NASA DEVELOP National Program
Pop-Up Project
Spring 2022

Mexico Disasters

Comparing Feasibility of Flood Detection Methods Using Google Earth Engine and Open Data Cube for Flood Mitigation in Mexico

DEVELOP Comparison Analysis Report

April 1st, 2022

Philip Casey (Project Lead)
John Willis
Zachary Silberman
Sean Nelsen

Advisors:
Dr. Brian Killough, Committee on Earth Observations (Science Advisor)
Dr. Kent Ross, NASA Langley Research Center (Science Advisor)

Abstract

In this project, NASA DEVELOP partnered with Mexico's National Institute of Statistics, Geography and Informatics (INEGI); the Committee on Earth Observation Satellites (CEOS) Systems Engineering Office; the United Nation's Regional Centre for Space Science and Technology Education for Latin American and the Caribbean (CRECTEALC); and the Commonwealth Scientific and Industrial Research Organization (CSIRO) to investigate the accuracy and effectiveness of different flood detection approaches in Mexico. We compared two flood detection methods for the Mexican states of Hildalgo, Tabasco, and Chiapas: CEOS's Open Data Cube (ODC) and NASA SERVIR's Hydrologic Remote Sensing Analysis for Floods (HYDRAFloods). These methods leveraged surface reflectance data and backscatter data from NASA's Landsat 8 Operational Land Imager (OLI) sensor and the European Space Agency's Sentinel-1 C-band Synthetic Aperture Radar (C-SAR) sensor. The team applied the two flood detection tools to flood event case studies in the areas of interest created by the European Commission's Joint Research Centre Global Surface Water Explorer. We found that ODC and HYDRAFloods were mostly aligned in their detection of flood waters, though ODC picked up more flooding in high-elevation areas because it does not include radiometric terrain correction. HYDRAFloods tended to show more flooding in scenes with a small amount of water due to its automated thresholding for surface water detection. The project produced a Jupyter notebook utilizing ODC, historical surface water maps, a case study analysis of the two methods, and a code video tutorial to support the expedient use of Earth observation data for disaster planning and response.

1. Introduction

1.1 Background Information

When Hurricane Eta hit Mexico in the fall of 2020, it caused severe flooding in the states of Chiapas and Tabasco, destroying thousands of homes and affecting tens of thousands of people. In 2021, the torrential rain caused by Hurricane Grace resulted in severe flooding and mudslides. When major flooding events like this occur in Mexico, decision-makers need to know the extent and duration of flooding in the areas they oversee. Accessing information on the severity, location, and extent of flooding impacts can support the response and recovery efforts necessary to save lives and livelihoods. By leveraging open-source flood detection methodologies, policymakers can quickly identify the impacts of flooding events, and make subsequent emergency management decisions. Currently, two tools that provide rapid flood detection capabilities are Open Data Cube (ODC) and Hydrologic Remote Sensing Analysis for Floods (HYDRAFloods).

1.2 Objectives

The Mexico Disasters DEVELOP team worked to understand how ODC and HYDRAFloods methodologies align and then compared their outputs to understand discrepancies in flood detection. Upon comparing the outputs, the team aimed to investigate what methodological parameters and geographic contexts caused differences in flood detection. The results of the comparison analysis can support refining the two flood detection methodologies to create effective flood impact informational resources and illuminate potential limitations when using these methodologies in specific contexts. The team selected three regions of interest in Mexico for analysis: Tula de Allende, Hidalgo; Villahermosa, Tabasco; and Salto de Agua, Chiapas (Figures 1 & 2, Table 1).

Regions of Interest

Figure 1. Case study locations in Mexico including Tula de Allende, Villahermosa, Tabasco, and Salto de Agua, Chiapas for the comparison analysis between to flood detection tools, HYDRAFloods and the ODC.

Figure 2: Regions of interest detailed (from left to right: Salto de Agua, Villahermosa, Tula de Allende)

Location	Lat/Lon	Event	Date
Villahermosa, Tabasco	17.9895, -92.9475	Hurricane Sally	October 4, 2020
Salto de Agua, Chiapas	17.453815, -92.240279	Hurricane Eta	November 24, 2020
Tula de Allende, Hidalgo	20.020185, -99.290747	Hurricane Ida	September 16, 2021

Table 1: Case study information for the comparison analysis between the flood detection tools, HYDRAFloods and the ODC.

2. Methodology

2.1 Open Data Cube

The Open Data Cube is built as a set of Python libraries and can be used in Jupyter notebooks or Google Colab. The ODC takes a region and timeframe which aggregate Sentinel-1 scenes to support the flood detection analysis. Users define a centroid using a latitude and longitude and a boundary box (in degrees) of their choice to delineate the region of interest. The time extent is then defined by the user; to expedite processing time and memory, a month is the suggested time period for collecting scenes. After, establishing the region and time extent, users can then observe the collection of Sentinel-1 C-band synthetic aperture radar (SAR) data within the notebook. The SAR imagery is sorted by acquisition date, pass direction, mission type, and orbit number. This information allows users to filter scenes using these parameters to better organize imagery for analysis. Additionally, filtering ensures that comparison analyses between scenes are conducted using the same pass direction (ascending or descending). Acquisitions with the same orbit number can come from two missions (Sentinel-1A and 1B with 6-day separation) or one mission (Sentinel-1A or 1B with 12-day separation). Once

the scenes for analysis are selected, ODC employs the steps described below to identify flood extent.

A speckle filtering algorithm is used to perform a common "block averaging" filter to average the surrounding pixels of any given pixel. Speckle filtering ameliorates issues inherent to SAR imagery when scenes are "grainy" in appearance due to speckle (statistical noise) and differences between water and land. ODC then takes the speckle filtered, single-date scene and creates a histogram plotting the statistical backscatter distribution of the VV and VH polarisation data. Pixels are interpreted as the difference between land pixels (higher values) and water pixels (lower values). The ODC uses the VH polarisation channel, which is recommended for detection of water, as VV polarisation can be impacted by wind and rain interference. The backscatter amplitude for VV and VH is then scaled to an 8-bit range of 0-255 to enhance visual output and maximize contrast in image products. ODC then produces a single-date flood extent map by selecting a scene, setting the threshold (users manually select a decibel threshold at which land and water will be differentiated). This threshold produces visualizations mapping water based on SAR backscatter intensity and plotting the pixels that are below the threshold in blue (water), and pixels that are at or above the threshold in the same color scale as the SAR distribution. ODC then provides a capability to analyze the changes in flood extent for two scenes representing different dates. The algorithm identifies flooding over a period of time by selecting a change threshold, which determines how much the decibel reading for a pixel needs to decrease for it to be reclassified from land to water, indicating flooding. The final product performs a comparison of the two dates and calculates the change in backscatter between those dates. Pixels with significant reduction in backscatter likely represent changes from land to water due to flooding. ODC maps all the variables created using Python packages, including numpy, xarray, pandas, and matplotlib. An example of the result for a scene in Chiapas for November 24, 2020 is shown in Figure 3.

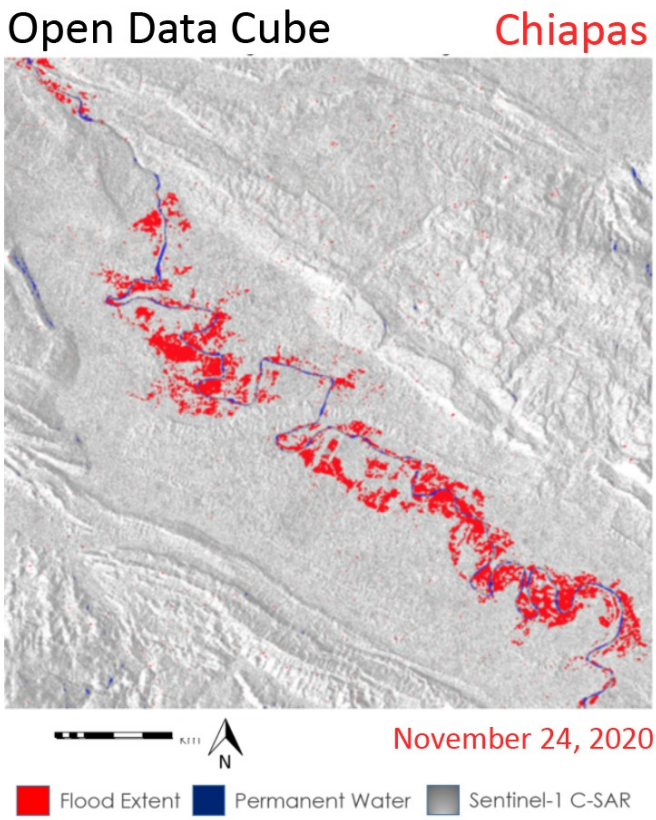

Figure 3: ODC flood extent map of Salto de Agua, Chiapas on November 24th, 2020 (Sentinel-1 C-band SAR, 2018)

2.2 HYDRAFloods

HYDRAFloods is a Google Earth Engine (GEE) Python API based application developed by NASA SERVIR that uses Sentinel-1 C-band SAR imagery to produce flood water maps. The first step to data acquisition preprocessing is defining our region of interest within Google Earth Engine, as well as the time period for which Sentinel-1 images will be pulled. To begin processing, HYDRAFloods implements radiometric terrain correction (RTC) to account for the effects of elevation variations on SAR imagery. HYDRAFloods RTC employs two elevation datasets from the Multi-Error-Removed Improved-Terrain (MERIT) Hydro: Global Hydrography digital elevation model (DEM) and the height above nearest drainage (HAND). After applying RTC, HYDRAFloods applies a speckle filtering algorithm using a gamma map processing algorithm to reduce geometric and radiometric noise in the SAR imagery. Then, a mask is applied on areas that are 20m or higher than the nearest body of water. After the elevation mask, HYDRAFloods applies an Edge Otsu algorithm to automatically determine a threshold for water values in the SAR data. The Edge Otsu method in HYDRAFloods requires an initial threshold where the algorithm will begin its search along the backscatter data to find the optimal threshold. Finally, HYDRAFloods leverages the European Commission's Joint Research Center (JRC) Yearly Water History dataset to mask permanent water in the scene from SAR detected flooding. By masking the processed SAR scene when a pixel has been classified as permanent water in the past five years,

the final product will only include SAR-detected flooding and remove permanent water from the output, and includes SAR data, historical permanent water, and flood extent as layers. HYDRAFloods uses the geemap package for displaying these layers and variables on an interactive Folium map. An example of the result for a scene in Chiapas for November 24, 2020 is shown in Figure 4.

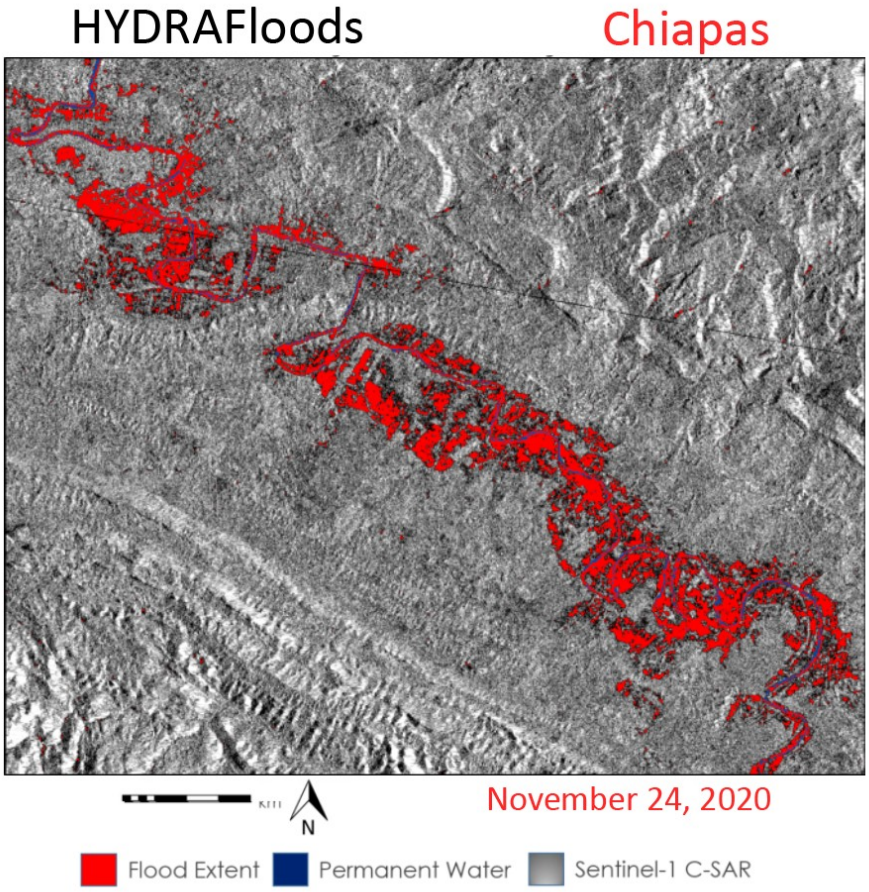

Figure 4: HYDRAFloods flood extent map single-date, Salto de Agua, Chiapas on November 24, 2020 (Sentinel-1 C-band, 2018).

2.3 Comparison Analysis

The NASA DEVELOP team applied a comparison analysis between the two flood detection methodologies to understand where the methodologies differed in illuminating flood extents. The analysis highlighted differences where the two algorithms matched and did not match. Given that ODC and HYDRAFloods differ in methodologies — most notably in RTC, Edge Otsu thresholding, and baseline water masking — the DEVELOP team worked to update the scripts so they could directly compare results. The team attempted to control as much as possible for RTC and thresholding techniques to gain insights on how these processing steps impact the alignment of the final flood extent maps of the two methodologies.

For the final comparison analysis, the technical components of the methodologies needed to align so a comparison analysis could occur. To attempt to match water

thresholds as closely as possible between the two tools, we set the initial threshold in HYDRAFloods' Edge Otsu method equal to the threshold value manually selected in ODC. This had the effect of bringing the HYDRAFloods threshold closer to the ODC threshold, though the Edge Otsu method still tended to select a higher (i.e., less conservative) threshold than the one chosen manually for ODC. As mentioned earlier, ODC's map products use solely Python visualization packages (matplotlib_pyplot), whereas HYDRAFloods uses Google Earth Engine mapping (geemap). To visualize HYDRAFloods results with the same pure-Python visualization tools used by ODC, the HYDRAFloods maps were exported to a GeoTIFF file in Google Drive and re-imported to the notebook as a GeoTIFF raster using the Python packages "rasterio" and "xarray." Once the HYDRAFloods GeoTIFF was loaded, both ODC and HYDRAFloods outputs were plotted with matplotlib. The raster images were then assigned water classification variables and brought into an image plot, with the following characteristics:

- If sum = 0, then no flooding (black)
- If ODC water class sum = 1 and HYDRAFloods water class = 0, then only ODC flooding (red)
- If HYDRAFloods water class sum = 1 and ODC water class = 0, then only HYDRAFloods flooding (green)
- If water class = 2, then both ODC and HYDRAFloods flooding (blue)

There were several adjustable parameters in the code that we decided to keep consistent between all analyses. In ODC, we chose a block size of 5 for speckle filtering; we did not adjust any of the ODC code in the backscatter scaling; and we selected a change threshold of –5 in ODC for every analysis. In both ODC and HYDRAFloods, we analyzed only the "VH" band of the SAR data. The initial threshold in HYDRAFloods' Edge Otsu algorithm was always set to the water threshold manually set in ODC. However, while we selected a manual water threshold of –22 for all the case studies, this was not pre-planned; in each case, we judged that threshold to be the best fit for the backscatter data we were looking at. With all these controls in play, we attempted to isolate a few core differences between ODC and HYDRAFloods:

- Edge Otsu (HYDRAFloods) versus manual water thresholding (ODC)
- JRC permanent water masking (HYDRAFloods)[1] versus manually-selected baseline water mask (ODC)
- Radiometric Terrain Correction (HYDRAFloods) versus analysis with uncorrected terrain (ODC)
- Differing speckle filtering algorithms

3. Results & Discussion

3.1 Case Study: Villahermosa, Tabasco (September-October 2020)

The team first compared the flood detection algorithms in Villahermosa, Tabasco during Hurricane Sally. The backscatter histograms below show the SAR backscatter values from two dates produced from an ODC script (Figures 5 and 6). Based on the backscatter histograms in Figures 5 and 6, we manually selected a water threshold of –22 decibels in ODC, which meant that the initial threshold

[1] HYDRAFloods does not require that JRC permanent water data be used to mask non-flood waters, but it was the default method for masking non-flood waters in the HYDRAFloods code our team received.

provided to HYDRAFloods' Edge Otsu method was also –22, but the final threshold determined by the Edge Otsu algorithm was –20.1.

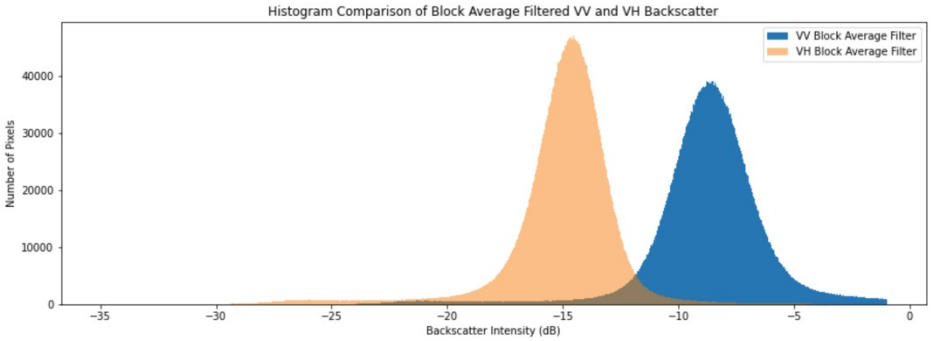

Figure 5: Backscatter histogram for the baseline date in Tabasco, September 16, 2020

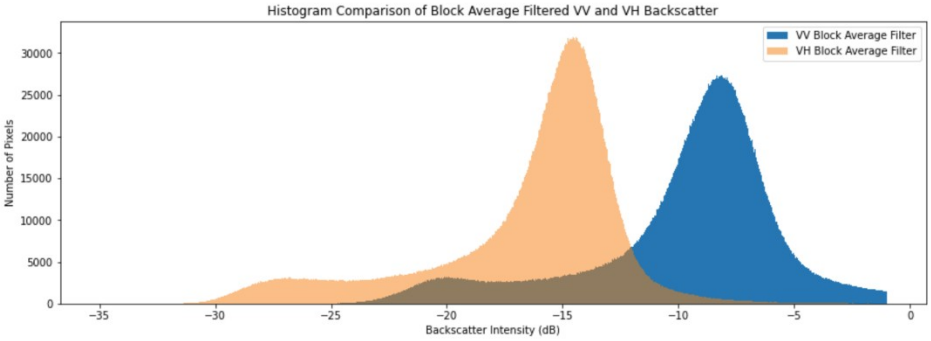

Figure 6: Backscatter histogram for the baseline date in Tabasco, October 4, 2020

In Figure 7, the images display the flood extent maps created by the two methodologies. Upon inspection, the maps differ in where flooding is detected using SAR. The SAR imagery also differs in contrast, where HYDRAFloods is a darker SAR scene and ODC is a lighter SAR scene; the reason for the difference is likely due to the RTC algorithm used in HYDRAFloods.

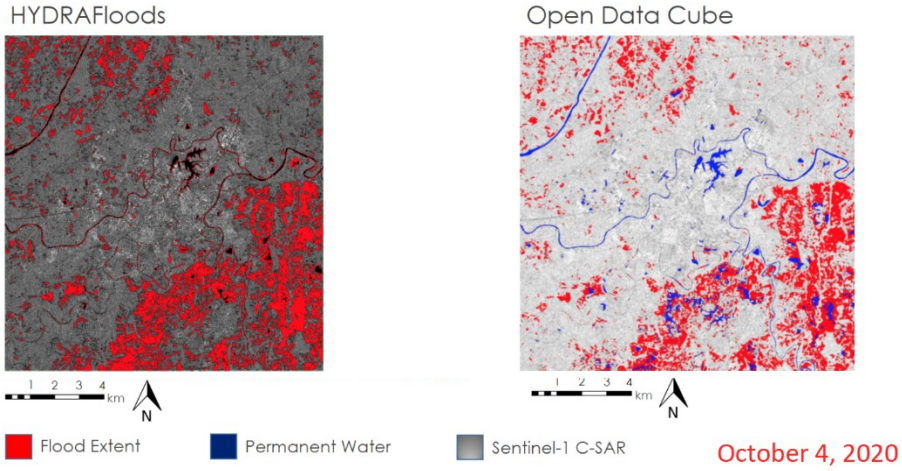

Figure 7: HYDRAFloods and ODC flood extent map single-date in Villahermosa, Tabasco. (Sentinel-1 C-band, 2018)

To get a detailed assessment of where the flood detection methodologies match, the team produced a specified comparison map. Below, the comparison map details where the algorithms agree and where they disagree (Figure 8). Black highlights agreement in methodologies where no water was detected, blue shows agreement in flood detection, green displays only HYDRAFloods detection, and red displays only ODC detection.

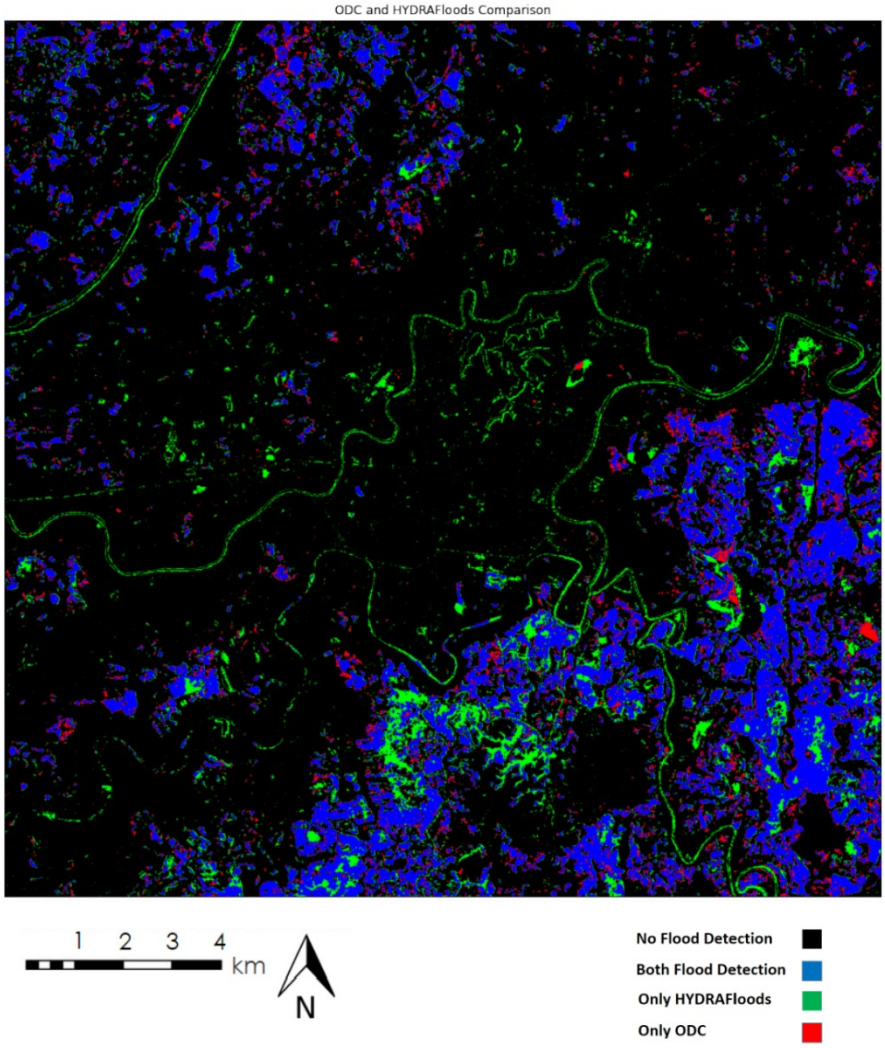

Figure 8: Comparison Map of Flood Extents between HYDRAFloods and ODC in Villahermosa, Tabasco on October 4, 2020 (Sentinel-1 C-band, 2018)

The comparison map above shows distinct differences in where the two tools detect flooding. HYDRAFloods detected flooding on the edges of water bodies, most notably on the Rio Carrazil and Rio Grijalva. This is likely due to inaccuracies in the JRC permanent water dataset, as well as differences in the water threshold values used by each tool. The HYDRAFloods threshold was higher and less

conservative than the ODC threshold, meaning more pixels would be counted as flooding. HYDRAFloods also detected more flooding overall in the region, which may point towards errors of commission in HYDRAFloods or errors of omission in ODC.

As can be seen in the table below (Table 2), there were about 1.5 times as many pixels counted as flooding by HYDRAFloods alone than pixels counted as flooding by ODC alone. With regard to general pixel classification, the two tools were mostly in agreement, with 93% of pixels classified the same (flooding or not flooding) by both methods. Of all pixels classified as flooding, the two methodologies agreed approximately 60% of the time.

Flood Detection	**Pixel Count**	**Percent of Total**
No Flood Detected	4388200	82.1%
Both ODC and HYDRAFloods Detected Flooding	571329	10.7%
Only HYDRAFloods Detected Flooding	236752	4.4%
Only ODC Detected Flooding	142887	2.6%

Table 2: Flood Detection Alignment between ODC and HYDRAFloods in Villahermosa, Tabasco on October 4, 2020

3.2 Case Study: Salto de Agua, Chiapas (October-November 2020)

Next, we performed the same analysis on Salto de Agua, Chiapas during Hurricane Eta, looking at flooding on November 24, 2020 compared to a baseline water level taken from October 25, 2020 (Figures 9 and 10). These histograms were produced using an ODC script. There was very little water in proportion to land in this study area, so neither of the backscatter histograms contained clear peaks for water backscatter. Nevertheless, we selected a manual threshold of –22db for ODC. The final threshold calculated by the Edge Otsu algorithm in HYDRAFloods was –20.6. As can be seen in the two flood extent maps below (Figure 11), flooding in the region of interest was concentrated in the valley through which the Rio Xanil flows.

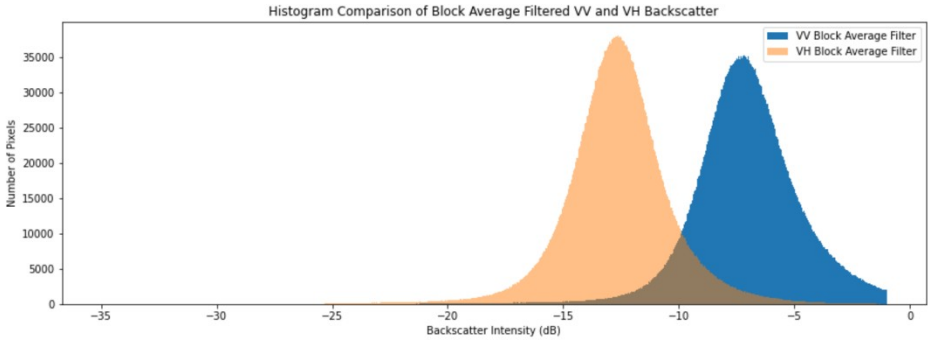

Figure 9: Backscatter histogram for the baseline date in Chiapas, October 25, 2020

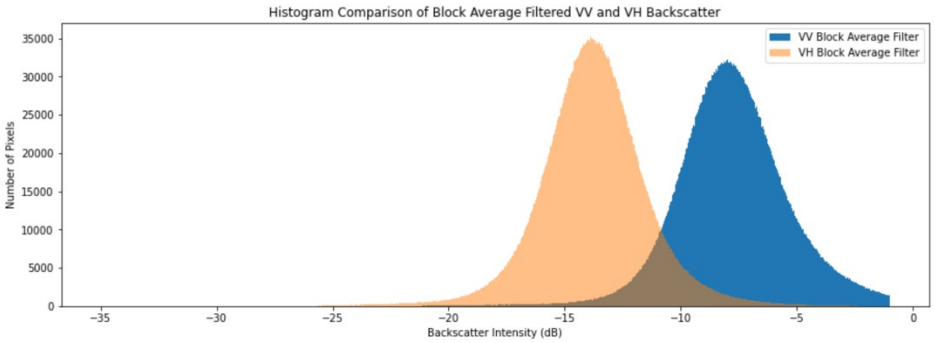

Figure 10: Backscatter histogram for the flood date in Chiapas, November 24, 2020

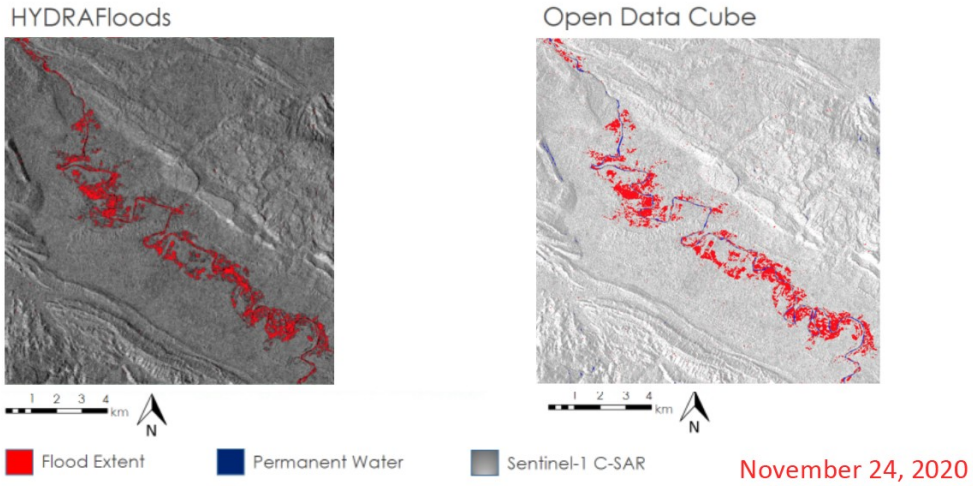

Figure 11: HYDRAFloods and ODC flood extent map single-date in Salto de Agua, Chiapas (Sentinel-1 C-band SAR, 2018)

In the comparison map below, the differences between the two methods become clearer. Again, we see issues with HYDRAFloods attributing permanent water edges as flooding, which may be due to inaccuracies in the JRC permanent water mask, or the more liberal water threshold selected by Edge Otsu for HYDRAFloods. Additionally, both HYDRAFloods and ODC detect small spots of flooding away from the Rio Xanil in higher elevation areas. This could be due to noise in the original SAR imagery that was not filtered out by either tool's speckle filtering. Given that these are high elevation areas, these spots could also result from shadows created by various peaks in the terrain. This latter explanation is also supported by the fact that ODC, which does not have terrain correction, seems to detect more flooding in these mountainous areas than HYDRAFloods, which employs RTC to counter the effects of elevation on SAR backscatter. Still, RTC cannot completely remove the effects of elevation, and HYDRAFloods still detects some flooding at higher elevations.

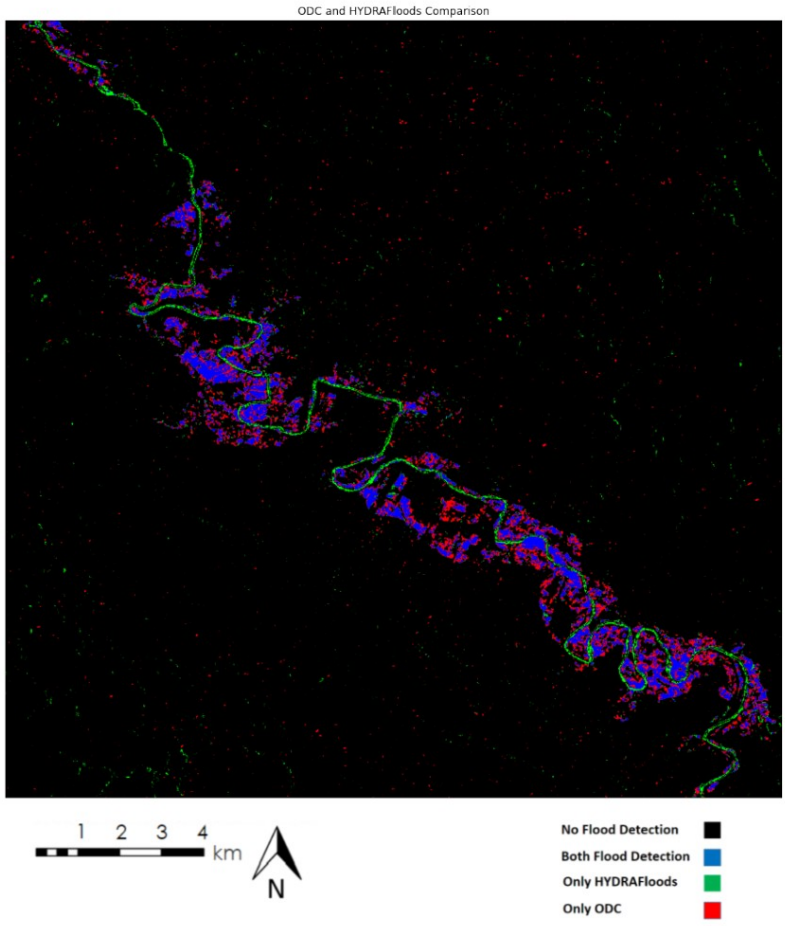

Figure 12: Comparison Map of Flood Extents between HYDRAFloods and ODC in Salto de Agua, Chiapas on November 24th, 2020 (Sentinel-1 C-band SAR, 2018)

As seen in the pixel count table below (Table 3), the two tools agreed 97% of the time in their classification of pixels in general. For all pixels classified as flooding, the two methodologies agreed approximately 44% of the time. ODC detected more than twice as much flooding as HYDRAFloods. It is especially interesting that ODC detected more flooding than HYDRAFloods in this example because the water threshold used by ODC was more conservative than the one used by HYDRAFloods. The flooding detected exclusively by ODC tended to be either at higher elevation or very close to pixels that were also identified as flooding by HYDRAFloods. While the lack of RTC in ODC may account for the portion of ODC-only flood pixels located in high elevation areas, our team does not currently have an explanation for ODC-only flood pixels situated along the Rio Xanil, very close to pixels identified as flooding by both tools.

Flood Detection	**Pixel Count**	**Percent of Total**
No Flood Detected	4860560	95.4%
Both ODC and HYDRAFloods Detected Flooding	101196	1.99%

Only HYDRAFloods Detected Flooding	40510	0.8%
Only ODC Detected Flooding	88745	1.74%

Table 3: Flood detection alignment between ODC and HYDRAFloods in Salto de Agua, Chiapas on November 24, 2020

3.3 Case Study: Tula de Allende, Hidalgo (September 2021)

Finally, we analyzed a period of flooding in Tula de Allende, Hidalgo following Hurricane Ida in September 2021. The histograms below show the backscatter values based on two separate dates, both created from an ODC script (Figures 13 and 14). Once again, there was no clear peak for water in the backscatter histograms for this study area, as the vast majority of the viewing area is made up of land rather than water. We chose a manual threshold of –22 for ODC, which served as the initial threshold for the Edge Otsu threshold calculation. The resulting threshold from the Edge Otsu calculation in HYDRAFloods was –20.2.

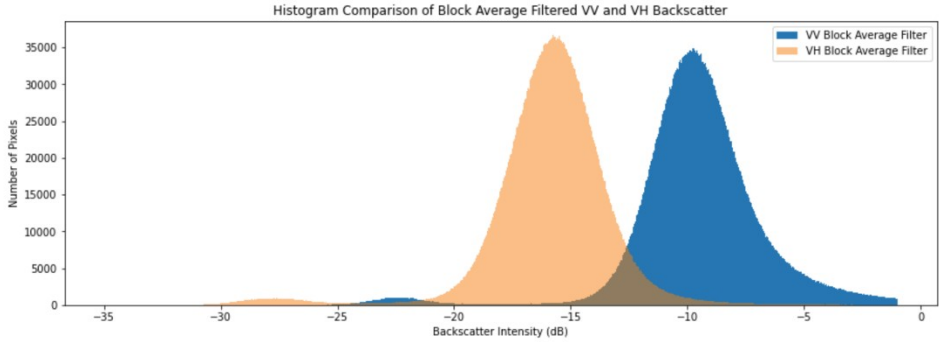

Figure 13: Backscatter histogram for the baseline date in Hidalgo, September 4, 2021.

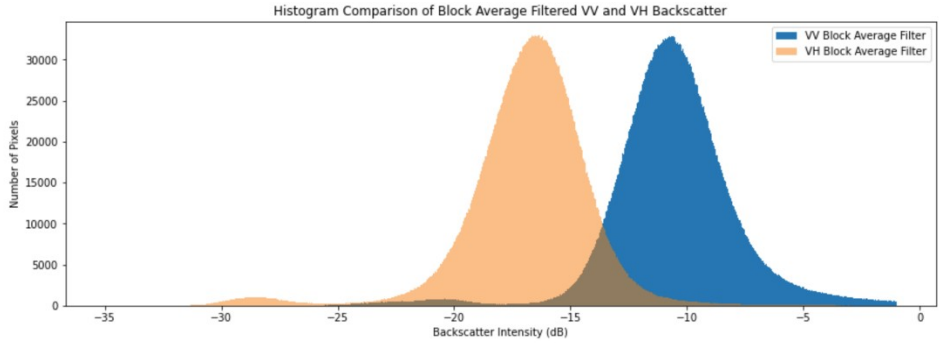

*Figure 14: B*ackscatter histogram for the flood date in Hidalgo, September 16, 2021.

As can be seen in the flood extent maps below (Figure 15), flooding in Tula de Allende during the study period was mostly spread out and sparse. The major exception to this is in the HYDRAFloods flood extent map, where there is a bright blotch of flood pixels in the Northwest, the outline of a waterbody in the South, and a section of double-bounce backscatter highlighted in the city center. The blotch of flood pixels in the Northwest is due to an error in the JRC permanent water mask, where a large swath of a water body is left unmasked. The flood detection in the city center may be due to the lower threshold selected by the Edge Otsu algorithm,

but it is difficult to derive any certainty for SAR backscatter in city centers, as the double-bounce scattering off of buildings and other perpendicular surfaces leads to unpredictable results.

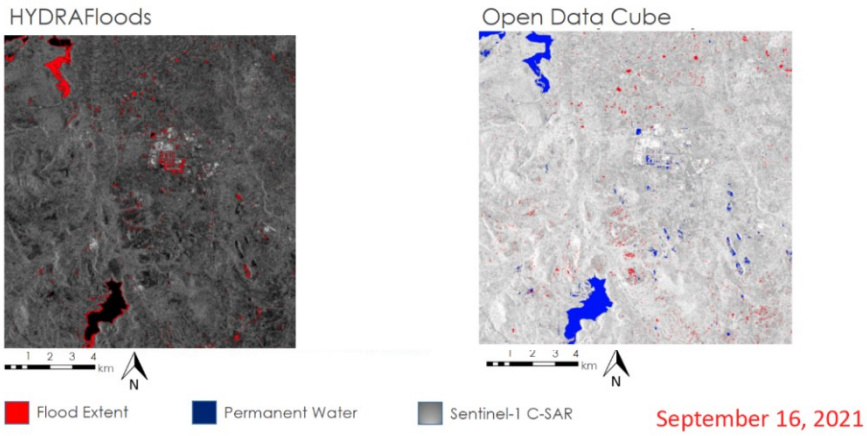

Figure 15: HYDRAFloods and ODC flood extent map single-date in Tula de Allende, Hidalgo (Sentinel-1 C-band SAR, 2018)

Most of the agreement between tools in the comparison image below seems to lie North of the city, where there are several blue spots. There are some ODC-only flood areas to the South of the city that appear substantial, but much of the image contains specks of red and green that could just as likely be noise as true flooding. Interestingly, while HYDRAFloods detected more flooding in the area, you can see that ODC seemingly detected more flooding in the higher elevations to the West—likely an issue with not correcting terrain.

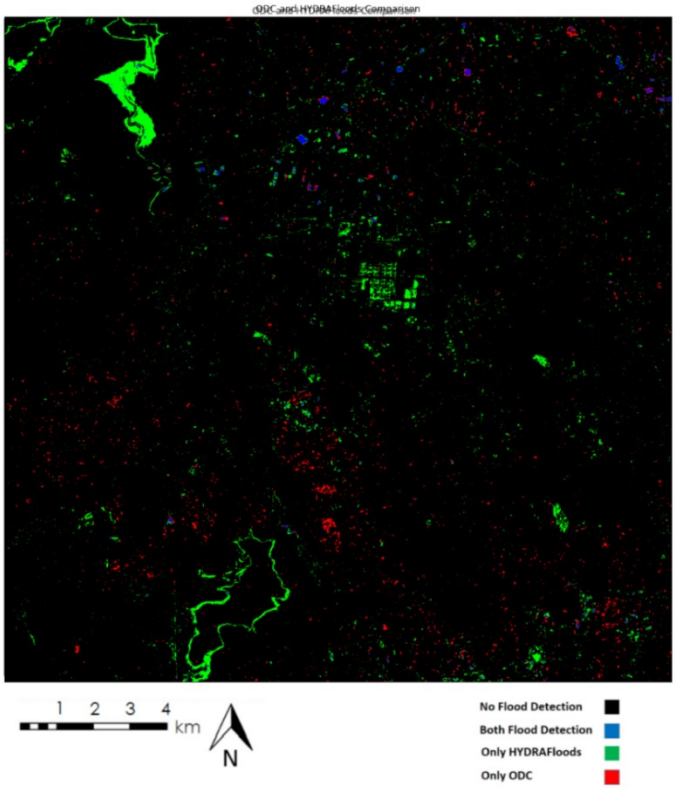

Figure 16: Comparison Map of Flood Extents between HYDRAFloods and ODC in Tula de Allende, Hidalgo on September 16, 2021 (Sentinel-1 C-band, 2018)

In this case study, there was very little flooding detected at all, with 97% of the study area identified by both tools as not flooded. Despite this agreement on what is not flooded in the region of interest, the two tools had almost no agreement about what was flooded. Only 7% of all pixels identified as flooding were counted as flooding by both tools. HYDRAFloods accounted for the majority of flooding identified in the region of interest, but most of these pixels are likely false positives for flooding, due to the issues in the JRC permanent water mask discussed above and the unreliability of SAR for flood detection in city centers, where a considerable portion of the HYDRAFloods-only flood pixels are located in this case study.

Flood Detection	**Pixel Count**	**Percent of Total**
No Flood Detected	4,949,143	97.1%
Both ODC and HYDRAFloods Detected Flooding	10,386	0.2%
Only HYDRAFloods Detected Flooding	96,659	1.9%
Only ODC Detected Flooding	43,332	0.8%

Table 4: Flood Detection Alignment between ODC and HYDRAFloods in Tula de Allende, Hidalgo on September 16, 2021

3.4 Future Work

Future work should include an analysis that would isolate variables in the different tools to find out exactly why the flood detection images differ from each other.

Performing flood analyses with matching water thresholds between the two tools could help determine the effect of the differing thresholds versus other variations in the tools' flood detection methods. It would also be helpful to see how much more the two tools' flood detection might align if the same baseline water mask were applied in both, rather than using the JRC permanent water mask for HYDRAFloods and a water mask from a manually-selected baseline date in ODC. A thorough comparison of RTC and non-RTC HYDRAFloods images could help determine the significance of terrain correction to flood detection. Our team completed two of these RTC-corrected flood maps, one in Hidalgo and the other in Chiapas, but time constraints did not warrant a full analysis of these maps here. In addition, comparing the flood detection maps to ground-truth measurements taken onsite during flooding events could help determine true and false positives and negatives for flooding. Finally, it would be worthwhile to test other thresholding algorithms available in HYDRAFloods, such as bmax-otsu and k-means, as these might yield different results for flood detection.

4. Conclusions

By comparing the ODC and HYDRAFloods flood detection algorithms in a consolidated Google Colab notebook, the NASA DEVELOP team was able to understand the differences in the methodologies and create a foundation to refine the algorithms to enhance their flood detection capabilities. In each case study, HYDRAFloods had issues attributing permanent water edges as flooding. This created errors in the over commission of flood pixels, likely due to inaccuracies in the JRC permanent water masks. The aim to use different geographies to better understand the performance of the tools helped illuminate that variability in the flood detection methodologies. In Tabasco, HYDRAFloods identified more pixels as flooding; in Chiapas, more pixels were counted as flooding by ODC. These differences show that the tools may perform better in different geographic contexts. Given that both methodologies use 2018 Sentinel-1 C-band SAR imagery to detect flooding, the expectation was that the algorithms should significantly align in locations of flood detection. However, it was only in the Tabasco case study that the two tools agreed on where flooding was the majority of the time. Thus, we found there to be more notable differences between the two tools than we anticipated.

This material contains modified Copernicus Sentinel data (2016 – 2021), processed by ESA.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Aeronautics and Space Administration.
This material is based upon work supported by NASA through contract NNL16AA05C.