Cairo University

Faculty of Economics and Political Science

Statistics Department

English Section

Prediction of Electricity Demand in the United States

Prepared by

Abdelghaffar Hosny Muhammed

Esraa Salah Elganzoury

Farah Hany Ahmed Lashin

Under the Supervision of

Dr. Eman Mahmoud

2022/2023

"The goal of forecasting is not to predict the future but to tell you what you need to know to take meaningful action in the present"

Consulting professor and technology forecaster,

Paul Saffo

Acknowledgements

We would like to pay our sincerest regards to our supervisor dr. Eman Mahomoud for the great help she provided us, since without her guidance, we won't be able to present this thesis.

We also want to give our appreciation to Dr. Fatma Elzanaty and Ms. Menna Gamal for their patience and guidance throughout the process. We want to show our gratitude as well to Dr. Abdelnasser Saad for supporting and reassuring us when we felt down.

We also want to thank our colleagues, teaching assistants, and professors in statistics department, as they were a huge support through our three years in the department and we learned a lot from them.

At last, we would like to express our gratitude, love, and respect to our family and friends who were here, caring for us, attending to our needs, and giving us love with no conditions.

Abstract

This paper aims at forecasting the electricity demand in the United States of America. Electricity demand can be identified as electricity consumption per hour in Kilowatts. The forecasting process is done using two models, SARIMA model and Prophet model, and comparing their results to find out which model is more suitable for the nature of the data provided. The data used in the analysis is hourly data for the months of April, May, June, and the first week ofJuly for the year 2021, the data is conducted from the United States Energy Information Administration's official website. The analysis for the paper was conducted using R-studio and MINITAB. Both the SARIMA model and the Prophet model produced close results, however, the SARIMA model had an advantage in terms of accuracy measures, as the SARIMA model showed a MAPE value of 3.47%, while the Prophet model showed a value of 3.95%. In terms of visual representation, it was apparent that the SARIMA model had a better performance in terms of how close the forecasted values are to the actual ones. It can be concluded that the SARIMA model is better on the short run, while the Prophet model is preferred on the long run, which is why obtaining a larger dataset that includes longer periods for the electricity consumption is recommended for future research using the Prophet model. It is also recommended to employ a hybrid model of both the SARIMA and the Prophet model, that way we can put both models' advantages to use.

Table of contents

List of Tables

List of Figures

List of Abbreviations

Abbreviation	Definition
ACF	Autocorrelation Function
ADF	Augmented Dickey-Fuller
AIC	Akaike Information Criterion
ANN	Artificial Neural Network
ARIMA	Autoregressive Integrated Moving Average
ARIMAX	Autoregressive Integrated Moving Average with Exogenous Input
BIC	Bayesian Information Criterion
ETS	Error Trend and Seasonality
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MPE	Mean Percentage Error
PACF	Partial Autocorrelation Function
PFM	Prophet Forecasting Model
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Networks
RNN-LSTM	Recurrent Neural Networks- Long Short-Term Memory
SARIMA	Seasonal Autoregressive Integrated Moving Average
SMA	Seasonal Moving Average

Chapter One: Introduction

1.1 Electricity and Its Sources

Electricity is considered the only energy source that isn't a fuel, but a power source, generated by other fuels, such as coal, fuel oil, natural gas, or even wood. Unlike any other fuel, electricity is first generated, and then transmitted to utilities of use.

In the United States of America, electricity consumption has increased exponentially since households' first years of use in 1910, and nationally in 1903. Household electricity demand continued to increase since 1910, with only a single exception during the Great Depression in 1933. Between 1944 and 1954, growth rates took place. Electricity was hence established as the single energy source that witnessed continuous growth rates with only a few exceptions, in addition to that, it wasn't substantially affected by wars, depressions, or the energy crisis in the 1970s. Hence, it has easily become the energy source of choice [16].

According to the U.S Energy Information Administration, there are many sources for producing electricity; some sources are clean and renewable and some are not. The total electricity generated in kilowatts in the United States in 2022 was about 4,243 billion kWh. Fossil fuel takes the highest share as it produces 60.2% of the produced electricity which include 39.8% of natural gas, 19.5% of coal, 0.6% of petroleum (divided into 0.4% liquids and 0.2%), and 0.3% of other gases. Nuclear energy takes about 18.2% of producing electricity. While renewables produce 21.5% of electricity in the USA; including 10.2% by wind, 6.2% by hydropower, 3.4% by solar energy (divided into 3.4% of photovoltaic and 0.1% of solar thermal), and 1.3% by biomass energy (divided into 0.9% wood, 0.2% landfill gas, and 0.1% municipal solid waste (biogenic)), 0.4% geothermal energy, and 0.1% other biomass waste. The pumped storage hydropower takes about 0.1% of the produced electricity. Othersources produce a small percentage which is 0.3% [26].

Focusing on the renewables, according to the U.S Energy Information Administration, these include: hydropower, which is done by using flowing water to spin a turbine which is connected to a generator, wind energy, which is done using wind turbines that convert energy to electricity, biomass, which can be converted to a gas that is burned in internal combustion engine generators, steam generators, or gas turbines, solar energy which is the energy extracted from the sun using small PV cells that converts the sunlight to electricity, and finally, geothermal which is the heat in the earth itself and converting it to electricity [27]. With the increase in electricity, the amount of energy increases. And if this energy is not generated from a renewable source, this would cause harm to the environment which will harm humans as a result. Forecasting electricity will help predict the amount of energy needed. This would allow governments to account for such demand by building stations for clean energy sources.

1.2 US Electricity Supply and Demand

Currently, electricity generation in the United States is mainly through using combustible fossil fuels which include coal and natural gas. Electricity is extracted through these fossil fuels by burning them and getting steam in boilers that turn them into electricity. In addition to that, electricity can be generated via nuclear power, which extracts steam from the radioactive elements' fission into an energy generator as well. Also, energy is generated by wind turbines or hydropower, which is known as mechanical energy, it also includes solar photovoltaic panels (PV), which convert light into energy immediately. Finally, geothermal energy, which uses underground heat to generate steam from which energy is generated. Usually, electricity produced must be used as soon as its generation as the required technologies for storing such energy are not yet available on a wide level. Selecting a certain source of energy generation in the US is directly affected by the cost of fuel, for many years, fossil fuels were known to provide the least cost, and therefore, coal and natural gas have accounted for two-thirds of power generation since 2000. Renewable sources, however, like wind and solar PV power, don't demand fuel, so the electricity they generate is solely dependent on the available wind and sunlight. Renewables prices have decreased in the last decade which caused an increase in their employment [3].

1.3 Study Problem

Forecasting electricity demand is an important issue, for which many methods were provided. These methods include statistical models such as the ARIMA model by Bilal (2016) in Turkey, the SARIMA model by Abodulaye et al. (2016) in the USA, the Prophet model by Alex (2020) in Australia, and the Artificial Neural Networks (ANN) by Ewa (2021) in Poland. This paper is focused on hourly American electricity demand data from 04/1/2021 to 06/30/2021, which will be used to forecast the second week of July, using the SARIMA model and the Prophet model, and finally, comparing the results.

The electricity demand in the USA is one of the highest demands in the world; which makes it very interesting to study the pattern of the demand more closely. In this study, a time series analysis is conducted to check the patterns, build a model, and forecast future electricity demand.

In this paper, the aim is to study the electricity data of the United States using two models, the prophet model (which was first used by Facebook, to predict daily data on the application), a forecasting model that works with time series data. The prophet model was specifically chosen as it has great advantages, that is: it is robust when it comes to missing values and changes in trend, and to outliers, meaning that influential values can't change or affect the analysis in a biased way [22]. In addition to that, we will also be applying the seasonal autoregressive integrated moving average model (SARIMA) model (Seasonal ARIMA), as electricity data do have seasonality within the day.

1.1 Some Used Measurements in Selection Criteria and Forecasting Accuracy

Some Selection Criteria Measurements:

Two measures are used in this paper for model selection, the The Akaike Information Criterion (AIC), and the The Bayesian Information Criterion (BIC), they each are a measure of the fit that depends on the residual sum of squares of the model, and hence, the lower its value is, the better. They are used to evaluate statistical models, especially in cases where multiple forms of the model are available, and we want to choose only one, we can then evaluate both the AIC and the BIC of each model and choose the model that possesses the lowest values for both. One of the criterion used in this paper to deem a model as a good fit, is searching for the one that has lower AIC and BIC values.

The Akaike Information Criterion (AIC) takes the form:

$$AIC = 2k + n + \ln(\frac{RSS}{n}) \tag{1}$$

Where:

"k" is the number of parameters or regressors, "n" is the number of time series data, "RSS" is the residual sum of squares.

The Bayesian Information Criterion (BIC) takes the form:

$$BIC = kln(n) - \ln(\frac{RSS}{n}) \tag{2}$$

Some Forecasting Accuracy Measures:

The following methods are used to evaluate how accurate the predictions of time series model are, and how far these forecasts are from the actual values, by looking at the difference between the actual value "y_i" and the forecasted one using the model "$\hat{y}_i$". This is why we aim to find the models with lower values for these models, and can conclude these models as better fits for the data.

Mean Absolute Error (MAE): which is the mean of the absolute value of the prediction error.

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{3}$$

Mean Absolute Percentage Error (MAPE): which is the mean absolute difference between actual and forecasted values divided by the actual values, in percentage.

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \left|\frac{y_i - \hat{y}_i}{y_i}\right| \ x\ 100 \tag{4}$$

Mean Percentage Error (MPE): which is the mean of the difference between the actual and forecasted values divided by the actual values in percentage.

$$MPE = \frac{1}{n}\sum_{i=1}^{n}\frac{y_i - \hat{y}_i}{y_i}\; x\; 100 \tag{5}$$

Root Mean Squared Error (RMSE): the square root of the mean of the squared differences between the actual values and forecasted values. [6]

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{6}$$

Statistics were done using Minitab and R 3.5.0 (R Core Team, 2018), the packages used are:
“readxl”, “seastests”, “forecast”, “ggplot2”, “TSstudio”, “read”, “zoo”, “ggsci”, “MASS”, “tseries”, “reshape2”, “reshape”, “prophet”, “readr”, “rlang”, “Rcpp”, “tidyverse”

1.2 Literature Review of Forecasting Electricity Demand

(1) Autoregressive Integrated Moving Average (ARIMA)

Mohamed (2004) employed the ARIMA model to forecast electricity demand in New Zealand, the Maldives, the USA, and the UK, using annual data from the year 2000 to 2015. In the United States, the ARIMA model gave more accurate results, and on the regional level, the ARIMA model was the best fit. Miao, (2015) used the ARIMA model in China to forecast electricity demand using annual data from the year 2000 to 2012, results showed that the ARIMA (1,1,1) has a very small relative error for electricity consumption, and hence has higher prediction accuracy.

Kao et al. (2020) predicted electricity consumption in Taiwan using the ARIMA model, on annual data from 1965 to 2014, they combined the ARIMA model with Support Vector Regression (SVR) models, the hybrid model provided more accurate results than the ARIMA model or the SVR model alone.
Fathin et al. (2021) employed the ARIMA model in Indonesia using monthly data from October 2004 to May 2014 to forecast electricity demand, results showed that ARIMA (8,2,0) has the least average error percentage for their data (5.3%), it also showed that ARIMA is the best fit for medium-term predictions.
Chodakowska et al. (2021) applied the ARIMA model in Poland for forecasting electricity load, using hourly data from 6 July 2020 to 27 September 2020, they identified multiple limitations in their study and recommended using different model classes for the ARIMA model, they also concluded that the forecasting accuracy of the model is affected by the noise level of the signals observed.

(2) Seasonal Autoregressive Integrated Moving Average (SARIMA)

Camara et al. (2016) applied the SARIMA model to forecast electricity consumption in the USA, they used quarterly data over the period from January 1973 to June 2015, they compared the SARIMA model with the Artificial Neural Network (ANN), they concluded that both SARIMA and ANN models can obtain good forecasts to data with seasonality like energy data, it was also found that both SARIMA and ANN models performances are not significantly different.

Chaturvedi et al. (2022) used the SARIMA model in India to forecast electricity demand, they used data from 2008 to 2017 to predict 24 months' data from 2017 to 2019, they conducted their analysis with 3 other models, the Long Short Term Memory Recurrent Neural Network model (LSTM RNN), the Indian's Central Energy Authority's (CEA) existing trend-based model, and Facebook (Fb) Prophet model, they found that both the CEA and Fb Prophet models have lower prediction errors than the SARIMA model and the LSTM RNN model.

(3) The Prophet Model

Chadalavada et al. (2020) employed the Prophet model to predict electricity demand in India, the model was trained on the data from January 11, 2016, to May 27, 2016, the interval was 10 minutes, for 4.5 months, it was compared with the ARIMA model, and the Prophet model was found to be easier and had higher accuracy, it was also found to be most suitable for their data.

Almazrouee et al. (2020) employed the Prophet model in Kuwait for monthly data from January 2015 to May 2020 in order to forecast electricity demand, the prophet models with single and multiple regressors were used and Holt-Winters models were employed as benchmarks for comparative analysis. Results showed that the Holt-Winters model was the best fit for the data among the applied models.

Leung (2020) applied the Prophet model in Australia to build a model that can accurately forecast the electricity demand, the data used was daily from January 1, 2015 to October 6, 2020. Results showed that the Prophet model simplified such forecasting and provided accurate results.

Henzel et al. (2022) used the Prophet model for daily data from 1 March 2021 to 28 October 2021 to accurately predict electricity demand, results showed that the Prophet model produced the best predictions for the next day energy demands.

Liang et al. (2023) used the Prophet model for hourly data from January 1, 2015, to August 3, 2018, in the USA to predict electricity demand, the prophet model decomposition showed higher level of accuracy in forecasting compared to other models used.

(4) Artificial Neural Network (ANN)

Pay (2006) utilized the ANN model in Taiwan, to forecast electricity demand using monthly data from

January 1990 to December 2002, the results were compared with the ARIMAX model, it was found that theforecasting ability of ANN model was higher than that of ARIMAX.

(5) Other Models

Ponraj et al. (2020) employed the ARIMA, RNN-LSTM (Recurrent Neural Networks-Long Short-Term Memory) to forecast electricity demand in France, using monthly data from December 2006 through November 2010, results showed that the ARIMA model's performance fell short in comparison to the RNN and LSTM models.

Pełka, (2023) used the ARIMA, ETS (Error Trend and Seasonality), and Prophet models, and created hybrids of each two models to forecast electricity demand in 35 European countries, using monthly data from 1998 to 2014. Results showed that using hybrids of each two models lead to better prediction results than using each model alone.

Chapter Two: Descriptive Analysis

2.1 Data Description

The electricity demand in the USA is used in this paper, conducted from The United States Energy Information Administration's official website. The data type is hourly, obtained from 04/1/2021 to 06/30/2021. This data is used to forecast the second week of July (which is already known) using the SARIMA model and the Prophet model in order to test and understand the prediction accuracy of the selected models. In this dataset, there are 2184 hourly observations from the months of April, May, June, and the first week of July.

In this chapter, the data will be investigated for seasonality and stationarity. Descriptive values and plots will be presented as well.

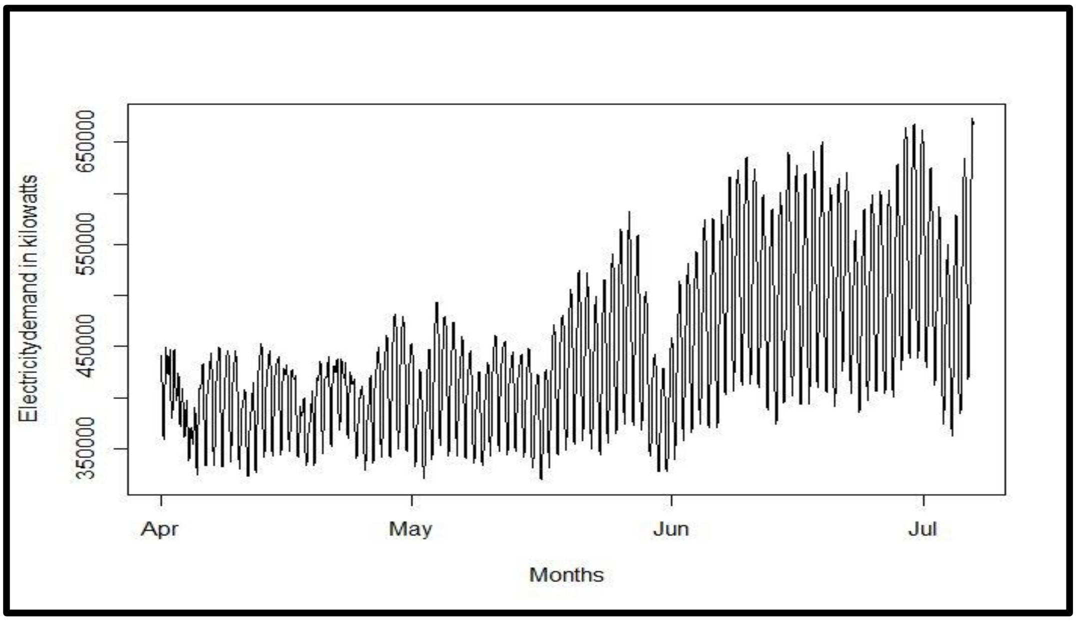

Figure (2.1): Time Series Plot for Daily Electricity Demand in the U.S in kilowatts

Figure (2.1) shows a repeated pattern in the electricity demand through the selected three months; which concludes that there might be seasonality and non-stationarity. That requires performing formal tests with reliable output.

2.2 Testing Seasonality

Testing seasonality can be done using the Kruskal-Wallis test, where the null hypothesis states that all days have the same mean. Under this hypothesis, the test statistic follows a Chi-Square distribution. When this hypothesis is rejected, it is assumed that the time series values differ significantly between periods. The hypotheses are:

H_0: There’s an absence of seasonality in the series

H_1: There’s seasonality in the series

Table (2.1): Kruskal-Wallis seasonality test results

Test Statistic	p-value
2036.57	0

From table (2.1), since the p-values < 0.05; the null hypothesis is rejected. This means that there is seasonality in the electricity demand.

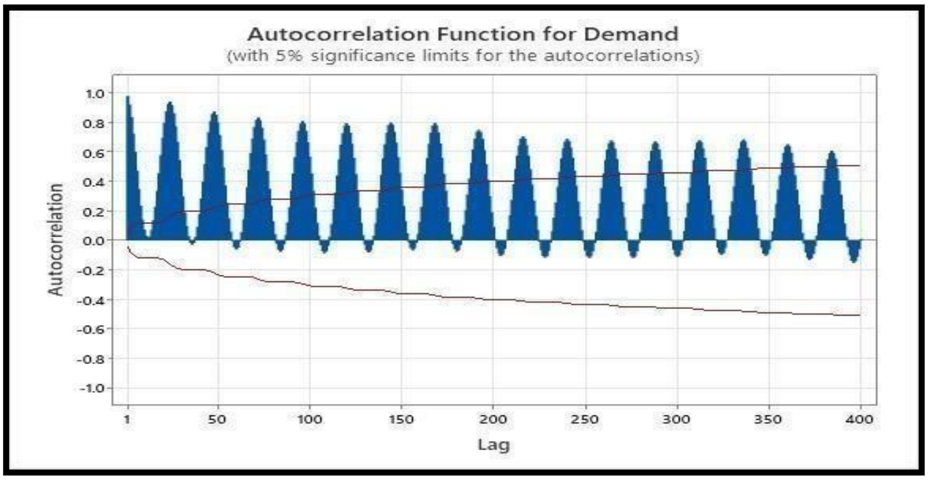

Figure (2.2): Autocorrelation Function Plot for USA Electricity Demand

Figure (2.2) shows daily seasonality as there are regular peaks every 24 lags, meaning a pattern repeats itself every 24 lag, and the lag represents an hour, this means there is daily seasonality, where this pattern repeats itself every 24 hours.

2.1 Studying Stationarity:

To ensure the conclusion from the time series plot, a formal test, the Dickey-Fuller test, can be used to test the null hypothesis that the series is non-stationary. If the test statistic is greater than that critical value, we can reject the null hypothesis and conclude that the series is stationary. The hypotheses are:

H_0: electricity demand is non-stationary

H_1: electricity demand is stationary

Table (2.2): Augmented Dickey-Fuller stationarity test results

Test Statistic	Critical Value	P-value
-1.59436	-2.86288	0.486

Table (2.2) shows that we fail to reject the null hypothesis, meaning that the series is not stationary. To solve this, some measures were taken; such as transforming the data using the logarithmic transformation and then taking the first seasonal difference.

2.2 Daily Progress of Electricity Demand:

Electricity demand varies between the different days of the week, as well as the different hours within the day, the following figure illustrates the demand fluctuation among the week.

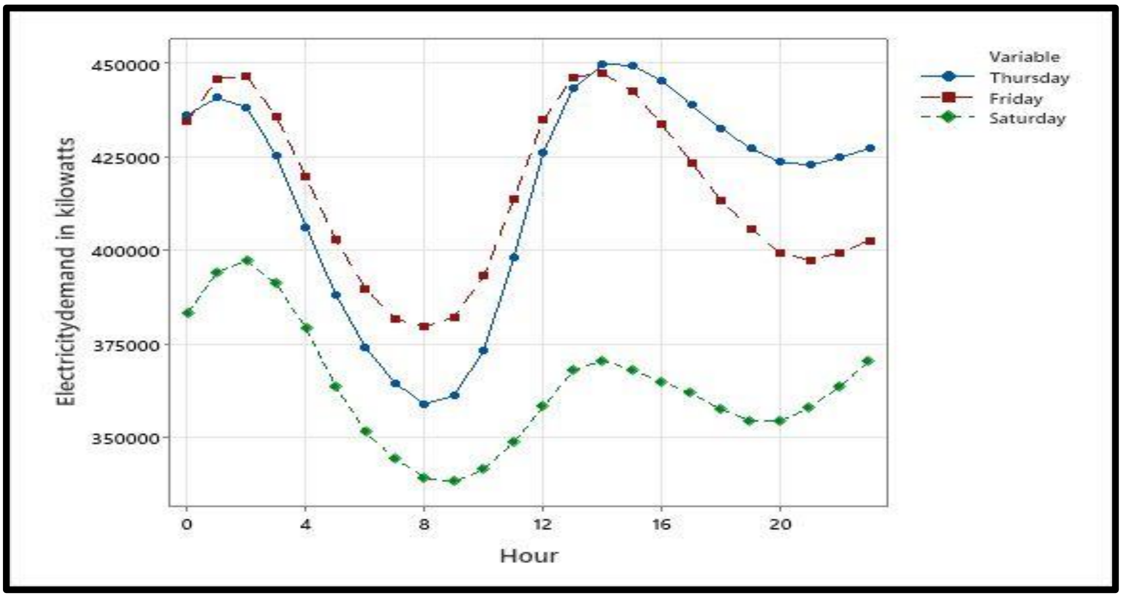

Figure (2.3): Hourly Electricity Demand on Thursday, Friday and Saturday in kilowatts

Figure (2.3) shows how demand differs between weekdays (Thursday and Friday) and weekends (Saturday), it can be seen that consumption on Thursday and Friday is relatively high, as the week days are usually when household members are home and hence consume electricity at almost all times. Whereas during the weekend, Saturday, it can be seen that a sharp decline compared to week days occurs, which can be explained as Saturdays are usually when household members aren’t present at home and prefer to spend the day outdoors.

2.3 Descriptive Statistics:

The summary statistics are as follows:

Table (2.3): Summary Statistics for hourly electricity demand in the United States in kilowatts

N	Mean	SE Mean	Standard deviation	Min	Q1	Median	Q3	Max
2184	440517	1627	76026	319497	386669	425003	476676	667644

As shown in table (2.3), the average electricity demand is 440517 kilowatts and the standard deviation is 76026. While the other measures of location -minimum, first quartile, median, third quartile, and maximum- are equal to 319497 kilowatts, 386669 kilowatts, 425003 kilowatts, 476676kilowatts, and 667644 kilowatts, respectively.

Chapter Three: The SARIMA Model Analysis

The aim of this chapter is applying the SARIMA model to electricity demand, finding the suitable parameters, and forecast the second week of July using the estimated model

3.1 The SARIMA Model

When the ARIMA model was introduced by Box and Jenkins [29], it was created in order to combine the autoregressive (AR) model, and the Moving Average (MA) model, and solve non-stationary time series by taking the differences (I). However, this model didn't take seasonality into consideration; hence, the SARIMA model which refers to seasonal ARIMA as an extension of the ARIMA model, was introduced. The SARIMA model includes two parts; the first one is similar to the ARIMA's and consists of non-seasonal elements (p, d, q), whilethe other one consists of seasonal elements (P, D, Q). Where "p" is the auto-regression order, "d" is the difference order, "q" is the moving average order, "P" is the Seasonal autoregression order, "D" is the seasonal difference order, "Q" the seasonal moving average order, and "s" is the number of time steps for a single seasonal period

A multiplicative seasonal ARIMA model can be expressed as [2]:

$$\phi_p(B)\Phi_p(B)(1-B)^d(1-B^s)^D Y_t = \Theta_q(B)\mathit{\Theta}_Q(B^s)\ \varepsilon_t$$

$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$$

$$\Phi_P(B^2) = 1 - \Phi_1 B - \Phi_2 B^{2s} - \cdots - \Phi_p B^{Ps}$$

$$\Theta_q(B) = 1 - \Theta_1 B - \Theta_2 B^2 - \cdots - \Theta_p B^q$$

$$\mathit{\Theta}_Q(B^2) = 1 - \mathit{\Theta}_1 B^s - \mathit{\Theta}_2(B^{2s}) - \cdots - \mathit{\Theta}_Q(B^{Qs}) \qquad (7)$$

Where; the " ϕ " is the autoregressive parameters to be estimated, " θ " is the moving average parameters to beestimated, "Y_t" is the electricity demand at time t, "ε_t" is the vector of random errors which is assumed follow a normal distribution with mean 0, "B" is the backshift operator and so $BY_t = Y_{t-1}$, s is the number of seasonal period; p, q and P, Q are the lags, capital letters refer to the seasonal part; d is the order of differencing, and D is the order of seasonal differencing.

In order to be able to model the series of electricity demand, the series has to be stationary. For this purpose, a natural logarithmic transformation was applied as well as a seasonal difference, as previously mentioned. To check the stationarity after transformation, the augmented Dickey-Fuller test (ADF) was reapplied, with the hypothesis:

H_0: electricity demand is non-stationary

H_1: electricity demand is stationary

Table (3.1): Augmented Dicky-Fuller stationarity test results after data transformation

Test statistic	Critical value	p-value
-8.60285	-2.86299	0.000

At $\alpha = 0.05$, we can reject H_0 and say that the electricity demand is stationary.

Once the stationarity assumption is satisfied, the SARIMA model can be employed. Stages of the SARIMA model are: (1) Model identification: since the series is now stationary, the first seasonal difference was taken to remove the effect seasonality, by doing this, the D parameter is determined and it equals 1. The model can now be identified based on the autocorrelation function (ACF) and the partial autocorrelation function (PACF). (2) Parameter estimation: as soon as the model is identified, the model parameters can now be estimated, the selection of these parameters will depend on the Akaike's information criterion (AIC)and Schwarz's Bayesian information criterion (BIC). (3) Diagnostic checking: after estimating the parameters and selecting them, these selected parameters will be tested for statistical significance.

(1) Model Identification:

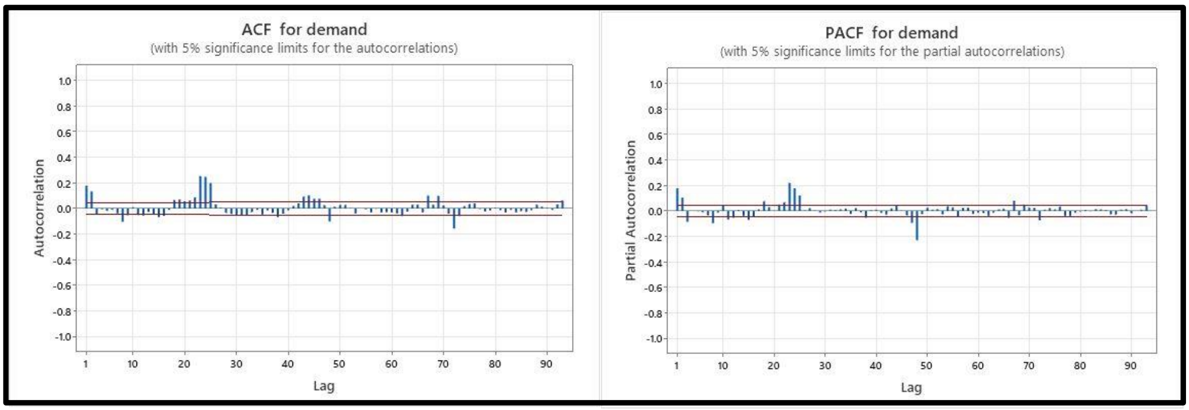

Figure (3.1): The ACF and PACF Plots for the USA Electricity demand after transformation

From figure (3.1) the PACF shows a clear spike at lag 1 and not much else until about lag 25. This is accompanied by a tapering pattern in the early lags of the ACF, a non-seasonal AR (1) may be a useful partof the model. Both show white noise which indicate that our data is stationary. The lags are multiples of 24(hourly). From the ACF plot, there's a cluster of negative spikes around lag 48 and then not much else. The PACF tapers in multiples S which equals 24, that is, the PACF has significant lags at 24, 48, 96, andso on, so, SMA (1) is suggested, where D=1. The SARIMA model of order (1,0,1) (1,1,1)24 can be an initial guess.

(2) Parameter estimation:

Table (3.2) shows the AIC and BIC for some models of orders, ranging from p=0,1,2, and q=0,1,2, P=0,1 and Q=0,1; where d = 0 and D = 1.

Table (3.2): AIC and BIC for different orders of the SARIA model

Model (d = 0, D = 1)	AIC	BIC
p = 2, q = 0, P = 0, Q = 1	42259.1	42282.0
p = 2, q = 1, P = 0, Q = 1	42484.7	42513.3
p = 2, q = 2, P = 0, Q = 0	42666.1	42694.8
p = 2, q = 0, P = 0, Q = 0	42711.6	42728.8
p = 2, q = 1, P = 0, Q = 0	42712.3	42735.3
p = 1, q = 2, P = 0, Q = 0	42840.3	42863.3
p = 1, q = 1, P = 0, Q = 0	43330.8	43348.1
p = 1, q = 0, P = 1, Q = 0	44243.7	44261.0
p = 1, q = 0, P = 0, Q = 1	44278.2	44295.5
p = 1, q = 0, P = 0, Q = 0	44281.6	44293.1
p = 2, q = 2, P = 0, Q = 1	44288.5	44323.0
p = 0, q = 2, P = 1, Q = 0	47503.4	47526.4
p = 0, q = 2, P = 0, Q = 0	47526.2	47543.4
p = 1, q = 1, P = 1, Q = 0	49200.3	49223.3
p = 1, q = 2, P = 1, Q = 0	49360.9	49389.6
p = 0, q = 0, P = 1, Q = 1	52453.2	52470.4
p = 0, q = 0, P = 0, Q = 1	52516.2	52527.7
p = 0, q = 0, P = 1, Q = 0	52646.5	52658.0
p = 2, q = 2, P = 1, Q = 0	68595.5	68630.0
p = 2, q = 1, P = 1, Q = 0	84776.1	84804.8
p = 2, q = 0, P = 1, Q = 0	84861.7	84884.7

From table (3.2), the SARIMA model of order (2, 0, 0) (0, 1, 1)24 shows the best fit as it has the smallest AIC and BIC. All the parameter coefficients of the model were estimated using maximum likelihood method. The SARIMA model can be written as follows:

$$(1 - 1.7169B)\ (1 - 0.7335B^{2})\ (1 - 0.7925B^{24})\ (1 - B^{1})\ (1 - B^{24})Y_t = 0.2506 + \varepsilon_{t-1} \qquad (8)$$

(3) Diagnostic checking:

To check the significance of the model parameters estimates, the t-test is conducted, where:

H_0: Coefficient is not significant

H_1: Coefficient is significant

Table (3.3): t-test of parameters' significance results

Type	Coefficient	SE coefficient	T-value	P-value
AR(1)	1.7169	0.0126	136.51	0.000
AR(2)	-0.7335	0.0126	-58.38	0.000
SMA(24)	0.7925	0.0130	61.20	0.000
Constant	0.2506			

From table (3.3), the p-values of all model coefficients are less than 0.05, and therefore we can reject H_o which means that all coefficients are significant. So, the model is appropriate.

The first two terms capture the autoregressive (AR) component of the model, while the third term represents the seasonal moving average (SMA) components. The constant term, 0.2506, is added to the right side of the equation.

3.2 Testing the Forecasting Power:

The second week of July was forecasted using the estimated model in order to evaluate the forecasting abilityof the model. The results are in table (3.4).

Table (3.4): Forecasting accuracy measures for the SARIMA model for one day, three days, and a week

Duration	MAPE	MAE	MPE	RMSE
One day	3.307801	18761.03	3.290144	24396.34
Three days	3.510059	18548.46	3.504173	21531.58
One week	3.474097	18079.58	1.164671	21540.77

From table (3.4), a MAPE value of 3.307801% for a day suggest that, on average, the forecasted valuesdeviate from the actual values by approximately 3.3%, for three days, the MAPE value is 3.5150059%, implying that the forecasted values deviate from the actual values by an average of 3.5%, finally, for a week, a MAPE value of 3.474097% suggest that, on average, the forecasted values deviate by approximately 3.47% from the actual values. In many forecasting applications, a MAPE below 5% is considered acceptable, so a value of 3.3%, 3.5%, and 3.47% would generally be considered good [23].

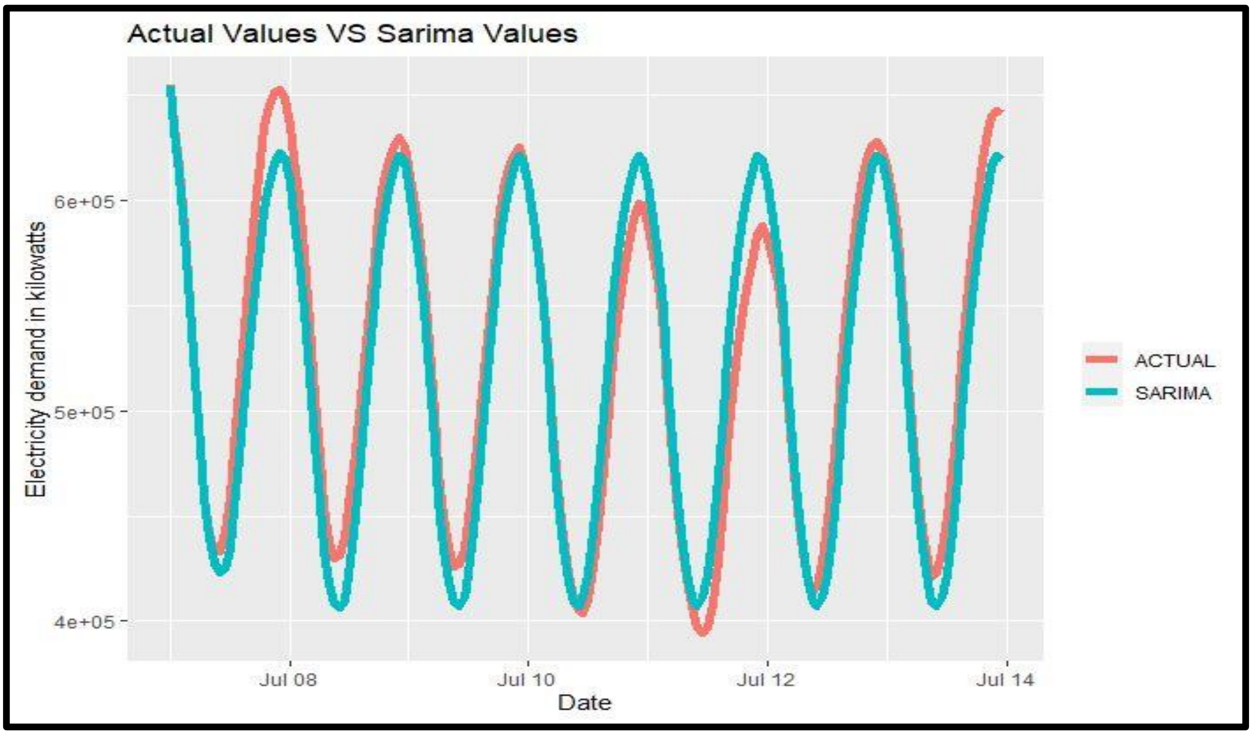

Figure (3.2): Actual vs Estimated electricity demand values for the second week of July using the SARIMA model

From figure (3.2) it can be seen that the forecasted values are close to the actual values of the second week of July; however, there is a noticeable difference between the forecasted and actual values of the electricity demand during the weekend days.

Table (3.5): 95% confidence interval values obtained by the SARIMA model for electricity demand by kilowatts compared to the actual values

Day and time	Forecasted value	Lower bound value	Upper bound value	Actual value
2021-07-07 15:00:00	505575.874	452382.091	558769.657	539058
2021-07-11 10:00:00	406977.247	344557.929	469396.566	396743
2021-07-13 16:00:00	527258.737	462361.099	592156.376	551249

As seen in table (3.5), three random days were selected from the full table of the actual values, the forecasted values and their 95% confidence intervals. It is shown that the forecasted values lie within the 95% confidence interval lower and upper bounds.

Chapter Four: The Prophet Model Analysis

The aim of this chapter is using the Prophet Forecasting Model to forecast the second week of July by dealing with the seasonality and the holiday effect to the model.

4.1 The Prophet Model

The Facebook Prophet Forecasting Model (PFM) is an additive model introduced in 2018 by Sean J. Taylor and Benjamin Letham who used the Prophet model on the data of the daily number of events created on Facebook [25]. This model is designed to take seasonality, trend and holiday effect into consideration, unlike the SARIMA model which does not consider the holiday effect. It takes the following form:

$$Y_t = (t) + s(t) + h(t) + \varepsilon_t \qquad (9)$$

Where, "Y_t" is the forecasted value, g(t) is the trend and the non-seasonal effect, s(t) is the seasonality (yearly, monthly, or daily), h(t) represents the holiday outliers, and the ε_t is the error term at time t. An essential parameter in the Prophet model is change-points, the number of change-points can be specified. The selection process for the number of change-points is done through plotting a large number of them, and then using L1 regularization to choose the few points that will be employed in the model, the L1 regularization is used to only choose the significant change-points. The L1 regularization formula is as follows:

$$L(x, y) = \sum_{i=1}^{n} (y_i - h_\Theta(x_i))^2 + \lambda \sum_{i=1}^{n} |\Theta_i| \qquad (10)$$

Where, the x and y represent the coordinates of the change-points, the term $\boldsymbol{\sum_{i=1}^{n} (y_i - h_\theta(x_i))^2}$ is the loss function, defined as the difference between the forecasted and actual values, squared. The term $\boldsymbol{\lambda \sum_{i=1}^{n} |\theta_i|}$ is added to ensure that the weights are smaller, to avoid overfitting, $\boldsymbol{\lambda}$ determines the weights, where a low value can cause a high bias and a high value can lead to under-fitting; a balance mustbe achieved, which is determined by the Prophet model [21].

4.2 Seasonality in The Prophet Model

The following plot shows the seasonality of the electricity demand, using the Prophet model.

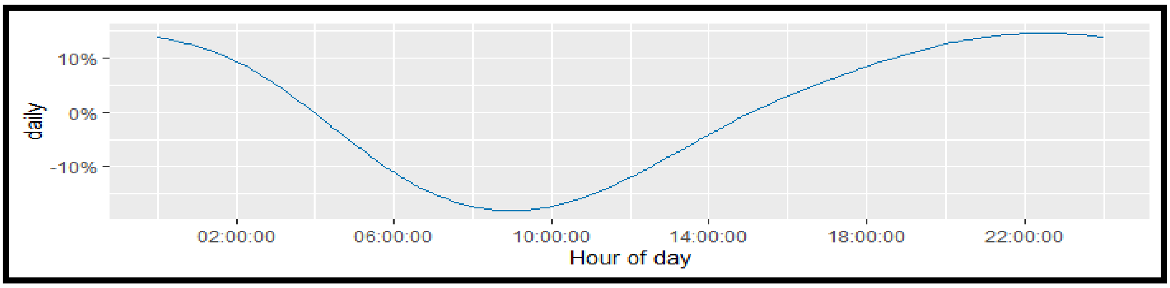

Figure (4.1): Daily Electricity Demand using the Prophet model

Figure (4.1) shows how the demand starts to decrease starting from 12:00AM until 9:00AM, where the demand is decreased by 10%, until it reaches 10:00AM, where the demand starts increasing again until it reaches its maximum at 10:00PM where the demand is increased by 10%.

To further explore the holiday effect, the model was run on 3 national holidays in the United States of America which are: Memorial Day, Juneteenth, and Independence Day, which had the following dates:

Table (4.1): National Holidays Dates in the United States of America

Holiday	Date
Memorial Day	2021-05-31
Juneteenth	2021-06-19
Independence Day	2021-07-05

The holiday effect is highlighted in the following plot:

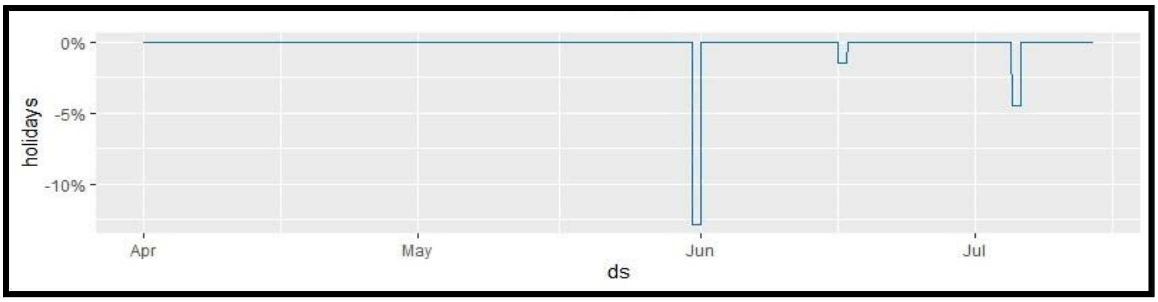

Figure (4.2): The holiday Effect on Electricity Demand using the Prophet model

Figure (4.2) shows that on Memorial Day there is a 10% decrease in electricity demand which suggest that the demand time series typically experiences a decrease of 10% during memorial day, which indicates that this holiday has a dampening effect on electricity demand. Furthermore, Independence Day shows a 5% decrease as well, also indicating that the electricity demand time series experiences a 5% decrease during this holiday, and that Independence Day also has a dampening effect on electricity demand, resulting in lower demand values. On Juneteenth, the plot shows a 1.5% decrease as well, implying a 1.5% decrease in the electricity demand during this holiday, which causes a dampening effect on the demand.

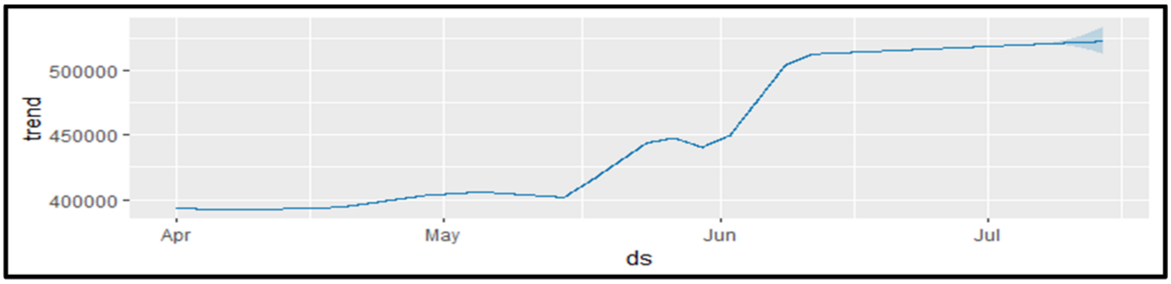

Figure (4.3): Electricity Demand time series Plot in kilowatts

Figure (4.3) shows the electricity demand throughout the months of April, May, June, July and the first week of July, it shows how the demand starts to increases from mid-May and continues to increase until the first week of July. It also shows how the demand is affected by the mentioned holidays, where varies on Memorial Day which is on May 31st, a decrease in electricity demand can be seen, on Independence day which is on July 5th, and on Juneteenth, June 19th, not much change can be seen, which is due to the fact that little change took place in the demand, with a value of change equal to -5% and -1.5%, respectively.

4.3 Forecasting

The Prophet model shows a MAPE value for a single day of 5.763543%, which suggests that on average, the forecasted values deviate by approximately 5.8% from the actual values, for three days, the MAPE is 4.012849%, implying that the forecasted values deviate from the actual values on average by approximately 4%, finally, for one week, the MAPE value is 3.958558%, which also suggests that on average, the forecasted values deviate from the actual ones by 4.9%.

Table (4.2): Forecasting accuracy measures for the Prophet model for one day, three days, and a week

Duration	MAPE	MAE	MPE	RMSE
One day	5.763543	34295.01	5.763543	40620.6
Three days	4.012849	23442.35	3.80376	29560.73
One week	3.958558	21258.41	0.7999412	27779.97

The forecasted model graph takes the form:

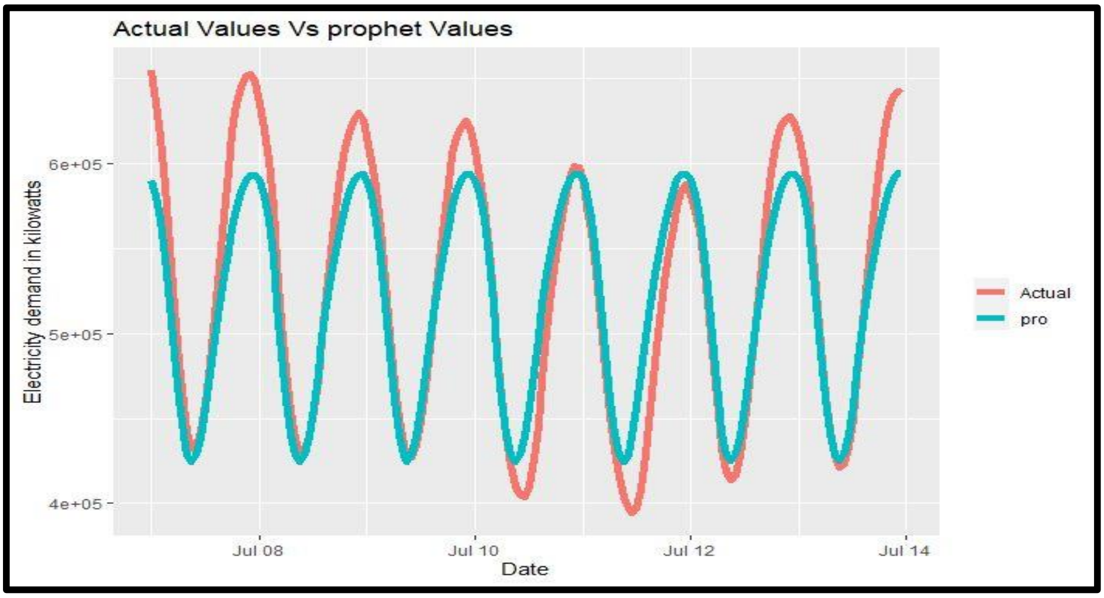

Figure (4.4): Actual vs Estimated electricity demand values for the second week of July using the Prophet model

From figure (4.4), it can be seen that the forecasted values are significantly deviated from the actual ones.

Table (4.3): 95% confidence interval values obtained by the Prophet model for electricity demand by kilowatts compared to the actual values

Day and time	Forecasted value	Lower bound value	Upper bound value	Actual value
2021-07-07 12:00:00	589776.026	553812.659	626942.293	654943
2021-07-09 1:00:00	581718.5879	544798.0442	618426.604	602120
2021-07-10 08:00:00	428462.133	391412.626	465345.879	422120

As seen in table (4.3), three forecasted values were randomly selected, two of the three values (the second and the third) lie in the 95% confidence interval, while only one doesn’t (the first one).

Chapter Five: Comparative Analysis

After applying both models, the Prophet and SARIMA, it is necessary to compare between them to find which one has better forecasting power. The comparison will include values of the MAPE, RMSE, and MAE, and the graphs for the fitted values against the actual ones.

5.1 Key Performance Indicator (KPI) values

Table (5.1): Comparing Key Performance values for the SARIMA model and the Prophet model

	SARIMA	Prophet
MAPE	3.474097%	3.94168%
RMSE	21540.77	27037.93
MAE	20790.93	21343.59

5.1.1 MAPE values

As mentioned previously both modes have a value of MAPE less than 5%. That indicates that both models are good for forecasting the USA's electricity demand. For the values of the MAPE; the SARIMA and the Prophet have the values of 3.474097% and 3.94168%, respectively. This shows that the SARIMA model has a higher forecasting quality than the Prophet model as it has a lower MAPE value. Although the SARIMA model is better, it is not very different from the Prophet model; since the difference between the two MAPE values is only 0.467583%.

5.1.2 RMSE values

The values for RMSE for the SARIMA and the Prophet are 21540.77 and 27037.93, respectively. This also indicates that the SARIMA model is better at forecasting electricity demand in the USA.

5.1.3 MAE values

The values of the MAE, 20790.93 and 21343.59 also show how better the SARIMA model is, because it has the lower value of MAE.

5.2 Fitted values from both models and actual values

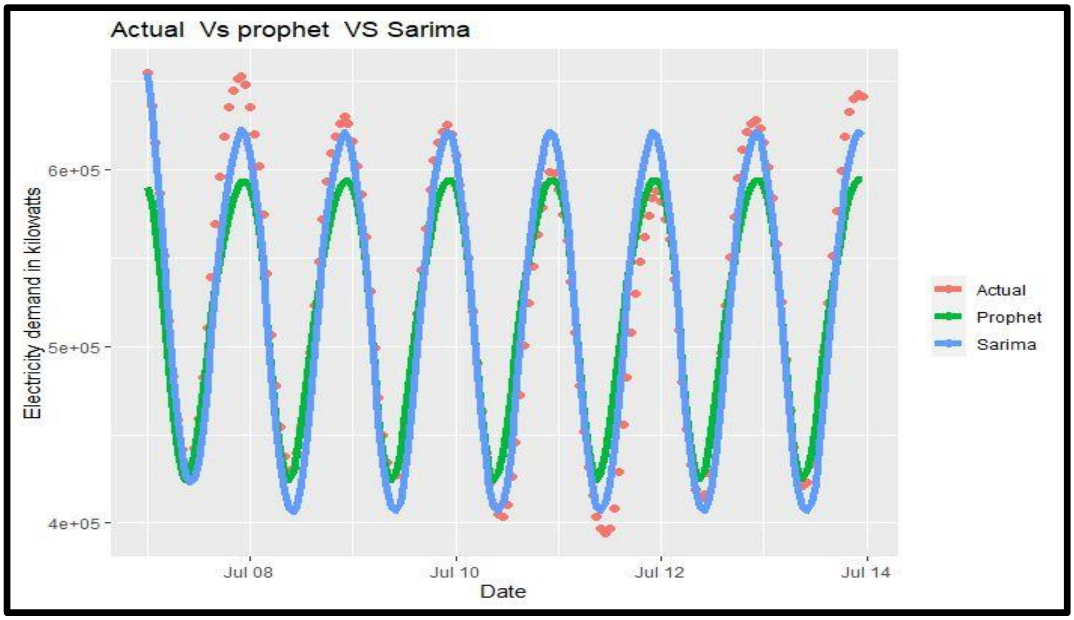

Figure (5.1): The SARIMA and the Prophet Models' Fitted Vs Actual Values

Figure (5.1) shows that the SARIMA model is a better fit for the electricity demand. Although the difference in Key Performance Indicators (KPI) values is not much, the graphs show a larger difference in the fitting process.

5.3 Conclusion and recommendations

To conclude, this paper's aim was to compare the forecasting accuracy of the selected two models- the SARIMA and the Prophet- using the data of electricity demand in the USA months of April, May, June, the first week of July, and using these months to forecast the second week of July. The results of all used criteria have shown that the SARIMA model was better than Prophet Model in terms of the forecasting accuracy, where the MAPE value for the SARIMA model is 3.47% while for the Prophet model it was 3.95%, this difference is numerically small, however, its actual magnitude is far more apparent visually, using the plots to compare both models, where it is shown that the SARIMA model captures the actual data points better, meanwhile the Prophet model is more diverted. It can also be concluded that the SARIMA model would perform better on the short run, while the Prophet model is preferred for the long run.

Recommendations include:

- Employing the used models on a larger dataset that includes more than 4 months, as well as expanding the time range of the data to include more historical observations. Increasing the amount of available data can improve the models' ability to capture patterns and make accurate forecasts.
- Combine the predictions of multiple models to potentially improve the accuracy of the forecasting. For example, we can use model averaging to leverage the strengths of both the SARIMA and the Prophet model.
- Explore more advanced time series forecasting models that may be suitable for the used data such as the Recurrent Neural Networks (RNN), the Artificial Neural Networks model (ANN), or The Holt-Winters model.

References

1- Almazrouee, A., Almeshal, M., Almutairi, et al. (2020). Forecasting of Electrical Generation Using Prophet and Multiple Seasonality of Holt-Winters Models: A Case Study of Kuwait. A*pplied Sciences*, 10(23):8412, pp. 9-14.

2- Camara, A., Feixing, W., & Xiuqin, L. (2016). Energy Consumption Forecasting Using Seasonal ARIMA with Artificial Neural Networks Models. International Journal of Business and Management, Vol. 11, pp. 240-242.

3- Congressional Research Service. (2021). U.S. Energy in the 21st Century: A Primer.

4- Chadalavada, R., Raghavendra, S. & Rekha, V. (2020). Electricity requirement prediction using time series and Facebook's PROPHET, Indian Journal of Science and Technology, vol. 13, pp. 4643-4644.

5- Chaturvedi, S., Rajasekar, E., Natarajan, McCullen, S. (2022). A comparative assessment of SARIMA, LSTM RNN and Fb Prophet models to forecast total and peak monthly energy demand for India, Energy Policy, Elsevier, vol. 168(C).

6- Chicco D., Warrens MJ., Jurman G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. PeerJ Comput. Sci. 7:e623 DOI 10.7717/peerj-cs.623, pp. 4-6.

7- Chodakowska, E., Nazarko, J. & Nazarko, L. (2021). ARIMA Models in Electrical Load Forecasting and Their Robustness to Noise, Energies 14(23), pp. 17-18.

8- Fathin, M., Widhiyasana, Y., & Syakrani, N. (2021). Model for Predicting Electrical Energy Consumption Using ARIMA Method. Advances in Engineering Research, volume 207, pp. 301-303.

9- Henzel, J., Wróbel, L., Fice, M. & Sikora, M. (2022). Energy Consumption Forecasting for the Digital-Twin Model of the Building, Energies,15(12), 431, pp. 18-19.

10- Hyndman, J., Athanasopoulos, G. (2014). Forecasting: principles and practice, 2nd Edition, University of Western Australia, pp. 84-87.

11- Kao, Y., Nawata, K. & Huang. C. (2020). Predicting Primary Energy Consumption Using Hybrid ARIMA and GA-SVR Based on EEMD Decomposition. Mathematics, *8*(10), 1722, pp. 17.

12- Leung, A. (2020). A Multivariate Model for Electricity Demand using Facebook Prophet, Medium.

13- Liang, S., Deng, T., Huang, A. et al. (2023). Energy consumption prediction using the GRU-MMattention-LightGBM model with features of Prophet decomposition, PLOS ONE 18(1), pp. 17.

14- M, S., Baby, M., & Ponraj A. (2020). Analysis of energy consumption using RNN-LSTM and ARIMA Model, Journal of Physics: Conference Series, 1716, pp. 11.

15- Miao, J. (2015). The Energy Consumption Forecasting in China Based on ARIMA Model. Atlantis Press, pp. 39-42.

16- Mohamed, Z. (2004), Forecasting Electricity Consumption: A Comparison of Growth Curves, Econometric and ARIMA Models for Selected Countries and World Regions. University of Canterbury, Electrical and Computer Engineering, pp. 331-335.

17- Morrison B. & Maas, B. (1992). Ninety Years of U.S. Household Energy History: A Quantitative Update, University of Minnesota, American Council for an Energy-Efficient Economy (ACEEE), pp. 10.126-10.128.

18- *Our World in Data*, (2022), Electricity Demand, 1990 to 2022 accessed 1/3/2023, {https://ourworldindata.org/grapher/electricity-demand?tab=chart&country=USA~GBR~FRA~DEU~IND~BRA}.

19- Pay, H. (2006). Comparing linear and nonlinear forecasts for Twain's electricity consumption, Energy 31(12).

20- Pełka, P. (2023). Analysis and Forecasting of Monthly Electricity Demand Time Series Using Pattern-Based Statistical Methods, Energies 16(2), pp. 19-20.

21- Rafferty, G. (2021). Forecasting Time Series Data with Facebook Prophet. Packt Publishing, pp. 24-26.

22- Shen, J.,Valagolam, D. & McCalla, S. (2020). Prophet forecasting model: a machine learning approach to predict the concentration of air pollutants (PM2.5, PM10, O3, NO2, SO2, CO) in Seoul, South Korea, PeerJ 8:e9961, pp. 5-6.

23- Swanson, David A. (2014). On the Relationship among Values of the same Summary Measure of Error when used across Multiple Characteristics at the same point in time: An Examination of MALPE and MAPE. Review of Economics and Finance, 5(1).

24- T., B., (2021). How to Remove Non-Stationarity in Time Series Forecasting, The Medium.

25- Taylor, S., Letham, B., (2018), Forecasting Scale, American Statistician, vol 72, pp 37-45

26- *U.S Energy Information Administration*, (March 2023), What is U.S. electricity generation by energy source?, accessed 1/3/2023 { https://www.eia.gov/tools/faqs/faq.php?id=427&t=3 }

27- *U.S Energy Information Administration*. (July 2022), Electricity explained in the United States, accessed 1/3/2023{ https://www.eia.gov/energyexplained/electricity/electricity-in-the-us.php }

28- Wang, J., Li, F., Cui, H. *et al.* (2022). Electricity consumption variation versus economic structure during COVID-19 on metropolitan statistical areas in the US. *Nat Commun* 13, 7122, pp. 1-2.

29- Young, W. (1977). The Box-Jenkins approach to time series analysis and forecasting: principles and applications, pp. 129-143.

Appendix

The SARIMA forecasted values and the actual values

Table (1): SARIMA Fitted Values and Actual Values

time	Actual	Forecasted
7/7. 12:00 AM	654943	653790.686
7/7. 01:00 AM	636293	633593.676
7/7. 02:00 AM	615384	611961.898
7/7. 03:00 AM	586346	583676.688
7/7. 04:00 AM	550996	550014.46
7/7. 05:00 AM	514454	514908.382
7/7. 06:00 AM	482817	483413.593
7/7. 07:00 AM	458695	458423.411
7/7. 08:00 AM	441983	439662.116
7/7. 09:00 AM	432551	427468.075
7/7. 10:00 AM	433090	423131.501
7/7. 11:00 AM	442172	424893.672
7/7. 12:00 PM	458932	434695.19
7/7. 01:00 PM	482596	453079.226
7/7. 02:00 PM	510223	477522.318
7/7. 03:00 PM	539058	505575.874
7/7. 04:00 PM	569090	533255.019
7/7. 05:00 PM	596277	557909.376
7/7. 06:00 PM	618662	578689.171
7/7. 07:00 PM	635420	595540.158
7/7. 08:00 PM	644899	607728.342
7/7. 09:00 PM	651194	617365.823
7/7. 10:00 PM	652980	622697.998
7/7. 11:00 PM	648059	619954.464
8/7. 12:00 AM	635189	608438.26
8/7. 01:00 AM	619997	591140.637
8/7. 02:00 AM	602181	572531.362
8/7. 03:00 AM	574821	547320.097
8/7. 04:00 AM	541448	516724.443
8/7. 05:00 AM	506744	484630.102
8/7. 06:00 AM	477633	456054.68
8/7. 07:00 AM	454414	433862.599
8/7. 08:00 AM	437653	417756.696
8/7. 09:00 AM	429602	408060.258
8/7. 10:00 AM	430806	406053.815
8/7. 11:00 AM	440035	409973.477
8/7. 12:00 PM	455041	421758.418

8/7. 01:00 PM	473380	441953.426
8/7. 02:00 PM	496221	468039.098
8/7. 03:00 PM	523287	497572.808
8/7. 04:00 PM	547628	526577.098
8/7. 05:00 PM	572207	552410.063
8/7. 06:00 PM	593228	574231.117
8/7. 07:00 PM	609346	591995.639
8/7. 08:00 PM	618813	604979.457
8/7. 09:00 PM	626326	615304.497
8/7. 10:00 PM	629981	621225.83
8/7. 11:00 PM	626308	618982.448
9/7. 12:00 AM	616169	607886.408
9/7. 01:00 AM	602120	590937.536
9/7. 02:00 AM	586141	572613.673
9/7. 03:00 AM	561793	547632.022
9/7. 04:00 AM	530960	517217.179
9/7. 05:00 AM	498811	485261.278
9/7. 06:00 AM	471117	456787.809
9/7. 07:00 AM	449438	434666.535
9/7. 08:00 AM	434419	418605.108
9/7. 09:00 AM	425890	408931.134
9/7. 10:00 AM	427073	406928.986
9/7. 11:00 AM	434423	410838.176
9/7. 12:00 PM	448698	422600.862
9/7. 01:00 PM	468183	442764.437
9/7. 02:00 PM	491490	468811.751
9/7. 03:00 PM	518552	498302.111
9/7. 04:00 PM	543155	527259.703
9/7. 05:00 PM	566847	553044.005
9/7. 06:00 PM	588342	574815.583
9/7. 07:00 PM	605424	592530.758
9/7. 08:00 PM	615051	605466.121
9/7. 09:00 PM	621640	615744.2
9/7. 10:00 PM	625089	621620.528
9/7. 11:00 PM	620287	619334.443
10/7. 12:00 AM	608078	608198.241
10/7. 01:00 AM	591452	591211.905
10/7. 02:00 AM	574306	572853.356
10/7. 03:00 AM	550029	547839.821
10/7. 04:00 AM	519588	517395.868
10/7. 05:00 AM	490550	485413.568
10/7. 06:00 AM	462958	456916.311
10/7. 07:00 AM	439750	434773.742
10/7. 08:00 AM	422120	418693.377
10/7. 09:00 AM	410322	409002.673

10/7. 10:00 AM	404695	406985.849
10/7. 11:00 AM	403675	410882.258
10/7. 12:00 PM	410219	422633.901
10/7. 01:00 PM	426062	442788.014
10/7. 02:00 PM	446059	468827.296
10/7. 03:00 PM	472748	498310.91
10/7. 04:00 PM	500152	527262.904
10/7. 05:00 PM	524629	553042.626
10/7. 06:00 PM	545522	574810.519
10/7. 07:00 PM	563373	592522.794
10/7. 08:00 PM	578377	605455.937
10/7. 09:00 PM	591158	615732.383
10/7. 10:00 PM	598567	621607.577
10/7. 11:00 PM	598017	619320.781
11/7. 12:00 AM	588367	608184.223
11/7. 01:00 AM	574448	591197.82
11/7. 02:00 AM	559721	572839.443
11/7. 03:00 AM	536563	547826.269
11/7. 04:00 AM	507930	517382.823
11/7. 05:00 AM	477857	485401.141
11/7. 06:00 AM	451428	456904.583
11/7. 07:00 AM	431588	434762.766
11/7. 08:00 AM	415432	418683.184
11/7. 09:00 AM	403580	408993.277
11/7. 10:00 AM	396743	406977.247
11/7. 11:00 AM	394037	410874.436
11/7. 12:00 PM	396902	422626.834
11/7. 01:00 PM	408423	442781.671
11/7. 02:00 PM	429083	468821.64
11/7. 03:00 PM	455493	498305.9
11/7. 04:00 PM	482190	527258.496
11/7. 05:00 PM	507690	553038.775
11/7. 06:00 PM	529847	574807.181
11/7. 07:00 PM	547922	592519.924
11/7. 08:00 PM	562102	605453.492
11/7. 09:00 PM	574002	615730.32
11/7. 10:00 PM	583829	621605.856
11/7. 11:00 PM	587707	619319.364
12/7. 12:00 AM	581902	608183.075
12/7. 01:00 AM	571921	591196.908
12/7. 02:00 AM	560493	572838.736
12/7. 03:00 AM	537965	547825.739
12/7. 04:00 AM	509388	517382.446
12/7. 05:00 AM	479502	485400.893
12/7. 06:00 AM	452958	456904.443

12/7. 07:00 AM	433292	434762.716
12/7. 08:00 AM	419333	418683.208
12/7. 09:00 AM	413573	408993.36
12/7. 10:00 AM	416411	406977.376
12/7. 11:00 AM	428508	410874.6
12/7. 12:00 PM	447937	422627.024
12/7. 01:00 PM	470913	442781.88
12/7. 02:00 PM	496190	468821.86
12/7. 03:00 PM	523338	498306.126
12/7. 04:00 PM	550221	527258.722
12/7. 05:00 PM	573429	553038.999
12/7. 06:00 PM	595259	574807.399
12/7. 07:00 PM	611232	592520.134
12/7. 08:00 PM	621342	605453.692
12/7. 09:00 PM	626010	615730.508
12/7. 10:00 PM	627985	621606.033
12/7. 11:00 PM	623411	619319.528
13/7. 12:00 AM	615363	608183.226
13/7. 01:00 AM	601012	591197.046
13/7. 02:00 AM	584074	572838.862
13/7. 03:00 AM	558066	547825.853
13/7. 04:00 AM	525142	517382.548
13/7. 05:00 AM	492160	485400.984
13/7. 06:00 AM	464092	456904.524
13/7. 07:00 AM	442889	434762.787
13/7. 08:00 AM	428157	418683.269
13/7. 09:00 AM	421094	408993.413
13/7. 10:00 AM	423357	406977.422
13/7. 11:00 AM	433898	410874.64
13/7. 12:00 PM	449832	422627.058
13/7. 01:00 PM	473077	442781.907
13/7. 02:00 PM	497079	468821.883
13/7. 03:00 PM	524363	498306.144
13/7. 04:00 PM	551249	527258.737
13/7. 05:00 PM	576485	553039.01
13/7. 06:00 PM	599580	574807.408
13/7. 07:00 PM	618750	592520.14
13/7. 08:00 PM	632433	605453.696
13/7. 09:00 PM	640214	615730.51
13/7. 10:00 PM	642831	621606.034
13/7. 11:00 PM	641177	619319.528

The Prophet model forecasted values and the actual values

Table (2): Prophet Values and Actual Values

time	Actual	Forecasted
7/7. 12:00 AM	654943	589776.026
7/7. 01:00 AM	636293	581406.08
7/7. 02:00 AM	615384	566569.247
7/7. 03:00 AM	586346	544601.563
7/7. 04:00 AM	550996	517046.252
7/7. 05:00 AM	514454	487624.832
7/7. 06:00 AM	482817	460873.896
7/7. 07:00 AM	458695	440428.413
7/7. 08:00 AM	441983	428116.988
7/7. 09:00 AM	432551	424221.477
7/7. 10:00 AM	433090	428257.031
7/7. 11:00 AM	442172	439380.5
7/7. 12:00 PM	458932	456182.577
7/7. 01:00 PM	482596	476423.584
7/7. 02:00 PM	510223	497384.832
7/7. 03:00 PM	539058	516828.014
7/7. 04:00 PM	569090	533816.672
7/7. 05:00 PM	596277	548655.743
7/7. 06:00 PM	618662	561984.342
7/7. 07:00 PM	635420	573840.834
7/7. 08:00 PM	644899	583510.047
7/7. 09:00 PM	651194	590156.848
7/7. 10:00 PM	652980	593459.123
7/7. 11:00 PM	648059	593488.772
8/7. 12:00 AM	635189	589934.532
8/7. 01:00 AM	619997	581562.334
8/7. 02:00 AM	602181	566721.512
8/7. 03:00 AM	574821	544747.922
8/7. 04:00 AM	541448	517185.205
8/7. 05:00 AM	506744	487755.876
8/7. 06:00 AM	477633	460997.75
8/7. 07:00 AM	454414	440546.771
8/7. 08:00 AM	437653	428232.037
8/7. 09:00 AM	429602	424335.477
8/7. 10:00 AM	430806	428372.115
8/7. 11:00 AM	440035	439498.571
8/7. 12:00 PM	455041	456305.162

8/7. 01:00 PM	473380	476551.607

8/7. 02:00 PM	496221	497518.485
8/7. 03:00 PM	523287	516966.891
8/7. 04:00 PM	547628	533960.112
8/7. 05:00 PM	572207	548803.168
8/7. 06:00 PM	593228	562135.347
8/7. 07:00 PM	609346	573995.024
8/7. 08:00 PM	618813	583666.833
8/7. 09:00 PM	626326	590315.418
8/7. 10:00 PM	629981	593618.579
8/7. 11:00 PM	626308	593648.234
9/7. 12:00 AM	616169	590093.037
9/7. 01:00 AM	602120	581718.588
9/7. 02:00 AM	586141	566873.777
9/7. 03:00 AM	561793	544894.282
9/7. 04:00 AM	530960	517324.158
9/7. 05:00 AM	498811	487886.921
9/7. 06:00 AM	471117	461121.604
9/7. 07:00 AM	449438	440665.129
9/7. 08:00 AM	434419	428347.085
9/7. 09:00 AM	425890	424449.477
9/7. 10:00 AM	427073	428487.198
9/7. 11:00 AM	434423	439616.642
9/7. 12:00 PM	448698	456427.747
9/7. 01:00 PM	468183	476679.629
9/7. 02:00 PM	491490	497652.139
9/7. 03:00 PM	518552	517105.768
9/7. 04:00 PM	543155	534103.552
9/7. 05:00 PM	566847	548950.594
9/7. 06:00 PM	588342	562286.353
9/7. 07:00 PM	605424	574149.214
9/7. 08:00 PM	615051	583823.619
9/7. 09:00 PM	621640	590473.988
9/7. 10:00 PM	625089	593778.035
9/7. 11:00 PM	620287	593807.696
10/7. 12:00 AM	608078	590251.542
10/7. 01:00 AM	591452	581874.842
10/7. 02:00 AM	574306	567026.042
10/7. 03:00 AM	550029	545040.641
10/7. 04:00 AM	519588	517463.11
10/7. 05:00 AM	490550	488017.965
10/7. 06:00 AM	462958	461245.457
10/7. 07:00 AM	439750	440783.487
10/7. 08:00 AM	422120	428462.133
10/7. 09:00 AM	410322	424563.477
10/7. 10:00 AM	404695	428602.282

10/7. 11:00 AM	403675	439734.713
10/7. 12:00 PM	410219	456550.332
10/7. 01:00 PM	426062	476807.652
10/7. 02:00 PM	446059	497785.793
10/7. 03:00 PM	472748	517244.644
10/7. 04:00 PM	500152	534246.992
10/7. 05:00 PM	524629	549098.02
10/7. 06:00 PM	545522	562437.359
10/7. 07:00 PM	563373	574303.404
10/7. 08:00 PM	578377	583980.405
10/7. 09:00 PM	591158	590632.559
10/7. 10:00 PM	598567	593937.49
10/7. 11:00 PM	598017	593967.158
11/7. 12:00 AM	588367	590410.047
11/7. 01:00 AM	574448	582031.096
11/7. 02:00 AM	559721	567178.307
11/7. 03:00 AM	536563	545187.001
11/7. 04:00 AM	507930	517602.063
11/7. 05:00 AM	477857	488149.009
11/7. 06:00 AM	451428	461369.311
11/7. 07:00 AM	431588	440901.845
11/7. 08:00 AM	415432	428577.182
11/7. 09:00 AM	403580	424677.478
11/7. 10:00 AM	396743	428717.365
11/7. 11:00 AM	394037	439852.784
11/7. 12:00 PM	396902	456672.917
11/7. 01:00 PM	408423	476935.674
11/7. 02:00 PM	429083	497919.446
11/7. 03:00 PM	455493	517383.521
11/7. 04:00 PM	482190	534390.433
11/7. 05:00 PM	507690	549245.446
11/7. 06:00 PM	529847	562588.365
11/7. 07:00 PM	547922	574457.594
11/7. 08:00 PM	562102	584137.191
11/7. 09:00 PM	574002	590791.129
11/7. 10:00 PM	583829	594096.946
11/7. 11:00 PM	587707	594126.62
12/7. 12:00 AM	581902	590568.553
12/7. 01:00 AM	571921	582187.35
12/7. 02:00 AM	560493	567330.572
12/7. 03:00 AM	537965	545333.36
12/7. 04:00 AM	509388	517741.015
12/7. 05:00 AM	479502	488280.054
12/7. 06:00 AM	452958	461493.165
12/7. 07:00 AM	433292	441020.203

12/7. 08:00 AM	419333	428692.23
12/7. 09:00 AM	413573	424791.478
12/7. 10:00 AM	416411	428832.448
12/7. 11:00 AM	428508	439970.855
12/7. 12:00 PM	447937	456795.502
12/7. 01:00 PM	470913	477063.697
12/7. 02:00 PM	496190	498053.1
12/7. 03:00 PM	523338	517522.398
12/7. 04:00 PM	550221	534533.873
12/7. 05:00 PM	573429	549392.872
12/7. 06:00 PM	595259	562739.37
12/7. 07:00 PM	611232	574611.783
12/7. 08:00 PM	621342	584293.978
12/7. 09:00 PM	626010	590949.7
12/7. 10:00 PM	627985	594256.402
12/7. 11:00 PM	623411	594286.082
13/7. 12:00 AM	615363	590727.058
13/7. 01:00 AM	601012	582343.604
13/7. 02:00 AM	584074	567482.837
13/7. 03:00 AM	558066	545479.72
13/7. 04:00 AM	525142	517879.968
13/7. 05:00 AM	492160	488411.098
13/7. 06:00 AM	464092	461617.019
13/7. 07:00 AM	442889	441138.561
13/7. 08:00 AM	428157	428807.278
13/7. 09:00 AM	421094	424905.478
13/7. 10:00 AM	423357	428947.532
13/7. 11:00 AM	433898	440088.927
13/7. 12:00 PM	449832	456918.087
13/7. 01:00 PM	473077	477191.72
13/7. 02:00 PM	497079	498186.754
13/7. 03:00 PM	524363	517661.275
13/7. 04:00 PM	551249	534677.313
13/7. 05:00 PM	576485	549540.298
13/7. 06:00 PM	599580	562890.376
13/7. 07:00 PM	618750	574765.973
13/7. 08:00 PM	632433	584450.764
13/7. 09:00 PM	640214	591108.27
13/7. 10:00 PM	642831	594415.858
13/7. 11:00 PM	641177	594445.544