Predicting Droughts in the Amazon Basin based on Global Sea Surface Temperatures

Dario Lepke

Introduction

Placeholder

Related work

Placeholder

EDA

EDA precipitation

In this section we want study the time series of precipitation in the Central Amazon Basin. The CHIRPS data set contains the **precipitation data**, created from in-situ and satellite measurements (Funk et al. (2015)). It can be downloaded for example from here [https://www.chc.ucsb.edu/data/chirps]. It contains observations **from 1981 to 2016**. Data comes on a **high resolution of 0.05 grid**.

Below the area of the Central Amazon Basin that is object of our study.

We will now inspect the precipitation values from three perspectives. Firstly the raw values without time or spatial dependency, second the mean and standard deviation for each spatial grid cell for the whole time series and then the mean and standard deviation for each grid cell but for each month in year separately.

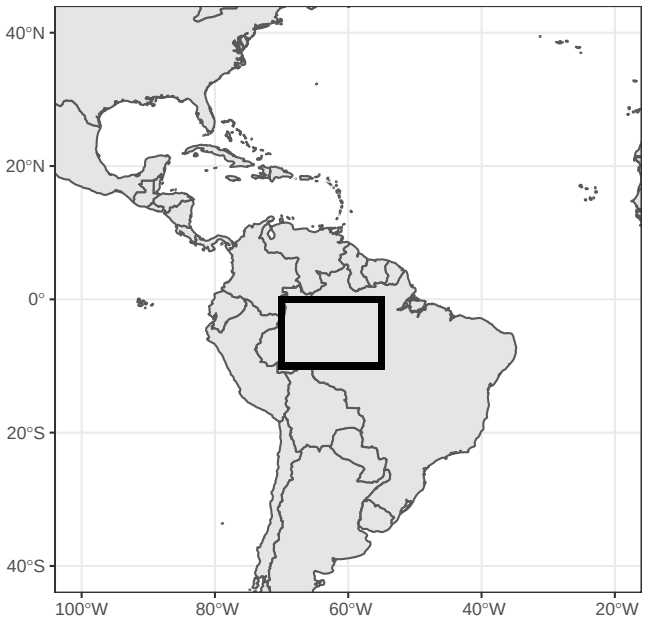
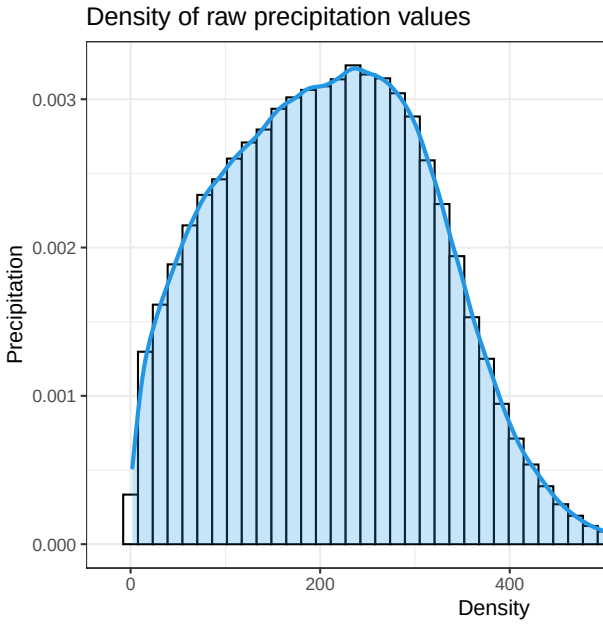


Figure 1: Localisation of the area under study. The central amazon basin (CAB) spanning across 0,-10 latitude and -70,-55 longitude

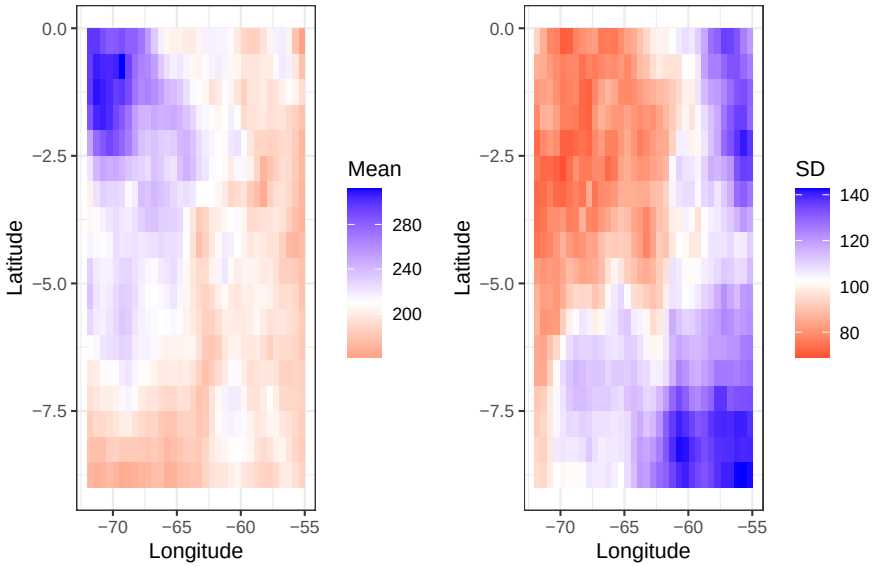


Firstly we inspect the precipitation values in general
Its form is a uni-modal, right-skewed density. The values range from 0 up to

`max(precip_dens_plot$data$value)`, but only few observations take these high values, forming a large tail.

Mean and SD at each location

Mean and SD of each location in the CAB for the complete time series

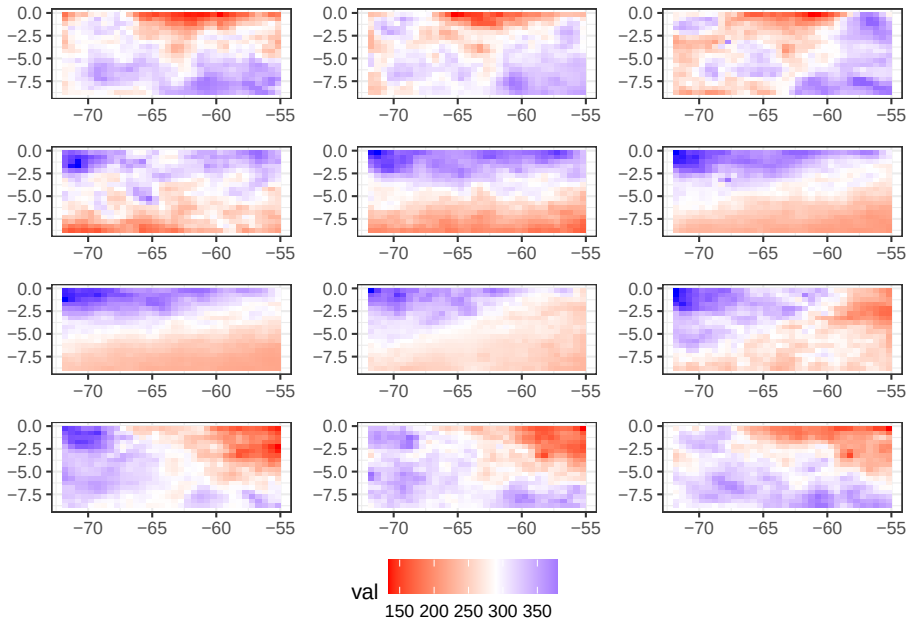


As we can see most locations have a mean precipitation of around 200 mm/month, over the whole time series. Regionally in the "upper left" corner of the Amazon Basin, mean precipitation is higher or equal to the mean. The reference point for "higher" is the mean of the location means. This region seems to be more or less spatially consistent. The rest of the region with lower mean precipitation has also some small areas where precipitation is again a little bit higher. For example in the upper right corner and on the bottom, right of the middle.

For the standard deviation we also see regional patterns. These patterns overlap with the regions of the mean but their magnitude is flipped. Meaning, in the upper left where we observe larger mean values we generally observe lower standard deviation and in the lower and upper right corners, higher standard deviations.

Question: There are obviously regional differences in magnitude of mean and standard deviations. Should we therefore NOT normalise the time series, prior to clustering? Mean and SD contain information and the variables are all measured on the same scale. (Consideration: seasonality might play a role, meaning that the values of different months come a different distribution)

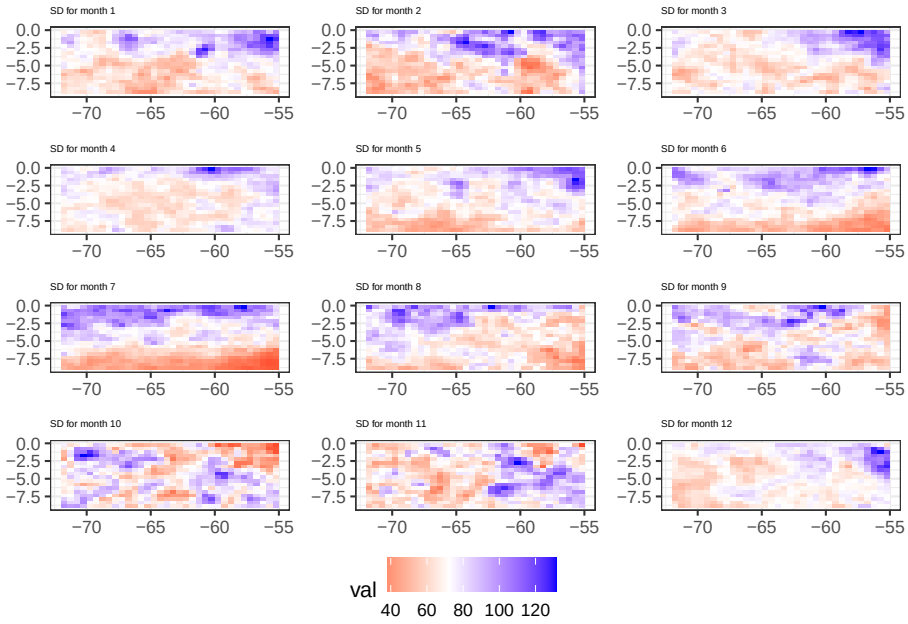
Means per month




```
## list()
```

We see spatial patterns of the mean evolving over time. For example: From May until August there is a spatial separation in two parts that dissolves in september. As expected there is a large seasonal component regarding the means.

SD per month



For the standard deviation we see as well large differences in values during different months of the year.

EDA SST

Mean at each location

SD at each location

Mean and SD at each location

Trend at each location

Means per month TEST

SD per month TEST

Trend per month TEST

Funk, Chris, Pete Peterson, Martin Landsfeld, Diego Pedreros, James Verdin, Shraddhanand Shukla, Gregory Husak, et al. 2015. "The Climate Hazards Infrared Precipitation with Stations-a New Environmental Record for Monitoring Extremes." *Scientific Data* 2 (1): 1–21.