Maria Sajer & Alex King
CptS 475
December 10, 2022

Climate Mapping Report

ABSTRACT

This project collects data from Google Earth Engine, specifically the ERA5_Hourly data set. There is also data collected from our coach for this project, Funso Oje, who collected his current coordinates along with a timestamp for each entry. The goal of this project is to gain a deeper understanding of what factors affect our climate the most, and how our climate has changed over the last twenty years. It appears that topsoil temperature can be a large factor in the fertility of the soil. Additionally, the temperatures are slowly rising, but the changes in pressure are more drastic throughout the year. The large swings in pressure can affect the soil in many ways other than its water content. The data reflects that in larger cities, the soil can often get too warm, while in less populated agricultural land the soil is at a more moderate temperature.

INTRODUCTION

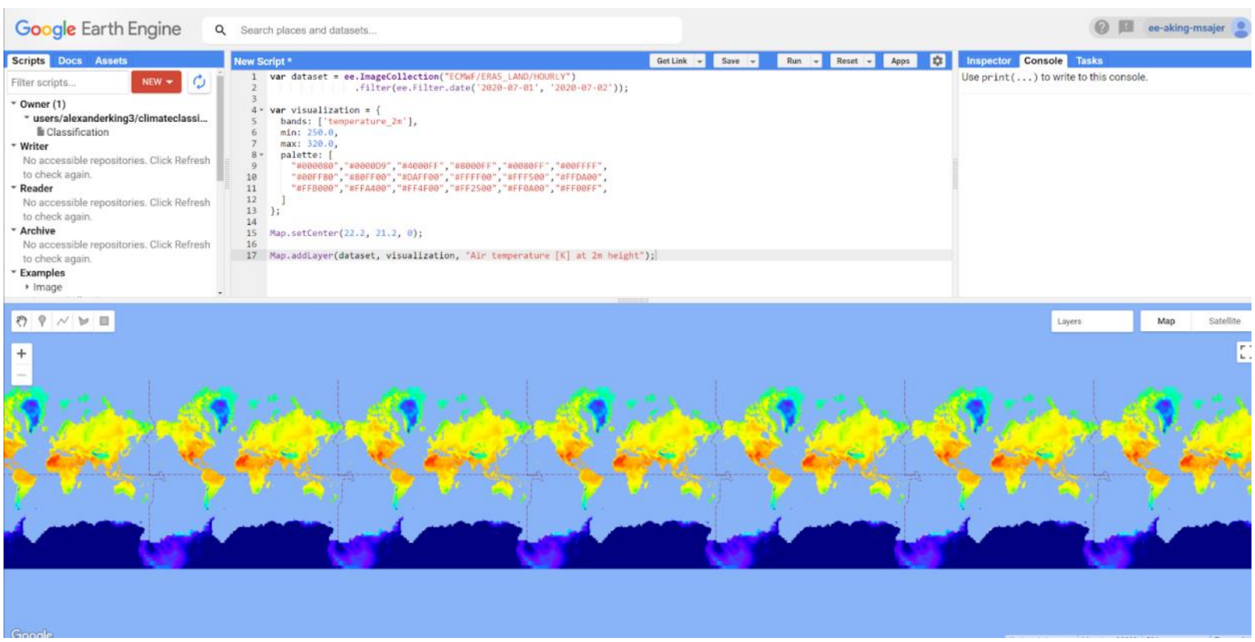
For the majority of human history, we have had little knowledge of our climate and how it can affect our lives. Over the past few decades however, we have made countless discoveries and it has shown us more than we could have ever expected. This has been possible with the power of satellites and the images they take of our planet. Our project utilizes satellite imagery, provided by the Google Earth Engine, to make new discoveries of our own. We are utilizing the data from these satellites, as well as other data they provide, combined with our own dataset given to us by Funso Oje, our coach for this project. The dataset we were given contains specific locations made up of coordinates in addition to the time the measurement was taken. The two locations these coordinates are taken at is Lagos, Nigeria and Pullman, WA, USA. This project analyzes many different qualities of each environment, and the stark contrast in these two locations provides some very definitive results.

PROBLEM DEFINITION

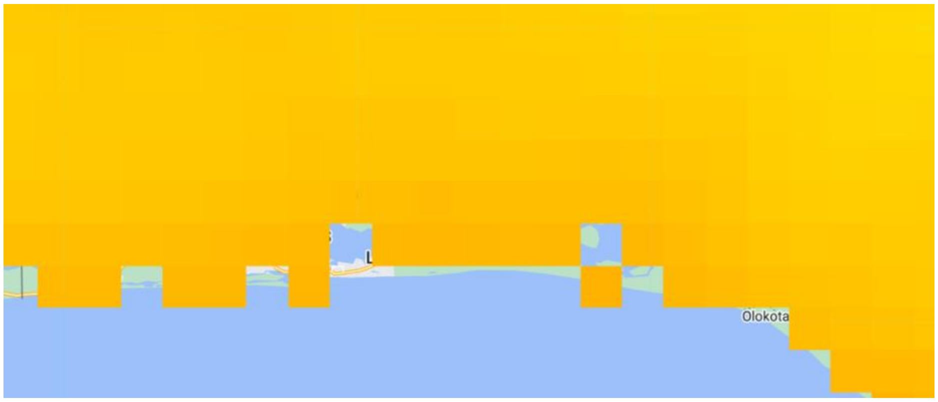
A major problem in the agriculture industry is our ever-changing climate. The weather is becoming more and more unpredictable, which can hurt farmers who want to protect their crops and soil. With research projects like this one, scientists can determine what causes soil to be rich, and what causes it to go bad. Additionally, discovering when crops need to be watered or cared for is a very important aspect of farming. We plan to discover what times of day soil becomes the most dry, and what other factors can lead to soil drying out. We are using the Google Earth Engine's satellite imagery database to gather information on specific areas at specific times. One would intuitively think that with greater temperatures, the more likely soil is to dry out, but we are going to determine what other factors may also cause this problem.

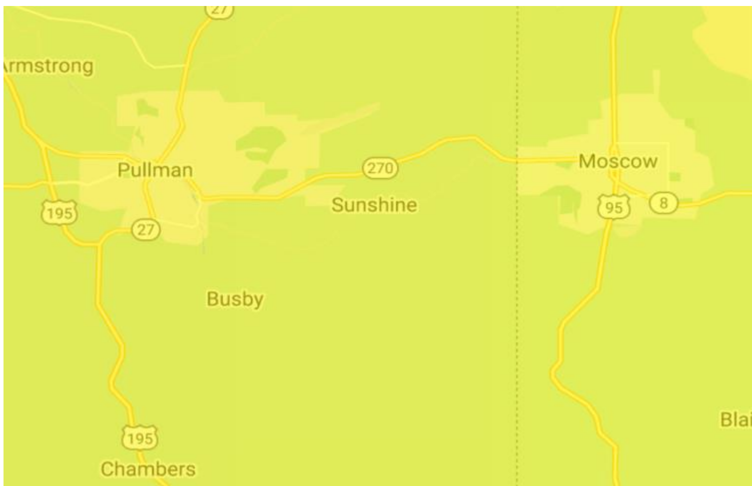
MODELS/ALGORITHMS/MEASURES

To begin our research, our first goal was to familiarize ourselves with the wide array of datasets that Google Earth Engine (GEE) has to offer. They all have different python scripts attached that allow us to view what it is they track. At the beginning, we had used the MODIS satellite. The satellite gives a mapping of the climate based on average temperatures in specific areas. However, this data was not very specific, and after some attempts at writing our own versions of the script we found we could not use this data. The MODIS satellite does not take images every hour, which is a large component of the locations dataset given to us. Thus, a different dataset was necessary. After discussing our options with Funso, we decided to use the ERA5_Hourly satellite dataset. This dataset takes images every hour every day and has a wide variety of measurements for us to look at. The default measurement, and mapping script focuses on the temperature two meters above the ground. The default script displays a map of every continent at that time with the coloring based on the temperature at that moment. As seen in the figure below, the Google Earth Engine has a python API we utilized to help visualize some of our scripts.



The editor allows us to zoom in and take a look at a more detailed level. These images on the right are of Lagos, Nigeria and Pullman, WA respectively. Clearly the images are not enough to get a full grasp of the climate in either location. Therefore we had to get creative in our research. We decided to look in depth at each location and have it return specific values of our choosing. Google Earth Engine allows users to run scripts of their own in Google Colab, a cloud-based Python notebook IDE.





ERA5 has a large number of measurements to access at each location. Some of the ones we have focused on are temperature at two meters, expected evaporation, skin reservoir content, surface pressure, and topsoil surface temperature.

IMPLEMENTATION ANALYSIS

As stated earlier, we are using the ERA5 hourly satellite imagery dataset combined with our locations dataset given to us by our coach. We are primarily utilizing the following measurements given by the datasets: longitude, latitude, date-time, temperature at 2m above the ground, surface pressure, skin reservoir content, and potential evaporation.

Our work is done by using the ERA5 hourly satellite dataset from Google Earth Engine, along with coordinates from Lagos, Nigeria and Pullman, Washington, each of which are associated with a date. To integrate both of these datasets, we first experimented with obtaining the temperature for a single coordinate in Google Earth Engine. We set the latitude and longitude and set the selected field to ‘temperature_2m’ to obtain the temperature for a single point on the map, as seen below. Here we can see that the temperature, in degrees Kelvin, is the output. This was the first step we took to begin our research on a deeper level.

```
var dataset = ee.ImageCollection("ECMWF/ERA5_LAND/HOURLY")
                .filter(ee.Filter.date('2020-07-01', '2020-07-02'));

var temp = dataset.select('temperature_2m')

var visualization = {
  bands: ['temperature_2m'],
  min: 250.0,
  max: 320.0,
  palette: [
    "#000080","#0000D9","#4000FF","#8000FF","#0080FF","#00FFFF",
    "#00FF80","#80FF00","#DAFF00","#FFFF00","#FFF500","#FFDA00",
    "#FFB000","#FFA400","#FF4F00","#FF2500","#FF0A00","#FF00FF",
  ]
};

Map.setCenter(22.2, 21.2, 0);

Map.addLayer(dataset, visualization, "Air temperature [K] at 2m height");

var geom = ee.Geometry.Point(30.0444, 31.2357)
Map.addLayer(geom)

var date = ee.Filter.date('2020-07-01', '2020-07-02')
var fields = ['temperature_2m']
var temp2 = dataset.select(fields).filter(date).first().reduceRegion({reduce
print(temp2)
```

```
Use print(...) to write to this console.

▾Object (1 property)                JSON
   temperature_2m: 295.8993835449219
```

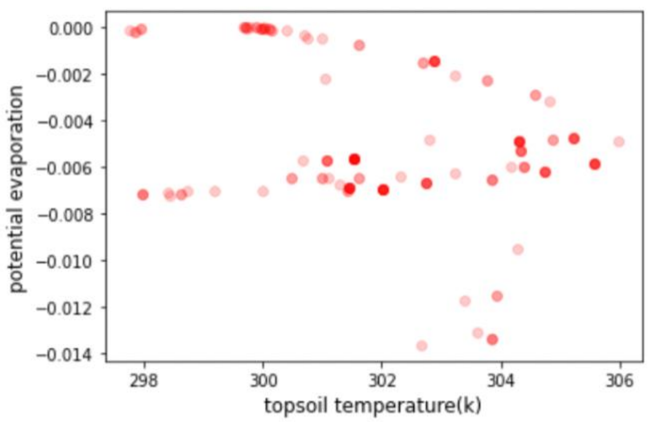
Our next task was to read the csv file containing the coordinates into our project to obtain the temperature, potential evaporation, soil temperature, surface pressure, and skin reservoir content for each point in the dataset. Once the dataset was read into an ee.FeatureCollection, we converted the ee.FeatureCollection to a Pandas dataframe in order to create plots with the data. The columns in the dataframe are temperature, potential evaporation, soil temperature, surface pressure, and skin reservoir content. However, there is a limitation while importing this information. Google Earth Engine only allows for five-thousand entries to be read at a time. Therefore, we needed to read in values at this interval, then append it to the pandas dataframe. There was another issue with this though; Google Colab has a limit on memory for users on a free subscription. Since this is a research project for a course, and not intended for publication we did not pay for extra memory and runtime. For this reason, we are only able to import about ⅕ of the dataset. (about 60k entries) The importing of the data still takes upwards of thirty minutes each time the runtime is disconnected.

Once the import is complete, the data must still be cleaned. Since the data we are joining the satellite imagery data with is composed of latitude and longitudes at a certain time, there are an overwhelming number of NaN entries because the satellite was not in position at that time. These were fairly easy to drop with the .dropna() function. However, this dropped our entry count to a miniscule 2888 entries. We attempted to drop the NaN entries at the same time we imported new data, but this took even more memory and caused us to have even less entries available. Ideally, we would not have limitations on memory, but Google Earth Engine does not permit access to their database from local machines without a paid subscription. With access to this subscription, we would be able to gain access to faster machines and it would allow us to process more data. Unfortunately, we only have access to the months of April and May over the years 2016-2019. With more memory and uninhibited access, we could expand this range to 2000-2020 and it would allow us to draw conclusions over year’s worth of data.

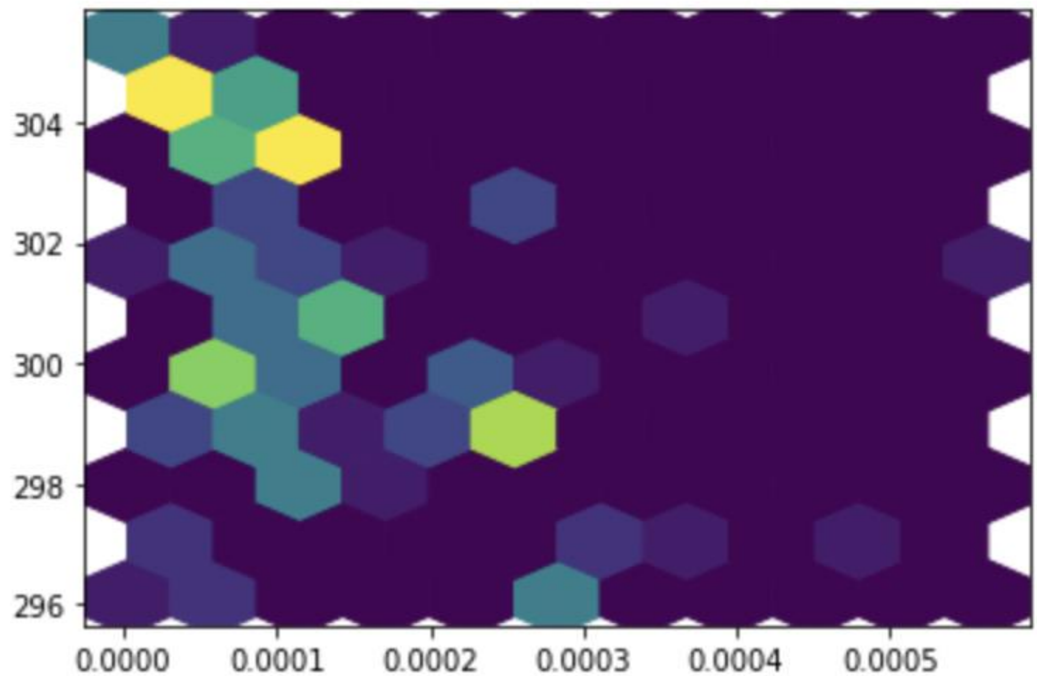
RESULTS AND DISCUSSION

A scatter plot is created by using the desired variable inputs. In the example below, soil temperature and potential evaporation are the variables being compared. Potential evaporation is the amount of evaporation that would occur if a sufficient water source were available. The temperature seems to have a slight impact on potential evaporation, as to be expected. There is variation seen on the plot due to the fact that the points are from both Lagos and Pullman, where the temperatures vary greatly. The soil in Lagos is much more humid than in Pullman, which makes the soil much less affected by temperature. In one location, as temperature increases, potential evaporation decreases. In the other location, as temperature increases, potential evaporation increases slightly. This may be because the climates in Lagos and in Pullman are very different.

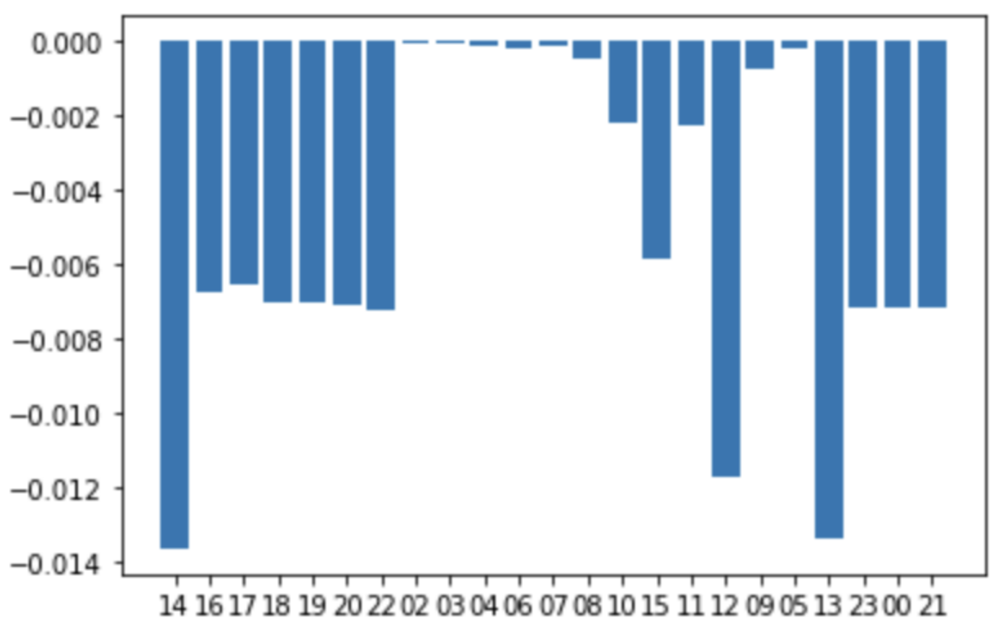
```
[ ]
    plt.scatter(df.soil_temperature_level_1, df.potential_evaporation, alpha=0.2,color="red")
    plt.ylabel('potential evaporation', fontsize=12)
    plt.xlabel('topsoil temperature(k)', fontsize=12)
    plt.show()
```



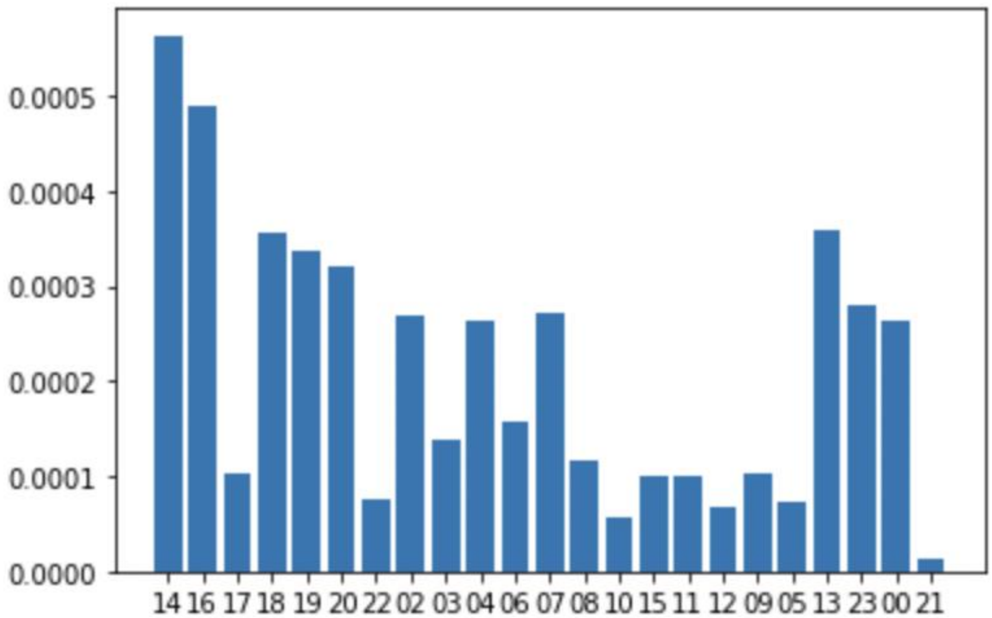
Another visualization we created correlates the skin reservoir content and the temperature two meters above ground. The skin reservoir content measures the amount of water in a vegetation canopy and/or in a thin layer of soil. The y-axis displays the temperature, and the x-axis displays the skin reservoir content, showing that there is a direct correlation between the two variables. The sections of the graph where the colors are brighter represents where there is a higher density of points. As the temperature decreases, the skin reservoir content increases.



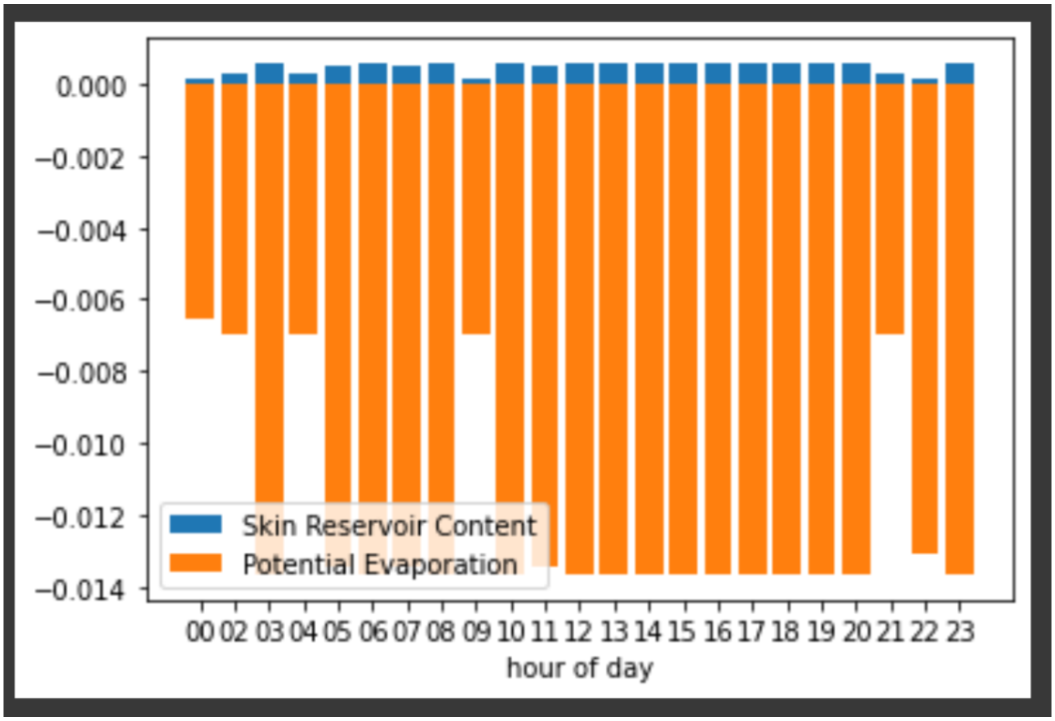
Another variable that we integrated into our data visualizations was time of day. Below is a bar graph with time of day, in hours, on the x-axis, and potential evaporation on the y-axis. This graph shows that potential evaporation is at its highest in the late afternoon. This can indicate to farmers that watering in the late afternoon is the most crucial time of day to ensure that crops don't dry out. In addition to agriculture, this can indicate that it will be the hottest time of day, as the ground will be at its highest temperature.



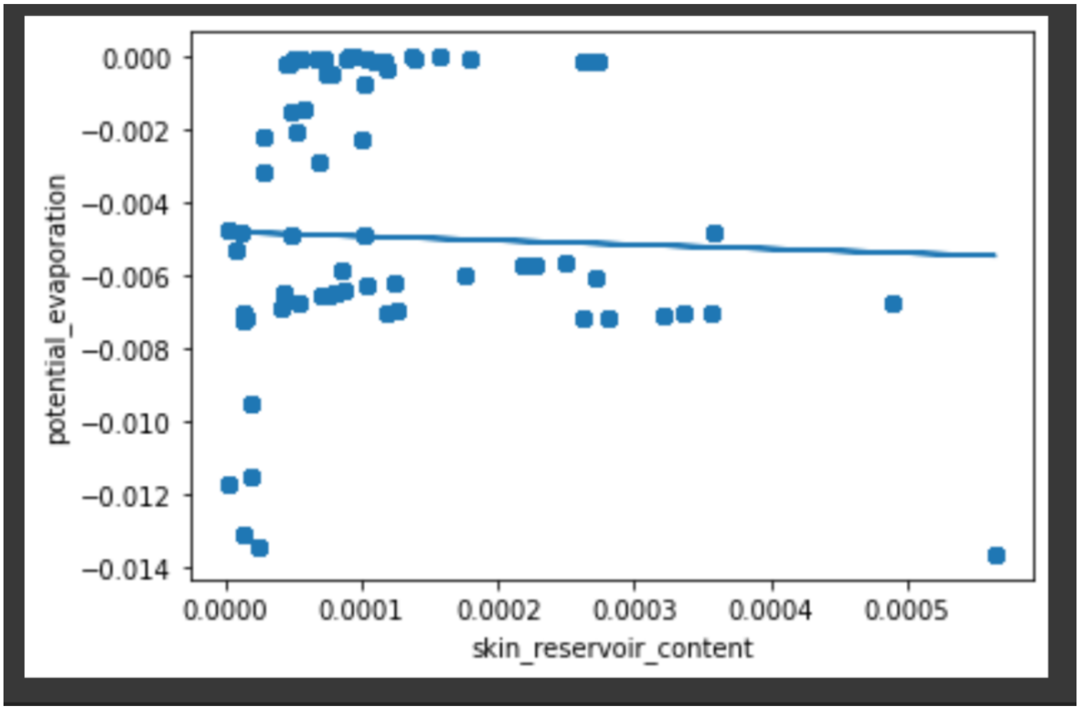
Another variable we compared regarding time of day is skin reservoir content. The bar graph below shows time of day on the x-axis and skin reservoir content on the y-axis. Skin reservoir content measures the amount of water in a vegetation canopy and/or in a thin layer of soil. Skin reservoir content is the highest at 2pm and at 4pm. This aligns with the potential evaporation around that time of day too.



In the figure below, these two measurements are plotted in the same chart.



As seen above, the potential evaporation is highest during the afternoon, while the skin reservoir content is highest as well. These two may seem like conflicting the results, but in reality, one is the result of the other. The skin reservoir content is the amount of liquid in the foliage on the ground, which would gain a greater content as the ground water evaporates in the afternoon. A small amount of the evaporated water from the ground condenses on the foliage above, thus creating a greater skin reservoir content.



The measurements show a lot of outlying values, but the line of best fit shows that as the skin reservoir content increases, the potential evaporation also grows. This may be a bit misleading because the line has a negative slope, but that is only because potential evaporation is measured by the amount of water lost, not the amount of vapor the air gains. This chart above coincides with our previous hypothesis. Clearly, to prove such a hypothesis would take a much more comprehensive report and further experimentation in a closed environment, but for our purposes the data from ERA5 is sufficient.

RELATED WORK

When choosing a dataset from Google Earth Engine to work with, many related datasets were available. The work we did on our dataset could be done on many other datasets with different data. For example, there was a snow cover dataset and a dataset with data about the ocean. Different variables could be examined on both datasets in a similar way to the way in which we conducted research and drew insight from data.

A project that uses satellite imagery to help understand the climate more accurately is the Nimbus AI project from the University of Hawaii. This project uses satellite imagery to make accurate weather forecasting and create what they call “nowcasting”. They can make accurate predictions for where the cloud cover will travel within the next hour with an image in ten-minute intervals. Their goal with the project is to be able to determine when solar panels will be at their peak for electrical output, as well as knowing how much energy can be stored day to day. The data can help Hawaiians, and many others, determine how much energy they can access and gain from solar panels, in an attempt to help Hawaii and everywhere else transfer to clean sustainable energy. (Sadowski. P,...)

CONCLUSION

In this report, we discovered that soil temperature can vary depending on many factors, but keeping it moist can be the most crucial part of keeping it fertile. The information we have collected suggests specific times of day to water crops, as well as ensuring soil does not get too hot or too warm. Soil quality is one of, if not most, important aspects of agriculture. Ensuring that crops are planted and maintained in a rich, fertile soil can be the difference in a good yield during harvest. Otherwise, the farmer may be in serious financial trouble, and on a grander scale so would the consumer. A smaller yield means greater costs for basic goods and can create a ripple effect throughout the entire economy. Ensuring consistency in soil quality is crucial with our ever-growing population. More people being born means more people need to eat, and the agricultural landscape must adapt. Doing similar studies could help prospective and current farmers determine where to expand their farms. Much of the land in the United States is uninhabited, and untouched. Much of this land can be converted into agricultural land if necessary, and a great way to determine that would be to analyze the climate in that area. If this project were to expand, one thing that should be prioritized is expanding the range of imagery. The database collects images dating back to the year 2000, but the limits in Google Colab have prevented us from processing this entire timeframe.

BIBLIOGRAPHY

“Era5-Land Hourly - Ecmwf Climate Reanalysis | Earth Engine Data Catalog | Google Developers.” *Google*, Google, https://developers.google.com/earth-engine/datasets/catalog/ECMWF_ERA5_LAND_HOURLY?hl=en#bands.

Google. (n.d.). Google accounts. Retrieved December 13, 2022, from https://code.earthengine.google.com/

Oje, F. (n.d.). locations. Pullman, WA; Lagos, Nigeria.

Sadowski, P., Galgon, G., Torre, G., & Hart, K. (n.d.). *Our technology*. Nimbus AI. Retrieved December 12, 2022, from https://nimbus.solar/about/

APPENDIX

https://colab.research.google.com/drive/1jOtjIUD0P5dibc0TgTywZS5eGcACp7yI?usp=sharing