**U.S. Electric Grid: Will it power the future?**

Developing Analytics Applications in R

RESEARCH PROJECT

By

**Ben Mayes**

M00019554

**CIS 663**

Summer 2025

Master of Science in Information Systems (MSIS) Program

# **Abstract**

Electricity is an essential component of modern-day life, both for comfort and industrialization. Americans consume (‘demand’) trillions of megawatt hours of electricity every year. Research suggests that the demand for electricity will increase over the next several decades. It is also suggested that electricity production (‘generation’) will need to increase in order to keep pace. This study asks the question, “Will forecasted electricity generation equal forecasted demand”? Monthly generation and demand data from the nation’s electricity providers were collected for the past 25 years. Using R program, time series models and forecasts were produced to predict both variables in the future. The sample means of both forecasts were compared to reveal that there is no statistical difference in the projected demand for and generation of electricity for the next five years. This study is limited as to the variables examined, and further research is required to gather a fuller scope of the issue.

# **Introduction**

This study compares the forecasted mean monthly electricity demand to the forecasted mean monthly electricity generation in the United States. Over the last two decades, electricity consumption has remained relatively flat, until most recently (Schipper & Hodge, 2025). Over the same period, electricity generation has seen a similar trend (Vailshery, 2025). Given the recent uptick in electricity demand and the expectation that these demands will continue to increase (Schipper & Hodge, 2025) there is concern that electricity generation will not be sufficient in coming years (Walton, 2024).

The demand for electricity in the United States hit an all-time high at over 4,170 Terawatt Hours consumed in 2024 (U.S National Power Demand Study, 2025). A rather flat growth period in electricity demand the country has seen for the past several decades appears to be over. A recent study by Grid Strategies suggests the nationwide electric demand is forecasted to increase by nearly 16% by 2029 (Wilson, Zimmerman, Gramlich, 2024). In recent decades, the country has built limited amounts of high-voltage power lines connecting different grid regions (Popovich, Plumer, 2023). This lack of infrastructure growth could become evident in years to follow.

While a high percentage of this growth is attributed to new data centers going online, the effects could be felt from industries to households. Electricity from the nation’s grid powers everyday life for over 350 million Americans. As nearly as vital as clean water, electricity is more than a commodity; it is a necessity. Hospitals, office buildings, industry and homes rely on power from the grid to sustain their lives and livelihood. The effects of a shortage in electricity in the US has the potential to nearly cripple a nation.

Using monthly data collected from the nation’s electricity providers and time series forecasting models, we attempt to forecast electricity demand and generation over the next five years and compare the results.

Forecasting electricity demand is a complex task involving many system variables (Energy projections). Large industrial growth, the housing market and other large-scale variables play a key role in forecasting demand. Likewise, for generation there are many aspects that can affect future models, such as infrastructure construction and federal regulations. It is important to note that this study only focuses on historical demand and generation data collected over the last 25 years.

# **Literature Review**

In January of 2025, the federal government declared a national emergency concerning energy (Declaring a National Emergency, 2025). Over the past two years, the 5-year load growth forecast has increased by almost a factor of five, from 23 Gigawatts to 128 Gigawatts (Wilson, Zimmerman, Gramlich, 2024). Electricity demand in the United States is expected to increase by 2% each year and by 50% by the year 2050 (Walton, 2025). Much of this growth is a result of new data centers and their high electricity demand (Wilson, Zimmerman, Gramlich, 2024).

In 2024, the net capacity of the electric sector was an estimated 1.2 terawatts. This number is expected to increase by more than 97% by 2050 (Electric power sector capacity, 2025). Grid expansion is critical to sustaining high demand growth. While electricity generation capacity is projected to double over the next 15 years, it still may not be sufficient to meet the rising demand (U.S. National Power Demand Study, 2025).

# **Research Question**

In this project, the following research question will be addressed:

* Does forecasted electricity generation equal forecasted electricity demand?

# **Theory**

The literature suggests a large spike in energy consumption in the years to follow. However, there is evidence to imply the nation’s power providers are on pace to meet this demand with new generation. The hypothesis is built on the belief that new electricity generation will meet the expected demand.

H1: Forecasted electricity demand is not statistically different from forecasted electricity generation.

# **Data**

The dataset was collected from Ember, an independent and open data source of electricity data from around the world. The collection includes monthly electricity generation and demand totals from January 2010 to December 2024. These monthly totals were taken directly from the U.S. Energy Information Administration (EIA) and include data from all 50 states. The EIA was established in 1978 and collects independent and impartial energy information from around the country. Both electricity generation and demand are measured in terawatt hours.

The initial dataset download included electricity data from over 100 countries. The categories included country, date of collection, collection type (generation and demand), unit of measurement and value. The dataset was filtered to only include the United States and monthly totals of electricity generation and demand from 2010 to 2024. Two new csv files were created, one for electricity generation by month in the United States and the other for electricity demand by month in the United States.

Data source: <https://storage.googleapis.com/emb-prod-bkt-publicdata/public-downloads/monthly_full_release_long_format.csv>

# **Methodology**

R program was used to conduct data analysis, using RStudio. The required libraires were installed and loaded in RStudio. Both csv files were loaded into R and the date columns for each file were formatted using as.Date(). The next step was to create a time series model based on generation and demand starting in 2010 and reported monthly. A plot diagram showing both generation and demand was generated. Then an Augmented Dickey-Fuller Test (ADF) was used to check the stationarity of both time series. Both were found to be stationary. Next, the data in both files were split for training and testing with 80% of the data used for training the model. A summary of both training and testing data was produced. An ARIMA model was developed using the auto.arima() function on the trained datasets. The auto function enables R to find the best fitting ARIMA model using a stepwise search algorithm. The summary and residuals of the ARIMA models were generated and recorded. Next, a forecast for both generation and demand was created using the forecast() function from the ARIMA models. The summary and box plot were created and recorded.

With the forecasts for both generation and demand created, a series of T-test were performed. The first T-test performed was to see if the forecasted mean value of demand was statistically different than the mean of the partitioned demand test values. The same test was performed on the forecasted mean value of generation and the partitioned generation test values. The final T-test performed to check if the forecasted mean value of generation was statistically different than the forecasted mean value of demand.

# **Results**

Any metric that is measured over regular time intervals forms a time series (Prabhakaran). A time series analysis in R breaks down components so they can be modeled and forecasted (Prabhakaran). The first step in modeling the generation and demand data was to create a time series analysis for both. The summary for both time series models are shown in Table 1 and graphed in Figure 1.



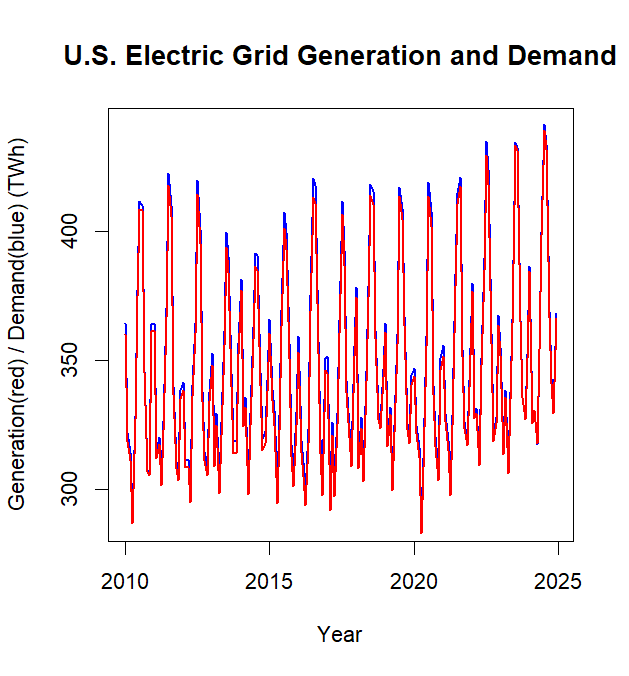


Figure 1 Time Series Plot of Generation and Demand

Stationarity refers to a dataset where the fundamental statistical characteristics of a time series — mean, variance, and autocorrelation – remain stable over time (Tate, 2023). An Augmented Dickey-Fuller Test (ADF) was performed on both time series models. The results for each revealed that the models were stationary, with a p-value < 0.05.

In time series analysis, training and testing are crucial for evaluating model performance. This process involves dividing the data into two sets: a training set used to train the model and a separate testing set to assess its ability to generalize unseen data. This split ensures the model is evaluated on data it hasn't seen before, providing a more realistic estimate of its predictive power in real-world scenarios. Our data was split into 80% train and 20% test. The summary of the test/train data for both models are shown in Table 2.



ARIMA (autoregressive integrated moving average) is a commonly used technique utilized to fit time series data and forecasting (Khan, 2017). This model is a combination of both auto regression and moving averages. In this analysis we used the auto.arima() function to find the best fitting ARIMA model using a stepwise search algorithm. After fitting the ARIMA models a L Jung Box test was performed on both. The results are shown in Table 3, and suggest that the autocorrelations are not significantly different from zero (p-value > 0.05), meaning there is no autocorrelation detected.



Using the fitted models, we created a forecast for both generation and demand for the next five years. The results are shown in Figure 2 and Figure 3. Dark blue indicates the trained data in the models. Red shows the testing data of the model. And light blue indicates the forecasts.

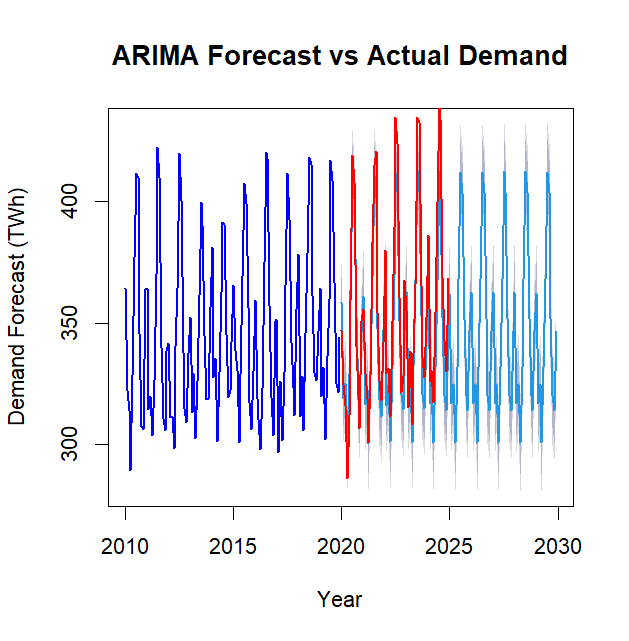


Figure 2 ARIMA Forecast of Demand

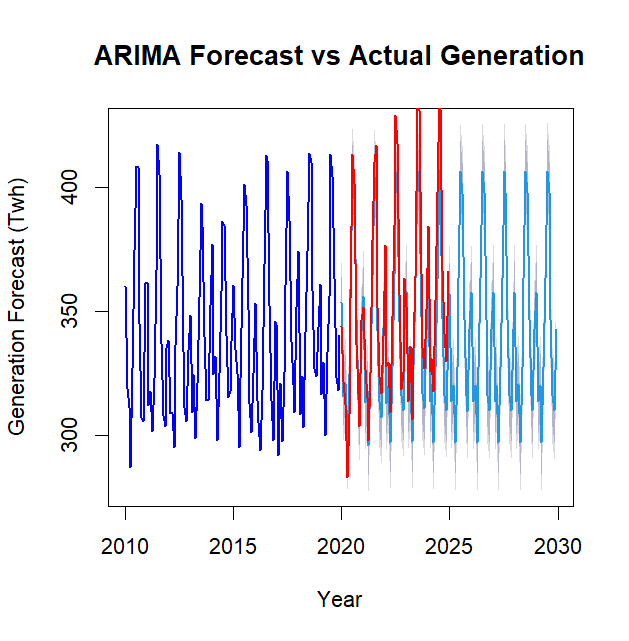


Figure 3 ARIMA Forecast of Generation

The two-sample T-test is a method to determine if there is a statistical difference between two means of two independent samples. In this study we performed three two-sample T-test on the data generated from the previous models. The first was to test the forecasted mean of demand against the mean test data for demand. The results indicate there is not enough evidence to reject the null hypothesis (HO: The forecasted demand values are not significantly different from the actual values), since t = -1.4638, d.f. = 103.06 and p-value = 0.1463 (>0.05).

The next two-sample T-test was performed on the forecasted mean of generation against the mean test data of generation. The results indicate there is not enough evidence to reject the null hypothesis (HO: The forecasted generation values are not significantly different from the actual values), since t = -1.7983, d.f. = 102.57 and p-value = 0.07507 (>0.05). These two test were only performed to verify the forecasted data was not statistically different from the test data.

The final two-sample T-test will test our original null hypothesis from our theory (H1: Forecasted electricity demand is not statistically different from forecasted electricity generation). This test compares the forecasted mean of demand to the forecasted mean of generation. The results indicate that there is not enough evidence to reject the null hypothesis since, t = 1.0512, d.f = 237.94, and p-value = 0.2942 (>0.05). The forecasted mean value of demand is 345.8819 and the forecasted mean of generation is 341.3208.

# **Implications**

While the results of this study suggest that there is not a statistically significant difference between electricity generation and demand within the next five years, more in-depth studies are required. As stated previously, this study only takes into account historical electricity data from the past 25 years. Many other variables can impact the demand for electricity and the ability to generate electricity. Future work could involve exploring more complex models, such as VAR or machine learning approaches, to capture non-linear relationships and improve forecasting accuracy. Additionally, incorporating external factors such as weather data, economic indicators, and policy changes could enhance the model's predictive power. Further analysis of regional differences in demand and generation patterns could also provide valuable insights for localized grid management strategies. This study is meant to be coupled with future research on the topic of electricity generation and demand.

# **Conclusion**

R is a robust program well suited to generate models and complete statistical analysis. By utilizing R programming and the various tools available we were able to answer our research question, “Does forecasted electricity generation equal forecasted demand?” A key component in the research was developing a forecast model. Using time series analysis, ARIMA model, and forecasting in R we were able to generate forecasted models for both electricity generation and demand. Using the embedded two-sample T-test in R, we were then able to test these two forecasted models against one another to check for statistical differences. The available resources in R enabled the research to be completed within one application.

# **References**

*DECLARING A NATIONAL ENERGY EMERGENCY*. (2025, January 20).  White House. Retrieved June 3, 2025, from <https://www.whitehouse.gov/presidential-actions/2025/01/declaring-a-national-energy-emergency/>

*Electric power sector capacity in the United States in 2024, with a forecast to 2050*. (2025, May 19).  Statistica. Retrieved June 2, 2025, from <https://www.statista.com/statistics/504369/electric-power-sector-capacity-in-the-us-forecast/>

*Energy projections*. (n.d.). IAEA. Retrieved June 1, 2025, from <https://www.iaea.org/topics/energy-projections>

Khan, R. (2017, December 13). *ARIMA model for forecasting– Example in R*. RPubs. Retrieved June 3, 2025, from <https://rpubs.com/riazakhan94/arima_with_example>

Popovich, N., Plumer, B. (2023). Why the U.S Electric Grid Isn’t Ready for the Energy Transition. NY Times. Retrieved April 2, 2025 from <https://www.nytimes.com/interactive/2023/06/12/climate/us-electric-grid-energy-transition.html>

Prabhakaran, S. (n.d.). *Time Series Analysis*. R Statistics. Retrieved June 2, 2025, from <https://r-statistics.co/Time-Series-Analysis-With-R.html>

Schipper, M. S., & Hodge, T. H. (2025, May 13). After more than a decade of little change, U.S. Electricity consumption is rising again. Retrieved June 1, 2025, from <https://www.eia.gov/todayinenergy/detail.php?id=65264>

Tate, A. (2023, December 1). *Understanding the Importance of Stationarity in Time Series*. Hex. Retrieved June 4, 2025, from <https://hex.tech/blog/stationarity-in-time-series/>

U.S National Power Demand Study. (2025). Retrieved April 2, 2025 from <https://cleanpower.org/wp-content/uploads/gateway/2025/03/US_National_Power_Demand_Study_2025_FINAL.pdf>

Vailshery, L. (2025, May 14). *Net electricity generation in the United States from 1990 to 2024, by energy source*. Statista. Retrieved June 3, 2025, from <https://www.statista.com/statistics/220174/total-us-electricity-net-generation-by-fuel/>

Walton, R. (2024, December 6). *Five-year US load growth forecast surges 456%, to 128 GW: Grid Strategies*. Utility Dive. Retrieved June 1, 2025, from <https://www.utilitydive.com/news/shocking-forecast-us-electricity-load-could-grow-128-gw-over-next-5-years-Grid-Strategies/734820/>

Walton, R. (2025, April 7). *US electricity demand will grow 50% by 2050, electrical manufacturer study finds*. Utility Dive. Retrieved June 2, 2025, from <https://www.utilitydive.com/news/us-electricity-demand-will-grow-50-by-2050-electrical-manufacturer-study/744575/>

Wilson, J., Zimmerman, Z., Gramlich, R. (2024). Strategic Industries Surging: Driving US Power Demand. Grid Strategies. Retrieve April 5, 2025 from <https://gridstrategiesllc.com/wp-content/uploads/National-Load-Growth-Report-2024.pdf>