
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**.

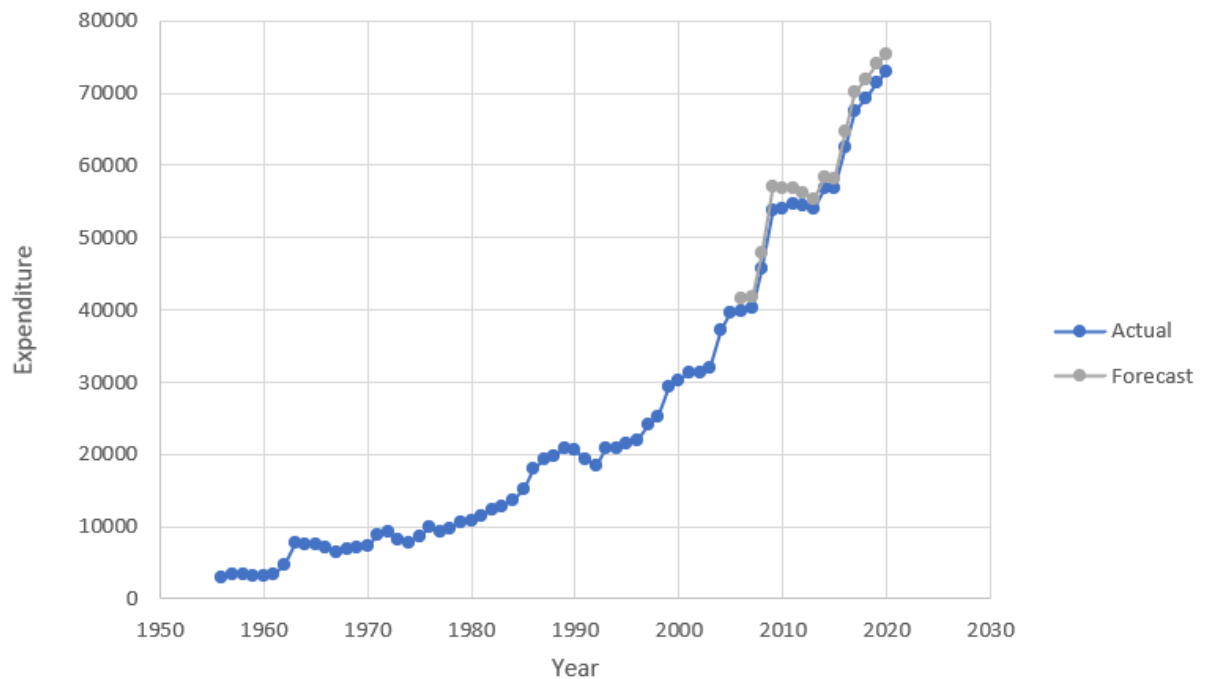


Figure 1: Forecast plot for Question 1

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

```
1: procedure HOLTMETHOD(data, alpha, beta)
2:    $n \leftarrow \text{length}(\text{data})$ 
3:   Initialize arrays: u1, u2, forecasts, L, b of size n
4:    $L[1] \leftarrow \text{data}[1]$ 
5:    $b[1] \leftarrow \text{data}[2] - \text{data}[1]$ 
6:   for t from 2 to n do
7:      $L[t] \leftarrow \alpha \times \text{data}[t] + (1 - \alpha) \times (L[t - 1] + b[t - 1])$ 
8:      $b[t] \leftarrow \beta \times (L[t] - L[t - 1]) + (1 - \beta) \times b[t - 1]$ 
9:      $\text{forecasts}[t] \leftarrow L[t - 1] + b[t - 1]$ 
10:     $u1[t] \leftarrow ((\text{forecasts}[t] - \text{data}[t]) / \text{data}[t - 1])^2$ 
11:     $u2[t] \leftarrow ((\text{data}[t] - \text{data}[t - 1]) / \text{data}[t - 1])^2$ 
12:  end for
13:   $ustatrain \leftarrow \sqrt{\frac{\sum_{i=1}^{90} u1_i}{\sum_{i=1}^{90} u2_i}}$ 
14:   $ustatetest \leftarrow \sqrt{\frac{\sum_{i=91}^{111} u1_i}{\sum_{i=91}^{111} u2_i}}$ 
15:  return L, b, ustatrain, ustatetest, forecasts
16: 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:

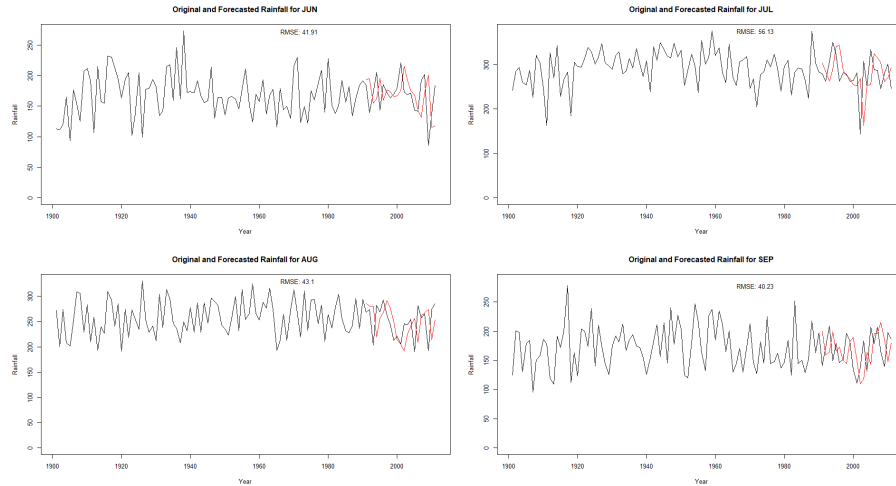


Figure 2: Holt's method for part 1

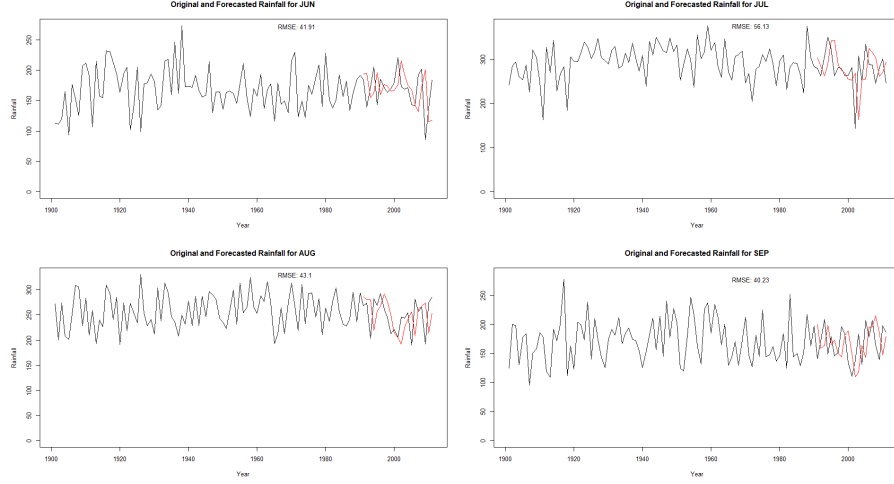


Figure 3: Holt's method for part 2

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

```

1: function ADAPTIVEETS(data, initial_alpha, beta)
2:    $n \leftarrow \text{length of } data$ 
3:   Initialize forecasts, a, m, u1, u2, alpha as numeric arrays of size n
4:    $forecasts[1] \leftarrow data[1]$ 
5:    $a[1] \leftarrow 0$ 
6:    $m[1] \leftarrow 0$ 
7:    $alpha[1] \leftarrow initial\_alpha$ 
8:   for t in 2 to n do
9:      $forecasts[t] \leftarrow \alpha[t-1] \times data[t-1] + (1 - \alpha[t-1]) \times (a[t-1] + m[t-1])$ 
10:     $a[t] \leftarrow \beta \times (data[t] - forecasts[t]) + (1 - \beta) \times a[t-1]$ 
11:     $m[t] \leftarrow \beta \times |data[t] - forecasts[t]| + (1 - \beta) \times m[t-1]$ 
12:     $\alpha[t] \leftarrow |a[t]/m[t]|$ 
13:     $u1[t] \leftarrow ((forecasts[t] - data[t])/data[t-1])^2$ 
14:     $u2[t] \leftarrow ((data[t] - data[t-1])/data[t-1])^2$ 
15:   end for
16:    $ustattrain \leftarrow \sqrt{\frac{\sum_{i=1}^{90} u1_i}{\sum_{i=1}^{90} u2_i}}$ 
17:    $ustatetest \leftarrow \sqrt{\frac{\sum_{i=91}^{111} u1_i}{\sum_{i=91}^{111} u2_i}}$ 
18:   return { forecasts, alpha, ustattrain, ustatetest, level, trend }
19: end function

```

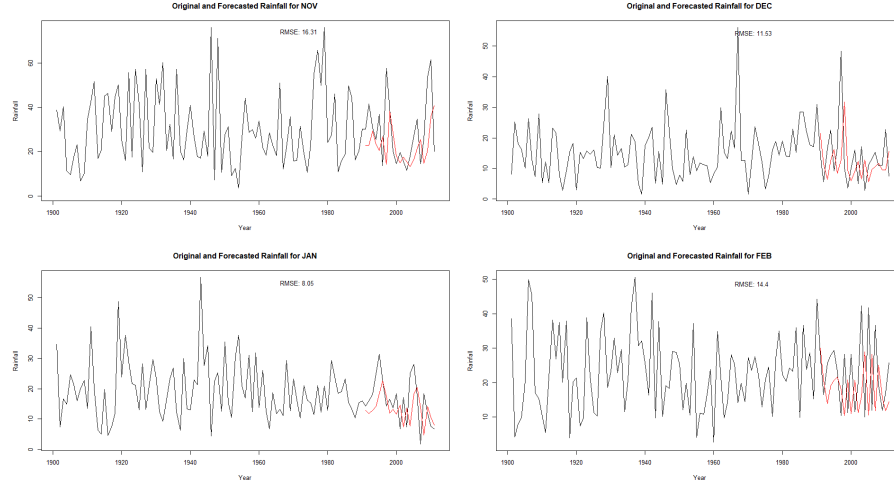


Figure 4: Adaptive method for part 1

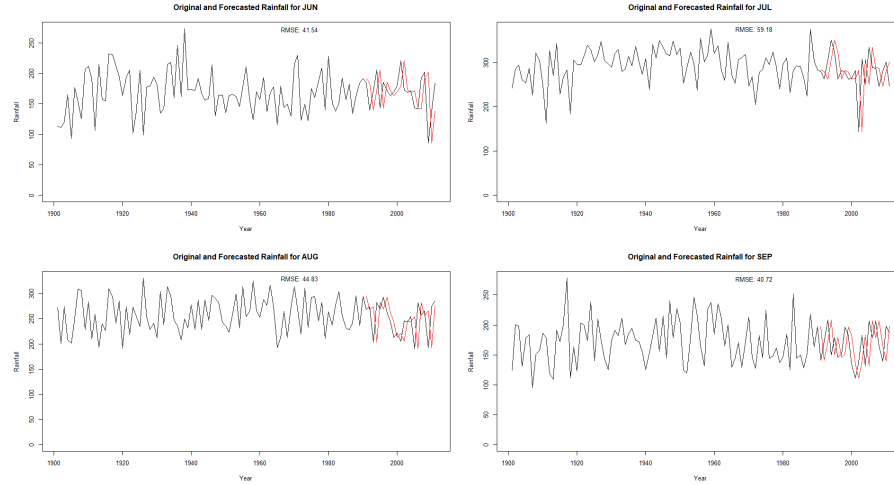


Figure 5: Adaptive method for part 2

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.