ARIMA modelling for industrial production in Spain

An Statistical and Economical Analysis

Iker Caballero

Time Series - Prof. Josep A. Sanchez

1 Introduction

Industrial-level data is very important for different economic agents, such as managers, investors, policy-makers and economists, as this kind of data allows to study specific lines of businesses, information about fluctuations and also gives an insight for the economic growth and the development of the industry in a given country. Because processing an using this data is quite difficult and not many agents have access, relevant economic institutions such the Federal Reserve Board, the European Comission and different national institutions publish an index which tries to aggregate this data, called the Industrial Production Index (IPI).

Of course, this index is presented as a time series, so it is interesting to apply time series analysis methodologies in order to obtain useful insights from both the statistical but also the economic point of view, as the analysis can also allow to obtain forecasts for future values, which is relevant for investors, managers and other economic agents. In order to exemplify the methodology and get interesting results, we will consider the IPI from the Spanish economy, an European country which is known for its international trade and tourism, but also because of its resilience to economic downturns. Due to these characteristics, its interesting to study the industrial sector, which might not be as known or studied from the international perspective.

Hence, the main goal of this project is to use IPI data from the Spanish Economy to apply the ARIMA methodology for analyzing time series data, proposed by Box and Jenkins. Moreover, the analysis will be extended by including the treatment of calendar effects and outliers and using the estimated models for prediction. This report will first start by presenting relevant aspects about the data and a first exploratory analysis. Then, it will proceed with the ARIMA modelling, the predictions and the outlier and calendar effects treatment. Finally, we will summarize the most important takes from the project.

2 About the Data

First of all, one must define clearly what the index we are working with is. The Industrial Production Index (IPI) is a short-term indicator measuring monthly evolution of productive activity in the different industrial branches of an economy. It measures the joint evolution of quantity and quality without taking into account the prices.

The IPI for the Spanish economy is extracted from the Instituto Nacional de Estadística (INE), and the index measures the productive activity of industrial branches contained in the National Classification of Economic Activities (excluding construction industry, which is quite relevant in this economy). The data that we are using is the IPI from 1990 to 2020, which uses 2016 as base year.

The time series for Spain's IPI is plotted in Figure 1. As one can see just from the graph, there seems to some positive trend from 1992 to 2008, when the Great Recession happened. Moreover, there is a clear seasonal pattern in the data, probably caused because of the summer holidays of July-August. The mean of this series is plotted as a red line, and one can see that it is above the base level, hence indicating that, on average, the industrial production has increased during the 1990-2020 period with respect to 2020.

Now that we have pointed out some aspects about the data, one can proceed to the exploratory analysis, where we delve into the different characteristics of this data and prepare the series for the ARIMA methodology.

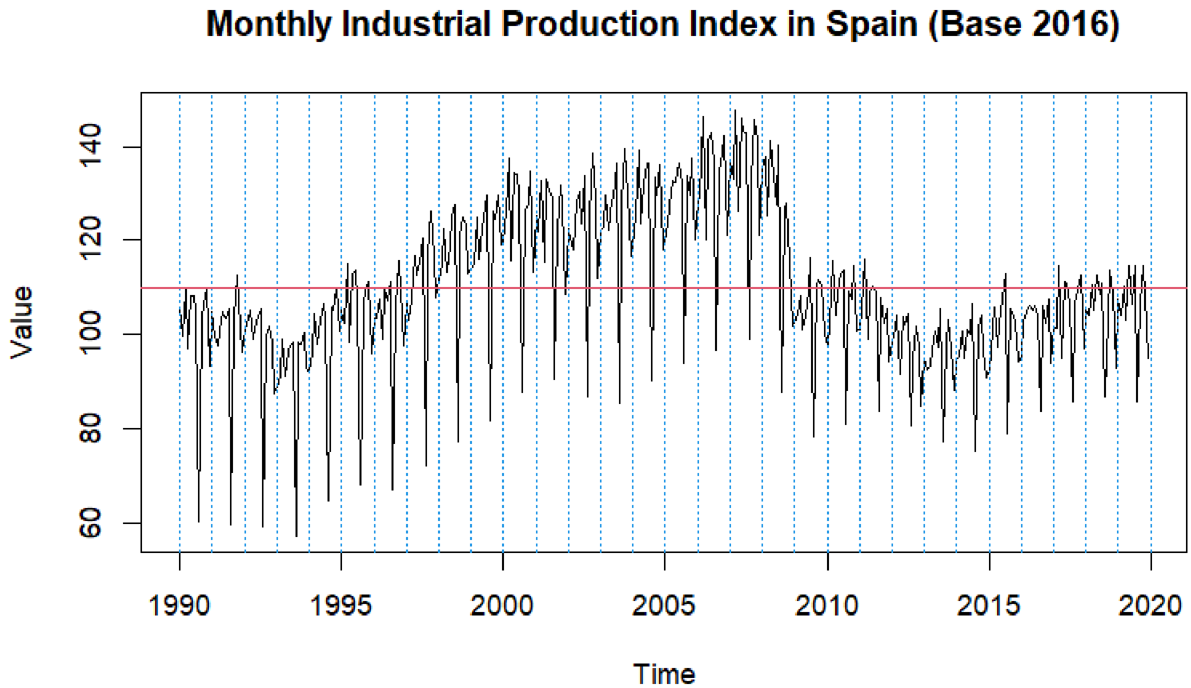


Figure 1: Spain's IPI Time Series

3 ARIMA modelling

Now that we have carried on an insightful analysis of the characteristics of the time series and we have transformed the time series, we can now apply the ARIMA methodology to model the data. This methodology consists in four primary steps in which we will divide the analysis: the first step is to transform the time series, the second is to identify the order of the model, the third step is to estimate the possible models and the final one is to validate them.

3.1 Transformations

Before delving into modelling, we first need to work with a well-suited time series, so we might need to apply some transformations. To do so, we look at four key aspects: the variability of the data, the seasonality pattern, the mean of the time series and the stationarity. These aspects will allow to obtain a stationary time series, which is the basic requirement for this methodology.

3.1.1 Variability

From the plot of Figure 1 one cannot easily observe a clear pattern. However, by looking carefully, one can see that the variability of the series seems to be greater before 2008 compared to the range from 2008 to 2020. The declines in the industrial production index seem to be sharper than those after 2008, but when looking at the series ignoring the declines, there seems to be no different pattern of variability.

In order to do a more accurate study of the variability pattern, we make use of a boxplot, which allows to study the variability of different groupings of observations graphically. The boxplot is represented at the left part in Figure 2, and we can see that for values from 90 to 120 of the index, the variability seems be constant, in the sense that boxes at the same level seem similar. However, when looking at high values of the index, variability is clearly different, being high compared to other values at the start or at the end of the time series.

Moreover, one can also see that variability is higher for higher values when considering a mean-variance plot, which is represented at the right part of Figure 2. In here, the line that represents the relation between the different mean values for the series and the variance has a clear positive slope, indicating exactly what we are pointing out.

Because a non-constant variance across the time series difficult for the modelling in the ARIMA framework, we decide to apply a logarithmic transformation of the data, which corresponds to the recommended Box-Cox transformation in this case. This will yield a constant variance time series, which is represented in Figure 3.

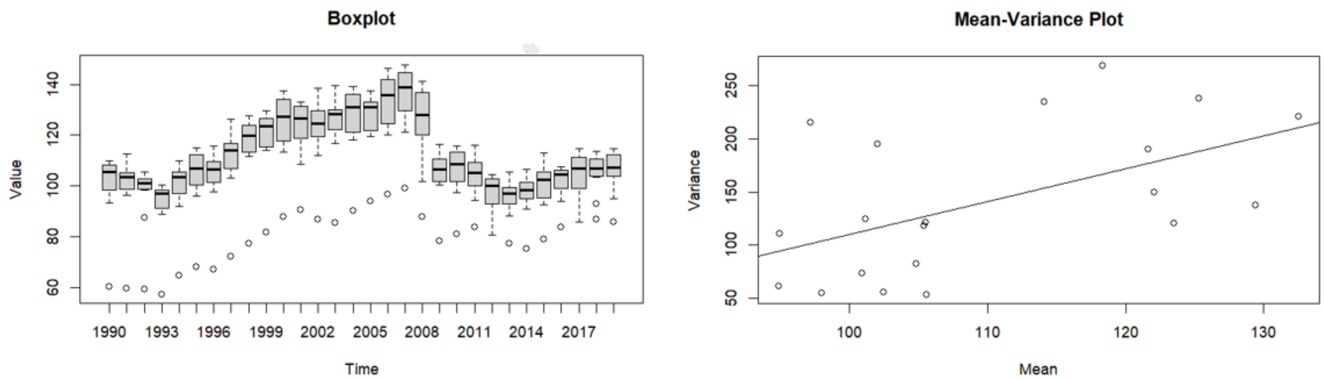


Figure 2: Spain's IPI Boxplot and Mean-Variance Plot

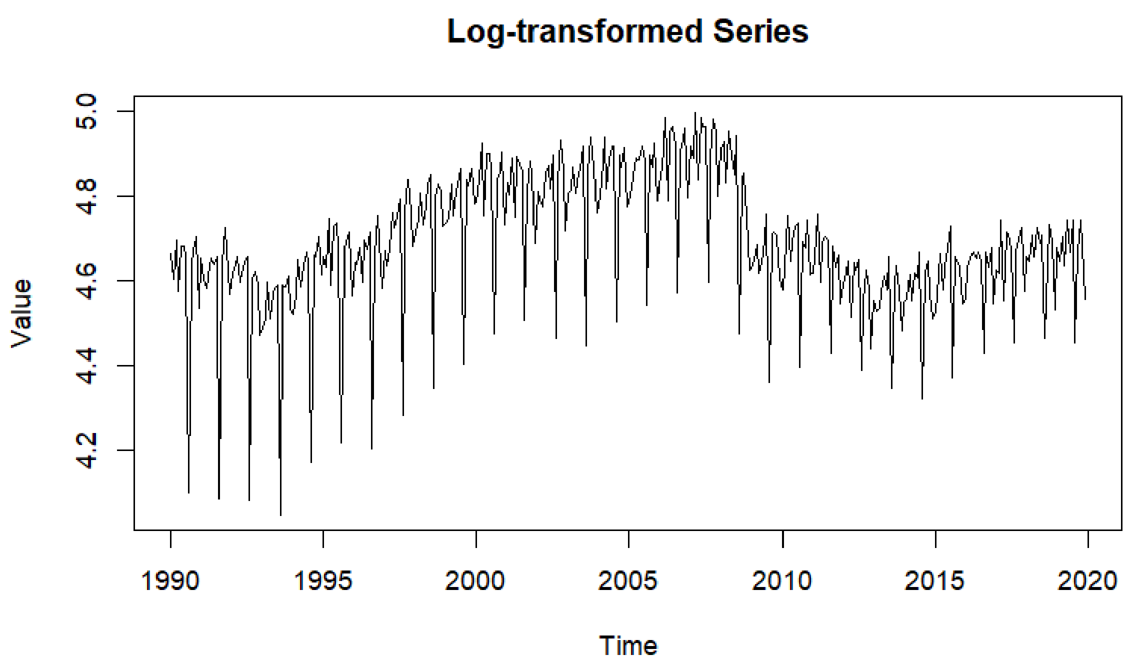


Figure 3: Spain's IPI Log-transformed series

3.1.2 Seasonality

Even though the seasonality pattern seems very clear, to assess it more accurately, one can use a monthplot, which allows to divide the time series for the different months and shows the level of the observations corresponding to each one. The monthplot is represented in Figure 4, and, as expected, one can see that the levels seem merely the same for each month except for August, where there is a huge decline in the index level.

Due to this decline, we can use seasonal differences in order to eliminate this seasonal pattern. In this case, we apply a seasonal difference of order 12 (for monthly patterns) to the log-transformed series. The results of applying this transformation is given in Figure 5.

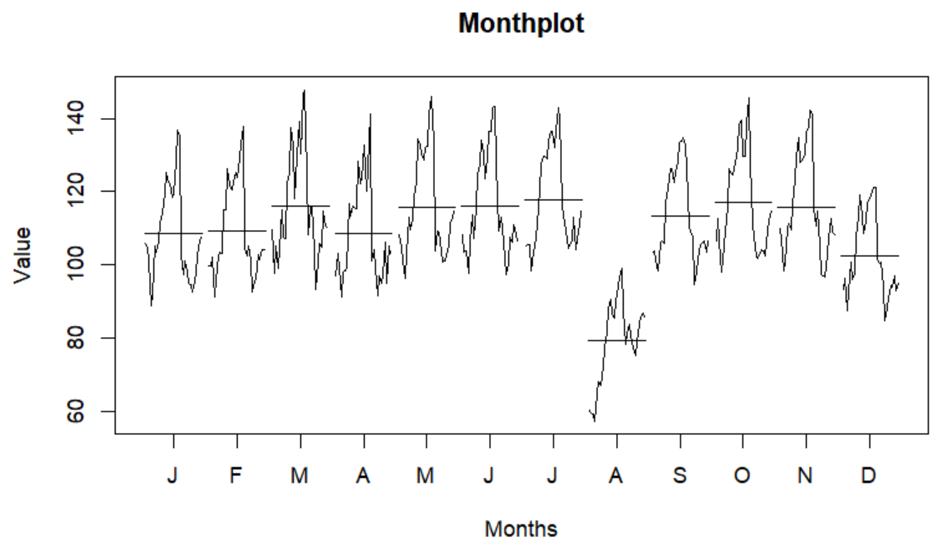


Figure 4: Spain's IPI Monthplot

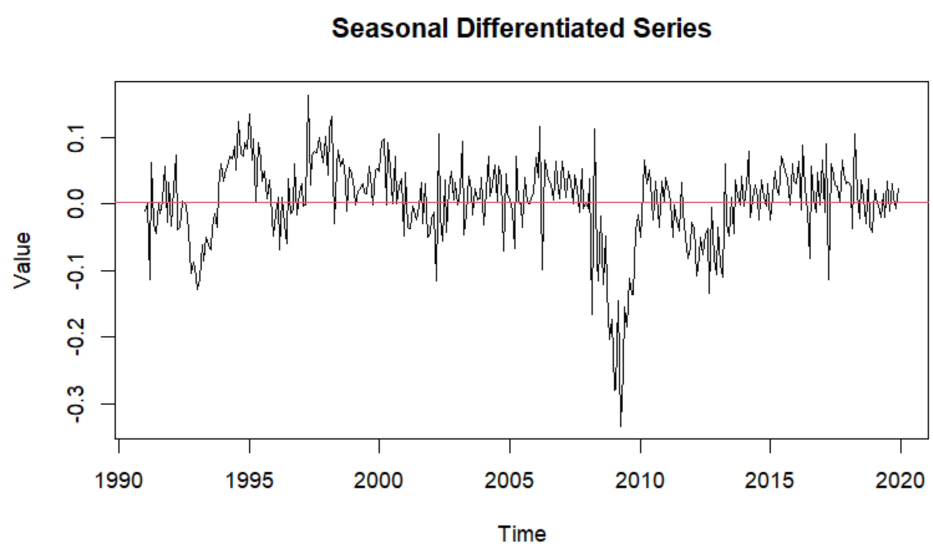


Figure 5: Spain's IPI Seasonal Differentiated series

3.1.3 Mean

Another aspect to check is the mean of the index. If one looks at the original time series, one can clearly see that the mean is seems to be more or less constant but at a positive level. However, when applying the logarithmic transformation and the seasonal difference, the resulting time series seems to have a constant mean around zero, which is depicted in Figure 5 with a red line. Therefore, there is no need to apply no change for the mean.

3.1.4 Stationarity

The fundamental characteristic to work within the ARIMA framework is the stationarity of the time series. This is not easily evaluated just by looking at the plotted time series, even though it is common for economic time series not to be stationary. One useful tool for assessing stationarity is the autocorrelation function (ACF) plot.

In Figure 6, the left panel represents the ACF plot for the original time series, while the right panel is the plot of the corresponding ACF for the transformed time series. By looking at the first, one can clearly see that the autocorrelation is persistent and significant between lags, but the right panel shows a decay of the autocorrelation at the first 12 lags, and then stabilizes more or less inside the significance bands. Consequently, even though one can conclude that the original time series is not stationary, the transformed one seems to be stationary, so we can work with it in the following sections. We highlight the clear fact that, looking at Figure 5, we can see a sharp negative value in 2008 that can be considered an outlier. This might hinder the fit of the following proposed models, but we deal with these kind of values afterwards. For the next section, we assume that there is stationarity.

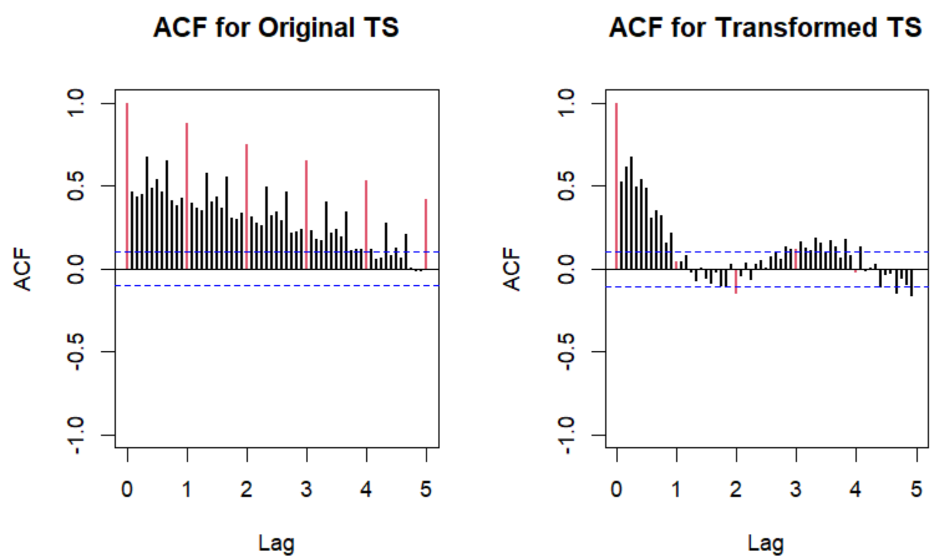


Figure 6: ACF for the Original and the Transformed Time Series

3.2 Identification

In order to identify the order of the model, we need to look at the autocorrelation (ACF) and the partial autocorrelation (PACF) functions. One normally looks at the significant lags in both functions in order to choose the order, but because of the presence of seasonal patterns, we also need to consider what we call the "seasonal lags". The presence of a seasonal pattern in the series indicates that the model that we should use is a Seasonal ARIMA or SARIMA, so we keep this in mind while identifying the model.

Both the ACF and the PACF of the transformed time series are shown in Figure 7. Regarding the seasonal lags, it is difficult to see a clear pattern, but we can claim that there are two possible scenarios: the ACF seasonal lags exhibits a decay while the PACF seasonal lags are cut (which would indicate a SAR model), and the other way around (which would indicate a SMA model). We also consider that the ACF has a significant second lag, while the PACF shows that just the first seasonal lag seems to be significant. Then, the possible patterns and the lags would suggest that the seasonal part could be modeled as an SAR(1) if the decay is in the ACF, an SMA(2) if the decay is in the PACF, or even a SARMA(1,1).

When it comes to the regular lags, the patterns are clearer; one can easily see that the ACF presents a decay while the PACF lags seem to be cut, which indicates an AR model for the regular part. In this case, we would clearly propose an AR(3), because of the cut produced in the third lag in the PACF. Consequently, we decide to consider three possible models for estimation and validation: a SARIMA(3,0,0)(1,1,0), a SARIMA(3,0,0)(0,1,2) and a SARIMA(3,0,0)(1,1,1) for the logarithmic transformation of the original series (because the seasonal difference is applied in the model). These models would have no mean, given that the mean of the transformed time series is zero (or at least it is approximately zero, so one can assume zero mean).

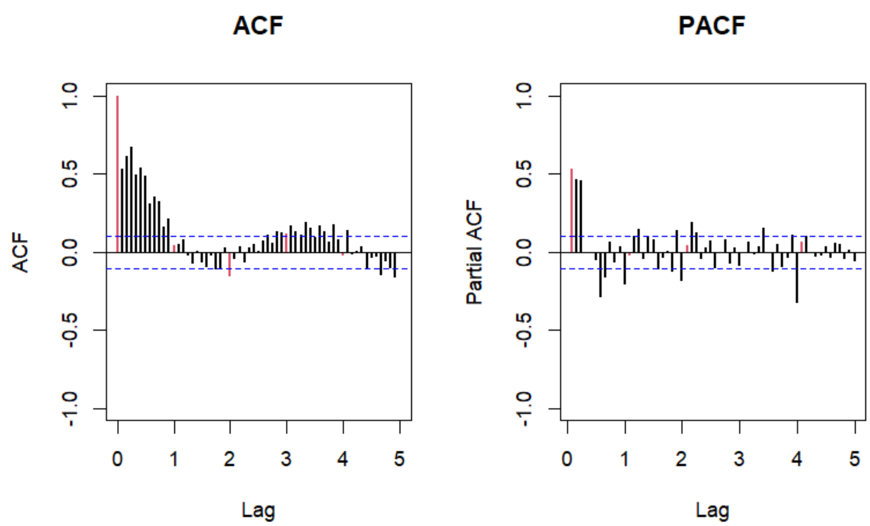


Figure 7: ACF & PACF for the Transformed Time Series

3.3 Estimation

In order to estimate the models, we use a maximum likelihood estimation procedure implemented in R, from the "stats" package. The procedure allows to obtain not just the estimated coefficients with their standard error, but also an AIC estimation for the models, which allows to compare different models based on a information criterion. We estimate the models using the time series when the logarithm and the seasonal differences are applied, so that we do not apply a seasonal difference inside the model (it is already applied) but the results of doing one way or another are equivalent.

```
Call:
arima(x = d12lserie, order = c(3, 0, 0), seasonal = list(order = c(1, 0, 0)),
    include.mean = F)

Coefficients:
         ar1     ar2     ar3     sar1
      0.0519  0.3242  0.5177  -0.3033
s.e.  0.0457  0.0423  0.0466   0.0522

sigma^2 estimated as 0.001751:  log likelihood = 609.37,  aic = -1208.74

Call:
arima(x = d12lserie, order = c(3, 0, 0), seasonal = list(order = c(0, 0, 2)),
    include.mean = F)

Coefficients:
         ar1     ar2     ar3     sma1     sma2
      0.1041  0.3439  0.4940  -0.4899  -0.1725
s.e.  0.0473  0.0433  0.0484   0.0651   0.0620

sigma^2 estimated as 0.001449:  log likelihood = 640.06,  aic = -1268.12

Call:
arima(x = d12lserie, order = c(3, 0, 0), seasonal = list(order = c(1, 0, 1)),
    include.mean = F)

Coefficients:
         ar1     ar2     ar3    sar1     sma1
      0.0946  0.3382  0.5113  0.1473  -0.7073
s.e.  0.0465  0.0426  0.0468  0.0713   0.0439

sigma^2 estimated as 0.001464:  log likelihood = 638.29,  aic = -1264.57
```

Figure 8: ACF & PACF for the Transformed Time Series

Figure 8 shows the estimation results for the three models, ordered in the same order presented above when identifying the different models. One interesting fact to notice here when looking at the SARIMA(3,0,0)(1,1,0) model is that all coefficients seem to be significant but the first one, which corresponds to the first coefficient of the AR part. This is shown in the following Figure 9, where the values of the t-statistic computed for each coefficient are shown.

```
     ar1       ar2       ar3      sar1
1.135828  7.658163 11.099453  5.808377
     ar1       ar2       ar3      sma1      sma2
2.200041  7.945157 10.204398  7.527822  2.783434
     ar1       ar2       ar3      sar1      sma1
2.036145  7.931923 10.915787  2.067144 16.120203
```

Figure 9: T-statistic values for the coefficients of each model (in the same order as before)

However, for the other two models, all the coefficients are significant. Additionally, the AIC measure indicates that these other two models have a higher fit than the first one, which might say that our model selection should consider just these two. We proceed to the validation of each of these models in order to discuss their characteristics and their fit to the data.

3.4 Validation

For the validation of the models, we create a custom function which we will call "validation", which allows to obtain different graphs and computes some hypothesis tests to assess if the models are adequate given the assumptions that the ARIMA methodology requires when modelling. Then, we will also look at the AR and MA infinite forms for the models and report some adequacy measures.

3.4.1 Residual Check

For the SARIMA(3,0,0)(1,1,0), the results obtained from the validation function are shown in Figure 10. In it, one can assess different assumptions through the residuals. The Q-Q plot and the histogram show that the residuals fit normality, even though there is some differences regarding the negative extremes of the distribution and the p-value of the Shapiro-Wilk test. However, this can be due to the presence of outliers, so in subsequent sections we will see the effect of treating them on normality.

Moreover, we can see that there are significant residual lags in both the ACF and the PACF typical from a noise process, as we expect 5% of the lags to be outside the confidence bands, and hence we can see that. The squared residuals ACF and PACF indicates whether there is the presence of volatility in the time series, and even though there are some significant lags, we consider there is no volatility present as there is no significant amount of lags above the confidence bands. Finally, the Ljung-Box plot shows how there is significant dependency for most lags, so that the observations are not independent (as it would be in a noise process).

We now turn to analyze the residuals for the SARIMA(3,0,0)(0,1,2). In this case, normality is also maintained because of the same arguments. The ACF and the PACF for the residuals seem to behave as a noise process, and there is a lack of volatility.The Ljung-Box statistics also show again that there might not be independence among the lags, so this case is very similar to the last one.

Finally, we must check the residuals of the last model proposed, the SARIMA(3,0,0)(1,1,1). This case can be evaluated exactly as the previous cases, given that the graphical and numerical results coincide with the analysis made earlier.

All in all, we can see that the assumptions of the residuals being a Gaussian white noise process seems to be adequate, but the non-independence of the lags contradicts the assumed independence of the observations of the process. In view of this, we would have to reject the validity of the three models, but we notice that the non-independence might be due to the presence of outliers and calendar effects in our time series. Therefore, we would need to treat them in order to obtained unbiased results and assess the true validity of the models.

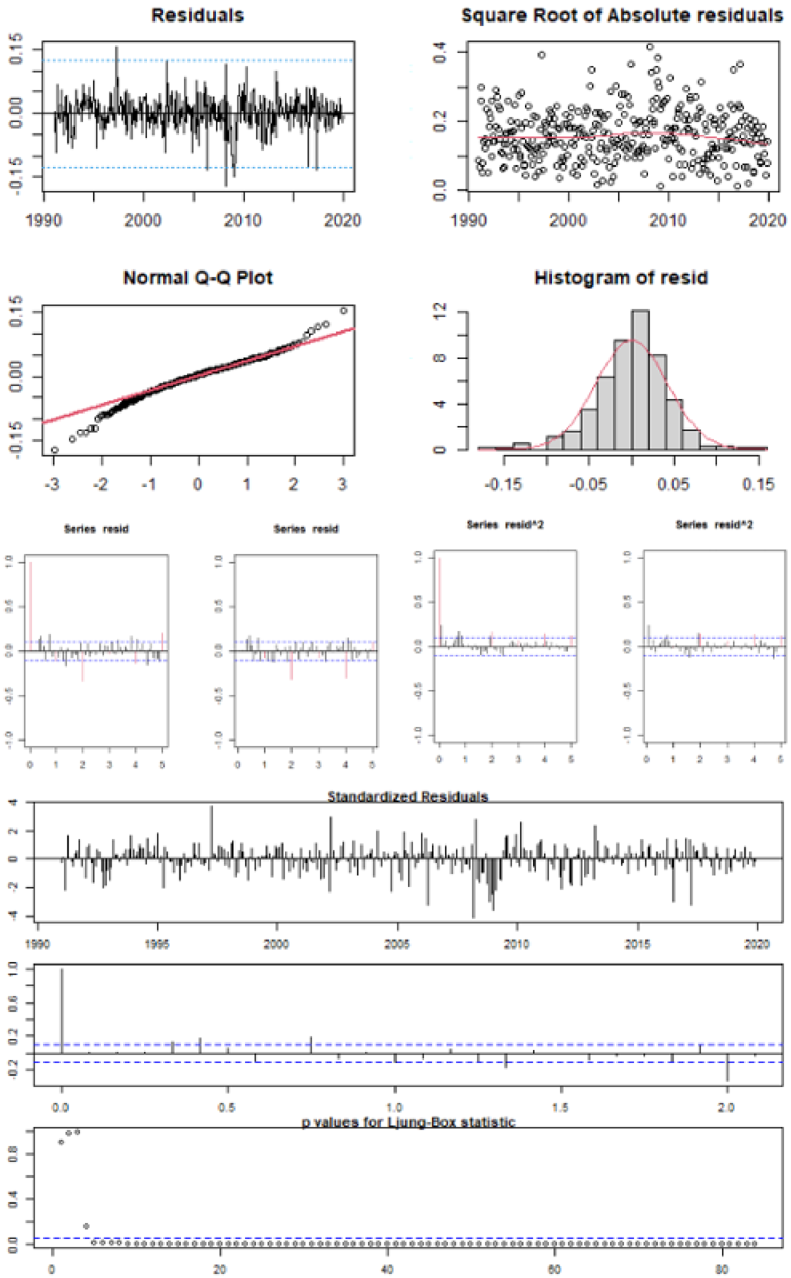


Figure 10: Validation Graphics and Results for Model 1

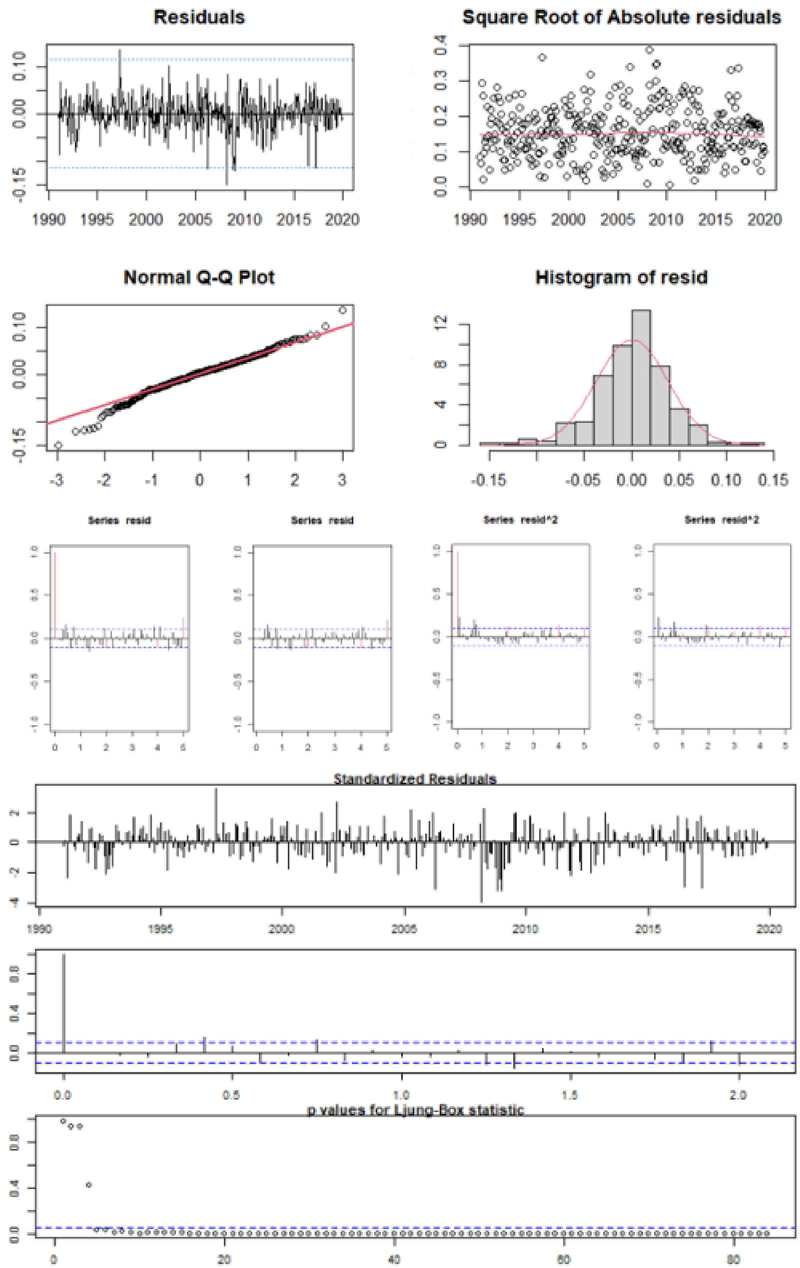


Figure 11: Validation Graphics and Results for Model 2

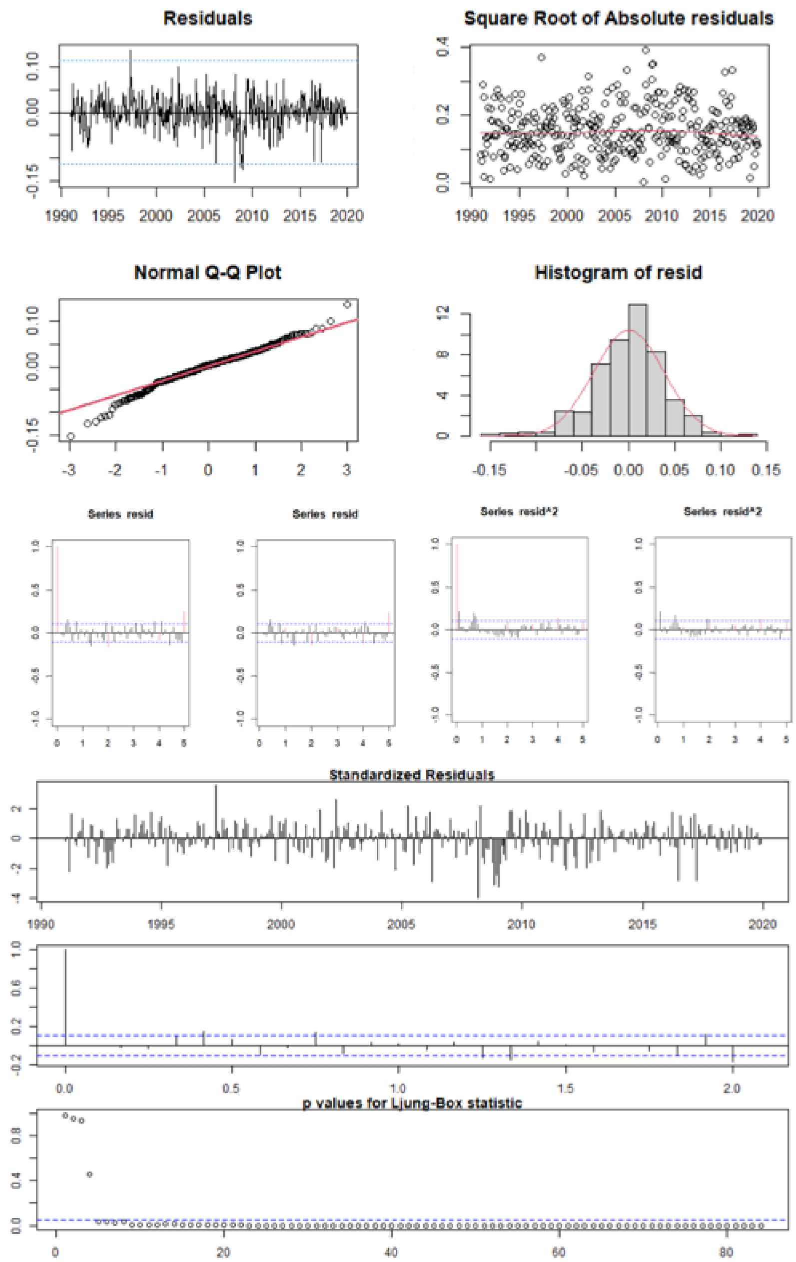


Figure 12: Validation Graphics and Results for Model 3

3.4.2 AR and MA infinite forms

We now proceed to check the AR and MA infinite forms for the three models and report adequacy measures. For the first model, the SARIMA(3,0,0)(1,1,0) would have the form

$$(1-\Phi_1 B^{12})(1-\phi_1 B-\phi_2 B^2-\phi_3 B^3)x_t = w_t \tag{1}$$

where X_T denotes the time series observations the different ϕs denote the coefficients of the AR part, the different θs denote the coefficients of the MA part, Φ and Θ are the parameters for the corresponding seasonal parts, B is the lag operator, and the w_t is an identical and independently distributed Gaussian white noise process. From this form, we can get the infinite AR and MA representations by dividing the sets coefficient terms by each other. To obtain the infinite MA form of the model, we just need to divide the coefficients multiplying w_t (in this case, 1) by the ones multiplying x_t, which yields

$$x_t = \frac{1}{(1-\Phi_1 B^{12})(1-\phi_1 B-\phi_2 B^2-\phi_3 B^3)} w_t \implies x_t = (1+\psi_1 B+\psi_2 B^2+...)w_t \tag{2}$$

The AR form would be obtained interchanging the coefficients that act as denominator. However, because this is a Seasonal AR model, the coefficients coincide with the π coefficients coincide with the ones in this model. The estimated coefficients for both forms of this model are shown in Figure 13, and because no parameters have an absolute value greater or equal to one, we can say that the model is both invertible and causal.

```
Psi-weights (MA(inf))

--------------------
      psi 1       psi 2       psi 3       psi 4       psi 5       psi 6       psi 7
psi 8
 0.05189804  0.32691853  0.55153520  0.16148859  0.35646200  0.35641134  0.21768043
0.31141009
      psi 9      psi 10      psi 11      psi 12      psi 13      psi 14      psi 15
psi 16
 0.27126820  0.22774759  0.26100174 -0.07549171  0.19861997  0.12096327  0.03159010
0.14369272
     psi 17      psi 18      psi 19      psi 20
 0.08032746  0.06711315  0.10392300  0.06874208

Pi-weights (AR(inf))

--------------------
       pi 1        pi 2        pi 3        pi 4        pi 5        pi 6        pi 7
pi 8
 0.05189804  0.32422512  0.51774213  0.00000000  0.00000000  0.00000000  0.00000000
0.00000000
       pi 9       pi 10       pi 11       pi 12       pi 13       pi 14       pi 15
pi 16
 0.00000000  0.00000000  0.00000000 -0.30332565  0.01574201  0.09834580  0.15704447
0.00000000
      pi 17       pi 18       pi 19       pi 20
 0.00000000  0.00000000  0.00000000  0.00000000
```

Figure 13: Estimated values for the $AR(\infty)$ and $MA(\infty)$ forms for Model 1

For the second model, the SARIMA(3,0,0)(0,1,2) has the following form:

$$(1-\phi_1 B-\phi_2 B^2-\phi_3 B^3)x_t = (1+\Theta_1 B^{12}+\Theta_2 B^{24})w_t \tag{3}$$

From this model, one can use the previous logic to obtain the $AR(\infty)$ and the $MA(\infty)$ forms. In this case, the $MA(\infty)$ form would be the following:

$$x_t = \frac{1 + \Theta_1 B^{12} + \Theta_2 B^{24}}{1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3} w_t \implies x_t = (1 + \psi_1 B + \psi_2 B^2 + ...) w_t \tag{4}$$

And the $AR(\infty)$ form would be

$$\frac{1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3}{1 + \Theta_1 B^{12} + \Theta_2 B^{24}} x_t = w_t \implies (1 + \pi_1 B + \pi_2 B^2 + ...) x_t = w_t \tag{5}$$

The estimated coefficients for both forms of this model are shown in Figure 14, and because no parameters have an absolute value greater or equal to one, we can say that the model is both invertible and causal.

```
Psi-weights (MA(inf))

--------------------
       psi 1        psi 2        psi 3        psi 4        psi 5        psi 6 
psi 7        psi 8 
 0.104089599  0.354713718  0.566690464  0.232382786  0.394281055  0.400882614 
0.292103820  0.363024788 
       psi 9       psi 10       psi 11       psi 12       psi 13       psi 14 
psi 15       psi 16 
 0.336261172  0.304129675  0.326614803 -0.185248979  0.243265723  0.122957477 
0.004944366  0.162964156 
      psi 17       psi 18       psi 19       psi 20 
 0.079400961  0.066747166  0.114752107  0.074119483 

Pi-weights (AR(inf))

--------------------
       pi 1        pi 2        pi 3        pi 4        pi 5        pi 6        pi 7 
pi 8        pi 9 
 0.10408960  0.34387907  0.49397422  0.00000000  0.00000000  0.00000000  0.00000000 
0.00000000  0.00000000 
      pi 10       pi 11       pi 12       pi 13       pi 14       pi 15       pi 16 
pi 17       pi 18 
 0.00000000  0.00000000 -0.48993436  0.05099707  0.16847817  0.24201495  0.00000000 
0.00000000  0.00000000 
      pi 19       pi 20 
 0.00000000  0.00000000 
```

Figure 14: Estimated values for the $AR(\infty)$ and $MA(\infty)$ forms for Model 2

For the final model, the SARIMA(3,0,0)(1,1,1) has the following form:

$$(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3)(1 - \Phi_1 B^{12}) x_t = (1 + \Theta_1 B^{12}) w_t \tag{6}$$

From this model, one can use the previous logic to obtain the $AR(\infty)$ and the $MA(\infty)$ forms. In this case, the $MA(\infty)$ form would be the following:

$$x_t = \frac{1 + \Theta_1 B^{12}}{(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3)(1 - \Phi_1 B^{12})} w_t \implies x_t = (1 + \psi_1 B + \psi_2 B^2 + ...) w_t \tag{7}$$

And the $AR(\infty)$ form would be

$$\frac{(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3)(1 - \Phi_1 B^{12})}{1 + \Theta_1 B^{12}} x_t = w_t \implies (1 + \pi_1 B + \pi_2 B^2 + ...) x_t = w_t \tag{8}$$

The estimated coefficients for both forms of this model are shown in Figure 15, and because no parameters have an absolute value greater or equal to one, we can say that the model is both invertible and causal.

```
Psi-weights (MA(inf))

--------------------
      psi 1       psi 2       psi 3       psi 4       psi 5       psi 6       psi 7
psi 8       psi 9
 0.09464089  0.34714302  0.57611322  0.22030827  0.39316160  0.40625401  0.28404332
0.36527666  0.33832822
     psi 10      psi 11      psi 12      psi 13      psi 14      psi 15      psi 16
psi 17      psi 18
 0.30076920  0.32963178 -0.25413973  0.24119406  0.10540555 -0.03838555  0.15532505
0.05560755  0.03816679
     psi 19      psi 20
 0.10182825  0.05097413

Pi-weights (AR(inf))

--------------------
       pi 1        pi 2        pi 3        pi 4        pi 5        pi 6        pi 7
pi 8        pi 9
 0.09464089  0.33818613  0.51125306  0.00000000  0.00000000  0.00000000  0.00000000
0.00000000  0.00000000
      pi 10       pi 11       pi 12       pi 13       pi 14       pi 15       pi 16
pi 17       pi 18
 0.00000000  0.00000000 -0.56002368  0.05300114  0.18939224  0.28631382  0.00000000
0.00000000  0.00000000
      pi 19       pi 20
 0.00000000  0.00000000
```

Figure 15: Estimated values for the $AR(\infty)$ and $MA(\infty)$ forms for Model 3

We can finally provide adequacy measures for the three models. These measures are the estimated σ^2, the AIC and the BIC. For comparing models, we prefer the model with the lowest Σ^2 (with the lowest variance for the white noise) and the lowest AIC and BIC. These measures are portrayed in Figure 16, and we can clearly see that the second model, the SARIMA(3,0,0)(0,1,2) is the one that has the best adequacy measures, and hence it is an indication that this might be a good selection to model this time series.

	Sigma2Z <dbl>	AIC <dbl>	BIC <dbl>
SARIMA(3,0,0)(1,0,0)	0.001750673	-1208.740	-1189.479
SARIMA(3,0,0)(0,0,2)	0.001449487	-1268.122	-1245.009
SARIMA(3,0,0)(1,0,1)	0.001464114	-1264.572	-1241.459

Figure 16: Adequacy measures for the different models

4 Predictions

Once we have estimated the models and have checked different aspects, we can focus on the prediction capability of the models and choose the one which is best for this purpose.

4.1 Forecasting Capability

Now we check both the stability of the models and their predictive capability. When looking at the SARIMA(3,0,0)(1,1,0), we can see that the model is stable, in the sense that the model does not vary in terms of significance, magnitude and sign of the estimated coefficients. Moreover, the predictions out-of-sample seem to be pretty close to the real values, but we must remember that the first coefficient of the regular AR part is not significant, which can make the model predictions not reliable. In Figure 17 one can see the estimation results of the first model, which does not include the 12 last observations, and the model taking into account the whole sample, while there is also a graph for the time series predictions.

The results of the other two models are similar: both models are stable and the predictions are quite similar (visually) to the observed values. The results for both models are represented in Figures 18 and 19. Note that in the graph of the time series, we have zoomed to the 2015-2020 window for the predictions to be easily represented, but the time series spans from 1990 to 2020. However, a more robust assessment of the predictive capability of both models is given by common metrics such as the RMSPE, the MAPE and others, which are presented in a table in Figure 20.

By looking at the metrics, we can observe how the first model, which is the SARIMA(3,0,0)(1,1,0) is the best model in numerical terms, as it has the lowest values for these measures. Nevertheless, we pointed out that not all the coefficients of this model are significant (the first one for the AR part is not), so it would be prudent to choose a model with significant coefficients. Hence, we can choose the third model, the SARIMA(3,0,0)(1,1,1), as it is the best among the two left in terms of the measures.

```
Call:
arima(x = lserie1, order = pdq, seasonal = list(order = PDQ, period = 12))

Coefficients:
         ar1     ar2     ar3     sar1
      0.0519  0.3242  0.5177  -0.3033
s.e.  0.0457  0.0423  0.0466   0.0522

sigma^2 estimated as 0.001751:  log likelihood = 609.37,  aic = -1208.74

Call:
arima(x = lserie2, order = pdq, seasonal = list(order = PDQ, period = 12))

Coefficients:
         ar1     ar2     ar3     sar1
      0.0537  0.3242  0.5162  -0.3026
s.e.  0.0467  0.0432  0.0476   0.0535

sigma^2 estimated as 0.001801:  log likelihood = 583.53,  aic = -1157.06
```

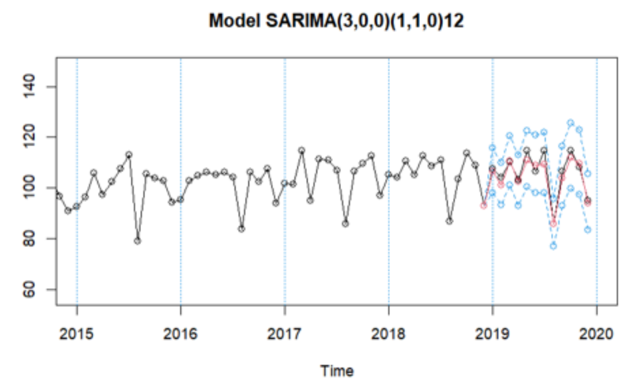


Figure 17: Stability and out-of-sample prediction for Model 1

```
Call:
arima(x = lserie1, order = pdq, seasonal = list(order = PDQ, period = 12))

Coefficients:
         ar1     ar2     ar3     sma1     sma2
      0.1041  0.3439  0.4940  -0.4899  -0.1725
s.e.  0.0473  0.0433  0.0484   0.0651   0.0620

sigma^2 estimated as 0.001449:  log likelihood = 640.06,  aic = -1268.12

Call:
arima(x = lserie2, order = pdq, seasonal = list(order = PDQ, period = 12))

Coefficients:
         ar1     ar2     ar3     sma1     sma2
      0.1088  0.3418  0.4911  -0.4823  -0.1816
s.e.  0.0484  0.0442  0.0494   0.0685   0.0655

sigma^2 estimated as 0.001486:  log likelihood = 613.7,  aic = -1215.41
```

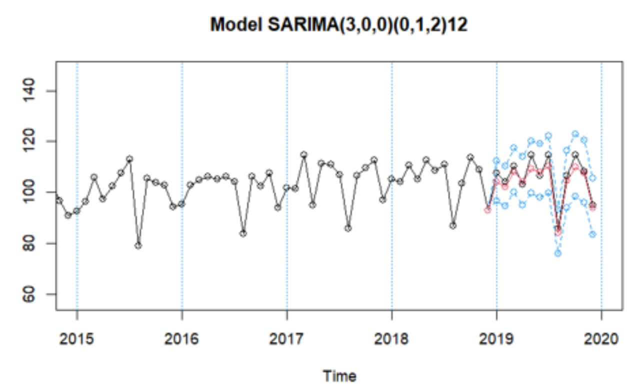


Figure 18: Stability and out-of-sample prediction for Model 2

```
Call:
arima(x = lserie1, order = pdq, seasonal = list(order = PDQ, period = 12))

Coefficients:
         ar1     ar2     ar3    sar1     sma1
      0.0946  0.3382  0.5113  0.1473  -0.7073
s.e.  0.0465  0.0426  0.0468  0.0713   0.0439

sigma^2 estimated as 0.001464:  log likelihood = 638.29,  aic = -1264.57

Call:
arima(x = lserie2, order = pdq, seasonal = list(order = PDQ, period = 12))

Coefficients:
         ar1     ar2     ar3    sar1     sma1
      0.0982  0.3373  0.5089  0.1459  -0.7079
s.e.  0.0475  0.0436  0.0478  0.0723   0.0444

sigma^2 estimated as 0.001502:  log likelihood = 611.81,  aic = -1211.62
```

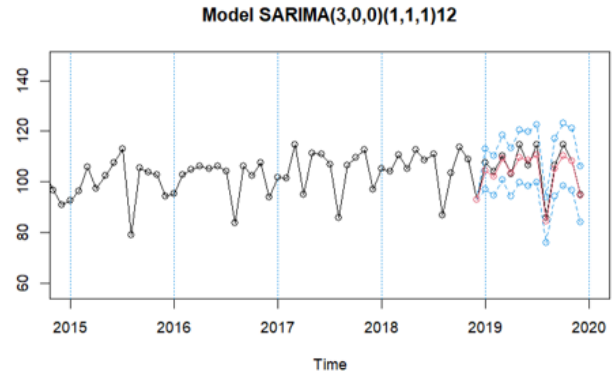


Figure 19: Stability and out-of-sample prediction for Model 3

	AIC	BIC	RMSE	MAE	RMSPE	MAPE	meanCI
SARIMA(3,0,0)(1,1,0)	-1208.740	-1189.479	2.493172	2.010205	0.02243398	0.01840424	19.92079
SARIMA(3,0,0)(0,1,2)	-1268.122	-1245.009	2.857626	2.449517	0.02578888	0.02262845	18.68248
SARIMA(3,0,0)(1,1,1)	-1264.572	-1241.459	2.645716	2.116395	0.02373318	0.01943488	18.84681

Figure 20: Predictive metrics for the models

4.2 Long-term Forecasts

We can use the third model, which we have selected as the best for prediction purposes, to obtain a forecast of the next 12 months (after 2020, the ending of our data) for the index. These predictions are represented in Figure 21 with their respective confidence intervals.

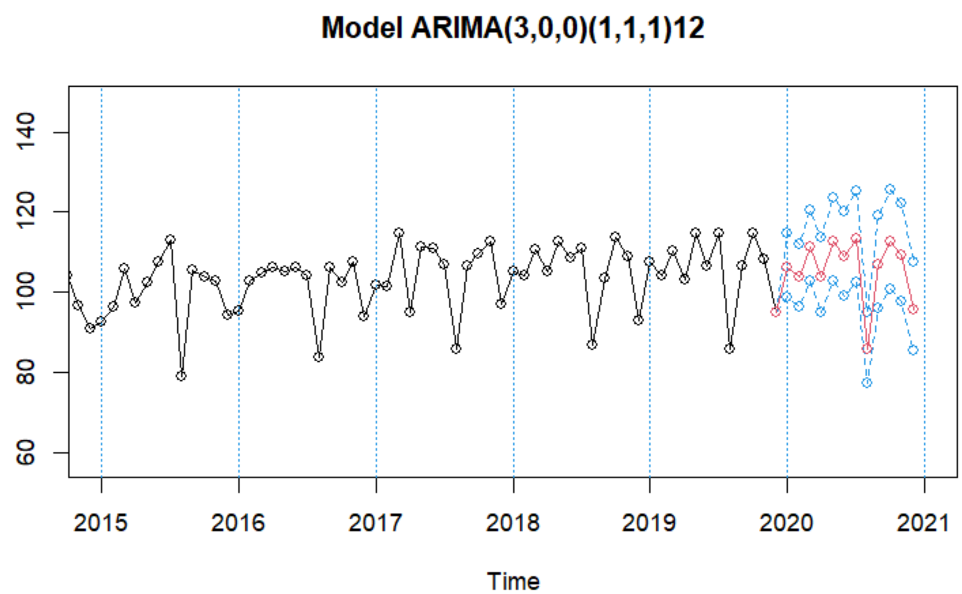


Figure 21: Long-term forecasts using Model 3

As we can see, the predicted evolution of the IPI is very similar to the one of the last year. It seems that the IPI level will oscilate between 80 and 120 (using 2016 as the base year) by looking at the confidence intervals. However, we now know that we should not expect the same behaviour, as the COVID-19 crisis effects began on 2020 and had very harmful effects on all industries and countries, especially in Europe, where strict lockdowns made production levels decrease.

5 Outlier Treatment

As we mentioned earlier, it might be the case that this time series contains outliers or is affected by calendar effects. For example, one can clearly see that the 2008 financial crisis has had an impact on the index, sharply reducing its level at the time and also lowering the posterior level. Therefore, we now proceed to the outlier detection and the calendar treatment for our last selected model (Model 3), and we seek to obtain better results when treating them.

5.1 Calendar Effects

First of all, we start by analyzing the impact of calendar effects in our time series. We consider two kinds of calendar effects: the trading days effect and the effect of eastern holidays. From an economic perspective, we should not expect the trading days effect to have a significant impact on the IPI time series, as there is no reason why the number of trading days in a year should affect production at the macroeconomic level. In contrast, Eastern holidays might have a significant effect given that, because of holidays, firms close or at least they do not have as much workers producing, and hence it reduces the monthly level.

Anyway, we consider both effects by including them into the model and we then discuss their significance. For this purpose, we have created a variable for both kind of calendar effects, introducing them through a seasonal ARIMAX model, which allows to include external variables in the model.

Additionally, we consider the effect of the crisis of 2008 in the same way as we consider the calendar effects. Based on the INE statistics, the crisis of 2008 for the Spanish economy span from 2008 to 2014, so we consider a variable for all the months compressed in this time interval, to see whether the effect is significant or not.

We show the results of the linearized model with the external variables accounting for calendar effects and the crisis in Figure 22. The image presents a comparison between the SARIMA(3,0,0)(0,1,2) including the three mentioned external variables and the model without the variable of the crisis. The results from the first model shows that the trading days and the Eastern holidays effects are significant in this model for the time series, while the variable of the crisis of 2008 is not. This allows us to reestimate the model ignoring the third variable, and now we obtained a model with better fit (in terms of the AIC) and significant coefficients (all of them).

```
Call:
arima(x = lserie, order = c(3, 0, 0), seasonal = list(order = c(1, 1, 1), period = 12),
    xreg = data.frame(wTradDays, wEast, cr))

Coefficients:
         ar1     ar2     ar3    sar1     sma1  wTradDays    wEast       cr
      0.4961  0.3102  0.1494  0.1079  -0.5678     0.0072  -0.0683  -0.0004
s.e.  0.0544  0.0575  0.0542  0.0851   0.0627     0.0004   0.0044   0.0141

sigma^2 estimated as 0.0006143:  log likelihood = 790.62,  aic = -1563.25

Call:
arima(x = lserie, order = c(3, 0, 0), seasonal = list(order = c(1, 1, 1), period = 12),
    xreg = data.frame(wTradDays, wEast))

Coefficients:
         ar1     ar2     ar3    sar1     sma1  wTradDays    wEast
      0.4965  0.3101  0.1491  0.1079  -0.5678     0.0072  -0.0683
s.e.  0.0530  0.0574  0.0536  0.0851   0.0627     0.0004   0.0043

sigma^2 estimated as 0.0006143:  log likelihood = 790.62,  aic = -1565.25
```

Figure 22: Results of the linearized model taking including and excluding the crisis variable

The total calendar effects (which is the sum of the effects of trading days and Eastern holidays) on the time series is represented in Figure 23.

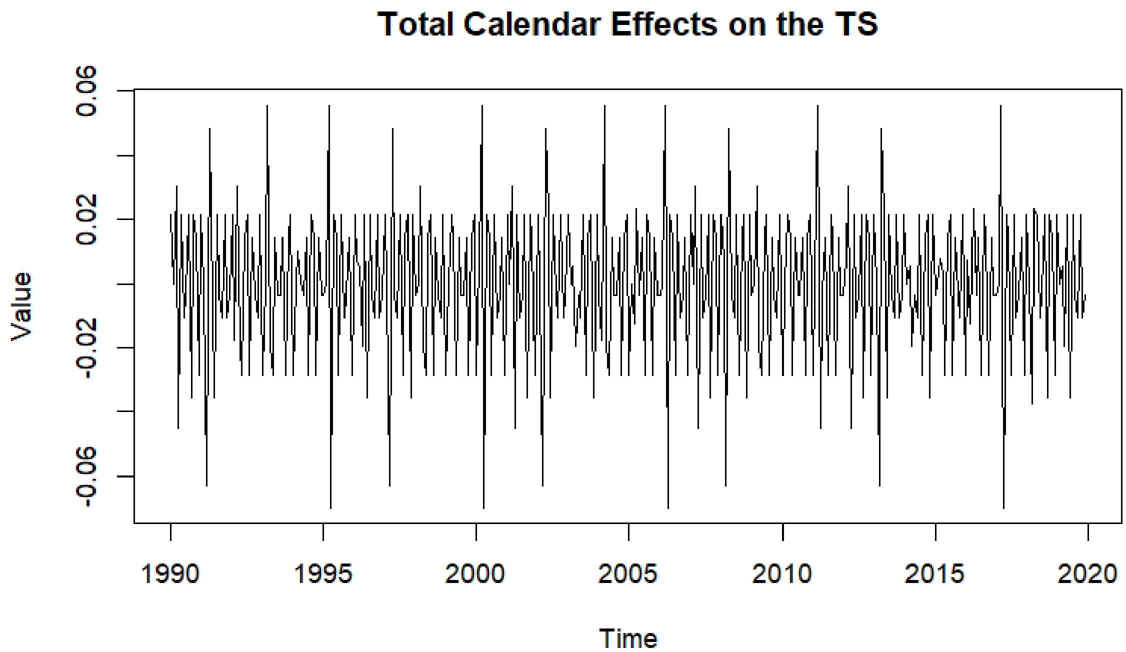


Figure 23: Total calendar effects estimated

5.2 Outlier Treatment

Once the calendar effects are analyzed, we can go and analyze the presence of outliers in this series. To do so, we create a function which allows to automatically detect outliers in our time series, and presents the type (among additive outliers, level shifts and transitory changes), the weight of each outlier and other metrics. We present the results of applying this function to our time series in Figure 24.

	Obs <int>	type_detected <chr>	W_coeff <dbl>	ABS_L_Ratio <dbl>	Fecha <chr>	perc.Obs <dbl>
9	15	AO	-0.07257312	2.992216	Mar 1991	92.99977
5	88	AO	0.09469381	3.689789	Abr 1997	109.93222
10	147	AO	-0.07021774	2.929404	Mar 2002	93.21908
7	148	AO	0.07723559	3.105423	Abr 2002	108.02966
3	184	AO	0.09968814	3.742840	Abr 2005	110.48263
6	219	LS	-0.07552708	3.648173	Mar 2008	92.72546
4	220	AO	0.09646880	3.689751	Abr 2008	110.12752
1	227	LS	-0.11127177	4.885023	Nov 2008	89.46956
8	241	AO	-0.07516431	3.060959	Ene 2010	92.75911
2	316	AO	0.10771478	3.946289	Abr 2016	111.37300

Figure 24: Outliers estimates

The results in Figure 24 show that type of outliers present in our data are mainly additive outliers, but there are also some important level shifts. Some of this outliers are linked to important economic events in Spain. We proceed to list some of the most relevant events:

- In 2008, some negative level shifts happened due to the financial crisis of 2008, which led to a decrease in economic activity worldwide.

- In late 2009 the European debt crisis shattered Europe, which already was in a recession due to the 2008 crisis. This might have had a negative effect on production in 2010 for Spain.
- During this last crisis, many structural reforms were introduced, which led Spain to a strong recovery of its GDP. This might explain the increase in the production level.

Figure 25 shows the effects of this outliers in the series in a more graphical way.

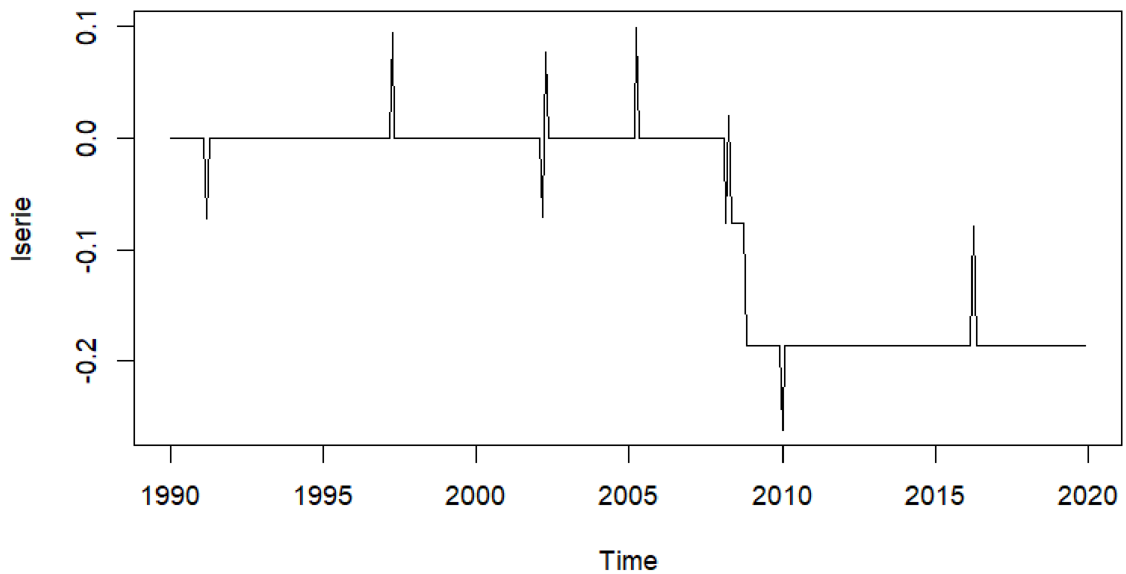


Figure 25: Outliers effect on the time series

5.3 Model Comparison

Finally, we can take both the calendar effects and the outlier effects in order to reestimate the models and compare the best one to the original estimation in terms of forecasting capability.

When we use the series taking into account both the calendar and the outliers' effects, we can obtain the following ACF and PACF plots in Figure 26. One can compare to the original ACF and PACF and see how the pattern is way clearer when using the linearized series, but we identify the same models as before: the SARIMA(3,0,0)(1,1,0), the SARIMA(3,0,0)(0,1,2) and the SARIMA(3,0,0)(1,1,1).

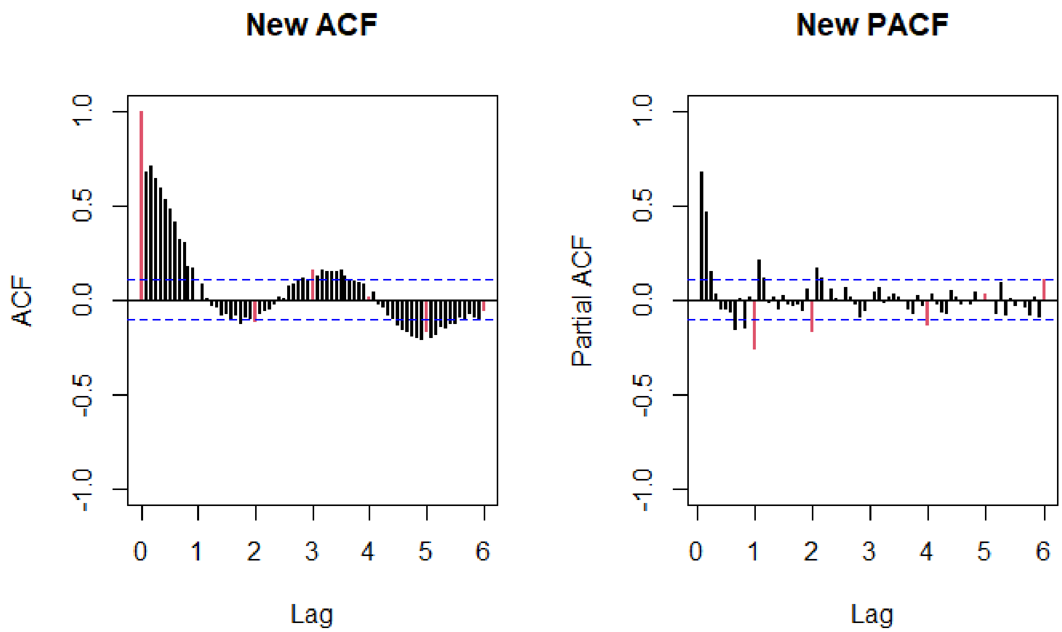


Figure 26: ACF and PACF taking into account calendar and outliers' effects

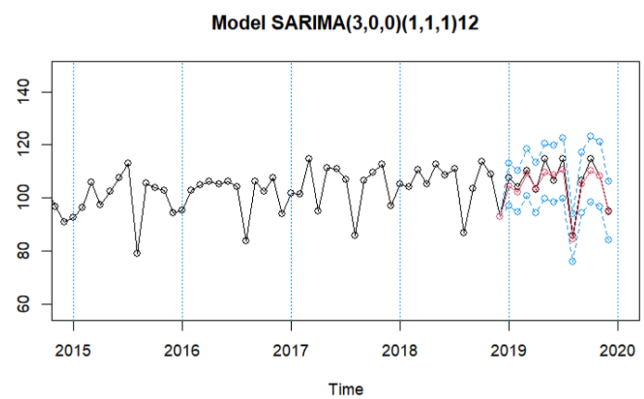
We estimate and check the validity of this models, and the results show that the SARIMA(3,0,0)(1,1,0) still suffers from the problem of independence of the residuals, but both the second and the third models are valid, and both fit the data equally good in terms of the AIC. We do not show the results of the validations, but we show the estimation results in Figure 27. To keep the analysis made above for the SARIMA(3,0,0)(0,1,2), we select this model again as the one we will use to forecast.

To proceed to the forecasting part, we first check that the newly estimated model is stable, and the results show it is. Hence, we can proceed as before: we remove the 12 last observations in order to carry on an out-of-sample prediction exercise and obtain different metrics to compare to the out-of-sample predictions of the original model. This comparison is made graphically in Figure 28, where we can observe the differences between the interval lengths and some differences of the different points.

However, a more precise analysis can be made by computing and comparing the different measures we used previously to evaluate the forecasts, shown in Figure 29. In here, the results seem to be a little bit surprising: in terms of RMSPE and MAPE, the original SARIMA(3,0,0)(0,1,2) (without taking into account the effects) is better than the new one. Even though the mean confidence interval length is lower for the new model, it seems that there is some kind of noise that affects the predictive capability of our new model (like added variables, for example), and hence the original model yields better results. That also explains why there is a better fit for the data in terms of AIC and BIC, but there is a worsening in measures that tkae into account the out-of-sample predictions.

```
Call:
arima(x = d12lnserie.lin, order = c(3, 0, 0), seasonal = list(order = c(1, 0,
    0)), include.mean = F)

Coefficients:
         ar1     ar2     ar3     sar1
      0.3311  0.3815  0.1880  -0.3716
s.e.  0.0530  0.0519  0.0528   0.0514

sigma^2 estimated as 0.0006977:  log likelihood = 769.32,  aic = -1528.65

Call:
arima(x = d12lnserie.lin, order = c(3, 0, 0), seasonal = list(order = c(0, 0,
    2)), include.mean = F)

Coefficients:
         ar1     ar2     ar3     sma1     sma2
      0.3475  0.3821  0.1997  -0.5151  -0.0042
s.e.  0.0531  0.0524  0.0528   0.0630   0.0579

sigma^2 estimated as 0.0006431:  log likelihood = 782.58,  aic = -1553.16

Call:
arima(x = d12lnserie.lin, order = c(3, 0, 0), seasonal = list(order = c(1, 0,
    1)), include.mean = F)

Coefficients:
         ar1     ar2     ar3    sar1     sma1
      0.3476  0.3819  0.1998  0.0049  -0.5213
s.e.  0.0530  0.0523  0.0527  0.0869   0.0682

sigma^2 estimated as 0.0006431:  log likelihood = 782.58,  aic = -1553.16
```

Figure 27: Estimation of the models taking into account calendar and outliers' effects



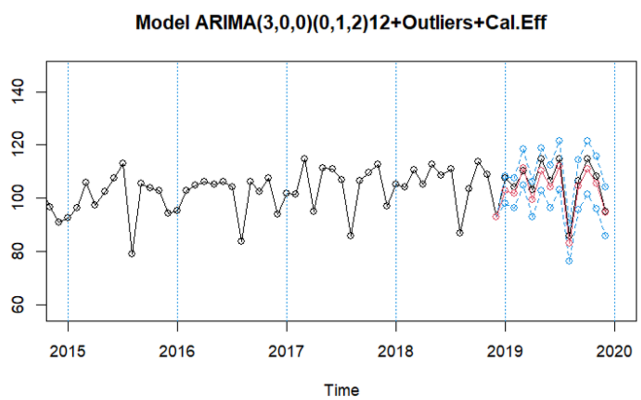


Figure 28: Out-of-sample forecasts of the treated model and the original model

	par <int>	Sigma2Z <dbl>	AIC <dbl>	BIC <dbl>	RMSE <dbl>	MAE <dbl>	RMSPE <dbl>	MAPE <dbl>	meanLength <dbl>
ARIMA(3,0,0)(0,1,2)12	5	0.0014641211	-1264.57	-1241.457	2.645716	2.116395	0.02373318	0.01943488	18.84681
ARIMA(3,0,0)(0,1,2)12+Outliers+Cal.Eff	17	0.0006121684	-1546.54	-1477.229	2.993877	2.746305	0.02807610	0.02582209	14.62825

Figure 29: Forecasts metrics comparison

6 Conclusions

To summarize the most important conclusions of the project, we list them in a brief way:

- The IPI time series needs a logarithmic transformation to achieve constant variance and a monthly seasonal difference to eliminate the seasonal pattern and obtain a stationary time series

- The identified models have been a SARIMA(3,0,0)(1,1,0), a SARIMA(3,0,0)(0,1,2) and a SARIMA(3,0,0)(1,1,1). The validation check of these models reveals that there is some dependence structure not captured by the models, even though other assumptions do hold.

- All of these models are both causal and invertible, and the values of the coefficients for the AR(∞) and the MA(∞) forms are given.

- The model that yields the best forecasts is the SARIMA(3,0,0)(0,1,2) in terms of oust-of-sample prediction metrics. The long-term forecasts for the next 12 months is pretty similar to the last 12 months.

- The model that yields the best forecasts is the SARIMA(3,0,0)(0,1,2) in terms of oust-of-sample prediction metrics. The long-term forecasts for the next 12 months is pretty similar to the last 12 months.

- The trading days' effects and the Eastern holidays' effects seem significant for this series, even though the first one does not have a direct economic reasoning on why.

- The crisis variable introduced resulted non-significant when checking calendar effects, but the outlier treatment shows that the crisis had the most substantial effect on the time series, shifting down the IPI level. Other outliers were detected, but just some have an economic reasoning and relative relevance in the time series.

- When taking into account both types of effects, we identified the same models as before. However, when carrying on out-of-sample prediction with this new model, the metrics seem to favour the original model and not this new one.