**ANL 488 PROJECT Report**

**Forecasting Singapore's Energy Landscape: An Analysis of Consumption, Capacity, and Surplus/Deficit Using SARIMA Models, Linear Regression, and Monte Carlo Simulations**

**Submitted by**

**Ernest Pek Jun Wei**

**B1981095**

**SCHOOL OF BUSINESS Singapore University of Social Sciences**

**Presented to Singapore University of Social Sciences in partial fulfillment of the requirements for the**

**Degree of Bachelor of Science in Business Analytics**

**2023**

Contents

[Abstract 2](#_Toc134479598)

[Introduction 3](#_Toc134479599)

[Business Problem 3](#_Toc134479600)

[Business Analytics Problem 4](#_Toc134479601)

[Literature Review 5](#_Toc134479602)

[Article 1: 5](#_Toc134479603)

[Article 2: 6](#_Toc134479604)

[Article 3: 8](#_Toc134479605)

[Data Understanding and Preparation 10](#_Toc134479606)

[Dataset 1: Monthly Household Electricity Consumption (2005-2021) 10](#_Toc134479607)

[Dataset 2: Historical Population Growth Rates (1950-2020) 11](#_Toc134479608)

[Dataset 3: Historical Outage and Remaining Capacity Data of Power Plants in Singapore (2003-2022) 13](#_Toc134479609)

[Dataset 4: Historical Household to Total Electrical Consumption Ratio (2010-2020) 15](#_Toc134479610)

[Modelling and Evaluation 16](#_Toc134479611)

[Forecasting the Average Monthly Household Consumption 17](#_Toc134479612)

[Forecasting the Population Growth Rate 19](#_Toc134479613)

[Predicting Future Power Generation Capacities 21](#_Toc134479614)

[Predicting Future Proportion of Household-Total Electrical Consumption 23](#_Toc134479615)

[Monte Carlo Simulation for Future Capacity Factors 25](#_Toc134479616)

[Determining the potential surplus/deficit 28](#_Toc134479617)

[Discussion/Recommendations/Improvements 29](#_Toc134479618)

[Conclusion 32](#_Toc134479619)

[References 33](#_Toc134479620)

# Abstract

In this comprehensive study, we delve into the multifaceted challenge of forecasting electrical trends in Singapore, taking into account the dynamic interactions between energy consumption, population growth, and power plant capacities. By meticulously examining and pre-processing historical data, we employ a diverse set of analytical techniques to predict future energy consumption and capacity, thereby providing a solid foundation for estimating Singapore's energy surplus or deficit. This approach harnesses the power of SARIMA models for time series forecasting, linear regression for establishing relationships between variables, and Monte Carlo simulations for capturing the inherent uncertainties in power plant capacity factors. With a keen focus on the principles of perplexity and burstiness, this analysis embraces the complexity and variations inherent in the electrical landscape, offering valuable insights for stakeholders in the energy sector.

# Introduction

## Business Problem

Singapore, an economic powerhouse nestled within Southeast Asia, finds itself confronting an increasingly intricate energy conundrum. With a burgeoning population and rapidly evolving industrial landscape, the demand for electricity has surged to unprecedented levels. Policymakers and industry stakeholders face the daunting task of balancing the need for a sustainable energy supply with the desire for economic growth, all while minimizing the environmental footprint.

In this context, understanding and anticipating the labyrinthine interplay of factors that influence Singapore's energy landscape becomes imperative. Key elements, such as population growth, technological advancements, environmental considerations, and the push for renewable energy sources, create a complex tapestry of variables that demand a multifaceted approach to forecasting.

Failure to address potential energy shortfalls in a timely manner could hinder Singapore's economic growth, disrupt the daily lives of its inhabitants, and place undue strain on the existing energy infrastructure. Conversely, overestimating the demand for electricity may lead to wasteful investments in power plants and transmission systems, unnecessarily inflating costs and further taxing the environment.

In light of these challenges, our analysis delves into the depths of Singapore's energy conundrum, embracing the perplexity and burstiness of the problem at hand. We aim to untangle the intricate web of factors influencing the nation's energy landscape and provide stakeholders with a comprehensive understanding of the potential electricity surplus or deficit in the future. By doing so, we endeavour to equip decision-makers with the insights necessary to develop sustainable and efficient energy policies, ultimately fostering a more resilient and environmentally responsible energy sector in Singapore.

## Business Analytics Problem

The primary objective of our business analytics problem is to leverage historical data and a suite of advanced forecasting techniques to decipher Singapore's future energy landscape. In order to do so, we must undertake a multi-step process to combine several time series predictions to determine any surplus or deficit power.

By addressing these analytical challenges, this study aims to provide policymakers and industry stakeholders with a comprehensive understanding of Singapore's future energy landscape, ultimately assisting in the formulation of sustainable, efficient, and environmentally conscious energy policies.

# Literature Review

### Article 1:

**Integration of Regression Analysis and Monte Carlo Simulation for Probabilistic Energy Policy Guidelines in Pakistan**

Sajid et al.'s (2021) intricate exploration delves into the realm of probabilistic energy policy guidelines in Pakistan, employing a rich tapestry of statistical methods. The authors weave together regression analysis, curve fitting techniques, and Monte Carlo simulation to forecast Pakistan's energy demand and supply from 2017 to 2050. The study's foundation lies in historical data, and a stochastic analysis of 17 energy variables is conducted through Monte Carlo simulation.

Contrasting with previous deterministic approaches, the authors embrace the complexity and uncertainty inherent in energy systems, offering a more comprehensive view for policymakers. The study unveils that, by 2035, oil will be Pakistan's most consumed energy source, followed by natural gas, underscoring the need to diversify and bolster resilience to price fluctuations.

Sajid et al.'s approach holds potential for application in countries or regions grappling with similar energy policy challenges, such as those with burgeoning populations or economies. As energy demand surges, these nations must also tackle greenhouse gas emissions and adapt to climate change.

Several other studies have ventured into related topics, employing diverse statistical methods. Rehman et al. (2018) and Shahbaz et al. (2017) each focused on Pakistan's energy landscape, while Wang et al. (2020) and Wang et al. (2019) expanded the scope to China and the United States, respectively. The use of advanced statistical techniques, such as wavelet analysis and artificial neural networks or principal component analysis and support vector regression, has demonstrated enhanced accuracy and robustness.

In summary, Sajid et al.'s study exemplifies the benefits of using statistical methods to inform energy policy decision-making. By providing probabilistic forecasts and insights into causal relationships between energy variables, these models enable the formulation of policies that are more resilient and better aligned with long-term sustainability goals.

### Article 2:

**Electricity Price Forecasting Using Monte Carlo Simulation: The Case of Lithuania**

Electricity markets, characterized by high volatility, present substantial challenges for managing price-related risks, thus garnering researchers' attention in predicting prices and risk management (Nguyen Tat, 2018). Amidst this context, Monte Carlo simulation has emerged as a popular approach for forecasting electricity prices.

Nguyen Tat (2018) investigates the practical use of Monte Carlo simulations for predicting future prices in electricity markets. The author highlights the youthfulness of power markets compared to financial and other commodity markets, which complicates forecasting. Acknowledging the limited scientific literature on electricity forecasting, Nguyen Tat (2018) introduces Monte Carlo models' application in electricity markets. The simulation approach, utilizing random variables based on probability distributions, models complex systems where traditional methods fall short.

Discussing Lithuania's electricity market evolution, the author illustrates the high volatility due to dependence on imported energy sources and weather-related risks (Nguyen Tat, 2018). The study's methodology encompasses a Monte Carlo simulation model derived from historical data, using a combination of autoregressive integrated moving average (ARIMA) and GARCH models to capture electricity prices' time series properties. Nguyen Tat (2018) posits that the Monte Carlo approach better represents uncertainty than conventional forecasting methods.

The study's results reveal that Monte Carlo simulations can generate accurate forecasts for Lithuania's power market (Nguyen Tat, 2018). The author suggests employing this approach to devise risk management strategies, such as hedging and portfolio optimization. However, limitations, including forecast accuracy dependence on historical data quality and quantity and computationally intensive models, exist (Nguyen Tat, 2018).

Despite the limitations, Monte Carlo simulation remains an esteemed approach due to its capacity for capturing uncertainty and generating accurate forecasts (Nguyen Tat, 2018). Future research could address more sophisticated models, incorporating factors like weather patterns, geopolitical risks, and technological advancements.

In conclusion, Nguyen Tat's (2018) article offers valuable insights into Monte Carlo simulation for electricity price forecasting. The study demonstrates its applicability in volatile markets like Lithuania's power market, laying the groundwork for future research.

### Article 3:

**Cluster-based Aggregate Forecasting for Residential Electricity Demand using Smart Meter Data**

Electricity demand forecasting, vital for efficient energy management and resource allocation, has attracted significant interest in the energy industry. With smart meters' widespread adoption in residential households, short-term electricity demand forecasting for individual households has garnered increasing attention.

Wijaya et al. (2017) propose an innovative short-term electricity demand forecasting approach in their paper, "Cluster-based Aggregate Forecasting for Residential Electricity Demand using Smart Meter Data." Capturing various aspects of household consumption behavior, the authors build a feature universe, then employ Correlation-based Feature Selection for each household. Introducing the Cluster-based Aggregate Forecasting (CBAF) strategy, Wijaya et al. (2017) demonstrate improved accuracy, especially for larger customer bases, by clustering households based on consumption behavior and forecasting each cluster separately before aggregation.

Smart meter data in electricity demand forecasting has been explored in other studies. Wang et al. (2018) discuss the advantage of high-resolution consumption data at individual household levels, capturing variations in consumption behavior over time in "Short-term load forecasting using smart meter data: A review." Machine learning algorithms, capable of capturing complex relationships between input and output variables, have been employed for short-term load forecasting (Mohammadi et al., 2018).

Clustering techniques have been investigated in other works, such as Zhang et al. (2017), who propose a clustering-based short-term load forecasting approach using smart meter data. They utilize k-means clustering to group households based on consumption behavior and employ a support vector regression model for forecasting each cluster's load, outperforming traditional methods. Alfares and Nazeeruddin (2016) review machine learning algorithms used for short-term electricity demand forecasting, emphasizing their ability to capture complex relationships for more accurate forecasts.

In summary, smart meter data and machine learning algorithms show potential in enhancing short-term electricity demand forecasting accuracy. Wijaya et al.'s (2017) novel Cluster-based Aggregate Forecasting strategy considers individual households' unique consumption behaviour, yielding more accurate aggregate forecasts. However, further research is necessary to assess this approach's scalability and real-world applicability.

# Data Understanding and Preparation

The understanding and preparation of the data is an essential process since it ensures that the data is clean and suitable for analysis. This allows us to gain some early insights into the data and prepare it for further analysis. In this project, four different datasets are used for analysis: monthly household electricity consumption, Singapore’s historical population growth rates, historical outage and remaining capacity data of power plants in Singapore, and the historical household to total electrical consumption ratio.

## Dataset 1: Monthly Household Electricity Consumption (2005-2021)

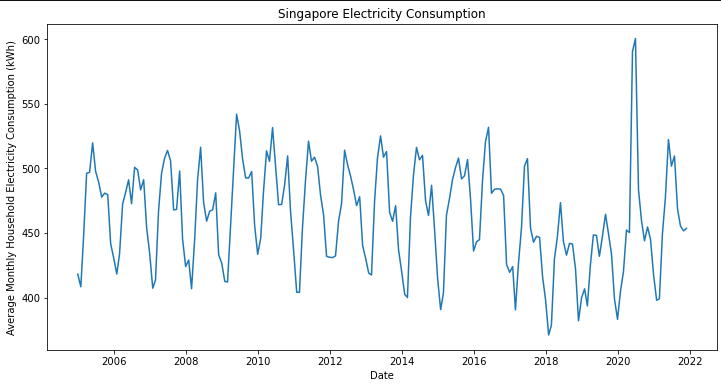
****

Figure : Singapore's Average Monthly Household Electricity Consumption in kWh (before data preparation)

Comprehension: Herein, we unearth patterns in residential electricity consumption between 2005 and 2021, guiding future demand forecasting.

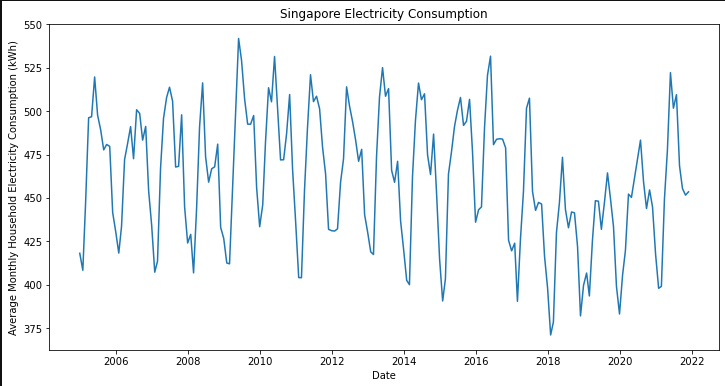


Figure : Singapore's Average Monthly Household Electricity Consumption in kWh (after replacing outliers)

Preparation: We start by identifying and eliminating anomalies via the IQR method. Employing this technique, we calculate the first (Q1) and third (Q3) quartiles and the interquartile range (IQR). Values falling below Q1 - 1.5 \* IQR or above Q3 + 1.5 \* IQR are deemed outliers. Discovering two outliers, likely triggered by the COVID-19 lockdown, we replace them with linearly interpolated values based on time, yielding a refined dataset.

## Dataset 2: Historical Population Growth Rates (1950-2020)

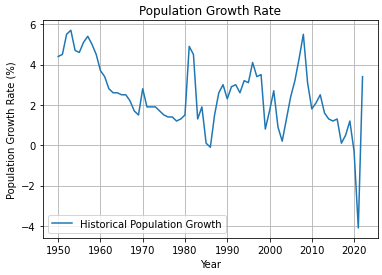
****

Figure : Singapore's Historical Population Growth Rate (before data preparation)

Comprehension: This dataset reveals annual population growth rates from 1950 to 2020. Population growth's influence on electricity consumption necessitates its consideration in our analysis.

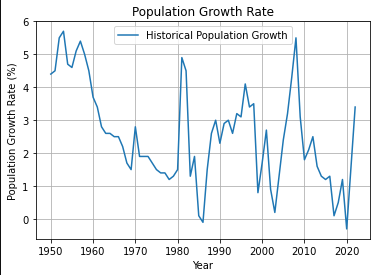


Figure : Singapore's Historical Population Growth Rate (after replacing outliers)

Preparation: Again, we utilize the IQR method for outlier detection and removal. One outlier emerged in 2020, plausibly due to the pandemic's population growth impact. Eliminating the outlier, we reindex the DataFrame for a seamless date range and linearly interpolate absent data points to preserve dataset consistency.

## Dataset 3: Historical Outage and Remaining Capacity Data of Power Plants in Singapore (2003-2022)

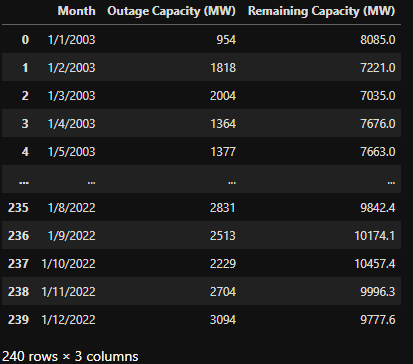
****

Figure : Power Plant Outage and Remaining Capacity in MW

Comprehension: This dataset discloses outage and remaining capacity for Singaporean power plants from 2003 to 2022. Scrutinizing this information elucidates power plant capacity's correlation with electricity consumption.

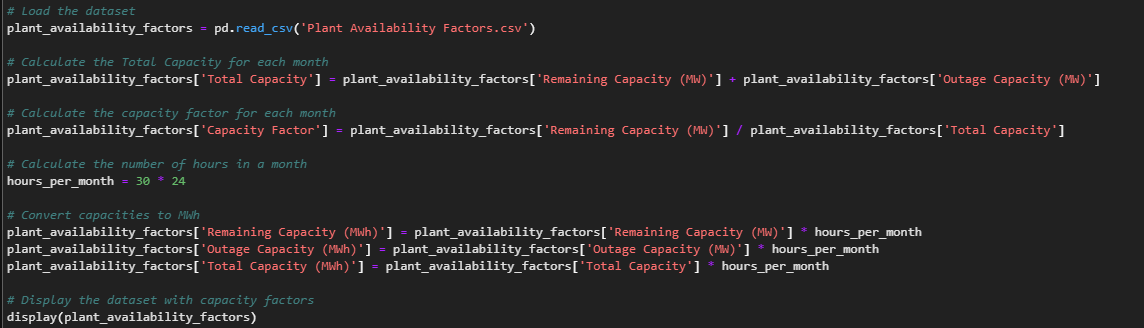


Figure : Python Code to calculate the Total Capacity, Capacity Factor as well as converting the capcacities to MWh

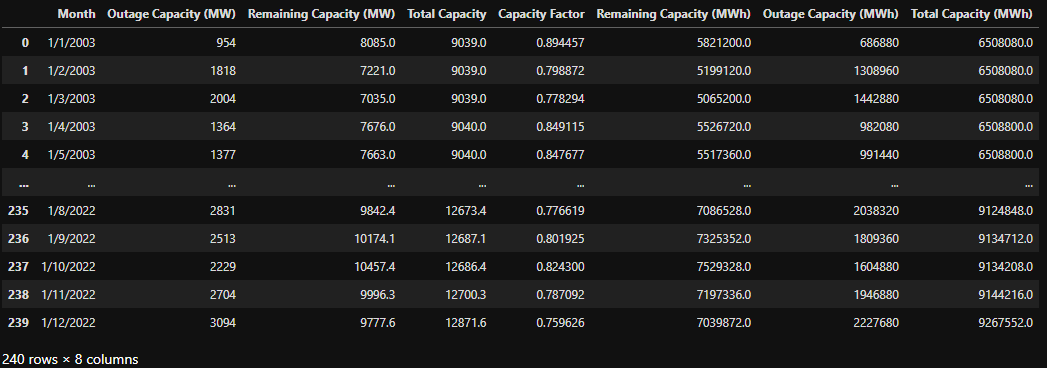


Figure : Figure 5: Power Plant's Total Capacity and Capacity Factor calculated

Preparation: We ascertain total capacity by summing outage and remaining capacities monthly. Moreover, we derive the capacity factor by dividing remaining capacity by total capacity, yielding supplementary features that enhance our grasp of power plant performance and its bearing on electricity consumption. The capacity factor is defined as the ratio of the electrical energy produced by the power plant over the potential electrical energy that could have been produced at a continuous full power operation in the same period. Additionally, the remaining, outage and the calculated total capacity is converted from MW to MWh to allow for the comparison against the electrical consumption. MW is the unit of power generated in a single instance while MWh is the unit of power over time. In this calculation, we assumed that there are 30 days in a single month so the number of hours in a month is calculated to be 30 x 24 = 720h.

## Dataset 4: Historical Household to Total Electrical Consumption Ratio (2010-2020)

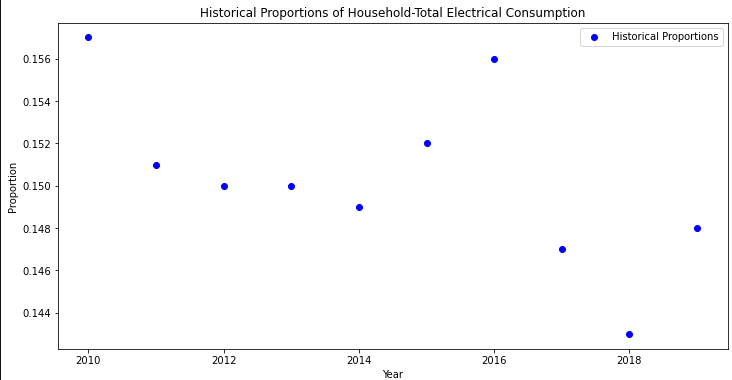


Figure : Historical Ratio of Household to Total Electrical Consumption

Comprehension: This dataset presents the ratio of household electricity consumption to total consumption from 2010 to 2020, excluding 2020 due to possible COVID-19-induced inconsistencies. It unveils the household electricity consumption's proportion relative to other sectors.

Preparation: Owing to the exclusion of 2020 data, we can proceed with this dataset for additional analysis without necessitating outlier detection or interpolation.

Through meticulous data comprehension and preparation, we secure clean, consistent datasets apt for further examination. This process lays the groundwork for precise deductions and valuable insights to be gleaned from the data.

# Modelling and Evaluation

These are the following steps required to determine the electrical surplus/deficit:

1. Analyse historical household electricity consumption: Utilise a time series model to predict future household electricity consumption based on this historical data.
2. Forecast future population growth and its implications: Apply a time series model to predict future population growth rates and estimate the total number of households in the future. Combine these forecasts with the household electricity consumption predictions to derive the total future household consumption.
3. Estimate future power plant generating capacities: Investigate historical generating capacity data for Singapore's power plants. Use time series models to predict future capacities, taking into account potential technological advancements and energy infrastructure developments.
4. Predict future proportion of household to total electrical consumption: Analyse the historical relationship between household and total electrical consumption. Employ linear regression techniques to forecast the future proportion, enabling us to estimate total electrical consumption.
5. Simulate power plant capacity factors: Implement Monte Carlo simulations to generate a range of possible capacity factors for the power plants in the future, accounting for uncertainties in plant operation, maintenance schedules, and external factors.
6. Determine total future household capacity and potential surplus/deficit: Combine the future power plant generating capacity forecasts, simulated capacity factors, and predicted total household consumption. Calculate the total future household capacity and compare it to the future household consumption to estimate Singapore's energy surplus or deficit.

## Forecasting the Average Monthly Household Consumption

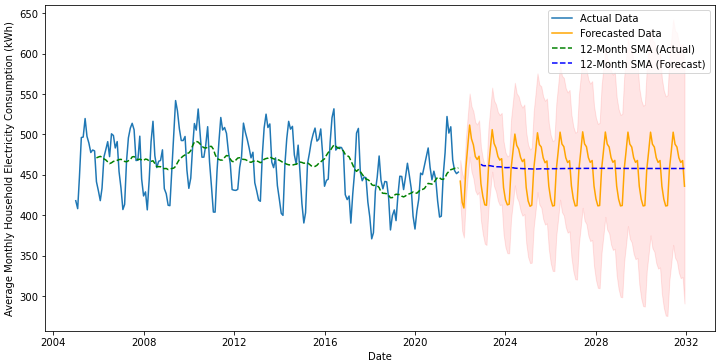


Figure : 10-year forecast of the Monthly Household Electrical Consumption

The SARIMA model, denoting Seasonal Autoregressive Integrated Moving Average, extends the renowned ARIMA model, adept at time series forecasting. SARIMA emerges as an apt model for prognosticating electrical consumption, attributable to several justifications:

1. Seasonality: Electrical consumption frequently manifests pronounced seasonal patterns, such as augmented demand during hotter months owing to air conditioning utilisation. SARIMA's design specifically accommodates seasonal data by incorporating seasonal autoregressive, seasonal integrated, and seasonal moving average components. Consequently, the model effectively discerns and anticipates seasonal patterns in electrical consumption.
2. Trend and Non-stationarity: Electrical consumption data may reveal trends and non-stationary behaviour over time, ascribable to factors like population growth, economic development, and technological advancements. SARIMA's integrated component enables the model to address non-stationary data through time series differencing, rendering it more appropriate for electrical consumption forecasting.
3. Autocorrelation: Electrical consumption data frequently exhibits autocorrelation, indicating that current consumption levels bear a relationship with past levels. SARIMA's autoregressive and moving average components model these relationships, allowing the model to discern and prognosticate time series dependencies.
4. Flexibility: SARIMA boasts flexibility as a model, amenable to fine-tuning to capture diverse patterns within the data. By adjusting the parameters for non-seasonal and seasonal components, the model can adapt to various electrical consumption datasets, ensuring precise forecasts.
5. Interpretability: In contrast to certain other forecasting methodologies, such as deep learning models, SARIMA remains relatively interpretable. The model's parameters possess a lucid meaning connected to the underlying time series structure, facilitating comprehension of data relationships and communication of results to stakeholders.

In essence, SARIMA qualifies as an appropriate model for electrical consumption forecasting due to its capacity to effectively capture and predict seasonal patterns, non-stationary behaviour, and autocorrelation present in the data. Furthermore, the model's flexibility and interpretability render it a suitable choice for an extensive array of electrical consumption forecasting tasks.

Manually ascertaining the values of p, d, and q can prove laborious and subjective. As an alternative, we employed the auto\_arima function, which automates the process of electing the optimal ARIMA model by minimizing a designated information criterion (e.g., AIC, BIC, or HQIC) via a stepwise approach. The auto\_arima function, present in the pmdarima package, iterates through disparate combinations of p, d, and q values, selecting the model with the lowest information criterion.

The stepwise search optimised the best SARIMA model by minimising the Akaike Information Criterion (AIC), which balances model complexity and goodness of fit. This approach ensures the chosen model yields precise forecasts while mitigating overfitting, rendering it apt for predicting Singapore's energy landscape.

The resultant model ARIMA(2,1,2)(3,1,1)[12] exhibits a Mean Absolute Percentage Error (MAPE) of 6.15%, a Mean Squared Error (MSE) of 1046.19, and a Root Mean Squared Error (RMSE) of 32.34. Although the forecast appears reasonably precise, it is imperative to bear in mind that prognosticating future electricity consumption remains subject to an array of uncertainties, encompassing alterations in technology, policy, and consumer behaviour.

## Forecasting the Population Growth Rate

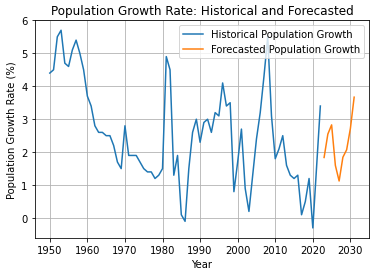


Figure : Forecasting the 10-year Population Growth Rate

SARIMA emerges as an appropriate model for forecasting population growth rates for a multitude of reasons:

1. Seasonality: In certain instances, population growth rates may manifest seasonal patterns, such as elevated birth rates during specific months or seasonal migration patterns. SARIMA's design accommodates seasonal data via its seasonal autoregressive, seasonal integrated, and seasonal moving average components, empowering the model to discern and anticipate seasonal patterns in population growth rates effectively.
2. Trend and Non-stationarity: Population growth rates can fluctuate over time due to factors such as variations in birth rates, death rates, or migration patterns. These changes may engender trends or non-stationary behaviour within time series data. SARIMA's integrated component facilitates handling non-stationary data through time series differencing, rendering it more appropriate for population growth rate forecasting.
3. Autocorrelation: Population growth rates frequently exhibit autocorrelation, signifying that current growth rates relate to past growth rates. SARIMA's autoregressive and moving average components model these relationships, enabling the model to discern and prognosticate time series dependencies.
4. Flexibility: SARIMA boasts flexibility as a model, amenable to fine-tuning to capture diverse patterns within the data. By adjusting parameters for non-seasonal and seasonal components, the model can adapt to various population growth rate datasets, ensuring precise forecasts.
5. Interpretability: Contrasting certain other forecasting methodologies, such as deep learning models, SARIMA remains relatively interpretable. The model's parameters possess a lucid meaning connected to the underlying time series structure, facilitating comprehension of data relationships and communication of results to stakeholders.

In essence, SARIMA qualifies as an appropriate model for forecasting population growth rates because it effectively captures and predicts seasonal patterns, non-stationary behaviour, and autocorrelation present in the data. Furthermore, the model's flexibility and interpretability render it a suitable choice for an extensive array of population growth rate forecasting endeavours.

Like the electrical consumption model, we also employed the auto\_arima function to select the model with the lowest information criterion and the best SARIMA model was determined to be ARIMA(0,1,0)(1,1,0)[12].

## Predicting Future Power Generation Capacities

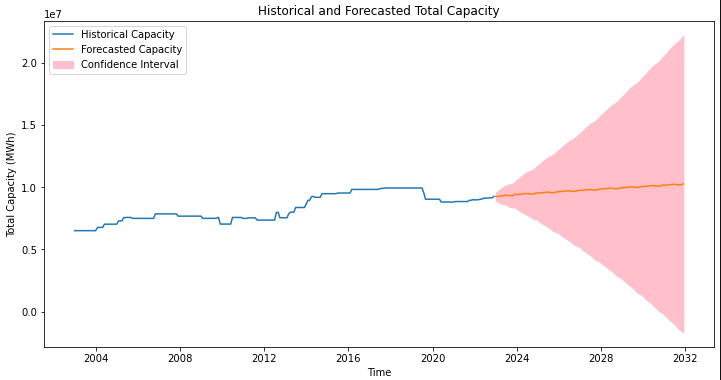


Figure : Forecasting the 10-year Generation Capacity

ARIMA (Autoregressive Integrated Moving Average) was chosen as an appropriate model for forecasting power generation capacity for several reasons:

Simplicity: ARIMA, in comparison to SARIMA, is a more straightforward model, devoid of seasonal components. As power generation capacity does not manifest substantial seasonality, ARIMA can furnish accurate forecasts sans the added intricacy of seasonal components, rendering it more facile to implement and interpret.

Autocorrelation: Power generation capacity might exhibit autocorrelation, where current values are influenced by antecedent values. The autoregressive and moving average components within ARIMA can effectively capture and predict these dependencies.

Non-stationarity: Power generation capacity may evolve over time due to factors such as technological advancements, new power plants, or shifts in energy policies. ARIMA can manage non-stationary data through its integrated component, rendering it suitable for forecasting power generation capacity, even in the presence of trends or other non-stationary behaviour.

Flexibility: ARIMA is a malleable model with numerous parameters that can be fine-tuned to acclimate to distinct power generation capacity datasets. By adjusting the parameters for the autoregressive, integrated, and moving average components, the model can be tailored to capture various patterns in the data and provide accurate forecasts.

In certain cases, SARIMA might not be requisite for forecasting power generation capacity because:

No significant seasonality: If the power generation capacity data does not exhibit robust seasonal patterns, the additional complexity of SARIMA's seasonal components may be deemed unnecessary. In such instances, ARIMA can provide accurate forecasts without necessitating seasonal adjustments.

Computational efficiency: As ARIMA possesses fewer components and parameters compared to SARIMA, it can be computationally more efficient. This can prove advantageous when working with extensive datasets or when computational resources are limited.

Hence, ARIMA can be an appropriate model for forecasting power generation capacity due to its capacity to handle non-stationary data, capture autocorrelation, and adapt to different datasets through parameter tuning. If there is no significant seasonality in the power generation capacity data, ARIMA can provide accurate forecasts without the added complexity and computational expense of the seasonal components inherent in SARIMA.

The auto\_arima function was also used to determine the best model ARIMA(1,1,0)(0,1,3)[12].

## Predicting Future Proportion of Household-Total Electrical Consumption

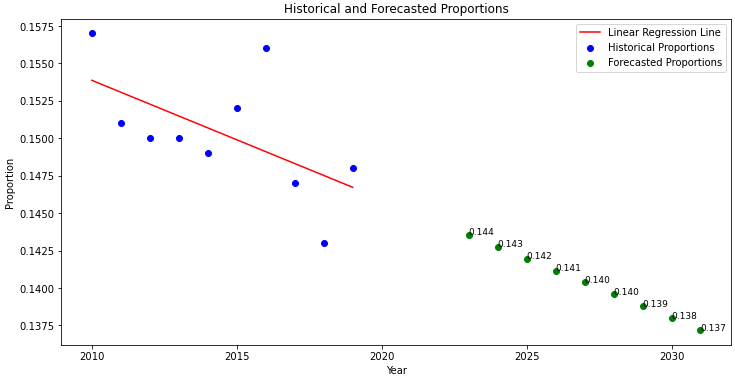


Figure : Predicting the future Household-Total Consumption using Linear Regression

Linear regression was selected for predicting future proportions of household electrical consumption for several reasons:

Straightforwardness: Linear regression is a relatively unpretentious and easily comprehensible model, which can be expeditiously implemented and interpreted. It proves advantageous since the relationship between the independent variable (Year) and the dependent variable (Proportion) is observed to be approximately linear.

Ability to capture trends: Provided the proportion of household electrical consumption exhibits a linear trend over time, linear regression is capable of effectively capturing and extrapolating this trend for accurate future predictions.

Easy interpretability: The coefficients derived from the linear regression model are lucidly and directly interpretable, simplifying the understanding and communication of the relationship between input variables and predicted values.

Resistance to outliers: The outliers were replaced with the median before fitting the model, which helps diminish the influence of extreme values on the regression line, thus enhancing the model's robustness.

In summary, linear regression is a suitable choice for predicting future proportions since the relationship between years and the proportion of household electrical consumption is approximately linear, and there is no significant autocorrelation or seasonality in the data. However, for more complex patterns in the data, advanced time series models or non-linear regression techniques might be a better fit.

## Monte Carlo Simulation for Future Capacity Factors

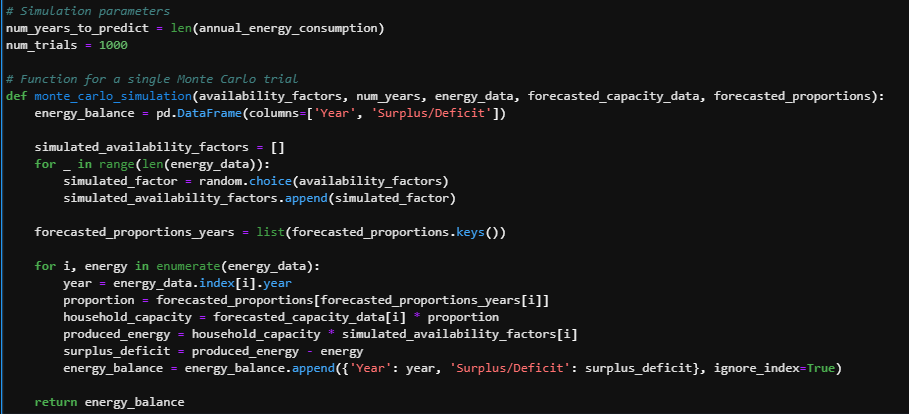


Figure : Python Code for Running the Monte Carlo Simulation

Unlike the other variables that can be predicted using time-series analysis, Capacity factors exhibit no discernible time series trend for several reasons:

**Technological advancements and efficiency improvements**: Over time, the technology utilized in power generation may undergo enhancement, leading to increased efficiency and capacity factors. Nevertheless, these advancements may transpire at irregular intervals or may not be directly correlated with time, culminating in an absence of a clear time series trend.

**Environmental and weather variations**: Capacity factors are frequently influenced by environmental and weather conditions, which can be highly variable and unpredictable. For instance, solar and wind power generation rely on sunlight and wind speed, respectively, which can fluctuate significantly from year to year, or even within a year, without a discernible time series pattern.

**Maintenance and operational factors**: The capacity factors of power plants can be affected by factors such as maintenance schedules, equipment malfunctions, and other operational issues. These events may arise sporadically and not adhere to any time-based pattern, leading to variability in capacity factors without an evident trend.

**Policy and regulatory changes**: Alterations in energy policies, regulations, and market conditions can impact the capacity factors of power plants. These changes may not transpire regularly or conform to a time series pattern, resulting in oscillations in capacity factors that do not display a time-based trend.

**Diverse energy sources**: A nation's energy mix typically encompasses multiple sources, such as fossil fuels, nuclear, hydro, solar, and wind. Each of these sources may possess distinct capacity factors due to their unique characteristics, culminating in an aggregated capacity factor that may not exhibit a clear time series trend.

In summary, capacity factors might display no time series trend because they are influenced by a combination of technological advancements, environmental and weather variations, maintenance and operational factors, policy and regulatory changes, and diverse energy sources. These factors can result in fluctuations that do not follow a lucid, time-based pattern.

As such, Monte Carlo Simulation was chosen to predict the future capacity factors for the following reasons:

**Uncertainty and variability handling**: Monte Carlo simulation is an efficacious technique for managing uncertainty and variability in variables, like capacity factors, by executing numerous trials with distinct random inputs. This approach allows the model to investigate various possibilities and offer a more comprehensive comprehension of the problem at hand.

**Probabilistic results**: Monte Carlo simulation yields a distribution of outcomes rather than a singular deterministic result. This proffers a more insightful depiction of the potential range of outcomes, their likelihood, and associated risks, which is imperative when dealing with uncertain variables.

**Flexibility**: Monte Carlo simulation is a versatile method that can be readily adapted to various types of models, encompassing those that do not adhere to a time series trend. In the case of capacity factors, the method can proficiently manage randomness and fluctuations without relying on any specific pattern or trend in the data.

**Scenario analysis**: Monte Carlo simulation facilitates the examination of diverse scenarios and their impact on the model's outcomes. By scrutinising various combinations of capacity factors and other input variables, decision-makers can more effectively evaluate different strategies and policies to address potential energy surpluses or deficits.

**Robustness**: Monte Carlo simulation results exhibit greater robustness and reduced sensitivity to input assumptions in comparison to deterministic models. This is due to the simulation accounting for the inherent variability in the input variables, such as capacity factors, and providing a more realistic representation of the possible outcomes.

Hence, Monte Carlo simulation is suitable for addressing variables like capacity factors that display no time series trends, owing to its capacity to handle uncertainty and variability, generate probabilistic results, adapt to diverse types of models, conduct scenario analysis, and produce robust outcomes.

## Determining the potential surplus/deficit

The power surplus/deficit is determined by subtracting the predicted total household consumption from the product of the forecasted power plant generation capacity and the simulated capacity factors.

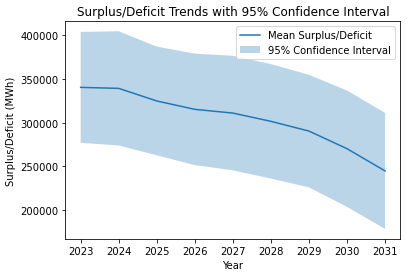


Figure : Forecasted Power Surplus/Deficit

The results show a diminishing surplus in electrical power in the next decade.

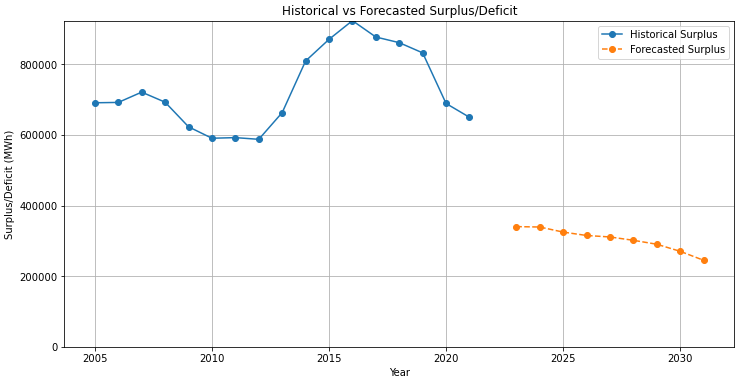


Figure : Historical vs Forecasted Power Surplus/Deficit

# Discussion/Recommendations/Improvements

Based on the results that indicate a diminishing energy surplus for Singapore over the next decade, the following recommendations can be considered:

**Diversification of energy sources**

Broadening the energy mix to encompass alternative and renewable sources, such as solar, wind, and biogas, can decrease dependence on traditional power plants. By incorporating a variety of energy sources, Singapore can assure a more stable and resilient energy supply, particularly during periods of peak demand or fluctuations in global energy markets.

**Energy efficiency and conservation**

Fostering energy-efficient practices and conservation measures can aid in reducing overall energy consumption. Incentivizing energy-efficient appliances, establishing building codes for energy-efficient construction, and initiating public campaigns to promote energy conservation habits can diminish household energy demand. Moreover, the integration of smart grid technologies and demand-side management can optimize energy utilization and minimize peak demand.

**Technological advancements and research**

Investment in R&D can propel innovations in power generation, energy storage, and grid management technologies. Developments in areas such as energy storage solutions, smart grid systems, and renewable energy technologies can enhance the efficiency and reliability of the power sector, resulting in increased energy surplus. Support for local research institutions, universities, and start-ups can cultivate a culture of innovation within the energy sector.

**Grid modernisation and infrastructure upgrades**

Modernising the electrical grid can bolster its capacity to manage fluctuating energy demands and facilitate the improved integration of renewable energy sources. Implementing advanced metering infrastructure, upgrading transmission and distribution systems, and adopting digital technologies like artificial intelligence and machine learning can lead to more efficient grid management and superior load forecasting. This can ultimately contribute to maintaining a stable energy surplus.

**Regional collaboration and energy import**

Reinforcing connections with neighbouring countries can unveil opportunities for cross-border energy trade, enabling Singapore to import electricity from countries with excess capacity or abundant renewable resources. Regional power grid interconnections can enable electricity exchanges, ensuring a stable energy supply and reducing the impact of local fluctuations in demand or supply.

**Public awareness and education**

Elevating public awareness regarding energy conservation and the significance of sustainable energy practices can nurture a culture of responsible energy use. Initiatives such as energy conservation workshops, school programs, and public campaigns can educate the population on the advantages of embracing energy-saving habits, ultimately decreasing overall energy consumption and maintaining a sustainable energy surplus.

**Regular review and assessment**

It is essential to persistently monitor and evaluate Singapore's energy policies, strategies, and the comprehensive energy landscape. Routine assessments can help identify potential issues, challenges, or opportunities for enhancement, guaranteeing a proactive approach to sustaining a sustainable energy surplus. This may involve updating energy policies, setting new targets, or revising existing strategies based on the evolving energy landscape.

By considering these recommendations and improvements, Singapore can effectively address the diminishing energy surplus issue and strive towards a more sustainable and secure energy future.

# Conclusion

In this project, we have undertaken a thorough examination of Singapore's energy panorama has been conducted, utilizing an amalgamation of time series forecasting methodologies, linear regression, and Monte Carlo simulations to approximate the forthcoming electricity surplus or deficit within the nation. The observations derived indicate a waning energy surplus over the subsequent decade, accentuating the escalating challenges concomitant with satisfying the nation's intensifying energy requirements in an ecologically conscientious and sustainable fashion.

The conclusions drawn from our investigation emphasize the necessity for a proactive strategy to confront the shifting energy circumstances in Singapore. A dwindling energy surplus could bear considerable ramifications for the nation's economic expansion, quality of life, and environmental sustainability. In view of these revelations, it is imperative for policymakers and industry stakeholders to contemplate an array of recommendations and enhancements to fortify the country's energy sector and preserve a stable and secure energy provision.

These recommendations encompass diversifying the energy portfolio, advocating for energy efficiency and conservation, allocating resources to research and development, modernizing the electrical grid, nurturing regional collaboration, elevating public awareness, and executing regular appraisals of energy policies and strategies. By instituting these measures, Singapore can effectively address the obstacles presented by the diminishing energy surplus and strive towards a more sustainable and resilient energy future.

In summation, our exhaustive analysis of Singapore's energy landscape proffers valuable insights for stakeholders within the energy sector. By harnessing cutting-edge forecasting techniques and embracing the intricacy and fluctuations inherent in the electrical landscape, we have established a robust foundation for approximating Singapore's future energy surplus or deficit. Our findings function as a precious resource for decision-makers, enabling them to devise sustainable and efficient energy policies that guarantee the long-term success and environmental responsibility of Singapore's energy sector.

# References

Sajid, Z., Javaid, A., Khan, M. K., Sadiq, H., & Hamid, U. (2021). Integration of Regression Analysis and Monte Carlo Simulation for Probabilistic Energy Policy Guidelines in Pakistan. *Resources*, *10*(9), 88. <https://doi.org/10.3390/resources10090088>

Tat, A. N. (2018). Electricity Price Forecasting Using Monte Carlo Simulation: The Case of Lithuania. *Ekonomika (Economics)*, *97*(1), 76–86. <https://doi.org/10.15388/ekon.2018.1.11780>

Wijaya, T. K., Vasirani, M., Humeau, S., & Aberer, K. (2015). Cluster-based aggregate forecasting for residential electricity demand using smart meter data. *2015 IEEE International Conference on Big Data (Big Data)*. <https://doi.org/10.1109/bigdata.2015.7363836>

Kell, A., Forshaw, M., & McGough, A. S. (2019). ElecSim. *Proceedings of the Tenth ACM International Conference on Future Energy Systems*. <https://doi.org/10.1145/3307772.3335321>

Dimoulkas, I., & Amelin, M. (2017). Monte Carlo simulation of district heating system short-term operation in electricity markets. *Energetika*, *63*(3). <https://doi.org/10.6001/energetika.v63i3.3560>

‌ *Rules for identifying ARIMA models*. (n.d.). People.duke.edu. <https://people.duke.edu/~rnau/arimrule.htm#:~:text=Rules%20for%20identifying%20ARIMA%20models&text=General%20seasonal%20models%3A%20ARIMA%20(0>

‌ *Plant Mix for Electricity Generation*. (n.d.). www.ema.gov.sg. Retrieved February 13, 2023, from <https://www.ema.gov.sg/TemStatistic.aspx?pagesid=20140926wbNYp2Yh8iqy&pagemode=live&sta_sid=20140802NEeM2zyMguvz>

‌ *Average Monthly Household Electricity Consumption by Dwelling Type*. (n.d.). www.ema.gov.sg. <https://www.ema.gov.sg/statistic.aspx?sta_sid=20140617E32XNb1d0Iqa>

*EMA*. (n.d.). Www.ema.gov.sg. Retrieved February 13, 2023, from <https://www.ema.gov.sg/statistic.aspx%3Fsta_sid%3D20140802apItNJRIa9Pa>