IME692A Assignment 1

Rajarshi Dutta
200762

Saubhagya Bawari
200889

1 Question 1

The given question contains data with two columns namely **Expenditure** and **Year**. The **Expenditure** data shows almost an exponentially increasing trend thus **Holt's linear trend model** fits quite properly to the data. On further optimization of the two **level** and **trend** hyper parameters, the **U-statistic** value of **0.933** with α being **0.97** and β being equal to **0.21**.

Here is the figure attached for the forecasted value for the next 15 years from **2005**.

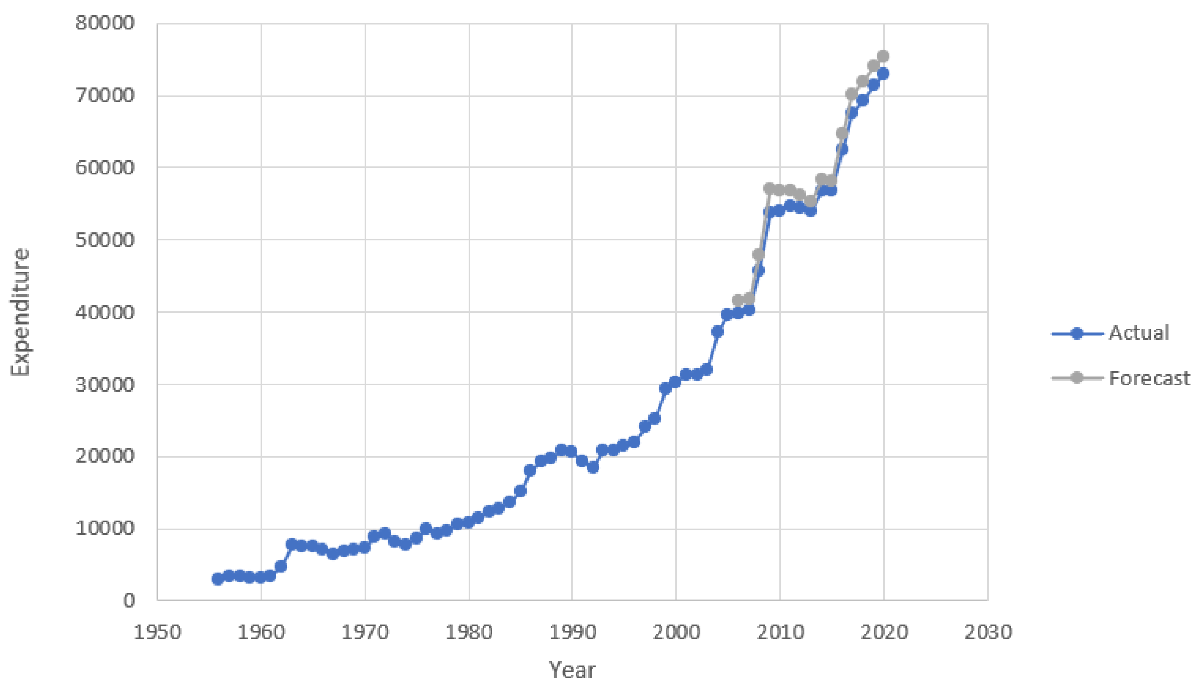

Figure 1: Forecast plot for Question 1

2 Question 2

The second question deals with multiple forecasting. For the first part, I have used two forecasting models like **Adaptive Exponential Smoothing** and **Holt Linear Method** for forecasting for the winter months namely **November, December, January, February** etc.

The pseudocodes for the R scripts are given below:

Preprint. Under review.

Algorithm 1 Holt's Linear Smoothing Method

procedure HOLTMETHOD(data, alpha, beta)
  $n \leftarrow$ length(data)
  Initialize arrays: u1, u2, forecasts, l, b of size n
  $\mathrm{l}[1] \leftarrow \mathrm{data}[1]$
  $\mathrm{b}[1] \leftarrow \mathrm{data}[2] - \mathrm{data}[1]$
  for t from 2 to n **do**
    $\mathrm{L}[t] \leftarrow \alpha \times \mathrm{data}[t] + (1-\alpha) \times (\mathrm{l}[t-1] + \mathrm{b}[t-1])$
    $\mathrm{b}[t] \leftarrow \beta \times (\mathrm{l}[t] - \mathrm{L}[t-1]) + (1-\beta) \times \mathrm{b}[t-1]$
    $\mathrm{forecasts}[t] \leftarrow \mathrm{L}[t-1] + \mathrm{b}[t-1]$
    $\mathrm{u1}[t] \leftarrow ((\mathrm{forecasts}[t] - \mathrm{data}[t])/\mathrm{data}[t-1])^2$
    $\mathrm{u2}[t] \leftarrow ((\mathrm{data}[t] - \mathrm{data}[t-1])/\mathrm{data}[t-1])^2$
  end for
  $\mathrm{forecasts}[n+1] \leftarrow \mathrm{l}[n] + \mathrm{b}[n]$
  $\mathrm{ustat} \leftarrow \sqrt{\sum \mathrm{u1}[1:90] / \sum \mathrm{u2}[1:90]}$
  return $\{$L, b, ustat, forecasts$\}$
end procedure

2.1 Holt's Linear Method

The plots for **Adaptive Exponential Smoothing** method for the two sets of data corresponding to the months are:

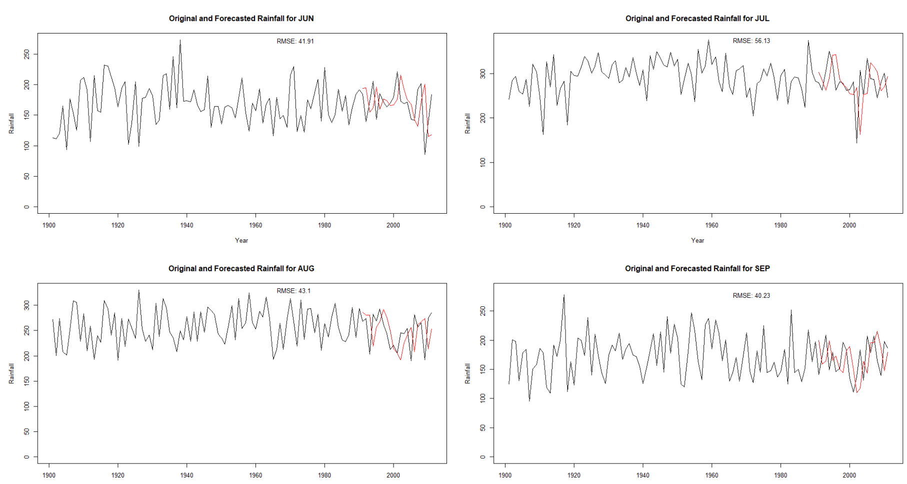

Figure 2: Holt's method for part 1

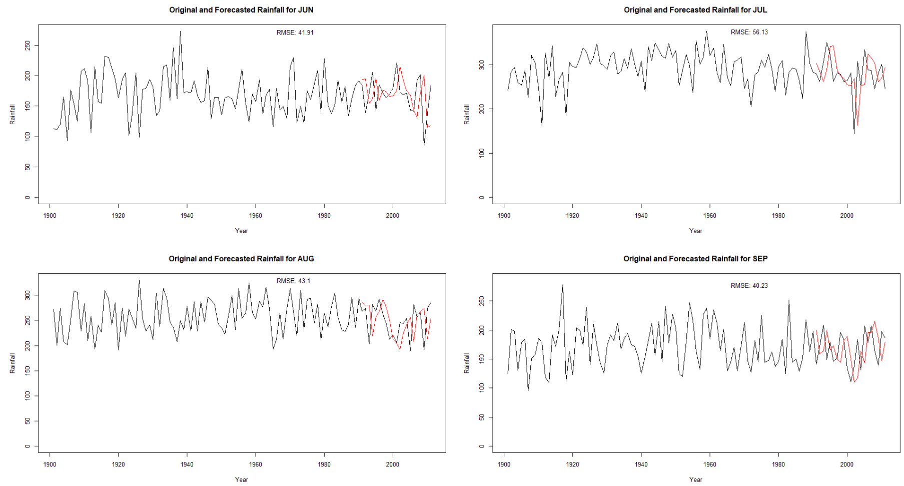

Figure 3: Holt's method for part 2

The model parameters α and β for modelling **winter months** are **0.5** and **0.35** whereas the α and β values for modelling **summer** months are **0.6** and **0.4**. The **RMSE** and **U statistic** values for the case of **summer** and **winter** months are mentioned in the tables:

Month	RMSE	Test U-statistic
November	17.858	0.921
December	13.033	0.901
January	8.961	0.551
February	13.77	0.730

Month	RMSE	Test U-statistic
June	41.907	0.939
July	51.130	0.913
August	43.098	0.920
September	40.226	0.995

2.2 Adaptive Exponential Smoothing

Algorithm 2 Adaptive Exponential Smoothing

function ADAPTIVEETS($data, initial\_alpha, beta$)
  $n \leftarrow$ length of $data$
  Initialize $forecasts, a, m, u1, u2, alpha$ as numeric arrays of size n
  $forecasts[1] \leftarrow data[1]$
  $a[1] \leftarrow 0$
  $m[1] \leftarrow 0$
  $alpha[1] \leftarrow initial\_alpha$
  for t in 2 to n **do**
    $forecasts[t] \leftarrow \alpha[t-1] \times data[t-1] + (1 - alpha[t-1]) \times (a[t-1] + m[t-1])$
    $a[t] \leftarrow \beta \times (data[t] - forecasts[t]) + (1 - beta) \times a[t-1]$
    $m[t] \leftarrow \beta \times |data[t] - forecasts[t]| + (1 - beta) \times m[t-1]$
    $\alpha[t] \leftarrow |a[t]/m[t]|$
    $u1[t] \leftarrow ((forecasts[t] - data[t])/data[t-1])^2$
    $u2[t] \leftarrow ((data[t] - data[t-1])/data[t-1])^2$
  end for
  ustattrain $\leftarrow \sqrt{\frac{\sum_{i=1}^{90} u1[i]}{\sum_{i=1}^{90} u2[i]}}$
  ustattest $\leftarrow \sqrt{\frac{\sum_{i=91}^{111} u1[i]}{\sum_{i=91}^{111} u2[i]}}$
  return { $forecasts, alpha,$ ustattrain, ustattest$, level, trend$ }
end function

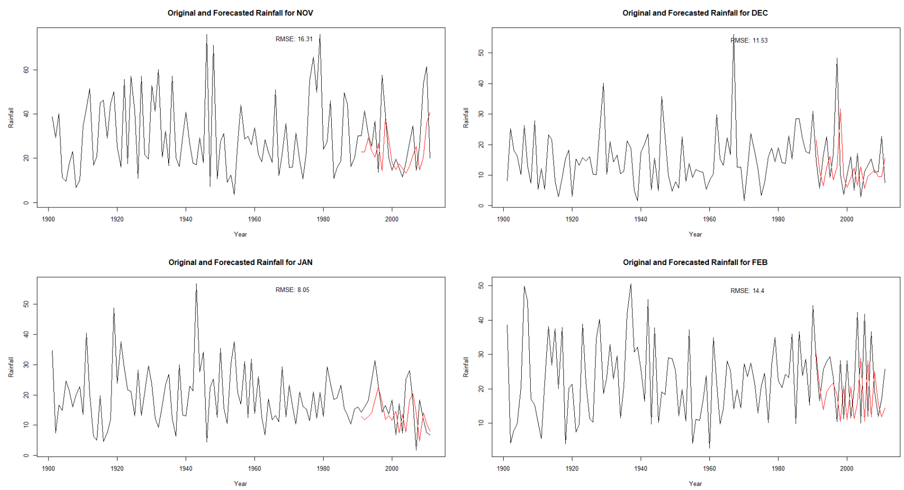

Figure 4: Adaptive method for part 1

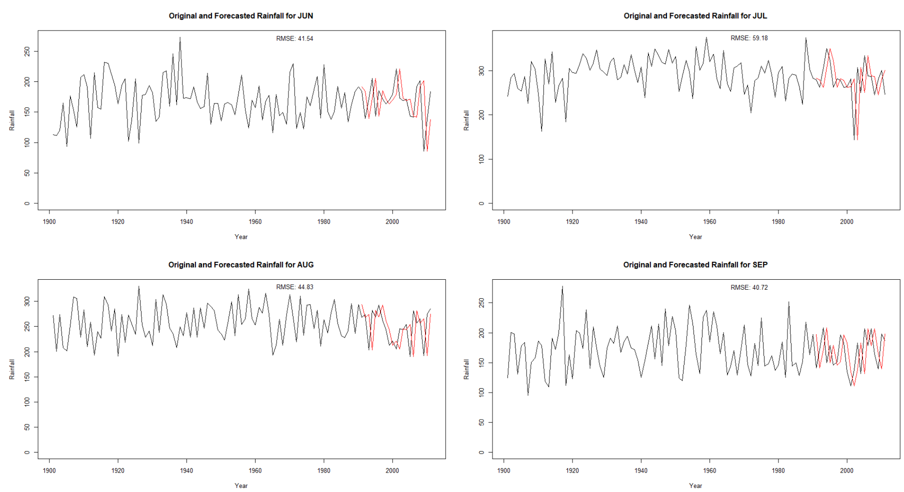

Figure 5: Adaptive method for part 2

The model parameters α and β for modelling **winter months** are **0.3** and **0.01** whereas the α and β values for modelling **summer** months are **0** and **1** (Naive Method). The **RMSE** and **U statistic** values for the case of **summer** and **winter** months are mentioned in the tables:

Month	RMSE	Test U-statistic
November	16.306	0.987
December	11.526	0.842
January	8.052	0.835
February	14.398	0.976

Month	RMSE	Test U-statistic
June	49.544	1
July	54.177	1
August	44.827	1
September	40.716	1

The last **Adaptive model** doesn't work well in case of the **summers** data since the **Test U statistic** always comes out to be greater than 1. Hence the **Naive approach** works better and the **RMSE** results are mentioned in the given table above.

2.3 Conclusion

The hyperparameters used for both the models (**holt's linear method** and **adaptive exponential method** are different for summer and winter months) since the summer data seems to be a bit noisier than the corresponding data for the winter months.