GALLOGLY COLLEGE OF ENGINEERING
DATA SCIENCE AND ANALYTICS INSTITUTE
The UNIVERSITY of OKLAHOMA

Comparative Time Series Forecasting of Electricity Production in Turkey: Drift, ARIMA, ETS, and Neural Networks

Author

Nolvenne Leama Tah

Date: December 10th, 2024

Table of Contents

Abstract

Accurate electricity production forecasting is crucial for effective energy planning, especially in a rapidly growing energy market like Turkey. This project aims to develop a reliable forecasting solution by analyzing 35 years monthly electricity dataset sourced from TEIAS. The dataset, encompassing key indicators such as energy imports, exports, gross generation, and transmitted energy, captures seasonal trends and variations in electricity usage. We employ three time series models such as NAIVE Drift, ETS, and ARIMA, to forecast electricity production and compare their performances to a neural network model using standard evaluation metrics. The Seasonal Naive model captures recurring seasonal patterns, while the Drift model predicts based on linear trends. ETS accounts for error, trend, and seasonality components, and ARIMA leverages historical values and lagged errors for precision. By identifying the best-performing model, this study provides practical insights for energy providers to make data-driven decisions, optimize electricity usage, and plan for a sustainable energy future. The results contribute to advancing forecasting techniques and supporting long-term energy strategies in Turkey.

Introduction

Electricity forecasting plays an important role in ensuring the stability and efficiency of energy systems. For a country like Turkey, which has witnessed significant growth in energy demand over the past decades, accurate forecasting is essential for effective energy planning and management. Electricity production and consumption in Turkey is influenced by various factors, including economic activities, population growth, and seasonal variations, such as higher demand during colder winter months and warmer summer periods. These variations make forecasting complex but crucial for maintaining a balance between supply and demand, avoiding overproduction or shortages, and minimizing costs.

This study addresses the problem of forecasting monthly electricity consumption in Turkey using a systematic approach. The goal is to evaluate and compare the performance of different time series forecasting models, including Drift method, ETS (Error, Trend, Seasonality), and ARIMA (Autoregressive Integrated Moving Average) and ANN. By analyzing a 35-year dataset of monthly electricity data, the study aims to identify the most effective model for forecasting electricity consumption patterns in Turkey.

The findings of this study are significant for energy providers and policymakers. Accurate forecasts can guide data-driven decisions, optimize production schedules, and enhance long-term energy planning. This is particularly important in Turkey, where ensuring a stable and sustainable energy supply is a national priority. By comparing traditional and simpler forecasting methods, this project also bridges a gap in the literature, offering practical insights into model performance in the context of Turkey's unique energy landscape. Ultimately, this work contributes to improving energy management strategies and supporting sustainable development.

Literature Review

Electricity forecasting is a crucial area of research, as accurate predictions are essential for effective energy management and planning. Numerous studies have explored various forecasting techniques to model electricity demand.

Traditional time series models such as ARIMA and ETS have been widely applied due to their effectiveness in handling trends, seasonality, and irregular variations in data. The Auto-Regressive Integrated Moving Average (ARIMA) model is one of the most widely used statistical methods in electricity forecasting. Shivam Sharma et al. (2023) demonstrated ARIMA's capability to handle both stationary and non-stationary data effectively. By modeling temporal dependencies and irregular variations, ARIMA remains a benchmark in short- and medium-term forecasting.

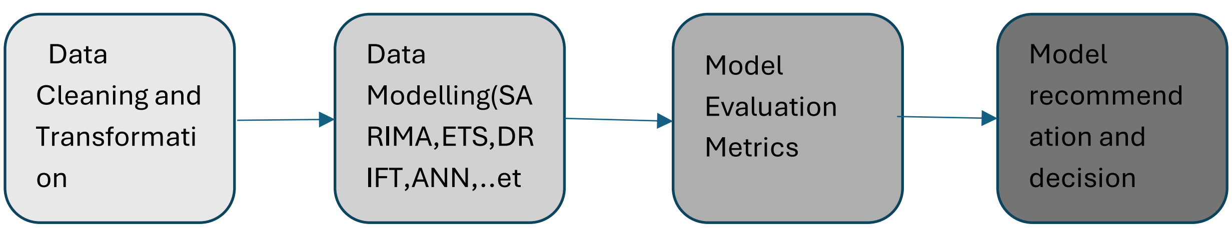
Similarly, the Error, Trend, Seasonal (ETS) model is highly regarded for its ability to explicitly incorporate error, trend, and seasonal components. According to M. Shen et al. (2022) ETS models outperform ARIMA in scenarios where data exhibit clear seasonal patterns, such as electricity demand influenced by weather variations. The study also emphasized the interpretability of ETS models, making them particularly useful for practical applications in utility management.

Recent advancements in machine learning have introduced techniques like Random Forest, XGBoost, and Artificial Neural Networks (ANNs). These methods are capable of capturing complex nonlinear relationships and high-dimensional interactions in data. In a comparative study, Mel Keytingan M. Shapi et al (2021) reported that models such as k-NN, SVM and ANN consistently outperformed traditional time series models in forecasting electricity consumption for metropolitan areas. ANNs, known for their flexibility and ability to learn hidden patterns, have also gained popularity in electricity forecasting.

Despite significant progress, research gaps remain. Many studies focus on individual models or advanced methods without adequately comparing their performance against simpler, more interpretable techniques like the Drift method. Moreover, limited attention has been given to country-specific contexts, such as Turkey, where electricity production is influenced by unique seasonal and economic factors. This study aims to address these gaps by systematically comparing ARIMA, ETS, Drift and ANN models on a comprehensive dataset of Turkey's monthly electricity production. By highlighting their strengths and weaknesses, this project provides valuable insights for practical applications in energy forecasting.

Approach

The approach for this project is divided into four sequential phases:



Meta Data

The dataset utilized in this project is a time series collected from Turkey, spanning a period of 35 years with monthly records of electricity-related data. The dataset includes the following variables: Date (Monthly), Gross Income, Population, Load (Hourly) in MWh, Immediate Load in MWh, Electricity Imports (GWh), Electricity Exports (GWh), Gross Production (GWh), Transmitted Energy

(GWh), Net Electricity Consumption (kWh), T.C. Electricity Consumption (Gross Demand) in kWh, and Lost Electricity (kWh).

Column Name	Min	1st Quartile	Median	Mean	3rd Quartile	Max
Date (Monthly)						
Gross_Income	2490454	3082549	4601099	4932825	6043319	8811568
Population	40446729	48360679	56959988	56820195	65446165	72752325
Load_hourly	2335	5188	10574	12690	18809	33191
Immediate_load	2125	5274	10795	12851	19049	33392
Import	0	36.42	61.65	106.39	164.25	490.3
Export	0	0	22.1	44.9	66.42	314.19
Gross_Production_GWH	1140	2395	5996	7311	10894	20621
Transmitted_energy_GWH	1118	2274	5158	6125	8828	18139
Net_Electricity_Consumption	1234000000	2300000000	4772000000	5936000000	8629000000	16360000000
Gross_Demand_KWh	1429000000	2803000000	5994000000	7409000000	11180000000	20450000000
Lost_Electricity_KWh	194300000	486600000	1217000000	1473000000	2396000000	4091000000
Date	1976-01-01	1984-09-23	1993-06-16	1993-06-16	2002-03-08	2010-12-01

Data Preparation:
After collecting the data, we proceeded to data preprocessing. Each model requires the dataset to be in a specific format before it can be applied. First, we checked for missing data, and fortunately, our dataset contained no missing values. Next, we conducted a visual exploration of the dataset using functions like auto plot () and ggplot() and box plot to identify outliers.

From the visual analysis, we observed that the data exhibits clear seasonality and a trend over time. To further analyze these patterns, we applied STL decomposition to extract the seasonal, trend, and remainder components. This step provided a clearer understanding of the trend patterns and the underlying seasonal effects.

Before applying the ARIMA model, we checked whether the dataset was stationary, as stationarity is a prerequisite for ARIMA. Our analysis revealed that the dataset was not stationary. To address this, we applied first-order differencing, which successfully transformed the data into a stationary series, making it suitable for ARIMA modeling.

Exploratory Data Analysis

Exploratory Data Analysis (EDA) is a crucial step in understanding the dataset and uncovering important patterns, trends, and anomalies that can inform further modeling decisions. In this task, we analyze the dataset to identify trends, seasonality, cyclic patterns, and potential outliers.

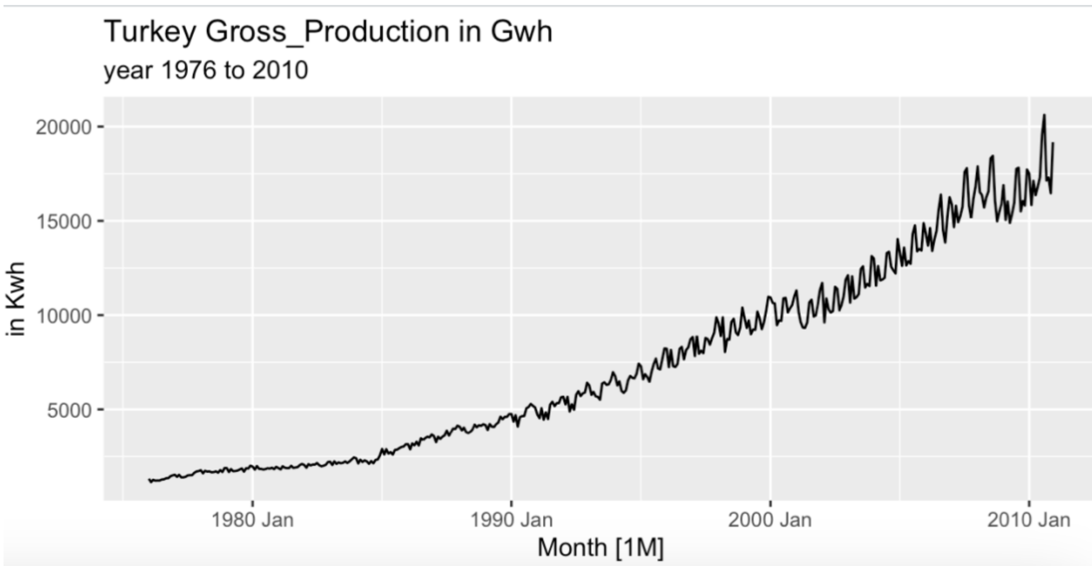


Fig-1: Gross Electricity Production Over Time

Observations: This graph visualizes the time series data of electricity production. It helps in identifying trends, seasonality, and any abrupt changes in production over time. Understanding the overall trend and seasonal patterns aids in selecting an appropriate time series model, such as SARIMA or ETS, which can account for seasonality. The graph shows a steady upward trend, indicating increasing electricity production over the years. Seasonal peaks and troughs are evident, which must be captured by the model to ensure accurate forecasting.

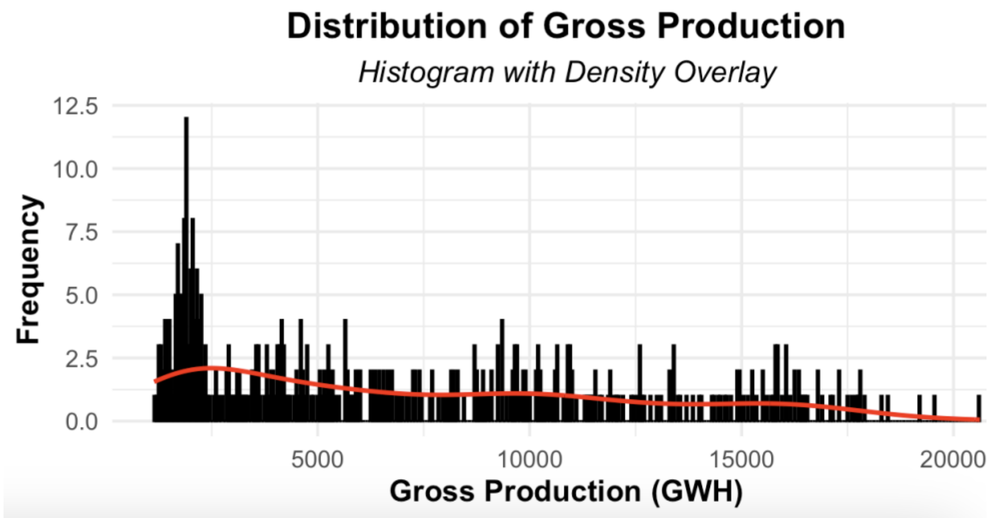


Fig-2: Distribution of Electricity Production

Observations: The histogram shows the frequency distribution of production values, helping us understand data skewness and variability. Skewness or heavy tails in the distribution can influence model assumptions, especially if a transformation is needed to stabilize variance. The data is positively skewed, with most production values concentrated in the lower range. This highlights potential variability in electricity generation over time.

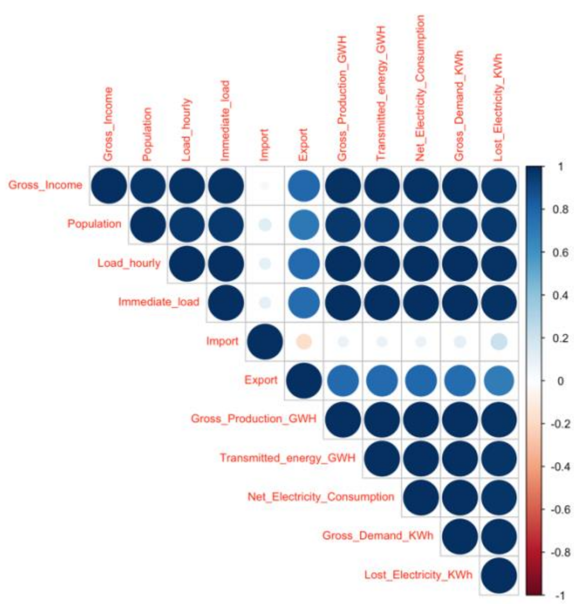


Fig-3: Correlation Matrix

Observations: The correlation matrix identifies relationships between different variables in the dataset, such as Gross Production, Demand, Imports, and Net Consumption. Strong correlations suggest variables that could serve as external regressors (e.g., economic factors or population) to improve model predictions. Gross Production has a high correlation with Net Electricity Consumption, indicating that these variables could complement each other in a regression or ARIMA model.

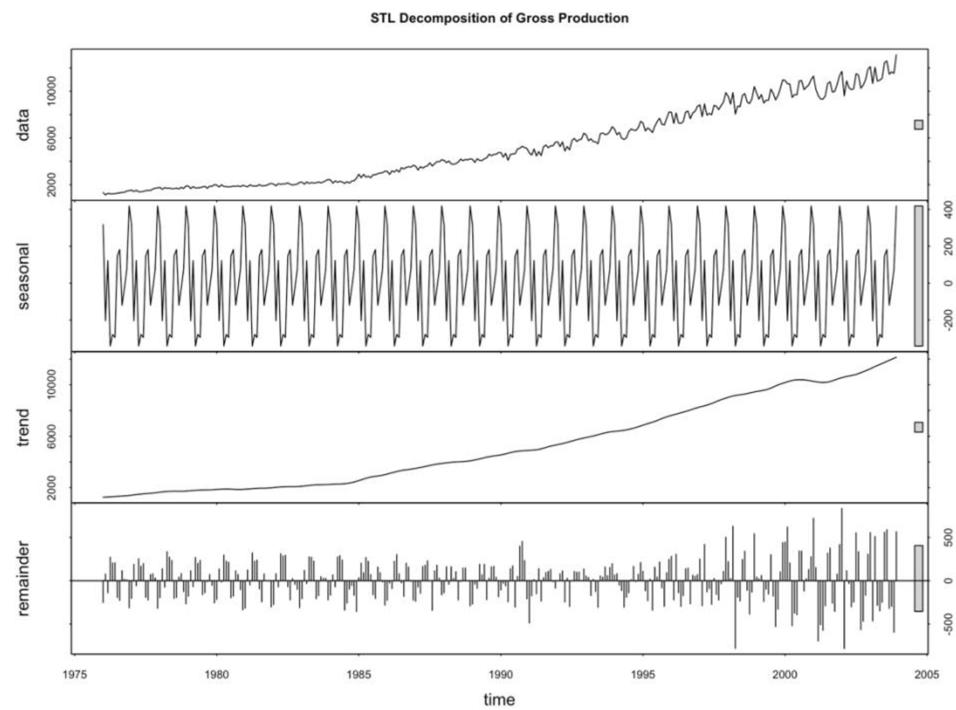


Fig-4: STL Decomposition

Observations: This graph decomposes the time series data into three components: trend, seasonality, and residuals. It allows us to analyze these components individually. STL decomposition helps in understanding if the data requires differencing for stationarity or if additional seasonal adjustments are needed in the model. The trend shows consistent growth, while the seasonal component exhibits predictable fluctuations. The residuals appear random, indicating that the model should focus on capturing the trend and seasonality.

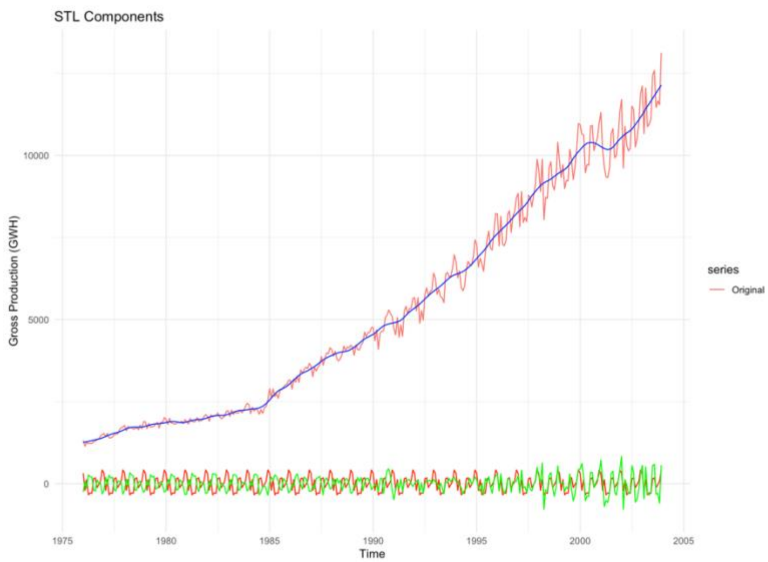


Fig-5: STL Components

Observation: This graph visualizes the original time series along with its decomposed components in a clearer format. Highlights the exact contribution of trend and seasonality to the original series. Helps verify if the seasonal pattern is consistent and strong. Understanding the contribution of trend and seasonality informs the selection of the appropriate modeling approach (e.g., including seasonal terms in ARIMA). The seasonal component is relatively consistent, while the trend is steadily increasing. This emphasizes the need for a model that accommodates both strong seasonality and an upward trend.

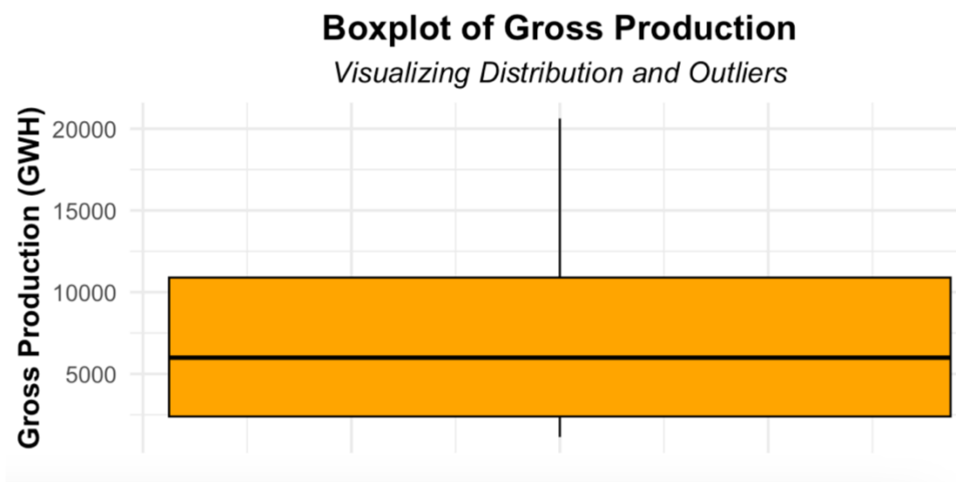


Fig-6: Box Plot for Outlier Detection

Observations: Box plots help identify outliers in the data, which could skew the model or lead to inaccurate forecasts. Addressing outliers ensures the model is not disproportionately influenced by extreme values, improving its generalizability.
While most data points fall within the expected range, a few outliers suggest unusual production spikes. These may need to be handled using transformations or robust models.

Missing Value Analysis

We carefully examined the dataset to identify any missing values across all variables. The results confirmed that the dataset is complete, with no missing values detected. This demonstrates that the data is already well-structured and does not require additional steps like imputation or deletion to address missing entries.

Having a fully complete dataset ensures a robust foundation for further analysis and modeling. It eliminates challenges associated with handling missing data, such as introducing bias during imputation or losing data through deletion, allowing us to proceed directly to the next stages of our project.

missing_values_summary

Variable	Missing_Values	Action_Taken
Date (Montly)	0	No Action Required
Gross_Income	0	No Action Required
Population	0	No Action Required
Load_hourly	0	No Action Required
Immediate_load	0	No Action Required
Import	0	No Action Required
Export	0	No Action Required
Gross_Production_GWH	0	No Action Required
Transmitted_energy_GWH	0	No Action Required
Net_Electricity_Consumption	0	No Action Required
Gross_Demand_KWh	0	No Action Required
Lost_Electricity_KWh	0	No Action Required
Date	0	No Action Required
Total Missing Values	0	No Action Required

Table-1: Missing Values Summary

Decomposition Analysis

Decomposition is essential to separate the time series data into its key components, allowing for a clearer understanding of the underlying trends, seasonal variations, and residual noise. For this project, we utilized the STL (Seasonal-Trend Decomposition using Loess) technique.

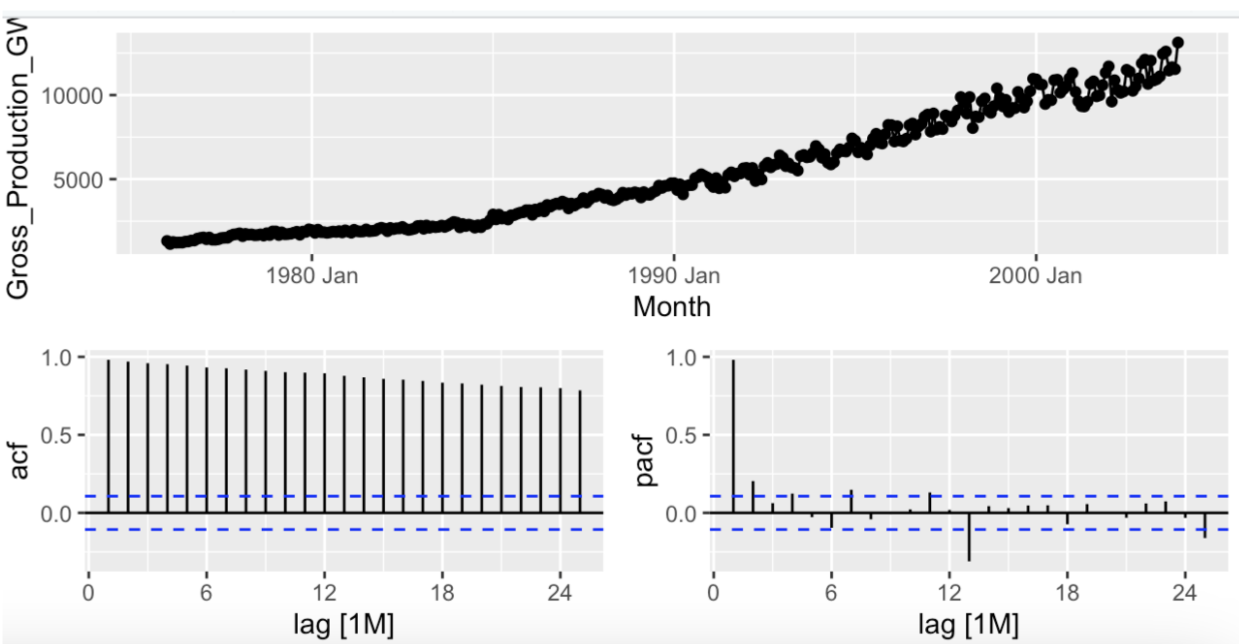
While Task 3.1 introduced Figures 4 and 5 to visually present the decomposition, Task 3.4 focuses on interpreting the results more deeply and exploring outliers or patterns that might have been missed during the initial EDA:

1. Outlier Identification
   The decomposition residuals highlight irregular variations that were not apparent in the raw data. These anomalies might represent one-off events such as economic fluctuations, weather anomalies, or unexpected demand spikes. These insights are crucial for improving model accuracy, as they can be addressed during preprocessing or modeling.
2. Seasonal Variability
   The seasonal component, isolated through decomposition, underscores the importance of capturing cyclical patterns for accurate forecasting. Unlike a simple visualization of the raw

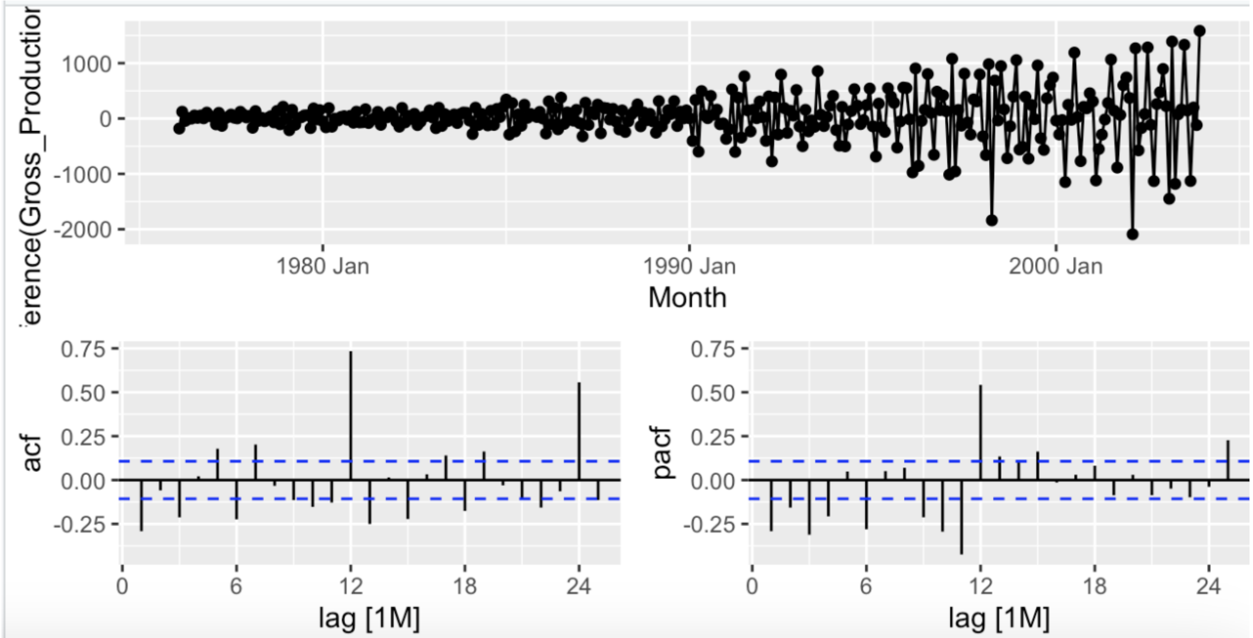
time series in Task 1, decomposition provides a quantifiable measure of seasonal behavior, which is invaluable for model training and evaluation.

3. Residual Analysis for Model Diagnostics
Beyond identifying outliers, the residuals allow us to test whether the remaining data exhibits patterns or autocorrelation that may violate assumptions of predictive models. This diagnostic step ensures that the models used later are robust and unbiased.
4. Validation of EDA Findings
The trend and seasonality components confirm observations from Task 1, such as the steady upward growth and periodic fluctuations. However, decomposition adds precision by numerically separating these components and clarifying their individual contributions.
Refining Forecast Models: By addressing anomalies and seasonality explicitly, decomposition ensures that models capture systematic patterns while treating residuals appropriately.
Supporting Statistical Assumptions: The analysis helps verify assumptions like stationarity and independence, which are critical for ARIMA, ETS, and other time series models.
Model Selection: With a clearer understanding of trends and seasonality, decomposition aids in choosing models (e.g., SARIMA for seasonality or ETS for trend-focused data).

Is the dataset stationary? From the image below we can clearly see that dataset is not stationary. We will apply first differencing.



After differencing:



Model Building:

The next step involves applying the forecasting models, but first, we split the dataset into two subsets: 80% for training and 20% for testing.

For the traditional forecasting models, such as ARIMA, ETS, and the Drift method, we use the train data subset to fit the models and the test data subset to forecast electricity production for the next 10 years.

SARIMA

Based on our visual exploration, we determined that SARIMA is appropriate for this dataset, as it represents monthly time series data. Initially, we manually configured an ARIMA model, followed by an automated selection of the best model. The best-fit model was identified as ARIMA (0,1,1) (1,1,1) [12], which captures the seasonal components effectively.

ETS: For the ETS model, we applied a similar procedure, but we specifically selected Holt's method for modeling. This choice was based on prior observations and practice, where Holt's method consistently delivered better performance for datasets with trend and seasonality.

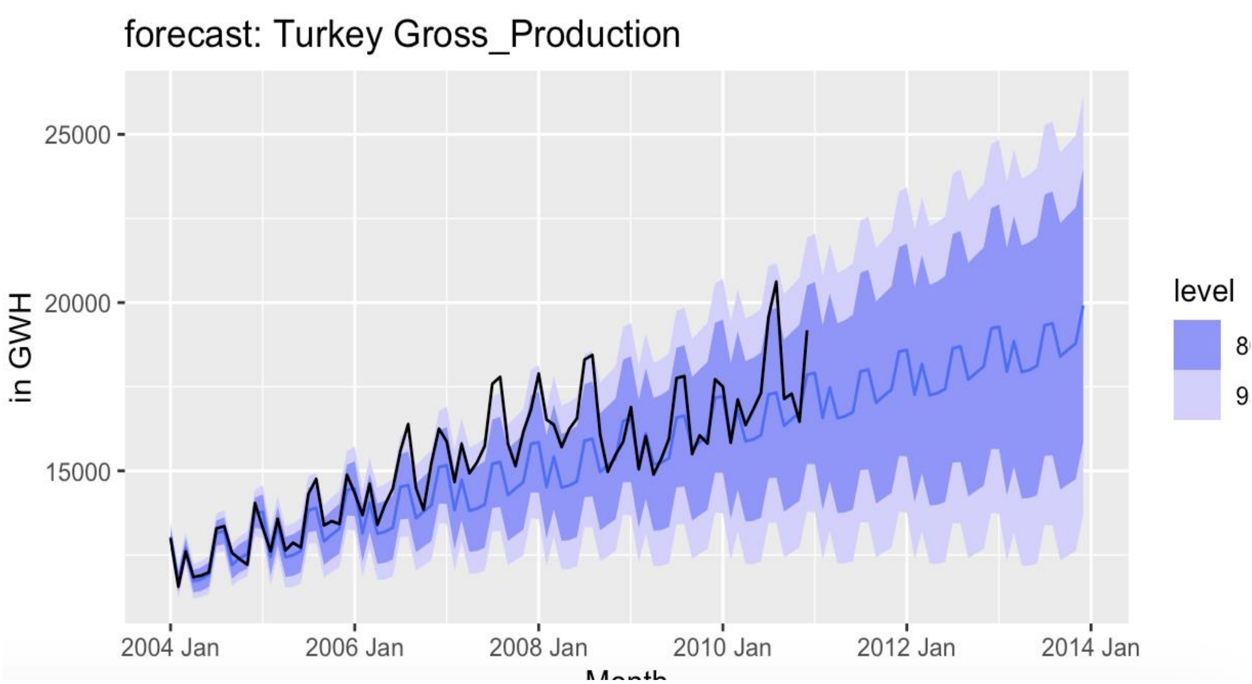
Drift Method: As a baseline, we utilized a standard forecasting approach, the Naive method, and enhanced it by incorporating the Drift method to account for the linear trend observed in the data. This adjustment helps improve the model's performance by extending the trend line into future forecasts.

ANN: Applying a neural network model (ANN) required additional preprocessing steps compared to the traditional forecasting models. First, we normalized the numeric columns in the dataset to ensure the neural network could process the data effectively. After normalization, we split the dataset into training and testing subsets. Next, we defined the neural network architecture using the neural net() function and trained the model on the train data. Finally, we used the trained model to make predictions on the test data. It is a 2 hidden layers neural network.

Results:

The analysis of our forecasting models provides insights into their performance, strengths, and areas for improvement. The results of each model are summarized below:

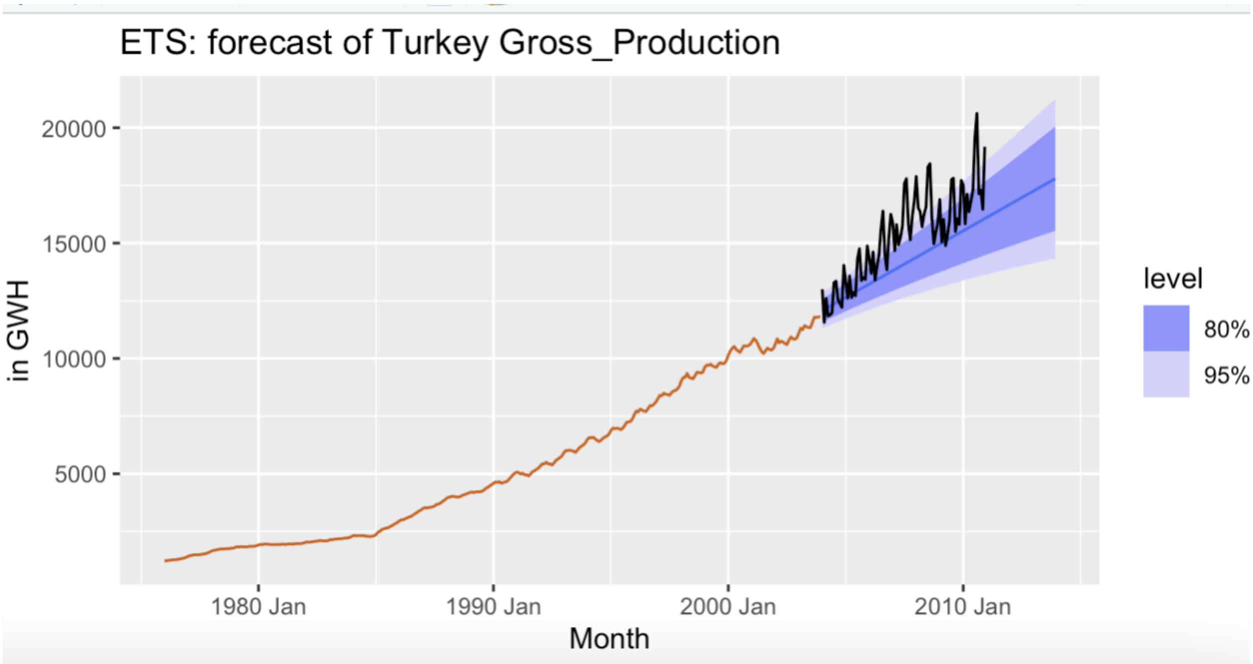
SARIMA: The SARIMA model demonstrated great performance in forecasting electricity production for the next 10 years. The forecast values were closely aligned with the actual data, indicating that the model effectively captured both the trend and seasonality of the dataset. Evaluation metrics confirm that SARIMA handled the variability in the data well and produced reliable forecasts even when applied to new datasets. Its ability to adapt to seasonal patterns and long-term trends makes it a strong candidate for accurate electricity production forecasting.




```
> print(accuracy_metrics)
# A tibble: 1 × 10
  .model .type    ME  RMSE   MAE   MPE  MAPE  MASE RMSSE  ACF1
  <chr>  <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
1 auto   Test   709. 1070.  786.  4.27  4.81   NaN   NaN 0.665
```

ETS

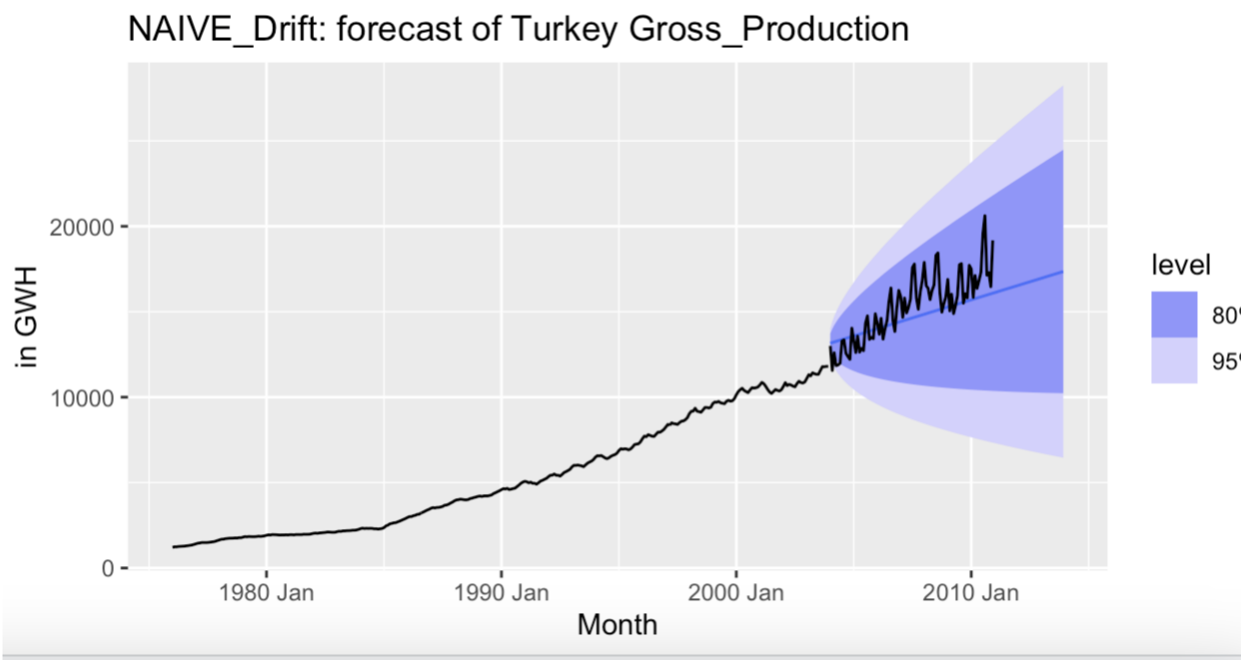
The initial forecast using the ETS model alone was less accurate and did not produce remarkable results. However, after adding Holt's method, we can observe a significant enhancement on the model's precision as depicted in the image below. While the evaluation metrics suggest that the ETS model could reliably forecast electricity production for the next 10 years, there is still room for improvement. We could use some optimization algorithm to get to more accurate and consistent predictions.




```
> print(accuracy_ETS)
# A tibble: 1 × 10
  .model        .type    ME  RMSE   MAE   MPE  MAPE  MASE RMSSE  ACF1
  <chr>         <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
1 Holt's method Test  1229. 1675. 1294.  7.38  7.92   NaN   NaN 0.476
```

NAIVE with Drift Method

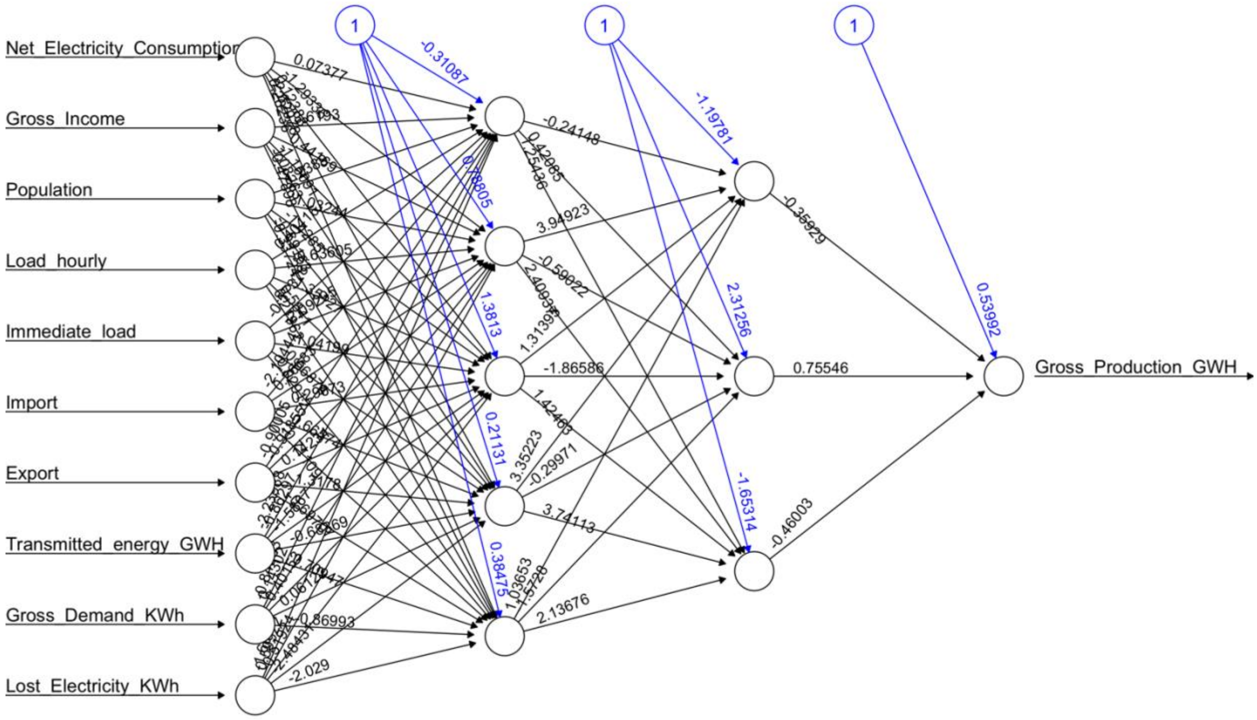
Frome the below image, we can observe that the Naive method with the Drift extension struggled to capture the complex patterns in electricity production. The forecast plot showed a noticeable decline in predicted values over the next 10 years. While the forecasted values stayed within the prediction interval (blue line), the downward trend highlights a limitation in the Drift method's ability to adapt to data with strong seasonal or non-linear trends.




```
> print(accuracy_Drift)
# A tibble: 1 × 10
  .model .type     ME  RMSE   MAE   MPE  MAPE  MASE RMSSE  ACF1
  <chr>  <chr>  <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
1 Drift  Test    675. 1462. 1116.  3.48  6.98   NaN   NaN 0.585
>
```

ANN

The Artificial Neural Network (ANN) outperformed all traditional forecasting models, achieving the lowest Mean Squared Error (MSE) of 0.0059. This demonstrates that the model was able to capture non-linear patterns and interactions within the dataset.

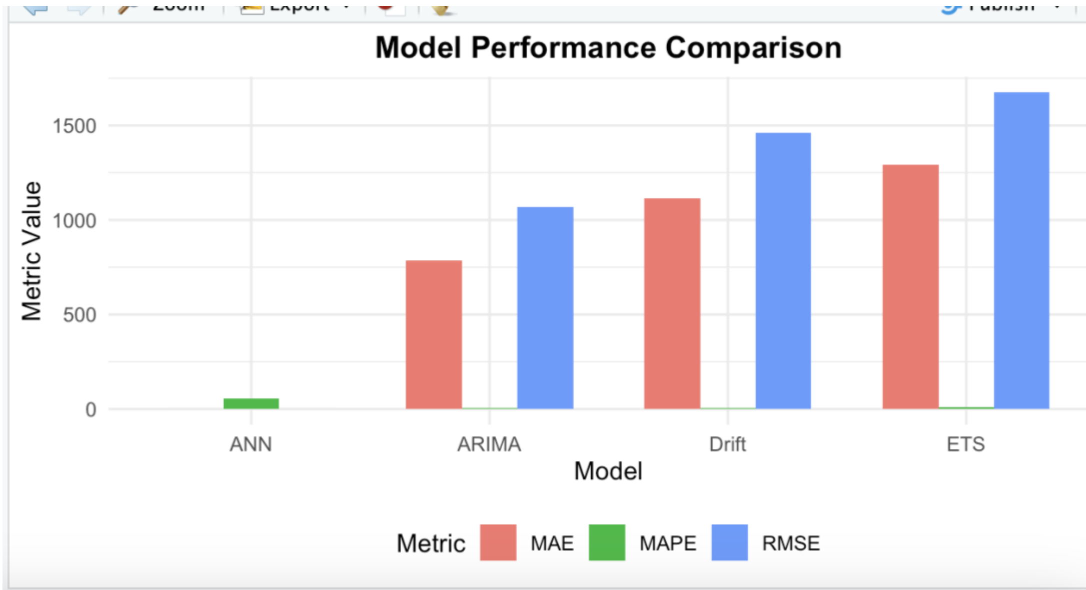



```
> print(metrics_df)
  Metric_1      Value_1
1      MSE 7.503606e-04
2      MAE 1.800874e-02
3     MAPE 1.318032e+01
4     MSLE 2.939178e-04
```

Model Evaluation:

Models were evaluated using performance metrics like RMSE, MAE, and MAPE to identify the best fit for the dataset. A bar chart of all evaluation metrics shows how each model performed compared to other models.

```
> print(comparison_table)
# A tibble: 7 × 13
  .model         .type    ME     RMSE      MAE    MPE  MAPE  MASE RMSSE   ACF1 Model Metric_1  Value_1
  <chr>          <chr> <dbl>    <dbl>    <dbl>  <dbl> <dbl> <dbl> <dbl>  <dbl> <chr> <chr>       <dbl>
1 auto           Test   709.  1.07e+3  7.86e+2   4.27  4.81   NaN   NaN  0.665 ARIMA NA         NA
2 Holt's meth…   Test  1229.  1.67e+3  1.29e+3   7.38  7.92   NaN   NaN  0.476 ETS   NA         NA
3 Drift          Test   675.  1.46e+3  1.12e+3   3.48  6.98   NaN   NaN  0.585 Drift NA         NA
4 NA             Test    NA   2.74e-2 NA        NA    NA       NA    NA NA     ANN   MSE       7.50e-4
5 NA             Test    NA  NA        1.80e-2  NA    NA       NA    NA NA     ANN   MAE       1.80e-2
6 NA             Test    NA  NA       NA        NA    13.2     NA    NA NA     ANN   MAPE      1.32e+1
7 NA             Test    NA  NA       NA        NA    NA       NA    NA NA     ANN   MSLE      2.94e-4
>
```



Conclusion

The objective of this project was to forecast electricity production in Turkey using a variety of models, including SARIMA, ETS, Naive with Drift, and Artificial Neural Networks (ANN). Through systematic analysis and evaluation, we explored each model's strengths, weaknesses, and applicability to the forecasting task. SARIMA provided a reliable model for capturing seasonality and trends, performing well for short- and medium-term forecasts. ETS with Holt's method, offered reasonable predictions. ANN delivered the best performance, with the lowest Mean Squared Error (MSE). This project demonstrates the effectiveness of combining traditional and modern forecasting techniques to address the challenges of electricity production forecasting.

Decision and Recommendation:

The following decisions can be made regarding their suitability for electricity forecasting:

- The ANN model showed superior performance with the lowest Mean Squared Error (MSE) and better adaptability to complex patterns in the dataset. This makes it the most reliable model for forecasting electricity production over the long-term horizon.
- The SARIMA model demonstrated strong results, particularly in capturing seasonal patterns and trends, making it a dependable choice for short-term forecasting.
- Although ETS with Holt's method improved over the basic ETS, its performance was not significant as SARIMA or ANN. However, its interpretability and simplicity make it a useful supplementary model, especially in scenarios requiring clear decomposition of trends and seasonality.
- The Naive method with Drift should be used solely as a baseline for comparison, as it failed to capture the intricate seasonal and trend-based variations effectively.

The following recommendations are proposed:

- We can combine forecasts from multiple models using ensemble methods. This can improve accuracy and robustness by leveraging the strengths of individual models.
- Periodically reevaluate the models with updated datasets to ensure their performance remains accurate over time.

Limitations

- The dataset used in this project is limited to data up to 2010, it does not capture recent consumption trends or advancements in electricity usage.
- While Neural Networks provided strong forecasting results, their implementation requires significant computational power and expertise in hyperparameter tuning, which may not be feasible in all contexts.
- The absence of external variables, such as weather or socio-economic factors, reduces the depth of the analysis. These variables often play a crucial role in shaping electricity demand and consumption patterns.

Further Works

- Future work can incorporate external variables such as weather conditions, economic growth rates, and industrial output, as these factors significantly affect electricity consumption. Integrating these variables may improve the model's accuracy and relevance.
- We plan to combine traditional models like ARIMA with advanced approaches such as Neural Networks to create hybrid models capable of addressing both linear and nonlinear patterns in electricity consumption.

- Using updated datasets, including data beyond 2010, and applying the methodology to other regions or countries can help assess the scalability and effectiveness of our models in different contexts.
-

Tasks

Is the data stationery? If not, find an appropriate difference that yields stationary data.

The dataset is not stationary, as observed from its increasing trend. After applying first-order differencing, the dataset becomes stationary, meeting the requirements for time series modeling like ARIMA.

What can you learn from the ACF and PACF graphs?

The ACF graph indicates significant lags, showing seasonality and autocorrelation patterns, while the PACF graph highlights the most influential lag terms. Together, they guide model selection, such as identifying AR and MA terms in ARIMA.

Identify a couple of ARIMA models that might be useful in describing the time series. Which of your models is the best according to their AIC values? Does the ARIMA function in R give the same model that you chose? If not, which do you think is better?
Candidate models include ARIMA (1,1,1) and ARIMA (0,1,1) (1,1,1) [12]. Based on AIC values, ARIMA (0,1,1) (1,1,1) [12] performs best. The auto. Arima function in R identifies a similar model, validating its suitability for the data.

Write the model in terms of the backshift operator, then without using the backshift operator.

- Backshift operator: $(1 - B)(1 - B^{12})Y_t = (1 - \theta_1 B)(1 - \Theta_1 B^{12})\epsilon_t$
- Without backshift: Differenced series is modeled using a combination of current and past errors with seasonal adjustments.

Estimate the parameters of your best model and do diagnostic testing on the residuals. Do the residuals resemble white noise? If not, try to find another ARIMA model which fits better.
The ARIMA (0,1,1) (1,1,1) [12] parameters were estimated, and residual diagnostics showed no significant autocorrelation, resembling white noise, indicating a good fit.

Forecast the 12 periods ahead of data using your preferred model. Create a plot of the series with forecasts and prediction intervals for the 12 periods ahead.
Using ARIMA (0,1,1) (1,1,1) [12], forecasts for 12 months were generated with 80% and 95% prediction intervals, highlighting seasonal patterns and trends in future electricity production.

Now try to identify an appropriate ETS model.
The ETS(A,Ad,A) model was chosen, leveraging additive error, dampened trend, and additive seasonality to capture the time series characteristics effectively.

Do residual diagnostic checking of your ETS model. Are the residuals white noise?

Residual diagnostics showed minimal autocorrelation and randomness, confirming that the residuals are close to white noise and the ETS model fits the data well.

Use your chosen ETS model to forecast the 12 periods ahead.
The ETS model forecasted electricity production for the next 12 months, demonstrating consistent seasonality and trend alignment with historical data.

Which of the two models (ETS and ARIMA) do you think is the best approach?
ARIMA performs better for capturing complex seasonal and trend dynamics, while ETS provides interpretable forecasts. For accuracy, ARIMA is superior, but ETS excels in simplicity and interpretability, depending on application needs.

Reference

1. Sharma, S., & Kumar Mishra, S. (2023). Electricity demand estimation using ARIMA forecasting model. In *Recent Developments in Electronics and Communication Systems* (pp. 677-682). IOS Press.
2. M. Shen, "Application of ARIMA and ETS Model in Fund Index Prediction," *2022 IEEE 2nd International Conference on Mobile Networks and Wireless Communications (ICMNWC)*, Tumkur, Karnataka, India, 2022, pp. 1-4,
3. Mel Keytingan M. Shapi, Nor Azuana Ramli, Lilik J. Awalin, Energy consumption prediction by using machine learning for smart building: Case study in Malaysia, Developments in the Built Environment, Volume 5, 2021, 100037, ISSN 2666-1659,
4. Kwangbok Jeong, Choongwan Koo, Taehoon Hong,"An estimation model for determining the annual energy cost budget in educational facilities using SARIMA (seasonal autoregressive integrated moving average) and ANN (artificial neural network)", Energy, Volume 71, 2014, Pages 71-79, ISSN 0360-5442.

5. Cristian-Dragos Dumitru, Adrian Gligor, Wind Energy Forecasting: A Comparative Study Between a Stochastic Model (ARIMA) and a Model Based on Neural Network (FFANN), Procedia Manufacturing, Volume 32, 2019, Pages 410-417, ISSN 2351-9789.
6. Box, G. E., & Cox, D. R. (1964). An analysis of transformations. Journal of the Royal Statistical Society: Series B (Methodological), 26(2), 211-243. (For discussing the Box-Cox transformation method)
7. Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: Principles and Practice (2nd ed.). OTexts. (A comprehensive guide to time series analysis and forecasting techniques, including ARIMA, ETS, and decomposition methods)
8. Chatfield, C. (2003). The Analysis of Time Series: An Introduction (6th ed.). CRC Press. (For foundational knowledge on time series decomposition and modeling)
9. Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). STL: A seasonal-trend decomposition procedure based on loess. Journal of Official Statistics, 6(1), 3-33. (For discussing the STL decomposition method used in the analysis)