# Annex

## Tourist arrivals in Spain

Nicolás Jiménez Muñoz and Aina Vila Arbusà

```
Series: d12lnseries
ARIMA(2,1,0)(2,0,0)[12]

Coefficients:
          ar1      ar2     sar1     sar2
      -0.6830  -0.3341  -0.4586  -0.2184
s.e.   0.0655   0.0647   0.0717   0.0770

sigma^2 = 0.0022:  log likelihood = 351.38
AIC=-692.76   AICc=-692.48   BIC=-675.93

Training set error measures:
                        ME       RMSE        MAE       MPE     MAPE      MASE         ACF1
Training set -0.0005475747 0.04635031 0.03499303 -18.34131 195.9945 0.5302564 -0.001666658
```

**Figure 1**: auto.arima() output for d12lnseries.

```
Call:
arima(x = d1d12lnseries, order = c(2, 0, 0), seasonal = list(order = c(2, 0,
    0), period = 12))

Coefficients:
          ar1      ar2     sar1     sar2  intercept
      -0.6830  -0.3341  -0.4586  -0.2184     -1e-04
s.e.   0.0655   0.0647   0.0717   0.0770      1e-03

sigma^2 estimated as 0.002158:  log likelihood = 351.39,  aic = -690.77

Call:
arima(x = d1d12lnseries, order = c(0, 0, 4), seasonal = list(order = c(1, 0,
    0), period = 12))

Coefficients:
          ma1     ma2     ma3      ma4     sar1  intercept
      -0.7366  0.0949  0.0733  -0.1299  -0.3589     -1e-04
s.e.   0.0699  0.0900  0.0915   0.0697   0.0653      7e-04

sigma^2 estimated as 0.002206:  log likelihood = 349.62,  aic = -685.23

Call:
arima(x = d1d12lnseries, order = c(2, 0, 0), seasonal = list(order = c(1, 0,
    0), period = 12))

Coefficients:
          ar1      ar2     sar1  intercept
      -0.7127  -0.3564  -0.3684    -0.0001
s.e.   0.0638   0.0635   0.0649     0.0012

sigma^2 estimated as 0.002251:  log likelihood = 347.48,  aic = -684.96

Call:
arima(x = d1d12lnseries, order = c(0, 0, 4), seasonal = list(order = c(0, 0,
    1), period = 12))

Coefficients:
          ma1     ma2     ma3      ma4     sma1  intercept
      -0.7086  0.0818  0.0706  -0.1444  -0.4749      0e+00
s.e.   0.0708  0.0878  0.0894   0.0689   0.0778      5e-04

sigma^2 estimated as 0.002119:  log likelihood = 353.23,  aic = -692.45
```

```
Call:
arima(x = d1d12lnseries, order = c(2, 0, 0), seasonal = list(order = c(0, 0,
    1), period = 12))

Coefficients:
          ar1      ar2     sma1  intercept
      -0.6863  -0.3341  -0.4642     -1e-04
s.e.   0.0655   0.0645   0.0760      9e-04

sigma^2 estimated as 0.002179:  log likelihood = 350.4,  aic = -690.8
```

**Figure 2**: model estimation with d1d12lnseries.

```
                                      intercept
"mod1 : Do not reject H0: Not significant"
                                      intercept
"mod2 : Do not reject H0: Not significant"
                                      intercept
"mod3 : Do not reject H0: Not significant"
                                      intercept
"mod4 : Do not reject H0: Not significant"
                                      intercept
"mod5 : Do not reject H0: Not significant"
```

**Figure 3:** intercept significance of each model.

```
Call:
arima(x = lnseries, order = c(2, 1, 0), seasonal = list(order = c(2, 1, 0),
    period = 12))

Coefficients:
          ar1      ar2     sar1     sar2
      -0.6830  -0.3341  -0.4586  -0.2184
s.e.   0.0655   0.0647   0.0717   0.0770

sigma^2 estimated as 0.002158:  log likelihood = 351.38,  aic = -692.76

Call:
arima(x = lnseries, order = c(0, 1, 4), seasonal = list(order = c(1, 1, 0),
    period = 12))

Coefficients:
          ma1     ma2     ma3      ma4     sar1
      -0.7365  0.0949  0.0733  -0.1299  -0.3589
s.e.   0.0699  0.0900  0.0915   0.0697   0.0653

sigma^2 estimated as 0.002206:  log likelihood = 349.61,  aic = -687.21

Call:
arima(x = lnseries, order = c(2, 1, 0), seasonal = list(order = c(1, 1, 0),
    period = 12))

Coefficients:
          ar1      ar2     sar1
      -0.7127  -0.3563  -0.3684
s.e.   0.0638   0.0635   0.0649

sigma^2 estimated as 0.002251:  log likelihood = 347.47,  aic = -686.95

Call:
arima(x = lnseries, order = c(0, 1, 4), seasonal = list(order = c(0, 1, 1),
    period = 12))

Coefficients:
          ma1     ma2     ma3      ma4     sma1
      -0.7086  0.0817  0.0706  -0.1445  -0.4750
s.e.   0.0708  0.0878  0.0894   0.0689   0.0777

sigma^2 estimated as 0.002119:  log likelihood = 353.23,  aic = -694.45
```

```
Call:
arima(x = lnseries, order = c(2, 1, 0), seasonal = list(order = c(0, 1, 1),
    period = 12))

Coefficients:
          ar1      ar2     sma1
      -0.6863  -0.3341  -0.4643
s.e.   0.0655   0.0645   0.0760

sigma^2 estimated as 0.002179:  log likelihood = 350.4,  aic = -692.79
```
**Figure 4**: model estimation with lnseries.

| | T_Value |
|---|---|
| | <dbl> |
| ar1 | -10.429090 |
| ar2 | -5.166715 |
| sar1 | -6.399923 |
| sar2 | -2.838294 |

**Figure 5**: mod1 significance test.

| | T_Value |
|---|---|
| | <dbl> |
| ma1 | -10.5324704 |
| ma2 | 1.0543841 |
| ma3 | 0.8008811 |
| ma4 | -1.8631577 |
| sar1 | -5.4956136 |

**Figure 6**: mod2 significance test.

| | T_Value |
|---|---|
| | <dbl> |
| ar1 | -11.175206 |
| ar2 | -5.612554 |
| sar1 | -5.678867 |

**Figure 7**: mod3 significance test.

| | T_Value |
|---|---|
| | <dbl> |
| ma1 | -10.0126429 |
| ma2 | 0.9309191 |
| ma3 | 0.7898606 |
| ma4 | -2.0967179 |
| sma1 | -6.1129571 |

**Figure 8**: mod4 significance test.

| | T_Value |
|---|---|
| | <dbl> |
| ar1 | -10.471170 |
| ar2 | -5.176650 |
| sma1 | -6.105575 |

**Figure 9**: mod5 significance test.

```
Call:
arima(x = lnseries, order = c(0, 1, 4), seasonal = list(order = c(1, 1, 0),
    period = 12), fixed = c(NA, NA, 0, NA, NA))

Coefficients:
          ma1     ma2  ma3      ma4     sar1
      -0.7508  0.1442    0  -0.0960  -0.3547
s.e.   0.0703  0.0700    0   0.0557   0.0650

sigma^2 estimated as 0.002213:  log likelihood = 349.3,  aic = -688.6

Call:
arima(x = lnseries, order = c(0, 1, 4), seasonal = list(order = c(0, 1, 1),
    period = 12), fixed = c(NA, 0, NA, NA, NA))

Coefficients:
          ma1  ma2     ma3      ma4     sma1
      -0.6735    0  0.1177  -0.1505  -0.4833
s.e.   0.0582    0  0.0717   0.0680   0.0776

sigma^2 estimated as 0.002127:  log likelihood = 352.77,  aic = -695.54
```

**Figure 10**: mod2 and mod4 re-estimation.

| | T_Value |
|---|---|
| | <dbl> |
| ma1 | -10.685880 |
| ma2 | 2.059399 |
| ma3 | 0.000000 |
| ma4 | -1.475624 |
| sar1 | -5.047720 |

**Figure 11**: mod2_2 significance test.

| | T_Value |
|---|---|
| | <dbl> |
| ma1 | -11.568549 |
| ma2 | 0.000000 |
| ma3 | 1.729705 |
| ma4 | -1.941014 |
| sma1 | -8.300908 |

**Figure 12**: mod4_2 significance test.

```
Call:
arima(x = lnseries, order = c(0, 1, 2), seasonal = list(order = c(1, 1, 0),
    period = 12))

Coefficients:
          ma1     ma2     sar1
      -0.7594  0.1008  -0.3523
s.e.   0.0665  0.0674   0.0650

sigma^2 estimated as 0.002245:  log likelihood = 347.8,  aic = -687.6

Call:
arima(x = lnseries, order = c(0, 1, 4), seasonal = list(order = c(0, 1, 1),
    period = 12), fixed = c(NA, 0, 0, NA, NA))

Coefficients:
          ma1  ma2  ma3      ma4     sma1
      -0.6441    0    0  -0.0815  -0.4861
s.e.   0.0546    0    0   0.0566   0.0776

sigma^2 estimated as 0.002154:  log likelihood = 351.42,  aic = -694.84
```

**Figure 13**: mod2 and mod4 second re-estimation.

| | T_Value |
| --- | --- |
| | <dbl> |
| ma1 | -11.417333 |
| ma2 | 1.496675 |
| sar1 | -5.418128 |

**Figure 14**: mod2_3 significance test.

| | T_Value |
| --- | --- |
| | <dbl> |
| ma1 | -11.804107 |
| ma2 | 0.000000 |
| ma3 | 0.000000 |
| ma4 | -1.493259 |
| sma1 | -8.590764 |

**Figure 15**: mod4_3 significance test.

```
	Shapiro-Wilk normality test

data:  resid(mod1)
W = 0.96488, p-value = 2.147e-05


	Shapiro-Wilk normality test

data:  resid(mod3)
W = 0.96408, p-value = 1.713e-05


	Shapiro-Wilk normality test

data:  resid(mod5)
W = 0.96165, p-value = 8.753e-06
```

**Figure 16**: Shapiro-Wilk test results for mod1, mod3, mod5.

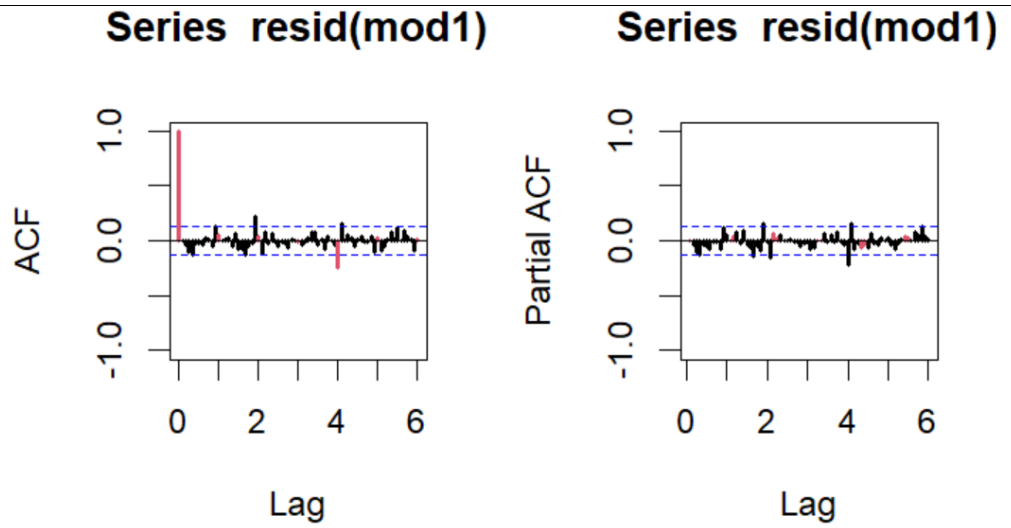


**Figure 17**: ACF / PACF of mod1 residuals.

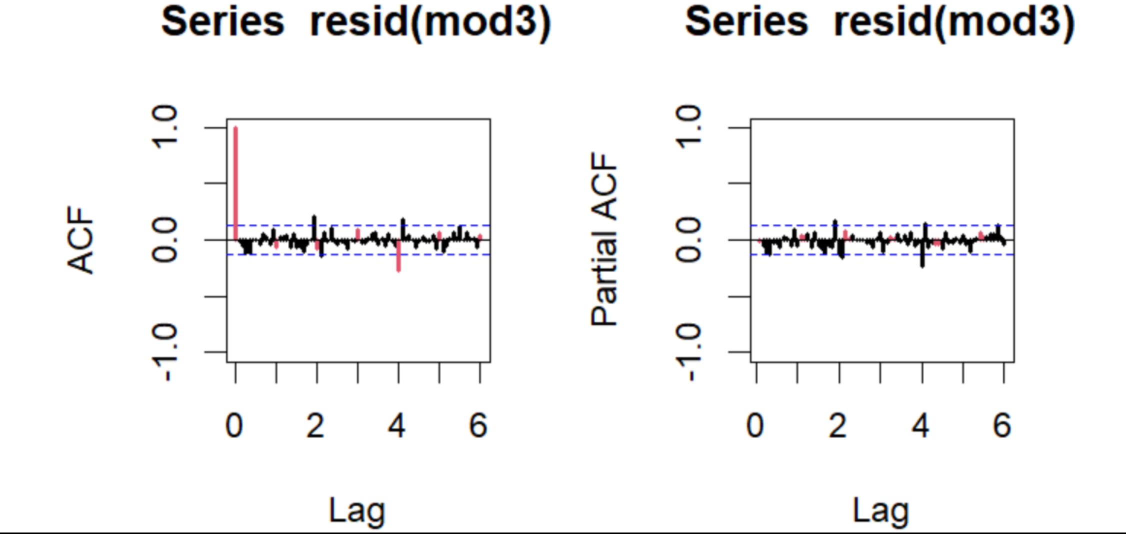


**Figure 18**: ACF / PACF of mod3 residuals.

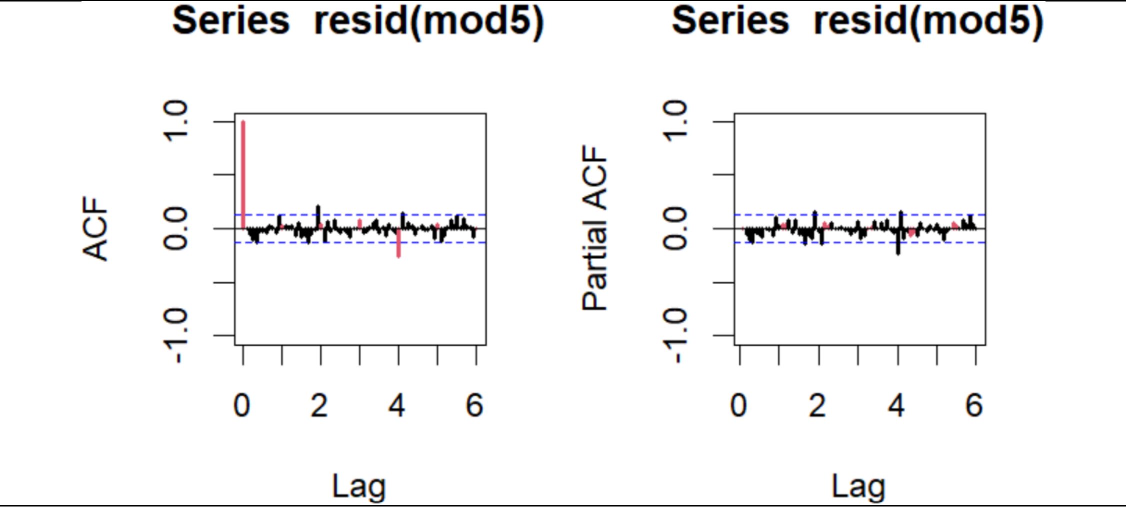


**Figure 19**: ACF / PACF of mod5 residuals.

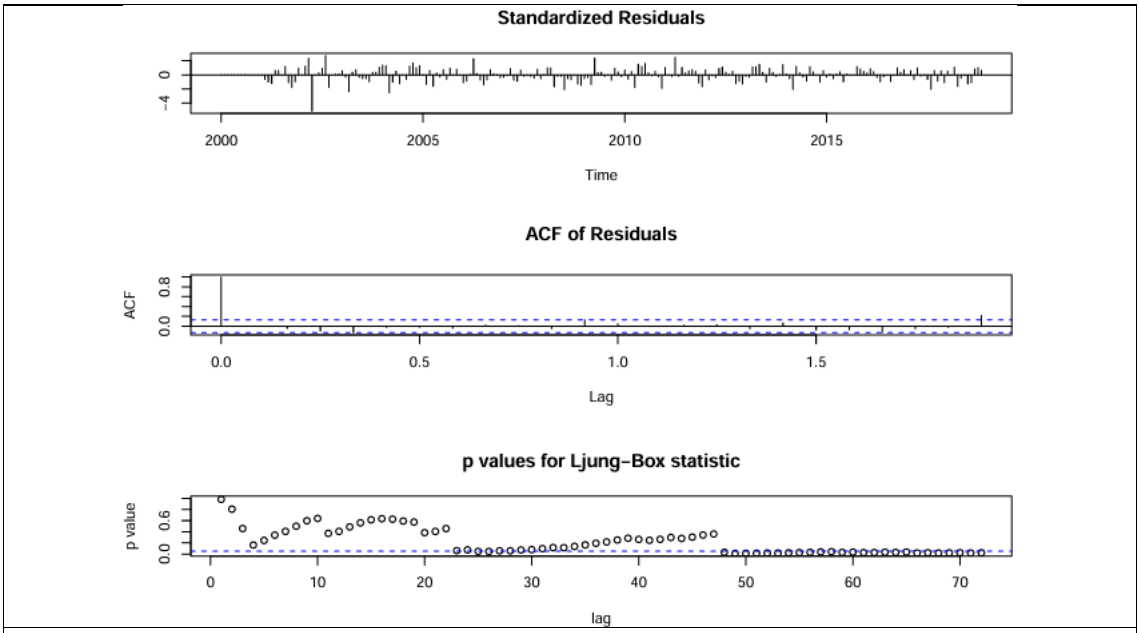


**Figure 20**: Ljung-Box test mod1.

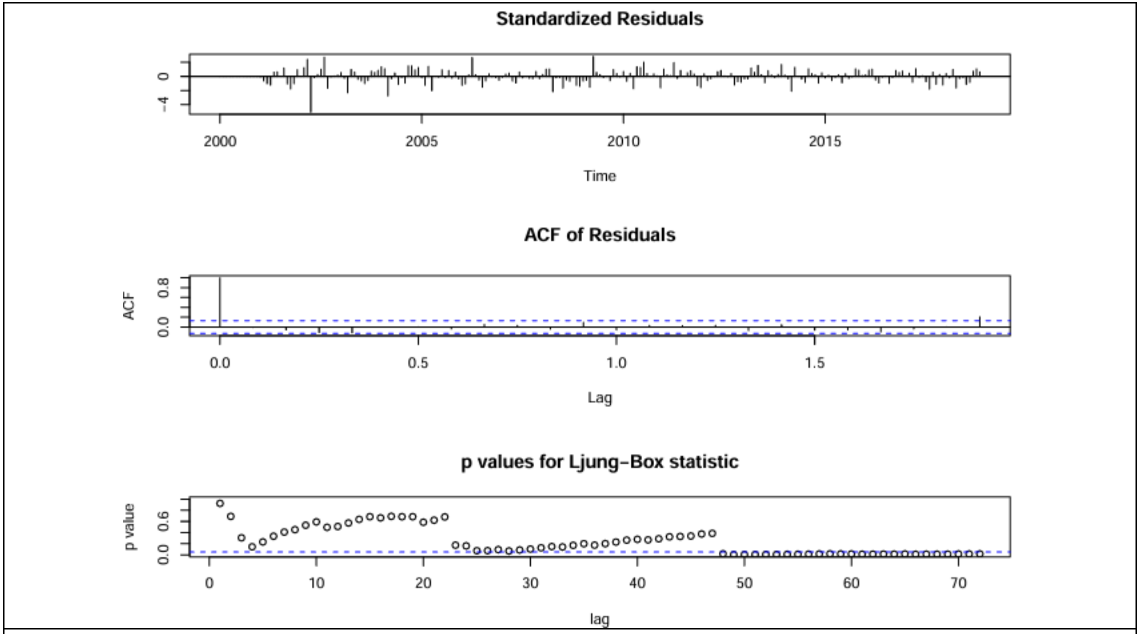


**Figure 21**: Ljung-Box test mod3.

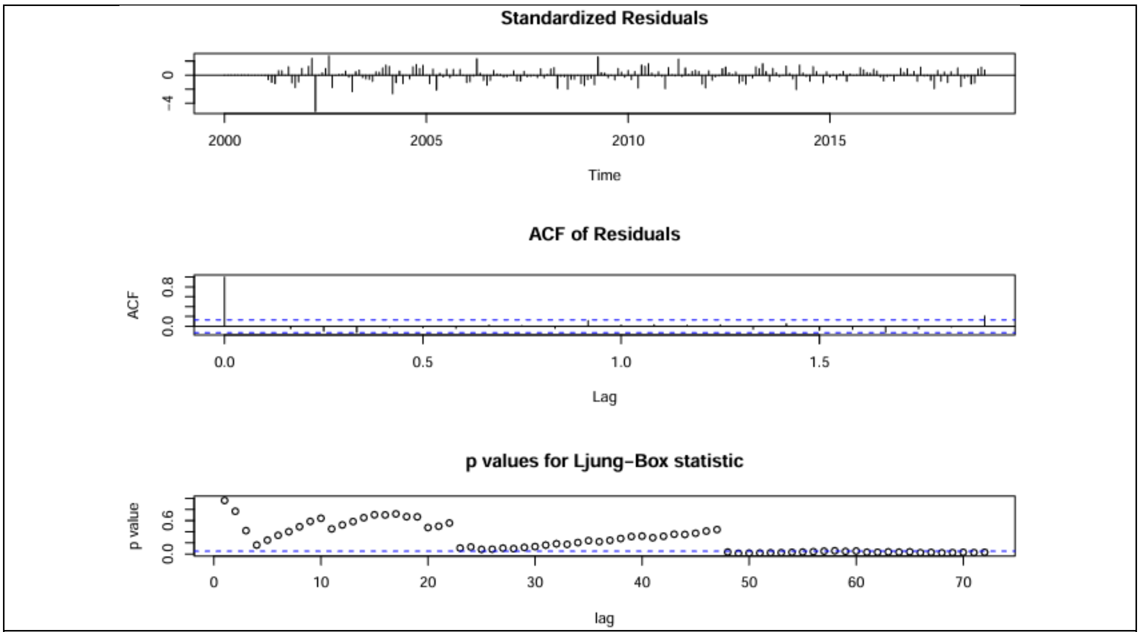


**Figure 22**: Ljung-Box test mod5.

```
 [1] 1.065443 1.065443 1.065443 1.065443 1.065443 1.065443 1.065443 1.065443 1.065443
1.065443 1.065443 1.065443
[13] 1.065443 1.065443 1.065443 1.065443 1.065443 1.065443 1.065443 1.065443 1.065443
1.065443 1.065443 1.065443
[25] 1.729988 1.729988
numeric(0)
```

**Figure 23**: modulus of the polynomial roots (mod1).

```
 [1] 1.086772 1.086772 1.086772 1.086772 1.086772 1.086772 1.086772 1.086772 1.086772
1.086772 1.086772 1.086772
[13] 1.675217 1.675217
numeric(0)
```

**Figure 24**: modulus of the polynomial roots (mod3).

```
[1] 1.730061 1.730061
 [1] 1.066025 1.066025 1.066025 1.066025 1.066025 1.066025 1.066025 1.066025 1.066025
1.066025 1.066025 1.066025
```

**Figure 25**: modulus of the polynomial roots (mod5).

| Date <date> | Forecast <dbl> | Lower_Bound <dbl> | Upper_Bound <dbl> |
|---|---|---|---|
| 2019-12-01 | 4.203541 | 3.837683 | 4.604279 |
| 2020-01-01 | 4.264251 | 3.875761 | 4.691682 |
| 2020-02-01 | 4.427394 | 3.990401 | 4.912242 |
| 2020-03-01 | 5.531977 | 4.921818 | 6.217777 |
| 2020-04-01 | 7.262775 | 6.417025 | 8.219992 |
| 2020-05-01 | 8.517321 | 7.465363 | 9.717513 |
| 2020-06-01 | 9.001774 | 7.827482 | 10.352235 |
| 2020-07-01 | 10.864516 | 9.383040 | 12.579900 |
| 2020-08-01 | 11.088786 | 9.511310 | 12.927892 |
| 2020-09-01 | 9.434078 | 8.039148 | 11.071053 |
| 2020-10-01 | 8.075854 | 6.839254 | 9.536042 |
| 2020-11-01 | 4.789325 | 4.031543 | 5.689541 |

**Figure 26**: mod1 2019 forecast.

```
Call:
arima(x = lnseries, order = c(2, 1, 0), seasonal = list(order = c(2, 1, 0),
    period = 12), xreg = data.frame(vEa, vTD))

Coefficients:
          ar1      ar2     sar1     sar2     vEa      vTD
      -0.5457  -0.2016  -0.2964  -0.0486  0.0737  -0.0025
s.e.   0.0676   0.0679   0.0728   0.0761  0.0082   0.0007

sigma^2 estimated as 0.001568:  log likelihood = 386.69,  aic = -759.38

Call:
arima(x = lnseries, order = c(2, 1, 0), seasonal = list(order = c(2, 1, 0),
    period = 12), xreg = data.frame(vEa))

Coefficients:
          ar1      ar2     sar1     sar2     vEa
      -0.5996  -0.2635  -0.2514  -0.0424  0.0768
s.e.   0.0660   0.0665   0.0718   0.0757  0.0083

sigma^2 estimated as 0.001654:  log likelihood = 381.08,  aic = -750.17

Call:
arima(x = lnseries, order = c(2, 1, 0), seasonal = list(order = c(2, 1, 0),
    period = 12), xreg = data.frame(vTD))

Coefficients:
          ar1      ar2     sar1     sar2      vTD
      -0.6160  -0.2407  -0.5099  -0.2455  -0.0033
s.e.   0.0686   0.0685   0.0722   0.0768   0.0008

sigma^2 estimated as 0.002013:  log likelihood = 358.58,  aic = -705.16
```

**Figure 27**: Calendar effects model proposals.

| | T_Value <dbl> |
|---|---|
| ar1 | -8.0745270 |
| ar2 | -2.9691097 |
| sar1 | -4.0722995 |
| sar2 | -0.6394446 |
| vEa | 9.0325030 |
| vTD | -3.5539561 |

**Figure 28**: t-test on the model with calendar effects.

```
Call:
arima(x = lnser_star, order = c(2, 1, 0), seasonal = list(order = c(1, 1, 0)
    period = 12))

Coefficients:
          ar1      ar2     sar1
      -0.5485  -0.2049  -0.2804
s.e.   0.0671   0.0671   0.0677

sigma^2 estimated as 0.001571:  log likelihood = 386.49,  aic = -764.97
```

**Figure 29**: final model definition with Calendar effect coefficients.

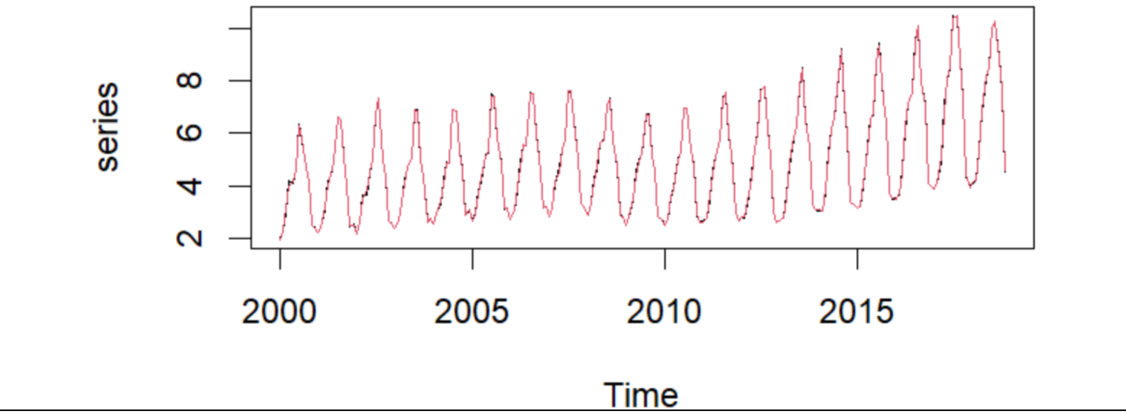


**Figure 30**: comparison between the original (Black) and the calendar effects adjusted series (red).

| Obs | type_detected | W_coeff | ABS_L_Ratio | date | PercVar |
|---|---|---|---|---|---|
| <int> | <chr> | <dbl> | <dbl> | <chr> | <dbl> |
| 27 | AO | 0.07317274 | 3.149224 | Mar 2002 | 107.59164 |
| 32 | AO | 0.10469745 | 4.164723 | Aug 2002 | 111.03746 |
| 50 | AO | 0.06951395 | 3.188600 | Feb 2004 | 107.19870 |
| 88 | TC | -0.07210436 | 3.096986 | Apr 2007 | 93.04338 |
| 124 | AO | -0.09086129 | 3.723054 | Apr 2010 | 91.31444 |
| 148 | AO | -0.06974023 | 3.129660 | Apr 2012 | 93.26361 |
| 160 | LS | 0.08609274 | 3.598692 | Apr 2013 | 108.99074 |

**Figure 31**: outlier detection on the calendar effects adjusted series.

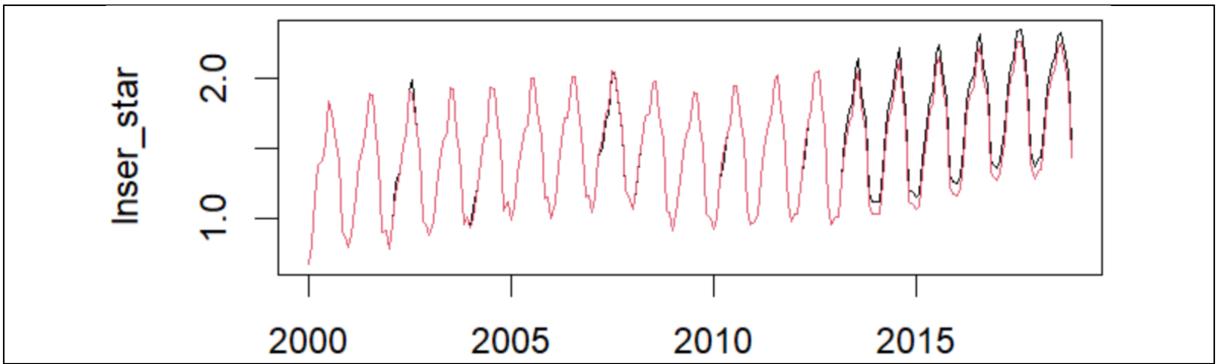


**Figure 32**: series adjusted with calendar effects (black) and series we would expect if outliers had not occurred in the adjusted series for calendar effects (red).

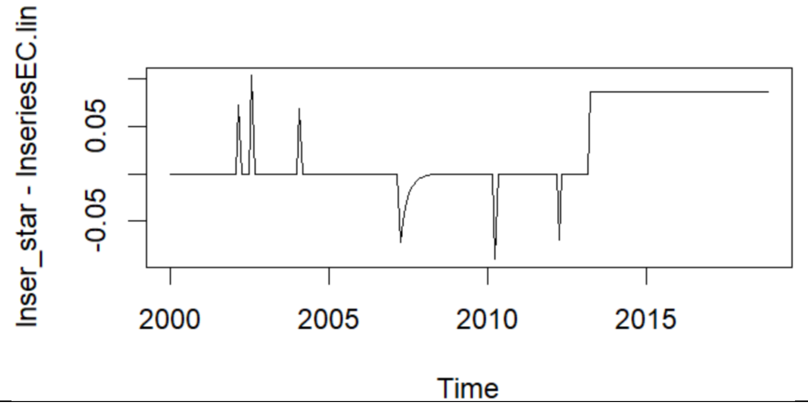


**Figure 33**: impact of outliers over time correction.

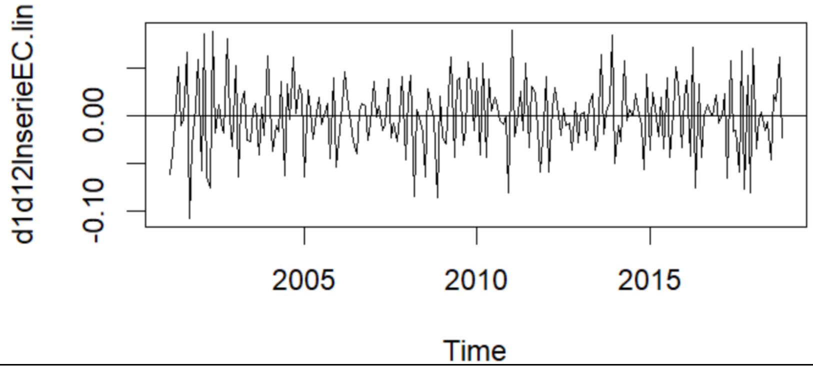


**Figure 34**: differenced linearized series.

```
Call:
arima(x = d1d12lnserieEC.lin, order = c(2, 0, 0), seasonal = list(order = c(1,
    0, 0), period = 12))

Coefficients:
          ar1      ar2     sar1  intercept
      -0.5053  -0.1814  -0.1802    -0.0002
s.e.   0.0678   0.0681   0.0721     0.0011

sigma^2 estimated as 0.001089:  log likelihood = 426.07,  aic = -842.14
                                          intercept
"mod3EaTD.lin : Do not reject H0: Not significant"
```

**Figure 35**: model estimation with the transformed differentiated series.

```
Call:
arima(x = lnseriesEC.lin, order = c(2, 1, 0), seasonal = list(order = c(1, 1,
    0), period = 12))

Coefficients:
          ar1      ar2     sar1
      -0.5052  -0.1814  -0.1803
s.e.   0.0678   0.0681   0.0721

sigma^2 estimated as 0.001089:  log likelihood = 426.05,  aic = -844.11
```

**Figure 36:** mod3EaTD.lin estimation.

| | T_Value |
|---|---|
| | <dbl> |
| ar1 | -7.453109 |
| ar2 | -2.662033 |
| sar1 | -2.499212 |

**Figure 37:** mod3EaTD.lin coefficients t-test.

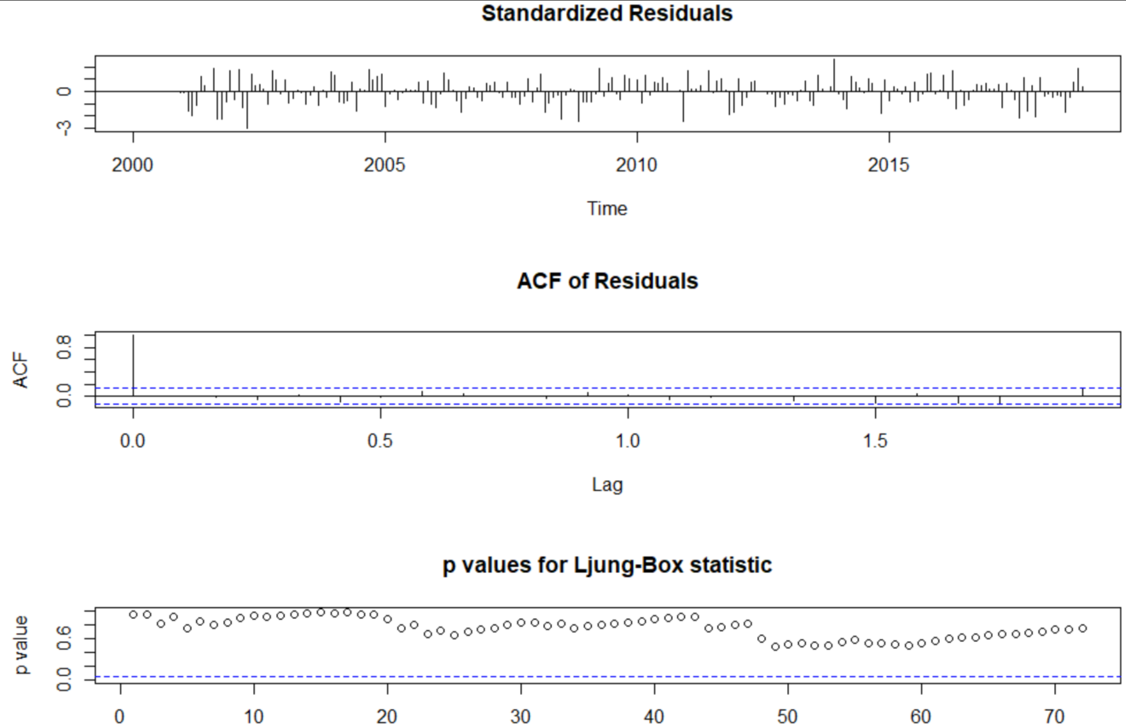


**Figure 38:** mod3EaTD.lin Ljung-Box test.

```
 [1] 1.153457 1.153457 1.153457 1.153457 1.153457 1.153457 1.153457 1.153457 1.153457
1.153457 1.153457 1.153457
[13] 2.347951 2.347951
numeric(0)
```

**Figure 39**: modulus of the polynomial roots (mod3EaTD.lin).