**Time Series Analysis**

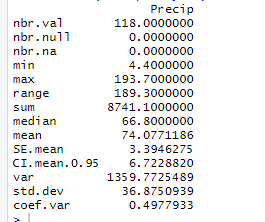
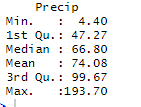
**Welland Rain**

**Data Transformation**

Data is transformed to the time series datatype

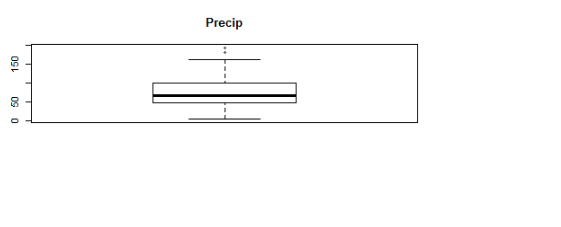
**Descriptive Data Analysis**

We look at the summary statistics



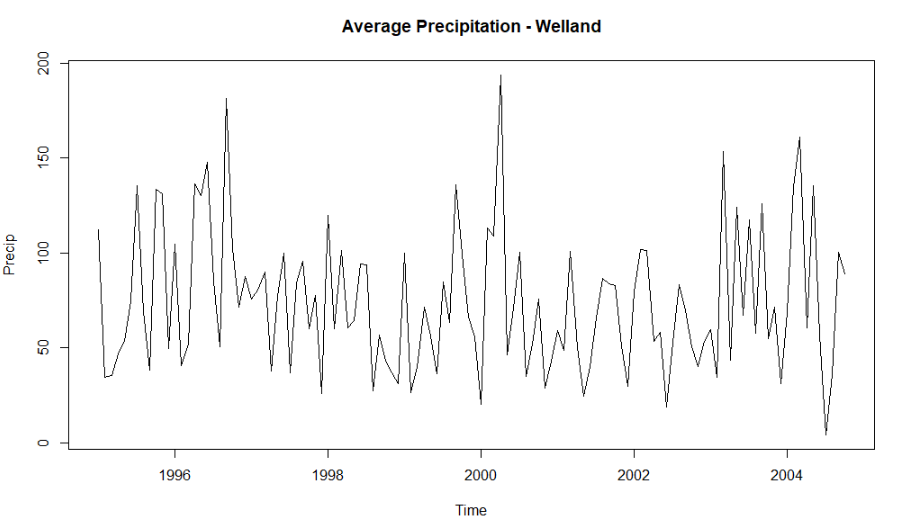
Data seems to be normally distributed. Average precipitation is 74.08, Minimum is 4.40 and Maximum is 193.70. There may be some outliers looking at the mean and max values.

Let’s see the boxplot to see if outliers are present.



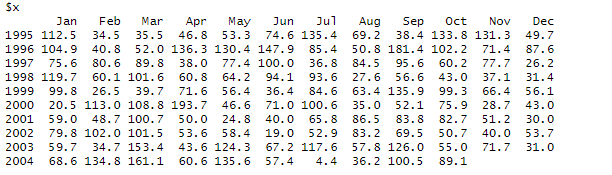
The boxplot confirms presence of few outliers. This could be some unsual real event rather than an error. However, this doesn’t seem to be an issue. We will leave them

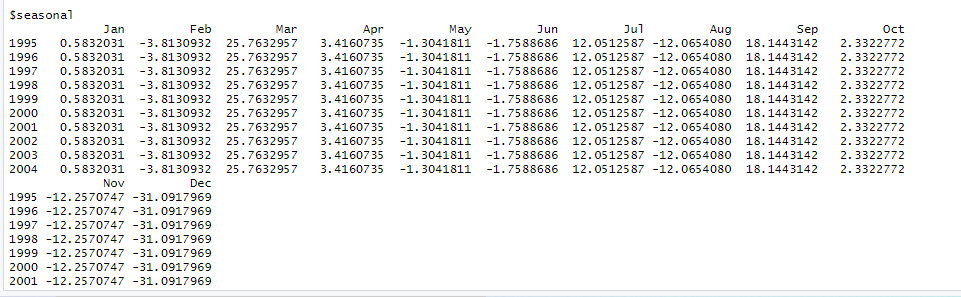
**Plotting the Time Series**

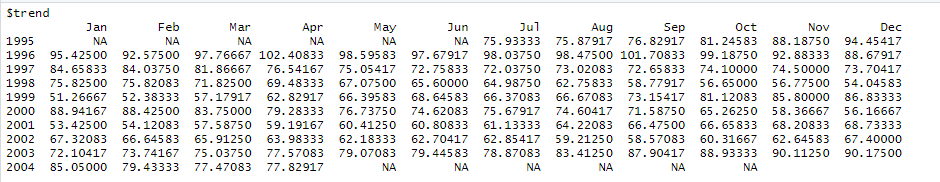


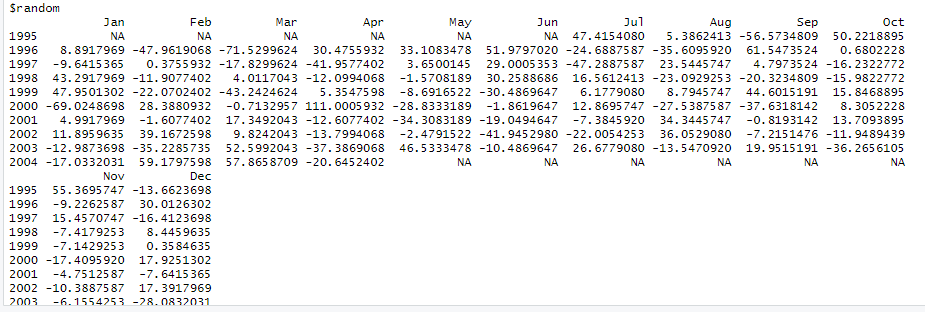
This is an additive model. The plot looks like a pattern and the range is fairly steady apart from the occasional spikes in precipitation around year 1997 and 2000 which is confirmed by the outlier above. There is also a time when the precipitation was too low around 2005.

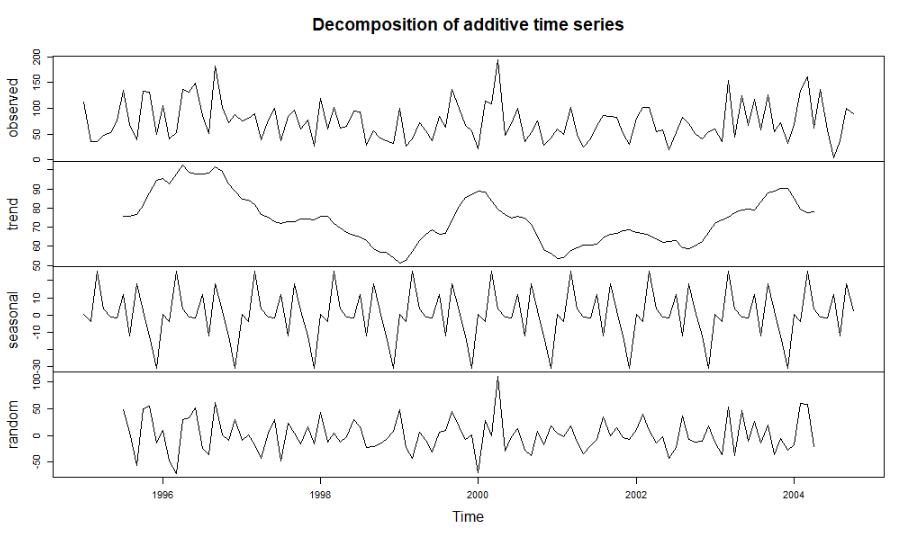
**Decomposing the Time Series**











**Observation**

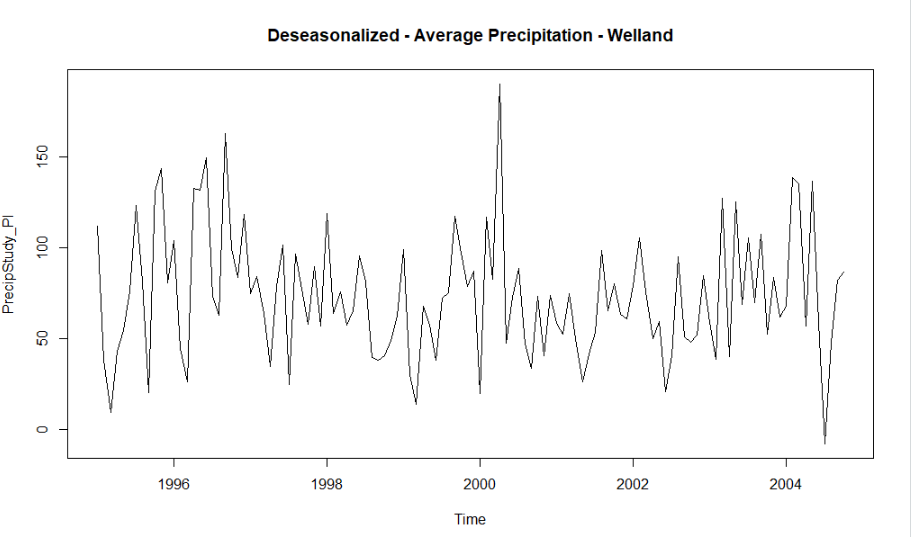
There seem to be no apparent trend as it looks reasonably constant over the years of the series

The residuals seem fairly constant in variability but shows a spike in variability in the middle years of the series.

The time series seem to be stationary. Looking at the seasonal, it seems that the statistical properties did not actually change over a time period. For example, looking at the plot, the mean and variance seem to be constant over time.

Secondly the p-value is 0.01 which is less than 0.05 indicating that it is stationary. This confirms the above.

**Deseasonalize the Precipitation Information**



Comparing the initial precipitation plot and the deseasonalized plot above, there seem to be no impact of season on the series. This shows that there is no season component as it looks the same with the normal average precipitation in Welland. Seasonality does not have any impact in the overall precipitation variation. This is understandable because there is precipitation all season whether it is winter, spring, summer and fall.

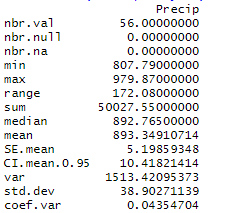
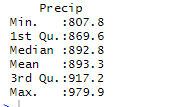
**Waterloo Precipitation**

**Data Transformation**

Data is transformed to the time series datatype

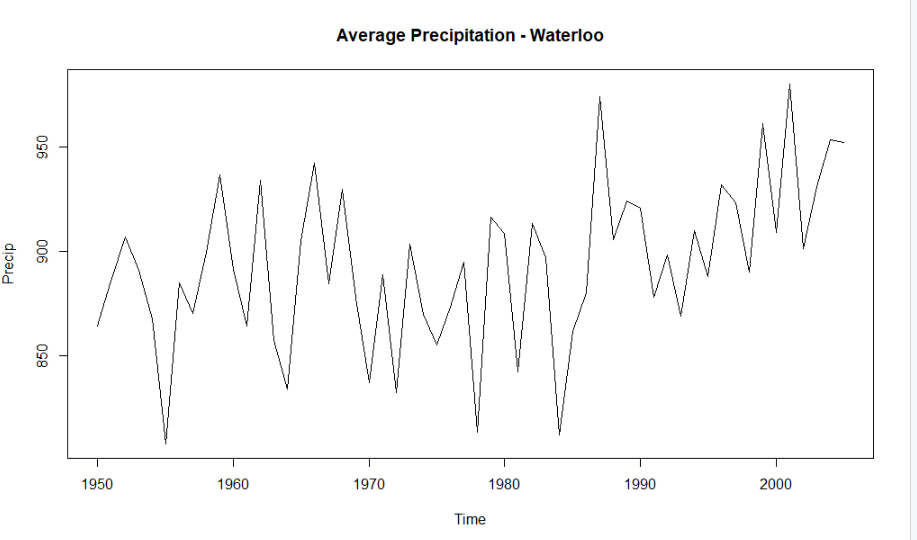
**Descriptive Data Analysis**

A look at the summary statistics



The distribution of the data seem to be balanced. The minimum precipitation is 807.8, mean is 893.3 while the maximum precipitation is 979.9. There don’t seem to be an outlier in the data

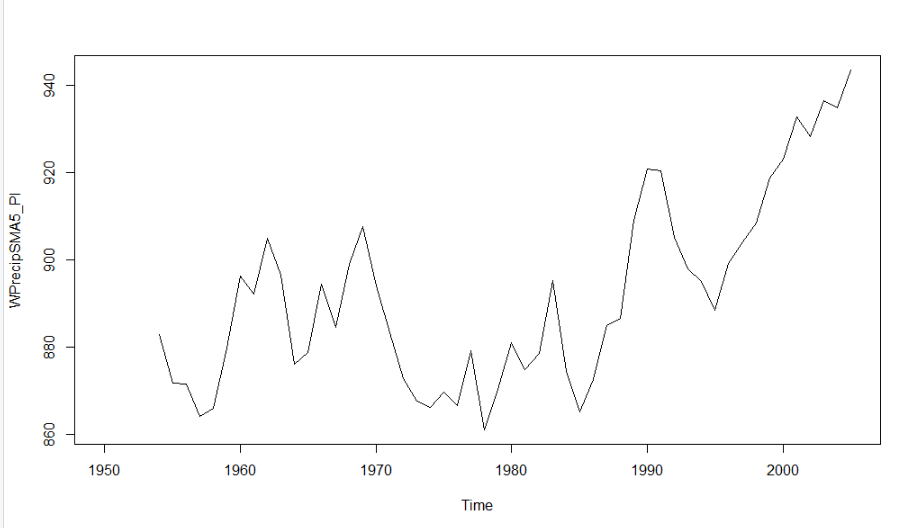
**Plotting the Time Series**



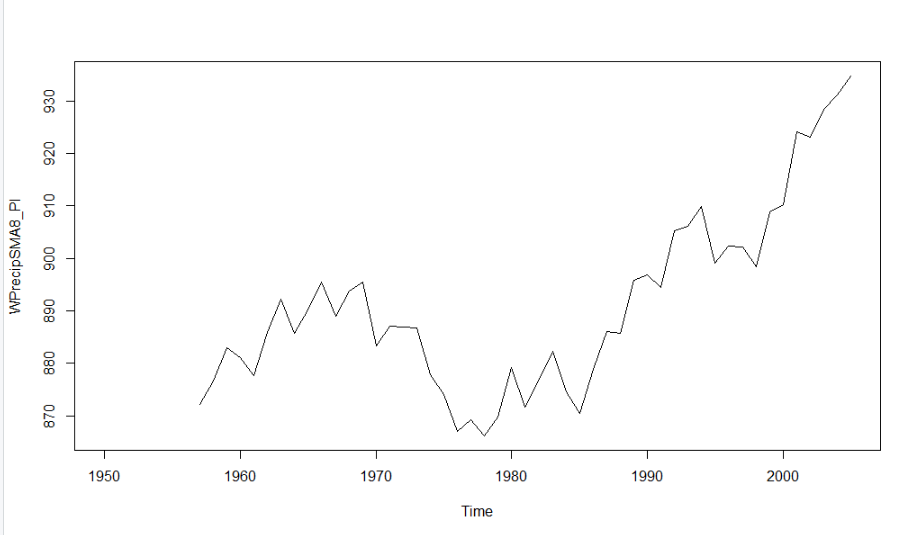
This seems to be an additive model showing some changing variability in the later years of the series. There seem to be trend forming from around 1993

**Smoothing the Precipitation Chart**

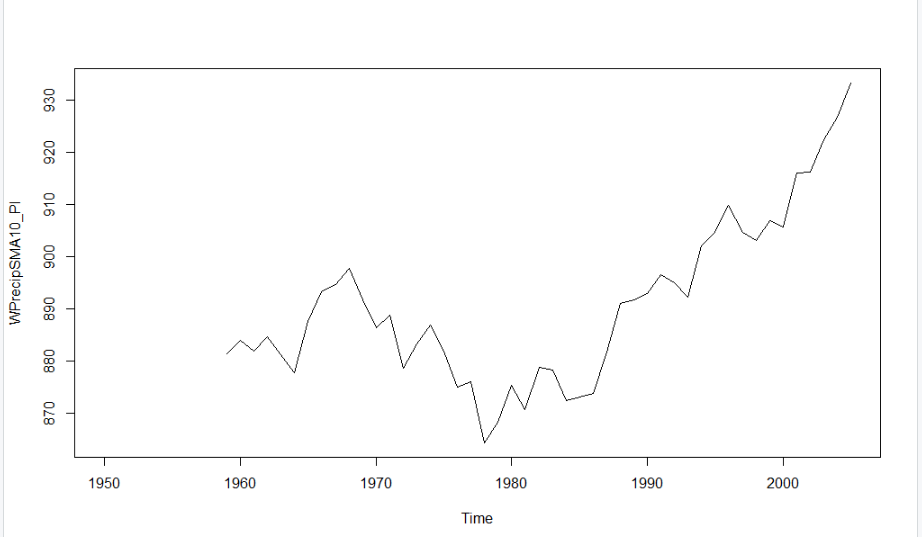
Let’s check for n= 5



Let’s see the chart for n= 8 as our moving average



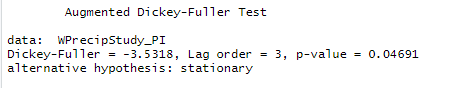
Let’s increase the moving average for n= 10 and see the trend



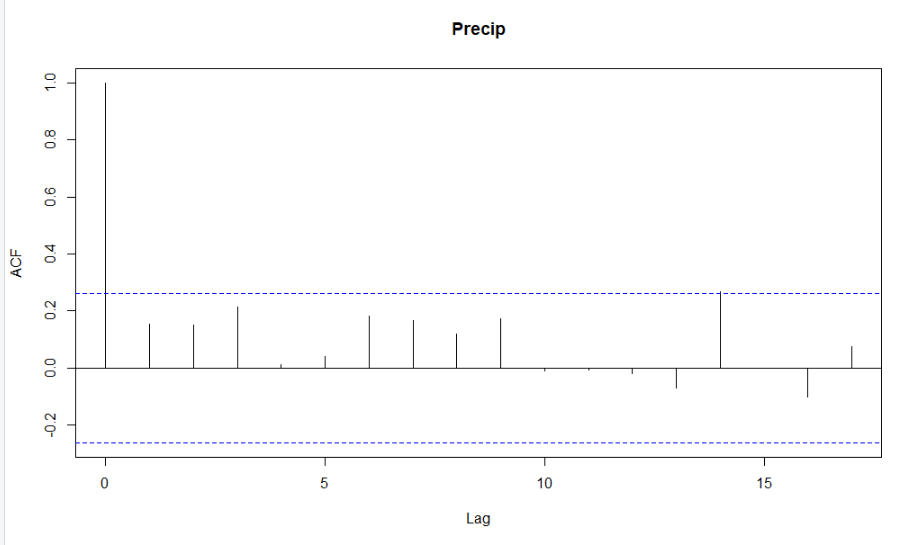
Looking at the charts for the 3 different moving averages, the third chart, n= 10 seems to better show the trend. It shows a downward trend towards the middle of the series and has been on the rise since then. We will work with this.

**Is the Time Series Stationary?**

Yes, it is. The p-value is 0.04 which is less than 0.05 suggesting that it is stationary

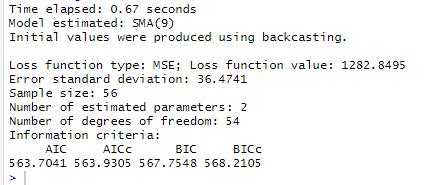


**Creating an Autocorrelation Chart**



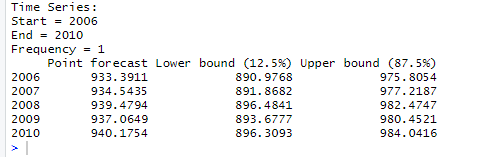
The lag at 0 is equal to 1 which is understandable as it shows a time series that is correlated with itself. However, none of the other lags are above the dashed lines showing that the lags are not statistically significant. This indicates that the precipitations are not correlated across the series. In addition, the plot does not suggest any hint of seasonality.

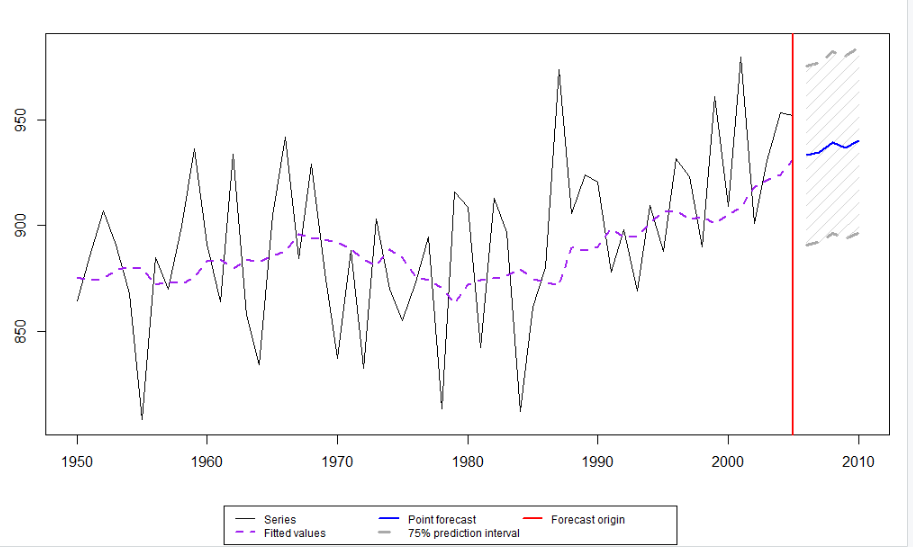
**Creating a Simple Moving Average Forecast**



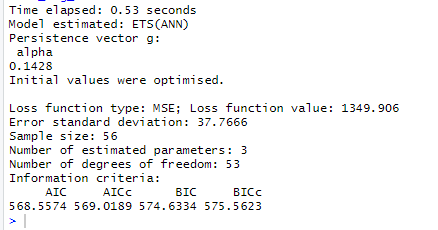
From the result above, the model seems to choose smoothing the series at 9 years moving average as the best fit for this model because this is where it found the lowest MSE. This is also the model with the best AIC as the SMA will choose models with the best AIC

Let’s do the forecast for 5 years and 75% prediction interval



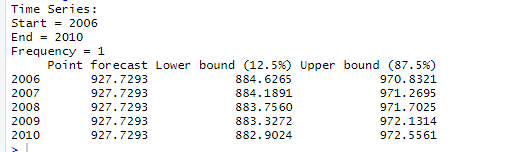


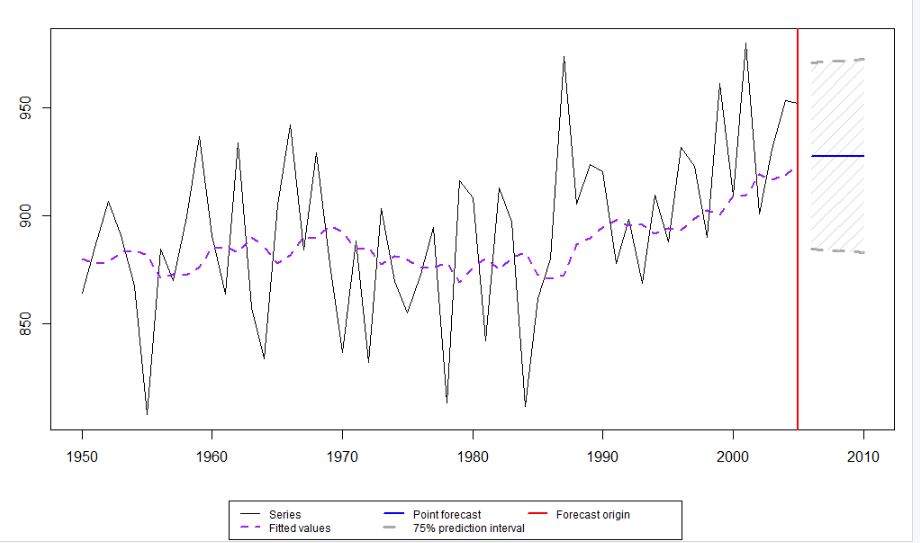
**Creating an Exponentially Smoothed Forecast**



This is the model with the best AIC as chosen by the ES function

Now, Let’s do the forecast for 5 years and 75% prediction interval





**Which Forecast is Better?**

Comparing the two forecasts above, the Simple Moving Average Model forecast is better for the following reasons:

1. Has an MSE value of approximately 1282 is lower when compared to the Exponential Smoothed forecast with MSE of 1349. This means that it fits better with the data we have seen historically. The lower the value of MSE, the better the forecast
2. It has a lower AIC of approximately 563 as against 568 for Exponential Smoothed forecast model. This means that it is a slightly better fit for the historical data. The lower the AIC, the better the forecast.
3. Looking at it visually, the smoothed moving average seems to make more sense and more reflective of the trend.