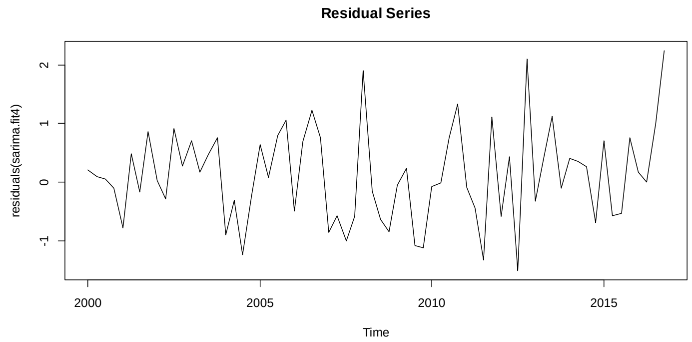
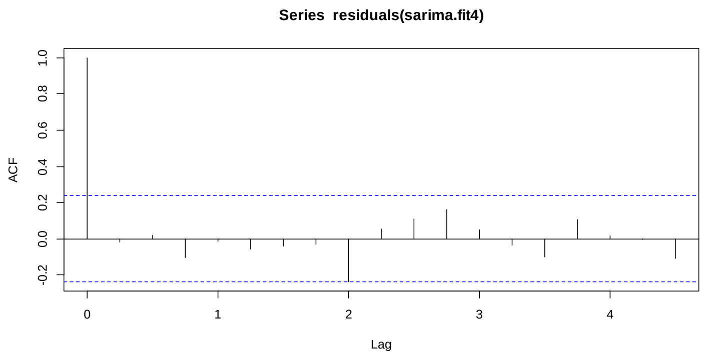
STATS 326
Applied Time Series
ASSIGNMENT FIVE
ANSWER GUIDE

Question One:

```
> sarima.fit4 = arima(red.CO2.ts,order=c(1,1,1),
  seasonal=list(order=c(0,1,1),period=4))
> sarima.fit4

Call:
arima(x = red.CO2.ts, order = c(1, 1, 1), seasonal = list(order =
c(0, 1, 1), period = 4))

Coefficients:
         ar1      ma1     sma1
      0.4830  -0.9199  -0.7776
s.e.  0.1471   0.0977   0.1024

sigma^2 estimated as 0.6873:  log likelihood = -80.69,  aic = 169.37

> plot.ts(residuals(sarima.fit4),main="Residual Series")
```




```
> acf(residuals(sarima.fit4))
```



The Residual Series shows reasonably random scatter about 0, although there are 3 large positive residuals (2008.1, 2012.4 and 2016.4). The autocorrelation function plot of the residuals shows no significant lags. The assumptions appear to be satisfied.

```
> SARIMA4.pred = predict(sarima.fit4,n.ahead=4)
> SARIMA4.pred
$pred
         Qtr1     Qtr2     Qtr3     Qtr4
2017 412.1403 411.4029 398.1207 408.2079

$se
          Qtr1      Qtr2      Qtr3      Qtr4
2017 0.8292973 0.9518559 0.9957232 1.0171728

> actual
2017.1 2017.2 2017.3 2017.4
413.75 412.42 398.47 408.44

> RMSEP.SARIMA4 = sqrt(1/4*sum((actual-SARIMA4.pred$pred)^2))
> RMSEP.SARIMA4
[1] 0.974876
```

The SARIMA$(1,1,1)\times(0,1,1)_4$ model had an RMSEP (0.97 ppm). The best predicting model from previous assignments was the Holt-Winters model with RMSEP of 0.96 ppm.

Question Two:

```
> sarima.fit4.full = arima(full.CO2.ts,order=c(1,1,1),
  seasonal=list(order=c(0,1,1),period=4))
> sarima.fit4.full

Call:
arima(x = full.CO2.ts, order = c(1, 1, 1), seasonal = list(order =
c(0, 1, 1),
    period = 4))

Coefficients:
         ar1      ma1     sma1
       0.549  -0.9153  -0.7732
s.e.   0.139   0.0863   0.0940

sigma^2 estimated as 0.6855:  log likelihood = -85.34,  aic = 178.68
```

$(1-\rho_1 \mathrm{B})(1-\mathrm{B})(1-\mathrm{B}^4)y_t = (1+\alpha_1 \mathrm{B})(1+\mathrm{A}_1 \mathrm{B}^4)\varepsilon_t$
$(1-\rho_1 \mathrm{B})(1-\mathrm{B}-\mathrm{B}^4+\mathrm{B}^5)y_t = (1+\alpha_1 \mathrm{B}+\mathrm{A}_1 \mathrm{B}^4+\alpha_1 \mathrm{A}_1 \mathrm{B}^5)\varepsilon_t$
$(1-\mathrm{B}-\mathrm{B}^4+\mathrm{B}^5-\rho_1 \mathrm{B}+\rho_1 \mathrm{B}^2+\rho_1 \mathrm{B}^5-\rho_1 \mathrm{B}^6)y_t = (1+\alpha_1 \mathrm{B}+\mathrm{A}_1 \mathrm{B}^4+\alpha_1 \mathrm{A}_1 \mathrm{B}^5)\varepsilon_t$
$y_t-(1+\rho_1)y_{t-1}+\rho_1 y_{t-2}-y_{t-4}+(1+\rho_1)y_{t-5}-\rho_1 y_{t-6} = \varepsilon_t+\alpha_1\varepsilon_{t-1}+\mathrm{A}_1\varepsilon_{t-4}+\alpha_1 \mathrm{A}_1\varepsilon_{t-5}$
$y_t = (1+\rho_1)y_{t-1}-\rho_1 y_{t-2}+y_{t-4}-(1+\rho_1)y_{t-5}+\rho_1 y_{t-6}+\varepsilon_t+\alpha_1\varepsilon_{t-1}+\mathrm{A}_1\varepsilon_{t-4}+\alpha_1 \mathrm{A}_1\varepsilon_{t-5}$
$y_t = 1.549y_{t-1}-0.549y_{t-2}+y_{t-4}-1.549y_{t-5}+0.549y_{t-6}+\varepsilon_t-0.9153\varepsilon_{t-1}-0.7732\varepsilon_{t-4}$
$\quad +0.70770996\varepsilon_{t-5}$

```
> full.CO2.ts
       Qtr1   Qtr2   Qtr3   Qtr4
2000 375.33 375.21 362.52 370.42
……
2015 406.55 405.82 392.31 402.96
2016 408.63 408.35 396.07 407.67
2017 413.75 412.42 398.47 408.44

> residuals(sarima.fit4.full)
             Qtr1          Qtr2          Qtr3          Qtr4
2000  0.2166968053  0.0968167496  0.0526563888 -0.0997203630
……
2015  0.7310728098 -0.6086097361 -0.5093977597  0.7965998392
2016  0.1372648171 -0.0279754929  0.9936960060  2.1601285641
2017  1.4076001116 -0.1055204121 -0.4018711569 -0.1136896868
```

$y_{t+1} = 1.549y_t-0.549y_{t-1}+y_{t-3}-1.549y_{t-4}+0.549y_{t-5}+\varepsilon_{t+1}-0.9153\varepsilon_t-0.7732\varepsilon_{t-3}$
$\quad +0.70770996\varepsilon_{t-4}$

```
> (1.549*408.44)-(0.549*398.47)+413.75-
(1.549*407.67)+(0.549*396.07)+(0.9153*0.1136896868)-
(0.7732*1.407600116)+(0.70770996*2.1601285641)
[1] 414.1696 (= 414.17 ppm)
```

$y_{t+2} = 1.549y_{t+1}-0.549y_t+y_{t-2}-1.549y_{t-3}+0.549y_{t-4}+\varepsilon_{t+2}-0.9153\varepsilon_{t+1}-0.7732\varepsilon_{t-2}$
$\quad +0.70770996\varepsilon_{t-3}$

```
> (1.549*414.1696)-(0.549*408.44)+412.42-
(1.549*413.75)+(0.549*407.67)+(0.7732*0.1055204121)+
(0.70770996*1.407600116)
[1] 413.725 (= 413.73 ppm)
```

$y_{t+3} = 1.549y_{t+2}-0.549y_{t+1}+y_{t-1}-1.549y_{t-2}+0.549y_{t-3}+\varepsilon_{t+3}-0.9153\varepsilon_{t+2}-0.7732\varepsilon_{t-1}$
$\quad +0.70770996\varepsilon_{t-2}$

```
> (1.549*413.725)-(0.549*414.1696)+398.47-
(1.549*412.42)+(0.549*413.75)+(0.7732*0.4018711569)-
(0.70770996*0.1055204121)
[1] 400.4971 ( = 400.50 ppm)
```

$y_{t+4} = 1.549y_{t+3}-0.549y_{t+2}+y_t-1.549y_{t-1}+0.549y_{t-2}+\varepsilon_{t+4}-0.9153\varepsilon_{t+3}-0.7732\varepsilon_t$
$\quad +0.70770996\varepsilon_{t-1}$

```
> (1.549*400.4971)-(0.549*413.725)+408.44-
(1.549*398.47)+(0.549*412.42)+(0.7732*0.1136896868)-
(0.70770996*0.4018711569)
[1] 410.667 ( = 410.67 ppm)
```

Question Three:

My brief was to predict the atmospheric concentration of carbon dioxide at Barrow in Alaska (in parts per million) for 2018.

We need to be a little careful with our predictions and their reliability as we have a Time Series with only 72 observations. However the model that is used is a good model so the predictions should be reasonably reliable.

I built several different models using the first seventeen years of data (2000 – 2016) and used each model to predict the 4 quarters of 2017. Each model's predictions were compared with the actual 2017 values to find the model that produced the most accurate predictions.

Once the best predicting model was found, it was re-run with all the data (2000 – 2017) and predictions done for 2018, as requested.

I predict the carbon dioxide concentration in the atmosphere for 2018 at Barrow in Alaska will be:

414.03 ppm for Quarter 1, 2018
413.42 ppm for Quarter 2, 2018
400.36 ppm for Quarter 3, 2018
410.63 ppm for Quarter 4, 2018