Energy Consumption in London: Predicting future trends across different Sectors

ABSTRACT

This paper analyses the energy consumption for the 32 boroughs of London and then uses that data to make ten years predictions using exponential smoothing of the average or sum of prior observations to make time series forecasts.

The paper uses a dataset from the London Datastore to predict ten years of energy consumption across the gas, petroleum products, electricity, bioenergy and wastes, Manufactured Fuels, Petroleum Products, and Coal fuel types. It goes through all the steps involved in the S.E.M.M.A. methodology to clean and obtain meaningful data from the raw dataset. The obtained data is now analyzed using the Holt-Winters method through visualization software like Microsoft Excel and Power Bi.

Many questions have arisen from these predictions: Which sector consumes and will consume the most fuel in the future? What type of fuel will be the most consumed in the future? How much will there be drops in consumption across different sectors across the London boroughs? This paper investigates the energy consumption of 32 boroughs of London across various sectors in more depth.

Keywords: Exponential Smoothing, Time Series Forecast, ARIMA, Holt-Winters, S.E.M.M.A.

1.0 Introduction

Each year, Londoners spend billions of Pounds on energy bills. The energy per capita is in thousands of kWh. We cannot overstate the importance of energy because it is one of the driving forces of the economy. Developed nations tend to have higher per capita than developing and under-developed nations. Population, services, economy, and housing growth will lead to more energy consumption.

Energy companies contributed 2.5 percent of gross value added (G.V.A) to the U.K. economy in 2021 and employed about 175,000 directly. Energy companies also accounted for 8% of total investments and 26.2 of industrial investments in the U.K. economy [1]. The energy crisis ravaging the United Kingdom has re-emphasized the importance of using past data set to predict the future trend. The energy crisis across London and the United Kingdom has drastically driven up the standard of living costs. Different factors have contributed to the surge in energy bills, and they are not limited to the Coronavirus pandemic and demand outweighing the supply of energy. The energy crisis has reiterated the need for consumption efficiency. Different programs and incentives have been put in place by the government to curtail the effects of the crisis.

Energy is vital to a country's economy and its citizens' well-being. Unsurprisingly, countries with the highest energy consumption per capita tend to be more developed and economically advanced. In the U.K., 404,000 households live in fuel poverty [2]. Researchers from Stanford University found a link between energy and poverty. They found how access to electricity affects the well-being of Uganda (a country in East Africa) and that countries with access to energy experienced economic well-being roughly double that of regions without access. [3]

This project started with an abstract and discussed the various objectives, goals, and deductions at the end of the project. The journal literature review shows a brief outline of energy in the UK and academic scholars' previous works and results. The introduction details the methods used to analyze the time series data and results and their advantages and disadvantages.

Methodology spoke in detail about the model(Holt-Winters) used in the project. The results thoroughly discussed the forecasts of the dataset with various visualizations—the project terminated with the conclusion, which spoke about final deductions and an explanation of the results obtained.

2.0 Literature Review

The bulk of energy used in London was majorly in the residential areas, with more than 44% of the energy consumed on average. Then closely followed by the Commercial and Industrial sectors, with more than 35% on average. On the other hand, the Rail sector had the lowest energy consumed between 2005 to 2015, with less than 0.2% on average. Between 2005 and 2015, gas and electricity accounted for more than 77% of total fuel used, with gas accounting for 42%.

Izzaamirah *et al.* [4] researched and forecasted the electricity consumption of Malaysia's residential sector from 2019 until 2032 and identified the best exponential smoothing model to make the predictions. They determined energy consumption patterns by using the dataset from 1997 to 2018. it was discovered, based on the research results, that Holt exponential smoothing was the best model for forecasting electricity consumption in Malaysia. Also, they discovered from their results that energy consumption comes mostly from household appliances and that the moving average helps to predict electricity consumption. They also made it known that ARMA (autoregressive-moving average) and ARIMA (autoregressive integrated moving average) are both suitable for energy predictions and studies.

While Wang Rongbin *et al.* [5] embarked on research into the study prediction of energy conservation in universities based on exponential smoothing, they concluded that it is necessary to use exponential smoothing for making statistical predictions of energy consumption. The combined use of both smoothed non-linear planning and linear planning methods resulted in little or inconsequential values of energy consumption for the next few years.

Tryggvi Jónsson *et al.* [6] argued that the Holt-Winters model might be more suitable for short-range prediction and can be susceptible to unforeseeable events. In their research, Y.W Lee et al. [7] forecasted the monthly electricity consumption using a time series analysis of a Malaysian school from January to December 2018 using data from January 2011 to December 2017. Their research showed that Holt-Winters Model was the best model for their forecast because it gives the minor mean absolute error (M.A.E.) and absolute percentage error (MAPE), thereby making a better forecast.

Eva Ostertagová *et al.* [7] pointed out in their paper that a simple exponential smoothing model is only suitable for non-seasonal patterns and short-term predictions.

Based on past journals and research in this project, the Holt-Winters model was preferred because it allows the three components (level, trend, and seasonality) to change over time, making the method

S.A Yeboah *et al.* [8] argued, based on their findings, that the ARIMA models are "efficient" and "robust" in forecasting energy consumption and are suitable for use by developing and developing countries for energy consumption forecasting. In a comparative study between Holt-Winters and ARIMA models, M Markovska *et al.* [9] concluded from their findings that the Holt-Winters model is more suitable for daily findings. In contrast, the ARIMA model is more accurate for weekly and monthly time series datasets.

Looking at both ARIMA and Exponential Smoothing models, Exponential Smoothing models are non-stationary (a time series where the statistical properties change over time). It describes how the components (error, trend, seasonality) change over time. ARIMA models are stationary (a time series where the statistical properties experience no change over time). [10]. A dataset in the ARIMA model will have to be converted to stationary if it is not already stationary. Both can analyze univariate (data that has only one component) data. ARIMA and exponential smoothing models use weighted sums or averages for their predictions.

This project will use Triple Exponential Smoothing or the Holt-Winters, model.

3.0 Methodology

3.0.1 Triple Exponential Smoothing (Holt-Winters Model)

Triple Exponential Smoothing is a way to predict and model the behavior of values over a particular period. It is a model developed in the 1960s when Chris Holt and Peter Winter worked together to improve Holt's method of capturing seasonality. The Holt-Winters model is a way to model the statistical properties (trend, error, and seasonality) [11]. The Holt-Winters uses values and datasets from the past to make typical values predictions for the present and future. The four equations that represent Holt-Winters are below.

$$s_t = a\frac{y_t}{I_{t-L}} + (1-a)(s_{t-1} + b_{t-1}) \quad (1)$$

$$b_t = \gamma(S_t - S_{t-1}) + (1-\gamma)b_{t-1} \quad (2)$$

$$I_t = \beta\frac{y_t}{S_t} + (1-\beta)I_{t-L} \quad (3)$$

$$F_{t+m} = (S_t + mb_t)I_{t-L+m} \quad (4)$$

Equations 1,2,3,4 are the overall smoothing, trend smoothing, seasonal smoothing, and forecast, respectively [12].

Where

- y is the observation
- S is the smoothed observation
- b is the trend factor
- I_t is the seasonal index
- F is the forecast at m periods ahead

- T is an index denoting a period

α, β, and γ are constants.

3.0.2 Sample, Explore, Modify, Model, and Assess (SEMMA)

This project applied SEMMA methodology for data mining. SEMMA was the most suitable for the project because of its size and straightforwardness. The following are the five stages of SEMMA:

- **Sample**: This stage aimed to select the variables that would be useful for the project. The data was then prepared and moved into Excel and Power Bi for further action.
- **Explore**: In this stage, univariate analysis was conducted on the dataset to identify the relationships between the data and the missing components. All factors influencing the study were analyzed using data virtualization software.
- **Modify**: During this stage, the data was parsed and cleaned with software like Microsoft Excel and Power Bi before being passed onto the modeling stage and explored if the data requires refinement and transformation.
- **Model**: Here, there was a time series application on the dataset using Triple Exponential Smoothing or Holt-Winters Model. In this stage, there was a projection for the targeted outcomes using the time-series analysis. Visualization software like Microsoft Excel and Power Bi was the best for modeling.
- **Assess**: This is the final stage of the SEMMA process to evaluate the model's usefulness and reliability in the chosen topic [13].

4.0 Results

4.1. Data Visualization

Visualization software like Microsoft Excel and Power Bi was the best for the project. There were various visualizations for the different fuel types used in London.

4.1.1 All Fuel Types

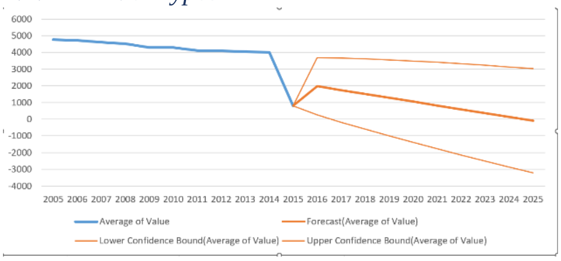

Fig. 1. Forecast of All Fuel Types for ten years between 2016-2025 using Microsoft Excel.

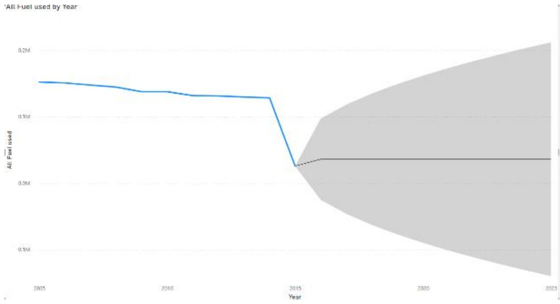

Figure 2: Power Bi Forecast of All Fuel Types for 2016-2025 using Power Bi.

The figures above show the visualizations of all fuel types used in London between 2005 and 2015 and the forecast for 2016-2025.

4.1.2 Coal

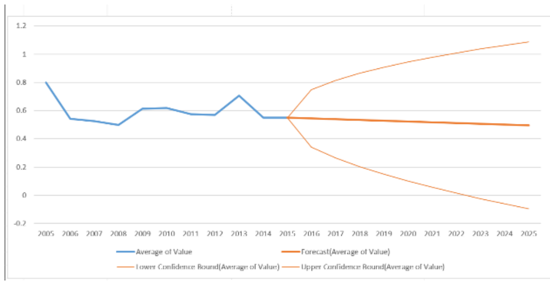

Figure 3: Forecast of Coal for ten years between 2016-2025 using Microsoft Excel.

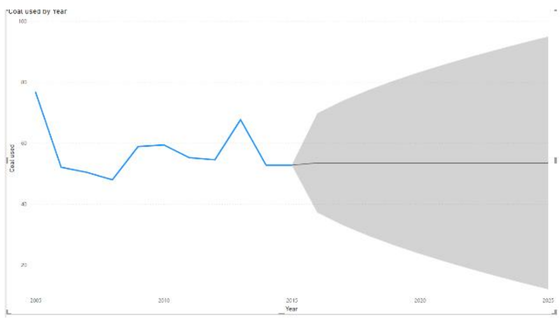

Figure 4: Power Bi Forecast of Coal Types for 2016-2025 using Power Bi.

4.1.3 Bioenergy and Wastes

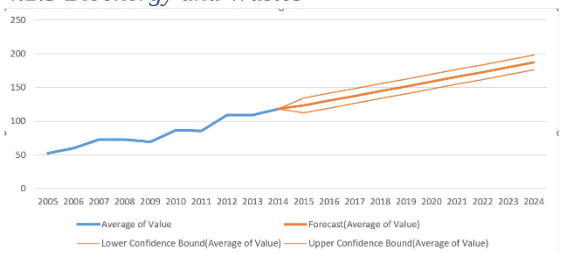

Figure 5: Forecast of Bioenergy and wastes for ten years between 2015-2024 using Microsoft Excel.

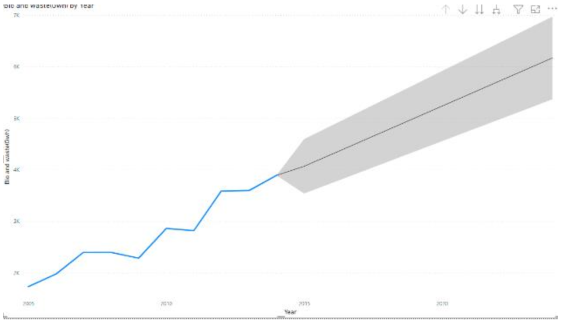

Figure 6: Power Bi Forecast of Bioenergy and wastes for 2015-2024 using Power Bi.

4.1.4 Electricity

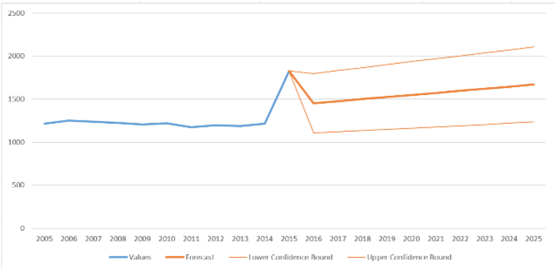

Figure 7: Forecast of Electricity for ten years between 2016-2025 using Microsoft Excel.

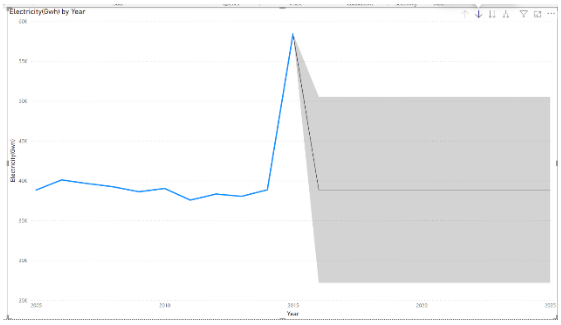

Figure 8: Power Bi Forecast of Electricity for 2016-2025 using Power Bi.

4.1.5 Gas

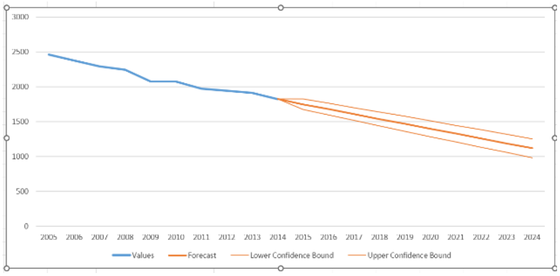

Figure 9: Forecast of Gas for ten years between 2016-2025 using Microsoft Excel.

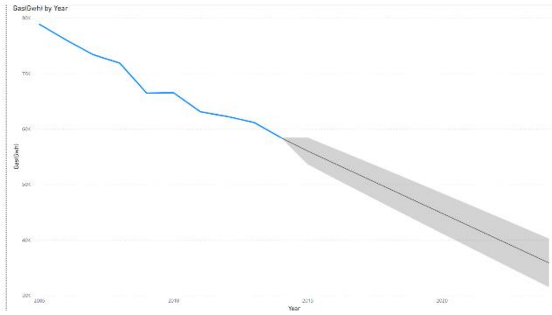

Figure 10: Power Bi Forecast of Bioenergy and wastes for 2016-2025 using Power Bi.

4.1.6 Manufactured Fuel

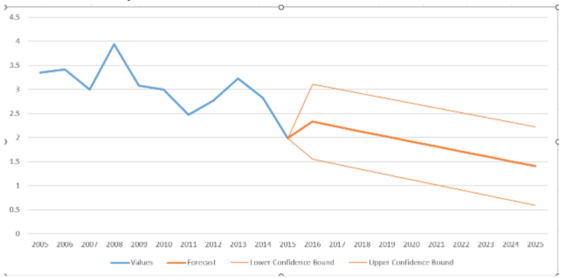

Figure 11: Forecast of Manufactured Fuels for ten years between 2016-2025 using Microsoft Excel.

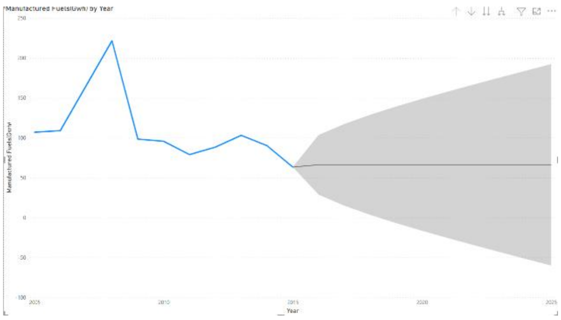

Figure 12: Power Bi Forecast of Manufactured Fuels for 2016-2025 using Power Bi.

4.1.7 Petroleum Products

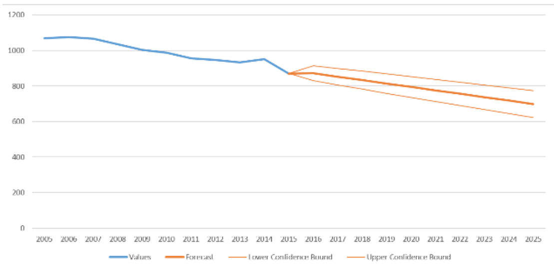

Figure 13: Forecast of Petroleum Products for ten years between 2016-2025 using Microsoft Excel.

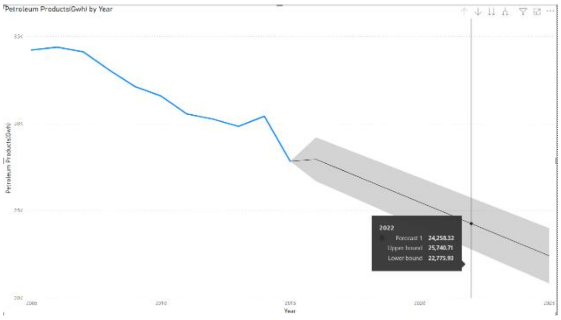

Figure 14: Forecast of Petroleum Products for ten years between 2016-2025 using Power Bi.

4.2 Data Analysis

Looking through the various visualizations', many deductions can be made. For all fuel types, the forecast tends to rise in 2016 and then fall sharply from 2017 to 2025.

In contrast, the Coal forecast is slightly different from all fuel types. The forecast falls linearly throughout the years. Gas, on the other hand, has been falling over the years. It is not surprising seeing the energy consumption of gas continue to fall over the next ten years.

Bioenergy and waste forecast is the direct conflict with gas. While the forecast for gas is in the fall, the forecast for bioenergy and waste is rising. Electricity forecasts witnessed a drop in the year 2016 but then fell in the year 2017. It then gradually rises from the year 2018 to 2025. The deduction from the graph is that the forecast is a direct reflection of the dataset.

Manufactured fuels differ from other fuel types because the dataset tends to rise and fall over the year. The forecast for synthetic fuels rises for the year 2016, then starts falling linearly from 2017 up to 2025. Using the ten years forecast, one expects the consumption of Manufactured Fuels to keep falling past the forecasts.

On the other hand, Petroleum Products witnessed a fall from 2005 to 2014, then fell sharply in 2015. The consumption then continues to fall over all the years. There will be a drop in the rate of consumption down the line.

From the visualizations, there will generally be a fall in the electricity consumption of the 32 boroughs of London across different sectors except for electricity, bioenergy, and waste. There tends to be a trend in the fuel types that will not experience a fall in consumption. This trend shall be talked about in detail in the conclusion section.

One thing to note about the differences between the Microsoft Excel and Power Bi visualizations is their shape and slope direction. These differences are primarily due to the dataset not being large enough and the model used. The lack of enough datasets has made the integrity of the visualizations questionable. Due to the lack of enough datasets over the years, one cannot rely on the forecast. For the forecast to give an accurate result, the dataset must contain different months and years. One can deduce more accurate results when we adjust seasonality if there is a more diverse range of datasets instead of being limited to a few years.

It is worth knowing that the project used a seasonality of 1 and a confidence level of 95% for the different visualizations on both Microsoft Excel and Power Bi. The specified seasonality has also not shown any effect on the results. With the change in seasonality to 2, there was a change in pattern in the visualizations.

The change in value in the forecast has proved to be an influencing factor in the results. The minimal change in forecasts has proven to be another reason the forecasts may be inaccurate.

5.0 Conclusion

The above dataset has shown that the energy consumption in London will reduce over the years across the different sectors except in electricity, bioenergy, and waste. The reduction is due to the available dataset and consumption pattern across these sectors.

One thing pointed out in the results is that electricity, bioenergy, and waste consumption will continue to experience a rise. This rise can be because the world is going to tilt more and more toward renewable energy.

One limitation of time series analysis models like the Holt-Winters is that even though it considers seasonality, they cannot make space for unforeseen circumstances and external factors. For example, the prediction does not take into account the issue of the Covid-19 pandemic. Another example is the issue of climate change which has made some countries abandon some fuel types. Climate change has, in turn, caused an unprecedented increase in fuel types and a decrease in other fuel types.

From the overall project, it is safe to say the integrity of the forecast may be questionable, and the visualizations give a straight-line pattern which may point to the insufficient dataset and the minimal change in value. Due to the previously given reasons, one begins to doubt the usefulness of the Holt-Winters Method in using the scanty dataset and long forecast periods to give a very accurate prediction. From various arguments made in the past, it is safe to say the Holt-Winters method may not be the right choice of model in this case.

References

[1] Department for Business, Energy & Industrial Strategy, “UK ENERGY IN BRIEF 2022,” National Statistics, 2022 July 28. [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1094025/UK_Energy_in_Brief_2022.pdf. [Accessed 30 November 2022].

[2] Energy for Londoners, “Energy for Londoners,” 2022. [Online]. Available: https://www.london.gov.uk/programmes-and-strategies/environment-and-climate-change/energy/energy-londoners. [Accessed 17 November 2022].

[3] M. Burke, N. Ratledge and e. al, “Stanford research ushers in a ‘new frontier’ in tackling global poverty,” nature, 16 November 2022. [Online]. Available: https://news.stanford.edu/2022/11/16/stanford-research-ushers-new-frontier-tackling-global-poverty/. [Accessed 22 November 2022].

[4] I. Ishak, N. S. Othman and N. H. Harun, “Forecasting electricity consumption of Malaysia’s residential sector: Evidence from an exponential smoothing model [version 1; peer review: 1 approved with reservations],” *F1000Research,* vol. I, p. 10, 2022.

[5] W. Rongbin, W. Zhang, W. Deng, R. Zhang and X. Zhang, “Study on Prediction of Energy Conservation and Carbon Reduction in Universities Based on Exponential Smoothing,” *Sustainability,* vol. 14, no. 19, p. 11903, 2022.

[6] T. Jónsson, P. Pinson, H. A. Nielsen and H. Madsen, “Exponential Smoothing Approaches for Prediction in Real-Time Electricity Markets,” *Energies,* vol. 7, no. 6, pp. 3710-373, 2014.

[7] E. Ostertagová and O. Ostertag, “THE SIMPLE EXPONENTIAL SMOOTHING MODEL,” *The 4th International conference on modelling of mechanical and mechatronic systems,* pp. 380-344, 2011.

[8] S. A. Yeboah, M. Ohene and T. Wereko, “Forecasting aggregate and disaggregate energy consumption using arima models: A literature survey,” *Journal of Statistical and Econometric Methods,* vol. 1, no. 2, pp. 71-79, 2012.

[9] Marija Markovska, A. Buchkovska and D. Taskovski, “Comparative study of ARIMA and Holt-Winters statistical models for prediction of energy consumption,” *XII International Conference,* vol. ETA16, pp. 1-6, 2016.

[10] A. Bert, “Week 4: ARIMA vs. ETS Models,” RPubs by RStudio, 20 7 2020. [Online]. Available: https://rpubs.com/andrea_bert/641847. [Accessed 25 November 2022].

[11] SolarWinds, “Holt-Winters Forecasting and Exponential Smoothing Simplified,” Orange Matter, 15 December 2019. [Online]. Available: https://orangematter.solarwinds.com/2019/12/15/holt-winters-forecasting-simplified/#:~:text=The%20Holt%2DWinters%20method%20uses,%E2%80%9Csmooth%E2%80%9D%20a%20time%20series.. [Accessed 25 November 2022].

[12] National Institute of Standards and Technology, “Triple Exponential Smoothing,” National Institute of Standards and Technology, [Online]. Available: https://www.itl.nist.gov/div898/handbook/pmc/section4/pmc435.htm. [Accessed 26 November 2022].

[13] N. HOTZ, “What is SEMMA?,” Data Science Process Alliance, 14 May 2021. [Online]. Available: https://www.datascience-pm.com/semma/. [Accessed 26 November 2022].

[14] Y. Lee, K. T. and Y. C. , “Forecasting Electricity Consumption Using Time Series Model,” *International Journal of Engineering & Technology,* vol. 7, no. 4.30, pp. 218-223, 2018.