CEE 260D
Jack Seagrist
Karthik Ramesh

Determining Correlation Strength between Soil Moisture and ET Across Varied Land Covers

Introduction

Soil moisture and evapotranspiration are two important hydrological variables that are used in a variety of calculations holding utmost importance in the agriculture sector. The two variables are also connected, with several studies estimating evapotranspiration by calculating the potential evapotranspiration and then applying a soil moisture function obtained from remotely sensed datasets as seen in Mahfouf et al., 1996 [1]. The evapotranspiration and soil moisture relationship has been studied in many different capacities, such as the study in *Vivoni et al. 2008* to understand the soil moisture control on evapotranspiration in monsoon-dominated ecosystems [2]. Our project will explore the correlation between remotely sensed soil moisture and in-situ evapotranspiration measurements and attempt to determine which factors govern the strength of the relationship.

Specifically, we will be exploring the relationship between ET and soil moisture across varying land cover types. Soil moisture varies in both the lateral and vertical directions, which can be attributed to controls from precipitation, climate, land-cover, and soil properties among other influences [3]. In addition, significant variations can be seen in soil moisture in areas with wet and dry seasons [4]. We will use data provided by the California Irrigation Management Information System (CIMIS) for in situ measurements of ET. In addition, we will explore the correlation between soil moisture and other collected CIMIS measurements, such as net radiation, to understand the relative strength of the relationship.

From our initial background research we expect land cover to have an influence on the relationship between ET and soil moisture. Findings from *Sazib et al 2018* investigating drought assessment using soil moisture found that the land cover type has a significant impact on the relationship between soil moisture anomalies and other drought indicators [5]. By looking at the relationship between CIMIS in situ measurements and soil moisture across varying land cover

---

[1] Mahfouf, J.-F., et al. (1996), Analysis of transpiration results from the RICE and PILPS Workshop, Global Planet. Change, 13, 73–88.
[2] Vivoni, E. R., Moreno, H. A., Mascaro, G., Rodriguez, J. C., Watts, C. J., Garatuza-Payan, J., and Scott, R. L. (2008), Observed relation between evapotranspiration and soil moisture in the North American monsoon region, *Geophys. Res. Lett.*, 35, L22403, doi:10.1029/2008GL036001.
[3] https://www.sciencedirect.com/science/article/pii/S0022169410007481?via%3Dihub#bi0005
[4] https://www.sciencedirect.com/science/article/pii/S0022169410007481?via%3Dihub#bi0005
[5] Sazib, N.; Mladenova, I.; Bolten, J. Leveraging the Google Earth Engine for Drought Assessment Using Global Soil Moisture Data. *Remote Sens.* **2018**, *10*, 1265.

regimes, we hope to identify the parameters which have the greatest influence on soil moisture and determine the underlying drivers of those relationships.

Data Sources

We plan to use two main sources of data for this project: soil moisture observations and in-situ CIMIS measurements.

We will obtain the soil moisture observations from the NASA-USDA data set on google earth engine [6]. This dataset is created with the SMAP Level 3 soil moisture observations and has a 0.25° x 0.25° spatial resolution.

We will use data provided by the California Irrigation Management Information System (CIMIS) for in situ measurements [7]. We plan to include all relevant metrics captured by the CIMIS station, such as ET, net radiation, average temperature and average soil temperature. The data provided from CIMIS will have a spatial resolution of 2.5km x 2.5km and will be downloaded in csv format.

CIMIS data download

From the CIMIS website, we were able to download CIMIS data for WY 2016 to 2019 across all stations. We included all recorded measurements in the csv files. With the data downloaded, we were able to cycle through the station ID's and obtain the station coordinates using the CIMIS Web API.

CIMIS Station Points

With google Earth Engine, we were able to import the CIMIS station coordinates and create a set of points corresponding to each station. With each point, we created a bounding box 2.5km x 2.5km which relates to the specified resolution of CIMIS measurements. The picture below shows the bounding box for the Arvin-Edison station.

[6] https://developers.google.com/earth-engine/datasets/catalog/NASA_USDA_HSL_SMAP_soil_moisture
[7] https://cimis.water.ca.gov/

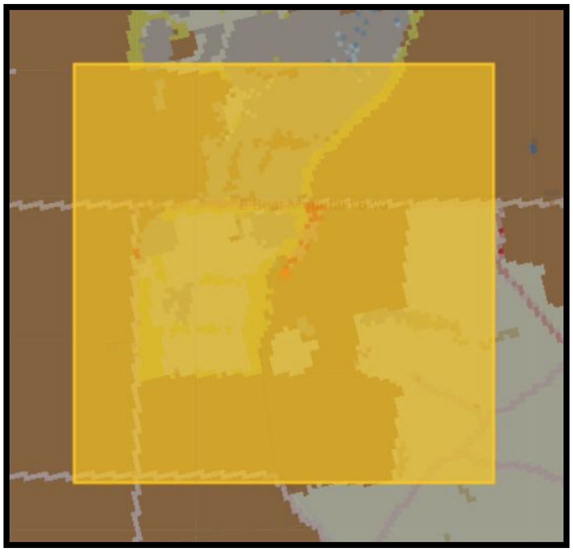

Fig 01 - Arvin-Edison CIMIS Station with 2.5 km x 2.5 km box

Time Series SMAP data

Using Google Earth Engine, we were able to import the SMAP dataset. Once the SMAP data was loaded, we were then able to create a script which generates a time series plot of a selected SMAP band for a single CIMIS station. The example below shows a time series output of surface soil moisture for Arvin-Edison across WY 19.

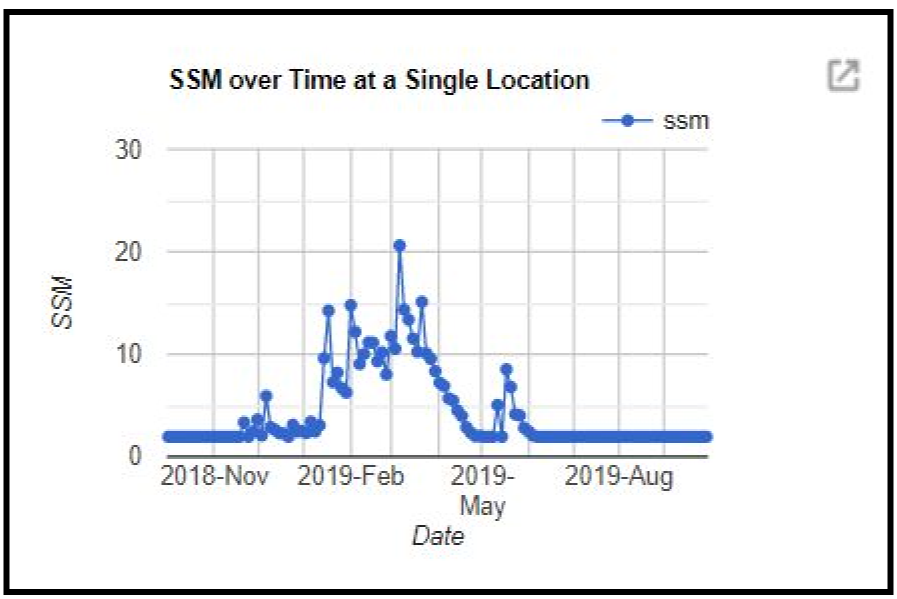

Fig 02 - Arvin-Edison GEE Time Series Output for SMAP Surface Soil Moisture

Analysis

The aforementioned data was collated from the respective sources for the following locations chosen in different parts of California:

- Arvin Edison - San Joaquin Valley
- Bishop - Inyo County
- Scott Valley - Northeastern Plateau
- Seeley - Imperial County

Table 1: Pearson correlation between SMAP soil moisture and CIMIS parameters for Arvin Edison

Parameter	Evapotranspiration	Net Radiation	Average Temperature	Average Soil Temperature
Soil Moisture	-0.4	-0.33	-0.39	-0.31

Table 2: Pearson correlation between SMAP soil moisture and CIMIS parameters for Bishop

Parameter	Evapotranspiration	Net Radiation	Average Temperature	Average Soil Temperature
Soil Moisture	-0.3	-0.27	-0.39	-0.46

Table 3: Pearson correlation between SMAP soil moisture and CIMIS parameters for Scott Valley

Parameter	Evapotranspiration	Net Radiation	Average Temperature	Average Soil Temperature
Soil Moisture	-0.68	-0.59	-0.73	-0.83

Table 4: Pearson correlation between SMAP soil moisture and CIMIS parameters for Seeley

Parameter	Evapotranspiration	Net Radiation	Average Temperature	Average Soil Temperature
Soil Moisture	-0.4	-0.37	-0.38	-0.43

From Tables 1-4, we can infer that there exists a negative correlation between soil moisture and the hydrological parameters under analysis. However, there is a difference in the value of Pearson R as observed above. Scott Valley seems to exhibit the strongest relationship between the datasets while the other three stations show almost similar relationships despite the varied geographic locations.

We also attempted to construct a model trained using the CIMIS parameters and SMAP dataset to predict the soil moisture values. Using the score function on python, we were able to observe the following levels of accuracy:

- Linear Regression : 27%
- Random Forest Regressor : 40%

The models especially fail for low levels of soil moisture (below 1mm) and the higher outliers (above 10mm).

This leads us to believe that additional factors such as land cover type, land-use pattern, vapor pressure, relative humidity could serve as driving factors for the retention of moisture in the soil. For the next part of this project, we aim to further investigate these relationships by grouping stations based on land cover type and will try to refine our analysis.

Current Challenges

In this section we discuss current challenges that we have encountered during our analysis. One important challenge to note is what appears to be inconsistencies with the data. There appear to be long stretches of time where there are unchanging data points. We believe this may be due to the fact that the actual changes occurring in soil moisture values are much lower than the capabilities of SMAP and therefore show a consistent value across the time period. As we can see from Fig, we have the same value of soil moisture for all the available years. Further investigation is needed to understand this pattern.

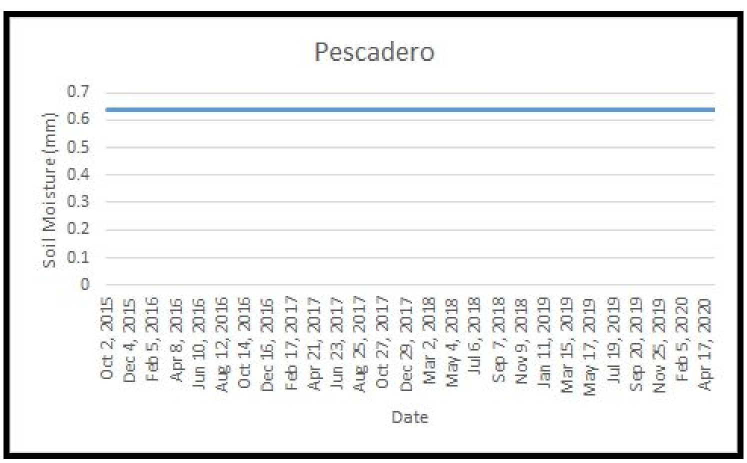

Fig 03 Timeseries of soil moisture at Pescadero

Future Work

The following section describes the work that is needed to complete the final project. The major outstanding task is to aggregate the land cover data in google earth engine. Our team will identify the majority land cover type at each CIMIS station to be used as the label for that correlation analysis. We currently have all of the data points and layers in place on google earth engine and simply need to identify the land cover type at each station which can be calculated by identifying the most repeated cover type in the box created around each CIMIS station.. Once that data has been collected, we will be able to understand the influence of land cover on the strength of the correlation.

Once this work has been completed, we will be able to continue with our analysis in better exploring the influence of additional factors such as slope, vapor pressure, relative humidity etc.

Work Breakdown

Karthik Ramesh and Jack Seagrist designed the project hypothesis, plan, and conducted the preliminary background research together. Jack performed the analysis to programmatically gather CIMIS station coordinates, the resulting land cover data from google earth engine, and generate a time series output of SMAP data at that station. Karthik joined the data outputs and provided the initial statistical analysis to determine the correlation between soil moisture and CIMIS parameters. Both team members helped create the report.