# HW7 Jin Kweon (3032235207)

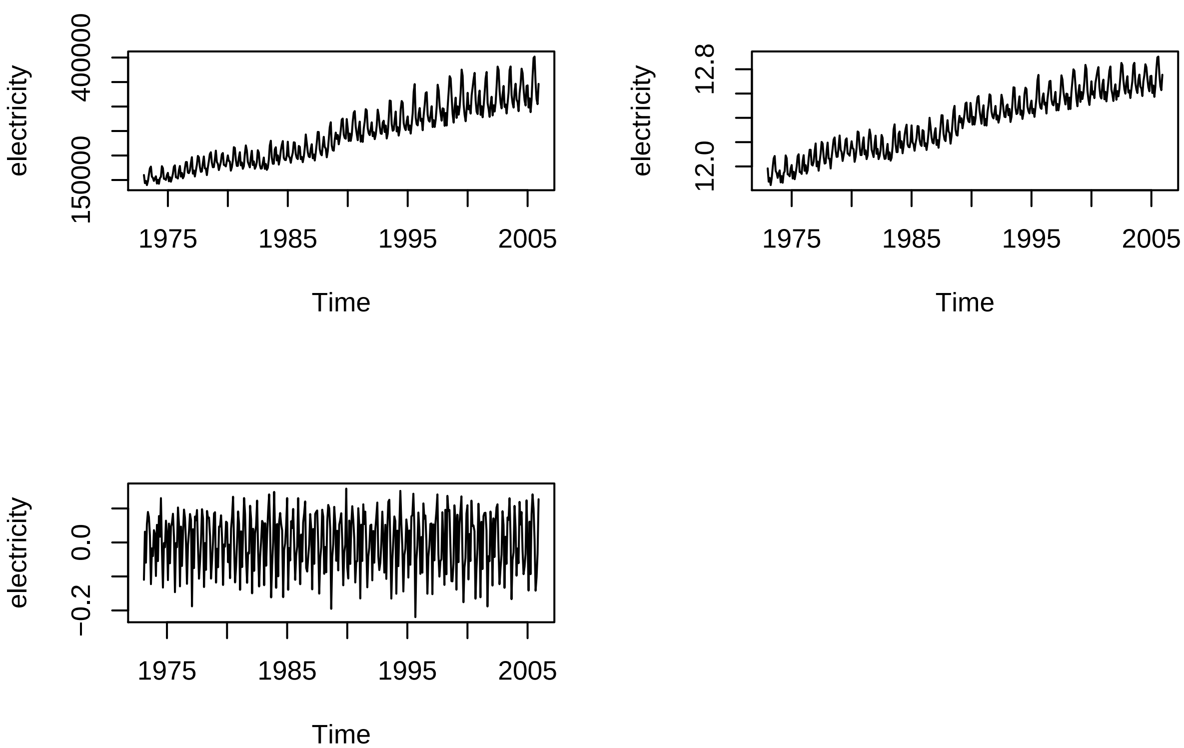
*Jin Kweon*

*4/10/2018*

## 1

### a) Construct a time series plot of the first difference of the logarithms of the electricity values. Does a stationary model seem warranted at this point?

```r
data("electricity")
#summary(electricity)
#str(electricity)

par(mfrow = c(2, 2))
plot.ts(electricity)
plot.ts(log(electricity))
plot.ts(diff(log(electricity)))
```



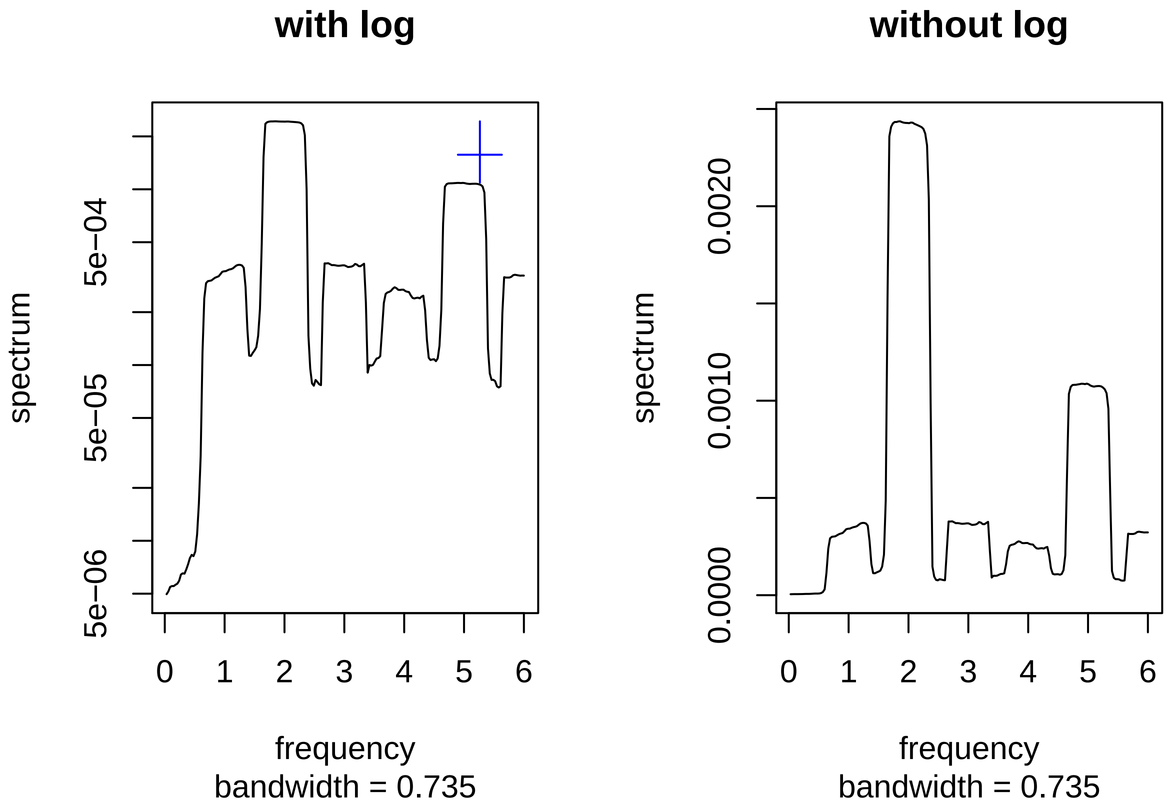

**Comment:**

The electricity production values gradually increase, and variance is also increasing as the time goes. (so, we have to apply log for power transformation)

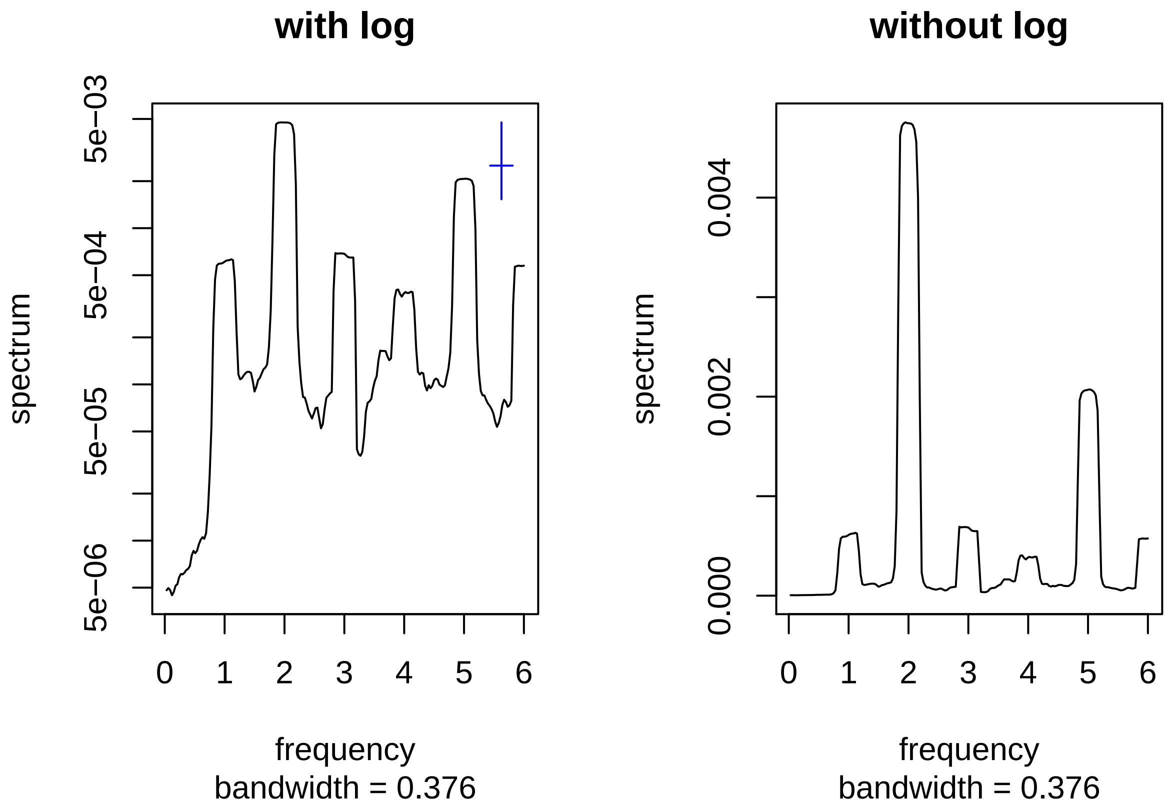
After taking log and do first differencing, the data looks stationary.

**b) Display the smoothed spectrum of the first difference of the logarithms using a modified Daniell spectral window and span values of 25, 13, and 7.**

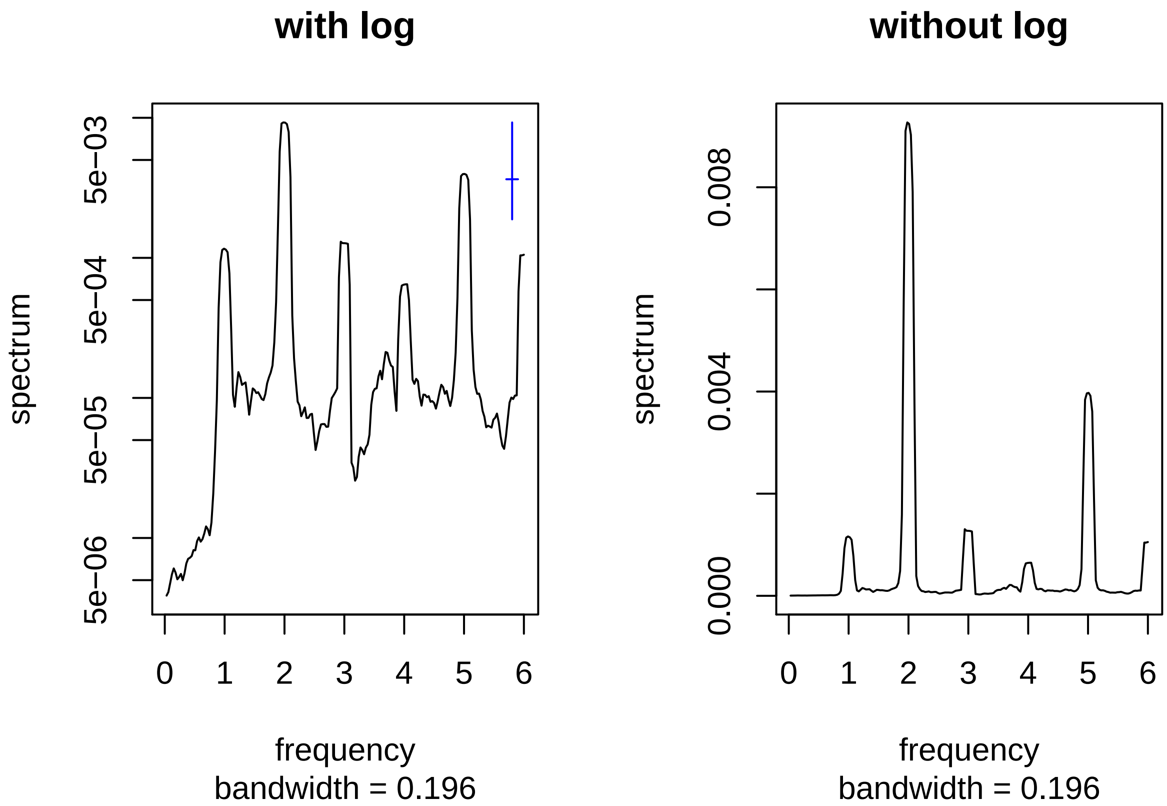
```
new1 <- diff(log(electricity))

#25
par(mfrow = c(1,2))
k = kernel("modified.daniell", 12)
mvspec(new1, kernel = k, main = "with log")
mvspec(new1, kernel = k, log = "no", main = "without log")
```




```
# mvspec(new1, spans = 25) -> same code but different way::: 12 * 2 + 1 = 25

#13
par(mfrow = c(1,2))
k = kernel("modified.daniell", 6)
mvspec(new1, kernel = k, main = "with log")
mvspec(new1, kernel = k, log = "no", main = "without log")
```




```r
#7
par(mfrow = c(1,2))
k = kernel("modified.daniell", 3)
mvspec(new1, kernel = k, main = "with log")
mvspec(new1, kernel = k, log = "no", main = "without log")
```



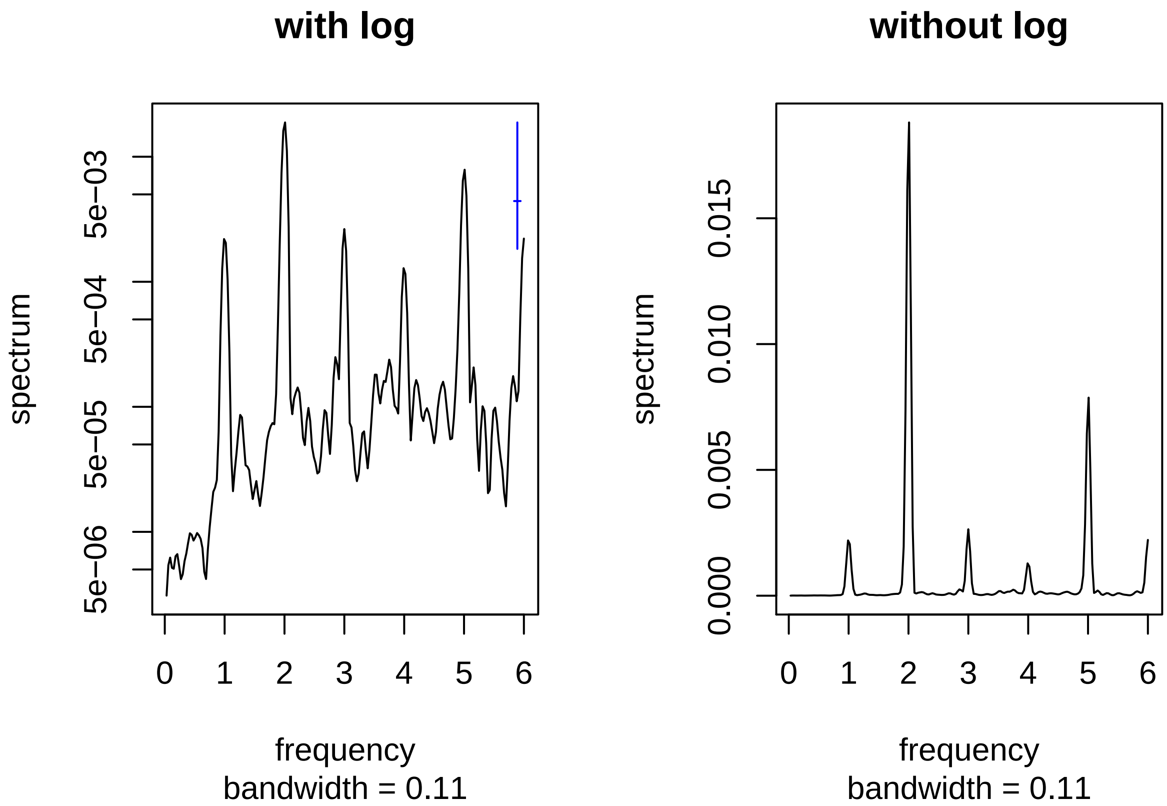

**Comment:**

**Note) I did not scale frequency to between 0 and 0.5, since it is not really necessary, but finding main peak/frequency is the most important. But, in part e), I scaled to 0 and 0.5!!!**

On pg 201, it says that spans is a vector of odd integers, given in terms of L = 2m + 1 instead of m.

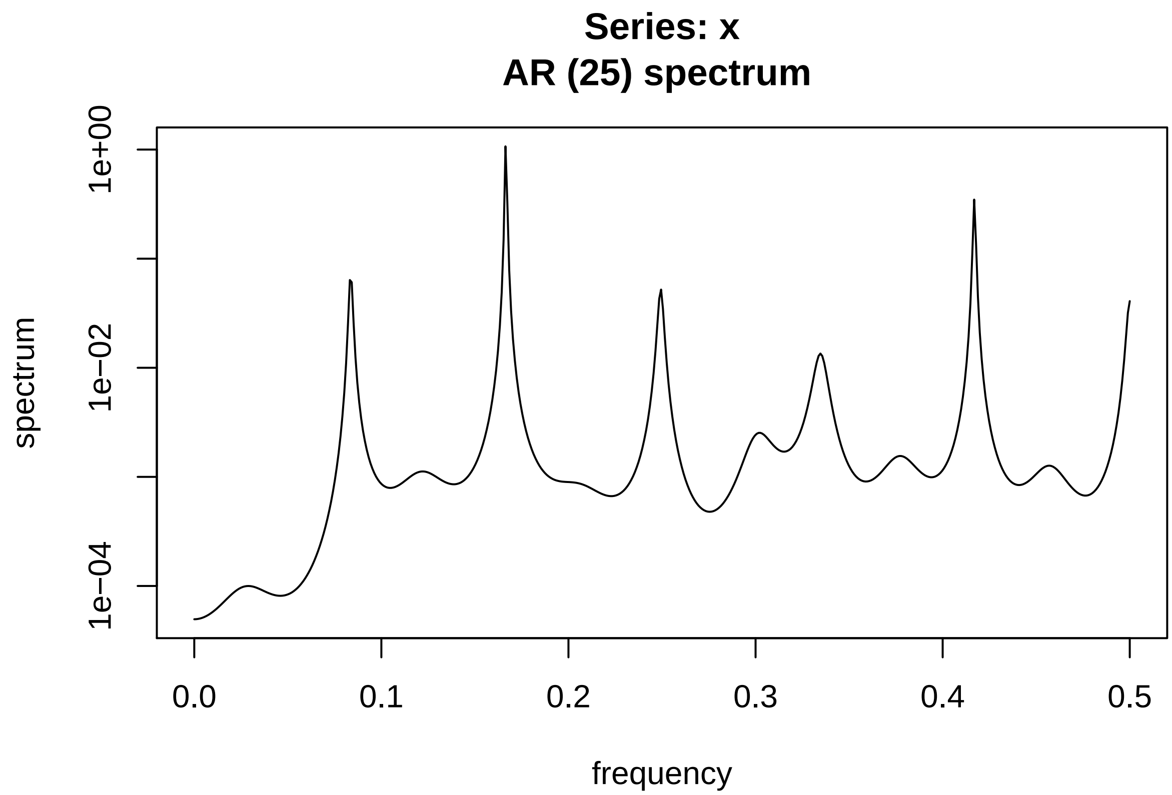
### c) Now use a spectral window that is a convolution of two modified Daniell windows each with span = 3. Also use a 10 percent taper.

```
par(mfrow = c(1,2))
mvspec(new1, taper = 0.1, spans = c(3, 3), main = "with log")
mvspec(new1, taper = 0.1, spans = c(3, 3), log = "no", main = "without log")
```




**d) Estimate the spectrum using an AR model with the order chosen to minimize the AIC. What order was selected?**

```
spec(new1, method = "ar")$method
```




```
## [1] "AR (25) spectrum "
```

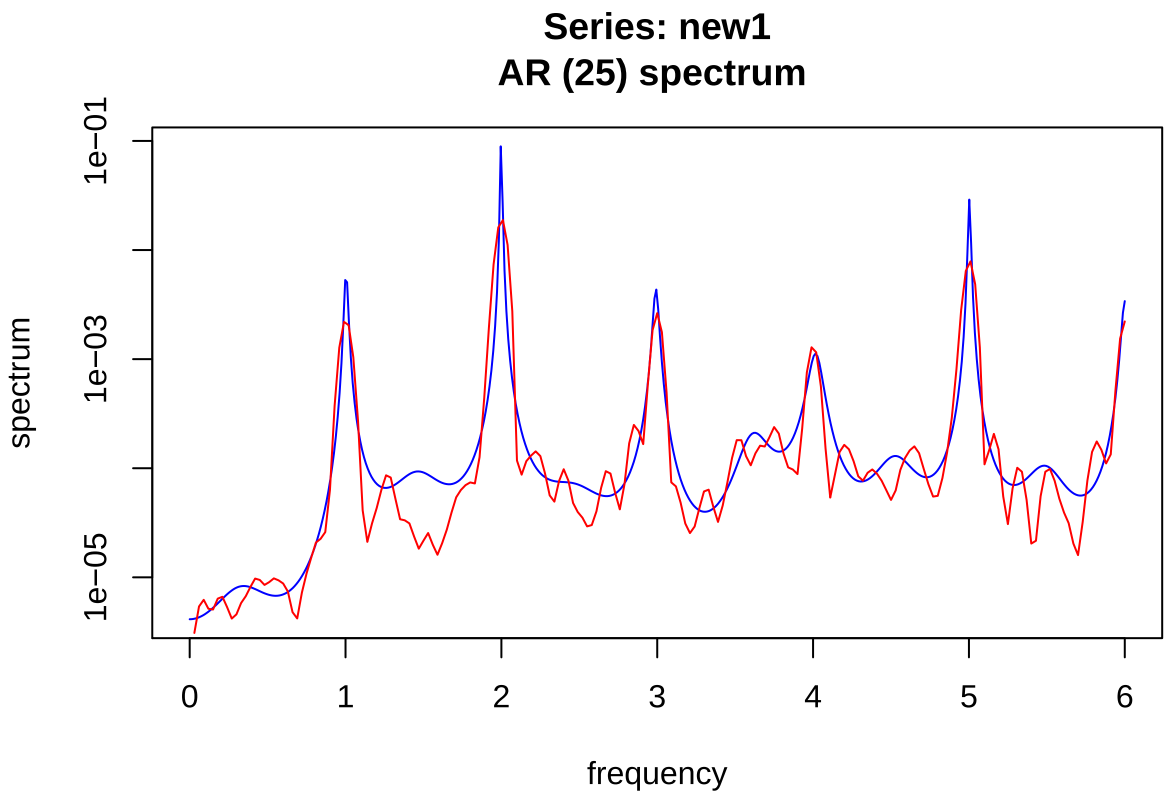
```r
#spec(new1, method = "ar", log = "no", main = "without log")
#spec(new1, method = "ar", main = "with log")

#spec.ar(new1) #other method
```

**Comment:**

The order will be 25.

### e) Overlay the estimates obtained in parts (c) and (d) above onto one plot. Do they agree to a reasonable degree?

```r
#x <- spec(new1, method = "ar")
spec.ar(new1, col = "blue")
spec.pgram(new1, taper = 0.1, spans = c(3, 3), col = "red", add = T)
```




```
#lines(c$freq / 12, c$spec + 0.0001, lty = "dotted", col = "red") -> other method
```

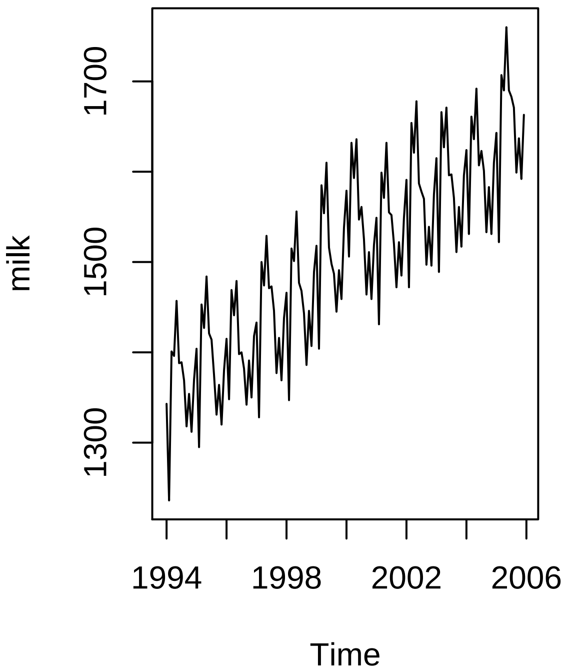
**Comment:**

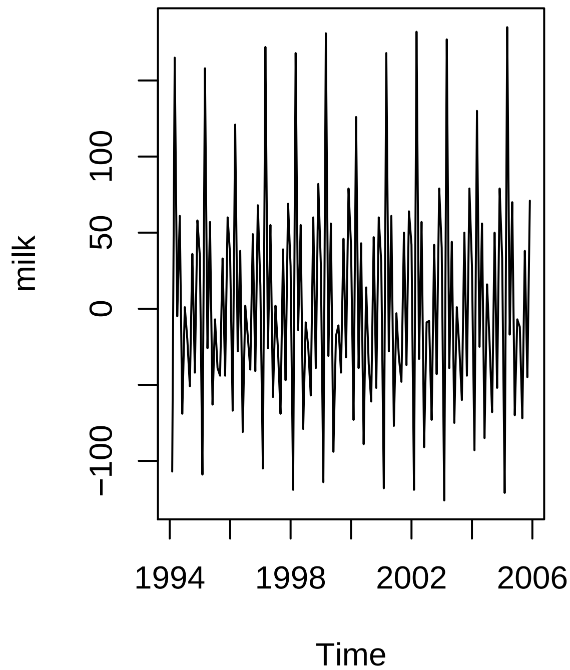
They agree with each other. (pattern looks similar)

## 2

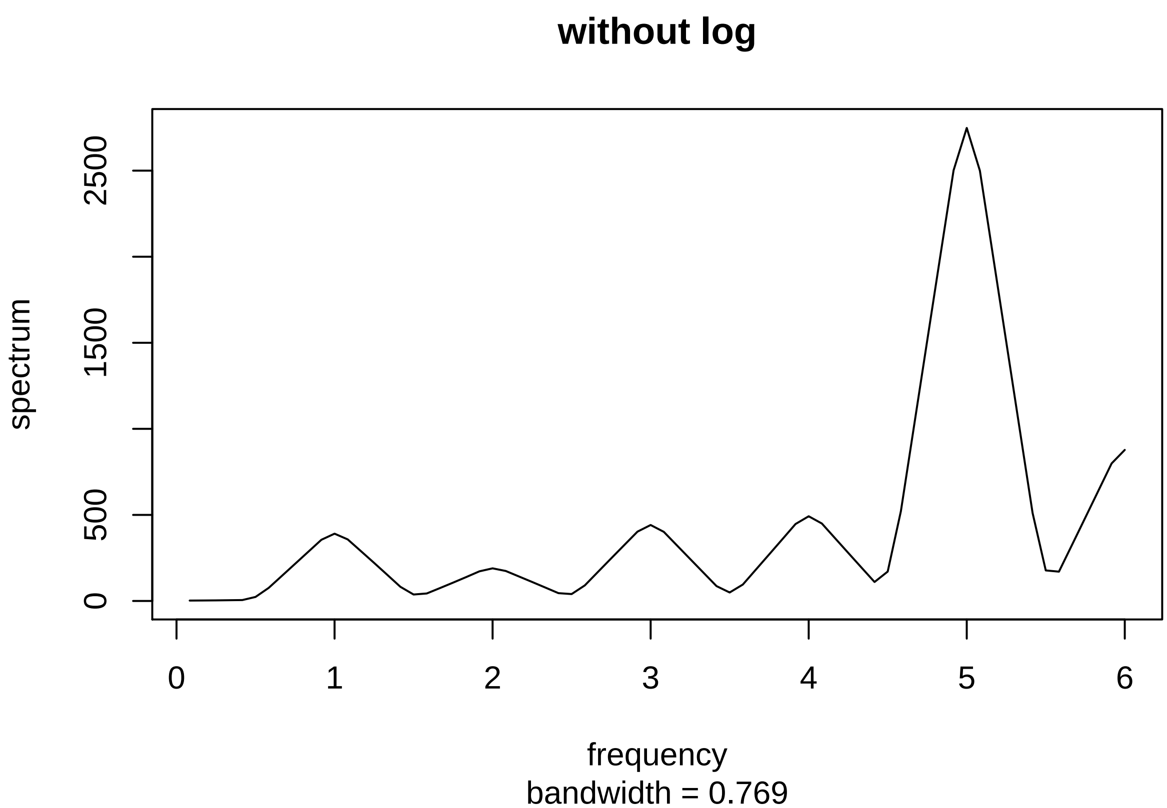
### a) Estimate the spectrum using a spectral window that is a convolution of two modified Daniell windows each with span = 7

```
data("milk")
new2 <- diff(milk)
par(mfrow = c(1,2))
plot.ts(milk)
plot.ts(new2)
```



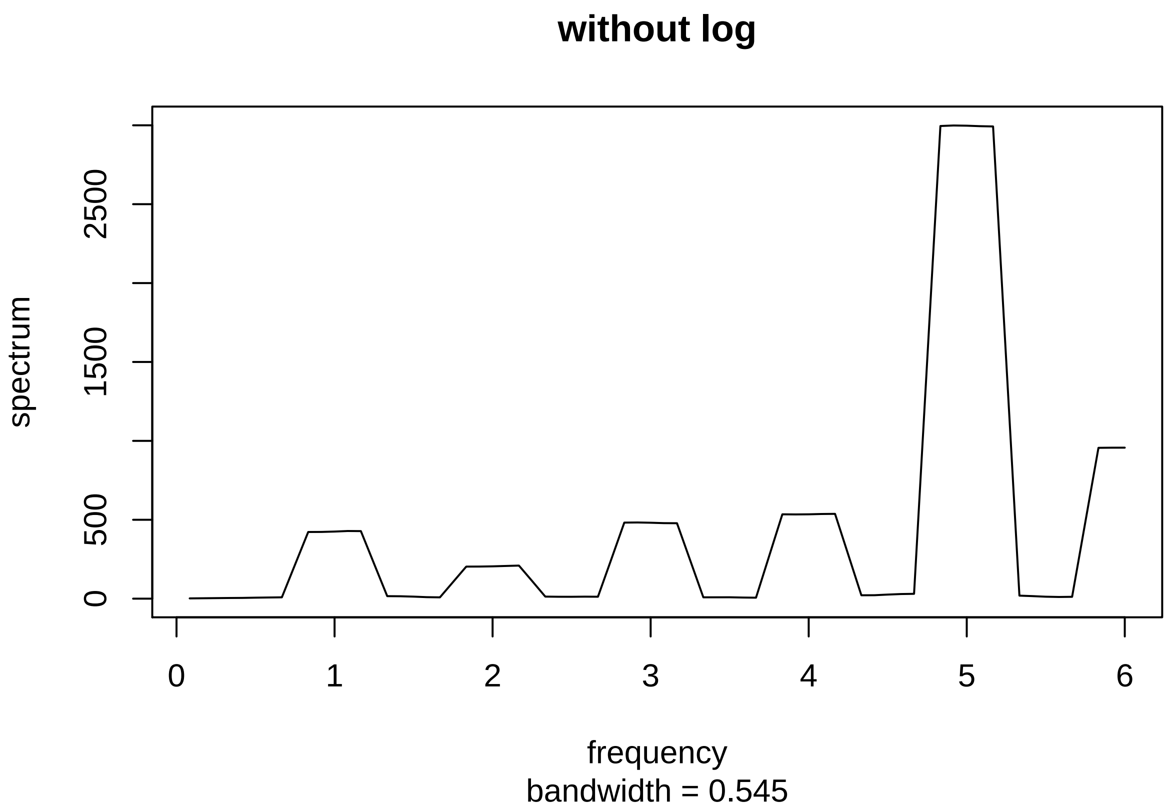




```r
#spec.pgram(milk, spans = c(7, 7))
#mvspec(new2, spans = c(7, 7), main = "with log")
par(mfrow = c(1,1))
k = kernel("modified.daniell", c(3,3))
mvspec(new2, spans = k, log = "no", main = "without log")
```



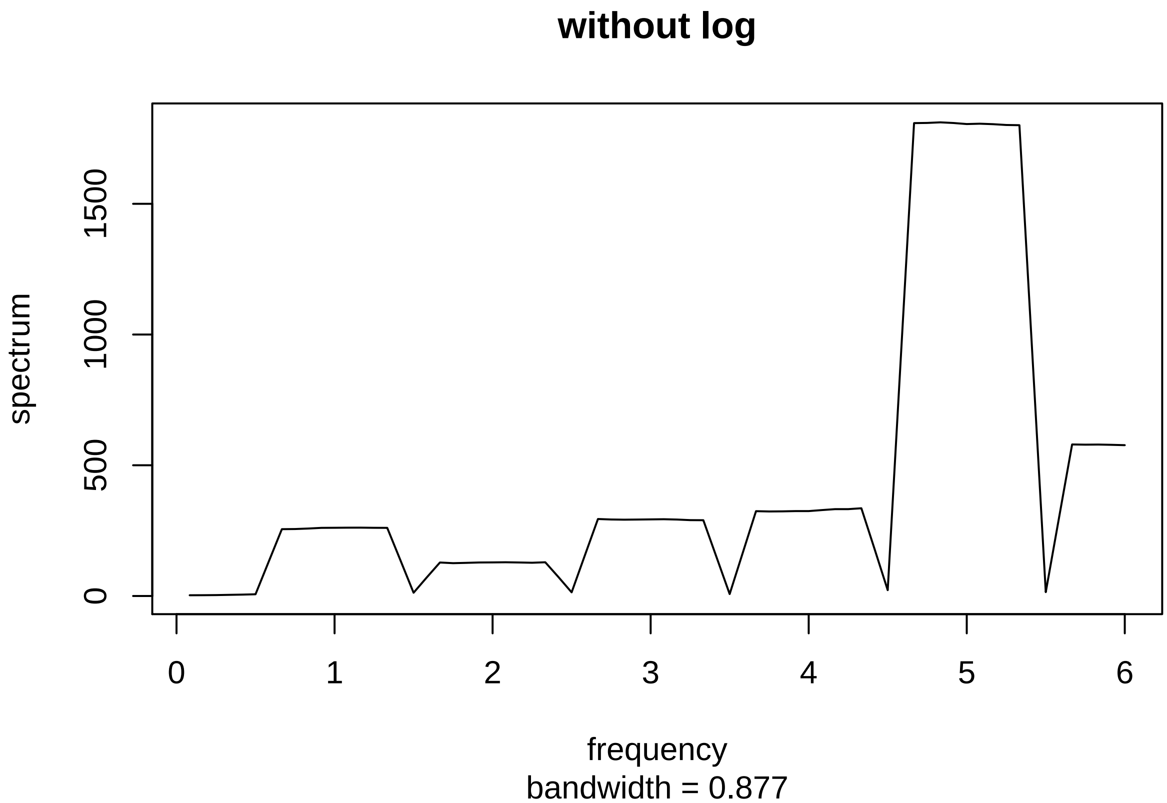
**b) Estimate the spectrum using a single modified Daniell spectral window with span = 7.**

```
#mvspec(new2, spans = 7, main = "with log")
par(mfrow = c(1,1))
k = kernel("modified.daniell", 3)
mvspec(new2, spans = k, log = "no", main = "without log")
```




**c) Finally, estimate the spectrum using a single modified Daniell spectral window with span = 11. Compare these results with those shown in parts (a) and (b).**

```
#mvspec(new2, spans = 11, main = "with log")
k = kernel("modified.daniell", 5)
mvspec(new2, spans = k, log = "no", main = "without log")
```




**Comment:**

Too much spanning makes me hard to find the main peak and frequency. (as it blurred out/averaged out)

### d) Among the three different estimates considered here, which do you prefer and why?

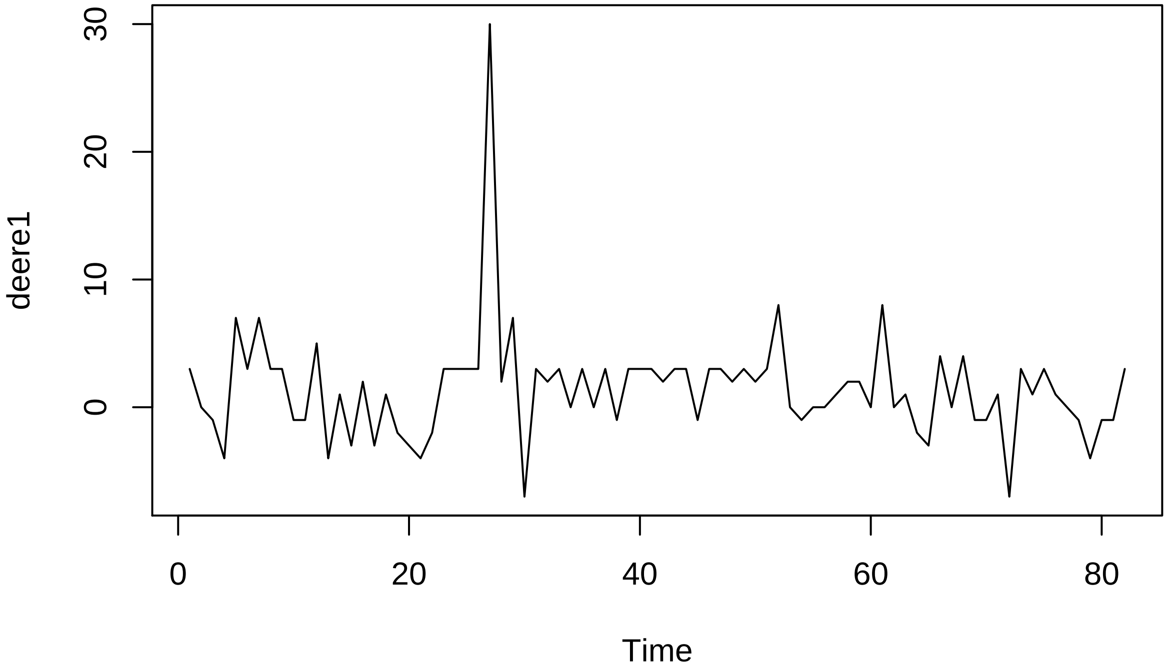
**Comment:**

I prefer part a)'s method, as the convolution really helped to find the main peak. (by giving only half weight to the end points)

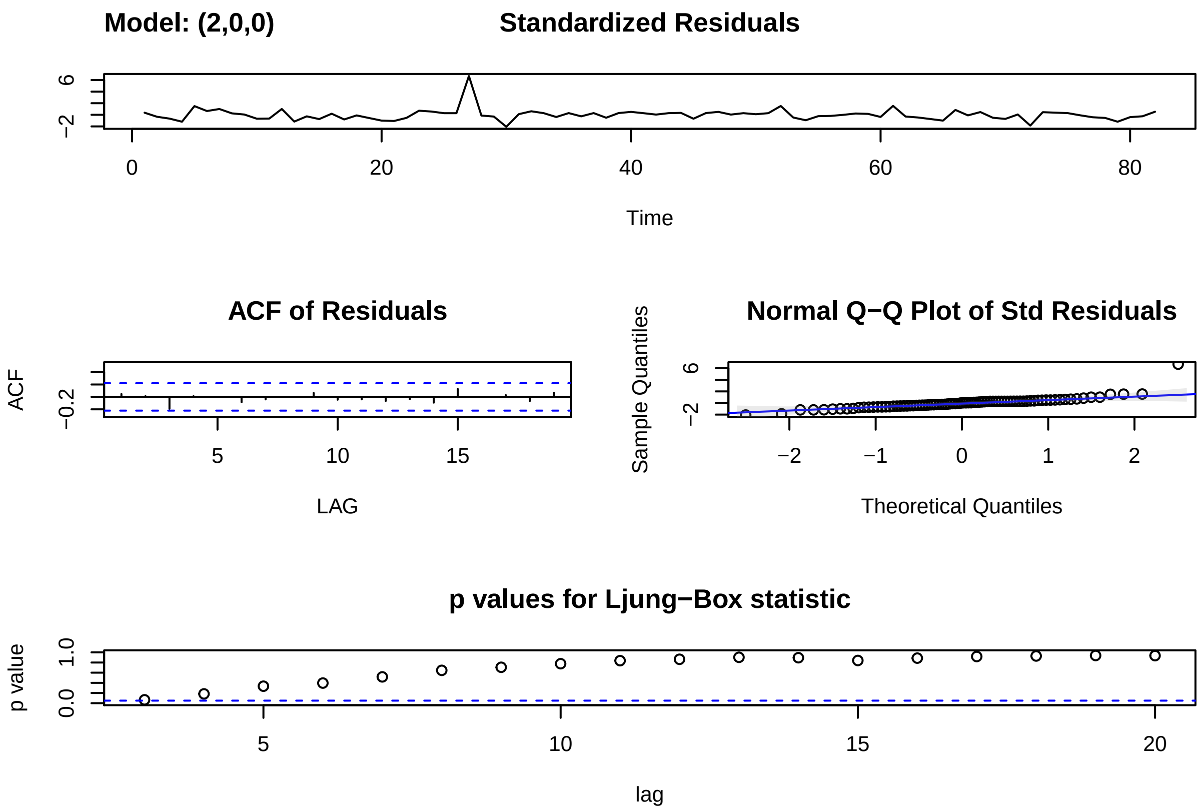
## 3

**a) Fit an AR(2) model to the full data set. Plot the standardized residuals from this model, and the sample ACF of the residuals. What do these diagnostics tell you?**

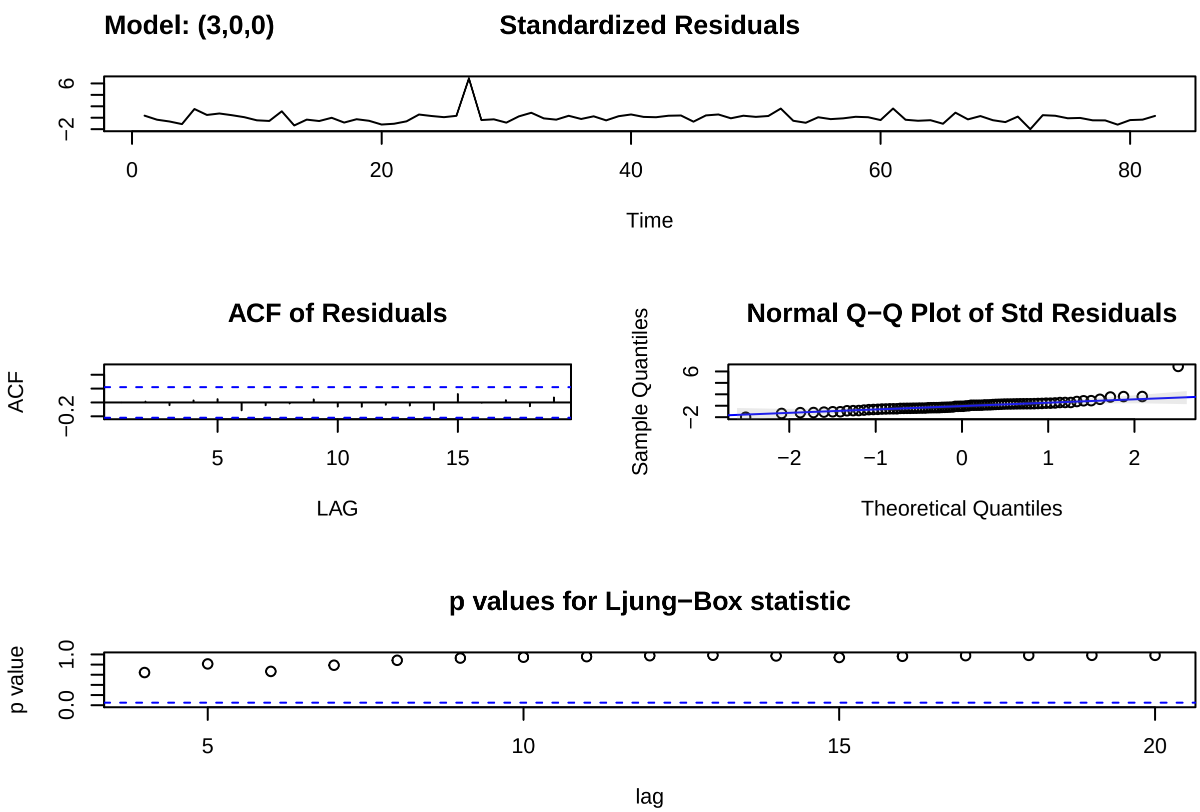
```r
data("deere1")

plot.ts(deere1)
```




```r
#eacf(deere1)
#spec(deere1, method = "ar") -> gives you AR(3)
sarima(deere1, 2, 0, 0)
```




```r
sarima(deere1, 3, 0, 0)
```




**Comment:**

Most of them looked fine except one unexpected peak at time t = 27. But, in overall, most of the diagnostics tell me that AR(2) model fits pretty well into this data set.

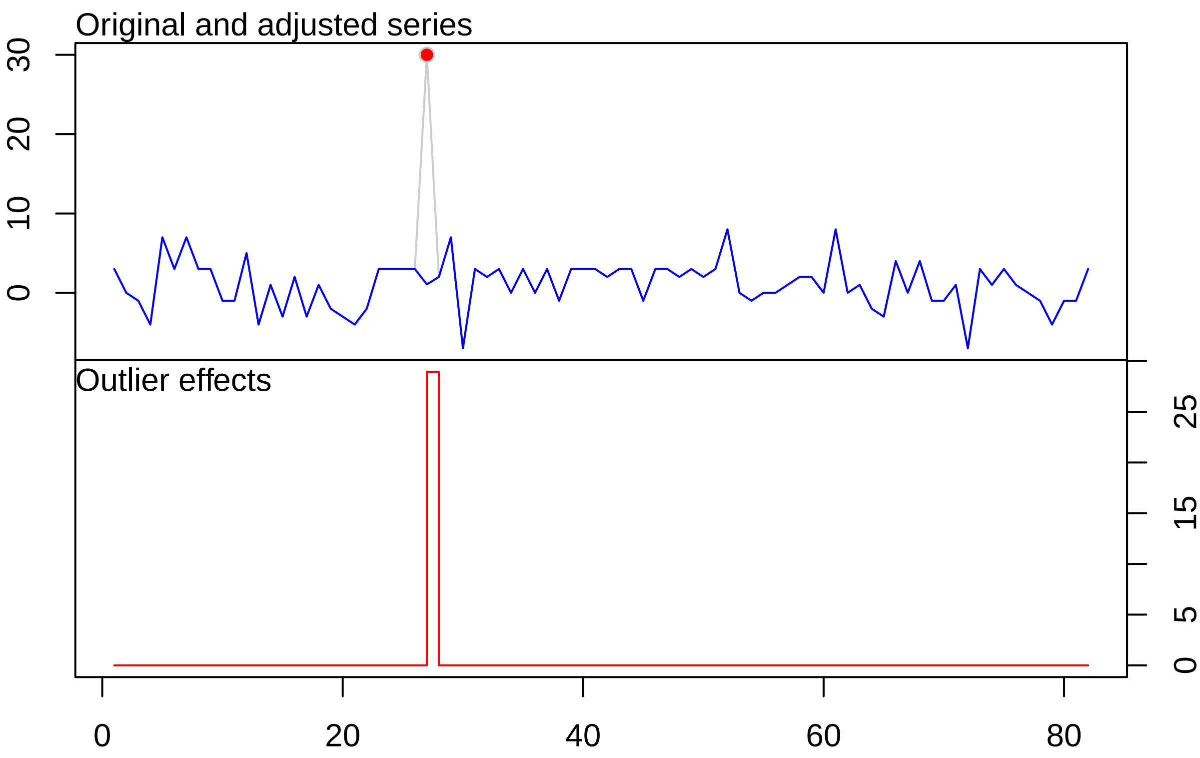
And, by the way, AR(3) model seems better fit.

### b) Detect either additive outliers and/or innovative outliers from the model in (a). What is your conclusion?

```
outs <- tso(deere1, types = c("TC", "AO", "LS", "IO", "SLS"))
tsoutliers(deere1)
```

```
## $index
## [1] 27
##
## $replacements
## [1] 2.5
```

```
plot(outs)
```

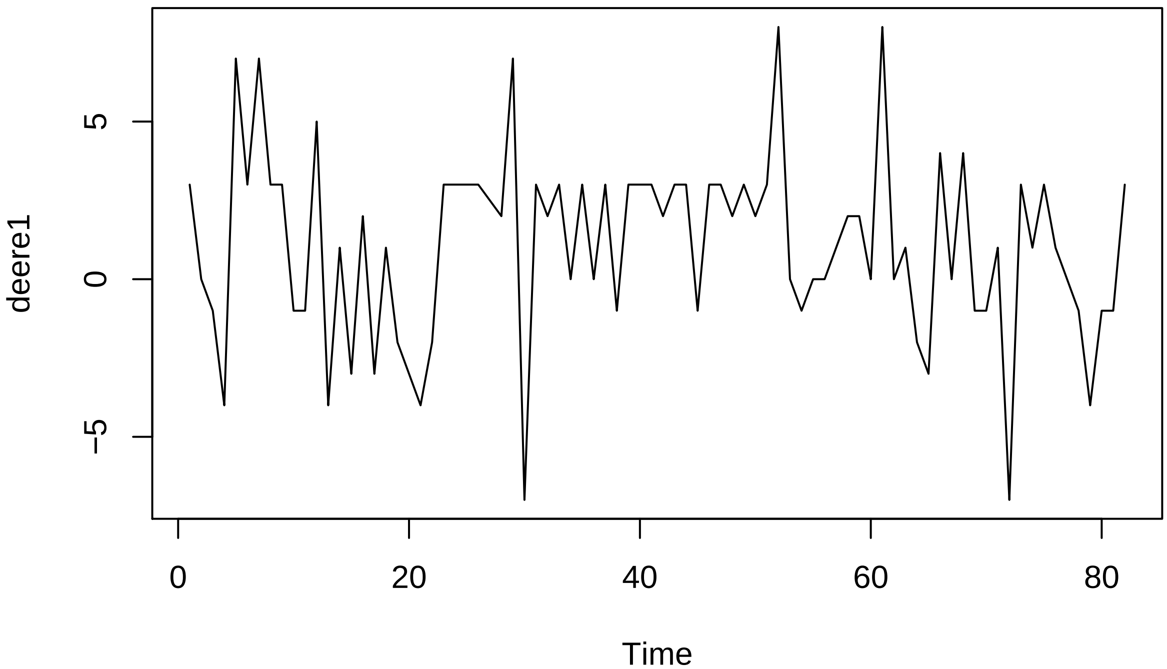



**Comment:**

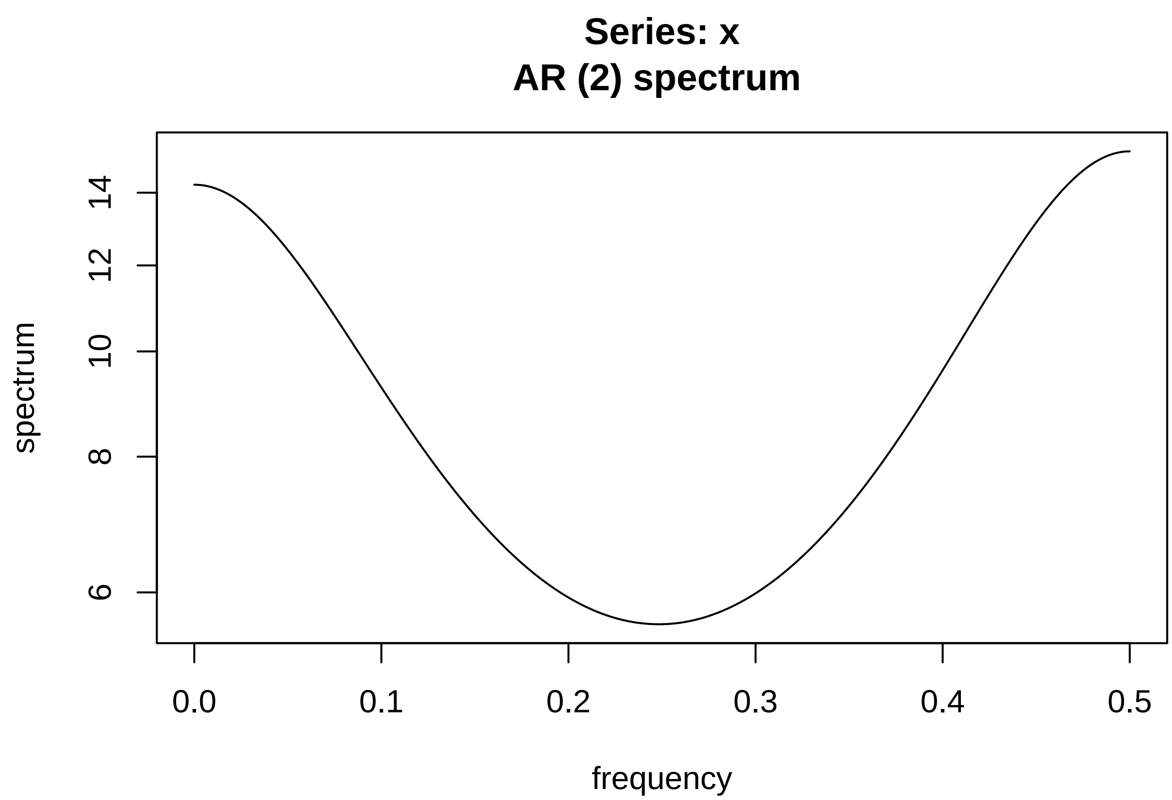
As I mentioned in part a), time t = 27 is an outlier, and coudl be replaced to 2.5 using tsoutliers function.

**c) Fit an AR(2) model that incorporates the most notable outlier into the model. Plot the standardized residuals from this model, and the sample ACF of the residuals. What do these diagnostics tell you? Compare the fitted model in part (a) to the fitted model in part (c).**

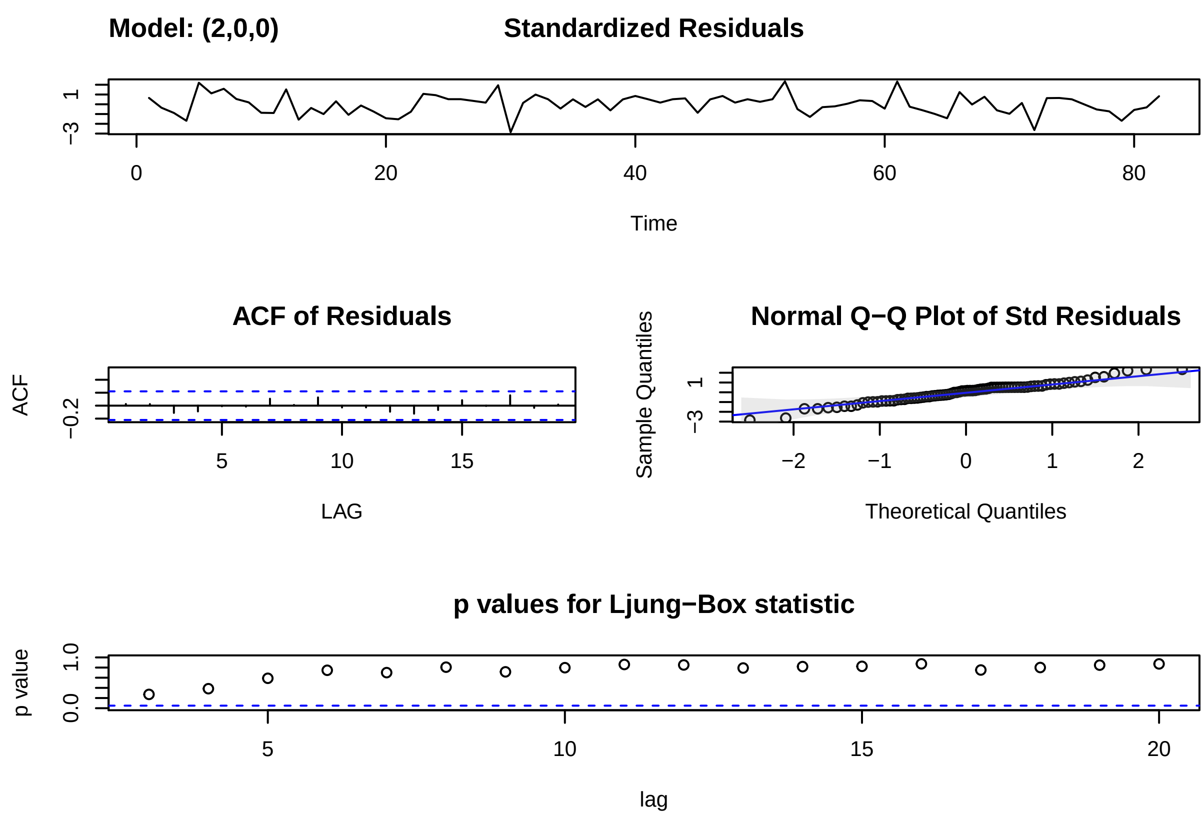
```
deere1[27] <- tsoutliers(deere1)$replacements
plot.ts(deere1)
```




```
spec(deere1, method = "ar")
```




```r
sarima(deere1, 2, 0, 0)
```




**Comment:**

It (model diagnostics) looks really good and AIC is lower, compared to part a).