**ANL 488 FINAL PROJECT REPORT**

**THE RELIABILITY OF RENEWABLE ENERGY TO REPLACE OIL AND GAS AS OUR ENERGY OF CHOICE**



**Submitted by**

**SAM CHENG JIA WEI**

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

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

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

**2023**

# ABSTRACT

Renewable energy has emerged as a promising replacement to the conventional use of fossil fuels such as oil and gas. While renewable energy holds significant promise, it faces challenges regarding its reliability and long-term sustainability. The aim of this study is to understand whether renewable energy can genuinely serve as a reliable and sustainable replacement for oil and gas.

This paper employs a multifaceted approach, utilising time series modelling, including ARIMA and Double Exponential Smoothing, to forecast renewable energy generation and evaluate the cost efficiency of renewable sources. We will examine the trends in electrical generation capacity and the Levelized Cost of Electricity (LCOE) of different energy sources, to understand the correlation and future implications of renewable energy.

The time series forecasting is modelled in Jupyter Notebook using Python3, utilising various datasets, such as Shares of Electricity Generation, Electrical Generations and Levelized Cost of Energies. Our results show that there is a substantial increase in the adoption of renewable energy sources. As renewable energy costs become more competitive and their share of global energy generation grows, they are emerging as an attractive alternative to traditional fossil fuels. Subsequently, we discuss the strategies aimed at achieving a sustainable energy future and potentially replace oil and gas.

**One-Sentence Summary:** Time series forecast analysis of renewable energy to project electrical generation and cost efficiency against traditional fossil fuels.

**Table of Contents**

ABSTRACT i

Chapter 1.0 1

Chapter 1.1 – Introduction 1

Chapter 1.2 – Business Problem 2

Chapter 1.3 – Business Analytics Problem 3

Chapter 2.0 Literature Review 3

Chapter 3.0 Data Understanding and Preparation 9

Chapter 3.1 Data Understanding 9

Chapter 3.2 Data Preparation 11

Chapter 4.0 Proposed Modelling and Evaluation 13

Chapter 4.1 Scope and Methods 13

Chapter 4.2 Modelling with ARIMA 14

Chapter 4.3 Modelling with Double Exponential Smoothing 23

Chapter 5.0 Results 30

Chapter 5.1 Discussions 36

Chapter 6.0 Recommendations and Conclusion 38

Chapter 6.1 Recommendations 39

Chapter 6.2 Conclusion 40

References 41

Appendices 42

Appendix A 42

Appendix B 44

Appendix C 46

Appendix D 47

Appendix E 48

Appendix F 49

Appendix G 57

# Chapter 1.0

## Chapter 1.1 – Introduction

The global energy production stands at a critical juncture, world governments are met with urgent imperatives to shift from fossil fuels to sustainable alternatives. Renewable energy has emerged as a promising replacement to the conventional use of fossil fuels such as oil and gas. Yet, we asked ourselves: can renewable energy truly replace oil and gas as our energy of choice? Is this transition reliable and sustainable?

Renewable energy is a natural source that replenishes itself in a certain time period. It has far lower emissions than burning fossil fuels and reduced negative externalities on the environment.

In light of global warming and the recent Fukushima Daiichi accident, the imperative to reduce climate change, greenhouse gas emissions, and ensuring reliable energy supply has forced governments to invest in renewable energy projects. However, transiting from fossil fuels to renewables are not without its challenges, many renewable energy projects face significant setbacks in achieving their intended targets. This project explores the reliability of renewable energy to replace oil and gas as our energy of choice, shedding light on the challenges and uncertainties that surround renewable energy sources.

Renewable energy does not achieve maximum effectiveness due to a combination of unpredictable weather conditions and suboptimal equipment performance. This has cast doubt over the reliability of renewable energy sources for long term sustainability. For instance, research conducted in China reveals that their wind farm generated a mere 37-45% of its technological potential compared to 54-61% in the United States (Huenteler et al., 2018). The discrepancy in its performance is due to the delays in grid connection and curtailment constraints in grid management. Furthermore, China’s wind farms suffer from suboptimal turbine model selection, poor wind farm siting, and low turbine hub heights (Huenteler et al, 2018). While these factors seem minor individually, their cumulative effect results in the underperformance of China’s wind farms.

Another reason why wind and solar projects fails to meet capacity utilisation target is due to wind generation curtailment and reduced irradiance for solar projects. Wind generation curtailment can occur when electricity supply exceeds demand on the grid, causing the prices go negative. Furthermore, the excess electricity generated threatens to overload the grid’s capacity. Hence, wind and solar farms are curtailed for economic or grid-capacity purposes. This recurring issue led to wind and solar projects consistently falling short on their performance expectations.

The objective of this project is to delve deeper into the reliability of renewable energy sources by conducting a time series forecast and understanding the trend of renewable energy in our society, as well as evaluating the cost efficiency of renewable energy as sustainable alternatives.

## Chapter 1.2 – Business Problem

The business problem at hand is multifaceted, presenting challenges in the energy generation capacity and the cost efficiency of renewable energy.

The first business problem is the lack of energy output from renewable sources. The energy output and its efficiency are primarily driven by uncontrollable factors such as climate change, wind speed, water current and the Earth’s mantle. Additionally, the capacity of renewable energy is further constrained by the limitation of available land and the geographical location of renewable energy plants. These pose significant challenges to ensure that consistent and reliable energy are generated from renewable energy sources.

The second business problem revolves around the high cost of renewable energy production. The cost of renewable energy production encompasses its initial installation cost, operations and maintenance (O&M), and the levelized cost of electricity (LCOE). Balancing these costs factors and keeping cost of production low, to remain economically competitive with conventional energy sources is a business problem that requires data analytics.

## Chapter 1.3 – Business Analytics Problem

The business analytics problem is to determine the most appropriate time series modelling to effectively forecast the growth and energy output of renewable energy sources. This entails a thorough model selection and fitting for the project. This proposal aims to establish a modelling technique that can reliably project future trend in renewable energy production. Furthermore, the model seeks to identify outliers within the dataset, ensuring a more accurate understanding of the implementation of renewable energy sources.

# Chapter 2.0 Literature Review

Abolhosseini, Heshmati, and Altmann (2014), stresses the importance of renewable energy sources in mitigating climate change. The authors argue that renewable energy technologies have the potential to reduce carbon dioxide (CO2) emissions significantly by replacing fossil fuels in both the power generation industry and transportation sectors. This assertion aligns with the broader consensus that renewable energy is quintessential for global efforts to combat climate change.

The study also highlights that renewable energy production and supply have been on the rise globally. This is due to technological advancements that have made renewable energy more affordable and economically competitive with fossil fuels. The decrease in cost of renewable energy, combined with its positive externalities, such as environmental benefits, enables renewables to be a viable alternative to conventional energy sources.

Hydropower, wind, solar, and geothermal energy sources are specifically mentioned in the study due to their substantial contributions to the electrical generations as renewable energy sources. Hydropower is recognised as the largest renewable energy source for power generation worldwide (Abolhosseini et al., 2014). Despite its large energy generation, hydropower faces challenges such as high initial investment and reallocation costs, as well as environmental concerns which hinders its widespread adoption.

Wind power has also seen significant growth, especially in countries like China, US, Germany, and Denmark. Wind power's advantages include rapid installation, relatively low investment and O&M cost, and zero fuel costs. However, intermittency of wind turbine and transmission costs remain a challenge for wind power.

Solar power technology has seen developments and cost reductions as of late. US adopted the highly efficient concentrated solar power technology and China's investment in solar power capacity has significantly lowered generation costs. The only issues include the land, material and chemical used and the affected aesthetics of buildings.

Geothermal energy, while a continuous and reliable source, is subject to its geological constraints. Geothermal energy draws heat energy from the Earth’s mantle for electricity generation, it offers an eco-friendly option from natural resources.

Abolhosseini, Heshmati, and Altmann (2014), also discusses the importance of energy efficiency in reducing overall energy consumption and CO2 emissions. Enhancing energy efficiency is considered an essential strategy to complement renewable energy adoption. Energy efficiencies are explored through various technologies, such as electric vehicles, combined heat and power (CHP), virtual power plants, and smart grids.

Electric vehicles (EVs) are noted as a potential solution to reduce emissions in the transportation sector. Transitioning to EVs on a large scale could result in substantial energy savings and emissions reductions, as battery, fuel cells, and hybrid types are viable option for both electricity storage and power generation. Additionally, the aid of smart grid technologies could further improve the large-scale use of EVs and enhance the efficiency of the EVs’ technology.

CHP technologies are recognised for their ability to improve efficiency by reusing and repurposing waste heat for heating buildings, thereby increasing overall energy efficiency. Virtual power plants and smart grids offer solutions to energy waste, reducing transmission losses and optimising load reductions.

A separate study by Kobos, Erickson, and Drennen (2005), discusses the impact of policies and economics in the renewable energy industry. The authors assert that government financial and institutional support are essential for fostering innovation and growth of renewable technologies.

The study noted that despite the substantial support for renewable technologies provided in the aftermath of the 1970s energy crises, the lack of a cohesive national system of innovation for renewable energy technologies in the US has hindered their progress. Instead, wind energy growth in the US is due to financial incentives and capital cost reductions than domestic technological innovation.

Solar photovoltaic technology, despite early goals of 20 years ago, struggled to gain traction. Oil prices declined during the 1980’s which results in a decline in political support for renewables and reduced research and development (R&D) funding (Kobos et al. 2005). The shortsighted approach to research and innovation has halted the progression of renewable energy goals.

The authors argue that the current state of renewable energy RD&D (research, development, and deployment) may not penetrate the market without sustained support from both the federal level and commercial marketplace. Therefore, adequate funding and appropriate energy policy planning are quintessential for fostering innovation and reducing costs.

Another study by Bull (2021), emphasises the potential socioeconomic, environmental, and community planning associated with the adoption of renewable energy sources.

In the US, the agricultural sector can diversify its income sources by growing crops specifically for energy production as biomass. Rural communities are expected to benefit from the renewable energy usage, leading to greater flexibility in energy choices and enhanced economic potential. This could potentially boost rural economies, create nonfarm jobs, and reduce dependency on oil imports.

Urban and suburban communities are expected to change by the restructuring of the electric industry, there will be more options for distributed energy resources and increased energy efficiency in buildings. These changes will be facilitated by advances in information technology and renewable energy technologies, leading to smart buildings and transportation systems.

In the context of international socioeconomic equity, renewable energy technologies are seen as tools for improving the quality of life for marginalised populations. These technologies provide electricity to areas without access to reliable energy sources, which reduces gender disparities and promotes entrepreneurship to manufacture, sell, and service renewable energy systems (Bull, 2021).

In the environmental perspective, Bull (2021) emphasises the substantial environmental damage caused by conventional energy production and use, including air pollution and CO2 emissions. Renewable energy is presented as a solution to these environmental challenges, especially to avert greenhouse gas emissions.

To summarise, the literature reviews highlight the reliability of renewable energy as a replacement for oil and gas. Renewable energy sources, including hydropower, wind, solar, biofuels, and geothermal power, have the potential to curtail CO2 emissions, mitigate climate change, while guaranteeing a stable energy supply. The authors employ various research methods to address the reliability of renewable energy as a substitute for fossil fuels. However, these approaches are met with challenges such as feasibility, initial investment costs, societal impacts and policy support must be addressed for wider adoption.

Kobos, Erickson, and Drennen employed a quantitative research method known as scenario modelling to assess the capital cost trend of wind and solar PV technologies. The analysis considers different target levels of cost effectiveness and how the implementation of subsidies can achieve these targets. Kobos et al. (2005) research highlights how coordinated R&D national policies are essential for achieving cost competitiveness and reducing the reliance on traditional fossil fuels.

Abolhosseini, Heshmati, and Altmann provides an overview of the renewable energy, focusing on hydro, wind, solar, and geothermal sources. The authors used a literature review methodology to understand the state of renewable energy technologies on reducing CO2 emissions and combating climate change. Their research identifies the potential and challenges of renewable energy, such as intermittency and policy support, that must be addressed for a wider adoption across the globe.

Bull (2021) used qualitative analytical approach to examine the socioeconomic, environmental, and economic aspects of renewable energy integration. The study explores the impact of renewable energy in transforming rural economies, community planning, international contexts, and the environment. Bull’s research summarised on the need to address environmental challenges but leaves us with an open-ended conclusion; that renewable energy adoption depends on the will of society.

In conclusion, these studies employ different research methods, such as literature review, quantitative analysis, and qualitative assessment, to address the reliability of renewable energy as a substitute for oil and gas. They collectively expand on the challenges and opportunities in the transition of renewable energy sources, as well as the need for comprehensive strategies to create a reliable and sustainable energy future. My project aims to add on to their research by utilising time series forecasting models such as ARIMA and Double Exponential Smoothing. These models can forecast renewable energy generation and its levelized cost of electricity, identifying the reliability of renewable energy sources in its eventual goal of replacing oil and gas as our energy of choice.

# Chapter 3.0 Data Understanding and Preparation

## Chapter 3.1 Data Understanding

Our goal is to complement the earlier mentioned research by making use of Our World in Data that has been meticulously collected, aggregated, and documented by Hannah Ritchie, Pablo Rosado, Edouard Mathieu, and Max Roser (2022). These comprehensive datasets encompass global renewable energy generation, fossil fuel generation, and the levelized cost of electricity (LCOE). We will be analysing data from as early as 1965 to 2022, applying time series forecasting on the extensive historical data on renewable energy production, measured in terawatt-hours (TWh) of electricity generation.

Electricity generation (TWh) represents the total electricity produced by various sources, including electricity plants, combined heat and power plants (CHP), and distributed generators. To put this into perspective, 1 TWh equates to 1,000 GWh or 1,000,000 MWh. The average household typically consumes approximately 800 to 1,000 kWh.

This electrical generation measurement is taken at the output terminals of these generation facilities, accounting for both on-grid and off-grid generation. It also accounts for the electricity self-consumed within energy sectors, not just the electricity fed into the grid.

A diagram of different types of power

Description automatically generated

Figure 1: Renewable energy sources

As illustrated in Figure 1, our data analysis will primarily concentrate on comparing the four major renewable energy sources: hydropower, wind, solar, bioenergy, and other renewables. Hydropower encompasses various forms of renewable hydropower, including pumped storage. Wind energy is divided into offshore and onshore wind energy. Solar energy includes both solar photovoltaic and concentrated solar power. Lastly, ‘bioenergy and others’ consist of solid biomass and geothermal.

The input data consists of continuous values, categorised as follows:

|  |  |  |
| --- | --- | --- |
| Shares of Electricity Generation | Electricity Generation / (TWh) | LCOE / (USD/kWh) |
| fossil\_share\_elec | Electricity from hydro | LCOE of hydro |
| renewables\_share\_elec | Electricity from solar | LCOE of solar |
| nuclear\_share\_elec | Electricity from wind | LCOE of wind |
|  | Electricity from Biofuels and other Renewables | LCOE of Biofuels & Others |
|  | Electricity from fossil fuels | LCOE of fossil fuels |

**Table 1:** Extracted data variables

In order to provide additional context for forecasting renewable energy generation, we employ a supplementary set of datasets containing various LCOE values. All datasets are measured against the categorical value 'Year'.

## Chapter 3.2 Data Preparation

The datasets used for this project were compiled and extracted from two primary sources: Our World in Data GitHub and IRENA (2022). These datasets were filtered to include only the necessary data, which includes:

* Shares of electricity generation from fossil fuels, nuclear and renewables.
* Electricity generation from hydro, solar, wind, biofuels and other renewables, and fossil fuels, measured in terawatt-hours (TWh).
* Levelized Cost of Electricity (LCOE) for hydro, solar, wind, biofuels & others, and fossil fuels, measured in USD per kilowatt-hours (USD/kWh).

The measured electricity generation and the shares of electricity generation data are sourced from Our World in Data GitHub, while the Levelized Cost of Electricity (LCOE) data is obtained from IRENA (2022). The LCOE of fossil fuels comprises of the combined-cycle gas turbine (CCGT) data from various countries, as there is no single weighted average for the world's LCOE of fossil fuels, given the differences between countries. The calculation of these averages is achieved using the AVERAGE() function, as depicted in Figure 2.

A screenshot of a graph

Description automatically generated

Figure 2: Combined-cycle gas turbine (CCGT)

Other variables, such as solar, wind, and biofuels & others, represent the combined average of two values and are grouped as follows:

* Wind energy combines both offshore and onshore wind energy.
* Solar energy combines both solar photovoltaic and concentrated solar power sources.
* Bioenergy and others encompass of solid biofuels and geothermal data.

For any missing geothermal data in 2011, it is filled using the median value from 2010 and 2012, as demonstrated in Figure 3.

A screenshot of a computer

Description automatically generated

Figure 3: LCOE of Geothermal

Subsequently, the datasets are consolidated and combined using the function AVERAGE() to determine the mean LCOE of variables; solar, wind, and biofuels & others, as shown in Figure 4. The datasets are then rounded to the nearest 3 significant figures.

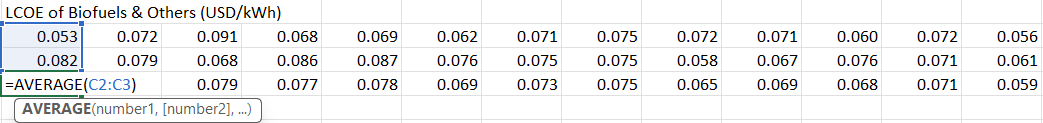


Figure 4: The combination of the LCOE of Biofuels and Geothermal

# Chapter 4.0 Proposed Modelling and Evaluation

## Chapter 4.1 Scope and Methods

Our objective is to analyse the historical trends in renewable energy production, make predictions about future renewable electrical generation, assess production costs, and identify areas for potential enhancement. Our work will make use of multiple datasets, with a primary emphasis on the world’s renewable electrical generation. We will employ Python3 to implement two forecasting techniques: ARIMA and Holt's Linear Exponential Smoothing (Double Exponential Smoothing). These forecasting methods are the most suited for predicting renewable electricity generation up to the year 2030.

We will be forecasting for a duration of 8 years, with forecasts frequencies on an annual basis. The year 2030 holds immense significance on a global scale due to the adoption of numerous eco-friendly initiatives by numerous nations, such as Singapore's Green Plan, the United Nations' Sustainable Development Goals with a focus on Climate Action and Affordable & Clean Energy (Goals 13 and 7), as well as China's strategic Action Plan. Following the forecasting process, we will employ Python3 to determine the level of accuracy in our forecast, using either the Mean Absolute Percentage Error (MAPE) or Mean Absolute Deviation (MAD).

By utilising Python3 to plot the graphs, we can discern distinctive trends with the global renewable energy data. A linear trend graph can be observed for hydropower and fossil fuels. Conversely, the data for wind, solar and biofuels & others exhibit an exponential trend, depicting their rapid and accelerating growth.

Based on the results, it becomes apparent to employ ARIMA modelling for forecasting hydropower and fossil fuels trends. Conversely, for wind, solar, and biofuels & others, an exponential smoothing provides a more suitable approach for accurate forecasting.

## Chapter 4.2 Modelling with ARIMA

This project imports from three open-source Python libraries: ‘Pandas’, ‘NumPy’, and ‘Matplotlib’. Pandas and NumPy work synergistically as the fundamental libraries for data manipulation and analysis. Matplotlib is used for the visualization of plotted graphs in Python.

A screen shot of a computer

Description automatically generated

Figure 5: ARIMA libraries

Figure 5 shows the implementation of Augmented Dickey-Fuller (ADF) Test, as well as the autocorrelation function (ACF) and partial autocorrelation function (PACF). These are imported from ‘statsmodels.tsa.stattools’ and ‘statsmodels.graphics.tsaplots’ respectively. The former assess the stationarity of the dataset. Its null hypothesis assumes non-stationary, and to reject this null hypothesis, data manipulation through n-order of differencing is applied to bring the time series to a constant mean and variance, thus stationary. The ACF and PACF are utilised to determine the order of moving average (MA) and autoregressive (AR) components, respectively. These tools help us understand the correlation among previous observations at different lags within the time series.

AutoRegressive Integrated Moving Average (ARIMA) is implemented via the use of ‘statsmodels.tsa.arima.model’. The orders (p, d, q) dictate the (AR), (I), (MA) component respectively and can build a model that best fit the patterns, trend, and seasonality of the electrical generation. The Mean Absolute Percentage Error (MAPE) and Akaike Information Criteria (AIC) are different measures for evaluating the performance of an ARIMA model. AIC score measures the degree of fit and complexity, ensuring against overfitting, and aids in the selection of the most suitable ARIMA model. AIC score varies with different values of p, d, and q in ARIMA(p, d , q), with lower AIC scores typically representing a better model fit.

The MAPE score measures the forecasting accuracy by calculating the average percentage difference between the predicted values and the actual observed values. Generally, a lower MAPE score, less than 10%, indicates a high degree of accuracy in the forecasting process.

*Hydropower - ARIMA*

A graph with a line and a line

Description automatically generated with medium confidence

Figure 6: Moving Average and Standard Deviation of Hydropower

Based on the graph above, hydropower stands as the leading source of electrical generation among the renewables in the world. Hydropower encompasses various forms, including reservoirs, run-of-river hydropower, and pumped storage. The utilisation of hydropower for electricity generation dates back to as early as 1965. The data exhibits a linear upward trend from 923.13 TWh to 4326.76 TWh, with an increasing Moving Average over the given period.

A screenshot of a computer

Description automatically generated

A diagram of a graph

Description automatically generatedA graph of a graph showing a line of water

Description automatically generated with medium confidence

Figure 7: ADF, ACF and PACF of Hydropower

Based on the results above, the p-value of the Augmented Dickey-Fuller (ADF) Test is significantly larger than 0.05, suggesting that the graph is non-stationary. Furthermore, the ACF is significant at lag 1 and has diminishing positive correlation on subsequent lags, while PACF is significant at lag 1 and shows no correlation on subsequent lags. Both the ACF and PACF do not display any signs of seasonality.

A graph showing the number of different levels of a graph

Description automatically generated with medium confidence

Figure 8: Moving Average and Standard Deviation of second-order differenced Hydropower

A screenshot of a computer

Description automatically generatedA screenshot of a computer code

Description automatically generated

Figure 9: ADF of first-order differencing (left) and second order differencing

In Figure 8, a second-order differencing is applied to the model, bringing the graph to stationarity. A first-order differencing may not be suitable despite the p-value of less than 0.05 and the ADF test statistic is lower than the 5% and 10% critical values. As shown in Figure 9, while the test statistic is higher than the 1% critical value, indicating some degree of stationarity in the data, it is not as robust as desired. Therefore, a second-order differencing is preferred. With a p-value is significantly less than 0.05 and the test statistics lower than the critical values, indicating that the graph can reject the null hypothesis and has achieved stationarity.

A graph of a diagram

Description automatically generated A graph of a diagram

Description automatically generated

Figure 10: ACF and PACF of second-order differenced Hydropower

At the second-order differencing, the ACF and PACF exhibit significant lags at 2 and 1 respectively. Since the dataset is a linear upward trend without seasonality, we will deploy ARIMA (p, d, q). ‘p’ represents the order of AR(q), ‘d’ is the order of differencing, and ‘q’ represents the order of MA(q). The significance of lag 1 in PACF suggest AR (1), and the significance of lag 2 in ACF suggests MA (2). Furthermore, the dataset achieve stationarity at second-order differencing.

Based on the results mentioned, we will apply an ARIMA (1, 2, 2) model.

The accuracy of the ARIMA (1, 2, 2) model can be assessed through Python. We will calculate the MAPE and evaluate the AIC and the coefficient of determination, as shown in Figure 11. The ARIMA (1, 2, 2) model has a MAPE score of 2.38% which is significantly lower than 10%. Furthermore, its AIC score is relatively low at 614 compared to other combinations of ARIMA (p, d, q). Lastly, the coefficient of determination stands at 0.979, indicating that the ARIMA model closely fits to the dataset.

For the code process of Hydropower ARIMA forecasting, please refer to Appendix A.

A screenshot of a computer

Description automatically generated

Figure 11: Summary of Hydropower ARIMA model

*Fossil Fuels – ARIMA*

A graph showing the growth of a number of individuals

Description automatically generated

Figure 12: Moving Average and Standard Deviation of Fossil Fuels

Based on the graph above, fossil fuel is at large and the primary source of electrical generation among in the world. The dataset of fossil fuel electrical generation ranges from 1985 to 2022. The data exhibits a linear upward trend from 6285.48 TWh to 17385.15 TWh, with an increasing Moving Average over the given period.

A screenshot of a computer

Description automatically generated

A graph of a graph showing the amount of fossil fuel

Description automatically generated A graph of a graph showing the amount of fossil fuel

Description automatically generated

Figure 13: ADF, ACF and PACF of Fossil Fuels

Based on the results above, the p-value of the Augmented Dickey-Fuller (ADF) Test is significantly larger than 0.05, suggesting that the graph is non-stationary. Furthermore, the ACF is significant at lag 1 and has diminishing positive correlation on subsequent lags, while PACF is significant at lag 1 and shows no correlation on subsequent lags. Both the ACF and PACF do not display any signs of seasonality.

A graph showing the growth of a number of percent

Description automatically generated with medium confidence

Figure 14: Moving Average and Standard Deviation of first-order differenced Fossil Fuels

A screenshot of a computer

Description automatically generated

A graph showing a graph of fossil fuel

Description automatically generated A graph of a graph showing the amount of fossil fuel

Description automatically generated

Figure 15: ADF, ACF and PACF of first-order differenced Fossil Fuels

In Figure 14, we applied first-order differencing to the model, achieving stationarity. A first-order differencing is sufficient for stationarity, as the p-value is significantly less than 0.05 and the test statistics are lower than the critical values, indicating that the graph can reject the null hypothesis and has achieved stationarity.

At the first-order differencing, the ACF and PACF do not exhibit significant lags at all. Since the data is stationary with no significant lags in the ACF and PACF, we will experiment with different orders of AR (p) and MA (q) components, start with ARIMA(0,1,0) as a baseline and gradually introduce AR and MA terms.

Based on the experiments mentioned, ARIMA (1, 1, 1) model has the lowest AIC, we will fit an ARIMA (1, 1, 1) model.

To assess the accuracy of the ARIMA (1, 1, 1) model. We will calculate the MAPE and evaluate the AIC and the coefficient of determination, as shown in Figure 16. The ARIMA (1, 1, 1) model has a MAPE score of 1.72% which is significantly lower than 10%. Furthermore, its AIC score is relatively low at 533 when compared to other ARIMA (p, 1, q) combinations. Lastly, the coefficient of determination stands at 0.914, indicating that the ARIMA model closely fits to the dataset.

For the code process of Fossil Fuels ARIMA forecasting, please refer to Appendix B.

A screenshot of a computer

Description automatically generated

Figure 16: Summary of Fossil Fuels ARIMA model

## Chapter 4.3 Modelling with Double Exponential Smoothing

When modelling for Holt’s Linear Exponential Smoothing, also known as Double Exponential Smoothing, we shall import from three open-source Python libraries: ‘Pandas’, ‘NumPy’, and ‘Matplotlib’. Pandas and NumPy for data manipulation and analysis, while Matplotlib is used for the visualization of plotted graphs in Python.

A screen shot of a computer

Description automatically generated

Figure 17: Double Exponential Smoothing libraries

Figure 17 shows the implementation of the autocorrelation function (ACF) and partial autocorrelation function (PACF). These are imported from ‘statsmodels.graphics.tsaplots’. The ACF measures the correlation between data points at different time lags, while PACF measures the correlation affected by the previous time lags. In an exponential growth, a strong positive autocorrelation is observed at lag 1 of an ACF plot. Subsequent lags may have diminishing positive correlations. Conversely, PACF will exhibit a robust positive autocorrelation at lag 1, with subsequent lags experiencing a sharp decline or demonstrating no correlation. This indicates a direct relationship between the current value and the previous one.

Since the datasets exhibits exponential growth, there is no need to test for stationarity, and instead, Double Exponential Smoothing (DES) is employed. DES is executed through the use of ‘statsmodels.tsa.holtwinters’. Double Exponential Smoothing is essentially Holt-Winters Exponential Smoothing without season but with trend. In this project, seasonality is removed through Python, ‘model = ExponentialSmoothing(df, trend='add', seasonal=None)’.

Similarly, to the modelling with ARIMA, we also use the Mean Absolute Percentage Error (MAPE) and Akaike Information Criteria (AIC) for evaluating the performance of a DES model. The AIC score assesses the degree of fit and complexity, preventing overfitting. Lower AIC scores in DES modelling results typically represents a suitable DES model fit. The MAPE score measures the forecasting accuracy by calculating the average percentage difference between the predicted values and the actual observed values. A lower MAPE score of less than 10%, signifies a high degree of accuracy in the forecasting process.

Even though MAPE is a useful tool for assessing forecasting accuracy, it may have limitations in some scenarios. For example, MAPE can be sensitive to extreme values and is not well-defined when the actual values are zero. As such, we include the Mean Absolute Deviation (MAD) in DES models to assess the forecasting accuracy. MAD measures the average deviation between each data points and the mean of the dataset. A lower MAD value indicates a suitable DES model fit, with smaller errors in its predictions.

*Solar power – Double Exponential Smoothing*

A graph with a line

Description automatically generated

Figure 18: Plot of Solar power electrical generation

A graph of a solar power

Description automatically generated A graph of a solar power

Description automatically generated

Figure 19: ACF and PACF of Solar power

Based on the graphs above, solar power exhibits an exponential growth pattern in the time series data. Although solar power has been used since 1983, however it did not experience significant adoption until 2007, leading to widespread global adoption. Solar power encompasses both solar photovoltaics and concentrated solar power.

The data shows a gradual increase from 0.003 TWh in 1983 to 6.92 TWh in 2007. After 2007, an exponential growth trend emerges, with power generation soaring from 11.36 TWh in 2008 to 1289.27 TWh in 2022.

The ACF exhibits a slow decay and is positive at lag 1, indicating a decreasing positive correlation between observations as the time lag increases, suggesting a trend is present. However, the PACF decreases sharply, while positive at lag 1, implying that the direct correlation of previous time lags diminishes rapidly. Both ACF and PACF show no signs of seasonality. These characteristics suggest an exponential growth trend with no seasonality. Therefore, we will deploy Double Exponential Smoothing (DES).

To assess the accuracy of the DES model. We will calculate the MAD instead of MAPE, as some original values are close to 0, which can result in extremely large MAPE calculations. The MAD score is 6.60, significantly less than 10, indicating the suitable goodness of fit for DES model. Furthermore, its AIC score is relatively low at 224.

For the code process related to Solar power DES forecasting, please refer to Appendix C.A screenshot of a computer

Description automatically generated

Figure 20: Summary of Solar power DES model

*Wind power – Double Exponential Smoothing*

A graph with a line drawn on it

Description automatically generated

Figure 21: Plot of Wind power electrical generation

A graph of a wind power

Description automatically generated A graph of a graph with blue lines

Description automatically generated with medium confidence

Figure 22: ACF and PACF of Wind power

Based on the graphs above, wind power exhibits an exponential growth pattern in the time series data. While wind power has been in use since as early as 1978, it did not achieve significant global adoption until 2000. Wind power comprises both offshore and onshore wind power.

The data exhibits a gradual increase from 0.003 TWh in 1978 to 31.16 TWh in 2000. After 2000, there is an exponential growth with power generation capacity increasing from 38.16 TWh in 2000 to 2139.23 TWh in 2022.

The ACF and PACF in wind power’s time series data share similar characteristics with those of Solar power. Both ACF and PACF show no signs of seasonality, suggesting an exponential growth trend with no seasonality. Therefore, we will deploy Double Exponential Smoothing (DES).

To assess the accuracy of the DES model, we will calculate the MAD and MAPE. Both MAD and MAPE depicts a score of 10.94 and 11.71% respectively. The MAD score is 10.94, and the MAPE is 11.71%, both of which are relatively low. This indicates that the DES model is a suitable fit for the time series data. Furthermore, the AIC score is low at 282.

For the code process related to Wind power DES forecasting, please refer to Appendix D.

A screenshot of a computer

Description automatically generated

Figure 23: Summary of Wind power DES model

*Biofuels & other Renewables – Double Exponential Smoothing*

A graph with a line drawn on it

Description automatically generated

Figure 24: Plot of Biofuels and other Renewables electrical generation

A graph of a graph with blue lines

Description automatically generated with medium confidence A graph of a graph with numbers

Description automatically generated with medium confidence

Figure 25: ACF and PACF of Biofuels and other Renewables

Examining the graphs above, Biofuels & other Renewables exhibits a slow exponential growth pattern. Biofuels & other Renewables have been utilised as early as 1965, and they have experienced a gradual increase in global adoption. Biofuels & other Renewables mainly comprises of solid biomass and geothermal sources. The data exhibits a slow yet increasing rate of electricity generation from 17.985 TWh in 1965 to 777.31 TWh in 2022.

The ACF and PACF in the time series data share similar characteristics with the preceding datasets for solar power and wind power. Both the ACF and PACF show no signs of seasonality, suggesting an exponential growth trend with no seasonality. Therefore, we will deploy Double Exponential Smoothing (DES).

To assess the accuracy of the DES model, we will calculate both MAD and MAPE. MAPE is suitable for this assessment, as the low data points of the dataset is larger than 0, allowing for the calculation of MAPE. The MAD score is 4.55, and the MAPE is 2.51%, both of which are significantly lower than 10. This suggests that the DES model is a suitable fit for the time series data. Furthermore, the AIC score is low at 245.

For the code process related to Biofuels & other Renewables DES forecasting, please refer to Appendix E.

A screenshot of a computer

Description automatically generated

Figure 26: Summary of Biofuels and other Renewables DES model

# Chapter 5.0 Results

The forecasted models are fitted against the original data in orange and the forecasted values are plotted in green up to the year 2030, we will discuss the following forecast results:

*Hydropower - ARIMA*

A graph showing the growth of water

Description automatically generatedA number of numbers on a white background

Description automatically generated

Figure 27: Plotted forecast and forecasted values of Hydropower

Hydropower is forecasted using an ARIMA (1,2,2) model. The fitted line generally fits the hydropower dataset, with the exception for an outlier in 1967. The forecast shows a rising trend in hydropower electricity generation. Within the forecast period, some years show relatively constant electrical generation, including 2024 and 2025, 2026 and 2027, and 2028 and 2029.

In 2023, it is projected to generate 4338.71 TWh, and by 2030, the forecasted generation is at 4759.30 TWh. An overall 9.69% increase in hydropower electrical generation.

*Fossil Fuels – ARIMA*

A graph showing the growth of fuel prices

Description automatically generated A white background with numbers and a black text

Description automatically generated with medium confidence

Figure 28: Plotted forecast and forecasted values of Fossil Fuels

Fossil fuels is forecasted using an ARIMA (1,1,1) model. Similarly, the fitted line generally fits the fossil fuels dataset, and the forecast shows an upward linear trend in fossil fuels electricity generation.

In 2023, fossil fuel is projected to generate 17680.91 TWh, and by 2030, the forecasted generation is at 19751.19 TWh. An 11.71% increase in fossil fuels electrical generation.

*Solar power – Double Exponential Smoothing*

**A graph of a solar power

Description automatically generated**A screenshot of a computer

Description automatically generated

Figure 29: Plotted forecast and forecasted values of Solar power

Solar power is forecasted using Double Exponential Smoothing (DES) model. Based on the graph above, the fitted line closely fits the solar power dataset, and the forecast shows a sharp increase in solar power electricity generation in the subsequent years.

In 2023, solar power is projected to generate 1538.48 TWh, and by 2030, the forecasted generation is at 3282.95 TWh. An outstanding 113.39% increase in solar power electrical generation, more than doubling the production level in 2023.

*Wind power – Double Exponential Smoothing*

A graph with a line

Description automatically generatedA screenshot of a computer

Description automatically generated

Figure 30: Plotted forecast and forecasted values of Wind power

Wind power is forecasted using Double Exponential Smoothing (DES) model. Similar to solar power, the fitted line closely fits the wind power dataset, and the forecast shows a sharp increase in the following years of wind power electricity generation.

In 2023, wind power is projected to generate 2429.10 TWh, and by 2030, the forecasted generation is at 4460.54 TWh. A spectacular 83.63% increase in wind power electrical generation, coming in just 298.76 TWh shy of the highest renewable energy generation, hydropower.

*Biofuels & other Renewables – Double Exponential Smoothing*

A graph with a line

Description automatically generatedA screenshot of a computer

Description automatically generated

Figure 31: Plotted forecast and forecasted values of Biofuels & other Renewables

Lastly, biofuels & other renewables have the lowest combined electricity generation compared to other sources. The original dataset displays a parabolic pattern, and we use a Double Exponential Smoothing (DES) model for forecasting. The fitted line fits closely to the dataset, and the forecast shows a linear increase in electricity generation.

In 2023, biofuels & other renewables are projected to generate 818.79 TWh, and by 2030, the forecasted generation is at 1084.95 TWh. This shows a 32.51% increase in electrical generation from biofuels & other renewables.

## Chapter 5.1 Discussions

This project forecast for electrical generation is essential in providing insights to the capacity of electrical production, yet it alone cannot determine the reliability of renewable energy as an oil and gas replacement. To fully assess reliability, we must also understand its economic feasibility.

The Levelized Cost of Electricity (LCOE) is a comprehensive metric that encompasses various form of costs, including the total installed costs, the lifetime capacity factor, Operation and Maintenance (O&M) costs, the economic lifespan of the project, and the cost of capital. LCOE provides valuable insights into renewable energy's potential as a reliable, cost-effective, and sustainable alternative.

We shall conduct a quick forecast using Excel to understand the LCOE of different energy sources. Please refer to Appendix F for the settings used.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Forecasted Levelized Cost of Electricity (LCOE) of / (USD/kWh) | | | | |
| Year | Hydropower | Solar power | Wind power | Biofuels & others | Fossil Fuels |
| 2022 | 0.0610 | 0.084 | 0.057 | 0.0589 | 0.167 |
| 2023 | 0.0534 | 0.067 | 0.037 | 0.0633 | 0.167 |
| 2024 | 0.0551 | 0.051 | 0.028 | 0.0634 | 0.168 |
| 2025 | 0.0568 | 0.034 | 0.019 | 0.0653 | 0.168 |
| 2026 | 0.0585 | 0.018 | 0.010 | 0.0555 | 0.168 |
| 2027 | 0.0602 | 0.001 | 0.001 | 0.0599 | 0.168 |
| 2028 | 0.0619 | -0.015 | -0.008 | 0.0601 | 0.168 |
| 2029 | 0.0636 | -0.032 | -0.017 | 0.0619 | 0.169 |
| 2030 | 0.0653 | -0.048 | -0.026 | 0.0522 | 0.169 |

Table 2: Levelized Cost of Electricity Forecast

Table 2 represents the forecasted LCOE for five energy sources: hydropower, solar power, wind power, biofuels & other renewables, and fossil fuels. To assess the financial impact of these projections, we will multiply the energy generation by the LCOE, expressed as TWh \* USD/kWh, considering that 1 TWh equals 1,000,000,000 kWh.

Hydropower costs increased approximately 48% from 2010 to 2022, reaching 0.061 USD/kWh. This increase is due to the rising installation cost as remote plants, with poor infrastructure, has higher logistical, civil engineering and grid connection costs. The LCOE is projected to rise to 0.0653 USD/kWh in 2030.

In 2022, hydropower generated 4,326.76 TWh, utilising the formula, the global cost of hydropower amounts to $263.93 billion. In 2030, hydropower is expected to have an LCOE of 0.0653 USD/kWh and a projected generation of 4,759.30 TWh, resulting in a forecasted global cost of $310.78 billion.

Solar power costs decreased by a significant 80% from 2010 to 2022, reaching 0.084 USD/kWh. The LCOE is projected to further decrease to -0.048 USD/kWh in 2030. However, negative cost is not possible, instead we will assume cost plateau and take the most reasonable cost of 0.018 USD/kWh.

In 2022, solar power generated 1,289.27 TWh, costing around $108 billion globally. However, in 2030, solar power is expected to have an LCOE of 0.018 USD/kWh and a projected generation of 3282.95 TWh resulting in a reduced estimated cost of $59 billion.

Wind power costs reduced by 60% from 2010 to 2022, reaching 0.057 USD/kWh. This can be attributed to the improved technology of wind turbine, having larger turbines, longer blades, and higher heights. The LCOE of wind power is projected to further decrease to -0.026 USD/kWh in 2030. Similar to solar power, negative cost is not possible, and we will assume cost plateau, taking the most reasonable cost of 0.010 USD/kWh.

Wind power generated 2,139.23 TWh in 2022, with a global cost of approximately $121.94 billion. In 2030, wind power is expected to have an LCOE of 0.010 USD/kWh and a projected generation of 4460.55 TWh, resulting in a forecasted global cost of $44.6 billion.

Biofuels & other renewables costs decreased by about 23% from 2010 to 2022, reaching 0.0589 USD/kWh, expected to further decrease to 0.052 USD/kWh by 2030.

Biofuels & others generated 777.31 TWh in 2022, incurring a global cost of about $45.78 billion. By 2030, an LCOE of 0.052 USD/kWh and a projected generation of 1,084.95 TWh are expected, resulting in an estimated cost of $56.41 billion.

Fossil Fuels experienced a gradual 39% cost increase from 2012 to 2022, reaching 0.167 USD/kWh and its LCOE is projected to rise to 0.169 USD/kWh by 2030.

For fossil fuels in 2022, the global energy generation reached 17,385.15 TWh, incurring a substantial cost of $2.9 trillion. By 2030, with an LCOE projected at 0.169 USD/kWh and a generation forecast of 19,751.19 TWh, the cost is estimated to rise to $3.33 trillion.

These projections have significant implications for the budgets allocated to energy generation expenses, as some energy sources are expected to become more costly in the future, while others are expected to become more cost-effective, making these energy sources an attractive alternative to traditional fossil fuels.

# Chapter 6.0 Recommendations and Conclusion

## Chapter 6.1 Recommendations

While forecasting models such as ARIMA and exponential smoothing provide valuable insights into calculating electrical generation capacity, and levelized cost of electricity offers an understanding of the economic feasibility of renewable energy. These methodologies are insufficient to determine if renewable energy is reliable to replace oil and gas.

A graph showing the growth of electricity generation

Description automatically generated

Figure 32: Shares of Electricity Generation

Figure 32 illustrates the global shares of electricity generation sources. As seen in the graph, fossil fuels continue to dominate, constituting approximately 60% of the world's electrical generation. This is supported by the analysis of upward trend in fossil fuel generation, suggesting that alternative energy sources would not be replacing oil and gas in the near future. However, a glimmer of hope emerges as renewable energy's share is on the rise, progressing from around 20% to 30% of global energy generation. Meanwhile, nuclear energy's share is in decline, and fossil fuel usage is gradually decreasing since 2010. This shift can be attributed to the concerns stemming from the Fukushima incident and the geopolitical influence of oil prices in the middle east. On the other hand, renewable energy is as an autonomous and free resource that can be harnessed. Continuous technological improvement has led to the reduction in the overall LCOE and an increase in production for renewables. This improved cost-effectiveness and increasing affordability encourage government bodies to consider adopting renewable energy as replacements for oil and gas.

## Chapter 6.2 Conclusion

Our analysis of renewable energy replacement for oil and gas has revealed its challenges and opportunities. The reliability of renewable energy sources faces obstacles such as unpredictable weather conditions and suboptimal equipment performance, which cast doubt on their long-term sustainability and concerns for the financial impact of these setbacks.

Our project utilises time series forecasting not only to estimate electrical generation capacity but to assess the cost efficiency of renewable energy as a sustainable alternative. Our forecast reveals that renewables electrical generation are on the rise and the Levelized Cost of Electricity (LCOE) suggests an increase in hydropower and fossil fuels costs, but a decrease in biofuels & others, solar and wind power costs. These trends indicate that fossil fuels will become more obsolete in the future, while generally renewables will become more cost-effective, making renewables an increasingly attractive alternative to traditional fossil fuels.

In light of these findings, we recommend comprehensive strategies to address the challenges surrounding renewable energy, such as fostering innovation through research and development, enhancing energy efficiency through technologies like electric vehicles and combined heat and power, and addressing environmental and socio-economic factors associated with renewable energy adoption. While renewable energy presents an encouraging path forward, it requires a multifaceted approach for the world to adopt it as a reliable and sustainable energy source, potentially replacing oil and gas in the future.

# References

Abolhosseini, S., Heshmati, A., & Altmann, J. (2014). A review of renewable energy supply and Energy Efficiency Technologies. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2432429

Bull, S. R. (2021, April 1). *Renewable energy today and tomorrow*. IEEE Journals & Magazine. https://ieeexplore.ieee.org/document/940290

Hannah Ritchie, Max Roser and Pablo Rosado (2022) - "Energy". Published online at OurWorldInData.org. Retrieved from: 'https://ourworldindata.org/energy' [Online Resource]

Huenteler, J., Tang, T., Chan, G., & Anadon, L. D. (2018). Why is China’s wind power generation not living up to its potential? *Environmental Research Letters*, *13*(4), 044001. https://doi.org/10.1088/1748-9326/aaadeb

IRENA (2022), Renewable power generation costs in 2022, International Renewable Energy Agency, Abu Dhabi

Kobos, P. H., Erickson, J. D., & Drennen, T. E. (2005). Technological Learning and renewable energy costs: Implications for US renewable energy policy. *Energy Policy*, *34*(13), 1645–1658. https://doi.org/10.1016/j.enpol.2004.12.008

# Appendices

Appendix A – Hydropower Forecast

A screenshot of a computer program

Description automatically generated

A screenshot of a computer program

Description automatically generated

A computer screen shot of a code

Description automatically generated

A computer screen shot of a computer code

Description automatically generated

Appendix B – Fossil Fuels Forecast

A screenshot of a computer code

Description automatically generated

A screen shot of a computer

Description automatically generated

A computer screen shot of a computer code

Description automatically generated

A computer screen shot of a computer code

Description automatically generated

Appendix C – Solar power Forecast

A screenshot of a computer code

Description automatically generated

A close-up of a white rectangular object

Description automatically generated

A computer screen shot of a computer code

Description automatically generated

A screen shot of a computer code

Description automatically generated

Appendix D – Wind power Forecast

A computer screen shot of a computer code

Description automatically generated

A white rectangular object with a black border

Description automatically generated

A screen shot of a computer code

Description automatically generated

A computer code with text

Description automatically generated with medium confidence

Appendix E – Biofuels & other Renewables Forecast

A screenshot of a computer code

Description automatically generated

A close-up of a computer code

Description automatically generated

A screenshot of a computer program

Description automatically generated

A screenshot of a computer

Description automatically generated

Appendix F – LCOE Forecasts

*Hydropower*

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a data table

Description automatically generated

*Solar power*

*A screenshot of a computer screen

Description automatically generated*

*A screenshot of a computer

Description automatically generated*

*A screenshot of a data table

Description automatically generated*

*Wind power*

*A screenshot of a graph

Description automatically generated*

*A screenshot of a computer

Description automatically generated*

*A screenshot of a data table

Description automatically generated*

*Biofuels & other Renewables*

*A screenshot of a computer screen

Description automatically generated*

*A screenshot of a computer

Description automatically generated*

*A screenshot of a data table

Description automatically generated*

*Fossil Fuels*

*A screenshot of a computer screen

Description automatically generated*

*A screenshot of a computer

Description automatically generated*

*A screenshot of a data table

Description automatically generated*

Appendix G – Shares of electricity generated

A screenshot of a computer code

Description automatically generated

*A graph showing the amount of electricity in the year

Description automatically generated*

*A computer screen shot of a computer code

Description automatically generated*

*A graph showing the number of electricity generation

Description automatically generated*