University of Potsdam

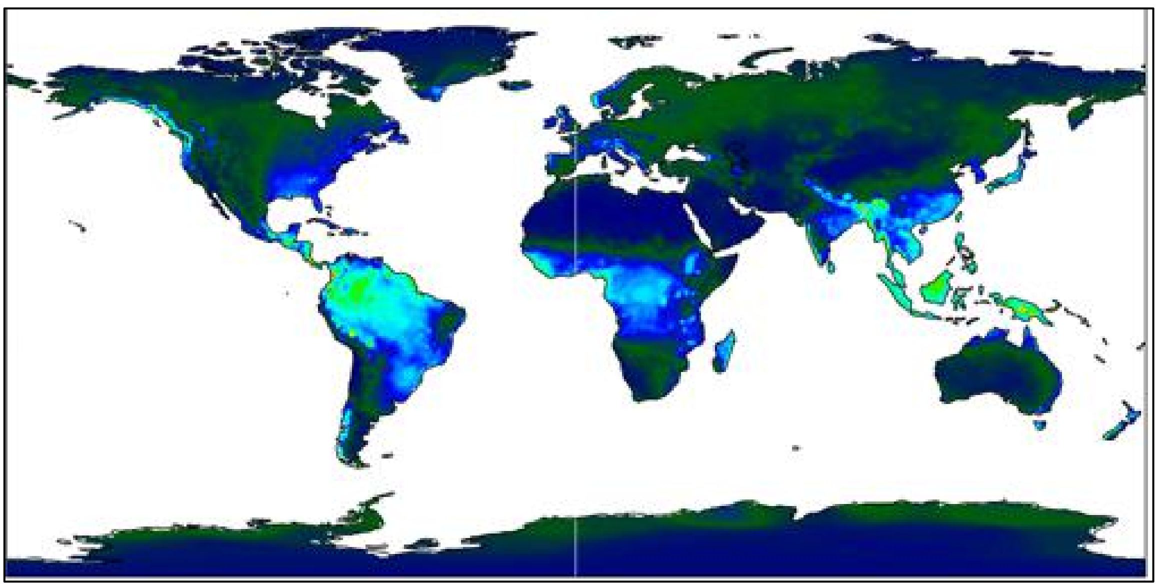
Faculty of science

Departement of Earth Science

GEW-DAP03 Big Data Analytics

RELATIONSHIP BETWEEN TERRESTRIAL AIR TEMPERATURE AND PRECIPETATION

Ifeany OKARO, Farid ATMANI



Academic year 2018/2019

Table of contents

1. Introduction

There have been many discussions about how much effect temperature has on precipitation. Some scientist believe its only one of several factors and has just a little effect while others believe it is the major driver of global precipitation and thus can be used to estimate the amount of precipitation in a place. This project is focused on exploring Terrestrial Air Temperature and Precipitation from1900-2017 Gridded Monthly Time Series obtained from metrology stations Big Data and Spatial analysis approaches will be used to explore the data on Python with then aim of finding trends in the data as well as possible associations or relationships between these two variables

2. Objectives

- To find the trends and variability in Temperaure dataover time
- To find the trends and variability in Precipitation data over time
- To find the trends and variability in Precipitation data over time
- To explore other anomalies that could occur in the time series data

3. Data Source

Terrestrial Air Temperature and Precipitation Data was downloaded from National Oceanic and Atmospheric Administration. This came in two HDF5 files which were both imported in python for analysis. The precipitation data is monthly total in cm, while the temperature data represents the monthly mean temperature for the 118 years period

Data 1

Source: www.esrl.noaa.gov

File format HDF5

Variables

Dataset_title: Terrestrial Air Temperature: 1900-2017 Gridded Monthly Time Series

Dimensions(sizes): lat(360), lon(720), time(1416)

Data 2

Source: www.esrl.noaa.gov

File format HDF5

Variables

Dataset_title: Terrestrial Precipitation: 1900-2017 Gridded Monthly Time Series

Dimensions(sizes): lat(360), lon(720), time(1416)

4. Methods

Spatial Analysis

To display the temperature and precipitation data on map, for the entire globe, the mean of all time (1900 -2017) was calculated on the time axis The resultant variable was projected on map using Cartopy

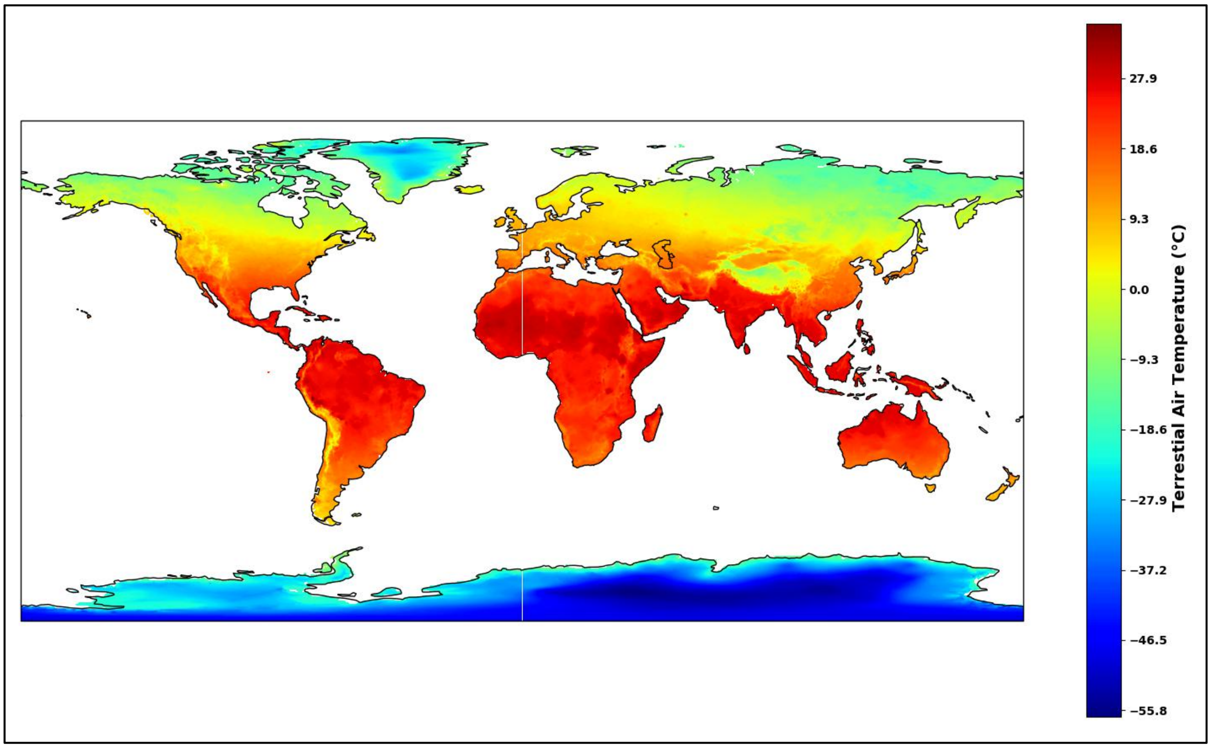


Figure 01: Global Mean Temperture Map (1900 - 2017)

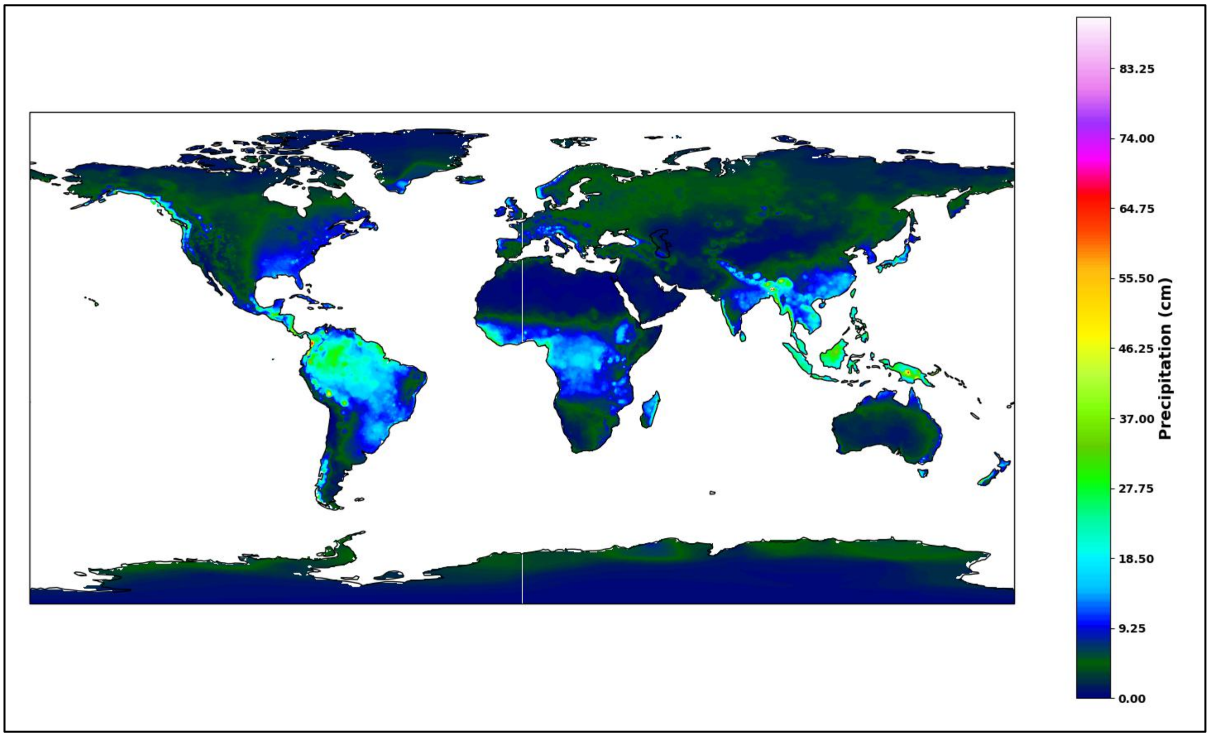


Figure 02: Global Mean Monthly PrecipitationMap (1900 – 2017)

Looking at both maps we noticed the area of high temperature and precipitation are located around the Tropics. This was an indication that there could be some relationship hence prompting more analysis on the data. To better visualize this observation, seasonality approach was adopted. The summer and winter months were indexed out and the mean values plotted on a map for better comparism.

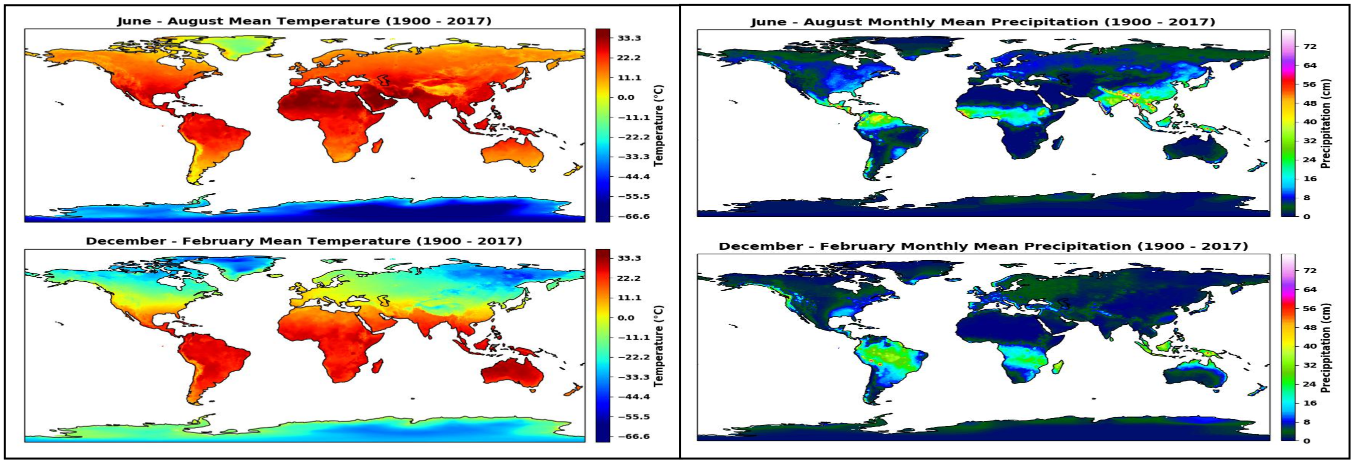


Figure 03: Seasonal Mean Comparism

From the maps, we noticed thr high precipitation areas changed with the seasons. The June to August mean showed high temperature and precipitation areas in the Northern hemisphere while the December to February mean showed this trend shift towards the southern hemisphere, however in all of these, the peak areas still revolves around the tropics and sub-tropics. We also looked at the standard deviation which showed much instability in temperature around the Arctic and Subpolarr regions, with precipitation varying mostly in the tropics and sub-tropics

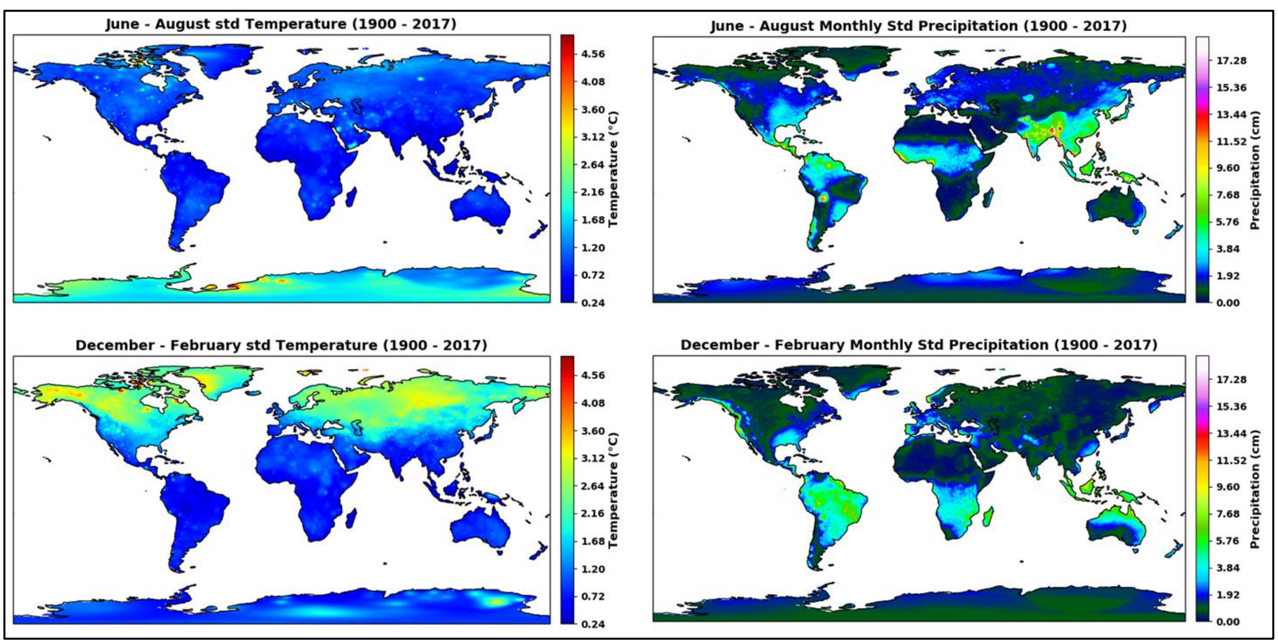


Figure 04: Seasonal Standard Deviation Comparism

5. Time Series Analysis

With our spatial analysis we were able to detect seasonal trends and variations in temperature and precipitation it is will be nice to look at the data on a time domain to explicitly identify what these trends signify. We plotted our whole data against time and it displayed a really weird result.

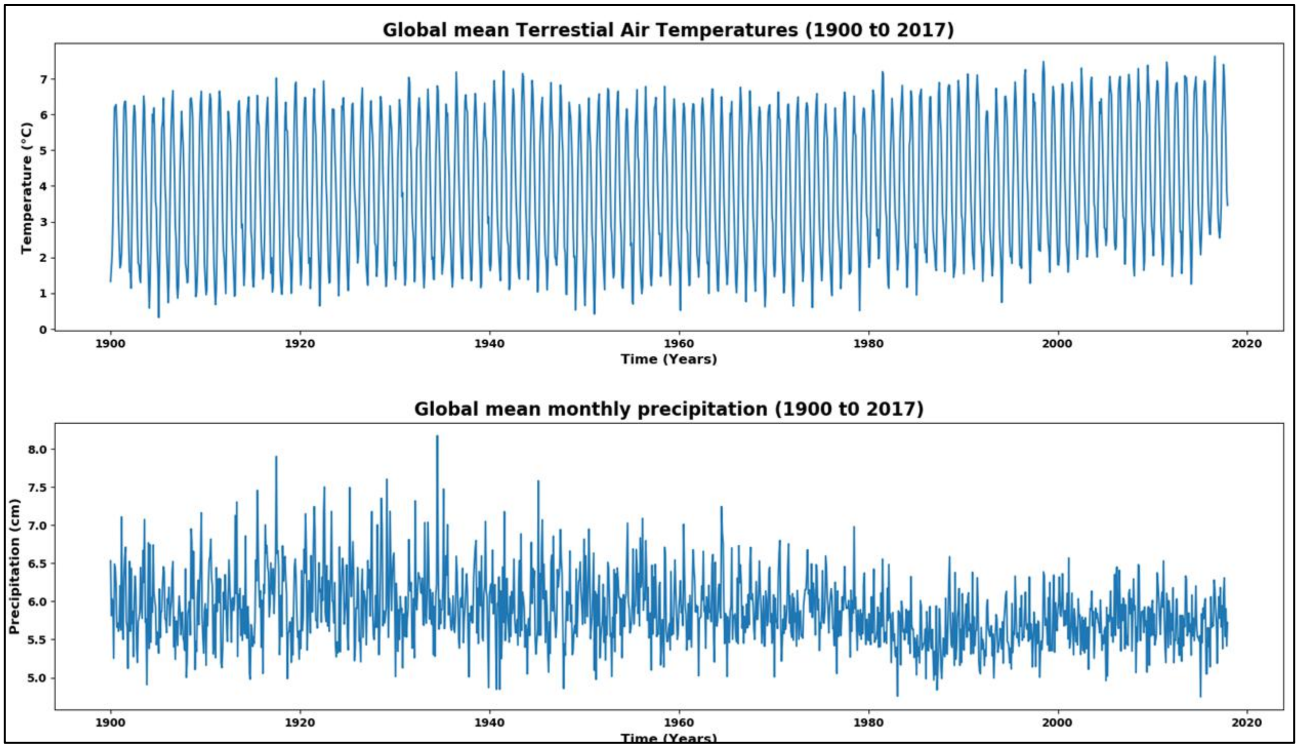


Figure 05: TIME SERIES OF THE TEMPERATURE AND PRECIPETATION

Our data showed so much noise and seasonality, which makes it difficult for any reasonable interpretation. Summer and Winter months were obtined by using double colon indexing (::) , for every year. This was plotted against the corresponding indexed time.

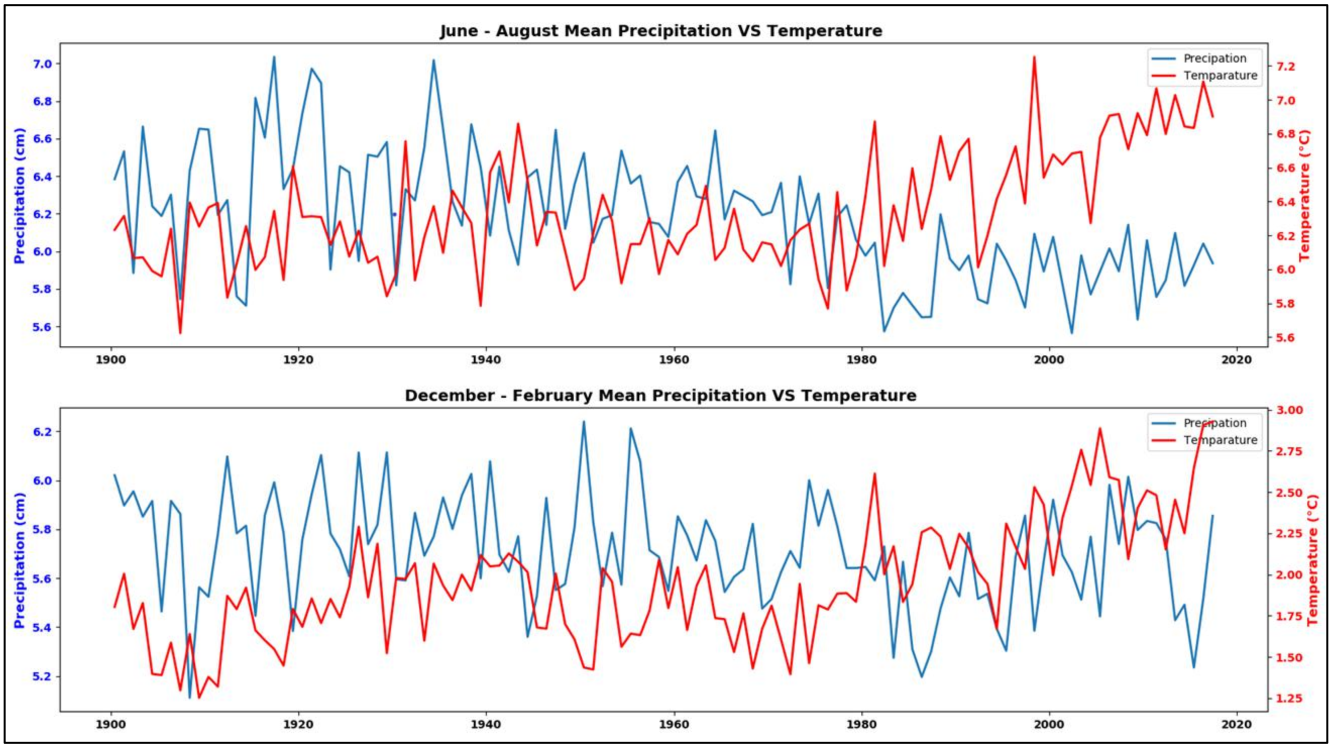


Figure 06: SEASONAL TIME SERIES OF THE TEMPERATURE AND PRECIPETATION

The seasonal plot showed some upward rising trend in the June to August temperature and a likely downward movement of the precipitation curve but the noise in the data prevented accurate interpretation. Savitzky_Golay filter model was applied to smoothen the data for better interpretation.

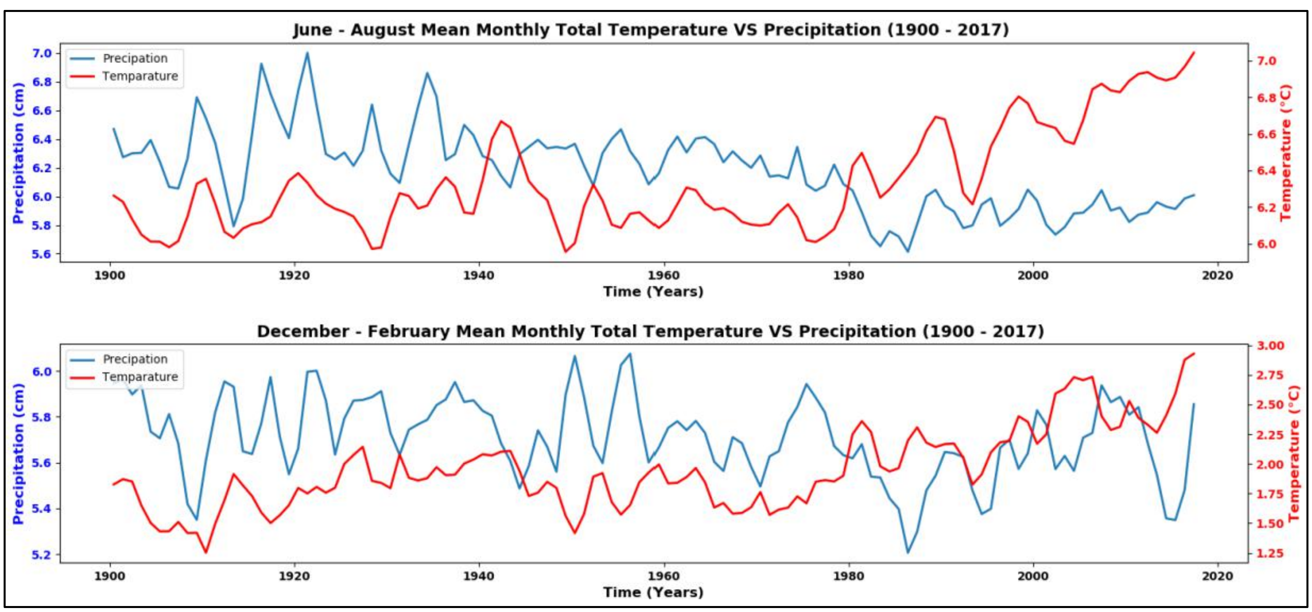


Figure 07: FILTRERED SEASONAL TIME SERIES OF THE TEMPERATURE AND PRECIPITATION

The filtered data showed a clearer picture of the trend. The temperature persistently rose from midway into the time series especially in the June to August months. The precipitation showed an opposite movement indicating possibility of a negative correlation. We decided to divide the time series of the summer months into two halves to focus more on the recent years. This is because we have a very long time series which may be affected by many factors such as accurate records, improved instruments etc.

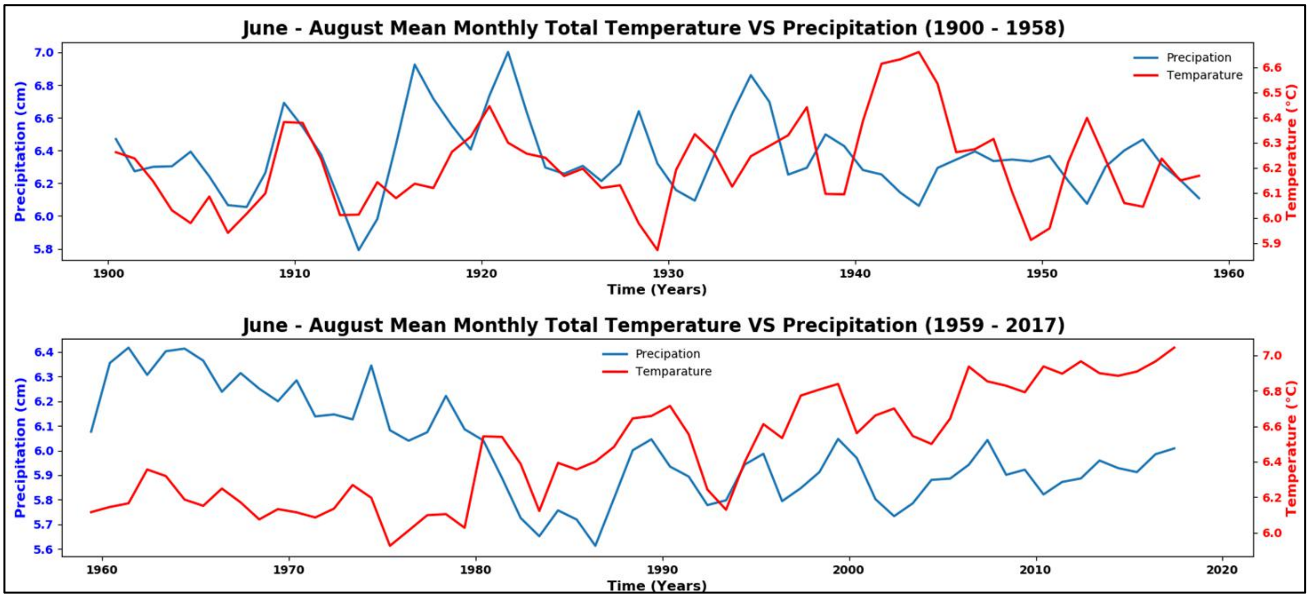


Figure 08: OLDER AND RECENT YEARS TIME SERIES OF THE TEMPERATURE AND PRECIPETATION

The 1959 to 2017 data showed what looks like a negative correlation between temperature and precipitation. To verify this we did a scatter plot which is usually the best way to visualize correlation. The scatter plot showed no corresponding pattern which was a surprising discovery in our analysis. To verify this, we used spearmans rank correlation coefficient which explores statistical dependence between the rankings of two variables. We also explored this correlation independently for these variables again time

.

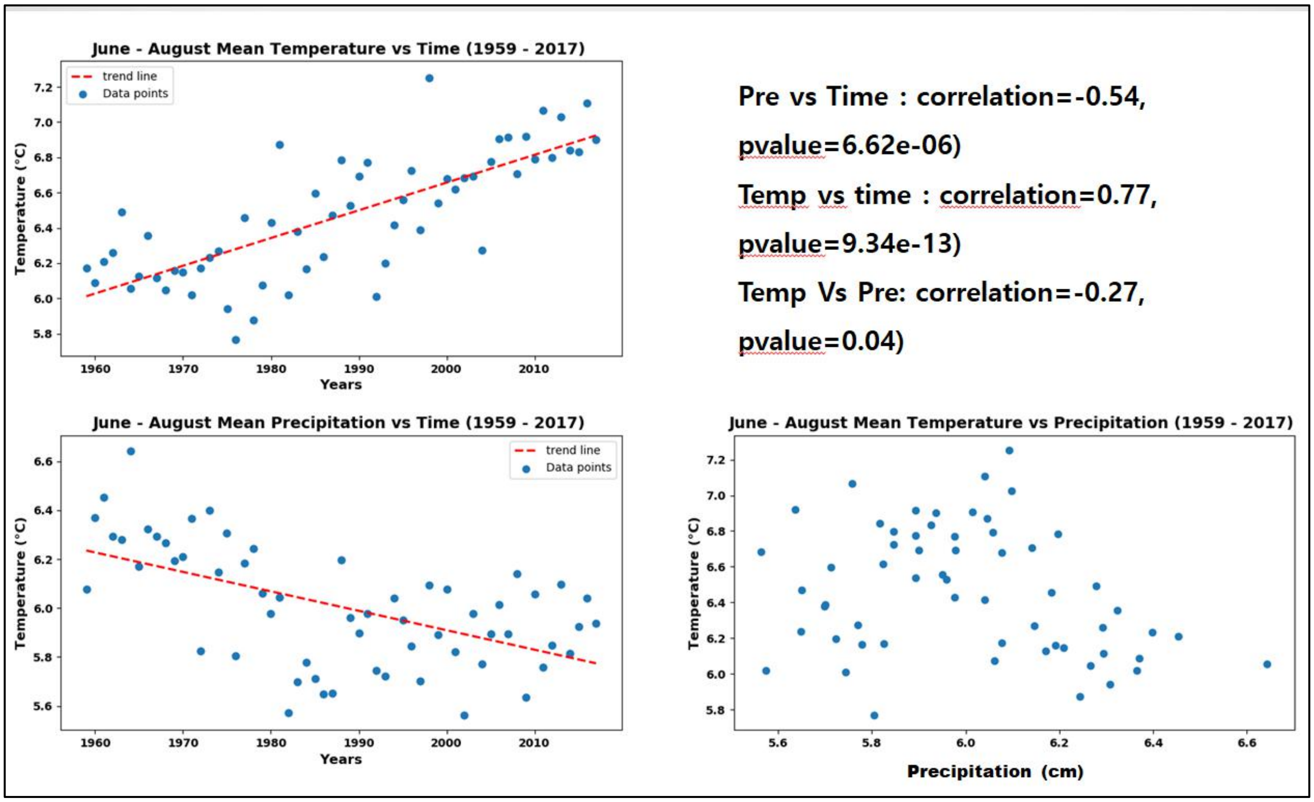


Figure 09: SCATTER PLOTS OF TEMPERATURE, PRECIPITATION AND TIME

The variables precipitation and temperature had some reasonable correlation with time but do not have with each other. As correlation coefficient goes fro -1 to 1 with values closer to the extreme representing higher correlations. The value of -0.27 obtained when correlating temperature with prepitation is much closer to 0 hence indicating no correlation. P values were used to test the statistical significance of our correlation which in this case is the probability of unrelated variables producing the same correlation , with low P values ($P < 0.05$), we can say that accept the results of our correlation. Considering the seasonality differences between Northern and Southern Hemisphere, we divided our data along these axis and tried to correlate only the Northerm hemisphere data but this still showed similar result, however we noticed unusual increase in the temperature range from 5.8 to 7.2 °C when considering the whole globe to 17 to 18 °C when considering only the northern hemisphere.
This prompted us to conduct a more localized correlation

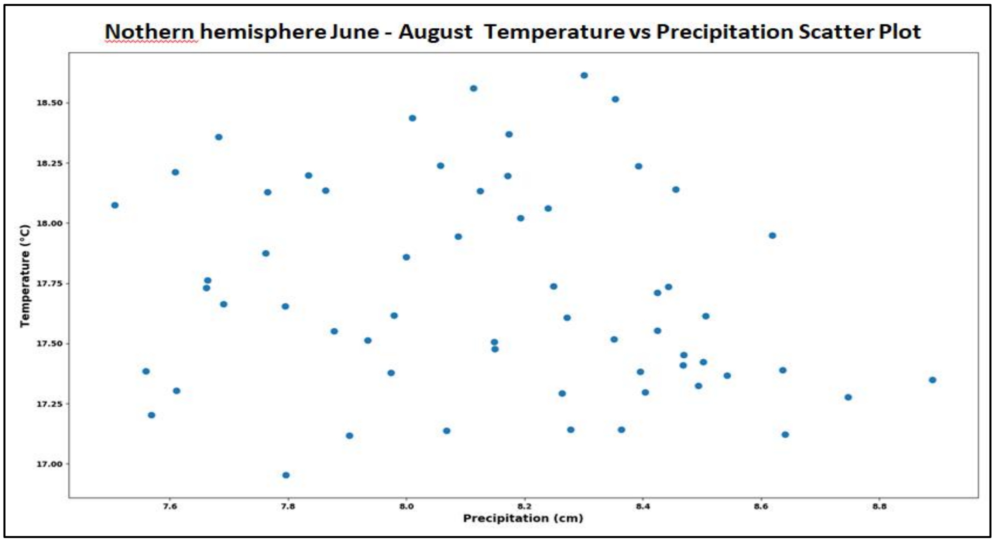


Figure 10: SCATTER PLOT OF TEMPERATURE AND PRECIPITATION

To better understand this trend. The monthly data was converted into yearly data and the time series plotted. We fitted a sixth order polyline to the plot which revealed the trend line of increasing temperature.

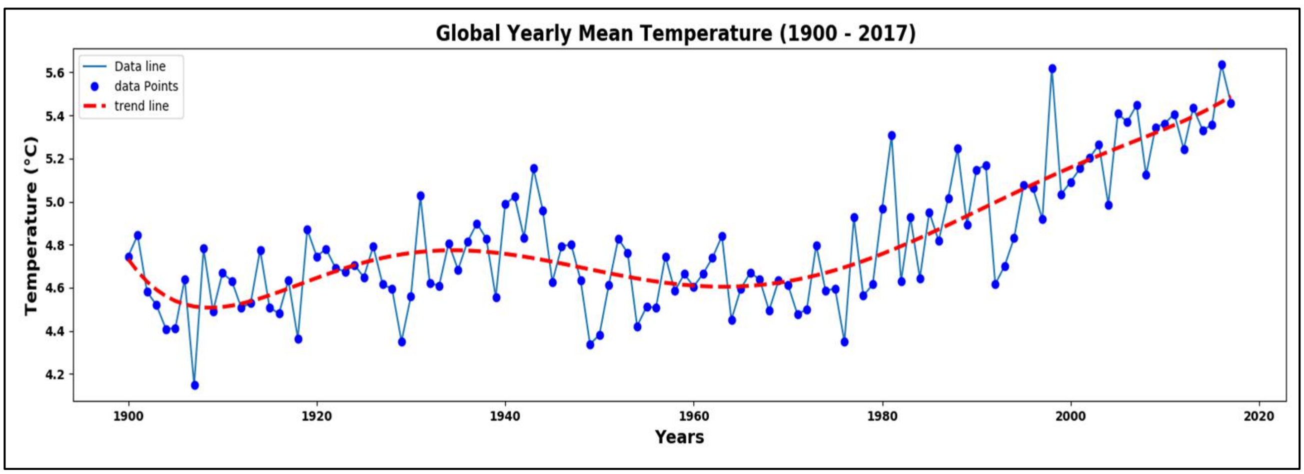


Figure 11: TREND FITTING TO THE GLOBAL YEARLY MEAN TEMPERATURE

The temperature data was further divided into Northern and Southern hemisphere, this showed similar trends but much faster yearly temperature increase in the Northern hemisphere. This observation further emphasized the need for geographic slicing to a localized study area with both nothern and southern hemisphere features.

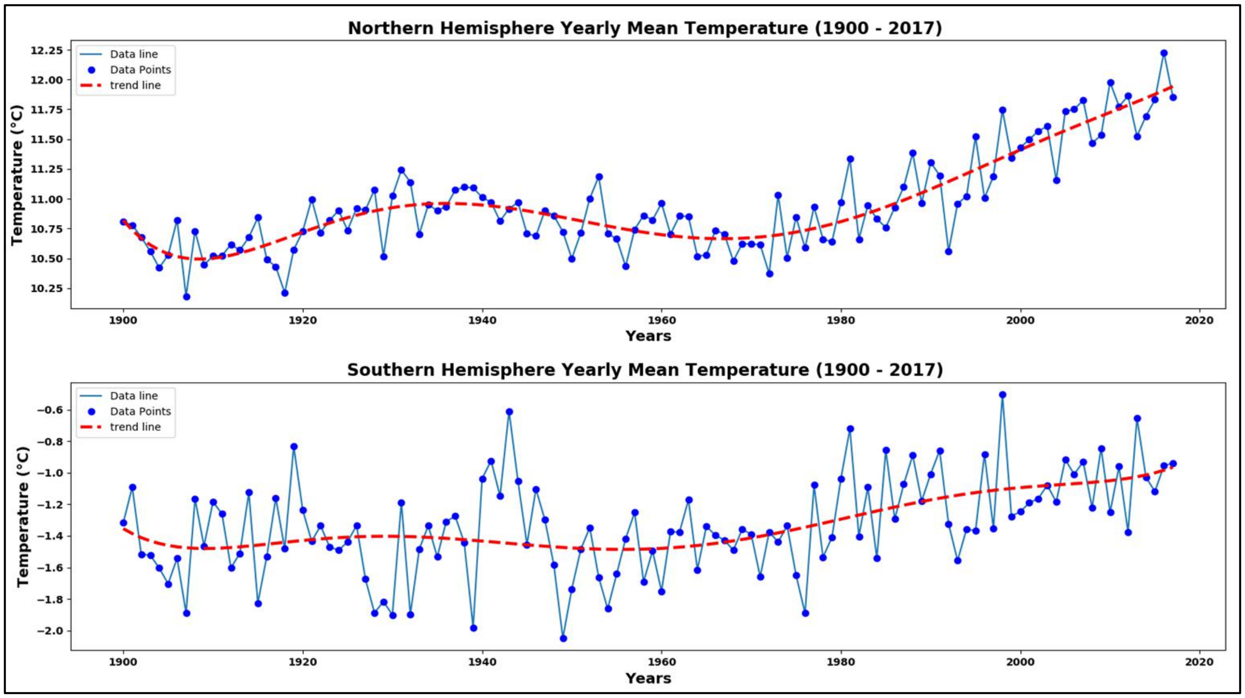


Figure 12: TREND FITTING TO THE NORTHERN AND SOUTHERN HEMISPHERE YEARLY MEAN TEMERATURE

6. Geographic Slicing of India

India being one of the places with the highest amount of rainfalls in the world was used for localized correlation. This region also cuts across the northern and southern hemisphere being a perfect sample for all seasons. The coordinates of india was converted to image pixels of our data and used to slice out the region.

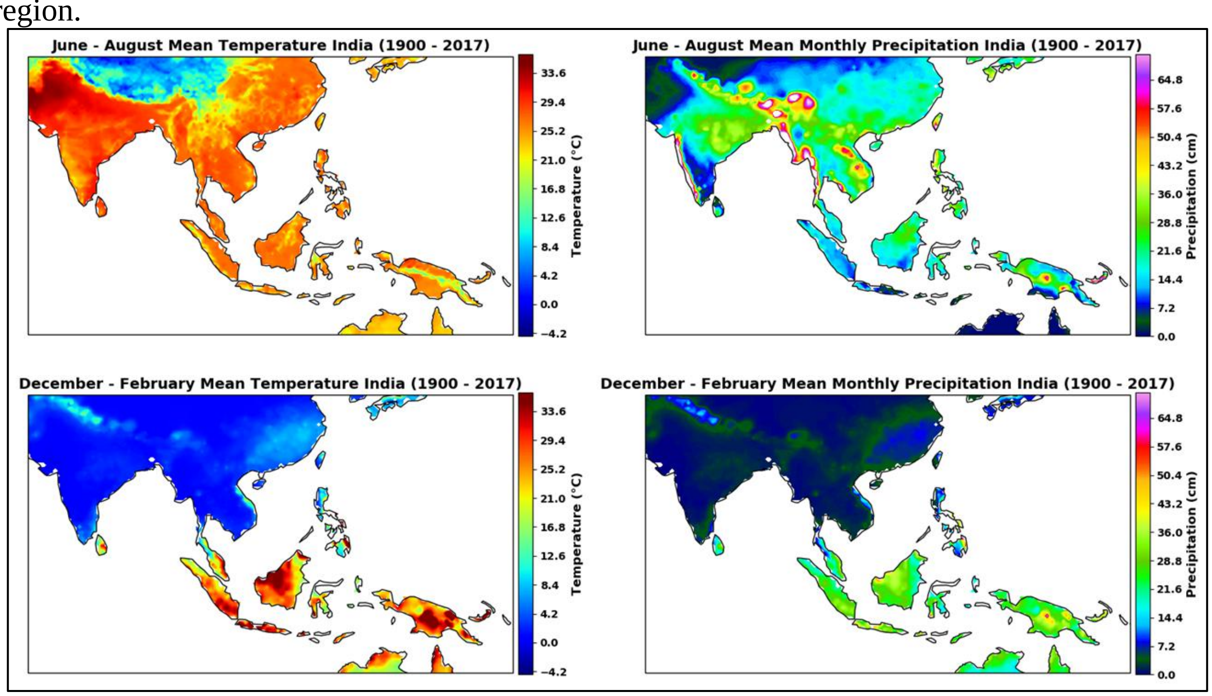


Figure 13: TEMPERATURE AND PRECIPITATION DURING SUMMER AND WINTER IN INDIA

The map shows a trend of high temperature and high precipitation in june – August at Northern hemisphere part of the area and same pattern in the summer months (December – February) of the southern hemisphere part. This indicated possibility of localized correlation. A scatter plot was used to verify this. The plot showed a pattern which sees precipitation rise at a certain temperature

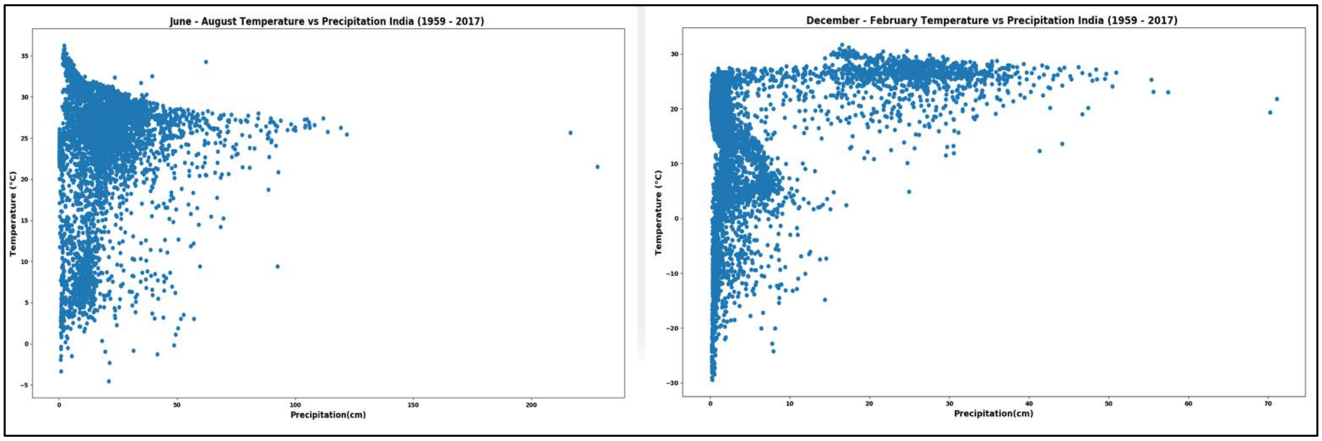


Figure 14: SCATTER PLOTS OF TEMERATURE AND PRECIPITATION DURING SUMMER AND WINTER IN INDIA

7. DISCUSSION

The precipitation and temperature data revealed some useful information. Precipitation is highest along the tropical and subtropical regions There is high seasonality variations between the northern and southern hemisphere, the winter months in Northern hemisphere corresponds to summer months in southern hemisphere. Another trend is the amplified warming of the northern hemisphere. The temperature of the Northern hemisphere have increased by approximately 1° C since 1900 while the southern hemisphere temperature has barely increased by 0.2 °C over the same period. The temperature and precipitation data does not correlate on a global scale but only have some association on a local scale because a lot of factors are also responsible for precipitation. For instance , on a global scale the temperature of Antarctica is considered which is a large area of cold surface. Also ocean currents play a vital role in precipitation. There are also other factors that influence precipitation such as atmospheric pressure and humidity.

8. LIMITATIONS

Though we were able to discover some relationships and anomalies. There were some limitations to this research which include;

- Limited time for such volume of project
- Atmospheric pressure Data was not considered
- Sea surface temperature was not considered
- Humidity data was not used

9. CONCLUSION

BigData analytics approach proved successful in exploring and finding the trends as well as relationships in the temperature and precipitation data. From our research we discovered an association between temperature and precipitation on a local scale. However, further research is required to understand this association and explore other possible factors affecting it.

10. REFERENCES

- Center for Climatic Research Department of Geography University of Delaware Newark, DE 19716 The University of Delaware Web site offers extensive documentation of this data set. Please address questions on the analysis method to Kenji Matsuura (kenjisan@udel.edu) University of Delaware**.**
  https://www.esrl.noaa.gov/psd/data/gridded/data.UDel_AirT_Precip.html