Final Code

Trelan Hakeem-Barron

4/22/2020

Load data:

```
##
## Attaching package: 'TSA'
```

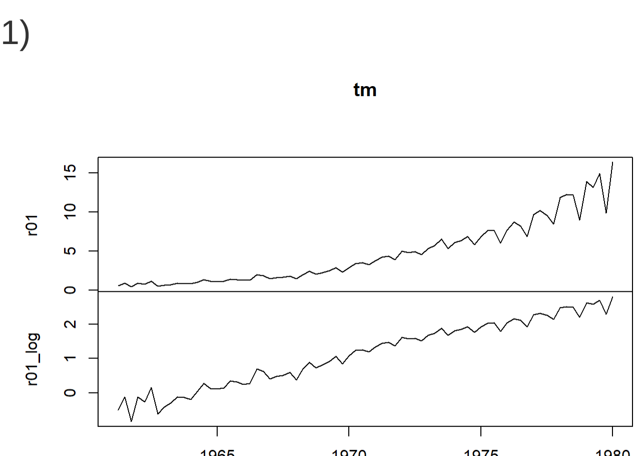
```
## The following objects are masked from 'package:stats':
##
##     acf, arima
```

```
## The following object is masked from 'package:utils':
##
##     tar
```

```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
```

```
## Registered S3 methods overwritten by 'forecast':
##   method             from
##   fitted.Arima       TSA
##   fitted.fracdiff    fracdiff
##   plot.Arima         TSA
##   residuals.fracdiff fracdiff
```

1)

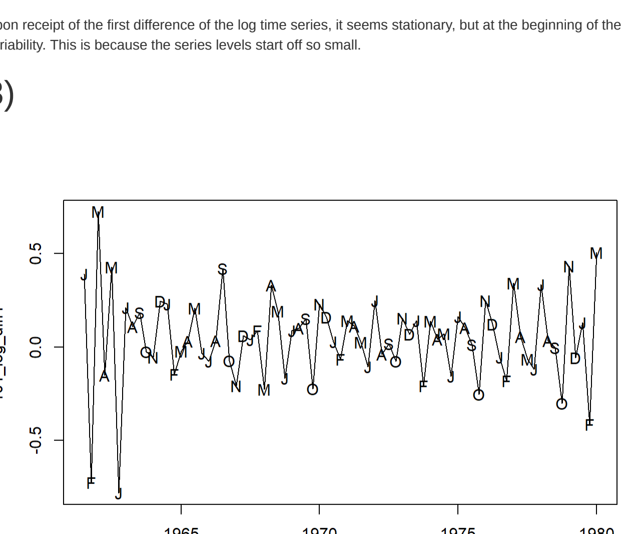


The initial time series is not stationary, since the level of the series increases (does not remain constant in time). The variance of the series also seems to be growing.

For the logarithm of the time series, the variance remains approximately constant in time, but the level of the series increases, which indicates its nonstationarity.

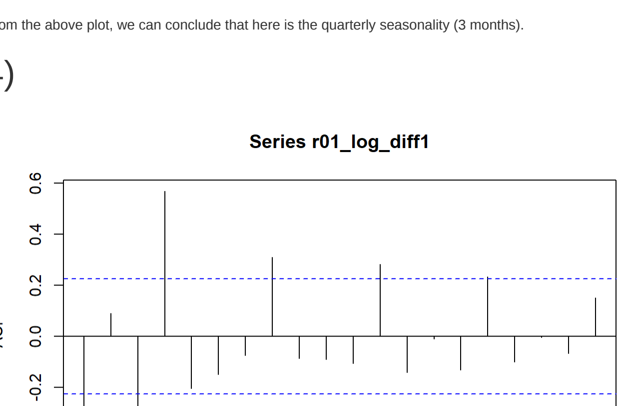
Therefore, the stationary model does not seem reasonable only with a logarithmic transformation.

2)



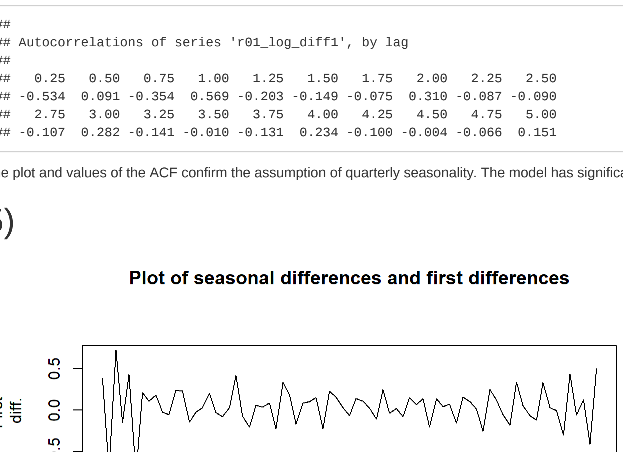
Upon receipt of the first difference of the log time series, it seems stationary, but at the beginning of the series there is something more than normal variability. This is because the series levels start off so small.

3)



From the above plot, we can conclude that here is the quarterly seasonality (3 months).

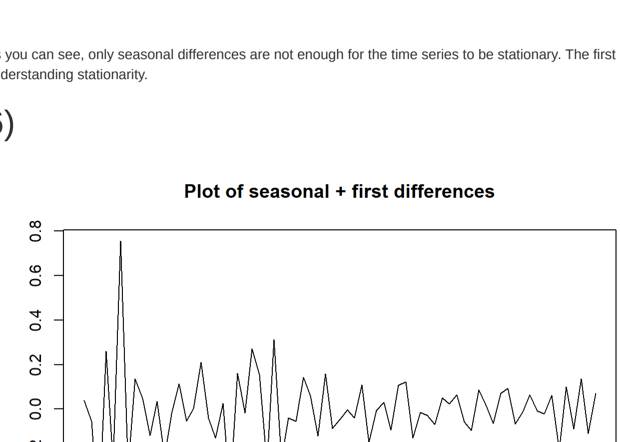
4)




```
##
## Autocorrelations of series 'r01_log_diff1', by lag
##
##   0.25   0.50   0.75   1.00   1.25   1.50   1.75   2.00   2.25   2.50
## -0.534  0.091 -0.354  0.569 -0.203 -0.149 -0.075  0.310 -0.087 -0.090
##   2.75   3.00   3.25   3.50   3.75   4.00   4.25   4.50   4.75   5.00
## -0.107  0.282 -0.141 -0.010 -0.131  0.234 -0.100 -0.004 -0.066  0.151
```

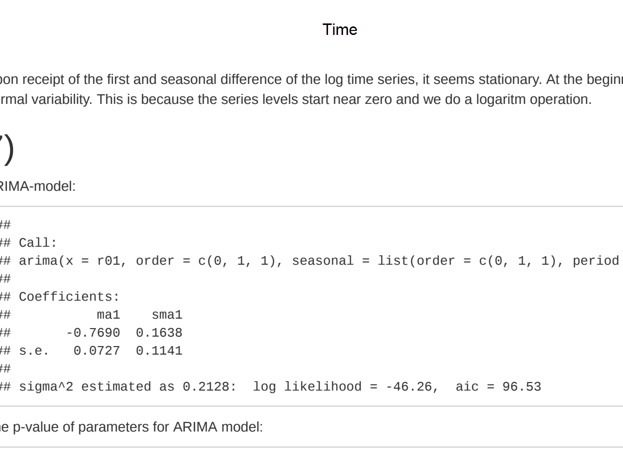
The plot and values of the ACF confirm the assumption of quarterly seasonality. The model has significant lags at 1, 2, 3, and 4.

5)



As you can see, only seasonal differences are not enough for the time series to be stationary. The first differences give the best result in understanding stationarity.

6)



Upon receipt of the first and seasonal difference of the log time series, it seems stationary. At the beginning of the series there is little big than normal variability. This is because the series levels start near zero and we do a logaritm operation.

7)

ARIMA-model:

```
##
## Call:
## arima(x = r01, order = c(0, 1, 1), seasonal = list(order = c(0, 1, 1), period = 4))
##
## Coefficients:
##           ma1    sma1
##       -0.7690  0.1638
## s.e.   0.0727  0.1141
##
## sigma^2 estimated as 0.2128:  log likelihood = -46.26,  aic = 96.53
```

The p-value of parameters for ARIMA model:

```
##       ma1      sma1
## 0.0000000 0.1510686
```

or by the other way:

```
## Loading required package: zoo
```

```
##
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
##
## z test of coefficients:
##
##       Estimate Std. Error  z value Pr(>|z|)
## ma1  -0.769028   0.072737 -10.5727   <2e-16 ***
## sma1  0.163833   0.114108   1.4358   0.1511
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Thus, the coeficient ma1 is significant (its p-value is less than the significance level $\alpha = 0.05$). But the coeficient sma1 is insignificant.

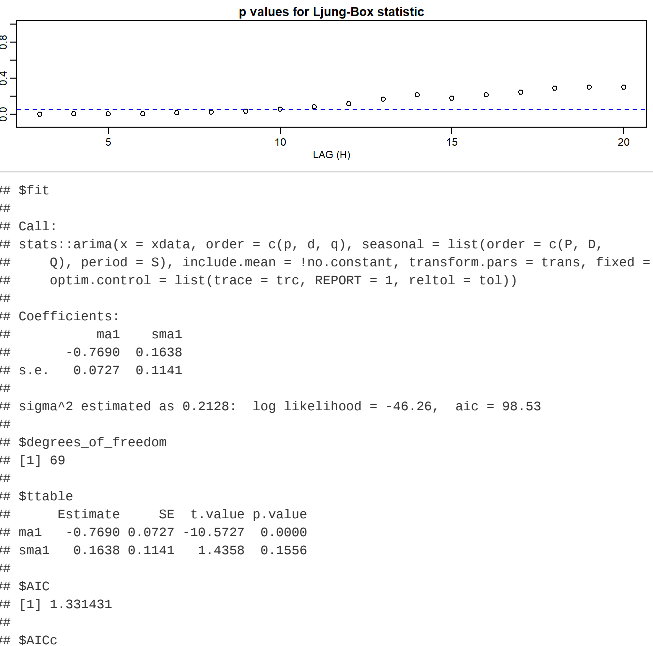
8)

Diagnostic Checking:

```
##
## Attaching package: 'astsa'
```

```
## The following object is masked from 'package:forecast':
##
##     gas
```

```
## initial  value -0.467445
## iter   2 value -0.749998
## iter   3 value -0.764115
## iter   4 value -0.769867
## iter   5 value -0.772012
## iter   6 value -0.773277
## iter   7 value -0.773347
## iter   8 value -0.773363
## iter   9 value -0.773363
## iter   9 value -0.773363
## iter   9 value -0.773363
## final  value -0.773363
## converged
## initial  value -0.767243
## iter   2 value -0.767320
## iter   3 value -0.767347
## iter   4 value -0.767348
## iter   4 value -0.767348
## iter   4 value -0.767348
## final  value -0.767348
## converged
```

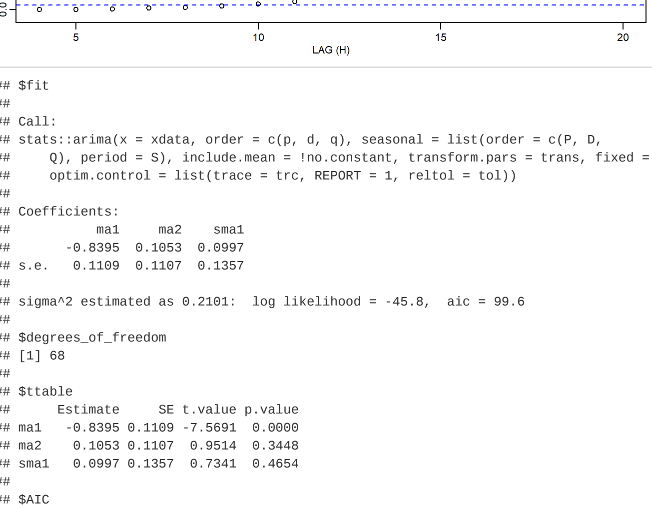



```
## $fit
##
## Call:
## stats::arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D,
##     Q), period = S), include.mean = !no.constant, transform.pars = trans, fixed = fixed,
##     optim.control = list(trace = trc, REPORT = 1, reltol = tol))
##
## Coefficients:
##           ma1    sma1
##       -0.7690  0.1638
## s.e.   0.0727  0.1141
##
## sigma^2 estimated as 0.2128:  log likelihood = -46.26,  aic = 98.53
##
## $degrees_of_freedom
## [1] 69
##
## $ttable
##      Estimate     SE  t.value p.value
## ma1   -0.7690 0.0727 -10.5727  0.0000
## sma1   0.1638 0.1141   1.4358  0.1556
##
## $AIC
## [1] 1.331431
##
## $AICc
## [1] 1.333715
##
## $BIC
## [1] 1.423161
```

From the diagnostic plots, it can be seen that the residues are evenly distributed relative to zero, do not correlate with each other (on the ACF plot, the black lines do not intersect the blue dashed lines), they are almost normally distributed (QQ plot). But on the plot of p-values of the test starting from point 10 we see the dependence in this time series.

9)

```
## initial  value -0.467445
## iter   2 value -0.680237
## iter   3 value -0.760041
## iter   4 value -0.774741
## iter   5 value -0.777812
## iter   6 value -0.779528
## iter   7 value -0.779812
## iter   8 value -0.779827
## iter   9 value -0.779828
## iter  10 value -0.779830
## iter  11 value -0.779830
## iter  11 value -0.779830
## iter  11 value -0.779830
## final  value -0.779830
## converged
## initial  value -0.773773
## iter   2 value -0.773836
## iter   3 value -0.773862
## iter   4 value -0.773866
## iter   5 value -0.773866
## iter   6 value -0.773866
## iter   6 value -0.773866
## iter   6 value -0.773866
## final  value -0.773866
## converged
```

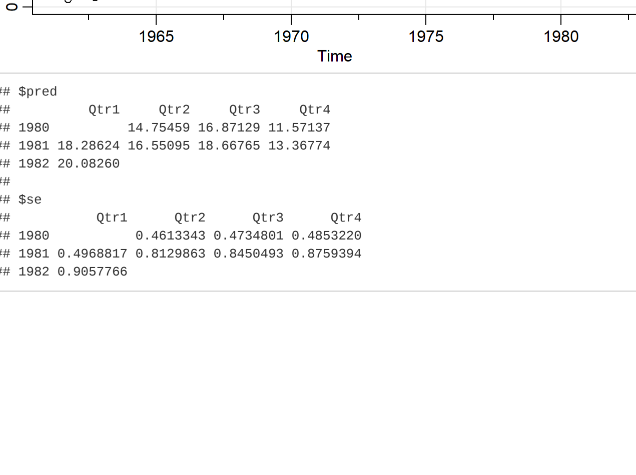



```
## $fit
##
## Call:
## stats::arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D,
##     Q), period = S), include.mean = !no.constant, transform.pars = trans, fixed = fixed,
##     optim.control = list(trace = trc, REPORT = 1, reltol = tol))
##
## Coefficients:
##           ma1     ma2    sma1
##       -0.8395  0.1053  0.0997
## s.e.   0.1109  0.1107  0.1357
##
## sigma^2 estimated as 0.2101:  log likelihood = -45.8,  aic = 99.6
##
## $degrees_of_freedom
## [1] 68
##
## $ttable
##      Estimate     SE t.value p.value
## ma1   -0.8395 0.1109 -7.5691  0.0000
## ma2    0.1053 0.1107  0.9514  0.3448
## sma1   0.0997 0.1357  0.7341  0.4654
##
## $AIC
## [1] 1.345949
##
## $AICc
## [1] 1.350583
##
## $BIC
## [1] 1.468257
```

When we compare these results with those reported in previous paragraph, we see that the estimates of θ_1 and θ have changed very little—especially when the size of the standard errors is taken into consideration. In addition, the estimate of the new parameter, θ_2, is not statistically different from zero. Note also that the estimate, the log-likelihood, and AIC have not changed much.

10)

Forecasts for the next two years of the series:




```
## $pred
##          Qtr1     Qtr2     Qtr3     Qtr4
## 1980          14.75459 16.87129 11.57137
## 1981 18.28624 16.55095 18.66765 13.36774
## 1982 20.08260
##
## $se
##           Qtr1      Qtr2      Qtr3      Qtr4
## 1980           0.4613343 0.4734801 0.4853220
## 1981 0.4968817 0.8129863 0.8450493 0.8759394
## 1982 0.9057766
```