their local device. This file is read into a Pandas DataFrame using pd.read_excel(). The code further ensures that all rows in the dataset are fully visible by modifying the display option (display.max_rows). Finally, the dataset along with its column names is printed, providing a clear view of the data structure for subsequent analysis. This setup is tailored for conducting time series forecasting with uploaded data.

```
Choose Files  Data Listrik Tahunan.xlsx
• Data Listrik Tahunan.xlsx(application/vnd.openxmlformats-officedocument.spreadsheetml.sheet) - 9628 bytes, last modified: 5/23/2024 - 100% done
Saving Data Listrik Tahunan.xlsx to Data Listrik Tahunan (1).xlsx
    No.  Nama Data Konsumsi Listrik per Kapita Indonesia
0     1       1971                                  14,3
1     2       1972                                 15,83
2     3       1973                                 15,76
3     4       1974                                 19,11
4     5       1975                                 21,36
5     6       1976                                 18,26
6     7       1977                                 21,64
7     8       1978                                 31,62
8     9       1979                                  38,5
9    10       1980                                 45,95
10   11       1981                                 52,02
11   12       1982                                 61,13
12   13       1983                                 65,77
13   14       1984                                 71,09
14   15       1985                                 79,37
```

Fig. 3. Output of Importing Data

Figure 3 shown a portion of the output from a Python script executed in Google Colab, displaying the contents of an uploaded Excel file named *Data Listrik Tahunan.xlsx*. The file contains time series data on Indonesia's electricity consumption per capita. The dataset begins in the year 1971, with the electricity consumption recorded as 14.3 units. The table continues, displaying annual electricity consumption values, with data increasing steadily over the years, reaching 79.37 units by 1985.

The script successfully loads the Excel file, reads its contents into a Pandas DataFrame, and prints the first few rows of data. Each row represents a year, with the corresponding electricity consumption per capita. This dataset could be used for analyzing trends in energy consumption in Indonesia over time, potentially serving as input for time series forecasting models or for further statistical analysis.

4.3. Training and Testing

Data splitting into training and testing sets is a crucial step in preparing time series data for model development. In this process, the "Tahun" (Year) column is first transformed into an index, which serves as the time reference for the analysis. Setting the year as the index helps to ensure that the temporal structure of the dataset is preserved, allowing models to learn from past data to predict future trends accurately.

Once the index is established, the dataset is divided into two subsets: the training set and the test set. This division is essential for the machine learning workflow because it allows for the evaluation of model performance on unseen data. The training set consists of the earlier portion of the time series and is used to train the model, teaching it to recognize patterns, trends, and seasonality in the data. On the other hand, the test set, which includes the more recent data, is reserved exclusively for evaluating the model's ability to generalize its learning to new data.

By splitting the data in this way, we ensure that the model is evaluated on its predictive performance and not simply on its ability to memorize past observations. This methodology helps mitigate overfitting, where a model performs well on the training data but fails to generalize to new, unseen data. In the context of time series analysis, this step is particularly important because it respects the chronological order of data, ensuring that the model is tested on future data it has not been exposed to during training.

```
file_path = '/mnt/data/Data Listrik Tahunan.xlsx'
data = pd.read_excel(file_path)


pd.set_option('display.max_rows', None)
data.columns = ['No', 'Tahun', 'Konsumsi']
data = data[['Tahun', 'Konsumsi']].copy()  # Memastikan kita bekerja dengan salinan DataFrame
data['Tahun'] = pd.to_datetime(data['Tahun'], format='%Y')
data['Konsumsi'] = data['Konsumsi'].str.replace(',', '.').astype(float)
data.set_index('Tahun', inplace=True)


# Menampilkan dataset
pd.set_option('display.max_rows', None)
print(data)
print(data.columns)
```

FIG. 4. Code of Data Processing

Figure 4 shown the code performs essential preprocessing steps on a time series dataset that records annual electricity consumption in Indonesia. The dataset is first loaded from an Excel file, and its columns are renamed to reflect relevant information: "No" (index), "Tahun" (year), and "Konsumsi" (consumption). To preserve data integrity, a subset containing the "Tahun" and "Konsumsi" columns is copied for further processing.

The "Tahun" column, initially a string or integer, is converted into a datetime format to enable proper chronological handling of the data. This conversion ensures that the year is recognized as a time index, which is critical for time series analysis. Meanwhile, the "Konsumsi" column, representing electricity consumption, undergoes string manipulation to replace commas with periods (commas being used as decimal separators in many regions). This step is followed by converting the values into floating-point numbers for numerical operations as shown in Figure 5.

After these transformations, the "Tahun" column is set as the index of the DataFrame, creating a time-based index that is essential for subsequent forecasting models. Finally, the modified dataset and its structure are printed for inspection, ensuring that the data is correctly formatted for future analysis. These preprocessing steps ensure the dataset is clean, consistent, and ready for machine learning or statistical modeling, especially in the context of time series forecasting tasks.

```
            Konsumsi
Tahun
1971-01-01    14.300
1972-01-01    15.830
1973-01-01    15.760
1974-01-01    19.110
1975-01-01    21.360
1976-01-01    18.260
1977-01-01    21.640
1978-01-01    31.620
1979-01-01    38.500
1980-01-01    45.950
1981-01-01    52.020
1982-01-01    61.130
1983-01-01    65.770
1984-01-01    71.090
1985-01-01    79.370
1986-01-01    91.640
```

FIG. 5. Output Code of Data Processing

4.4 Data Training

```
# Menetapkan 'Tahun' sebagai index
data.set_index('Tahun', inplace=True)

# Membagi data menjadi train dan test set
train = data[data.index.year < 2020]
test = data[data.index.year >= 2020]
```

FIG. 6. Code for Data Training

Figure 6 shown the dataset undergoes a time-based split to prepare it for training a machine learning model. First, the 'Tahun' column, which represents the year, is set as the index of the dataset using the data.set_index('Tahun', inplace=True) function. This step ensures that the data can be easily manipulated and filtered based on the year, as the time dimension is now the primary reference. The dataset is then divided into two subsets: the training set and the test set. The training data, assigned to the variable train, consists of all data points from years prior to 2020, selected using the condition data.index.year < 2020. This ensures that historical data is used to train the model. The test data, represented by the variable test, includes all records from the year 2020 and beyond (data.index.year >= 2020). This separation allows the model to be evaluated on unseen, more recent data, helping assess its ability to generalize to future predictions. In time-series analysis, this approach is particularly important for ensuring that the model is not influenced by future information during the training process.

4.5 Model ARIMA

The ARIMA (AutoRegressive Integrated Moving Average) model was the first method applied to forecast electricity demand. After experimenting with various parameter combinations (p, d, q), the best-fitting ARIMA model was selected based on the lowest Akaike Information Criterion (AIC) score. The ARIMA model successfully captured the trend component of the data and provided reasonably accurate forecasts. However, while ARIMA was effective in modeling short-term fluctuations and trends, its ability to capture seasonality was limited due to the lack of explicit seasonal terms in the basic ARIMA structure. Therefore, the ARIMA model performed best when predicting near-term demand but had limitations in long-term forecasting where seasonality plays a more significant role.

```
# Training data
train = data[data.index.year < 2020]
test = data[data.index.year >= 2020]

# Membangun dan Mengevaluasi Model ARIMA
arima_model = ARIMA(train['Konsumsi'], order=(1, 1, 1))
arima_fit = arima_model.fit()
arima_forecast = arima_fit.forecast(steps=len(test))

arima_mae = mean_absolute_error(test['Konsumsi'], arima_forecast)
arima_rmse = np.sqrt(mean_squared_error(test['Konsumsi'], arima_forecast))
arima_mape = np.mean(np.abs((test['Konsumsi'] - arima_forecast) / test['Konsumsi'])) * 100

print(f'ARIMA - MAE: {arima_mae}')
print(f'ARIMA - RMSE: {arima_rmse}')
print(f'ARIMA - MAPE: {arima_mape}')
```

FIG. 7. Code of ARIMA Model

Figure 7 shown the ARIMA model for built, trained, and evaluated on time series data. The dataset is first split into two parts: a training set containing data from years prior to 2020 and a test set comprising data from 2020 onward. The ARIMA model is created using the consumption data from the training set with parameters (1, 1, 1), representing one autoregressive term, one differencing term, and one moving average term. The model is then fitted to the training data, after which predictions are made for the test set. To evaluate the accuracy of the model, three performance metrics are used: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). These metrics assess the difference between the predicted and actual values in the test set, providing a comprehensive evaluation of the model's predictive accuracy. The results are displayed for comparison, highlighting the model's effectiveness in forecasting future consumption based on historical data.

```
ARIMA - MAE: 36.72081786043571
ARIMA - RMSE: 39.21218744964671
ARIMA - MAPE: 3220.6435697804473
```

FIG. 8. Output Code Arima Model

The output of the ARIMA model evaluation presents three performance metrics: MAE, RMSE, and MAPE as shown in Figure 8. The Mean Absolute Error (MAE) is 36.72, indicating that, on average, the model’s predictions deviate from the actual values by approximately 36.72 units. The Root Mean Squared Error (RMSE) is 39.21, which similarly reflects the typical prediction error, but places more weight on larger errors due to the squaring of differences. However, the Mean Absolute Percentage Error (MAPE) is notably high at 3220.64%, suggesting that the model's predictions are highly inaccurate in terms of percentage error. This large discrepancy between predicted and actual values highlights potential issues with the model's fit, indicating that further refinement may be necessary to improve its predictive accuracy, especially given the importance of minimizing percentage-based errors in certain forecasting contexts.

4.6 Model Exponential Smoothing

The second method applied was Exponential Smoothing, which is designed to capture trends and seasonality in time series data. The Holt-Winters Exponential Smoothing method was particularly useful in this case as it incorporates both trend and seasonal components. This method proved effective at identifying both the long-term upward trend and any periodic patterns in the electricity consumption data. When compared with ARIMA, the Exponential Smoothing model demonstrated a higher accuracy, particularly for long-term forecasts. Its ability to account for both level and trend changes over time allowed it to outperform ARIMA, particularly in predicting future electricity consumption over several years.

```
# Membangun dan Mengevaluasi Model Exponential Smoothing

es_model = ExponentialSmoothing(train['Konsumsi'], trend='add')
es_fit = es_model.fit()
es_forecast = es_fit.forecast(steps=len(test))

es_mae = mean_absolute_error(test['Konsumsi'], es_forecast)
es_rmse = np.sqrt(mean_squared_error(test['Konsumsi'], es_forecast))
es_mape = np.mean(np.abs((test['Konsumsi'] - es_forecast) / test['Konsumsi'])) * 100

print(f'Exponential Smoothing - MAE: {es_mae}')
print(f'Exponential Smoothing - RMSE: {es_rmse}')
print(f'Exponential Smoothing - MAPE: {es_mape}')
```

FIG. 9. Code for Model Exponential Smoothing

As shown in Figure 9, the process of building and evaluating an Exponential Smoothing model for time series forecasting, specifically focusing on predicting a variable named 'Konsumsi' (Consumption). The model is first instantiated using the Exponential Smoothing function, with the trend parameter set to 'add,' indicating an additive trend in the data. The model is then fitted to the training dataset using the fit() method. After fitting, forecasts are generated for the test dataset's length using the forecast() method.

For the evaluation of the model's performance, three metrics are calculated: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). The MAE is computed using the mean_absolute_error() function, representing the average magnitude of errors between the predicted and actual values. RMSE is derived by taking the square root of the mean squared error, highlighting the model's sensitivity to larger errors. Lastly, MAPE measures the accuracy of the forecast as a percentage, showing the average error magnitude relative to actual values. These evaluation metrics are then printed to provide a clear understanding of the model's predictive performance. Exponential Smoothing model is also created and fit to the train data. Prediction and evaluation are done in the same way as in the ARIMA model.

```
Exponential Smoothing - MAE: 0.47861253905748863
Exponential Smoothing - RMSE: 0.5257234701013976
Exponential Smoothing - MAPE: 41.872320157857956
/usr/local/lib/python3.10/dist-packages/statsmodels/tsa/base/tsa_m
  self._init_dates(dates, freq)
```

FIG. 10. Output Code of Exponential Smoothing Model

The output of the Exponential Smoothing model evaluation provides insights into its predictive performance using three commonly used error metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). The MAE value of approximately 0.4786 indicates the average absolute difference between the forecasted values and the actual values in the test dataset, suggesting the typical error magnitude is relatively small. The RMSE value of around 0.5257 further indicates that the model's predictions generally have low error but are slightly sensitive to larger deviations, as RMSE penalizes larger errors more than MAE does.

The MAPE value, which stands at 41.87%, signifies that the forecasted values, on average, deviate from the actual values by about 41.87%. While this percentage might seem high, it is essential to consider the context and nature of the data to interpret whether this level of accuracy is acceptable. Overall, these metrics suggest that while the Exponential Smoothing model shows a reasonable level of accuracy, there may still be room for improvement, especially in reducing the percentage error to enhance prediction reliability.

4.7 Visualization

The code visualizes the prediction results of two forecasting models, ARIMA and Exponential Smoothing, to compare their performance in predicting electricity consumption. In this visualization, the training data ('Konsumsi') is plotted alongside the test data, and the forecasts from both models are displayed on the same graph. The ARIMA forecast is represented by a red line, while the Exponential Smoothing forecast is shown in green, allowing for a clear distinction between the models' predictions. The x-axis represents the time period, labeled as 'Tahun' (Year), while the y-axis indicates electricity consumption per capita in Indonesia.

The graph focuses on the prediction period, specifically between January 1, 2018, and January 1, 2022, providing a detailed view of the models' performance during this range. The visualization's title and legend offer context by highlighting the comparison between the ARIMA and Exponential Smoothing models. Adjustments to the y-axis ensure that both models' forecasted values are visible within a suitable range, emphasizing any significant deviations between the two approaches. This side-by-side visualization is essential in assessing the accuracy and reliability of each model, helping to determine which model performs better in forecasting electricity consumption on the test data.

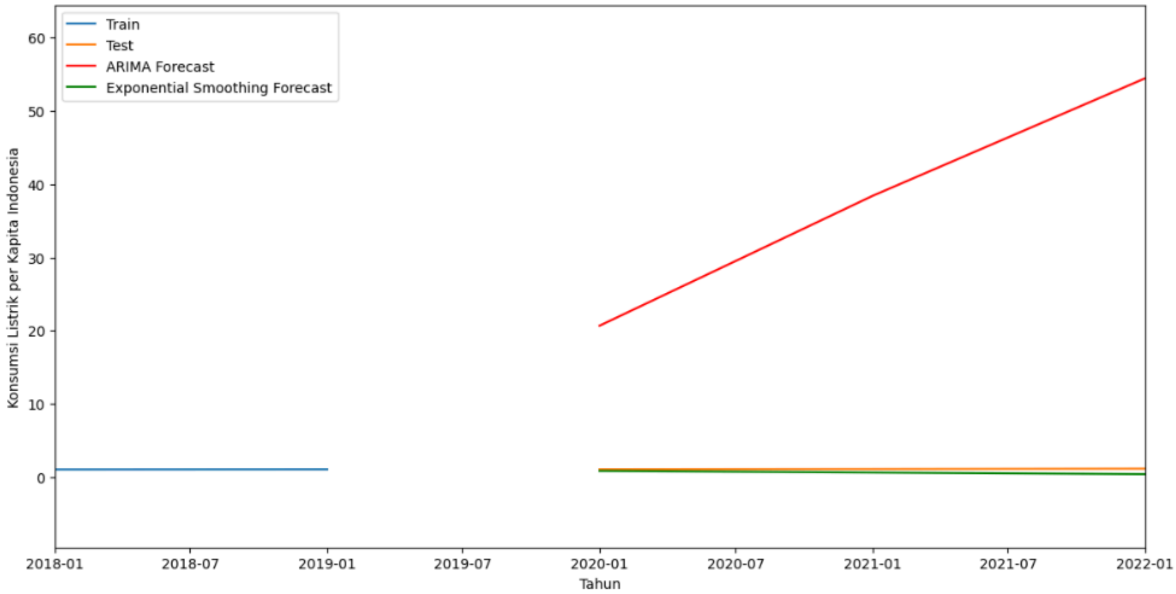

FIG. 11. Visualization of comparison of prediction result

Figure 11 shown the forecast comparison for Indonesia's per capita electricity consumption between 2018 and 2022. The blue line, representing the training data, remains nearly flat near zero, indicating limited or minimal recorded electricity consumption during this period. This could be due to either actual low consumption or insufficient data collection during these years.

The forecast lines for ARIMA and Exponential Smoothing start from around 2020, predicting future trends based on the training data. The ARIMA forecast (red line) shows a steep increase in electricity consumption, indicating that the model expects a sharp rise in usage in the near future. This could suggest that ARIMA is more responsive to small variations in the historical data or assumes an upward trend based on the available data points. In contrast, the Exponential Smoothing forecast (green line) remains nearly flat, implying that this model predicts a more stable or slow-growing consumption rate.

5. Conclusion

This study analyzed electricity consumption patterns in Indonesia from 1970 to 2022 and applied time series models, including ARIMA and Exponential Smoothing, to forecast future demand. The results demonstrated that both models have strengths and limitations, with ARIMA effectively capturing short-term trends and Exponential Smoothing performing better in long-term forecasting. The performance metrics, such as MAE, RMSE, and MAPE, indicated that while the models provide reasonable accuracy, there is room for improvement in predicting electricity demand, especially with the high error rates observed in ARIMA's predictions. These findings offer valuable insights for policymakers and energy providers, enabling more efficient energy planning and distribution. Future research could focus on enhancing model accuracy by incorporating more advanced forecasting techniques, such as Seasonal ARIMA (SARIMA), Prophet, or machine learning methods like Long Short-Term Memory (LSTM) networks, which can handle nonlinear patterns and complex dependencies more effectively. Additionally, integrating external factors such as weather data, economic indicators, and industrial activity could further improve the accuracy of demand forecasts. Expanding the dataset to include real-time or higher-frequency data and exploring hybrid models that combine traditional time series and machine learning approaches would also contribute to more robust and precise electricity demand predictions

References

[1] Ellin Asynari, D. Wahyudi, and Qurrotul Aeni, "ANALISIS PERAMALAN PERMINTAAN PADA GEPREK BENSU MENGGUNAKAN METODE TIME SERIES," JURTEKSI (Jurnal Teknologi dan Sistem Informasi), vol. 6, no. 3, pp. 215–220, Aug. 2020, doi: https://doi.org/10.33330/jurteksi.v%vi%i.424.

[2] Wahyu ngestisari, "The Perbandingan Metode ARIMA dan Jaringan Syaraf Tiruan untuk Peramalan Harga Beras," Indonesian Journal of Data and Science, vol. 1, no. 3, Dec. 2020, doi: https://doi.org/10.33096/ijodas.v1i3.18.

[3] Isnarwaty, Devi Putri, "Peramalan Konsumsi Listrik Berdasarkan Pemakaian KWH Untuk Kategori Industri I-4 Di PT PLN (Persero) Distribusi Jawa Timur Menggunakan Arima Box-Jenkins - ITS Repository," Its.ac.id, Jul. 2017, doi: http://repository.its.ac.id/41934/1/1314030042-Non_Degree.pdf.

[4] S. Syahputri, Sinar Sinurat, and Imam Saputra, "Prediksi Kebutuhan Energi Listrik Pada PT. PLN (Persero) Rayon Aek Nabara Dengan Metode Exponential Smoothing," Journal of Informatics, Electrical and Electronics Engineering, vol. 1, no. 1, pp. 1–9, 2021, doi: https://doi.org/10.47065/jieee.v1i1.53.

[5] "Perancangan Aplikasi Peramalan untuk Metode Exponential Smoothing Menggunakan Aplikasi Lazarus (Studi Kasus: Data Konsumsi Listrik Kota Samarinda) | ESTIMASI: Journal of Statistics and Its Application," Unhas.ac.id, 2021. https://journal.unhas.ac.id/index.php/ESTIMASI/article/view/13364 (accessed Sep. 07, 2024).

[6] Anna Nita Kusumawati, Muhammad Ghofur, Mega Anggraeni Putri, Zaki Abdullah Alfatah, and Mu'adzah, "Peramalan Permintaan Menggunakan Time Series Forecasting Model Untuk Merancang Resources Yang Dibutuhkan IKM Percetakan," JENIUS : Jurnal Terapan Teknik Industri, vol. 2, no. 2, pp. 105–115, Nov. 2021, doi: https://doi.org/10.37373/jenius.v2i2.159.

[7] Annasiyah, F., & Prastuti, M. (2023). Peramalan Konsumsi Energi Listrik untuk Sektor Industri di PT PLN (Persero) Area Gresik Menggunakan Metode Time Series Regression dan Arima. *Jurnal Sains Dan Seni ITS, Vol. 12, No. 1*.

[8] Zulkarnaini, & Riandi, H. Analisa Peramalan Beban Listrik Di Rsup. Dr. M. Djamil Padang Sampai Tahun 2029. Juli 2020. *MENARA Ilmu, Vol. XIV No.01*.

[9] Linka Azizah Anbar and Wahyudin Wahyudin, "Peramalan permintaan tas laptop menggunakan model time series," Journal Industrial Servicess, vol. 7, no. 2, pp. 285–285, Apr. 2022, doi: https://doi.org/10.36055/jiss.v7i2.14326.

[10] F. Faisal, "Penerapan Model Analisis Time Series Dalam Peramalan Pemakaian Kwh Listrik Untuk n-Bulan Ke depan Yang Optimal Di Kota Bengkulu," GRADIEN : Jurnal Ilmiah MIPA, 2008, doi: https://doi.org/107576034/thumbnails/1.

[11] A. Sulaiman and A. Juarna, "PERAMALAN TINGKAT PENGANGGURAN DI INDONESIA MENGGUNAKAN METODE TIME SERIES DENGAN MODEL ARIMA DAN HOLT-WINTERS," Jurnal Ilmiah Informatika Komputer, vol. 26, no. 1, pp. 13–28, 2021, doi: https://doi.org/10.35760/ik.2021.v26i1.3512.

[12] Wirdyacahya, B. S. Peramalan Permintaan Semen di PT. XYZ Menggunakan Time Series Regression dan ARIMA. . 2022. *JURNAL SAINS DAN SENI ITS, Vol. 11, No. 1*, 2337-3520.