

빅데이터론 HW#4 16012993 이준희

```
library("TTR")
```

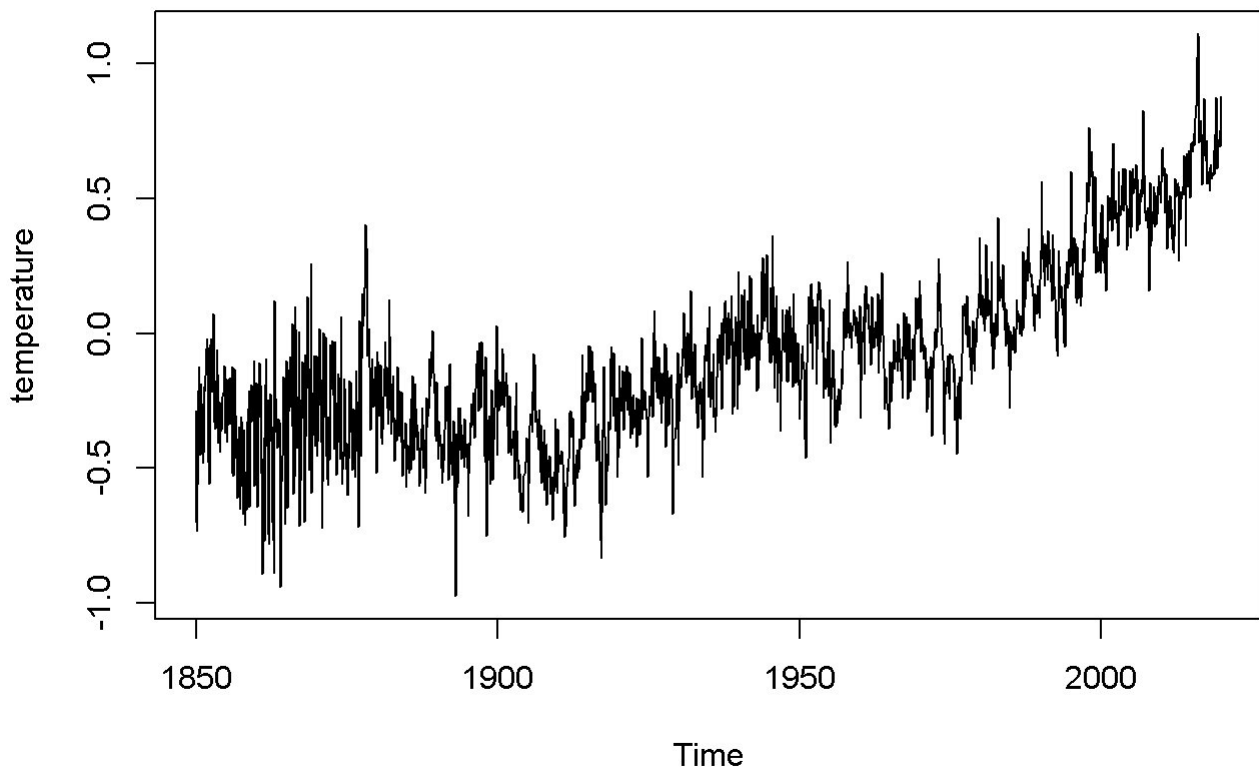
```
## Warning: package 'TTR' was built under R version 3.6.3
```

```
temperature.df <- read.csv("Global Temperature Anomalies.csv")  
temperature.ts<- ts(temperature.df$anomaly_value)
```

```
temperature<-ts(temperature.ts,start=c(1850,1),end=c(2019,12),freq=12)
```

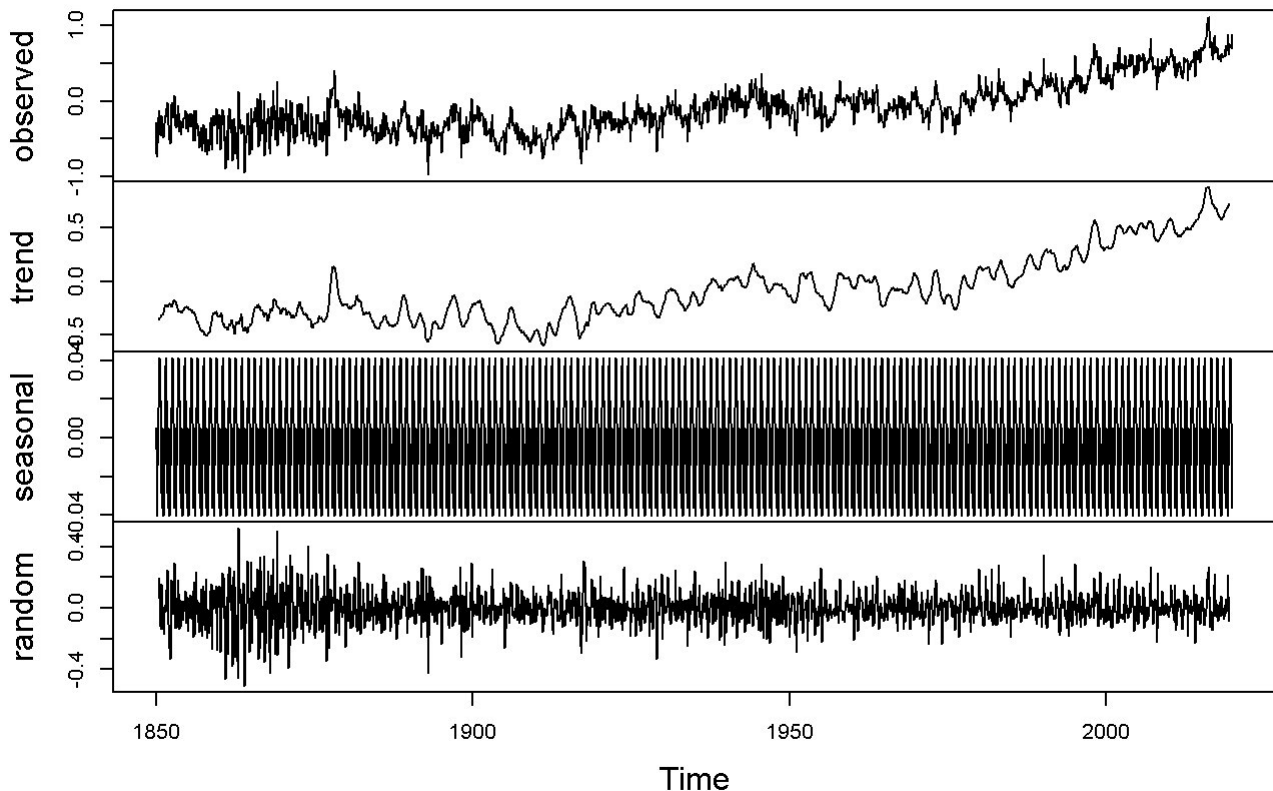
```
plot.ts(temperature,main="Global Temperature Change")
```

Global Temperature Change



```
temperaturecomponents <- decompose(temperature)  
#temperaturecomponents$seasonal  
#temperaturecomponents$trend  
#temperaturecomponents$random  
#temperaturecomponents$type  
#temperaturecomponents$figure  
plot(temperaturecomponents)
```

Decomposition of additive time series



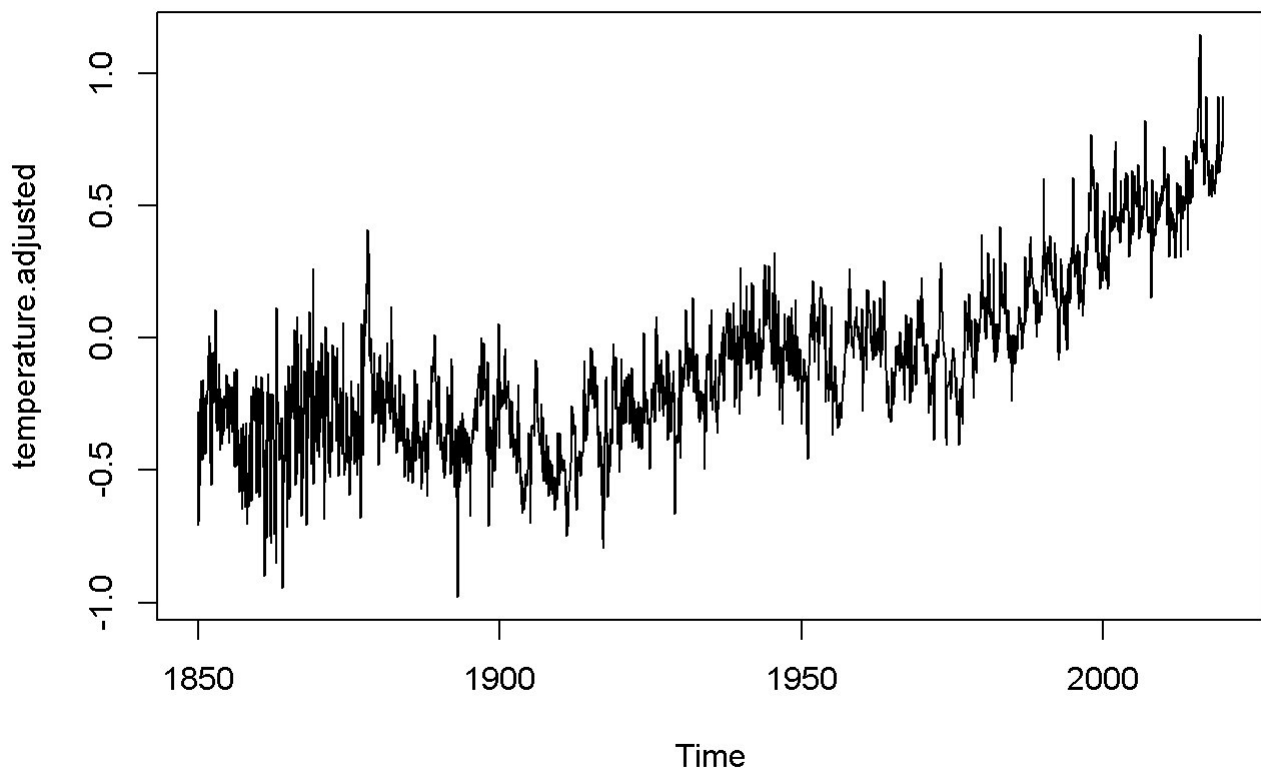
```
temperature.adjusted <- temperature - temperaturecomponents$seasonal
plot(temperature.adjusted,main ="Adjusted global temperature change")
```

1850년부터 2020년까지 150년간의 지구 기온 변화 데이터를 해석했습니다. 50년 단위로 세 구간으로 분할을 해서 간략히 설명하면 다음과 같습니다. 우선 1850~1900년에는 1차 산업혁명이 일어나면서 증기 및 석탄산업이 활발지는 동시에 오랜 기간의 중세 온난기가 막을 내리면서 소빙기에 잠시 도입했습니다. 그래서 기온의 변화가 상대적으로 오락가락했으며 이 영향으로 그 당시 농업에 막대한 피해가 생겼습니다. 이러한 자연적 현상과 산업의 발전으로 점차 기술중심의 사회가 자리 잡으면서 더욱 기온의 변화의 변동폭이 컸습니다.

1900~1950년대에는 제1,2차 세계대전 및 6.25전쟁 등 국제적 분쟁이 많이 발발했고 대량생산 원리인 2차산업혁명이 일어나서 1800년대보다 더 많은 온실가스를 방출하게 되었습니다. 특히 1940년대부터 CO₂의 방출량이 폭발적으로 증가하면서 지구 온난화 현상을 가속화 시켰습니다. 나아가 2000년대의 3차산업혁명을 통해서 석유,화학 등의 자원이 더욱 많은 분야에 활용되면서 우리의 삶이 풍요로워지고 전세계적으로 경제 및 무역이 활성화되었지만 간혹 지나친 자원의 남용으로 생태계 및 오존층 파괴 등의 환경적 이슈들이 수면 위로 올라왔습니다.

데이터의 분할된 요소를 보면 다음과 같습니다. 우선 계절성 요인을 보면 항상 여름과 겨울의 변화율이 크게 나타나는 정직한 양극화의 흐름을 보여주고 있습니다. 두 계절의 차이가 점차 커지는 것을 감안했을 때 계절성 요인은 굉장히 정직한 흐름을 보입니다. 트렌드를 보면 기온 자체의 증감 폭의 이동이 점차 상향선으로 이동합니다. 즉, 변동폭의 위치가 대략 일정하게 상승하고 있습니다. 이것은 점차 높아지는 기본 온도 기준으로 변동폭을 이를 것을 알려줍니다.

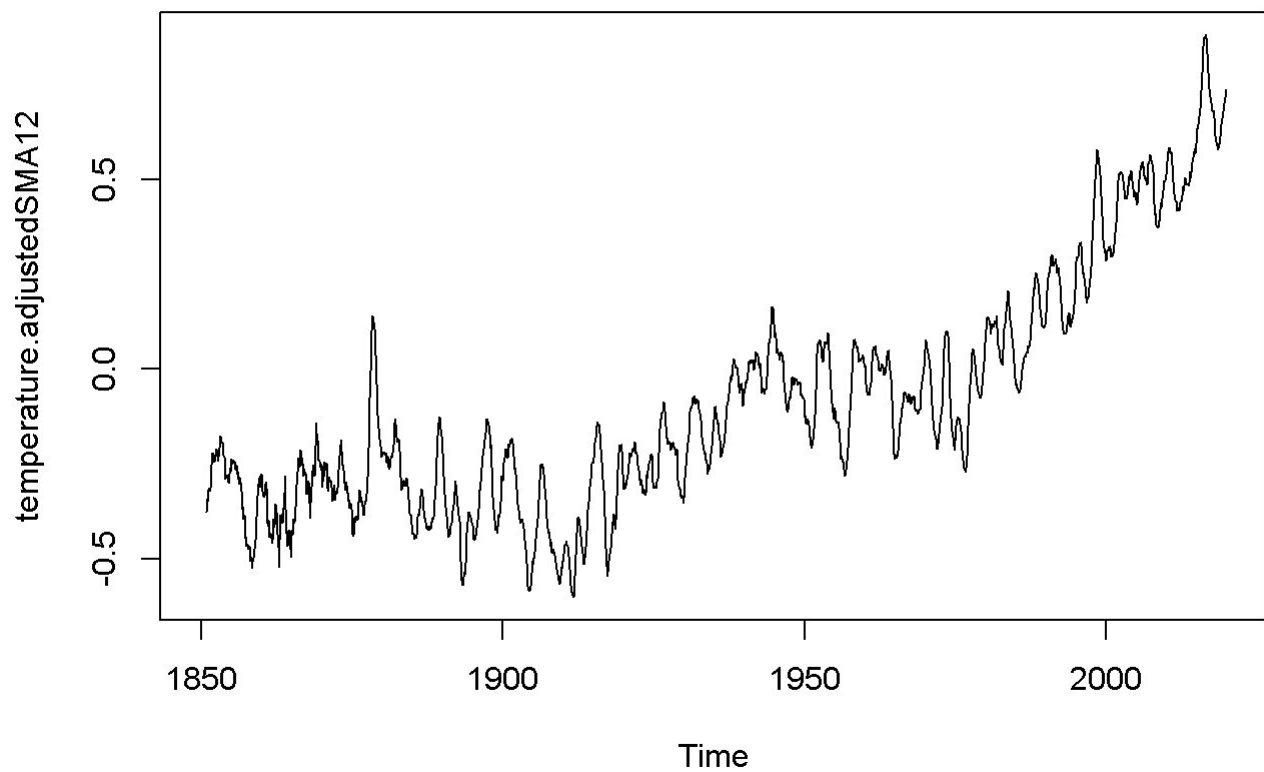
Adjusted global temperature change



```
temperature.adjustedSMA12 <- SMA(temperature.adjusted,n=12)
plot.ts(temperature.adjustedSMA12,main="Clearer trend of global temperature change")
```

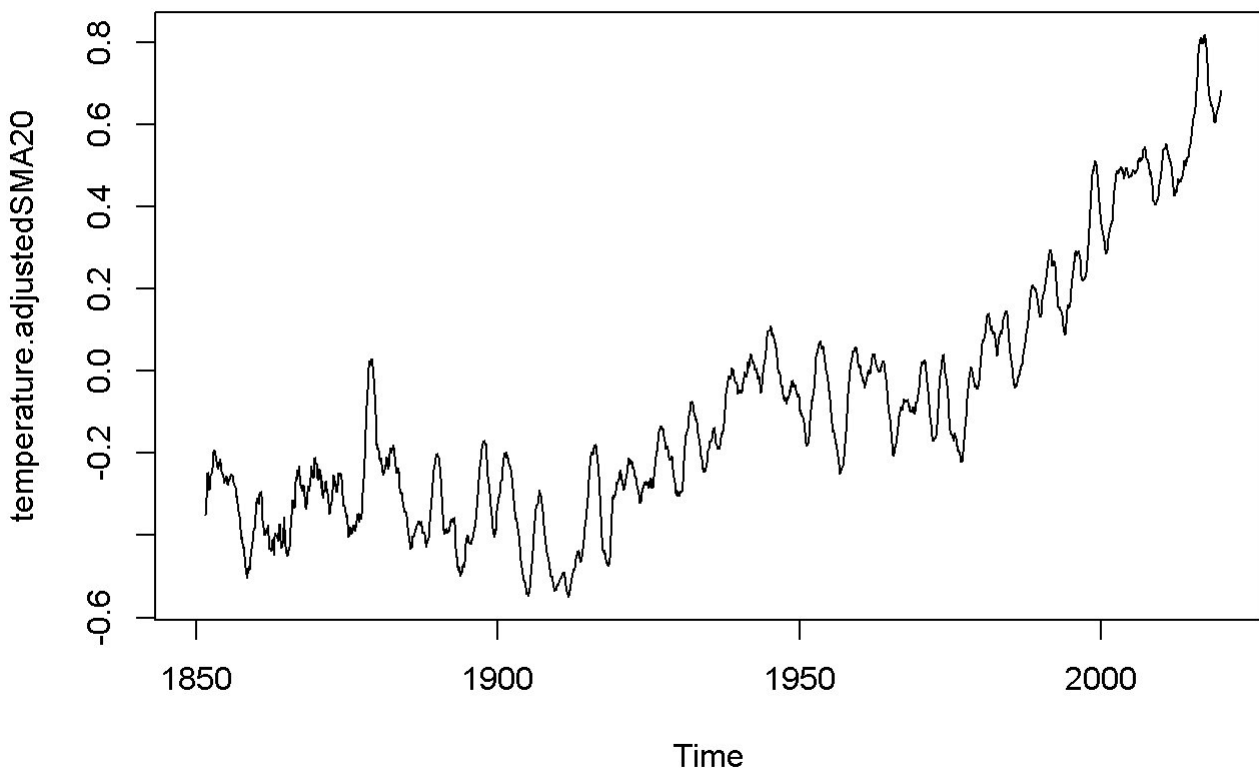
계절성요인과 랜덤요인을 제거 후 트렌드를 보다 잘 부각하기 위해서 그래프 조정을 했지만 150년의 흐름을 담고 있고 기온의 변동폭을 고려했을 때 그렇게 성과가 좋지 못했습니다. 그래서 이동평균법을 적용하여 흐름을 최대한 부드럽게 표현하기 위해서 각각 12년, 20년을 기준으로 SMA를 적용했습니다. 20년 기준의 SMA를 적용한 그래프를 보면 확실히 추세선이 전보다 눈에 더 잘 들어오는 것을 볼 수 있습니다.

Clearer trend of global temperature change



```
temperature.adjustedSMA20 <- SMA(temperature.adjusted,n=20)  
plot.ts(temperature.adjustedSMA20,main="Clearer trend of global temperature change")
```

Clearer trend of global temperature change



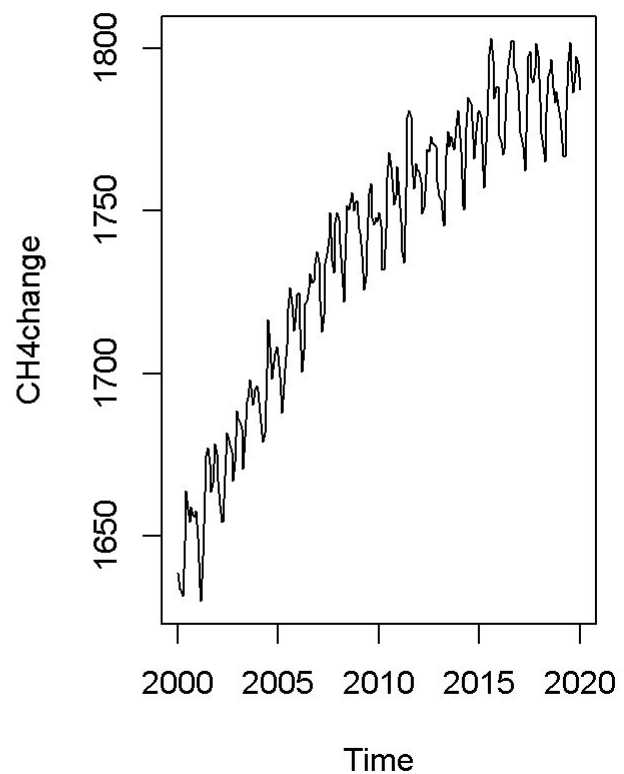
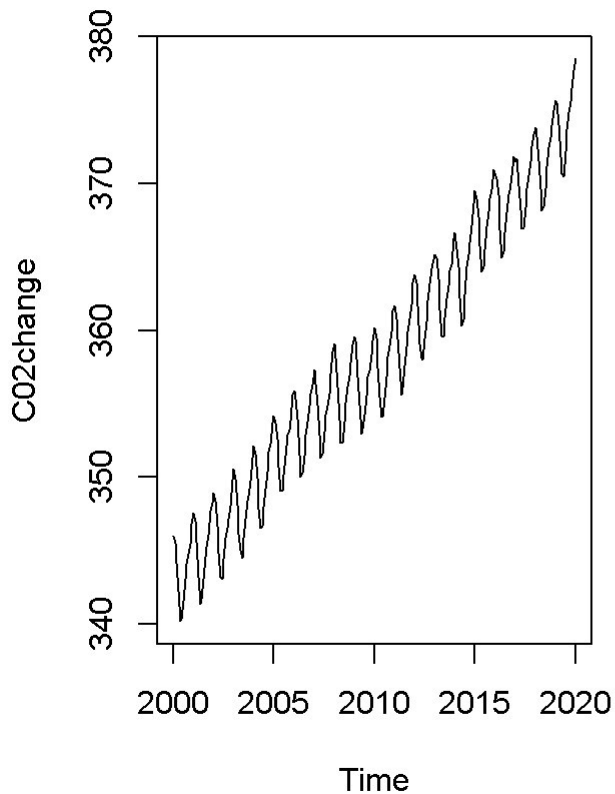
```
par(mfrow=c(1,2))

climatechange <- read.csv("climate_change.csv")

climatechangeC02 <- ts(climatechange$C02)
C02change <- ts(climatechangeC02,start=c(2000,1),end=c(2020,1),freq=12)
plot.ts(C02change)

climatechangeCH4 <- ts(climatechange$CH4)
CH4change <- ts(climatechangeCH4,start=c(2000,1),end=c(2020,1),freq=12)
plot.ts(CH4change)
```

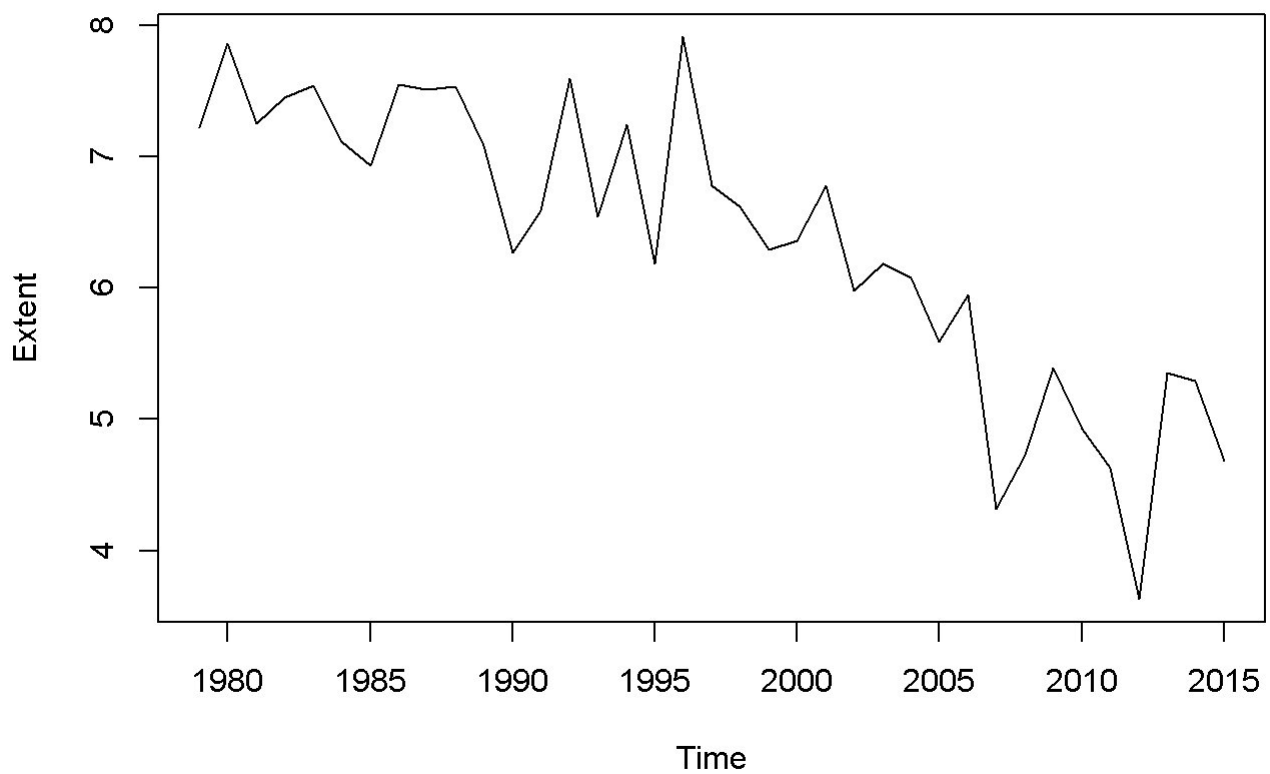
다음은 온난화의 주범이라고 불리는 CO₂(이산화탄소)와 CH₄(메탄)의 방출량을 표현했습니다. 꾸준히 증가하는 두 요소는 인공온실가스의 방출로 인하여 지난 50년간 지구 평균 온도 상승에 많은 영향을 끼쳤습니다. 대기의 속성을 고려했을 때, 각종 산업활동으로 오존층이 지속적으로 파괴되어 왔으며 이것이 CO₂와 CH₄의 방출 증가를 초래한 것으로 알려져 있습니다. 특히 CH₄는 열을 잘 방출하는 요소라는 점에서 기온에 영향을 미쳤기 때문에 최근 매년 특정 시기에 역대 가장 더운 여름 날씨 라는 말이 나오고 있고 계속되는 오존층의 파괴와 불균형으로 최근 3~5월 계절에 맞지 않은 이상기후가 종종 발생했습니다.



```
par(mfrow=c(1,1))

seaice <- read.csv("SeaIce.csv")
Extent <- ts(seaice$Extent,start=c(1979),end=c(2015))
plot.ts(Extent)
```

기온 상승과 관련해서 전세계 해수의 얼음 및 빙하의 농도 데이터를 불러왔습니다. 먼저 얼음의 증감 흐름을 보면 2000년대부터 기하급수적으로 감소하고 있습니다. HoltWinters로 exponential smoothing을 한 결과, $a=0.4$ 인 점을 통해서 데이터의 흐름이 SSE를 최소화하기 위해서 현재에 더 집중하는 시계열 데이터라는 점을 알 수 있습니다. 흐름의 오차 또한 그렇게 큰 수치는 아니며 Holt Winters filtering의 시계열 그래프를 보면 실제 흐름과 예측된 흐름 간에 큰 차이가 없고 이동평균기법을 적용시킨 흐름처럼 추세선이 부드럽게 나타납니다.



```
sealiceExtentforecast <- HoltWinters(Extent,beta=FALSE,gamma=FALSE)
sealiceExtentforecast
```

```
## Holt-Winters exponential smoothing without trend and without seasonal component.
##
## Call:
## HoltWinters(x = Extent, beta = FALSE, gamma = FALSE)
##
## Smoothing parameters:
##  alpha: 0.4009073
##  beta  : FALSE
##  gamma: FALSE
##
## Coefficients:
##      [,1]
## a 4.861559
```

```
sealiceExtentforecast$fitted
```

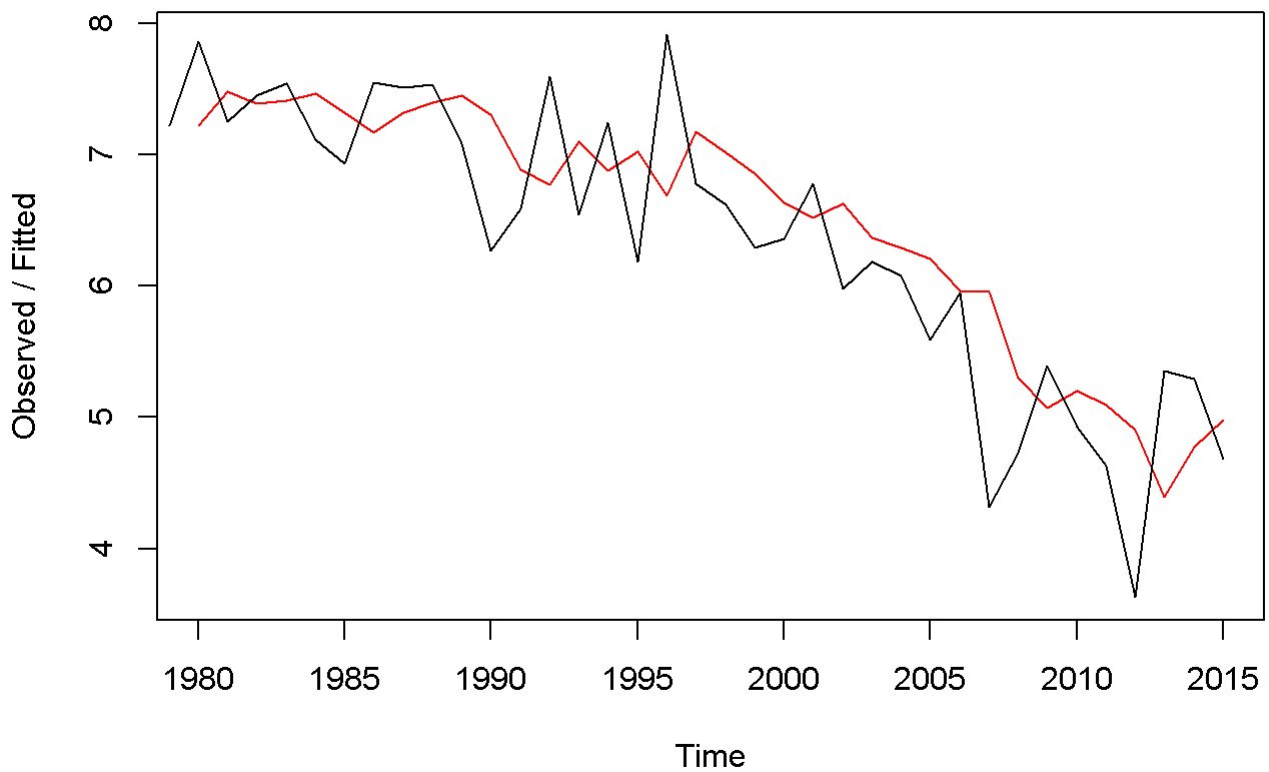
```
## Time Series:
## Start = 1980
## End = 2015
## Frequency = 1
##      xhat      level
## 1980 7.220000 7.220000
## 1981 7.476581 7.476581
## 1982 7.385743 7.385743
## 1983 7.411504 7.411504
## 1984 7.463019 7.463019
## 1985 7.321491 7.321491
## 1986 7.164539 7.164539
## 1987 7.319073 7.319073
## 1988 7.395617 7.395617
## 1989 7.449492 7.449492
## 1990 7.301360 7.301360
## 1991 6.887880 6.887880
## 1992 6.768458 6.768458
## 1993 7.097820 7.097820
## 1994 6.874186 6.874186
## 1995 7.020843 7.020843
## 1996 6.683743 6.683743
## 1997 7.175358 7.175358
## 1998 7.016856 7.016856
## 1999 6.857754 6.857754
## 2000 6.630137 6.630137
## 2001 6.521837 6.521837
## 2002 6.625337 6.625337
## 2003 6.366616 6.366616
## 2004 6.291801 6.291801
## 2005 6.206888 6.206888
## 2006 5.959573 5.959573
## 2007 5.955735 5.955735
## 2008 5.299957 5.299957
## 2009 5.071457 5.071457
## 2010 5.199163 5.199163
## 2011 5.091254 5.091254
## 2012 4.906334 4.906334
## 2013 4.394642 4.394642
## 2014 4.777652 4.777652
## 2015 4.983056 4.983056
```

```
seaiceExtentforecast$SSE
```

```
## [1] 13.51795
```

```
plot(seaiceExtentforecast)
```


Holt-Winters filtering



```
logExtent <- log(Extent)
plot.ts(logExtent)
```

비록 차이가 적지만 로그를 씌웠을 때의 흐름을 보면 2000년 기준 전후의 얼음 및 빙하의 증감 차이가 더 극단적으로 나타납니다. 마찬가지로 알파 값은 기준과 거의 차이가 없으며 exponential smoothing을 통한 시계열 그래프를 통해서 부드러운 추세선을 확인할 수 있습니다.

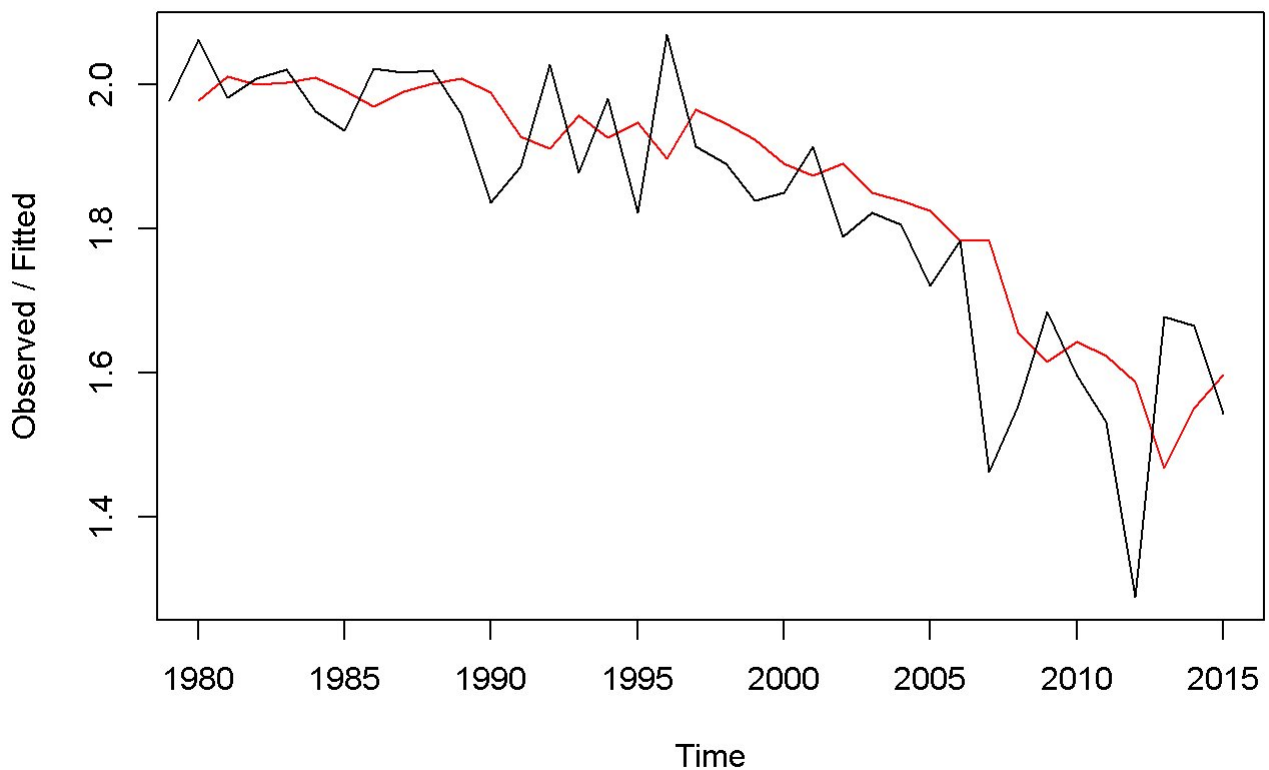


```
logseaExtentforecast <- HoltWinters(logExtent,beta=FALSE,gamma=FALSE)
logseaExtentforecast
```

```
## Holt-Winters exponential smoothing without trend and without seasonal component.
##
## Call:
## HoltWinters(x = logExtent, beta = FALSE, gamma = FALSE)
##
## Smoothing parameters:
##  alpha: 0.397776
##  beta : FALSE
##  gamma: FALSE
##
## Coefficients:
##      [,1]
## a 1.575709
```

```
plot(logseaExtentforecast)
```

Holt-Winters filtering



```
Area <- ts(seaice$Area,start=c(1979),end=c(2015))  
plot.ts(Area)
```

실제 면적 변화의 흐름을 보면 앞에서의 증감 흐름과 비슷하게 나타납니다. 당연히 얼음 및 빙하를 포함한 해수의 증감은 그것을 포함한 해수의 면적에 영향을 끼칩니다. 알파값과 SSE 역시 크게 다르지 않은 값을 가졌고 예측값 또한 실제값의 흐름에서 이동평균기법을 적용한 것처럼 부드럽게 나타납니다. 로그를 적용했을 때도 증감의 데이터처럼 2000년기준으로 전후의 차이가 더 극단적으로 나타나는 점에서 유사합니다.



```
seaiiceAreaforecast <- HoltWinters(Area,beta=FALSE,gamma=FALSE)
seaiiceAreaforecast
```

```
## Holt-Winters exponential smoothing without trend and without seasonal component.
##
## Call:
## HoltWinters(x = Area, beta = FALSE, gamma = FALSE)
##
## Smoothing parameters:
##  alpha: 0.3876864
##  beta  : FALSE
##  gamma: FALSE
##
## Coefficients:
##      [,1]
## a 3.411177
```

```
seaiiceAreaforecast$fitted
```

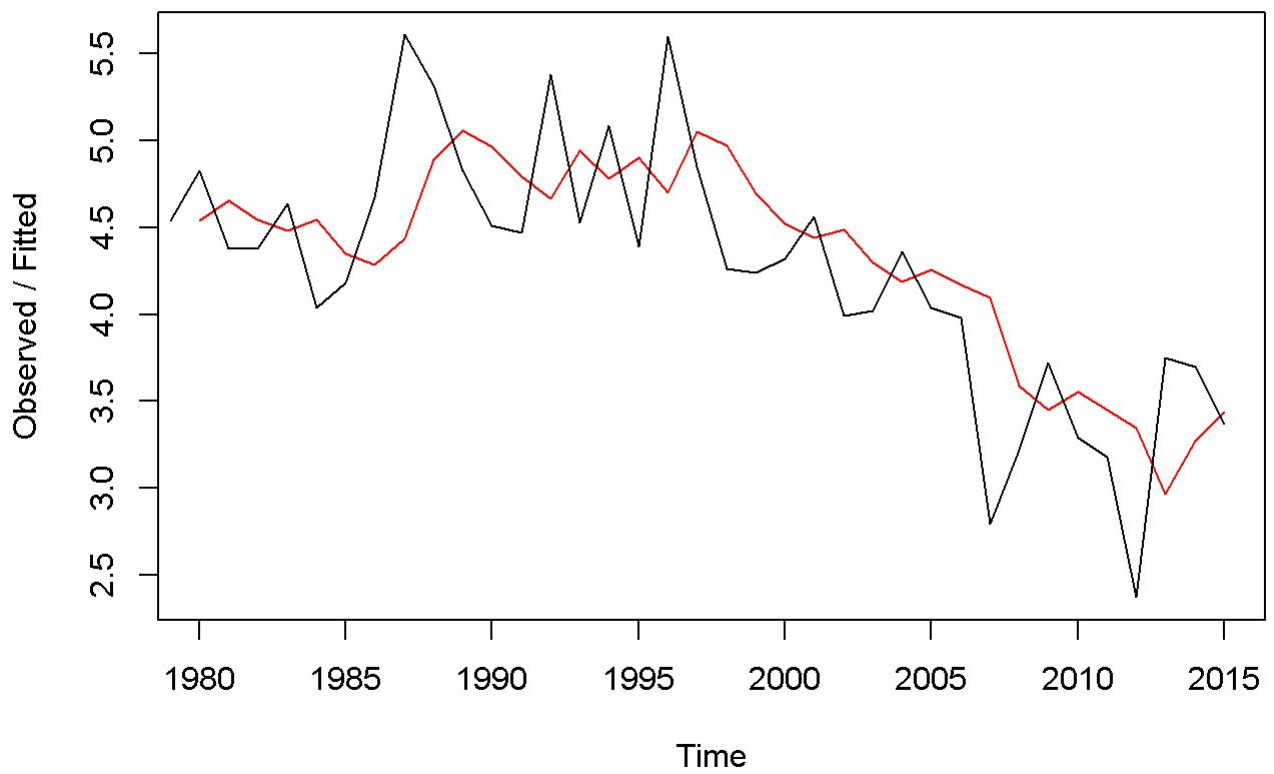
```
## Time Series:
## Start = 1980
## End = 2015
## Frequency = 1
##      xhat      level
## 1980 4.540000 4.540000
## 1981 4.652429 4.652429
## 1982 4.546812 4.546812
## 1983 4.482141 4.482141
## 1984 4.543341 4.543341
## 1985 4.348203 4.348203
## 1986 4.282993 4.282993
## 1987 4.433030 4.433030
## 1988 4.889325 4.889325
## 1989 5.056292 5.056292
## 1990 4.968562 4.968562
## 1991 4.790784 4.790784
## 1992 4.666420 4.666420
## 1993 4.943065 4.943065
## 1994 4.782926 4.782926
## 1995 4.901974 4.901974
## 1996 4.703489 4.703489
## 1997 5.051054 5.051054
## 1998 4.973108 4.973108
## 1999 4.696646 4.696646
## 2000 4.519610 4.519610
## 2001 4.442224 4.442224
## 2002 4.487884 4.487884
## 2003 4.294861 4.294861
## 2004 4.188301 4.188301
## 2005 4.254867 4.254867
## 2006 4.171566 4.171566
## 2007 4.097298 4.097298
## 2008 3.590477 3.590477
## 2009 3.446848 3.446848
## 2010 3.552745 3.552745
## 2011 3.450882 3.450882
## 2012 3.345865 3.345865
## 2013 2.967535 2.967535
## 2014 3.270886 3.270886
## 2015 3.437248 3.437248
```

```
seaiceAreaforecast$SSE
```

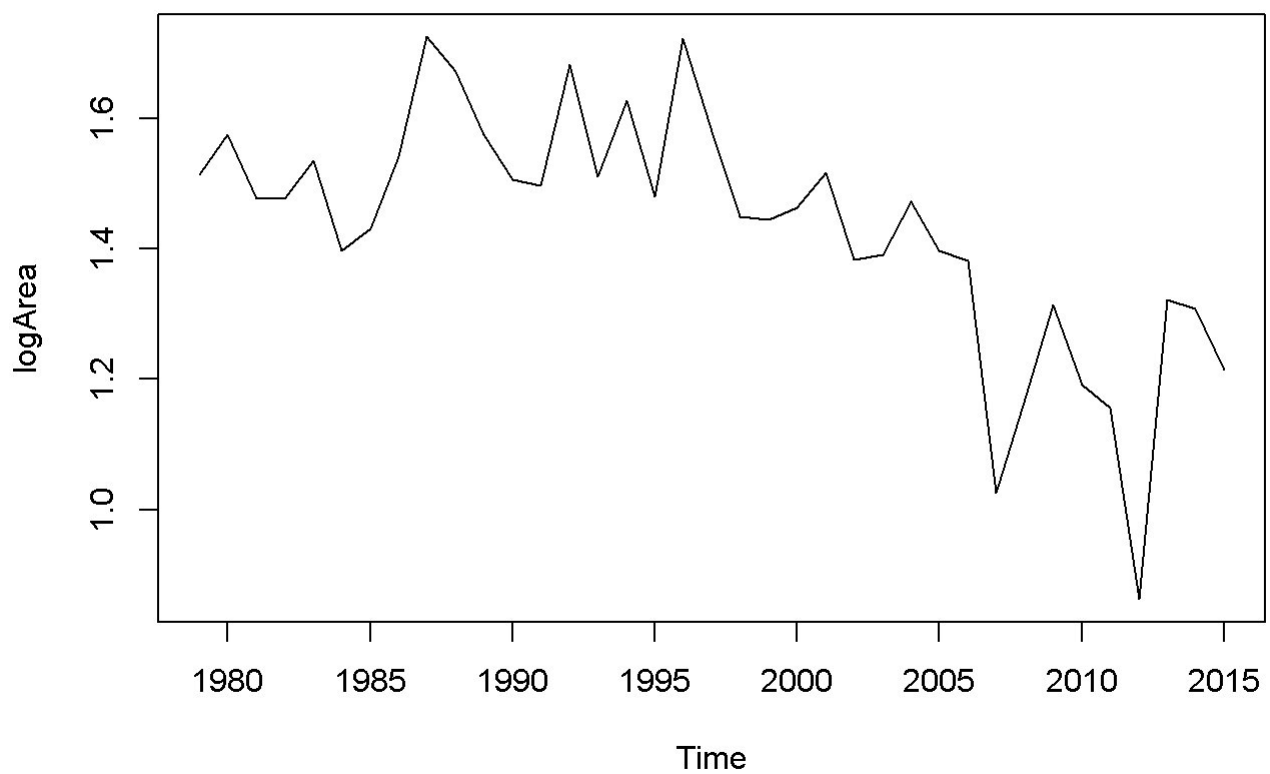
```
## [1] 9.481074
```

```
plot(seaiceAreaforecast)
```

Holt-Winters filtering



```
logArea <- log(Area)
plot.ts(logArea)
```

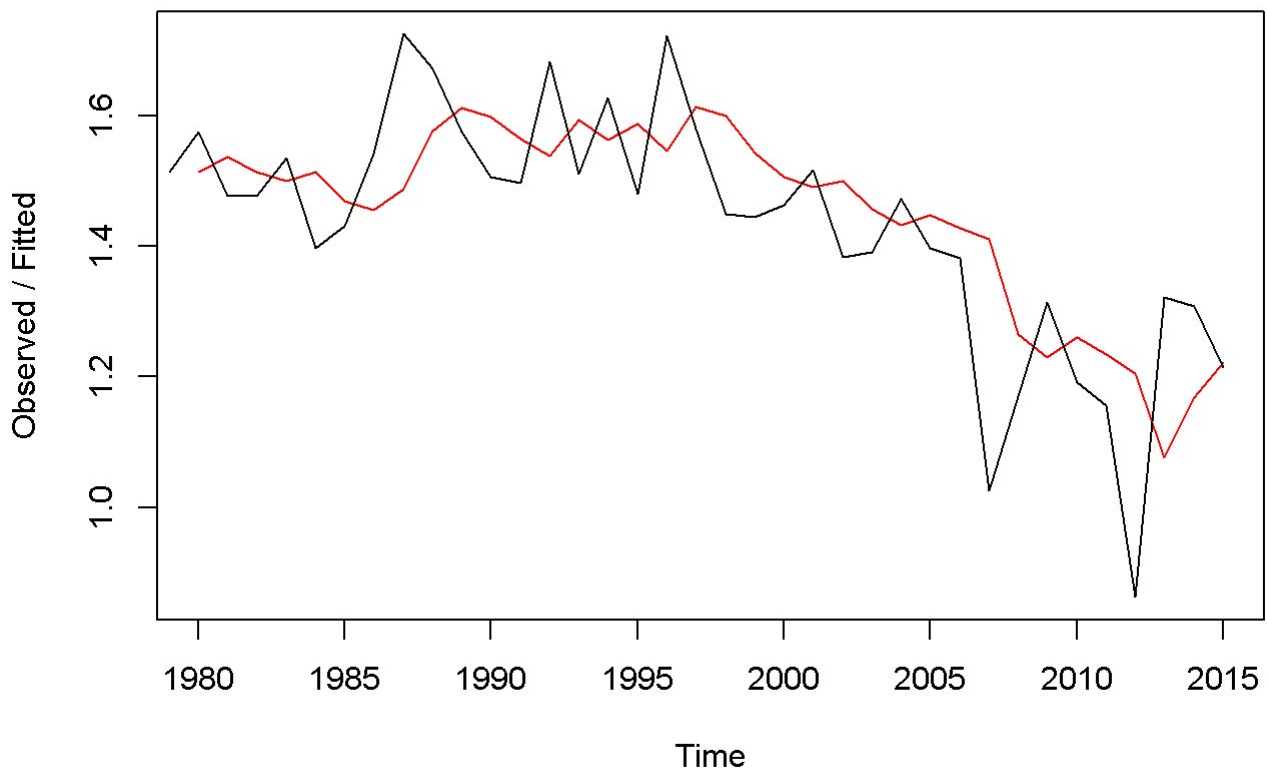


```
logseaAreaforecast <- HoltWinters(logArea,beta=FALSE,gamma=FALSE)
logseaAreaforecast
```

```
## Holt-Winters exponential smoothing without trend and without seasonal component.
##
## Call:
## HoltWinters(x = logArea, beta = FALSE, gamma = FALSE)
##
## Smoothing parameters:
##  alpha: 0.3766288
##  beta : FALSE
##  gamma: FALSE
##
## Coefficients:
##      [,1]
## a 1.218928
```

```
plot(logseaAreaforecast)
```

Holt-Winters filtering



기온변화와 CO₂, CH₄ 방출량의 시계열 데이터를 통해서 해수면에서의 얼음 및 빙하의 증감과 면적의 흐름을 보면, 점차 해수면이 상승할 것으로 예상할 수 있습니다. 실제로 수십년 이내에 해양과 접한 세계 도시 곳곳이 결국 물에 잠길 것이라는 연구 결과가 많았습니다.

Exponential smoothing으로 각 예측된 흐름과 실제 데이터의 흐름의 차이가 매우 유사하다는 점을 통한 빙하면적의 감소를 보면, 그만큼 환경오염이 더 심해질 것이라는 예측이 이전부터 계속되었다는 것을 알 수 있습니다.

물론, 수십억 년 간 지속적으로 환경을 바꾸는 지구의 자연스러운 운동의 영향이 생각보다 큽니다. 그 이유는, 생각을 해보면 먼 옛날부터 꼭 석유 화학 및 각종 유해물질로 인하여 폐름기, 백악기에 많은 생명들이 죽은 것이 아니었기 때문입니다. 이것은 지구가 오랜 주기로 한두번씩 스스로 정화를 통해서 새로운 생태계를 조성하는 것을 알려주는 생태계 사건들입니다. 하지만 그렇다고 그저 현재 지구 온난화를 방치하고 있기에는 당장 우리를 포함한 생명체들에게 가장 좋지 않고 각종 이상기후 현상 등으로 생각지도 못한 자연재해가 발생할 수 있기 때문에 조금 더 조심하고 빠른 시일 내에 무해물질인 대체에너지의 개발이 필요할 것 같습니다.