```r
# Load the required packages
require(ggplot2)
require(tidyverse)
require(MASS)
require(mgcv)
require(dplyr)
require(magrittr)
require(factoextra)
require(reshape2)
require(knitr)

require(mnormt)
require(readr)
require(sf)
require(tmap)
require(geoR)
require(maptools)
require(gstat)

require(forecast)
require(lubridate)
require(dlm)
```

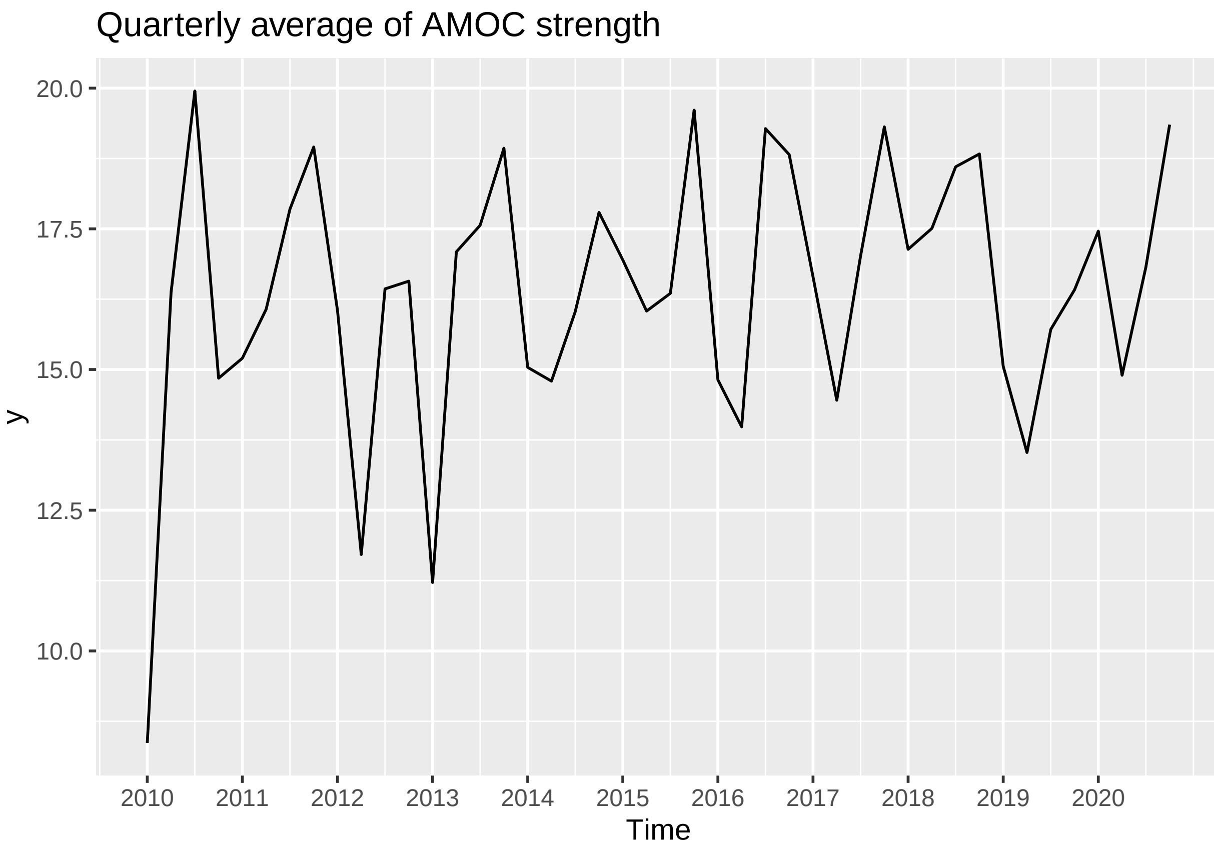
Question 2

```r
amoc <- read_csv("AMOCdata.csv")

# Convert Date column to Date format
amoc$Date <- as.Date(amoc$Date, format = "%d/%m/%Y")
```

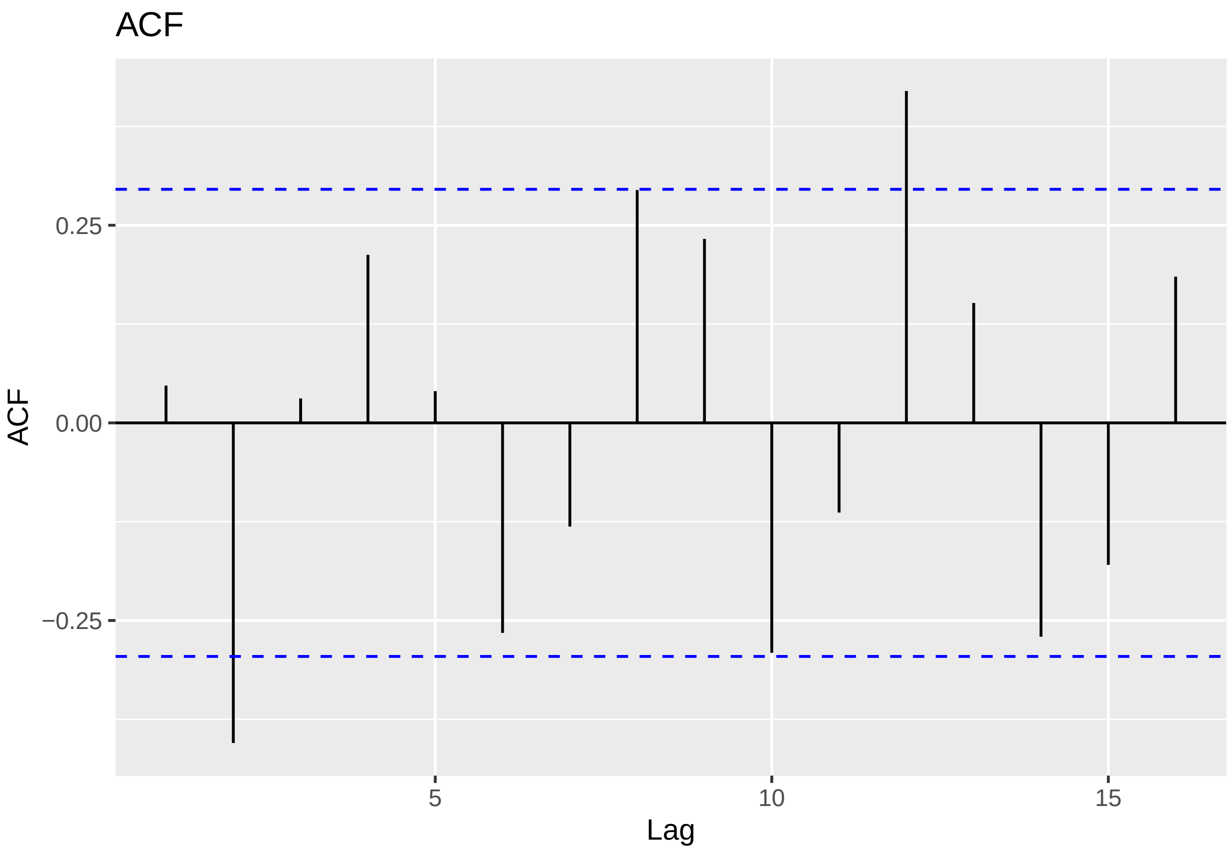
2a

Average the data to quarterly means and plot the quarterly average

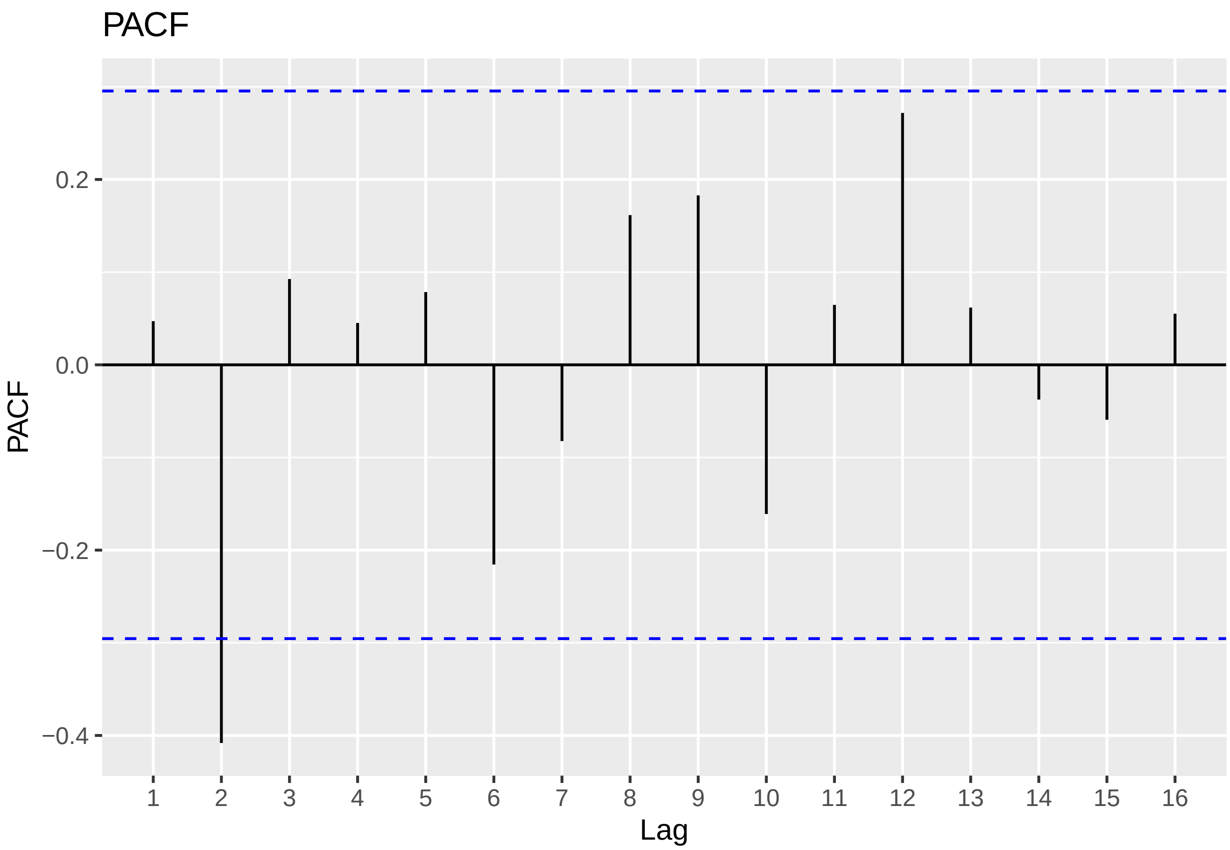
```r
# Compute the Quarterly mean
amoc_qtr <- aggregate(amoc$Strength, by = list(Year = amoc$Year,
    Quarter = amoc$Quarter), mean)

# Arrange the Quarterly mean table
names(amoc_qtr)[3] <- "Strength"
amoc_qtr <- amoc_qtr %>%
    arrange(Year, Quarter)

# Convert to time-series object
amoc_qtr_ts <- ts(amoc_qtr$Strength)

ggplot(data = data.frame(Time = 1:length(amoc_qtr_ts), y = as.numeric(amoc_qtr_ts)),
    aes(x = Time, y = y)) + geom_line() + scale_x_continuous(breaks = seq(1,
    length(amoc_qtr_ts), by = 4), labels = amoc_qtr$Year[seq(1,
    nrow(amoc_qtr), by = 4)]) + ggtitle("Quarterly average of AMOC strength")
```




```
ggAcf(amoc_qtr_ts) + ggtitle("ACF")
```




```
ggPacf(amoc_qtr_ts) + ggtitle("PACF")
```



The quarterly average looks stationary with a constant mean (around 16.5). There is no strongly linear trend like decreasing or increasing trend. There can be a seasonal pattern, that the strength is low in the 1st and 2nd quarters every year, while peaking in the 3rd and 4th quarters. The seasonal variation can be decreasing over time (as in 2019 the difference between peak and bottom is much less than 5 while in 2012 and early years, it was around 5)

It can be seen from the three plots that ACF and PACF are not decaying from lag 1 but quite fluctuating suddenly overtime. Both ACF and PACF can be cutoff at lag 2 but lag 1 is nearly 0. This can be due to the lack of data points (we only have 44 points after quarterly averaging).

We might try fitting multiple possible models to see which performs best.

2b

As there is no difference needed, firstly, we will fit an ARMA model without a seasonal term. We will try possible models with order 2, 3 and 4, and if there is no sensible fit, we will raise the order later.

```r
set.seed(1234)
# order 1
model_d0_a1 <- Arima(amoc_qtr_ts, order = c(1, 0, 0))
model_d0_m1 <- Arima(amoc_qtr_ts, order = c(0, 0, 1))

# order 2
model_d0_a1m1 <- Arima(amoc_qtr_ts, order = c(1, 0, 1))
model_d0_a2m0 <- Arima(amoc_qtr_ts, order = c(2, 0, 0))
model_d0_a0m2 <- Arima(amoc_qtr_ts, order = c(0, 0, 2))

# order 3
model_d0_a1m2 <- Arima(amoc_qtr_ts, order = c(1, 0, 2))
```

```r
model_d0_a2m1 <- Arima(amoc_qtr_ts, order = c(2, 0, 1))
model_d0_a3m0 <- Arima(amoc_qtr_ts, order = c(3, 0, 0))
model_d0_a0m3 <- Arima(amoc_qtr_ts, order = c(0, 0, 3))

# order 4
model_d0_a2m2 <- Arima(amoc_qtr_ts, order = c(2, 0, 2))
model_d0_a0m4 <- Arima(amoc_qtr_ts, order = c(0, 0, 4))
model_d0_a4m0 <- Arima(amoc_qtr_ts, order = c(4, 0, 0))
model_d0_a1m3 <- Arima(amoc_qtr_ts, order = c(1, 0, 3))
model_d0_a3m1 <- Arima(amoc_qtr_ts, order = c(3, 0, 1))

# check model statistics
model_d0_a1
```

```
## Series: amoc_qtr_ts
## ARIMA(1,0,0) with non-zero mean
##
## Coefficients:
##          ar1     mean
##       0.0665  16.3878
## s.e.  0.1788   0.3726
##
## sigma^2 = 5.572:  log likelihood = -99.2
## AIC=204.41   AICc=205.01   BIC=209.76
```

```r
model_d0_m1
```

```
## Series: amoc_qtr_ts
## ARIMA(0,0,1) with non-zero mean
##
## Coefficients:
##          ma1     mean
##       0.5108  16.3418
## s.e.  0.1771   0.5060
##
## sigma^2 = 5.245:  log likelihood = -98.02
## AIC=202.05   AICc=202.65   BIC=207.4
```

```r
model_d0_a1m1
```

```
## Series: amoc_qtr_ts
## ARIMA(1,0,1) with non-zero mean
##
## Coefficients:
##           ar1     ma1     mean
##       -0.4204  0.7718  16.3721
## s.e.   0.2390  0.1466   0.4067
##
## sigma^2 = 5.045:  log likelihood = -96.64
## AIC=201.29   AICc=202.31   BIC=208.43
```

```r
model_d0_a2m0
```

```
## Series: amoc_qtr_ts
## ARIMA(2,0,0) with non-zero mean
##
```

```
## Coefficients:
##          ar1      ar2     mean
##       0.0990  -0.5565  16.4298
## s.e.  0.1576   0.1488   0.2113
##
## sigma^2 = 4.321:  log likelihood = -93.45
## AIC=194.9   AICc=195.92   BIC=202.04
```

```
model_d0_a0m2
```

```
## Series: amoc_qtr_ts
## ARIMA(0,0,2) with non-zero mean
##
## Coefficients:
##          ma1      ma2     mean
##       0.1426  -0.4450  16.4287
## s.e.  0.1386   0.1277   0.2230
##
## sigma^2 = 4.538:  log likelihood = -94.41
## AIC=196.82   AICc=197.84   BIC=203.95
```

```
model_d0_a1m2
```

```
## Series: amoc_qtr_ts
## ARIMA(1,0,2) with non-zero mean
##
## Coefficients:
##          ar1     ma1      ma2     mean
##       0.0230  0.1275  -0.4485  16.4289
## s.e.  0.3051  0.2420   0.1348   0.2224
##
## sigma^2 = 4.651:  log likelihood = -94.41
## AIC=198.81   AICc=200.39   BIC=207.73
```

```
model_d0_a2m1
```

```
## Series: amoc_qtr_ts
## ARIMA(2,0,1) with non-zero mean
##
## Coefficients:
##           ar1      ar2     ma1     mean
##       -0.0669  -0.5475  0.2187  16.4255
## s.e.   0.2740   0.1555  0.2883   0.2300
##
## sigma^2 = 4.366:  log likelihood = -93.15
## AIC=196.31   AICc=197.88   BIC=205.23
```

```
model_d0_a3m0
```

```
## Series: amoc_qtr_ts
## ARIMA(3,0,0) with non-zero mean
##
## Coefficients:
##          ar1      ar2     ar3     mean
##       0.1626  -0.5690  0.1464  16.4227
## s.e.  0.1729   0.1479  0.1708   0.2409
##
```

```
## sigma^2 = 4.35:  log likelihood = -93.09
## AIC=196.17   AICc=197.75   BIC=205.1
```

```
model_d0_a0m3
```

```
## Series: amoc_qtr_ts
## ARIMA(0,0,3) with non-zero mean
##
## Coefficients:
##          ma1      ma2      ma3     mean
##       0.1650  -0.4453  -0.0288  16.4293
## s.e.  0.2224   0.1279   0.2254   0.2214
##
## sigma^2 = 4.649:  log likelihood = -94.4
## AIC=198.8   AICc=200.38   BIC=207.72
```

```
model_d0_a2m2
```

```
## Series: amoc_qtr_ts
## ARIMA(2,0,2) with non-zero mean
##
## Coefficients:
##          ar1      ar2      ma1     ma2     mean
##       0.0787  -0.9982  -0.0254  0.9998  16.4015
## s.e.  0.0285   0.0067   0.0900  0.1155   0.2684
##
## sigma^2 = 3.379:  log likelihood = -89.46
## AIC=190.91   AICc=193.18   BIC=201.62
```

```
model_d0_a0m4
```

```
## Series: amoc_qtr_ts
## ARIMA(0,0,4) with non-zero mean
##
## Coefficients:
##          ma1      ma2      ma3     ma4     mean
##       0.1533  -0.5189  -0.0073  0.1528  16.4197
## s.e.  0.1667   0.1686   0.1748  0.1377   0.2416
##
## sigma^2 = 4.63:  log likelihood = -93.8
## AIC=199.6   AICc=201.87   BIC=210.3
```

```
model_d0_a4m0
```

```
## Series: amoc_qtr_ts
## ARIMA(4,0,0) with non-zero mean
##
## Coefficients:
##          ar1      ar2     ar3     ar4     mean
##       0.1597  -0.5376  0.1355  0.0684  16.4182
## s.e.  0.1727   0.1665  0.1725  0.1716   0.2570
##
## sigma^2 = 4.441:  log likelihood = -93.01
## AIC=198.02   AICc=200.29   BIC=208.72
```

```
model_d0_a1m3
```

```
## Series: amoc_qtr_ts
```

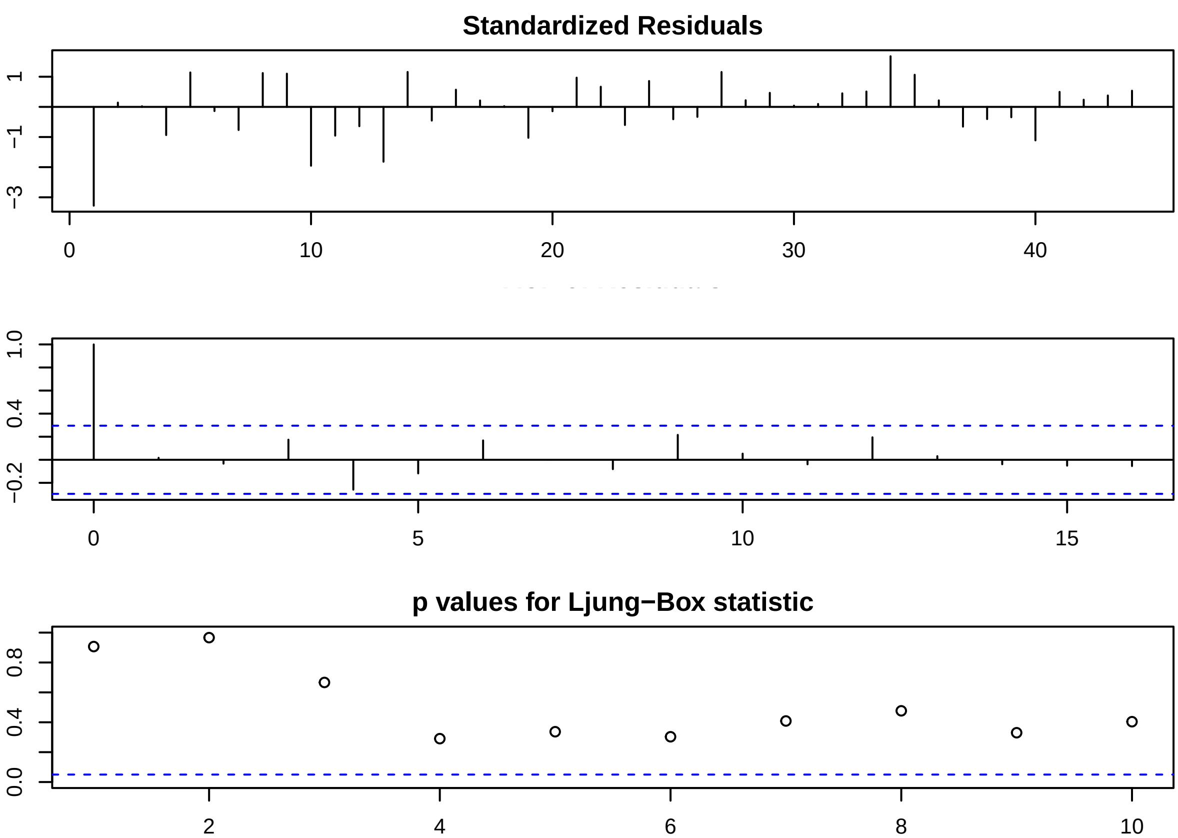
```
## ARIMA(1,0,3) with non-zero mean
##
## Coefficients:
##           ar1     ma1      ma2      ma3     mean
##       -0.5284  0.7214  -0.3646  -0.3072  16.4299
## s.e.   0.9228  0.8649   0.2077   0.3545   0.2195
##
## sigma^2 = 4.733:  log likelihood = -94.25
## AIC=200.5   AICc=202.77   BIC=211.21
```

```
model_d0_a3m1
```

```
## Series: amoc_qtr_ts
## ARIMA(3,0,1) with non-zero mean
##
## Coefficients:
##          ar1      ar2     ar3      ma1     mean
##       0.4092  -0.5931  0.2864  -0.2467  16.4191
## s.e.  0.6330   0.1651  0.3580   0.6291   0.2537
##
## sigma^2 = 4.449:  log likelihood = -93.03
## AIC=198.06   AICc=200.34   BIC=208.77
```

From the statistics, the model (2,0,2) (named as a2m2) seems the best with highest log-likelihood (-89) and lowest AIC (190). We will check the residuals of this model, and compare with the result from auto.arima() to see if it's the best.

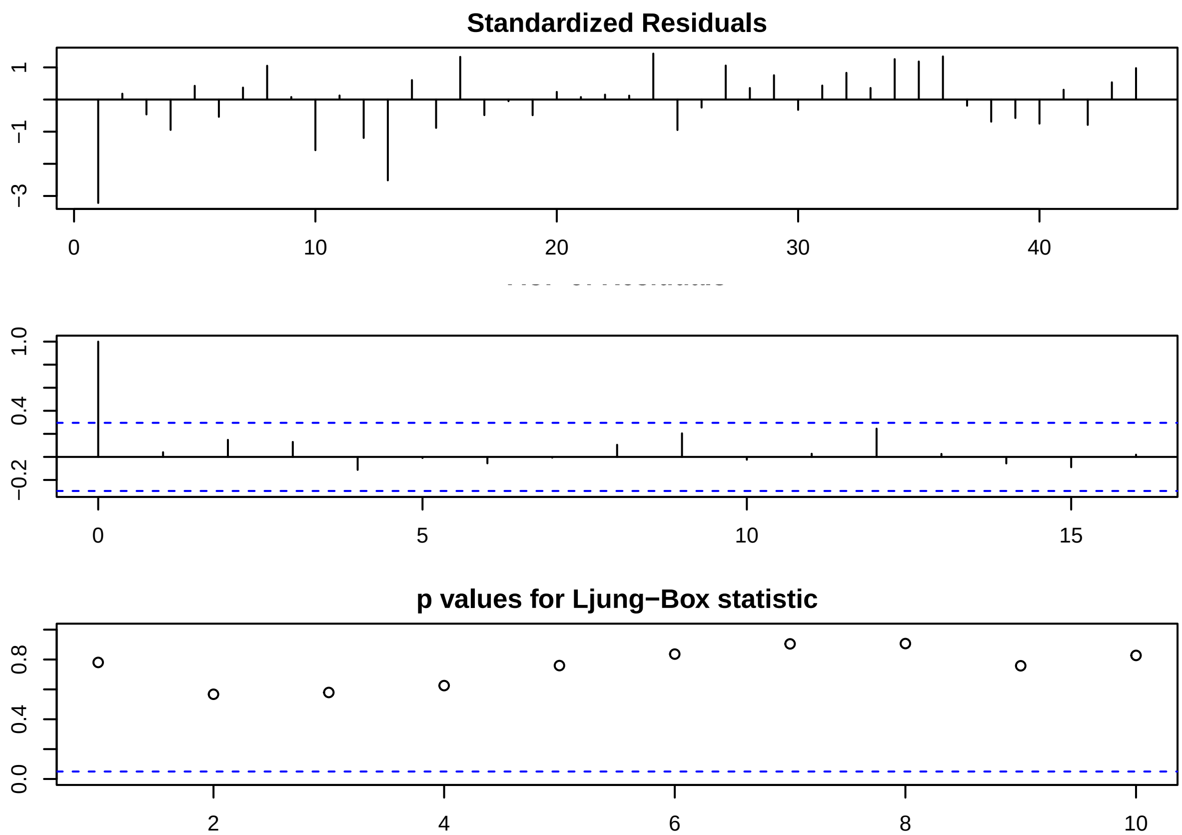
```
par(mar = c(3, 3, 2, 2))
tsdiag(model_d0_a2m2)
```



```r
model_auto_d0 <- auto.arima(amoc_qtr_ts)
model_auto_d0
```

```
## Series: amoc_qtr_ts
## ARIMA(2,0,0) with non-zero mean
##
## Coefficients:
##          ar1      ar2     mean
##       0.0990  -0.5565  16.4298
## s.e.  0.1576   0.1488   0.2113
##
## sigma^2 = 4.321:  log likelihood = -93.45
## AIC=194.9   AICc=195.92   BIC=202.04
```

```r
par(mar = c(3, 3, 2, 2))
tsdiag(model_auto_d0)
```



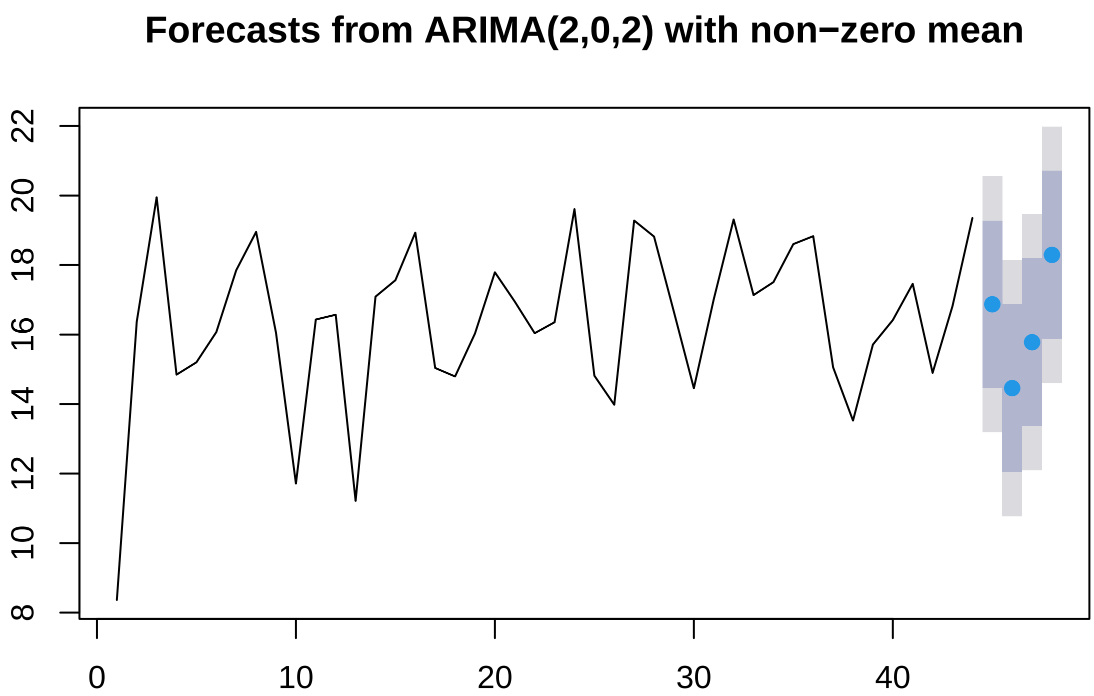
The auto.arima returns (2, 0, 0) model - named as a2m0 as the best model based on AICc (by default). However, it can be seen that the AICc of a2m2 (193) is slightly lower than that of a2m0 (195). Thus, in terms of AIC, AICc or log likelihood, a2m2 seems better, despite that it has higher order (which has been added into AIC).

The variance of a2m2 is 3.4, much lower than that of the auto a2m0 (4.3).

We can see that the residuals from model a2m2 looks quite random, but it seems to overestimate or underestimate the strength in periods. The positive or negative residuals tend to be grouped of around 3 or 4 consecutive quarters, especially in the last 3 years. This pattern of residuals can be seen in the auto ARIMA model. We will add seasonal term later, so now just compare the two models on other characteristics.

ACF plots and p-values of a2m0 look acceptable - most p values are larger than 0.05, thus we have evidence to conclude that there is no evidence of autocorrelation. The p values of a2m2 are higher than 0.05 up to lag 4. As we already know that there is a seasonal component that we did not incorporate into the model, we can make a good guess that the seasonal cycle is 4 quarter. Thus, although a2m2 has higher order and more complex, we will choose a2m2 as the best model to forecast.

```
# forecast for the next 4 quarters
pred_4q <- forecast(model_d0_a2m2, 4)
plot(pred_4q)
```



2c

We will model AMOC strength using a Dynamic Linear Model with a seasonal component. From the above part, we see that the mean is quite constant. However, the cyclic trend seems funneling.

We will set up the order of dlmModPoly = 1, frequency of dlmModSeas = 4. The plot below shows the decomposition of components of dlm models. As can be seen, the trend looks flat-off above 15 and the seasonal cycles are slightly funneling.

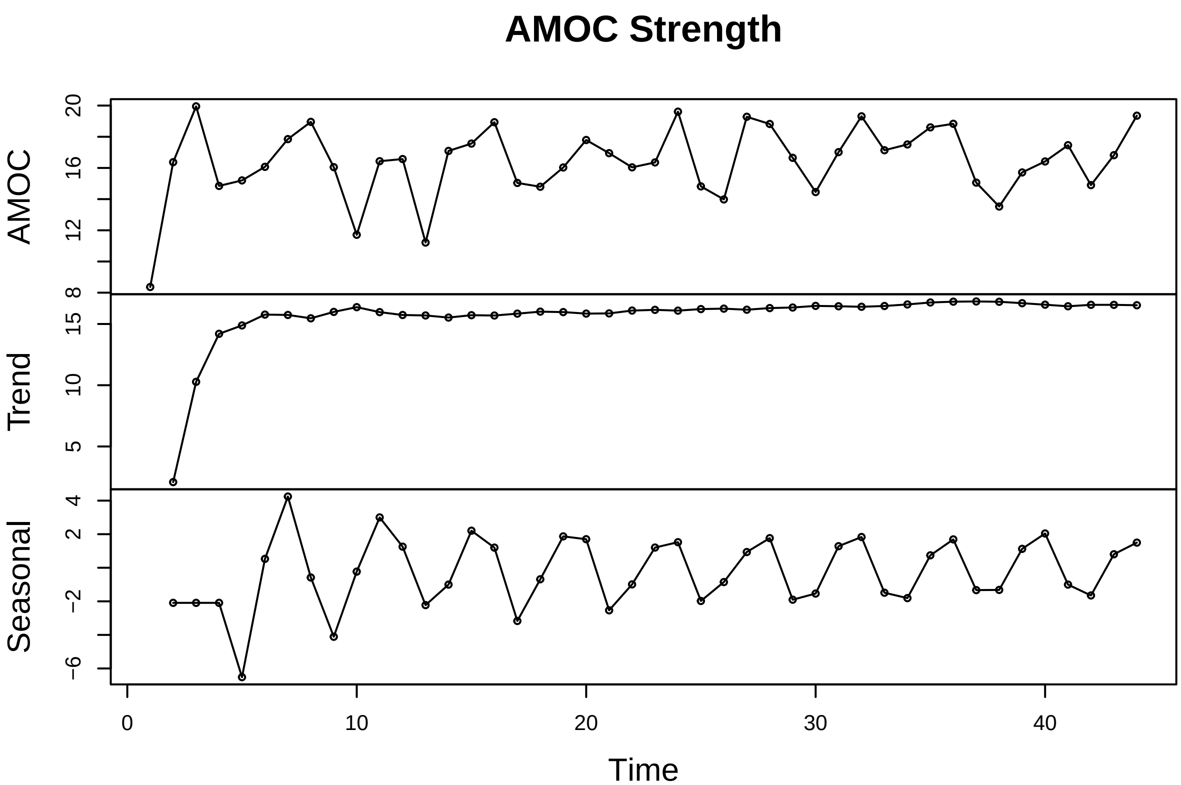
We will use this dlm to forecast the next 4 quarters.

```
# Build DLM
buildFun <- function(x) {
    dlmModPoly(order = 1, dV = exp(x[1]), dW = exp(x[2])) + dlmModSeas(frequency = 4,
        dV = 0, dW = c(exp(x[3]), rep(0, 2)))
}

# Fit model
fit <- dlmMLE(amoc_qtr_ts, parm = c(0, 0, 0), build = buildFun)
fitted_model <- buildFun(fit$par)

# calculate the hidden stats
pred_dlm <- dlmFilter(amoc_qtr_ts, mod = fitted_model)
summary(pred_dlm)
```

```
##      Length Class  Mode
```

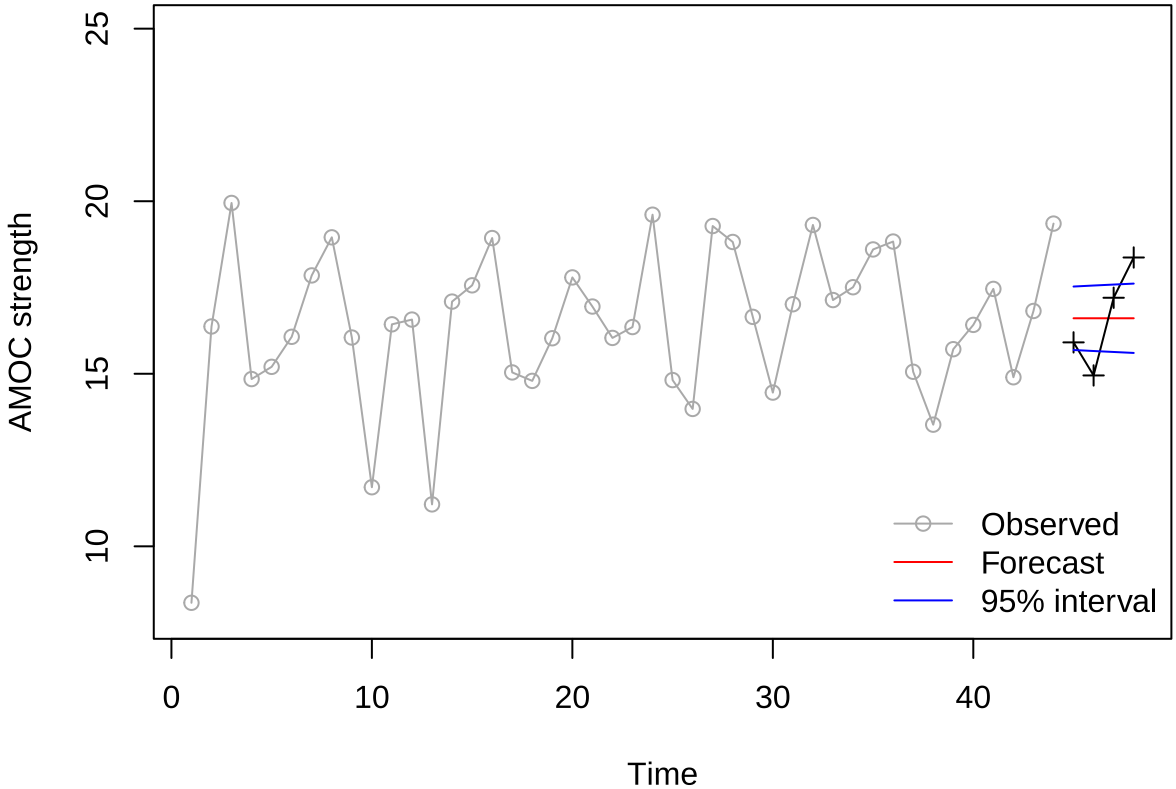
```
## y    44    ts     numeric
## mod  10    dlm    list
## m   180    mts    numeric
## U.C  45    -none- list
## D.C 180    -none- numeric
## a   176    mts    numeric
## U.R  44    -none- list
## D.R 176    -none- numeric
## f    44    ts     numeric
```

```
# Decomposition of original series
x <- cbind(amoc_qtr_ts, dropFirst(pred_dlm$a[, c(1, 2)]))
x <- window(x, start = c(1, 1))
colnames(x) <- c("AMOC", "Trend", "Seasonal")
plot(x, type = "o", main = "AMOC Strength")
```



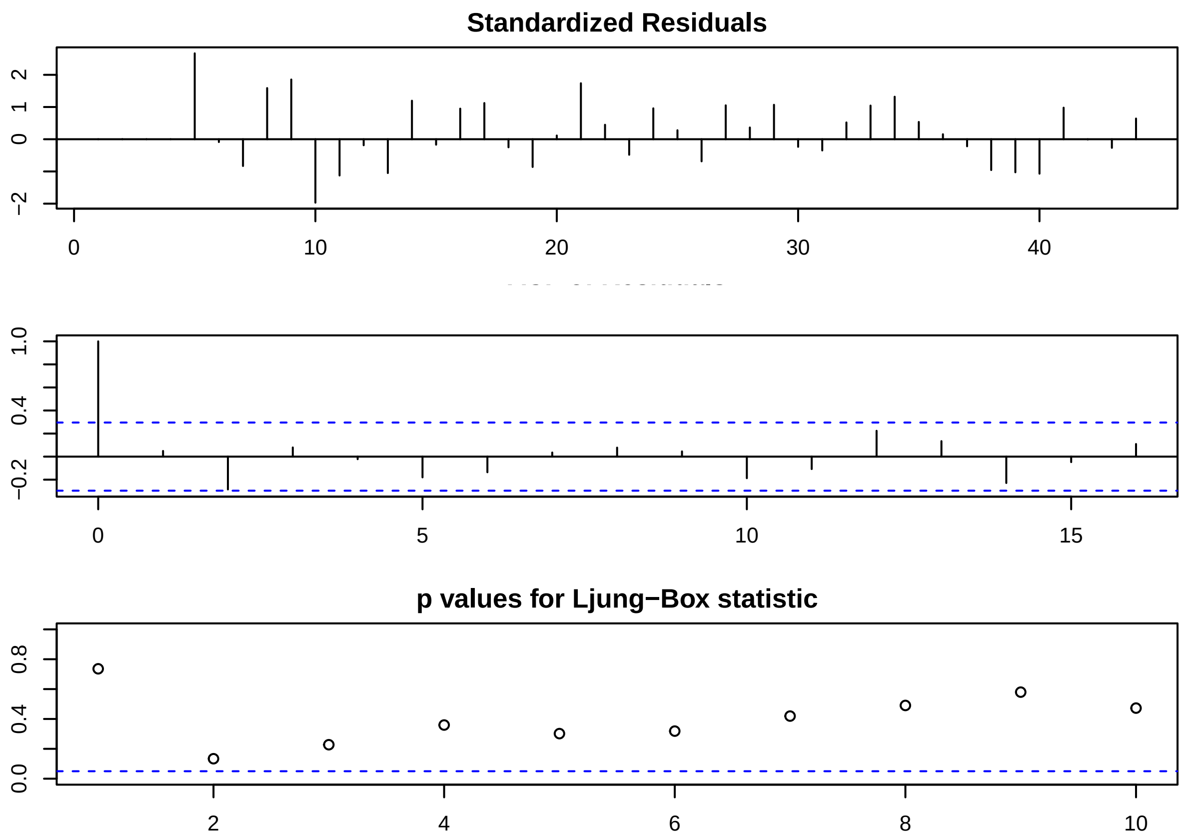

```
# Forecast
AMOC_fc <- dlmForecast(pred_dlm, nAhead = 4)

# Plot the predictions
sqrtR <- sapply(AMOC_fc$R, function(x) sqrt(x[1, 1]))
pl <- AMOC_fc$a[, 1] + qnorm(0.025, sd = sqrtR)
pu <- AMOC_fc$a[, 1] + qnorm(0.975, sd = sqrtR)
x <- ts.union(window(amoc_qtr_ts, start = c(1, 1)), AMOC_fc$a[,
    1], AMOC_fc$f, pl, pu)

par(mar = c(4, 4, 2, 2))
plot(x, plot.type = "single", type = "o", pch = c(1, NA, 3, NA,
    NA), col = c("darkgrey", "red", "black", "blue", "blue"),
    ylab = "AMOC strength", ylim = c(8, 25))
```

```r
legend("bottomright", legend = c("Observed", "Forecast", "95% interval"),
    bty = "n", pch = c(1, NA, NA), lty = 1, col = c("darkgrey",
        "red", "blue"))
```



We will need to check the residual diagnostic of dlm model. As can be seen from plots below, the residuals are quite randomly distributed around 0, within distance of 2. The ACF of residuals is reasonable. p values of is larger than 0.05 until lag 2, which means no evidence of autocorrelation at lag 1. From the residuals checking, this model is a good fit.

```r
# Residual plot of DLM
par(mar = c(3, 3, 2, 2))
tsdiag(pred_dlm)
```

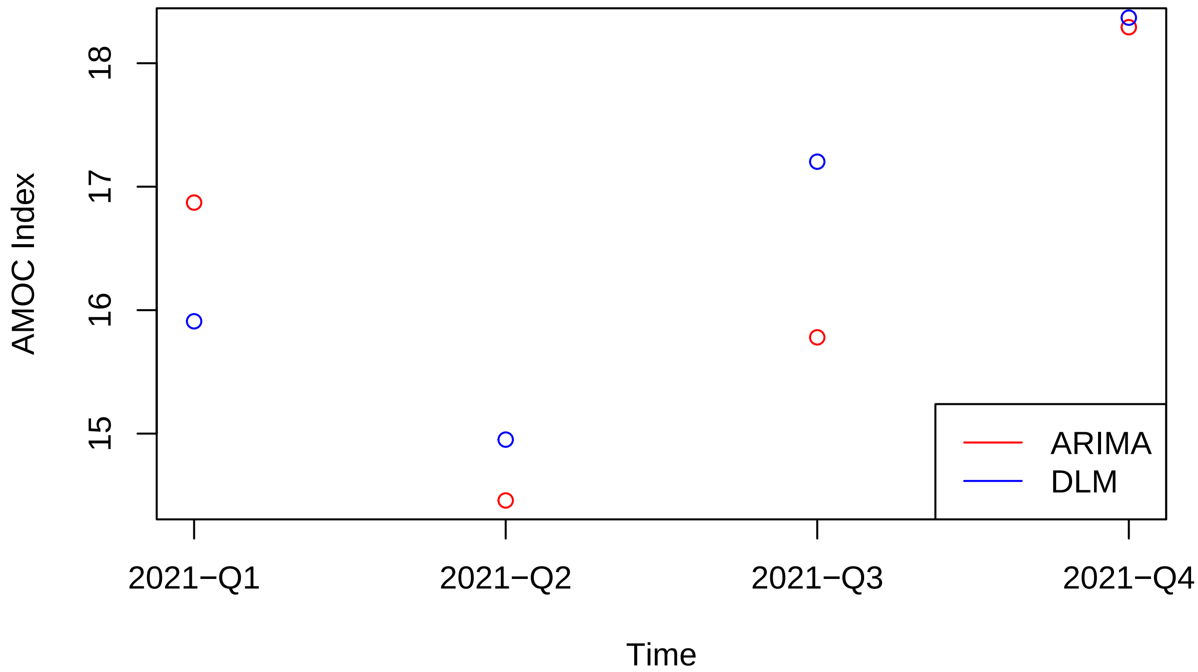


2d

The forecasts from dlm look quite different from the forecasts from ARIMA, except for the Q4 prediction. The predicted AMOC strength in 2nd and 3rd quarters by dlm model is higher than those by ARIMA model. Meanwhile, Q1 prediction of dlm is lower than that of ARIMA.

As dlm predicts higher bottom point (Q2 prediction), dlm seems to see a shrinking variation in seasonal trend of the AMOC strength, while ARIMA (without seasonal component) did not capture that. However, ARIMA(2,0,2) did capture the cyclic trend quite well, even without the seasonal difference.

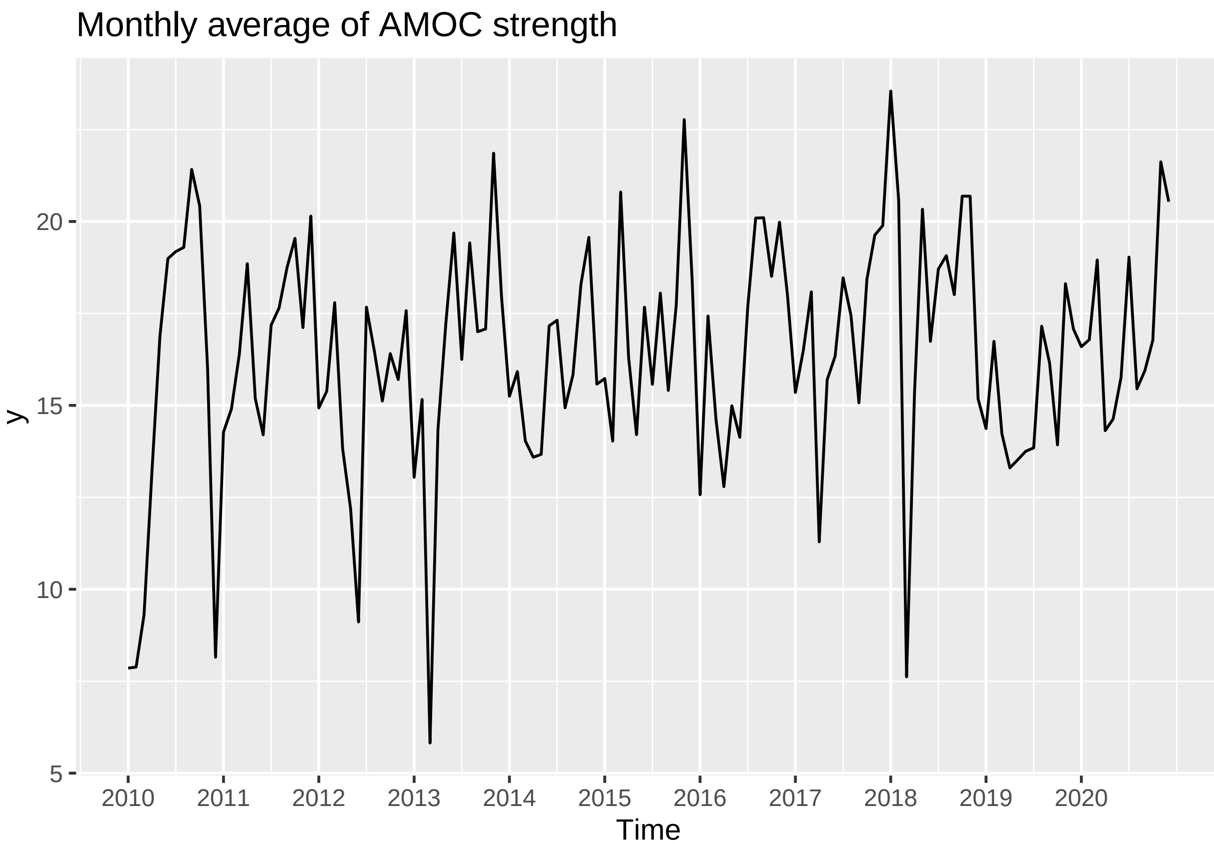
```
plot(pred_4q$mean[1:4], col = "red", xlab = "Time", ylab = "AMOC Index",
    xaxt = "n")
points(AMOC_fc$f[1:4], col = "blue")
axis(1, at = 1:4, labels = c("2021-Q1", "2021-Q2", "2021-Q3",
    "2021-Q4"))
legend("bottomright", legend = c("ARIMA", "DLM"), lty = 1, col = c("red",
    "blue"))
```



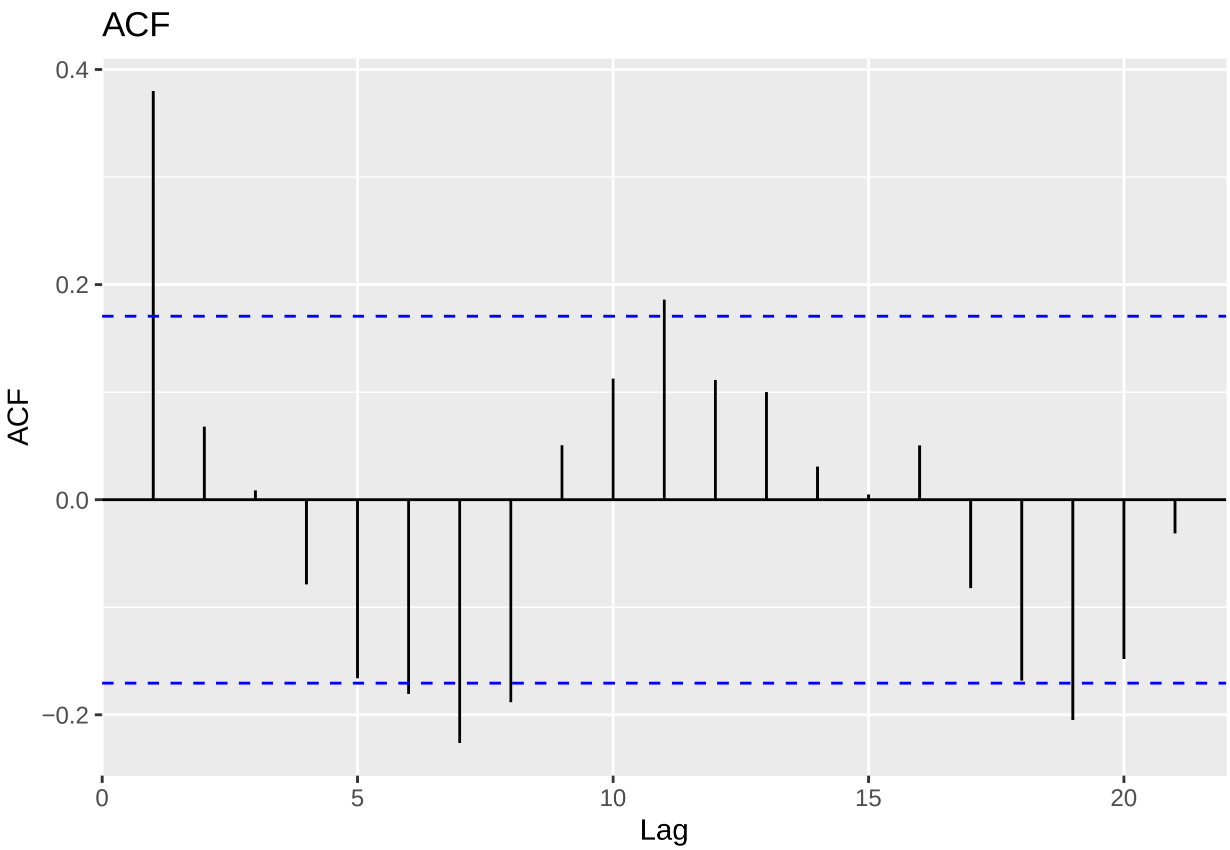
2e

Return to the original data, and calculate monthly averages instead Find an appropriate 1) ARMA/ARIMA/SARIMA model 2) a DLM for this monthly dataset, and use each to predict the AMOC strength for the next 12 months.

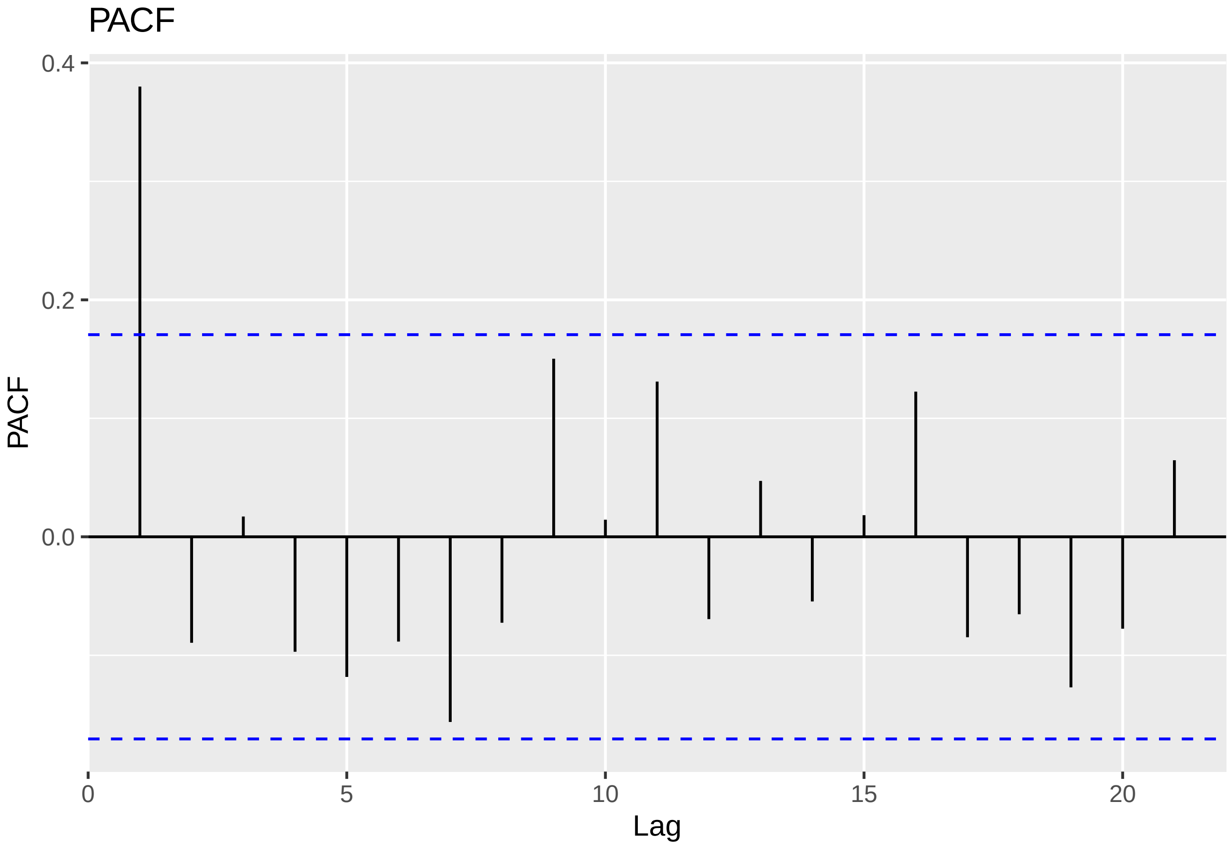
```
# Compute the Monthly mean
amoc_m <- aggregate(amoc$Strength, by = list(Year = amoc$Year,
    Month = amoc$Month), mean)

# Arrange the Quarterly mean table
names(amoc_m)[3] <- "Strength"
amoc_m <- amoc_m %>%
    arrange(Year, Month)

# Convert to time-series object
amoc_m_ts <- ts(amoc_m$Strength)

# Plot the time series of monthly average
ggplot(data = data.frame(Time = 1:length(amoc_m_ts), y = as.numeric(amoc_m_ts)),
    aes(x = Time, y = y)) + geom_line() + scale_x_continuous(breaks = seq(1,
    length(amoc_m_ts), by = 12), labels = amoc_m$Year[seq(1,
    nrow(amoc_m), by = 12)]) + ggtitle("Monthly average of AMOC strength")
```




```
ggAcf(amoc_m_ts) + ggtitle("ACF")
```




```
ggPacf(amoc_m_ts) + ggtitle("PACF")
```

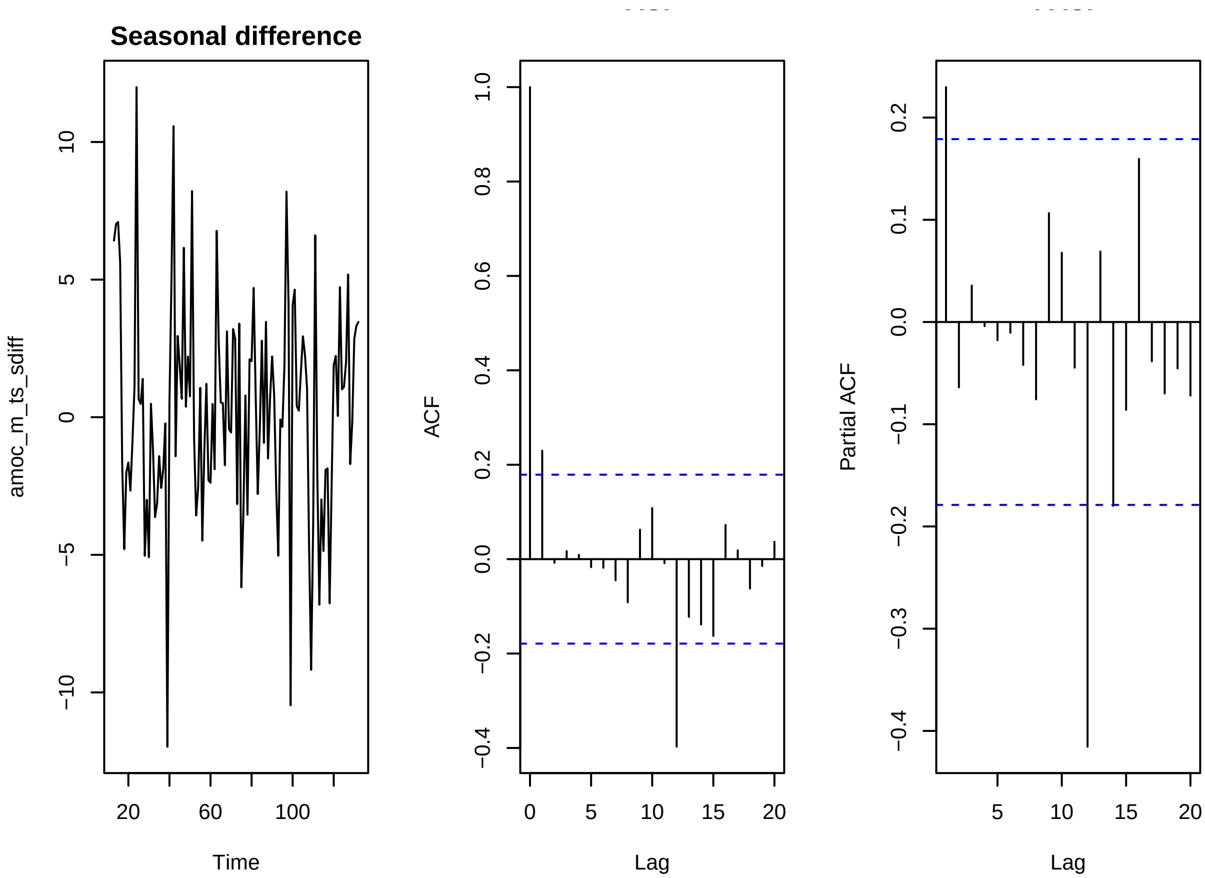


From the monthly plot, there is a yearly trend in the AMOC over the period. ACF has spikes nearly every 12 lags (at 7 and 19, at 1 and 11), indicating a seasonal trend. There was some sudden plunges in some few years when AMOC strength dropped to an unprecedented low, however, it quickly recovered to the yearly strength in the previous years.

After seasonal difference of lag 12, the AMOC strength looks stationary, thus, it does not need any further differencing. We will check ACF and PACF.

In ACF, the significant spike at lag 1 might indicate a non-seasonal MA(1) component, and the significant spike at lag 12 suggests a seasonal MA(1) component. Consequently, we begin with an ARIMA(0,0,1)(0,1,1)[12] model, indicating a seasonal difference, and non-seasonal and seasonal MA(1) components. Note: by the logic, if we analyse the PACF, we can also start with AR(1) in both non-seasonal and seasonal components.

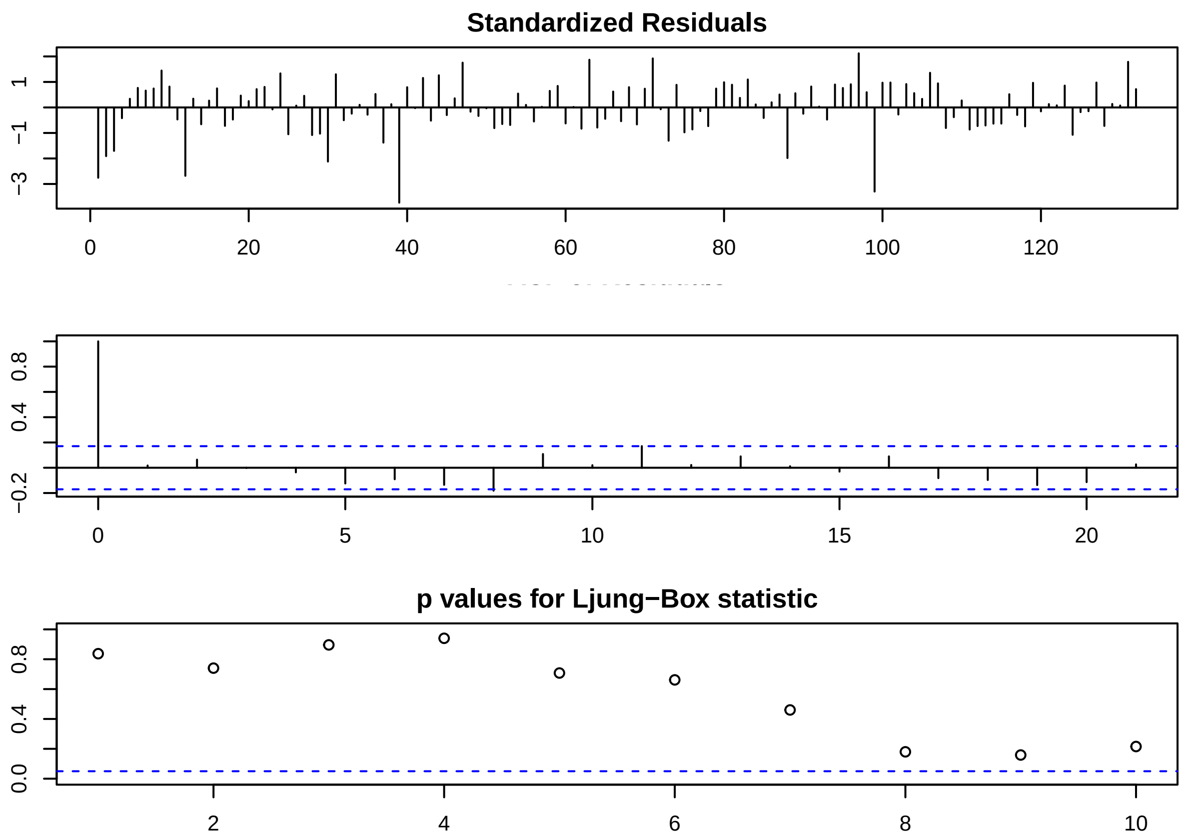
```
# Plot the SARIMA with seasonal difference
amoc_m_ts_sdiff <- diff(amoc_m_ts, lag = 12)

par(mfrow = c(1, 3), mar = c(4, 4, 2, 2))
plot(amoc_m_ts_sdiff, main = "Seasonal difference")
acf(amoc_m_ts_sdiff, main = "ACF")
pacf(amoc_m_ts_sdiff, main = "PACF")
```



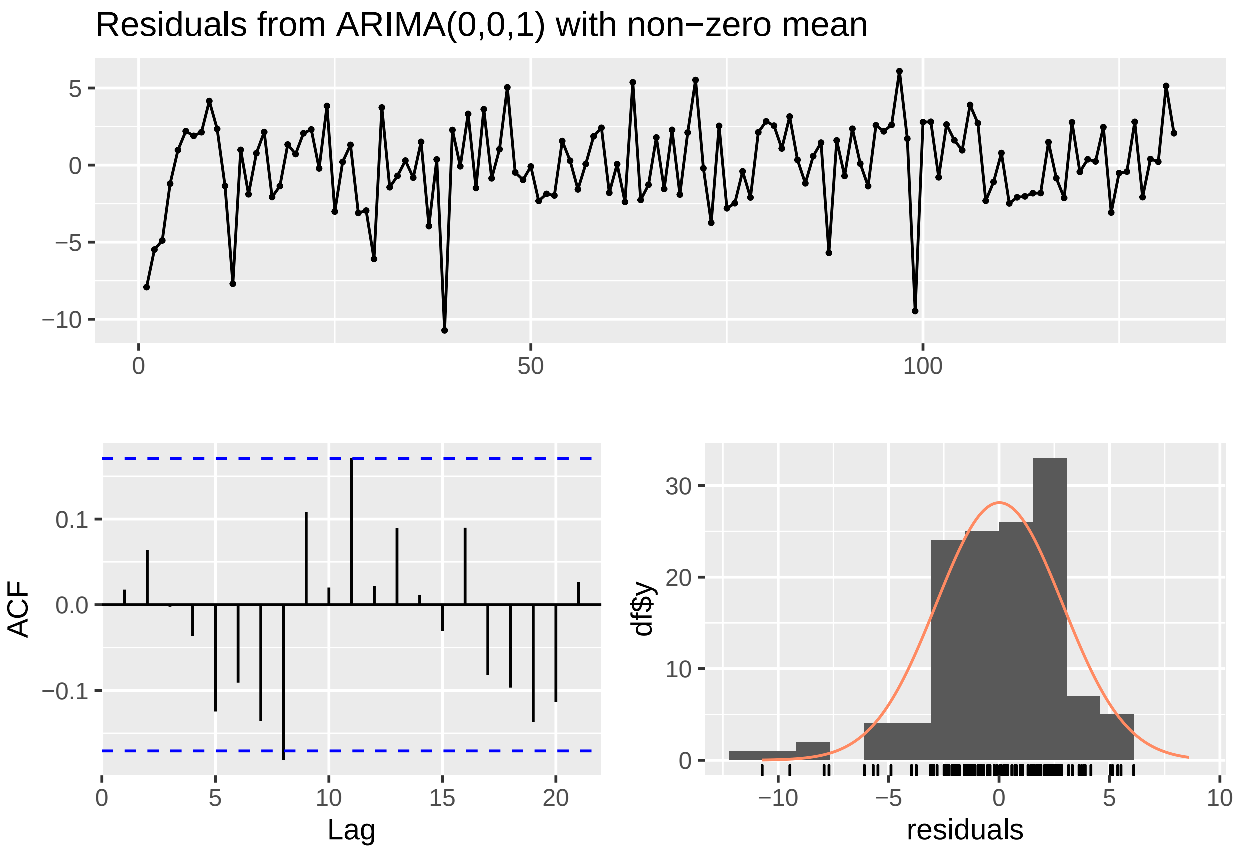

```
# Test first SARIMA model
model_1 <- Arima(amoc_m_ts, order = c(0, 0, 1), seasonal = c(0,
    1, 1), include.mean = TRUE)
model_1
```

```
## Series: amoc_m_ts
## ARIMA(0,0,1) with non-zero mean
##
## Coefficients:
##          ma1     mean
##       0.4040  16.3971
## s.e.  0.0755   0.3477
##
## sigma^2 = 8.254:  log likelihood = -325.69
## AIC=657.38   AICc=657.57   BIC=666.03
```

```
par(mar = c(3, 3, 2, 2))
tsdiag(model_1)
```



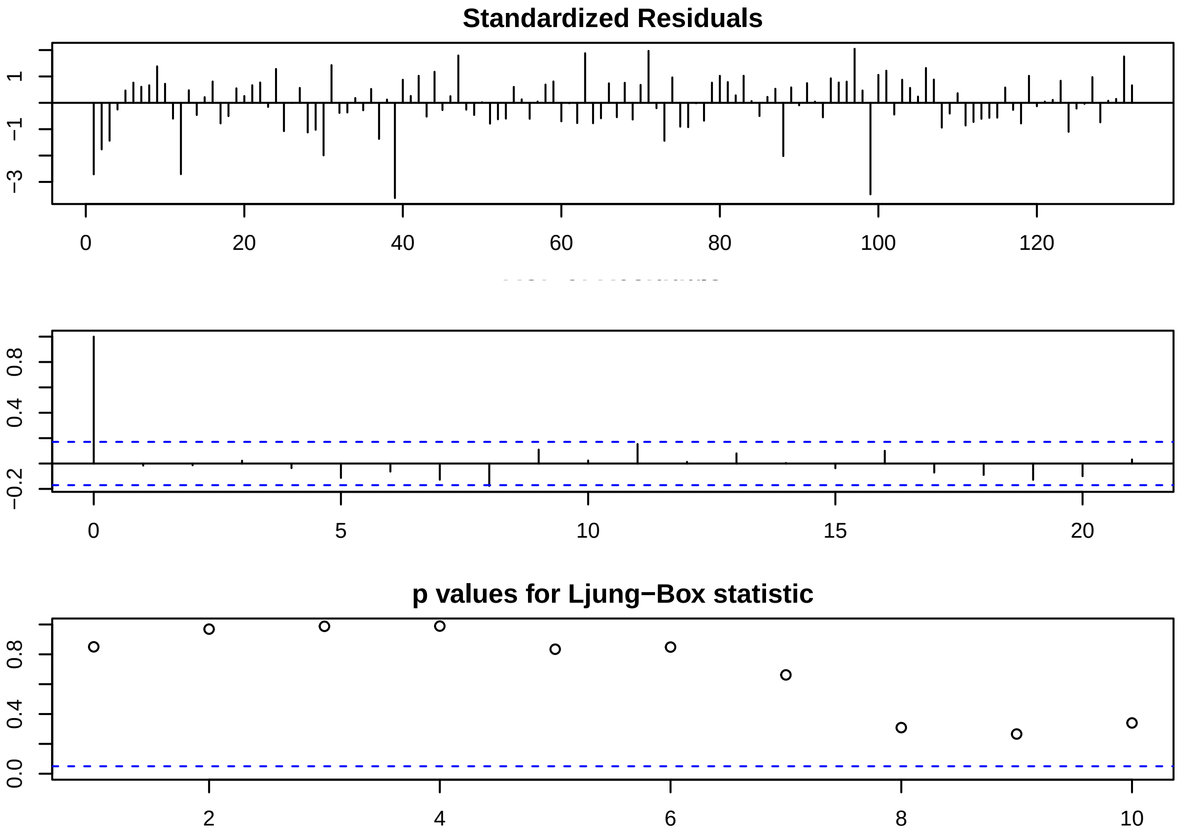

```
checkresiduals(model_1)
```



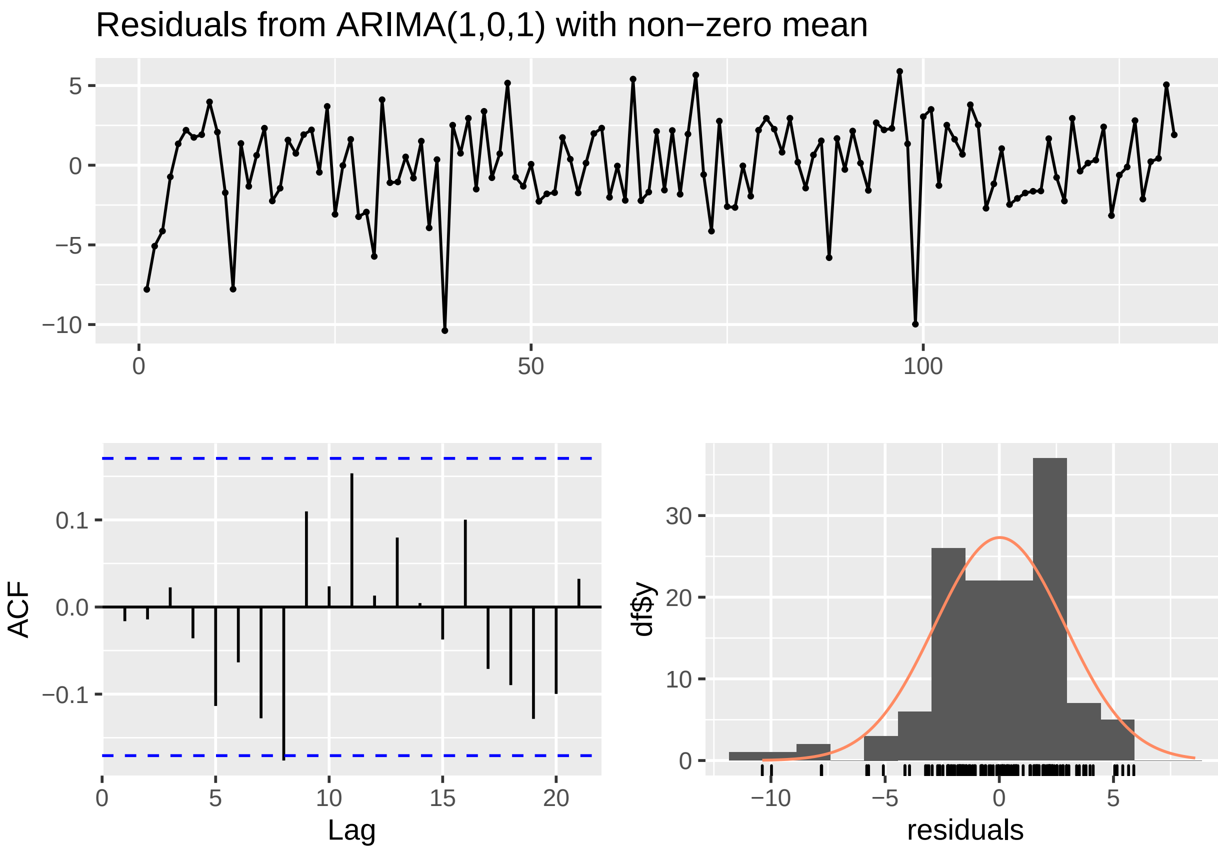

```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(0,0,1) with non-zero mean
## Q* = 13.15, df = 9, p-value = 0.1559
##
## Model df: 1.   Total lags used: 10
```

The residuals look random. From the ACF plot of residuals, we can see that there is a considerable spike at lag 8 and 11. This might suggest some missing non-seasonal components.

```
# Test some SARIMA model Add AR(1)
model_2 <- Arima(amoc_m_ts, order = c(1, 0, 1), seasonal = c(0,
    1, 1), include.mean = TRUE)
par(mar = c(3, 3, 2, 2))
tsdiag(model_2)
```

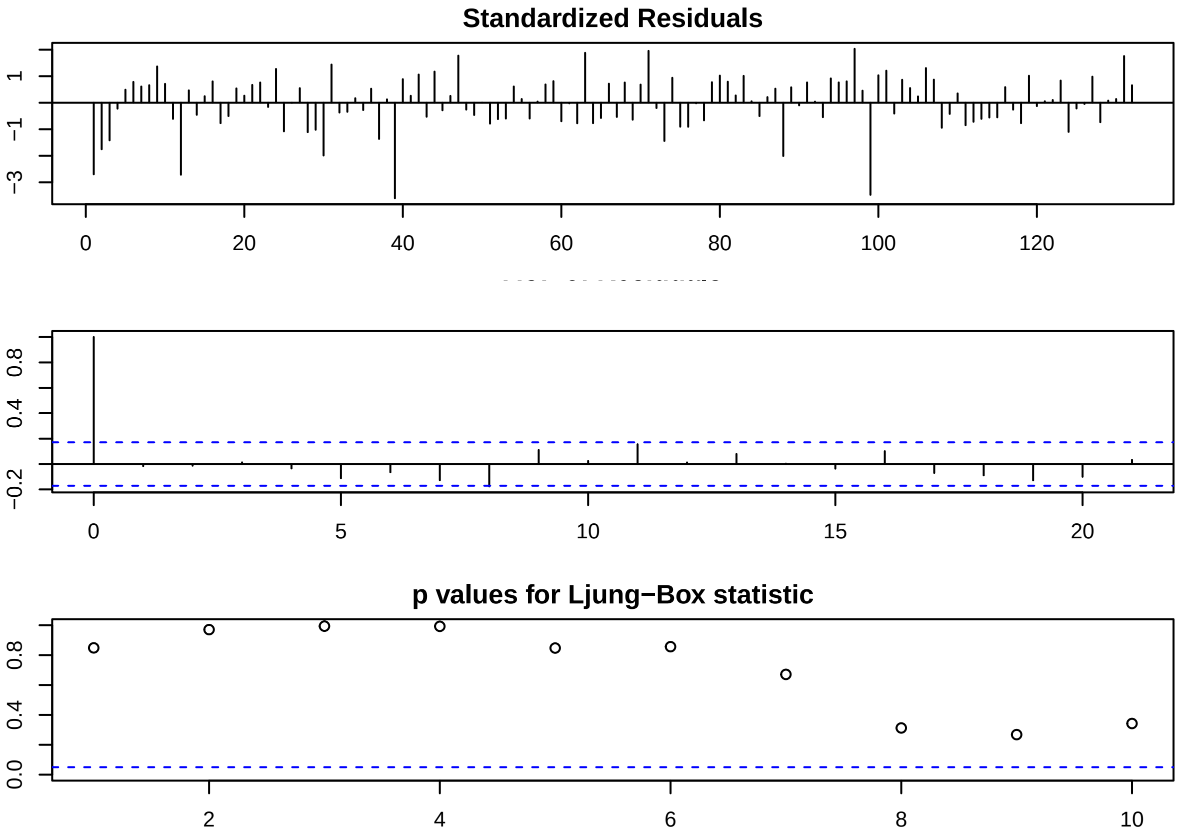



```
checkresiduals(model_2)
```

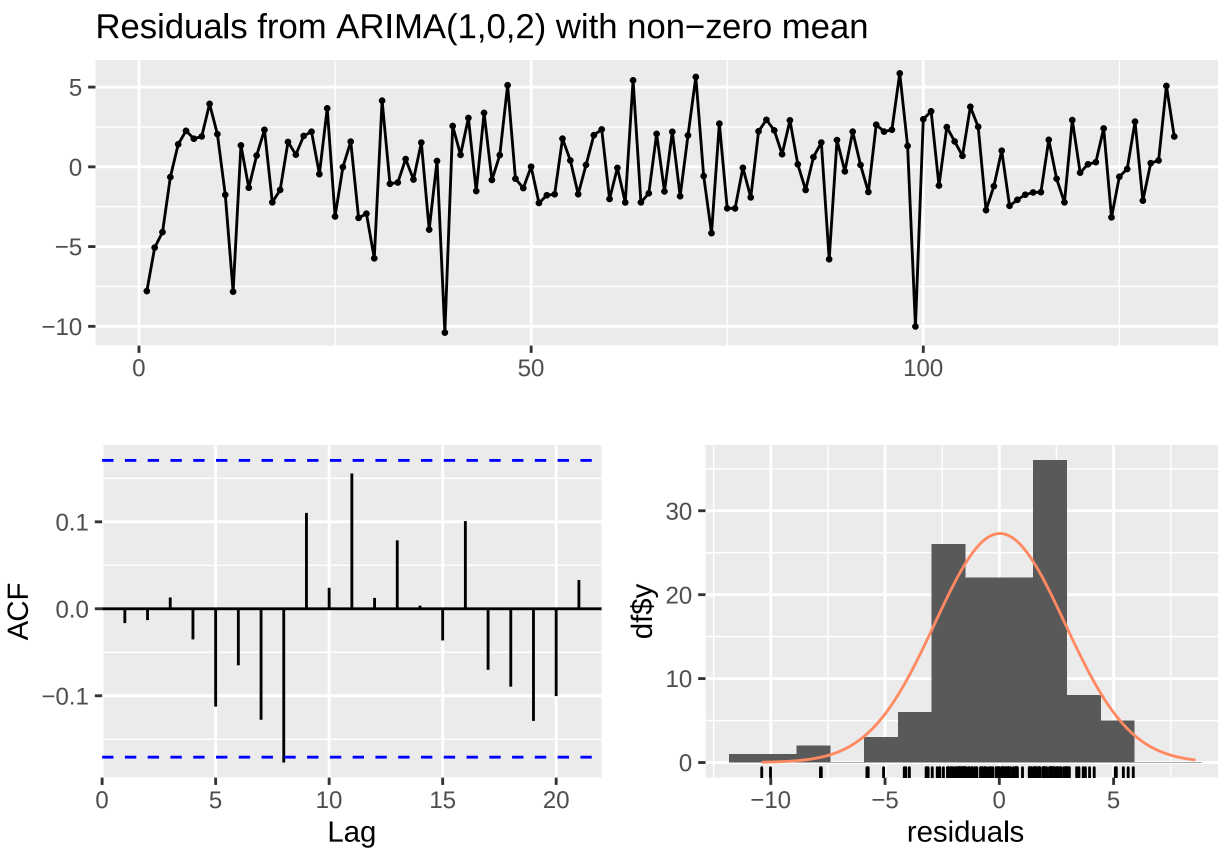



```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(1,0,1) with non-zero mean
## Q* = 11.225, df = 8, p-value = 0.1893
##
## Model df: 2.   Total lags used: 10
```

```
# Raise MA(2) in non-seasonal component
model_3 <- Arima(amoc_m_ts, order = c(1, 0, 2), seasonal = c(0,
    1, 1), include.mean = TRUE)
par(mar = c(3, 3, 2, 2))
tsdiag(model_3)
```

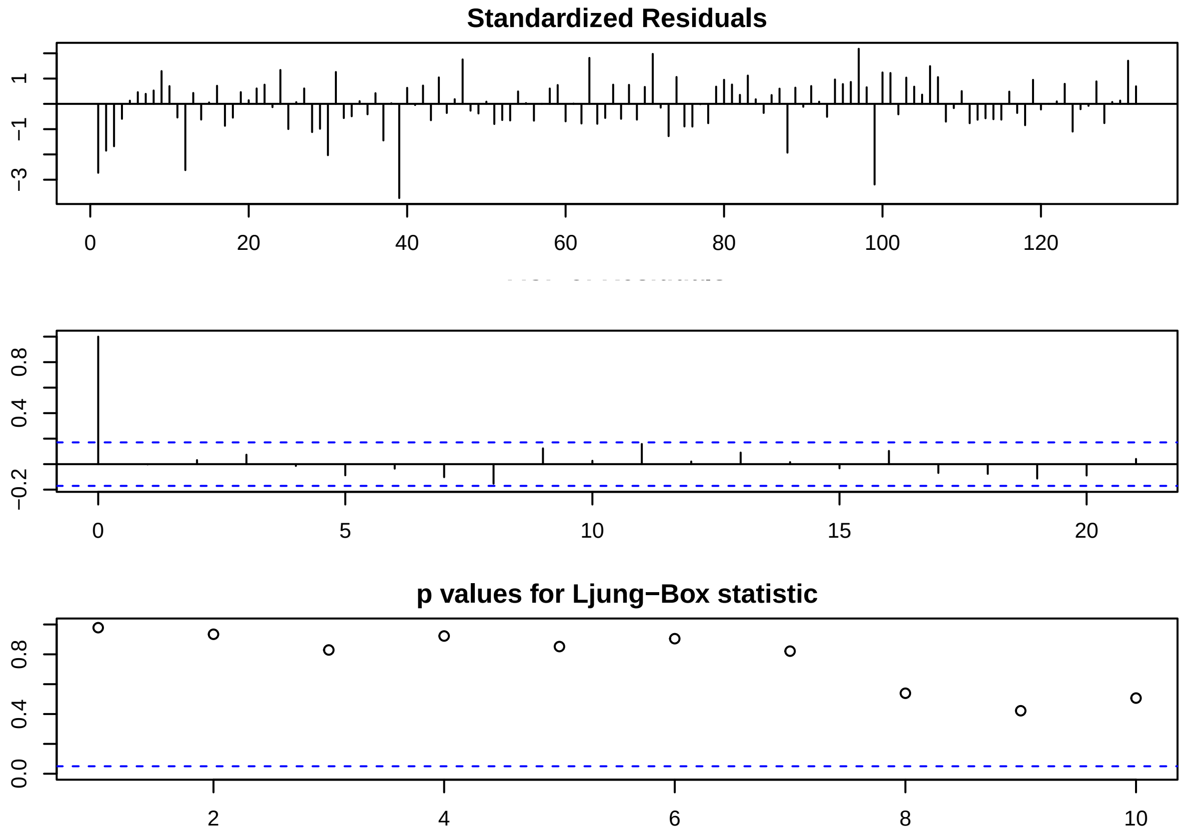



```
checkresiduals(model_3)
```

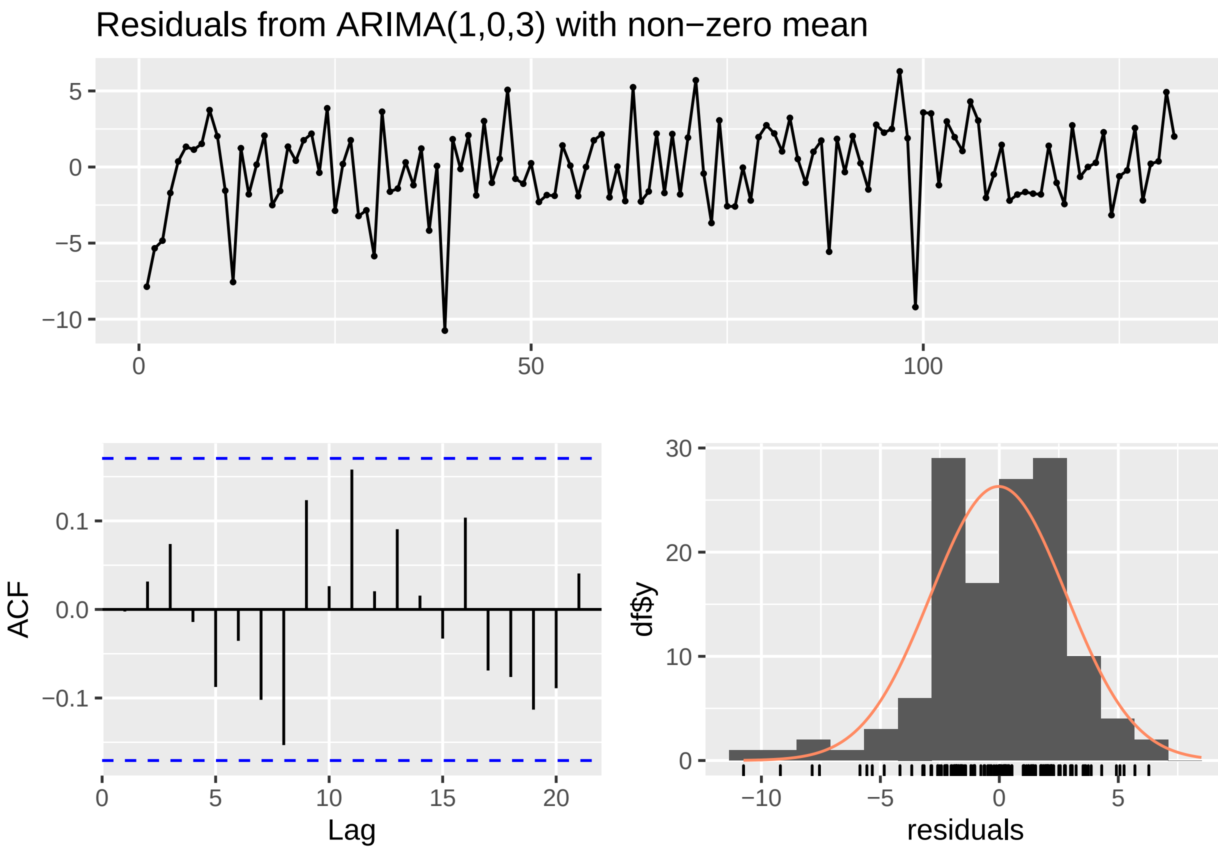



```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(1,0,2) with non-zero mean
## Q* = 11.199, df = 7, p-value = 0.1302
##
## Model df: 3.   Total lags used: 10
```

```
# more uniformity, thus, we can continue raise MA(3)

# Raise MA(3)
model_4 <- Arima(amoc_m_ts, order = c(1, 0, 3), seasonal = c(0,
    1, 1), include.mean = TRUE)
par(mar = c(3, 3, 2, 2))
tsdiag(model_4)
```




```
checkresiduals(model_4)
```




```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(1,0,3) with non-zero mean
## Q* = 9.2684, df = 6, p-value = 0.159
##
## Model df: 4.   Total lags used: 10
```

```
model_2
```

```
## Series: amoc_m_ts
## ARIMA(1,0,1) with non-zero mean
##
## Coefficients:
##          ar1     ma1     mean
##       0.2181  0.2174  16.3926
## s.e.  0.2318  0.2313   0.3838
##
## sigma^2 = 8.263:  log likelihood = -325.26
## AIC=658.52   AICc=658.83   BIC=670.05
```

```
model_3
```

```
## Series: amoc_m_ts
## ARIMA(1,0,2) with non-zero mean
##
## Coefficients:
##          ar1     ma1      ma2     mean
##       0.3347  0.1015  -0.0504  16.3910
```

```
## s.e.  0.5006  0.5013   0.2139   0.3894
##
## sigma^2 = 8.325:  log likelihood = -325.23
## AIC=660.47   AICc=660.94   BIC=674.88
```

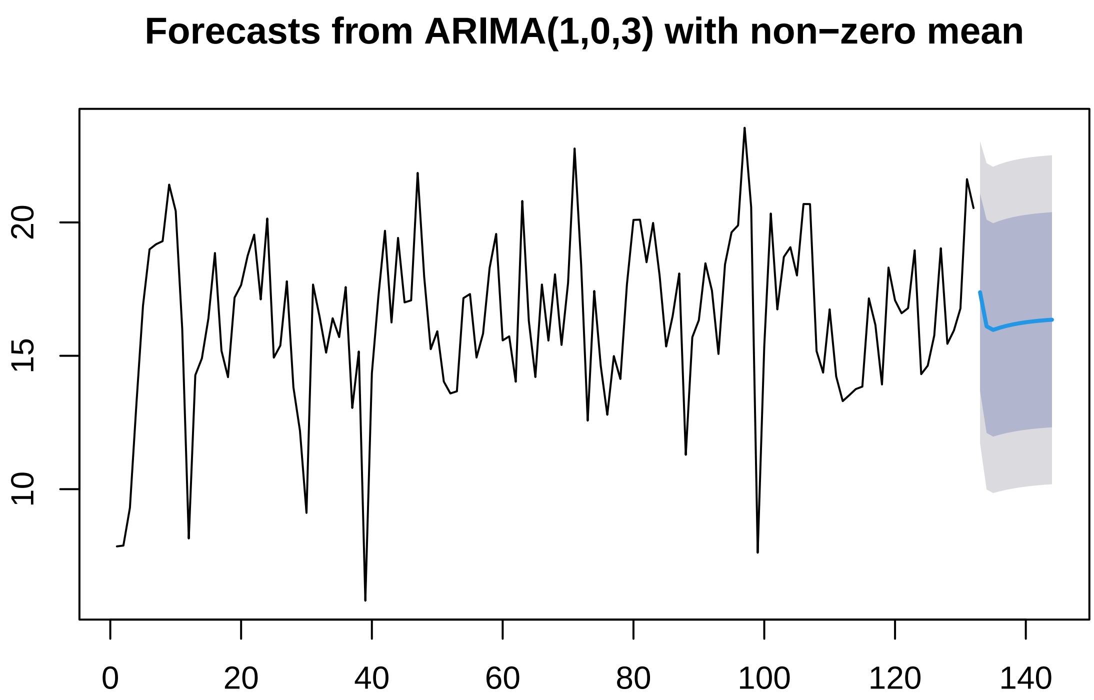
```
model_4
```

```
## Series: amoc_m_ts
## ARIMA(1,0,3) with non-zero mean
##
## Coefficients:
##          ar1      ma1      ma2      ma3     mean
##       0.8234  -0.4131  -0.3182  -0.0958  16.4314
## s.e.  0.1525   0.1664   0.1005   0.0804   0.2490
##
## sigma^2 = 8.332:  log likelihood = -324.79
## AIC=661.59   AICc=662.26   BIC=678.88
```

After adding some non-seasonal components, only model_4 has no spike in their residuals ACF. The residuals plot of model_4 is normally distributed.

We will predict the next 12 months of AMOC with model SARIMA(1,0,3)(0,1,1)[12].

```
pred_12m <- forecast(model_4, 12)
plot(pred_12m)
```



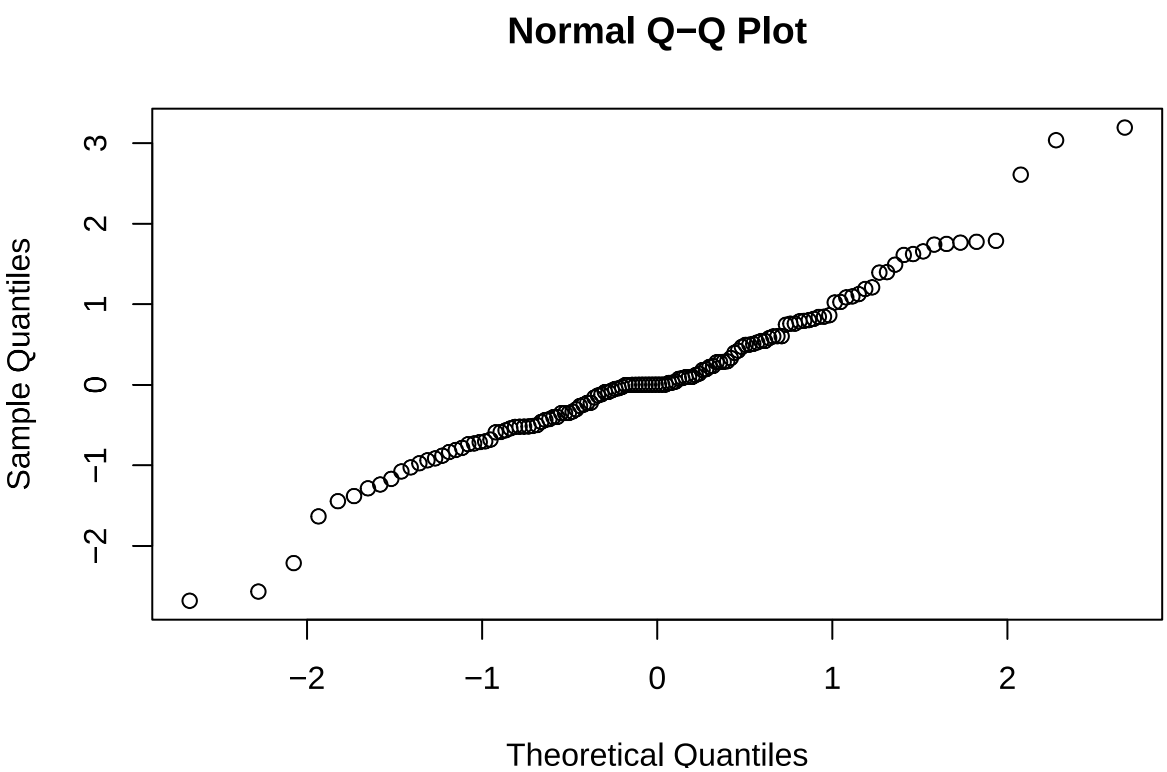
Next we will build DLM for the monthly average.

```
# Build DLM
buildFun2 <- function(x) {
    dlmModPoly(order = 1, dV = exp(x[1]), dW = exp(x[2])) + dlmModSeas(frequency = 12,
        dV = 0, dW = c(exp(x[3]), rep(0, 10)))
}

# Fit model
```

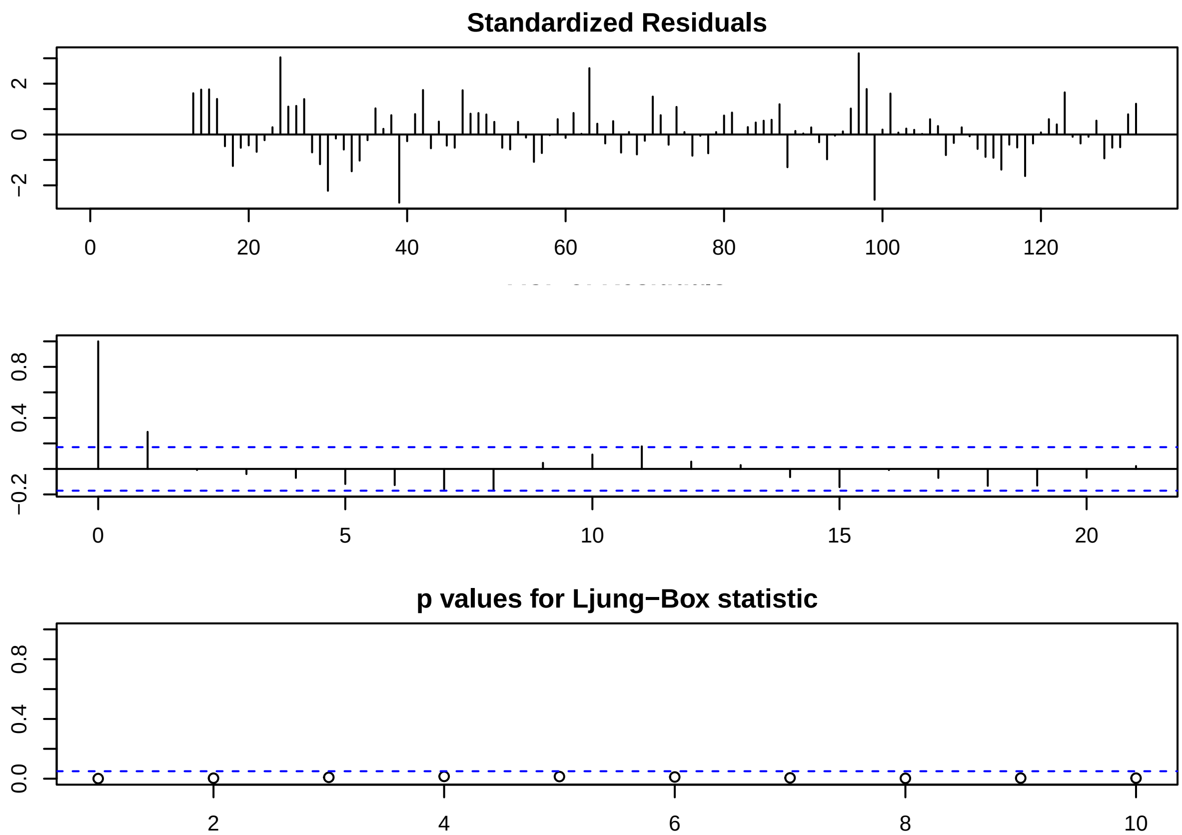
```r
fit2 <- dlmMLE(amoc_m_ts, parm = c(0, 0, 0), build = buildFun2)
fitted_model2 <- buildFun2(fit2$par)

# calculate the hidden stats
pred_dlm2 <- dlmFilter(amoc_m_ts, mod = fitted_model2)
summary(pred_dlm2)
```

```
##     Length Class  Mode
## y    132   ts     numeric
## mod   10   dlm    list
## m   1596   mts    numeric
## U.C  133   -none- list
## D.C 1596   -none- numeric
## a   1584   mts    numeric
## U.R  132   -none- list
## D.R 1584   -none- numeric
## f    132   ts     numeric
```

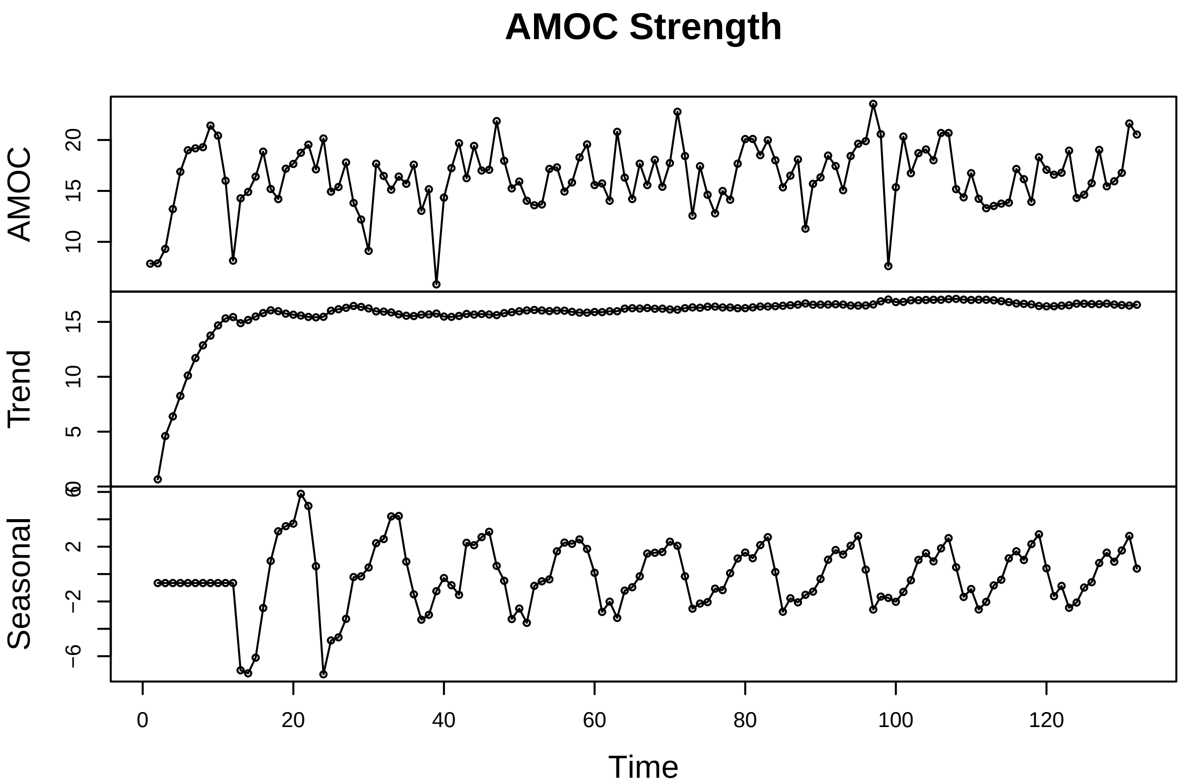
```r
# check residuals
res <- residuals(pred_dlm2, sd = FALSE)
qqnorm(res)
```




```r
par(mar = c(3, 3, 2, 2))
tsdiag(pred_dlm2)
```

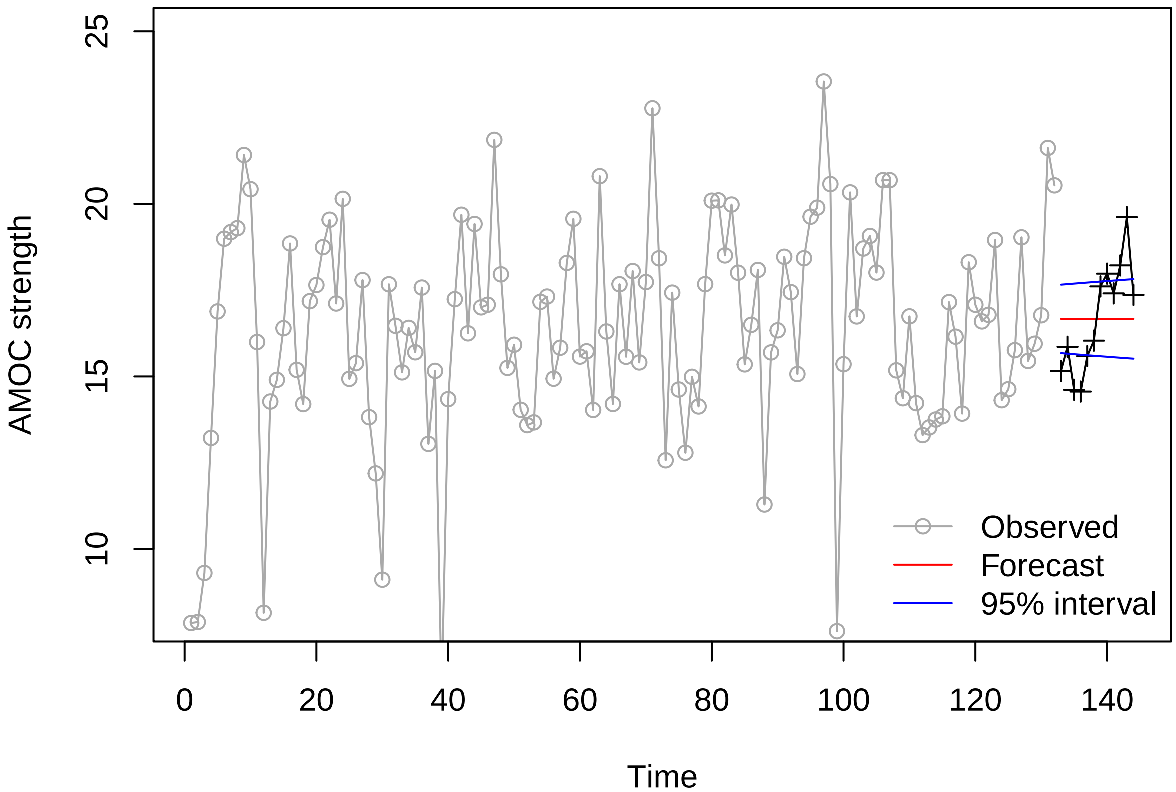



```r
# Decomposition of original series
x <- cbind(amoc_m_ts, dropFirst(pred_dlm2$a[, c(1, 2)]))
x <- window(x, start = c(1, 1))
colnames(x) <- c("AMOC", "Trend", "Seasonal")
plot(x, type = "o", main = "AMOC Strength")
```



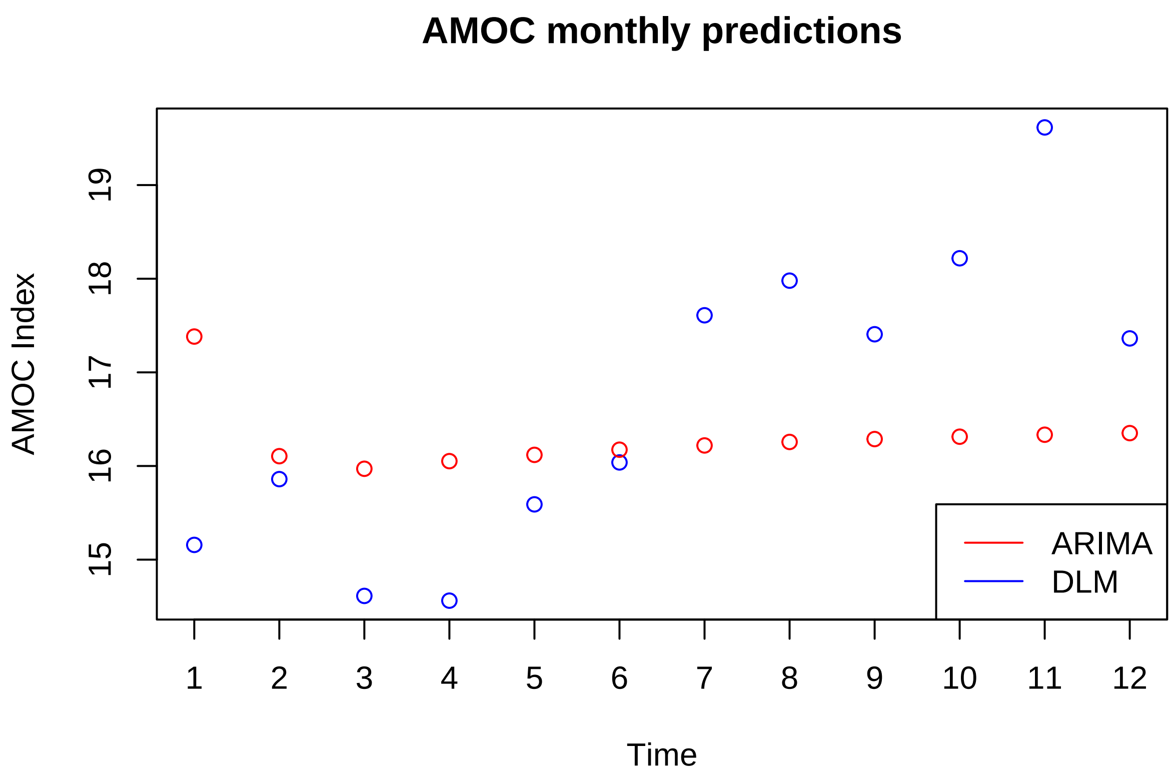
Check the residuals, we can see that there is still a significant spike at lag 1, and all the p-value are equal to 0. This indicates that there is significant autocorrelation remaining in the residuals, although the residuals look normal along the Q-Q line. The model fails the Ljung-Box test, it can be used for forecasting, however, the prediction intervals may not be accurate as there is correlated residuals.

```
# Forecast
AMOC_fc_2 <- dlmForecast(pred_dlm2, nAhead = 12)

# Plot the predictions
sqrtR <- sapply(AMOC_fc_2$R, function(x) sqrt(x[1, 1]))
pl <- AMOC_fc_2$a[, 1] + qnorm(0.025, sd = sqrtR)
pu <- AMOC_fc_2$a[, 1] + qnorm(0.975, sd = sqrtR)
x <- ts.union(window(amoc_m_ts, start = c(1, 1)), AMOC_fc_2$a[,
    1], AMOC_fc_2$f, pl, pu)

par(mar = c(4, 4, 2, 2))
plot(x, plot.type = "single", type = "o", pch = c(1, NA, 3, NA,
    NA), col = c("darkgrey", "red", "black", "blue", "blue"),
    ylab = "AMOC strength", ylim = c(8, 25))
legend("bottomright", legend = c("Observed", "Forecast", "95% interval"),
    bty = "n", pch = c(1, NA, NA), lty = 1, col = c("darkgrey",
        "red", "blue"))
```

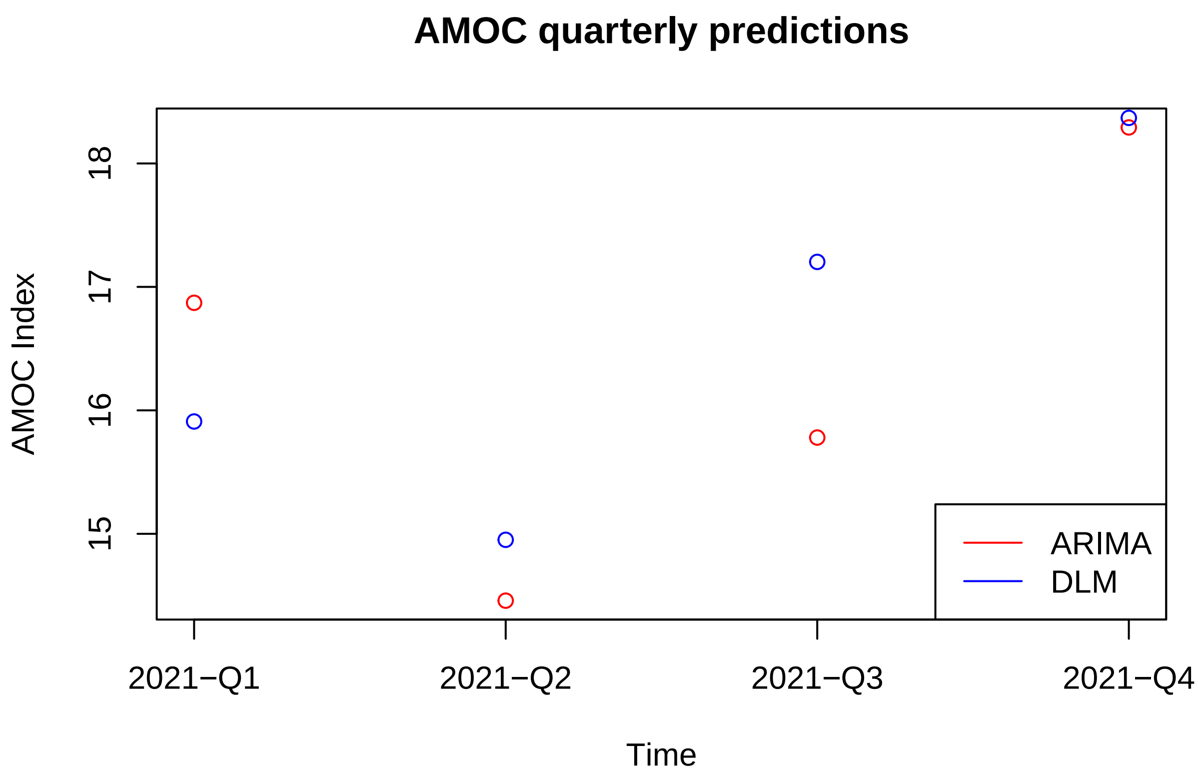


2f

```
plot(AMOC_fc_2$f[1:12], col = "blue", xlab = "Time", ylab = "AMOC Index",
    main = "AMOC monthly predictions", xaxt = "n")
points(pred_12m$mean[1:12], col = "red")
axis(1, at = 1:12)
legend("bottomright", legend = c("ARIMA", "DLM"), lty = 1, col = c("red",
    "blue"))
```




```
plot(pred_4q$mean[1:4], col = "red", xlab = "Time", ylab = "AMOC Index",
    main = "AMOC quarterly predictions", xaxt = "n")
points(AMOC_fc$f[1:4], col = "blue")
axis(1, at = 1:4, labels = c("2021-Q1", "2021-Q2", "2021-Q3",
    "2021-Q4"))
legend("bottomright", legend = c("ARIMA", "DLM"), lty = 1, col = c("red",
    "blue"))
```



With ARIMA model, the monthly predictions are much different from quarterly predictions. The quarterly results follow an yearly ups and downs, however, monthly predictions only slightly and gradually increase from February 2021 to December 2021. Meanwhile, monthly predictions from DLM are along with quarterly predictions, low in the first two quarters and higher in the later quarters.

(https://otexts.com/fpp2/seasonal-arima.html)