The Impact of COVID-19 on Flight Traffic in Denmark.

A Forecasting Analysis

Exam paper
Predictive Analytics
[CDSCO1003E]

MSc Business Administration and Data Science

Author:
Michelle Judith Sara von Huth (mivo22ab) (158397)

The Impact of COVID-19 on Flight Traffick in Denmark.

A Forecasting Analysis

Michelle Judith Sara von Huth
Copenhagen Business School MSc. Business Administration and Data Science

Abstract

This paper focuses on forecasting traffic in Denmark under a hypothetical non-pandemic scenario and compares the predictive performance of two forecasting models: ARIMA and ETS. The study utilizes empirical data from 2001 to 2022 on departing passengers from public airports in Denmark. Exploratory data analysis techniques were applied to understand the time series patterns. The data were transformed and checked for stationarity, considering the effects of seasonal and first differences. Structural break tests were conducted to identify changes in data patterns. The analysis revealed the presence of a structural break in 2008, leading to a division of the dataset. The selected period after the break, 2009-2019, were used for modelling and forecasting. The ETS model was identified as the most suitable model, particularly the guessed ETS(M, Ad, A) model. The findings offer insights into future growth and development trends in the aviation sector.

1 Introduction

The aviation industry is critical to the global tourism economy. Yet, it is particularly vulnerable to a wide range of global events and situations, such as economic shifts, natural disasters, health crises such as pandemics, and political disruptions. With globalization causing an increase in international travel and business, the aviation sector finds itself in need of reliable air traffic estimates. Such forecasts are critical in a variety of ways, including assisting with strategic investment planning, identifying prospective new airline routes, and evaluating airport capacity. (Nhamo et al., 2020)

The Covid-19 pandemic brought forth an unexpected challenge, disrupting the previously uninterrupted growth trajectory observed in the industry, which has prompted significant inquiries and concerns. Thus, the central focus of this paper revolves around the following research question: *"What is the most appropriate forecasting technique, specifically between ARIMA and ETS models, for accurately predicting air traffic in Denmark under a hypothetical non-pandemic scenario?"*

As said, this investigation will use both ARIMA and ETS models to predict air traffic, as well as a SNAIVE model as a benchmark model. To select the model with the best predictive performance, this paper will follow an empirical strategy to forecast air traffic in Denmark. In addition, it is anticipated that this research's findings will offer a data-driven method for predicting future growth and development trends in the wake of a crisis.

2 Data

The data for this study was gathered from the FLYV92 report available on the StatBank Denmark Transportation portal, provided by Statistics Denmark. This report delivers monthly figures of "*departing passengers from major, manned, public airports by type of transport and flight*". The dataset included a 22-year span from January 2001 to March 2023 with a monthly frequency. To predict how air traffic would behave in the absence of a pandemic, the dataset was truncated to January 2020, which provided 264 observations. To make use of the data easier, it is presented in increments of 1,000 passengers.

The time period of the dataset was carefully examined to ensure a thorough understanding of the patterns, fluctuations, and potential outliers that might have affected Danish aviation traffic during this time. To acquire a deeper understanding of the subtleties of the dataset, an exploratory data analysis (EDA) was conducted after this stage. The following section of the report will go into further detail about the EDA and the data transformation.

2.1 EDA & Data Transformation

The objective of the EDA is to gain a deeper understanding of the time series, comprising the identification of missing values, outliers and data types. The first step in this exploration was to convert the flight traffic data into a time series object, or tsibble, with specific consideration to the 'month' as the index. To create a non-pandemic model of flight traffic in Denmark, data only up until January 2020 were utilized. (Hyndman & Athanasopoulos, 2018)

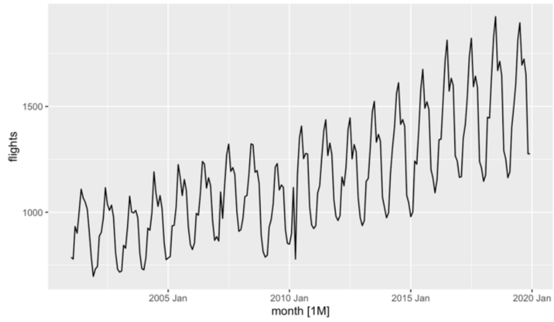


Figure 1: Time series of the monthly flight traffic

Visual inspection is a crucial part of the EDA. In Figure 1, the time series of monthly flying passengers in Denmark is plotted, which immediately reveals an upward trend and a clear seasonal component. Additionally, there are no apparent outliers or missing values in the data, alleviating the need for any preliminary data cleaning. To further highlight the seasonality in the flight data, two distinct visual representations were crafted. The subseries plot (Figure 2a) in conjunction with the seasonal plot (Figure 2b) underscores that the peak period of travel occurs from June through August, thereby suggesting a 12-month cyclical pattern. (Hyndman & Athanasopoulos, 2018)

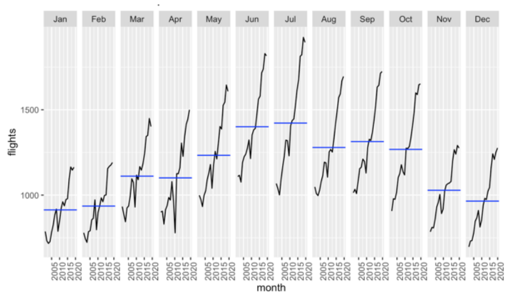


Figure 2a: Subseries plot

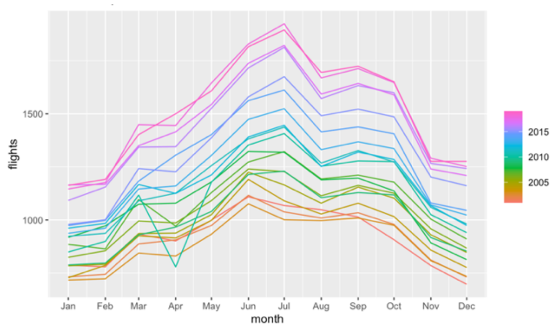


Figure 2b: Seasonal plot

Furthermore, the autocorrelation function (ACF) plot (Figure 2c) also detects seasonality, as the longest lags are observed at 12, 24, 36 and so on. Additionally, the examination of lags in the autocorrelation functionreveals the presence of a trend, as evidenced by a gradual decay from left to right. Given the results of the initial EDA, it is plausible that both a trend and seasonality components are present, influencing the structure of the series. (Chen, 2021)

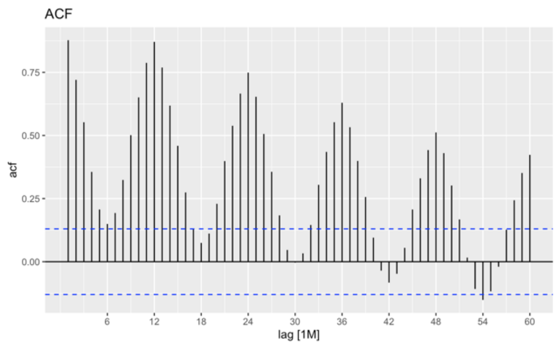


Figure 2c: ACF

A SEATS decomposition was performed as the final stage of the exploratory data analysis (EDA). As shown in Figure 3, this analysis reveals an upward trend and increasing seasonality over the examined time period. Although there are no immediate indications of structural breaks in the time series data, such as sudden drops, it is important to validate this observation and ensure the accuracy of our future forecasts. Therefore, a structural break tests will be conducted later in the analysis to mitigate the risk of drawing incorrect conclusions. Additionally, the irregular component captures fluctuations that cannot be attributed to either the trend or seasonality. By considering these analyses collectively, valuable insights can be gained to anticipate flight traffic in the absence of the Covid-19 pandemic.

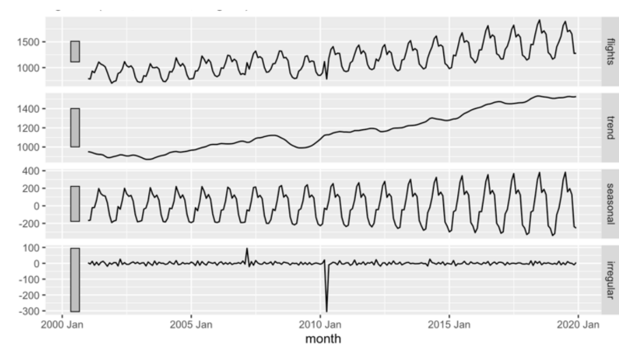


Figure 3: SEAT Decomposition

After performing exploratory data analysis, the Guerro lambda value was calculated to be -0.28, indicating a potential need for a Box-Cox transformation to stabilize the variance of the data. However, upon applying the transformation and observing the results in Figure 4, it was noted that the transformation did not significantly affect the variance of the data. To ensure confidence in this decision, models were tested on both the transformed and non-transformed data, and it was discovered that the models produced identical outputs regardless of whether the data was transformed or not. As a result, it was determined that there was no need for further transformation and the original data was used for subsequent analysis. Please refer to Appendix A for a SEATS decomposition and ACF of the Box-Cox transformed data. (Osborne, 2010)

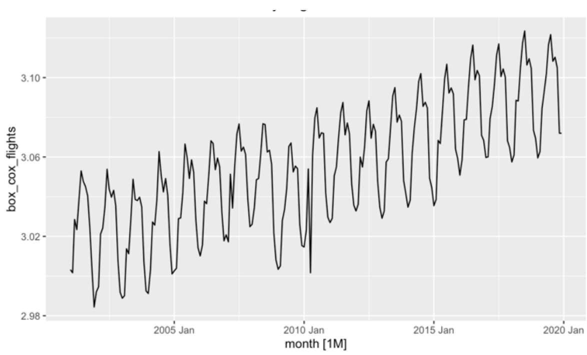


Figure 4: Box-Cox Transformation

2.2 Structural Breaks

After visually inspecting the flights data, an examination for possible structural breaks was conducted. As shown in Figure 5, the Quandt likelihood ratio (QLR) test was conducted, which provided visual evidence of a structural break. The breakpoint revealed that the structural break took place in year 2008, suggesting a difference in the pattern before and after this year.

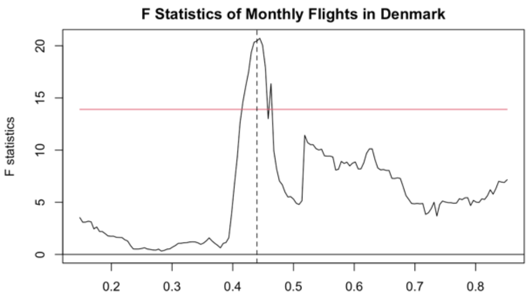


Figure 5: QLR Test

The QLR test plot showed that the curve went above the red line, indicating the presence of a structural break in the data. The associated p-value of 0.00267 indicated that the null hypothesis of no structural break could be rejected. This suggested that there is likely at least one structural break in the data, indicating a difference in the pattern before and after the break. To further investigate the presence of a structural break, the Chow test was performed. (Muthuramu & Maheswari, 2019)

The Chow test yielded a p-value of 0.54348, which surprisingly indicates no strong statistical evidence to reject the null hypothesis of no structural break. The p-value was greater than the commonly used significance level of 0.05, suggesting no significant difference in the relationship between the two segments of the data. Since the QLR test can handle different types of variations in real-world data, while the Chow test assumes the same level of variation for all observations, the QLR test is considered more reliable in this particular situation. (Muthuramu & Maheswari, 2019).

To confirm the presence of structural breaks, a cumulative sum (cusum) test was conducted. The cusum analysis identified potential break points at indices around 100. These break points indicated significant shifts or changes in the underlying data patterns. As shown in Figure 6, visual inspection of the cusum graph revealed a spike in the cusum crossing the threshold represented by the red line, further confirming the presence of structural breaks in the data.

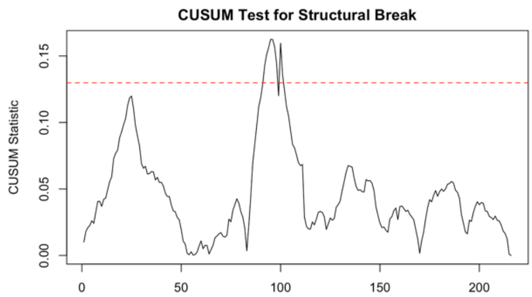


Figure 6: Cumulative sum Test

Based on these findings, the decision was made to split the data into two subsets, using only the data after the structural break for further analysis. Moving forward, the subsequent analysis and modelling will focus on the dataset that takes place after the identified structural break in 2008, specifically from 2009 to 2019, allowing for a more accurate representation of the underlying dynamics and patterns in the flights data. Following the identification of the structural break, Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots were generated using only the data after the structural break, as shown in Figure 7a and 7b below.

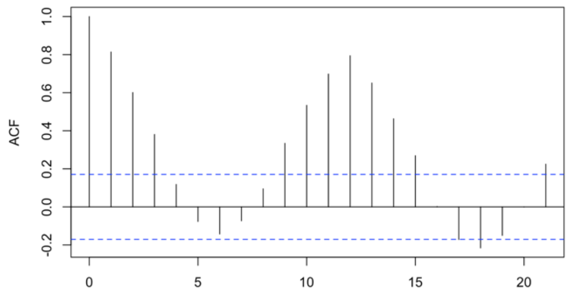


Figure 7a: ACF (data after structural break)

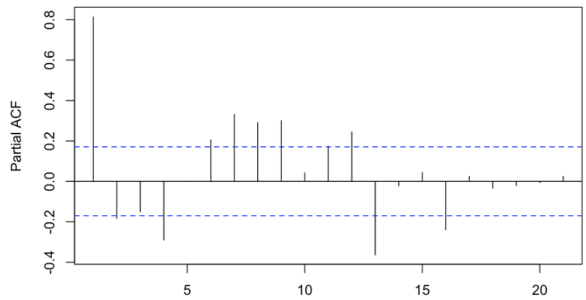


Figure 7b: PACF (data after structural break)

3 Checking for Stationarity

To conduct an effective time series analysis, it is necessary that the data series are stationary. A stationary time series is one in which all statistical properties, including mean, variance, and autocorrelation, remain constant over time. A visual inspection of the time series data is one of the first stages involved in determining stationarity. By plotting the data, one can determine if there are obvious patterns or changes over time. In this particular case, three distinct plots are considered: one highlighting seasonal differences, one highlighting first differences, and one that combines both seasonal and first differences. (Schlitzer, 1995)

Observing these plots allows one to draw preliminary conclusions regarding the stationarity of the data. For example, after incorporating seasonal differences (Figure 8a), the variance appeared to vary over time, indicating non-stationarity. Similarly, the plot of first differences (Figure 8b) revealed robust seasonality, indicating once more the possibility of non-stationarity. However, after integrating first and seasonal differences (Figure 8c), the plot began to resemble white noise, an indicator of stationarity. However, in order to determine whether these assumptions are somewhat accurate, ADF and KPSS tests needs to be conducted. (Schlitzer, 1995)

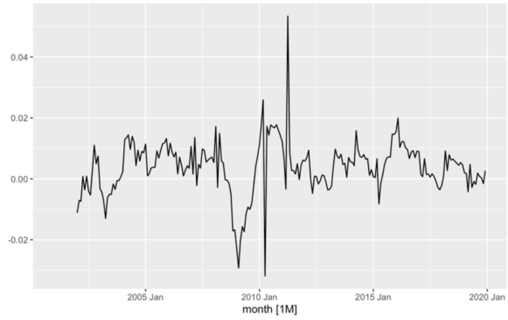


Figure 8a: Seasonal differences

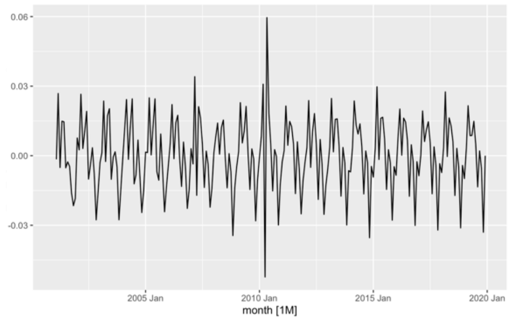


Figure 8b: First differences

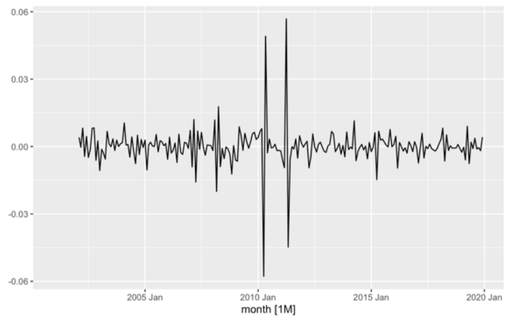


Figure 8c: First and seasonal differences

3.1 ADF & KPSS Test

ADF tests with the type's (trend, drift, none) was employed to assess the presence of a unit root. The test results indicated non-stationarity based on the tau3 test statistic being greater than the critical value corresponding to the 10% significance level in the ADF test with type "trend". In other words, we fail to reject the null hypothesis of a unit root, indicating non-stationarity of the data. In the ADF test with type "drift", the tau2 test statistic did not reject the null hypothesis of a unit root, also suggesting non-stationarity. However, the phi1 test statistic rejects the null hypothesis of no drift, indicating the presence of drift in the data. Furthermore, in the ADF test with type "none", the tau1 test statistic did not reject the null hypothesis of a unit root, again pointing towards non-stationarity.

In addition to the ADF test, the KPSS test was employed with two different types: "tau" (trend and drift) and "mu" (drift). The KPSS test with type "tau" indicated that the test statistic was smaller than the critical value, suggesting that we do not reject the null hypothesis of stationarity, which implies potential stationarity in the data, without a unit root or a deterministic trend. On the other hand, the KPSS test with type "mu" resulted in a test statistic that exceeded the critical value, leading to the rejection of the null hypothesis of no drift. This implies the presence of drift in the data and suggests non-stationarity.

Overall, the ADF test suggests potential non-stationarity, and the KPSS "tau" test indicates the possibility of stationarity, these conflicting findings call for further analysis to clarify the stationarity properties. See Appendix B for the complete results of the ADF and KPSS tests conducted without any differencing. (Kwiatkowski et al., 1992)

3.2 Addressing Non-Stationarity

In the initial analysis, the results of the ADF and KPSS tests yielded contrasting outcomes, potentially due to differing test assumptions and underlying mechanisms. To gain deeper insights, the effects of first differences and seasonal differences were explored to assess their impact on the stationarity of the time series data. The comprehensive results can be found in Appendix C and D.

In the subsequent phase, first differences were applied to the time series and repeated the ADF and KPSS tests. The test statistics in all ADF tests were greater than the critical value at the 10% significance level, indicating the presence of a unit root and non-stationarity once again. However, the KPSS tests, both "tau" and "mu", yielded smaller test statistics than the critical value, suggesting no evidence of a unit root and the time series is stationary. Given that the ADF tests did not yield stationarity with first differences alone, further transformations were needed. As a result, a combination of seasonal and first differences was performed to achieve the stationarity of the time series.

Finally, both first differences and seasonal differences were applied to the time series, and the ADF and KPSS tests were run again. The ADF tests with seasonal and first differences provided evidence supporting the stationarity of the time series. The rejection of the null hypothesis of a unit root was indicated by the test statistics tau3, tau2, and tau1 falling below the critical values. Furthermore, significant results from the phi2 and phi3 tests suggested the presence of deterministic trend and drift components. Overall, these findings support the conclusion that the time series is stationary after applying seasonal and first differences. Also, in the KPSS tests with seasonal and first difference with types "tau" and "mu", the test statistics were both smaller than the critical values, leading to the failure to reject the null hypothesis and suggesting stationarity in the time series. (Kwiatkowski et al., 1992)

In conclusion, the application of first differentiation alone did not achieve stationarity in the time series data. However, the comprehensive ADF and KPSS analyses confirmed the stationarity of the differenced data when both first differences and seasonal differences were applied. For further details, please refer to Appendix E and F.

4 Methodology

This section will describe the methodology used to model the time series and forecast flight traffic in Denmark. Following the data transformation and the discovery of stationarity and a structural break, the objective is to identify the most applicable ARIMA and ETS models for this forecasting. The subsequent subsections will elaborate on these procedures. The dataset was divided into two sets: a training set and a test set, following a 20/80 split. The analysis focused on the data after the structural break in 2008, specifically using the period from 2009 to 2019. The training set consisted of data from 2009 to 2017, while the test set included data from 2018 to 2019.

4.1 ARIMA

The interpretation of the ACF and PACF graphs guided the identification of the optimal ARIMA model. Based on data exploration and analysis, it appears that the most applicable model is the seasonal ARIMA(p,d,q)(P,D,Q) model, where lowercase letters represent non-seasonal components and uppercase letters represent seasonal components. With ARIMA models, only stationary time series must be utilized. (Hyndman & Khandakar, 2008)

First of all, given the application of one seasonal differencing and one first differencing, the values assigned to d and D are both set to 1. The ACF plot indicated significant autocorrelation at lag 1, suggesting the potential inclusion of an autoregressive AR(1) model. Additionally, significant autocorrelation at lag 12 suggested the presence of seasonal behavior, indicating the possible inclusion of a seasonal moving average (SMA) term. This led to the consideration of an ARMA(1,1) model or an AR(1) model with a seasonal MA(1) term, resulting in the initial candidate model ARIMA(1,1,1)(0,1,0)[12].

In contrast, the PACF plot showed significant partial autocorrelation at lag 1, which diminished and became non-significant beyond lag 1. This suggested the inclusion of an autoregressive term in the model, such as an AR(1) model. However, the non-significant spikes at lags 2, 3, 4, 6, and 7 indicated the possibility of residual autocorrelation. This indicated that an AR(1) model or an ARMA(1,0) model might be appropriate for capturing the partial autocorrelation patterns in the data. As an alternative candidate model, the ARIMA(1,1,0)(0,1,1)[12] model was considered. Finally, the auto ARIMA model (1,0,1)(1,1,1)[12] was automatically suggested, incorporating a non-seasonal difference of 1 and a seasonal difference of 1. It included AR and MA seasonal components with values of p = 1 and q = 1.

Table 1: ARIMA Models

	AIC	AICc	BIC
auto - ARIMA(1,0,1)(1,1,1)[12] w/ drift	-681	-680	-665
model1 - ARIMA(1,1,1)(0,1,0)[12]	-626	-652	-618
model2 - ARIMA(1,1,0)(0,1,1)[12]	-653	-626	-645

As shown in Table 1, model1 ARIMA demonstrates the best trade-off between model fit and complexity, as it has the lowest AIC, AICc, and BIC values among the three models. This suggests that model1 ARIMA achieves a better balance in capturing the data patterns while maintaining simplicity.

However, when evaluating the p-values of the Ljung-Box test for the three models, notable differences were observed. The Auto ARIMA model exhibited a significantly higher p-value compared to the other two models. A higher p-value suggests a better fit of the model and indicates a lack of significant autocorrelation in the residuals. In this case, the Auto ARIMA model had a p-value of 0.99, while model1 had a p-value of 0.018 and model2 had a p-value of 0.20.

Furthermore, upon a careful visual inspection of the autocorrelation structure of the residuals (Figure 9), it becomes apparent that the Auto ARIMA model demonstrates a superior fit, as evidenced by the presence of smaller spikes and closer alignment to the horizontal reference line in the ACF plot. Although there is some discernible discrepancy in the innovation residuals, as they do not align perfectly in a straight line, it is important to note that the model still captures a substantial portion of the data's variation. Moreover, the histogram plot of the residuals reveals a slight departure from a perfectly symmetric distribution and displays some skewness. However, even with these deviations from a perfectly normal distribution, the residuals still maintain a shape that closely resembles a normal distribution with a mean of zero. This suggests that the model effectively captures a considerable amount of the underlying data patterns, despite the presence of some remaining imperfections in the fit. For a more detailed analysis of the residuals, please refer to Appendix G, which includes the respective residual graphs for model1 and model2.

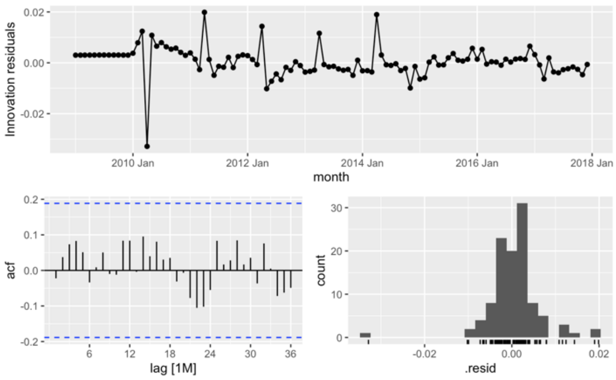


Figure 9: Residuals of ARIMA(1,0,1)(1,1,1) model

However, when considering the forecast accuracies, as seen in Table 2, then the model2 ARIMA has overall the lowest values. These metrices indicates that model2 outperforms the other two models in terms of prediction accuracy and capturing the underlying patterns in the data. Nevertheless, given these very contradictory results, the last test was to visually explore the accuracy of the predictions by plotting the forecasted values from each model alongside the test data.

Table 2: Forecast accuracies on ARIMA Models

	Type	ME	RMSE	MAE	MAPE	MASE	RMSSE	ACF1
auto	*Test*	-0,0047	0,0061	0,0051	0,1660	0,6755	0,6130	0,7267
model1	*Test*	0,0051	0,0057	0,0053	0,1727	0,7040	0,5741	0,1141
model2	*Test*	-0,0010	0,0038	0,0032	0,1036	0,4219	0,3786	0,5629

Figure 10a shows the prediction by the automatically selected ARIMA model, while Figures 10b and 10c display the forecast generated by model1 and model2, respectively. Upon examination of the graphs, it is evident that the automatically selected ARIMA model closely captures the underlying seasonality pattern. Considering all these factors, it is recommended to select the Auto ARIMA model as the preferred choice. It demonstrates superior performance in the Ljung-Box test, the autocorrelation structure of the residuals, the ACF plot, the forecast accuracy, and provides a good visual fit when examining the forecast graphs on the test data.

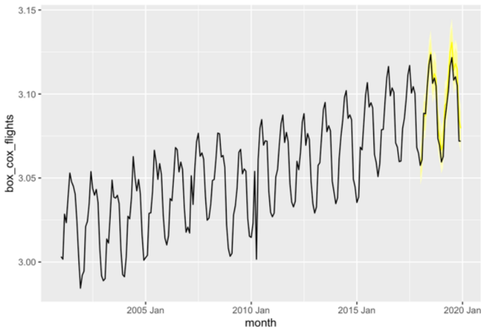


Figure 10a: Forecast with auto ARIMA

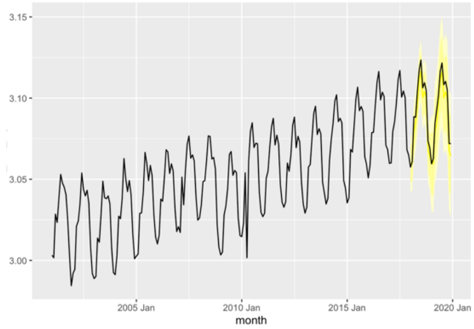


Figure 10b: Forecast with model1

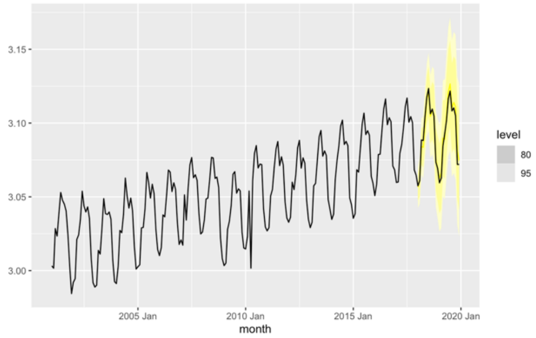


Figure 10c: Forecast with model2

4.2 ETS

During the examination of the ETS's (Error, Trend, Seasonality) models to capture underlying patterns and generate forecasts, three exponential smoothing models were analyzed: "auto_ets," "guess1_ets," and "guess2_ets". For all three models, the residuals were observed to oscillate around the zero line, and the ACF plots revealed no significant spikes, indicating the absence of substantial autocorrelation. Please refer to Appendix H for each model's residual's graph.

The selected models were ETS(A, Ad, A) and ETS(M,Ad,A) for "guessed1" and "guessed2" respectively. The letters "A" and "M" represent additive and multiplicative error components, "Ad" denotes additive trend, and "A" signifies additive seasonality. The auto-selected model suggested the ETS(A, A, A) model. As seen in Table 3, Comparing the AIC, AICc, and BIC values, the guessed2 model outperformed the other two models in terms of goodness of fit. It had the lowest AIC and AICc values among the three models, indicating a better fit to the data. Furthermore, the guessed2 model had a slightly lower BIC value as well, suggesting that it achieves a better trade-off between model complexity and goodness of fit. These findings indicate that the guessed2 model, with an additive error component, additive trend, and additive seasonality (ETS(M, Ad, A)), provided a more suitable representation of the underlying patterns in the data. (Jain & Mallick, 2017)

To compare the models further, the Ljung-Box test and a forecasting accuracy test of the models were conducted. All three models had similar p-values (approx. 0.997), indicating that there is no evidence to reject the null hypothesis of independence for any of the models, suggesting that they adequately capture the autocorrelation structure of the time series.

Going further, when considering the evaluation of forecast accuracy values in Table 4, the guessed1 and guessed2 models shows to perform better than the auto model. Both guessed1 and guessed2 models have lower values for these metrices, indicating that they have lower prediction errors and provide more accurate forecasts compared to the auto model. Taken together, the evaluation metrics reinforce the finding that the guessed2 model (ETS(M, Ad, A)) is the most appropriate model for generating forecasts for the flight traffic data. Please refer to Appendix H for the residuals plots of all three ETS models.

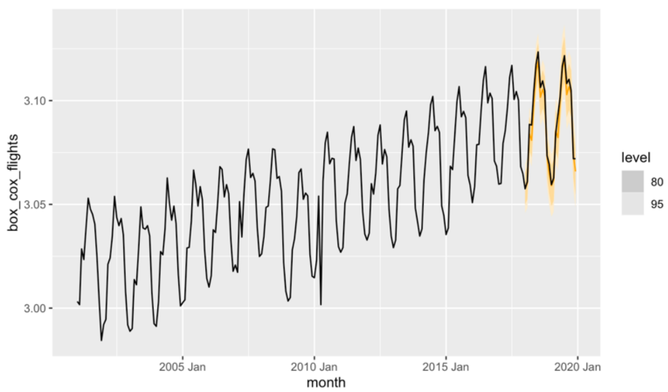


Figure 11: Forecast with ETS(M, Ad, A) model.

Table 3: ARIMA Models

	AIC	AICc	BIC
auto - ETS(A,A,A)	-583	-576	-537
guessed1 - ETS(A,Ad,A)	-582	-575	-534
guessed2 - ETS(M,Ad,A)	-581	-573	-533

Table 4: Forecast accuracies on ETS Models

	Type	ME	RMSE	MAE	MAPE	MASE	RMSSE	ACF1
auto	*Test*	-0,0030	0,0052	0,0042	0,1366	0,5561	0,5173	0,5243
guessed1	*Test*	0,0030	0,0042	0,0033	0,1073	0,4375	0,4193	-0,0317
guessed2	*Test*	0,0030	0,0042	0,0033	0,1083	0,4418	0,4179	-0,0610

4.3 Seasonal Naïve

Lastly, the seasonal naive (SNAIVE) model was included as a benchmark in the analysis. The SNAIVE model exhibited a relatively large variance, indicating that its predictions deviated significantly from the actual data points. This suggests that the SNAIVE model may not have effectively captured the underlying patterns and dynamics of the data, resulting in less accurate forecasts compared to the more advanced ARIMA and ETS models. In the accuracy assessment of the forecasts, the SNAIVE model exhibited higher errors compared to the ETS and ARIMA models, as expected due to its simplicity. Please refer to Appendix I for the residuals plot, forecast accuracy metrices and the forecast plots for SNAIVE.

5 Conclusion

In conclusion, the primary aim of this study was to identify the optimal model for forecasting flight traffic and assess its potential growth in the absence of a pandemic. Various ARIMA and ETS models were evaluated, and the most suitable models were compared in terms of their accuracy in predicting future data. The auto ARIMA model (1,0,1)(1,1,1)[12]) demonstrated great accuracy in capturing seasonal fluctuations in the flight traffic data. For example, it had a relatively low root mean square error (RMSE) of 0.0061, indicating that its forecasts were close to the actual values.

On the contrary, when evaluating the performance of the ETS(M,Ad,A) model, it exhibited superior forecast accuracy values compared to the auto ARIMA model. Notably, the ETS(M,Ad,A) model achieved a notably lower RMSE of 0.0042, signifying enhanced overall prediction accuracy. Furthermore, the ETS model demonstrated a lower MAE of 0.0033, indicating superior capability in capturing the average magnitude of errors. Additionally, the ACF1 value of -0.061 observed for the ETS model suggests a robust autocorrelation in the residuals, potentially implying a more favourable fit to the time series data.

Overall, based on the comparison of these models, the ETS(M,Ad,A) model demonstrated the best forecasting accuracy in capturing the seasonal fluctuations of the air traffic data, followed by the Auto ARIMA model. Future research endeavours could concentrate on predicting the duration of the pandemic's impact and the timeline for air traffic to re-gain pre-pandemic levels. Finally, Figure 12 provides a visual representation of the actual air traffic patterns during and after the pandemic.

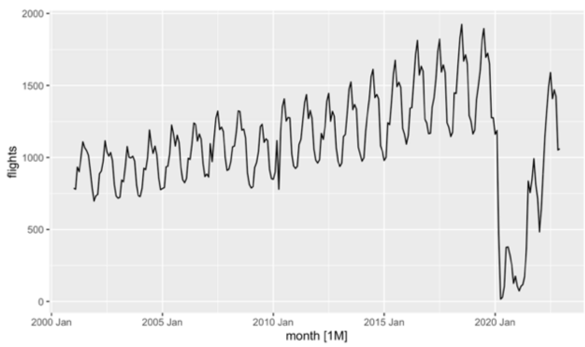


Figure 12: The air traffic in Denmark during the pandemic

6 References

Chen, Y. (2021). An analytical process of spatial autocorrelation functions based on Moran's index. PLoS One, 16(4), e0249589.

Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: Principles and Practice. OTexts.

Hyndman, R. J., & Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for R. Journal of Statistical Software, 27(1), 1–22. [[DOI]](https://doi.org/10.18637/jss.v027.i03)

Jain, G., & Mallick, B. (2017). A study of time series models ARIMA and ETS. Available at SSRN 2898968.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? Journal of Econometrics, 54(1-3), 159–178. [[DOI]](https://doi.org/10.1016/0304-4076(92)90104-Y)

Muthuramu, P., & Maheswari, T. U. (2019). Tests for structural breaks in time series analysis: A review of recent development. Shanlax International Journal of Economics, 7(4), 66-79.

Nhamo, G., Dube, K., Chikodzi, D. (2020). COVID-19 and Implications for the Aviation Sector: A Global Perspective. In: Counting the Cost of COVID-19 on the Global Tourism Industry. Springer, Cham. https://doi.org/10.1007/978-3-030-56231-1_4

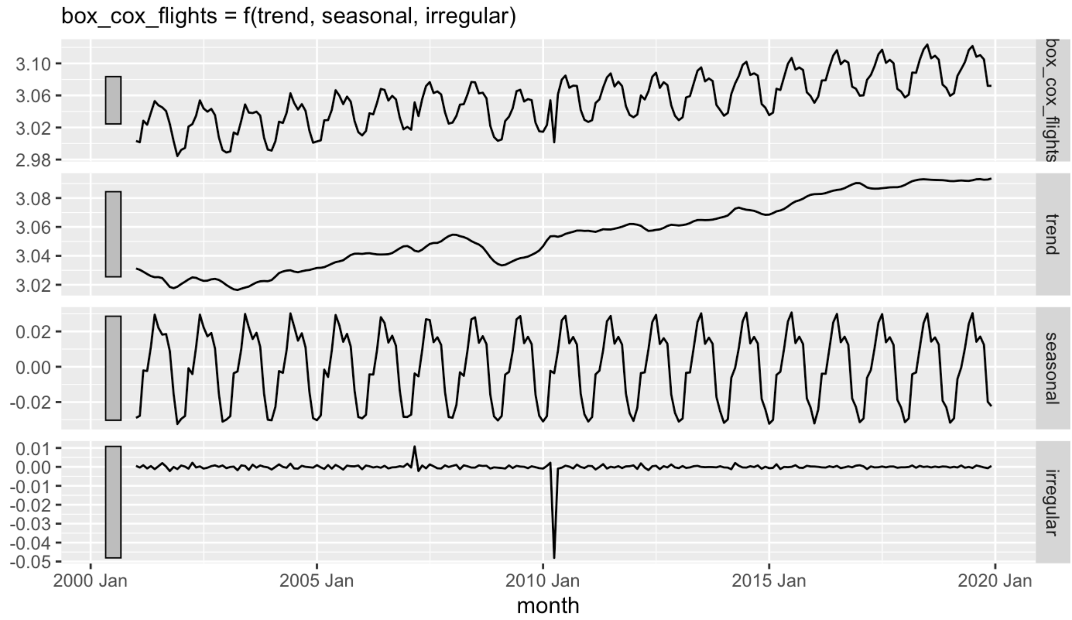
Osborne, J. (2010). Improving your data transformations: Applying the Box-Cox transformation. Practical Assessment, Research, and Evaluation, 15(1), 12.

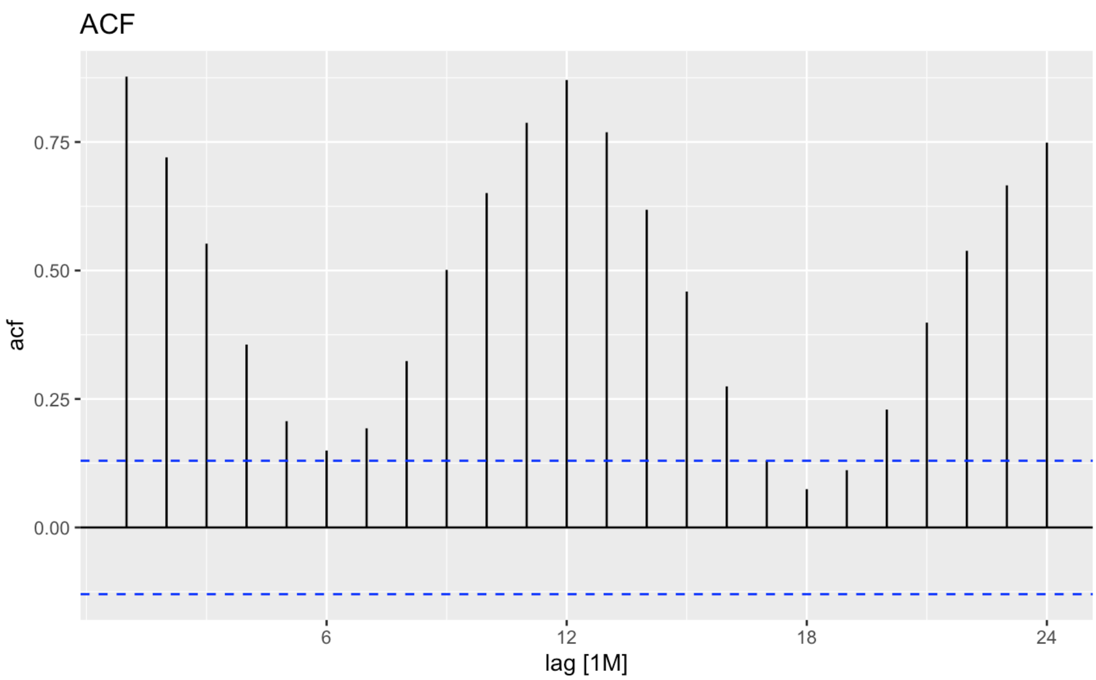
Schlitzer, G. (1995). Testing the stationarity of economic time series: further Monte Carlo evidence. Ricerche Economiche, 49(2), 125-144.

Statistics Denmark. (2023, June 11). FLYV92: Departing passengers from major, manned, public airports by type of transport and flight [Data table]. Retrieved from www.statbank.dk/FLYV92

7 Appendices

Appendix A – SEATS Decomposition and ACF of Box-Cox transformed data.





***Appendix B** – ADF and KPSS without differencing.*

```
Residual standard error: 0.007088 on 200 degrees of freedom
Multiple R-squared:  0.7985,    Adjusted R-squared:  0.7844
F-statistic:  56.6 on 14 and 200 DF,  p-value: < 2.2e-16


Value of test-statistic is: -2.144 5.031 2.3485

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -3.99 -3.43 -3.13
phi2  6.22  4.75  4.07
phi3  8.43  6.49  5.47
```

a) ADF trend

```
Residual standard error: 0.007152 on 201 degrees of freedom
Multiple R-squared:  0.7938,    Adjusted R-squared:  0.7804
F-statistic: 59.51 on 13 and 201 DF,  p-value: < 2.2e-16


Value of test-statistic is: -0.1375 5.114

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.46 -2.88 -2.57
phi1  6.52  4.63  3.81
```

b) ADF drift

```
Residual standard error: 0.007135 on 202 degrees of freedom
Multiple R-squared:  0.7939,    Adjusted R-squared:  0.7806
F-statistic: 59.84 on 13 and 202 DF,  p-value: < 2.2e-16


Value of test-statistic is: 3.2017

Critical values for test statistics:
      1pct  5pct 10pct
tau1 -2.58 -1.95 -1.62
```

c) ADF none

```
Value of test-statistic is: 0.0376

Critical value for a significance level of:
                10pct  5pct 2.5pct  1pct
critical values 0.119 0.146  0.176 0.216
```

d) KPSS tau

```
Value of test-statistic is: 3.2009

Critical value for a significance level of:
                10pct  5pct 2.5pct  1pct
critical values 0.347 0.463  0.574 0.739
```

e) KPSS mu

__Appendix C__ – ADF and KPSS with first differencing.

```
Residual standard error: 0.04441 on 188 degrees of freedom
Multiple R-squared:  0.5145,	Adjusted R-squared:  0.4784
F-statistic: 14.23 on 14 and 188 DF,  p-value: < 2.2e-16


Value of test-statistic is: -4.5955 7.0605 10.5907

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -3.99 -3.43 -3.13
phi2  6.22  4.75  4.07
phi3  8.43  6.49  5.47
```

a) ADF trend

```
Residual standard error: 0.04433 on 189 degrees of freedom
Multiple R-squared:  0.5138,	Adjusted R-squared:  0.4803
F-statistic: 15.36 on 13 and 189 DF,  p-value: < 2.2e-16


Value of test-statistic is: -4.5778 10.4782

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.46 -2.88 -2.57
phi1  6.52  4.63  3.81
```

b) ADF drift

```
Residual standard error: 0.04523 on 190 degrees of freedom
Multiple R-squared:  0.4912,	Adjusted R-squared:  0.4564
F-statistic: 14.11 on 13 and 190 DF,  p-value: < 2.2e-16


Value of test-statistic is: -3.4207

Critical values for test statistics:
      1pct  5pct 10pct
tau1 -2.58 -1.95 -1.62
```

c) ADF none

```
Value of test-statistic is: 0.078

Critical value for a significance level of:
                10pct  5pct 2.5pct  1pct
critical values 0.119 0.146  0.176 0.216
```

d) KPSS tau

```
Value of test-statistic is: 0.1464

Critical value for a significance level of:
                10pct  5pct 2.5pct  1pct
critical values 0.347 0.463  0.574 0.739
```

e) KPSS mu

Appendix D –*ADF and KPSS with first and seasonal differencing.*

```
Residual standard error: 0.04384 on 187 degrees of freedom
Multiple R-squared:  0.8445,	Adjusted R-squared:  0.8329
F-statistic: 72.57 on 14 and 187 DF,  p-value: < 2.2e-16


Value of test-statistic is: -6.1494 12.6142 18.9109

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -3.99 -3.43 -3.13
phi2  6.22  4.75  4.07
phi3  8.43  6.49  5.47
```

a) ADF trend

```
Residual standard error: 0.04376 on 188 degrees of freedom
Multiple R-squared:  0.8443,	Adjusted R-squared:  0.8335
F-statistic: 78.42 on 13 and 188 DF,  p-value: < 2.2e-16


Value of test-statistic is: -6.1378 18.8467

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.46 -2.88 -2.57
phi1  6.52  4.63  3.81
```

b) ADF drift

```
Residual standard error: 0.04365 on 189 degrees of freedom
Multiple R-squared:  0.8443,	Adjusted R-squared:  0.8336
F-statistic: 78.82 on 13 and 189 DF,  p-value: < 2.2e-16


Value of test-statistic is: -6.1526

Critical values for test statistics:
      1pct  5pct 10pct
tau1 -2.58 -1.95 -1.62
```

c) ADF none

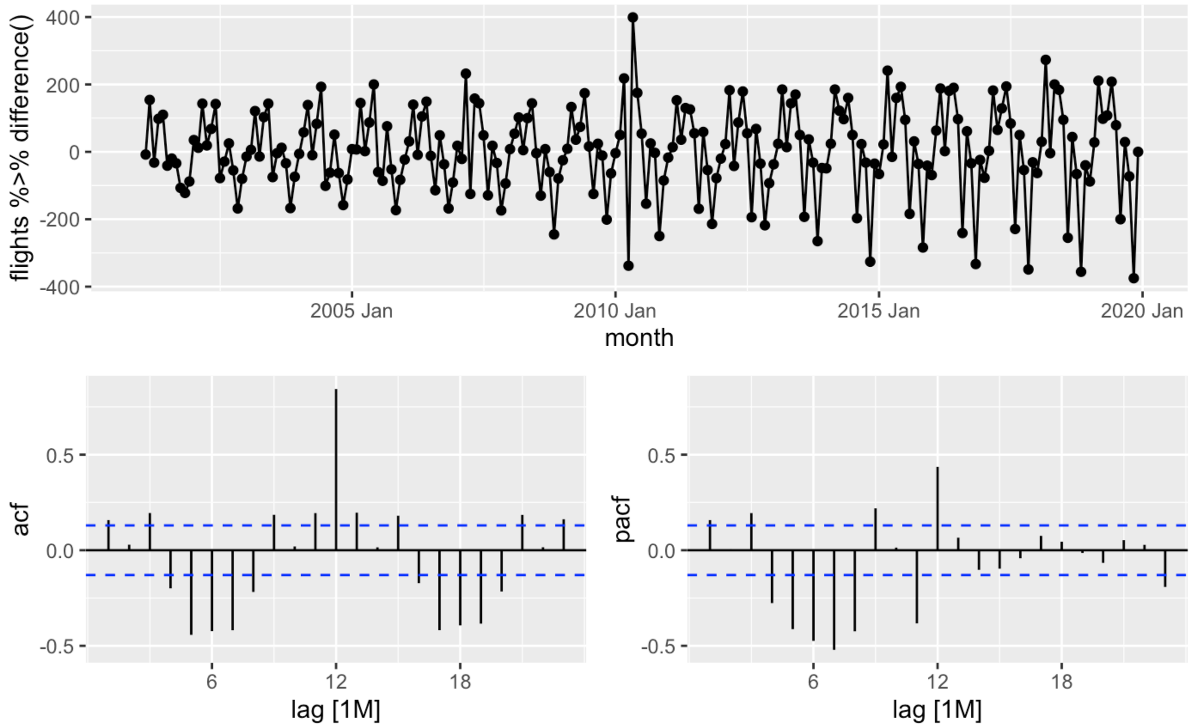
```
Value of test-statistic is: 0.0202

Critical value for a significance level of:
                10pct  5pct 2.5pct  1pct
critical values 0.119 0.146  0.176 0.216
```

d) KPSS tau

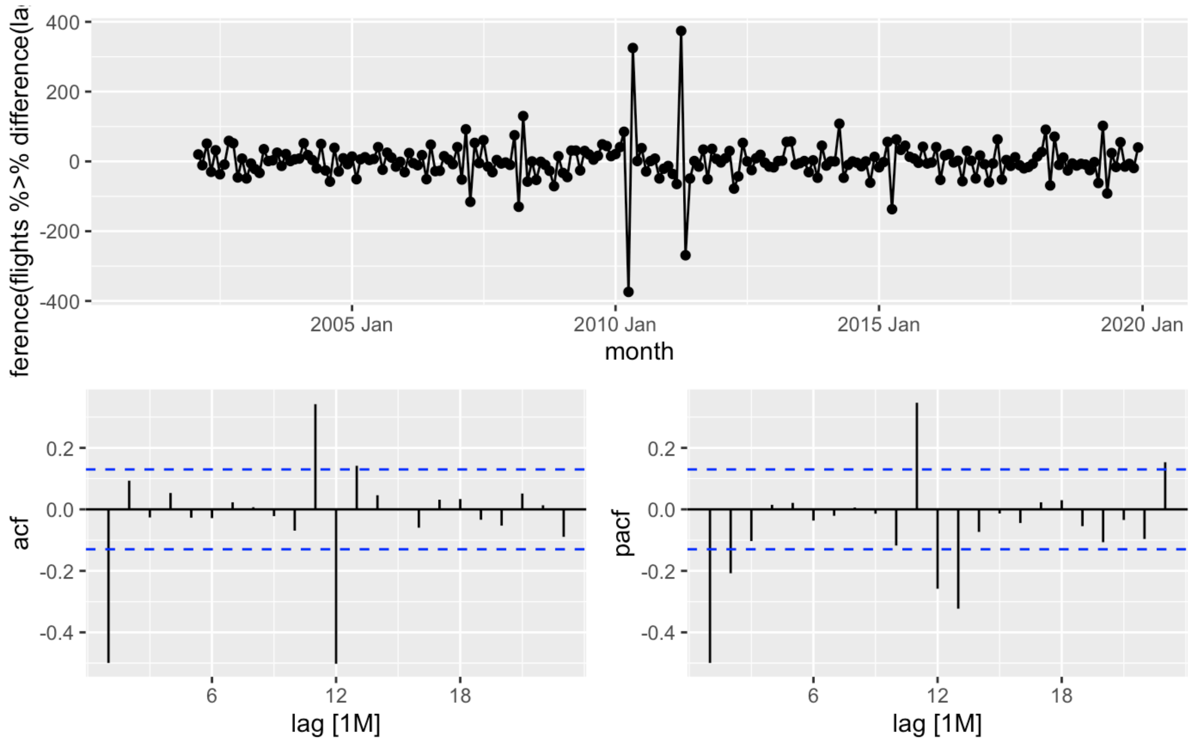
```
Value of test-statistic is: 0.0323

Critical value for a significance level of:
                10pct  5pct 2.5pct  1pct
critical values 0.347 0.463  0.574 0.739
```

e) KPSS mu

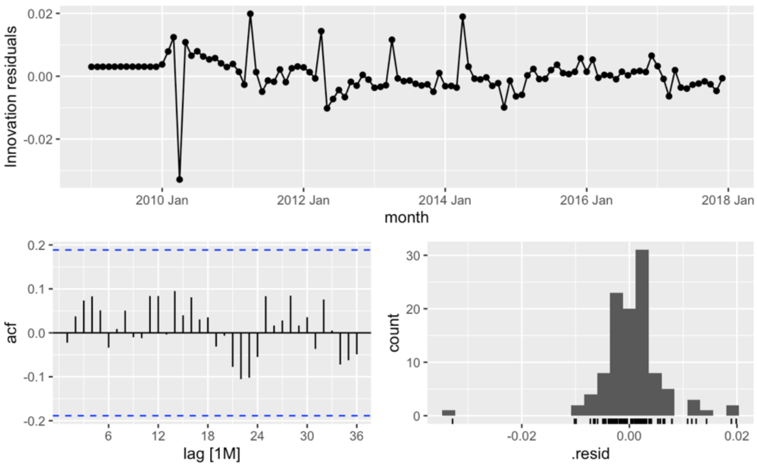
Appendix E – Residuals on data with first difference.



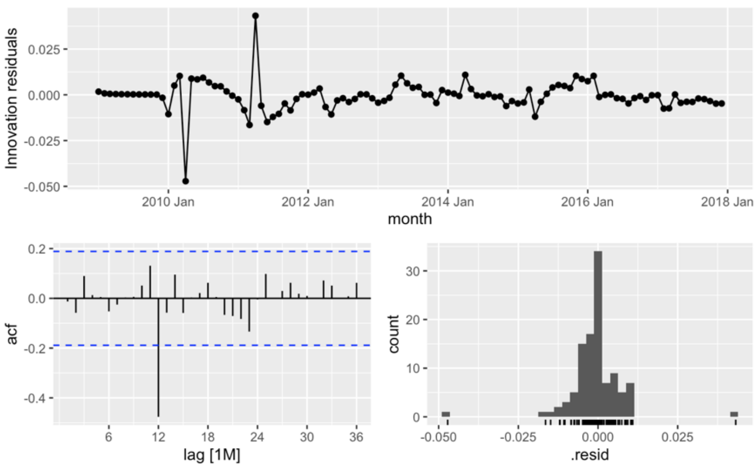
Appendix F – Residuals on data with first and seasonal differencing.



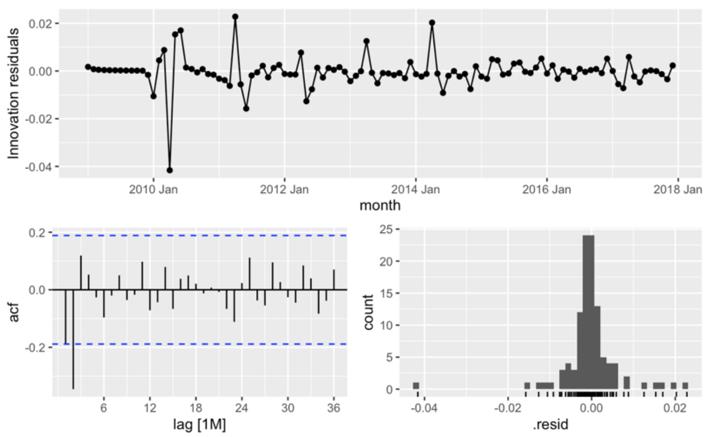
Appendix G – *Residuals on IRMA models*



a) ARIMA auto (1,0,1)(1,1,1)[12] w/drift

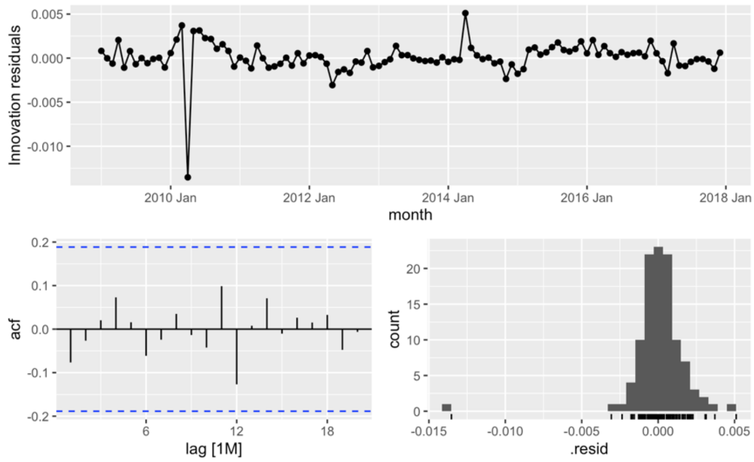


b) ARIMA model1 (1,1,1)(0,1,0)[12]

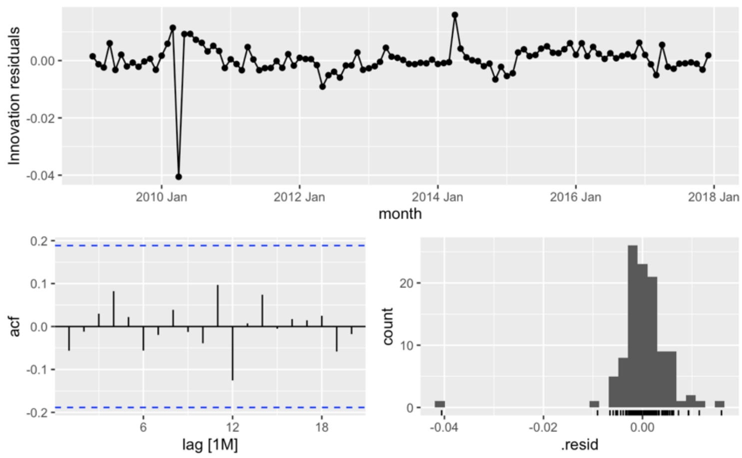


c) ARIMA model2 (1,1,0)(0,1,1)[12]

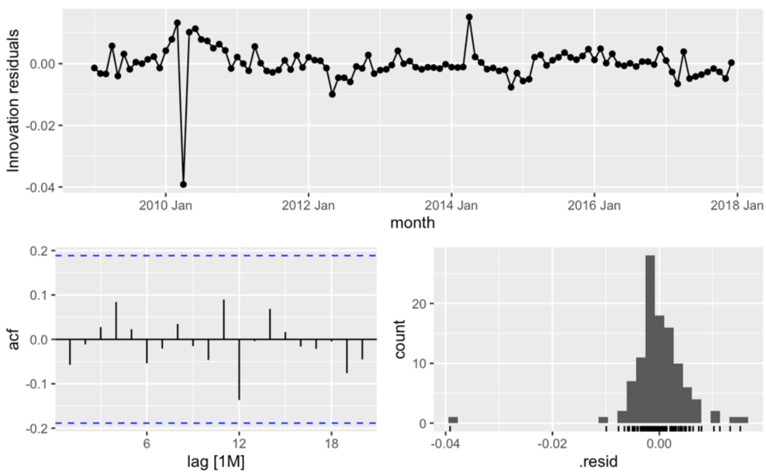
Appendix H *– Residuals on ETS models.*



a) ETS auto

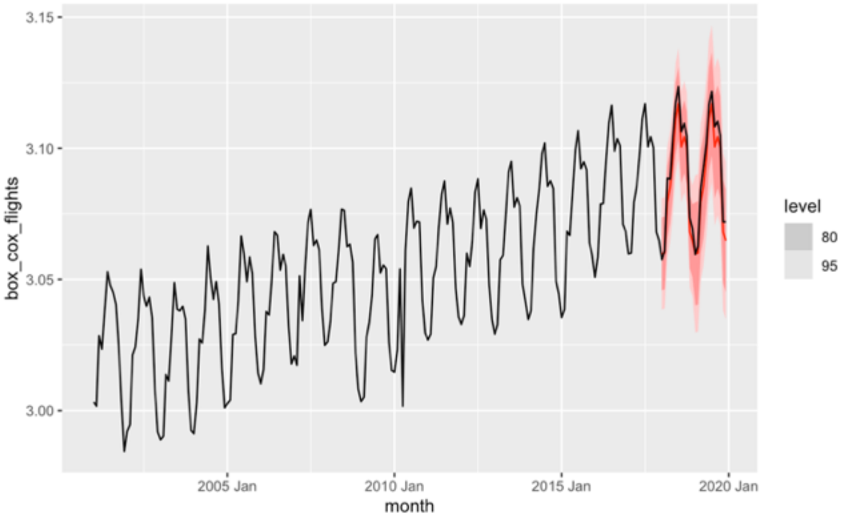


b) ETS(A,Ad,A)

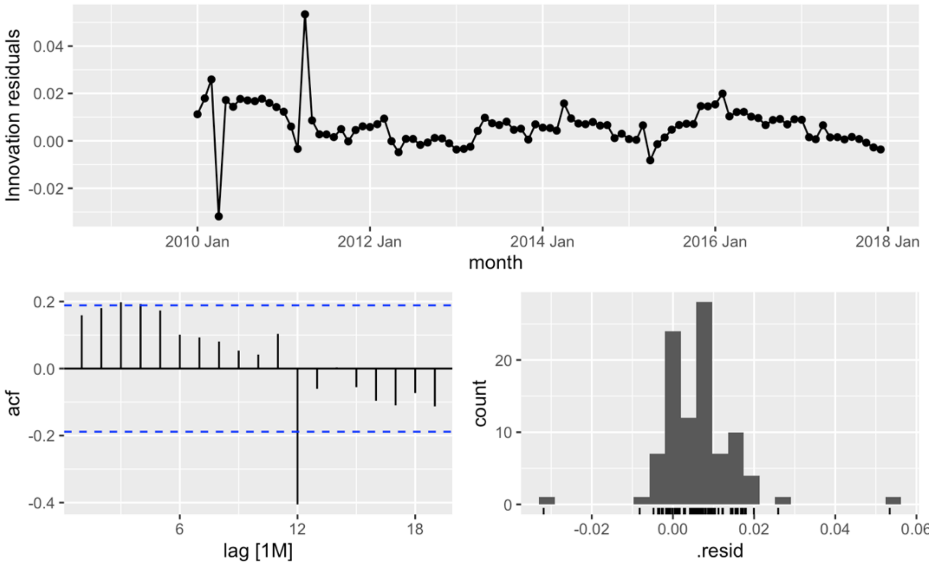


c) ETS(M,Ad,A)

__Appendix I__ – SNAIVE model



a) Forecast with the SNAIVE model.



b) Residuals on SNAIVE model.

.model	.type	ME	RMSE	MAE	MPE	MAPE	MASE	RMSSE	ACF1
drift	Test	0.020201	0.029032	0.024395	0.648838	0.785720	3.211209	2.883883	0.696909
mean	Test	0.026663	0.033962	0.028424	0.857678	0.915238	3.741535	3.373581	0.696894
naive	Test	0.027369	0.034520	0.028895	0.880537	0.930398	3.803560	3.428974	0.696894
snaive	Test	0.004819	0.005459	0.005026	0.155527	0.162287	0.661597	0.542232	0.101722

c) Forecast accuracy on SNAIVE model.