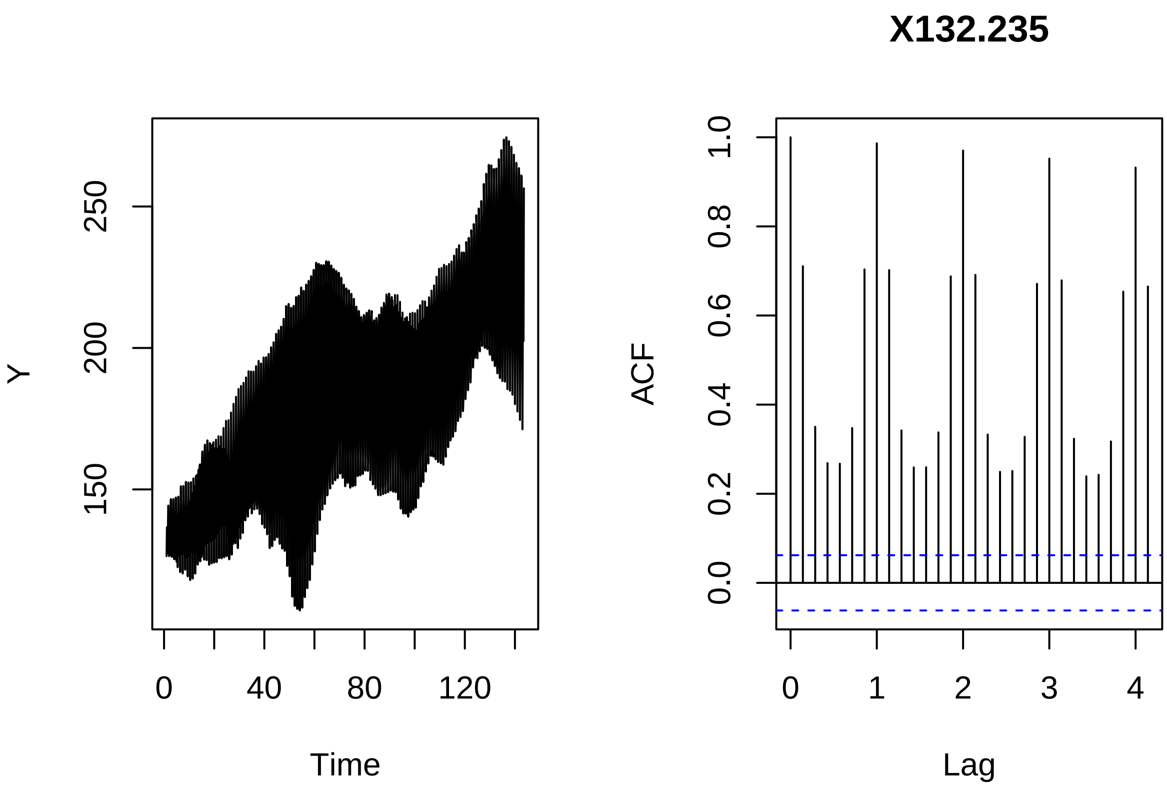
Q1)

a)

```
data <- read.csv("ElectricityConsumption.csv")
train_days <- 1:(130*7)
data.train <- data[train_days, ]
data.test <- data[-train_days, ] # how to keep the original time index

Y <- ts(data, frequency = 7)
par(mfrow=c(1,2))
# ts.plot(Y[0:200])
# ts.plot(Y[200:400])
ts.plot(Y)
#ts.plot(diff(Y))
acf(Y)
```




```
#acf(diff(Y))
```

From the timeseries plot, the process is *not stationary* because:

- Mean is not constant. The process is not centered around a constant mean, the mean seems to change with time.
- Variance is not constant. The variance seems like it could be constant in intervals $[0, 250]$ and $[475, 800]$ (roughly) but it definitely is not constant in the remaining time intevals.
- The autocovariance function is not scale-invariant. The sections with the darker lines seen in intervals $[0, 250]$ and $[675, 100]$ indicate that the process contracts in these regions and it expands in the others.

b)

Since log(Y) does not seem to stabilize the variance, we will try additive models only. But the best way is to try both.

```
Y.train = ts(data.train, frequency = 7)
Y.test = ts(data.test, start = 131, frequency = 7)

# Alpha must be TRUE.
hw1 <- HoltWinters(Y.train, gamma=FALSE)
hw2 <- HoltWinters(Y.train, gamma=FALSE, beta=FALSE)
hw3 <- HoltWinters(Y.train, beta=FALSE) # error
hw4 <- HoltWinters(Y.train)

HW.1.predict = predict(hw1, n.ahead=89)
HW.2.predict = predict(hw2, n.ahead=89)
HW.3.predict = predict(hw3, n.ahead=89)
HW.4.predict = predict(hw4, n.ahead=89 , prediction.interval = TRUE , level=0.95)
## if you don't specify n.ahead it fits the entire data


APSE.1 = mean((HW.1.predict[,1]-Y.test)^2)
APSE.2 = mean((HW.2.predict[,1]-Y.test)^2)
APSE.3 = mean((HW.3.predict[,1]-Y.test)^2)
APSE.4 = mean((HW.4.predict[,1]-Y.test)^2)
```

Combination 1

$$(\alpha, \beta, \gamma) = (1, 1, 0)$$

```
APSE.1
```

```
## [1] 893.4296
```

Combination 2

$$(\alpha, \beta, \gamma) = (1, 0, 0)$$

```
APSE.2
```

```
## [1] 827.1388
```

Combination 3

$$(\alpha, \beta, \gamma) = (1, 0, 1)$$

```
APSE.3
```

```
## [1] 64.21847
```

Combination 4

$$(\alpha, \beta, \gamma) = (1, 1, 1)$$

```
APSE.4
```

```
## [1] 60.33638
```

Combination #4 produces the best model based on the lowest APSE. So pick this one.

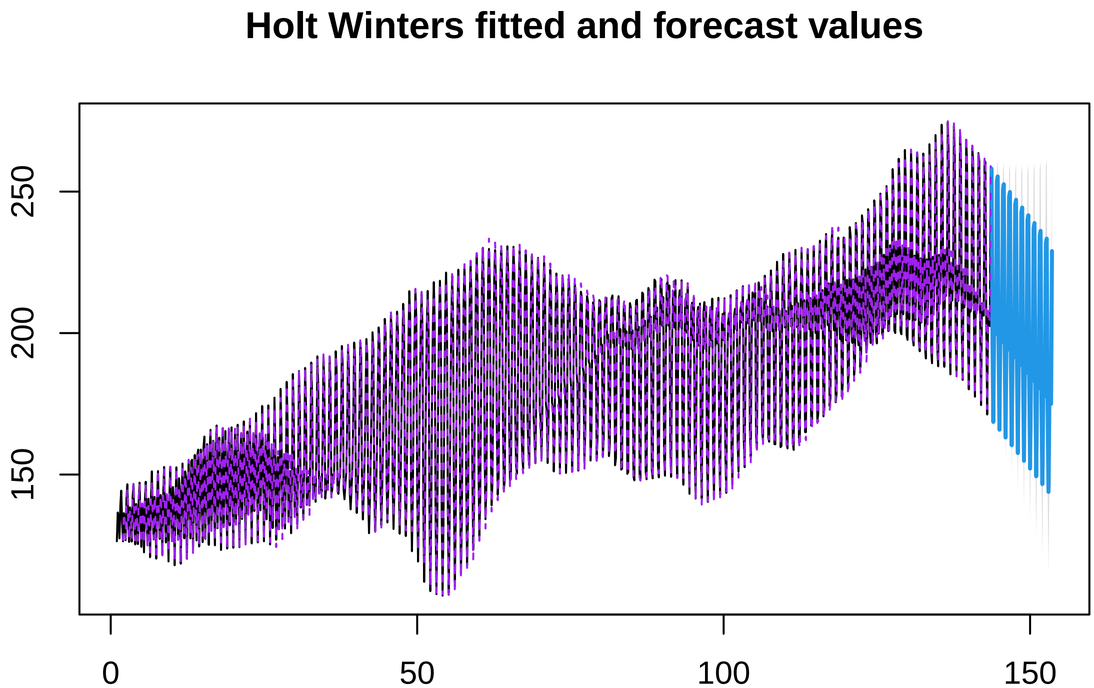
c)

```
library(forecast)
```

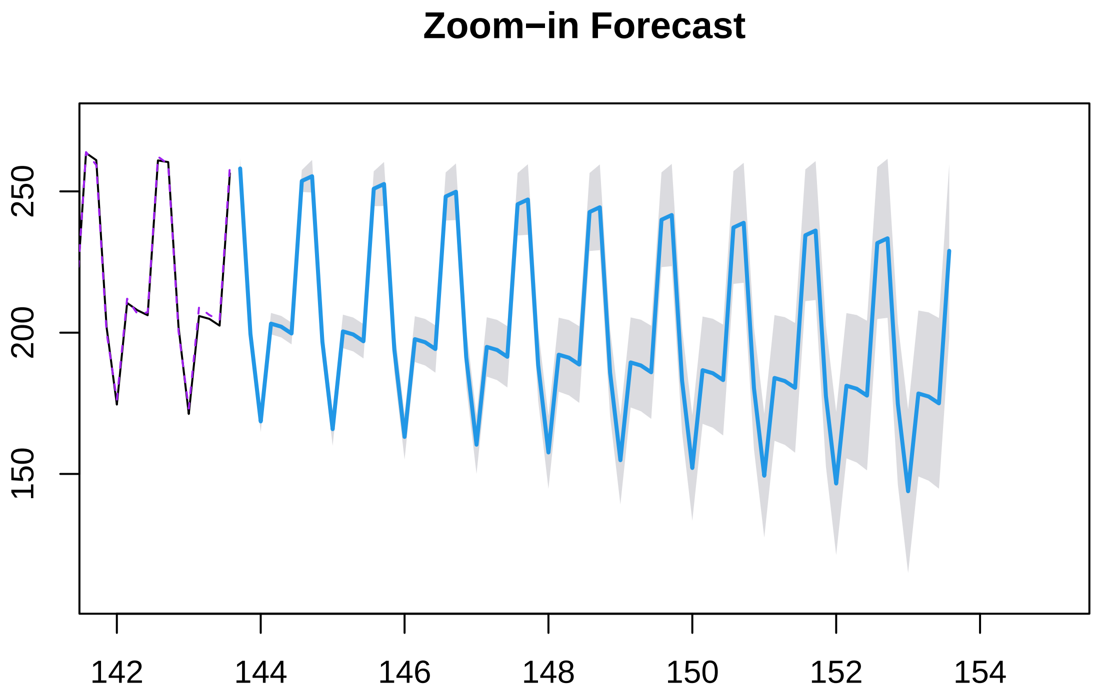
```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
```

```
hw <- HoltWinters(Y)
#HW.predict = predict(hw, n.ahead=70 , prediction.interval = TRUE , level=0.95)
HW <- forecast(hw, h=70, level = 95)

#visualize our predictions:
plot(HW, main = "Holt Winters fitted and forecast values")
lines(HW$fitted, lty=2, col="purple")
```

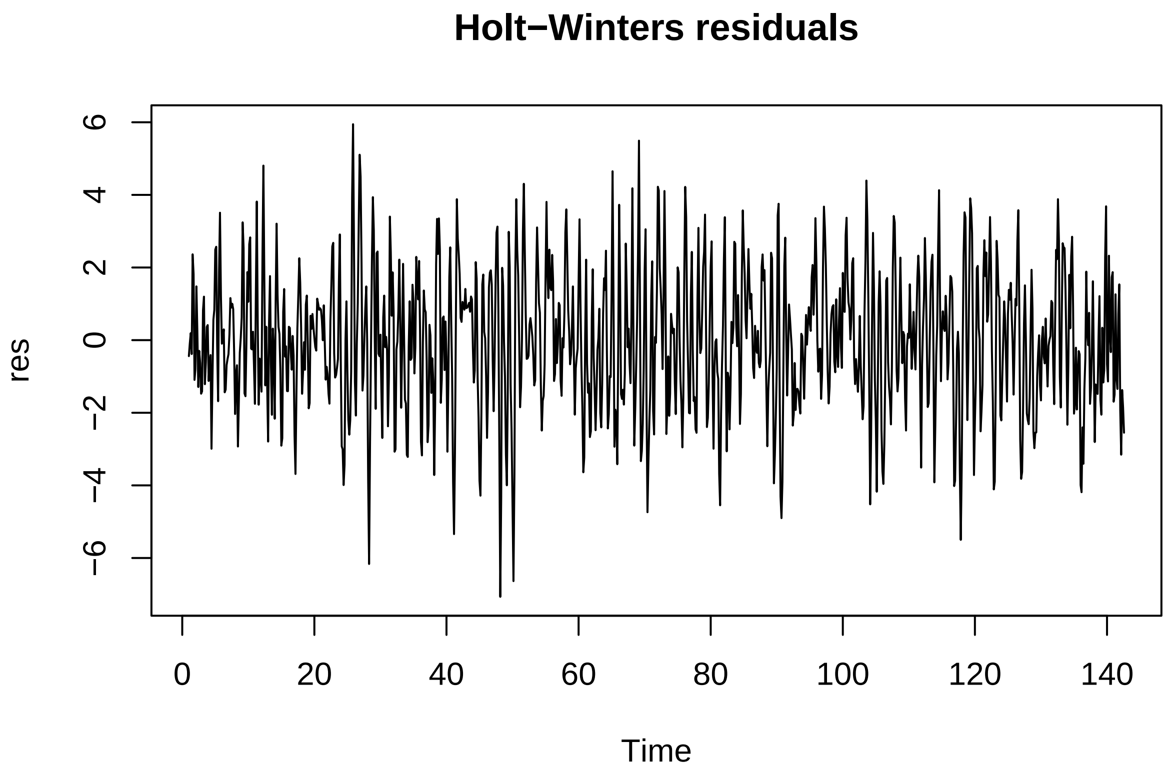



```
plot(HW, xlim=c(142, 155), main = "Zoom-in Forecast")
lines(HW$fitted, lty=2, col="purple")
```

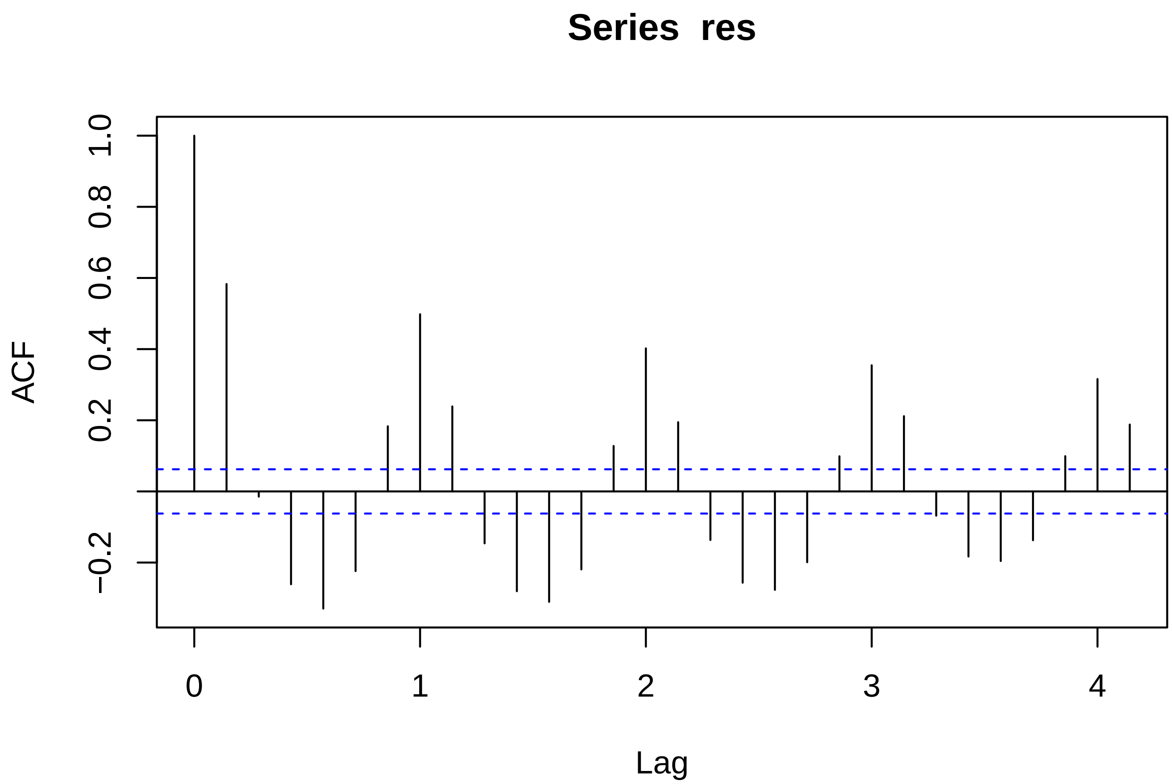

Zoom−in Forecast
250
200
150
142
144
146
148
150
152
154

d)

```r
res <- ts(HW$residuals[-seq(7),], frequency = 7) # removing NA's from lags in HW model
plot(res, main="Holt-Winters residuals")
```




```r
acf(res)
```




```
#par(mfrow=c(2,1))
#acf(diff(diff(res)), na.action = na.pass)
#plot(decompose(res))
```

Comments on stationarity:

- Visually, the residuals seem randomly distributed.
- They seem to have mean zero as they are distributed around it.
- The variance is finite and seems constant even if there is about three peaks slightly larger than the others.
- The ACF shows a periodic pattern so the residuals are not stationary.

Comments on White Noise

- It is not white noise because there seems to be a remaining pureley periodic pattern.