Accessing Global Precipitation Data and Analyzing Total Precipitation Trends over 40 years

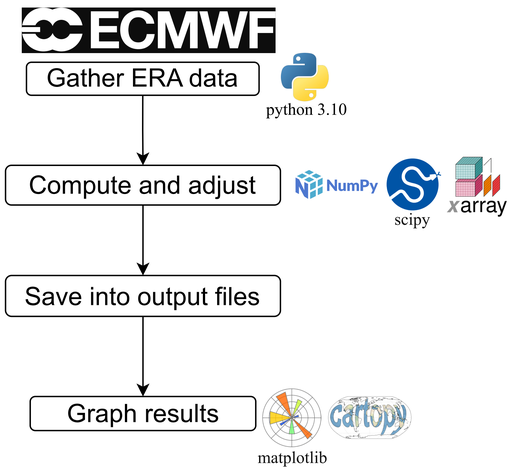
Leo Ling[1], Josiah Harrison[2], Eugene Clothiaux[2]
[1]State College Area High School, [2]Pennsylvania State University
Corresponding Author: Leo Ling (lxl24@scasd.org)

Wet places wetter, dry places dryer.

Abstract
Precipitation is an integral part of the worlds climate. One state-of-the-art extensive dataset is the European Centre for Medium-Range Weather Forecasting (ECMWF) Interim Re-analysis (ERA). Using linear regression, we characterize the precipitation trends globally over 34 years. We find that the wet places are getting wetter and the dry places are getting drier.

Introduction
Over recent years, many places have experienced severe droughts and torrential rains. This motivated our study to characterize the trends in precipitation over the past 34 years. We use the ERA for this purpose. The ERA spans 40 years but only 1984-2017, i.e., 34 years, were used in our analysis. Precipitation was accumulated every 3 hours over each day. To analyze these data, we retrieve it in the netCDF file format, which is the standard for multidimensional data in the atmospheric sciences. We then apply linear regression to each location to ascertain how the precipitation has been changing globally.

Methods
We use python since it has data analysis and graphing libraries. From python, matplotlib and cartopy are the libraries best suited for our purpose. We need xarray to handle reading and writing netCDF files. numpy is used for speeding up calculations. The last library we use is scipy for linear regression. Below is a flowchart showing how these libraries work together.



Implementation
First, we make a result array organized by year, longitude, and latitude. For each year, we get all of the files in the year and then open them as one dataset. We then unpack the values with xarray. We compute the sum of each year and then set the corresponding value in the result array. Next, we create a slope and intercept array organized by longitude and latitude. For each longitude and latitude, we take the 34 years, linearly regress them, and store them in their respective output arrays.

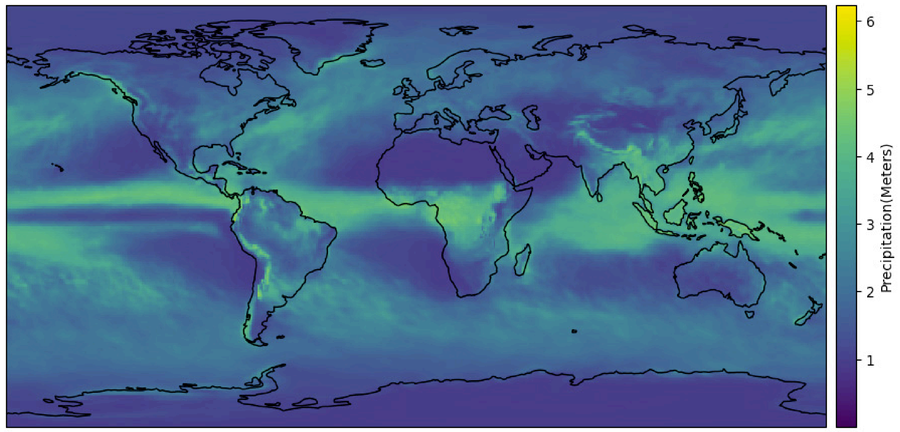


Figure 1. Total *Precipitation in 1984*

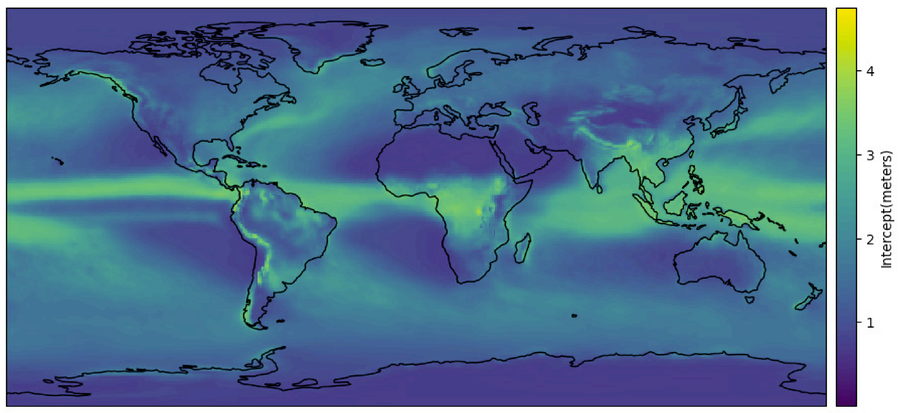


Figure 2. *Value of intercept from linear regression over 40 years*

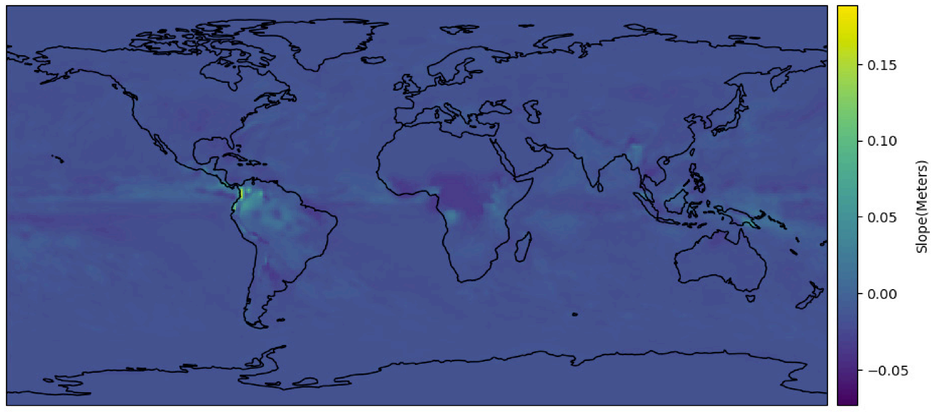


Figure 3. *Value of slope from linear regression over 40 years*

Full Charts & All Code

Discussion
In Figure 2, high values of precipitation are evident along the inter tropical convergence zone (ITCZ) near the equator. Therefore, the equatorial regions have high precipitation. Figure 2 of the intercepts captures the precipitation patterns in Figure 1, as it should. Inspecting Figure 3, we find these regions are rising in precipitation. Subtropical high pressure belts explain why areas below and above the equator don't get as much precipitation.

Conclusion
In Figure 1, we see the locations that have precipitation lie near the equator and oceans. In Fig ure 2, we see that the ocean and places with high intercepts have high slopes. In Figure 3, we can see that 1984 matches this trend quite nicely. This affects climate, influencing global temperatures, climate change, and natural disasters.

Partially related works
The World Meteorological Organization (WMO) leads much of the world in addressing climate change to snowfall. Author 1 wrote 2 programs that retrieve the latest weather data from all capitals around the world. One is more technical while the other is more high level. They use wget and pandas to produce CSV output files. They're a great introduction to anybody looking to learn pandas and contain in depth explanations for every line of code.

Acknowledgements
Thanks to the Environmentors and all coauthors.
Stack overflow also helped :)