Evaluating ECOSTRESS and Landsat 8 EVI for mapping irrigated areas in Ahuachapán, El Salvador

by Monica Quezada

Capstone Project Report

CRS CATHOLIC RELIEF SERVICES

Raíces AHUACHAPÁN

UC DAVIS UNIVERSITY OF CALIFORNIA

TABLE OF CONTENTS

COLLABORATORS

This research study was a collaboration between Catholic Relief Services RAICES program and UC Davis International Agricultural Development.

DONORS

This research was made possible through funding from the Henry A. Jastro Graduate Research Award, The Global Fellowships for Agricultural Development (GFAD) program, and the Blum Center for Developing Economies Poverty Alleviation through Sustainable Solutions (PASS) project awards program.

ACKNOWLEDGEMENTS

There are various people to thank for helping me bring this project to fruition. I was very privileged to have Kate Scow, Distinguished Professor of Soil Science and Microbial Ecology and Chair of the International Agricultural Development Group, as my advisor for this project. She was instrumental in helping me navigate the ups and downs of this journey. My other committee members, Isaya Kisekka, Associate Professor of Agricultural Water Management and Irrigation Engineering, and Mallika Nocco, Extension Specialist in Soil-Plant-Water-Relations, also provided invaluable insights and helped me move forward during certain moments when I felt I had hit a dead end.

Paul Hicks, the Co-Director of the RAICES program, and Adam Keough, Regional Project Monitoring and Evaluation Consultant, facilitated this collaboration by helping me get acquainted with RAICES goals and activities, finalizing project objectives, and connecting me with other RAICES staff. Francisco Casares, agronomist and RAICES irrigation manager, and Bernardo Romero, who works for the Ministry of Agriculture, helped me understand some of the current irrigation practices and regulations in El Salvador, and Manuel Pacheco, Data Manager for RAICES, was my primary contact for data needs.

Last but not least, Chippie Kislik, currently an Environmental Science, Policy, and Management PhD student at UC Berkeley, also devoted many hours helping me better understand potential remote sensing research questions as well as how to navigate ArcGIS Pro and Google Earth Engine. She was a very helpful sounding board and assisted me in interpreting my results.

TABLE OF FIGURES

ABSTRACT

Climate change in the form of increasingly unpredictable rains and increasing temperatures is posing a threat to farmer livelihoods. El Salvador, in particular, faces the prospect of increased dry periods during July and August and increasing rain intensity which will disproportionately affect its 400,000 subsistence farmers. Irrigation offers a potential for smallholder farmers to adapt to these increasingly challenging environmental conditions while stabilizing or increasing yields. However unregulated irrigation expansion can also pose a threat to freshwater resources. Understanding the extent of smallholder irrigation is important for regional water resource managers and for organizations working on these issues. Remote sensing using publicly available datasets offers a low-cost way to map and monitor irrigated lands and balance agricultural development and productivity with water resources management. While various irrigation mapping studies have been carried out in the U.S. and other temperate areas, only a handful of studies have been conducted in developing countries and/or tropical climates. To our knowledge, there are no published studies that use ECOSTRESS data to map smallholder irrigation. This project was a collaboration between Catholic Relief Services' RAICES program and UC Davis. The objective of this study was to map irrigation in the Agua Caliente watershed in northern Ahuachapán in El Salvador. We looked at time series for two data products, ECOSTRESS' Evapotranspiration layer and Landsat 8 Enhanced Vegetation Index (EVI), in order to determine whether or not we could identify a threshold. Based on the map we produced, which classifies irrigated annual crop areas which have an EVI $\geq$ 0.27, 5,523 hectares were classified as irrigated in the dry season. We discuss some of the weaknesses of this analysis and recommendations for future work in this area.

INTRODUCTION

Around 80% of the world's agriculture, accounting for two thirds of the world's food, is rainfed. In Latin America, that number is even higher, almost 90% (Wani *et al*., 2009). Climate change, in the form of increasingly unpredictable rains and increasing temperatures, is posing an increasing threat to farmer livelihoods. The United Nations Food and Agriculture Organization (FAO) has classified the Dry Corridor in Central America, which includes El Salvador, Guatemala, Honduras, and Nicaragua, as one of the most vulnerable regions in the world due to its high climate variability, weak institutions and the low socioeconomic status of many of its inhabitants. El Salvador, in particular, faces the prospect of increased dry periods during July and August and increasing rain intensity (Dorrance *et al*., 2011) which will disproportionately affect its 400,000 subsistence farmers (FAO, 2021). To make matters worse, El Salvador is the most densely populated country in Central America and has the lowest water reserves in the region (Lakhani, 2019). El Salvador is expected to face a severe water crisis in the future unless drastic measures are taken to improve and protect its water supplies (EFE Noticias, 2016).

Irrigation offers a potential for smallholder farmers to adapt to these increasingly challenging environmental conditions while stabilizing and increasing yields. However, irrigation also puts tremendous demands on fresh water resources in many countries (Johansson, 2005). In some countries, it accounts for 70% or more of water withdrawn from rivers, lakes and groundwater sources (Cai & Rosegrant, 2002). If not adequately monitored and managed, expanding irrigation can pose a threat to freshwater resources. Aside from the threat of depleting water resources, it can lead to increased pollution if expanding irrigation is coupled with increased fertilizer and chemical use. Unregulated irrigation can also result in conflicts between water users (de Fraiture & Giordano, 2014). Understanding the extent of smallholder irrigation is critical for regional water resource managers and for organizations working on these issues.

The Catholic Relief Services' (CRS) RAICES program works at the nexus of these two issues, agricultural development and natural resource management. CRS has an established presence and network in the region with a large initiative on Water-Smart Agriculture in Latin America. RAICES, which stands for Restorative Agriculture In Critical EcoSystems, is a 12-year program aimed at implementing water-smart agriculture at the landscape level. The program was launched in the department of Ahuachapán in western El Salvador and is currently planning expansions to other departments. One of their major goals is to expand irrigation technologies to smallholder farmers while mitigating negative environmental impacts. To fulfill this goal, CRS has partnered with the Agriculture Development Bank to create a $250,000 loan fund. However, being able to expand irrigation in an environmentally conscious way is hampered by the lack of publicly available data for irrigated areas in Ahuachapán.

Remote sensing using publicly available datasets (of which there are many) offers a low-cost way to map and monitor irrigated lands and balance agricultural development and productivity with water resources management. While remote sensing can be used at multiple scales, it is particularly useful at the regional level. Global maps of irrigated areas contain information for El Salvador, but these are often outdated or the resolution is too coarse to adequately identify smallholder irrigation. FAO's AQUASTAT provides an estimate of irrigated areas at the department level, but lacks estimates at the local level. Furthermore, this data is based on a 1993 World Bank study on private irrigation (FAO, 2021). Thenkebail *et al. (*2009) produced the "first satellite-sensor-based" Global Irrigated Area Map (GIAM) with 10 km resolution using climate data, ground truth data and remote sensing data from multiple satellites. This was definitely a leap forward as it captured 43% more irrigated area than what had been reported in FAO statistics (Meier *et al*., 2018). However, its spatial resolution of 10 km is still much too coarse to capture smallholder plots that are often smaller than one hectare.

While various irrigation mapping studies have been carried out in the U.S. and other temperate areas, outside of India, only a handful of studies have been conducted in developing countries and/or tropical climates (South Africa, Burkina Faso, Ethiopia) (Cai *et al*., 2015; Chaix-Bar *et al*., 2019; Vogels, 2019). For example, one study carried out in Burkina Faso, which has a similar climate to El Salvador, reported a much higher irrigated area than what had been reported in previous irrigation maps (Chaix-Bar *et al*., 2019). This reflects a common pattern in which irrigation on smallholder farms is often under-reported. Additionally, while many studies have used the Normalized Difference Vegetation Index (NDVI), which captures greenness or healthy vegetation, from Landsat or other satellite datasets, very few have tested the efficacy of ECOSTRESS, as it is a relatively new dataset.

The objective of this study is to map irrigation in the Agua Caliente watershed in northern Ahuachapán in El Salvador. This was done via analysis of two publicly available satellite datasets, ECOSTRESS and Landsat 8, along with ground truthing points collected from CRS RAICES participant sites and precipitation data collected from CRS RAICES weather stations. The goal was to benchmark current irrigation areas and practices in Ahuachapán, El Salvador, to inform expansion of irrigated areas. By mapping these data, we hope to provide information for RAICES to determine best irrigation practices based on geographic characteristics.

APPROACH

Study area

Our study area was the Agua Caliente watershed situated in northern Ahuachapán in western El Salvador (Figure 1). This is the primary area where the RAICES program launched and they have set up 3 weather stations in the region to collect climate data. It is an area of about 20,300 hectares with elevation ranging from 0-1600 meters in the southern region (Figure 2). The area has a Koppen Classification of Aw or Equatorial/Tropical savanna with a dry winter (McKnight & Hess, 2000). The average temperature is ~74° F and the average rainfall is 1747 mm. The area has a distinct dry season from November to April along with a short drought period in the rainy season known as the "canicula". Figure 3 shows a seasonal calendar for a typical year in Central America created by the Famine Early Warning System Network (FEWS NET, 2013).

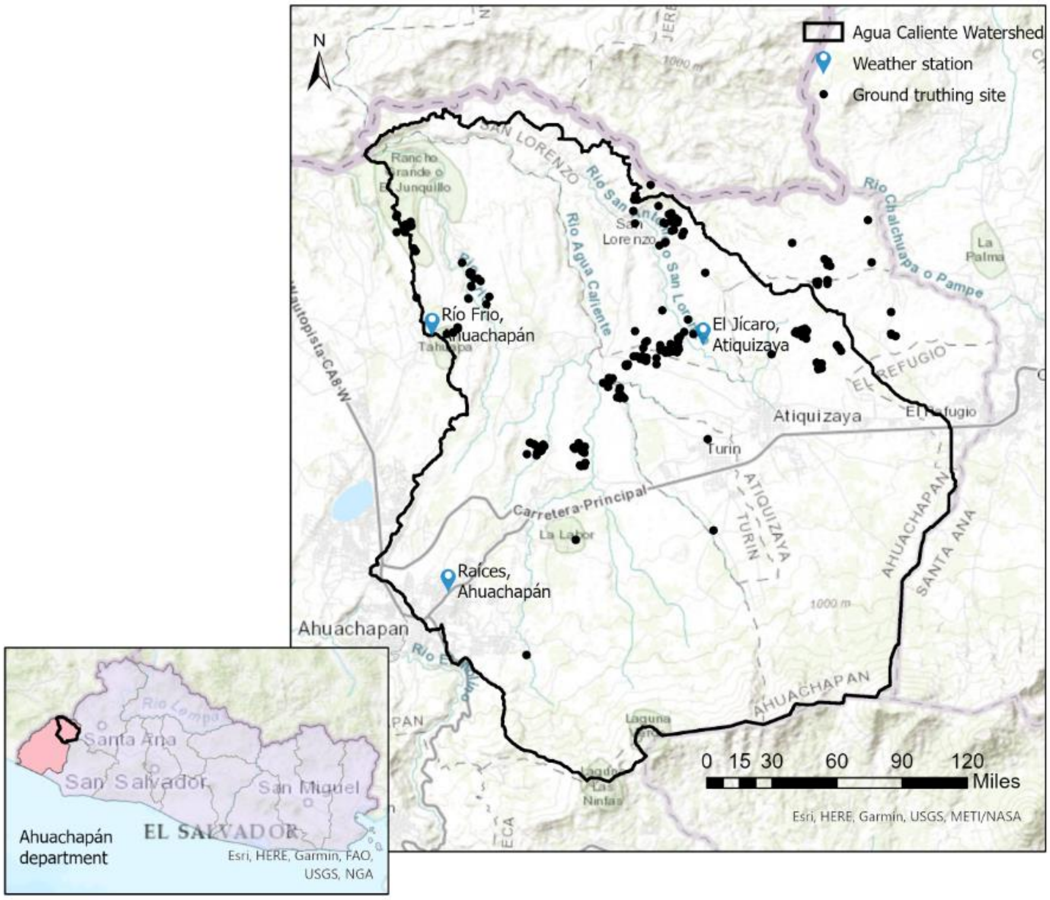


Figure 1: Map of Study Area in northern Ahuachapán

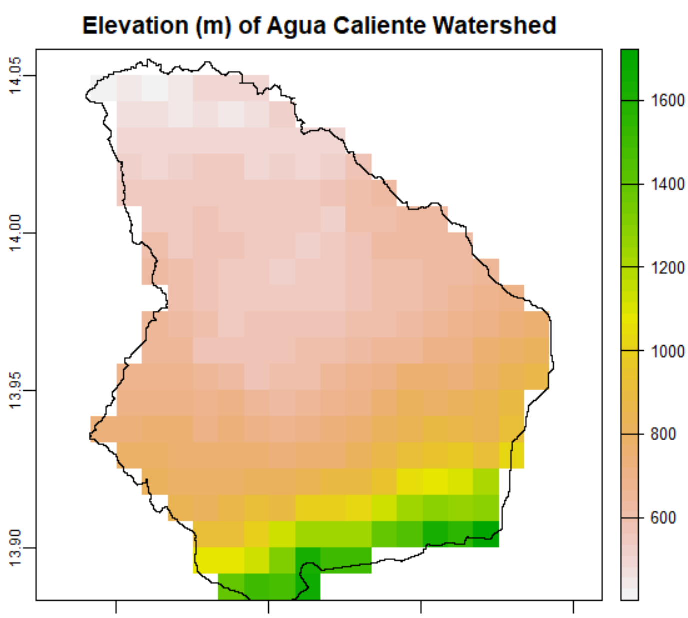


Figure 2: Elevation range of Agua Caliente Watershed

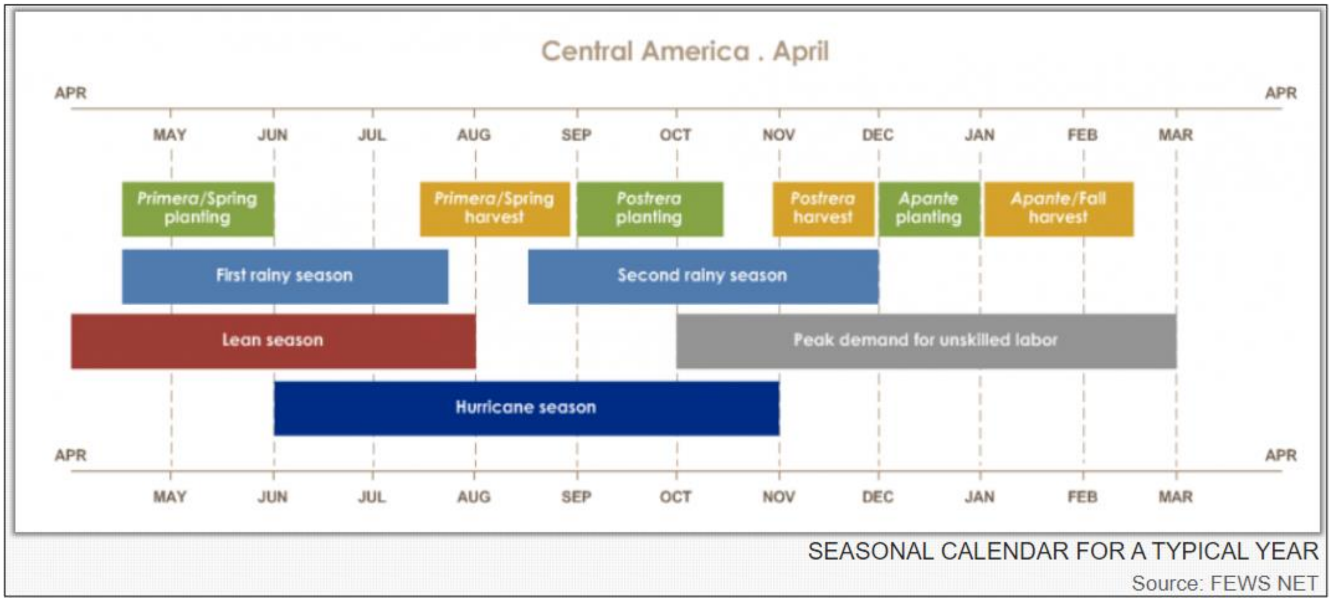


Figure 3: FEWS NET Seasonal Calendar for a Typical Year in Central America

Information on farm characteristics and irrigation trends came from various sources: irrigation permit data from the Ministry of Agriculture and Livestock, conversations with RAICES staff, and data collected from ground truthing sites.

When the government of El Salvador issues irrigation permits, it collects a variety of data including georeferenced water uptake points, water source, total property area, irrigated area in manzanas (Central American unit of measurement, 1 manzana (mz) = 7,025.79 square meters), distribution system (sprinkler, gravity, drip), hours of irrigation, frequency of irrigation, and crops irrigated. Unfortunately, these data are not publicly available and only a limited set of data were made available for this study. Requests on behalf of CRS to the Ministry of Agriculture for more recent irrigation permit data were not fruitful. While more recent data would have been helpful in verifying our results, it would still have fallen short in capturing all irrigation and where it is occurring due to two limitations. For one, there is not 100% compliance in terms of registering for irrigation permits. Second, while irrigation permit data indicates the amount of irrigated area associated with each water uptake point, farm locations are not geo-referenced so the exact location of these plots is unknown. Figure 4 shows the water uptake points and their associated irrigated area.

Irrigated area (mz) from georeferenced water uptake points 2014-2015

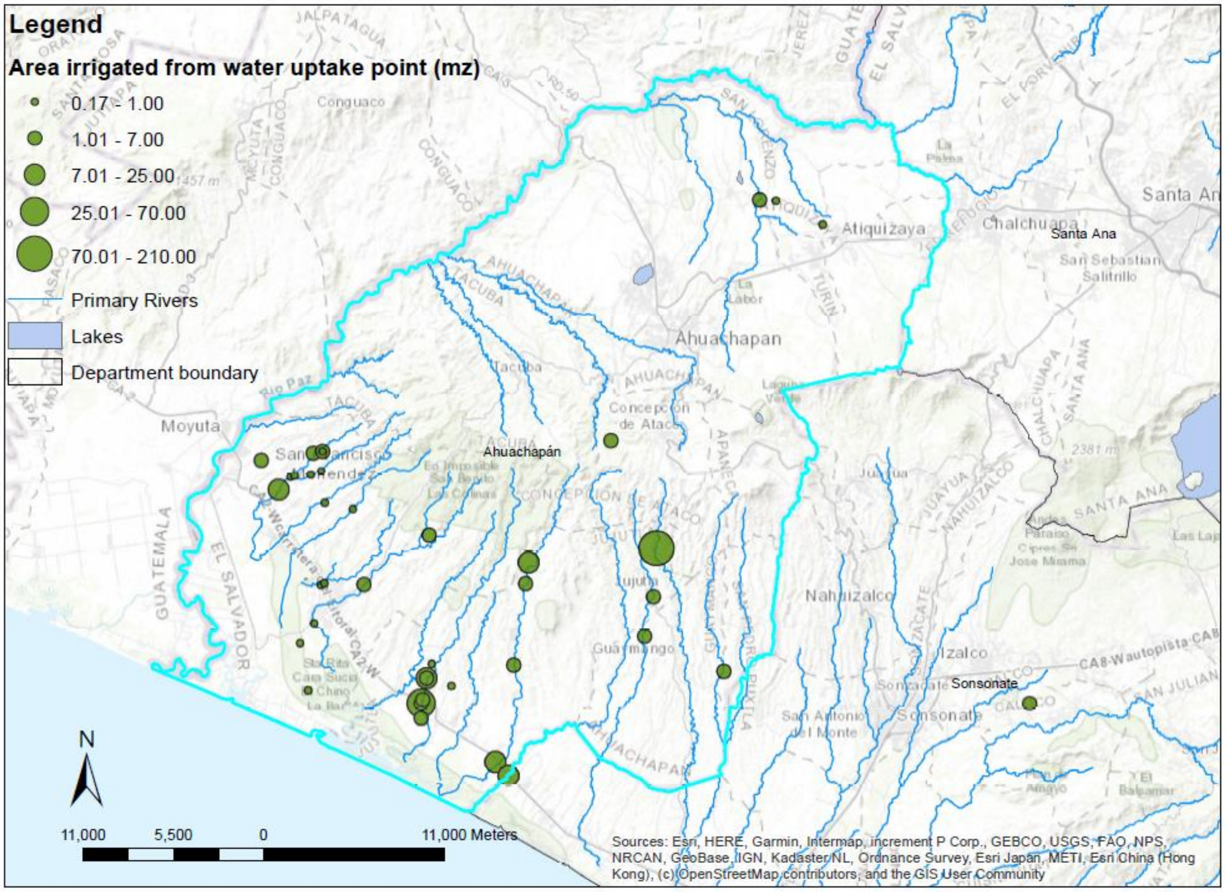


Figure 4: 2014-2015 Irrigated area (mz) from georeferenced water uptake points

While this data had its limitations, it did serve to understand farm characteristics and irrigation trends. The 2014-2015 Ahuachapán irrigation permit data that we received revealed that out of 67 permits, 85% of permits had a farm size of 25.5 hectares (ha) or less. 75% of the permits had an irrigated area size between 0.25 and 2.5 ha. Another insight from this data was that the most common water source is rivers (88% of permits) and the most common distribution system is gravity fed furrow irrigation (84% of permits). Other distribution systems were drip, sprinkler and a mix. Other water sources were wells, and smaller bodies of water such as a brook.

Depending on the source, the average farm size and irrigated area differed slightly. According to RAICES program staff, the average farm size is around 3 ha. This may be because the population with whom CRS works differs from the general farmer population in terms of farm size.

Data from ground truthing points for irrigated sites, provided by CRS, indicated that the average irrigated area ranges between 0.35 ha up to 8 ha depending on the crop (Table 1). Percent irrigated area for these sites is unknown as we do not have data for the total area of these farms.

Table 1: Average irrigated area by ground truthing point crop classes

Ground truthing points crop class	Count	Average Irrigated Area (ha)
Hort	27	1.61
Hort/Fruit trees	21	1.40
Fruit trees	21	0.85
Grains/Hort/Fruit trees	4	0.78
Agroindustrial	4	2.80
Grains/Hort	3	0.35
Grains/Pasture	2	0.85
Fruit trees/Pasture	2	0.85
Coffee	1	2.11
Sugarcane	1	8.00
Pasture	1	0.70
Coffee/Fruit trees	1	0.70
Sugarcane/Hort/Fruit trees	1	4.20
Grains/Fruit trees	1	0.35
Grains	1	0.35
Total	**91**	**1.38**

The crops grown in the region include: sugarcane, coffee (grown in higher elevation areas), basic grains such as maize, beans, sorghum, and horticultural crops such as elote (corn eaten fresh), cassava, lemon, ayote, bell pepper, cucumber, radishes, pipian, tomatoes, etc. There are also fruit trees such as banana, guayaba, jocote, papaya, mango, orange, and cacao.

There were some issues with the categorizing system that CRS used for the ground truthing crop classes. For example, a Hort/Fruit trees entry was categorized as Fruit trees and another entry was categorized as Hort/Fruit trees when crops grown included basic grains. There were also a few cases where the specific crops grown were not mentioned. Lastly, there were crops included in the list (i.e. jack bean) that are used as cover crops.

Methods

Common methods for irrigation mapping include thresholding and machine learning. Thresholding works well in arid climates where irrigated areas are clearly visible and can be easily identified. Machine learning involves classification algorithms such as Random Forest and Classification and Regression Tree (CART) to name a few. These methods use all available bands as input data to train the model. They then determine which factors are most important for the classification. NDVI is often used as an input into a classification algorithm. NDVI range and maximum may also be used. While machine learning has the advantage of incorporating all available data, it also requires a bit more knowledge and expertise. Thus, I opted for testing whether the simpler method of thresholding could be applied.

To Identify whether thresholding is viable, it is helpful to have a time series of the rainfed and irrigated points. According to Ozdogan *et al*. (2010), the most common method for regional scale irrigation mapping is to use medium to coarse resolution data throughout the year to do a time-series analysis. He also states that this method requires knowledge of the growth pattern of crops grown in the area.

Given that planting dates may vary from year to year, it is preferable to have multiple images from at least two consecutive seasons. However, it is also possible to map irrigation using a single image so long as the peak greenness window can be identified. This can be done by using coarse spatial resolution data to identify the peak greenness window and then using high resolution data to do the mapping (Ozdogan *et al*., 2010).

Lastly, when creating a map, it is necessary to do an accuracy assessment at the end to produce a measure of uncertainty. This may be in the form of a confusion matrix which shows the number of sites that were correctly and incorrectly classified, or a measure of error, such as the Root Mean Squared Error (RMSE). With classification algorithms, 80% of the ground truthing points are used to train the model and 20% are reserved to test the accuracy of the model.

To create the time series for each satellite dataset, four data inputs were compiled: ground truthing points, precipitation data, ECOSTRESS images and Landsat 8 EVI 8-day composite images for two years May '19 – April '20 and May '20 – April '21. Google Earth Engine (GEE), ArcGIS Pro and Excel were the programs used.

I chose to evaluate ECOSTRESS due to the fact that two of the three crop classes in the LULC map (fruit trees and sugarcane) were perennial crops. While approaches using greenness-based indexes work well for annual crops which have a green up and senescence, we posited that it was better to have a measure more directly tied to water. The ECOSTRESS mission offers readily available mapped ET data. Landsat 8 EVI was also used to see how it compared with the ECOSTRESS results.

Creating the time series charts for ECOSTRESS involved downloading images and uploading them to Google Earth Engine as an Image Collection. It also required uploading the ground truthing points into GEE as Assets. While ArcGIS works fine for examining one image, GEE is preferable when looking at a collection of images over a season. Once I had created the ECOSTRESS Image Collection, I used code to extract values for each of the ground truthing points. This data was then exported into a CSV file and I used Excel to plot and analyze those values. While it is possible to create these charts within GEE, I wanted to examine the data closely and it was also easier to manipulate and plot the values in Excel. Daily precipitation was also included in these charts to help distinguish rain events versus irrigation events. This also allowed us to analyze the ET trends and EVI trends for irrigated vs. rainfed points and to see whether there was a "thresholding window". The final map was created by setting a threshold in ArcGIS Pro using Raster Calculator in Image Analyst Tools.

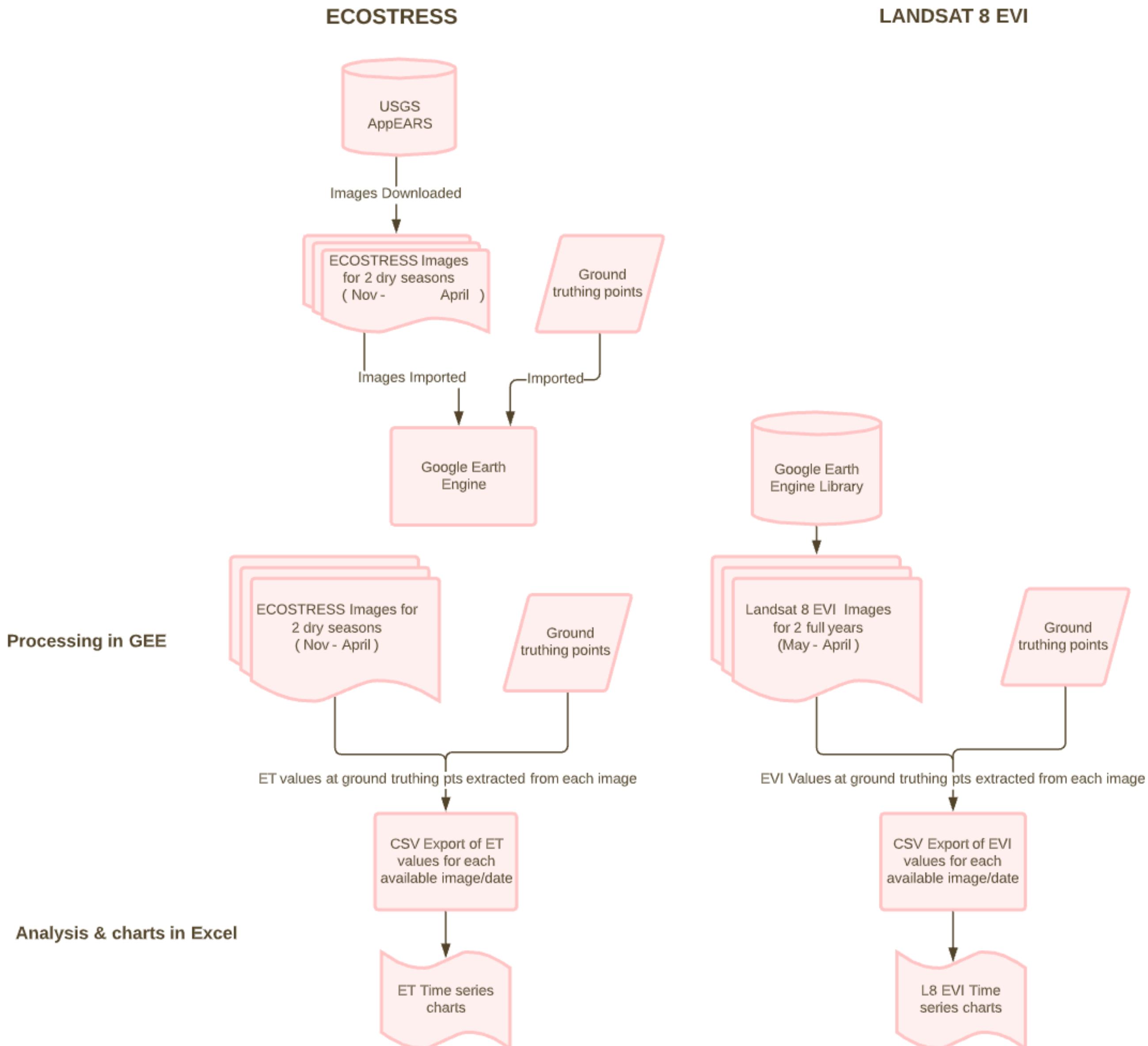


Figure 5: Flow-chart of time series creation process

Data

Land Use Land Cover Map

CRS, in collaboration with the International Tropical Agriculture Center (CIAT), developed a Land use/Land Cover (LULC) map of the Ahuachapán department using satellite imagery, ground truthing and field validation with drones to monitor changes in vegetative cover. Figure 6 shows the LULC map which has 15 classes: water, 4 types of forest (deciduous, gallery, mangrove, and evergreen), 6 crop classes (coffee, sugarcane, fruit trees, annual crops/mosaic of pasture and crops, palm trees, monoculture plantations), beaches, bare soil, urban, and secondary vegetation. CRS provided me with a file of the LULC map in vector/shapefile and raster/geotiff format. This made it possible to isolate the land cover class(es) of interest and ignore areas that had been classified as forest/non-agricultural land.

The map can be accessed at the following link: https://raicesahuachapan-crsorg.hub.arcgis.com/app/da27af3558714612bb53cb7ea104739c.

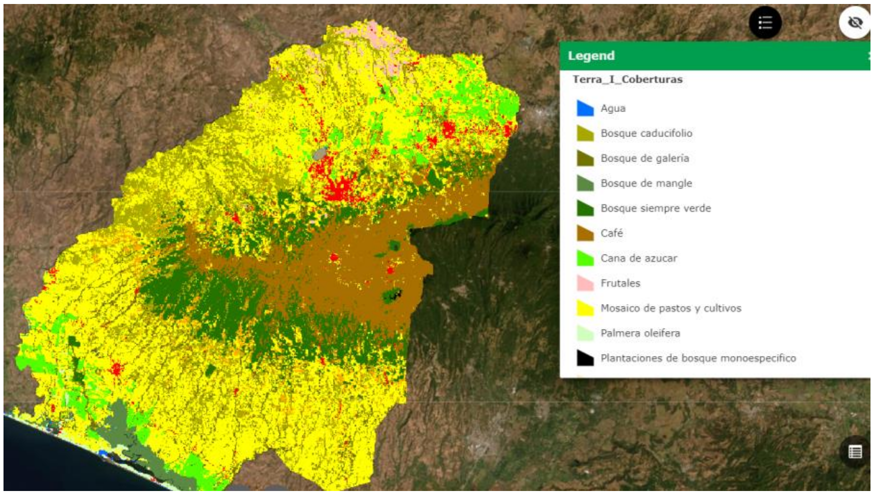


Figure 6: Land Use Land Cover Map developed by CRS and CIAT

Ground truthing points

Ground truthing points of irrigated and rainfed fields for sugarcane, fruit trees and annual crops were requested from CRS. However, due to time and resource constraints, new ground truthing points were not collected. Instead, CRS consolidated existing ground truthing points from farmers they work with (164 points) and from ground truthing points for another remote sensing project (46 points), into one data set for a total of 93 irrigated and 117 rainfed points. This made for a total of 210 points across 15 crop classes.

Irrigated points contained the following details: ID, Year, Crop(s) cultivated, Crop class, GPS point, Intervention site name, Canton, Municipio, Type of irrigation (sprinkler, drip, or furrow), and Date when the site was registered into the system. Data for the irrigated points was registered between September and October of 2020. The exact date when irrigation was installed and used on these plots is unknown, but we do know that areas identified as irrigated were using irrigation in the 2019-20 and 2020-21 dry seasons. Rainfed points had less details (GPS, Crop(s) cultivated, and Crop class) and we do not know when exactly this data was collected. Ground truthing point data was consolidated into an Excel document, uploaded into ArcGIS and converted to a shapefile. Once in shapefile format, it was uploaded into GEE as an Asset.

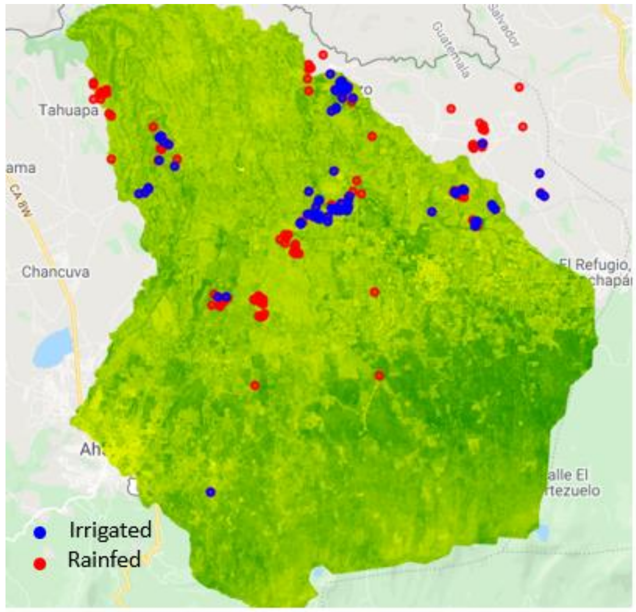


Figure 7: Ground truthing points plotted on mean EVI visualization of Agua Caliente watershed in Google Earth Engine.

While we had a seemingly even split of rainfed and irrigated ground truthing points, they were unevenly distributed across crop classes which made it hard to compare irrigated vs. rainfed for a particular crop class (Table 2, upper table). The Fruit trees class was the only crop class that had an even mix of irrigated vs. rainfed points. This is likely due to certain crops being more likely to be irrigated versus others. For example, there were 53 points of rainfed basic grains but only 5 irrigated basic grains fields. Conversely, horticultural crops are largely irrigated so there were 25 points of irrigated horticultural fields but only 1 rainfed.

Due to the uneven split between irrigated and rainfed, I decided to focus on crop classes that had 15 or more ground truthing points. Points from coffee areas were excluded from the analysis as it is a unique cropping system that is grown in higher elevation and sometimes shade grown, which creates limitations for remote sensing applications. This left us with 147 points across 5 crop classes. The 5 crop classes were: rainfed grains, irrigated horticulture, rainfed fruit trees, irrigated fruit trees, and irrigated Hort/Fruit trees which is a combination of horticulture and fruit trees. Table 2 shows the initial ground truthing points organized by crop class and irrigation status and 5 crops classes we chose to focus on.

*Table 2: Ground truthing points by Crop class and Irrigation Status**

Crop Class	Irrigated	Rainfed
Agroindustrial	3	5
Coffee	1	26
Coffee/Fruit trees	1	
Fruit trees	24	21
Fruit trees/Pasture	2	
Grains	4	53
Grains/Fruit trees	1	5
Grains/Hort	3	1
Grains/Hort/Fruit trees	4	
Grains/Pasture	2	
Hort	25	1
Hort/Fruit trees	17	2
Pasture	1	1
Sugarcane	4	2
Sugarcane/Hort/Fruit trees	1	
Grand Total	**93**	**117**

Crop Class	Irrigated	Rainfed
Fruit trees	24	21
Grains	4	53
Hort	25	1
Hort/Fruit trees	17	2
Grand Total	**70**	**77**

*Highlighted classes were selected for time series analysis

Table 3: Land Cover classes, Ground Truthing Crop classes, and Crops

Land Cover class	Ground Truthing Crop Class	Crops
Annual crops	Rainfed Grains	Maize, beans, sorghum
	Irrigated Horticulture	Elote, cassava, lemon, ayote, bell pepper, cucumber, radishes, pipian, tomatoes
Fruit trees	Rainfed Fruit trees	Jack bean*, guayaba, banana, cacao, jocote
	Irrigated Fruit trees	Jack bean*, guayaba, papaya, orange, lime, banana, mango, níspero, zapote
	Irrigated Hort/Fruit trees	Jack bean, ayote, bell pepper, elote, guayaba, lime, cucumber, pipian, banana, radishes, tomatoes, cassava, mango, papaya, orange

*Jack bean is used as a cover crop.

Daily Precipitation

CRS provided precipitation data from the three weather stations in the Agua Caliente watershed which they constructed. Data collection for these stations began on July 8, 2019. Precipitation in millimeters (mm) is shown in Table 4 for our study period. While there were some dates that had no precipitation data, these were generally reliable datasets. Table 5 shows the dates with missing precipitation data.

Table 4: Summary of precipitation data across 2 years

Station	Precipitation (mm) July 8, 2019 – April 30, 2020	Precipitation (mm) July 8, 2020 – April 30, 2021
El Jicaro	807	921.6
Raices	1040.4	1024.6
Rio Frio	814.8	645

Table 5: Missing precipitation data by station July 2019 - April 2021

Station	Missing Dates 2019	Missing Dates 2020	Missing Dates 2021
El Jicaro	Nov 12- 26 (15 days)	April 9-April 15 (7 days), October 26-Nov 2 (8 days)	No missing days from Jan- April
Raices	July 8	Feb 22- Mar 9 (17 days)	No missing days from Jan- April
Rio Frio	July 8-17 (10 days), Nov 8	Sept 9 - 30 (22 days)	No missing days from Jan- April

ECOSTRESS data

NASA's ECOsystem Spaceborne Thermal Radiometer Experiment on Space Station (ECOSTRESS) was launched on June 29, 2018 and began operations in August 20,2018. As indicated in its name, it is an instrument on the International Space Station (ISS), which has a temporal resolution of once a week, sometimes more frequently depending on the location (the revisit time varies). ECOSTRESS data has a 70-meter spatial resolution, which translates into half-hectare pixels. The ECOSTRESS mission measures plant temperature to understand water needs and plants' response to water stress. ECOSTRESS uses Land Surface Temperature and Emissivity to derive Evapotranspiration (ET). Other data products include Water Use Efficiency and the Evaporative Stress Index (ESI) (LP DAAC, 2021).

Landsat 8 Collection 1 Tier 1 8-Day Enhanced Vegetation Index (EVI) Composite

The Landsat 8 Collection 1 Tier 1 composites are one of many "derived datasets" in Google Earth Engine. It uses computed Top-of-Atmosphere (TOA) reflectance. The composites are said to consist of images/scenes from each 8-day period, however, images were 16 days apart. Spatial resolution for these images is 30 meters (Google Earth Engine, 2021).

Link to collection : https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LC08_C01_T1_8DAY_EVI

Depending on the vegetation index selected, it can be used to highlight healthy plants or to point out areas experiencing vegetation stress. In highly vegetated areas, NDVI tends to saturate easily. In this case, EVI is a good option as it is more sensitive to differences in highly vegetated areas. Additionally, it corrects for atmospheric conditions, canopy background noise, and solar incidence angle, allowing for easier comparison of different dates. EVI makes use of the near infrared band, the red band and the blue band.

$$EVI = 2.5 * \frac{(NIR - RED)}{(NIR + 6 * RED - 7.5 * BLUE + 1)}$$

(USGS, 2021)

EVI values range from -1 to 1 with healthy vegetation falling between 0.2 and 0.8.

RESULTS

ECOSTRESS

To create the ECOSTRESS time series, I downloaded ECOSTRESS L3_daily product (ET-PT-JPL) which is daily evapotranspiration for 2 dry seasons (Nov- April '19-'20 and '20-'21) from the United States Geological Survey (USGS) Application for Extracting and Exploring Analysis Ready Samples (*App*EARS). They have 4 levels of data products. The level indicates the amount of processing required to produce it. The L3_daily product (ET-PT-JPL) is in Watts per meter squared (W m^{-2}). Other Level 3 products which are available but were not used were: ET instantaneous (W m^{-2}) and ET components (% ET soil, % ET canopy, % Et interception) . There are also Level 4 products which include the Evaporative Stress Index and Water Use Efficiency whose values range from 0 to 1.

Upon downloading the data, I discovered ECOSTRESS images for my region were not available weekly. I downloaded an additional season to see how image availability compared over 3 dry seasons. Below are the dates of available images for each of the 3 dry seasons:

- 2018-2019: 10 images available
    - 1/12/2019, 1/22/2019, 1/26/2019, 2/02/2019, 2/25/2019, 2/28/2019, 3/04/2019, 3/07/2019, 3/11/2019
- 2019-2020: 9 images available
    - 11/21/2019, 11/25/2019, 11/29/2019, 12/07/2019, 12/11/2019, 03/20/2020, 03/24/2020, 04/10/2020, 04/18/2020
- 2020-2021: 8 images available
    - 1/28/2021, 2/09/2021, 2/13/2021, 3/21/2021, 3/25/2021, 3/29/2021, 4/02/2021, 4/06/2021

There did not seem to be any consistency in when sites were revisited. In some years, there were no images in the first two months; in one year there was a 3-month gap between images.

The availability of an image on a given date did not guarantee an ET measurement as some areas may have been covered with clouds and masked (Figure 8).

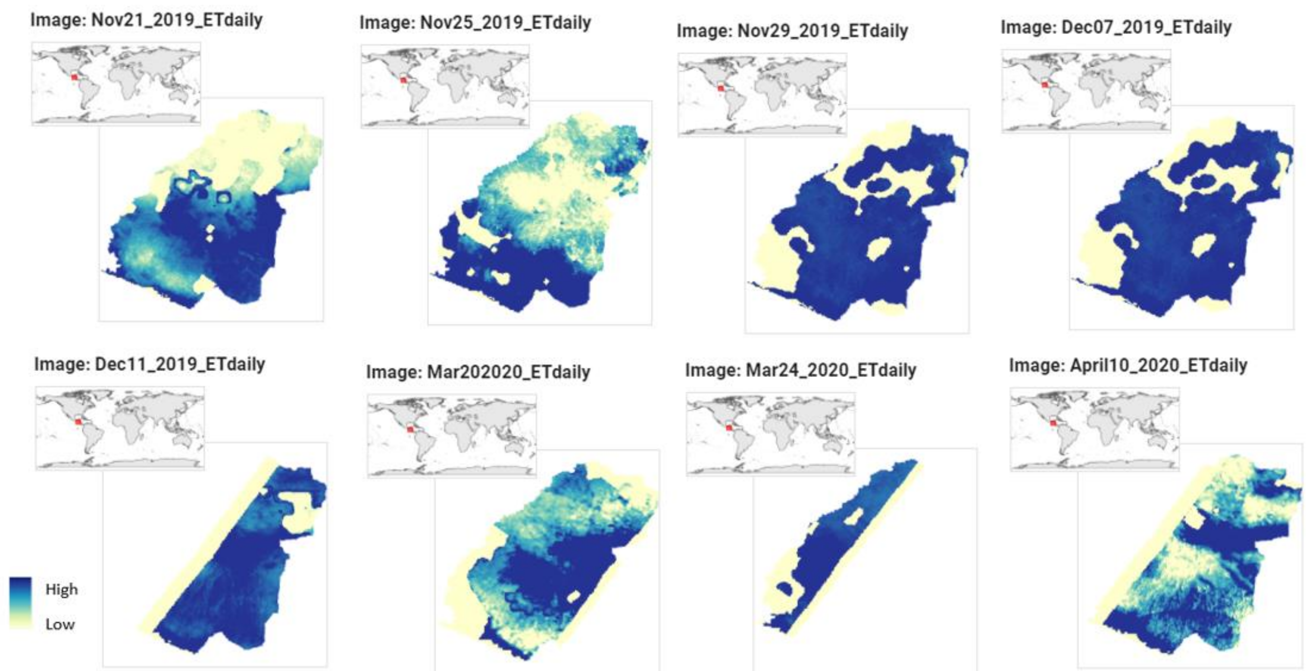


Figure 8: Collection of ECOSTRESS images from '19-'20 dry season visualized in GEE

I uploaded the ECOSTRESS images into Google Earth Engine as an Image Collection. I used code adapted from Spatial Thoughts (https://spatialthoughts.com/2020/04/13/extracting-time-series-ee/), a geospatial science blog, to extract the ET value for each of the ground truthing points from each image and then exported these values into an Excel document. Once in Excel, I could more easily manipulate the values in a pivot table and plot values by crop class and/or irrigation status. ET values were converted from W m^{-2} to mm/day by dividing them by 28.4 as 1 mm per day = 28.4 W m^{-2}. This made it easier to plot the ET values with daily precipitation.

Below are links to the code I developed:

ECOSTRESS Code: https://code.earthengine.google.com/196cce9a67ab736ee43376231979f703

Landsat 8 Code: https://code.earthengine.google.com/3c333a7b4642c3bfd9d0881ea4ba7992

Figure 9 is the time series for ECOSTRESS Daily ET from November 1, 2019 – April 30, 2020 comparing irrigated and rainfed sites across all crop classes. Available precipitation data (data collection began July 8, 2019) is also plotted alongside ET values. As can be seen, aside from a date in April, the points often overlap, indicating that we cannot apply a threshold ET to distinguish rainfed and irrigated fields using ECOSTRESS images. There are two reasons we believe this was the case. One is that ECOSTRESS's 70-meter resolution may be too coarse to capture smallholder irrigated plots which are sometimes as low as 0.35 hectares (Table 1). Another reason may be due to some ground truthing points being too close to each other and falling within the same pixel. When this happens an irrigated and rainfed site will have equal ET values.

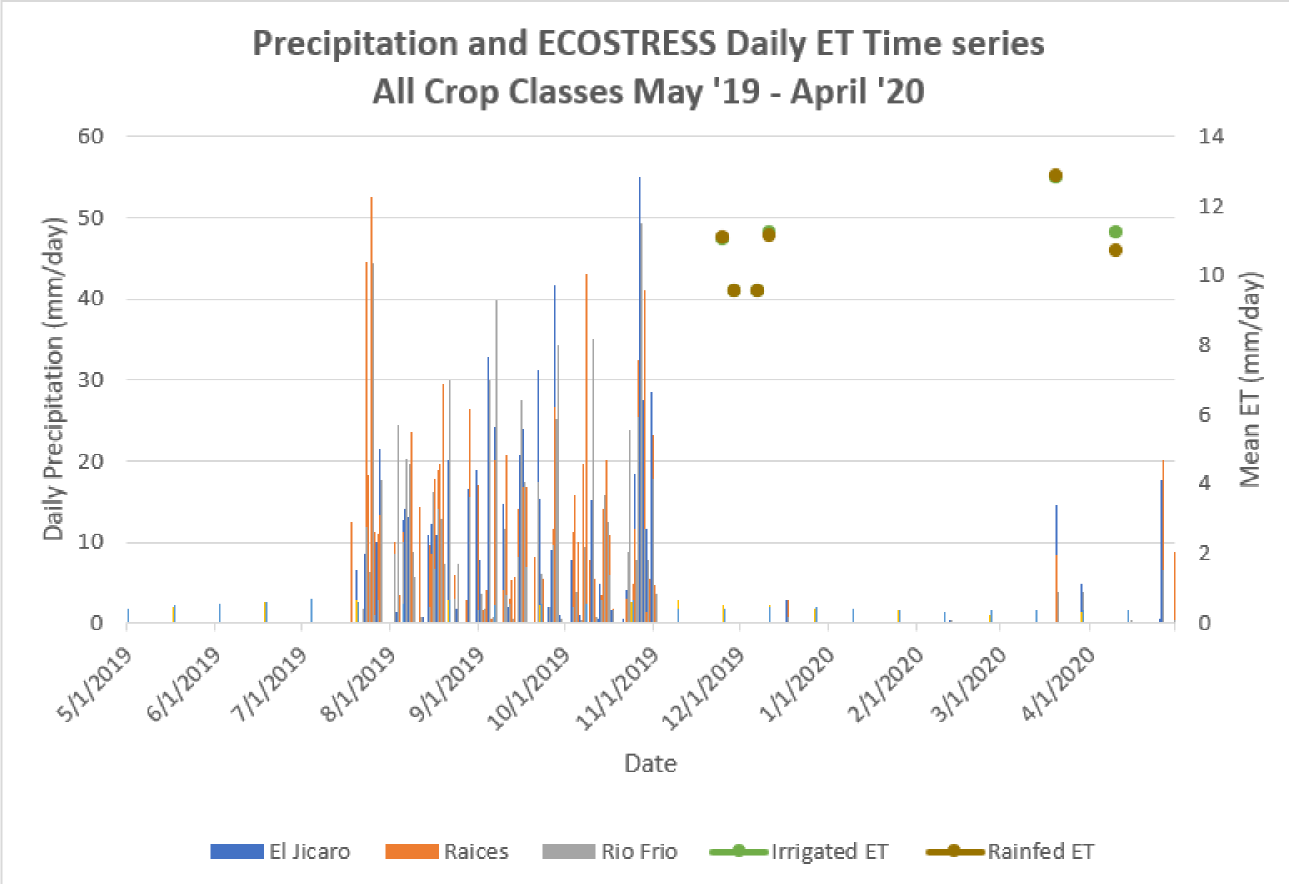


Figure 9: Time series for ECOSTRESS Daily Irrigated vs. Rainfed ET

The next step was to see if we could identify irrigated areas when looking at specific crop classes. I considered fruit trees as they had a large number of both irrigated (24) and rainfed (21) sites, unlike other crop classes. Figure 10 is the time series for ECOSTRESS Daily ET for irrigated and rainfed fruit trees. As can be seen, even when specifying crop class, values for irrigated and rainfed fruit trees still overlap.

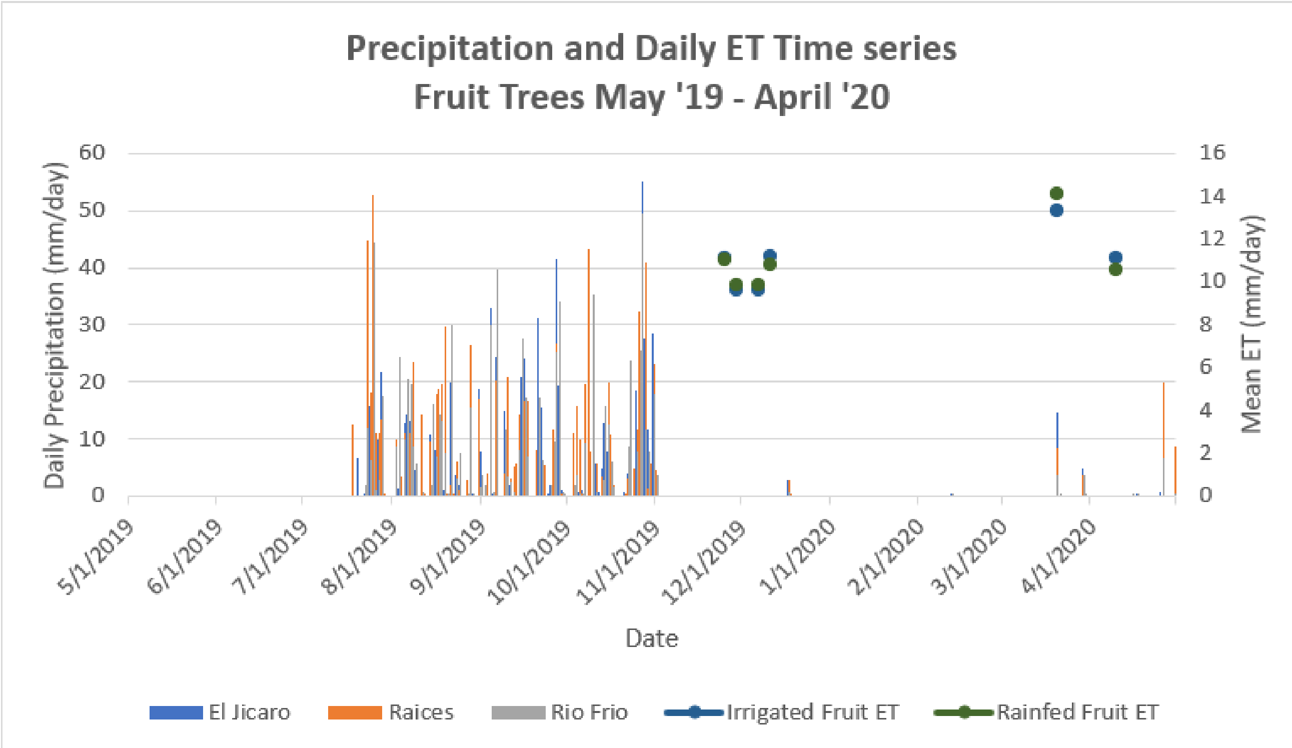


Figure 10: Time series for ECOSTRESS Daily Irrigated vs. Rainfed ET for Fruit Trees

<u>Landsat 8 EVI</u>

While ECOSTRESS evapotranspiration data provides a more direct measure of water applied, it seems to lack the spatial and temporal resolution required for identifying smallholder irrigation in our study region. Studies have shown greenness indexes can be used to identify irrigation in a variety of environments (Ozdogan *et al*., 2010) so I decided to compare the ECOSTRESS results to Landsat 8 which has an EVI 8-day composite product readily available in Google Earth Engine. While the temporal resolution was actually 16 days, not 8, it offered more images/data points compared to ECOSTRESS. The Landsat 8 EVI image collection was imported using code and the same code that was used to extract ET values for each of the ground truthing sites was adapted to extract EVI values. The values for 2 full years (May ’19 – April ’20 and May ’20 – April ’21) were then exported into an Excel document. The EVI time series were plotted with daily precipitation.

For Landsat EVI analysis, I chose to focus on the 5 crop classes that had 15 or more ground truthing points (those highlighted in gold in Table 2). The 5 crop classes were: rainfed grains, irrigated horticulture, rainfed fruit trees, irrigated fruit trees, and irrigated Hort/Fruit trees. I chose to organize those 5 crop classes according to the land cover classes in the LULC map. Rainfed grains, irrigated horticulture, and irrigated fruit trees/horticulture were grouped into Annual Crops and irrigated and rainfed fruit trees were grouped into Fruit Trees (Table 3). The availability of the LULC meant we could potentially identify thresholds for irrigated areas within each of these land cover classes. We expected that the irrigated crop classes would have a higher EVI, reflecting healthier vegetation and less stress compared to rainfed crop classes.

There was uncertainty over whether to sort irrigated fruit/horticulture with annual crops or with fruit trees. Determining where it would fit best would require having a better understanding of what these fields look like and whether they are majority horticultural crops or majority fruit trees. My guess was that there were trees interspersed with horticulture so I treated this category as an annual crop.

Figure 11 shows the time series for the 3 annual crop classes for May '19 - April '20. This time period captures the rainy season (May - October) and the dry season (November - April). Aside from a few dates when the rainfed grains EVI was higher than that of irrigated horticulture, the irrigated crop class generally has a higher EVI, as expected. However, there does not seem to be a clear trend between the crop classes until later in the dry season. The EVI values deviate from each other starting in late February. The date with the greatest deviation is February 26, 2020. This implies that identification of irrigation activity may be most successful using images from late February to April.

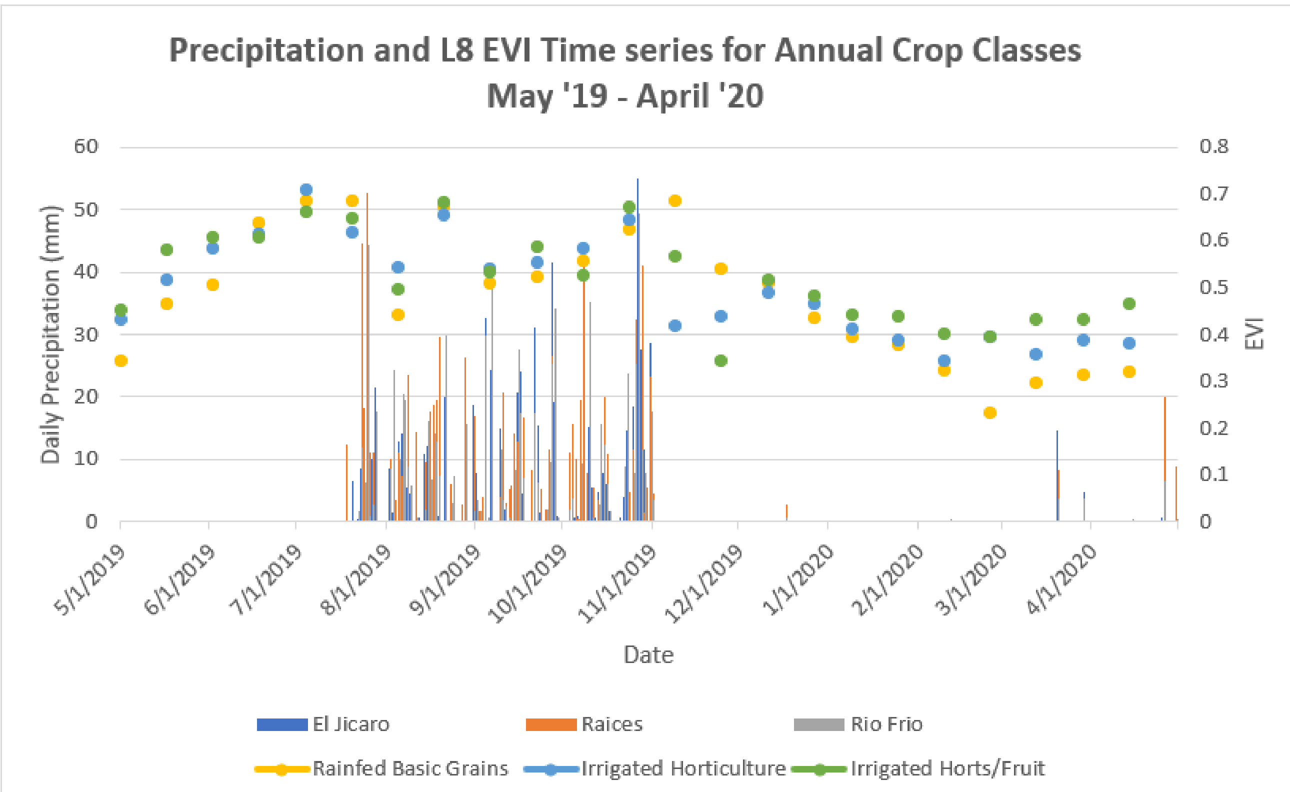


Figure 11: Time series for Landsat 8 EVI Irrigated vs. Rainfed ET for Annual Crop Classes '19 - '20

The annual crop classes for the '20-'21 season show the same pattern (Figure 12). The time series points start to deviate from each other in the late dry season, late February to April. In both years we see that the peak greenness periods occur around July and November. However, the dry season values in both charts show a consistent trend between crop classes whereas rainy season values do not exhibit a trend. Thus, the dry season period would be a better window for thresholding.

Another thing to note is that while the EVI values separate in the late dry season (Irrigated horticulture/Fruit consistently has the highest EVI followed by Irrigated Horticulture and then Rainfed grains), all the EVI values shift up in '20 - '21. Thus, a threshold chosen based on one year's data may fail to capture irrigation occurring in other years.

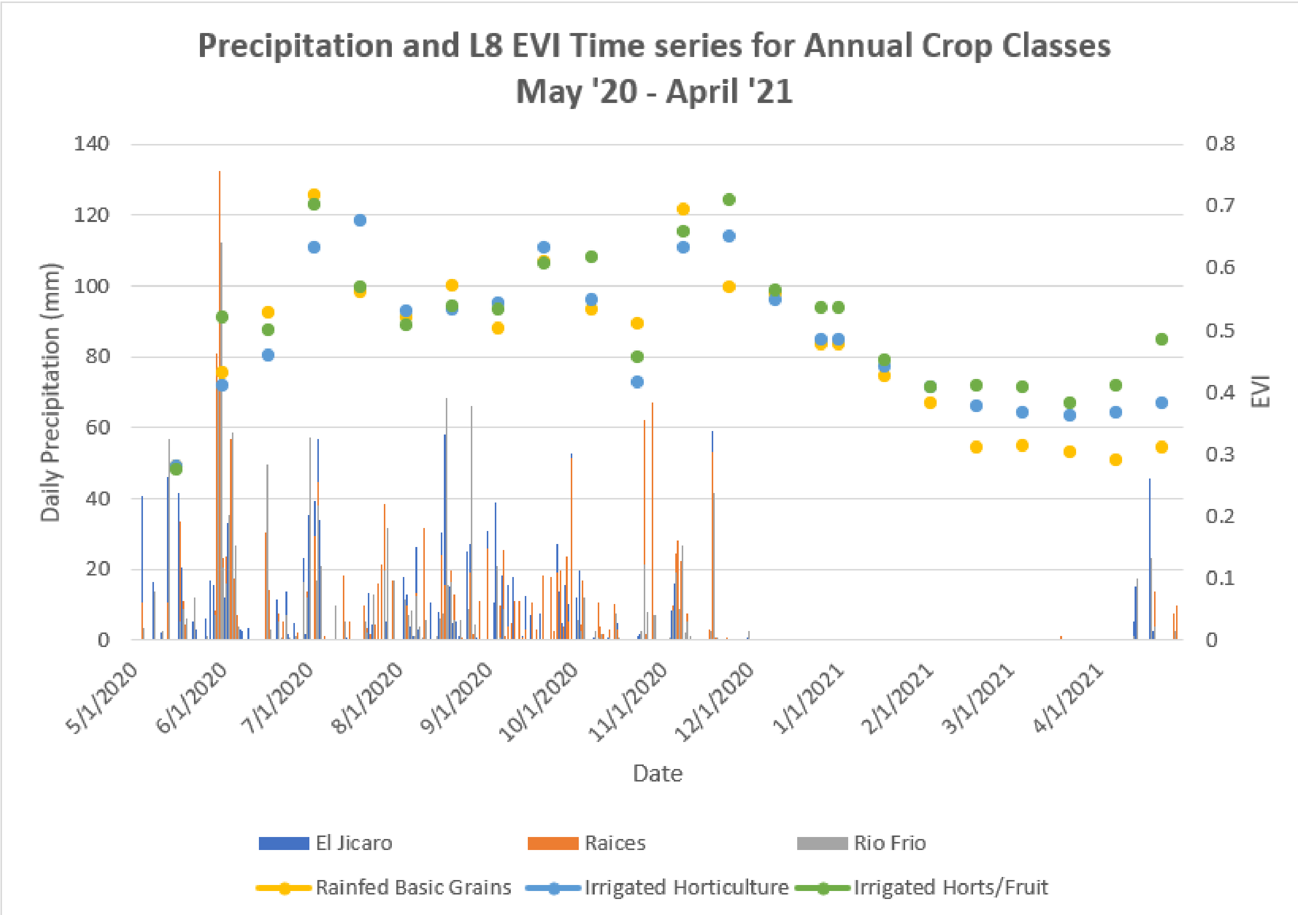


Figure 12: Time series for Landsat 8 EVI Irrigated vs. Rainfed for Annual Crop Classes '20 -'21

When we look at the time series for the fruit tree classes for '19 -'20 in Figure 13, we see no consistent trend between irrigated vs. rainfed crop classes in the rainy season. In the dry season, EVI values for the irrigated fruit tree class are generally higher than that of rainfed EVI. However, the values are very close together, at times overlapping with each other. The overlapping occurs even more in the '20 - '21 dry season (Figure 14) and there are instances where the rainfed fruit EVI values are a bit higher than irrigated fruit EVI values. We see the same pattern in the '20 – '21 season. Aside from the February 26, 2020 data point, the irrigated and rainfed fruit tree EVI values do not deviate from each enough to make thresholding a viable approach for estimating irrigated fruit tree areas.

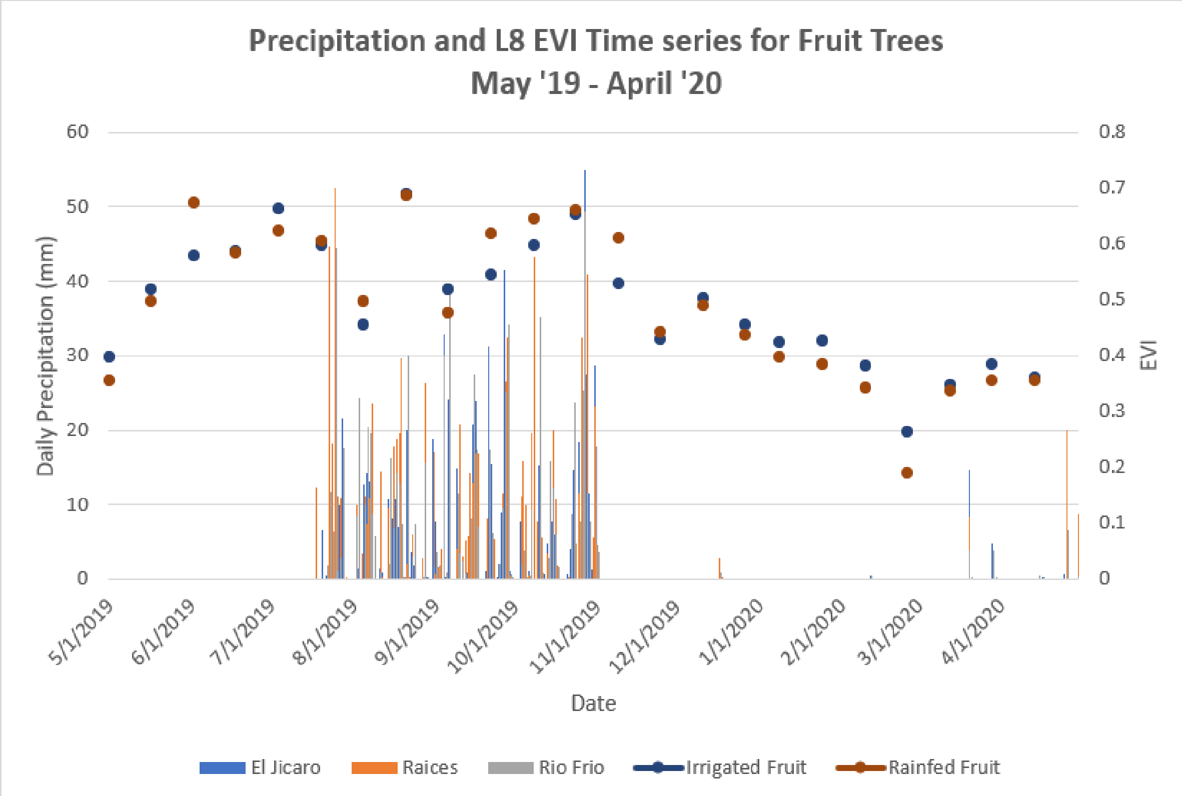


Figure 13: Time series for Landsat 8 EVI Irrigated vs. Rainfed for Fruit Trees '19 -'20

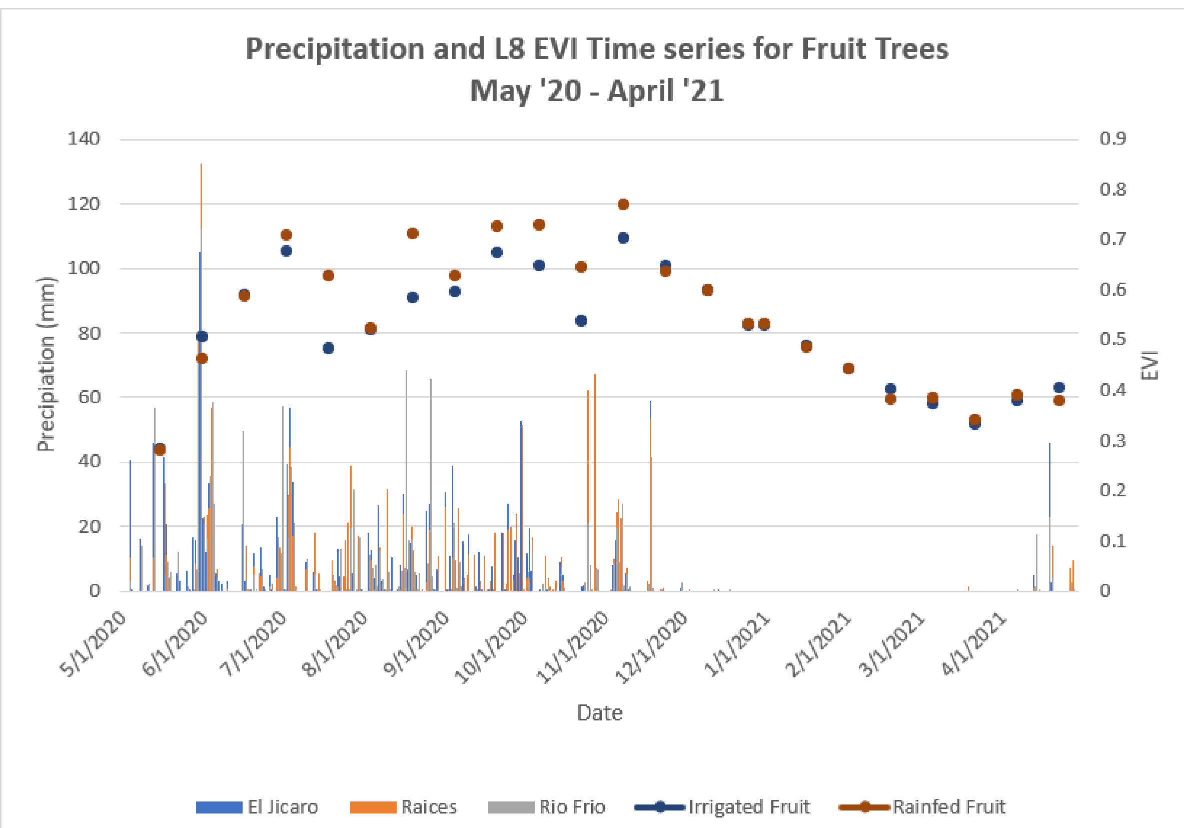


Figure 14: Time series for Landsat 8 EVI Irrigated vs. Rainfed for Fruit Trees '20 -'21

The date where crop class EVI values had the greatest difference was February 26, 2020. Those points represent averages of each of the classes. Plotting the distribution of the values between classes though a boxplot allows us to further assess whether we could apply a threshold to classify irrigated areas. Figure 15 shows the EVI distribution on February 26, 2020. We can see that the 75th percentile of rainfed grains is around EVI = 0.27 and that the majority of irrigated sites have an EVI above that so this appears to be a good choice for a threshold but we do not know the accuracy of choosing such a threshold.

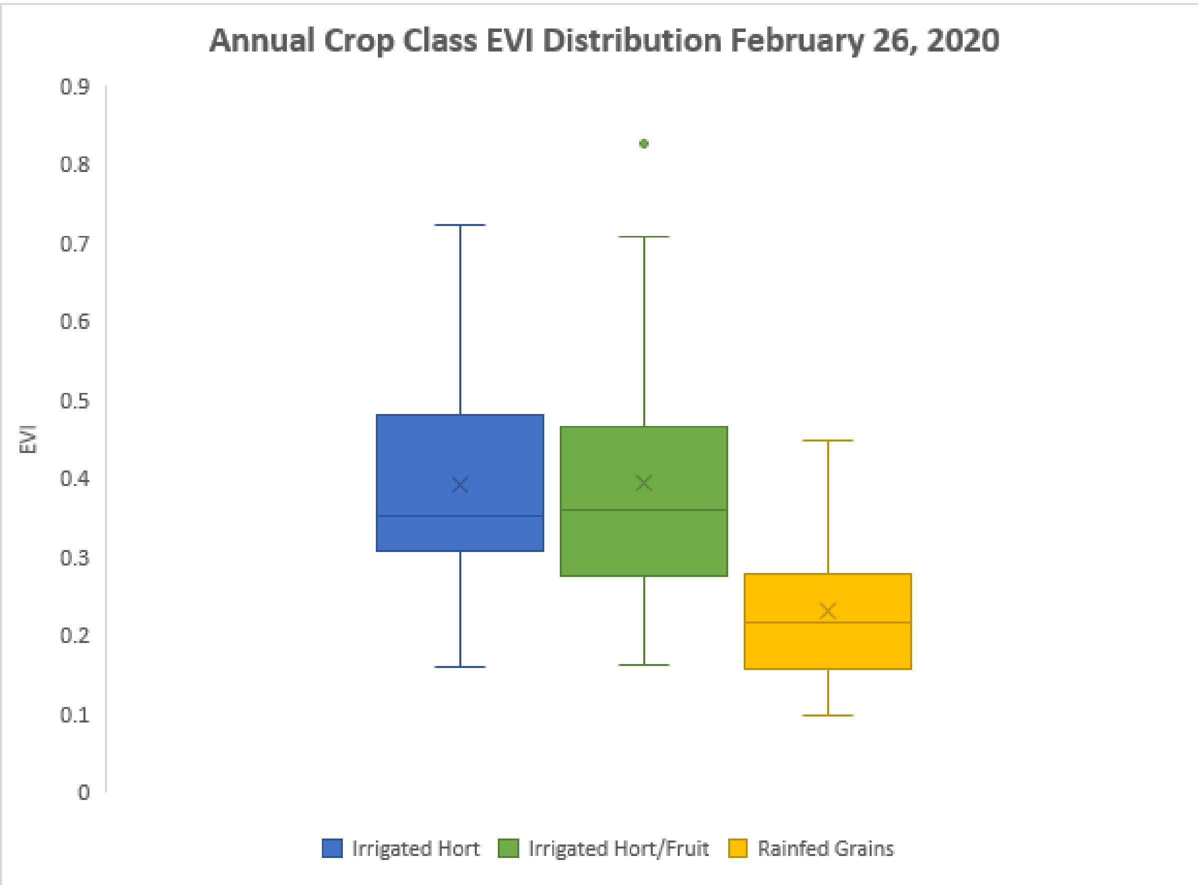


Figure 15: Box plot of EVI distribution for annual crop class

In order to get a measure of accuracy, we need to construct a confusion matrix. A confusion matrix is used to evaluate the number/percentage of correctly and incorrectly classified cases. It is generally used for a classification algorithm but it can easily be applied here. It is generally created with a subset of ground truthing points (those that are not used in training the model), but in this case we used all ground truthing points to assess accuracy.

Table 6 shows the confusion matrix for this boxplot. “Irrigated and Rainfed GT Sites” refers to the actual status of the ground truthing sites and “Classified Irrigated and Rainfed” refers to how the sites would be classified based on the threshold we select. If we apply a threshold of EVI $\geq$ 0.27 (the 75th percentile of rainfed grains) all points with EVI lower than 0.27 that would be classified as rainfed and all points with EVI greater than or equal to 0.27 would be classified as irrigated. The confusion matrix in Table 6 shows this threshold would result in 86% (36/42) of all irrigated sites and 72% (38/45) of all rainfed sites correctly classified. The cells highlighted in green are those that would be correctly classified while those highlighted in red are those would be incorrectly classified. Accuracy is then calculated by summing the number of correctly classified irrigated and rainfed sites and dividing by the total number of sites. Thus, a threshold of EVI $\geq$ 0.27 would result in an accuracy of 77% (36 correctly classified irrigated sites + 38 correctly classified rainfed sites/ 96 total sites).

Table 6: Annual Crops Confusion Matrix

Annual Crops Confusion Matrix (Threshold EVI $\geq$ 0.27)		
	Classified Irrigated	Classified Rainfed
Irrigated GT Sites (42)	86% (36)	17% (7)
Rainfed GT Sites (53)	28% (15)	72% (38)
	51	45

When we create a boxplot to examine the distribution for fruit tree EVI values (Figure 16), we see the interquartile ranges overlap much more so choosing a threshold of EVI=0.22 (the 75th percentile of rainfed fruit) would likely yield a lower accuracy.

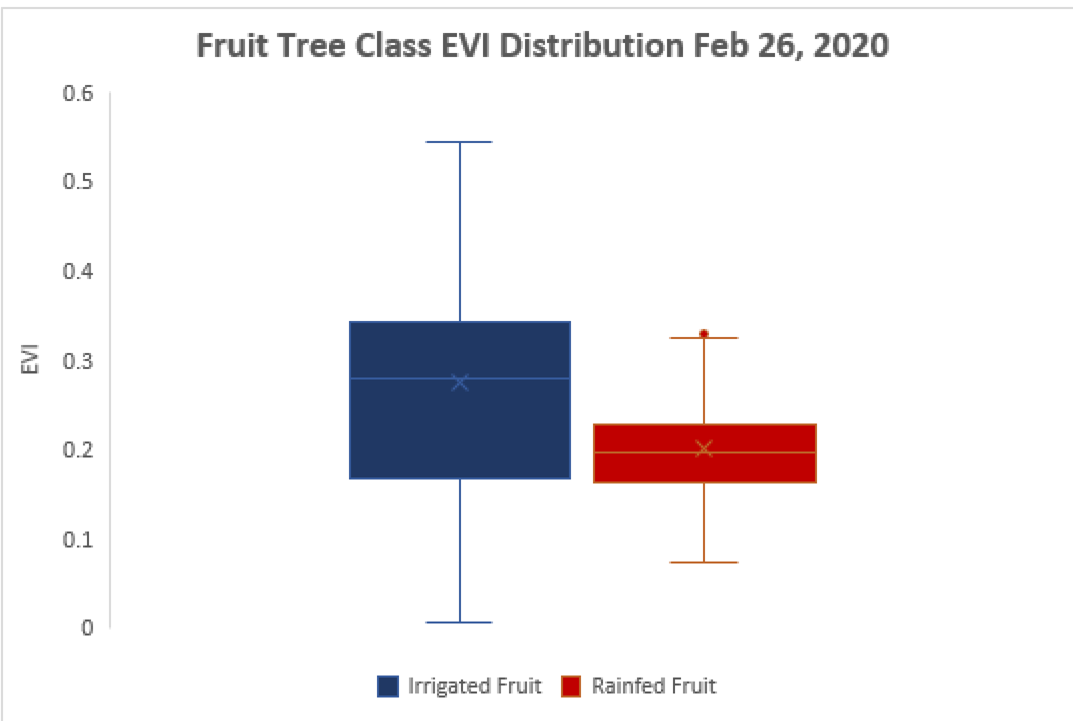


Figure 16: Box plot of EVI distribution for fruit tree crop class

Table 7 shows the Fruit Tree Confusion Matrix. Applying a threshold of EVI$\geq$ 0.22 would result in 54% (13/24) of irrigated fruit tree sites and 67% (14/21) of rainfed fruit tree sites correctly classified. Thus, our accuracy is 60 % (13 correctly classified irrigated sites + 14 correctly classified rainfed sites/ 45 total sites). As we concluded earlier, thresholding would not be a good option for estimating irrigated fruit tree area.

Table 7: Fruit Tree Confusion Matrix

Fruit Tree Confusion Matrix (Threshold EVI $\geq$ 0.22)		
	Classified Irrigated	Classified Rainfed
Irrigated Fruit GT Sites (24)	54% (13)	46% (11)
Rainfed Fruit GT Sites (21)	33% (7)	67% (14)
	20	25

The last step was to create a map based on the annual crop class threshold of EVI $\geq$ 0.27. The map in Figure 17 was made by combining the EVI band from Landsat 8 EVI Composite from February 26, 2020 with the annual crops' raster from the LULC map. Areas that were classified as annual crops and that had an EVI $\geq$ 0.27 were classified as irrigated. 61,368 30-meter pixels were classified as irrigated. This translates into an estimated irrigated area of 5,523 hectares. This is considerably higher than the total irrigated area from the irrigation permit data, which was only 329.5 hectares and was for the entire department of Ahuachapán. Areas south of the main highway "Carretera Principal" are underrepresented in the ground truthing points so it would help to collect points from those areas to inform the accuracy assessment.

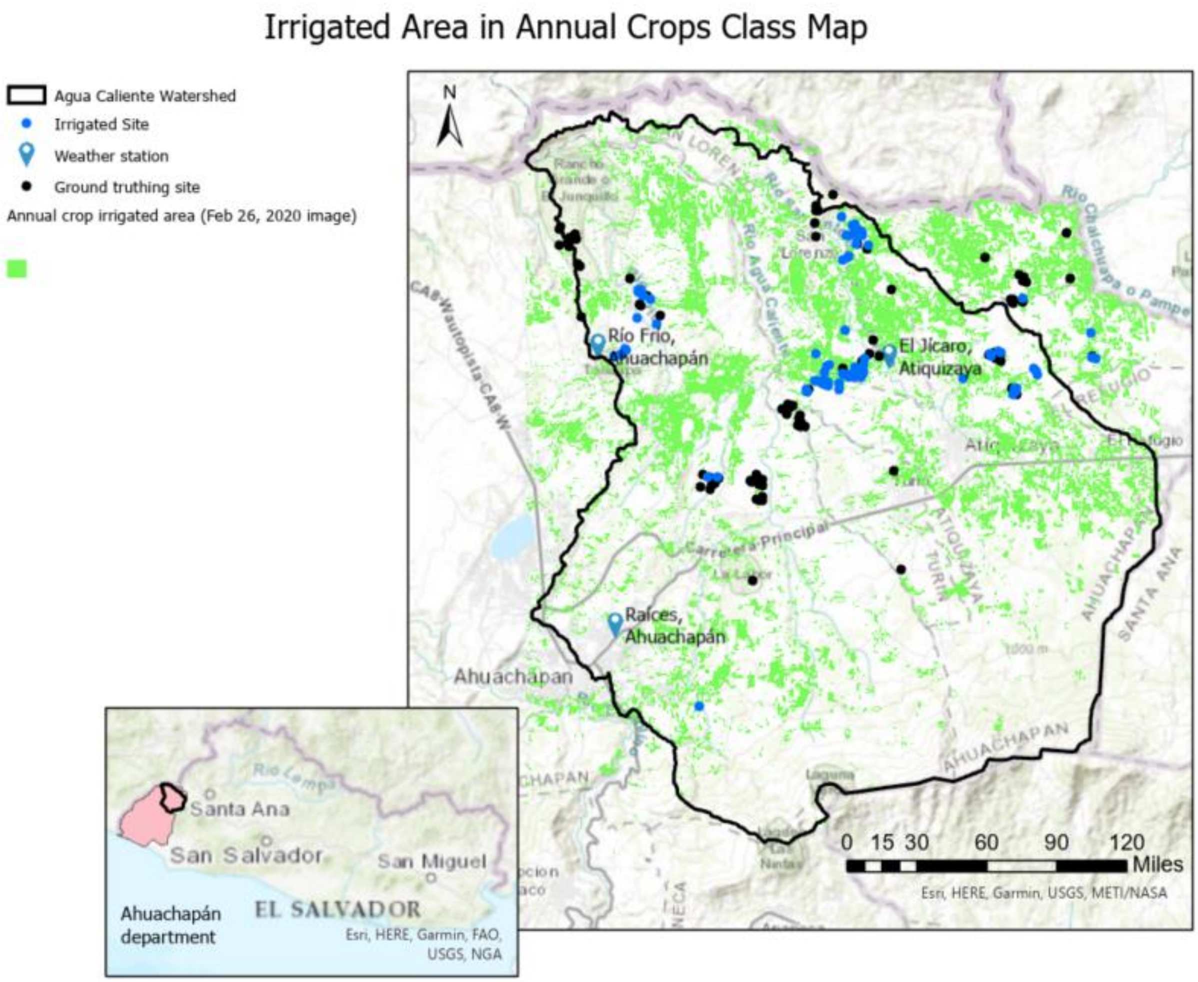


Figure 17: Irrigated Area in Annual Crops Class Map

DISCUSSION

Our objective of mapping irrigated areas in the Agua Caliente watershed was achieved with a few exceptions. We determined that applying a threshold of EVI $\geq$ 0.27 to areas classified as annual crops yields a 77% accuracy. We also produced a map of irrigated annual crop areas in the Agua Caliente watershed that can guide future ground truthing point acquisition. The limitation of this approach is that the threshold must be chosen for a given dry season and should not be blindly applied across different years. While our methodology has not been fully automated, the GEE code for Landsat 8 EVI can be adapted to new time periods and new CRS expansion areas.

Additionally, we consolidated available information on irrigation trends in Ahuachapán and provided recommendations for future irrigation mapping efforts which is a considerable step forward from where this project began. Hopefully future efforts to gain access to the irrigation permit data will be successful as this would help in assessing accuracy.

While there is promise in the ECOSTRESS data products for irrigation mapping and water resource management, given the inconsistent image availability and relatively coarse resolution, it may be a little early to use these products for identifying smallholder irrigation in El Salvador. While we had high hopes for the ECOSTRESS data, the temporal resolution along with our clustered ground truthing points, made it difficult to obtain any useful insights. Landsat EVI data proved more helpful in selecting a threshold for mapping.

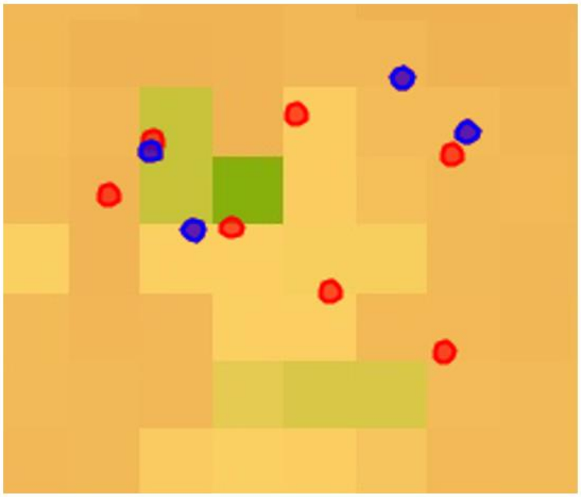
Some of the major challenges included working with few ECOSTRESS data points and becoming familiarized with satellite datasets enough to understand some of their limitations. We believe that ECOSTRESS low temporal resolution is due to ECOSTRESS focusing on certain areas more than others. For example, it acquires images over the U.S., key biomes, European and South Asian agricultural areas and certain FLUXNET validation sites (LP DAAC, 2021). Thus, it may take a few years for ECOSTRESS to provide more consistent images of certain regions.

Another limitation is that the GEE Landsat derived datasets use Top-of-Atmosphere images versus Bottom-of-Atmosphere images (Google Earth Engine, 2021). Thus, these images are not atmospherically corrected and thus may not have a cloud mask. This may explain why rainy season EVI values (when we would expect more clouds) do not have a consistent trend compared to dry season EVI values. It may be preferable to use Landsat 8 Surface Reflectance product (which has been atmospherically corrected), apply a cloud mask to crop out areas blocked by clouds, and then calculate EVI.

Recommendations

Below are a number of recommendations that might help future irrigation mapping projects based on what I learned in my project.

1. *Provide polygon data for the irrigated and rainfed fields versus point data and/or ensure there is only 1 site per field or per farm.* If it must be point data, ensure GPS points are taken at the center of fields versus at edges so that it is clear which field the GPS point corresponds to. As can be seen in Figure 18, there were instances where there were irrigated and rainfed ground truthing points within the same plot/field.



Ground truthing points over ECOSTRESS Min ESI '19-'20 visualization in GEE

Ground truthing points over natural color satellite imagery

– Irrigated - Rainfed

Figure 18: Close up of ground truthing points with ECOSTRESS data and with high resolution natural color imagery

2. *Ensure that points are approximately a half hectare apart so that irrigated and rainfed ground truthing points do not fall within the same pixel.* ECOSTRESS's 70 m spatial resolution means the pixels are approximately a half hectare (4,900 m^2). Figure 18 shows instances when an irrigated and rainfed point fell within the same pixel or appeared to belong to the same field. This may have been a factor in the overlap we saw between irrigated and rainfed ET in the ECOSTRESS time series. The 70 m spatial resolution may be too coarse to distinguish between smallholder plots.

3. *Provide a greater number of ground truthing sites/points, including from non-CRS farms, to get a more representative sample of irrigated areas in the landscape.*

4. *Compare results from thresholding approach to a Random Forest supervised classification if it is possible to improve ground truthing data. This will mean collecting polygons versus points, collecting more ground truthing points, and increasing the number of sites from underrepresented classes (i.e. irrigated grains, rainfed horticulture).* It is important to have sites for irrigated and rainfed fields for a given agricultural class. Use all available Landsat bands as input data to train the model. The model will then determine which factors are most important for the classification. The advantage to this method is that it can easily be applied to new areas and time periods and will provide a confusion matrix to assess accuracy.

5. *Provide a seasonal calendar for crops in the region that details the time period around when planting and harvesting for the crop classes occurs and an irrigated cropping calendar that shows what crops are irrigated and when this generally occurs.* This would provide valuable

information on the crop phenology of annual irrigated crops and can help with time series analysis.

6. *For future remote sensing projects, take new ground truthing points early on or have that data readily available so that project proposals can be structured around the existing ground truthing data. Also, if possible, include CIAT staff in project visioning as they have remote sensing expertise in El Salvador.*

7. *Strengthen collaboration and help researchers in contacting the Ministry of Agriculture. The Ministry irrigation permit data could be used to help validate maps of irrigated areas and/or to choose future ground truthing sites*. Sampling could be targeted, for example, to areas around rivers. Based on my understanding from the 2015 irrigation permit data, rivers are the most common water source. As was done in the study in the Limpopo province of South Africa (Cai *et al*. 2015), transect lines could be adapted to distribution of irrigated areas.

8. *Review Annex 1. Technical Guide on Open Data Kit Field Survey in Cai et al. 2015*. This section details recommendations on taking ground truthing (GT) points and has a GT form to collect GT point data which can be adapted.

9. *Include socioeconomic as well as biophysical factors in determining irrigation expansion areas.* Looking forward, any irrigation expansion efforts should determine what socioeconomic variables are most important in successful adoption of irrigation. This data can be collected through farmer surveys and/or interviews. Both of the studies in South Africa and Burkina Faso included non-biophysical factors in their data sets. The Cai *et al*. study in South Africa conducted a survey that covered demographics, economics, farming, marketing, crops grown, challenges and water source. The Chaix-Bar *et al.* study in Burkina Faso compared the remote sensing product to the Targeting Agricultural Water Management Interventions (TAGMI) result which, despite its shortcomings, included socioeconomic factors.

10. *Ensure irrigation expansion does not come at expense of community water needs by evaluating whether or not potential new areas have a threatened water supply.*

REFERENCES

Cai, X., Magidi, J., Nhamo, L., & Koppen, B. (2017). *Mapping Irrigated Areas in the Limpopo Province, South Africa*. IWMI Working Paper - 172. International Water Management Institute (IWMI). 37p. https://doi.org/10.5337/2017.205

Cai, X., & Rosegrant, M. W. (2002). Global Water Demand and Supply Projections: Part 1. A Modeling Approach. *Water International*, *27* (2), 159-169. https://doi.org/10.1080/02508060208686989

Chaix-Bar, M., Barron, J., Cinderby, S., Cambridge, H., & Morris, J. (2019). *Small-Scale Irrigation Mapping (SSIM) as a tool for improving and validating irrigated area maps: Contextual approach and lessons learnt in Burkina Faso*. CGIAR Research Program on Water, Land and Ecosystems (WLE). 8p. https://cgspace.cgiar.org/bitstream/handle/10568/107801/SSIM_Technical%20Report_Burkina%20Faso%20January2019.pdf?sequence=1&isAllowed=y

de Fraiture, C., & Giordano, M. (2014). Small private irrigation: A thriving but overlooked sector. *Agricultural Water Management*, *131*, 167–174. https://doi.org/10.1016/j.agwat.2013.07.005

Dorrance, J., Hauber, M., Hurley, M., & Khananusit, S. (2011). *Impact of Climate Change on Rural Salvadorian Farmers*. [Unpublished Capstone]. George Washington University. https://elliott.gwu.edu/sites/g/files/zaxdzs2141/f/downloads/acad/ids/climate-change-el-salvador.pdf

EFE Noticias. (2016, May 12). La vida en El Salvador será inviable en 80 años por la crisis de agua, dice un estudio. Retrieved July 23, 2021 from https://www.efe.com/efe/america/sociedad/la-vida-enel-salvador-sera-inviable-80-anos-por-crisis-de-agua-dice-un-estudio/20000013-2923438

FAO (Food and Agriculture Organization of the United Nations). (2021). El Salvador. World Agriculture Watch. Retrieved July 23, 2021, from http://www.fao.org/world-agriculture-watch/our-program/slv/en/

FAO (Food and Agriculture Organization of the United Nations). (2021). Global Maps of irrigated areas - El Savlador. AQUASTAT - FAO's Global Information System on Water and Agriculture. Retrieved July 23, 2021 from http://www.fao.org/aquastat/en/geospatial-information/global-maps-irrigated-areas/irrigation-by-country/country/SLV

FEWS NET (Famine Early Warning Systems Network). (2013, December). Central America and Caribbean—Seasonal Calendar. Retrieved July 23, 2021, from https://fews.net/central-america-and-caribbean/seasonal-calendar/december-2013

GEE (Google Earth Engine). (2021). Landsat 8 Collection 1 Tier 1 8-Day EVI Composite. Earth Engine Data Catalog. Retrieved June 2, 2021 from https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LC08_C01_T1_8DAY_EVI

Johansson, Robert C. (2005). Micro and Macro-Level Approaches for Assessing the Value of Irrigation Water. *Policy Research Working Paper*, *No. 3778*. World Bank. https://openknowledge.worldbank.org/handle/10986/8471

Lakhani, Nina. (2019, July 30). Living without Water: The Crisis Pushing People out of El Salvador. *The Guardian*. https://www.theguardian.com/globaldevelopment/2019/jul/30/el-salvador-water-crisis-privatization-gangs-corruption

LP DAAC (Land Processes Distributed Active Archive Center). (2021). ECOSTRESS Overview. Retrieved July 25, 2021. https://lpdaac.usgs.gov/data/get-started-data/collection-overview/missions/ecostress-overview/

Meier, J., Zabel, F., & Mauser, W. (2018). A global approach to estimate irrigated areas – a comparison between different data and statistics. *Hydrology and Earth System Sciences*, *22*(2), 1119–1133. https://doi.org/10.5194/hess-22-1119-2018

McKnight, T. L. & Hess, D. (2000). *Physical geography: A landscape appreciation*. Upper Saddle River, N.J. : Prentice Hall. http://archive.org/details/physicalgeographmckn

Ozdogan, M., Cervantes, C., Yang, Y., & Allez, G. (2010). Remote Sensing of Irrigated Agriculture: Opportunities and Challenges. *Remote Sensing*, *2*(9), 2274–2304. https://doi.org/10.3390/rs2092274

Thenkabail, P., Biradar, C., Noojipady, P., Dheeravath, V., Li, Y. J., Velpuri, N. M., Reddy, G. P. O., Cai, X., Gumma, M., & Turral, H. (2008). A Global Irrigated Area Map (GIAM) Using Remote Sensing at the End of the Last Millennium. International Water Management Institute (IWMI). https://doi.org/10.13140/RG.2.1.4945.0087

USGS (United States Geological Survey). (2021). Landsat Enhanced Vegetation Index. Retrieved July 23, 2021, from https://www.usgs.gov/core-science-systems/nli/landsat/landsat-enhanced-vegetation-index?qt-science_support_page_related_con=0#qt-science_support_page_related_con

Vogels, M. F. A. (2019, May 15). *Monitoring smallholder irrigation using remote sensing and object-based image analysis: Case studies in Eastern Africa* [Dissertation]. Utrecht Studies in Earth Sciences; Faculteit Geowetenschappen, Universiteit Utrecht. http://dspace.library.uu.nl/handle/1874/379900

Wani, S., Rockstrom, J., & Oweis, T. (2009). Rainfed Agriculture: Unlocking the Potential. CAB International. https://www.iwmi.cgiar.org/Publications/CABI_Publications/CA_CABI_Series/Rainfed_Agriculture/Protected/Rainfed_Agriculture_Unlocking_the_Potential.pdf

APPENDIX 1. Irrigation Mapping Studies conducted in developing countries and/or tropical climates

Country/Region	Year	Data Used	Paper	Methodology	# Ground Truthing points collected
Ethiopia/Horn of Africa	2019	Sentinel-2	Vogels Dissertation (Chapter 3: Irrigated-cropland mapping using GEOBIA), 2019	"novel method based on object-based image analysis (OBIA) and Sentinel-2 imagery". Segmentation using eCognition, NDVI difference.	"3000 objects were randomly selected to create a training and validation set for the classification. These objects were thematically labeled by visual interpretation" into 1 of 5 classes.
Burkina Faso	2019	Landsat 8	Chaix-Bar *et al*. Report	Calculated NDVI, Identified surface water and created a 1.5 km buffer, Time series analysis of NDVI, Informed Unsupervised classification (Select NDVI thresholds according to climate and cropping calendar), Accuracy Assessment	1,407 GT points were taken (619 for irrigated areas and 788 for other land covers) and 25 reservoirs or dams were visited. Transect crossing all three agro-ecological zones, 5 focal areas.
South Africa	2015	Landsat 8+MODIS	Cai *et al.* IWMI Working paper	Threshold value NDVI >0.14 used to separate fields in the dry season. No accuracy assessment provided.	455 GT points collected from field visits in 2014, 2015, 2016 using Open Data Kit (ODK) Collect.