Time Series Analysis

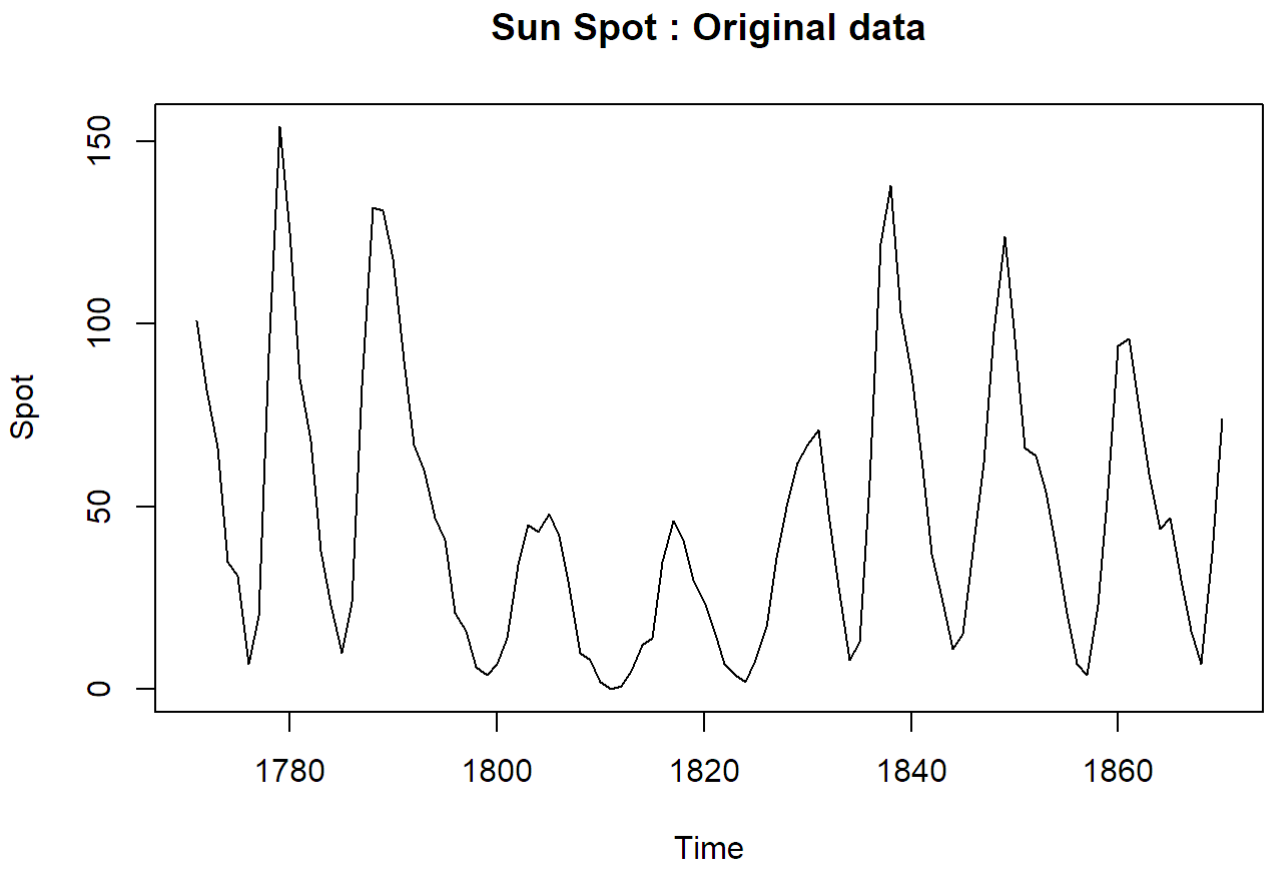
Homework2 201811526 이은주

```
spot = read.csv('spot.csv')
str(spot)
```

```
## 'data.frame':    100 obs. of  1 variable:
##  $ Spot: int  101 82 66 35 31 7 20 92 154 125 ...
```

1. Draw time series graph of sun spot. Do you see any cycle or seasonal effect?

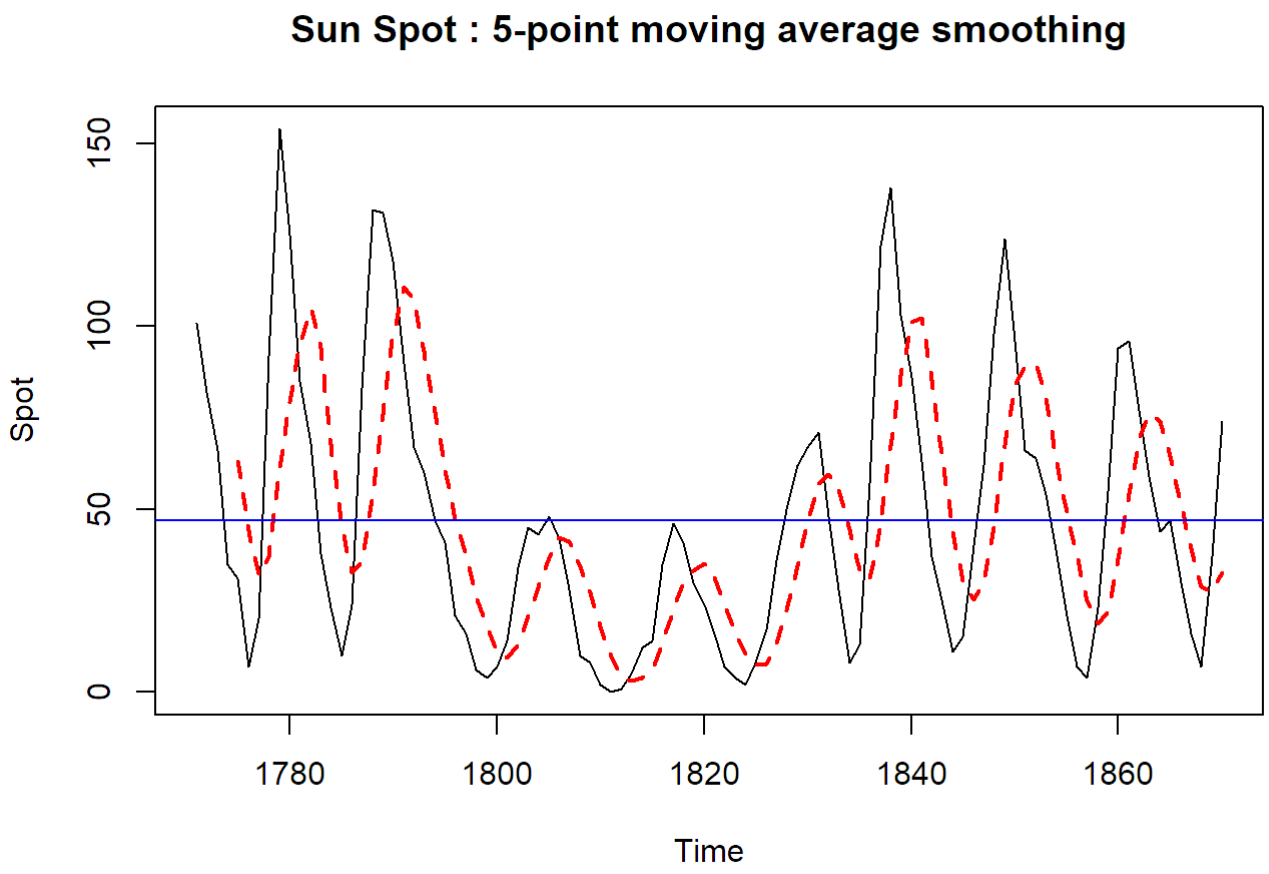
```
spot.ts = ts(spot, start=c(1771), frequency=1)
plot(spot.ts, main = "Sun Spot : Original data")
```



Sun spot의 graph에 따르면, 10년을 주기로 cycle effect가 있음을 볼 수 있다.

2. try 5-point moving average smoothing. Draw the plot of original graph in black, 5 point MA smoothing in red, mean value in blue.

```
m5 = filter(spot.ts, filter=rep(1/5,5), method="convolution", sides=1)
plot(spot.ts, main = "Sun Spot : 5-point moving average smoothing") +
  lines(m5, col="red", lty=2, lwd=2) +
  abline(h=sum(spot)/100, col="blue")
```

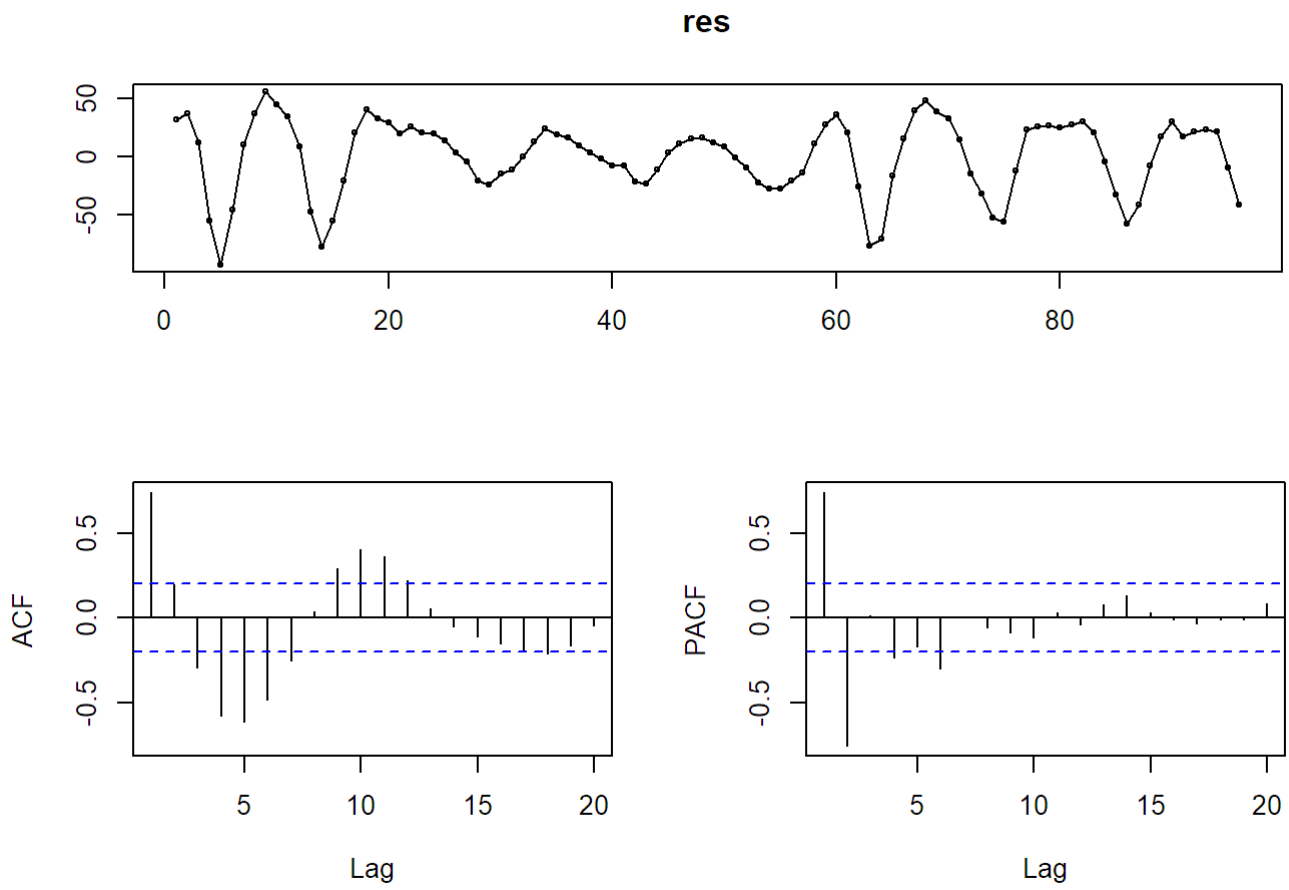



```
## integer(0)
```

Moving average smoothing을 통해 original data보다는 smoothing해진 것을 알 수 있다.

3. Check the residual plot, check the stationary and the test the independence assumption. Carefully interpret the residual analysis.

```
res = m5[-1:-4,] - spot.ts[-1:-4,]
tsdisplay(res)
```



Residual Plot : 완전히 random하지 않고 여전히 어떠한 pattern이 있는 것으로 보인다.
Stationary : ACF, PACF를 보았을 때, 유의수준 밖으로 벗어나는 값이 있음을 알 수 있다.
즉, 정상성을 갖추고 있다고 할 수 없다.

```
Box.test(res, type="Box-Pierce")
```

```
##
##  Box-Pierce test
##
## data:  res
## X-squared = 52.188, df = 1, p-value = 5.044e-13
```

H0 : Residual is independent
-> p-value가 5.044e-13으로 유의수준 0.05보다 작은 것을 알 수 있다.
-> 즉, 귀무가설을 기각할 수 있으므로 잔차가 자기상관이 있다고 할 수 있다.

4. Fit the simple exponential smoothing with alpha=0.1 and with the optimized alpha. If you think we need a trend, or seasonal, or both try them. Please address all the moeling and show how you find the best exponential smoothing model for spot data.

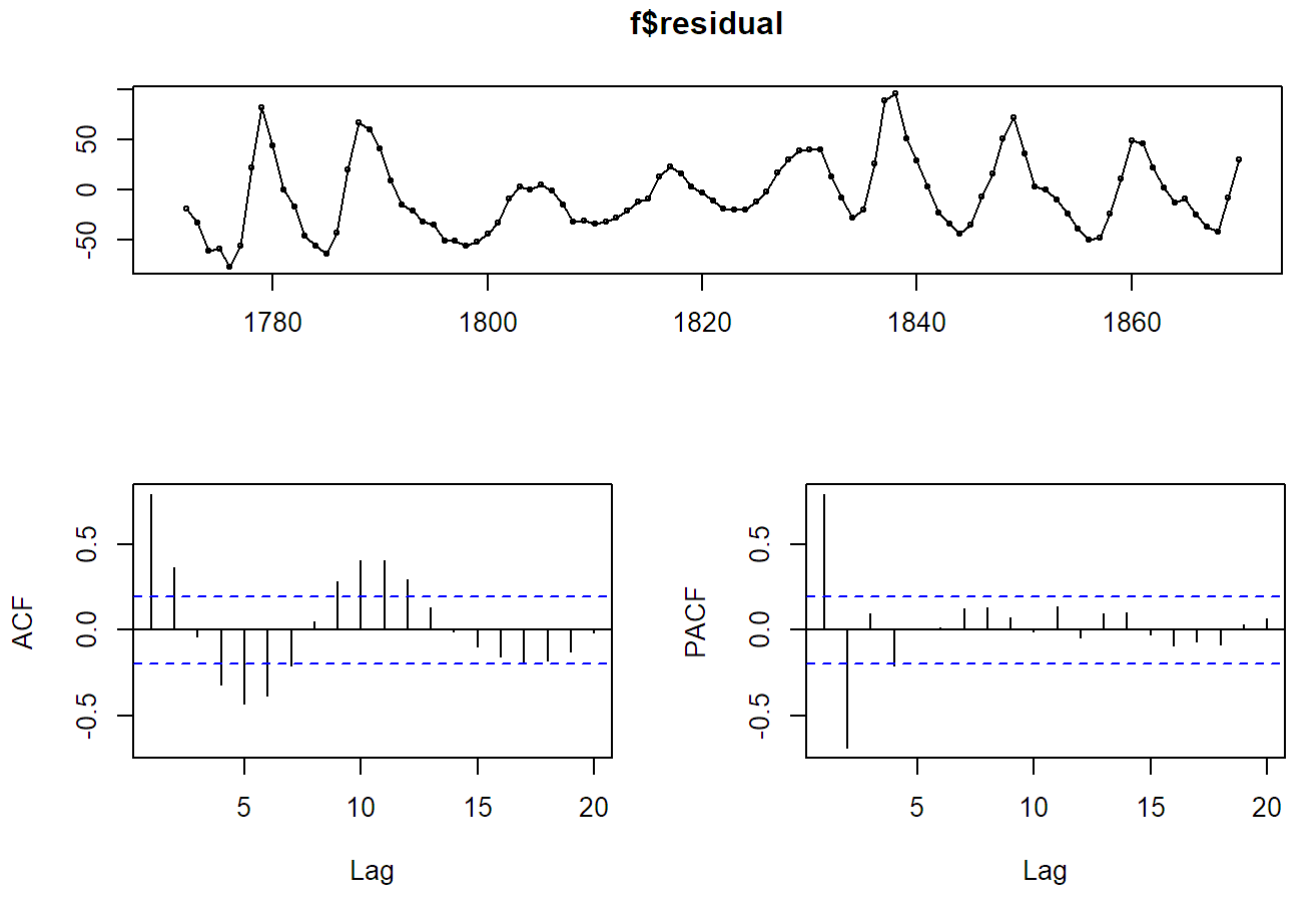
	alpha = 0.1	optimized alpha
X	ho model	ha model
trend	ho_trend model	ha_trend model

- 1년 주기의 data이므로 seasonal은 필요하지 않다.
- 총 4가지의 모델을 residual plot, stationary, test the independence를 비교하여 best model을 뽑도록 하였다.
- find 함수 : residual plot, stationary, test the independence 출력

```
find = function(x){
  f = forecast(x)
  tsdisplay(f$residual)
  Box.test(f$residual, type="Box-Pierce")
}
```

1) ho model

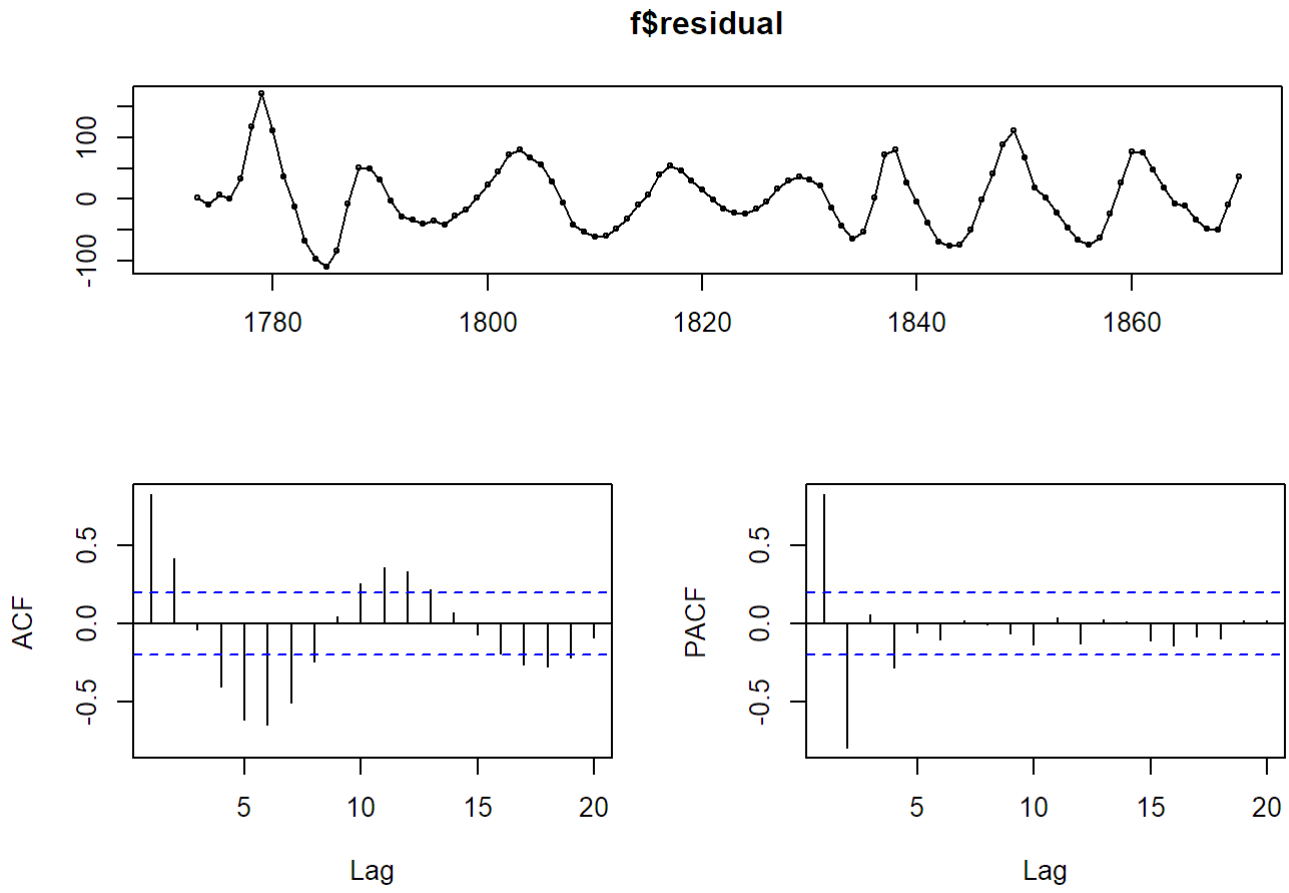
```
ho = HoltWinters(spot.ts, alpha=0.1, beta=F, gamma=F) ; find(ho)
```



```
##
##  Box-Pierce test
##
## data:  f$residual
## X-squared = 61.663, df = 1, p-value = 4.108e-15
```

2) ho_trend model

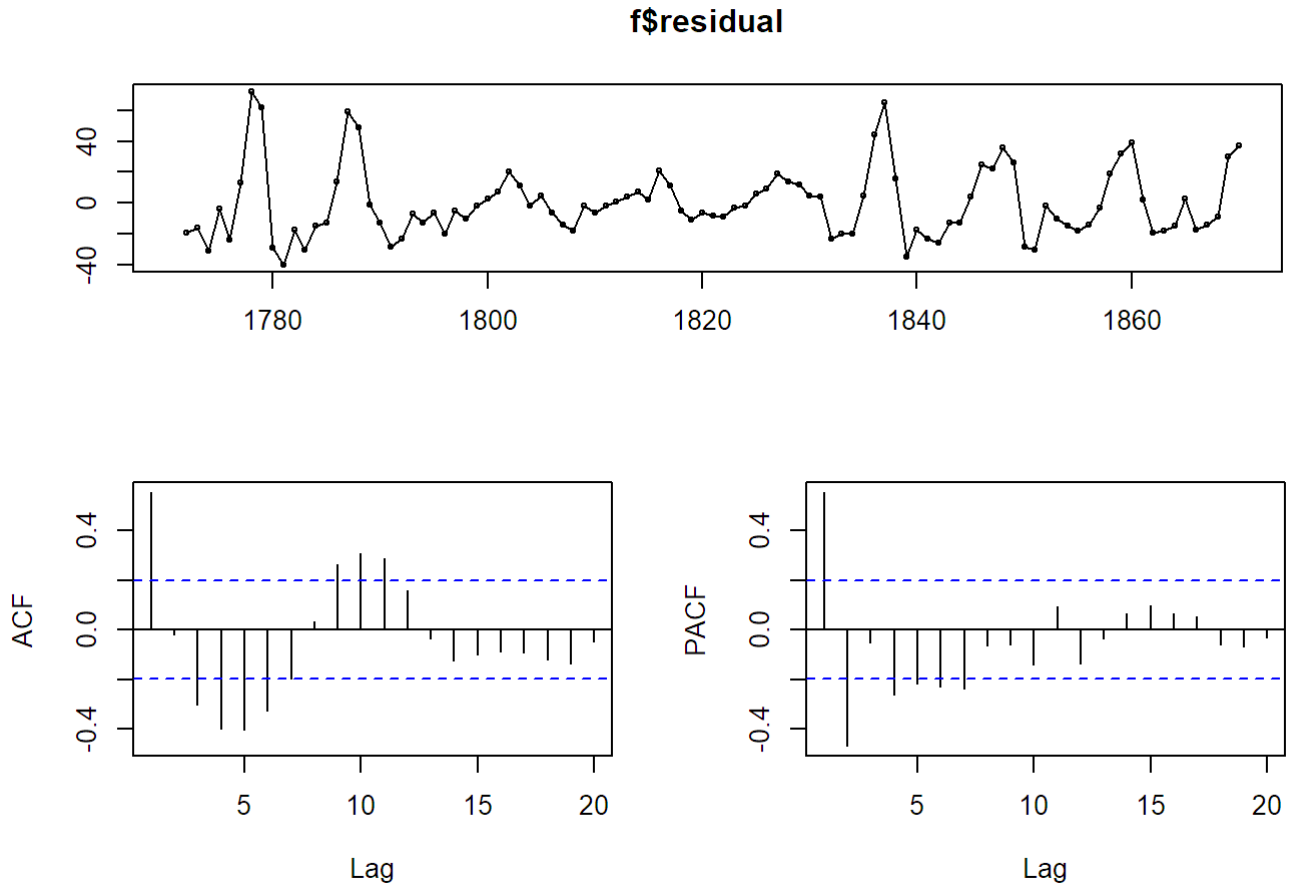
```
ho_trend = HoltWinters(spot.ts, alpha=0.1, beta=T, gamma=F) ; find(ho_trend)
```




```
##
##  Box-Pierce test
##
## data:  f$residual
## X-squared = 65.829, df = 1, p-value = 4.441e-16
```

3) ha model

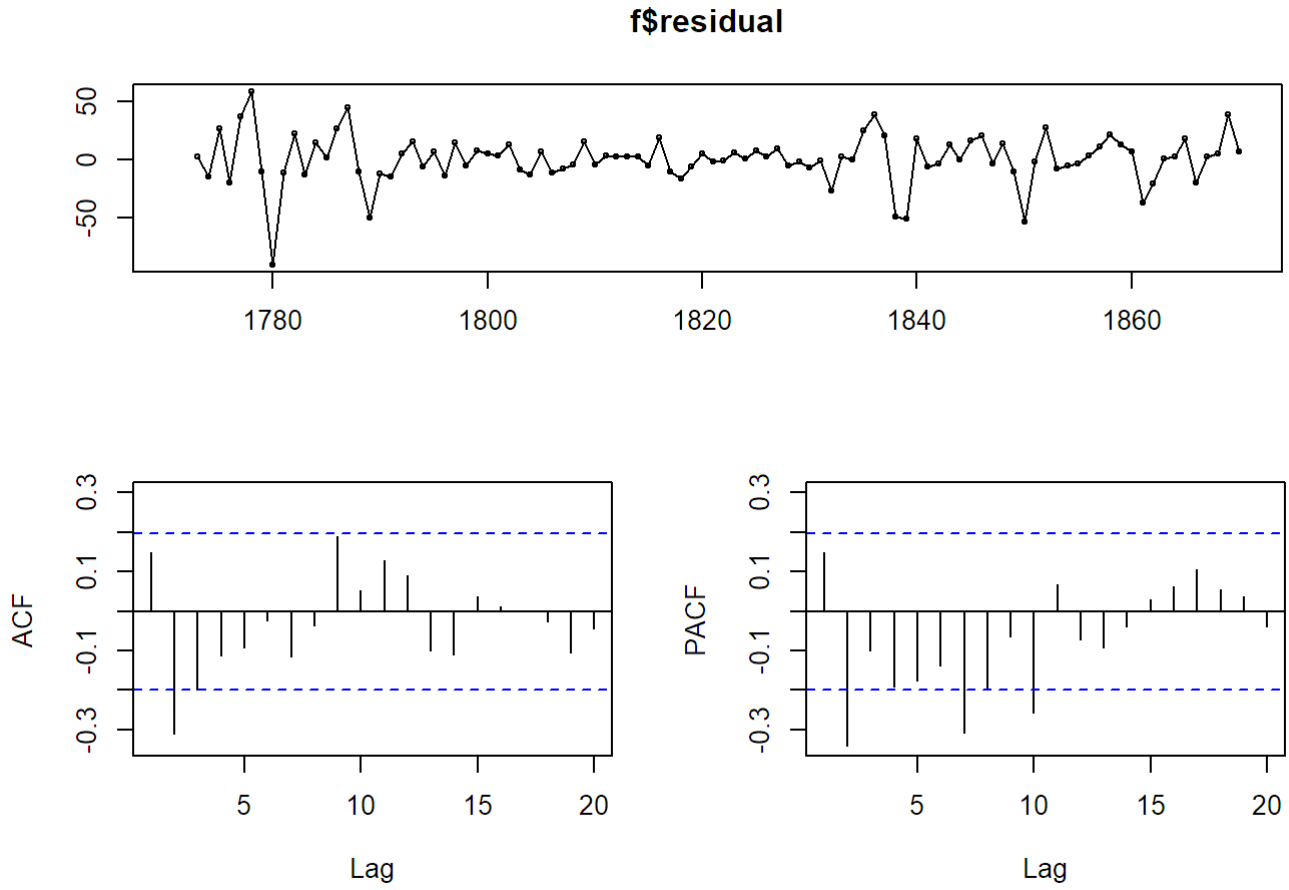
```
ha = HoltWinters(spot.ts, beta=F, gamma=F) ; find(ha)
```




```
##
##  Box-Pierce test
##
## data:  f$residual
## X-squared = 30.077, df = 1, p-value = 4.153e-08
```

4) ha_trend model

```
ha_trend = HoltWinters(spot.ts, beta=T, gamma=F) ; find(ha_trend)
```




```
##
##  Box-Pierce test
##
## data:  f$residual
## X-squared = 2.2051, df = 1, p-value = 0.1376
```

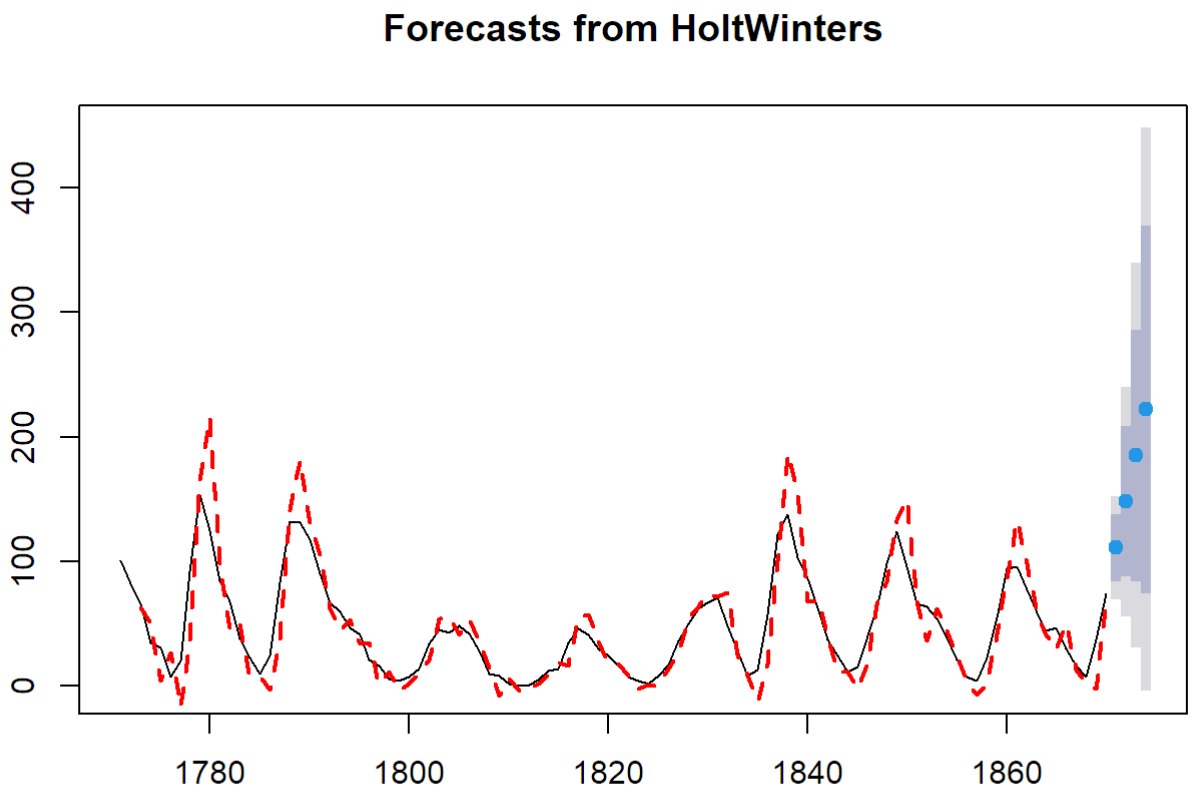
- 총 4가지 모델 중, optimized alpha를 적용시키고 trend를 고려하였을 때, independence test에서 p-value가 0.1376으로 0.05보다 커 잔차가 자기상관이 없다고 할 수 있으며 residual plot도 어느 정도 random해보인다. ACF와 PACF도 대체로 유의수준 내에 있다.
- 즉, ha_trend model이 best exponential smoothing model이라고 할 수 있다.

5.From your best model, find the forecast of next 4 points.

```
ha_trend = HoltWinters(spot.ts, beta=T, gamma=F)
fa_trend = forecast(ha_trend, h=4)
fa_trend
```

```
##      Point Forecast    Lo 80    Hi 80     Lo 95    Hi 95
## 1871       111.0013 84.01153 137.9911 69.724006 152.2786
## 1872       148.0030 87.65455 208.3514 55.708014 240.2980
## 1873       185.0047 84.02310 285.9863 30.566677 339.4427
## 1874       222.0064 74.18476 369.8280 -4.067271 448.0800
```

```
plot(fa_trend)
lines(fa_trend$fitted, col="red", lty=2, lwd=2)
```



다음 sun spot은 111, 148, 185, 222로 예측된다.