Mapping Agriculture Areas Using Synthetic Aperture Radar in the Context of Flood-Related Disaster Events

Brooke Kubby[1], Jordan Bell[2], Franz Meyer[1], and Alex Lewandowski[1]

[1] University of Alaska Fairbanks, Fairbanks, Alaska
[2] NASA Marshall Space Flight Center, Huntsville, Alabama

PRESENTED AT:

AGU FALL MEETING
Online Everywhere | 1–17 December 2020

INTRODUCTION AND BACKGROUND ON COEFFICIENT OF VARIATION

The Hindu-Kush Himalaya (HKH) is a highly mountainous geographic region of Asia which represents a significant source of food production. In recent years, climate variability (particularily flooding) has threatened agroindustry and food security throughout the region.

However, there is a gap in the data regarding the location of active agriculture in HKH region. Data gaps exists for several reasons, including 1) wet, cloudy climate and severe weather events render optical data ineffective, 2) a lack of government resources for on-the-ground yearly crop assessments, 3) inconsistency in type and location of fields from one year to the next require frequent surveys, 4) insufficient data processing capacity to process satellite imagery, and 5) a lack of geospatial education and expertise in the region all contribute to this data gap.

The NASA-ISRO SAR (NISAR) mission is a joint US and Indian satellite projected to launch in 2022. The satellite will use synthetic aperture radar (SAR) remote sensing, which is able to penetrate through clouds, detect changes in surface deformation, distinguish between various types of features, and provide global coverage every 12 days.

By conducting a time series analysis of the coefficient of variation (CoVar) on this imagery, temporal changes in the diffusion pattern of active crops can be pin-pointed and mapped.

Coefficient of Variation is defined as:

$$CV = \frac{std(\sigma)}{mean(\sigma)}$$

CoVar information can then be paired with flood, hail, and other disaster maps to more accurately determine the effects of disasters on crop yields, which is part of a larger flood mapping project called HydroSAR currently working on developing automatic data processing of SAR data in preparation for the launch of NISAR. Ideally, the final data products from this project will provide valuable information to HKH nations that will aid in decisions affecting the food security of this region.

Previous work by Paul Siqueira and Tracy Whelen provides a proof-of-concept for using CoVar to detect crops in the United States, using the USDA Cropland Data Layer (CDL) as a reference to determine the accuracy of their methods. Current data gaps provide opportunity to explore critical regional HKH hydrology processes associated with food production cycles. To that end, this poster suggests methods to improve upon and refine existing proof-of-concept academic code and data analysis process for agriculture mapping using the CoVar method, and make these SAR crop-detecting algorithms available to end users in the HKH.

METHODOLOGY

The Alaska Satellite Facility (ASF) utilizes the Jupyter Open SAR Lab (OSL) to host python notebooks. Processing and calculations were conducted in this cloud environment. Sentinel 1 images were selected and initially processed using ASF Data Search and HyP3.

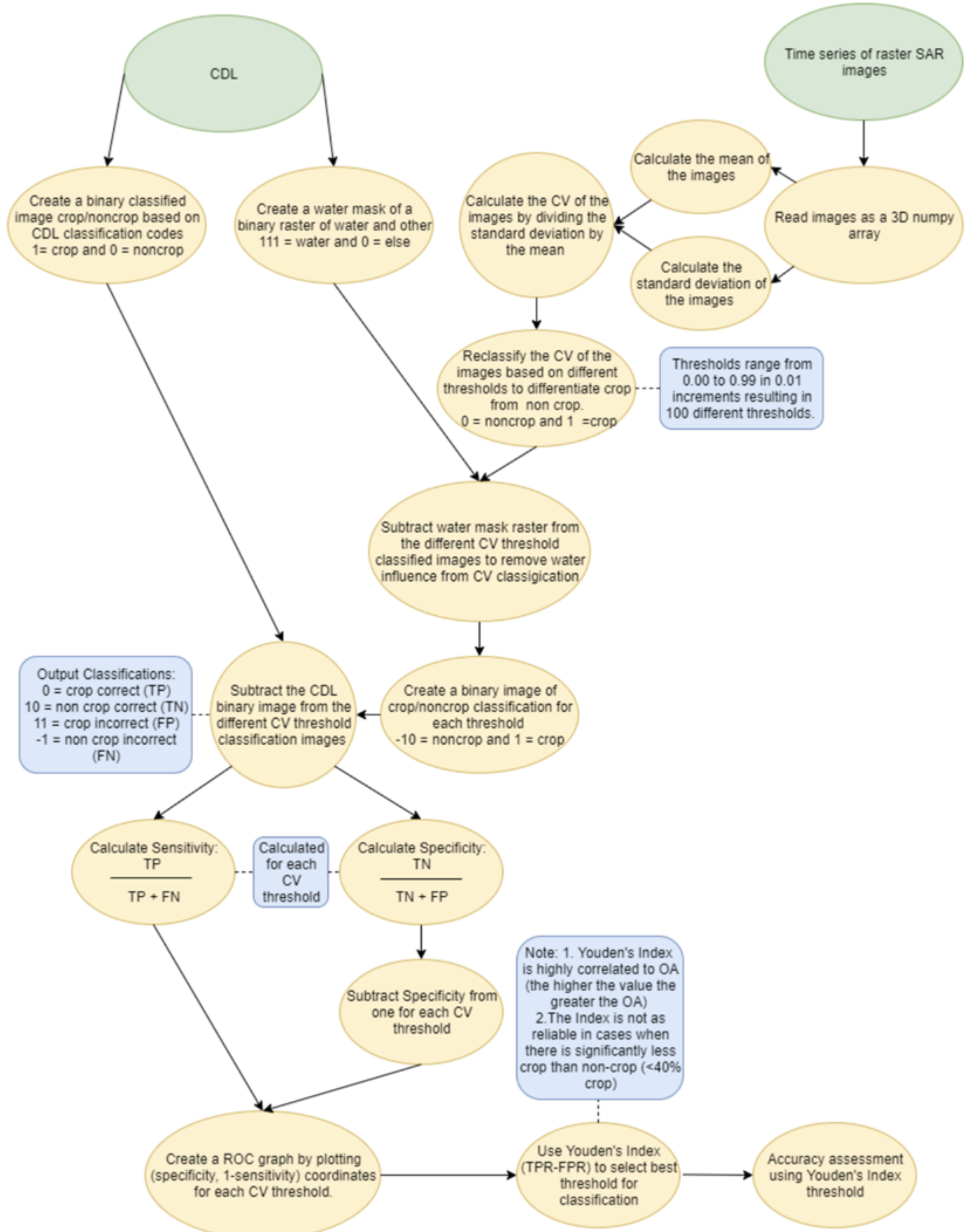
Three preliminary test sites were selected, with Sentinel 1 data from an agricultural area Nebraska, Bangladesh, and western India. Initial processing, including the creation of a time series stack and cropping to the AOI, were conducted in OSL. The python code to calculate CV, developed by the Microwave Remote Sensing Laboratory (MIRSL) at the University of Massachusetts- Amherst,

was originally developed to take in the Cropland Data Layer (CDL) as a reference layer, and was later modified to accommodate different agricultural maps for Bangladesh and India.

The reference layers were then re-classified into a binary crop/no crop layer, with an additional water mask layer. The SAR imagery was prepared as a time series, and then the mean and standard deviation of that time series was calculated to find the coefficient of variation. This is a gradational value, and the goal of this work is to identify the ideal cut-off value with which to most accurately classify crops and non crops.

The program then ran through various CV values in increments of 0.01 and created a Receiver Operating Characteristic (ROC) curve to compare these values to the reference layer and find the most accurate threshold. The statistic Youden's Index was used to determine the ideal threshold on the curve to use for best classification results. In addition to classified Geotiffs and images (shown in the slidedeck below), the program also provides accuracy statistics. Future work will be focused on additional masks and processing to increase the efficiency and efficacy of the CV threshold.

The figure below (by Shannon Rose) shows the data flow for the notebook:


CDL
Time series of raster SAR images
Create a binary classified image crop/noncrop based on CDL classification codes 1= crop and 0 = noncrop
Create a water mask of a binary raster of water and other 111 = water and 0 = else
Calculate the mean of the images
Calculate the CV of the images by dividing the standard deviation by the mean
Read images as a 3D numpy array
Calculate the standard deviation of the images
Reclassify the CV of the images based on different thresholds to differentiate crop from non crop. 0 = noncrop and 1 =crop
Thresholds range from 0.00 to 0.99 in 0.01 increments resulting in 100 different thresholds.
Subtract water mask raster from the different CV threshold classified images to remove water influence from CV classigication
Output Classifications: 0 = crop correct (TP) 10 = non crop correct (TN) 11 = crop incorrect (FP) -1 = non crop incorrect (FN)
Subtract the CDL binary image from the different CV threshold classification images
Create a binary image of crop/noncrop classification for each threshold -10 = noncrop and 1 = crop
Calculate Sensitivity: TP / TP + FN
Calculated for each CV threshold
Calculate Specificity: TN / TN + FP
Subtract Specificity from one for each CV threshold
Note: 1. Youden's Index is highly correlated to OA (the higher the value the greater the OA) 2.The Index is not as reliable in cases when there is significantly less crop than non-crop (<40% crop)
Create a ROC graph by plotting (specificity, 1-sensitivity) coordinates for each CV threshold.
Use Youden's Index (TPR-FPR) to select best threshold for classification
Accuracy assessment using Youden's Index threshold

CASE STUDY RESULTS AND EXPLANATIONS

1. Nebraska Case, 2018

The ideal CV threshold for the Nebraska case was found to be 0.56. At this threshold we achieved an overall accuracy rate of 81.59% when compared to the CDL. 73.98% of crops were classified correctly using the coefficient of variation. In actuality, the accuracy was likely higher because the CDL is imperfect. Most notably, the CDL includes miles and miles of farm roads which are below the 30 m resolution of SAR- these areas are classified by SAR as crops, and by the CDL as non-crops.

The perfect ROC curve would have a true positivity rate (sensitivity) of 1.0, and a false positive rate (specificity) of 0.0. At a CV threshold of 0.56, the ROC curve in the slide deck shows a sensitivity > 80%, but has a high false positive rate of about 40%. Further work will focus on decreasing this false positive rate, both by processing the reference data further and by adding an adaptive, more accurate water mask to the SAR time series stacks.

2. Bangledesh Case, 2018

3. India (West Bengal) Case, 2018

SUMMARY

Synthetic Aperture Radar shows great promise as a viable means for identifying and mapping active agricultural extent, especially in flood-prone areas. SAR is an appropriate remote sensing tool due to its abilities in cloud penetration, accurate deformation mapping, and determination of specific feature types. Issues associated with temporal resolution, water contamination of images, and data volume are currently being addressed by upcoming NISAR mission and associated science teams, including HydroSAR.

Coefficient of Variation, even in its rough preliminary stages, is able to identify agriculture with surprisingly high degrees of accuracy. Future work, involving the integration of an adaptive water mask for each image in the SAR time series, should improve this accuracy and make SAR a go-to tool for identifying agriculture pre and post flooding events.

FURTHER/FUTURE WORK:

Future work will focus heavily on increasing the true positivity rate (sensitivity) and decreasing the false positive rate (specificity). The larger HydroSAR project is working on accurately identifying and mapping flooding events in the HKH region. By combining the products from HydroSAR with the CoVar method of identifying agriculture, we can create a highly adaptative and dynamic water mask that masks out water differently for each image within a time series. This dynamic water masking will likely decrease the number of false positive detections.

Additionally, we would like to try adding a speckle and/or smoothing filter to the CDL in order to remove the (mostly dirt) roads from the reference layer, which are too small of a resolution for current SAR imagery.

ACKNOWLEDGEMENTS AND REFERENCES

Initial work on coefficient of variation was conducted by Paul Siqueira and Tracy Whelen[1] of the Microwave Remote Sensing Laboratory; University of Massachusetts, Amherst, MA

The initial notebook used for these case studies was created by Shannon Rose, Microwave Remote Sensing Laboratory (MIRSL) at the University of Massachusetts- Amherst.

The crop reference layer for Bangladesh and India was provided by Peter Potapov, Co-Director of Global Land Analysis and Directory (GLAD) and a research professor at the University of Maryland.

[1] Tracy Whelen and Paul Siqueira, "Coefficient of Variation for Use in Crop Area Classification across Multiple Climates," International Journal of Applied Earth Observation and Geoinformation 67, no. February (2018): 114–22, https://doi.org/10.1016/j.jag.2017.12.014.

ABSTRACT

Flooding in key agricultural areas can lead to widespread damage to crops and can have significant impacts on livelihoods and food availability in the affected areas. Access to actionable near real-time information on surface water extent, flood water depth, and impacted agricultural lands is often the main limitation of disaster management systems. This is especially true in many developing countries where in-situ observations for monitoring the impacts of rainfall and inundation on agriculture are often lacking.

Remote sensing is a promising mechanism to document agricultural activity on a regional scale and assess the impact of flooding events on crop health and crop productivity. However, cloud cover associated with many of these events limits monitoring of vegetation with optical systems. Data from Synthetic Aperture Radar (SAR) sensors such as Sentinel-1 or the upcoming NASA-ISRO SAR (NISAR) can supplement optical analyses by providing weather- and illumination-independent observations that are sensitive to vegetation structure and have proven capability in the mapping of surface water extent.

This paper introduces SAR-based algorithms and value-added data products that enable the monitoring of agricultural areas in the context of major flooding events. The work presented here is part of the HydroSAR project, which is building a cloud-based SAR data analysis service for the mapping of hydrological disasters and their impact on population and agriculture.

We will present the algorithms behind a SAR-based information product that is delineating regions of agricultural activity at the 1ha scale. It is derived from a statistical analysis of SAR time series data based on the coefficient of variation (CoV) metric. Changes in radar cross section throughout the crop cycle result in high CoV values, such that crop lands can be extracted from CoV via a thresholding algorithm. We will present first performance assessments of this algorithm for areas in the U.S. and in the Hindu-Kush Himalaya region.

We will also show how we intersect agriculture extent information with flood extent data to determine the extent and duration of flood impacts on crop lands. We will show examples of crop inundation maps for recent flooding events along the Nebraska River and for selected areas in eastern India and Bangladesh.