HW4

7.1

A1C is a parameter that is used to detect if someone has turned diabetic or not. A1C test measures the blood sugar level over a period - usually 3 months. This is different than a random blood sugar test which measures the blood sugar level at that instant, for example in the morning before eating anything. If doctors suspect that someone has turned diabetic, then they will prescribe an A1C test to see the blood sugar level over a period.

If I have the daily A1C data for a patient for a few months, I will use exponential smoothing to predict the future A1C values, so that if the forecasted values are greater than a threshold a proactive treatment plan can be generated. I will choose an alpha value less than 0.5, preferably 0.2 as this would put more emphasis on the previous estimate rather than on the current value which makes sense - as any significant increase or decrease in the current value could have been contributed by chance. In short, I don't want random changes to affect my forecast hence I will put more emphasis on the prior estimates than the current one.

7.2

I used 1) HoltWinters method and 2) EST function to exponentially smooth the data. Then I used CUSUM method to find out when there is a significant drop in the temperature for each year. A significant drop in the temperature would indicate the unofficial end of summer. Next, I took the average of temperatures until the end of unofficial summer for each year. I then plotted these averages over the years. As shown below in both cases it is not possible to definitively conclude that the summer has gotten progressively worse.

In both HoltWinters, and EST runs I performed trials with different values of alpha, beta and gamma. Similarly, in my CUSUM runs I tested the data for different C and threshold values.

Also, I plotted the fitted values and the original values to find out that when I used EST the match was very close. However, when I used HoltWinters method, the fitted values didn't match the original values.

It must be noted that the above results are for exponential smoothing runs that include seasonality and trends - i.e. including alpha, beta and gamma.

To further substantiate my findings, I also repeated the above by exponentially smoothing the data using HoltWinters without seasonality and trend (using only alpha), and including only trend i.e. excluding seasonality.

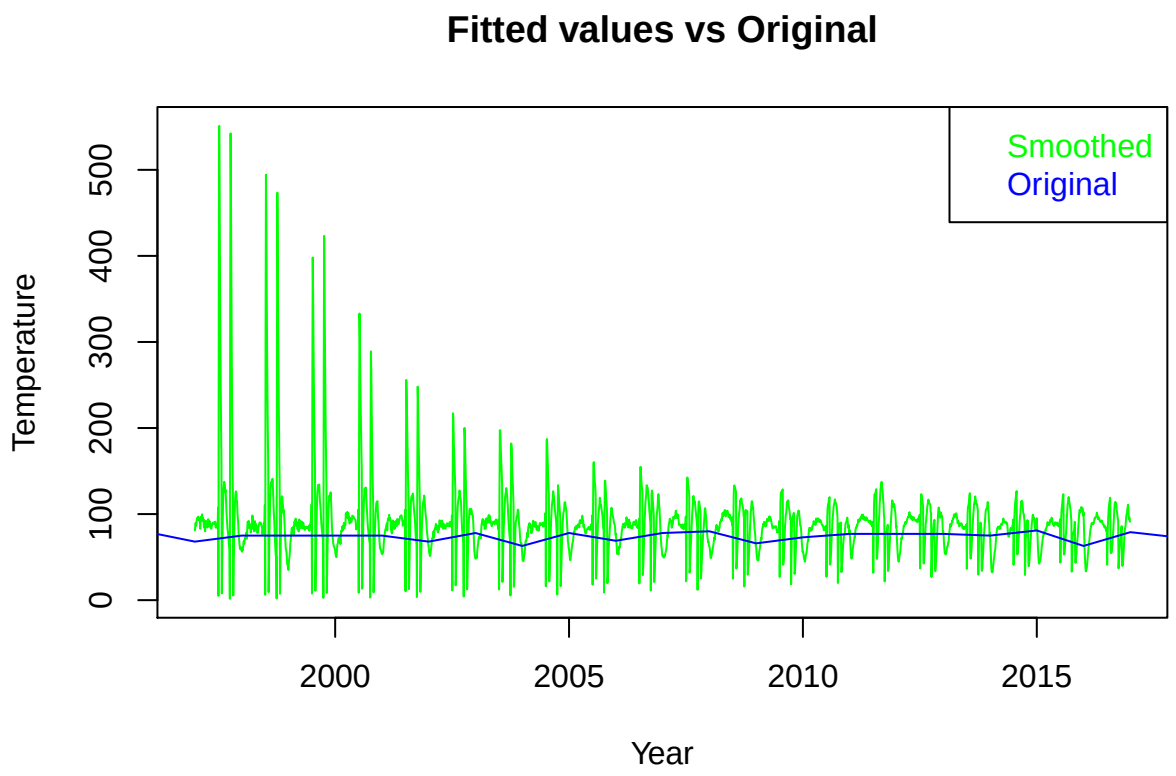
Overall, even after all these runs, I don't think there is strong evidence that suggests that summer in Atlanta has gotten progressively warmer.

```
library("plyr")
library("forecast")
```

```
## Warning: package 'forecast' was built under R version 3.4.4
```

```
temp = read.table("temps.txt",header=TRUE, check.names=FALSE)

t1 = temp[,2:20]
```

```
t1 = as.vector(t(t1))
myts = ts(temp, start=1996, frequency=123)

mod = HoltWinters(myts, alpha = 0.3, beta= 0.1, gamma=0.1, seasonal = "multiplicative")
val = fitted(mod)[,1]
mat = matrix(unlist(val),nrow = 123, ncol = 20, byrow=TRUE)

plot(fitted(mod)[,1], col = "green", xlab = "Year", ylab="Temperature", main="Fitted values vs Original
lines(t1, col="blue")
legend( "topright", c("Smoothed", "Original"),
text.col=c("green", "blue"))
```




```
c = 2.0
t = 200
ends = c()
s = rep(0, nrow(temp))
for (j in 1:ncol(mat)-1)
  {
    a = mean(as.numeric(mat[,j]))
    l = c()
    for (i in 1:nrow(mat))
      {
        d =  a - as.numeric(mat[i,j]) - c

        if (i == 1)
```

```
          s[i] = max(0, d)
        else
          s[i] = max(0, d+s[i-1])

        if (s[i] > t)
          {
            cat("The temperature shifts downward in", colnames(temp[j+2]),"is: ", as.character(temp[i,1]
            ends = c(ends, i)
            break
          }

      }

  }
```

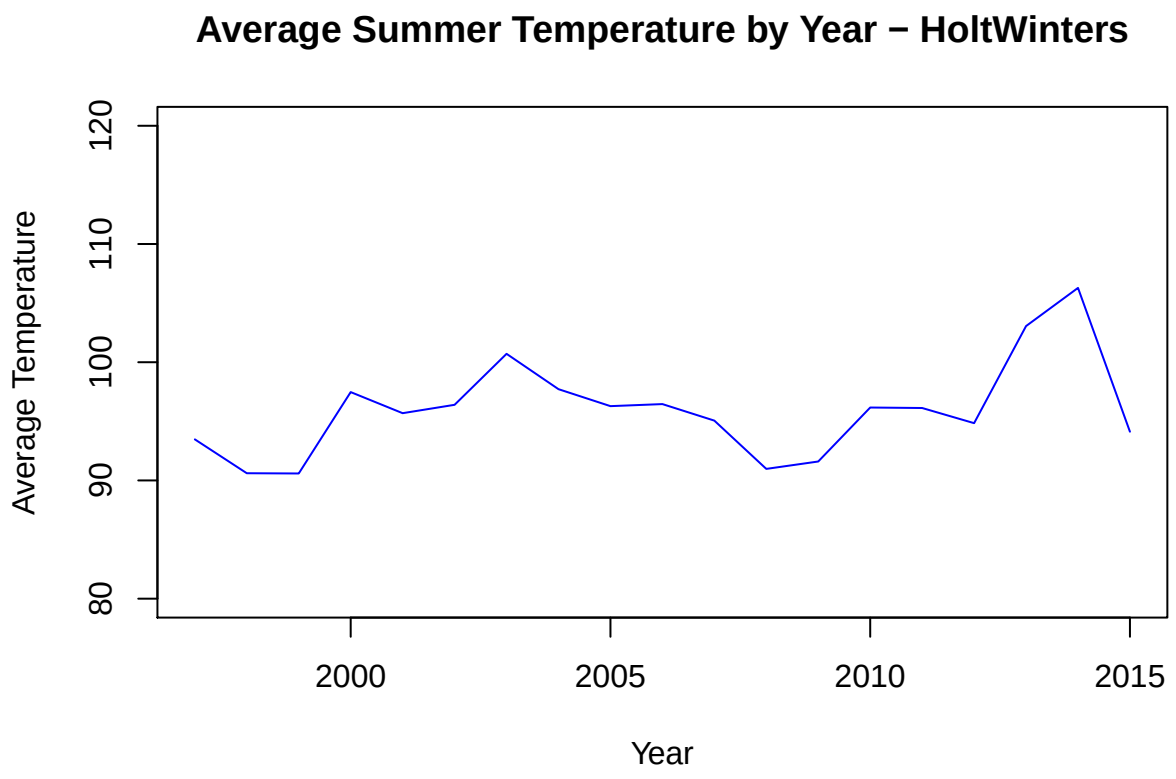
```
## The temperature shifts downward in 1997 is:  8-Oct
## The temperature shifts downward in 1998 is:  14-Oct
## The temperature shifts downward in 1999 is:  17-Oct
## The temperature shifts downward in 2000 is:  5-Sep
## The temperature shifts downward in 2001 is:  28-Sep
## The temperature shifts downward in 2002 is:  11-Sep
## The temperature shifts downward in 2003 is:  25-Aug
## The temperature shifts downward in 2004 is:  16-Sep
## The temperature shifts downward in 2005 is:  17-Sep
## The temperature shifts downward in 2006 is:  17-Sep
## The temperature shifts downward in 2007 is:  22-Sep
## The temperature shifts downward in 2008 is:  25-Sep
## The temperature shifts downward in 2009 is:  23-Sep
## The temperature shifts downward in 2010 is:  22-Sep
## The temperature shifts downward in 2011 is:  22-Sep
## The temperature shifts downward in 2012 is:  22-Sep
## The temperature shifts downward in 2013 is:  23-Aug
## The temperature shifts downward in 2014 is:  6-Aug
## The temperature shifts downward in 2015 is:  21-Sep
```

```
avg = c()
year = c()
maxi=c()
mini = c()
for (i in 1:length(ends))
{
  temp_var = mean(mat[1:ends[i]-1,i])


  avg = c(avg,temp_var)

  cat("The average summer temperature in", colnames(temp[i+2]), "is", temp_var, "\n")
  year = c(year,colnames(temp[i+2]))
}
```

```
## The average summer temperature in 1997 is 93.4741
## The average summer temperature in 1998 is 90.60914
## The average summer temperature in 1999 is 90.58854
```

```
## The average summer temperature in 2000 is 97.47308
## The average summer temperature in 2001 is 95.6899
## The average summer temperature in 2002 is 96.39801
## The average summer temperature in 2003 is 100.7104
## The average summer temperature in 2004 is 97.71887
## The average summer temperature in 2005 is 96.28504
## The average summer temperature in 2006 is 96.459
## The average summer temperature in 2007 is 95.05942
## The average summer temperature in 2008 is 90.97729
## The average summer temperature in 2009 is 91.60222
## The average summer temperature in 2010 is 96.16791
## The average summer temperature in 2011 is 96.12639
## The average summer temperature in 2012 is 94.84605
## The average summer temperature in 2013 is 103.0572
## The average summer temperature in 2014 is 106.2859
## The average summer temperature in 2015 is 94.1209
```

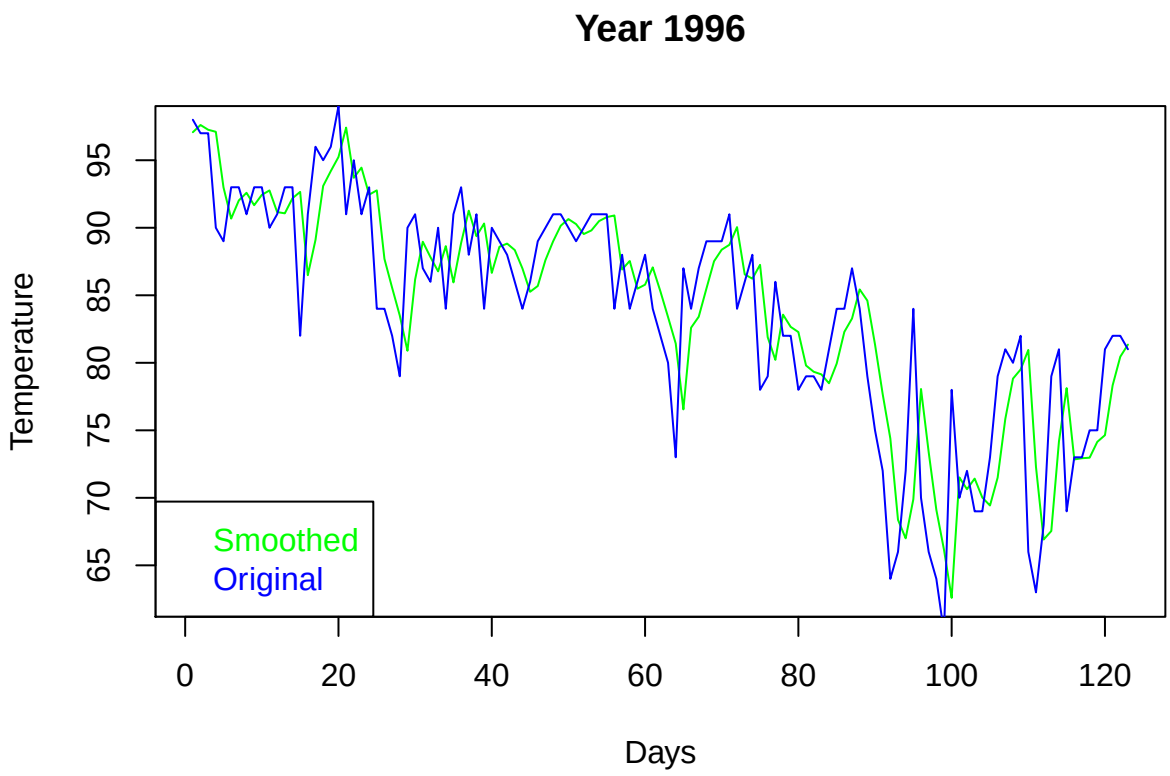
```
plot(year, avg, xlab='Year', ylab='Average Temperature', main='Average Summer Temperature by Year - Hol
```

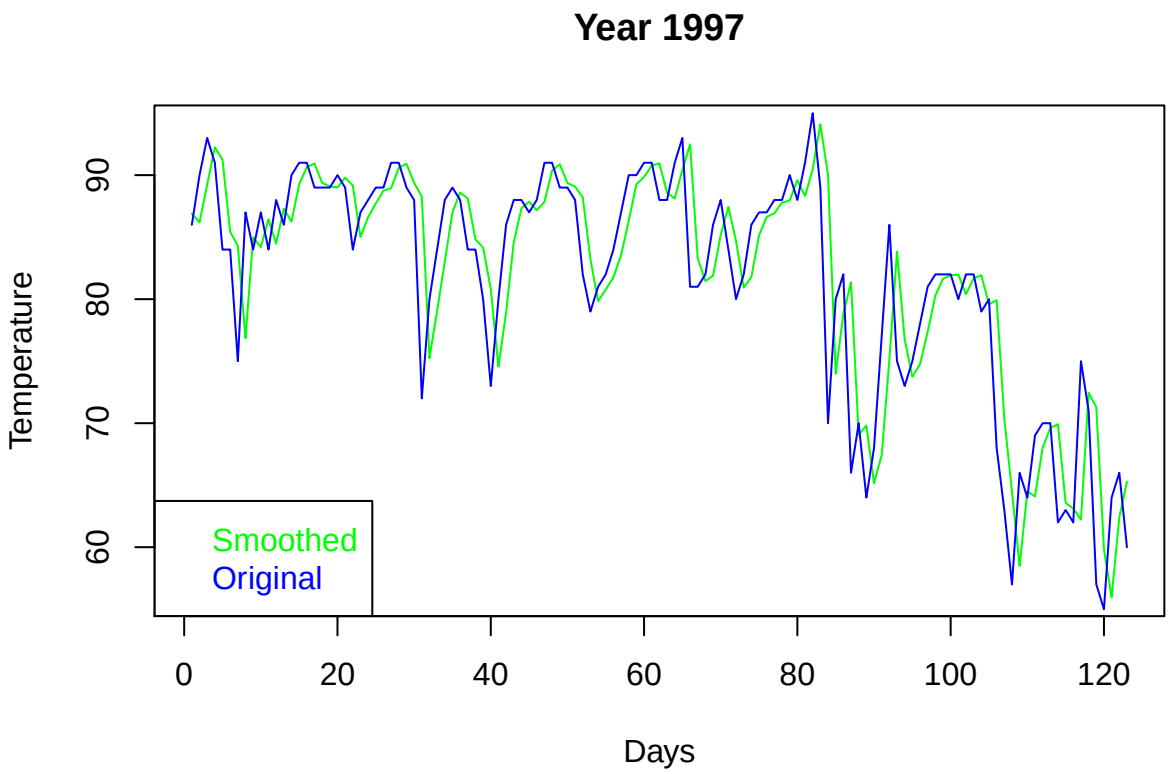


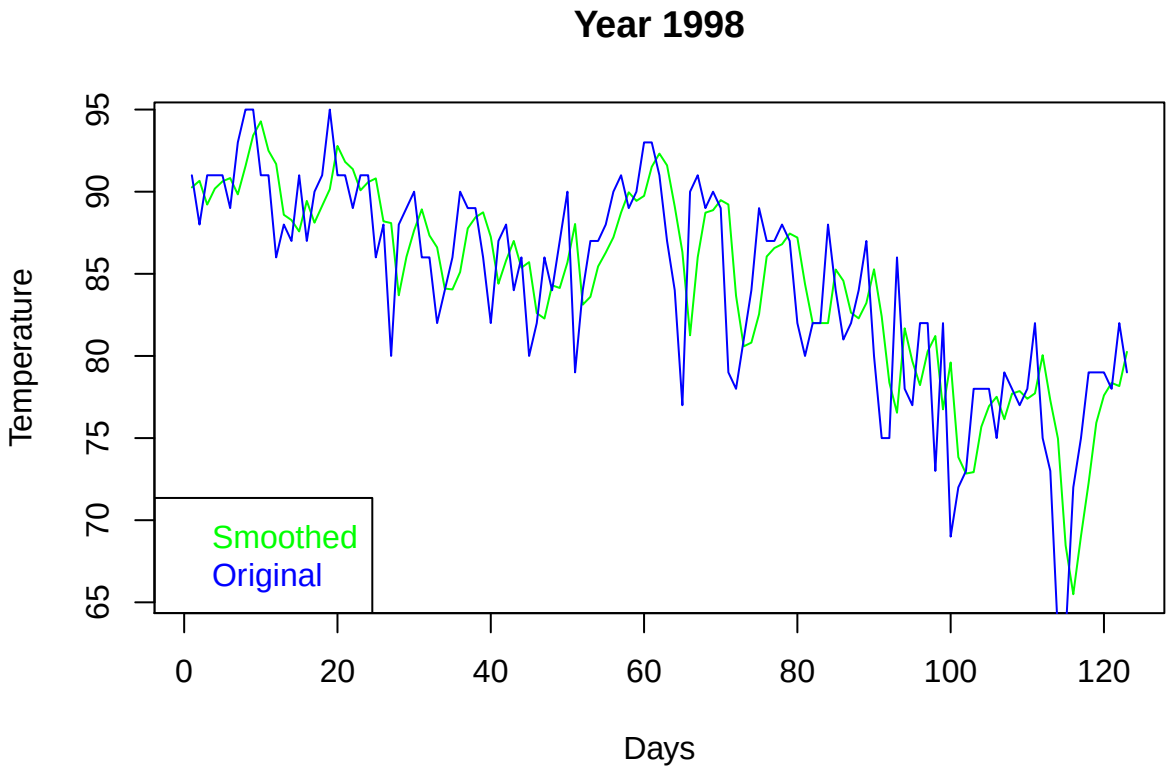
7.2 with EST Method

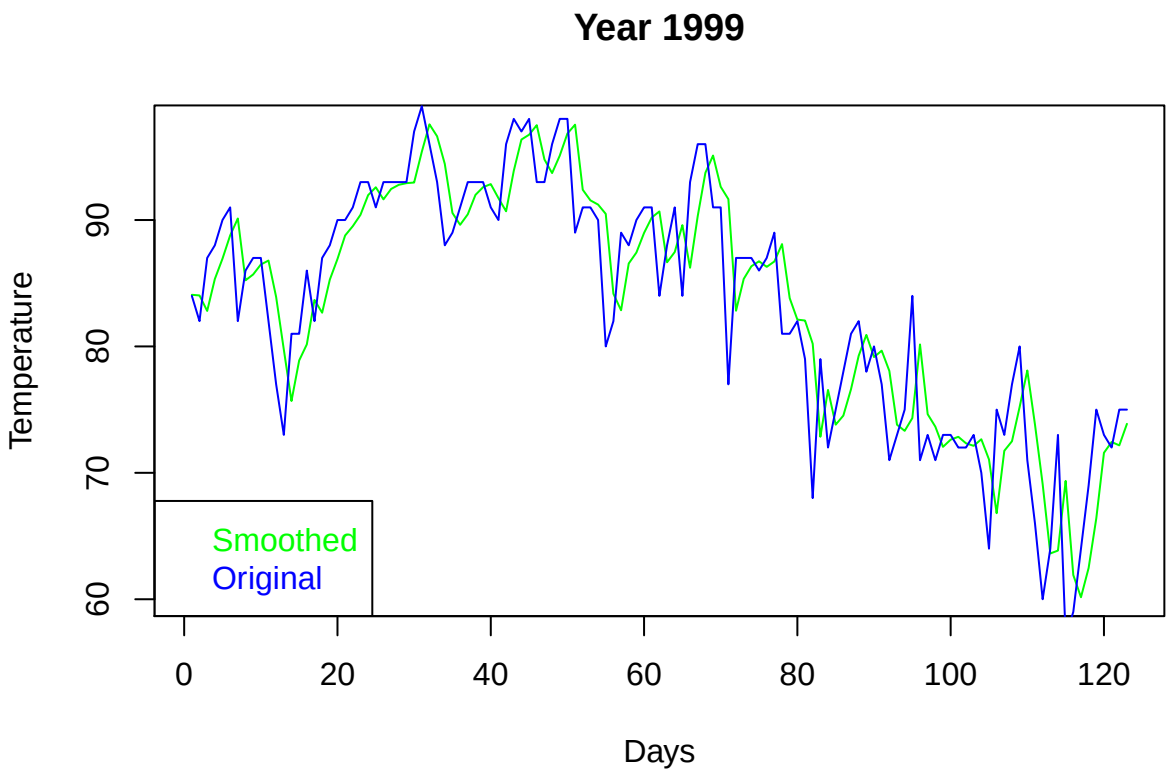
```
new_mat = c()
```

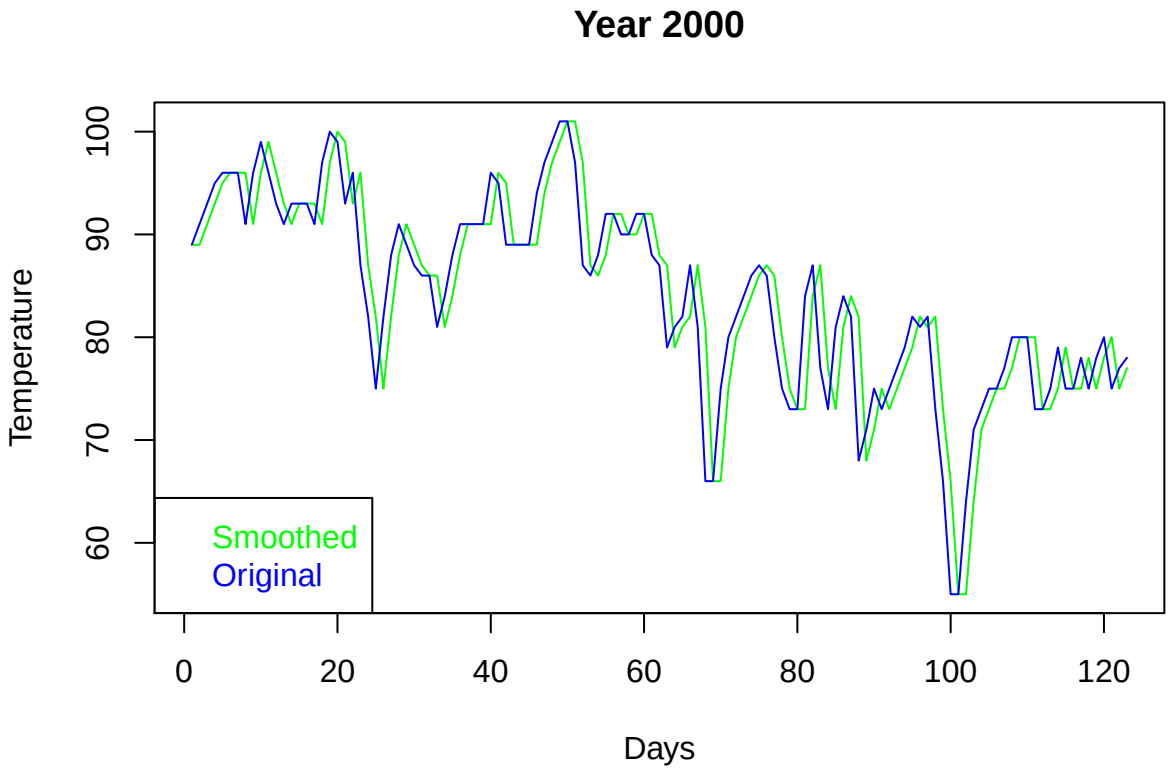
```r
for (i in 2:ncol(temp))
  {

    esmod = ets(temp[,i],model="ZZZ", damped = NULL, alpha=NULL, beta=NULL, gamma = NULL)
    res = fitted.values(esmod)

    new_mat = c(new_mat, res)
    plot(res, col="green", xlab="Days", ylab="Temperature", main=paste("Year", colnames(temp[i])))
    lines(as.numeric(temp[,i]), col="blue")
    legend( "bottomleft", c("Smoothed", "Original"),
    text.col=c("green", "blue"))

}
```

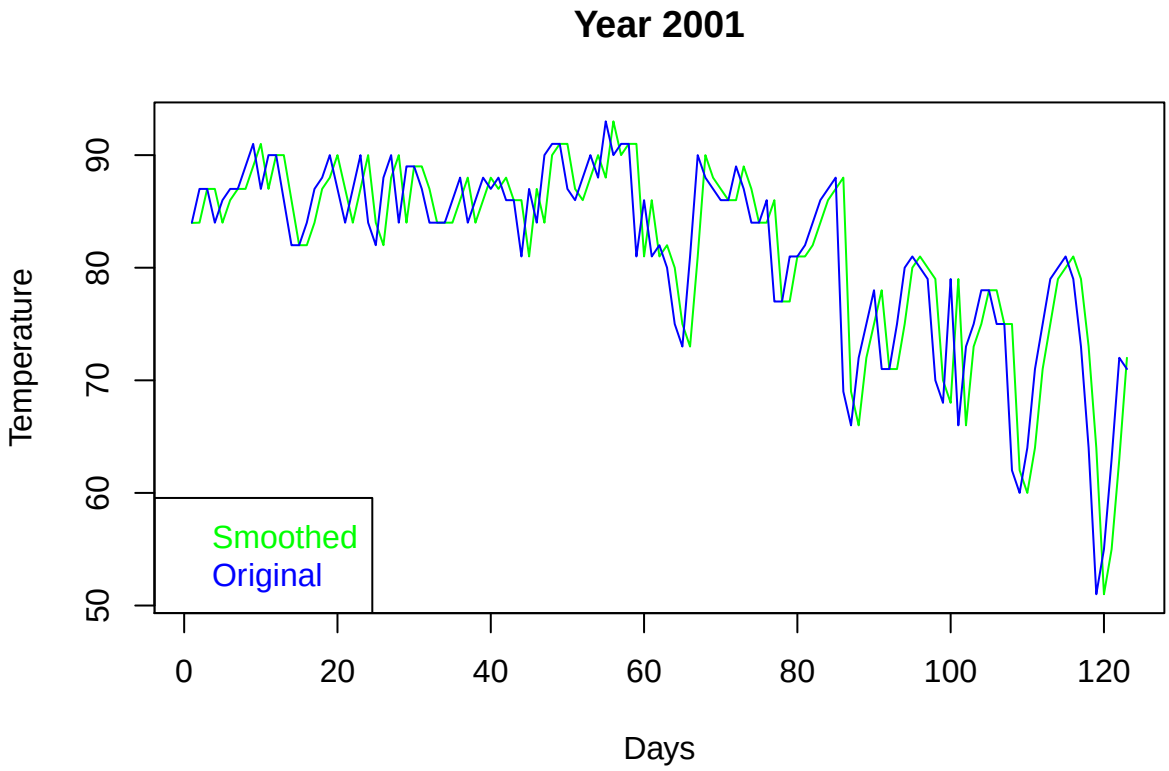


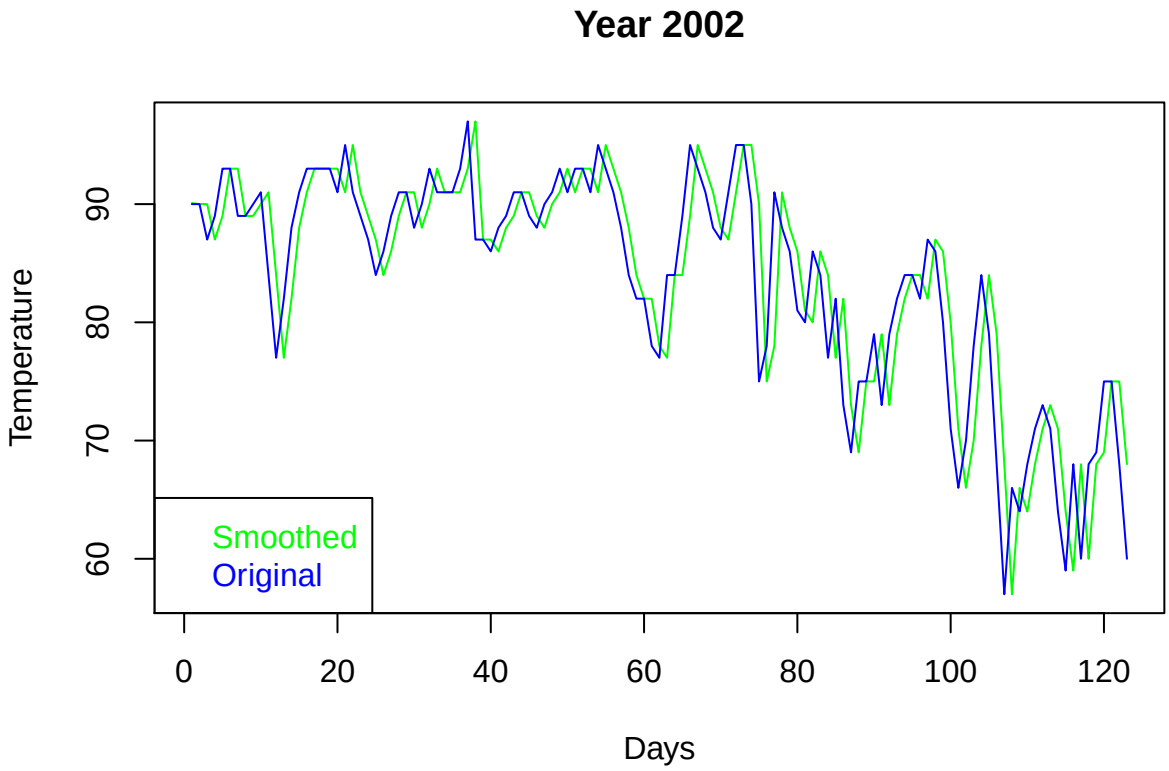

Year 1997
Temperature
60
70
80
90
0
20
40
60
80
100
120
Days
Smoothed
Original

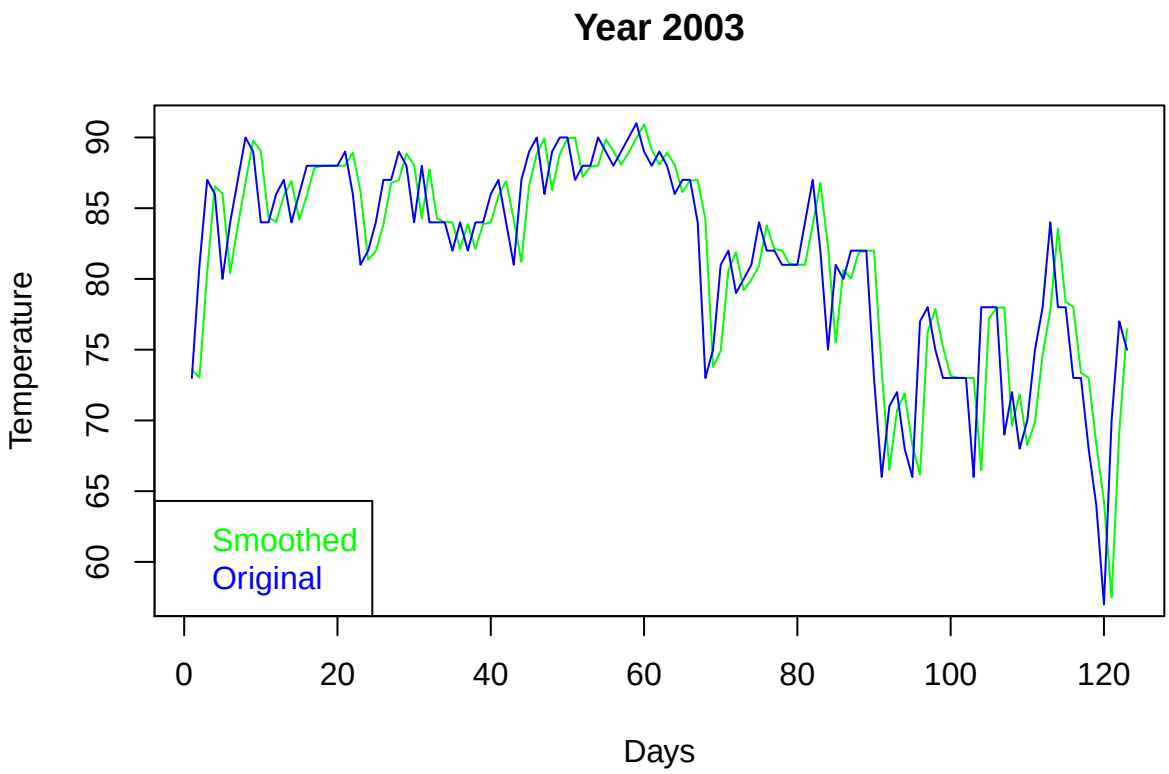

Year 1998
Temperature
65
70
75
80
85
90
95
0
20
40
60
80
100
120
Days
Smoothed
Original

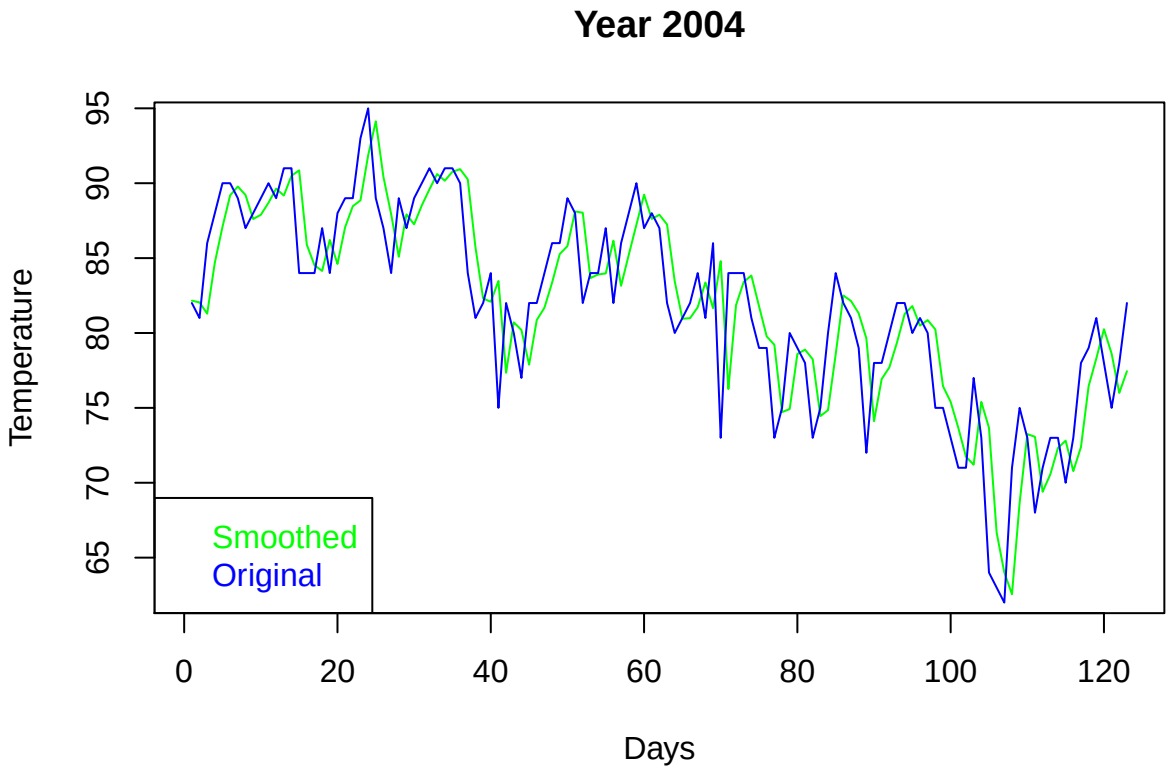

Year 1999
Temperature
60
70
80
90
0
20
40
60
80
100
120
Days
Smoothed
Original

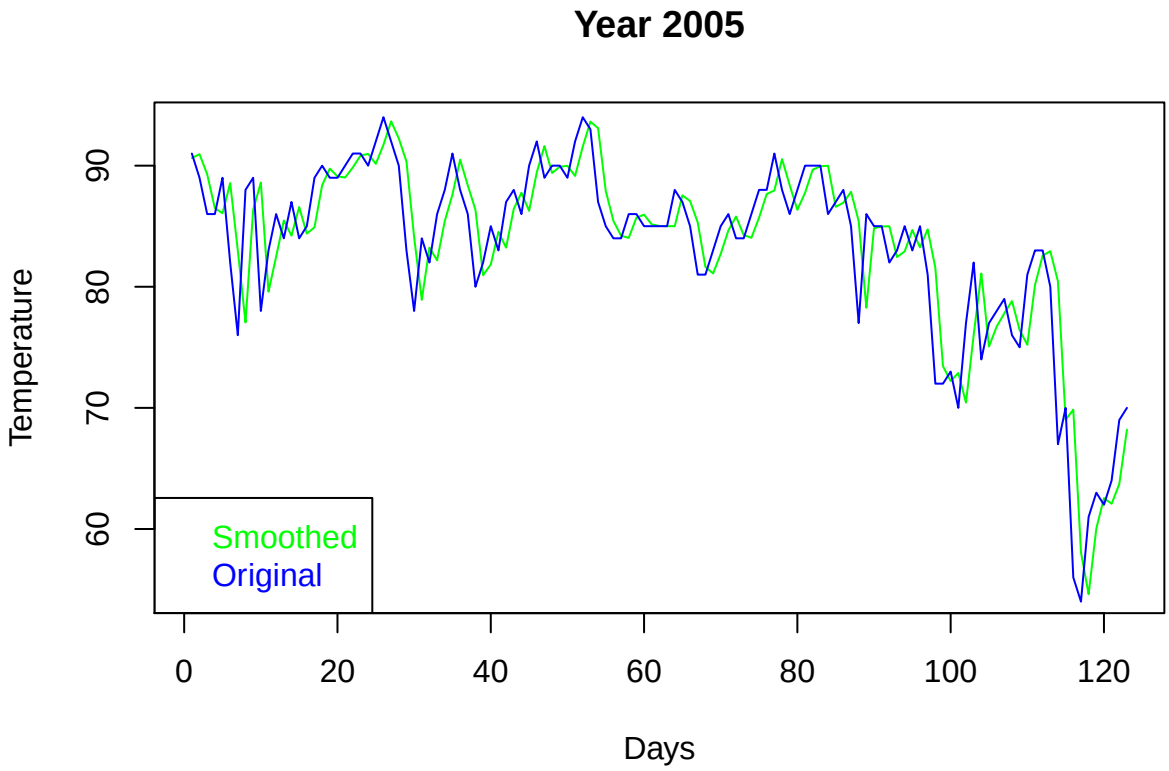

Year 2000
Temperature
60
70
80
90
100
0
20
40
60
80
100
120
Days
Smoothed
Original

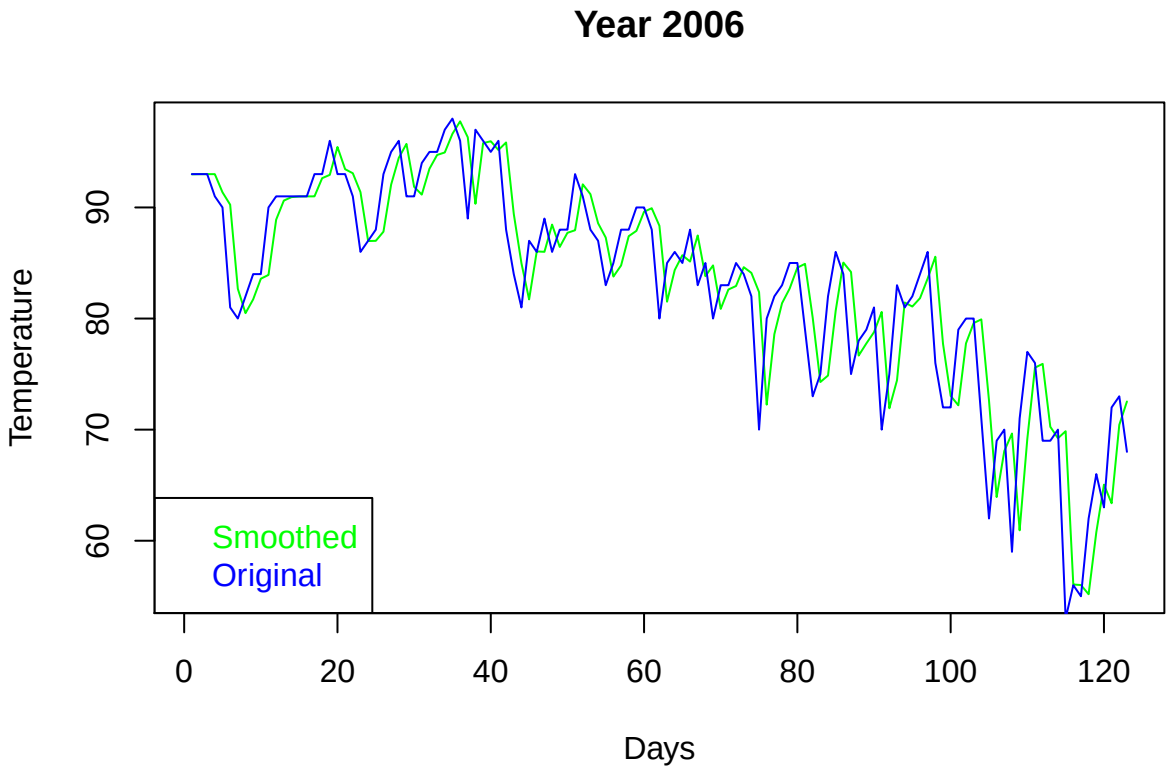

Year 2001
Temperature
50
60
70
80
90
0
20
40
60
80
100
120
Days
Smoothed
Original

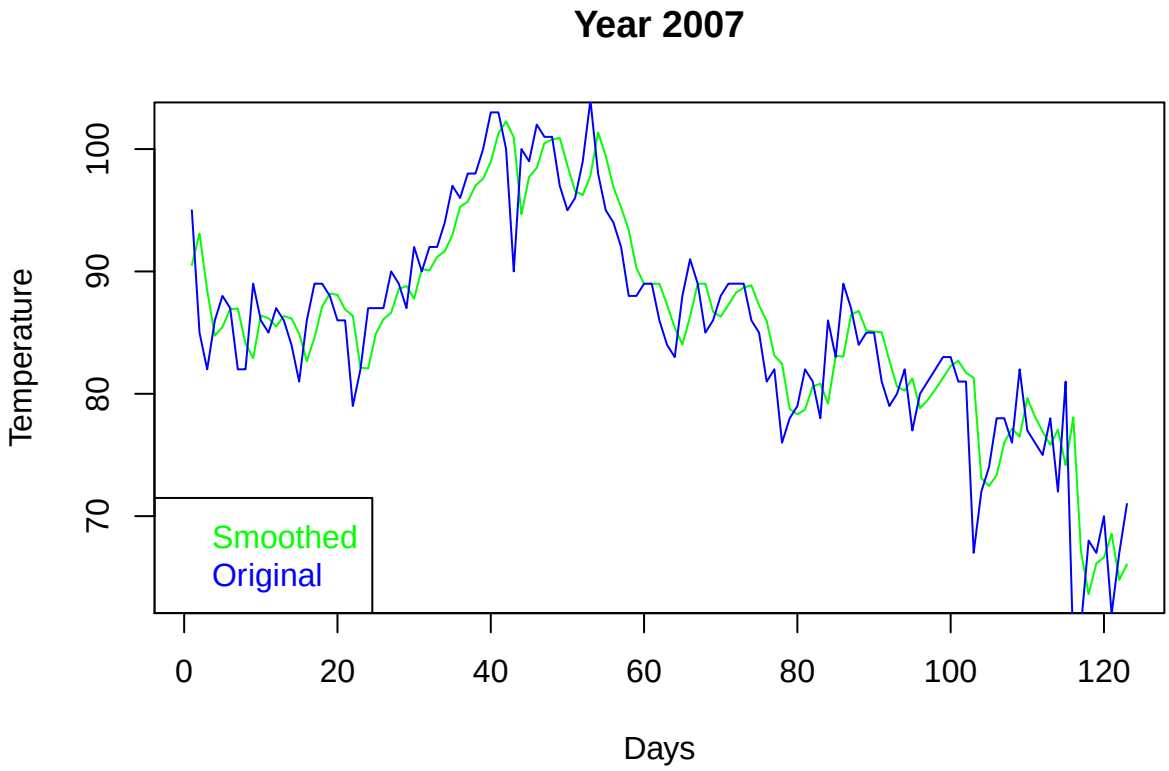

Year 2002
Temperature
60
70
80
90
0
20
40
60
80
100
120
Days
Smoothed
Original

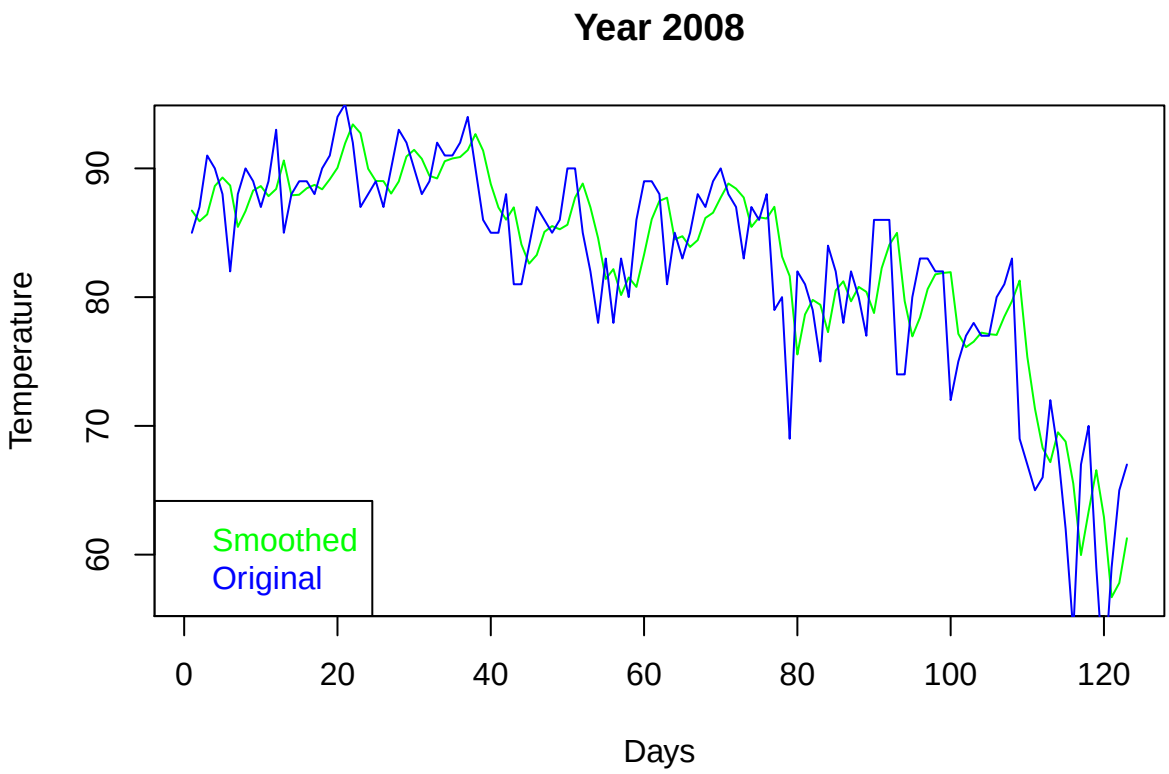

Year 2003
Temperature
60
65
70
75
80
85
90
0
20
40
60
80
100
120
Days
Smoothed
Original

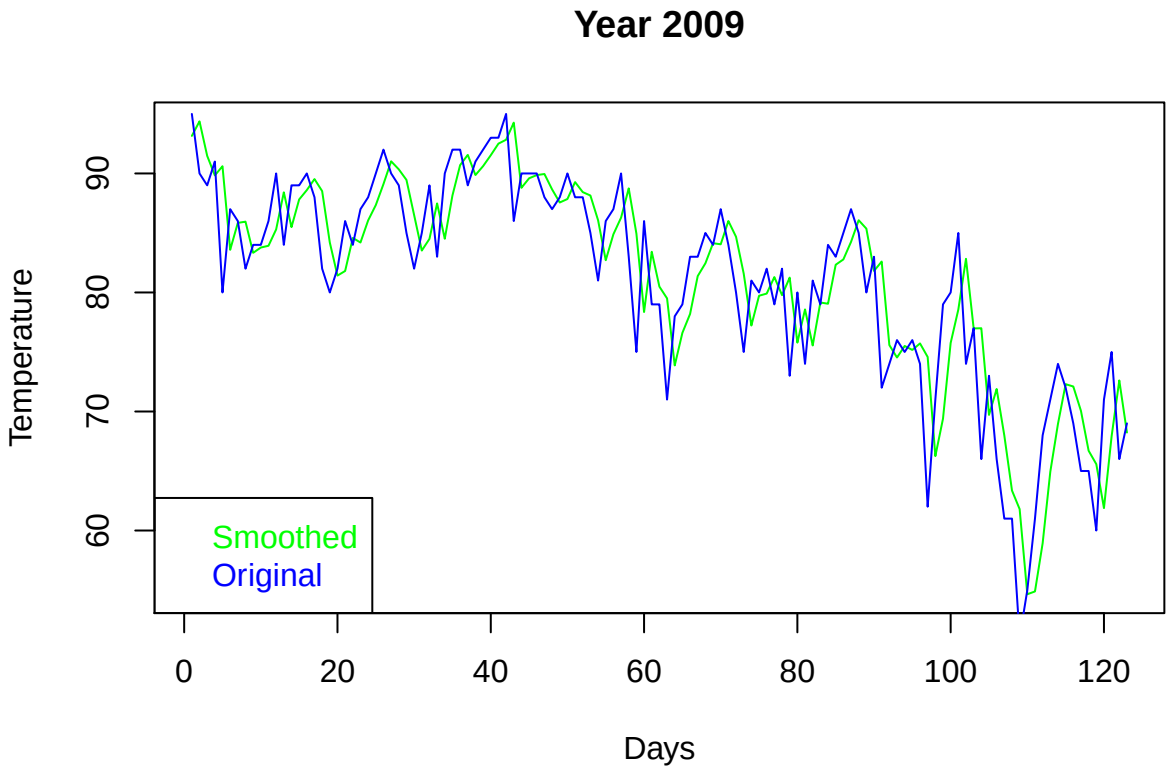

Year 2004
Temperature
65
70
75
80
85
90
95
0
20
40
60
80
100
120
Days
Smoothed
Original

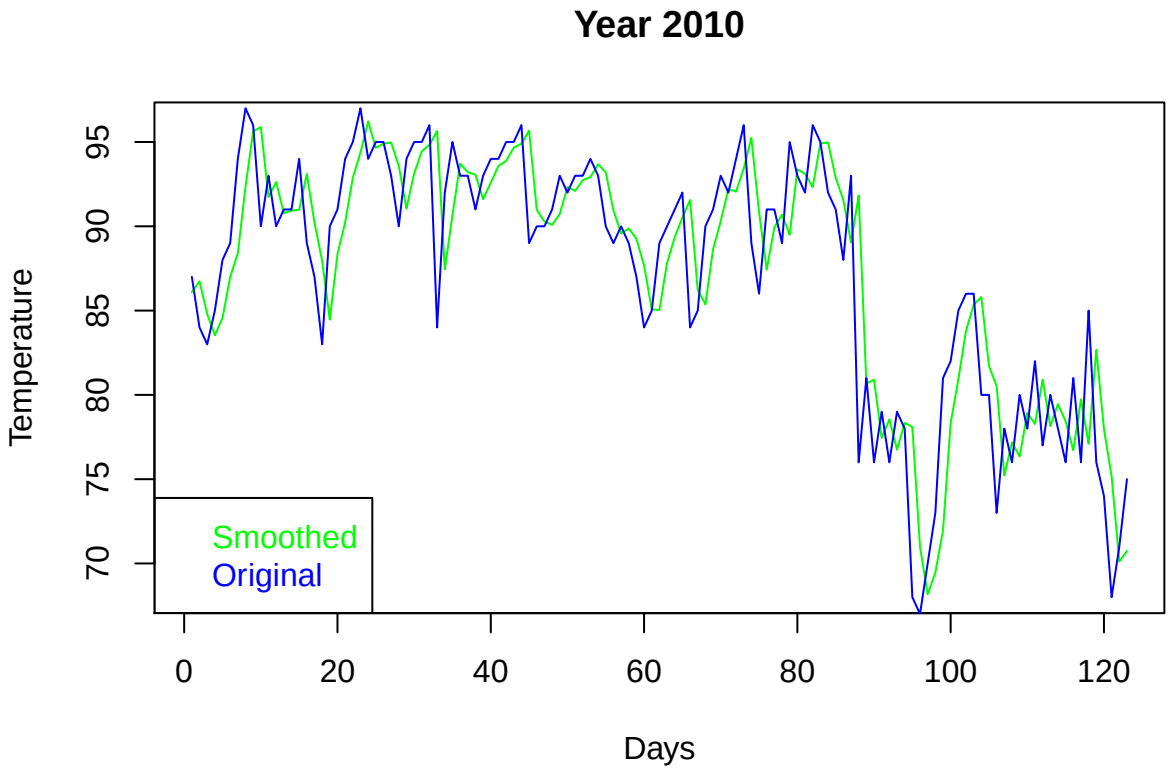

Year 2005
Temperature
60
70
80
90
0
20
40
60
80
100
120
Days
Smoothed
Original

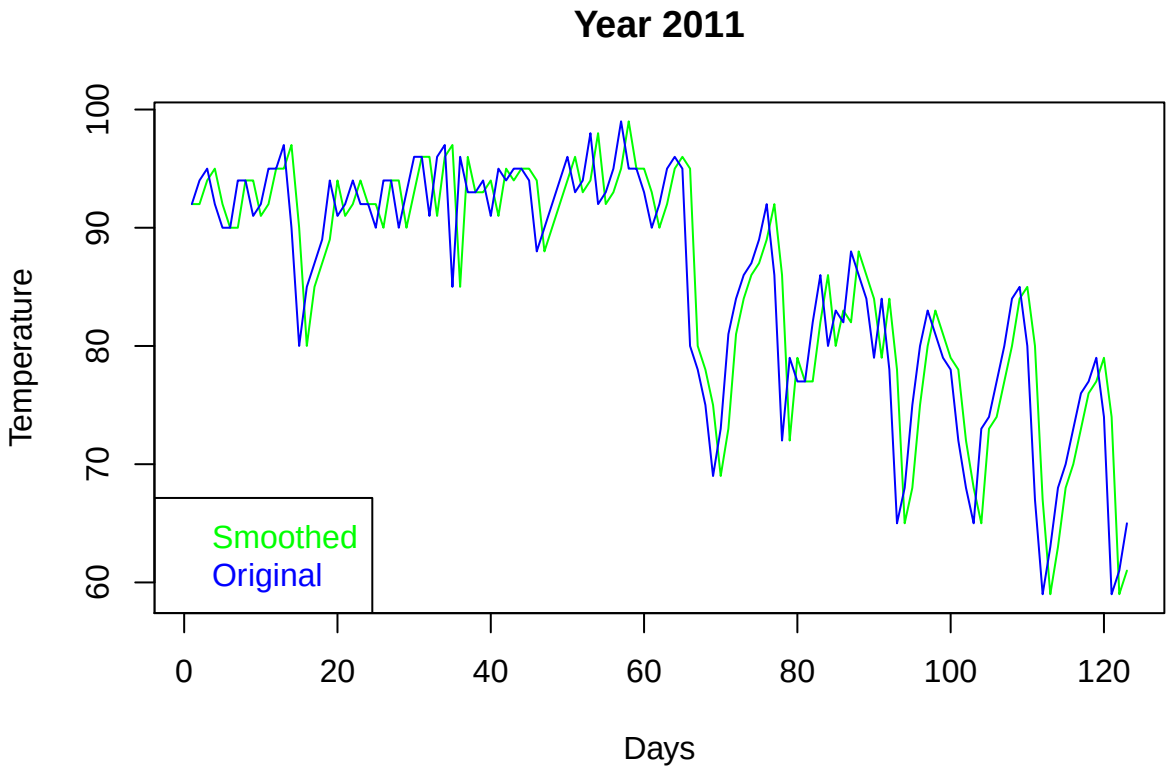

Year 2006
Temperature
60
70
80
90
0
20
40
60
80
100
120
Days
Smoothed
Original

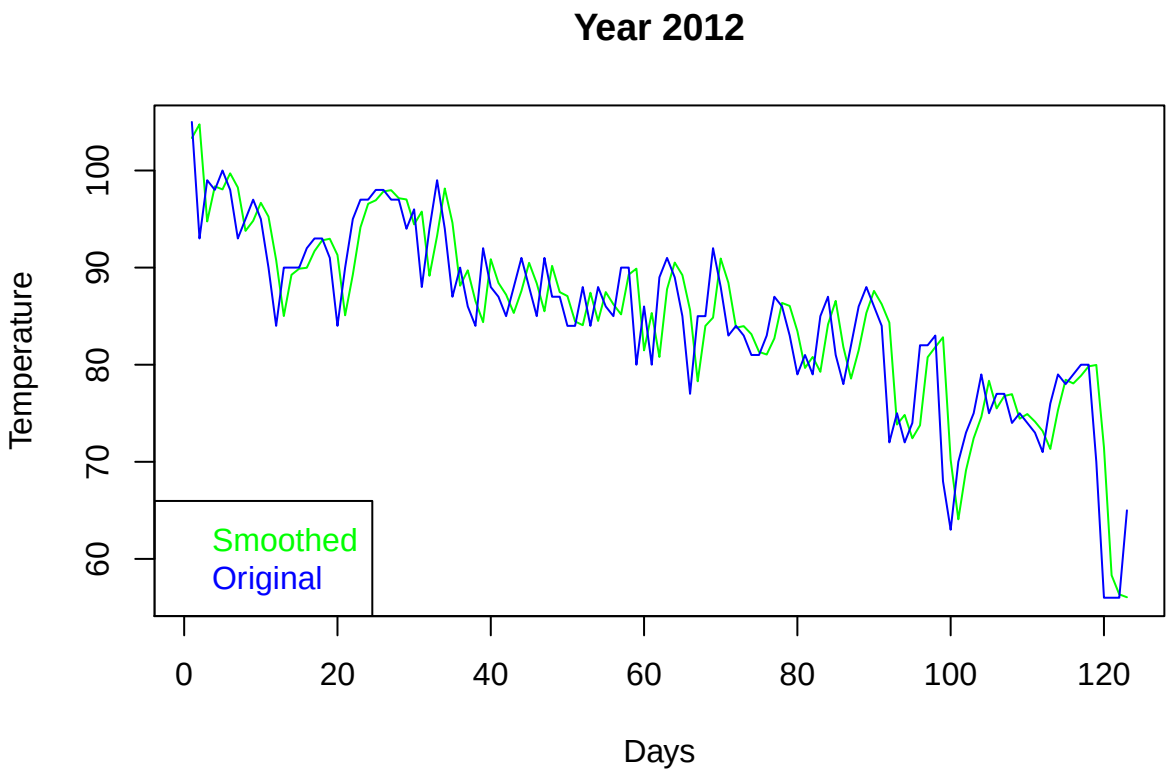

Year 2007
Temperature
70
80
90
100
0
20
40
60
80
100
120
Days
Smoothed
Original

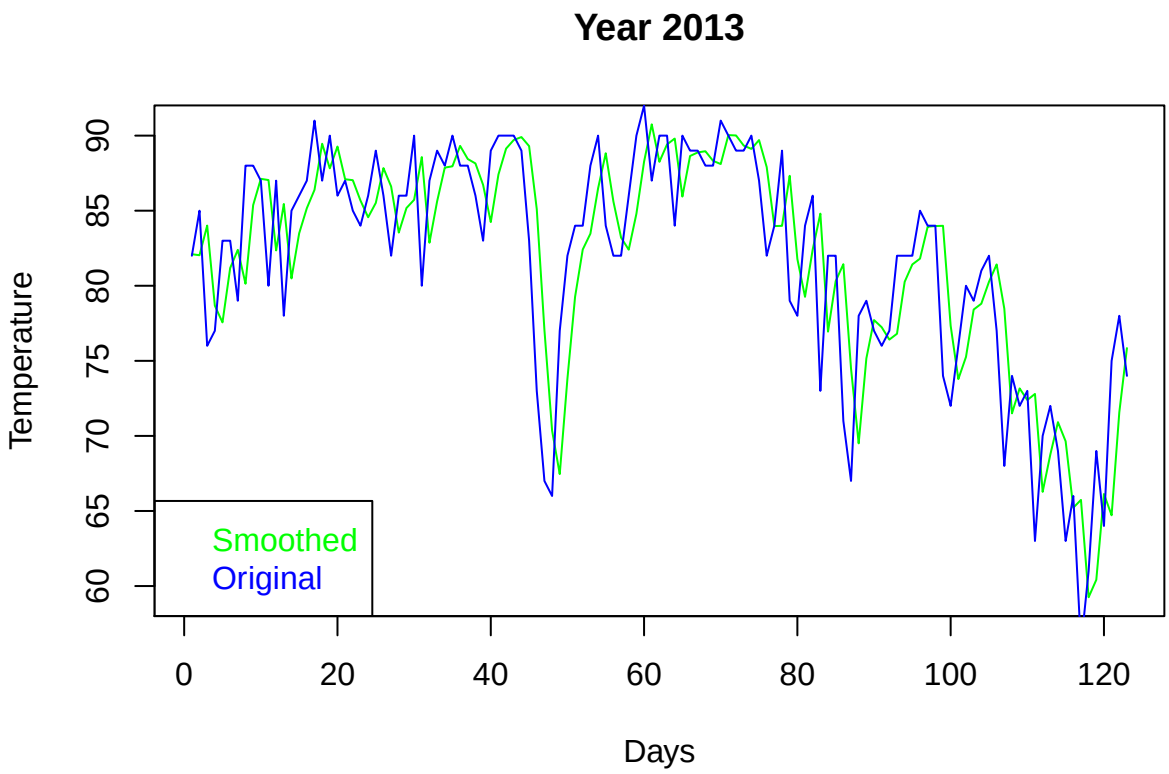

Year 2008
Temperature
90
80
70
60
Smoothed
Original
0
20
40
60
80
100
120
Days

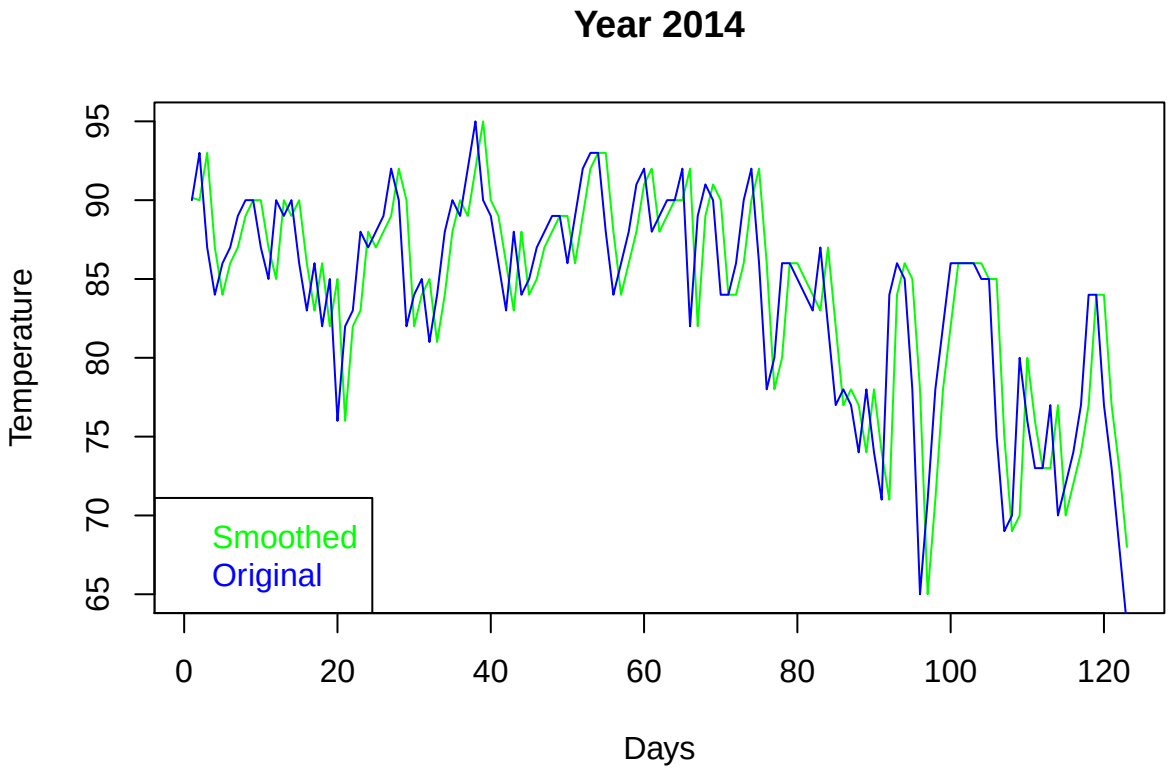

Year 2009
Temperature
60
70
80
90
0
20
40
60
80
100
120
Days
Smoothed
Original

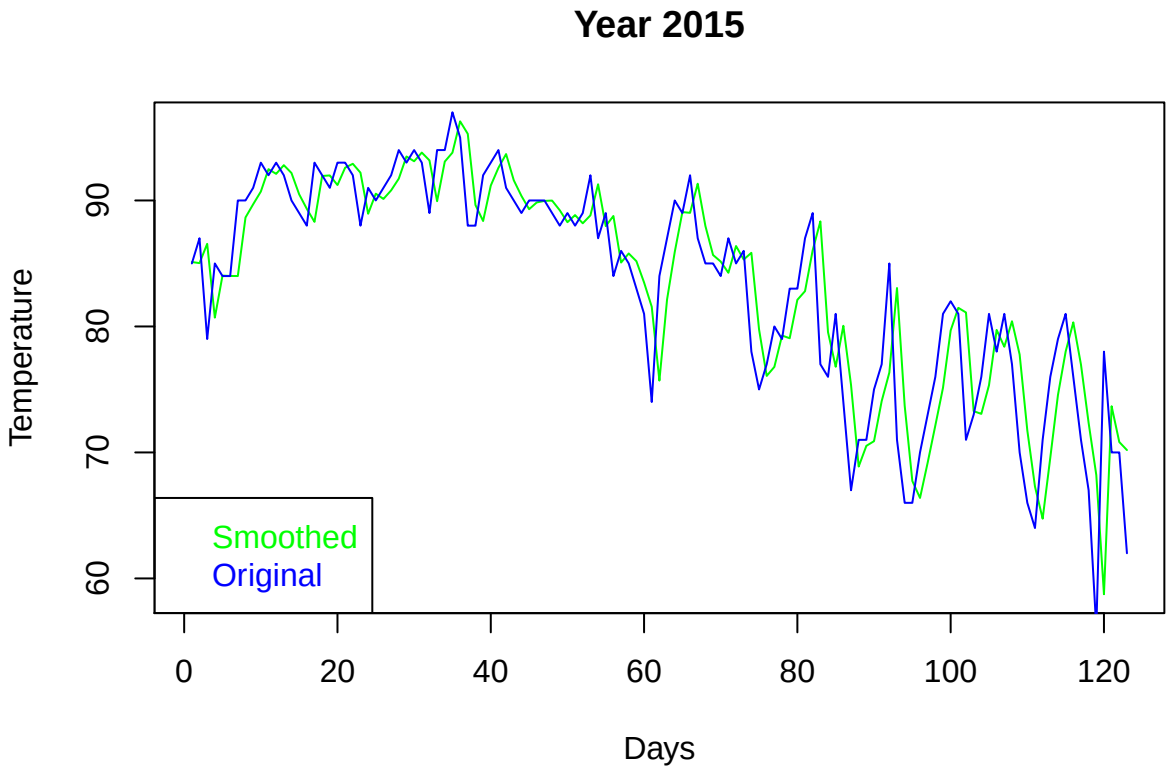

Year 2010
Temperature
70
75
80
85
90
95
0
20
40
60
80
100
120
Days
Smoothed
Original


Year 2011
Temperature
Days
Smoothed
Original
60
70
80
90
100
0
20
40
60
80
100
120


Year 2012
Temperature
60
70
80
90
100
0
20
40
60
80
100
120
Days
Smoothed
Original


Year 2013
Temperature
60
65
70
75
80
85
90
0
20
40
60
80
100
120
Days
Smoothed
Original


Year 2014
Temperature
65
70
75
80
85
90
95
0
20
40
60
80
100
120
Days
Smoothed
Original




```
new_mat = matrix(unlist(res),nrow = 123, ncol = 20, byrow=TRUE)

new_mat_v = as.vector(t(new_mat))

c = 0.5
t = 32
ends = c()
s = rep(0, nrow(temp))
for (j in 1:ncol(new_mat))
  {
    a = mean(as.numeric(new_mat[,j]))
    l = c()
    for (i in 1:nrow(new_mat))
      {
        d =  a - as.numeric(new_mat[i,j]) - c

        if (i == 1)
          s[i] = max(0, d)
        else
          s[i] = max(0, d+s[i-1])

        if (s[i] > t)
          {
            cat("The temperature shifts downward in", colnames(temp[j+1]),"is: ", as.character(temp[i,1]
            ends = c(ends, i)
            break
```

```
          }


      }

  }
```

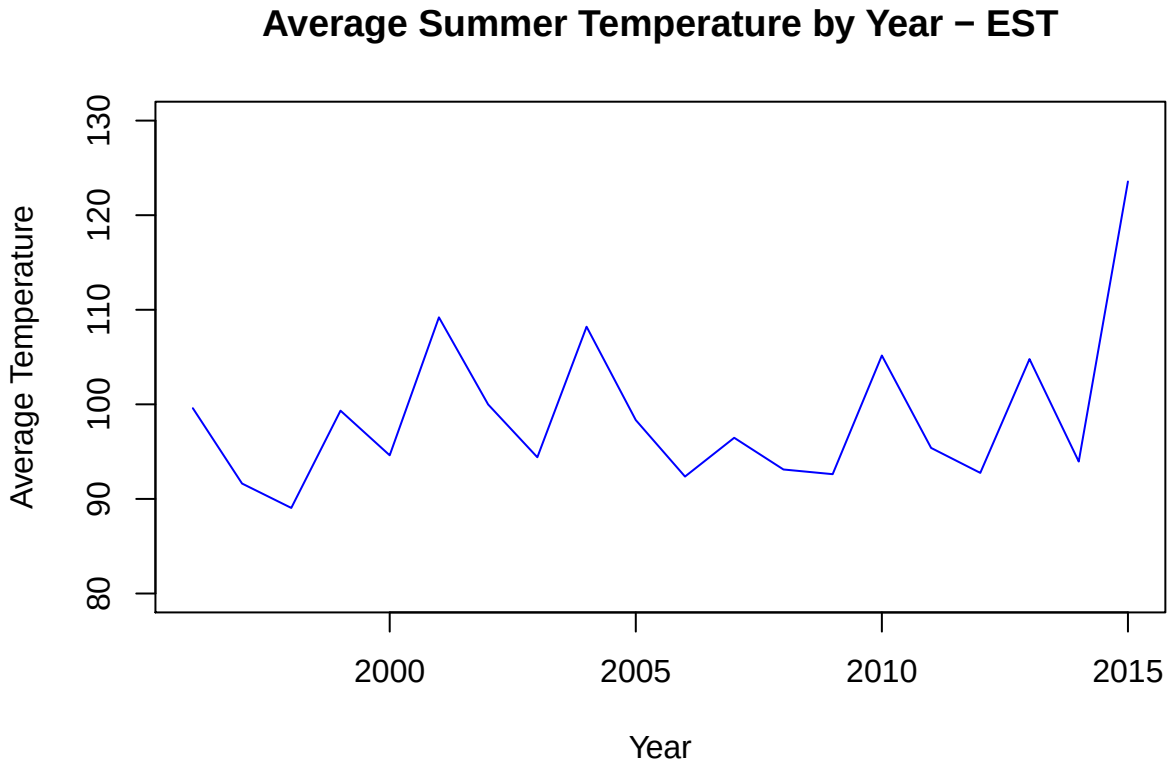
```
## The temperature shifts downward in 1996 is:  19-Aug
## The temperature shifts downward in 1997 is:  1-Oct
## The temperature shifts downward in 1998 is:  13-Jul
## The temperature shifts downward in 1999 is:  25-Aug
## The temperature shifts downward in 2000 is:  7-Oct
## The temperature shifts downward in 2001 is:  19-Jul
## The temperature shifts downward in 2002 is:  31-Aug
## The temperature shifts downward in 2003 is:  13-Oct
## The temperature shifts downward in 2004 is:  25-Jul
## The temperature shifts downward in 2005 is:  6-Sep
## The temperature shifts downward in 2006 is:  19-Oct
## The temperature shifts downward in 2007 is:  31-Jul
## The temperature shifts downward in 2008 is:  12-Sep
## The temperature shifts downward in 2009 is:  25-Oct
## The temperature shifts downward in 2010 is:  6-Aug
## The temperature shifts downward in 2011 is:  18-Sep
## The temperature shifts downward in 2012 is:  31-Oct
## The temperature shifts downward in 2013 is:  12-Aug
## The temperature shifts downward in 2014 is:  24-Sep
## The temperature shifts downward in 2015 is:  12-Jul
```

```
avg = c()
year = c()

for (i in 1:length(ends))
{
  temp_var = mean(mat[1:ends[i]-1,i])


  avg = c(avg,temp_var)

  cat("The average summer temperature in", colnames(temp[i+1]), "is", temp_var, "\n")
  year = c(year,colnames(temp[i+1]))
}
```

```
## The average summer temperature in 1996 is 99.59135
## The average summer temperature in 1997 is 91.62754
## The average summer temperature in 1998 is 89.04654
## The average summer temperature in 1999 is 99.33317
## The average summer temperature in 2000 is 94.61146
## The average summer temperature in 2001 is 109.199
## The average summer temperature in 2002 is 99.9906
## The average summer temperature in 2003 is 94.40284
## The average summer temperature in 2004 is 108.2092
## The average summer temperature in 2005 is 98.33724
## The average summer temperature in 2006 is 92.36435
```

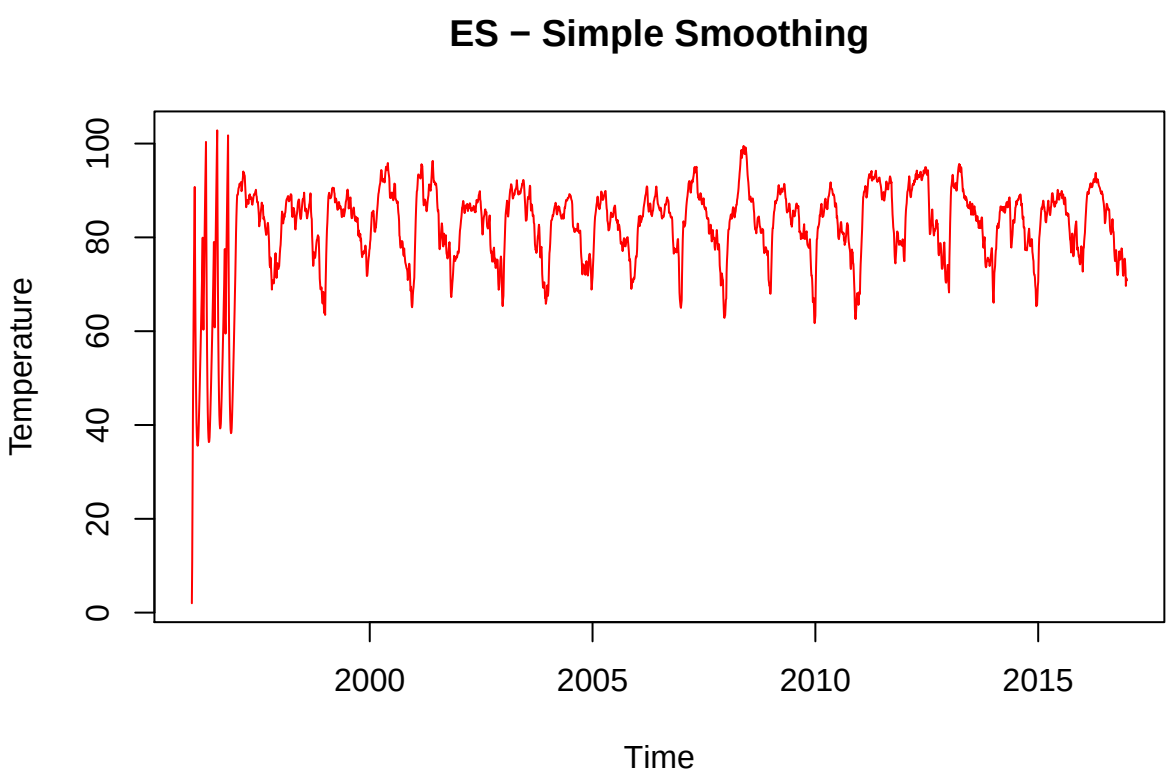
```
## The average summer temperature in 2007 is 96.46432
## The average summer temperature in 2008 is 93.11243
## The average summer temperature in 2009 is 92.61752
## The average summer temperature in 2010 is 105.1622
## The average summer temperature in 2011 is 95.3946
## The average summer temperature in 2012 is 92.75144
## The average summer temperature in 2013 is 104.7882
## The average summer temperature in 2014 is 93.95396
## The average summer temperature in 2015 is 123.5601
```

```
plot(year, avg, xlab='Year', ylab='Average Temperature', main='Average Summer Temperature by Year - EST
```



HoltWinters - simple exponential smoothing

```
mod1 = HoltWinters(myts, seasonal="mult", alpha=0.2, beta=FALSE, gamma=FALSE)
plot(fitted(mod1)[,1], col="Red", xlab="Time", ylab="Temperature", main="ES - Simple Smoothing")
```




```
val = fitted(mod)[,1]
mat = matrix(unlist(val),nrow = 123, ncol = 20, byrow=TRUE)




c = 2.0
t = 200
ends = c()
s = rep(0, nrow(temp))
for (j in 1:ncol(mat)-1)
  {
    a = mean(as.numeric(mat[,j]))
    l = c()
    for (i in 1:nrow(mat))
      {
        d =  a - as.numeric(mat[i,j]) - c

        if (i == 1)
          s[i] = max(0, d)
        else
          s[i] = max(0, d+s[i-1])

        if (s[i] > t)
          {
            cat("The temperature shifts downward in", colnames(temp[j+2]),"is: ", as.character(temp[i,1]
```

```
            ends = c(ends, i)
            break
          }

      }

  }
```

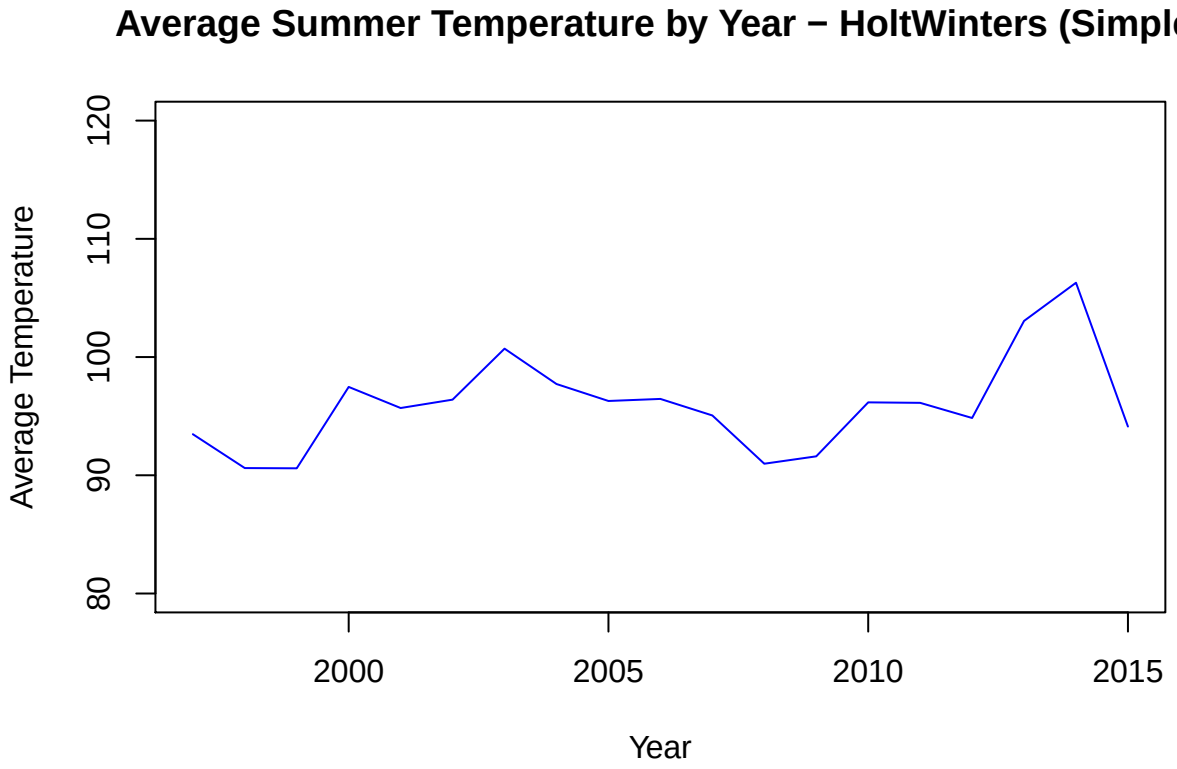
```
## The temperature shifts downward in 1997 is:  8-Oct
## The temperature shifts downward in 1998 is:  14-Oct
## The temperature shifts downward in 1999 is:  17-Oct
## The temperature shifts downward in 2000 is:  5-Sep
## The temperature shifts downward in 2001 is:  28-Sep
## The temperature shifts downward in 2002 is:  11-Sep
## The temperature shifts downward in 2003 is:  25-Aug
## The temperature shifts downward in 2004 is:  16-Sep
## The temperature shifts downward in 2005 is:  17-Sep
## The temperature shifts downward in 2006 is:  17-Sep
## The temperature shifts downward in 2007 is:  22-Sep
## The temperature shifts downward in 2008 is:  25-Sep
## The temperature shifts downward in 2009 is:  23-Sep
## The temperature shifts downward in 2010 is:  22-Sep
## The temperature shifts downward in 2011 is:  22-Sep
## The temperature shifts downward in 2012 is:  22-Sep
## The temperature shifts downward in 2013 is:  23-Aug
## The temperature shifts downward in 2014 is:  6-Aug
## The temperature shifts downward in 2015 is:  21-Sep
```

```
avg = c()
year = c()
maxi=c()
mini = c()
for (i in 1:length(ends))
{
  temp_var = mean(mat[1:ends[i]-1,i])


  avg = c(avg,temp_var)

  cat("The average summer temperature in", colnames(temp[i+2]), "is", temp_var, "\n")
  year = c(year,colnames(temp[i+2]))
}
```

```
## The average summer temperature in 1997 is 93.4741
## The average summer temperature in 1998 is 90.60914
## The average summer temperature in 1999 is 90.58854
## The average summer temperature in 2000 is 97.47308
## The average summer temperature in 2001 is 95.6899
## The average summer temperature in 2002 is 96.39801
## The average summer temperature in 2003 is 100.7104
## The average summer temperature in 2004 is 97.71887
## The average summer temperature in 2005 is 96.28504
## The average summer temperature in 2006 is 96.459
```

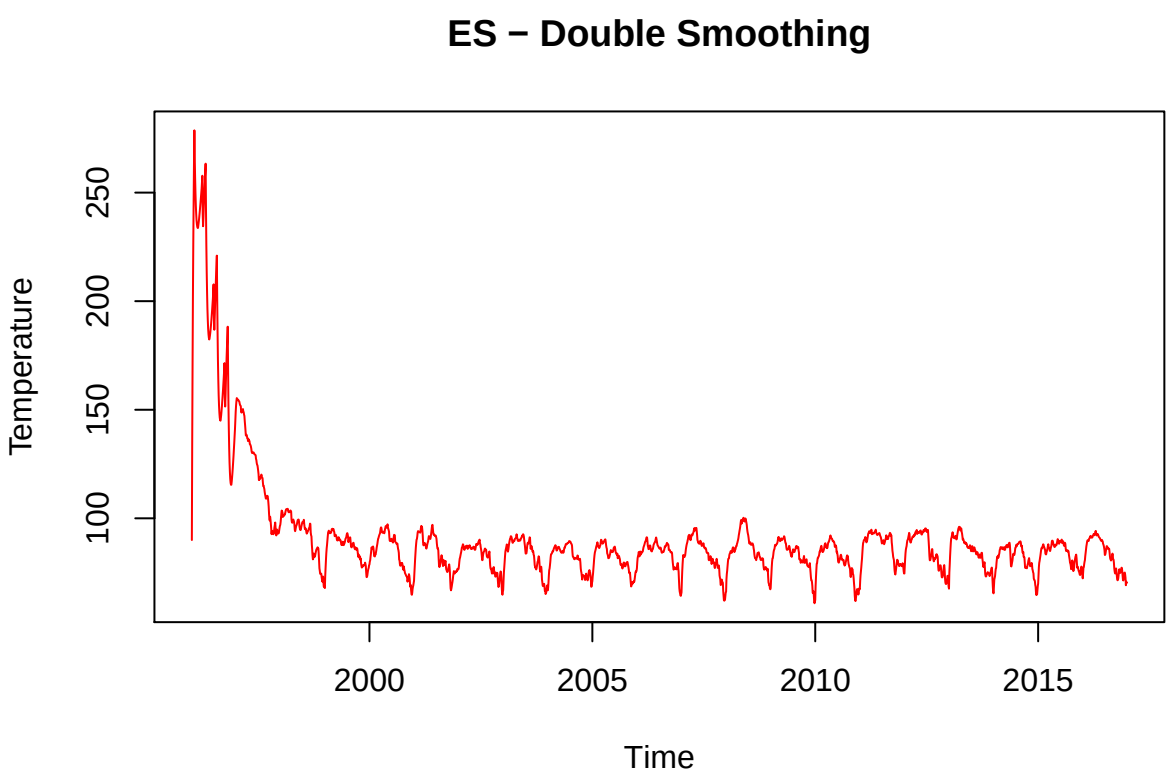
```
## The average summer temperature in 2007 is 95.05942
## The average summer temperature in 2008 is 90.97729
## The average summer temperature in 2009 is 91.60222
## The average summer temperature in 2010 is 96.16791
## The average summer temperature in 2011 is 96.12639
## The average summer temperature in 2012 is 94.84605
## The average summer temperature in 2013 is 103.0572
## The average summer temperature in 2014 is 106.2859
## The average summer temperature in 2015 is 94.1209
```

```
plot(year, avg, xlab='Year', ylab='Average Temperature', main='Average Summer Temperature by Year - Holt
```



HoltWinters - Exponential Smoothing (double - smoothing with trend)

```
mod1 = HoltWinters(myts, seasonal="mult", alpha=0.2, beta=0.01, gamma=FALSE)
plot(fitted(mod1)[,1], col="Red", xlab="Time", ylab="Temperature", main="ES - Double Smoothing")
```




```
val = fitted(mod)[,1]
mat = matrix(unlist(val),nrow = 123, ncol = 20, byrow=TRUE)




c = 2.0
t = 200
ends = c()
s = rep(0, nrow(temp))
for (j in 1:ncol(mat)-1)
  {
    a = mean(as.numeric(mat[,j]))
    l = c()
    for (i in 1:nrow(mat))
      {
        d =  a - as.numeric(mat[i,j]) - c

        if (i == 1)
          s[i] = max(0, d)
        else
          s[i] = max(0, d+s[i-1])

        if (s[i] > t)
          {
            cat("The temperature shifts downward in", colnames(temp[j+2]),"is: ", as.character(temp[i,1]
```

```
            ends = c(ends, i)
            break
          }

      }

  }
```

```
## The temperature shifts downward in 1997 is:  8-Oct
## The temperature shifts downward in 1998 is:  14-Oct
## The temperature shifts downward in 1999 is:  17-Oct
## The temperature shifts downward in 2000 is:  5-Sep
## The temperature shifts downward in 2001 is:  28-Sep
## The temperature shifts downward in 2002 is:  11-Sep
## The temperature shifts downward in 2003 is:  25-Aug
## The temperature shifts downward in 2004 is:  16-Sep
## The temperature shifts downward in 2005 is:  17-Sep
## The temperature shifts downward in 2006 is:  17-Sep
## The temperature shifts downward in 2007 is:  22-Sep
## The temperature shifts downward in 2008 is:  25-Sep
## The temperature shifts downward in 2009 is:  23-Sep
## The temperature shifts downward in 2010 is:  22-Sep
## The temperature shifts downward in 2011 is:  22-Sep
## The temperature shifts downward in 2012 is:  22-Sep
## The temperature shifts downward in 2013 is:  23-Aug
## The temperature shifts downward in 2014 is:  6-Aug
## The temperature shifts downward in 2015 is:  21-Sep
```

```
avg = c()
year = c()
maxi=c()
mini = c()
for (i in 1:length(ends))
{
  temp_var = mean(mat[1:ends[i]-1,i])


  avg = c(avg,temp_var)

  cat("The average summer temperature in", colnames(temp[i+2]), "is", temp_var, "\n")
  year = c(year,colnames(temp[i+2]))
}
```

```
## The average summer temperature in 1997 is 93.4741
## The average summer temperature in 1998 is 90.60914
## The average summer temperature in 1999 is 90.58854
## The average summer temperature in 2000 is 97.47308
## The average summer temperature in 2001 is 95.6899
## The average summer temperature in 2002 is 96.39801
## The average summer temperature in 2003 is 100.7104
## The average summer temperature in 2004 is 97.71887
## The average summer temperature in 2005 is 96.28504
## The average summer temperature in 2006 is 96.459
```

```
## The average summer temperature in 2007 is 95.05942
## The average summer temperature in 2008 is 90.97729
## The average summer temperature in 2009 is 91.60222
## The average summer temperature in 2010 is 96.16791
## The average summer temperature in 2011 is 96.12639
## The average summer temperature in 2012 is 94.84605
## The average summer temperature in 2013 is 103.0572
## The average summer temperature in 2014 is 106.2859
## The average summer temperature in 2015 is 94.1209
```

```
plot(year, avg, xlab='Year', ylab='Average Temperature', main='Average Summer Temperature by Year - Hol
```