Advanced Statistical Modelling
Time Series

Yifan Yao, Dinara Kurmangaliyeva

December 2024

1 Introduction

1.1 Dataset

This analysis focuses on the dataset **Turismos**, which captures the monthly production of cars in Spain, measured in thousands of units. The data span from 1994 onward and are organized as time series. The "Turismos" dataset was sourced from Spain's Ministry of Industry, Trade, and Tourism's statistical portal

1.2 Goal

This project aims to apply the Box-Jenkins ARIMA methodology, incorporating extensions to address calendar effects to the analysis and prediction of the monthly car production time series in Spain. The study seeks to model the underlying patterns, account seasonality, and generate accurate forecasts of future car production trends.

2 Time Series Analysis

2.1 Data Exploration

The data exhibits a clear seasonal pattern with sharp fluctuations throughout each year in Figure 1. There is an upward trend in production from 1995 to the mid-2000s, followed by a decline and then recovery leading up to 2020.

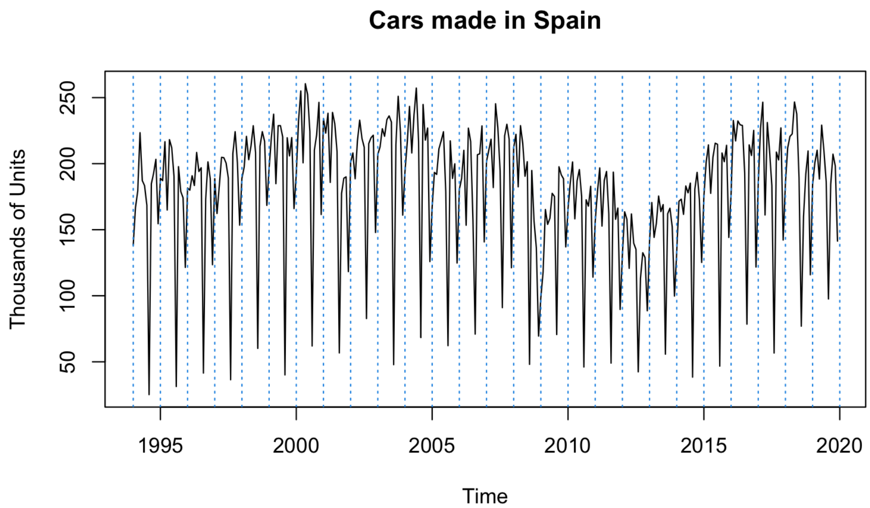

Figure 1

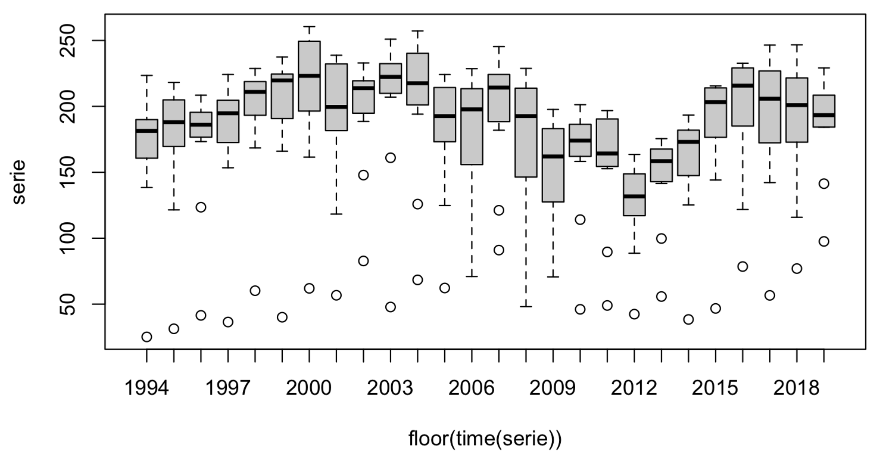

Figure 2

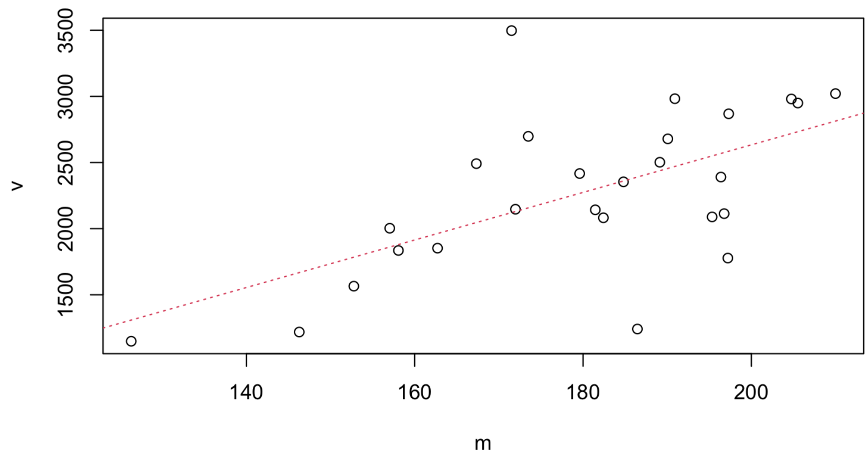

Figure 3

2.2 Transformation

From the visualizations (Figure 2) and the observed relationship between mean and variance (Figure 3), we can conclude that the time series has a **non-constant variance**. The variance is not constant over time, as indicated by the increasing spread of data values in later years. In addition, the variance appears to increase as the mean increases, suggesting a positive dependency between the two. In order to tackle non-constant variance, we apply transformations like logarithms to reduce heteroscedasticity:

```r
# Transform the series into log scale
lnserie <- log(serie)
```

The first transformation is completed. As we can observe from the transformed boxplot(Figure 4), the logarithmic transformation reduces the scale of larger values, effectively compressing the variability range. The relation(Figure 5) between mean and variance now appears to be more constant over time. However, we should keep in mind that the transformation does not eliminate the trend or seasonality in the data.

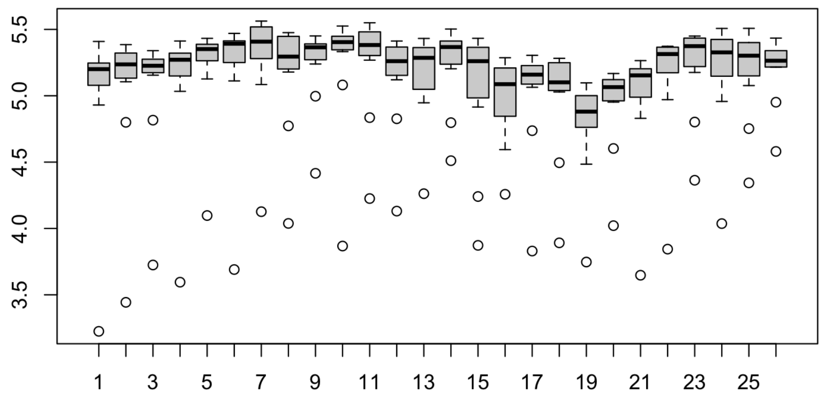

Figure 4

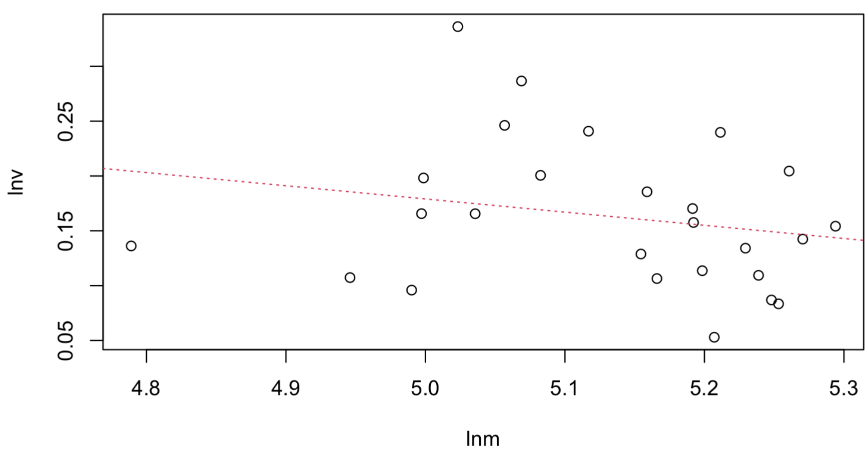

Figure 5

Is seasonality present? Yes, seasonality is present, and has an obvious downfall every august across all years which can be seen in the following monthplot (Figure 6 and Figure 7). We can assume that each august production is significantly decreased because many company employees go on vacation during that time.

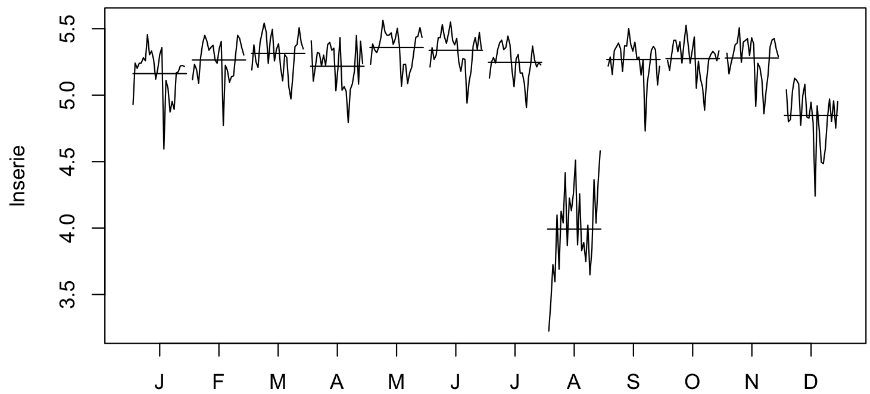

Figure 6

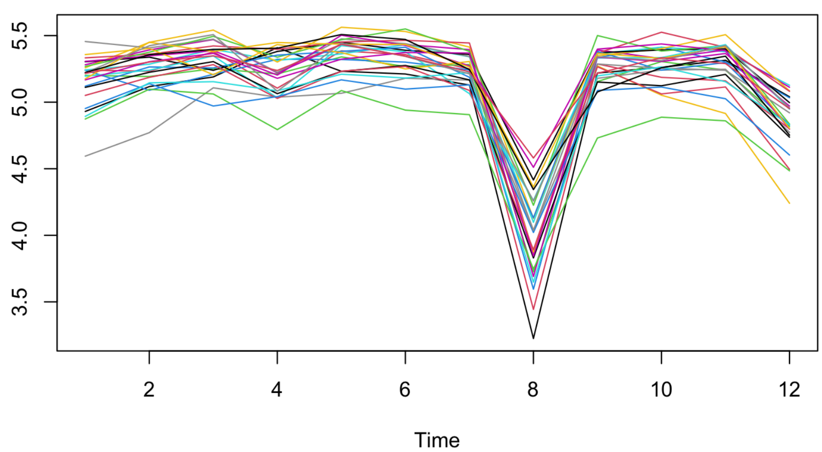

Figure 7

In order to make the data more stationary we remove both seasonal patterns and long-term trends. Seasonal differencing with a lag of 12 removes the repeating patterns caused by seasonality:

```
d12lnserie<-diff(lnserie,lag=12)
```

From the plots(Firgure 8, 9, 10), it seems the residuals no longer show a clear recurring pattern specific to a particular time, which suggests the B12 transformation successfully reduced the August seasonal effect.

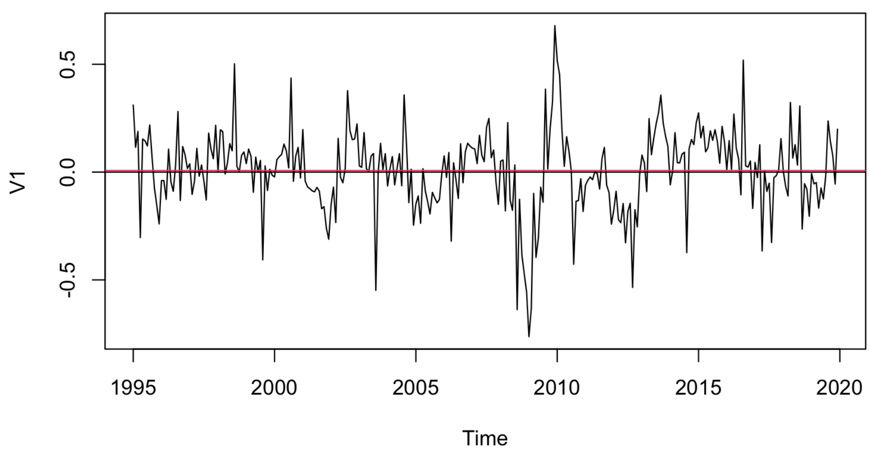

Figure 8

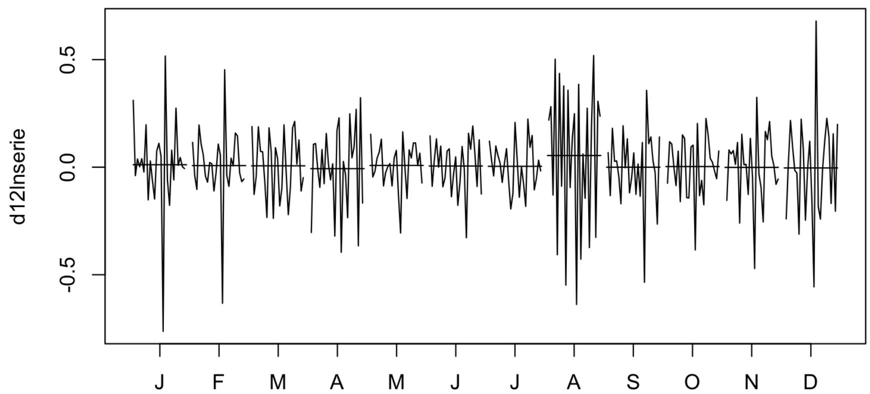

Figure 9

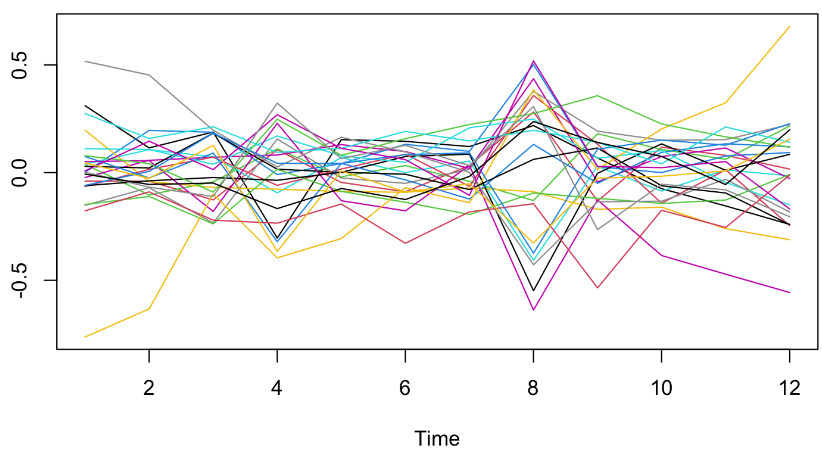

Figure 10

Is the mean constant? In our analysis, we check whether the mean and variance of the data remain stable after applying successive differencing. The mean of d12lnserie is approximately 0.007353 as calculated, and as shown in

Figure 8. Mean appears to be relatively constant.

We check if we need further transformations to achieve less variability in series. The variance of lnserie is 0.164901. This is the variance of the log-transformed data. The variance after the first differencing (d12lnserie) is 0.03583941, which is lower than the variance of lnserie. This suggests that differencing has reduced the overall variability in the series, which is expected as differencing helps remove trends or long-term components. We tried regular differencing (d1d12lnserie), which is slightly higher than the variance of d12lnserie.

Transformation	**Variance**
ln(serie)	0.164901
d_{12} ln(serie)	**0.03583941**
$d_1 d_{12}$ ln(serie)	0.04014573

Table 1: Variance comparison across transformations

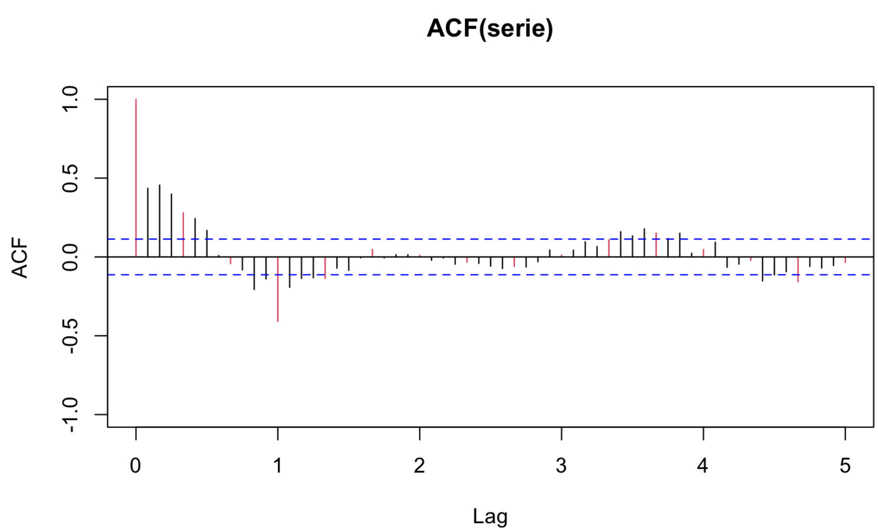

Figure 11

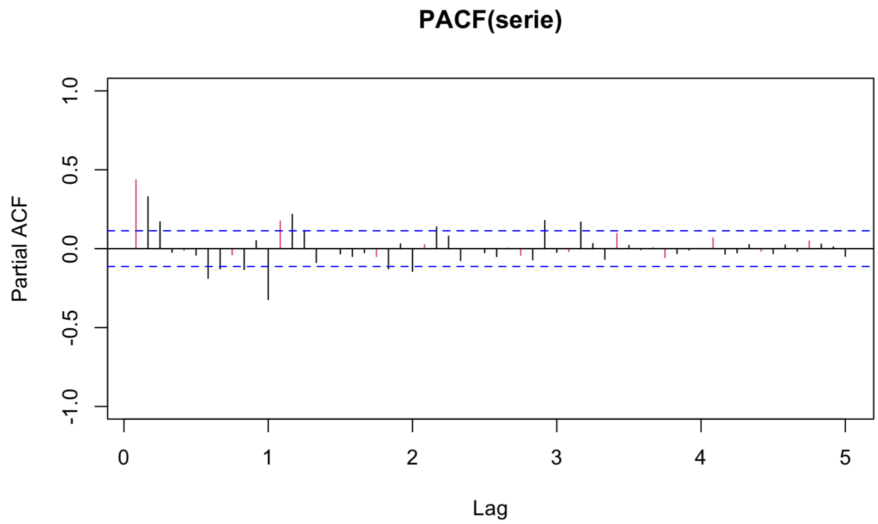

Figure 12

Is the series stationary? The ACF (Figure 11) drops quickly and becomes insignificant (crosses the confidence bands) after a few lags. The PACF (Figure 12) shows significant spikes only at the first few lags and then quickly becomes insignificant (inside the confidence bands). This behavior is also consistent with a stationary series, as there is no strong dependency on higher-order lags. Based on the plots, the series appears to be stationary, as the ACF and PACF both indicate a lack of long-term trends or seasonal dependencies.

Given observations above, we conclude that data is stationary after applying seasonal difference and log transformation. We do not do any further transformations.

Final transformation: $W_t = (1 - B^{12}) \log X_t$

2.3 Estimation

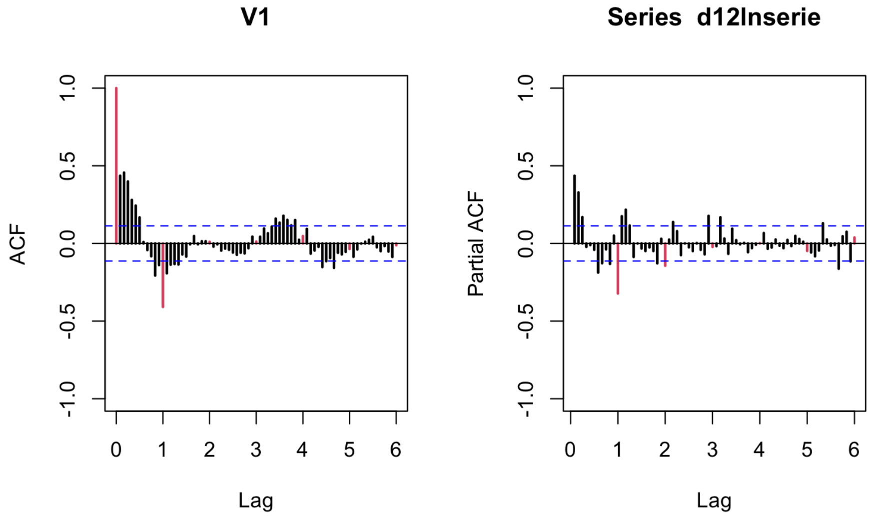

Figure 13

Based on Figure 13, we propose three models. Model 1 and Model 2 were identified after analyzing lags in the ACF and PACF. Model 3 can be seen as a parsimonious version of the ARIMA(3,0,0) and ARIMA(0,0,6) models, it offers a balanced approach with minimal parameters.

Models

Model 1: ARIMA(3,0,0)(0,1,1)$_{12}$

$$(1-\phi_1 B-\phi_2 B^2-\phi_3 B^3)(1-B^{12})\log X_t=(1+\Theta B^{12})Z_t$$

Model 2: ARIMA(0,0,6)(0,1,1)$_{12}$

$$(1-B^{12})\log X_t=(1+\theta_1 B+\theta_2 B^2+\theta_3 B^3+\theta_4 B^4+\theta_5 B^5+\theta_6 B^6)(1+\Theta B^{12})Z_t$$

Model 3: ARIMA(1,0,1)(0,1,1)$_{12}$

$$(1-\phi_1 B)(1-B^{12})\log X_t=(1+\theta_1 B)(1+\Theta B^{12})Z_t$$

Models with Coefficients

Model 1:

$$(1-0.2302B-0.2949B^2-0.2604B^3)(1-B^{12})\log X_t=(1-0.6868B^{12})Z_t$$

Model 2:

$$(1-B^{12})\log X_t = (1+0.2306B+0.3329B^2+0.3308B^3+0.1566B^4+0.2014B^5+0.2190B^6)(1-0.6372B^{12})Z_t$$

Model 3:

$$(1-0.9286B)(1-B^{12})\log X_t = (1-0.6160B)(1-0.7012B^{12})Z_t$$

2.4 Validation

2.4.1 Residual Analysis

Model 1 In Figure 14, residuals appear random around zero, with constant variance and approximate normality (some deviations at the tails).

Descriptive Statistics for the Residuals can be found in Appendix A.

- **Normality tests** (Shapiro-Wilk, Anderson-Darling, Jarque-Bera) reject normality (likely due to test sensitivity).
- **Homoscedasticity test** (Breusch-Pagan) rejects homoscedasticity (likely due to test sensitivity).
- **Durbin-Watson test** shows no significant first-order autocorrelation.
- **Ljung-Box test** shows no significant autocorrelation at various lags.

In Figure 16, standardized residuals appear random, and the ACF shows no significant autocorrelations.

In Figure 16, across all lags tested, the Ljung-Box p-values are consistently high, suggesting that the residuals are independent and there is no evidence of significant autocorrelation. This is a good outcome, indicating that the model has successfully captured the structure of the data, leaving residuals that are essentially random.

Overall, we believe that Model 1 exhibits normality, independence, and constant variance across observations.

Model 2 In Figure 18, residuals appear random around zero, with constant variance and approximate normality (some deviations at the tails).

Descriptive Statistics for the Residuals can be found in Appendix B.

- **Normality tests** (Shapiro-Wilk, Anderson-Darling, Jarque-Bera) reject normality (likely due to test sensitivity).
- **Homoscedasticity test** (Breusch-Pagan) rejects homoscedasticity (likely due to test sensitivity).
- **Durbin-Watson test** shows no significant first-order autocorrelation.

- **Ljung-Box test**: While most lags show no significant autocorrelation, lags 12 and 24 have p-values just above 0.04, very close to the 0.05 significance level. This suggests possible autocorrelation at these seasonal lags.

In Figure 20, standardized residuals appear random, and the ACF shows no significant autocorrelations.

In Figure 20, we can observe that p-values for Ljung-Box statistic, for lower lags (1 to 4), the model appears to fit well, with no significant autocorrelation in the residuals. However, there is evidence of significant autocorrelation at lag 12 and lag 24, as the p-values are less than 0.05 for these lags. This suggests that the model may not have fully captured some seasonal or periodic structure in the data.

Overall, we believe that Model 2 follows a normal distribution and maintains constant variance; however, it potentially displays autocorrelation at higher lags.

Model 3 In Figure 22, the residuals seem random and centered around zero, showing consistent variance and mostly normal distribution, though there are some small deviations at the tails.

Descriptive Statistics for the Residuals can be found in Appendix C.

- **Normality tests** (Shapiro-Wilk, Anderson-Darling, Jarque-Bera) reject normality (likely due to test sensitivity).
- **Homoscedasticity test** (Breusch-Pagan) rejects homoscedasticity (likely due to test sensitivity).
- **Durbin-Watson test** shows no significant first-order autocorrelation.
- **Ljung-Box test** crucially, shows significant autocorrelation at lags 3, 4, 12, 24, 36 and 48 (p-values are less than 0.05).

In Figure 24, standardized residuals appear random, and the ACF shows no significant autocorrelations.

In Figure 24, we observe that the p-values for the Ljung-Box statistic at lower lags (1 to 2) indicate a good model fit, showing no significant autocorrelation in the residuals. However, for lags beyond 2, the p-values fall below 0.05, indicating significant autocorrelation. This suggests that the model has not fully captured some seasonal effects.

Overall, we believe that Model 3 follows a normal distribution and maintains constant variance; however, it potentially displays autocorrelation.

Validation Plots for Model 1:

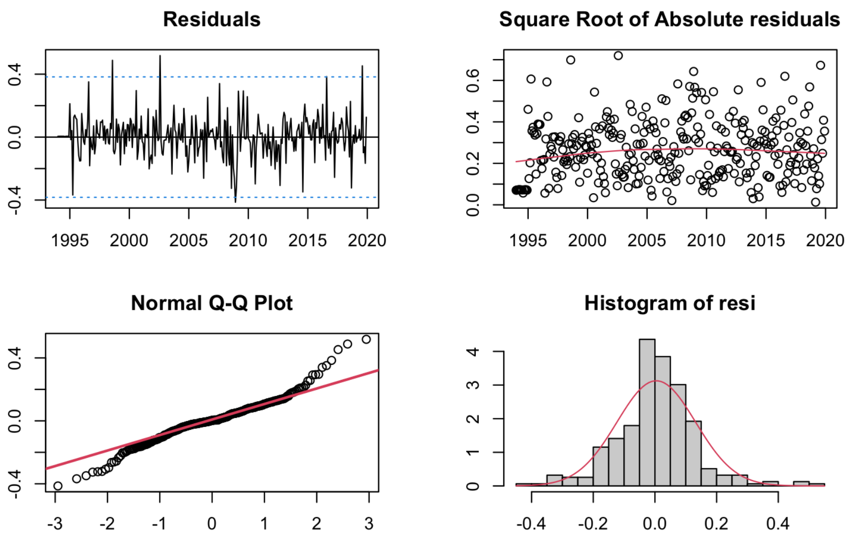

Figure 14

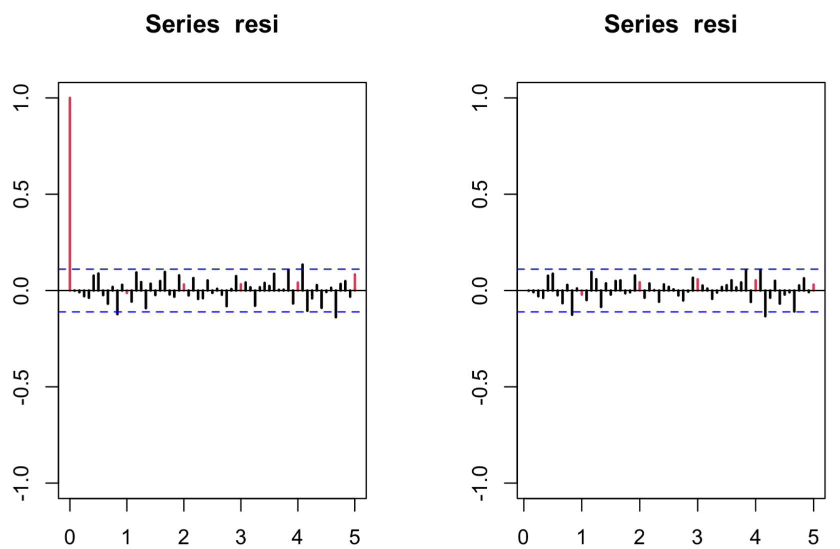

Figure 15

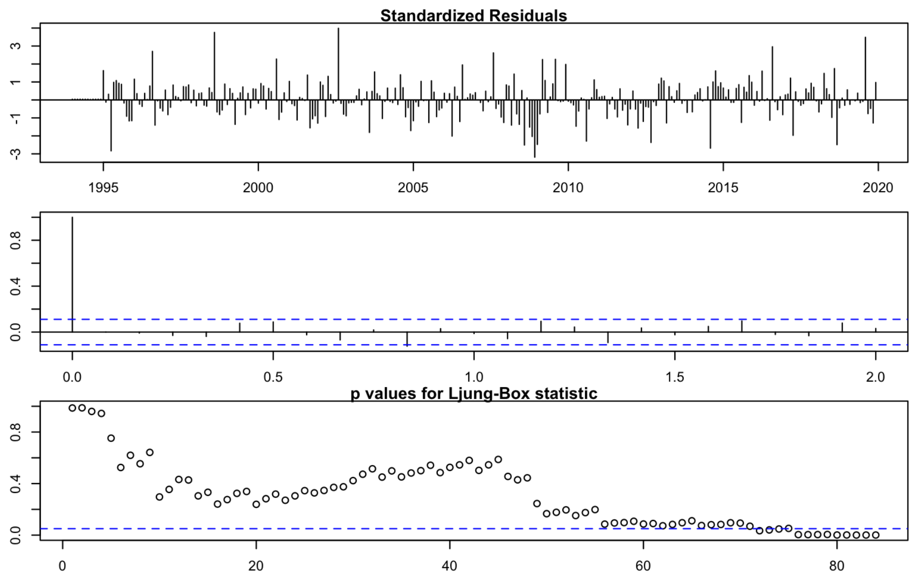

Figure 16

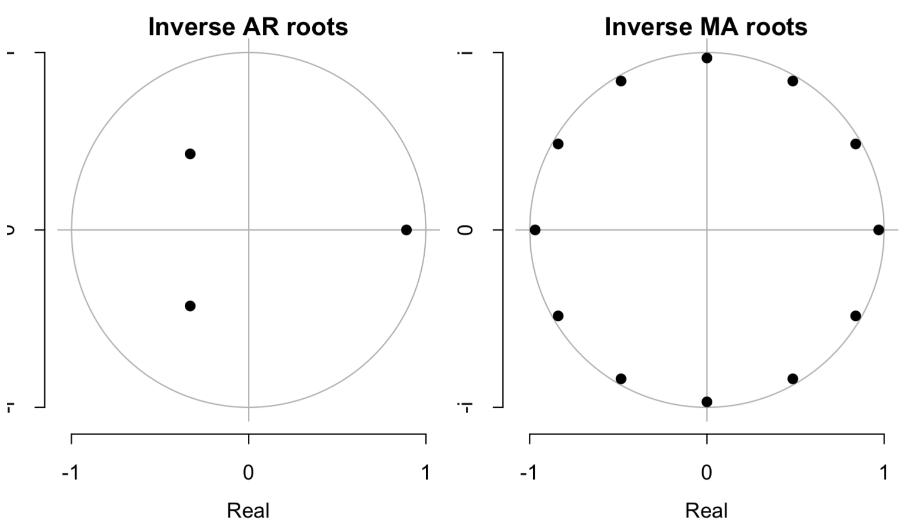

Figure 17

Validation Plots for Model 2:

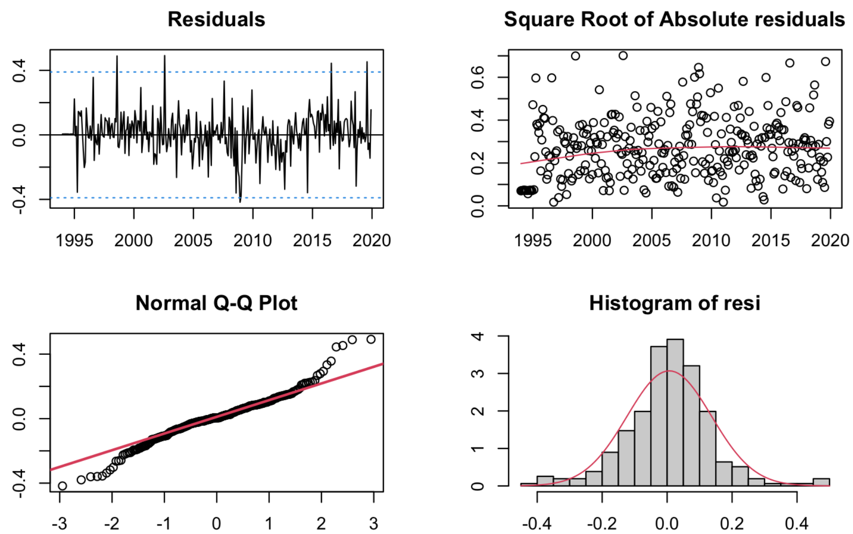

Figure 18

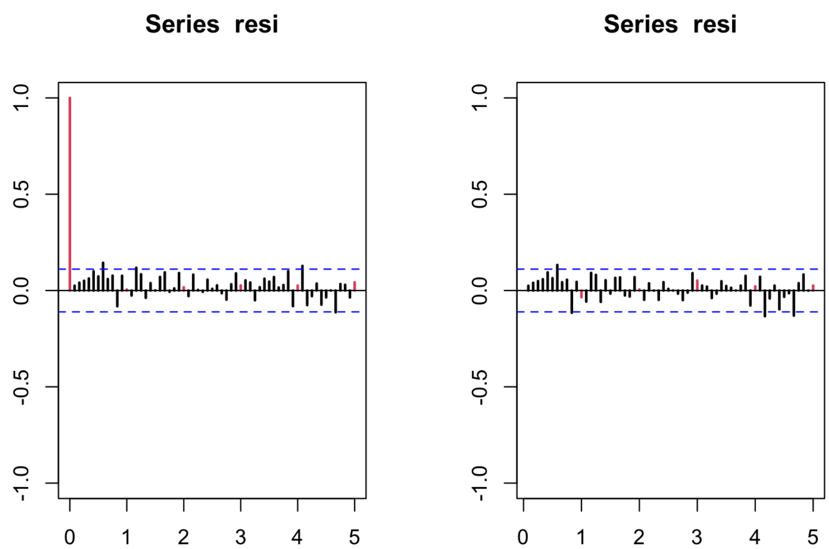

Figure 19

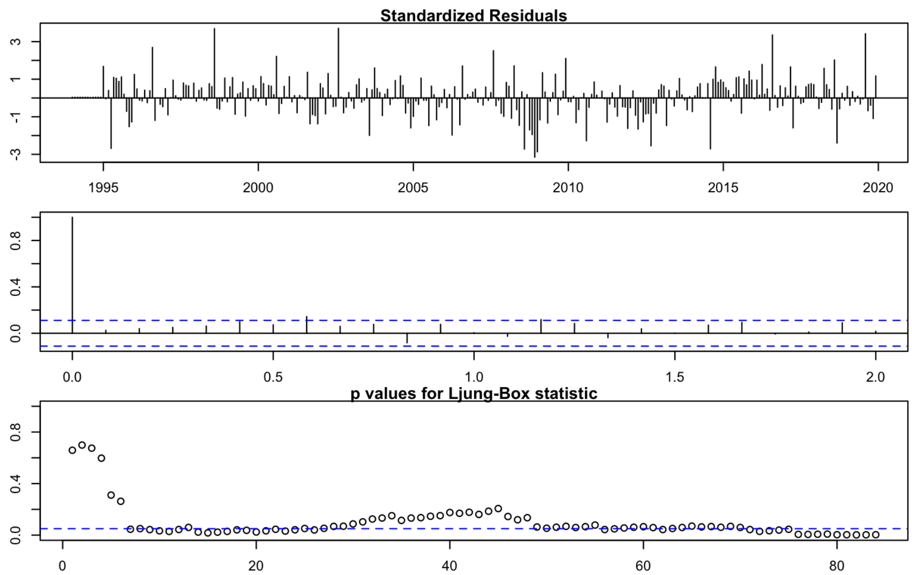

Figure 20

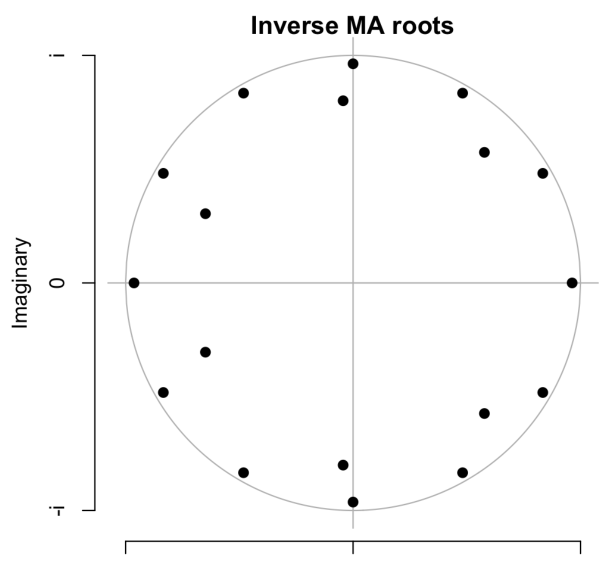

Figure 21

Validation Plots for Model 3:

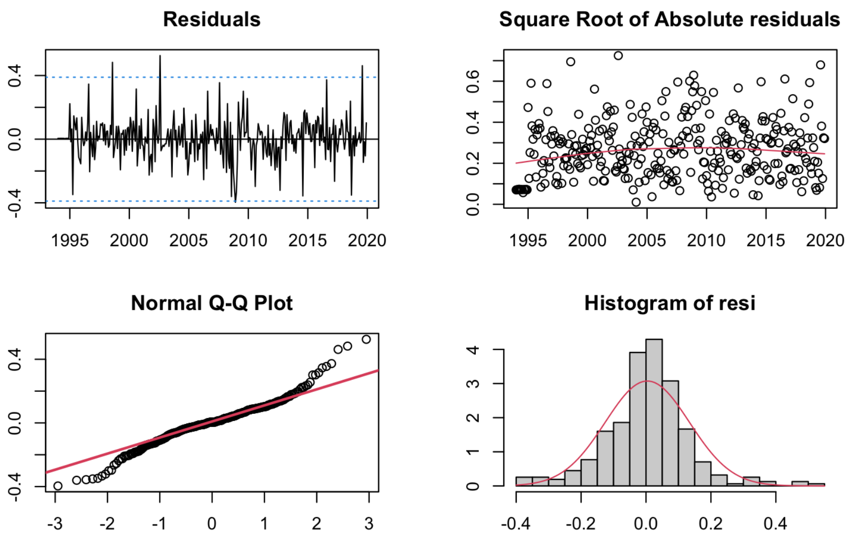

Figure 22

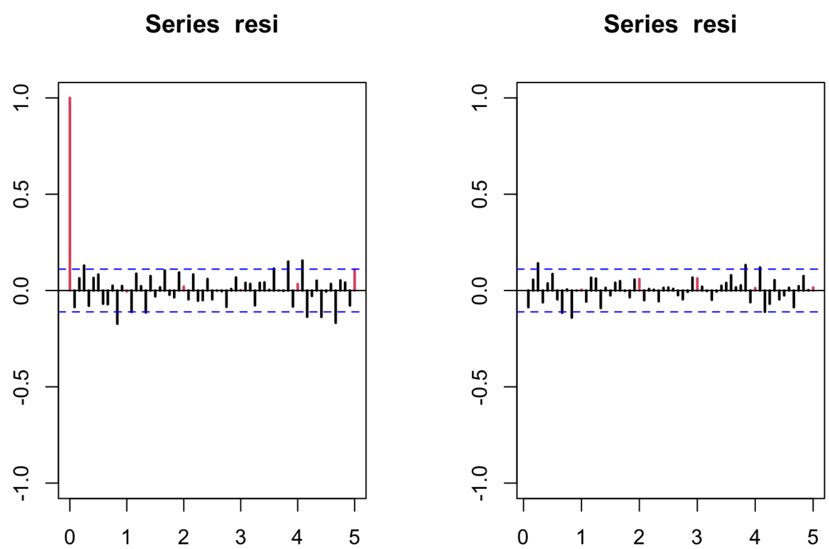

Figure 23

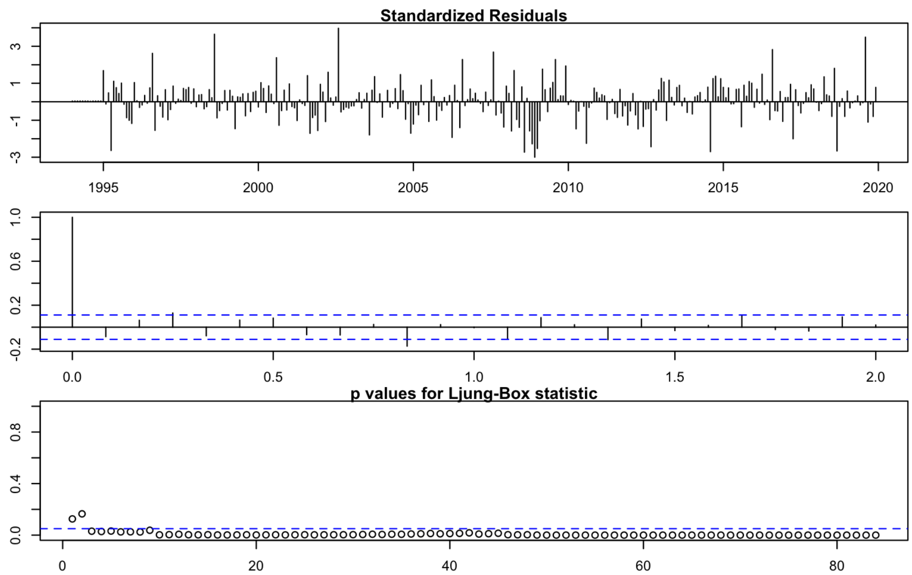

Figure 24

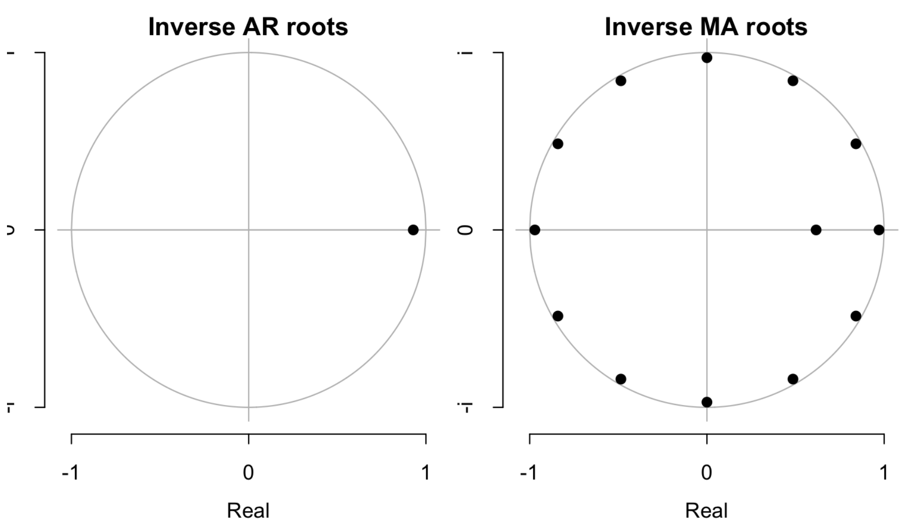

Figure 25

2.4.2 Analysis of AR and MA Infinite Models(Causality and Invertibility)

Model 1 For an AR part to be causal, all roots of the characteristic polynomial must lie outside the unit circle (i.e., have an absolute value greater than 1). In this case, all roots (1.1234, 1.848768, 1.848768) are greater than 1. As we can see in Figure 17, the inverse AR roots lie in the circle. Therefore, the AR part of the model is causal.

For an MA part to be invertible, all roots of the characteristic polynomial must lie outside the unit circle. Here all roots are 1.0318, which is greater than 1. As we can see in Figure 17, the inverse MA roots lie in the circle. Therefore, the MA part of the model (both regular and seasonal) is invertible.

As observed in Table 2 and Table 3, the Psi and Pi weights diminish over time, signifying causality and invertibility. However, it is noteworthy that the Pi weight at lag 24 is larger, which may be related to seasonal effects. This observation aligns with previous analyses and suggests model stability.

Table 2: Psi-weights for Infinite Order Moving Average (MA(∞)) in Model 1

psi 1	psi 2	psi 3	psi 4	psi 5	psi 6
0.23022066	0.34787210	0.40840855	0.25655905	0.27009123	0.24419654
psi 7	psi 8	psi 9	psi 10	psi 11	psi 12
0.20267823	0.18900854	0.16687506	0.14693583	0.13225882	-0.56959985
psi 13	psi 14	psi 15	psi 16	psi 17	psi 18
-0.05386706	-0.14591458	-0.19782062	-0.10259722	-0.11995287	-0.10938812
psi 19	psi 20	psi 21	psi 22	psi 23	psi 24
-0.08727398	-0.08358764	-0.07346672	-0.06429037	-0.05823337	-0.05149723

Table 3: Pi-weights for Infinite Order Autoregressive (AR(∞)) in Model 1

pi 1	pi 2	pi 3	pi 4	pi 5	pi 6
0.2302207	0.2948705	0.2604359	0.0000000	0.0000000	0.0000000
pi 7	pi 8	pi 9	pi 10	pi 11	pi 12
0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	-0.6868359
pi 13	pi 14	pi 15	pi 16	pi 17	pi 18
0.1581238	0.2025277	0.1788767	0.0000000	0.0000000	0.0000000
pi 19	pi 20	pi 21	pi 22	pi 23	pi 24
0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	-0.4717435

Model 2 In Figure 21, the model is invertible (MA roots outside the unit circle). There are no AR terms, so causality is not applicable.

Looking at the Psi-weights in Table 4 and Pi-weights in Table 5, we observe a decay over time, suggesting invertibility. Similar to previous observations, the Pi weight at lag 24 is notably large, which may be attributed to seasonal effects. This aligns with prior analyses, reinforcing the stability of Model 2.

Table 4: Psi-weights for Infinite Order Moving Average (MA(∞)) in Model 2

psi 1	psi 2	psi 3	psi 4	psi 5	psi 6
0.23061319	0.33293338	0.33075479	0.15655891	0.20137021	0.21897816
psi 7	psi 8	psi 9	psi 10	psi 11	psi 12
0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	-0.63722575
psi 13	psi 14	psi 15	psi 16	psi 17	psi 18
-0.14695266	-0.21215372	-0.21076547	-0.09976337	-0.12831829	-0.13953852
psi 19	psi 20	psi 21	psi 22	psi 23	psi 24
0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000

Table 5: Pi-weights for Infinite Order Autoregressive (AR(∞)) in Model 2

pi 1	pi 2	pi 3	pi 4	pi 5	pi 6
0.230613187	0.279750941	0.189461703	-0.056548300	0.022699360	0.079668709
pi 7	pi 8	pi 9	pi 10	pi 11	pi 12
-0.143721084	-0.091446467	0.008932882	0.071260999	0.012325823	-0.640935806
pi 13	pi 14	pi 15	pi 16	pi 17	pi 18
0.168622505	0.177494946	0.096684184	-0.054905787	0.021732242	0.059896448
pi 19	pi 20	pi 21	pi 22	pi 23	pi 24
-0.090691559	-0.055955742	0.009769632	0.044642815	0.002338096	-0.410783349

Model 3 In Figure 25, the model 3 is stationary and invertible because all the Inverse AR roots and Inverse MA roots are within the unit circle.

In Table 6 and 7, the weights decrease over time, indicating causality and invertibility. Like Models 1 and 2, the Pi 24 is significantly larger due to seasonal effects, as previously analyzed. Overall, Model 3 is stable.

Table 6: Psi-weights for Infinite Order Moving Average (MA(∞)) in Model 3

psi 1	psi 2	psi 3	psi 4	psi 5	psi 6
0.31263886	0.29032975	0.26961256	0.25037370	0.23250767	0.21591651
psi 7	psi 8	psi 9	psi 10	psi 11	psi 12
0.20050926	0.18620143	0.17291457	0.16057583	0.14911755	-0.56268581
psi 13	psi 14	psi 15	psi 16	psi 17	psi 18
-0.09061517	-0.08414910	-0.07814444	-0.07256825	-0.06738996	-0.06258119
psi 19	psi 20	psi 21	psi 22	psi 23	psi 24
-0.05811555	-0.05396858	-0.05011752	-0.04654126	-0.04322019	-0.04013611

Table 7: Pi-weights for Infinite Order Autoregressive (AR(∞)) in Model 3

pi 1	pi 2	pi 3	pi 4	pi 5	pi 6
0.312638862	0.192586693	0.118634113	0.073079051	0.045016965	0.027730617
pi 7	pi 8	pi 9	pi 10	pi 11	pi 12
0.017082162	0.010522675	0.006482007	0.003992940	0.002459666	-0.699647548
pi 13	pi 14	pi 15	pi 16	pi 17	pi 18
0.220144058	0.135609552	0.083535985	0.051458475	0.031698610	0.019526461
pi 19	pi 20	pi 21	pi 22	pi 23	pi 24
0.012028372	0.007409521	0.004564293	0.002811621	0.001731969	-0.490562249

2.4.3 Stability and Predictive Capability Evaluation

In this analysis, the stability of the proposed seasonal ARIMA models—modA and modB—is evaluated by reserving the last 12 observations of the time series data. Specifically, modA is fitted using data that includes these last 12 observations, whereas modB is fitted using the dataset excluding these observations. The objective is to compare the coefficients of both models to assess how the inclusion versus exclusion of the last 12 months affects the model parameters. This comparative analysis helps in understanding the sensitivity of the model to recent data variations and the robustness of the model's coefficients over different sample periods. By examining the changes in the coefficients, insights can be gained into the temporal stability and predictive reliability of each model.

Model 1 As observed in Figure 26 and Table 8, the predictions closely approximate the actual numbers. Furthermore, the coefficients of models Mod A and Mod B are not only similar in sign but also very close in value, with only slight differences in the second decimal place. These findings suggest that Model 1 is highly stable.

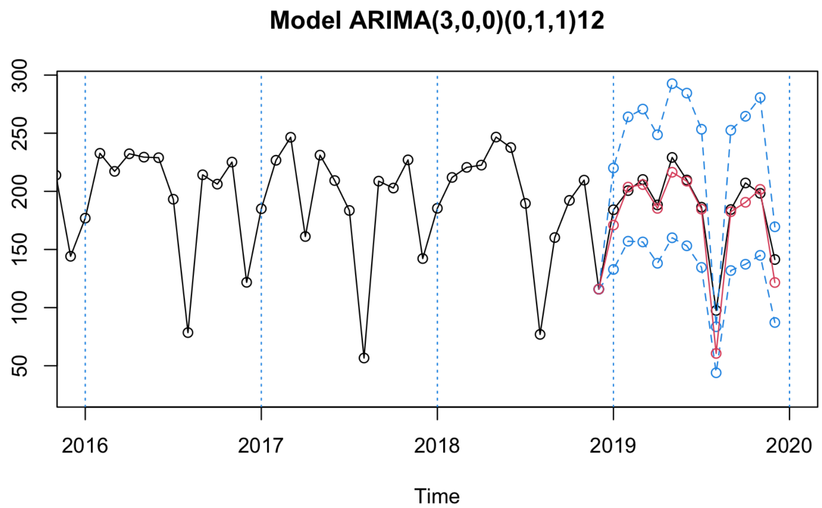

Figure 26

Table 8: Model 1 Coefficients of ARIMA Models ‘modA‘ and ‘modB‘

Model	**ar1**	**ar2**	**ar3**	**sma1**
		modA Coefficients		
Value	0.2302	0.2949	0.2604	-0.6868
s.e.	0.0560	0.0551	0.0562	0.0487
		ModB Coefficients		
Value	0.2400	0.2903	0.2671	-0.7061
s.e.	0.0571	0.0562	0.0571	0.0463

Model 2 In Figure 27, the predictions are somewhat accurate, showing only slight discrepancies. Additionally, in Table 9, the signs of the coefficients are consistent, and the values are very close, with only minor differences in the second decimal place, indicating that Model 2 is highly stable. However, it is important to note that these differences are small, but are larger than those observed in Model 1.

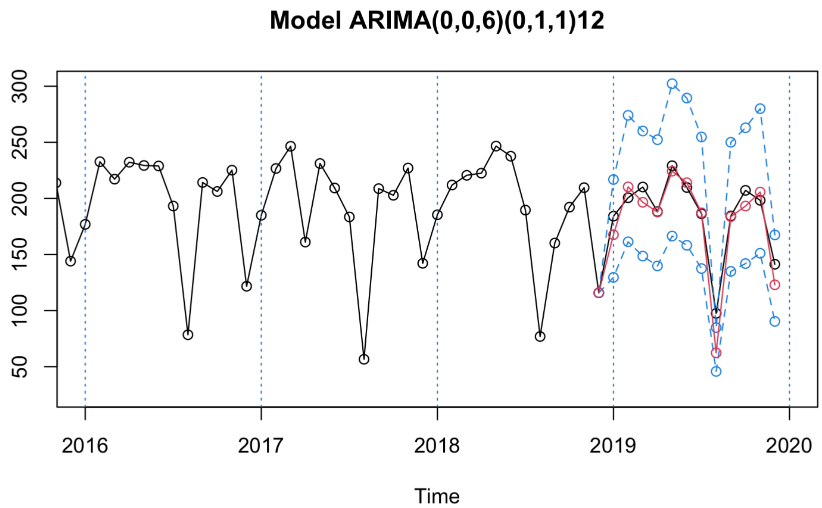

Figure 27

Table 9: Model 2 Coefficients of ARIMA Models ‘modA‘ and ‘modB‘

Model	**ma1**	**ma2**	**ma3**	**ma4**	**ma5**	**ma6**	**sma1**
			modA Coefficients				
Value	0.2306	0.3329	0.3308	0.1566	0.2014	0.2190	-0.6372
s.e.	0.0625	0.0638	0.0648	0.0597	0.0631	0.0678	0.0548
			ModB Coefficients				
Value	0.2441	0.3516	0.3648	0.1617	0.1854	0.2432	-0.6632
s.e.	0.0638	0.0637	0.0647	0.0614	0.0606	0.0698	0.0532

Model 3 In Figure 28 and Table 10, Model 3’s predictions are accurate. Not only are the signs of the coefficients the same, but the numbers are also very close. Compared to Model 2, Model 3 is better because the differences in the coefficients of Model 2 are slightly larger.

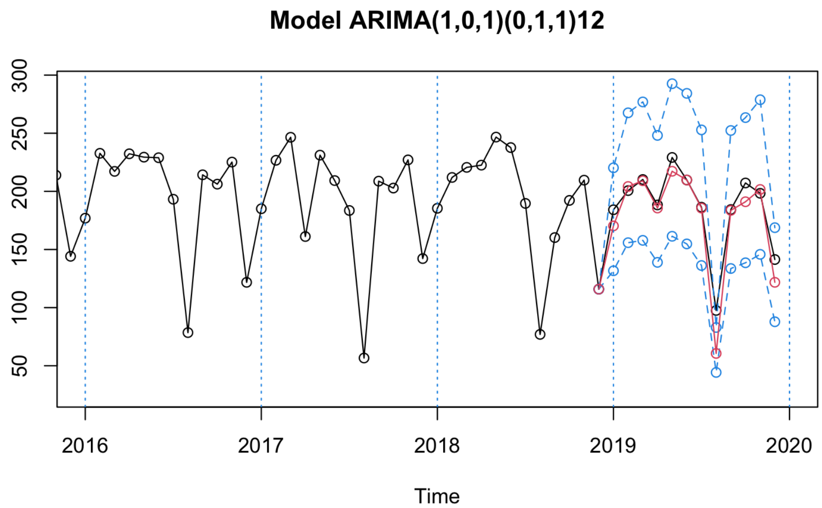

Figure 28

Table 10: Model 3 Coefficients of ARIMA Models ‘modA‘ and ‘modB‘

Model	ar1	ma1	sma1
modA Coefficients			
Value	0.9286	-0.6160	-0.7012
s.e.	0.0289	0.0542	0.0491
ModB Coefficients			
Value	0.9296	-0.6071	-0.7206
s.e.	0.0287	0.0549	0.0465

2.4.4 Selection of Optimal Forecasting Model

Based on the data presented in Table 11, Model 1 appears to be the better choice compared to Model 2 due to a lower AIC and BIC, which are -354.8938 and -336.3749, respectively. Lower values in both of these criteria generally indicate a better fit of the model to the data, balancing model complexity and goodness of fit. Additionally, the slightly lower Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) in Model 1 further support its superiority in terms of prediction accuracy, even though the differences are quite marginal.

Table 11: Comparison of ARIMA Models

Model	**par**	**Sigma2**	**AIC**	**BIC**	**RMSE**	**MAE**	**RMSPE**
ARIMA(3,0,0)(0,1,1)$_{[12]}$	4	0.01688502	-354.8938	-336.3749	14.22449	9.805151	0.1222426
ARIMA(0,0,6)(0,1,1)$_{[12]}$	7	0.01755732	-338.4078	-308.7776	14.19040	10.47256	0.1183757
ARIMA(1,0,1)(0,1,1)$_{[12]}$	3	0.01746909	-346.6285	-331.8134	14.22449	9.216918	0.1218861

Overall, Model 1 is causal, invertible, and shows no signs of residual autocorrelation. Despite some deviations from normality and homoscedasticity according to formal tests, these are likely minor and the model can be considered adequate for forecasting purposes, pending out-of-sample validation. Model 2 is invertible and generally shows good residual behavior. The standardized residuals and ACF plot confirm the randomness and lack of significant autocorrelation in the residuals. The Ljung-Box test, considered in its entirety through the p-value plot, does not indicate significant autocorrelation. The formal normality and homoscedasticity tests are still rejected, but as previously discussed, this is likely due to the tests' sensitivity. Model 3 shows random residuals centered around zero, suggesting consistent variance and approximate normality, though minor deviations exist. Tests for normality and homoscedasticity are rejected, likely due to their sensitivity. While the Durbin-Watson test shows no significant autocorrelation, the Ljung-Box test reveals significant seasonal autocorrelation at lags 12 and 24.

Based on the analysis above and the evaluation of prediction accuracy and all adequacy measures, Model 1 is slightly better than Model 2 and Model 3. Therefore, Model 1 is identified as the optimal forecasting model.

2.5 Prediction

As illustrated in Figure 29 and Table 12, the predictions for 2020 are presented along with their confidence intervals.

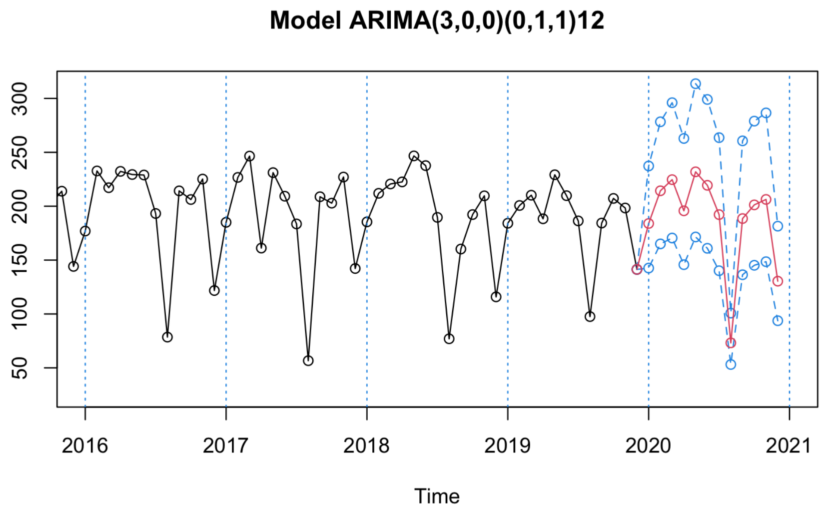

Figure 29: Final Prediction

Table 12: Prediction Monthly Data in 2020

Month	tl1	pr1	tu1
Dec 2019	141.36400	141.36400	141.3640
Jan 2020	142.59293	183.95313	237.3102
Feb 2020	165.01772	214.30542	278.3144
Mar 2020	170.47125	224.64588	296.0369
Apr 2020	145.73734	195.72627	262.8618
May 2020	171.43637	231.89284	313.6691
Jun 2020	160.93564	219.37912	299.0462
Jul 2020	140.12612	192.19701	263.6175
Aug 2020	53.12838	73.17676	100.7906
Sep 2020	136.39483	188.54170	260.6255
Oct 2020	145.17373	201.23515	278.9457
Nov 2020	148.50243	206.28974	286.5640
Dec 2020	93.72635	130.42278	181.4869

A Model 1 Descriptive Statistics for Residuals

A.1 Normality Tests

- **Shapiro-Wilk normality test:**
  Data: `resi`
  $W = 0.9614$, p-value $= 2.365 \times 10^{-7}$

- **Anderson-Darling normality test:**
  Data: `resi`
  $A = 3.239$, p-value $= 3.951 \times 10^{-8}$

- **Jarque Bera Test:**
  Data: `resi`
  $X^2 = 69.227$, df $= 2$, p-value $= 8.882 \times 10^{-16}$

A.2 Homoscedasticity Test

- **Studentized Breusch-Pagan test:**
  Data: `resi` $\sim$ `I(obs - resi)`
  $BP = 63.288$, df $= 1$, p-value $= 1.786 \times 10^{-15}$

A.3 Independence Tests

- **Durbin-Watson test:**
  Data: `resi` $\sim$ `I(1:length(resi))`
  $DW = 1.9993$, p-value $= 0.4748$
  Alternative hypothesis: true autocorrelation is greater than 0.

- **Ljung-Box test:**

Lag	Statistic	p-value
1	2.822741×10^{-4}	0.9865954
2	2.457423×10^{-2}	0.9877881
3	0.3006077	0.9599141
4	0.7550554	0.9443699
12	12.17097	0.4320488
24	26.99509	0.3046819
36	35.71434	0.4820668
48	48.70003	0.4446744

B Model 2 Descriptive Statistics for Residuals

B.1 Normality Tests

- **Shapiro-Wilk normality test:**
  Data: `resi`
  $W = 0.96208$, p-value $= 2.942 \times 10^{-7}$

- **Anderson-Darling normality test:**
  Data: `resi`
  $A = 3.0447$, p-value $= 1.174 \times 10^{-7}$

- **Jarque Bera Test:**
  Data: `resi`
  $X^2 = 61.108$, df $= 2$, p-value $= 5.373 \times 10^{-14}$

B.2 Homoscedasticity Test

- **Studentized Breusch-Pagan test:**
  Data: `resi` $\sim$ `I(obs - resi)`
  $BP = 59.127$, df $= 1$, p-value $= 1.478 \times 10^{-14}$

B.3 Independence Tests

- **Durbin-Watson test:**
  Data: `resi` $\sim$ `I(1:length(resi))`
  $DW = 1.9466$, p-value $= 0.2979$
  Alternative hypothesis: true autocorrelation is greater than 0.

- **Ljung-Box test:**

Lag	Statistic	p-value
1	0.1951259	0.65868417
2	0.7163975	0.69893416
3	1.5321608	0.67486778
4	2.7683592	0.59730673
12	21.4989198	0.04353483
24	37.2400059	0.04140534
36	45.5395482	0.13242378
48	58.8672871	0.13524952

C Model 3 Descriptive Statistics for Residuals

C.1 Normality Tests

- **Shapiro-Wilk normality test:**
  Data: `resi`
  $W = 0.96045$, p-value $= 1.749 \times 10^{-7}$

- **Anderson-Darling normality test:**
  Data: `resi`
  $A = 3.3745$, p-value $= 1.85 \times 10^{-8}$

- **Jarque Bera Test:**
  Data: `resi`
  $X^2 = 64.233$, df $= 2$, p-value $= 1.132 \times 10^{-14}$

C.2 Homoscedasticity Test

- **Studentized Breusch-Pagan test:**
  Data: `resi` $\sim$ `I(obs - resi)`
  $BP = 63.125$, df $= 1$, p-value $= 1.94 \times 10^{-15}$

C.3 Independence Tests

- **Durbin-Watson test:**
  Data: `resi` $\sim$ `I(1:length(resi))`
  $DW = 2.1708$, p-value $= 0.9274$
  Alternative hypothesis: true autocorrelation is greater than 0.

- **Ljung-Box test:**

Lag	Statistic	p-value
1	2.342953	0.125850766
2	3.601771	0.165152581
3	8.910397	0.030506241
4	10.888066	0.027851224
12	27.619745	0.006285618
24	48.146257	0.002420947
36	59.152196	0.008850240
48	79.423263	0.002910022