STAT 443: Assignment 3

Wenxuan Zan (61336194)

21 March, 2022

Question 1

Question 2

Question 3

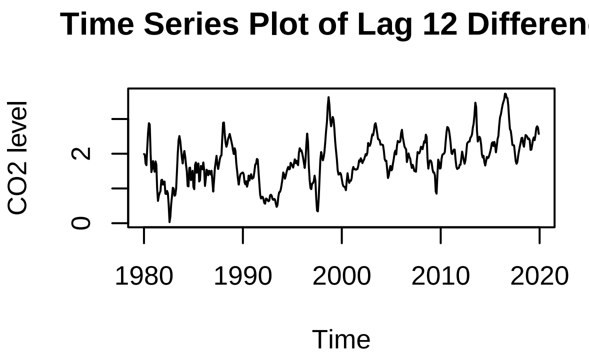
a)

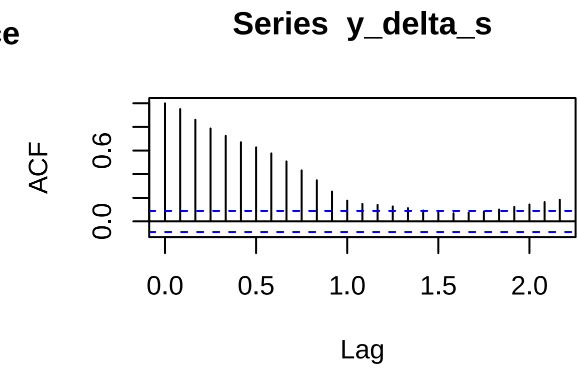
```
q3data <- read.csv("co2_mm_gl.csv", header = TRUE, skip = 55)
co2ts <- ts(q3data[,4], start = c(1979,1), frequency = 12)
training <- window(co2ts,start = c(1979,1),end = c(2019,12),frequency=12)
testing <- window(co2ts,start = c(2020,1),end = c(2022,10),frequency=12)
```

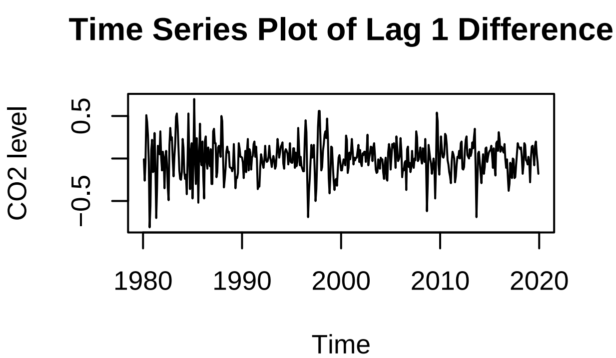
b)

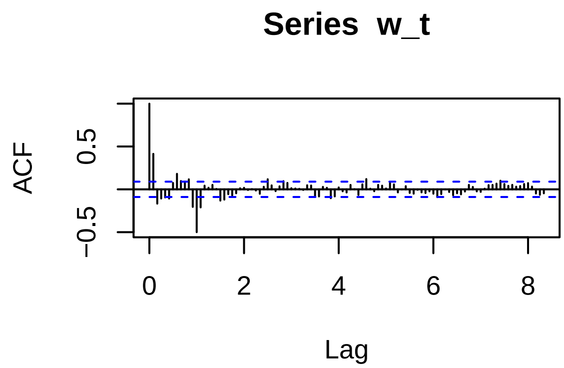
```
par(mfrow = c(2,2))
s=12
y_delta_s = diff(training, lag = s, difference = 1)
plot(y_delta_s,
     ylab = "CO2 level",
     main = "Time Series Plot of Lag 12 Difference")
acf(y_delta_s)

#
w_t = diff(y_delta_s, lag = 1, difference = 1)
plot(w_t,
     ylab = "CO2 level",
     main = "Time Series Plot of Lag 1 Difference")
acf(w_t, lag.max = 100)
```









i) Using $s = 12$, the time series plot indicates an upward trend, and is now void of seasonal variation. The ACF plot has a slow exponential decay which reflects the positive temporal dependence observed in the differenced series.

ii) Looking at the time series plot of lag 12 difference, the time series stil possesses a upward trend. Therefore to remove the trend component, we difference the time series again at lag 1.

iii)

$$d = 1, \;\; D = 1, \;\; s = 12$$

iv)

$$p = 0, \;\; q = 1, \;\; P = 0, \;\; Q = 0$$

v)

```
p = 0
P = 0
q = c(0:5)
Q = c(0:5)
d = 1
D = 1
s = 12
sequence = c(rep(0,6),rep(1,6),rep(2,6),rep(3,6),rep(4,6),rep(5,6))
mods <- data.frame(p = rep(0,36),
                   d = rep(1,36),
                   q = sequence,
                   P = rep(0,36),
                   D = rep(1,36),
                   Q = rep(0:5,6),
```

```r
                AIC = rep(0,36))
for (i in 1:36) {
  q = mods$q[i]
  Q = mods$Q[i]
  mod <- arima(training,
               order=c(p, d, q),
               seasonal=list(order=c(P, D, Q),period=s))
  mods$AIC[i] <- mod$aic
}
print(mods[which.min(mods$AIC),])
```

```
##    p d q P D Q       AIC
## 20 0 1 3 0 1 1 -681.2112
```

By compute AIC for all combinations of $q \in \{0, ..., 5\}$ and $Q \in \{0, ..., 5\}$, It appears the

$$SARIMA(0, 1, 3) \times (0, 1, 1)_{12}$$

has the lowest AIC, $AIC = -681.21$.