

Time Series, Assignment 1

Petersen Trentin (PTRTRE004)

September 22, 2024

Abstract

Within this report are the findings and conclusions made for a project conducted for Eskom, with a strong focus on the forecasting of peak daily electricity demand and the peak daily total proportion of energy unavailable due to unplanned outages (UCLF+OCLF). Using historical data from April 1, 2019 to September 13, 2023, a multitude of forecasting techniques such as ARIMA, SARIMA and GARCH models were employed to capture the trend, seasonality and volatility in the data. The analysis aims to produce valuable insights into the energy supply issues faced by South Africa. These forecasts could assist Eskom in the mitigating load shedding and unplanned outages which will ultimately result in improved national grid stability/reliability.

Section 1 - Introduction

Over the past decade, the demand for electricity in South Africa has only increased which has proven itself to be a tough challenge for Eskom, South Africa's primary supplier of power. Unplanned outages caused by ageing infrastructure and poor management combined with the sharp fluctuations in the demand for electricity result in hindrances in meeting the energy demands of South Africa as well as causing alarming rates of instability in the national power grid. This has led to an increase of the amount of load shedding hours in South Africa, even reaching a new record of consecutive days of load shedding in 2023 which had a detrimental impact on the economy of South Africa and the lives of many people. The aim of this project is to produce robust forecasts for the peak daily electricity demand as well as peak daily total proportion of energy which is unavailable due to all unplanned outages (which will be referred to as Peak Daily UCLF+OCLF). The robust forecasting of these metrics is vital to Eskom's success in reducing the frequency of load shedding and improving the stability and reliability of the national grid.

Eskom states that there are two peak electricity demand periods within a day: 6am to 9am and 5pm to 9pm. The focus of this report will be on the latter. During this period of the day is when most households in South Africa use electric appliances; at the same time there may also still be some commercial or industrial activities which overlap with this increased household demand. This spike in electricity demand, puts a lot of strain on the grid and the ageing infrastructure.

Forecasts will be made using a variety of different time series models. To capture trends and/or seasonal patterns to obtain point forecasts as well as prediction intervals, AR/MA/ARMA/ARIMA and SARIMA models will be fit to the series. ARCH/GARCH models will be used to model and forecast the volatility in the series of the peak daily demand and Peak Daily UCLF+OCLF. The forecasts produced will be a blend of short-term and medium-term forecasts which will contribute to operational decisions, grid management and resource allocation for maintenance and power generation.

The data used in the analysis is named 'ESK6816.csv' and is sourced from Eskom [4]. The dataset contains hourly observations beginning 01/04/2019 — 00 : 00 and ending 31/03/2024 — 23 : 00. For the purpose of this analysis the two primary variables of interest are *RSA.Contract.Demand* and *Total.UCLF.OCLF*. *RSA.Contract.Demand* represents the hourly average demand that Eskom must fulfill. *Total.UCLF.OCLF* represents the total proportion of Eskom's plant capacity that unavailable due to unplanned outages.

Section 2 - Questions/Hypotheses

2.1 - Hypotheses

It is hypothesized that the seasons of the year as well as the economic activity will have a large impact on the daily peak demand, causing an annual pattern to emerge. A weekly pattern is also expected due to differing demands on weekends and weekdays. Therefore seasonal models such as SARIMA models will outperform non-seasonal models such as ARIMA or ARMA models.

A slight positive linear trend in the energy unavailability due to unplanned outages is expected due to ageing infrastructure but there will be no seasonality in the series, energy unavailability is not affected by the

workdays vs. weekends/holidays.

The volatility in the energy unavailability due to unplanned generation outages is higher in the peaks of winter and summer than in transitional seasons such as spring and autumn.

2.2 - Questions

How can forecasts be produced that can be used to identify strategies that aim to minimize load shedding as well as energy unavailability during the peak demand period 5pm - 9pm.

Do the residuals of Eskom's forecasts say anything about the model/method that was used to produce the forecasts and how do forecasts align with the observed peak daily demand data?

How is the total energy unavailability due to unplanned generation affected by extreme weather conditions.

How well do ARCH/GARCH models capture the volatility in the peak daily demand series and the energy unavailability series?

Section 3 - EDA

3.1 - Data and variable names/classes

The data is read in from the ESK6816.csv dataset. The raw data consists of 22 variables and 43848 observations. Of the 22 variables we are only interested in the *Date*, *RSA.Contractd.Demand*, *RSA.Contractd.Forecast* and *Total.UCLF.OCLF*.

The data is filtered to only include the hourly observations from 5pm-9pm, this is a crucial period of the day for electricity demand and unplanned outages. The choice of filtering is applicable to both variables as they have a significant interaction during this period of the day which can have a large impact on the stability of the grid. The table below presents summary statistics of the data for the two variables of interest - *RSA.Contractd.Demand* and *Total.UCLF.OCLF*.

3.2 - Summary statisitcs and missing observations

Table 1: Summary statistics - RSA.Contractd.Demand & Total.UCLF.OCLF

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
Demand	18783.38	26702.1	28447.4	28360.85	30023.18	35004.75	1000
Energy Loss	5982.00	10532.0	12962.0	12891.27	15142.00	21308.00	1000

Table 1 shows that there are 1000 NA observations, on inspection the data shows that from the 14th of September onwards the datasets only contains values for the *Date* and *RSA.Contractd.Forecast* variables. The choice was made to truncate the data such that the final observation in the dataset had the Date and time: 2023/09/13 – 21 : 00. Thereafter the dataset contained no missing data.

Table 2: The first and last two observations with missing data.

Date	RSA.Contractd.Forecast	RSA.Contractd.Demand
2023-09-14	30836.54	NA
2023-09-14	32437.46	NA
2024-03-31	24694.54	NA
2024-03-31	22910.12	NA

3.3 - Plotting the raw data

The time plots for the *RSA.Contract.Demand* and *Total.UCLF.OCLF* variables are shown below. MA-720 (Monthly) smoothing was applied to expose the trend within the raw data.

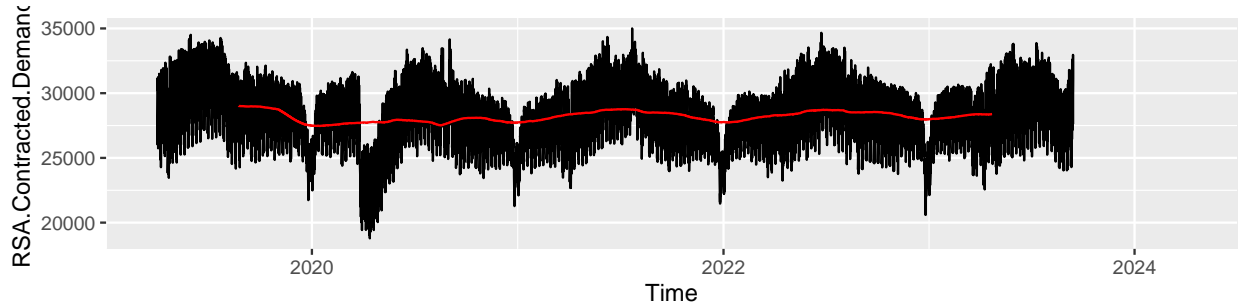


Figure 1: Time plot of the raw demand data.

The figure shows that there is some annual seasonality/cyclicality in the demand data. This can be logically explained with the seasons of the year and their correlation with electricity demand. In summer the demand is lower, then it increases until it peaks in mid-winter when the days are shorter and colder, requiring more electricity for lighting and heating. Every year there is also a sharp decline in December/January, which is investigated further in Section 3.4. The demand time plot also shows the dramatic effect that Covid had on the demand in 2020 making it an outlier amongst the other 4 years (2019, 2021, 2022, 2023).

The demand time plot has no clear trend and the variance appears to be fairly constant (and quite high) with the exception of the Covid era and the December/Christmas period which both exhibit a spike in variance.

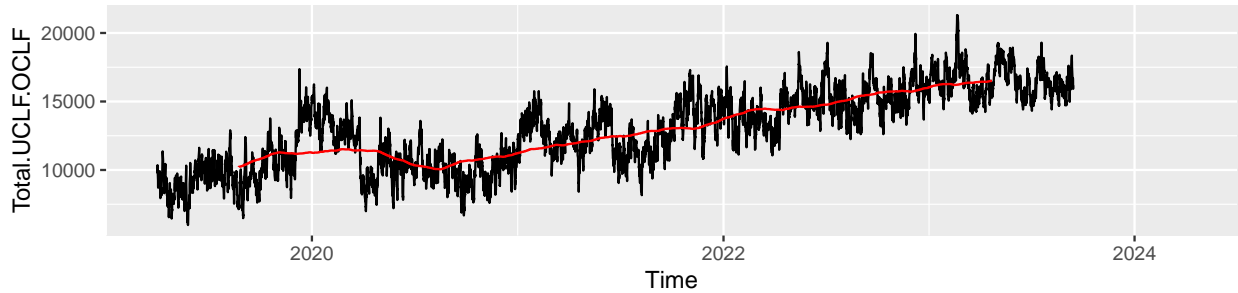


Figure 2: Time plot of the raw energy loss data.

The time plot for the proportion of total energy unavailable due to unplanned outages presents a clear positive linear trend (which can likely be attributed to the ageing infrastructure) but lacks any noticeable seasonality. The variance also appears to be constant for the majority of the time. There doesn't seem to be much correlation between the demand and the energy lost due to unplanned outages with the exception of the Covid era.

3.4 - Examining major decreases in December

It was mentioned in Section 3.3 that there appeared to be a drastic decrease in the demand around December/January every year. The time plot in the Appendix shows that the sharp decrease is due to Christmas and New years celebrations/holidays. During these times businesses will often close or scale down as people take/are given leave and thus the peak daily demand drastically decreases.

3.5 - Data Cleaning

Figure 1 shows how the demand in 2020 differed from the years 2019 and 2021-2023. Events such as the Covid pandemic are irregular occurrences and thus it is sensible to truncate the data such as to not include demand data which was greatly affected by the Covid pandemic. By comparing 2020 to 2021, it appears that by July 2020 the seasonal pattern seemed to have stabilized, therefore a choice was to truncate the data such that it begins in September of 2020 which will give us 4 full years of data. Thus observations in the final cleaned demand data will begin 2020/09/14 – 17 : 00 and run until 2023/09/13 – 21 : 00.

Much like for the demand data the effect of the Covid pandemic is reflected in the energy unavailability data. However unlike for demand data, there is no clear annual pattern or predictable trend for energy loss due to unplanned generation outages. This is most likely because it has a stronger dependence on current circumstances such as infrastructure failures, maintenance issues or other events that might affect generation capacity. Thus for forecasting the peak daily UCLF+OCLF, the data that should be used should be recent.

To ensure that the data is relevant whilst also ensuring that the forecasting models will be able to capture the underlying patterns in the data, the data is truncated to include only the past 2 years and will run from 2021/09/14 to 2023/09/13.

3.6 - Aggregating data

In order to convert the hourly demand data into daily data such that the peak daily demand can be forecasted, the hourly data is aggregated by selecting the maximum demand observed during the 5pm-9pm period each day.

The hourly energy unavailability data is converted to daily data by selecting the maximum UCLF+OCLF observed during the 5pm-9pm period each day. By aggregating the hourly data to daily data a lot of the detail within the series is maintained which makes it ideal for short-term and medium-term forecasts whilst reducing the noise present in the hourly data. Additionally by choosing to aggregate the UCLF+OCLF data in this way, the forecasts can be analysed together with the peak daily demand data, providing a more comprehensive set of forecasts for Eskom to use.

3.7 - Decomposition of transformed data

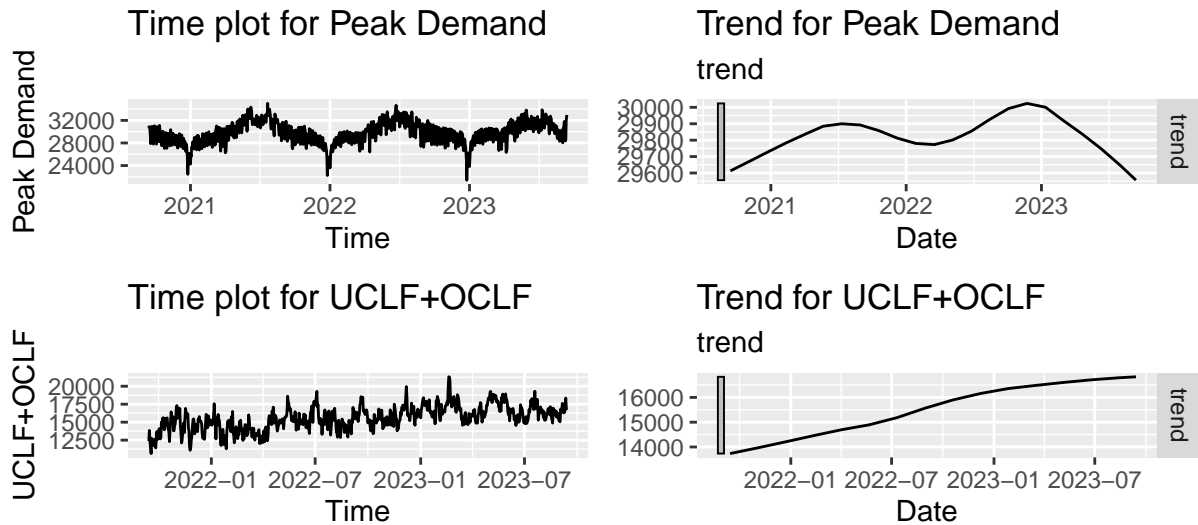


Figure 3: Time plots of cleaned/transformed data with trend plots

Figure 3 exhibits a non-linear trend in the peak daily demand data, this implies that the peak daily demand for electricity follows a more complex/cyclical pattern which could be influenced by exogenous variables. In contrast the peak daily UCLF+OCLF has a strong positive linear trend, indicating that there is a constant increase in peak daily UCLF+OCLF over time which as mentioned previously is likely due to ageing infrastructure and other factors.

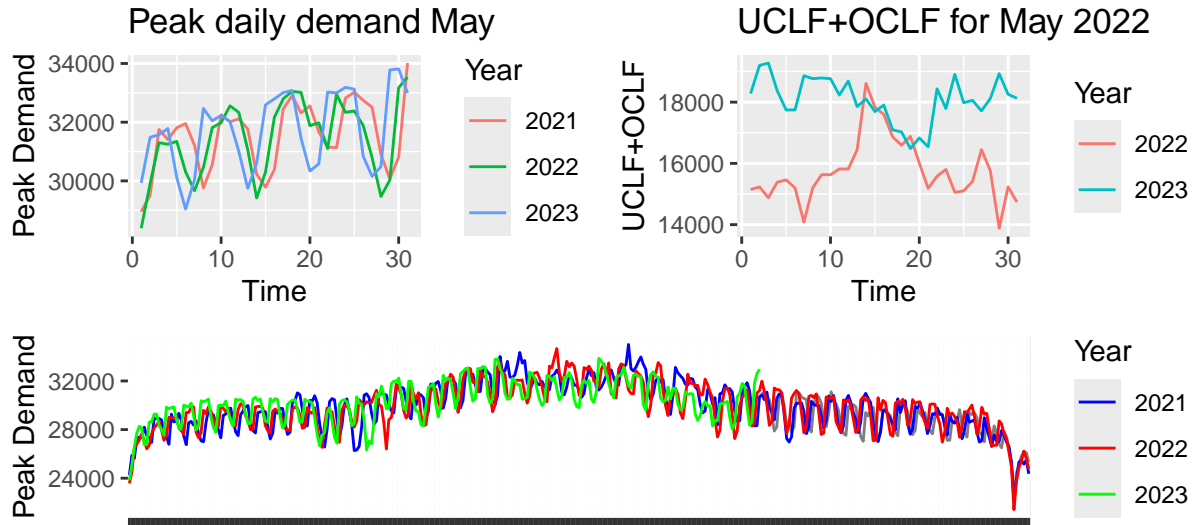


Figure 4: Weekly and annual seasonality of cleaned/transformed data.

The figure above shows that there is strong weekly seasonality in the peak daily demand data which is consistent across 2021-2023. There does not appear to be any weekly seasonality for the peak daily UCLF+OCLF data.

There is clear annual seasonality pattern for the daily peak demand data. Because only a single year of data is used for the peak daily UCLF+OCLF, it is not sensible to plot the annual seasonality.

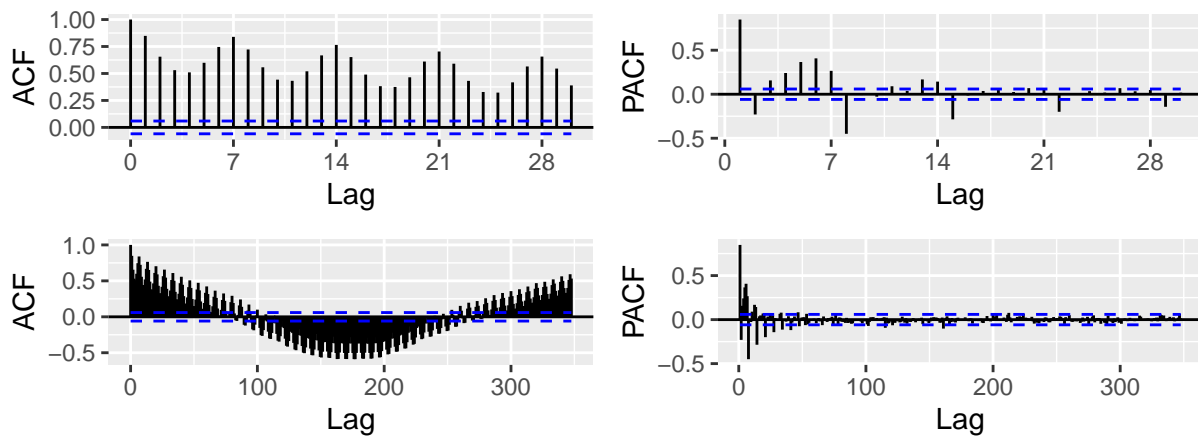


Figure 5: ACF and PACF plots of Peak Daily Demand data.

Figure 5 shows significant autocorrelation in the series up to relatively high lags in both the PACF and ACF. There are clear spikes in autocorrelation every 7 lags in both the ACF and PACF which represent the strong weekly seasonality identified in Figure 4. Figure 5 also confirms the presence of some trend in the data through the decaying pattern of the autocorrelation in the ACF, this pattern emerges as if there is a trend then observations that are close to each other in time are close in size and are therefore correlated.

The cyclicity in the ACF is representative of the annual seasonality/cyclicity in the peak daily demand data. From the ACF plot as well as the supporting time plots it can be inferred that the annual period is roughly 365 days. Figures 4 and 5 exhibit the presence of trend and strong seasonality in the peak daily demand series, therefore it is non-stationary.

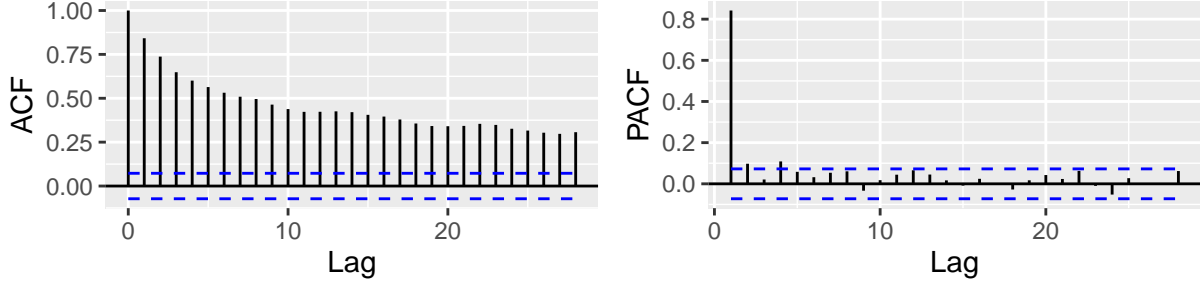


Figure 6: ACF and PACF plots of Peak Daily UCLF+OCLF data.

The slow decay of autocorrelation in the ACF plot in Figure 6 above reflects the presence of a trend in the series. This suggests that the series is non-stationary and that detrending or differencing the series will be required to make the data stationary before modelling of the data takes place.

The large spike in autocorrelation shown in the PACF plot might indicate an AR(1) process. If the spike remains after the data has been made stationary then there is strong evidence that an AR(1) component is required in the model.

Section 4 - Peak Daily Demand Model Fitting

For the purpose of the model fitting, forecasting and comparison, the data will be split up into training/test sets. The forecast horizon was set to 92 days (3 months), a decision which was based on the aims of the project which include producing short-term forecasts which will be used in handling load shedding and medium-term forecasts which will be used in grid management decisions. A forecast horizon of 3 months will also show if the models are accurately capturing the annual seasonal pattern in the data. Therefore, the training set will consist of all observations from 2020/09/14 to 2023/06/13 and the test set will be made up of the last 3 months from 2023/06/14 to 2023/09/14.

4.1 - Differencing for Non-seasonal Models

Figures 3,4 and 5 provide evidence that the peak daily demand series is non-stationary as the series does not fluctuate around a mean of zero and there is a clear presence of seasonality. Therefore for the non-seasonal models (AR, MA, ARMA and ARIMA) the series is differenced using first differences until it appears stationary.

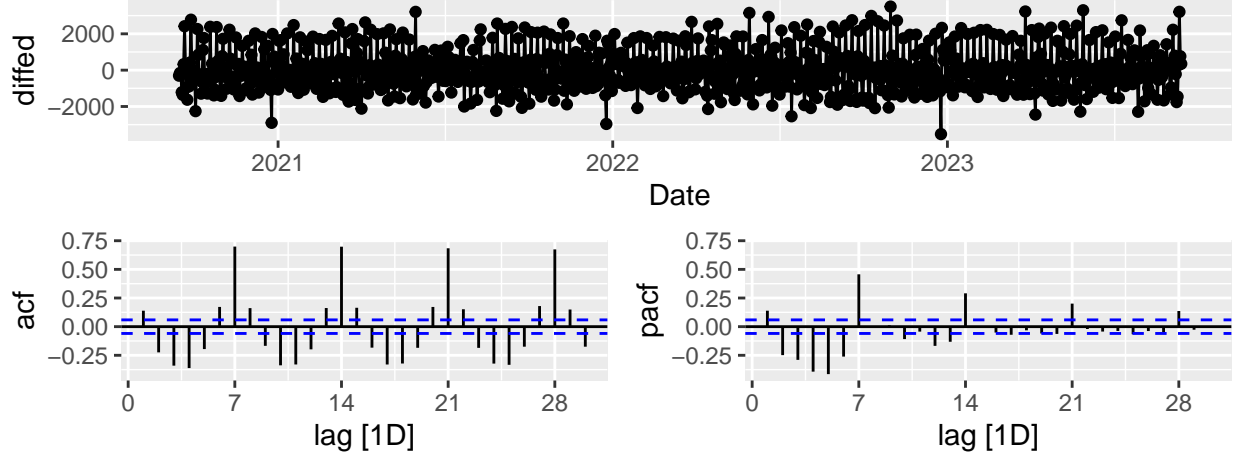


Figure 7: Differenced demand data and ACF/PACF plots (Non-seasonal).

The figure above shows data that is representative of stationary series as it fluctuates around a zero-mean and displays no heteroskedasticity.

An Augmented Dickey-Fuller test can be conducted to test for the presence of a unit root in the series.

ADF test:

H_0 : There is a unit root present in the series, i.e., the series is non-stationary.

H_1 : There is no unit root present in the series, i.e., the series is stationary.

Table 3: ADF Test Results

Statistic	P.Value	Alternative
-12.262	0.01	stationary

As shown in the table above, $p - value = 0.01$, this means that there is sufficient evidence that the series is stationary and the null hypothesis is rejected.

4.2 - AR/MA/ARMA/ARIMA (Non-seasonal) Models

To begin, simple model autoregressive, moving-average, ARMA and ARIMA models will be developed, these are relatively straight forward to fit however, given the seasonality present in the data as shown in Section 3.7, it is expected that the non-seasonal AR, MA, ARMA and ARIMA models will not fit the data adequately as they do not take the seasonal components of the data into consideration. For the sake of comprehensiveness and providing a baseline for forecasting models, non-seasonal models are included in the analysis.

Because of the theoretical limitations of the non-seasonal models, the model fitting and selection process was drastically simplified, putting the focus on the seasonal models/methods in Section 4.4. The non-seasonal models were fit using the same simple methodology:

1. Inspect the ACF and PACF plots of the stationarized data in Section 4.1.
2. Determine the relevant orders for p, d and q for a candidate model.
3. Fit multiple models by varying the relevant orders such as $p = 5, 6, 7, 8$ for an $AR(p)$ model.
4. Choose the best model by comparing AICc values. (Appendix 1.2)

Table 4: Best model specifications chosen using AICc.

AR(8)	MA(7)	ARMA(5,6)	ARIMA(5,1,4)
-------	-------	-----------	--------------

A residual analysis was conducted on the models listed above. This was comprised of interpreting the plots of the residuals (Appendix 1.3) and the results of the Ljung-Box test for autocorrelation of the residuals.

The hypotheses for the Ljung-Box test are as follows:

H_0 : There is no autocorrelation in the residuals for a fixed number of lags[3].

H_1 : There is autocorrelation present in the residuals at some lag/s l in the residuals.

Table 5: Ljung-Box Test Results

Model	Test Statistic	P Value	AC Present (Y/N)	First Significant AC at Lag
AR(8)	106.2382	0	Y	7
MA(7)	398.1555	0	Y	6
ARMA(5,6)	143.3960	0	Y	2
ARIMA(5,1,4)	272.8049	0	Y	5

The Ljung-Box test results above and the residual plots in Appendix 1.3 confirm the expectation that AR, MA, ARMA and ARIMA models will not be appropriate for modelling the given series as all models display significant autocorrelation in the residuals which indicates that there is information that the models are failing to capture. All of the non-seasonal models showed significant autocorrelation at lag 7 which indicates that the strong weekly seasonality was not captured by the models and was still present in the series. Hence the forecast plots are not closely analyzed as the non-seasonal models are inadequate for modelling this series, however, the plots can be found in Appendix 1.4.

4.3 - Differencing for Seasonal Models

As mentioned in Section 4.1 the peak daily demand data is clearly non-stationary as there is strong weekly seasonality in the series. This prompts the use of a model such as a SARIMA which takes into account the seasonal components of a series. Seasonal differencing ($m = 7$) was applied to try and make the data stationary.

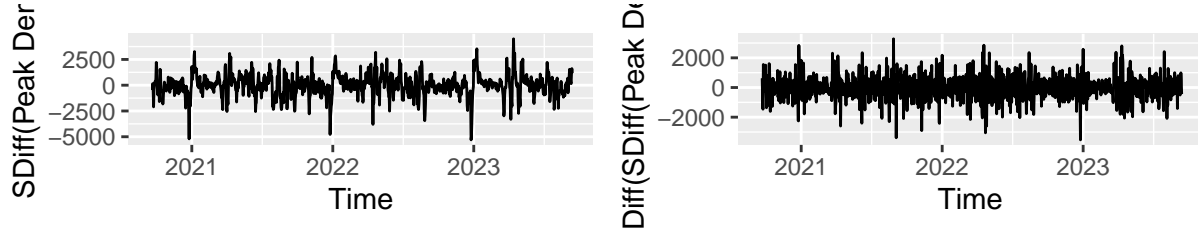


Figure 8: Differenced demand data and ACF/PACF (Seasonal)

The plot below shows that the data still somewhat exhibits the annual seasonal pattern, therefore a first difference will also be applied to the series. The seasonally and first differenced is shown in the figure below. The series appears to have been stationarized.

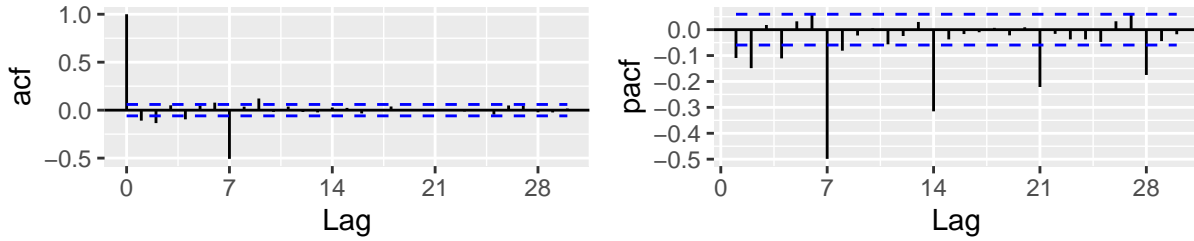


Figure 9: ACF/PACF of seasonal and first Differenced data.

Once again an ADF test (as per Section 4.1) can be performed to determine if the differenced series has reached stationarity.

Table 6: ADF Test Results

Statistic	P.Value	Alternative
-14.593	0.01	stationary

The p-value of the ADF test is $p = 0.01$, this means that there is sufficient evidence that the series is stationary and the null hypothesis is rejected.

4.4 - SARIMA (Seasonal-ARIMA) Models

Section 4.2 presented the shortcomings of applying non-seasonal models to the series as they failed to provide a good fit to the series and produced inaccurate forecasts. The focus is now on $SARIMA(p, d, q)(P, D, Q)[M]$ seasonal models. By adding seasonal components $(P, D, Q)[M]$ to the already existing $ARIMA(p, d, q)$ model framework, the aim is to capture the seasonality present in the data which should enable the development of more robust/accurate forecasts. The orders of p, d, q, P, D, Q were selected using the ACF and PACF plots in Section 4.3.

4.4.1 - SARIMA(4,1,4)(2,1,1)[7]

The spikes at lag 7 and lag 14 in the PACF plot in Figure 9 suggest that if the seasonal period is chosen to be $m = 7$ then start with $P = 2$ for the seasonal AR component as the initial candidate model. The ACF plot in Figure 9 suggests $Q = 1$ for the seasonal MA component due to the significant spike at lag 7. The significant spikes at lag 4 in the ACF/PACF of Figure 9 suggest that $p = 4$ and $q = 4$ be used as the non-seasonal AR and MA components.

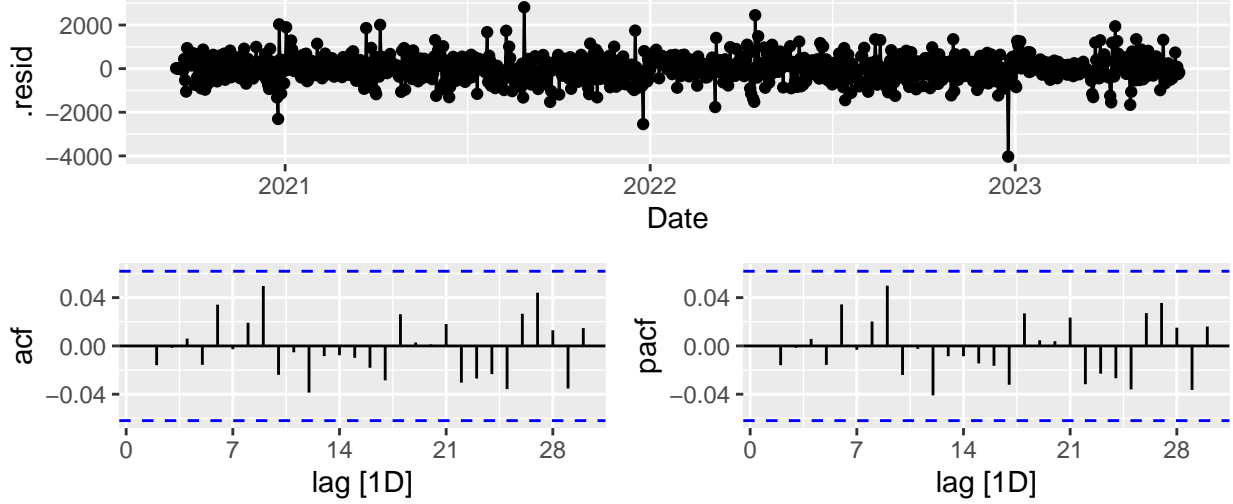


Figure 10: SARIMA(4,1,4)(2,1,1)[7] residual plot.

Table 7: Ljung-Box Test Results

LB Statistic	P-value
6.87	0.009

The residuals of the $SARIMA(4, 1, 4)(2, 1, 1)[7]$ model suggest that the model is a good fit to the data as there does not appear to be any significant autocorrelation in the ACF/PACF plots. This is supported by the results of the Ljung-Box test which show $p - value \gg 0.05$ indicating that there is sufficient evidence that there is no significant autocorrelation in the residuals and H_0 is not rejected.

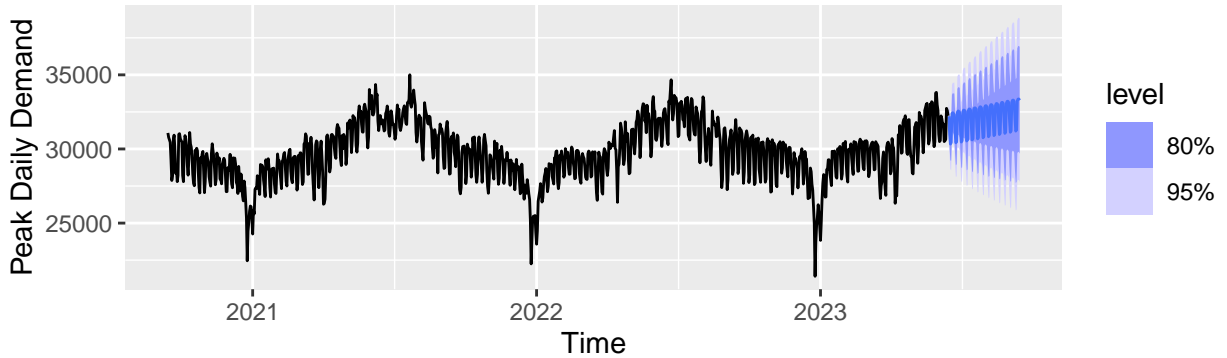


Figure 11: Forecasts for SARIMA(4,1,4)(2,1,1)[7] model.

Whilst the $SARIMA(4, 1, 4)(2, 1, 1)[7]$ model incorporates seasonal differencing as well as autoregressive and moving average components the forecasts suggest that the model is not a good fit. While the model appears to capture the weekly seasonality, it falls short in capturing the annual seasonality that is present in the data, thus the forecasts fail to align with what is expected of the forecasted values given the historical data. This suggests the need for an additional component which will capture the annual seasonal pattern.

4.4.2 - $SARIMA(3,0,3)(0,1,0)[365]$

An alternative approach to modelling the series is to fit another SARIMA model but set the seasonal period to $m = 365$ in order to try and capture the annual seasonality that the $SARIMA(4, 1, 4)(2, 1, 1)[7]$ model failed to capture. It is noted that this will very likely affect the model's ability to capture the weekly seasonality. The choice of the p and q parameters arises from seasonally differencing the data with $m = 365$ and then applying a further first difference which produced the ACF/PACF plots shown in the figure below.

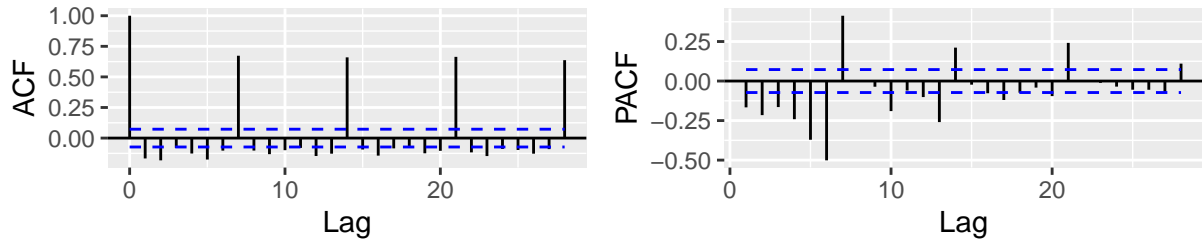


Figure 12: ACF/PACF of Seasonally and First Differenced data ($m=365$)

The ACF/PACF plots above suggest the $p = 7$ and $q = 7$, however, due to the high computational complexity of fitting this model with a $m = 365$, $p = 3$ and $q = 3$ are chosen as the orders for model. The model fit to the series is specified as a $SARIMA(3, 0, 3)(0, 1, 0)[365]$.

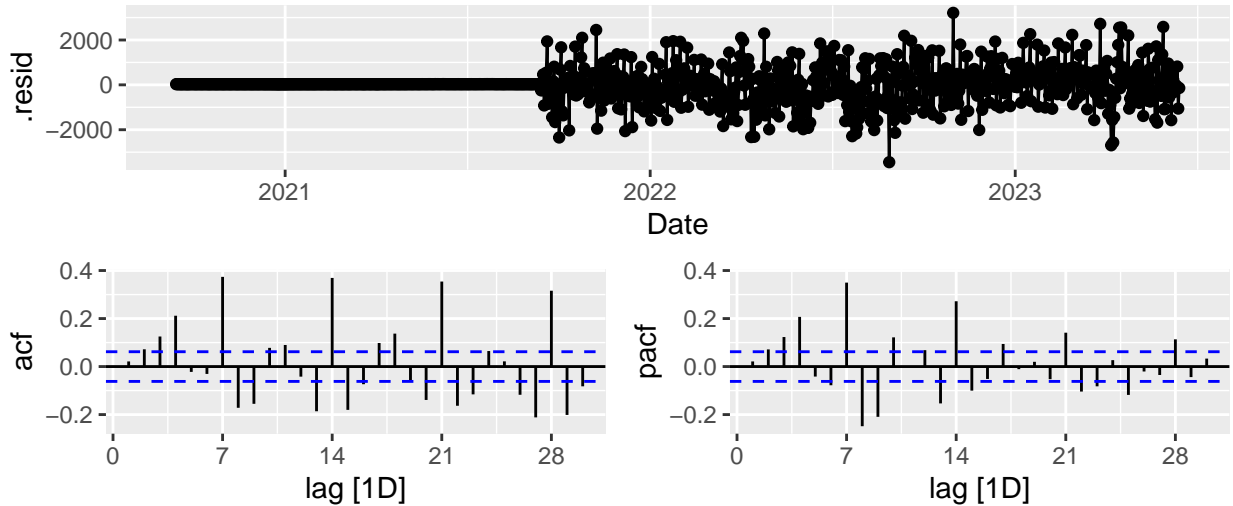
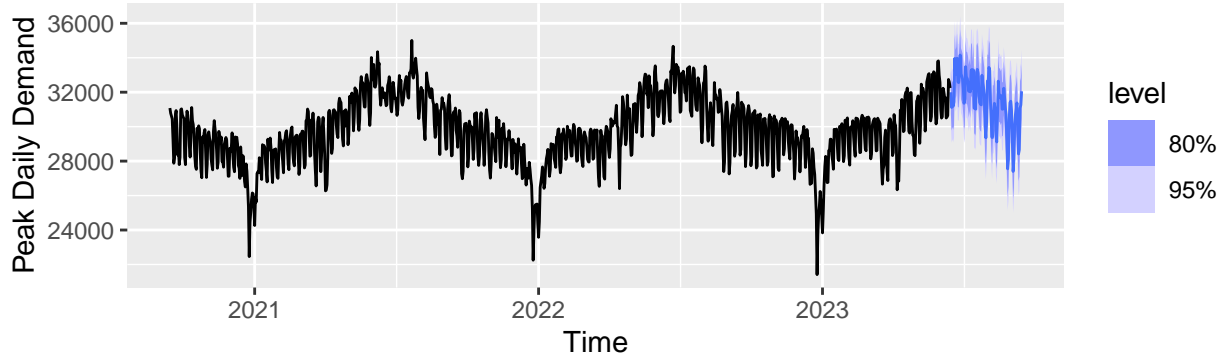


Figure 13: $SARIMA(3,0,3)(0,1,0)[365]$ residual plot.

Table 8: Ljung-Box Test Results

LB Statistic	P-value
454.603	0

The significant autocorrelation at lags 2,3,4,7 in the ACF and PACF plots of the residuals suggest that the model is an inadequate fit to the data and that there is information in the series that the model has failed to capture. It is clear that by neglecting the weekly seasonality by choosing $m = 365$ the model has failed to capture the weekly seasonality which can be seen by the spikes in autocorrelation in the ACF and PACF at lags that are a multiple of 7.

Figure 14: Forecasts for $SARIMA(3,0,3)(0,1,0)[365]$ model.

The forecasts for this represent the annual pattern in the series fairly well as they don't just fluctuate around a constant mean but seem to follow the complex patterns seen in the historical data, however, the strong autocorrelation present in the residuals at lag 7 make it clear that this model is not appropriate for the series. So whilst the forecasts appear more promising than the $SARIMA(4,1,4)(2,1,1)[7]$ model, further model fitting and exploration is required to find a model that can account for multiple/complex seasonalities.

4.4.3 - $SARIMA(7,1,7)(2,1,1)[7] + \text{Fourier}(365)$

A possible solution to handling a series with complex/multiple seasonality is fitting a dynamic harmonic regression model with an ARIMA error structure [4]. Put simply, Fourier terms are added to a SARIMA model to account for one of more of the seasonalities present in the data. Therefore, the $SARIMA(7,1,7)(2,1,2)[7] + \text{Fourier}(p = 365, k = 7)$ model is proposed (where p is the period of the Fourier term and k is the order of fourier terms). This specification uses the Fourier terms to try capture the annual seasonality in the data as the period of the annual seasonality is estimated to be 365 days, k is chosen to be 7 as this allows for a more flexible fit which will help avoid the model from underfitting. The non-seasonal AR and MA components were also increased to $p = 7$ and $q = 7$ to try and capture some of the autocorrelation at higher lags.

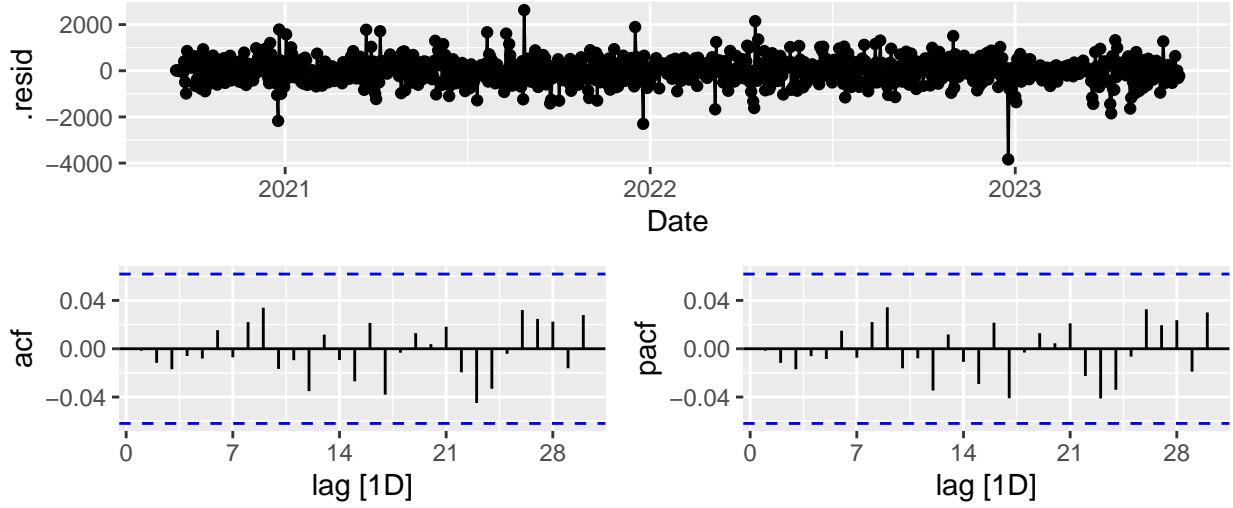


Figure 15: SARIMA(7,1,7)(2,1,1)[7]+Fourier(365) residual plot.

Table 9: Ljung-Box Test Results

LB Statistic	P-value
4.347	0.824

The residuals for the $SARIMA(7, 1, 7)(2, 1, 1)[7] + Fourier(p = 365, k = 7)$ model suggest that the model is a good fit to the data as there is no significant autocorrelation shown in the ACF or PACF plots. This is supported by the results of the Ljung-Box test for which the $p - value = 0.824$ which as per the hypotheses stated in Section 4.2 means that there is insignificant evidence to reject H_0 , therefore, it can be concluded that there is insignificant autocorrelation in the residuals and that the model is capturing the available information and underlying structures in the series.

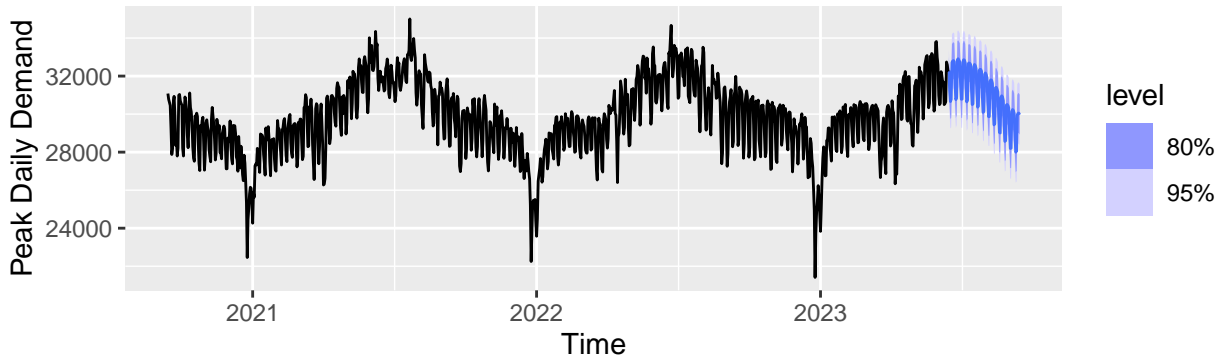


Figure 16: Forecasts for SARIMA(7,1,7)(2,1,1)[7]+Fourier(365) model.

The forecasts support the claim that the model is a good fit to the data as the model appears to capture both the weekly and annual seasonality and produce forecasts similar to what we'd expect given the historical data. The addition of the Fourier term seems to have helped capture the annual seasonal component, allowing the SARIMA component to focus on capturing the weekly seasonality.

4.5 - Benchmark Model - Seasonal Naive

Models built for forecasting should always be compared to a simple benchmark model as additional complexity introduced when using a model such as an ARIMA or SARIMA model must be justified [4]. The more complex model should significantly outperform and present improvements over the simpler forecasting model/method.

In the context of this problem, the seasonal naive forecasting method presents itself as the most appropriate benchmark method for forecasting the series due to the presence of strong seasonality in the series.

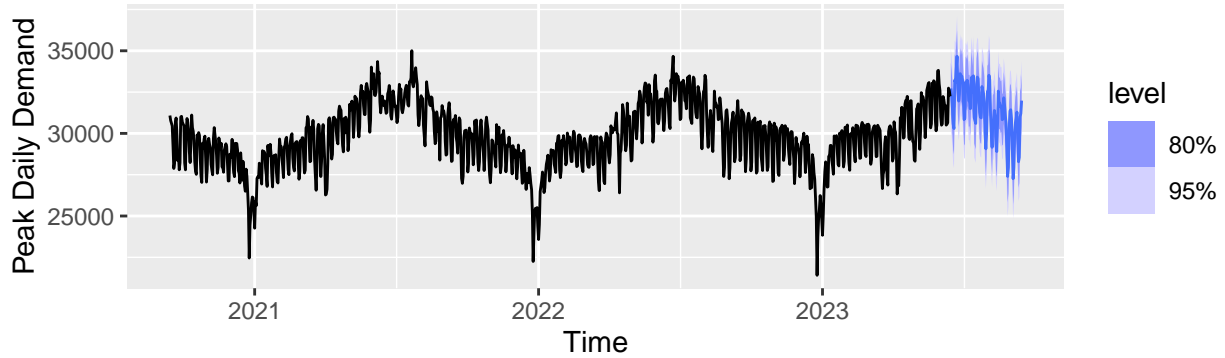


Figure 17: Forecasts for Seasonal Naive Forecasting Method.

The forecasts produced by the Seasonal Naive model appear to capture both the weekly and annual seasonalities present in the series. The accuracy of the forecasts will be compared to the other models in Section 4.7.

4.6 - Eskom's Forecasts

To determine the quality of the forecasts made by Eskom, a residual analysis is conducted using the residuals calculated by using the forecasts of the Peak Daily Demand provided by Eskom and the observed values of the Peak Daily Demand provided by Eskom.

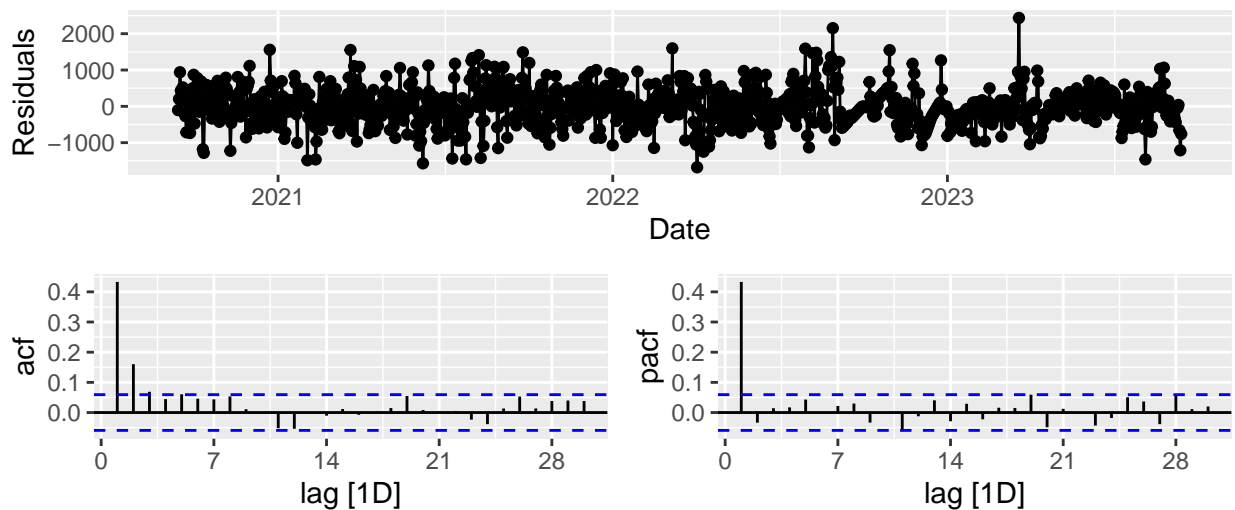


Figure 18: Residuals for Eskom's Forecasts.

The residuals for the forecasts made by Eskom appear to be stationary, however, they show that there is strong autocorrelation at lag 1 and 2 in the ACF and at lag 1 in the PACF. This suggests that the model/method that Eskom used to produce the forecasts failed to capture all of the information available in the data and that observations 1 time period (day in this case) apart are highly correlated. A Ljung-Box test cannot be performed as the model which Eskom used to produce these forecasts is not specified, however, the presence of autocorrelation in the residuals is clear.

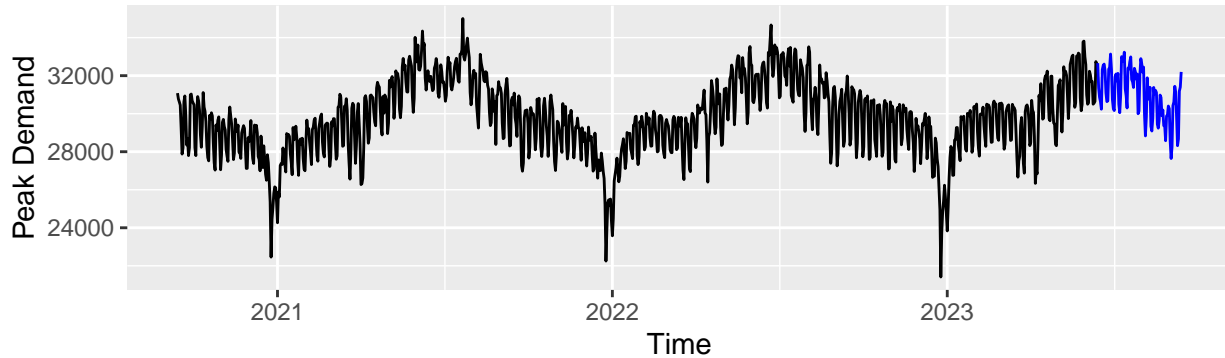


Figure 19: Eskom's Forecasts for Peak Daily Demand.

The forecasts made by Eskom are in line with what we'd expect from the series as they take the shape of the historical weekly and annual seasonality.

4.7 - Comparison of Mean Forecasting Models/Methods

Table 10: Forecasting accuracies of models/methods for Peak Electricity Demand

	RMSE	MAE
AR(8)	1490.474	1178.205
MA(7)	1428.732	1163.137
ARMA(5,6)	1430.187	1123.635
ARIMA(5,1,4)	1393.841	1114.426
SARIMA(4,1,4)(2,1,1)[7]	1783.741	1376.094
SARIMA(3,0,3)(0,1,0)[365]	1061.571	865.834
SARIMA(7,1,7)(2,1,1)[7]+Fourier(365)	1111.120	855.044
Seasonal Naive	1019.768	857.335
Eskom's Forecasts	435.362	332.929

The forecasted values for the peak demand of electricity were then used to calculate the accuracy of the forecasts by comparing them to the observed demand values in the test set. The table above shows the RMSE and MAE for the forecasts produced using AR, MA, ARMA, ARIMA and SARIMA models as well as the Seasonal Naive forecasting method and the forecasts provided by Eskom.

The findings in Table 5 and Section 4.2 allow for the non-seasonal models to be ruled out of contention, due to poor model fit, autocorrelated residuals and forecast quality/accuracy.

The $SARIMA(3,0,3)(0,1,0)[365]$ model had the best forecast accuracy on the test set, however, there was significant autocorrelation present in the residuals as found in Section 4.4.2 make it unreliable as there is clearly information that was not captured by the model. The performance on a single test set does not determine that the model generalizes well to all cases.

The most promising model was the $SARIMA(7, 1, 7)(2, 1, 1)[7] + Fourier(365)$ model. The residual analysis in Section 4.4.3 suggested that the model was a good fit and the forecasts aligned well with what is expected based on the historical data from the series. The forecast accuracy was also relatively good and could be improved through further tweaking of the model parameters.

The Seasonal Naive forecasting method performed very well on the test set, however, the simplicity of this method could result in poor generalization to other test sets and more rigorous testing such as Cross-Validation would be necessary to determine the general performance of this method.

Eskom's forecasts showed excellent accuracy, however, the presence of significant autocorrelation at lag 1 as discovered in the residual analysis in Section 4.6, suggests that whatever model/method was used to produce these forecasts failed to capture all of the available information. This might prove to be a problem in the long term for providing accurate and robust forecasts.

To produce robust forecasts, the $SARIMA(7, 1, 7)(2, 1, 1)[7] + Fourier(365)$ model should be fine-tuned to improve its forecast accuracy. If it turns out that the seasonal naive method consistently outperforms the optimized $SARIMA(7, 1, 7)(2, 1, 1)[7] + Fourier(365)$ model, even after rigorous testing, then the seasonal naive method could be preferred due to its simplicity, interpretability and its relatively low computational complexity for creating forecasts.

4.8 - ARCH and GARCH models

For the ARCH and GARCH models, the focus is on creating models used in forecasting future volatility based on recent patterns of volatility in the data. Thus the choice was made to only use the final year of observations in order to mitigate noise from previous years and ensure that the volatility patterns are more relevant to the current data.

The returns of the data are calculated as such:

$$\log(Return) = \log\left(\frac{x_t}{x_{t-1}}\right)$$

This transformation makes the series more stationary and scales the series which allows us to better investigate the volatility clustering and the evolution of volatility in the series. This transformation is also what enables the use of the ARCH/GARCH models.

The returns are then also squared, this further exposes clusters of volatility and ensures that the negative and positive returns contribute equally to the variance that is being modeled. Squaring the returns also amplifies the larger shocks in the returns which allow the GARCH model to identify the periods of high volatility.

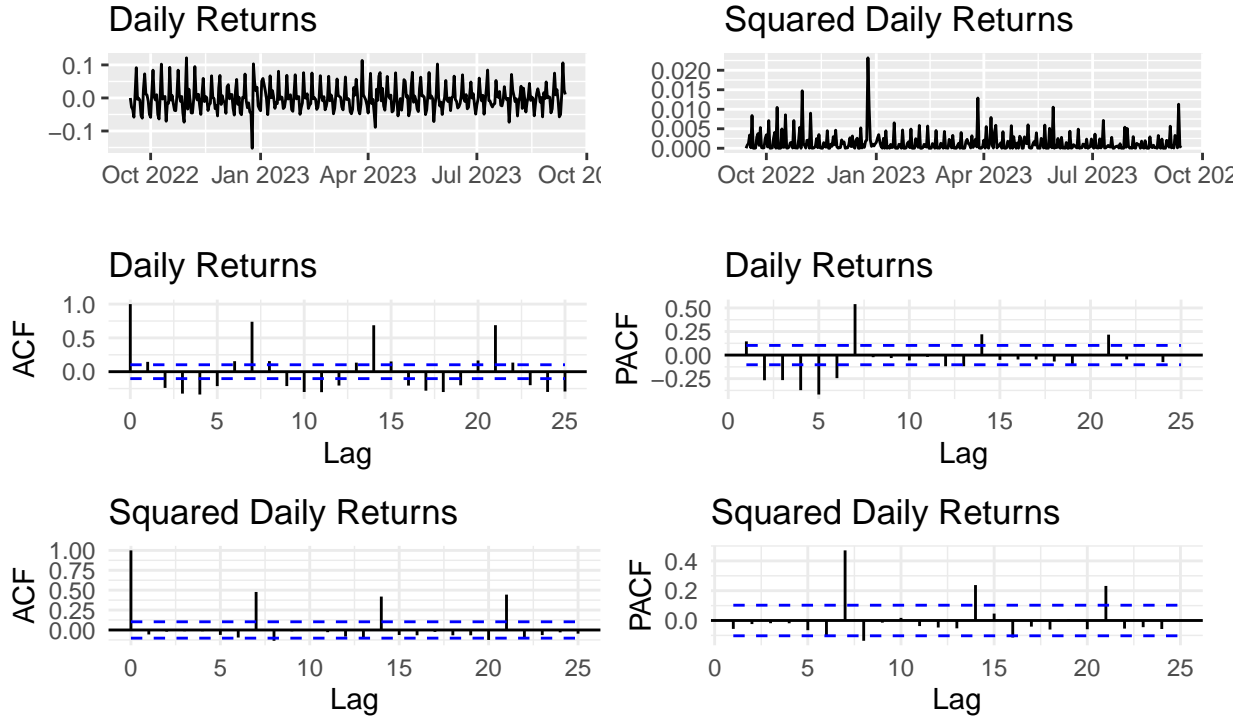


Figure 20: Time and ACF plots of returns/squared returns. (Electricity Demand)

The figure above shows that the squared daily returns appear to have significant autocorrelation, as shown by the spikes in the ACF/PACF plots at lags which are a multiple of 7, this suggests a $GARCH(p, q)$ model, however, $ARCH(p)$ models will also be fitted. To confirm the presence of ARCH (AutoRegressive Conditional Heteroskedasticity) effects, a formal test can be conducted.

ARCH-LM test:

H_0 : No presence of ARCH effects in the series.

H_1 : Presence of ARCH effects in the series.

Table 11: Results of ARCH-LM test.

Test.Statistic	P.Value
91.12	0

The $p - value < 0.001$, therefore there is strong evidence that there are ARCH effects present in the returns series and H_0 is rejected.

4.8.1 - ARCH

To determine the best $ARCH(p)$ model 3 models were fitted, using the PACF plot of the squared returns to determine the order of p [5]. The PACF of the squared returns in Figure 20 show that there is only significant autocorrelation at lags which are a multiple of 7. A baseline $ARCH(1)$ was fitted alongside two more complex models, the $ARCH(7)$ and $ARCH(14)$ models.

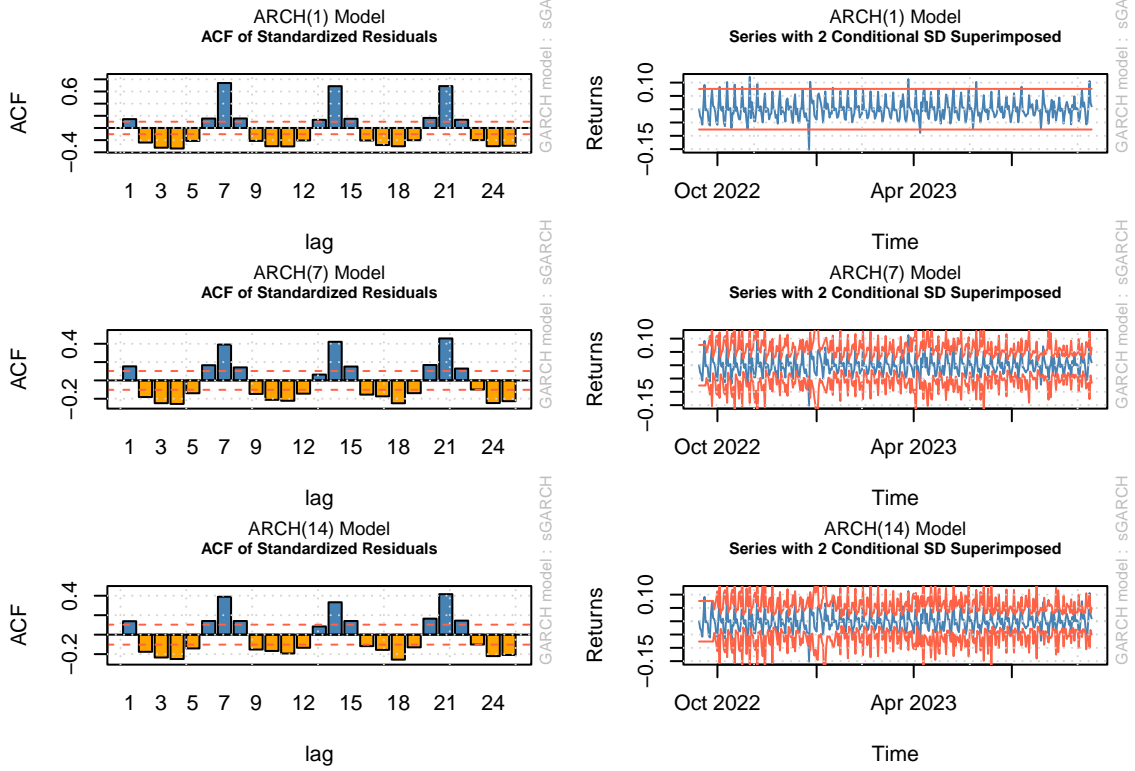


Figure 21: Standardized Residuals and Superimposed conditional SD for ARCH models.

The residuals and model fits in the figure above show that the $ARCH(1)$ severely under fits the series. This results in significant autocorrelation in the standardized residuals at lag 1 as well as poor model fit as shown by the bands around the log return series which fail to capture the volatility in the series.

The $ARCH(7)$ and $ARCH(14)$ models provide an improvement on the $ARCH(1)$ model, as shown by the plot of the conditional standard deviation which is superimposed on the returns series. The model appears to capture and mimic the patterns of the volatility in the series. With all this said, the $ARCH(7)$ and $ARCH(14)$ models still exhibit significant autocorrelation in the standardized residuals at low lags, which suggest that the models are not adequately modelling the volatility in the series.

4.8.2 - GARCH

The $GARCH$ models were fitted with an $ARMA$ mean component giving us $GARCH(p, q) - ARMA(p, q)$ which were fitted to the returns data. The orders for the $ARMA(p, q)$ component were selected using the ACF and PACF of the returns [5] as shown in Figure 20 and the orders for the $GARCH(p, q)$ were selected using the ACF/PACF plots of the squared returns.

The significant autocorrelation at lag 7 and multiples of 7 is due to the strong weekly seasonality in the series, to try and capture this autocorrelation $p = 7$ and $q = 7$ were chosen for the $ARMA(p, q)$ component for all the models. (Note: Lower orders of p and q were also tested but always resulted in a model which was underfitting.)

Much like in Section 4.8.1 a baseline $GARCH(1, 1) - ARMA(7, 7)$ was fitted alongside two other models which better represent the autocorrelation present in the ACF/PACF of the squared returns. The two more complex models were: $GARCH(7, 7) - ARMA(7, 7)$ and the $GARCH(14, 14) - ARMA(7, 7)$.

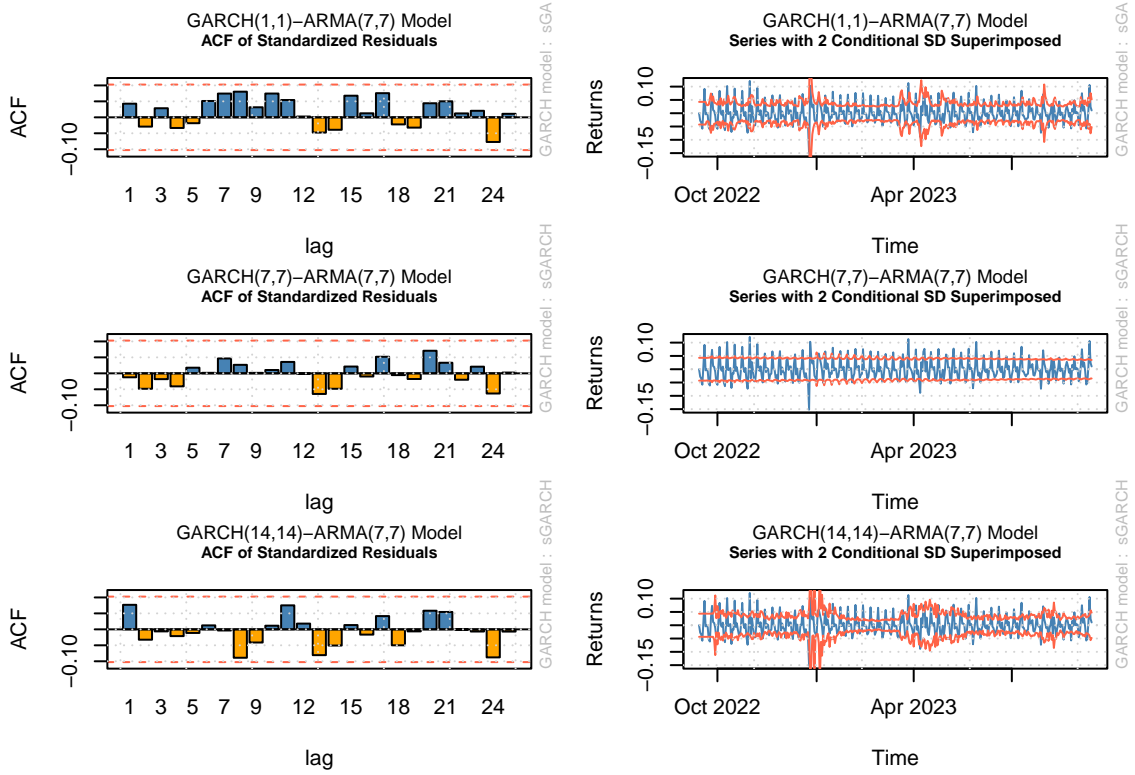


Figure 22: Standardized Residuals and Superimposed conditional SD for GARCH models.

The figure above shows that the inclusion of the past variance component helped with modelling the autocorrelation present in the returns as there is no significant autocorrelation in the standardized residuals of any of the three models. The model fit for all three models is decent but show that the models are underfitting the volatility in the series. The more complex $GARCH(7,7)$ and $GARCH(14,14)$ models are not worth the extra complexity as the $GARCH(1,1)$ produces a model fit similar to the $GARCH(14,14)$ model whilst still maintaining the quality of the residuals.

4.9 - Comparison of Volatility Forecasting Models

The ARCH models fit in Section 4.8.1 exhibited either significant autocorrelation in the residuals, poor model fit (by failing to accurately capture the volatility) or both. This is likely due to the lack of a mean component to account for the seasonality in the series. Therefore, none of the ARCH models fitted are appropriate for the capturing the volatility in the series. The GARCH models performed better, specifically in producing residuals which indicate good model fit. The most appropriate of the fitted models based on a balance of model fit and model complexity would be the $GARCH(1,1) - ARMA(7,7)$ model.

I would recommend either seasonally adjusting the data and then implementing the $GARCH(1,1) - ARMA(7,7)$ model or alternatively use a SARIMA model which explicitly incorporates the seasonality present and thus, is much more likely to provide an appropriate fit to the data.

Section 5 - Peak Daily UCLF+OCLF (Energy unavailability) Model Fitting

The analysis in this section follows the same workflow as established in Section 4 which includes similar time series forecasting models and model evaluation techniques. The workflow will be implemented with a focus on forecasting the peak daily UCLF+OCLF due to unplanned generation outages and therefore the necessary adjustments will be made to the methodology based on the differing characteristics of the data as discussed in Section 3.7.

The same training/test split and forecast horizon will be used as specified in Section 4 (test set is the last 3 months of observations). This will provide an adequate amount of data to train and test the model on and the 3-month forecast horizon is suitable as it balances forecasting accuracy with the model/s ability to capture recent trends whilst mitigating the effects of noise. This results in robust forecasts which can be used for informing critical decisions regarding resource allocation and maintenance scheduling which are vital in minimizing unplanned outages causing energy unavailability.

Figure 3 in Section 3.7 shows that there is a strong positive linear trend in the series as well as a lack of any sort of seasonality. Based on time series theory as well as the methodology and findings of Section 2, the use of a seasonal model will not be appropriate due to the lack of seasonality in the series, which is reflective of the use of non-seasonal models for the series in Section 4. Therefore an AR/MA/ARMA or ARIMA model will be fit to the series based on the findings from differencing/detrending the data.

Figure 3 shows that the series could possibly be trend stationary as the series appears to fluctuate around a deterministic trend, which can then be made stationary by simply removing the trend component of the series [6]. To test for trend stationarity a KPSS test as well as an ADF test can be performed to assist in confirming any uncertainties.

KPSS test:

H_0 : The series is trend stationary.

H_1 : The series is non-stationary and may have a unit-root.

Table 12: Results of the KPSS and ADF tests.

Test.Type	Truncation.Lag	Lag.Order	Test.Statistic	P.value
KPSS	6	NA	0.171	0.029
KPSS	19	NA	0.098	0.100
ADF	NA	8	-6.138	0.010

5.1 - Detrending the series

The KPSS test with the Truncation Lag parameter set to ‘short’ suggests that the series is not trend stationary ($p - value = 0.029$, reject H_0), this result conflicts with the other test results. The KPSS test with the Truncation Lag parameter set to ‘long’ suggests that the series is trend stationary ($p - value = 0.1$, fail to reject H_0), which is supported by the result of the ADF test which indicates that the series has no unit-root and is either stationary/trend stationary.

Based on the results of the tests as well as the suspicion of trend-stationarity, the series was detrended in an attempt to achieve stationarity.

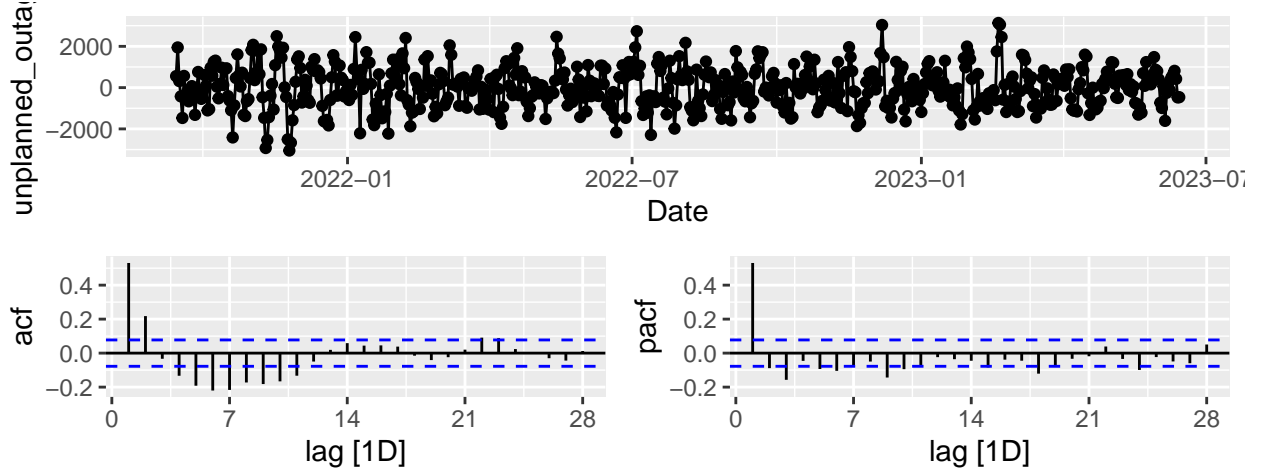


Figure 23: Differenced peak daily UCLF+OCLF data and ACF/PACF plots (Non-seasonal).

The figure above displays that detrending the data was adequate in achieving stationarity of the series as the series fluctuates around a constant mean and the variance seems to be fairly constant and time-independent.

The ACF/PACF plots indicate that the model should include both, autoregressive and moving-average components. Findings and discussions in Section 4 highlighted that fitting the correct type of model is crucial for model fit and validity, therefore, the use of simple AR or MA models can be disregarded as both components are required. ARIMA models can then also be disregarded as the model does not require any differencing as it is already stationary. An ARMA framework is appropriate for modelling and forecasting the total peak daily UCLF+OCLF due to unplanned generation outages.

5.2 - ARMA model fitting

Figure 23 in Section 5.1 exhibits a significant spike autocorrelation at lag 1 for both ACF and PACF, therefore, $p = 1$ and $q = 1$ are chosen as the initial candidate model. Both the ACF and PACF exhibit significant autocorrelation at lag 2 and therefore $p = 2$ and $q = 2$ were chosen for the second candidate model. Finally, in order to try and capture autocorrelation up until higher lags, $p = 3$ and $q = 4$ were chosen. The table below shows some results from fitting the models.

Table 13: AICc, RMSE and Ljung-Box test results for different ARMA(p,q) models.

Model	AICc	RMSE	Ljung-Box Stat	P-value
ARMA(1,1)	10400.91	1526.134	36.060	0.000
ARMA(2,2)	10312.97	1498.979	7.479	0.279
ARMA(3,4)	10313.14	1497.886	2.901	0.407

Based on the results in the table above, the ARMA(2,2) was the chosen model to forecast the peak daily UCLF+OCLF. The choice was made based on:

1. ARMA(2,2) has the second best RMSE
2. The Ljung-Box test results show that there is no significant autocorrelation in residuals up until a fixed number of lags ($p - value > 0.05$, fail to reject H_0).
3. The AICc is the lowest of the three models which suggests that the ARMA(3,4) model is not worth the extra complexity.

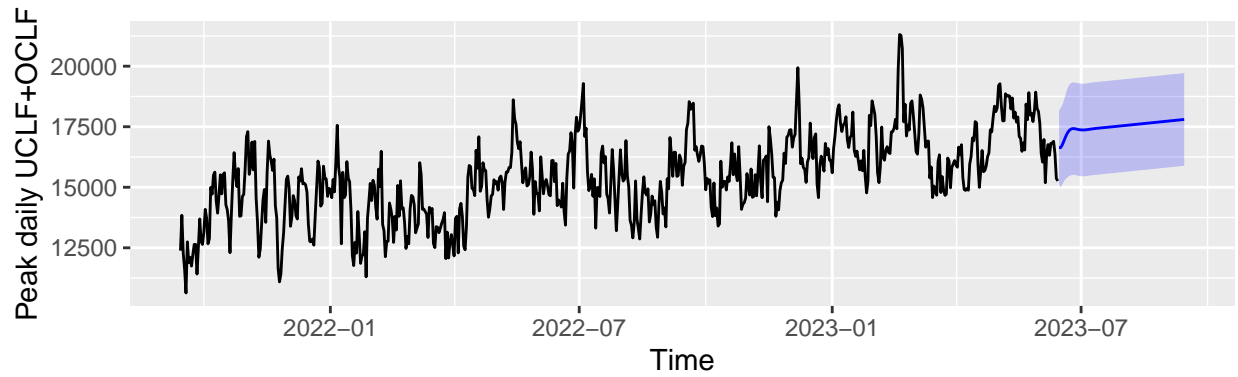


Figure 24: Forecasts for ARMA(2,2) model.

The figure above shows that the forecasts made using the ARMA(2,2) model successfully capture the trend in the data but fail to pick up on the smaller fluctuations in the series.

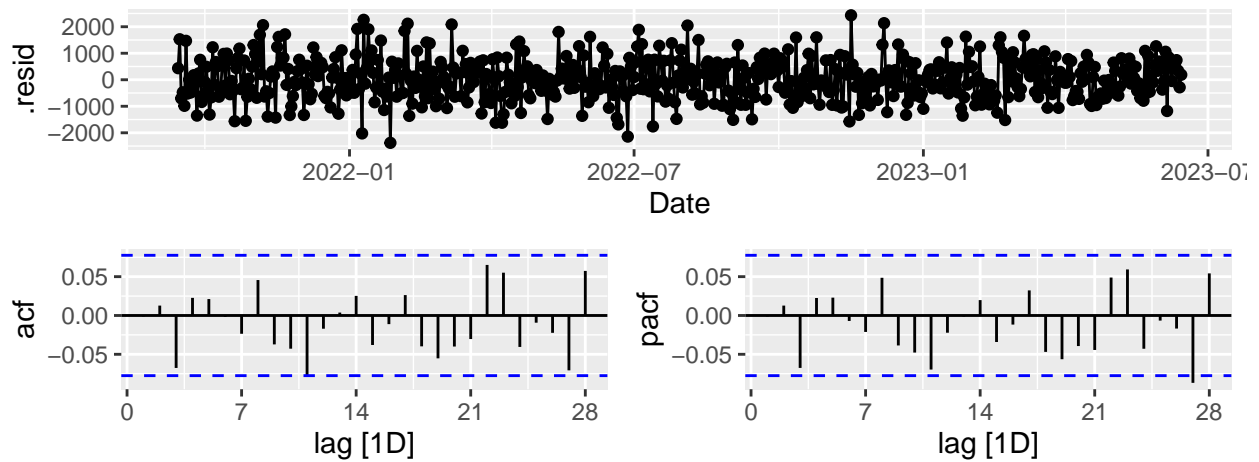


Figure 25: ARMA(2,2) residual plot.

The residuals resemble a white noise process and the lack of significant autocorrelation in the ACF/PACF plots confirm that the residuals are independent over time and that the model is capturing all available information.

5.3 - Benchmark Model/Method

As discussed in Section 4.5, new models should always be compared to a benchmark forecasting model/method [4]. The Drift forecasting method seems appropriate for this series due to the presence of a linear trend. The forecasts plotted below are akin to the forecasts produced by the $ARMA(2,2)$ model.

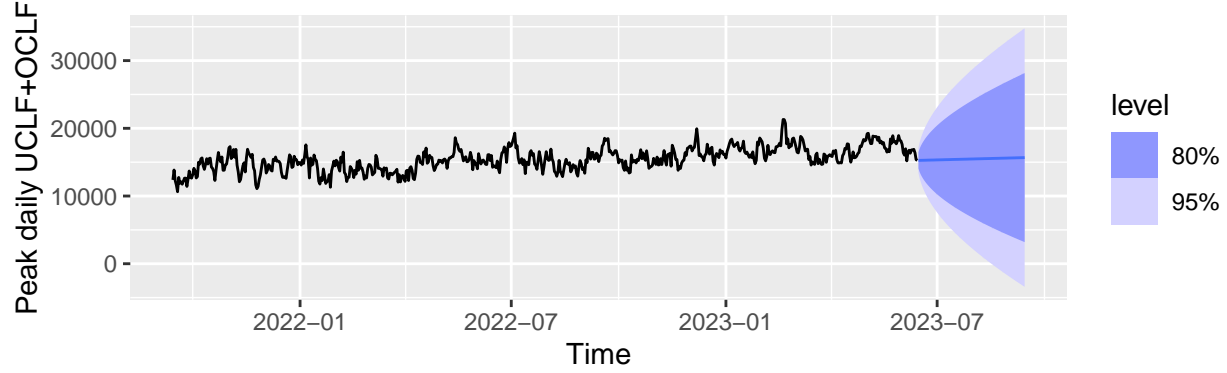


Figure 26: Forecasts for Drift Forecasting Method

5.4 - Comparison of Mean Forecasting Models/Methods

Table 14: Forecasting accuracies of models/methods for Peak daily UCLF+OCLF

	RMSE	MAE
ARMA(2,2)	1498.979	1310.714
Drift	1255.003	988.372

The ARMA(2,2) model is outperformed by the benchmark Drift forecasting method on the test set. As per Section 4.7 this may be the result of the limitations of only using a single test. Thus I would once again recommend performing more thorough testing of both the ARMA(2,2) model and the Drift forecasting method as additional data becomes available. This will enable more robust and reliable conclusions to be made about which forecasting tool to use.

5.5 - ARCH and GARCH

The returns for peak daily UCLF+OCLF due to unplanned generation outages were calculated (as in Section 4.8) and the relevant plots are shown below. For the same reasons as discussed in Section 4.8, only the last year of observations will be used for the ARCH/GARCH modelling.

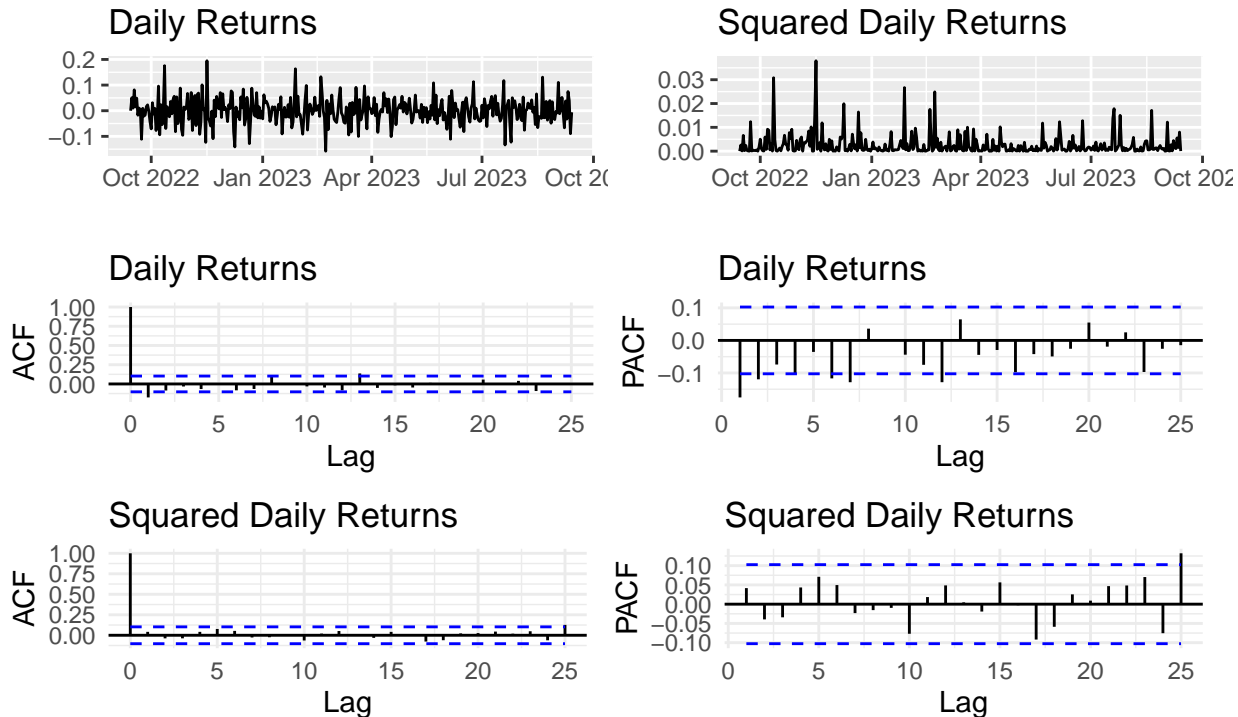


Figure 27: Time and ACF plots of returns/squared returns. Peak daily UCLF+OCLF

The ACF/PACF in the figure above display a lack of autocorrelation in the squared daily returns, this implies that the volatility of the daily returns does not exhibit the kind of clustering and time variation that ARCH/GARCH models are designed to capture.

To test for ARCH effects such as the previously mentioned volatility clustering and time-varying variance, an ARCH-LM test can be conducted (As specified in Section 4.8).

Table 15: Results of ARCH-LM test.

Test.Statistic	P.Value
5.177	0.638

The result of the ARCH-LM test indicates that there is insignificant evidence to suggest that there are ARCH effects present in the series ($p - val = 0.638$, fail to reject H_0), this supports the notion that the implementation of ARCH/GARCH models to capture the volatility of the series will produce inaccurate results and poor quality residuals.

To emphasize this, an $ARCH(1)$ and $GARCH(1,1) - ARMA(2,1)$ are fitted to the returns. The ACF/PACF plots of the squared returns show significant autocorrelation at lag 25 which is most likely spurious, however, for the sake of exploration a $GARCH(25,25) - ARMA(2,1)$ was fitted.

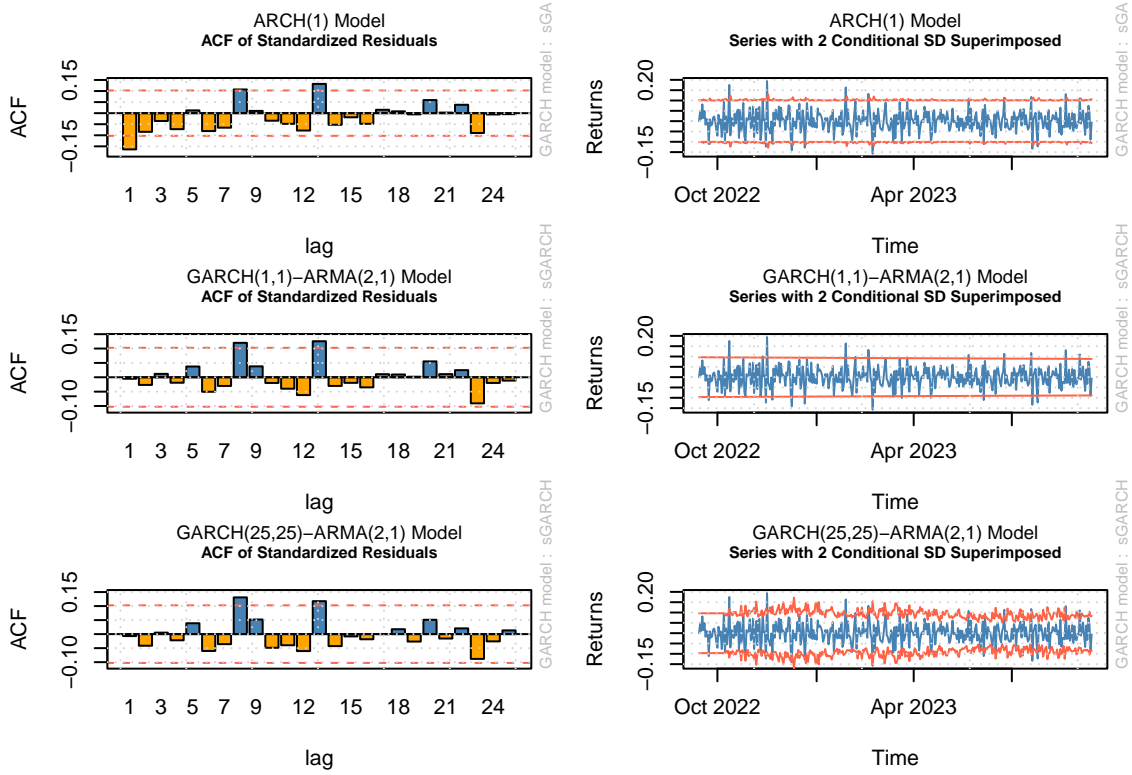


Figure 28: Standardized Residuals and Superimposed conditional SD for ARCH/GARCH models.

Both the $ARCH(1)$ and $GARCH(1,1) - ARMA(1,1)$ exhibit inadequate fit as the superimposed conditional standard deviation is fitted as straight bands, which confirms that the models are unable to capture the volatility of the series. The standardized residuals for the $ARCH(1)$ show significant autocorrelation at lag 1 whilst for the $GARCH(1,1) - ARMA(2,1)$ and $GARCH(25,25) - ARMA(2,1)$ the first significant spike in autocorrelation is at lag 8. Although the $GARCH(25,25) - ARMA(2,1)$ looks as if it is better capturing the volatility it is essentially performing the same as the other two models, as the conditional standard deviation bands are just fluctuating around some mean and not actually capturing the spikes in the returns and therefore not capturing the volatility.

5.6 - Comparison of Volatility Forecasting Models

According to the tests and investigation conducted in Section 5.5 the returns of the peak daily UCLF+OCLF exhibit no significant ARCH effects (which includes volatility clustering and time-varying variance) and therefore the use of ARCH/GARCH models to model the volatility in the series is not appropriate as presented in the discussion of the models in Section 5.5. I would recommend using alternative methods or models such as an Exponential Weighted Moving Average (EWMA) to produce robust forecasts of the volatility of the returns of the peak daily UCLF+OCLF.

Section 6 - Discussion and Limitations

As hypothesized, clear weekly and annual seasonal patterns were present in the daily peak demand series. This led to the SARIMA models outperforming the non-seasonal models as they were able to capture the seasonality in the series. The strong seasonality present in the series proved to play a large role in modelling the volatility of the series and thus the basic ARCH/GARCH-ARMA models that were fitted were found to be inadequate, presenting the need for a model that directly accounts for the seasonality in the returns.

The peak daily UCLF+OCLF was found to have a positive linear trend and no clear weekly or annual seasonality (as hypothesized) which implies that it is not greatly affected by the changing seasons or weekends/holidays. Therefore a non-seasonal ARMA model was found to be the best model based on forecasting accuracy and model complexity.

The volatility of the peak daily UCLF+OCLF was hypothesized to be higher in the peaks of winter and summer than in transitional seasons such as spring and autumn. The analysis in Section 5.5 suggests that there is insufficient evidence to support this claim. As a matter of fact the ARCH/GARCH models did not adequately capture the volatility in both the peak daily demand and the peak daily UCLF+OCLF.

The forecasts provided by Eskom were very accurate on the test set and align almost exactly how you'd expect given the historical data. Although as discussed in Section 4.6 the residuals of Eskom's forecasts suggest that the model/method they used was still not capturing all available information. Unfortunately the methodology of how Eskom produced these forecasts was not provided and thus can't be improved upon.

The forecasts of the peak daily demand as well as the peak daily UCLF+OCLF are limited by the use of time series models that do not include other exogenous variables. This could be addressed through the use of the relevant SARIMAX or ARIMAX models which incorporate other predictors into the model which can provide more insight into what impacts the variables that are being forecasted.

The ARMA model used to forecast the peak daily UCLF+OCLF will struggle to adapt if the trend changes over time which is quite likely due to the assignment of funds and resources to maintaining Eskom power plants being an ongoing discussion.

Section 7 - Conclusions

In this analysis, AR/MA/ARMA/ARIMA and SARIMA models, were developed in order to produce forecasts for each the values of the peak daily electricity demand as well as the peak daily UCLF+OCLF. The models were evaluated based on a residual analysis and forecasting accuracy (RMSE). ARCH and GARCH models were also used to capture the volatility of both the peak daily electricity demand and the peak daily UCLF+OCLF.

It was shown that choice of the $SARIMA(7, 1, 7)(2, 1, 1)[7] + Fourier(365)$ model to produce robust forecasts of the peak daily demand is satisfactory based on the results obtained in the analysis, however, as mentioned in Section 4.7, if the Seasonal Naive method continues to outperform the SARIMA model, then the Seasonal Naive should be used due to its simplicity, interpretability and low computational complexity. The fine-tuning and further improvement of the SARIMA model by extending it to include exogenous variables should be taken into strong consideration. For forecasting the volatility, the results determined that a model such as a GARCH-SARIMA model should be used due to the strong weekly seasonality in the data.

Robust forecasts of the peak daily UCLF+OCLF can be produced using either the $ARMA(2, 2)$ model or the Drift forecasting method as discussed in the analysis. As more data becomes available the best choice between the two will present itself. The analysis showed that to produce robust forecasts for the peak daily UCLF+OCLF, other methods or models such as EWMA should be implemented as ARCH/GARCH are not suited to modelling the time-independent volatility in this series.

By anticipating electricity demand and energy unavailability effective strategies can be developed to assist in mitigating load shedding and enhancing grid stability whilst optimizing resource allocation and plant maintenance.

References

- [1] Electricity Patterns Explained, Eskom, <https://www.eskom.co.za/eas/power-alert-2/#:~:text=In%20South%20Africa%20>
- [2] Demand Side, Eskom, <https://www.eskom.co.za/dataportal/demand-side/>
- [3] Ljung-Box Test, MathWorks, <https://www.mathworks.com/help/econ/lbqtest.html>
- [4] Hyndman, R.J., & Athanasopoulos, G., Forecasting: Principles and Practice (3rd ed.). OTexts, <https://otexts.com/fpp3/>
- [5] STA 510 Applied Time Series Analysis - ARCH/GARCH models, Penn State Eberly College of Science, <https://online.stat.psu.edu/stat510/lesson/11/11.1>
- [6] Trend-Stationary vs. Difference-Stationary Processes, MathWorks, <https://www.mathworks.com/help/econ/trend-stationary-vs-difference-stationary.html>

Appendix

1.1

Table 16: The variables of the ESK6816 dataset and their variable types.

Variable Name	Variable Type
Date	Date
RSA.Contract.Demand	numeric
RSA.Contract.Forecast	numeric
Total.UCLF.OCLF	numeric

1.2

Table 17: AICc for different AR(p) models.

Model	AICc
AR(5)	16308.39
AR(6)	16233.53
AR(7)	15998.21
AR(8)	16000.12

Table 18: AICc for different MA(q) models.

Model	AICc
MA(4)	16560.14
MA(5)	16515.28
MA(6)	16512.58
MA(7)	16338.71

Table 19: AICc for different ARMA(p,q) models.

Model	AICc
ARMA(4,4)	16066.38
ARMA(5,4)	16069.72
ARMA(5,5)	15948.33
ARMA(5,6)	15836.45
ARMA(5,7)	15837.25

Table 20: AICc for different AR(p) models.

Model	AICc
ARIMA(3,1,3)	16256.71
ARIMA(3,1,4)	16066.39
ARIMA(3,1,5)	16066.64
ARIMA(4,1,3)	16123.80
ARIMA(4,1,4)	16066.38
ARIMA(5,1,3)	16079.65
ARIMA(5,1,4)	16069.72
ARIMA(5,1,5)	15948.33

1.3

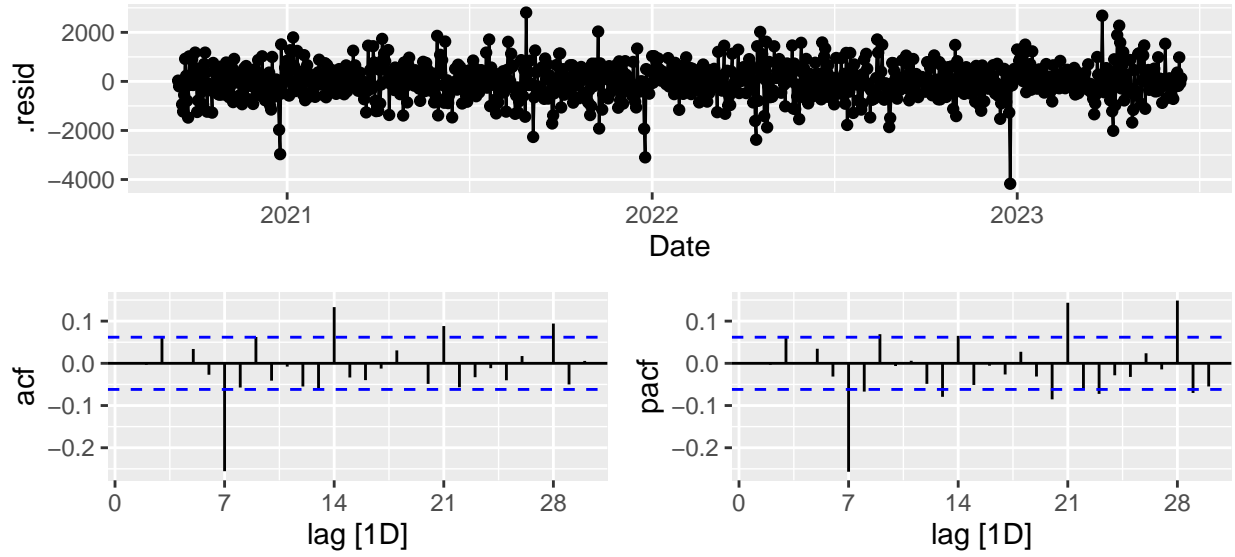


Figure 29: Residual plot for AR(8) model

1.4

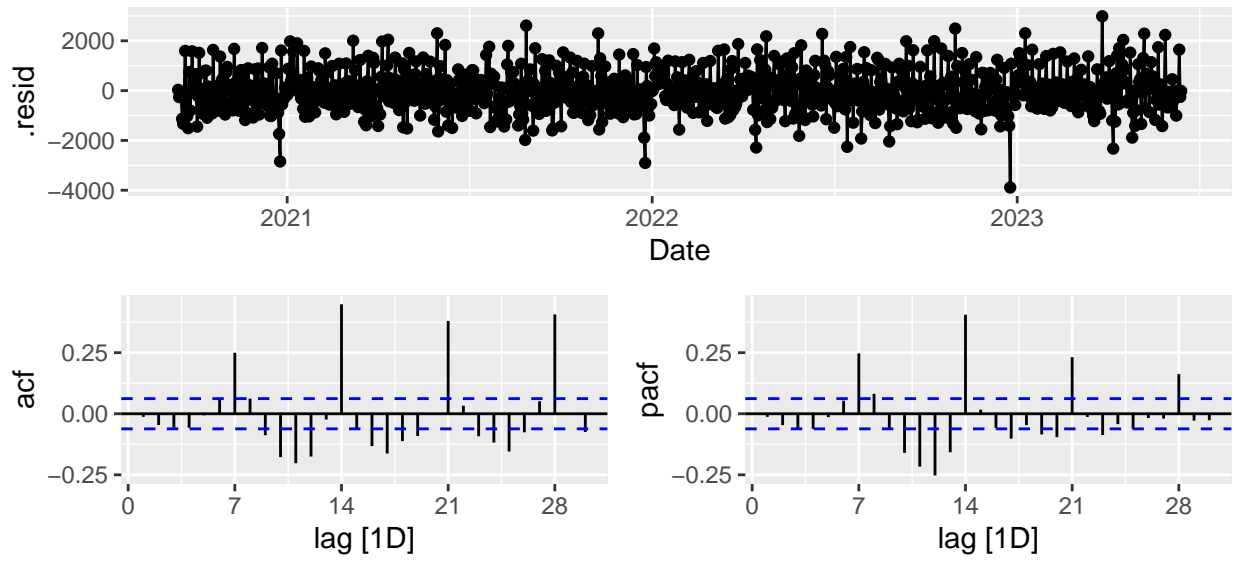


Figure 30: Residual plot for MA(7) model

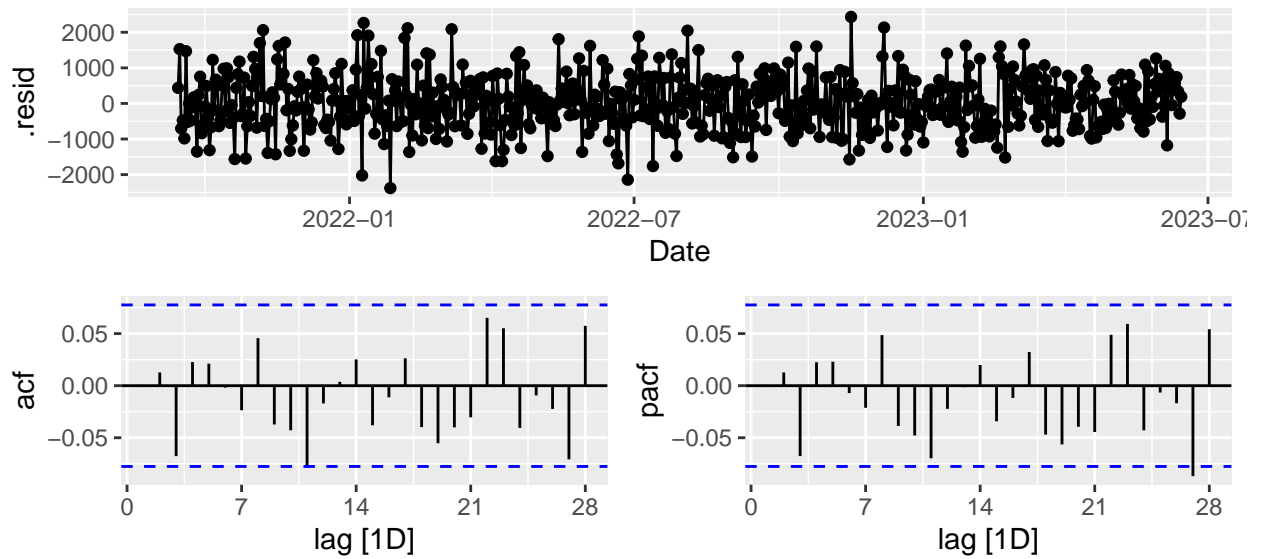


Figure 31: Residual plot for ARMA(5,6) model

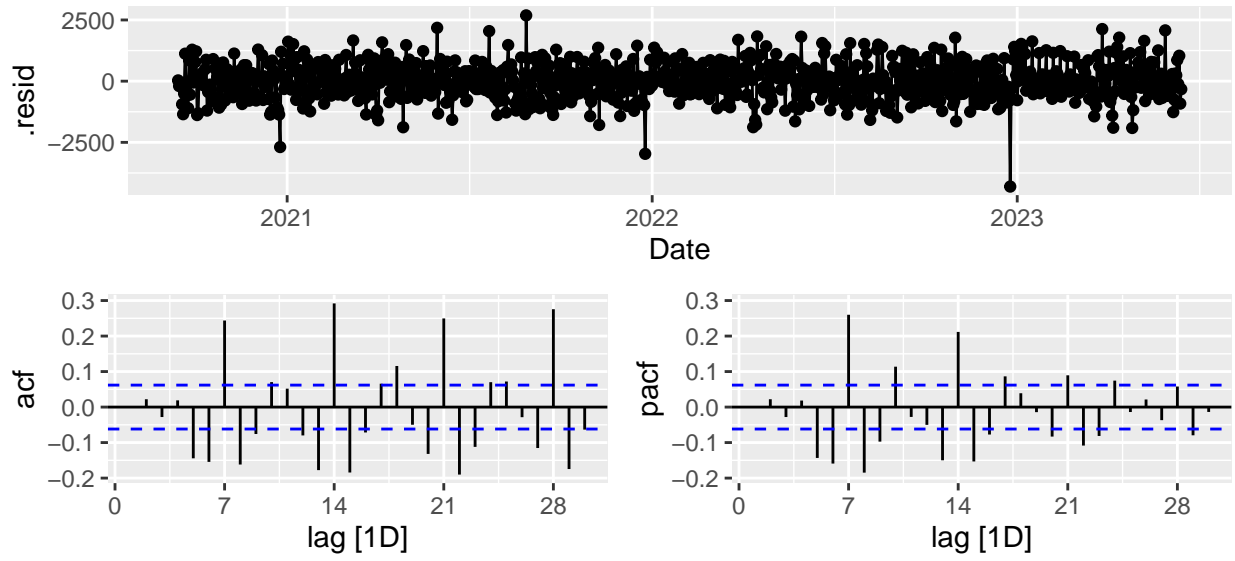


Figure 32: Residual plot for ARIMA(5,1,4) model

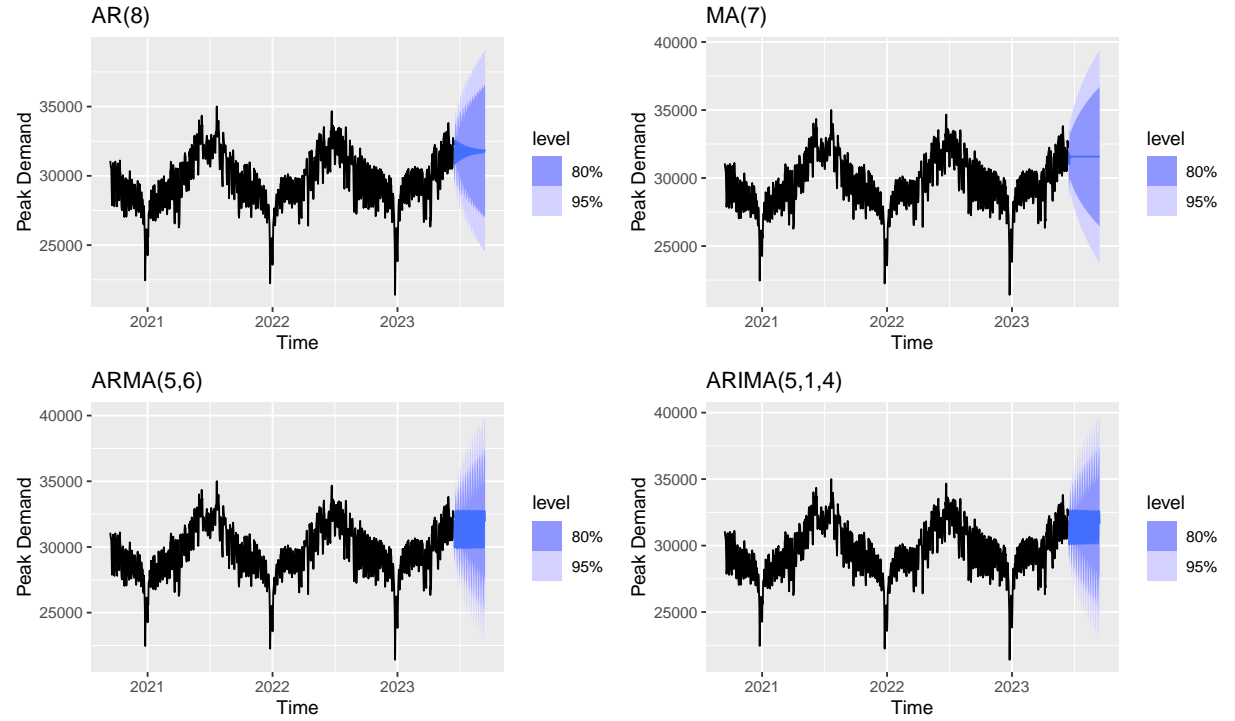


Figure 33: Forecasts for non-seasonal models