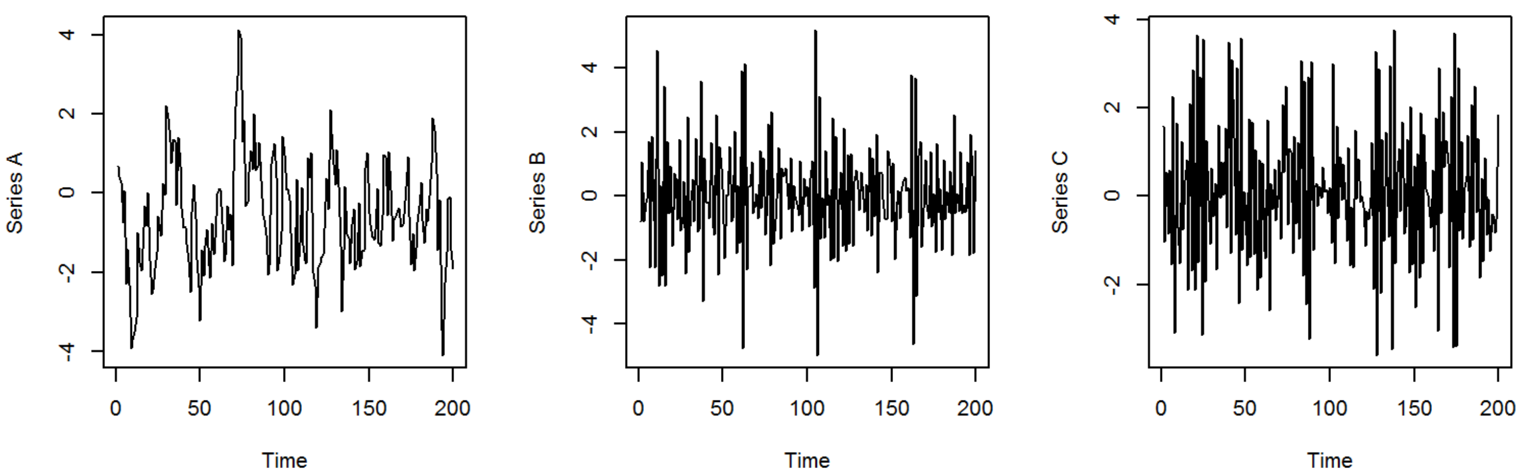
Question 5

a) plot the series

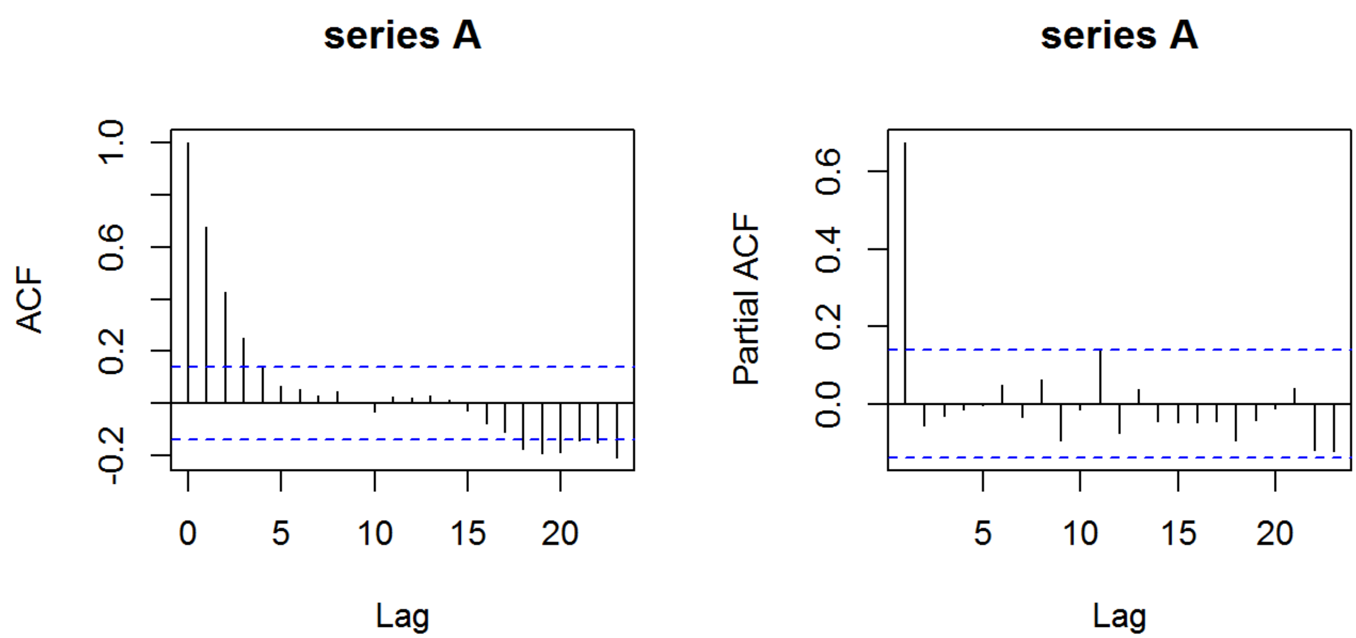
```
## read in data
data = read.table("SeriesforA2.txt", header = TRUE)
seriesA <- data[, 1]
seriesB <- data[, 2]
seriesC <- data[, 3]
par(mfcol=c(1,3))
plot(as.ts(seriesA), ylab = "Series A")
plot(as.ts(seriesB), ylab = "Series B")
plot(as.ts(seriesC), ylab = "Series C")
```



As we can see from the plots, data in series A has the least variation, and data in series C has the most. Also, they all have mean 0.

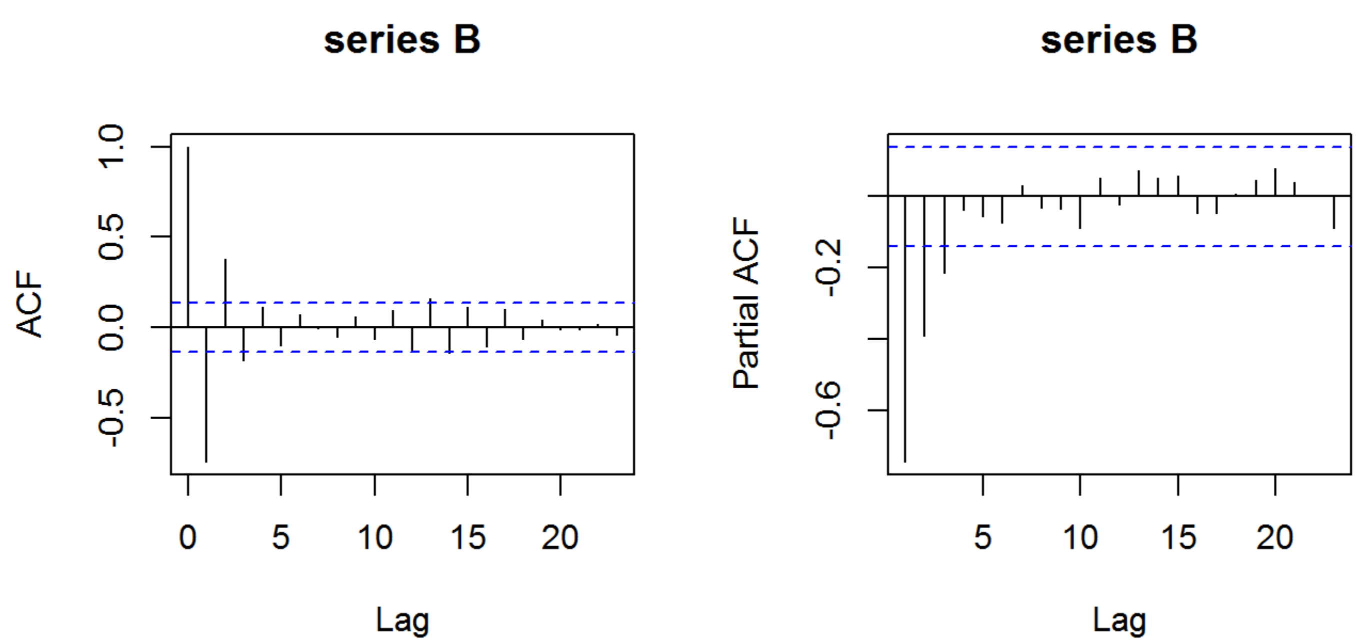
b) SACF & SPACF graphs

```
par(mfrow=c(1,2))
acf(seriesA, main = "series A")
pacf(seriesA, main = "series A")
```



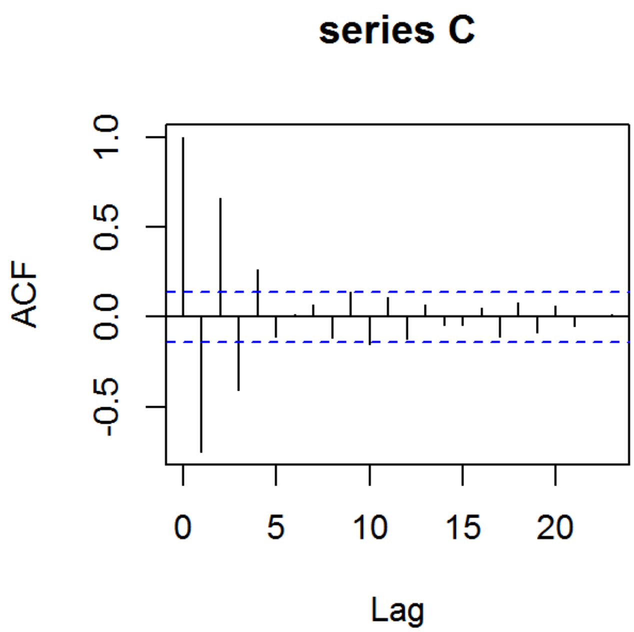
From the graph we see that ACF trend of series A has an exponential decay. So it would not have some AR components. Also, the PACF graph only has one spike at $lag = 1$,so it might be pure AR. Suggest $p = 1, q = 0; p = 2; q = 0; p = 1, q = 1$

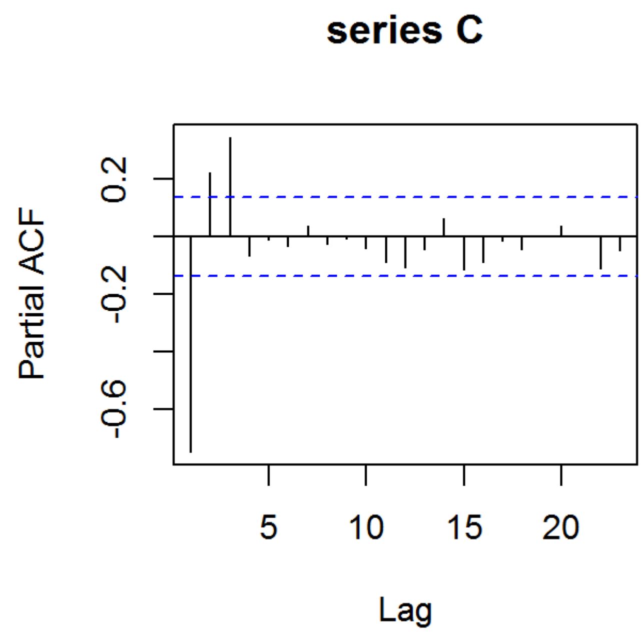
```
par(mfrow=c(1,2))
acf(seriesB, main = "series B")
pacf(seriesB, main = "series B")
```



From the ACF graph we know that B also has AR components. And the exponential decay in PACF suggests that it should be an ARMA process. Suggest $p = 1, q = 1; p = 2; q = 2; p = 1, q = 2; p = 2, q = 1$

```
par(mfrow=c(1,2))
acf(seriesC, main = "series C")
pacf(seriesC, main = "series C")
```





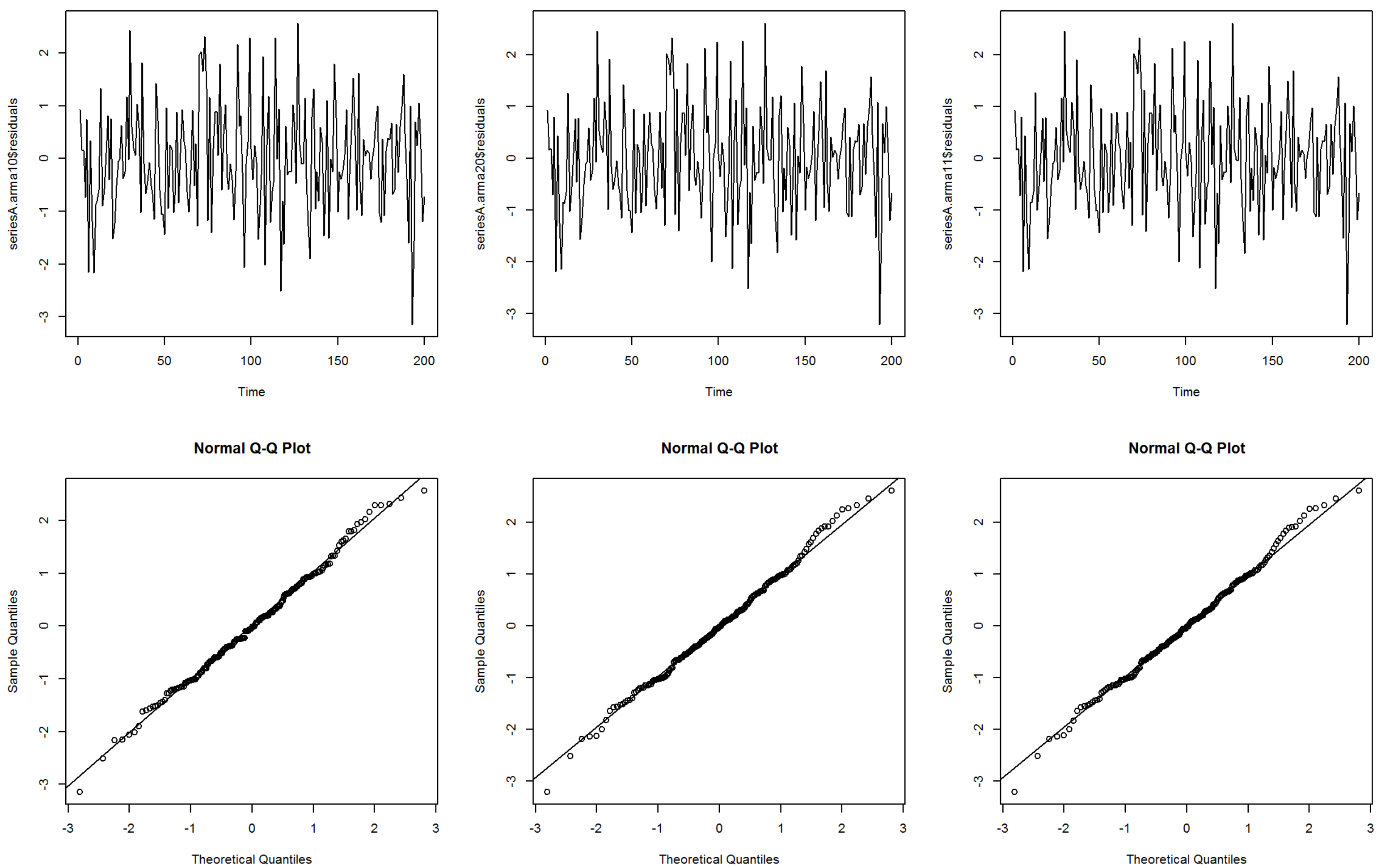
The ACF plot for series C also has a decay trend, and a spike at $lag = 2$. Suggest $p = 1, q = 1; p = 2; q = 2; p = 1, q = 2; p = 2.q = 1$.

c) ARIMA fit with suggested values

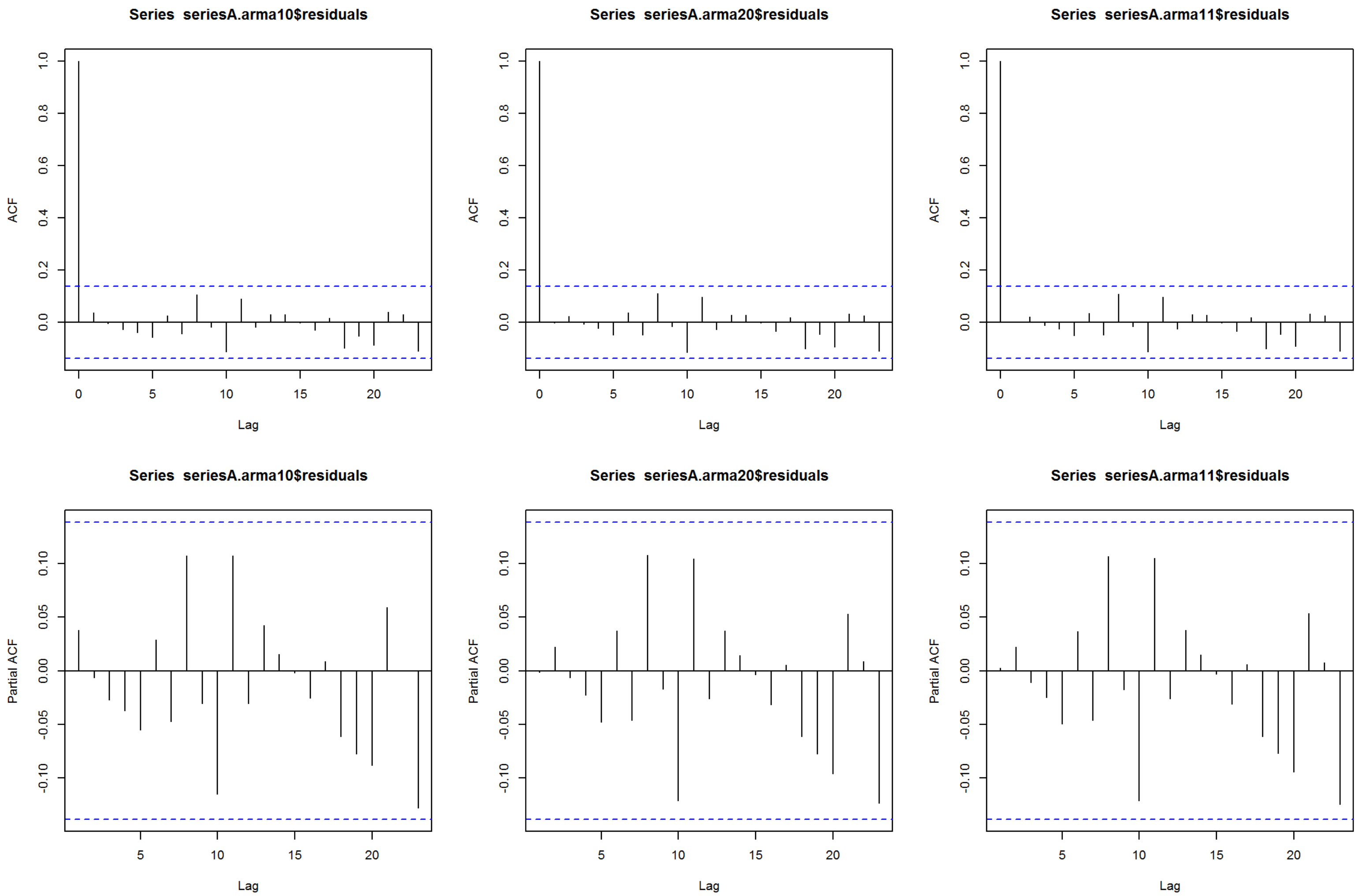
```
# then define the plausible arima value
seriesA.arma10 <- arima(seriesA, order = c(1, 0, 0), method = "ML")
seriesA.arma20 <- arima(seriesA, order = c(2, 0, 0), method = "ML")
seriesA.arma11 <- arima(seriesA, order = c(1, 0, 1), method = "ML")
seriesB.arma12 <- arima(seriesB, order = c(1, 0, 2), method = "ML")
seriesB.arma21 <- arima(seriesB, order = c(2, 0, 1), method = "ML")
seriesB.arma11 <- arima(seriesB, order = c(1, 0, 1), method = "ML")
seriesB.arma22 <- arima(seriesB, order = c(2, 0, 2), method = "ML")
seriesC.arma20 <- arima(seriesC, order = c(2, 0, 0), method = "ML")
seriesC.arma22 <- arima(seriesC, order = c(2, 0, 2), method = "ML")
seriesC.arma21 <- arima(seriesC, order = c(2, 0, 1), method = "ML")
seriesC.arma12 <- arima(seriesC, order = c(1, 0, 2), method = "ML")
```

d) diagnosis

```
## residual plots for seriesA models
par(mfrow=c(2,3))
plot(seriesA.arma10$residuals)
plot(seriesA.arma20$residuals)
plot(seriesA.arma11$residuals)

qqnorm(seriesA.arma10$residuals)
qqline(seriesA.arma10$residuals)
qqnorm(seriesA.arma20$residuals)
qqline(seriesA.arma20$residuals)
qqnorm(seriesA.arma11$residuals)
qqline(seriesA.arma11$residuals)
```




```
par(mfrow=c(2,3))
acf(seriesA.arma10$residuals)
acf(seriesA.arma20$residuals)
acf(seriesA.arma11$residuals)
pacf(seriesA.arma10$residuals)
pacf(seriesA.arma20$residuals)
pacf(seriesA.arma11$residuals)
```




```
shapiro.test(seriesA.arma10$residuals)$p.value
```

```
## [1] 0.7663966
```

```
shapiro.test(seriesA.arma20$residuals)$p.value
```

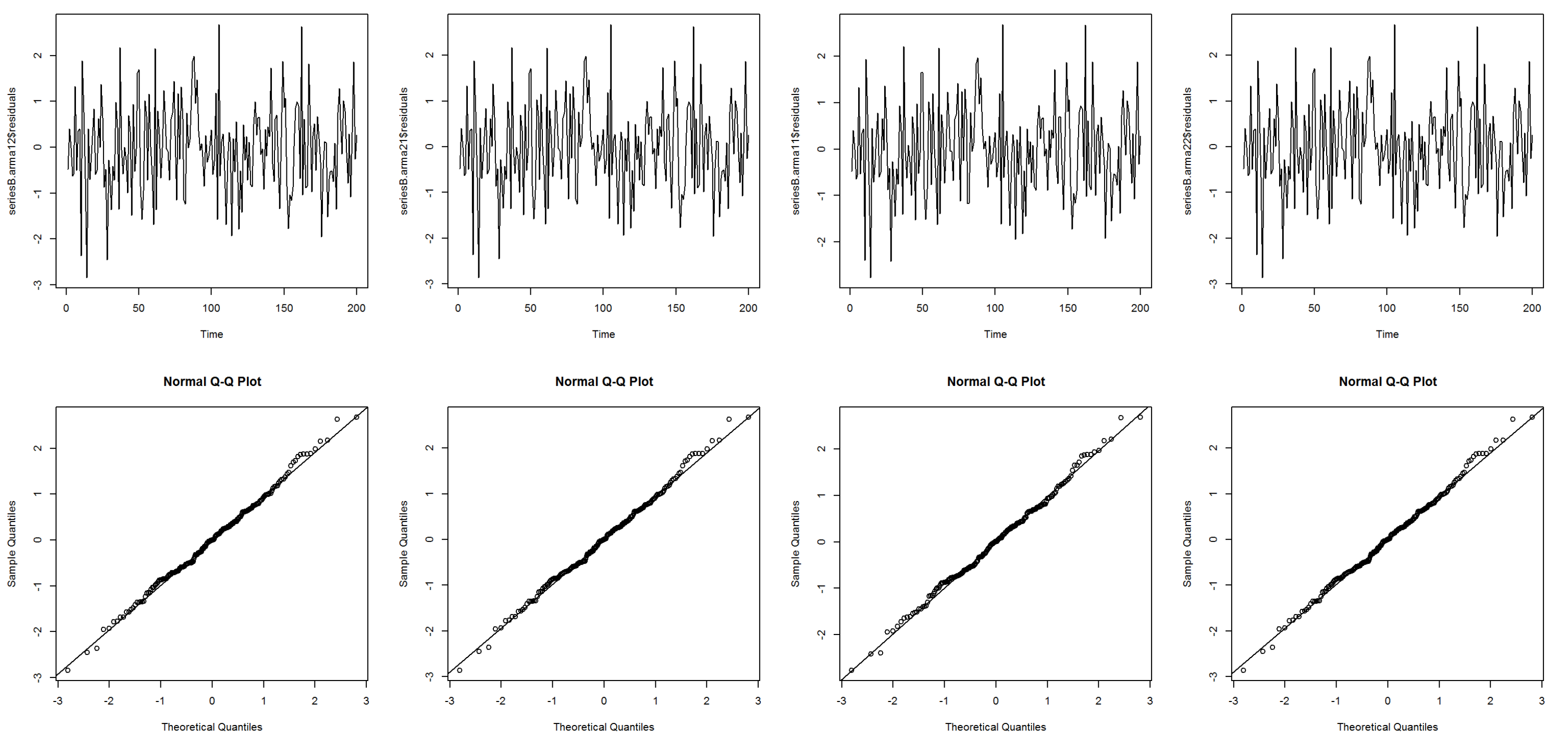
```
## [1] 0.7989902
```

```
shapiro.test(seriesA.arma11$residuals)$p.value
```

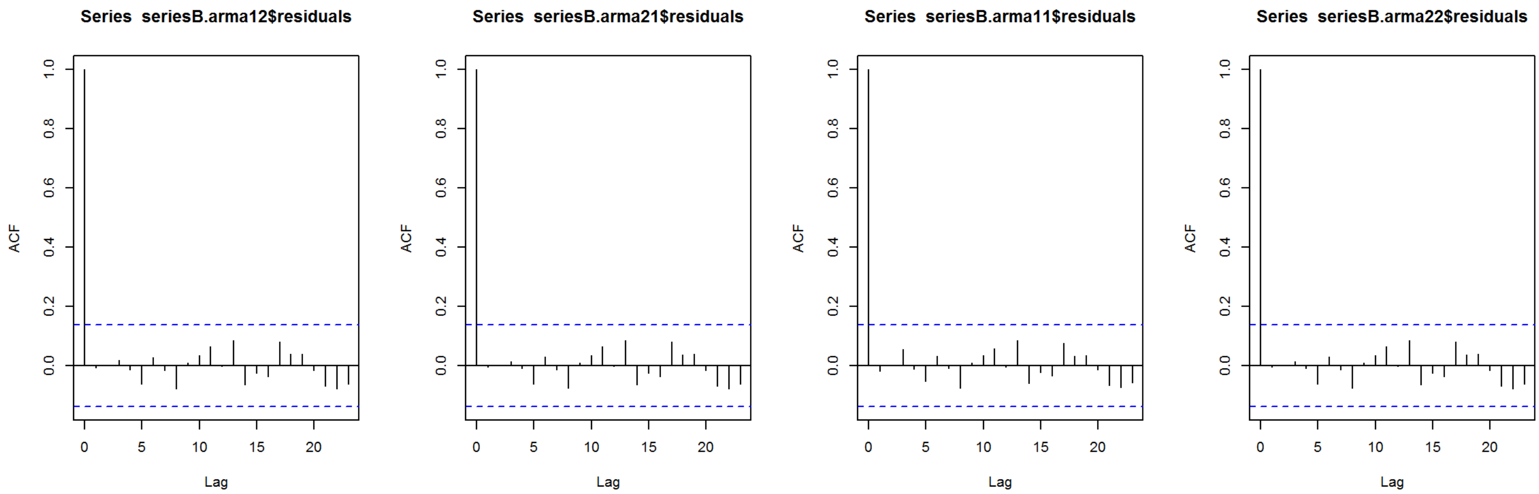
```
## [1] 0.7953363
```

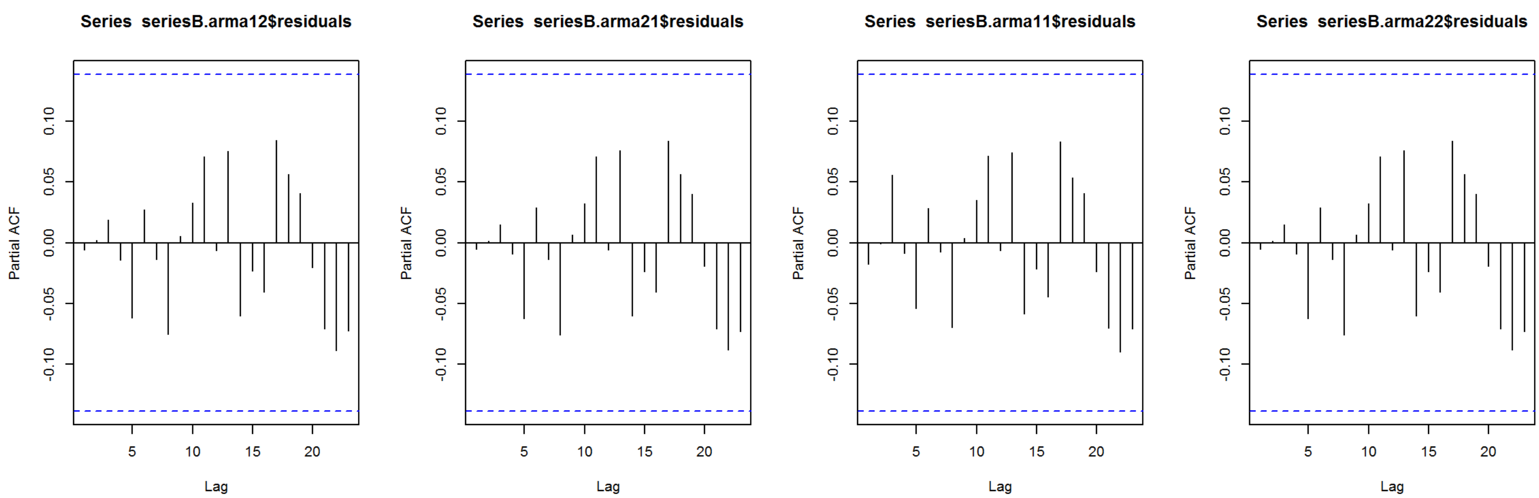
All models seem apporiate. Their residuals all look like white noise and form a stright line on the Q-Q plot. The ACF plot shows one spike on when $h = 0$, and no spike on PACF. The shapiro test also shows that the residuals are insignificant.

```
## residual plots for seriesB models
par(mfrow=c(2,4))
plot(seriesB.arma12$residuals)
plot(seriesB.arma21$residuals)
plot(seriesB.arma11$residuals)
plot(seriesB.arma22$residuals)

qqnorm(seriesB.arma12$residuals)
qqline(seriesB.arma12$residuals)
qqnorm(seriesB.arma21$residuals)
qqline(seriesB.arma21$residuals)
qqnorm(seriesB.arma11$residuals)
qqline(seriesB.arma11$residuals)
qqnorm(seriesB.arma22$residuals)
qqline(seriesB.arma22$residuals)
```




```
par(mfrow=c(2,4))
acf(seriesB.arma12$residuals)
acf(seriesB.arma21$residuals)
acf(seriesB.arma11$residuals)
acf(seriesB.arma22$residuals)
pacf(seriesB.arma12$residuals)
pacf(seriesB.arma21$residuals)
pacf(seriesB.arma11$residuals)
pacf(seriesB.arma22$residuals)
```






```
shapiro.test(seriesB.arma12$residuals)$p.value
```

```
## [1] 0.9022241
```

```
shapiro.test(seriesB.arma21$residuals)$p.value
```

```
## [1] 0.9083403
```

```
shapiro.test(seriesB.arma11$residuals)$p.value
```

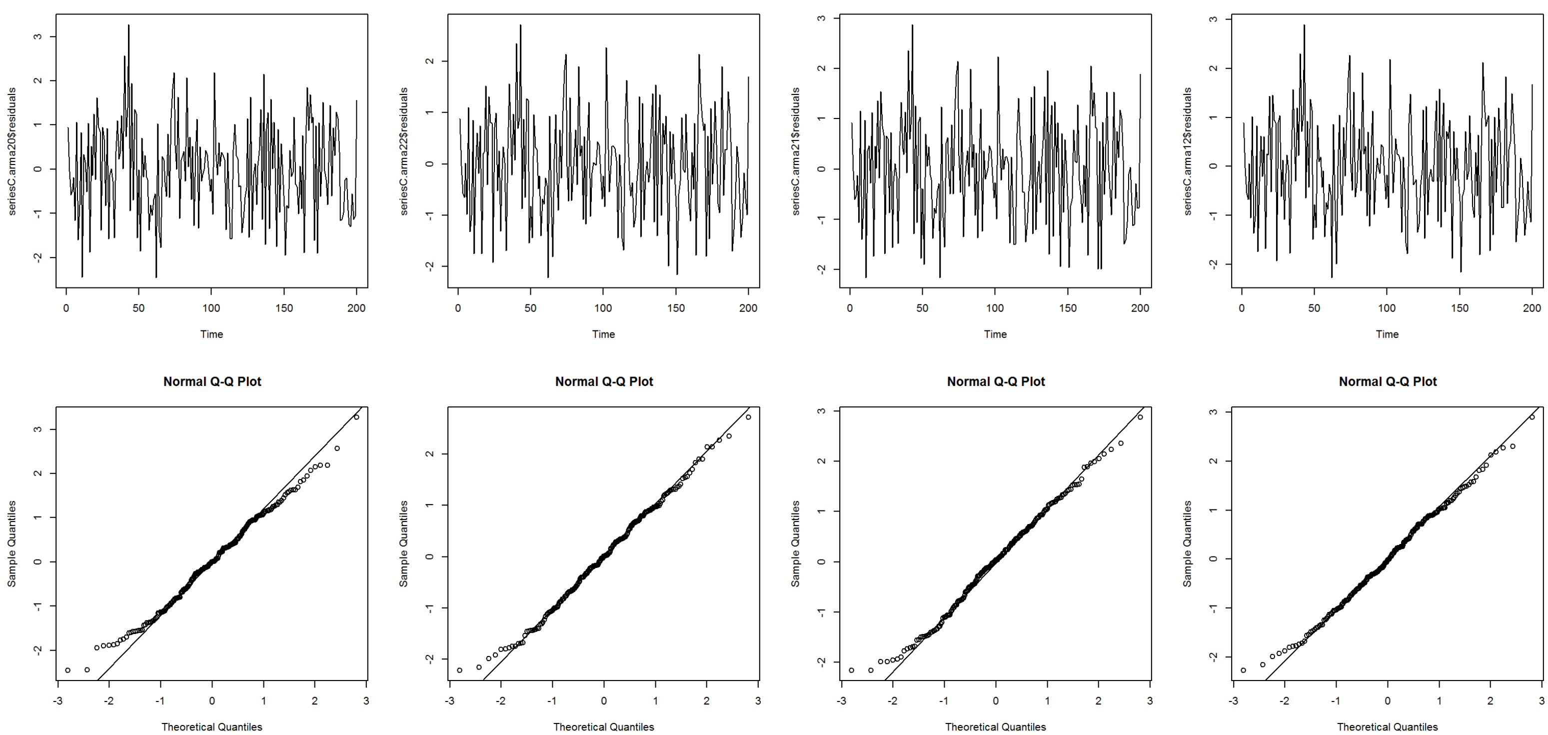
```
## [1] 0.7958741
```

```
shapiro.test(seriesB.arma22$residuals)$p.value
```

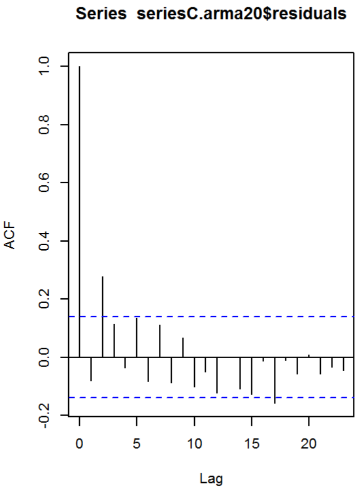
```
## [1] 0.9082675
```

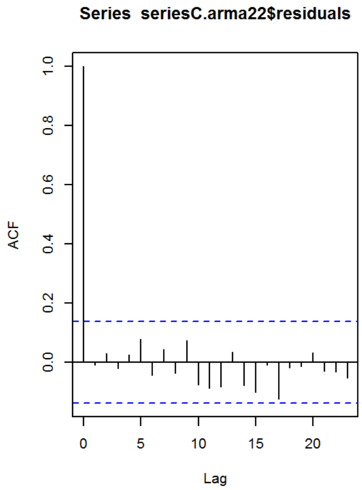
Similar to a), all models seem apporiate. Their residuals all look like white noise and form a stright line on the Q-Q plot. The ACF plot shows one spike on when $h = 0$, and no spike on PACF. The shapiro test also shows that the residuals are insignificant.

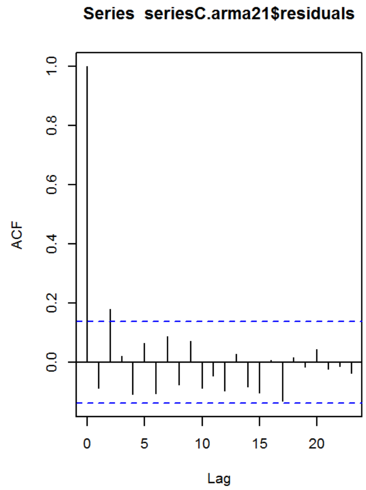
```
par(mfrow=c(2,4))
plot(seriesC.arma20$residuals)
plot(seriesC.arma22$residuals)
plot(seriesC.arma21$residuals)
plot(seriesC.arma12$residuals)

qqnorm(seriesC.arma20$residuals)
qqline(seriesC.arma20$residuals)
qqnorm(seriesC.arma22$residuals)
qqline(seriesC.arma22$residuals)
qqnorm(seriesC.arma21$residuals)
qqline(seriesC.arma21$residuals)
qqnorm(seriesC.arma12$residuals)
qqline(seriesC.arma12$residuals)
```

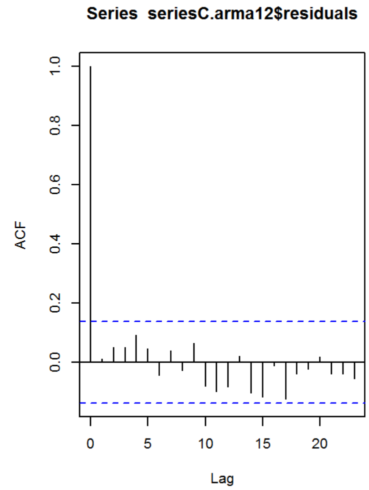


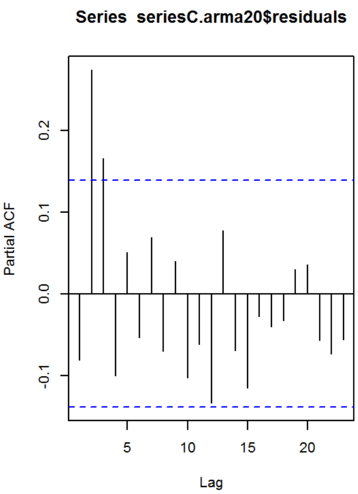

```
par(mfrow=c(2,4))
acf(seriesC.arma20$residuals)
acf(seriesC.arma22$residuals)
acf(seriesC.arma21$residuals)
acf(seriesC.arma12$residuals)
pacf(seriesC.arma20$residuals)
pacf(seriesC.arma22$residuals)
pacf(seriesC.arma21$residuals)
pacf(seriesC.arma12$residuals)
```

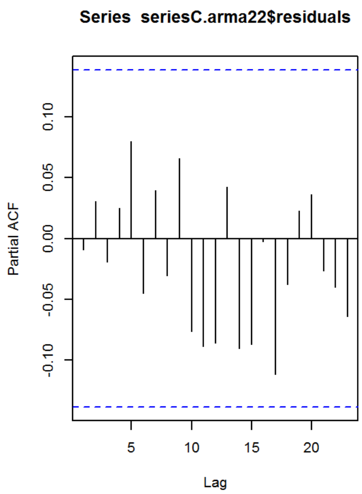


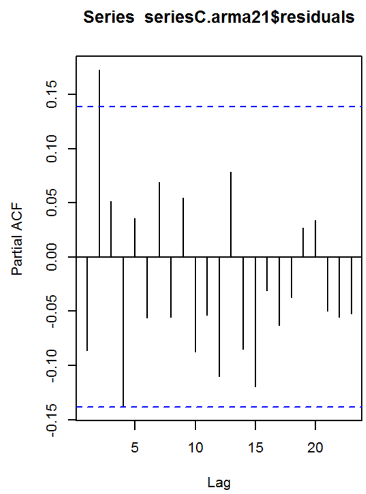


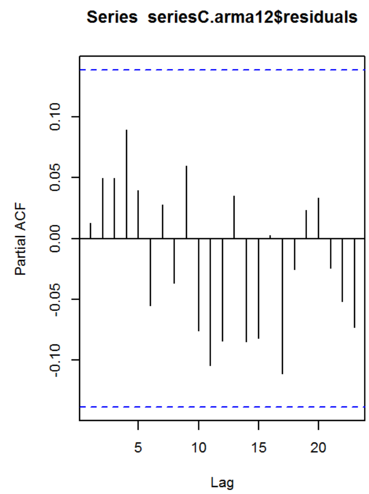













```
shapiro.test(seriesC.arma20$residuals)$p.value
```

```
## [1] 0.419945
```

```
shapiro.test(seriesC.arma22$residuals)$p.value
```

```
## [1] 0.61759
```

```
shapiro.test(seriesC.arma21$residuals)$p.value
```

```
## [1] 0.3503084
```

```
shapiro.test(seriesC.arma12$residuals)$p.value
```

```
## [1] 0.7206355
```

The arma20 and arma 12 model might be rejected, as the model does not seem to fit well. There is a heavy lower tail for the two models compared to the other ones. And there are some spikes other than $h = 0$ on the ACF&PACF plots. The shapiro test looks okay, as all the models indicate insignificant residuals.

e) conclusion

```
seriesA.arma10$aic # this one has the smallest aic among all seriesA models
```

```
## [1] 580.859
```

```
seriesA.arma20$aic
```

```
## [1] 582.2541
```

```
seriesA.arma11$aic
```

```
## [1] 582.3103
```

```
seriesB.arma12$aic
```

```
## [1] 567.3239
```

```
seriesB.arma21$aic
```

```
## [1] 567.2958
```

```
seriesB.arma11$aic # this one has the smallest aic among all seriesB models
```

```
## [1] 565.6496
```

```
seriesB.arma22$aic
```

```
## [1] 569.2957
```

```
seriesC.arma22$aic # this one has the smallest aic for seriesC models
```

```
## [1] 577.1532
```

```
seriesC.arma21$aic
```

```
## [1] 588.3907
```

Therefore, we choose $AR(1)$ $(i.e., ARMA(1,0))$ for seriesA, $ARMA(1,1)$ for seriesB, and $ARMA(2,2)$ for seriesC.