**NASA DEVELOP National Program**



International Research Institute for Climate and Society (IRI)

Palisades, NY

*Summer 2015*

Malawi Disasters II

Applications of Flood Definitions and NASA Earth Observations to Create a Flood Forecasting Methodology

 **Technical Report**

Rough Draft – June 25, 2015

Andrew Kruczkiewicz

Helen Cen

Brigitte Moneymaker

Dr. Pietro Ceccato, Lead Environmental Monitoring Program, IRI (Science Advisor)

Previous Contributors:

Jerrod Lessel

Alex Sweeney

# I. Abstract

In January 2015, extended periods of extreme rainfall caused a series of flood events throughout Malawi resulting in the displacement of over 230,000 residents and caused 276 fatalities. In order for local authorities and humanitarian agencies to provide post-disaster relief, these organizations often rely on remotely sensed satellite data to evaluate initial disaster impact and design response programs. In partnership with the Malawi Red Cross, this project aimed to expand on the findings from Spring 2015 by adding enhanced ground-truth data (locations of shelter sites of internally displaced people (IDPs) and origins of IDPs) into the initial analysis from the previous research, second using knowledge gained by communication with project partners, local authorities and from a visit to the study region to define regions by predominate flood type and third, by integrating European Space Agency (ESA) remotely sensed data to explore the potential predictive capabilities of soil moisture for flash flood detection. In addition to data from NASA sensors (MODIS, TRMM, SSM-I and AMSU-A data), this project incorporated ASCAT data from ESA and the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT). The results of this study will increase the ability to forecast and monitor flood events, benefiting organizations involved with disaster preparedness and relief efforts in Malawi and potentially allowing for more efficient action, including prepositioning of pre-flood resources, response operations and allocation of emergency flood relief efforts.

**Keywords**

Malawi, Floods, Inundation, Remote Sensing, Disaster Management, Soil Moisture, Precipitation, Early Warning

# II. Introduction

The African country of Malawi experiences a seasonal rainy season stretching from October to April, which provides about 95% of its annual precipitation (Malawi Meteorological Services, 2006). In addition to this high seasonality, about 20% of Malawi’s land cover is comprised of surface water from Lake Malawi, one of the Great African Lakes (Balbo et al., 2013). These unique features contribute to a countrywide vulnerability to riverine floods and flash floods. In January 2015, extended periods of extreme rainfall caused a series of flood events throughout the Central and Southern Regions of Malawi, which resulted in the displacement of over 230,000 residents and led to 276 fatalities (Malawi Government, 2015).

Following the flood events, disaster managers in Malawi, including the Malawi Red Cross, had the ability to access various flood maps, supposedly indicating which areas in Malawi were experiencing floods. This information was to be integrated into flood response programming, including allocation of food and resources in the areas of greatest impact. While some maps were similar, other varied significantly by indicating floods in areas that were not affected, or lacking flood signal altogether in known disaster areas.

The project partners for this project were the Red Cross Red Crescent Climate Centre (RCRCC) and the Malawi Red Cross (MRC) National Society. The most pressing concern in disaster risk management, as expressed by our project partners is first, the ability to identify affected areas in expedited fashion and second, to increase preparedness actions by the development of a framework for flood early warning.

The goal of this project was threefold: first we inter-compared each flood maps and validated them using ground truth data; second, we explored the ability of each map to detect certain types of floods, namely riverine floods and flash floods; third, we explored the predictive capacity of environmental variables for various flood types, using boundaries of spatial clusters designated by predominate soil moisture behavior.

The outputs of this project will contribute to the development a framework for forecasting and monitoring floods, with unique methods depending on flood type. This information will be produced in a format that can be integrated into decision making by disaster management organizations, including the MRC.

# III. Methodology

**Validation of Flood Detection Products**

In the previous term of DEVELOP, flood maps were qualitatively analyzed for potential discrepancies between the spatial distribution of flood signal compared to one another and with ground truth data. We selected 7 flood maps for the comparison and validation processes: the Dartmouth Flood Observatory (DFO) flood map, NASA Goddard Space Flight Center MODIS Near Real-Time Global Flood Mapping Project 3 Day Composite Flood Water product (NRT-GFM), TerraSAR-X flood map, RADARSAT flood map, RADARSAT-2 flood map, the University of Maryland Global Flood Monitoring System Flood Detection (GFMS-FD) product and GFMS Inundation 1KM (GFMS-I) product.

In this term, the MRC provided enhanced ground truth data which included names and locations of flooded villages as indicated by the ‘place of origin’ dataset within the data provided by internally displaced persons (IDPs). These new data allowed for the previously studied flood maps to be validated at a more precise level than before, using locations of assumed flooded villages as ground truth data points rather than locations of displacement sites.

The most recent, higher spatial resolution ground truth data included the following variables:

Shelter site ID

Name of shelter

Shelter location

Site start date

GPS coordinates of the site

Site Status (Open or Closed)

Name of flooded villages; origin of internally displaced person (IDP)

Location of flooded villages; origin of internally displaced person (IDP)

Type of shelter used

Land Ownership of the site

Survey date

Total number of IDP households

Total number of IDP individuals

Using this information, an enhanced, ground truth Shelter Location Layer was produced as a vector file in ArcGIS.

**Assessment of Validated Flood Detection Products**

The location of villages assumed to be flooded was determined from an enhanced qualitative comparison with the Malawi Villages Layer obtained from the Malawi Spatial Data Portal (MASDAP). Afterwards, if there was still some uncertainty as to the location of the village, the valid location was determined by analyzing boundaries of political districts of descending influence (provided in the dataset); starting with the Region in which the flooded village was located, followed by the Traditional Authority Area, Town, and ultimately, Village. Latitude/longitude coordinates of the origin villages were processed as a vector file using ArcGIS. This layer, deemed the Flooded Village Layer, along with the new Shelter Location Layer, replaced the spatially coarse and temporally outdated ground truth data layer from the previous term, affording the opportunity for a higher spatial resolution and more up to date analysis to be completed.

Next, both the Shelter Location Layer and Flooded Village Layer were placed above each flood detection product used in previous project to assess the product’s flood signal through qualitative analyses. The following satellite based products were used for this analysis:: DFO, NRT-GFM, TerraSAR-X, RADARSAT, RADARSAT-2, GFMS-FD, and GFMS-I.

Several techniques were used to isolate the maximum flood extent and georeference each product. For TerraSAR-X, RADARSAT, RADARSAT-2, and the DFO, data were received in a georeferenced format with the maximum flood extent data already included, so no further processing was needed. NRT-GFM data were downloaded in a daily time step, so each day of January 2015 was overlaid in chronological order to compile a maximum flood extent map in order to process the data in a format comparable to the other products. Next, each of the GFMS-FD raster data files were reprocessed and georeferenced in QGIS, and subsequently polygonized (using the QGIS raster to vector conversion tool) to isolate flood signal from the background noise. Afterwards, all eight GFMS-FD polygonized layers were stacked on top of each other in chronological order to generate a single day maximum flood extent map. Similarly, the eight GFMS-I images in raster format were georeferenced in QGIS and stacked in chronological order to generate a single map of maximum flood extent.

The satellite products were then grouped into three categories based on spectral characteristics of the primary flood-detecting sensor. Group 1 consisted of products exploiting the Synthetic Aperture Radar sensor: TerraSAR-X, RADARSAT and RADARSAT-2. Group 2 included products using MODIS: DFO and NRT-GFM. Group 3 comprised of products produced by University of Maryland’s Global Flood Monitoring System using TRMM/Global Precipitation Monitoring (GPM) precipitation data coupled with a hydrologic model: GFMS-FD and GFMS-I.

**Flood Type Disaggregation**

Part two of the project explored the relationship between flood type and remotely sensed flood detection signal. From communication with the MRC, the Malawi Department of Climate Change and Meteorological Services (DCCMS), and Andrew Kruczkiewicz’s interaction with local communities during a ground-truth mission, it was determined that the Malawi floods in January 2015 consisted of both a series of flash floods and an episode of large scale riverine flooding. During Kruczkiewicz’s two trips to Malawi after the flood period, areas impacted were visited, with special attention placed on interacting with community members regarding the timing of flood onset, the types of flooding, and damage patterns. From this trip and subsequent discussions with MRC and DCCMS, each impacted community (flooded place of origin) was placed into one of three categories: 1. Impacted by flash floods, 2. Impacted by riverine floods, 3. Impacted by a combination of the two.

Based on the flood type analysis, each flood map was re-assessed to gauge skill. Our results show that some products perform better than others depending on which type of flood has occurred.

**Soil Moisture and Precipitation Products**

Part three of the project examined the predictive capacity of soil moisture by analyzing the connections between antecedent flood conditions relative to the spatial and temporal distribution of flash floods and riverine floods, for the study period from December 1st 2014, to January 31st 2015.  Building on specific flood type definitions established in step 2 of this project, we explored the potential to use changes in soil moisture as a predictor for change in risk levels for certain flood types.

The Soil Water Index data provided by Copernicus Global Land Surface soil moisture data were acquired from the ESA/EUMETSTAT produced by the ASCAT Sensor onboard MetOp-A and MetOp-B satellites. Daily soil moisture was downloaded for December 1st, 2014 to January 31st, 2015, at a spatial resolution of 0.1˚ (~11 km) (Kidd et al., 2013). The data were obtained as a GeoTiff file and loaded in ArcGIS. The data have a standard range from 0-200, but are halved to yield percent saturation, i.e. a value of 200 means 100% soil saturation (Kidd et al., 2013).

An unsupervised iso-cluster analysis was conducted with ENVI software to visualize the behavior of soil moisture. The analysis was conducted for an approximately 30-day period before and after the floods to account for natural variability of soil moisture. This analysis led to the delineation of areas with similar soil moisture behavior characteristics (Figure 1a). In addition, these clusters aligned with the distribution of flood types in January 2015, revealing three distinct regions (Figure 1a); The Blantyre region, in which flash floods were the predominant type experienced, the Shire River region, which experienced riverine floods, and the Phalombe region, where both riverine floods and flash floods were experienced. In addition to having distinct soil moisture patterns, each region also has a unique topographical profile, which may impact risk for a certain flood type. The topography of Blantyre (flash floods) corresponds to an urban, hilly landscape while in comparison the Shire River region (riverine floods) is predominantly flat. Lastly, the Phalombe region (experienced flash floods and riverine floods) has both relatively flat areas and areas of complex topography (Figure 1b).

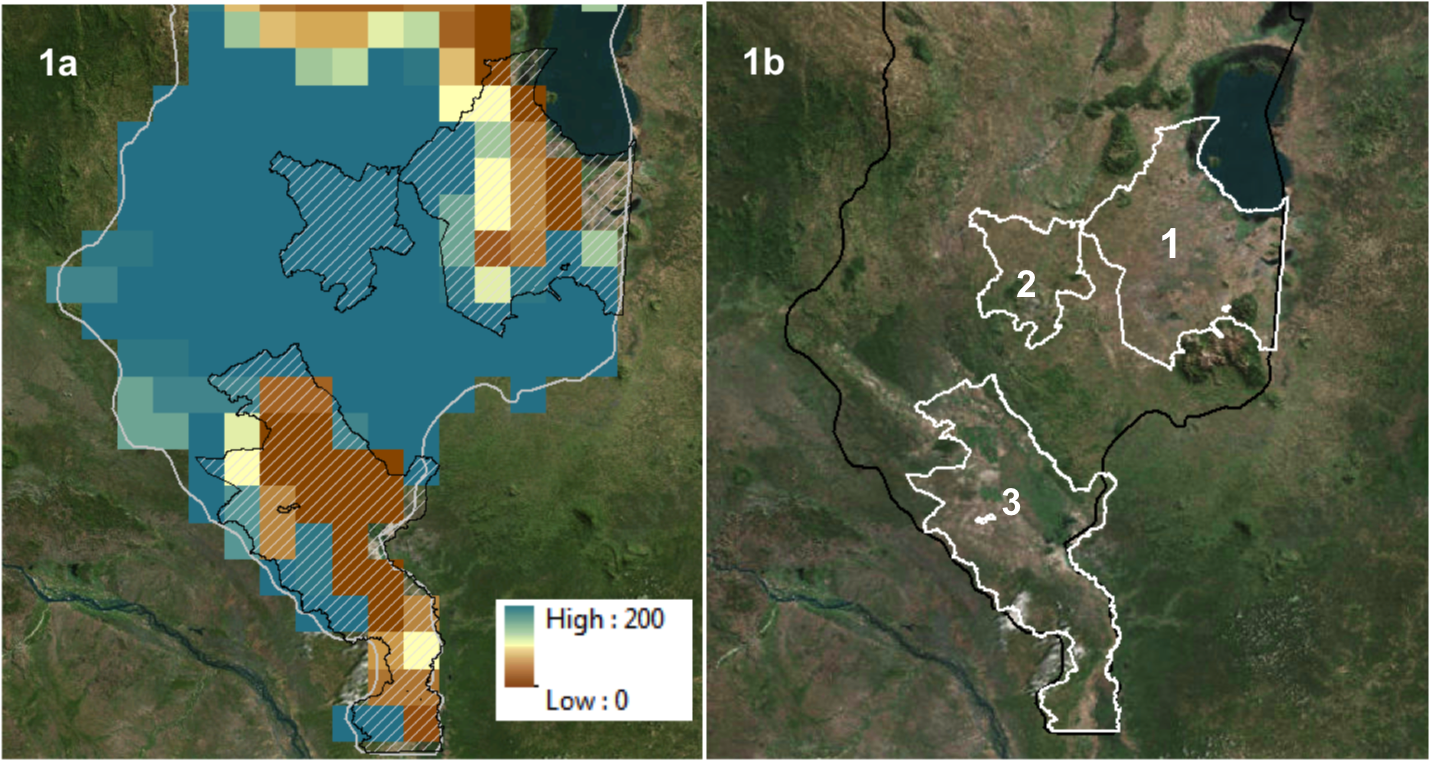


Figure 1b. Regions as defined by similarity in behavior of soil moisture. 1. Phalombe, 2. Blantyre and 3. Shire River on top of ArcGIS imagery base layer.

Figure 1a: Soil Water Index for January 20th, 2015 showing three regions selected to have similar soil moisture characteristics using an unsupervised iso cluster analysis.

Soil moisture was also compared to two monitoring-based precipitation datasets the Tropical Rainfall Measuring Mission (TRMM) and the Climate Prediction Center Morphing Method (CMORPH). Additionally one forecast-based product, International Federation of Red Cross and Red Crescent Society’s Six-Day Extreme Rainfall Forecast (IFRC) was used to assess the predictability of extreme precipitation that lead to floods.

TRMM, CMORPH and IFRC data were downloaded from International Research Institute Data Library as GeoTiff files from December 1st, 2014 to January 31st, 2015 (Blumenthal et al., 2014). TRMM precipitation maps are produced daily at grid values presented at 0.25ºx0.25º latitude--longitude (Cylindrical Equal Distance) global array of points; data range from 0-253 mm/day, value >3 means no precipitation (Huffman and Bolvin, 2014). CMORPH rainfall estimates are produced daily at an 8km spatial resolution. The data is derived from the passive microwaves aboard the SSM/I, AMSU-B, and AMSR-E and TMI on board NASA's Aqua and TRMM spacecraft; data range from 2-252 mm/day, where a value of 2 means no rainfall (Joyce et al, 2004). IFRC Heavy Rainfall Forecast maps are based on a daily ensemble mean precipitation forecast, produced by the National Oceanic and Atmospheric Administration (NOAA). These forecasts are produced at 1° x 1° spatial resolution using the National Centers for Environmental Prediction's (NCEP) Global Ensemble Forecasting System (GEFS) model; data range from 0-253, where values 86-168 means heavy rainfall and converts to top 10% rainfall experienced in the location, values 169-235 is very heavy rainfall which is top 5% rainfall experienced in the location, and values 236-253 represents extremely heavy rainfall which is top 1% experienced in the location (Barnston et al., 2003).

**Time Series**

Using ArcGIS, a time series analysis was conducted over each region to evaluate the relationship between soil moisture, rainfall and flood incidence. First, shapefiles for each region were created using the Traditional Authority Layer obtained from MASDAP. The zonal statistics tool was then used to average soil moisture and precipitation in each selected region on a daily time step from December 1st 2014, to January 31st 2015. Four separate time series (SWI, TRMM, CMORPH and IFRC) were then generated for each region using excel. Figures 2-4 show the comparison between precipitation and soil moisture in each selected region.

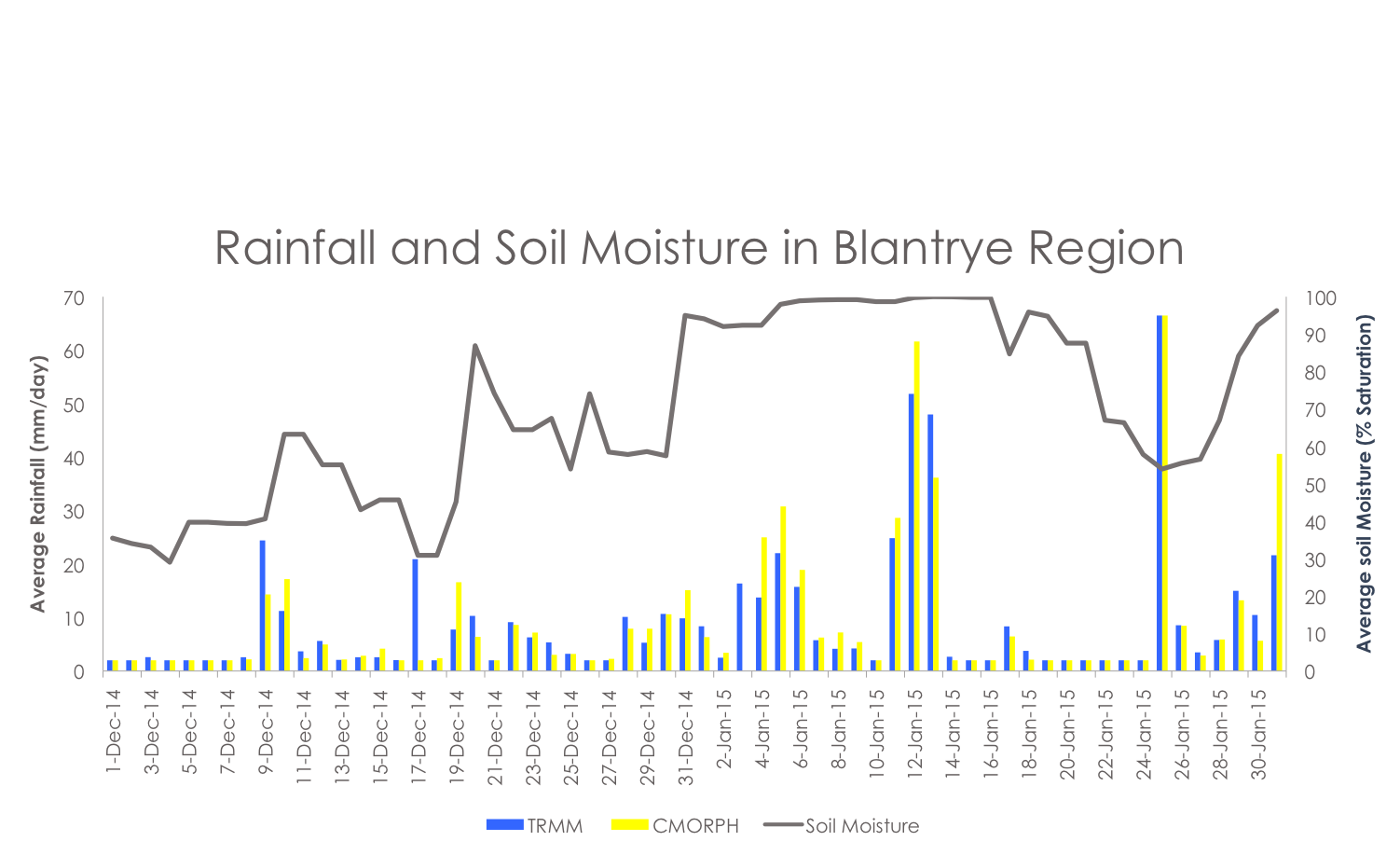
Figure 2. 

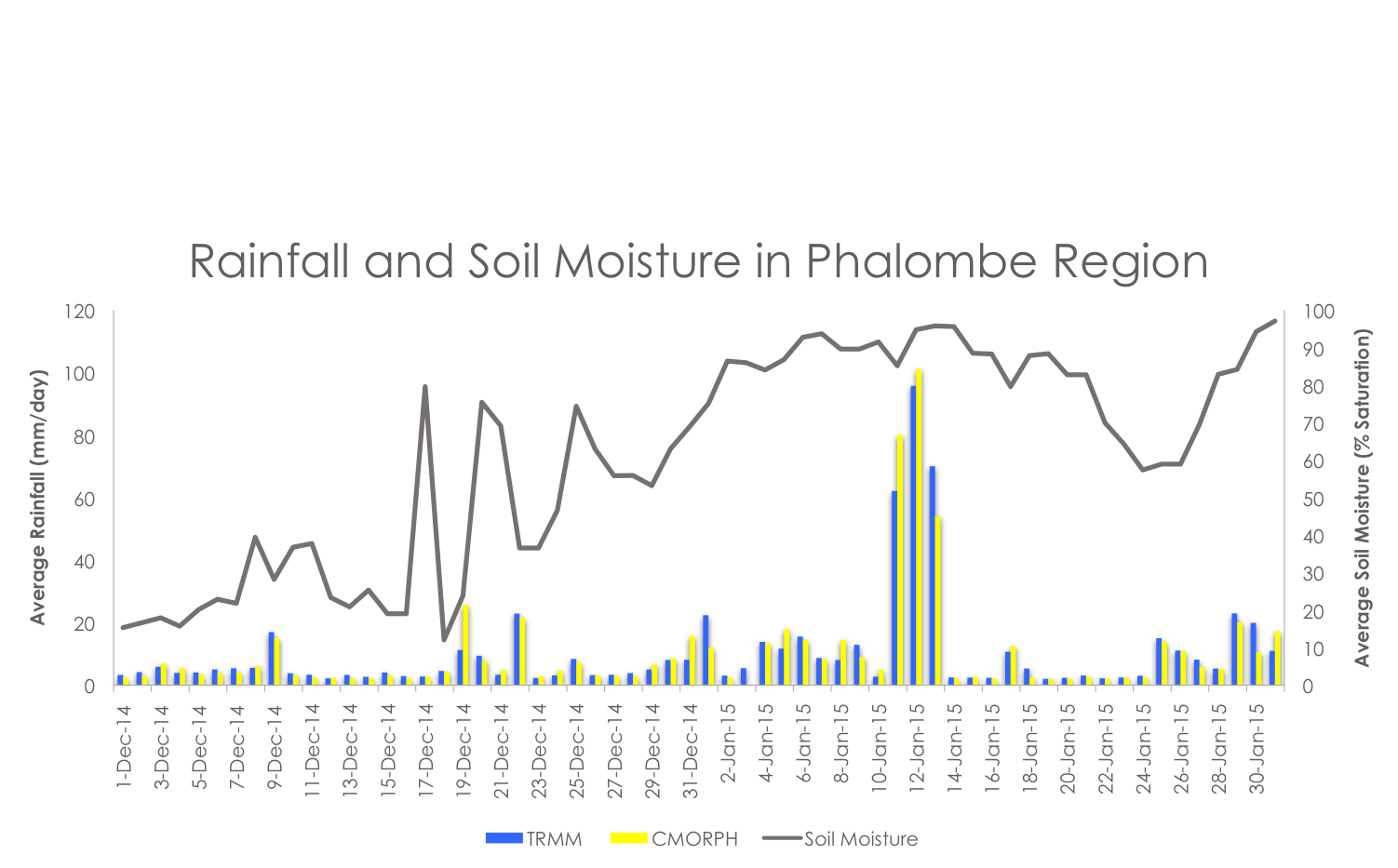
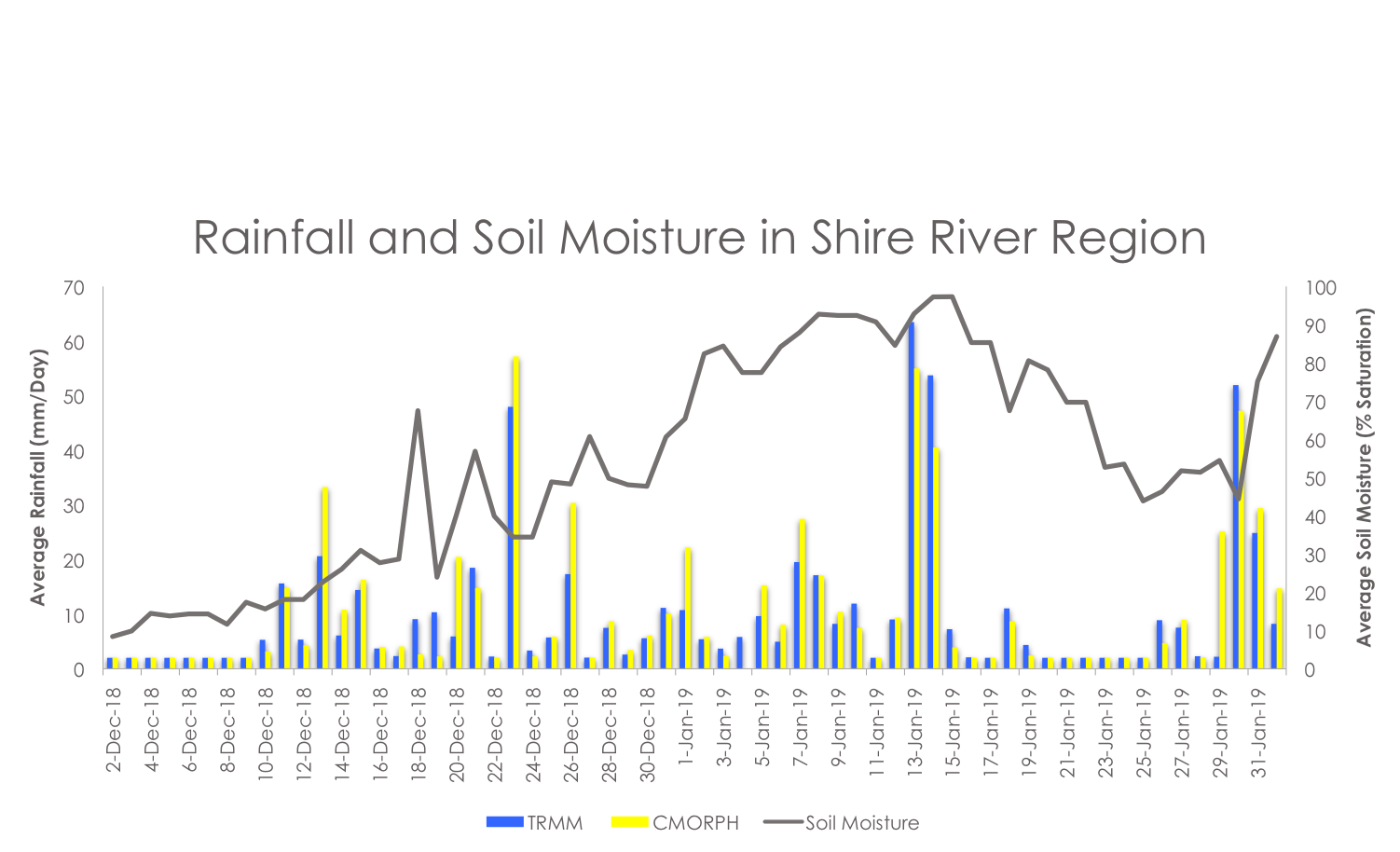
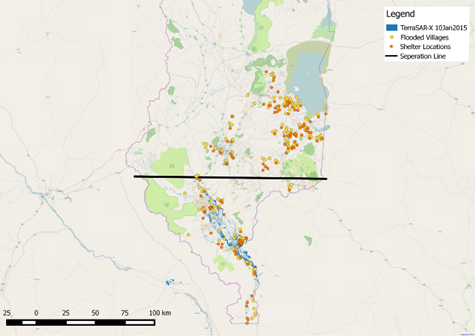
Figure 3.

Figure 4.

****

# IV. Results & Discussion

**Flood Detection Products**

From the qualitative analysis, there was a distinction between products that produced a maximum inundation signal north and south of a latitudinal line approximately through Chikwawa. This delineation approximately corresponded to the shift between the region that experienced flash floods (to the North) and exclusively riverine floods (to the South).

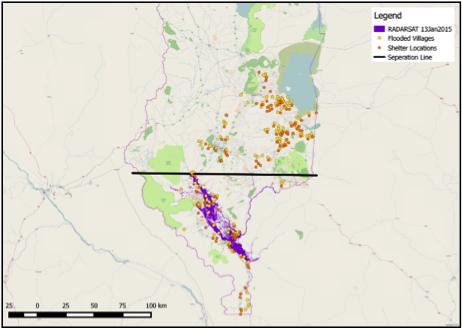
Group 1 satellite products only detected a flood signal in the southern cluster. TerraSAR-X (Figure 2) and RADARSAT (Figure 2) are relatively better at detecting riverine floods. Group 2 products had flood signals in the Northern and Southern clusters. When the spatial relationship between product flood signal and the community impact data were considered, DFO (Figure 4) performed the best. NRT-GFM did not perform well in detecting flood signal near shelter sites in the northern cluster. Group 3 products had flood signal in both northern and southern clusters, but it should be noted that the products are produced in various spatial resolution, with GFMS-FD at 30 km and GFMS-I with 1km resolution. GFMS-FD failed at distinguishing between areas that were flooded versus non-flooded as it exhibited a large spatial extent of false flood signal across most of the study region. GFMS-I performed fair at indicating areas of inundation, but also produced a false signal across non-flooded areas.

Figure 2. RADARSAT product on January 13, 2015 representing flooded area as a purple color. Shelter sites are indicated with orange and flooded villages are in yellow. The black line was drawn to assist spatial analysis.

Figure 3. TerraSAR-X product on January 10, 2015 representing flooded area as a blue color. Shelter sites are indicated with orange and flooded villages are in yellow. The black line was drawn to assist spatial analysis.

To create the flooded village layer, 201 flooded village locations were georeferenced based on ground truth data provided by MRC. Additionally, it should be noted that depending on the base layer used, some flooded village locations and shelter sites appeared to be located on water bodies, indicating either a continuous flux of water levels, the existence of ephemeral water bodies or potential errors in either the data or the base maps. For example, when using Open Street Map as a base layer, some villages appear to be on land when using Google Satellite as base layer, they appear to be located on the water.

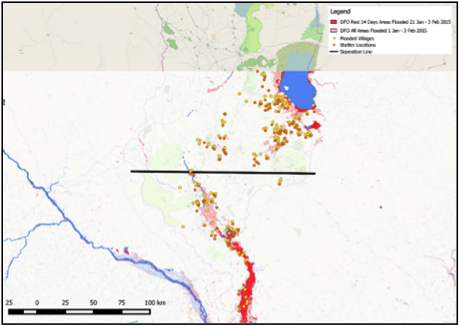
It should also be noted that the analysis was based on a simplified metric of detection vs. non-detection. False alarms, or areas indicated as flooded that were not actually flooded, did not play a large role into what we deemed as a product being ‘skillful’ or ‘performing best’. Further analysis would demand a more quantitative, statistical method to factor in metrics such as false alarm ratio (FAR), probability of detection (POD) and critical success index (CSI), an example of which can be found in Khan et al. 2011.

Figure 4. DFO maximum flood extent map from January1- February 3, 2015, with current flooded conditions indicated in red and all areas flooded in light red. Shelter sites are indicated with the color orange and flooded villages are in yellow. The black line was drawn to assist spatial analysis.

**Soil Moisture and Precipitation Products**

The first selected region (the Phalombe district and surrounding areas) corresponded to the area in Malawi where both flash floods and riverine floods occurred, the second region (the Blantyre district and surrounding areas) had mostly flash floods, and the majority of the third region (the Shire River region surrounding areas) experienced riverine floods.

All three time series for the respective selected regions showed an increase in percent saturation starting around December 31st, 2014. Further, average precipitation in all three time series showed peak rainfall on January 12th and 13th, 2015. On those dates, average soil moisture was also at or near full saturation. Due to the positive relationship between the average rainfall and percent soil moisture saturation, no distinction could be made between flash floods and riverine floods using solely these two types of data. Distinctive difference lies on how quickly soil moisture percent saturation increased around December 31st, 2015. Over the Blantyre region, soil moisture increased on a greater slope until saturation and maintained high levels longer relative to other regions. Further research is necessary, but this behavior may explain heightened flash flood risk. Moreover, the IFRC six-day extreme rainfall forecast accurately forecasted top 1% rainfall experienced in the location for the peak rainfall dates (Tables 1 - 3), but it is important to note that the 1% signal also appeared over non-flooded areas. This suggests that disaster managers can potentially use this product to better anticipate and prepare for flooding events, but will need to understand the limitations.

Table 1. 

Table 2.



Table 3. 

# V. Conclusions

For part one of this project, the flood detection products analysis (TerraSAR-X, RADARSAT, DFO, NRT-GFM and GFMS-I) depicted a wide variance in flood detection signal. When qualitatively compared to ground truth data, TerraSAR-X, RADARSAT, and DFO performed better than others when considering floods in the broader sense.

Upon communication with project partners on ground and from in-person observation by Kruczkiewicz, it was determined that different regions experienced different types of floods. We assessed skill of detection based on flood type and found that detection skills changed relative to the analysis in part 1. Our findings indicated that DFO has the most potential at identifying areas of flash floods while TerraSAR-X, RADARSAT and DFO were most skillful at detecting riverine floods.

For part three of this project, our analysis revealed that the relationship between soil moisture, rainfall and flooding is similar for the two regions (Shire River area and Phalombe area), which included riverine floods. However, within the Blantyre region where flash floods occurred, soil moisture was elevated at a faster rate and for a longer period of time prior to flooding. This finding may indicate that sustained positive anomalies of soil moisture may be used as a predictive indicator of heightened flash flood risk in some areas of Malawi. Furthermore, it is possible that these outputs may support the integration of soil moisture monitoring into a framework for flash flood early warning systems.

It is expected that the results of this study will increase the ability to monitor different types of flood events, directly benefitting organizations involved with disaster relief efforts in Malawi; and potentially allowing for a faster response and better allocation of emergency flood relief efforts. The next steps after this project include developing a multi-faceted early warning system for floods, with the goal of identifying areas of increased flash flood risk, exploring the thresholds related to this risk, and linking these directly to emergency preparedness actions.

# VI. Acknowledgments

Thank you to Dr. Pietro Ceccato for his guidance and counsel on this project. Thank you to Erin Coughlan de Perez, Hastings Kandaya, and Keiko Seito for their help on this project.

This project falls under the NASA Applied Sciences national application area of Disasters. This material [LMC1] is based upon work supported by NASA through contract NNL11AA00B and cooperative agreement NNX14AB60A.

# VII. References

Balbo, S., Boccardo, P., Dalmasso, S., and Pasquali, P. A public platform for geospatial data sharing for disaster risk management. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.* 2013. 40(5):27-28.

Barnston, A. G., S. J. Mason, L. Goddard, D. G. Dewitt, and S. E. Zebiak, 2003: Multimodel ensembling in seasonal climate forecasting at IRI. *Bull. Amer. Meteor. Soc.*, 84, 1783-1796.

Blumenthal, M.B., Mell, M., del Corral, J., Cousin, R., Khomyakov, I. IRI data library, enhancing accessibility of climate knowledge. *Earth Perspectives.* 2014. 1(19):1-12.

Huffman, G.J., and Bolvin, D.T. Real-Time TRMM multi-satellite precipitation analysis data set documentation. 2014. 1-46. <ftp://[trmmopen.gsfc.nasa.gov/pub/merged/V7Documents/3B4XRT\_doc\_V7.pdf](http://trmmopen.gsfc.nasa.gov/pub/merged/V7Documents/3B4XRT_doc_V7.pdf)> Accessed 3 August 2014.

Joyce, R.J., J.E. Janowiak, P. A. Arkin and P. Xie, 2004: CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution. *J. Hydrometeorology*. (5). 487-503.

Kidd, R., Makhmara, H., Paulik, C. (2013): BioPar product user manual soil water index (SWI) version 2.0. <<http://land.copernicus.eu/global/products/swi?qt:swi_characteristics=5>> Accessed 3 August 2015.

Malawi Government. Malawi 2015 floods post disaster needs assessment report. 1-111. <<http://www.preventionweb.net/english/professional/publications/v.php?id=45172&a=email&utm_source=pw_email>> Accessed 3 August 2015.

Malawi Meteorological Services. Climate of Malawi. Ministry of Natural Resources, Energy, and Environment. Department of Climate and Meteorological Service. 2006. <<http://www.metmalawi.com/climate/climate.php>> Accessed 8 Aug 2015.

# IV. Content Innovation

Interactive Plot View

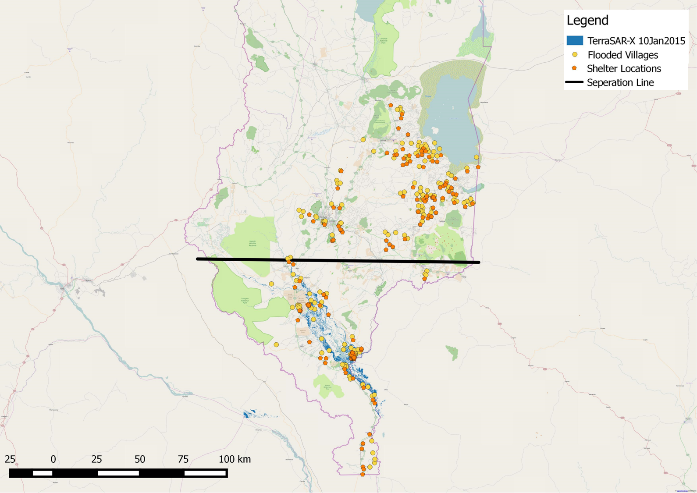
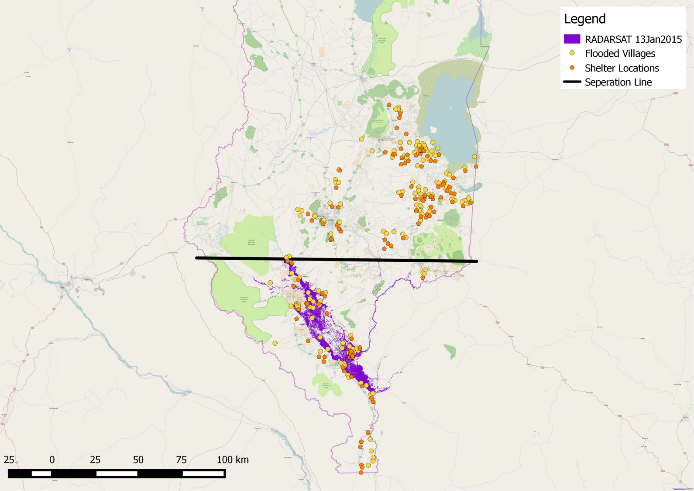
Blantyre Region Soil Moisture for January 2015

file name: 2015Sum\_IRI\_Content\_Innovation\_MapView.csv

Interactive Map Viewer

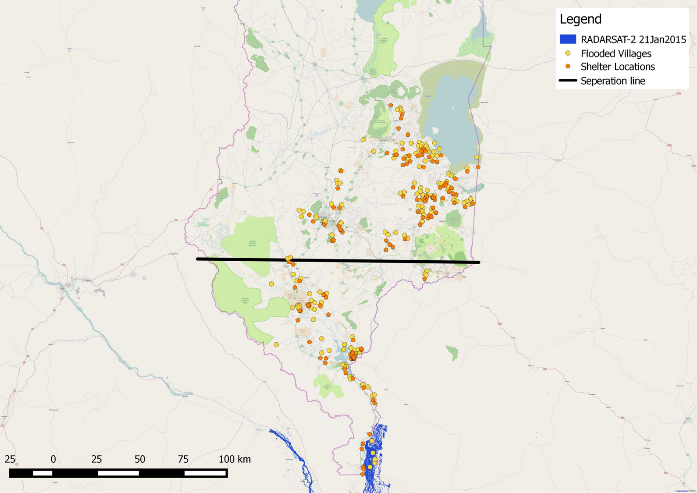
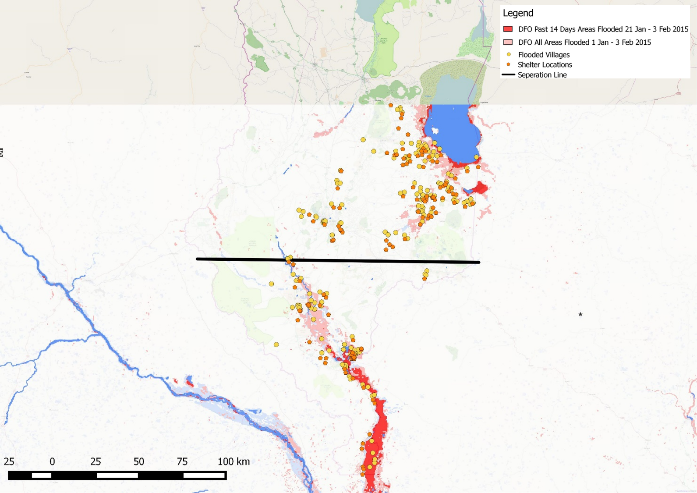
Soil Water Index for January 20th 2015 over southern Malawi.

file name: 2015Sum\_IRI\_Content\_Innovation\_iPlot

**Appendix A.** Comparison and validation of all seven satellite flood detection products.

RADARSAT flood detection product on January 13, 2015 represented with the color bright purple. Shelter sites are indicated with the color orange and flooded villages are in yellow. The black line was drawn to assist spatial analysis.

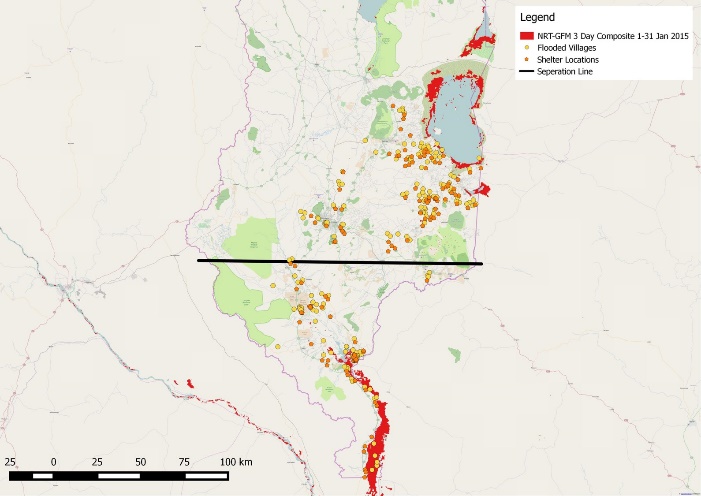
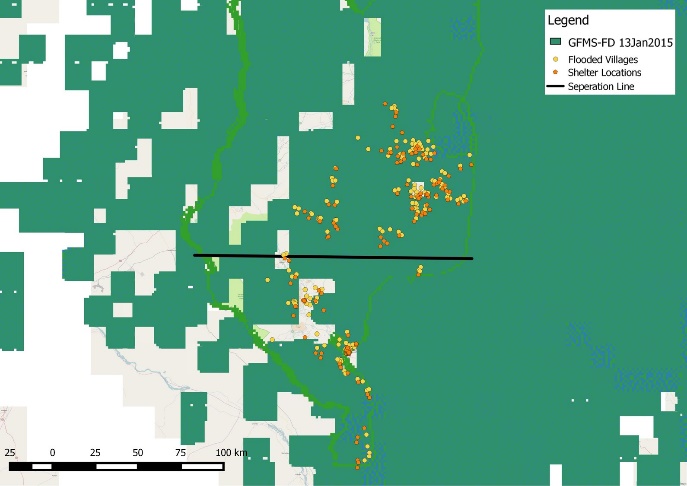
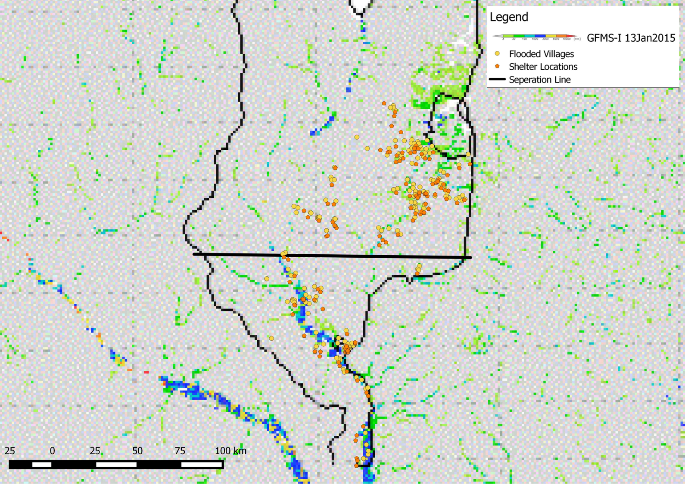
TerraSAR-X flood detection product on January 10, 2015 represented in the color bright blue. Shelter sites are indicated with the color orange and flooded villages are in yellow. The black line was drawn to assist spatial analysis.



DFO maximum flood extent map from January1- February 3, 2015, with current flood indicated in the color red and all areas flooded in light red. Shelter sites are indicated with the color orange and flooded villages are in yellow. The black line was drawn to assist spatial analysis.

RADARSAT-2 flood detection product on January 21, 2015 represented in the color navy blue. Shelter sites are indicated with the color orange and flooded villages are in yellow. The black line was drawn to assist spatial analysis.

Eight GFMS-FD flood detection maps overlaid together in chronological order for maximum flood extent on January 13, 2015, represented in navy green color. Shelter sites are indicated with the color orange and flooded villages are in yellow. The black line was drawn to assist spatial analysis.



Eight GFMS-I flood detection maps overlaid together in chronological order for maximum flood extent on January 13, 2015, represented in rainbow color spectrum. Shelter sites are indicated with the color orange and flooded villages are in yellow. The black line was drawn to assist spatial.

NRT-GFM maximum flood extent map for January 2015, represented in bright red. Shelter sites are indicated with the color orange and flooded villages in yellow. The black line was drawn to assist spatial analysis.

**Appendix B.** Validation of IFRC 6 Day Extreme Rainfall Forecast (Blantrye Area)



**Appendix C.** Validation of IFRC 6 Day Extreme Rainfall Forecast (Phalombe Area)



**Appendix D.** Validation of IFRC 6 Day Extreme Rainfall Forecast (Shire River Area)