ARC Flood Extend Depiction Algorithm Performance Document

AFED Version V05R00
Document Revision R00

Amospheric and Environmental Research, Inc. (AER)
131 Hartwell Avenue
Lexington, MA 02421-3126 USA
Phone: +1-781-761-2288

With contributions by:
John F. Galantowicz, Jeff Picton, and Ben Root

31 August 2018

Revision History

AFED Version	Document Revision	Date dd/mm/yyyy	Description
V01	R00	12/05/2015	Initial draft version for Model Draft 1 delivery.
V02	R00	11/09/2015	Updated version for Model Draft 2 delivery
V02	R02	21/12/2015	Phase I deliverable 4: Performance analysis of model: on-going throughout Phase 1; Updated for AFED V02R02
V02	R02 Amended	02/02/2016	Includes additional commentary for interpretation of results and other changes for corrections and clarity.
V02	R02 Amended	02/26/2016	Added section 2.5 on methods for analyzing temporal representation of the data.
V03	R00	29/03/2016	Updated for Final Historical Model (AFM Phase 1)
V03R01	R00	16/05/2016	Test cases added for AMSR2 and SSM/I. Updated for AFED V03R01.
V03R01	R02	07/06/2016	Updated for SSM/I results with atmospheric correction.
V03R01	R03	02/08/2016	Added case studies of false positives, false negatives, and flood event duration. Corrected and clarified in response to ARC comments. Renamed document and revised terminology to distinguish AFED from the ARC River Flood Model (AFM-R)
V03R01	R04	23/08/2016	Added performance summary and recommendations section. Added sections on Grande Comore flood event false negative and coastal false positive. Made corrections.
V03R01	R05	23/01/2017	Appended *AER Memo: AFED near real time processing system performance* as Appendix A.
V04R01	R00	03/13/2018	Updated for AFED V04R01 with algorithm changes performed under contract QRSA-040-16 Task Order 2. Reorganized performance test results to put AFED results earlier in document and flooded fraction results later (Appendix A). Revised recommendations for future algorithm changes to account for changes already completed. Added Appendix B on differences between V04R00 and V04R01. Removed former Appendix A “AER Memo: AFED near real time processing system performance”, which is moot because, as of AFED V04R00, the NRTPS and HPS use identical software.
V05R00	R00	08/31/2018	Updated for AFED V05R00 with algorithm changes performed under contract QRSA-040-16 Task Order 3. Added Appendix C on V05R00-V04R01 differences.

Table of Contents

1 Introduction

The African Risk Capacity (ARC) flood extent depiction (AFED) is a daily representation of temporarily flooded areas of Africa. The AFED algorithm derives this depiction by processing historical satellite sensor observations acquired from a number of similar space-borne microwave sensors that together provide a nearly continuous daily record since 1992. The purpose of this algorithm performance document (APD) is to summarize the AFED algorithm performance following methods used at AER on other remote sensing algorithm development programs (e.g., Galantowicz, 2001).

A performance assessment provides quantitative and qualitative estimates of errors, inconsistencies, omissions, and other algorithm qualities important for understanding the limits to which the algorithm products can be relied upon in applications. The most reliable method for evaluating AFED errors is through comparison with high-quality independently derived flood maps with similar spatial sampling characteristics (hereafter referred to as *validation* data). This document includes results from algorithm tests with validation data covering well-documented flood events, including descriptions of the events and the sources and derivation methods of the validation data used. Because trustworthy flood extent data are a rarity, only a relatively small sample of algorithm results can ever be objectively tested with validation data. As a result, the document also includes performance measures derived from statistical, qualitative, and graphical analyses of AFED results and other intermediate algorithm products. Performance metrics include:

1) scene classification confusion matrices that identify the percentages of flooded and unflooded scene points that are correctly and incorrectly classified with respect to validation data,
2) minimum detectable flooded fraction statistics constituting an AFED algorithm error model per location and day-of-year,
3) measures of consistency between algorithm products, integrated over flood event durations, and compared from year-to-year,
4) graphical representations of algorithm products evaluated side-by-side against flood imagery from independent sources that provide evidence for the algorithm's ability to capture flood extent and duration on a daily basis and, conversely, reliably represent the absence of flooding, and
5) flooded fraction time series analyses that show the degree to which algorithm results are consistent over time, provide a representation of flood duration, and relate to minimum detectable flooded fraction limits.

The AFED algorithm products discussed here include the AFED product itself and the flooded fraction (FF) product from the microwave algorithm component, which is an intermediate step leading to AFED derivation. All AFED algorithm products and methods are described in detail in the *ARC Flood Extent Depiction Algorithm Description Document* (hereafter ADD; Galantowicz et al., 2018).

This version of the APD provides performance indications for historical AFED results from algorithm version 5 revision 0 (V05R00) covering the period 1998 to 2017 using data from four satellite sensors: (1) TMI[1] (sensor for the 1998-2014 period), (2) AMSR-E[2] (July 2002 through

[1] Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (http://trmm.gsfc.nasa.gov/)
[2] Advanced Microwave Scanning Radiometer-EOS (http://nsidc.org/data/amsre).

October 2011), (3) AMSR2[3] (2012-present), and GMI[4] (2014-present). The SSM/I[5] sensor, which was used in V04R01 and earlier algorithm versions, is no longer included in V05R00. We have retained some V04R01 SSM/I results for reference.

2 Performance testing approach

Performance testing relies foremost on comparisons of AFED algorithm products to similar representations of well-documented flood events in Africa derived from independent sources. In addition to these rigorous tests, we perform case-by-case analyses to characterize AFED performance with regard to false positive and false negative flood detection and flood duration representation. We also analyze flooded fraction temporal and spatial characteristics to assess the continent-scale and long-term self-consistency of this key AFED algorithm component. This section describes: (1) the development of flood event validation scenes from independent sources, the derivation of flooded fraction at microwave sensor resolution from validation scenes, and how validation scenes are used to evaluate AFED and flooded fraction; (2) AFED algorithm test limitations, error budgeting, and self-consistency testing; (3) the effects of data availability and excluded data on temporal representation; (4) case-by-case false positive and false negative tests; and (4) flood duration tests.

2.1 Development of flood event validation scenes from independent sources

Independent flood event depictions at resolutions comparable to AFED can be derived from a number of satellite sources including visible/infrared sensors (e.g., MODIS[6], VIIRS[7], Landsat, AVHRR[8]) or imaging radars (e.g., PALSAR[9], JERS-1[10]). Although many significant events have already been analyzed and archived on public sites (e.g., Dartmouth Flood Observatory[11], NRT Global MODIS Flood Mapping[12]), we have chosen to derive new flood depictions for use as AFED validation datasets in this document. Although more time consuming, direct derivation of validation scenes from sensor data is preferable to use of third-party data for several reasons: it allows for control of flood map resolution; it provides control and documentation of external factors like cloud cover that may not be adequately represented otherwise; and it helps to insure that a consistent methodology is used across all test events.

2.1.1 *Flood events*

In this document we report in-depth analyses of 20 flood events that occurred in Africa between 2000 and 2013. Table 1 summarizes each event, including the abbreviated case name, impacted countries, nearby water bodies, day and time of the validation data analyzed, and brief

[3] Advanced Microwave Scanning Radiometer 2, Global Change Observation Mission 1st-Water (http://www.jaxa.jp/projects/sat/gcom_w/index_e.html)
[4] GPM Microwave Imager, Global Precipitation Measurement (http://pmm.nasa.gov/)
[5] Special Sensor Microwave/Imager (http://www.ncdc.noaa.gov/oa/rsad/ssmi/swath/index.html)
[6] Moderate Resolution Imaging Spectrometer (http://modis.gsfc.nasa.gov/).
[7] Visible Infrared Imaging Radiometer Suite (http://npp.gsfc.nasa.gov/viirs.html).
[8] Advanced Very High Resolution Radiometer (http://www.class.ngdc.noaa.gov/saa/products/search?datatype_family=AVHRR)
[9] Phased Array L-band Synthetic Aperture Radar (http://www.eorc.jaxa.jp/ALOS/en/obs/palsar_strat.htm).
[10] Japanese Earth Resources Satellite-1 (http://www.eorc.jaxa.jp/JERS-1/en/)
[11] http://floodobservatory.colorado.edu/
[12] http://oas.gsfc.nasa.gov/floodmap/

comments about the flood impact. The analyzed date of each flood event was chosen to be near the date of maximum flood extent when cloud cover was also low at the time of the validation scene observations.

Table 1: Analyzed flood events.

Case name	Countries	Nearby water bodies	Date	Min lat.	Max lat.	Min lon.	Max lon.	Reported impacts
Shebelle2006	Ethiopia	Shabele River	2006/11/1	3.25	7.5	42	46.25	60 dead
InnerNigerDelta2007	Mali	Bani and Niger Rivers	2007/9/15	13	16.25	-6.25	-2.75	1M+ affected
InnerNigerDelta2009	Mali	Inner Niger Delta	2009/9/18	13	16.25	-6.25	-2.75	160 dead, 600K affected
Limpopo2000	Mozambique	Limpopo River	2000/3/4	-26	-22	31	35.6	700 dead
Limpopo2013	Mozambique	Limpopo River	2013/1/25	-26	-22	31	35.6	150K affected
Pungue2008	Mozambique	Pungue River	2008/1/3	-20.25	-18.25	33.5	35.25	13K+ displaced
Zambezi2007	Mozambique	Shire and Zambezi Rivers	2007/2/10	-19	-16	34	36.75	29 dead, 60K displaced
Zambezi2008	Mozambique	Shire and Zambezi Rivers	2008/2/4	-19.25	-15.25	34	36.75	95K displaced
Caprivi2003	Namibia, Zambia	Zambezi River	2003/4/28	-18.25	-16.75	23	25.75	10K+ displaced
Caprivi2009a	Namibia, Zambia	Zambezi River	2009/3/27	-19	-16.5	23	25.75	92 dead, 40 year flood, thousands displaced
Caprivi2009b	Namibia, Zambia	Chobe and Zambezi Rivers	2009/4/14	-19	-16.5	23	25.75	
Etosha2009a	Namibia	Etosha Pan	2009/3/23	-20	-16	13.75	18	92 dead
Etosha2009b	Namibia	Cuvelai River	2009/3/28	-20	-16	13.75	18	
Niger2012	Nigeria, Benin, Niger	Niger River	2012/9/25	9.75	14.25	0.4	7	7M affected
Senegal2005	Senegal	Senegal River	2005/10/3	14.5	17.75	-15	-10.75	Thousands displaced
Fitri2007	Chad	Batha River and Lake Fitri	2007/9/18	12	13.5	16.75	18.75	1M+ affected
Kyoga2007	Uganda	Lake Kyoga, Bisina; Okok, Okere, Kelim Rivers	2007/9/18	0.25	3	31.75	35.5	10 dead, 1M+ affected
Barotse2003	Zambia, Angola	Lunega Pinto	2003/3/27	-16.75	-10.25	20.5	24	10K+ displaced, famine
Barotse2007	Zambia, Angola	Zambezi River	2007/3/4	-16.5	-14	20.75	24	Dozens dead
Kafue2008	Zambia	Kafue River	2008/2/14	-16.75	-13	24.5	29.5	Exceptional flooding

2.1.2 *Deriving flood event validation scenes from MODIS imagery*

This section describes derivation of binary (flooded/unflooded) flood maps from Moderate Resolution Imaging Spectroradiometer (MODIS) imagery. The MODIS instrument was carried aboard the NASA Terra (1999-present) and Aqua (2002-present) satellites. We used data from

both instruments and followed the same processing steps. The approach is general and could be applied to other similar sensor data available before 1999 (e.g., Landsat, AVHRR).

For each event list in Table 1, we downloaded the swath granules of MODIS reflectance data (MOD02 or MYD02 dataset) that intersected the affected region. We used the NASA-provided MODIS reprojection tool (MRT) to reproject and bilinearly interpolate MODIS swath data from its native 250-m or 500-m resolution to AFED 3-arcsecond (3″; ~90 m at the equation) geographic World Geodetic System (WGS) 1984 grid. We stored the reprojected data in a set of GeoTIFF files containing one sensor band per file.

Next, we produced a *false color* image of the flood scene designed to enhance the contrast between flooded and unflooded area. Figure 1 shows the false color image of the *Caprivi2009a* flood event along the Zambezi River in Namibia on March 27, 2009. The false color image displays MODIS band-1 reflectances (620-670 nm at 250 m resolution) as red intensities, band-2 reflectances (841-876 nm at 250 m resolution) as green intensities, and *band7* reflectances (2015-2155 nm) as blue intensities. Water can be seen as shades of brown and black and land as shades of green. When present in false color images, clouds are seen as white, gray, or magenta.

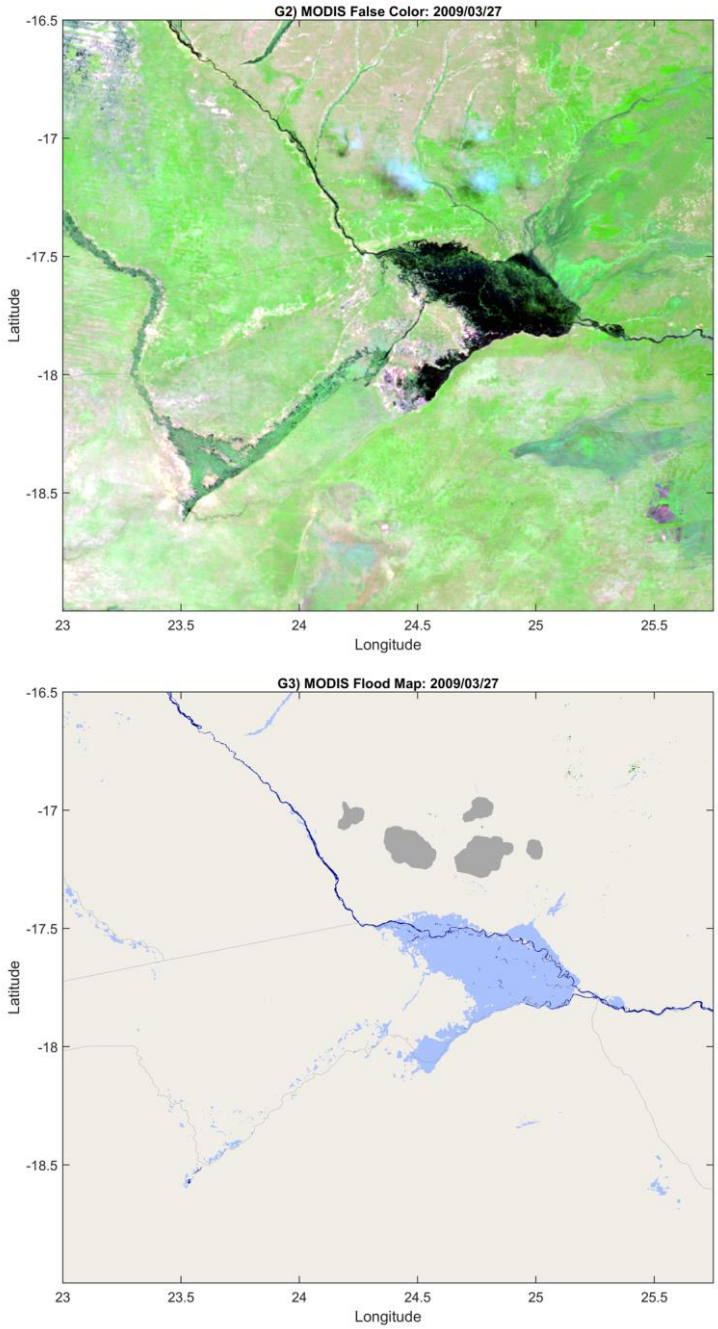


Figure 1: MODIS false color image of Caprivi2009a flood event (top) and AER MODIS-based flood map of Caprivi2009a flood event (bottom), in which light blue is flood water, dark blue is persistent open water, beige is dry land, green is tree cover, and gray is cloud mask.

In order for direct comparison with AFED, we convert false color image data into *binary* flood representations in which pixels are classified as "water" or "not water." Pixels judged to represent cloud or other conditions obscuring the surface (discussed below) are identified as "undetermined" and are to be left out of further analyses. Image classification is accomplished using a custom MATLAB graphical user interface (GUI) that processes and analyzes MODIS *band7* and normalized difference vegetation index (NDVI) imagery. NDVI is computed from

250-m reflectance data as (*band2* – *band1*)/(*band2* + *band1*); *band7* measures 2015-2155 nm reflectance at 500 m resolution. *Band7* and NDVI are used in this analysis because water is typically more distinctly identifiable from reflectance imagery at these wavelength combinations than from true-color imagery. The GUI allows an image analyst to estimate a MODIS-based binary flood map by tuning thresholds on NDVI and *band7* until the resulting binary flood map most effectively balances in-floodplain flood detection, which is more likely to be correct during a flood event, with out-of-floodplain detection, which is generally more likely to be false positive. False positives can be trigger by variety of factors including land cover type and conditions, clouds, and terrain and tree shadows.

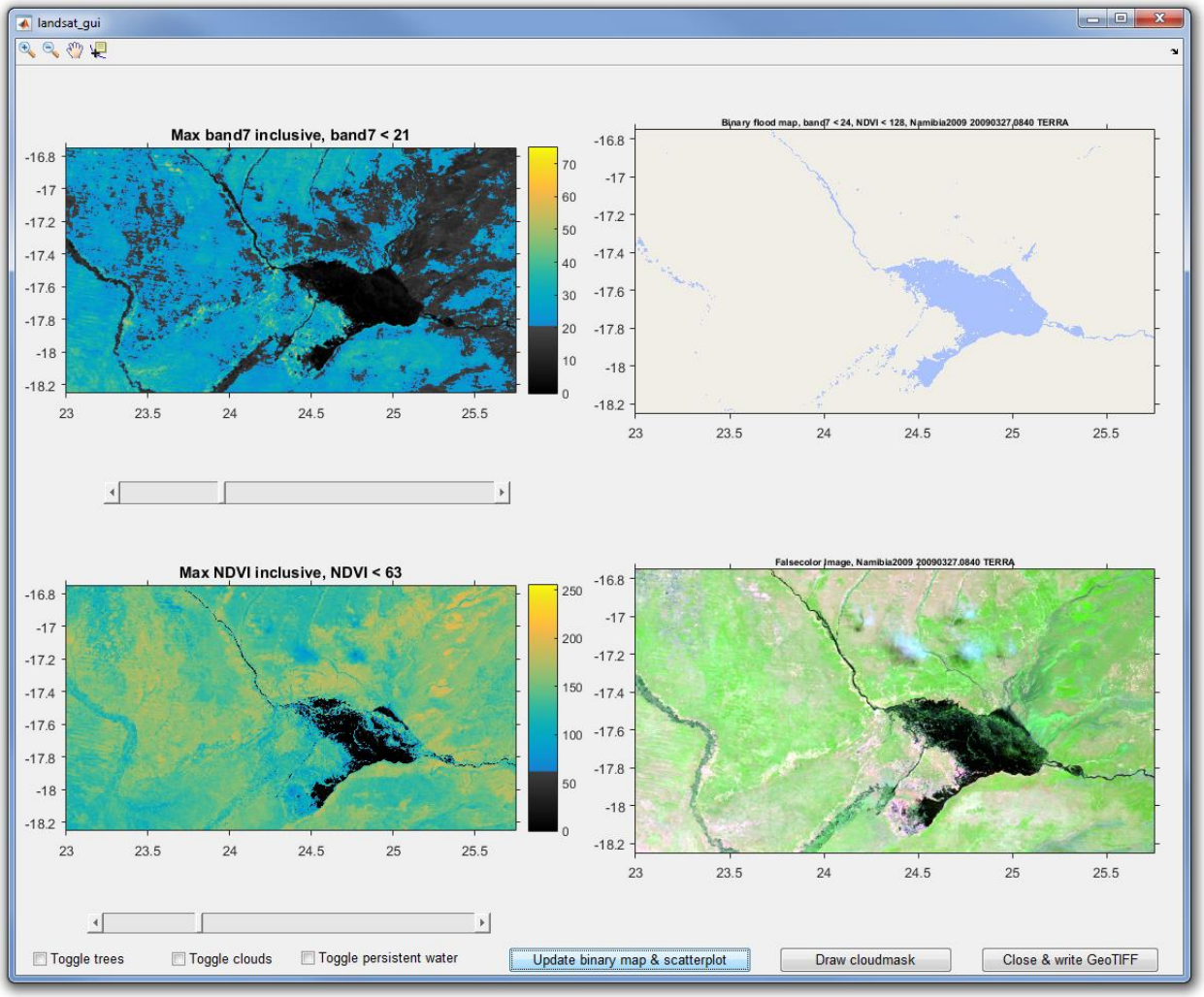


Figure 2: Screen capture of MATLAB flood mapping GUI being used to process MODIS imagery.

Figure 2 depicts the GUI for the *Caprivi2009a* case at the point where the analyst has defined the final MODIS *band7* (upper left image) and NDVI (lower right image) thresholds. (Note that the GUI display scales NDVI and *band7* reflectance values of 0-1 to 0-255.) The GUI shows the resulting flood map in the upper right and the false-color image in the lower right. The final MODIS-derived flood map is shown in Figure 1 with cloud, persistent water, and tree masks included (discussed below). Note the correspondence between areas classified as flooded and the

dark near-black areas of the false-color image. Figure 3 depicts the image analysis as a scatterplot of the image data on an NDVI vs. *band7* axes. Data falling below and to the left of the red line has been classified as flooded based on the *band7* and NDVI threshold values selected by the analyst in the process described above. The threshold values define the inflection point in the red line and the diagonal line segment connects the inflection point to a fixed intercept point at 200 on the NDVI axis. Test results showed that inclusion of the region between the NDVI threshold and the diagonal line segment on the scatterplot flagged more flood plain points as flooded without substantially increasing the flagging of non-flood plain points.

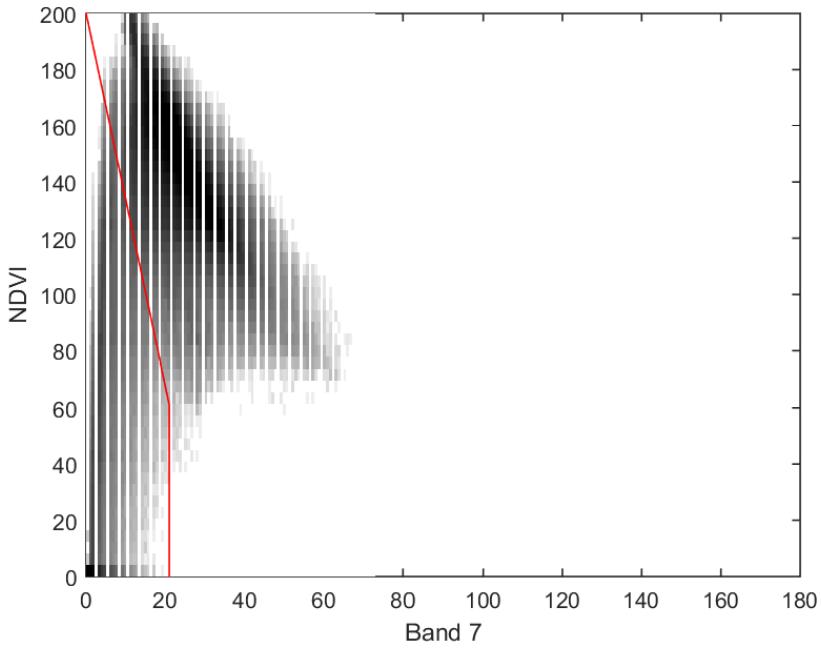


Figure 3: Scatter plot of Caprivi2009a case data showing relationship between NDVI and MODIS band 7 along with line (red) delineating flooded (below and left of line) and non-flooded pixels. NDVI and *band7* reflectance values of 0-1 are scaled to 0-255.

Because clouds block the view of the ground at MODIS *band7* and NDVI wavelengths, special effort was taken to choose flood scenes with clear sky conditions. For cases with unavoidable cloud, the MODIS level-2 250-m resolution cloud mask product (MOD35_L2 or MYD35_L2) is used to identify pixels that are likely to be cloud, cloud shadow, sun-glint, snow, or spurious artifacts that potentially interfere with water detection. The image analyst also used the GUI to manually edit the cloud mask as needed, e.g., when small clouds and cloud shadows were too numerous for a region of the image to be reliably used. Because trees and other vegetation can cause water to be incorrectly classified as land in MODIS imagery, we resampled the Hansen et al. (2014) tree database to the 3-arcsecond grid to create a tree cover mask (as described in the ADD). The following rules for cloud and tree cover are used in the flood classification:

- Any pixel flagged as cloud is marked as “missing” data in the MODIS image classification.
- Any pixel flagged as tree covered and classified as “not-water” in the MODIS image is reclassified as “missing”.

- Pixels flagged as tree covered but classified as water in the MODIS image are to remain classified as water because tree cover is presumed to be a cause of false-negatives, not false positives.

The AFED Land-Water Mask (LWM, described in the ADD) is not used in MODIS image classification, and, as a result, some areas masked as water in the LWM may be incidentally classified as land in MODIS scenes. This type of inconsistency has no effect on AFED performance analyses because all LWM water grid cells are excluded from AFED by definition and therefore are excluded from the analyses here.

Some residual ambiguity can be expected in the classification of flood imagery particularly for marginal areas that are either partially vegetation covered water or spatial mixtures of water and dry land at the native MODIS resolution (250-500 m). A single, robust flood mapping rule that consistently produces accurate flood maps cannot be defined because of variable atmospheric and solar illumination conditions and background dry land spectral signatures. For example, we tested MODIS-derived flood maps against maps derived from same-day 30-m Landsat imagery. After classifying the Landsat imagery at 30-m resolution (Figure 4) we upscaled to compute water fraction at 250- and 500-m resolution and used these data to evaluate MODIS classification results. We found that MODIS false positives were relatively common when compared to the land-water fraction defined from Landsat classification (Figure 5, top) but that MODIS false negatives were relatively uncommon (Figure 5, bottom). Nevertheless, the Landsat and MODIS flood maps consistently agree with respect to the overall location and extent of flooding (Figure 4). Therefore, for the purpose of AFED performance evaluation, we assume that a MODIS pixel classified as flooded is at least 50% water covered and that the distribution of partially water-covered pixels is balanced between false-positives where water cover is less than 50% and false negatives where water cover is greater than 50%.

Section 3.3 includes the MODIS-derived flood maps for all validation cases listed in Table 1.

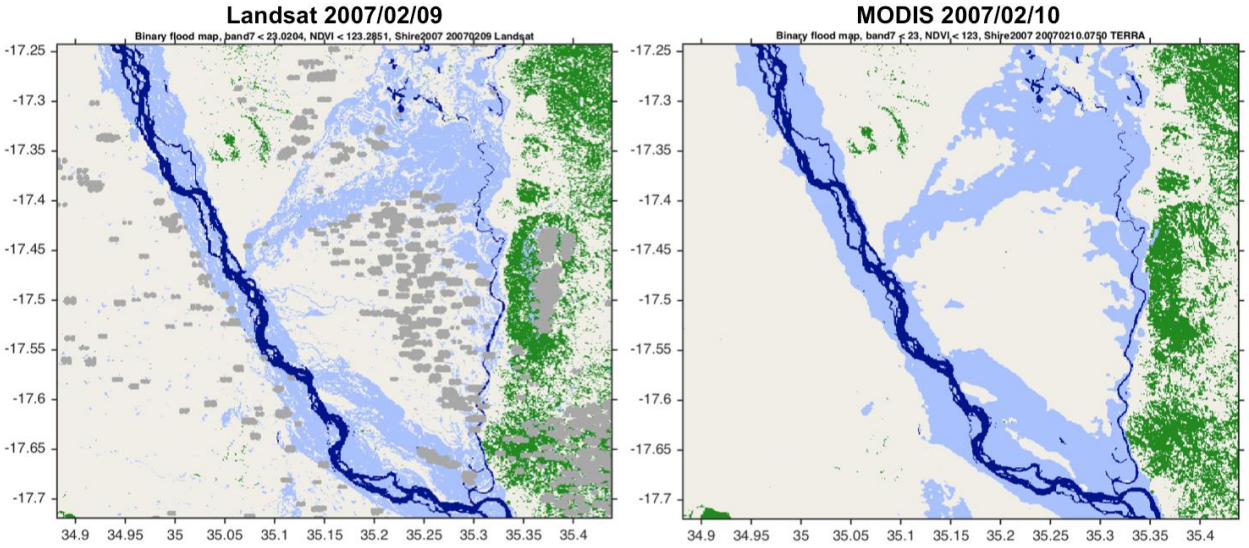


Figure 4: Landsat- (left) and MODIS-derived flood maps. Light blue: flood; dark blue: persistent water mask; gray: clouds; green: tree cover.

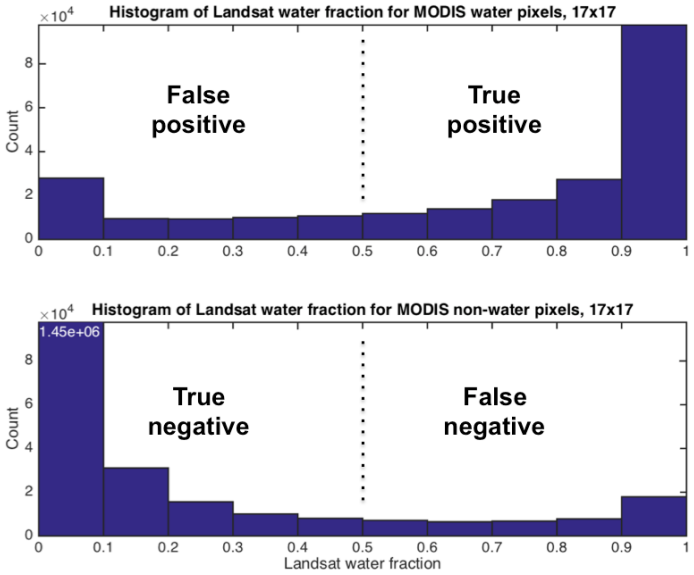


Figure 5: Histograms counting number of MODIS pixels classified as water (top) and land (bottom) at varying water fractions calculated from 17x17 groups of Landsat pixels (~500 m) coincident with the MODIS pixels. Data are from a larger scene including the subregion shown in Figure 4.

2.1.3 *Validation scenes from other sources*

No validation scene from sources other than MODIS are included in this version of the APD. Additional sources could provide an alternative view of AFED performance with different validation data uncertainty characteristics. For example, Landsat imagery gives higher-resolution (30 m) flood depictions with sensor bands similar to MODIS but with additional challenges such as 16-day temporal sampling that makes it more difficult to find cloud-free scenes during flood events. Radar imagery can provide validation scenes regardless of cloudiness but present a different set of interpretation challenges than the optical-infrared sensors and radar scene acquisition has historically been irregular. By focusing on MODIS for validation scenes, we were able to identify, acquire, and interpret a larger set of scenes representing different flood types and terrain conditions than would have been possible if we had also analyzed scenes from other sources.

2.2 Deriving flooded fraction at microwave sensor resolution from test scenes

Following production of the MODIS-derived flood maps for each event, we derived flooded fraction (FF), cloud fraction, and other metrics for evaluation of algorithm results at microwave sensor resolution. Each of these quantities is derived using our footprint averaging process, which effectively upscales mapped quantities from a higher-resolution 3-arcsecond grid (e.g., the pixel grid implicit in the MODIS-based binary flood map in Figure 1) to the lower-resolution of microwave sensor data. For each upscaled quantity, the footprint averaging process uses a two-dimensional Gaussian footprint weighting function with a half-peak diameter identical to that of the microwave data (i.e., 50-km for SSM/I and 22-km otherwise). The weighting function is applied once for each point on the same 5-arcminute grid used by the AFED microwave flooded fraction algorithm. Figure 6 shows the resulting flooded fraction data derived by this process from the MODIS-based binary flood map for the *Caprivi2009a* event (Figure 1).

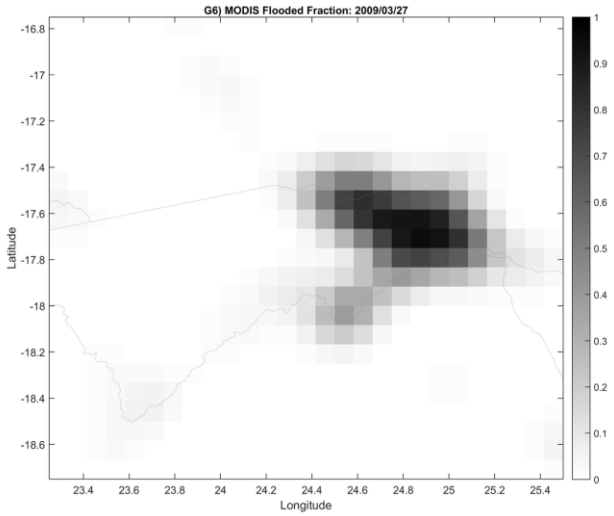


Figure 6: Footprint weighted flooded fraction from MODIS flood map of Caprivi2009a case.

2.3 Evaluating AFED and flooded fraction using test scenes

This section describes the methodologies used for evaluating AFED and flooded fraction results using the validation scene data described in the sections above. Test results based on validation scenes are presented in section 3. Note that all tests treat validation scenes as representations of "true" flood extent, which disregards the degree to which uncertainties in the validation scenes themselves affect the algorithm performance results.

AFED evaluation

We use the confusion matrix (also known as the contingency table) to quantify the performance of the AFED binary flood maps with respect to validation scene maps. The confusion matrix expresses the rate at which points in one of the validation scene groups (water or non-water) are also classified as water or non-water in the AFED product:

- The **true positive rate** (or "hit-rate") is the percentage of pixels classified as water by the AFED algorithm, given that they are classified as water in the validation scene.
- The **false-positive rate** (or "false-alarm" rate) is the percentage of pixels classified as water by the AFED algorithm given that they are classified as not-water in the validation scene.
- The **false negative rate** is the percentage of pixels classified as not-water by the AFED algorithm, given that they are classified as water in the validation scene.
- The **true negative rate** is the percentage of pixels classified as not-water by the AFED algorithm, given that they are classified as not-water in the validation scene.

Additional statistics provided include:

- The **accuracy**, defined as the percentage of all pixels for which the AFED algorithm classification agrees with the validation scene.
- The percentage of all pixels in the scene classified as water by the AFED algorithm.
- The percentage of all pixels in the scene classified as water in the validation scene.
- The total number of pixels in the scene classified as water by the AFED algorithm.
- The total number of pixels in the scene classified as water in the validation scene.
- The total number of pixels in the scene.

In computing these statistics, we restrict the data to the portion of the validation scene that is cloud free and has a validation scene flooded fraction of at least 5%. This restriction isolates the flood-affected area and excludes large, unflooded areas that would otherwise dominate the sample. We provide aggregate statistics over all validation scenes but because of the limited number and geographic variability of the scenes these aggregate statistics should not be interpreted as representative of overall performance.

Flooded fraction evaluation

We use several metrics to quantify the similarity of the validation scene and microwave algorithm-based flooded fractions. All cases exclude points where the validation scene is more than 5% cloud covered:

- Pearson correlation coefficient: Correlation of validation scene flooded fraction (FF_V) versus microwave-based FF (FF_M) computed for n FF points, where.

$$\rho(FF_V, FF_M) = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{FF_{Vi}-\mu_V}{\sigma_V}\right)\left(\frac{FF_{Mi}-\mu_M}{\sigma_M}\right)$$

  where μ_V and μ_M are the sample mean of FF_V and FF_M, respectively, and σ_V and σ_M are their standard deviations.
- RMSE, overall: Root mean squared error for n FF points in the domain, where

$$RMSE = \frac{1}{n}\sqrt{\sum_{i=1}^{n}(FF_{V\,i} - FF_{M\,i})^2}$$

- RMSE, validation scene FF > 5%: RMSE for the set of points with validation scene FF greater than 5%. This final RMSE is a better indicator of microwave algorithm performance in the vicinity of a flood because it excludes the large portion of flooded fraction points in each validation scene with near-zero values.

We also report the fraction of the scene remaining after accounting for validation scene cloud masking. This provides a sense for the overall quality of the validation scene: scenes with lower "unmasked" fractions are more widely cloud-contaminated and therefore less reliable indicators of AFED performance. As noted above, in computing these statistics we restrict the data to the portion of the comparison domain where the flooded fraction is at least 5%.

2.4 AFED algorithm supplementary analyses

We ran several types of supplementary analyses to mitigate the limitations of AFED testing with MODIS scenes and provide a more complete picture of overall algorithm performance. Although the set of MODIS test scenes in Table 1 is large, several factors limit the ability of such tests to completely characterize the end-to-end performance of the AFED algorithm. First, the validation data (e.g., from MODIS and the other data sources used in the analyses) have their own error profiles related to resolution, interference from vegetation and clouds/atmospheric effects, and classification ambiguities (e.g., van Leer and Galantowicz, 2011). Second, high-quality validation imagery is rare, which means that: (1) long-duration flood events may only be imaged once and short-duration events are unlikely to be imaged at all, (2) images may not cover the entire flood-affected area, (3) flooding in a region may be imaged at only one time of year (and AFED algorithm performance may depend on time of year), and (4) many areas may never be covered by validation imagery. Results from the following analyses—included in section 3

and Appendix A—address some of these deficiencies by assessing elements of AFED algorithm behavior not otherwise tested.

- Included in section 3:
  - *Day-to-day AFED consistency during flood events*: We expect day-to-day AFED evolution during a flood event to exhibit several qualities, for example: 1) the highest number of flooded days per pixel should be concentrated in central parts of the flood plain, depressions, and/or slow-draining areas of wetlands and 2) flooded areas should propagate of from high to low river reaches. To test for these qualities, we assemble AFED maps spanning the evolution of individual flood events into image sequences and inspected them by stepping through the images one-by-one and by viewing them as animations. We also compute the cumulative number of days flooded during selected flood events.
  - *Tests of consistency between AFED and flooded fraction:* This test assesses whether AFED is consistent with the flooded fraction inputs used to produce it. To conduct this test, we computed flooded fraction from AFED using the upscaling process discussed in section 2.2. We then created difference maps comparing this upscaled AFED flooded fraction to the flooded fraction from the microwave algorithm; that is, we compared the large-scale flooded fraction of the AFED product to the input flooded fraction from which it was derived. The test shows whether the algorithm is behaving as designed, i.e., producing aggregate AFED flooding corresponding to microwave-derived flooded fraction in time and location without spatial artifacts from the flooded-fraction-to-AFED downscaling process.

- Included in Appendix A
  - *Minimum detectable flooded fraction*: The microwave flooded fraction algorithm applies the minimum detectable flooded fraction (MDFF) threshold as a final step before production of the daily flooded fraction product and the beginning of the process that downscales flooded fraction to produce AFED. The AFED algorithm computes MDFF for each day of year as described in the ADD. The MDFF algorithm uses statistical analysis of day-to-day historical variability (at approximately the 3-sigma level) and spatial analysis to compute the threshold level above which flooded fraction is deemed to be most reliable. MDFF is computed locally for each day of year and therefore encompasses dependence on seasonal vegetation cover, meteorology, and other geographic factors. We review maps of the annual minimum, mean, and maximum MDFF as an indicator of areas and seasons where AFED is more/less likely to positively detect flooding than trigger false positives (lower/higher MDFF).
  - *Year-to-year flooded fraction consistency tests:* The AFED algorithm uses daily flooded fraction data derived from microwave data to compute the AFED product. It is important that the flooded fraction be consistently derived over time so that sensitivity to flooding and algorithm error characteristics do not change over the historical AFED analysis period or between the historical period and future near-real-time AFED processing. To test year-to-year

consistency, we count the number of days per year that flooded fraction exceeds MDFF. We assess the data for year-over-year trends or anomalous years that may indicate a change in AFED algorithm performance not attributable to natural variation.

2.5 Temporal representation: Microwave data availability and excluded data

When flood conditions are evolving day-by-day, an understanding of the timing of remotely sensed observations is needed to assess the accuracy of the resulting flood maps in depicting maximum flood extents. The AFED microwave algorithm makes daily flooded fraction estimates from the best and most timely data available at each point in the algorithm's 5-arcminute microwave data grid. Daily changes in satellite paths over Africa, sensor operational status, and local rainfall and environmental conditions affect the number and timing of observations available for use. The AFED algorithm produces a daily set of metadata files that record information about sensing and environmental conditions useful for concurrent performance assessment. This section describes methodologies that can be employed by AFED users to characterize the temporal representation of daily flooded fraction and AFED products from these metadata.

To reduce the effect of random measurement error, the microwave algorithm computes a daily flooded fraction estimate from a set of individual initial estimates made on the nominal AFED product date and the two preceding days. Typically, there are two to three passes per satellite each day for up to seven separate initial flooded fraction values in a three-day period[13] When two satellite are operating simultaneously (e.g., TMI and AMSR-E or GMI and AMSR2) there may be up to 13 total passes per day. The AFED flooded fraction algorithm—described in detail in the ADD—takes a weighted average of all high-quality initial flooded fraction values available in a three-day period to arrive at the daily flooded fraction estimate used for AFED. At times, no high-quality data are available from a three-day window, and in that case the algorithm proceeds using the last completed daily flooded fraction estimate from a prior day.

With this algorithm logic in mind, the following statistics are needed to characterize the temporal representation of the daily flooded fraction product. Each statistic is available from metadata product files described in more detail in the ADD:

1. **Total N metadata --** *Total number of instantaneous estimates used in a daily estimate:* With valid values from 0 to 13, this metric helps to indicate the degree to which (a) random measurement error may have been reduced and (b) shorter-period changes in flooded fraction may have been smoothed over by the averaging process. The metric also indicates (with a value of 0) when a prior daily estimate was used because no high-quality data were available within the 3-day algorithm window.
2. **Effective Time metadata --** *Effective date-time of flooded fraction estimate, computed as the weighted average of the observation times of all instantaneous flooded fraction values used in a daily estimate:* This statistic represents the day-to-day variation in effective timing of the final estimate within the three-day algorithm window; if a three-day estimate is unavailable, the statistic gives the effective time of the last valid estimate. This statistic is provided as a time offset in decimal days

[13] The AFED algorithm uses data from satellites that cross Africa traveling south-to-north (*ascending pass*) in one part of the day and north-to-south (*descending pass*) in another part of the day. The algorithm collects the ascending pass and descending pass data from each sensor in groups spaced about twelve hours apart from each other.

relative to 0000 UTC on the AFED product date. For example, an effective date-time value of -0.25 days for 20-Oct-2012 would be equivalent to 1800 UTC on 19-Oct-2012.

3. **Reason metadata --** *Explanation for why data was unavailable or excluded:* This metric enumerates the primary reason data points were unavailable or excluded for each satellite overpass each day, e.g., because the satellite did not pass over an area or rain obscured the surface during the pass.

In addition to these location-specific statistics, a daily schedule is included with historical AFED products indicating which sensor provided data to the AFED algorithm each day for the entire domain.

2.6 Case-by-case false negative and false positive tests

A false negative occurs when flood impacts are reported but AFED does not indicate flooding at the same time and place as the report. A false positive occurs when AFED indicates flooding but there is no independent evidence from the area that flooding occurred. False positives and false negatives are difficult to test definitively because independent flood data with reliable spatial and temporal specificity are rarely available. For example, without accurate validation data it is difficult to conclude whether the AFED algorithm failed to detect a large flood (i.e., one whose flooded fraction should have exceeded the minimum detection threshold, MDFF, but did not) or the flood itself was too small or brief to be detected.

To give some insight to AFED performance with regard to false positives and false negatives, we evaluated seven apparent false negative and three apparent false positive cases reported by ARC originally with regard to AFED version V03R01 (Table 2). For AFED V04R01[14], we reanalyzed these cases two ways. First, we acquired up to twice-daily MODIS data over the time of the reported event and for periods before and after and made side-by-side comparisons of MODIS false-color imagery (described in section 2.1.2) to AFED maps covering an identical area on the same date. We assessed the MODIS images for signs of flooding and how well they corroborated AFED results temporally and spatially. Second, we selected points on the microwave data grid within the area where flooding was indicated either by the MODIS imagery or AFED. For these locations, we plotted the time series of microwave data and flooded fraction results to put the case reports in temporal context and investigate the degree to which the results may have been affected by AFED algorithm components—i.e., the dry-land end-member definition, false positive flash detection, and the MDFF—or the characteristics of the flood signal in the microwave data itself, including day-to-day and year-to-year noise and signal timing.

In addition to the cases in Table 2, for AFED V04R01 we conducted two separate studies focused on coastal areas. In the first study, we analyzed brief AFED flooding recurrent in several years in the coastal zone in Cote d'Ivoire and Ghana and similar flooding in Namibia. These studies were based on ARC observations that the spatial pattern of AFED V03R01 flooding (i.e., just along the coasts) and its occurrence only in certain years suggested that the events were false positives. In the second study, we evaluated a false negative associated with a flood report on Grande Comore in April 2012.

[14] Case studies were not revisited for V05R00. Other tests indicate that overall V05R00 flood detection performance is similar to V04R01 and therefore the case study results are likely to remain applicable for V05R00.

Table 2: Apparent false negative and false positive cases analyzed with AFED V04R01 data

Case	Type	Key Location	Area Investigated	Reported Dates
Ouagadougou, Burkina Faso, Sept. 2009	False negative	Ouagadougou: 12.375° N, 1.542° W	12-13.5° N 2.3-0.6° W	Early Sept. 2009
Tintane, Maritania, August 2007	False negative	Tintane: 16.375° N, 10.208° W	16-17° N 10.5-9.5° W	7 Aug. 2007 - 5 Sept. 2007
Agadez, Niger, Sept. 2009	False negative	Agadez: 16.958° N, 7.958° E	15.5-19.0° N 5.3-8.3° E	Early Sept. 2009
Niamey, Niger, August 2010	False negative	Niamey: 13.524° N, 2.042° E	11.25-14° N 1.7-4.8° E	10 Aug. 2010
Busia District and the River Nzoia, Kenya, 2003 & 2008	False negative	Near Bujwane: 0.125° N, 34.042° E	0-0.5° N 33.9-34.3° E	29 Aug. - 12 Sept. 2003 13 Nov. 2008
Tana & Sabaki Rivers, Kenya, 2006	False negative	Near Wenje: 1.792° S, 40.125° E	3.8-1.1° S 39.6-31.4° E	April-May 2006
Greater Accra, Ghana, 2010	False negative	In Greater Accra: 5.708° N, 0.208° W	5.3-6.2° N 0.7° W - 0.8° E	22 Jun. 2010
Isiolo region, Kenya, 2009	False positive	In Isiolo County: 1.292° N, 39.208° E	0-3.5° N 37-40 ° E	October 2009
North Eastern region, Kenya, 2011	False positive	In Wajir County: 2.708° N, 39.292° E	1.5-5 ° N 37.5-41 ° E	May 2011
North Western region, Kenya, 2011	False positive	In Turkana County: 3.792° N, 34.708° E	2-4.75 ° N 33.75-36.5 ° E	May 2011

2.7 Flood duration tests

Flood duration is typically measured in terms of the time period that a river exceeds a defined stage height at a documented stream gauge location. For a spatially distributed flood representation like AFED we can use two types of duration measures. First, we can examine the flooded fraction at a microwave grid point and measure the duration as the period of time between the dates at which flooded fraction first exceeds and later falls below the MDFF threshold. This measure is complicated by the fact that flooded fraction represents flooding over an area that may include different streams and different sections of the same stream that are at different flood stages. Also, flooded fraction may cross the MDFF threshold several times during an apparent flood event, making a definitive duration measurement difficult. Second, we can examine daily AFED maps and determine flood duration regionally based on the dates when flooding first and last appears in the region. This method can be complicated by wetlands and other depressed areas that frequently impound water, falsely indicating that disastrous flooding is longer lasting.

For V0R01[15], we assessed flood duration performance in two case studies (Table 3). For each case, we analyzed flood duration from flooded fraction at selected grid points and side-by-side comparisons of daily AFED maps to MODIS false-color imagery as described in section 2.6.

[15] Flood duration studies were not revisited for V05R00. Other tests indicate that overall V05R00 flood detection performance is similar to V04R01 and therefore the case study results are likely to remain applicable for V05R00.

Table 3: Flood duration test cases

Case	Area Investigated	Dates
Shire River, Malawi and Mozambique, 2015	17.8 – 15.8° S 34.65 – 35.6° E	Dec. 2014 – Jun. 2015
Upper Zambezi River, Angola and Zambia, 2007	16.75 – 10.25° S 20.5 – 24.0° E	Nov. 2006 – Oct. 2007

3 Performance test results

This section presents results of performance tests for the AFED product as well as the flooded fraction produced as an intermediate product by the microwave algorithm. The tests are designed to provide evidence of the accuracy with which the algorithm captures the extent and duration of historical flood events large enough to have required international assistance.

Section 3.1 presents self-consistency test results in which the data are evaluated for realism in their temporal and spatial depiction of flood events and non-flood conditions. (Flooded fraction self-consistency results are in Appendix A.) Section 3.2 presents statistical results covering the flood event test cases described above. Section 3.3 presents graphical results for each of the flood event test cases, including comparison of AFED maps to the independently derived flood maps, comparison of microwave algorithm and validation scene flooded fraction results (including results without application of the minimum detectable flooded fraction limit), and comparison of microwave flooded fraction results to flooded fraction derived by upscaling the AFED product.

3.1 Self-consistency test results

Self-consistency tests use AFED and other algorithm data products to assess how well the algorithm performs relative to criteria inherent in the AFED algorithm design. These criteria may include spatial and temporal statistical measures that reveal processing artifacts, i.e., product features that are unlikely to be the result of natural conditions. This section includes AFED self-consistency tests. Tests of upscaled AFED consistency with microwave-derived flooded fraction are provided in section 3.4. Other flooded fraction self-consistency tests are covered in Appendix A.

3.1.1 *Day-to-day AFED consistency during flood events*

Figure 7 through Figure 15 show the V05R00 AFED cumulative number of days flooded for 11 selected flood events. (AFED animations of these events are available separately as GIF or mp4 format files). In most cases we selected days when little or no AFED flooding was evident as the beginning and end dates for each event. We observed the following characteristics:

- Down-river flood propagation (evident in animations): Benue River 2006, White Volta 2007, Inner Niger Delta 2007, Shebelle River 2006.
- Flooding concentrated in river flood plains: Benue River 2006, Bago River 2007, White Volta River 2007, Oti River 2007, Senegal River 2007.
- Flooding concentrated in river flood plains and known wetlands: Inner Niger Delta 2007, Zambezi and Shire Rivers 2007 and 2008.
- Flooding in flood plains and nearby depressions: Senegal River 2003, Betsiboka and Mahajamba Rivers 2007.
- Flooding in flood plains accompanied by possible false-positive flooding in surrounding areas: Senegal River 2003, Shebelle River 2006.

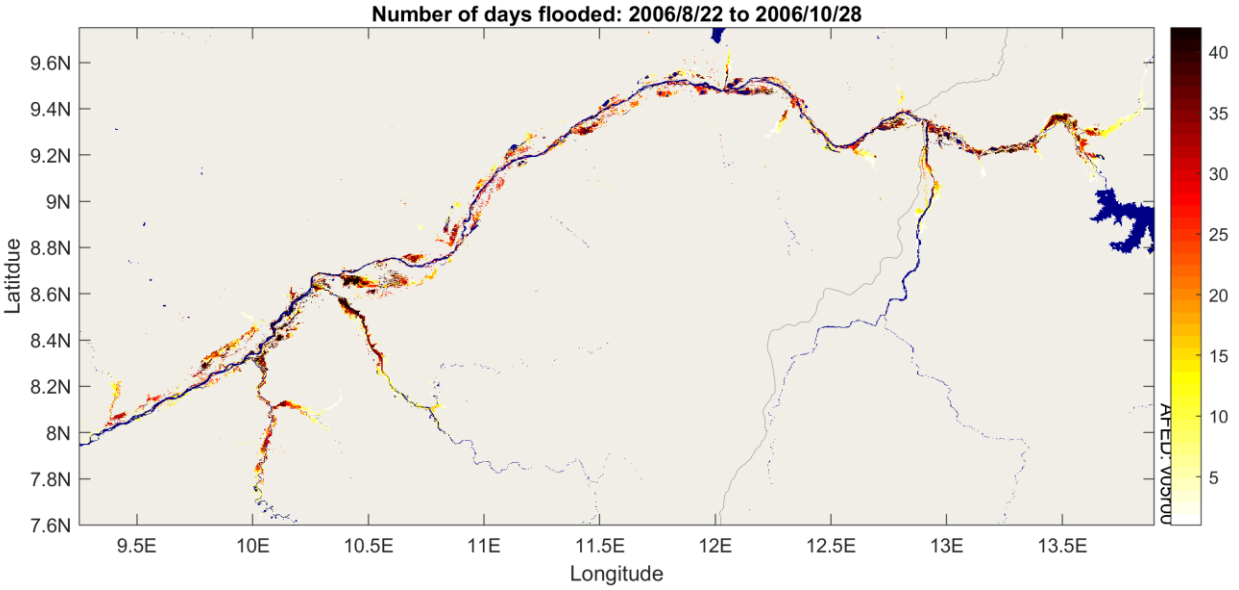


Figure 7: Number of days flooded, Benue River, Nigeria and Cameroon.

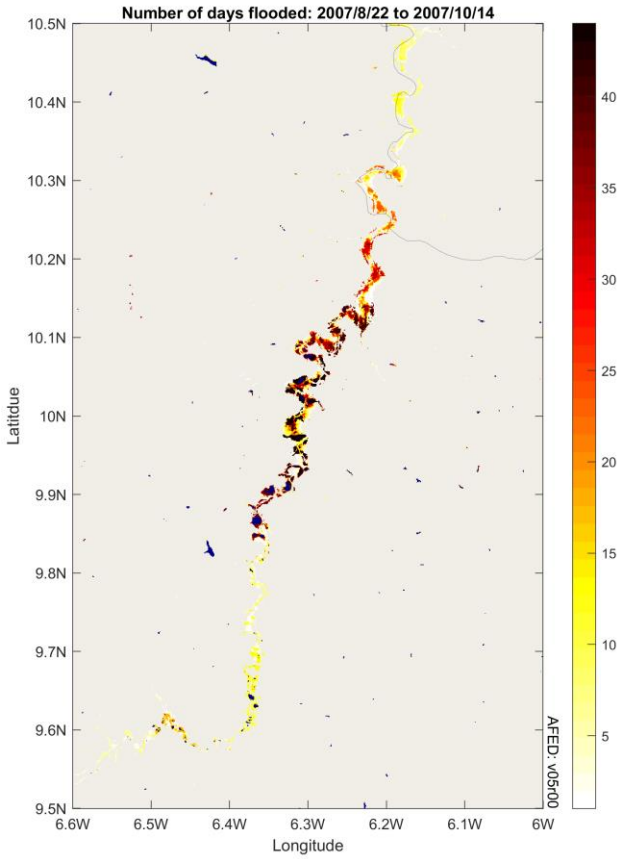


Figure 8: Number of days flooded, Bagoe, Cote d'Ivoire.

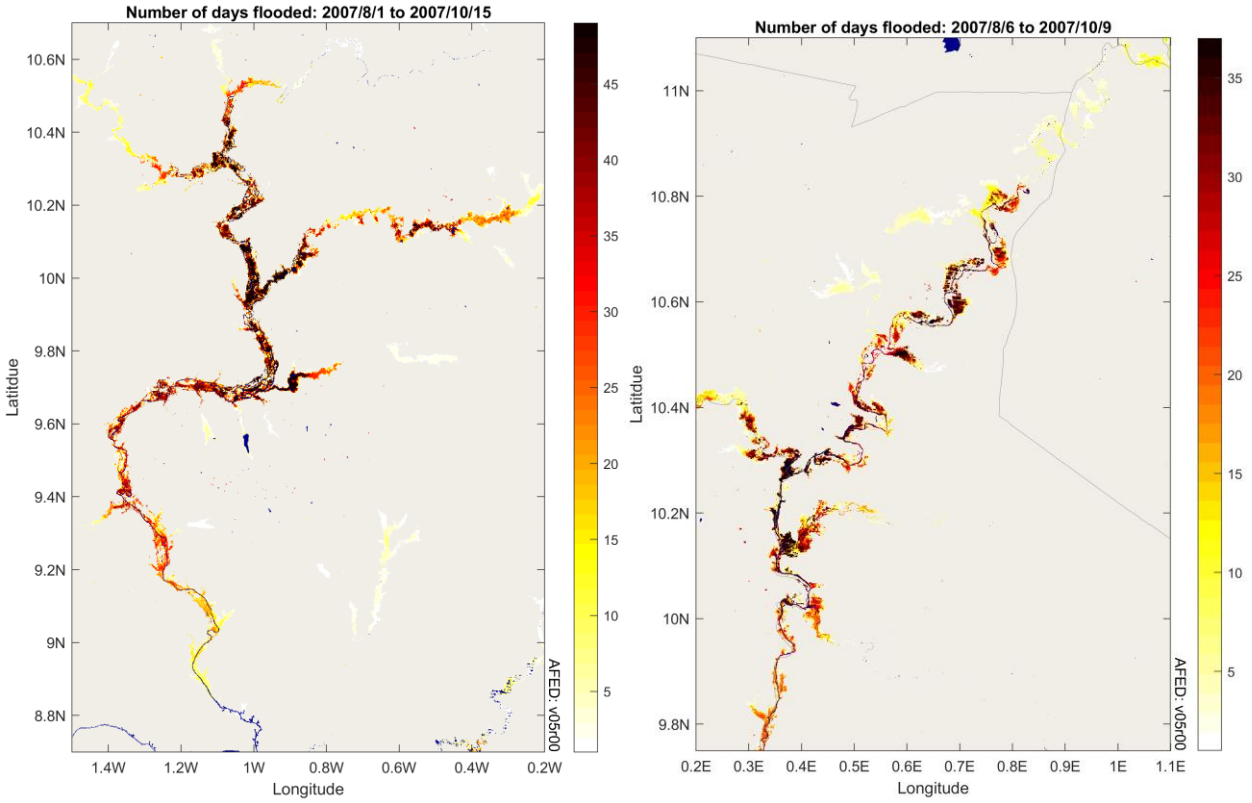


Figure 9: Number of days flooded. Left: White Volta River, Ghana. Right: Oti River, Togo.

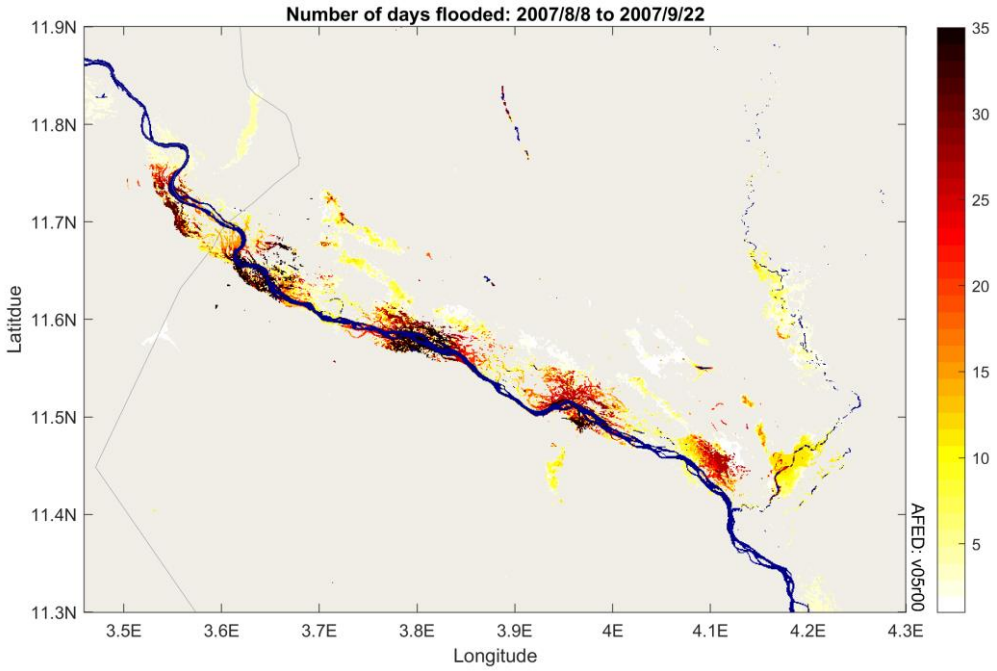


Figure 10: Number of days flooded, Niger River, Niger and Nigeria.

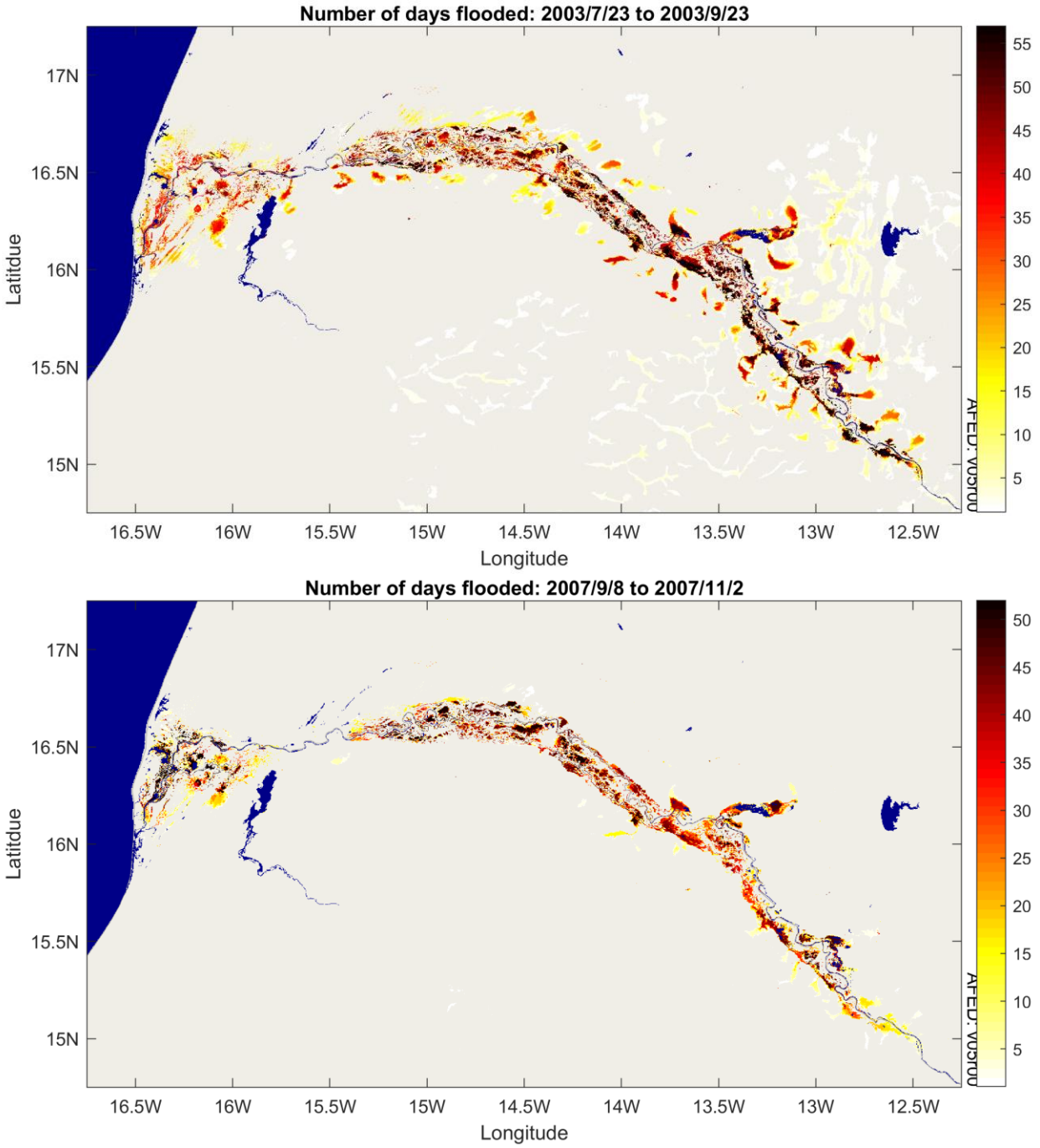


Figure 11: Number of days flooded, Senegal River, Senegal and Mauritania. Top: 2003. Bottom: 2007.

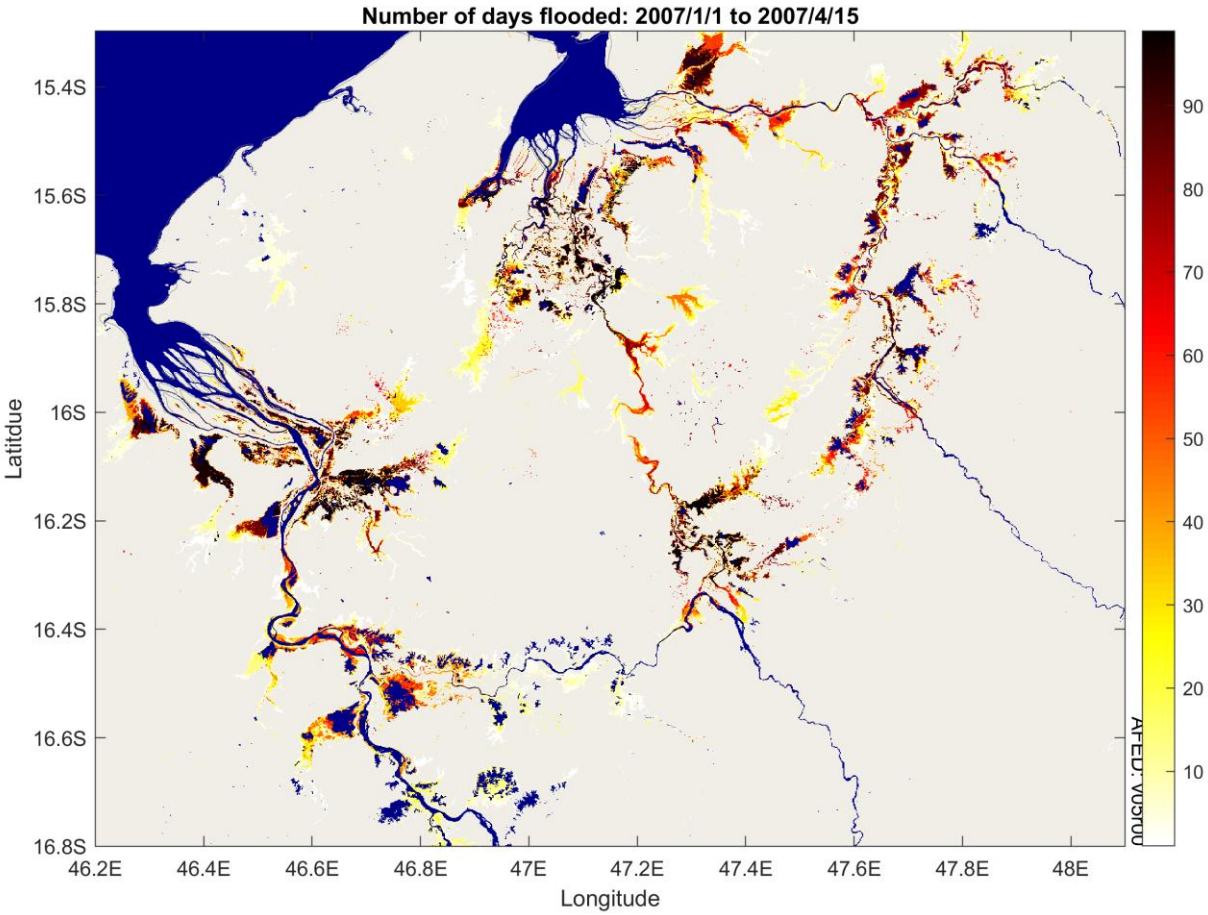


Figure 12: Number of days flooded, Betsiboka and Mahajamba Rivers, Madagascar.

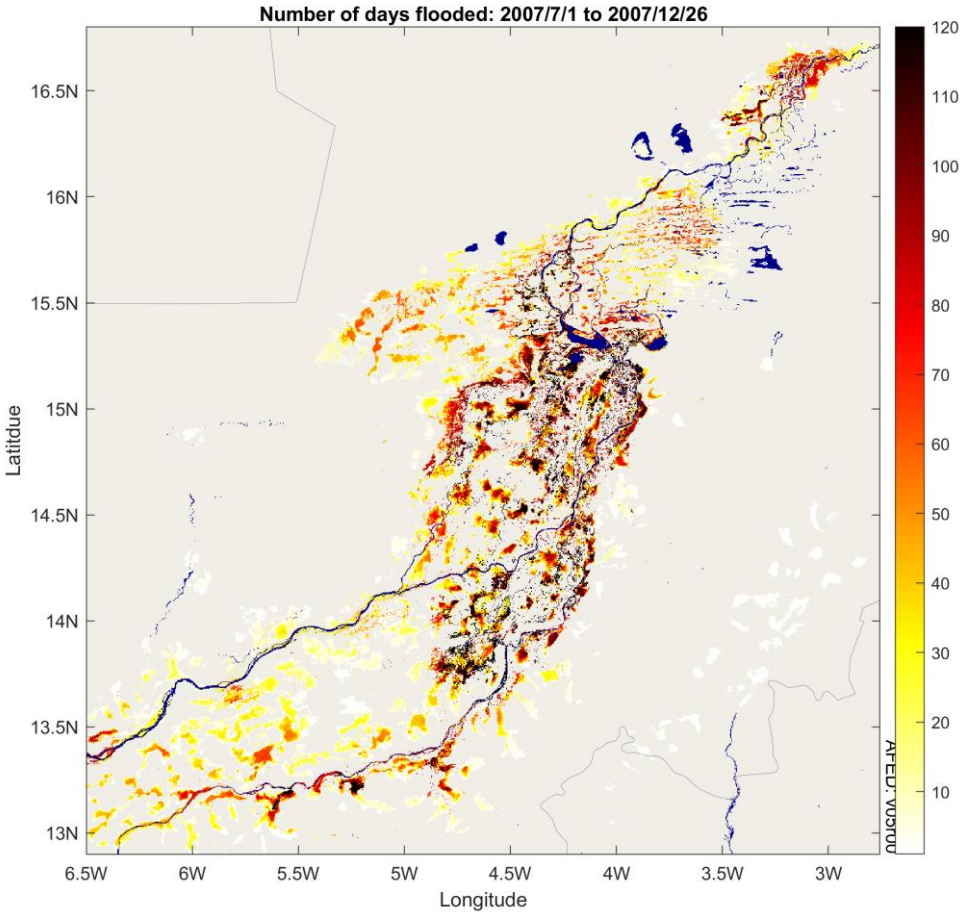


Figure 13: Number of days flooded, Inner Niger Delta, Mali.

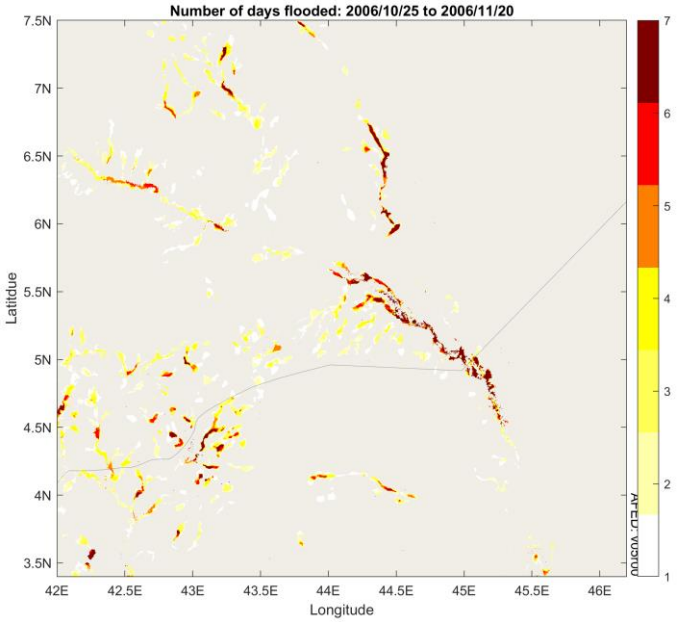


Figure 14: Number of days flooded, Shebelle River, Ethiopia and Somalia.

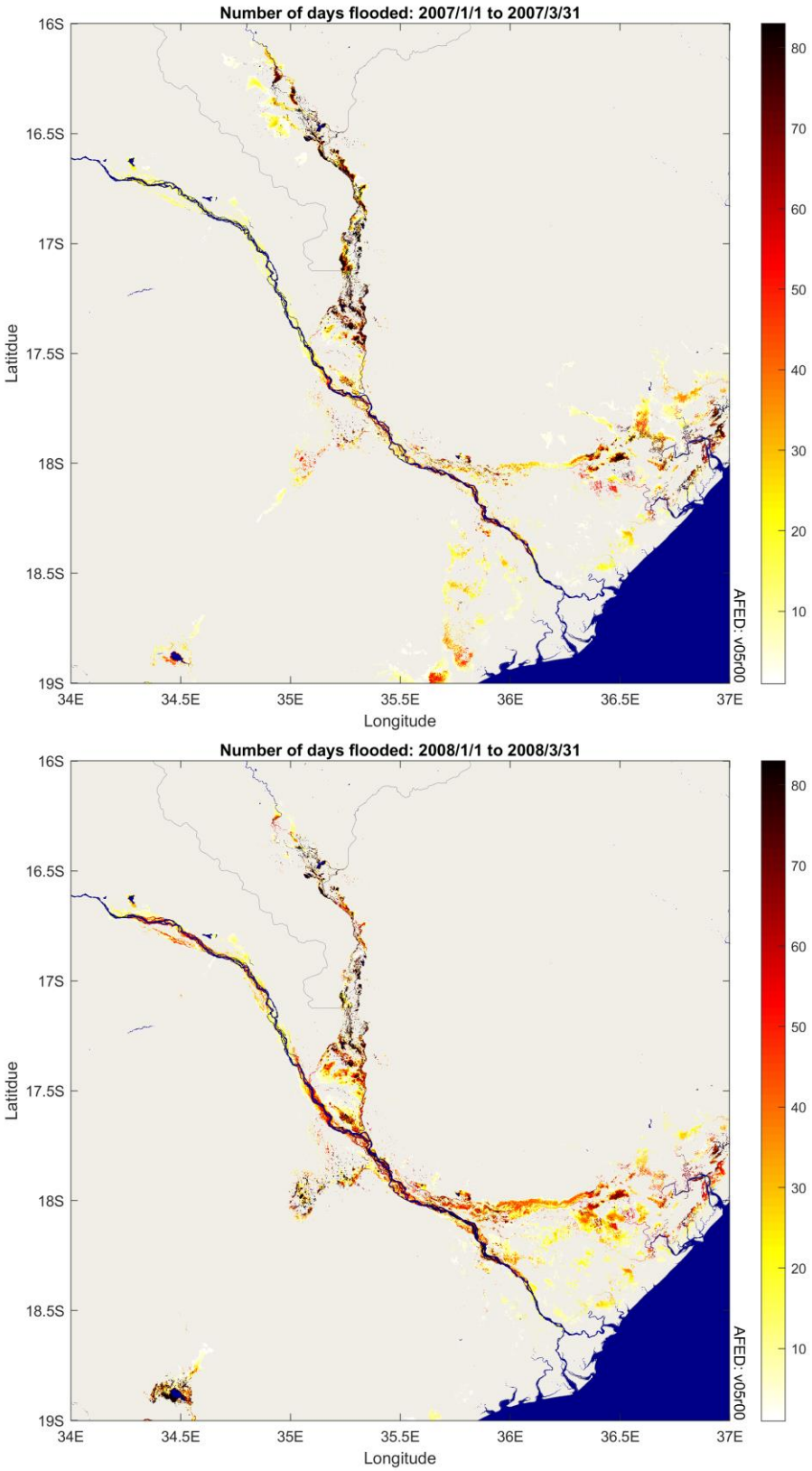


Figure 15: Number of days flooded, Zambezi and Shire Rivers, Mozambique and Malawi. Top: 2007. Bottom: 2008.

3.2 Flood event test cases: statistical results

3.2.1 *Flooded fraction*

Table 4 gives the results of statistical tests comparing microwave algorithm flooded fraction results to flooded fraction derived from the validation scenes listed in Table 1. Scatter plot representations of these data are included in section 3.4. Refer to section 2.3 for explanation of the evaluation methodology. AFED version V04R01 RMSE statistics are given for reference.

Table 4: Statistical comparisons between the MODIS-based flooded fraction and the microwave algorithm-based flooded fraction for each of the events in Table 1. Results are restricted to cells with <5% cloud mask.

Event name	Pearson correlation coefficient	RMSE (overall)	RMSE (validation scene FF >5%)		Cloud free area (%)
			V05R00	V04r01	
Shebelle2006	0.65	0.047	0.061	0.071	45
InnerNigerDelta2007	0.88	0.112	0.159	0.155	44
InnerNigerDelta2009	0.88	0.119	0.217	0.217	32
Limpopo2000	0.99	0.020	0.077	0.013	6
Limpopo2013	0.96	0.029	0.076	0.077	94
Pungue2008	0.98	0.068	0.112	0.109	11
Zambezi2007	0.86	0.039	0.083	0.074	82
Zambezi2008	0.96	0.053	0.089	0.082	27
Caprivi2003	0.99	0.038	0.075	0.063	56
Caprivi2009a	0.99	0.023	0.053	0.050	92
Caprivi2009b	0.99	0.031	0.076	0.069	90
Etosha2009a	0.94	0.052	0.098	0.097	73
Etosha2009b	0.95	0.041	0.075	0.076	90
Niger2012	0.56	0.039	0.133	0.132	81
Senegal2005	0.75	0.017	0.066	0.066	97
Fitri2007	0.62	0.154	0.262	0.262	100
Kyoga2007	0.93	0.038	0.108	0.109	48
Barotse2003	0.85	0.284	0.380	0.379	60
Barotse2007	0.93	0.245	0.310	0.327	48
Kafue2008	0.94	0.181	0.236	0.231	4

3.2.2 *AFED*

Table 5 gives the confusion matrix components for the validation scenes listed in Table 1 and Table 6 gives aggregate statistics on total area flooded. Refer to section 2.3 for explanation of the evaluation methodology. Table 5 also gives V04R00 true positive statistics for reference.

Table 5: Confusion matrix statistics for each of the flood events in Table 1.

Event name	True Positive (%)		False Negative (%)	False Positive (%)	True Negative (%)	Accuracy (%)
	V05R00	V04R01				
Shebelle2006	44	31.9	56	10	90	84
InnerNigerDelta2007	39	40.3	61	8	92	79
InnerNigerDelta2009	38	38.0	62	8	92	77
Limpopo2000	77	61.2	23	6	94	87
Limpopo2013	69	67.7	31	4	96	87
Pungue2008	67	66.2	33	4	96	74
Zambezi2007	54	56.5	46	3	97	78
Zambezi2008	67	68.5	33	6	94	76
Caprivi2003	80	82.0	20	4	96	91
Caprivi2009a	79	80.1	21	3	97	92
Caprivi2009b	76	78.3	24	4	96	90
Etosha2009a	61	59.8	39	6	94	86
Etosha2009b	72	72.0	28	5	95	89
Niger2012	27	27.8	73	0	100	88
Senegal2005	51	48.2	49	6	94	88
Fitri2007	21	20.8	79	0	100	89
Kyoga2007	55	53.7	45	1	99	90
Barotse2003	25	24.5	75	3	97	60
Barotse2007	39	36.5	61	6	94	62
Kafue2008	55	55.4	45	6	94	74

Table 6: Total flooded area statistics for each of the flood events in Table 1.

Event name	AFED Pixels Flooded (%)	MODIS Pixels Flooded (%)	No. AFED Pixels Flooded	No. MODIS Pixels Flooded	Total No. Pixels
Shebelle2006	14	14	160,341	154,054	1,112,424
InnerNigerDelta2007	16	25	580,235	918,405	3,642,294
InnerNigerDelta2009	16	26	534,961	868,990	3,335,318
Limpopo2000	32	37	507,251	721,602	2,195,272
Limpopo2013	20	24	605,935	740,187	3,076,626
Pungue2008	30	42	250,777	350,336	844,091
Zambezi2007	14	21	259,479	398,570	1,921,041
Zambezi2008	23	28	474,069	579,504	2,061,951
Caprivi2003	27	31	234,577	266,463	868,857
Caprivi2009a	22	25	311,153	358,159	1,417,816
Caprivi2009b	27	31	332,377	388,662	1,240,716
Etosha2009a	18	23	780,324	965,762	4,233,232
Etosha2009b	22	24	605,720	679,471	2,812,198

Niger2012	5	16	130,607	472,372	2,898,398
Senegal2005	12	13	112,483	126,579	955,185
Fitri2007	3	14	45,354	214,363	1,486,657
Kyoga2007	12	21	119,830	202,768	984,485
Barotse2003	11	39	1,292,848	4,500,275	11,631,758
Barotse2007	20	43	2,450,652	5,370,693	12,558,021
Kafue2008	30	48	626,813	1,008,827	2,116,872

3.3 Flood event test cases: graphical results

This section catalogs graphical results for each of the flood event test cases listed in Table 1. There are several graphics for each case:

G1. Unenhanced imagery of the flood event acquired on a single date from the independent source. True-color images are included for MODIS validation scenes.

G2. Enhanced imagery from the same source as graphic G1 highlighting areas most likely to be flooded. False-color images are included for MODIS validation scenes (discussed in section 2.4.2).

G3. A comparison flood map derived from the same source as G1 and G2, including cloud, tree, and persistent water masks (discussed in section 2.4).

G4. The AFED flood map on the same map grid as G1, G2, and G3, including the persistent water mask (but not the tree and cloud masks). Compare this map to G3 to assess the overall spatial correspondence between AFED and the independently derived flood map.

G5. The cumulative number of flooded days computed from AFED results for +/-30 days around the test scene date. This graphic provides context for the single-day AFED product and may reveal, for example, whether the AFED algorithm captures the maximum flood extent of the validation map over time.

G6. A comparison flooded fraction map derived from the flood map in G3 (discussed in section 2.5).

G7. Flooded fraction from the microwave algorithm without application of the minimum detectable water fraction (MDFF). Compare this map to G6 and G8 to assess where the MDFF limit may have an effect on the AFED product.

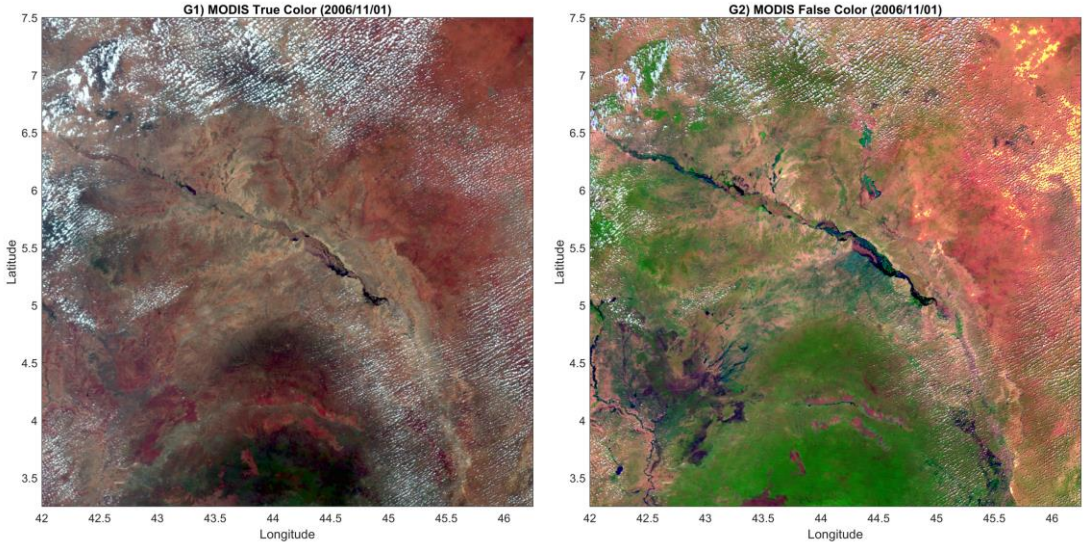
G8. Flooded fraction from the microwave algorithm on the same map grid as G6. Compare this map to G6 to assess the spatial correspondence between the microwave flooded fraction and the flooded fraction from the independently derived flood map.

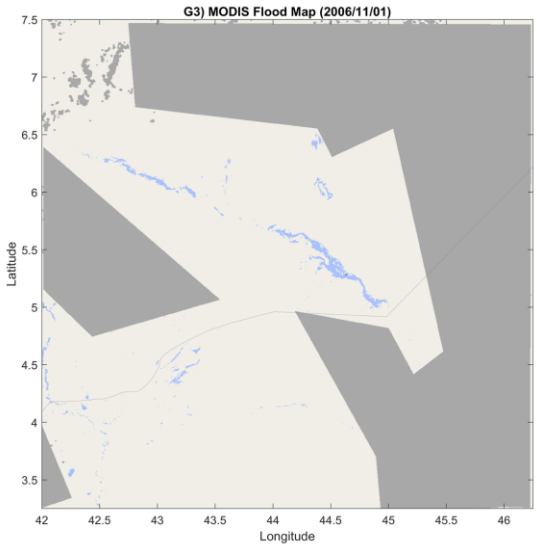
G9. Difference between upscaled AFED flooded fraction (discussed in section 2.5) and flooded fraction from the microwave algorithm. Use this map to assess the consistency between AFED and the flooded fraction data used by the AFED algorithm to derive it. Ideally the differences should be zero.

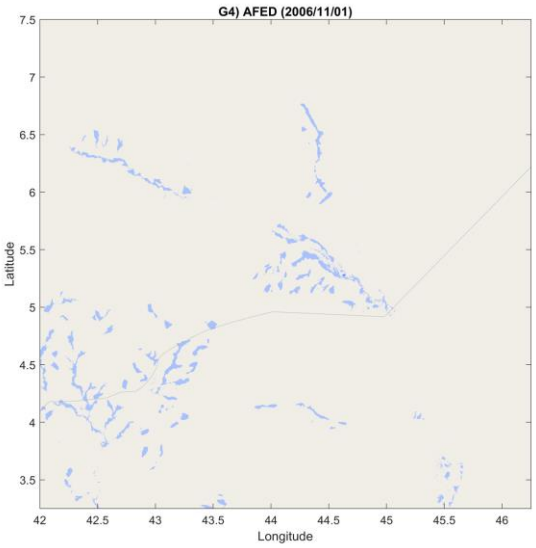
G10. Scatter plot comparing microwave flooded fraction data (G7 and G8) to independent flooded fraction data (G6). Microwave algorithm FF data falling below the MDFF limit are plotted in gray to indicate that the AFED algorithm resets these values to zero. Use this plot to assess the correspondence between microwave algorithm FF and the independent data at different values of FF.

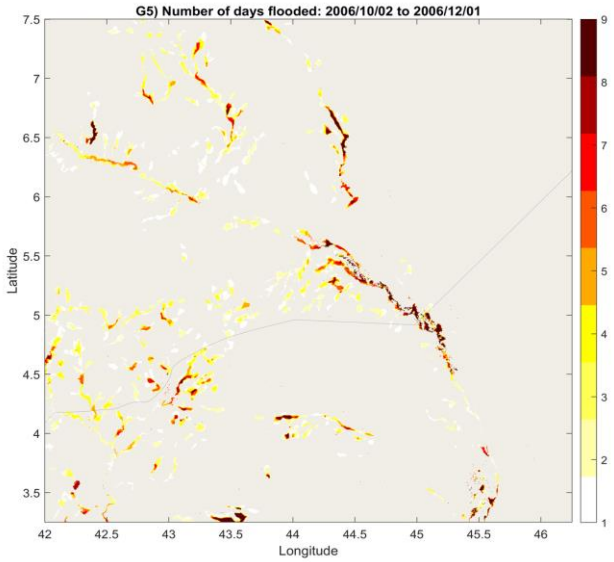
A short discussion of results for each case is included after the graphics. Section 3.5 provides a summary discussion of the overall test results.

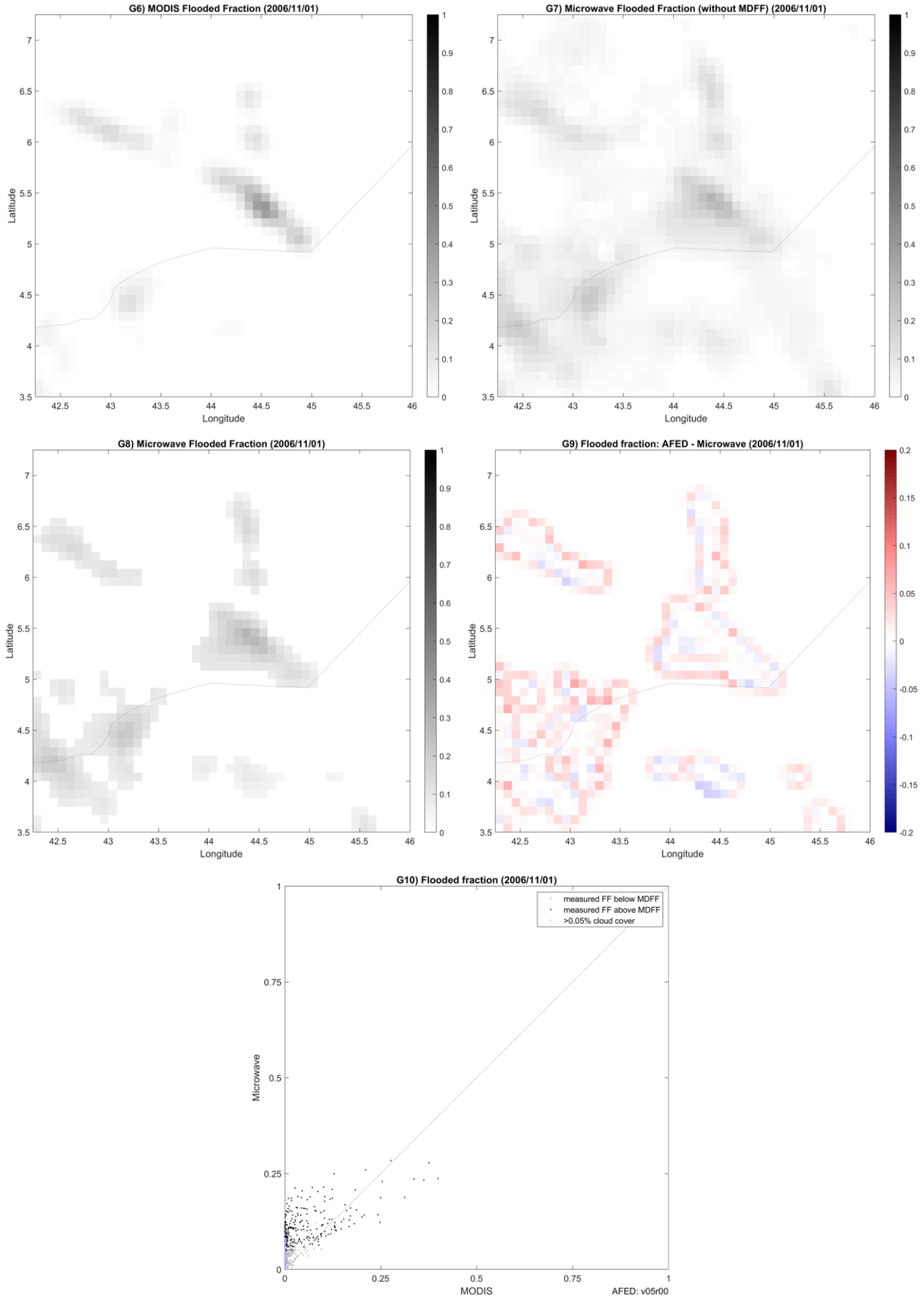
Case: Shebelle2006









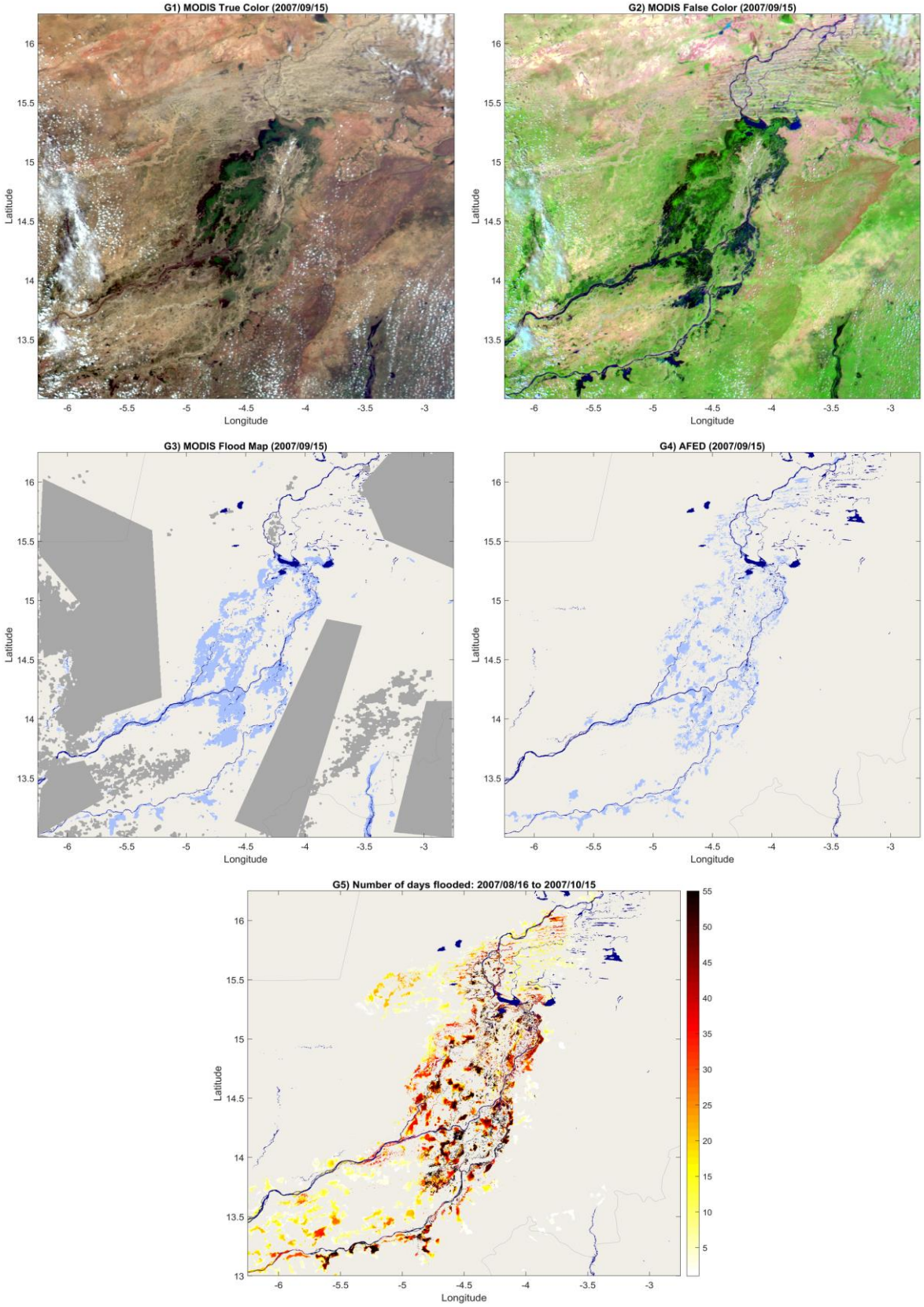


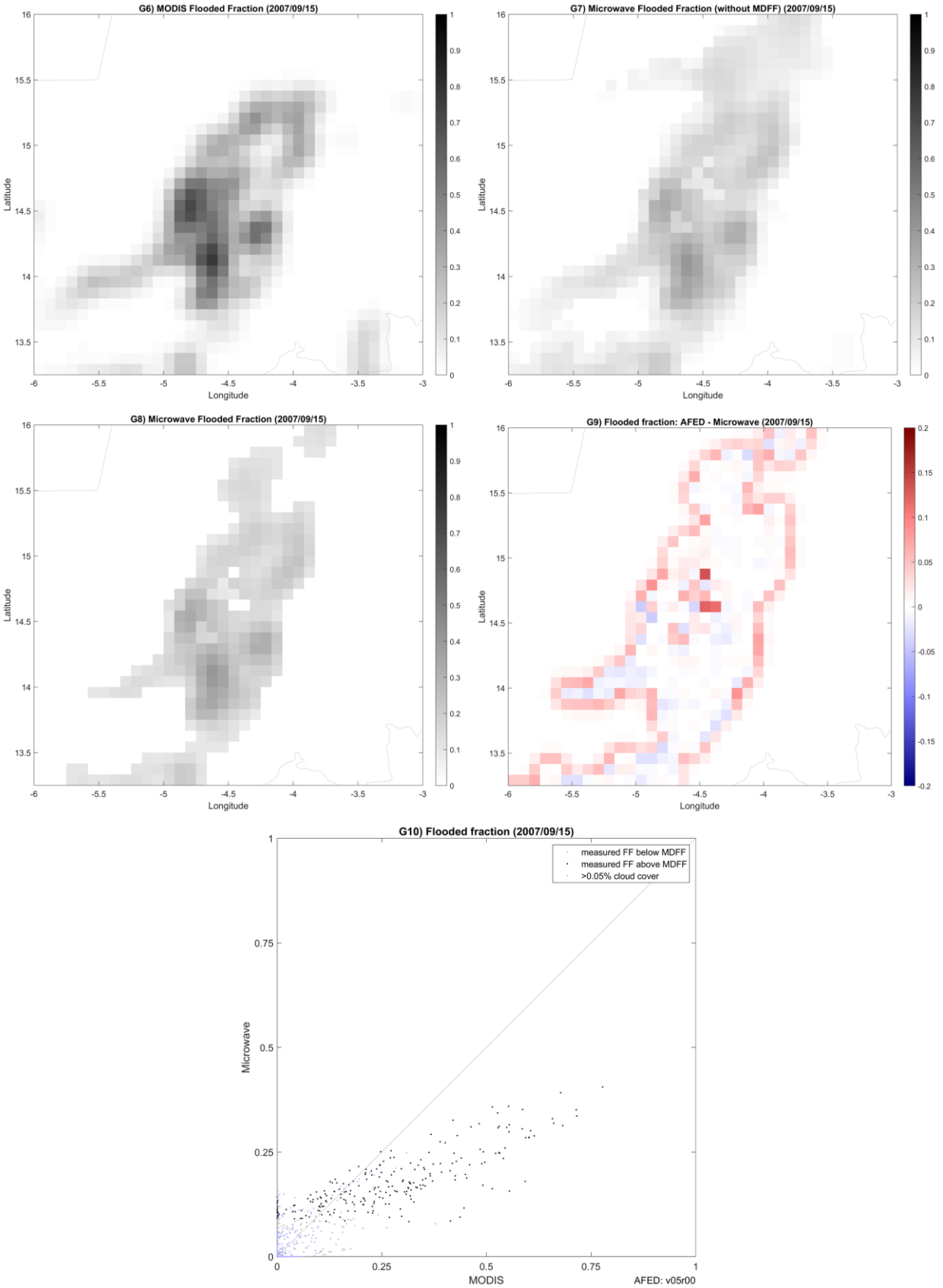
First row: G1) MODIS true color image and G2) MODIS false color image. Second row: G3) MODIS-derived flood map and G4) AFED flood map; light blue is flood water, dark blue is persistent open water, beige is dry land, and green is tree cover. Third row: G5) AFED cumulative number of days flooded for +/-30 days around the test scene date. Fourth row: G6) footprint-weighted flooded fraction derived from MODIS flood map and G7) microwave flooded fraction before applying MDFF. Fifth row: G8) microwave flooded fraction after applying MDFF and G9) difference between footprint-weighted AFED flooded fraction and microwave flooded fraction. Sixth row: G10) microwave-derived flooded fraction vs. MODIS-derived flooded fraction.

Shebelle2006 Discussion:

The single-day AFED result (G4) is a moderately good match to the MODIS flood map in this case. AFED depicts flooding in portions of two main Shebelle River floodplain segments in Ethiopia (running left to right) and in an area to the north where flooding is also seen in the MODIS image. AFED includes a number of areas of apparent false positives just south of the Shebelle and in a larger area to the southwest straddling the Ethiopia-Somalia border. The MODIS false color image is ambiguous in this second area and there may be indications of flooding or saturated soils. Flooded fraction (G7 and G8) includes most areas with the highest amount of MODIS image flooding The number of days flooded map (G5) shows that AFED also depicts flooding in the floodplain segment crossing the Ethiopia-Somalia border fairly well relative to the single-day validation map (G3).

Case: InnerNigerDelta2007





First row: G1) MODIS true color image and G2) MODIS false color image. Second row: G3) MODIS-derived flood map and G4) AFED flood map; light blue is flood water, dark blue is persistent open water, beige is dry land, and green is tree cover. Third row: G5) AFED cumulative number of days flooded for +/-30 days around the test scene date. Fourth row: G6) footprint-weighted flooded fraction derived from MODIS flood map and G7) microwave flooded fraction before applying MDFF. Fifth row: G8) microwave flooded fraction after applying MDFF and G9) difference between footprint-weighted AFED flooded fraction and microwave flooded fraction. Sixth row: G10) microwave-derived flooded fraction vs. MODIS-derived flooded fraction.

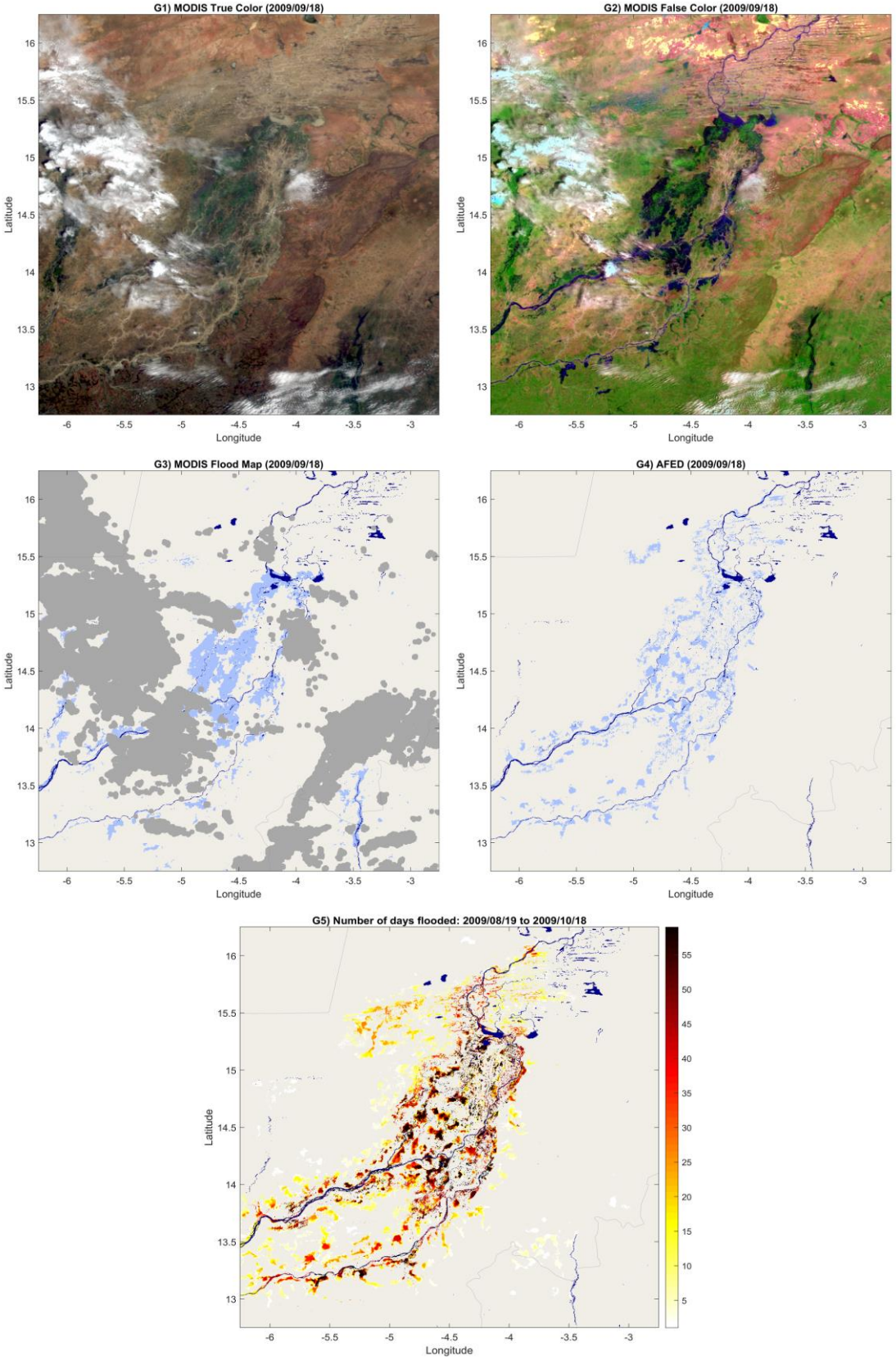
InnerNigerDelta2007 Discussion

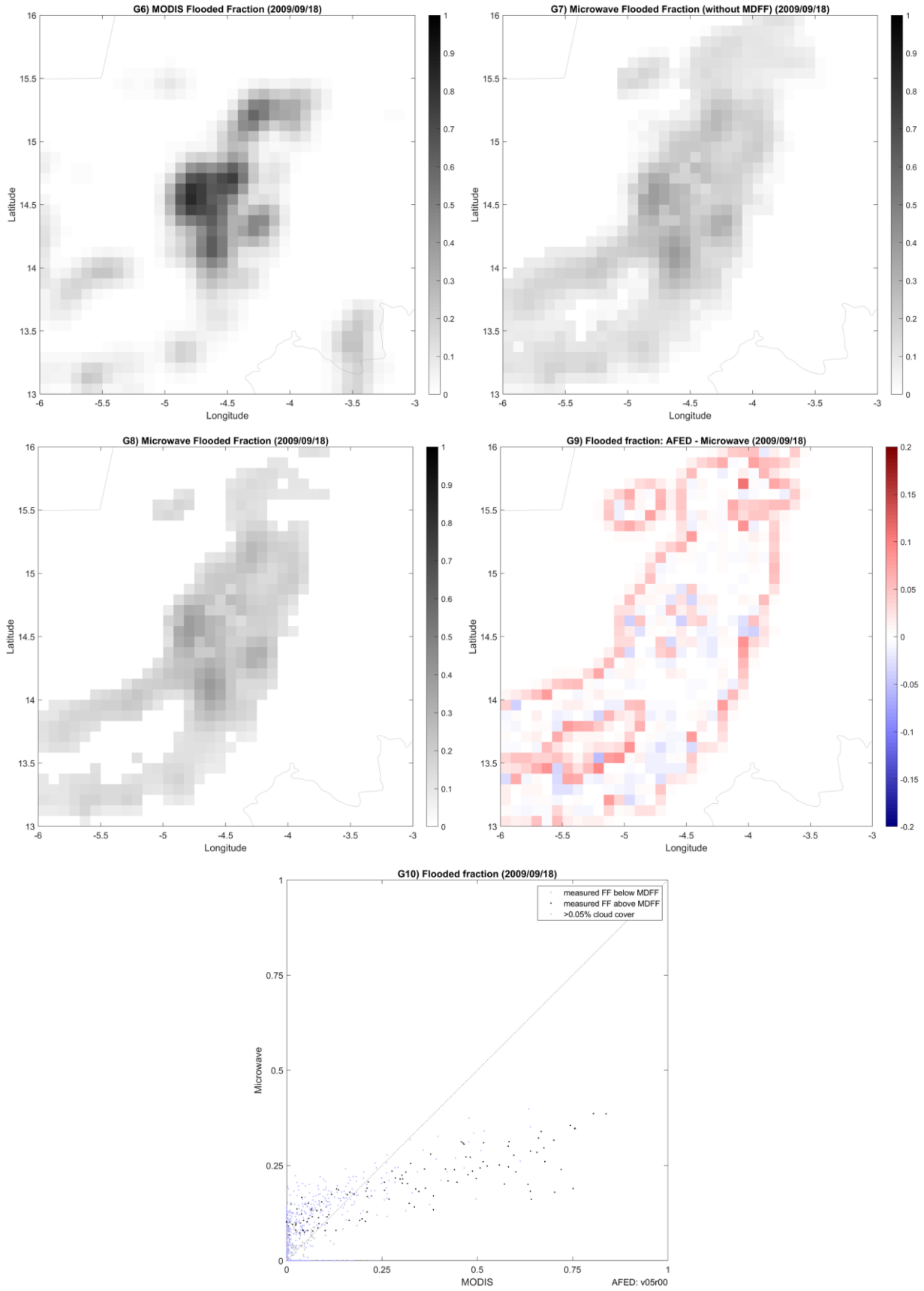
The AFED flood map comparison (G1-G4) shows good overall spatial correspondence between the AFED and MODIS flood maps with some disagreement at finer scales. The introduction of the MERIT elevation model (Yamazaki et al., 2017) in V05R00 has eliminated the "striping" artifact seen in prior AFED versions, which was traceable to artifacts in the SRTM elevation (Farr et al., 2017).

The flooded fraction (FF) maps (G6-G9) show that the microwave-based FF follows the larger-scale spatial pattern of the validation data, including the east-west extent and to some degree the FF variations within the flooded area. The MDFF-trimmed AFED FF (G8) eliminates some false positive areas (e.g., to the north and southwest) as well as some true-positive areas (e.g., part of the branch at 14° N latitude). G9 shows that the downscaling algorithm produces a flood map that accurately represents the large-scale features of the flooded fraction inputs; the areas of positive bias (red shades) on periphery of the flooded areas are a result of the MDFF-trimming process (discussed in the summary) and are not a reason for concern.

The FF scatter plot (G10) shows that there is good correlation in this case between the microwave-derived FF and validation scene (MODIS) flooded fractions. However, the AFED FF has a negative bias trend that increases with FF, a pattern that is repeated in several other cases and may be attributable to overestimation of flooding in MODIS image analysis (discussed in the summary).

Case: InnerNigerDelta2009



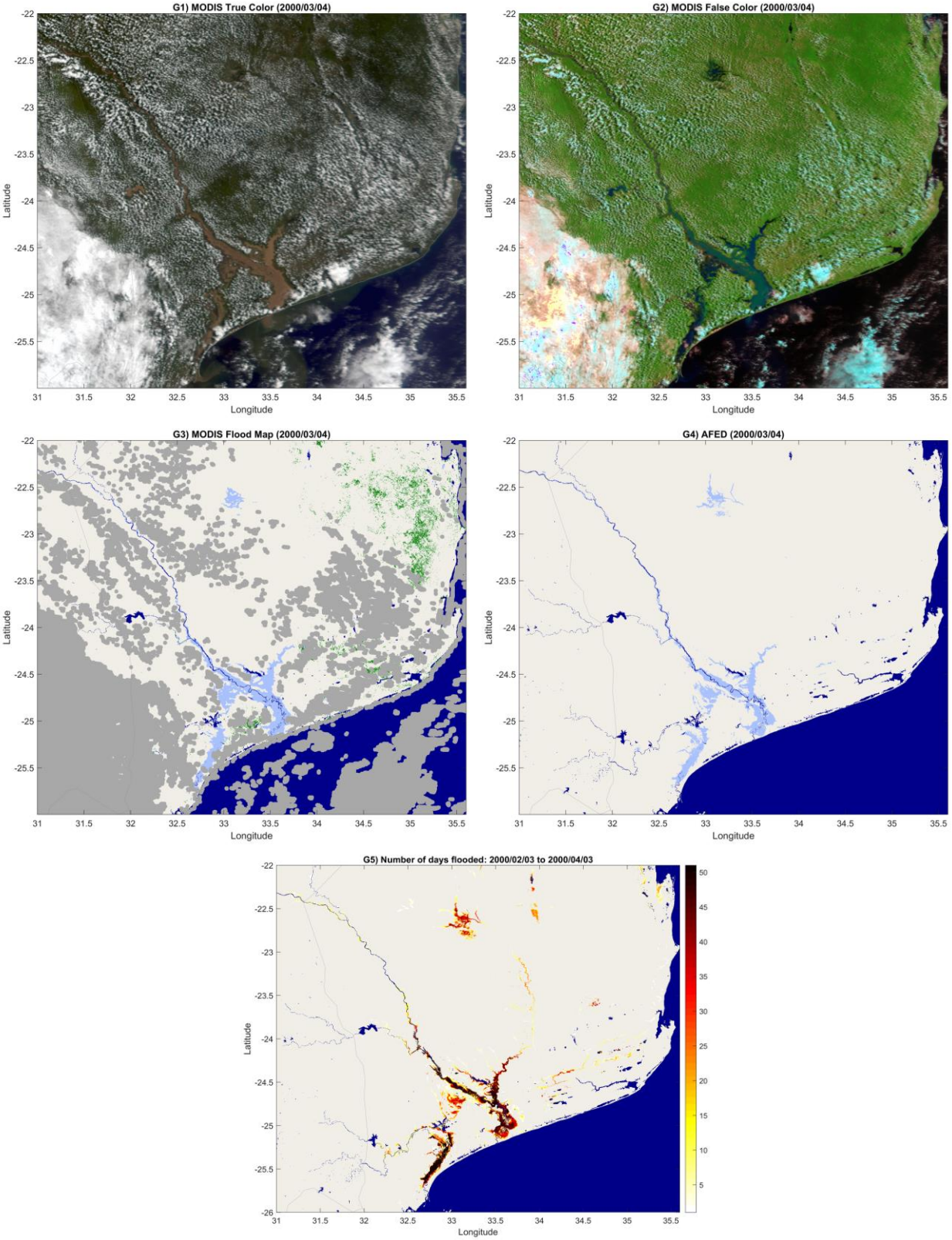


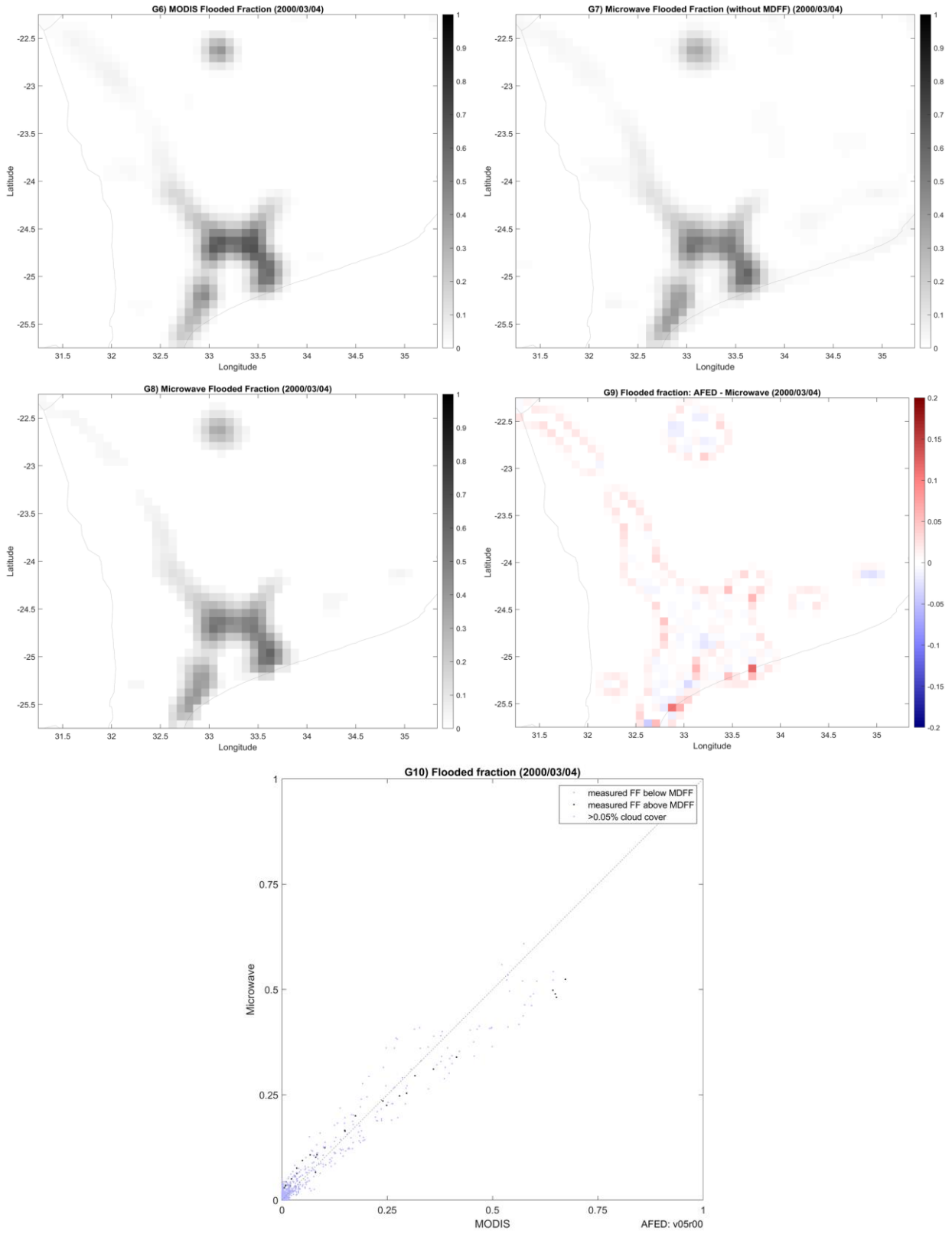
First row: G1) MODIS true color image and G2) MODIS false color image. Second row: G3) MODIS-derived flood map and G4) AFED flood map; light blue is flood water, dark blue is persistent open water, beige is dry land, and green is tree cover. Third row: G5) AFED cumulative number of days flooded for +/-30 days around the test scene date. Fourth row: G6) footprint-weighted flooded fraction derived from MODIS flood map and G7) microwave flooded fraction before applying MDFF. Fifth row: G8) microwave flooded fraction after applying MDFF and G9) difference between footprint-weighted AFED flooded fraction and microwave flooded fraction. Sixth row: G10) microwave-derived flooded fraction vs. MODIS-derived flooded fraction.

InnerNigerDelta2009 Discussion:

MODIS test scene confidence is low for this case because of widespread cloud contamination (visible in G1). When the most cloud-contaminated areas are excluded, the flooded fraction results (G10) are similar to those for the InnerNigerDelta2007 case covering the same domain, with negative bias that increases with flooded fraction amount. Areas of positive flooded fraction bias (G8, e.g., the southwest corner of the domain) are difficult to analyze further because of the cloud contamination in that part of the test scene. See InnerNigerDelta2007 case above for further discussion of performance at finer spatial scales.

Case: Limpopo2000



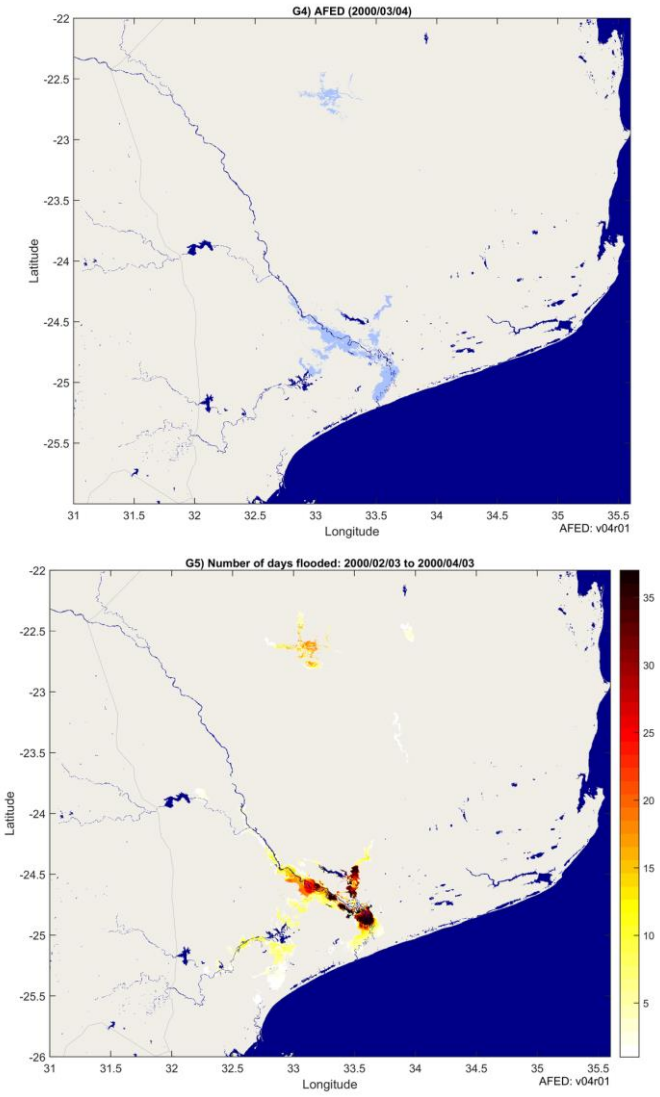


First row: G1) MODIS true color image and G2) MODIS false color image. Second row: G3) MODIS-derived flood map and G4) AFED flood map; light blue is flood water, dark blue is persistent open water, beige is dry land, and green is tree cover. Third row: G5) AFED

cumulative number of days flooded for +/-30 days around the test scene date. Fourth row: G6) footprint-weighted flooded fraction derived from MODIS flood map and G7) microwave flooded fraction before applying MDFF. Fifth row: G8) microwave flooded fraction after applying MDFF and G9) difference between footprint-weighted AFED flooded fraction and microwave flooded fraction. Sixth row: G10) microwave-derived flooded fraction vs. MODIS-derived flooded fraction.

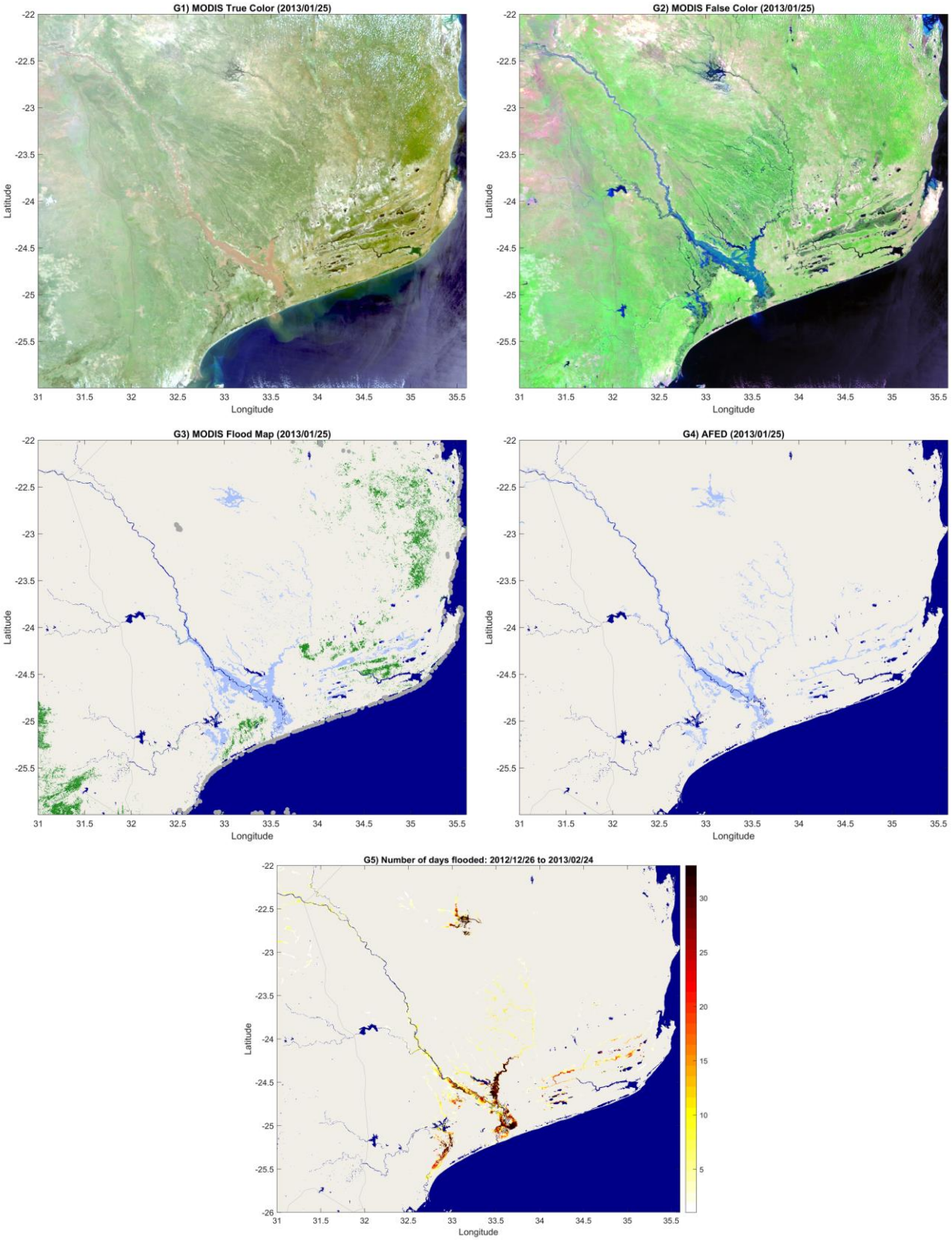
Limpopo2000 Discussion:

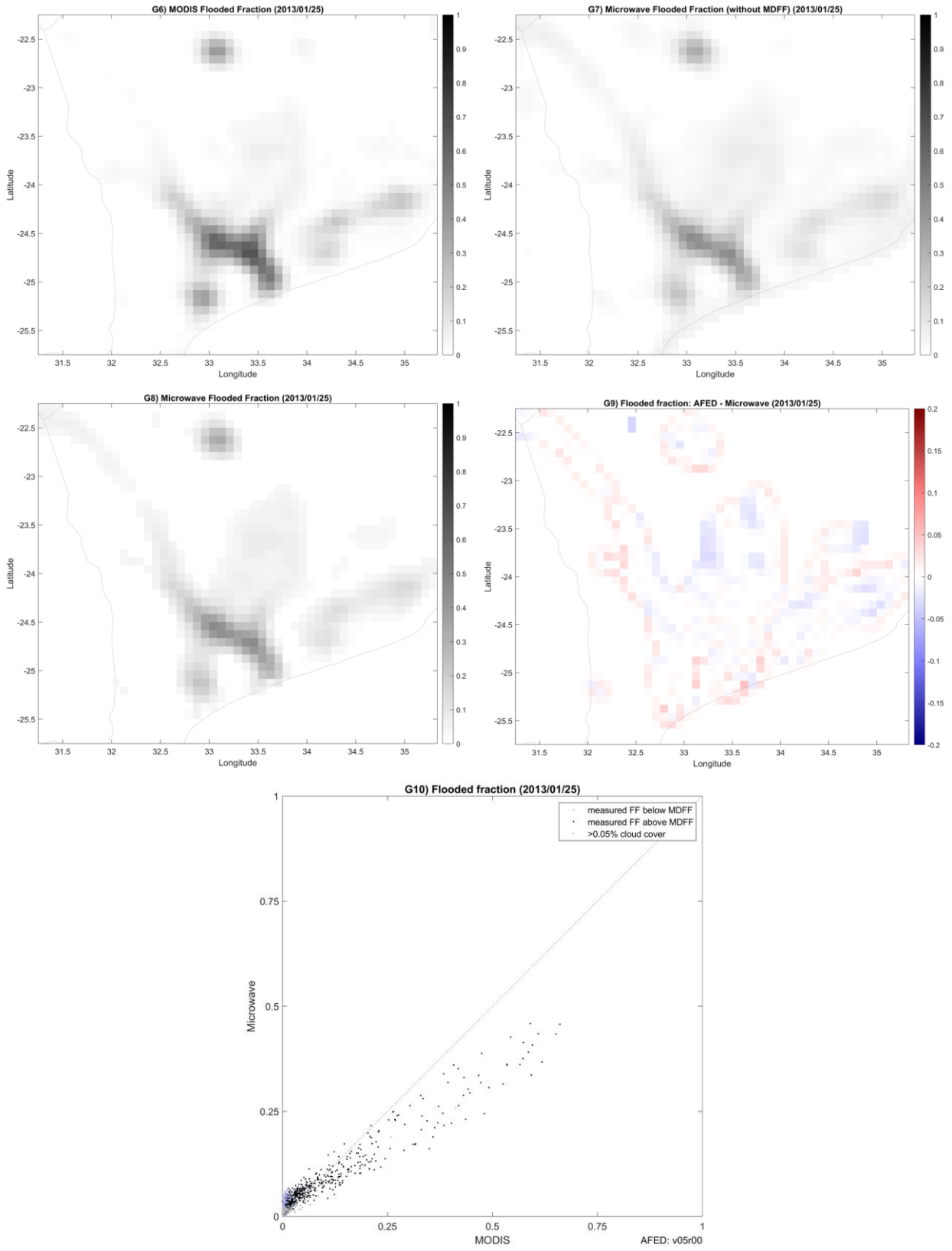
AFED (G4) performs well and is in excellent agreement with the MODIS flood map (G3) throughout the area shown. As of V05R00, AFED flooding extends farther down the Limpopo almost to the coastline and includes the areas of flooding extending to southwest and northeast of the Limpopo flood plain. (G4 and G5 V04R01 graphics are provided below for reference.) Furthermore, V05R00 improves or eliminates two spatial artifacts seen in V04R01. The first is in the flooded area extending just southwest of the main Limpopo River floodplain. The MODIS flood map shows near-contiguous flooding connected to the Limpopo's flood plain, while AFED does not. AFED depicts water cover in this area better in V05R00 than V04R01 but spatial discontinuities remain. This is likely due to the hydrological flow direction data underlying the downscaling process. Flow direction data denotes the most likely (e.g., downhill) direction of water flow but in this and similar situations—e.g., wide, flat flood plains—extreme flooding may create backflow that goes against the typical flow direction. The second feature is seen in the flooding in the north-center of the test scene. MODIS imagery indicates flooding in a single compact area while AFED also depicts flooding that spreads along two streamlines running to the north and west. The total amount of water cover is similar in the two maps, but assumptions in the AFED downscaling process cause it to spread water cover along streamlines rather than further expanding the central area of flooding.



Graphics G4 and G5 for V04R01 Limpopo2000 results based on 50-km SSM/I data.

Case: Limpopo2013





First row: G1) MODIS true color image and G2) MODIS false color image. Second row: G3) MODIS-derived flood map and G4) AFED flood map; light blue is flood water, dark blue is

persistent open water, beige is dry land, and green is tree cover. Third row: G5) AFED cumulative number of days flooded for +/-30 days around the test scene date. Fourth row: G6) footprint-weighted flooded fraction derived from MODIS flood map and G7) microwave flooded fraction before applying MDFF. Fifth row: G8) microwave flooded fraction after applying MDFF and G9) difference between footprint-weighted AFED flooded fraction and microwave flooded fraction. Sixth row: G10) microwave-derived flooded fraction vs. MODIS-derived flooded fraction.

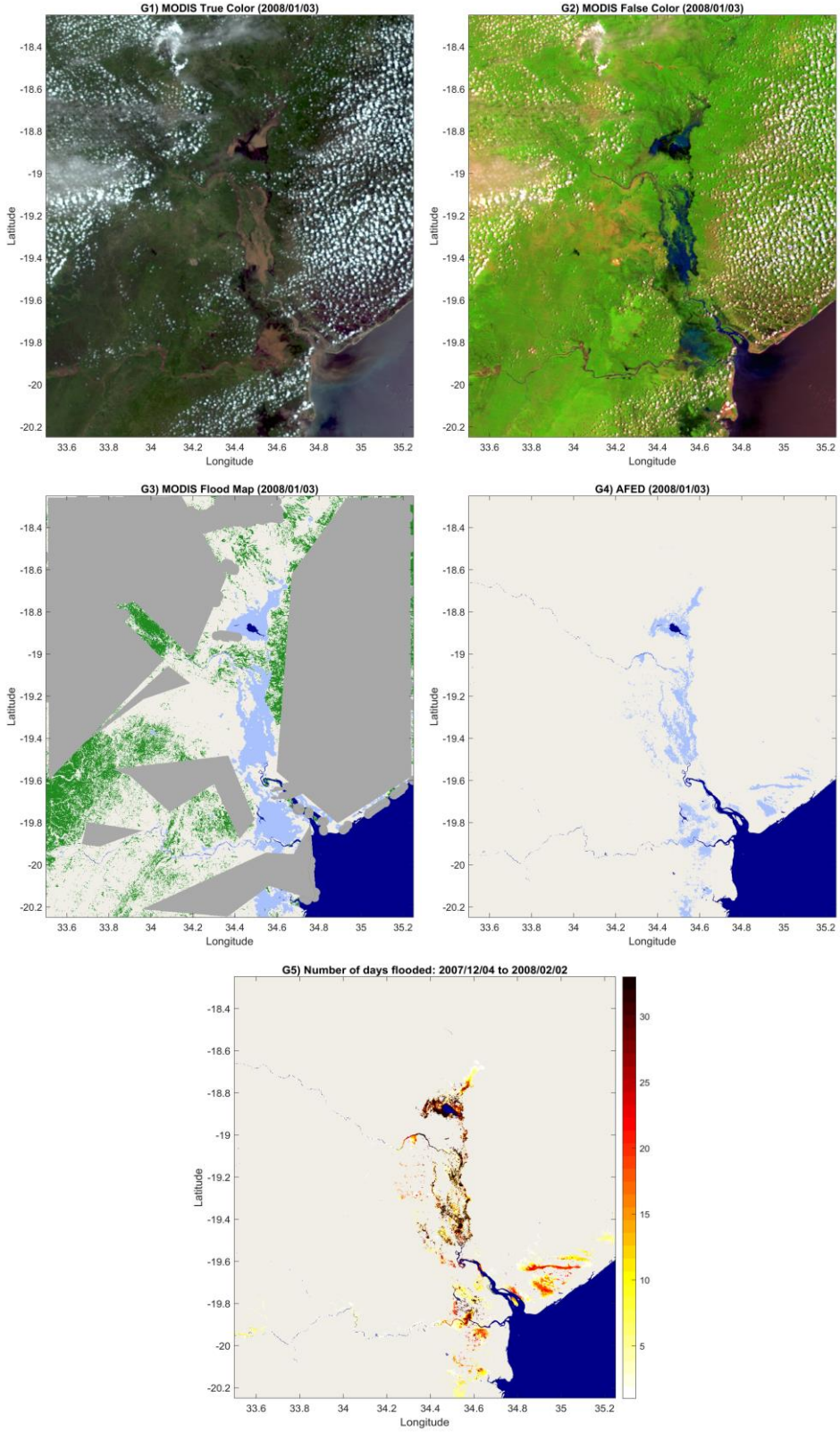
Limpopo2013 Discussion:

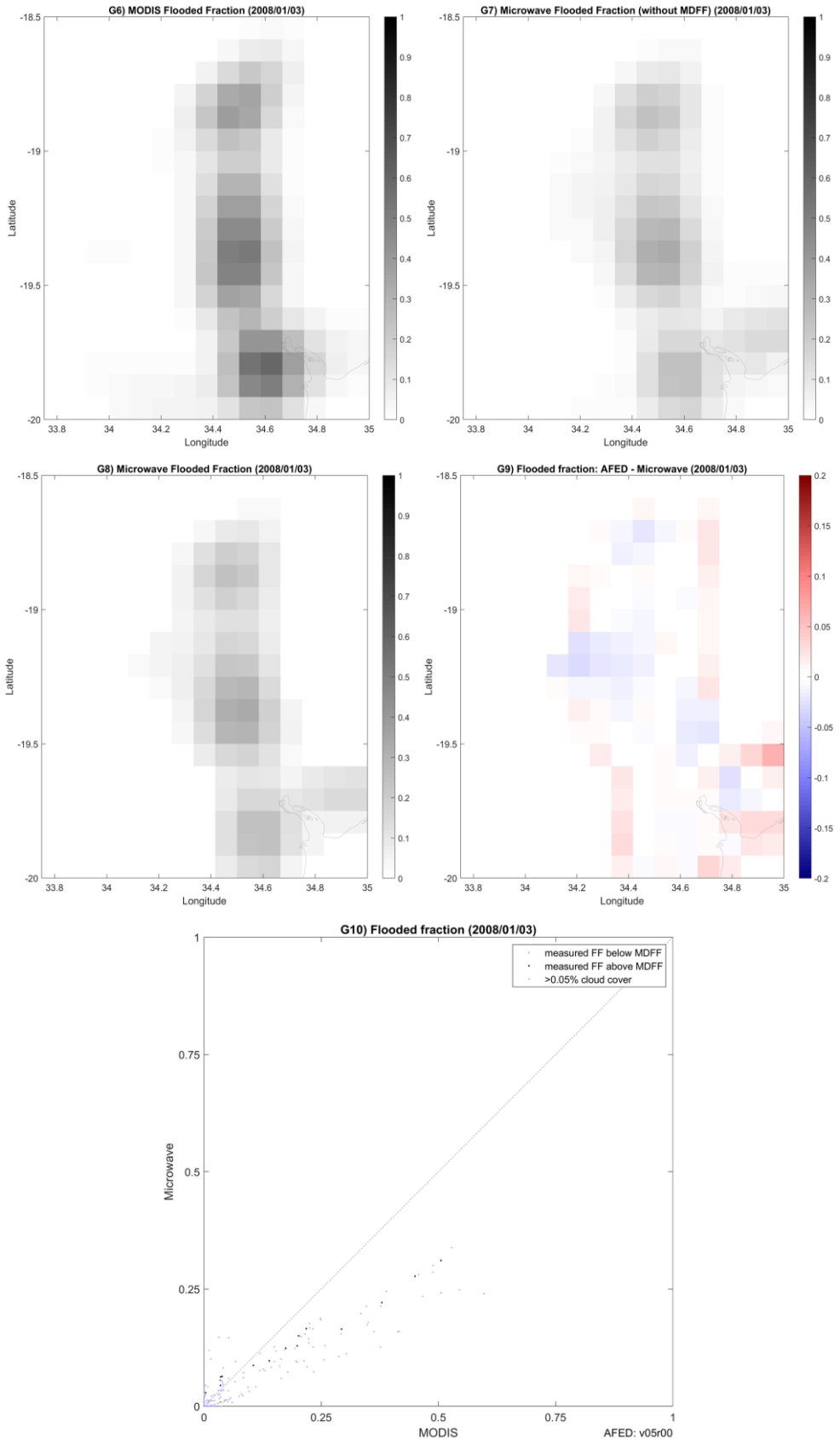
This case covers the same area as Limpopo2000. The two floods differ in two key respects: (1) the 2000 flood in the Limpopo floodplain was much larger than that in 2013 and (2) the 2013 flood includes a broad area of low-density flooding to the northeast of the Limpopo.

AFED (G4) is in good agreement with the MODIS flood map (G3) in the main Limpopo floodplain and the areas with lower-density flooding to the northeast of the Limpopo. Flooded fraction is slightly underestimated (G10), so the MODIS flood map depicts more extensive flooding than AFED. The two downscaling artifacts noted in the discussion of the Limpopo2000 case are present to a degree here as well.

The consistent results between this case and Limpopo2000 provide evidence that the algorithm (1) performs consistently over time and between sensor types (TMI vs. AMSR2) and (2) is able to differentiate flood extent details between large floods in the same area.

Case: Pungue2008





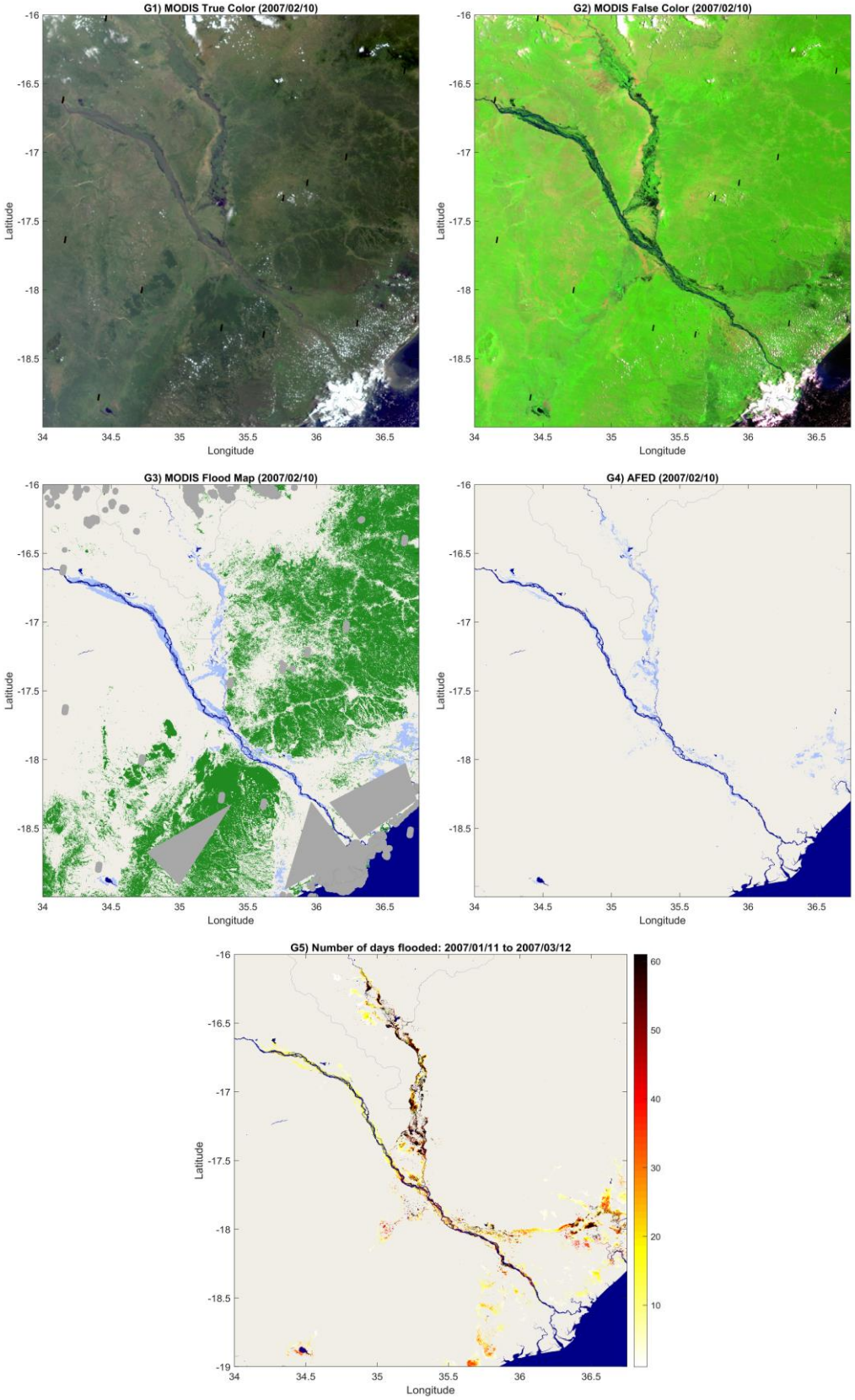
First row: G1) MODIS true color image and G2) MODIS false color image. Second row: G3) MODIS-derived flood map and G4) AFED flood map; light blue is flood water, dark blue is persistent open water, beige is dry land, and green is tree cover. Third row: G5) AFED cumulative number of days flooded for +/-30 days around the test scene date. Fourth row: G6) footprint-weighted flooded fraction derived from MODIS flood map and G7) microwave flooded fraction before applying MDFF. Fifth row: G8) microwave flooded fraction after applying MDFF and G9) difference between footprint-weighted AFED flooded fraction and microwave flooded fraction. Sixth row: G10) microwave-derived flooded fraction vs. MODIS-derived flooded fraction.

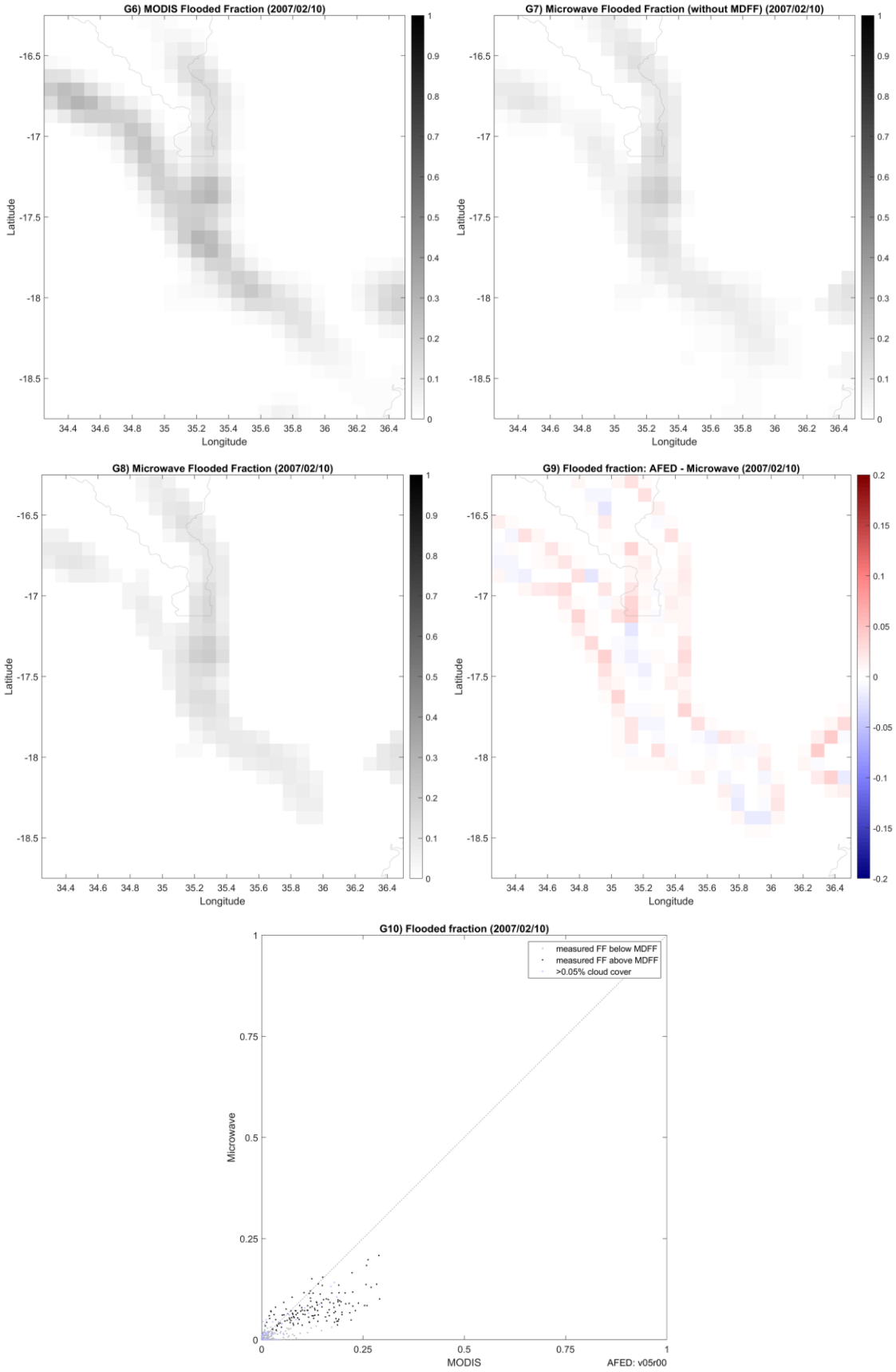
Pungue2008 Discussion:

Similar to other cases, the AFED (G4) shows good spatial correspondence to the MODIS validation imagery (G3), but with a somewhat lower total flooded area. Some of the coastal areas cannot be directly validated due to cloud contamination in the MODIS image, but AFED in these appears to correspond to flooding apparent between clouds in the false color image (G2). The cumulative number of days flooded map (G5) shows that AFED likely never reaches the full flood extent indicated by the MODIS map.

The FF scatter plot (G10) shows that there is good correlation in this case between the microwave-derived FF and validation scene (MODIS) flooded fractions. However, the AFED FF has a negative bias trend that increases with FF, a pattern that is repeated in several other cases (discussed in the summary).

Case: Zambezi2007



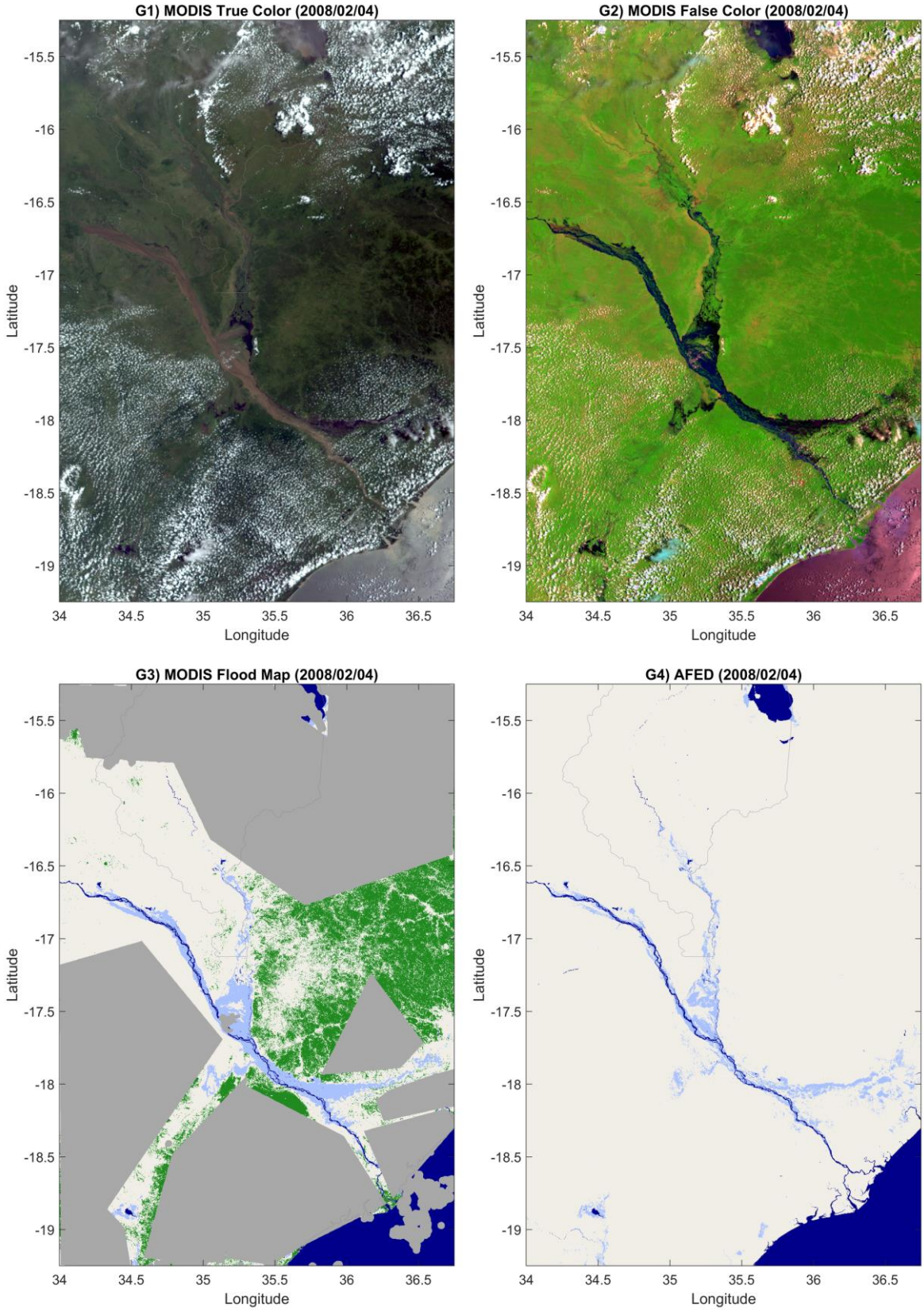


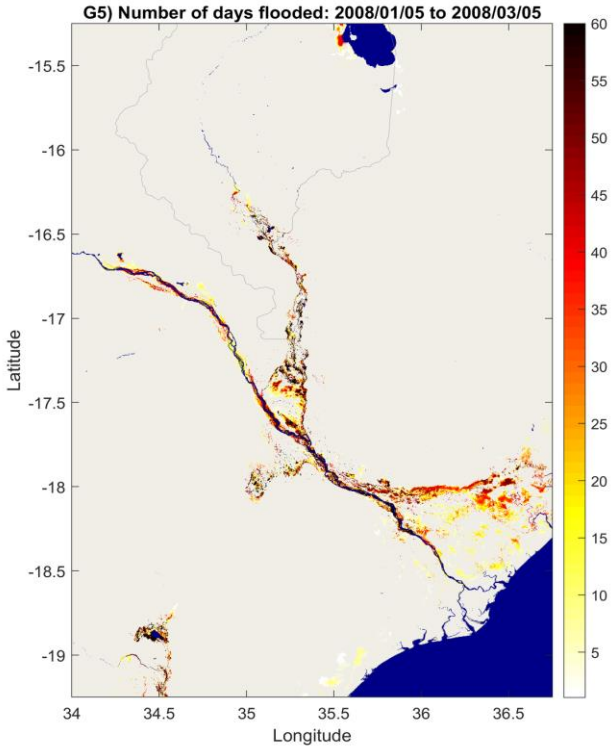
First row: G1) MODIS true color image and G2) MODIS false color image. Second row: G3) MODIS-derived flood map and G4) AFED flood map; light blue is flood water, dark blue is persistent open water, beige is dry land, and green is tree cover. Third row: G5) AFED cumulative number of days flooded for +/-30 days around the test scene date. Fourth row: G6) footprint-weighted flooded fraction derived from MODIS flood map and G7) microwave flooded fraction before applying MDFF. Fifth row: G8) microwave flooded fraction after applying MDFF and G9) difference between footprint-weighted AFED flooded fraction and microwave flooded fraction. Sixth row: G10) microwave-derived flooded fraction vs. MODIS-derived flooded fraction.

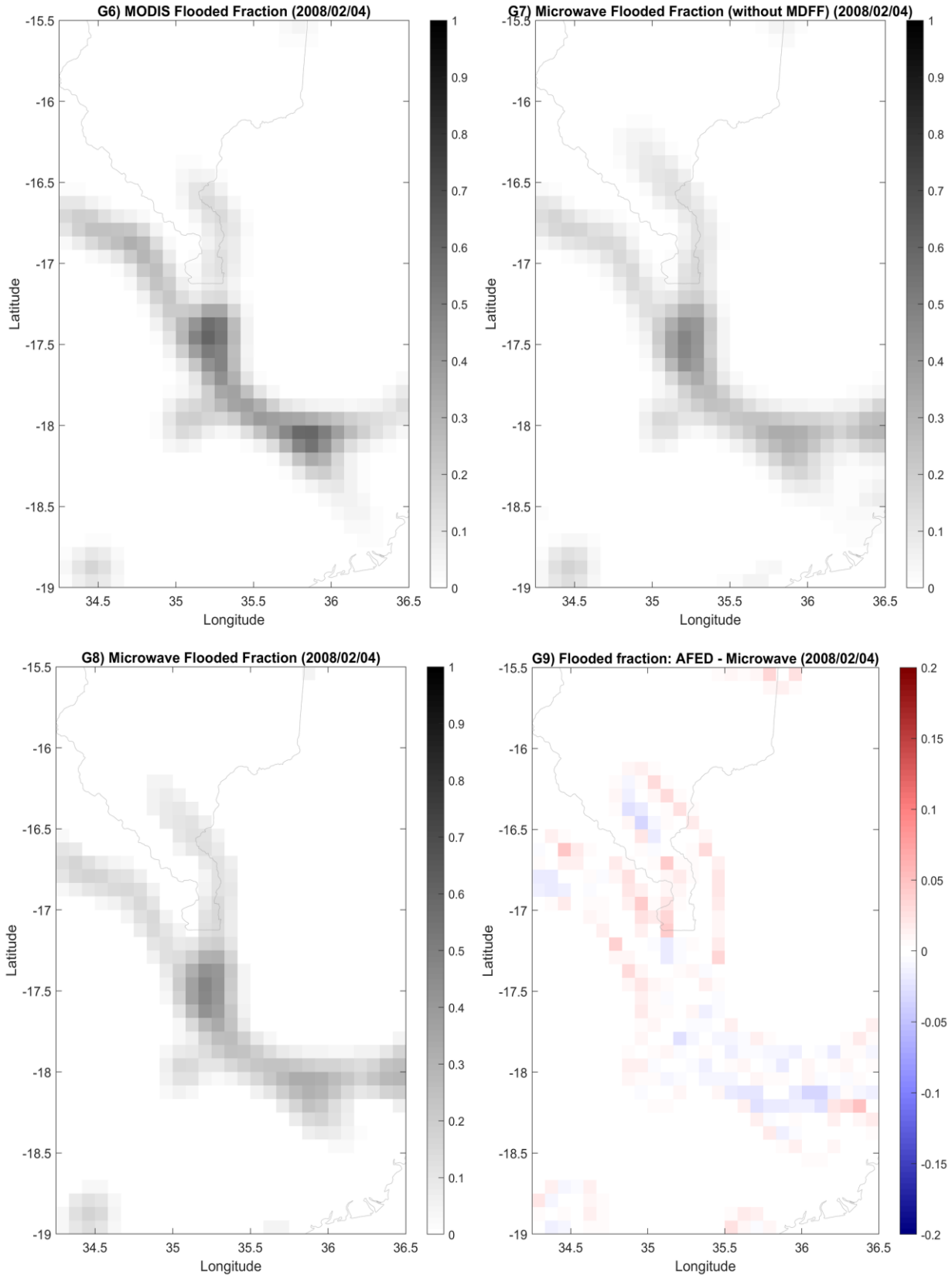
Zambezi2007 Discussion:

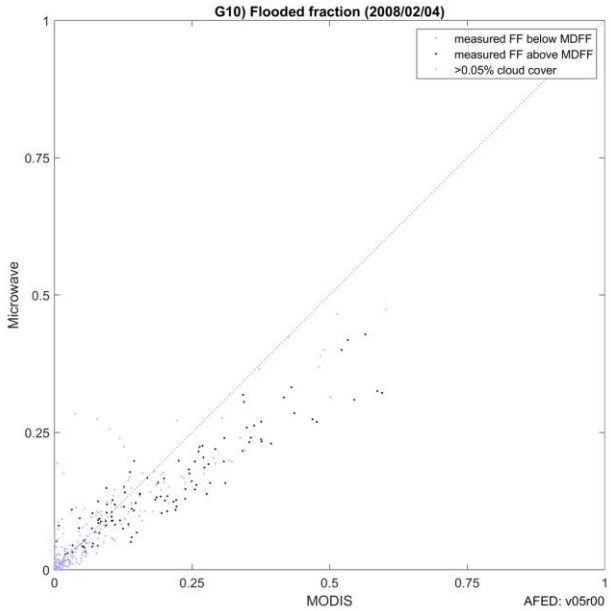
This case and Zambezi2008 provide two views of flooding in different years along the same river segments. Flooding was more extensive in the 2008 case than in 2007, with maximum FF of about 0.5 in 2008 and 0.25 in 2007. Correspondence between flooded fraction from the AFED algorithm (G7) and upscaled validation data (G6) is generally good except for areas where the AFED algorithm is negatively biased in both cases, e.g., on the western segment of the Zambezi River. The AFED algorithm does not flag this segment as seasonal wetlands and as a result it uses a more conservative (i.e., protective against false positives) dry-land end-member algorithm there that may underestimate flooding.

Case: Zambezi2008







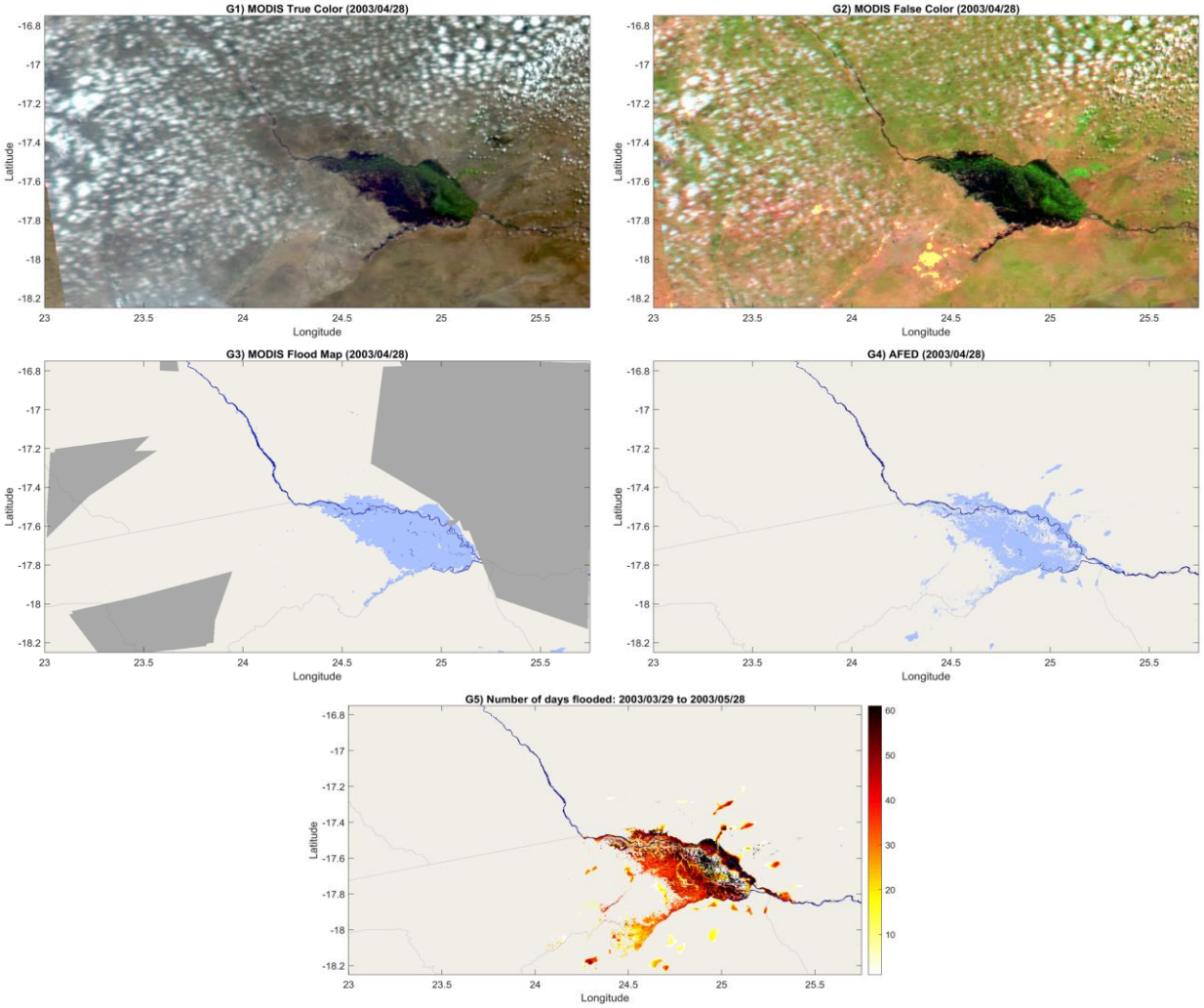


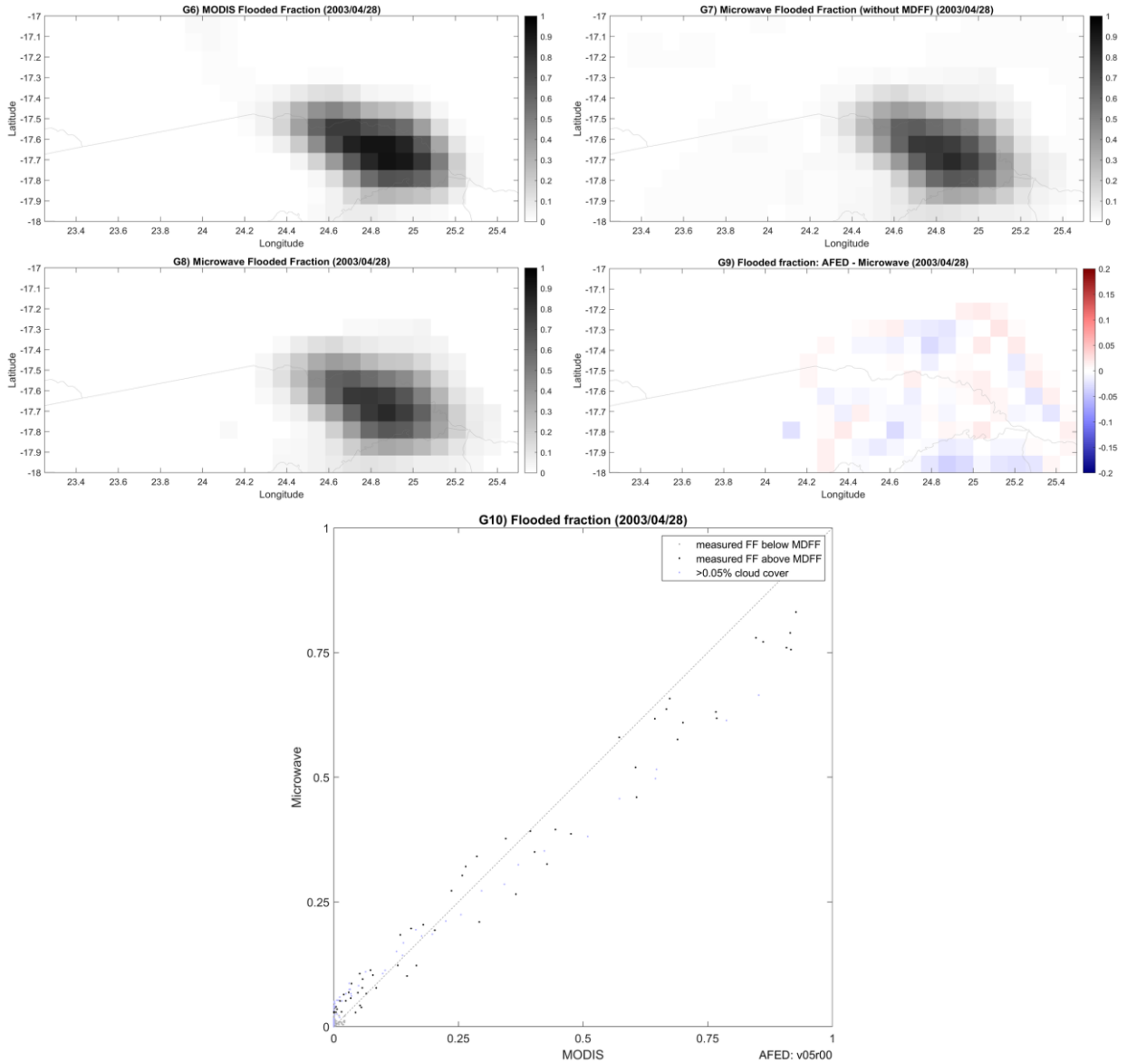
First row: G1) MODIS true color image and G2) MODIS false color image. Second row: G3) MODIS-derived flood map and G4) AFED flood map; light blue is flood water, dark blue is persistent open water, beige is dry land, and green is tree cover. Third row: G5) AFED cumulative number of days flooded for +/-30 days around the test scene date. Fourth row: G6) footprint-weighted flooded fraction derived from MODIS flood map and G7) microwave flooded fraction before applying MDFF. Fifth row: G8) microwave flooded fraction after applying MDFF and G9) difference between footprint-weighted AFED flooded fraction and microwave flooded fraction. Sixth row: G10) microwave-derived flooded fraction vs. MODIS-derived flooded fraction.

Zambezi2008 Discussion:

This case is discussed along with Zambezi2007 above.

Case: Caprivi2003



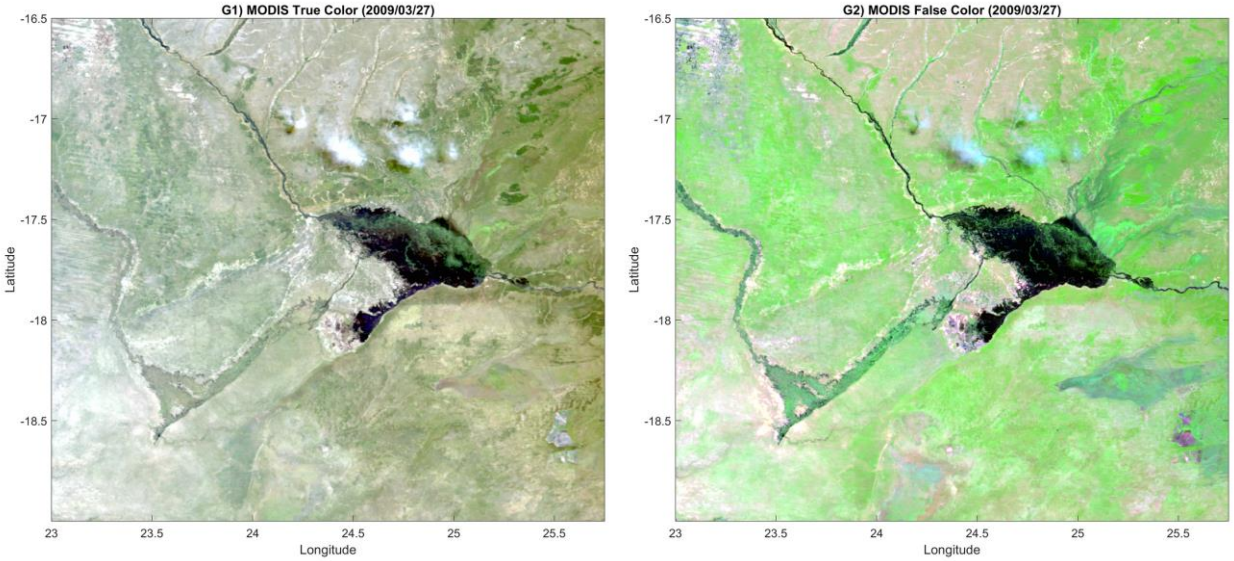


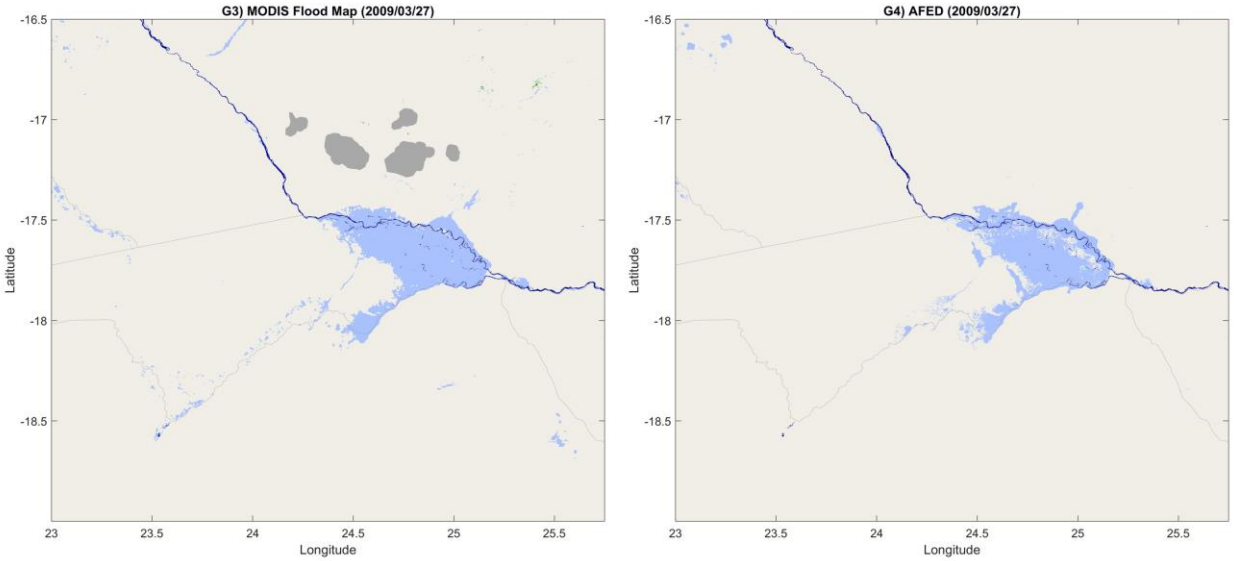
First row: G1) MODIS true color image and G2) MODIS false color image. Second row: G3) MODIS-derived flood map and G4) AFED flood map; light blue is flood water, dark blue is persistent open water, beige is dry land, and green is tree cover. Third row: G5) AFED cumulative number of days flooded for +/-30 days around the test scene date. Fourth row: G6) footprint-weighted flooded fraction derived from MODIS flood map and G7) microwave flooded fraction before applying MDFF. Fifth row: G8) microwave flooded fraction after applying MDFF and G9) difference between footprint-weighted AFED flooded fraction and microwave flooded fraction. Sixth row: G10) microwave-derived flooded fraction vs. MODIS-derived flooded fraction.

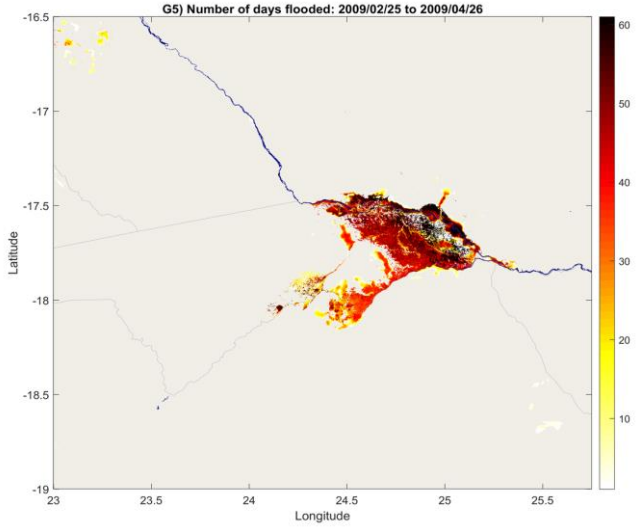
Caprivi2003 Discussion:

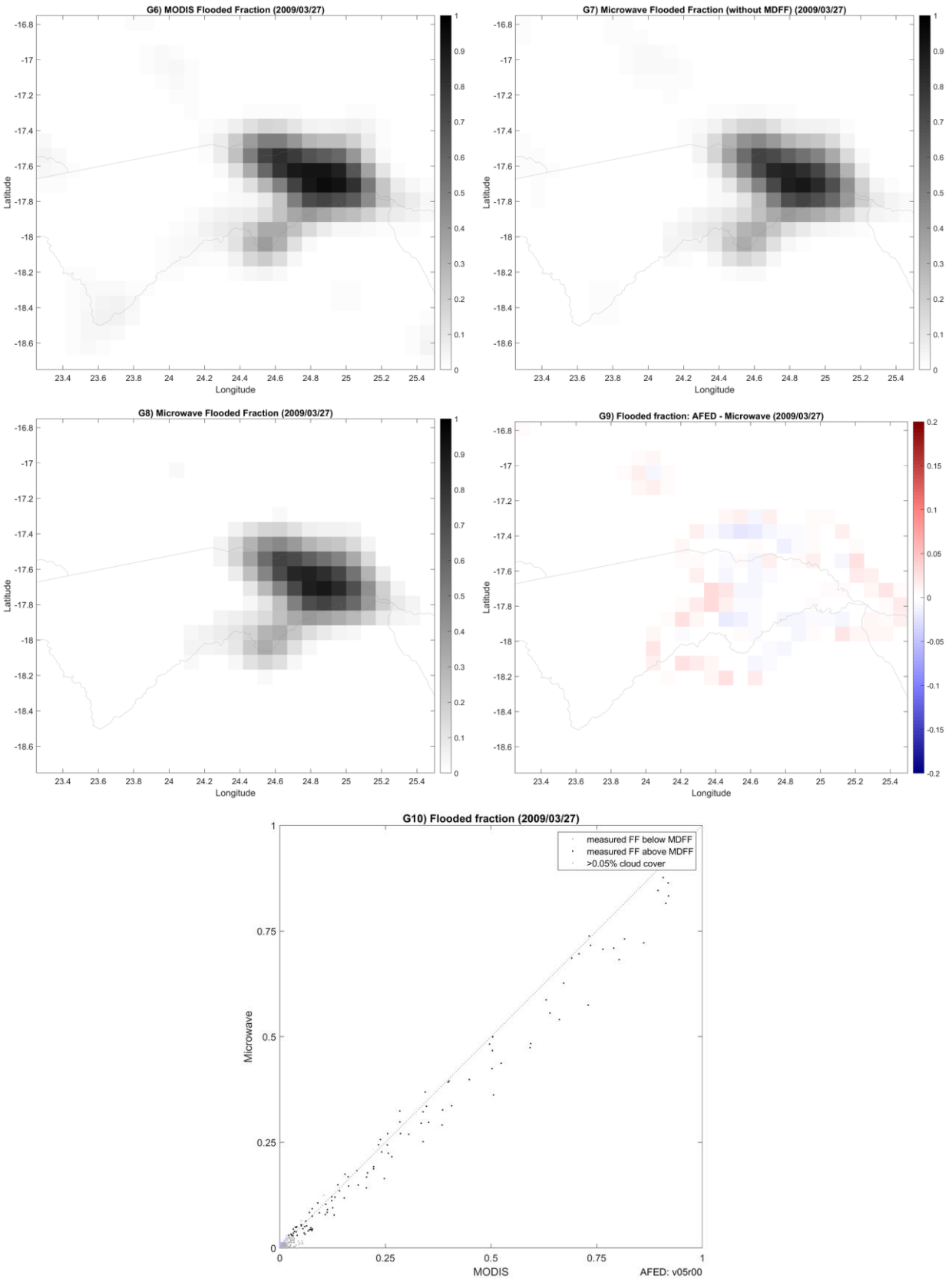
This case is very similar to Caprivi2009b and Caprivi2009a and the consistency provides additional evidence that algorithm performance is consistent from year-to-year. The correspondence of AFED and flooded fraction results to the validation data is excellent.

Case: Caprivi2009a









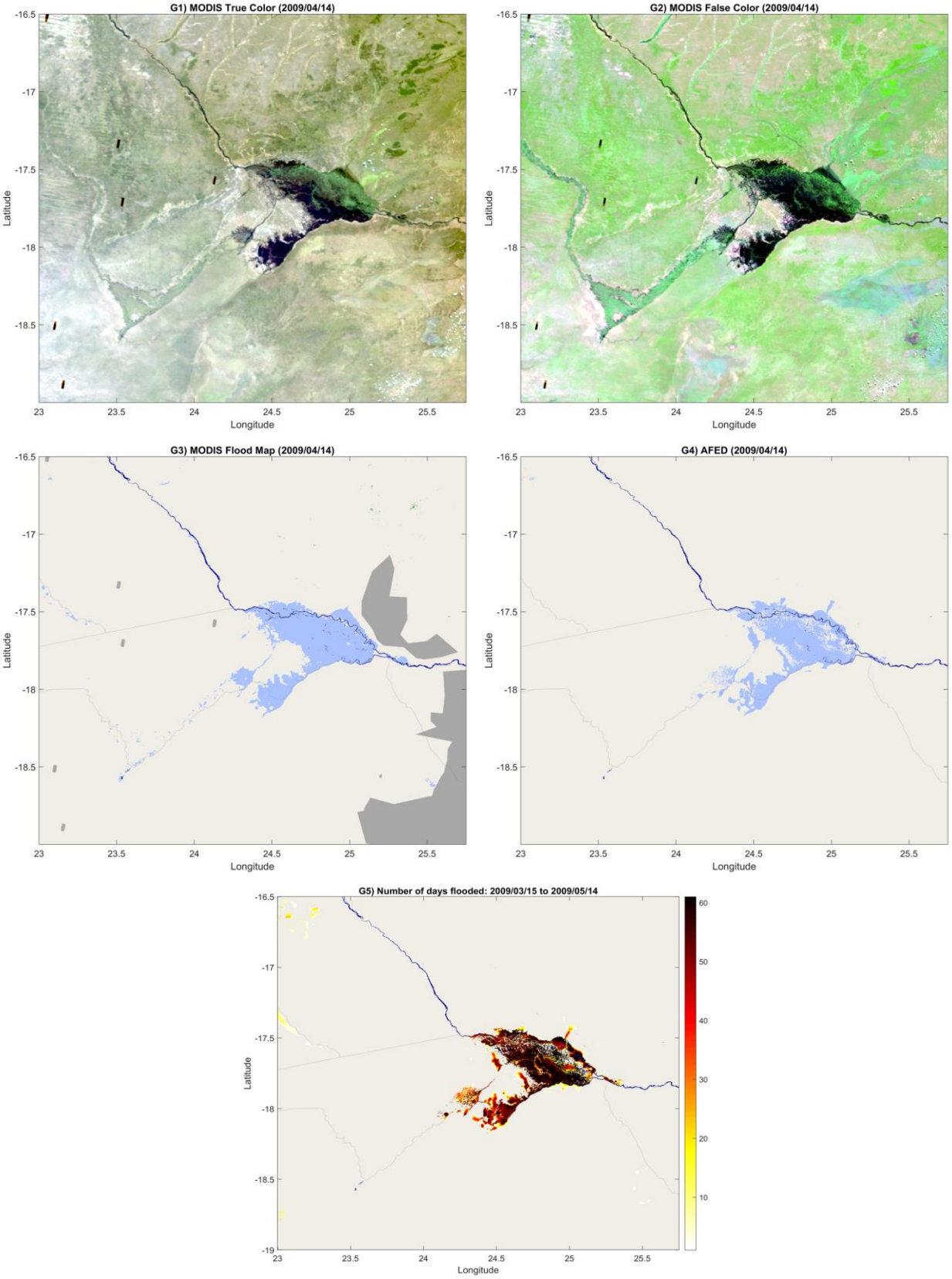
First row: G1) MODIS true color image and G2) MODIS false color image. Second row: G3)

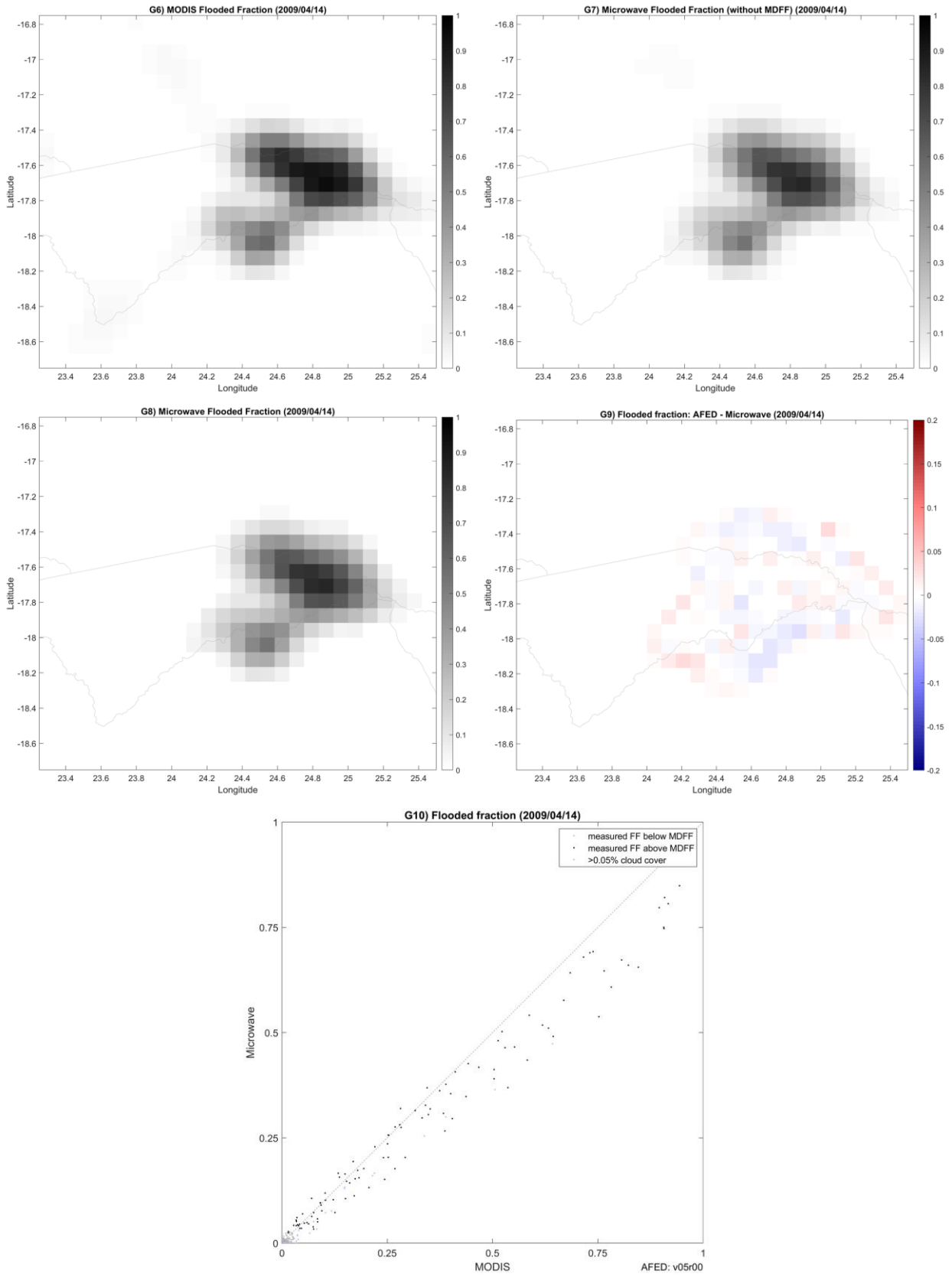
MODIS-derived flood map and G4) AFED flood map; light blue is flood water, dark blue is persistent open water, beige is dry land, and green is tree cover. Third row: G5) AFED cumulative number of days flooded for +/-30 days around the test scene date. Fourth row: G6) footprint-weighted flooded fraction derived from MODIS flood map and G7) microwave flooded fraction before applying MDFF. Fifth row: G8) microwave flooded fraction after applying MDFF and G9) difference between footprint-weighted AFED flooded fraction and microwave flooded fraction. Sixth row: G10) microwave-derived flooded fraction vs. MODIS-derived flooded fraction.

Caprivi2009a Discussion:

This case is very similar to Caprivi2009b and Caprivi2003 and the consistency provides additional evidence that algorithm performance is consistent over time. The correspondence of AFED results to the validation data is excellent. AFED flooded fraction underestimates MODIS flooded fraction by a small degree almost everywhere in the floodplain, which may be at least partially attributable to over-estimation of flooding by the MODIS image analysis process.

Case: Caprivi2009b





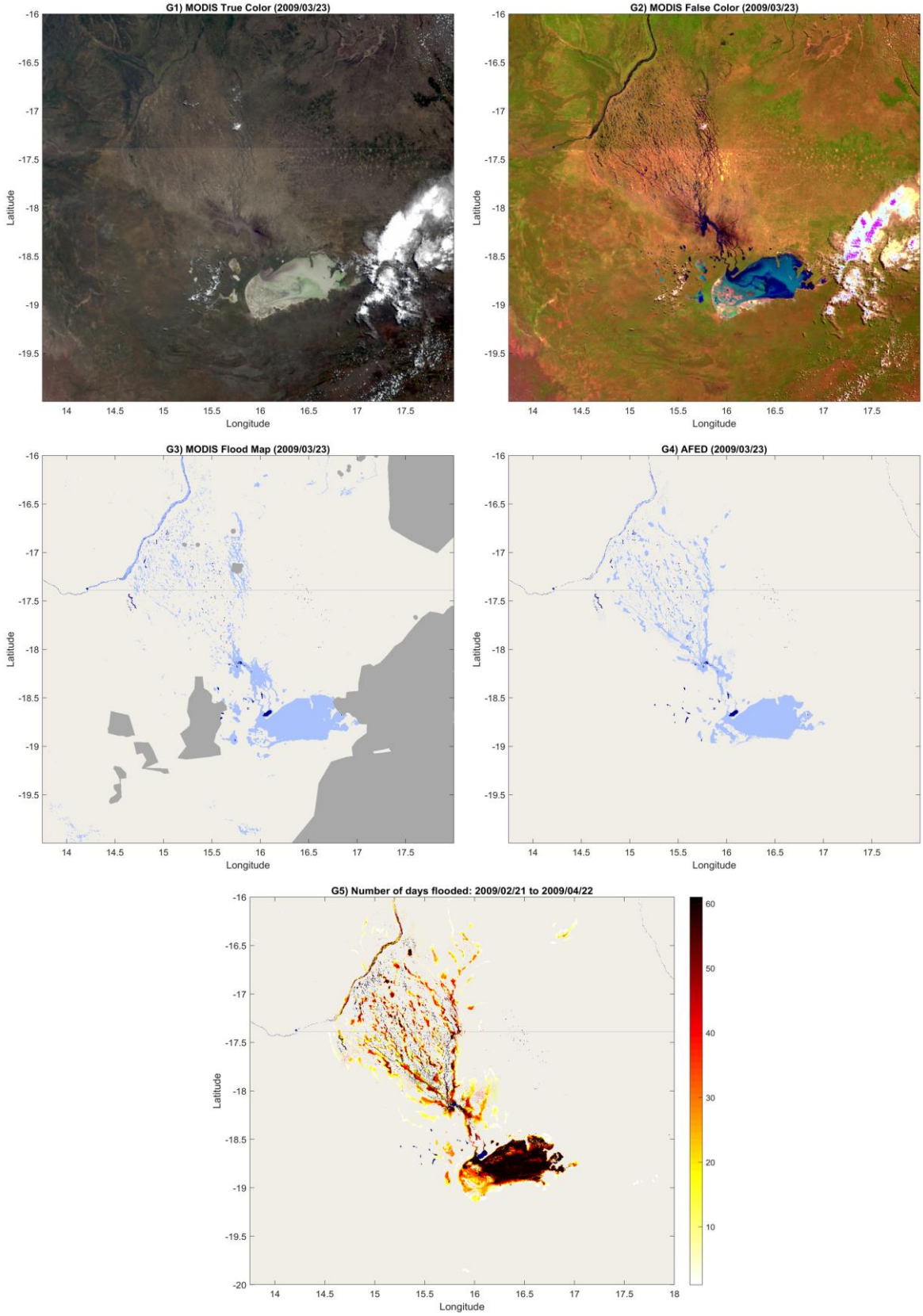
First row: G1) MODIS true color image and G2) MODIS false color image. Second row: G3)

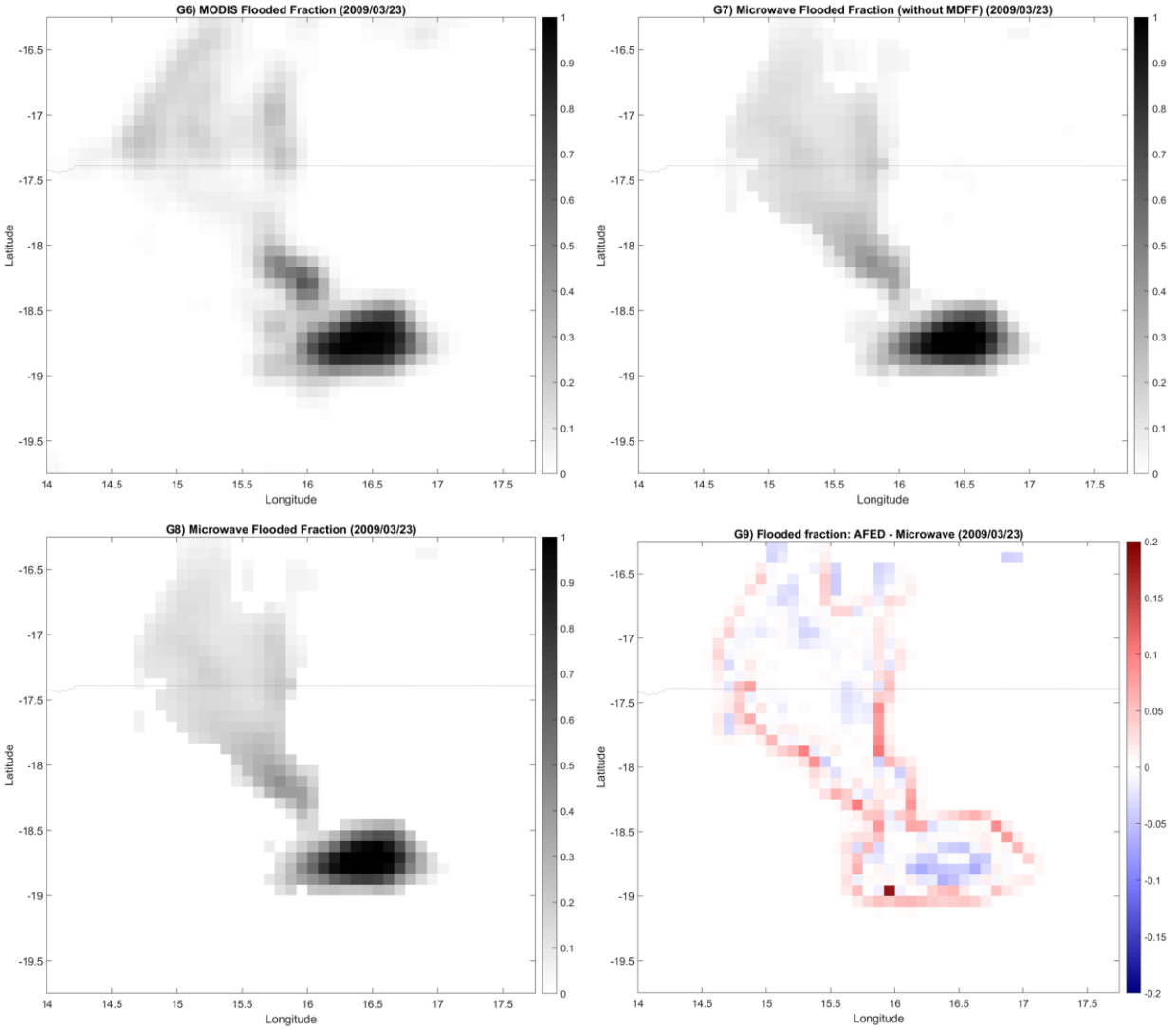
MODIS-derived flood map and G4) AFED flood map; light blue is flood water, dark blue is persistent open water, beige is dry land, and green is tree cover. Third row: G5) AFED cumulative number of days flooded for +/-30 days around the test scene date. Fourth row: G6) footprint-weighted flooded fraction derived from MODIS flood map and G7) microwave flooded fraction before applying MDFF. Fifth row: G8) microwave flooded fraction after applying MDFF and G9) difference between footprint-weighted AFED flooded fraction and microwave flooded fraction. Sixth row: G10) microwave-derived flooded fraction vs. MODIS-derived flooded fraction.

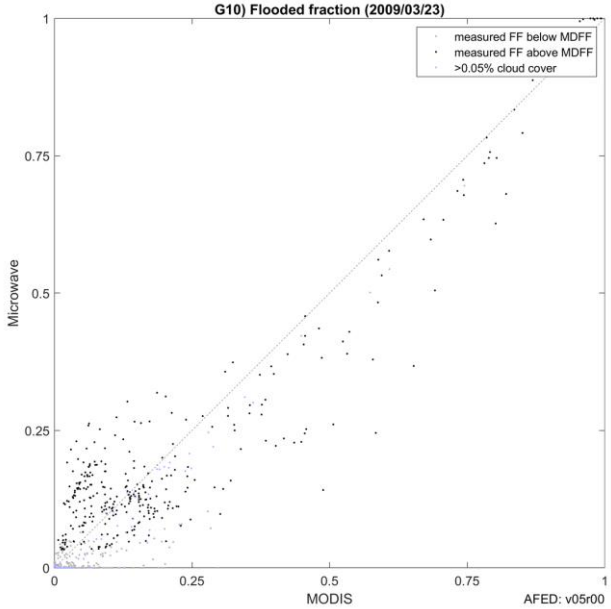
Caprivi2009b Discussion:

AFED algorithm performance is good in all aspects for this case, which is similar to cases Caprivi2009a and Caprivi2003. The flooded fraction values agree in all respects (G6-G10) and AFED (G3) accurately captures the significant features of the flood as depicted in the MODIS flood map.

Case: Etosha2009a





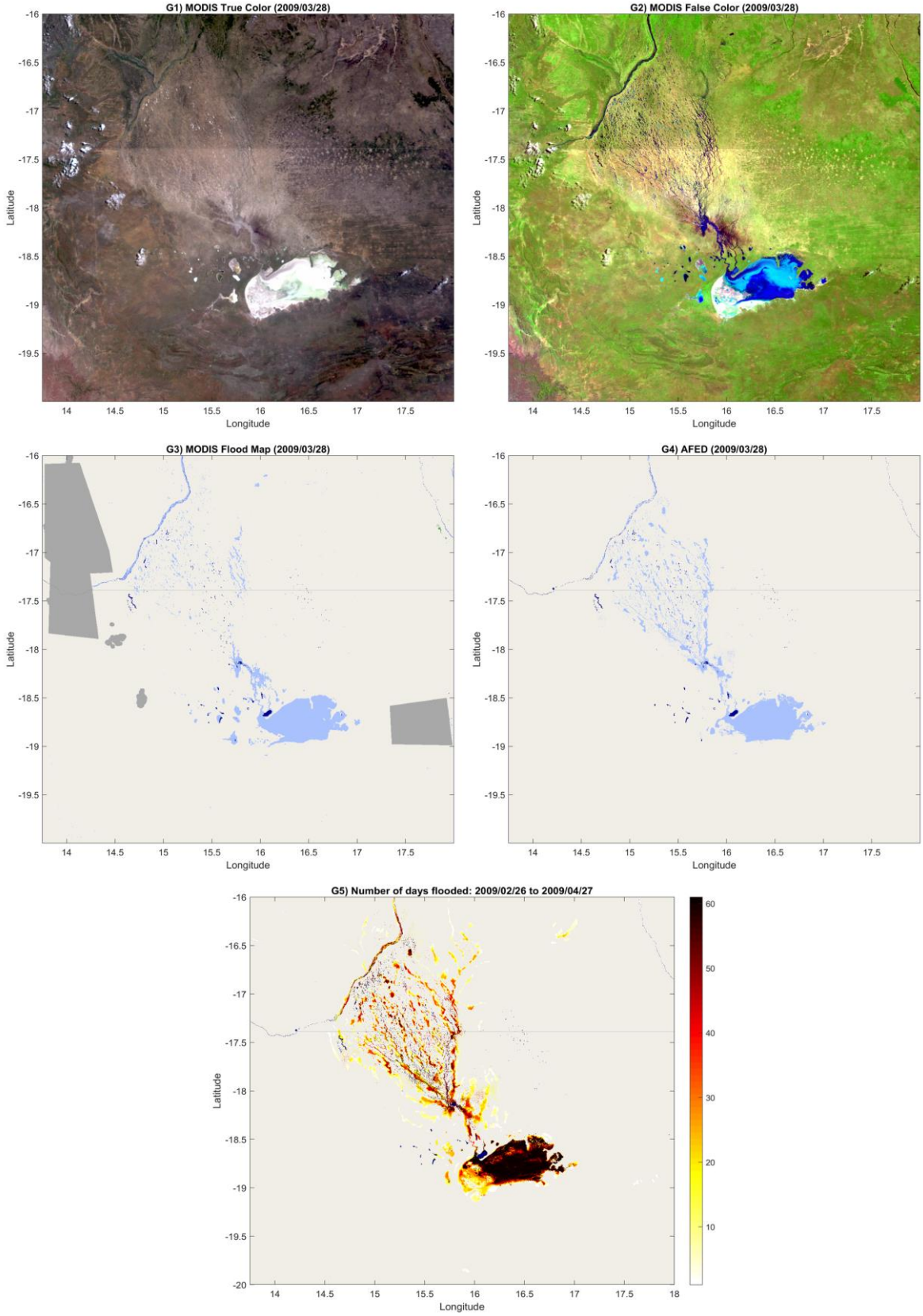


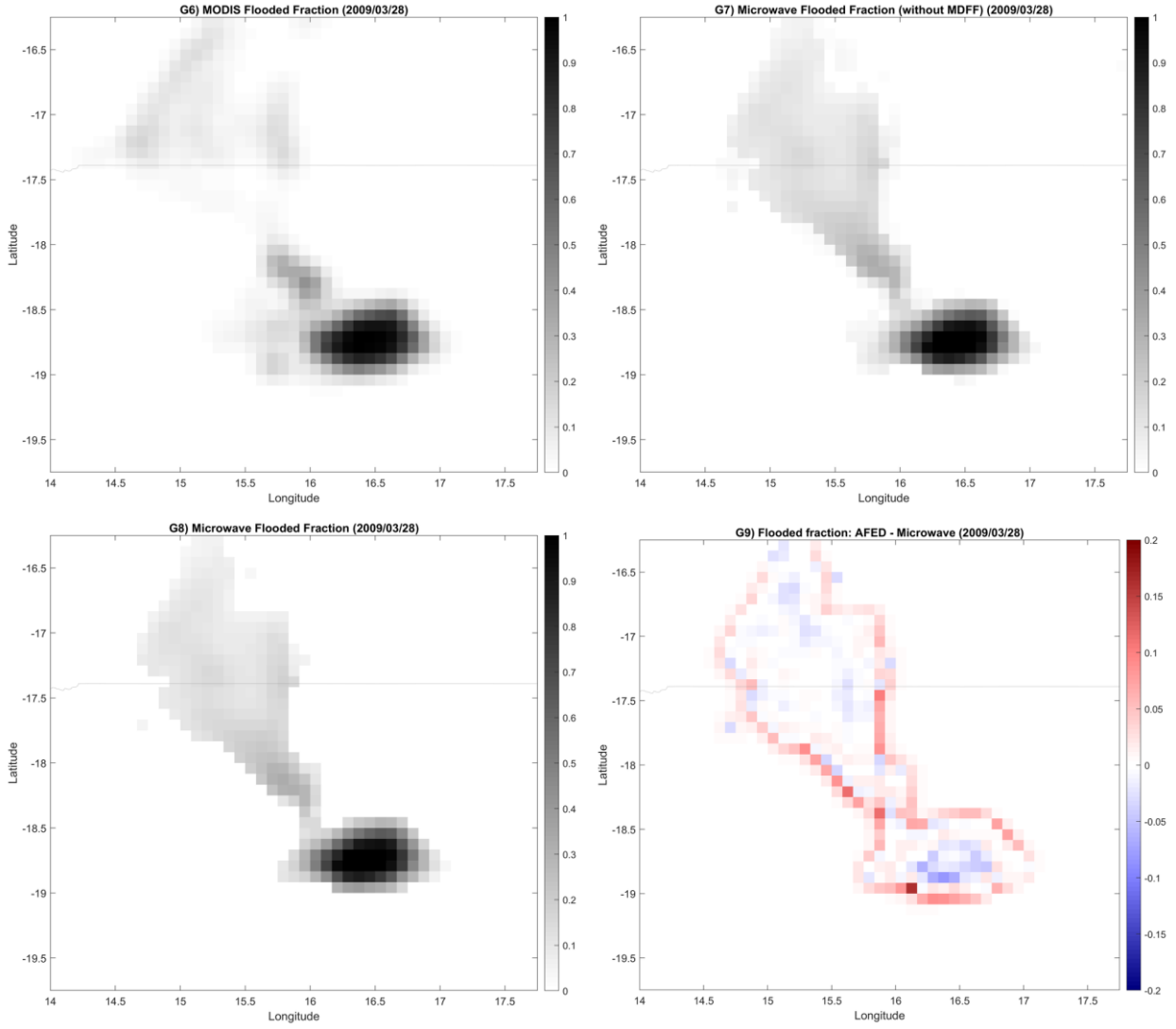
First row: G1) MODIS true color image and G2) MODIS false color image. Second row: G3) MODIS-derived flood map and G4) AFED flood map; light blue is flood water, dark blue is persistent open water, beige is dry land, and green is tree cover. Third row: G5) AFED cumulative number of days flooded for +/-30 days around the test scene date. Fourth row: G6) footprint-weighted flooded fraction derived from MODIS flood map and G7) microwave flooded fraction before applying MDFF. Fifth row: G8) microwave flooded fraction after applying MDFF and G9) difference between footprint-weighted AFED flooded fraction and microwave flooded fraction. Sixth row: G10) microwave-derived flooded fraction vs. MODIS-derived flooded fraction.

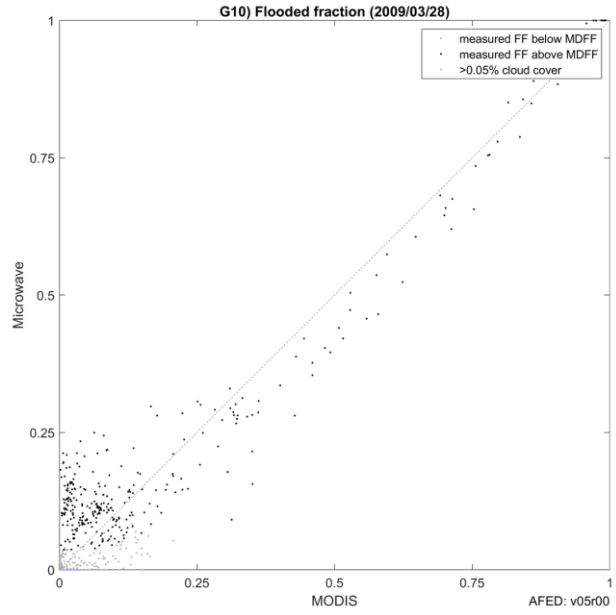
Etosha2009a Discussion:

This case is very similar to Etosha2009b (discussed further below) and the consistent results provide evidence that algorithm performance is consistent over time.

Case: Etosha2009b







First row: G1) MODIS true color image and G2) MODIS false color image. Second row: G3) MODIS-derived flood map and G4) AFED flood map; light blue is flood water, dark blue is persistent open water, beige is dry land, and green is tree cover. Third row: G5) AFED cumulative number of days flooded for +/-30 days around the test scene date. Fourth row: G6) footprint-weighted flooded fraction derived from MODIS flood map and G7) microwave flooded fraction before applying MDFF. Fifth row: G8) microwave flooded fraction after applying MDFF and G9) difference between footprint-weighted AFED flooded fraction and microwave flooded fraction. Sixth row: G10) microwave-derived flooded fraction vs. MODIS-derived flooded fraction.

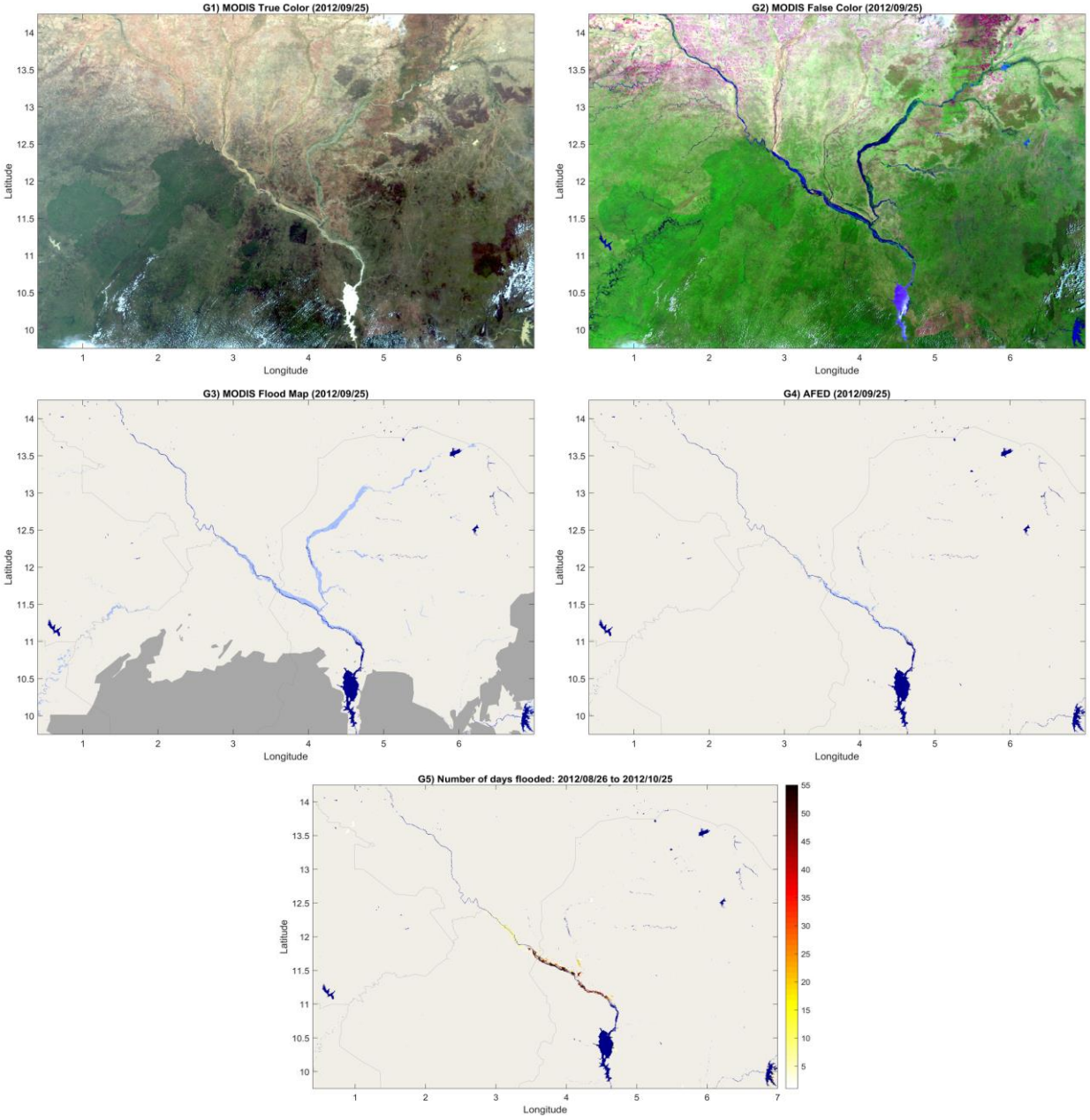
Etosha2009b Discussion:

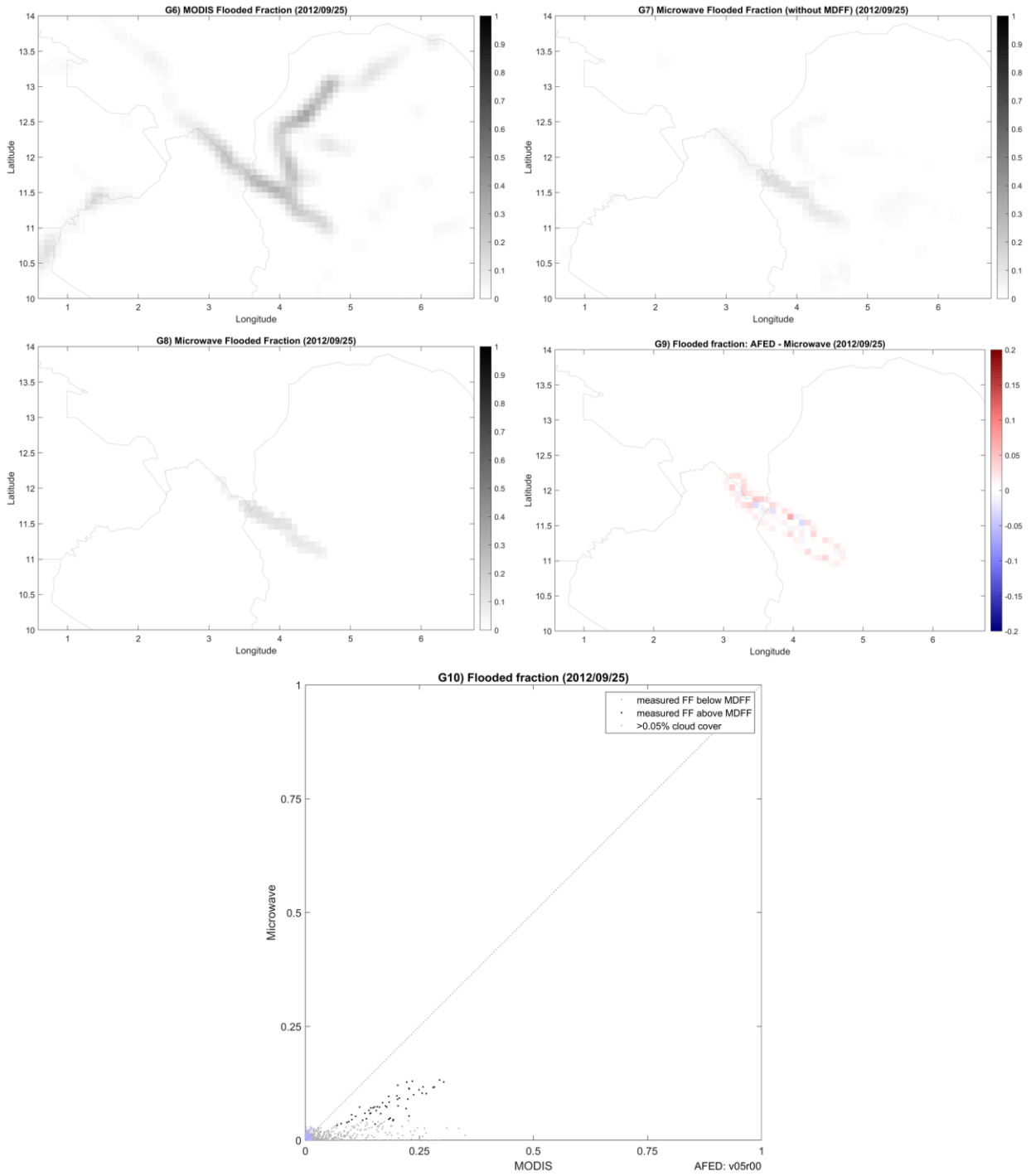
This case is very similar to Etosha2009a. The AFED algorithm appears to represent flooding fairly well overall and particularly well (1) where flooded fraction is highest in and around the Etosha Pan and (2) along the section of river in the northwest corner of the domain.

The distinct vegetation cover difference between Angola and Namibia is apparent in true- and false-color MODIS images (G1 and G2) and this difference appears to affect MODIS sensitivity to water. The effect is most obvious in the MODIS flooded fraction graphic (G6) where flooded fraction values are distinctly lower in Namibia than in Angola. The divide possibly has an effect in the microwave-derived FF results before (G7) and more so after (G8) applying the MDFF (G8). This is likely due to denser vegetation cover north of the border providing a more consistent background microwave signature, and thus a lower MDFF. Regardless, the difference is relatively minor and it seems reasonable to conclude that the AFED algorithm results are as accurate in Namibia as they are in Angola where they are in closer agreement with the validation data. Negative biases in the MODIS data would also explain the cluster of positively biased microwave algorithm FF points in G10.

AFED V05R00 eliminates the striping pattern attributable to SRTM elevation model artifacts that was apparent in V04R01 maps of this area.

Case: Niger2012





First row: G1) MODIS true color image and G2) MODIS false color image. Second row: G3) MODIS-derived flood map and G4) AFED flood map; light blue is flood water, dark blue is persistent open water, beige is dry land, and green is tree cover. Third row: G5) AFED cumulative number of days flooded for +/-30 days around the test scene date. Fourth row: G6) footprint-weighted flooded fraction derived from MODIS flood map and G7) microwave flooded fraction before applying MDFF. Fifth row: G8) microwave flooded fraction after applying MDFF and G9) difference between footprint-weighted AFED flooded fraction and microwave

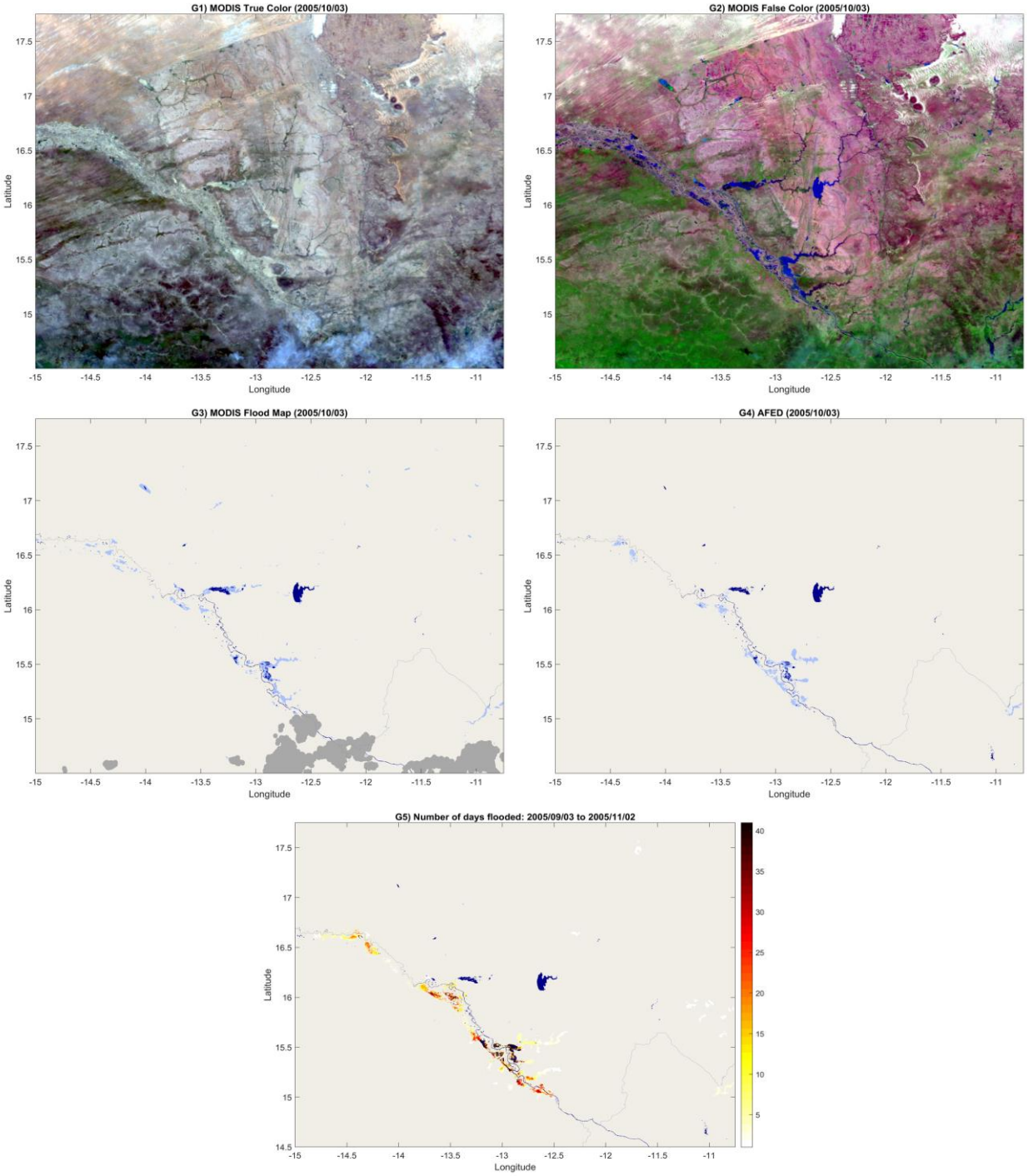
flooded fraction. Sixth row: G10) microwave-derived flooded fraction vs. MODIS-derived flooded fraction.

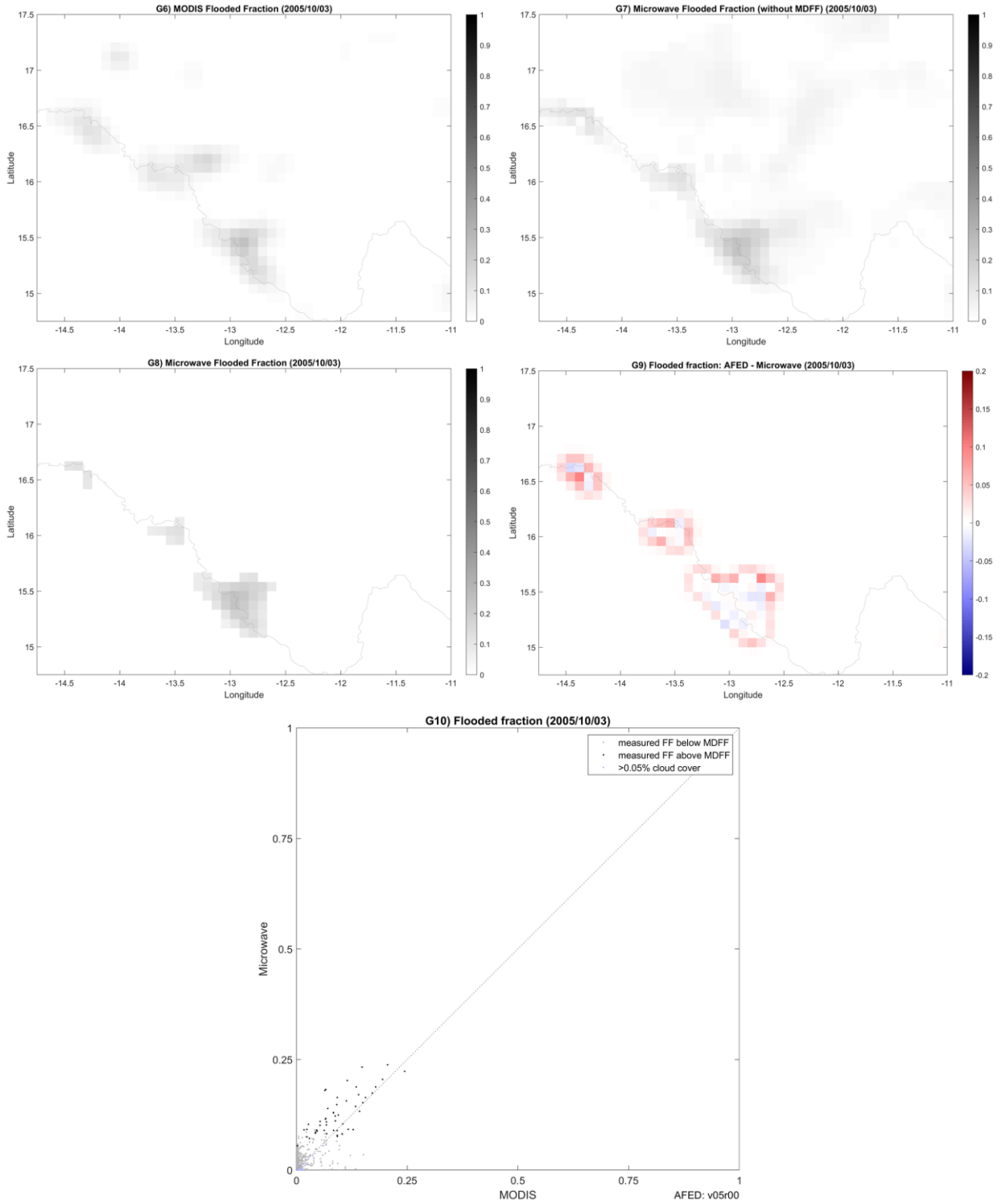
Niger2012 Discussion:

AFED (G4) significantly underestimates flooding compared to the MODIS flood map (G3) in this case. The algorithm detects flooding but it is restricted to the stretch of the Niger River in the region center (G4 and G7/G8); AFED is minimal or absent along the tributaries and the rest of the Niger River where MODIS shows flooding (G3). The microwave flooded fraction underestimates flood extent both before (G7) and after (G8) application of MDFF. The cumulative flood extent (G5) does show more flooding, but still not as much as the MODIS scene.

The negative bias may be attributable in part to the high dry-land end-member values, Q_{dry}, that occur in arid, less vegetated areas. This is particularly true in the northeast part of this scene where MODIS shows flooding along the Sokoto River in Nigeria. Sensitivity to flooding is lower where Q_{dry} is high because the contrast between Q_{dry} and the water end-member is reduced. Year-to-year variability in dry-land conditions in less vegetated areas may also be a contributing factor because the dry-land end-member algorithm adapts to the variability by reducing the weight of historical-average Q conditions in the Q_{dry} estimation. Ideally the algorithm would revert to historical-average Q or Q_{dry} derived from NDVI and f_{tree} during flooding but this is not always possible because it would lead to higher false-positive rates by raising the flooded fraction when flooding is not present.

Case: Senegal2005





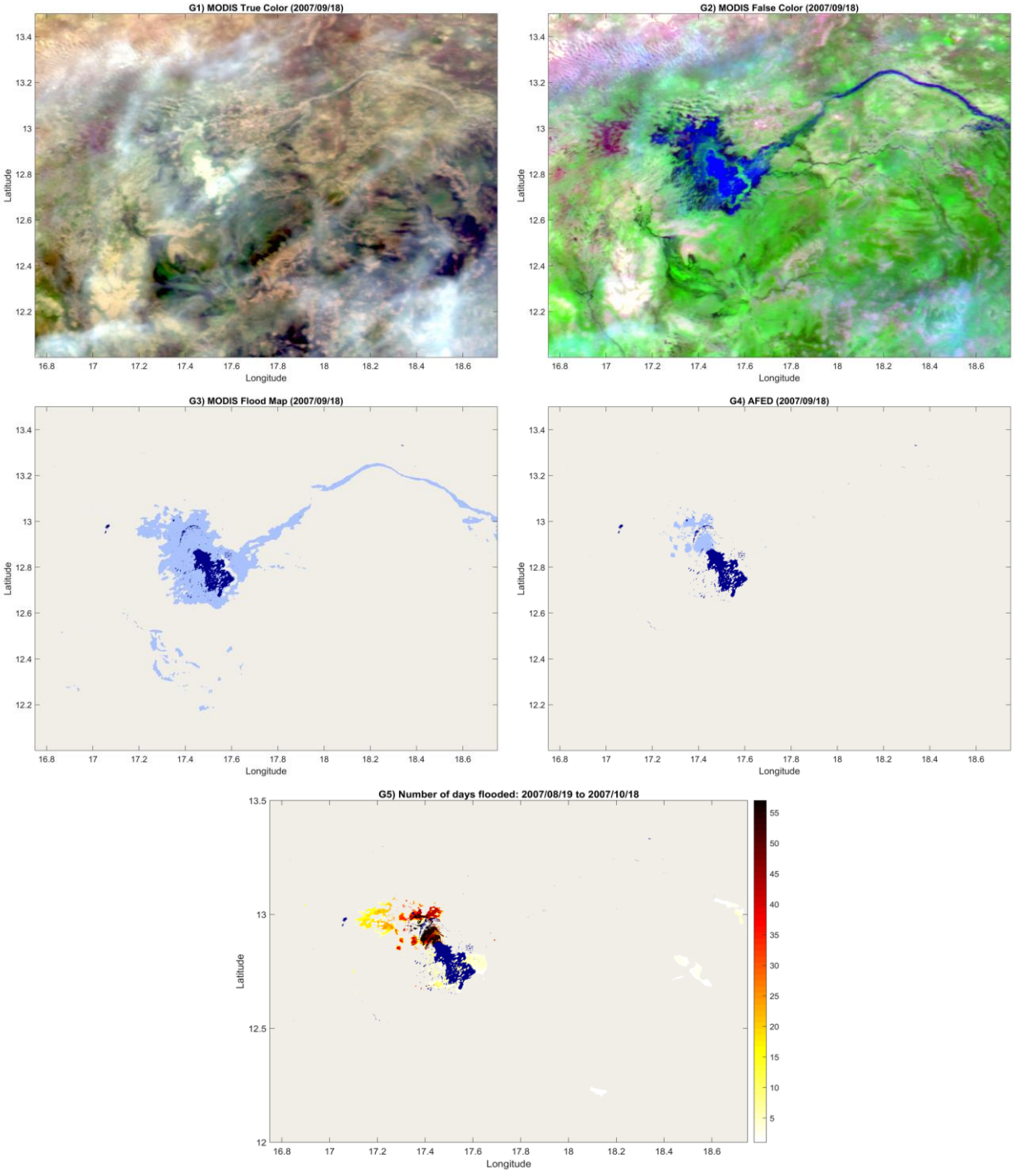
First row: G1) MODIS true color image and G2) MODIS false color image. Second row: G3) MODIS-derived flood map and G4) AFED flood map; light blue is flood water, dark blue is persistent open water, beige is dry land, and green is tree cover. Third row: G5) AFED cumulative number of days flooded for +/-30 days around the test scene date. Fourth row: G6) footprint-weighted flooded fraction derived from MODIS flood map and G7) microwave flooded fraction before applying MDFF. Fifth row: G8) microwave flooded fraction after applying

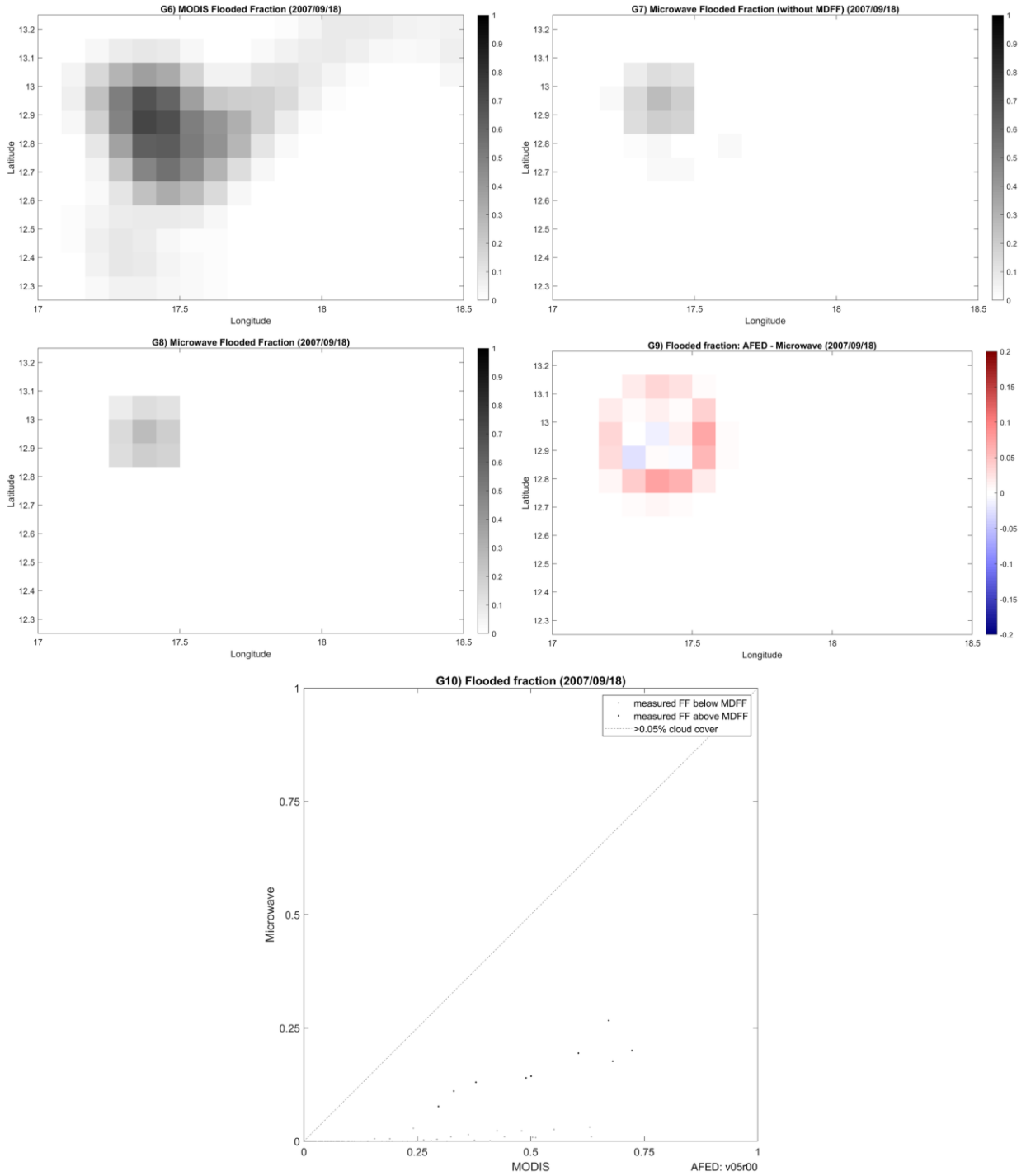
MDFF and G9) difference between footprint-weighted AFED flooded fraction and microwave flooded fraction. Sixth row: G10) microwave-derived flooded fraction vs. MODIS-derived flooded fraction.

Senegal2005 Discussion:

Flooded fraction results (G10) indicate fairly good agreement between the AFED and validation scenes. However, significant areas of lower-level flooding are excluded from AFED when MDFF is applied. AFED flooding (G4) is appropriately confined to the flood plain and apparent false positives outside the flood plain occur only rarely during the entire year (G5). There are areas of discrepancy between the details of the AFED and validation flood maps, which are in part the result of unevenness, braiding, and backwater areas within the Senegal floodplain that are difficult for the elevation model to represent accurately and may change over time.

Case: Fitri2007





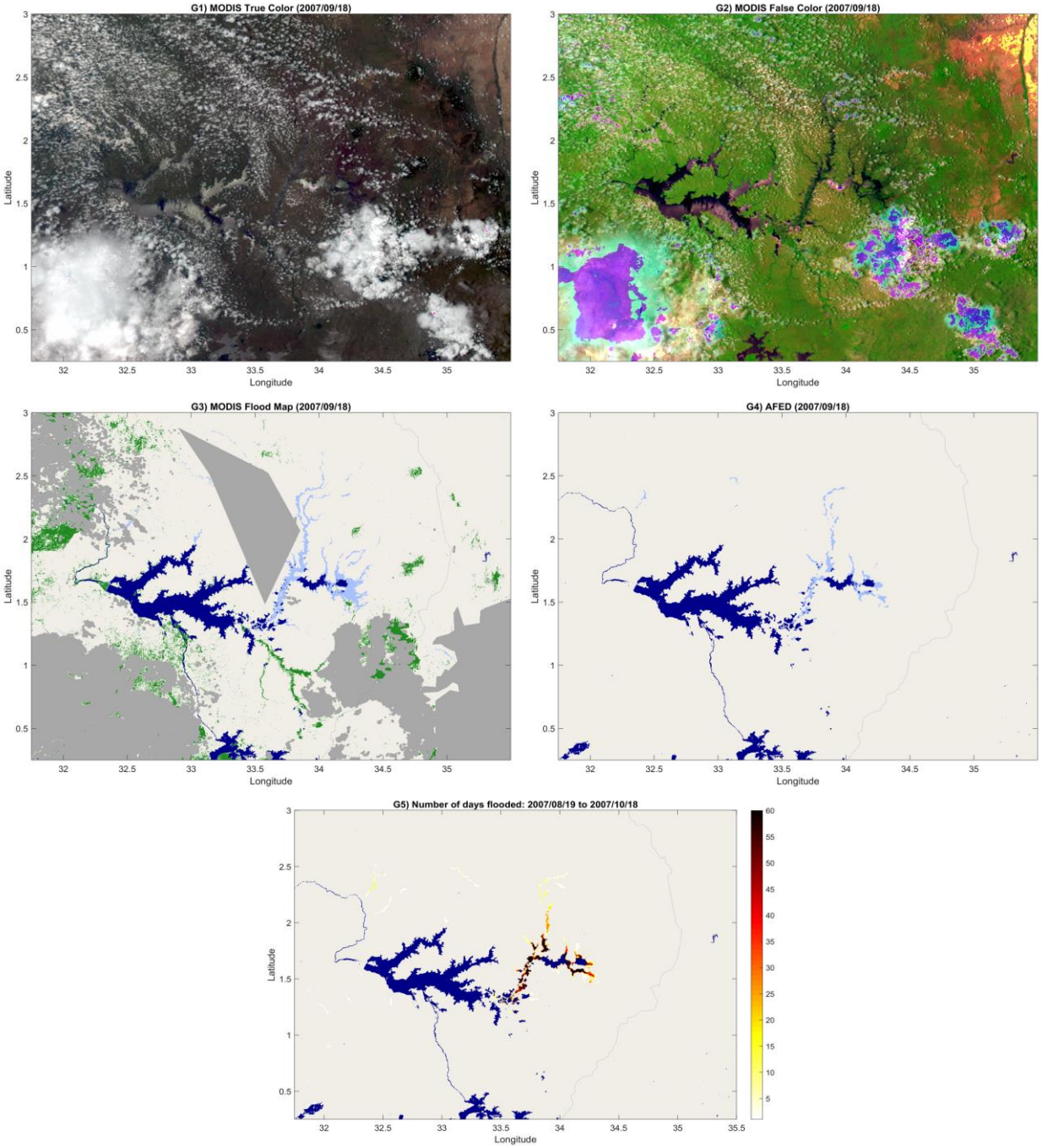
First row: G1) MODIS true color image and G2) MODIS false color image. Second row: G3) MODIS-derived flood map and G4) AFED flood map; light blue is flood water, dark blue is persistent open water, beige is dry land, and green is tree cover. Third row: G5) AFED cumulative number of days flooded for +/-30 days around the test scene date. Fourth row: G6) footprint-weighted flooded fraction derived from MODIS flood map and G7) microwave flooded fraction before applying MDFF. Fifth row: G8) microwave flooded fraction after applying MDFF and G9) difference between footprint-weighted AFED flooded fraction and microwave

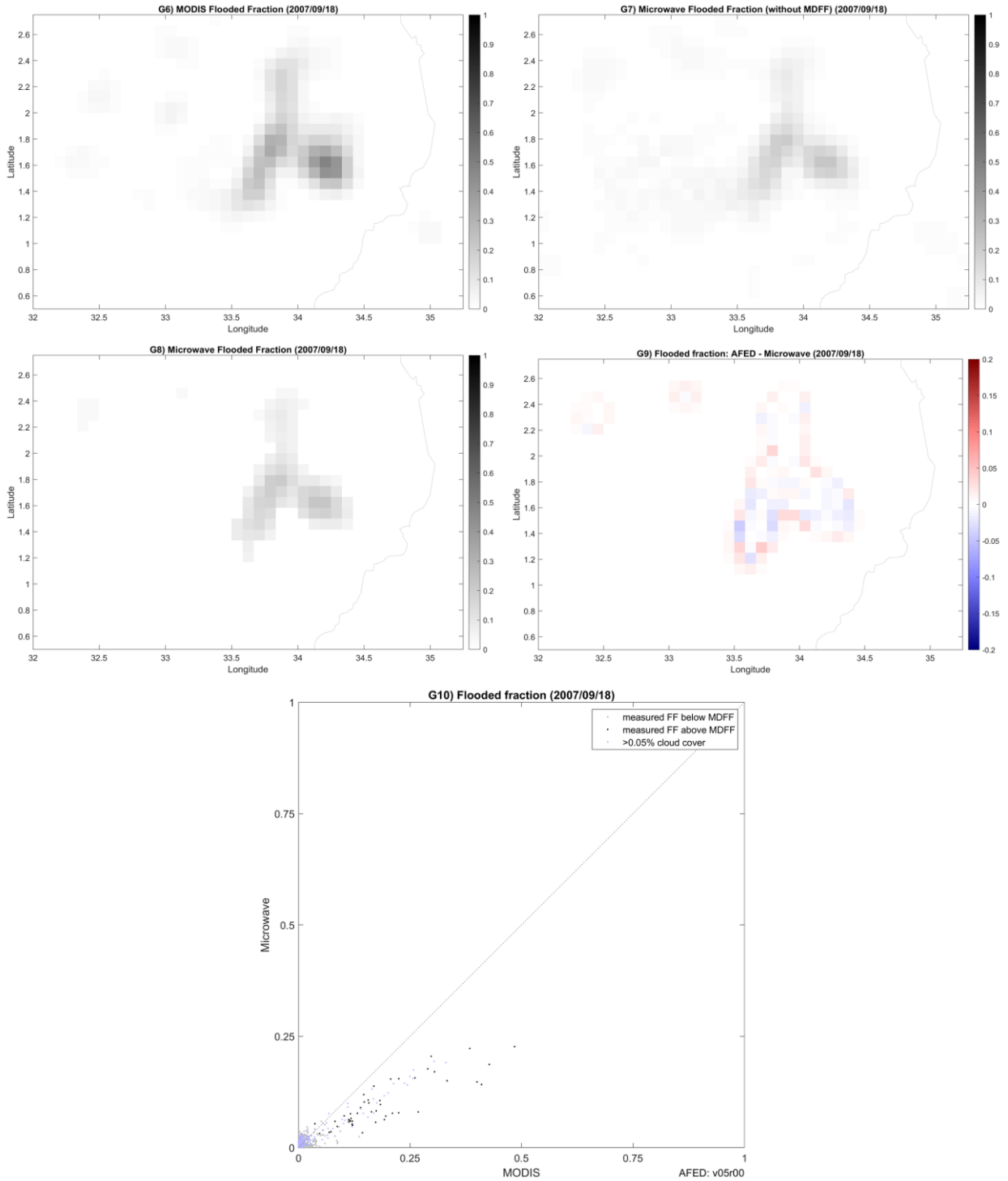
flooded fraction. Sixth row: G10) microwave-derived flooded fraction vs. MODIS-derived flooded fraction.

Fitri2007 Discussion:

Flooding appears to be widely underestimated in this case. The floodplain extending to the east of Lake Fitri may be too narrow to be detectable, although the flooded fraction graphics (G7 and G8) indicate that flooding is not being detected in the area even prior to MDFF application. High dry-land end-member—discussed above in the Niger2012 case—may be a factor. As of V04R01, the AFED algorithm no longer flags the areas around Lake Fitri as wetlands, which means that V04R01 dry-land end-member is higher than V04R00. Although the change in wetlands flagging was perhaps not ideal for this case, there were significant reductions in V04R01 apparent false positives in other areas flagged as wetlands in V04R00 but not V04R01. (Appendix B discusses changes between V04R00 and V04R01.) There is considerable unflagged cloud or smoke contamination that may be affecting the accuracy of the MODIS flood extent analysis. Additional time-series analyses of this case are warranted to determine if, e.g., adjustments to the dry-land end-member may improve the flood depiction or if the removal of short-term false-positive “flashes” is negating the flood signal.

Case: Kyoga2007





First row: G1) MODIS true color image and G2) MODIS false color image. Second row: G3) MODIS-derived flood map and G4) AFED flood map; light blue is flood water, dark blue is persistent open water, beige is dry land, and green is tree cover. Third row: G5) AFED cumulative number of days flooded for +/-30 days around the test scene date. Fourth row: G6) footprint-weighted flooded fraction derived from MODIS flood map and G7) microwave flooded fraction before applying MDFF. Fifth row: G8) microwave flooded fraction after applying MDFF and G9) difference between footprint-weighted AFED flooded fraction and microwave

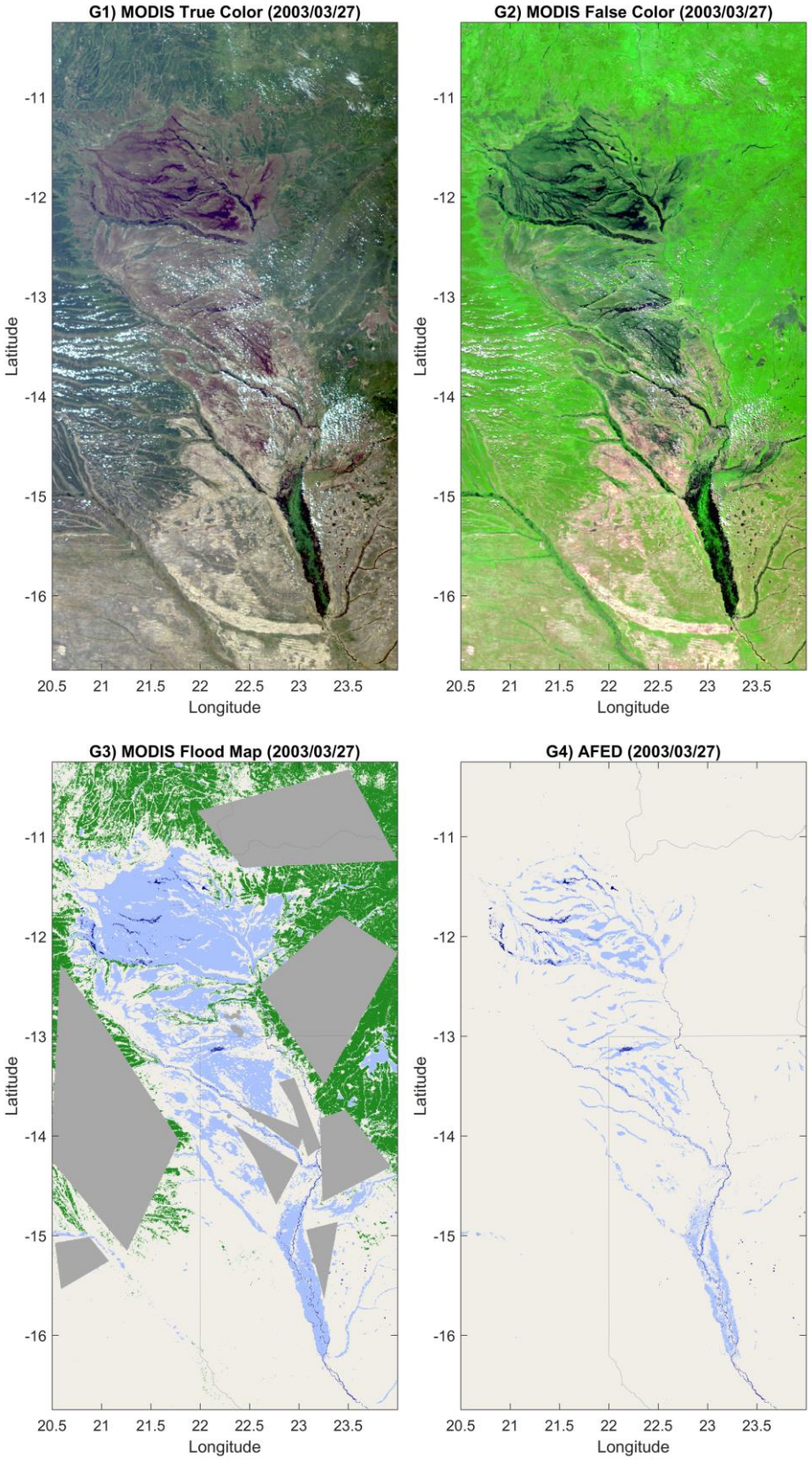
flooded fraction. Sixth row: G10) microwave-derived flooded fraction vs. MODIS-derived flooded fraction.

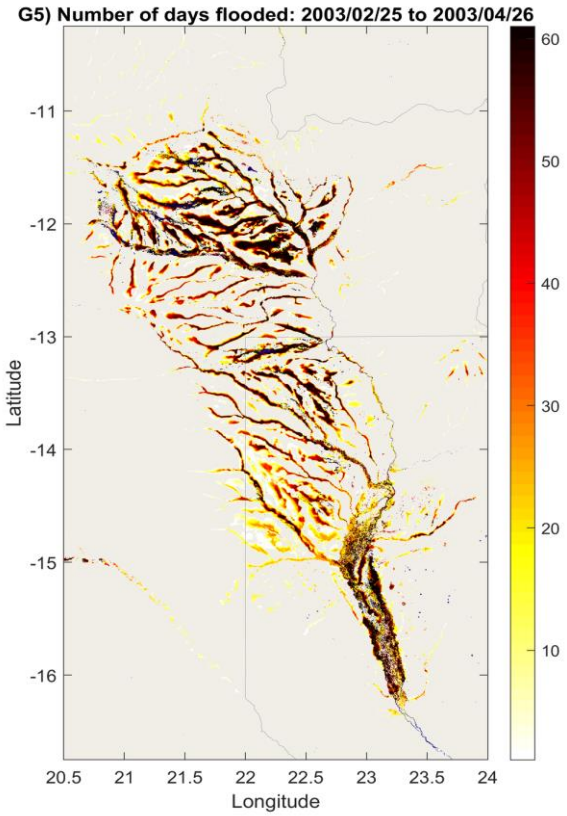
Kyoga2007 Discussion:

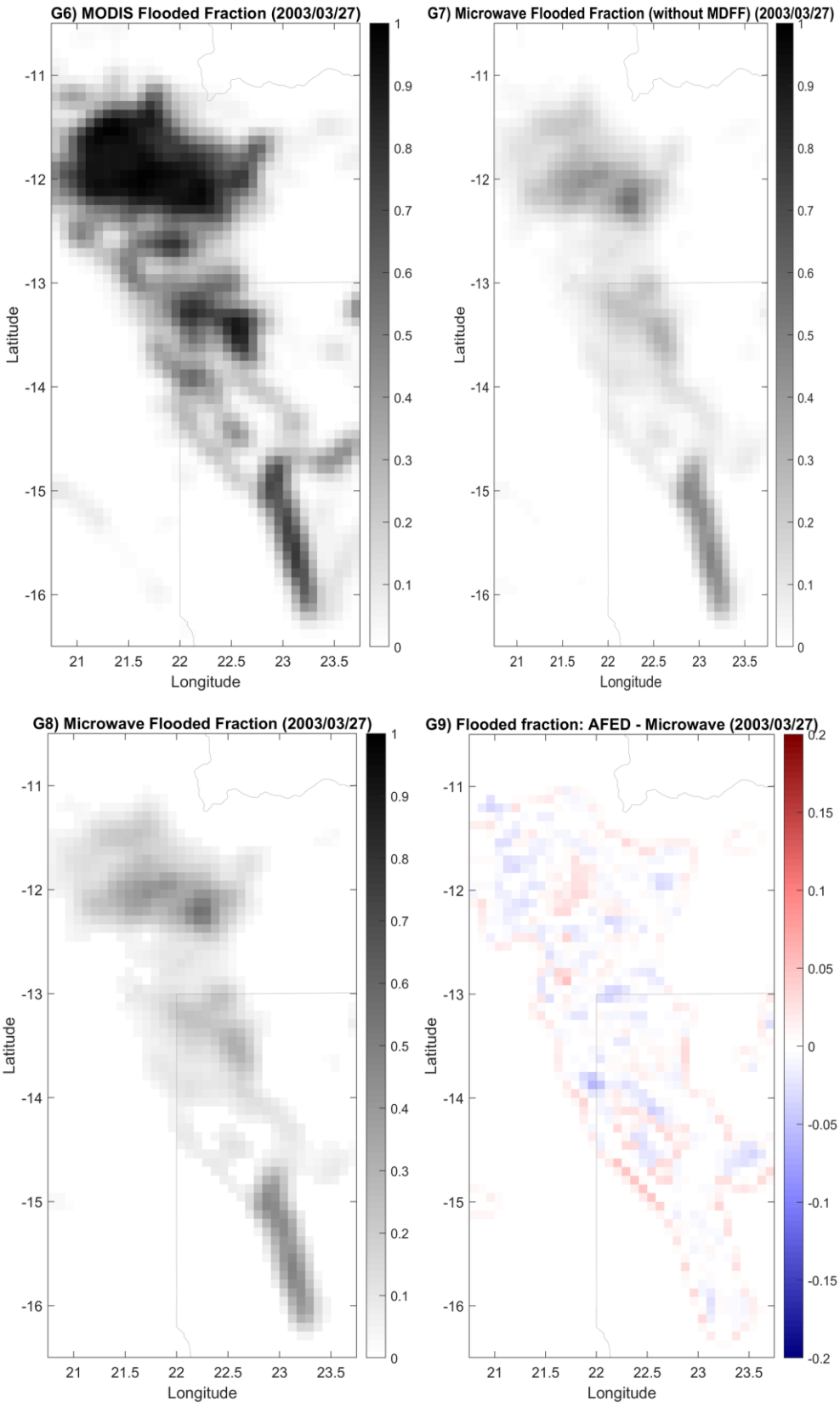
The AFED (G4) shows good spatial correspondence to the MODIS validation imagery (G3), but with a somewhat lower amount of flooding. The cumulative number of days flooded (G5) shows that during this period AFED likely never reach the extent of flooding indicated by the validation map. V05R00 improves slightly upon V04R01, with better representation of flood extent details at the northern end of the flooded region.

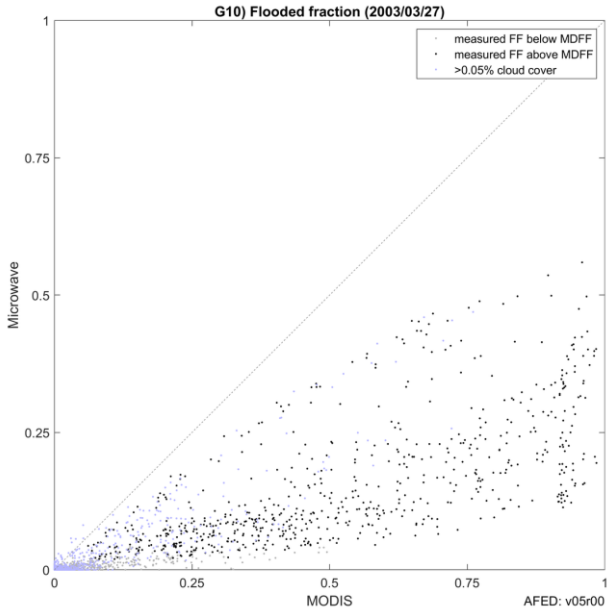
The FF scatter plot (G10) shows that there is good correlation in this case between the microwave-derived FF and validation scene (MODIS) flooded fractions. However, the microwave algorithm FF has a negative bias trend that increases with FF, a pattern that is repeated in several other cases (discussed in the summary).

Case: Barotse2003







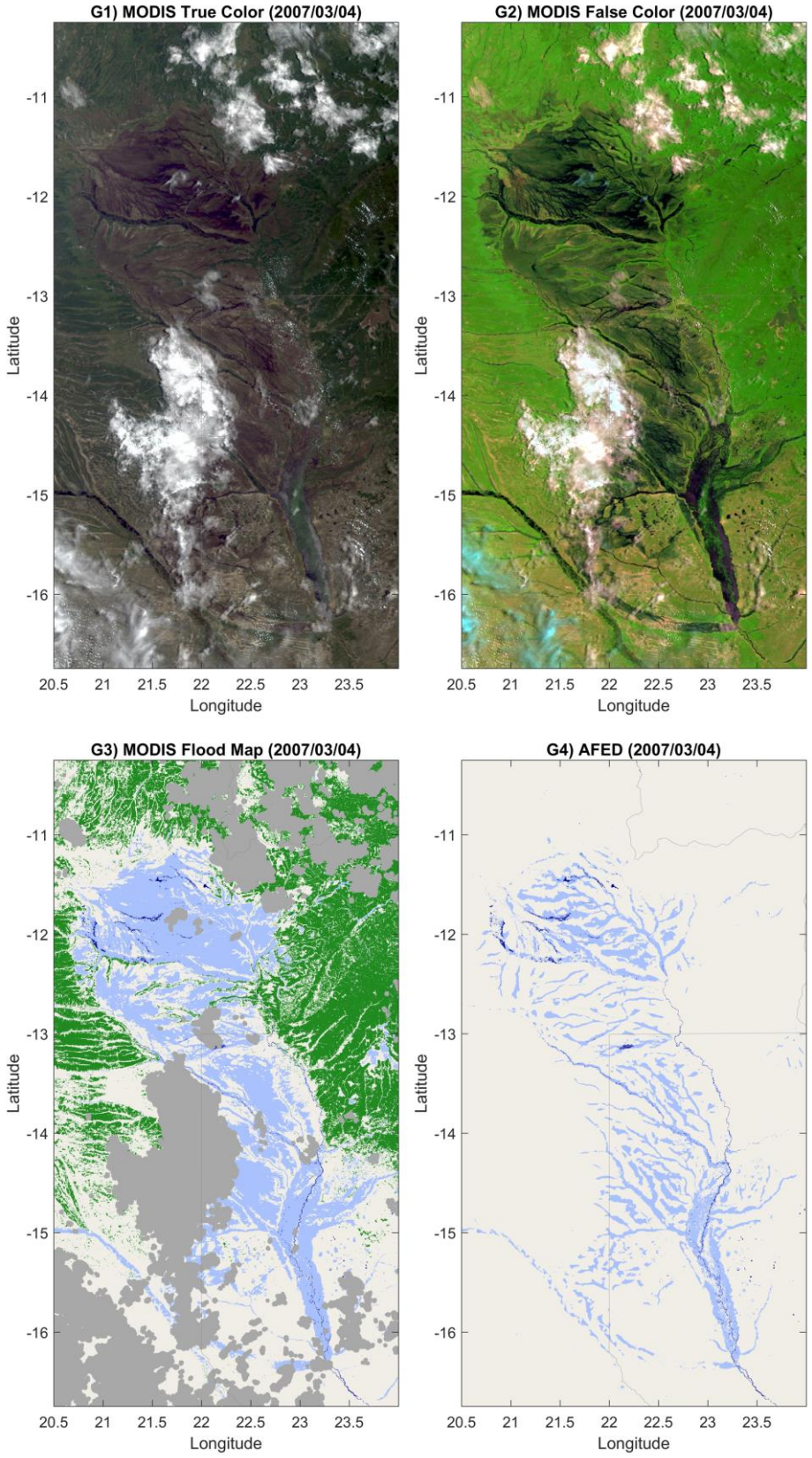


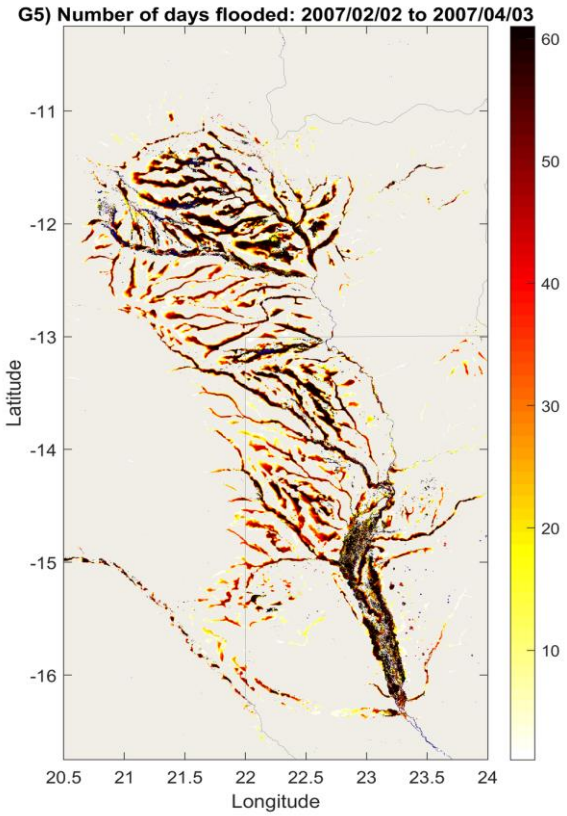
First row: G1) MODIS true color image and G2) MODIS false color image. Second row: G3) MODIS-derived flood map and G4) AFED flood map; light blue is flood water, dark blue is persistent open water, beige is dry land, and green is tree cover. Third row: G5) AFED cumulative number of days flooded for +/-30 days around the test scene date. Fourth row: G6) footprint-weighted flooded fraction derived from MODIS flood map and G7) microwave flooded fraction before applying MDFF. Fifth row: G8) microwave flooded fraction after applying MDFF and G9) difference between footprint-weighted AFED flooded fraction and microwave flooded fraction. Sixth row: G10) microwave-derived flooded fraction vs. MODIS-derived flooded fraction.

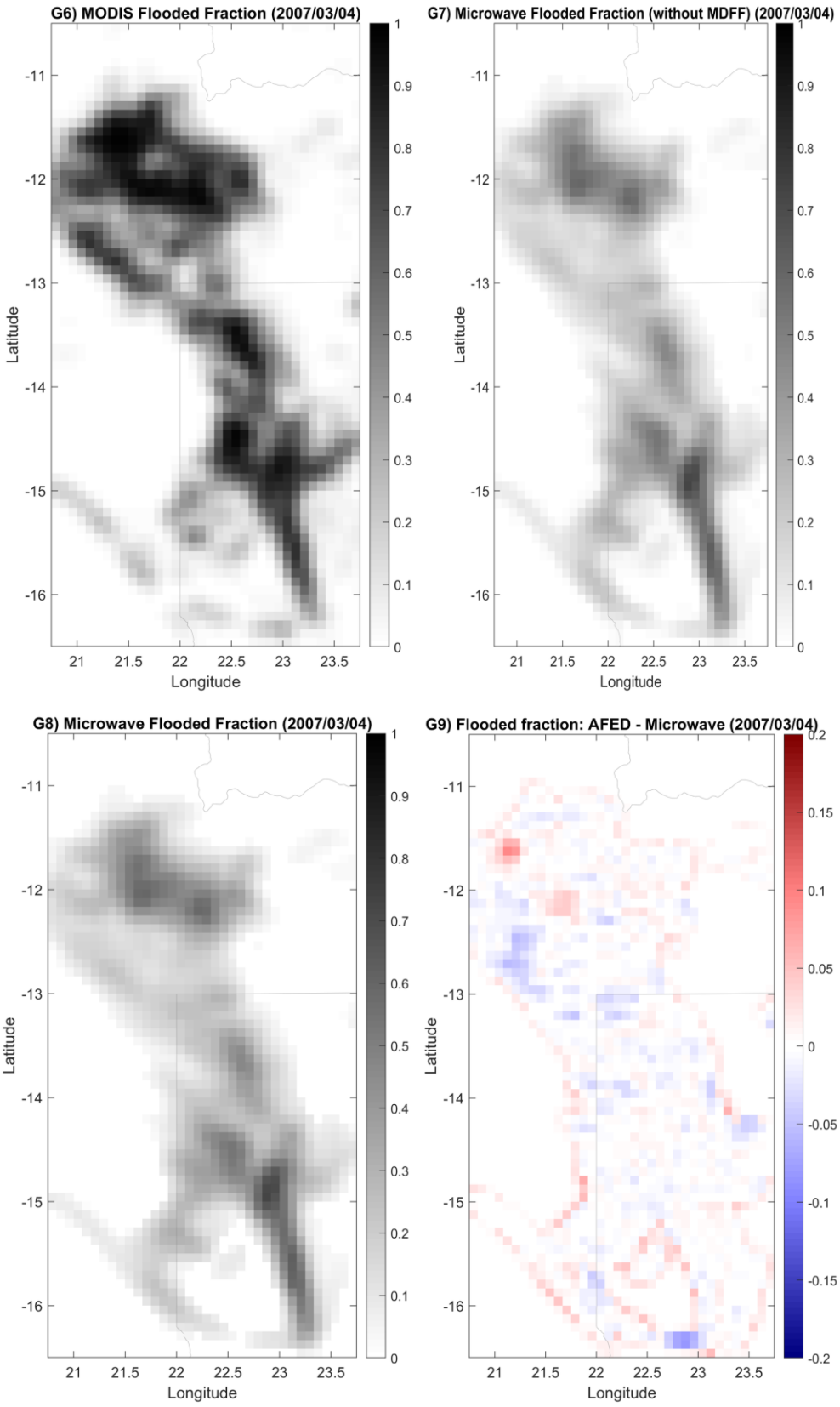
Barotse2003 Discussion:

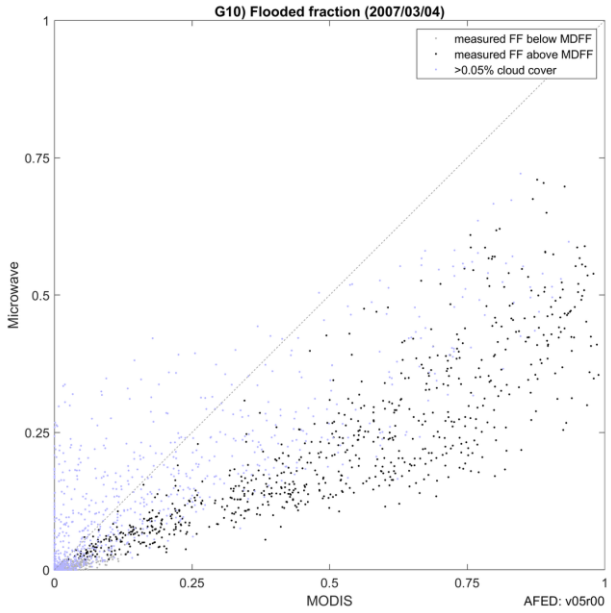
This case is discussed along with the Barotse2007 case below.

Case: Barotse2007









First row: G1) MODIS true color image and G2) MODIS false color image. Second row: G3) MODIS-derived flood map and G4) AFED flood map; light blue is flood water, dark blue is persistent open water, beige is dry land, and green is tree cover. Third row: G5) AFED cumulative number of days flooded for +/-30 days around the test scene date. Fourth row: G6) footprint-weighted flooded fraction derived from MODIS flood map and G7) microwave flooded fraction before applying MDFF. Fifth row: G8) microwave flooded fraction after applying MDFF and G9) difference between footprint-weighted AFED flooded fraction and microwave flooded fraction. Sixth row: G10) microwave-derived flooded fraction vs. MODIS-derived flooded fraction.

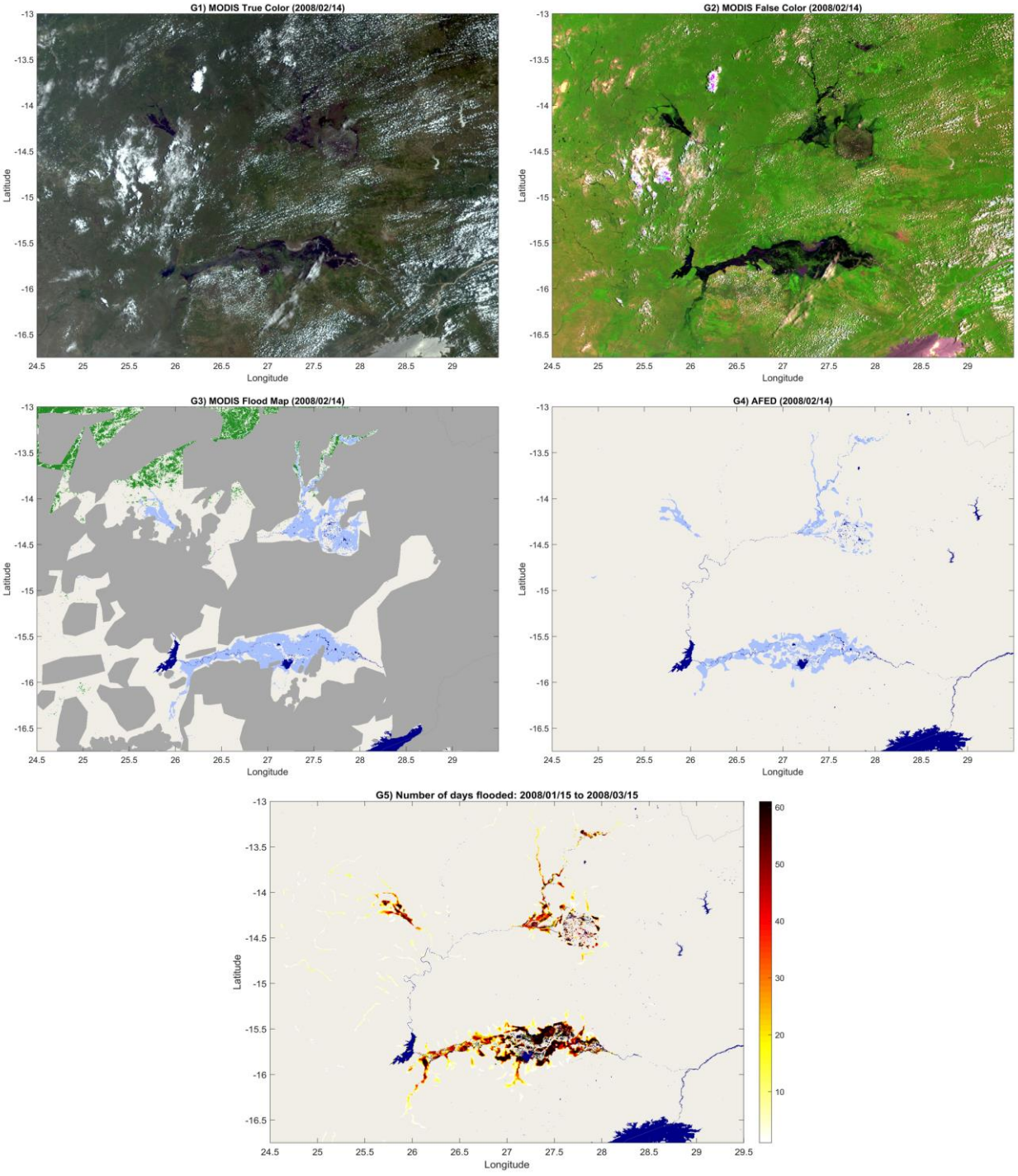
Barotse2007 Discussion:

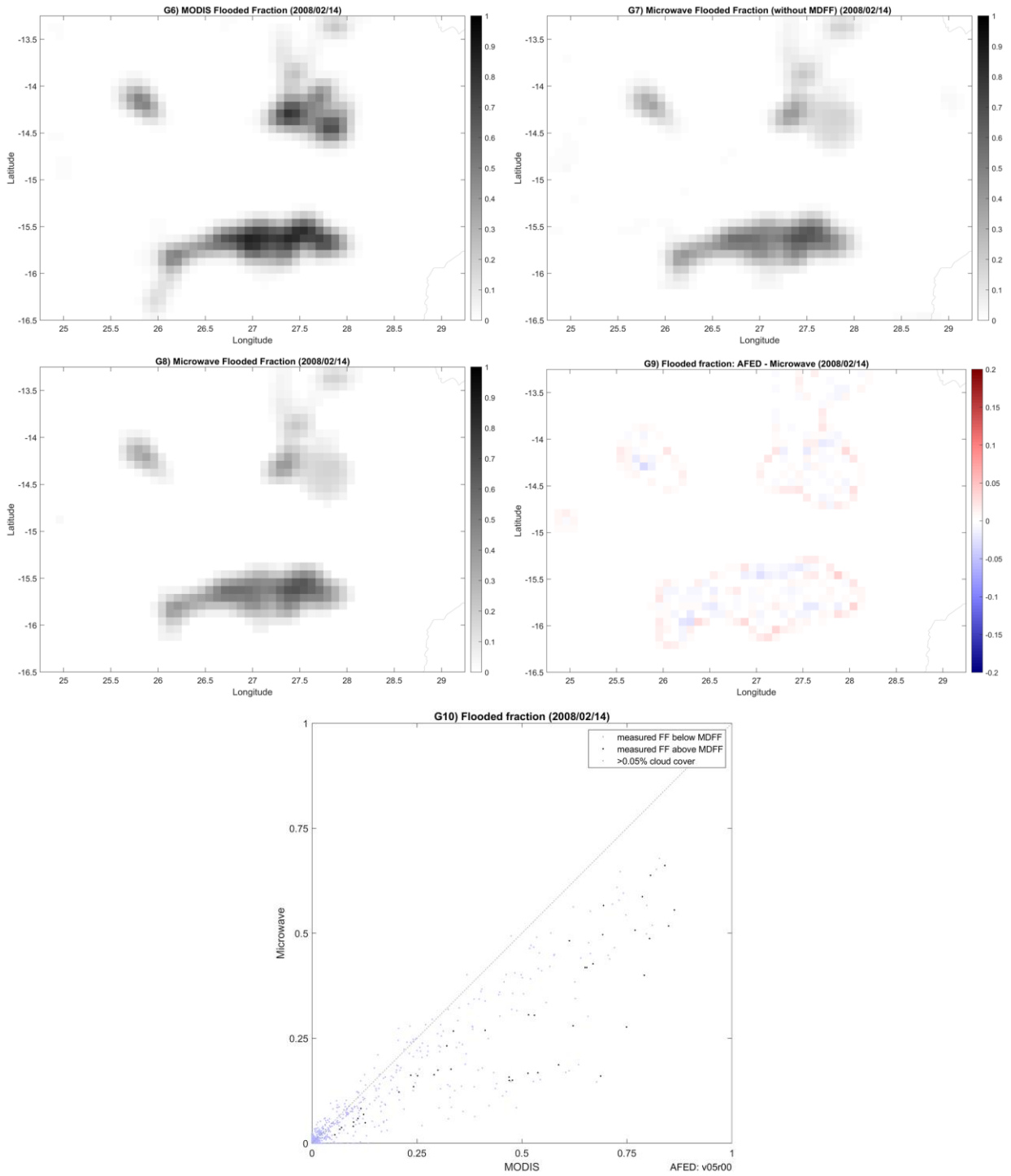
This case and Barotse2003 cover the large seasonal wetlands area straddling the Zambia-Angola border. Overall correspondence of the AFED product to the validation data is good in both cases although there is a significant negative bias trend in flooded fraction, some of which is likely attributable to overestimation of flooding in MODIS image analysis (discussed in the summary).

AFED depiction in this area benefits significantly from the iterative downscaling approach. The iterative algorithm minimizes the differences between upscaled AFED results and the microwave-derived flooded fraction data going into the downscaling algorithm, resulting in the concurrence seen in the flooded fraction graphics (G6-G9).

As of V04R00, the addition of GSW to the relative floodability database has improved AFED correspondence to the MODIS image in various areas such as the Barotse floodplain (the narrower area in the southwest corner) and the Cuando River floodplain (running northwest to southeast in the southwest corner).

Case: Kafue2008





First row: G1) MODIS true color image and G2) MODIS false color image. Second row: G3) MODIS-derived flood map and G4) AFED flood map; light blue is flood water, dark blue is persistent open water, beige is dry land, and green is tree cover. Third row: G5) AFED cumulative number of days flooded for +/-30 days around the test scene date. Fourth row: G6) footprint-weighted flooded fraction derived from MODIS flood map and G7) microwave flooded fraction before applying MDFF. Fifth row: G8) microwave flooded fraction after applying MDFF and G9) difference between footprint-weighted AFED flooded fraction and microwave

flooded fraction. Sixth row: G10) microwave-derived flooded fraction vs. MODIS-derived flooded fraction.

Kafue2008 Discussion:

AFED does a good job depicting flooding the Kafue Flats (lower flood region) but does not do quite as well in the northeast region. Overall flooded fraction is underestimated with the largest discrepancies in the northeast region. Nevertheless, the three main regions of flooding are well-depicted—including many smaller features in the northeast—and no large-scale false positives are apparent. Like many other cases, the distribution of floodwater in relatively flat areas more closely matches the MODIS image due to the addition of GSW data to the relative floodability database as of AFED V04R00. This can be seen most clearly in the number-of-days-flooded graphic, where main channels are less apparent (e.g., compared to V03R01) due to the displacement of floodwater into surrounding areas.

3.4 Flood event test cases: summary discussion

- We can conclude several things from the set of cases in which AFED closely agrees with the validation data (e.g., Caprivi2003, Caprivi2009a, Caprivi2009b, Kafue2008). First, the microwave data is generally accurate with respect to geolocation and measurement consistency spatially and over time. Second, the flooded fraction algorithm can be very accurate in some conditions. Third, the downscaling algorithm can accurately represent flooded areas at a much higher resolution than the flooded fraction inputs. These results build confidence that AFED algorithm concept is fundamentally sound and that errors in other cases are the result of the particular circumstances of the event, the surrounding land cover conditions, or other factors not always present.
- As of V05R00, no cases have the distinct striping pattern characteristic of the SRTM V3 elevation model. Otherwise, the use of the MERIT elevation data in V05R00 has improved flood map details and realism without fundamentally changing the aggregate flood depiction.
- In several cases there is a negative bias trend in the FF scatter plot (G10, e.g., InnerNigerDelta2007, Kyoga2007, InnerNigerDelta2009, Pungue2008, Barotse2007, Caprivi2003). This bias trend may be attributable to either (a) positive bias in our derivation of the validation scene flood map or (b) local variation in the flooded land end-member in microwave flooded fraction algorithm. We know from other test cases (e.g., Caprivi2009b, Etosha2009b, Caprivi2009a, Caprivi2003) that the FF bias trend is not always present. Therefore, it may be the result of local conditions such as the amount, type, or distribution of vegetation, which can have effects on both the microwave and MODIS scene observations. The microwave flooded fraction algorithm assumes that open water is a good representation of flooded-land conditions. However, if floodwater is partially obscured by vegetation then the actual flooded-land end-member will be lower than the open water value, leading to negative FF biases like those seen in these cases. The AFED algorithm could be changed to account for vegetation amount but it would be difficult to assess the impact (e.g., on overall false positive rates) because of the lack of accurate flood event validation and vegetation data for the full range of land cover conditions. For example, in some cases (notably Barotse2007 and Barotse2003) vegetation may be having a significant impact on our interpretation of MODIS scenes, leading to an over-estimation of flood extent in the validation image. Nevertheless, a conservative adjustment to the flooded-land end-member that accounts for vegetation amount may be justified by these results and the physical argument that some floodwater is likely to be obscured by vegetation when it is present.
- In some cases, MDFF trimming on the margins of the flooded area leads to relatively large differences between the MDFF-trimmed FF and upscaled AFED results (graphic G9, e.g., InnerNigerDelta2007, Etosha2009b, Etosha2009a, InnerNigerDelta2009). These are cases where MDFF is relatively large which means that the FF discontinuity at the location of the trim is likely to be relatively large as well. This discontinuity is a natural consequence of the use of MDFF to eliminated false positives due to random noise in AFED FF data. However, positive differences in graphic G9 on the flooded area margins in many of the test cases indicate that the downscaling algorithm is depicting the flooded area in way that is consistent with the

footprints of the non-zero FF values. As a result, we do not believe that these differences indicate a problem with algorithm performance.

3.5 Case-by-case false negative and false positive test results for AFED V04R01

3.5.1 *False negative cases*

3.5.1.1 *Ouagadougou, Burkina Faso, Sept. 2009*

- *Event summary from ARC and other sources*: Unprecedented rainfall of 1 September 2009 in Ouagadougou — when 263 mm fell within 12 hours — left more than 150,000 people affected within and around Ouagadougou.
- *Event information sources*:
  - http://www.irinnews.org/news/2009/09/04-0
  - http://earthobservatory.nasa.gov/NaturalHazards/view.php?id=40127
  - http://reliefweb.int/map/burkina-faso/burkina-faso-flooding-ouagadougou-03-sep-2009
  - http://www.sim.org/index.php/content/flood-relief-for-burikna-faso
  - http://www.crs.org/burkina-faso/ouagadougou-floods/
  - http://reliefweb.int/map/burkina-faso/overview-flood-affected-sectors-ouagadougou-burkina-faso-09-sep-2009
  - http://www.unoosa.org/pdf/pres/stsc2010/tech-54.pdf
  - http://www.dartmouth.edu/~floods/2007159.html
- *Location map indicating microwave grid points selected for time series analyses*: Results for a subset of these grid points are given below.

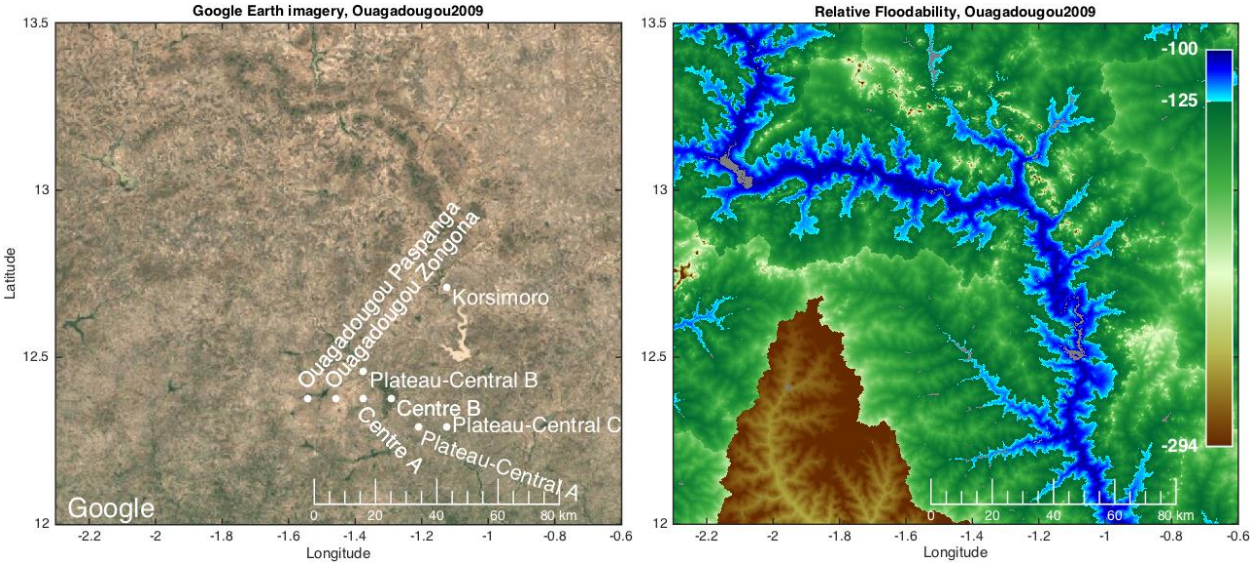
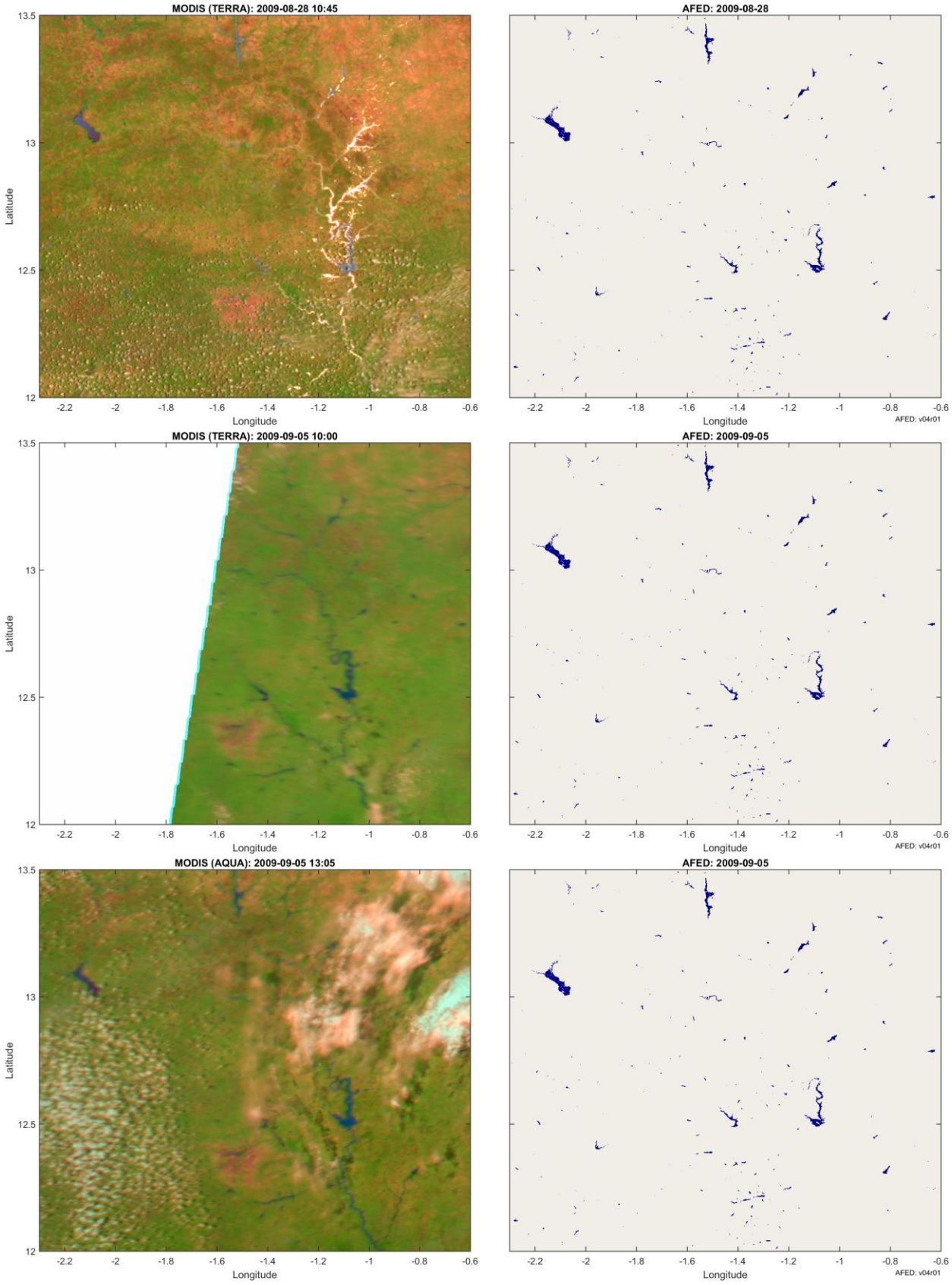
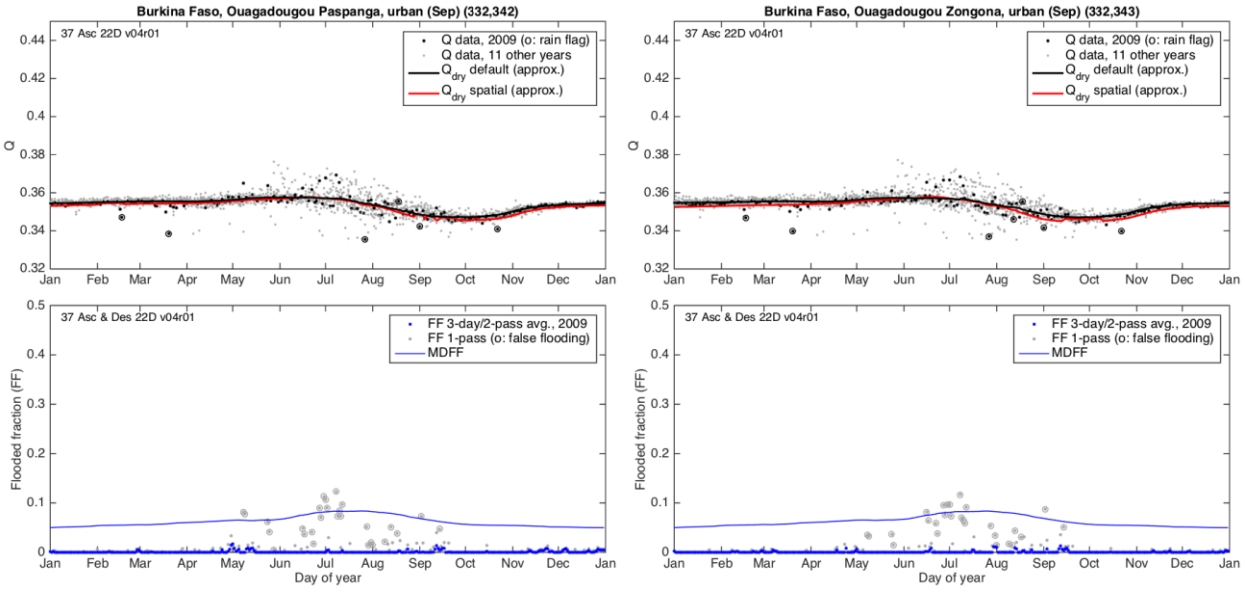


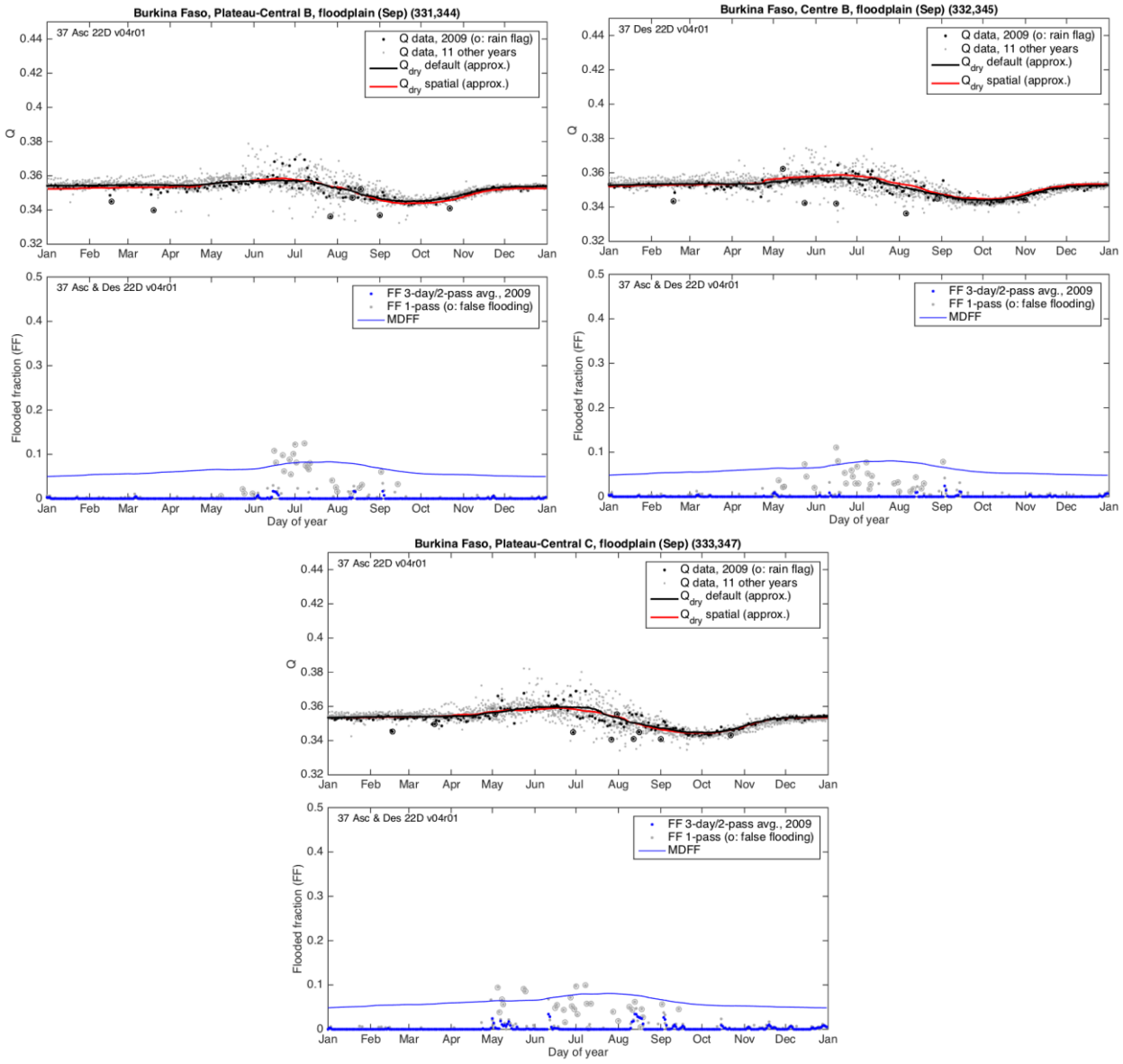
Figure 16: Google imagery with microwave data grid points (left) and relative floodability (RF) with blue tones indicating larger river flood plain areas within 25 m of the maximum RF. Distance scale represents approximate east-west distance at the middle latitude.

- *Side-by-side comparisons of MODIS false-color imagery to AFED*: Images were selected based on proportion of cloud free area and the degree to which the image can be interpreted to infer the timing and location of flooded or non-flooded conditions.



- *Time series analyses*: There are two panels (top and bottom) for each grid point and calendar year indicated. The top panel shows the microwave data, *Q*, for the indicated sensor band (37 GHz), satellite pass (ascending or descending), and reference footprint shape (22D or 22 km); black and gray dots indicate *Q* data for the current year and other years in the historical record, respectively. Circled dots indicate data excluded by the AFED algorithm due to rain detection. The lines are approximate values of the algorithm's seasonally-varying *Q* dry-land end-member, Q_{dry}: the black line is the default Q_{dry} based on time-series analyses and the red line is Q_{dry} from spatial analyses, both of which are described in the ADD. *Q* typically increases with flooding but is also affected by seasonal land cover change, soil moisture, rain, and other factors. The algorithm derives the single-pass flooded fraction according to the amount that *Q* data exceeds Q_{dry}. The bottom panel shows the microwave flooded fraction results from the AFED algorithm. Gray dots indicate single-pass flooded fraction results for both the ascending and descending passes; circled dots indicate data excluded by the final flooded fraction algorithm due to false positive detection. Blue dots indicate the flooded fraction estimate derived by averaging all available single-pass results over a 3-day period. Finally, the solid blue line is the local seasonally-varying MDFF; the AFED algorithm sets flooded fraction values below MDFF to zero to reduce false positives due to factors other than flooding that affect *Q* day-to-day.





Discussion:

The 2009-08-28 MODIS image shows pre-event conditions, with sun glint reflecting off water in the river systems to the east of Ouagadougou and but otherwise dry conditions. (Note that sun glint makes rivers appear wider than they would otherwise because sensor pixels can be affected by stray light from neighbors.) Heavy cloud cover in the MODIS imagery on 2009-09-01 (not shown) corresponds to the timing of the downpour in Ouagadougou. The first day afterward with relatively clear views of the city 2009-09-05 (two images). The rivers and reservoirs in the eastern half of the image are clearly delineated. While Ouagadougou itself is darker than on 2009-08-28, the area lacks the distinct dark lines or patches that distinguish flooding in the other cases. The rivers to the southeast of Ouagadougou are more distinct but the flooded area is too small to lead to flooded fraction values greater than MDFF in these areas.

The time series plots for Ouagadougou (Paspanga and Zongona areas) show that rain was flagged on 2009-09-01 (*Q* plots) and that a false positive "flash" was flagged on 2009-09-02 (FF

plots). Otherwise, there is no signal of the flood event in the urban area. This could be due in part to two factors. First, Q_{dry} in the urban area and the surrounding terrain is relatively high compared to more vegetated areas (e.g., in Malawi, section 3.6.1), which reduces the differential effect of flood water on Q. Second, the urban environment includes structures that reduce the overall visible surface area of flood water on the ground. This effect is exacerbated by the approximately 53° viewing angle of microwave sensors, which means that sides and tops of buildings are seen, further reducing the net contribution of the surface. However, even in the absence of structures it is not clear that the flooded area was extensive enough for long enough to be detectable above the MDFF. MDFF is relatively high in this area due to Q variability day-to-day and year-to-year (represented in the Q plots by aggregate data from 11 other years).

Three time series plots covering areas just to the east of Ouagadougou proper are also shown: Plateau-Central B, Centre B, and Plateau-Central C. These points are in areas where relative floodability is higher than it is in Ouagadougou. Plateau-Central B is similar to the more urban points, with little to no signal of flooding beyond the flash detected on 2009-09-02. The other two points show a small flood event in the days following 2009-09-02 but the flooded fraction remains below MDFF. These points show many similar periods of elevated FF during the year that are not known to be associated with reported flood events. As a result, any change to the AFED algorithm to be more sensitive to the reported 2009-09-01 event would be likely to increase the number of false positives at other times.

Figure 17 illustrates how flood plain size can affect the flooded fraction and the ability of the AFED algorithm to detect flooding. The top graphic is a simulated depiction of flooded fraction at 22-km resolution when all rivers of stream order greater than four are flooded to a relative floodability depth of 10-m, a reasonable but arbitrary limit selected to illustrate the effect of flood plain width. Flooded fraction at 10-m relative floodability depth goes up to 1 in some areas, many of which are wetlands (i.e., Barotse Flood Plain, Niger Inner Delta, Congo Flood Plain, etc.). The bottom graphic shows where the simulated flooded fraction at 10-m depth exceeds MDFF; this is one way to represent the flood plains where flooding is most detectable from 22-km microwave data. From this graphic we see that there are a few wide flood plains in Burkina Faso; for example, the White Volta river flood plain runs north-south through the center of the country and can be seen as the major flood plain area in Figure 16. Most areas of the country—including Ouagadougou—are outside the biggest flood plains.

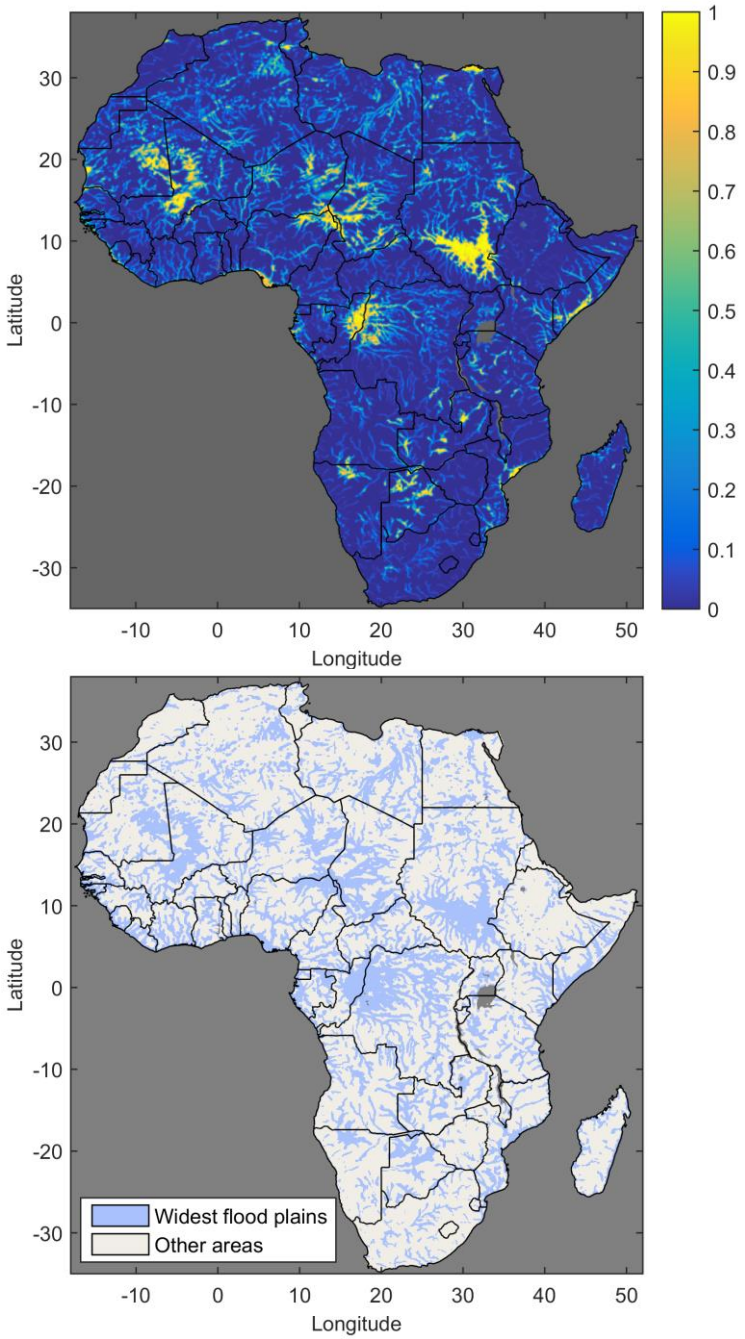


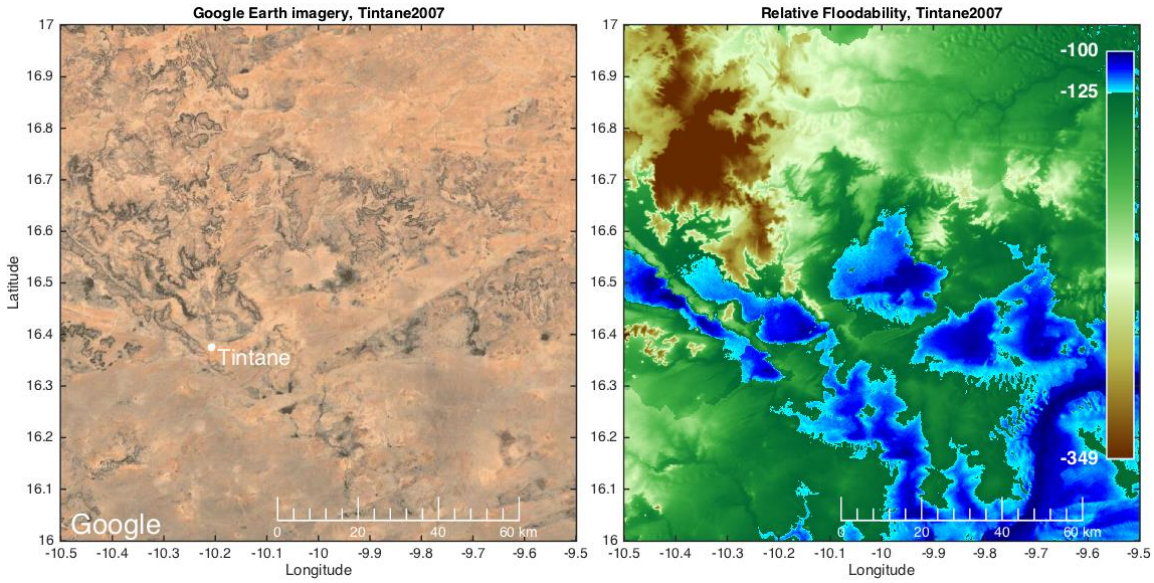
Figure 17: Top: Simulated flooded fraction with flooding up to a10-m relative floodability threshold depth in rivers of stream order greater than four. Bottom: Map of wide flood plains where the simulated flooded fraction at 10-m relative floodability depth exceeds the local MDFF.

3.5.1.2 *Tintane, Maritania, August 2007*

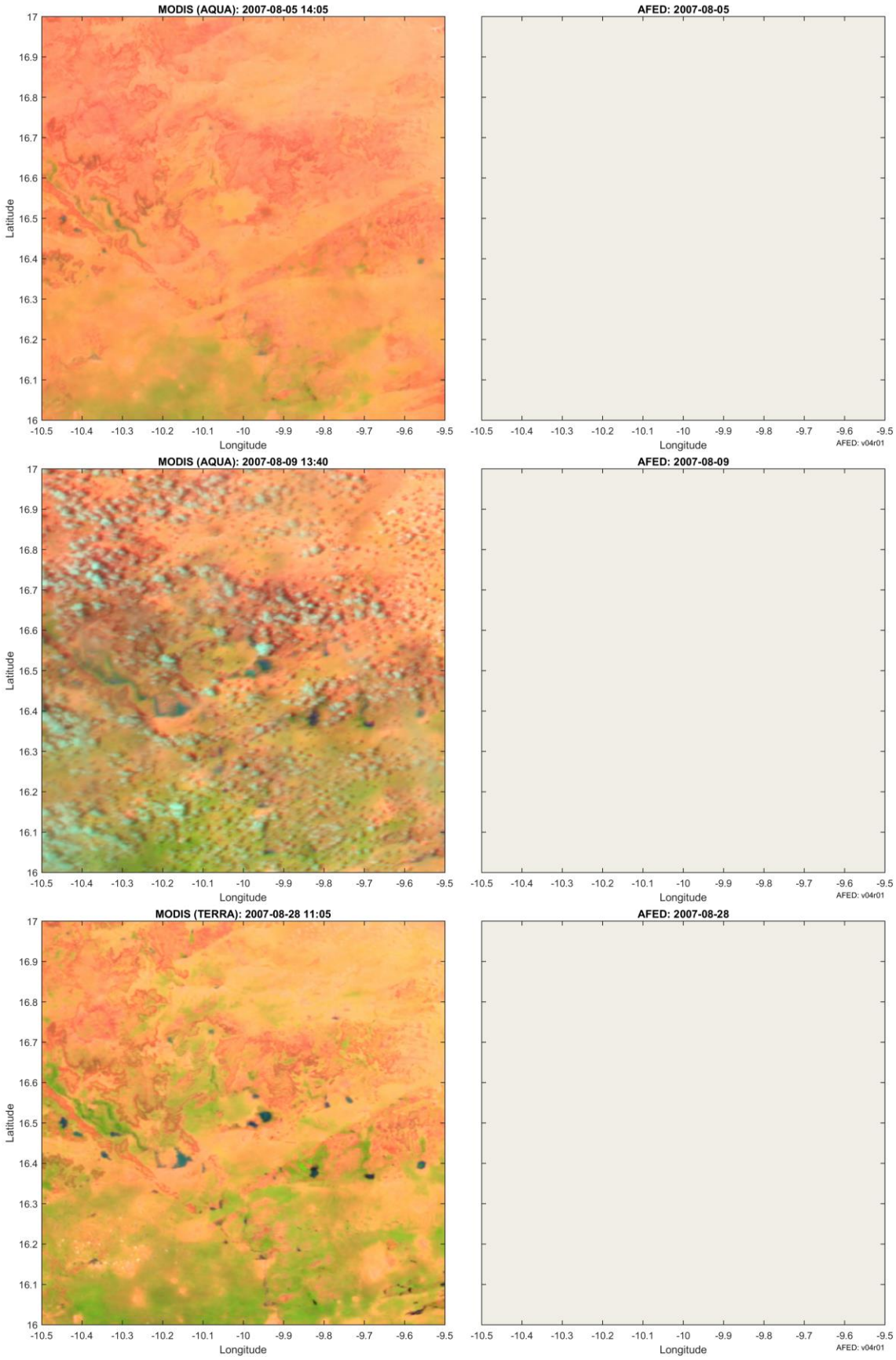
- *Event summary from ARC*: After a long and difficult dry season, the inhabitants of Tintane waited expectantly for the arrival of the rain season. But they did not anticipate the flash floods that engulfed the town in August 2007. Human lives were lost and homes, livestock and crops were destroyed, as heavy rainfall and intensive

flooding battered their town for days on end. As water levels rose, so did the number of families rendered homeless by the tide. Villagers braved the hazardous conditions in makeshift boats to rescue relatives and personal belongings. Families retreating from the floodwaters huddled in tents haphazardly pitched on higher ground. WFP staff from Aioun (Algeria) were on the scene without delay to assist the authorities in evacuating victims, storing and distributing relief supplies, and collecting information.

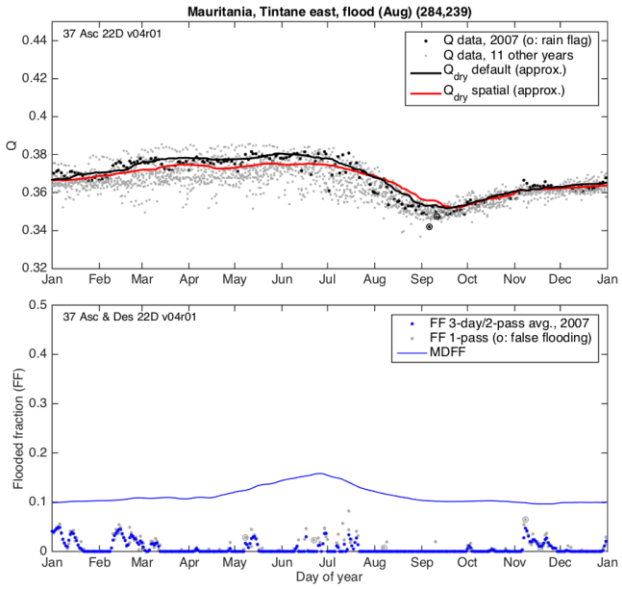
- *Event information sources*:
  - *http://home.wfp.org/sprs/donor_reports/reports/store/07/10359.0.pdf*
  - www.dartmouth.edu/~floods/Archives/2007sum.htm.webloc: Reports flood dates from 2007-08-07 to 2007-09-05.
- *Location map indicating microwave grid points selected for time series analyses*: Results for a subset of these grid points are given below. See Figure 16 for figure caption.



- *Side-by-side comparisons of MODIS false-color imagery to AFED*: Images were selected based on proportion of cloud free area and the degree to which the image can be interpreted to infer the timing and location of flooded or non-flooded conditions.



- *Time series analyses*: Plots are described in section 3.5.1.1.



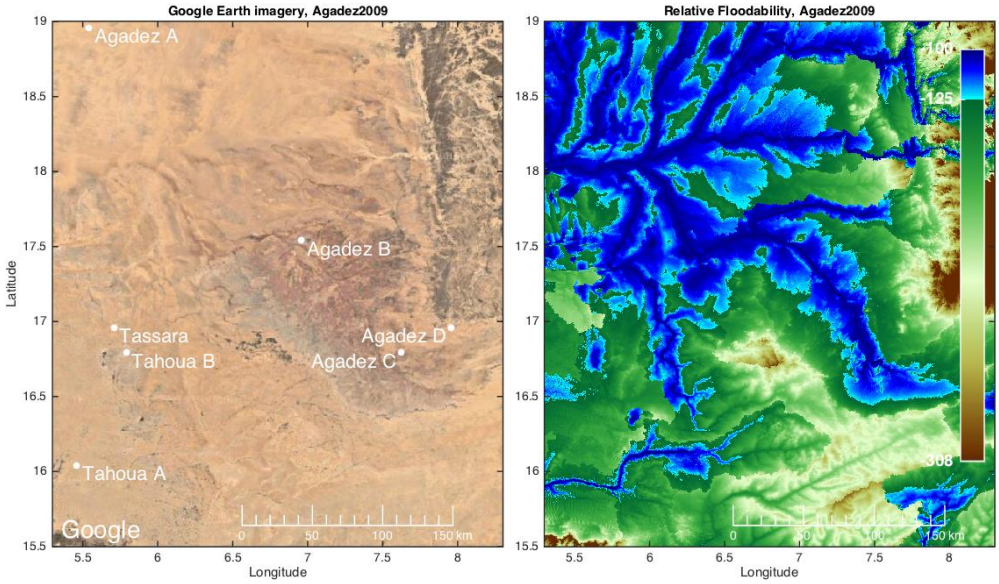
Discussion:

The MODIS image sequence shows apparently dry conditions over the region on 2007-08-05. The images on 2007-08-07 (not shown) show heavy cloud cover corresponding to the timing of the reported rainstorm. The next clear-sky image on 2007-08-09 shows a large area of local flooding in the northeast of the Tintane area as well as several other flood pools in the surrounding region. These pools can more clearly be seen in the image on 2007-08-28. (The dark band running away from Tintane to the northwest is not a flood plain; it is at a higher elevation than Tintane, which can be seen in the relative floodability map.) We used Google Earth to estimate that the size of the apparent flooded area in Tintane on 2007-08-28 is about 12 km^2. This represent about 0.02 flooded fraction at 22-km resolution, which is well below the MDFF of 0.12 in this area at this time of year as well as being below the median MDFF of 0.05 across all of Africa. AFED shows no flooding in this region for any day in August 2009.

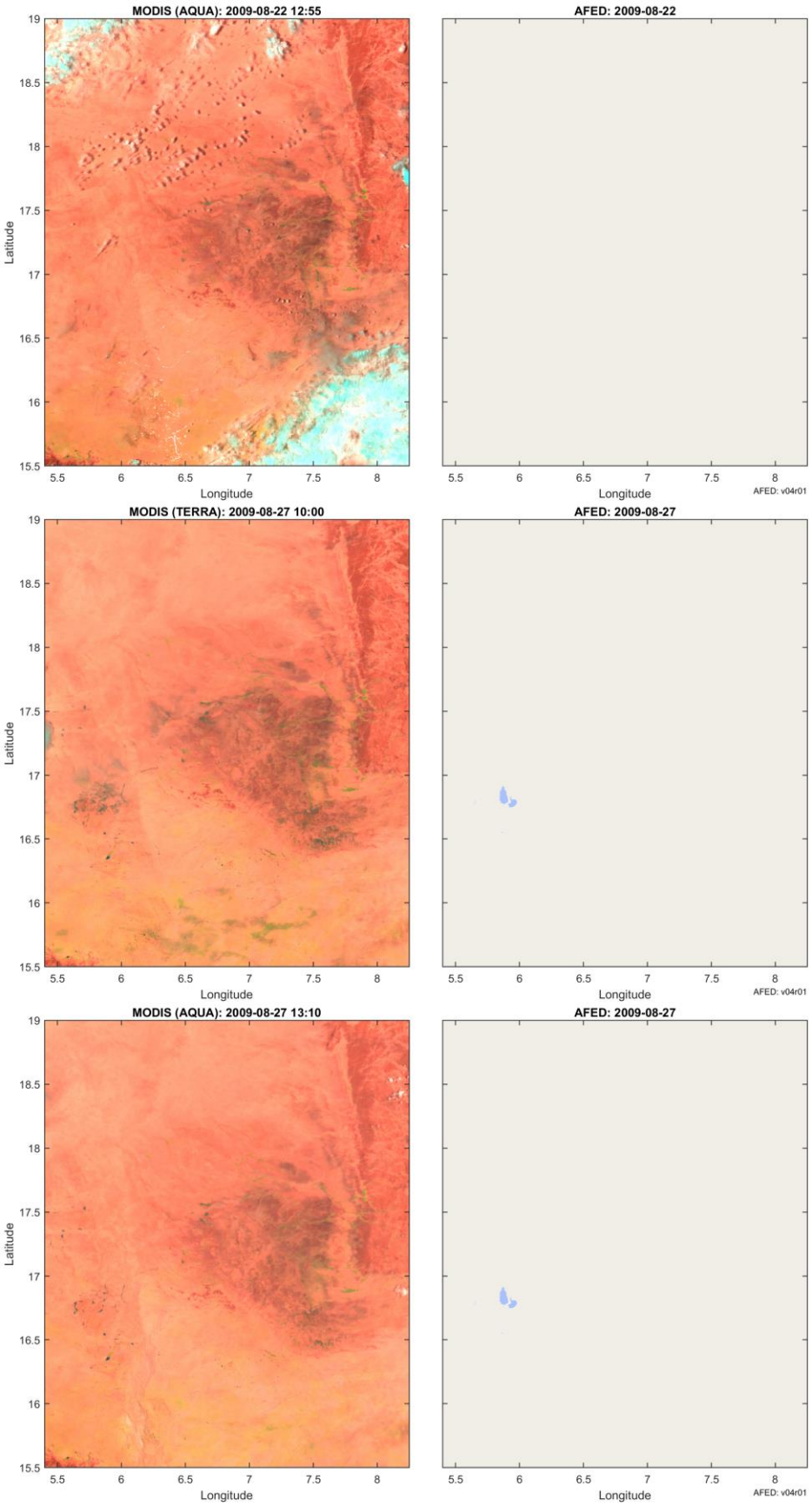
The time series plot shows a flash event detected on 2007-08-08 (FF plot) one day after the day of the reported downfall. Otherwise, the AFED algorithm detects very little flooded fraction signal. The *Q* plot shows why MDFF is relatively high in this region: there is large year-to-year and day-to-day variability in *Q* in the months leading up to the wet season. Following the start of the wet season, *Q* decreases dramatically in response to vegetation green-up from July to September. The August 2007 flood occurred during this transitional period. If Q_{dry} had been lower during the flood the algorithm may have shown a more significant flooded fraction response; however, if the true flooded fraction was in fact below MDFF the flooding would not have appeared in AFED even if the algorithm were able to estimate Q_{dry} perfectly.

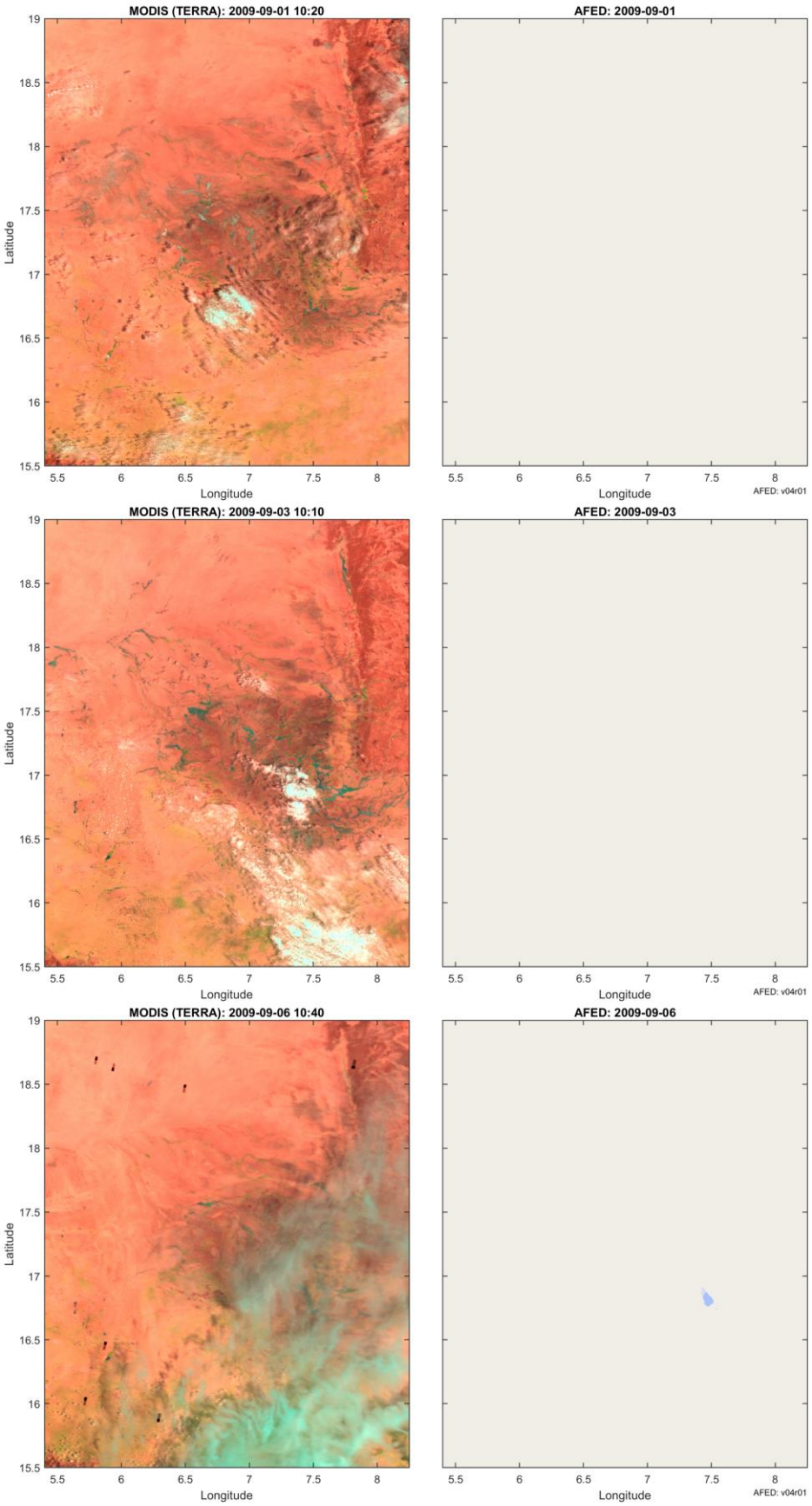
3.5.1.3 *Agadez, Niger, Sept. 2009*

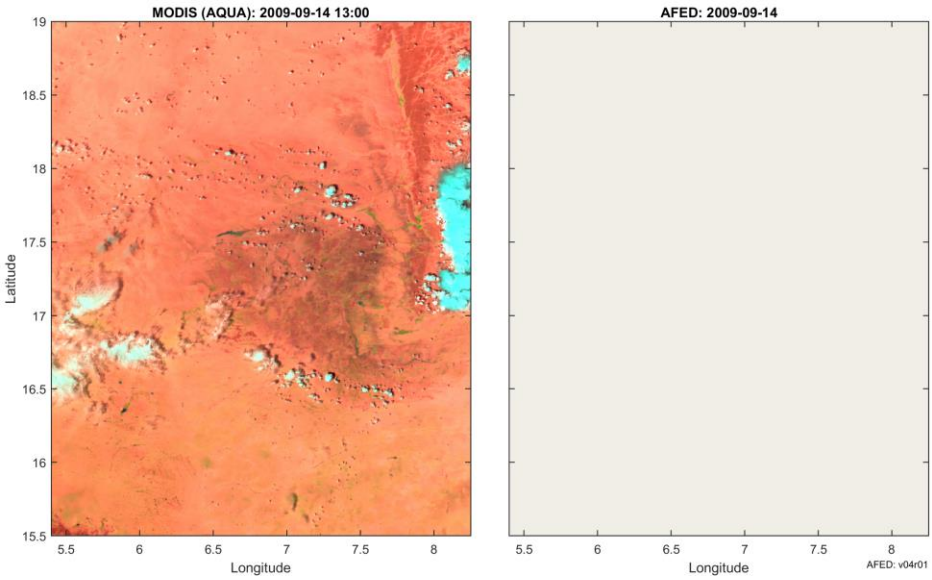
- *Event summary from ARC*: **"**In September 2009 torrential rains caused flooding in valleys and low-lying areas of the Department of Agadez. This caused the death of several persons, damaged and destroyed infrastructure, swept away crop areas, food stocks and cattle. In response to this crisis, WFP launched an emergency operation (EMOP) aiming at ensuring adequate food access and consumption for floods victims in Agadez region." [ARC Ref. 1] Response continued into 2010, the following areas were affected: "Agadez urban area and the rural communes of Dabaga, Tchirozerine, Ingall and Tabelot." [ARC Ref. 2]
- *Event information sources*:
  - ARC Ref. 1: *http://home.wfp.org/sprs/donor_reports/reports/store/09/200071.pdf*
  - *ARC Ref. 2:* *http://home.wfp.org/sprs/donor_reports/reports/store/10/200071.pdf*
  - http://floodobservatory.colorado.edu/Archives/MasterListrev.xls: Notes flood period from 2009-09-21 to 2009-10-15.
- *Location map indicating microwave grid points selected for time series analyses*: Results for a subset of these grid points are given below. See Figure 16 for figure caption.



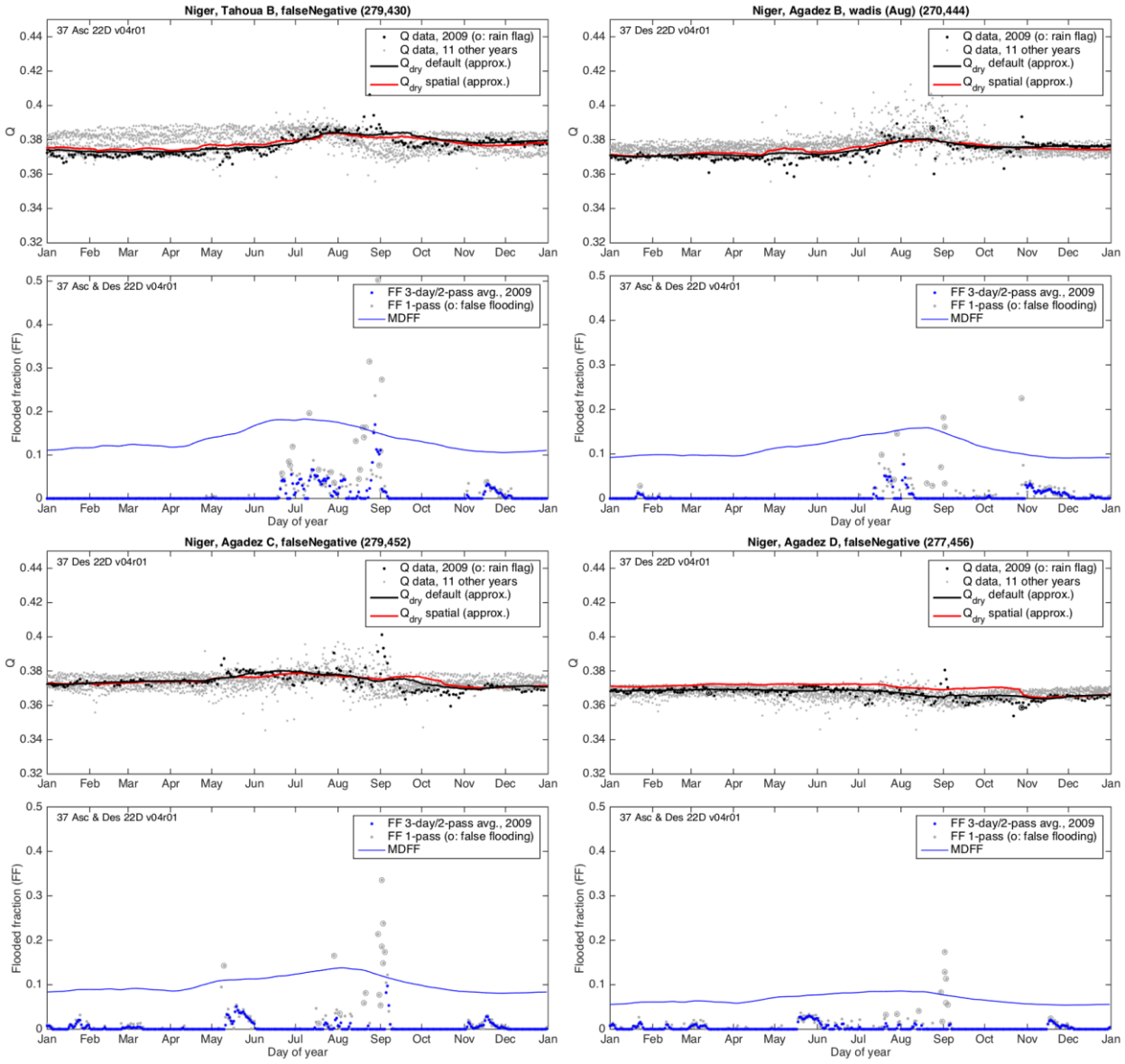
- *Side-by-side comparisons of MODIS false-color imagery to AFED*: Images were selected based on proportion of cloud free area and the degree to which the image can be interpreted to infer the timing and location of flooded or non-flooded conditions.







- *Time series analyses*: Plots are described in section 3.5.1.1.



Discussion:

Beginning from the mostly clear-sky MODIS image on 2009-08-22, the area appears dry overall except for a pattern of small ponds and streams visible only in sun glint in the area 15.5-16° N and 6-6.5°E. The presence of surface water suggests that the rainy season is already well underway. A flooded area with general AFED-MODIS agreement appears on 2009-08-27 around 16.8° N, but the morning and afternoon MODIS images on 08-27 indicate rapidly drying conditions and by 2009-08-30 (not shown) there is no long any indication of flooding in the area. On 2009-09-01 and 03 there are patches of apparent water cover in the MODIS image throughout the region around points Agadez C, B, and D; there is no corresponding AFED flooding in this area. On 2009-09-06 AFED has a patch of flooding at 7.5°E while some flooded

areas are still apparent in the MODIS imagery but shrinking; by 2009-09-14 they are mostly gone.

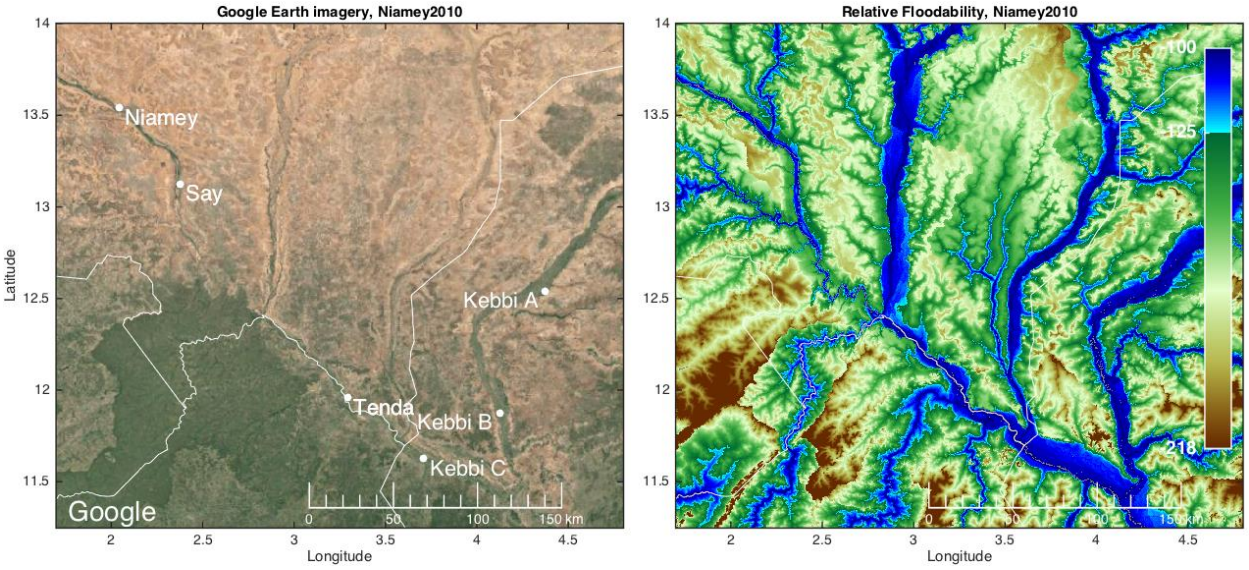
The time series plot for point Tahoua B helps to illustrate the challenges of detecting floodwater in this area. The event that AFED correctly identifies on 2009-08-27 appears in the Tahoua B FF time series as the peak of a longer period of elevated FF beginning around 2009-08-20 and ending 2009-09-10. During and before this period there are many points flagged as false positives (circled points on the FF plot) and the MDFF is at or near its maximum value. The *Q* plot shows that there is large year-to-year variation in dry-season *Q* (i.e., high *Q* noise) and Q_{dry} is very high even compared to the Tintane case (i.e., low available *Q* signal with which to detect water against dry-land background conditions). These factors make water difficult for the microwave algorithm to detect.

The *Q* time series pattern at points Agadez B, C, and D over the period from 08-30 to 09-08 is similar to that of the earlier Tahoua B event but FF never exceeds MDFF at these points. The AFED algorithm flags all the extreme FF values at Agadez B and D as false positives and the few points at Agadez C that are not flagged as false positives are below MDFF.

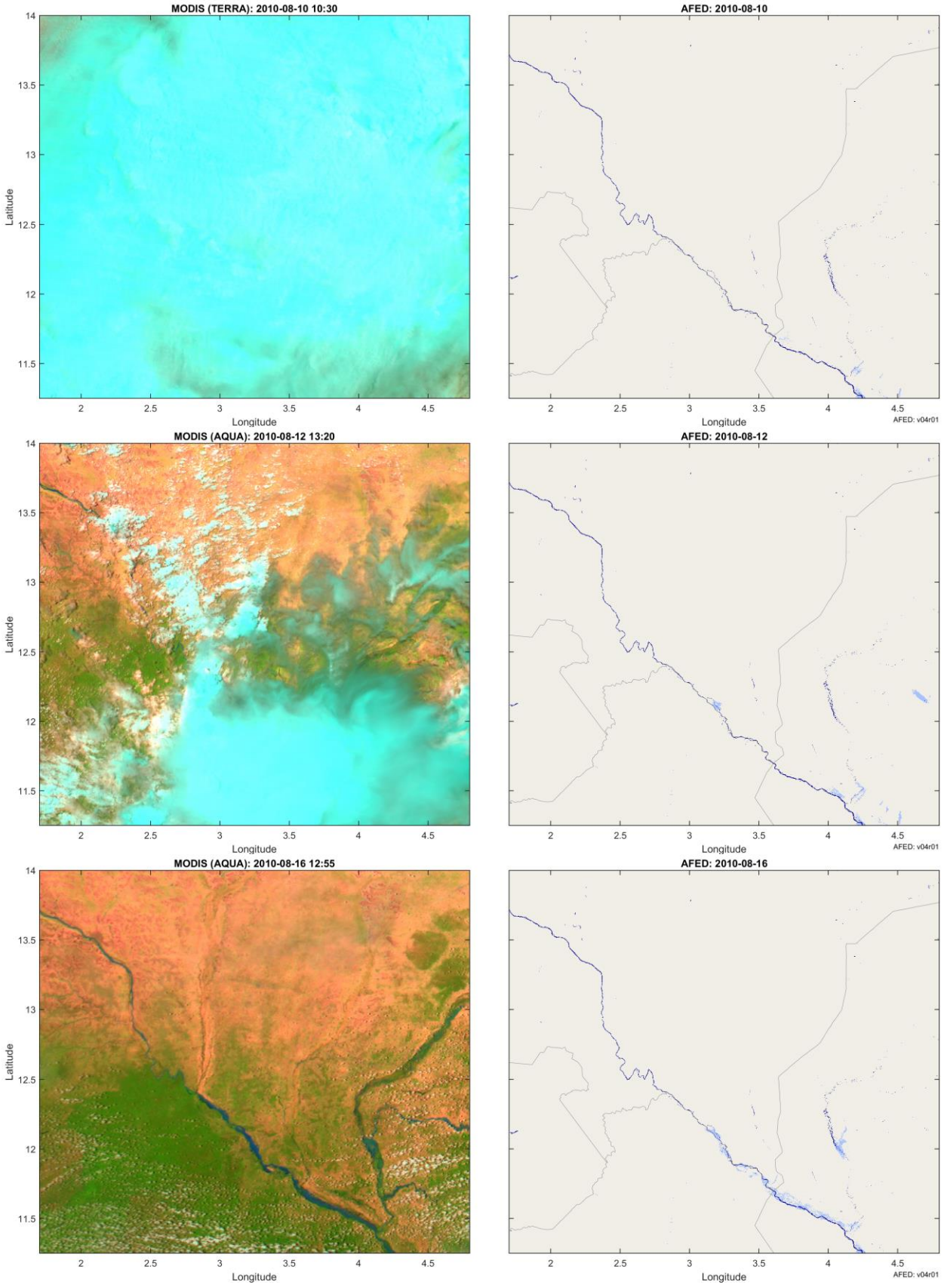
In summary, the inability of the AFED algorithm to detect flooding around Agadez in this period is primarily due to the flagging of false positives. The temporal-spatial signature of the event in the microwave data (i.e., large and brief) causes the AFED algorithm to flag many observations during the event that otherwise would have produced FF values above MDFF. The event is extensive but water cover is distributed in smaller pools over a larger area so no one area has FF above MDFF over an extended period (as seen in the Tahoua B event). Changes to the false positive detection algorithm to avoid flagging during this event would likely lead to an increase in false positives overall.

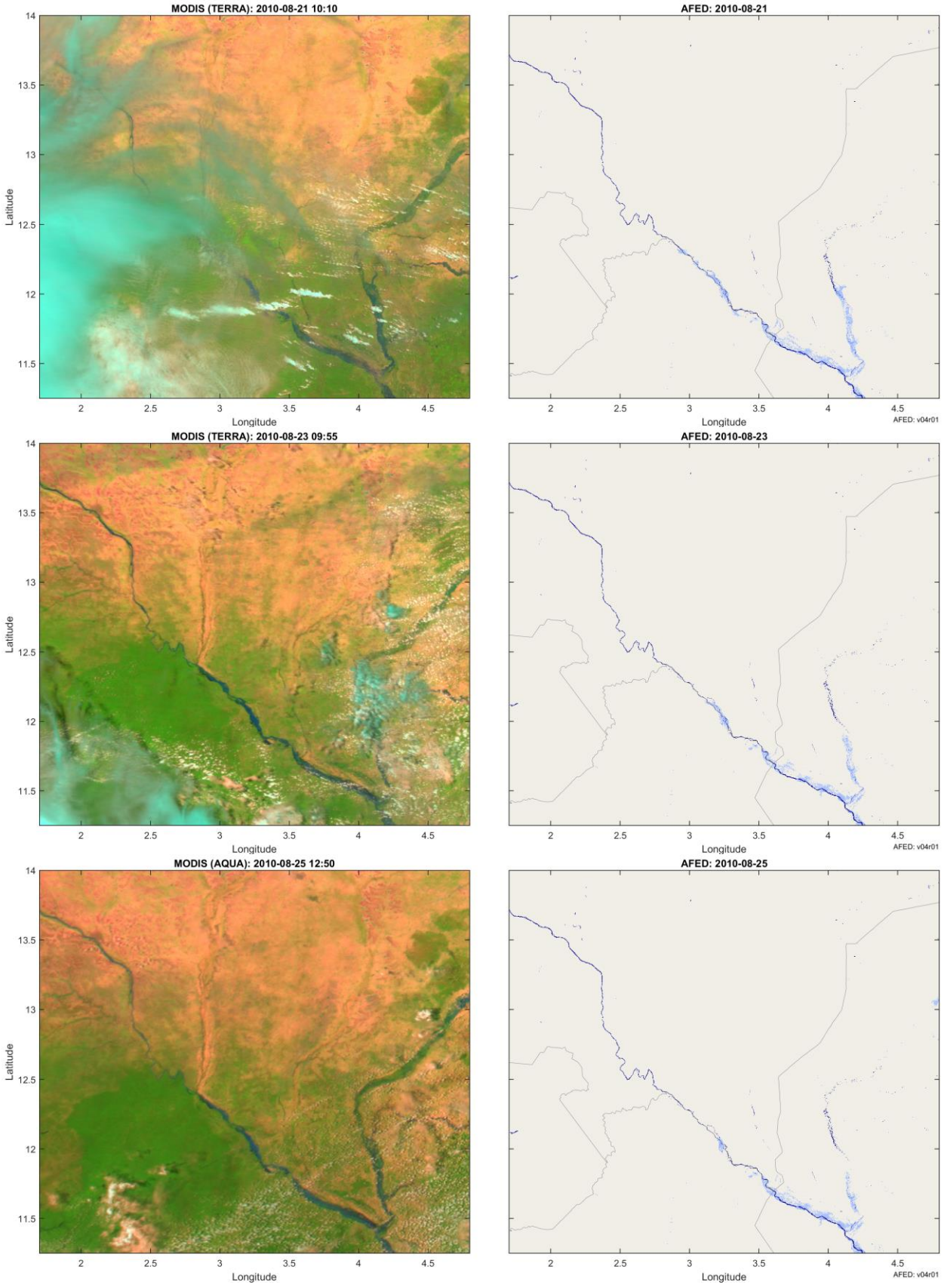
3.5.1.4 *Niamey, Niger, August 2010*

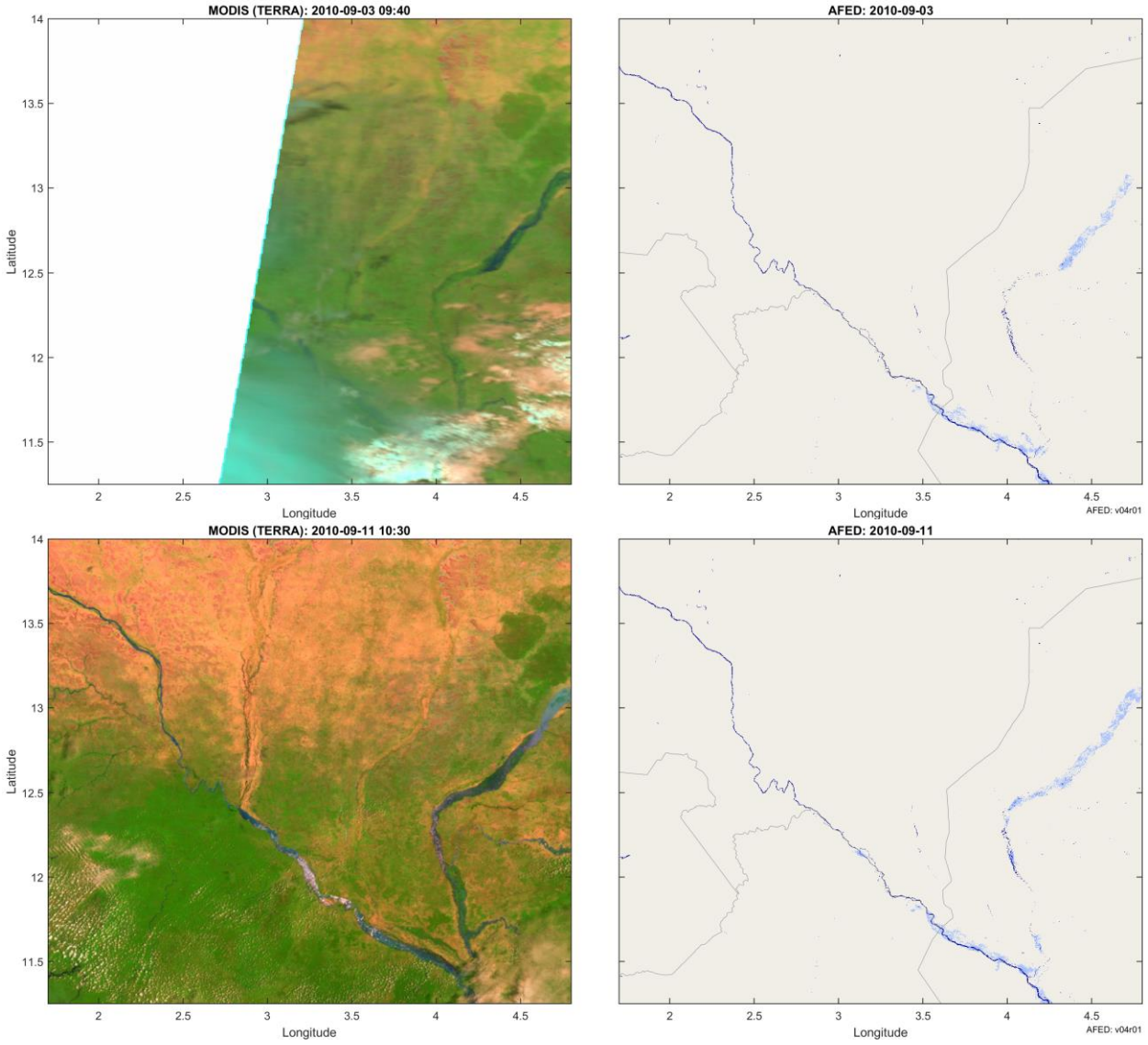
- *Event summary*: The torrential rains in Niger that began in July 2010 caused heavy flooding in Tillaberi, Tahoua, Maradi, Agadez and Zinder regions. This situation was worsened by heavy rains in August in Niamey, the capital city. The number of families in affected regions was estimated at 6,046 families (56,859 people). (http://www.ifrc.org/docs/appeals/10/MDRNE006dfr.pdf)
- *Other event information sources*:
  - http://www.bbc.com/news/world-africa-10929144
  - https://www.dartmouth.edu/~floods/Currentt.htm
- *Location map indicating microwave grid points selected for time series analyses*: Results for a subset of these grid points are given below. See Figure 16 for figure caption.



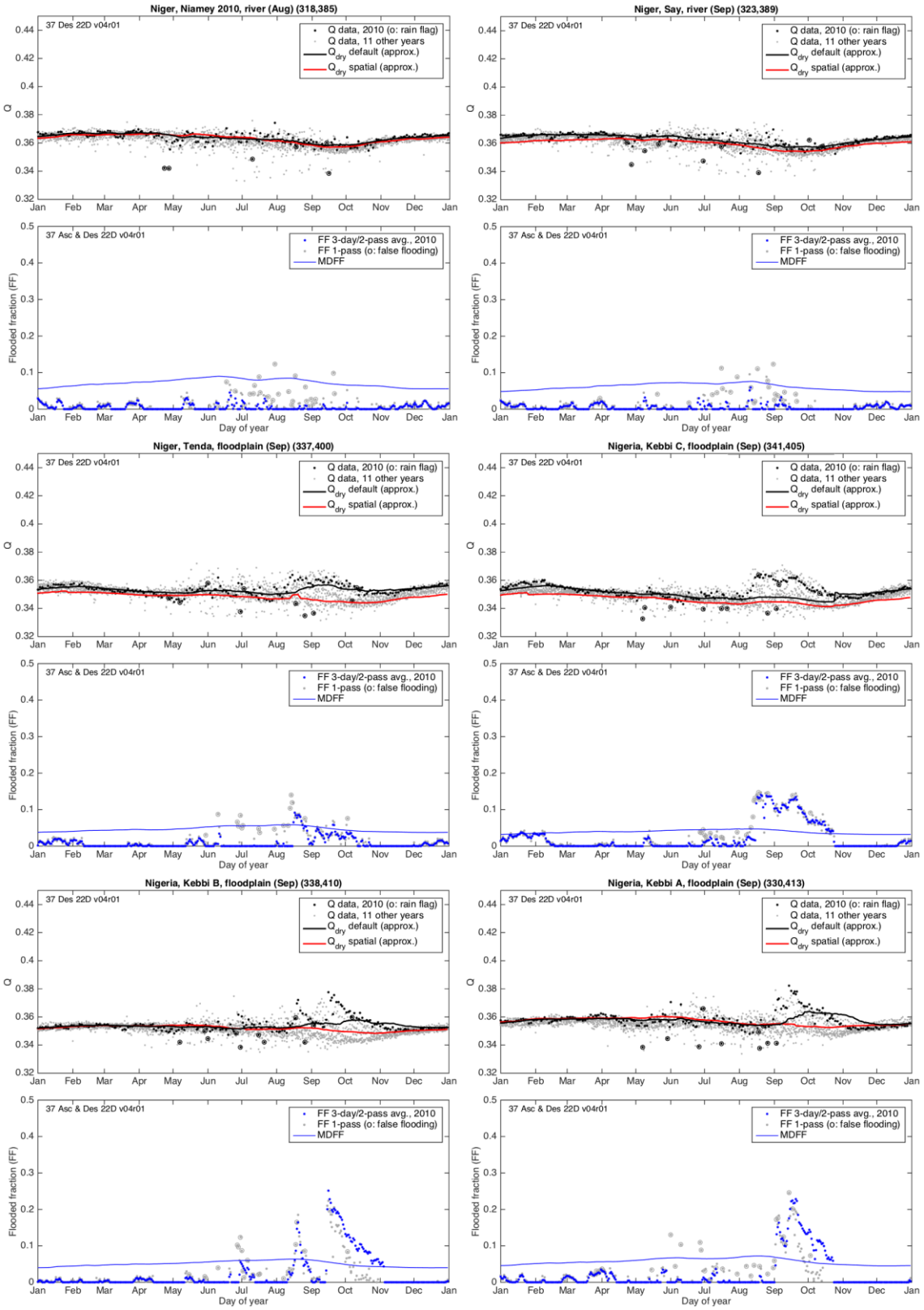
- *Side-by-side comparisons of MODIS false-color imagery to AFED*: Images were selected based on proportion of cloud free area and the degree to which the image can be interpreted to infer the timing and location of flooded or non-flooded conditions.







- *Time series analyses*: Plots are described in section 3.5.1.1.



Discussion:

The first MODIS image shows heavy cloud cover on 2010-08-10 corresponding to reported heavy rain and flooding in Niamey. AFED shows some flooding on the Niger far downstream from Niamey in Nigeria but none at Niamey itself. On 08-12 and 08-14 MODIS has are clear views of Niamey, which indicates that the width of the flooded area is relatively small; the narrow Niger River flood plain at Niamey is the most likely reason that the flooding at Niamey was not detected by the microwave flooded fraction algorithm.

The remaining images show the progression of flooding south of Niamey on the Niger and on the Sokoto River, which flows into the Niger from the northeast in Nigeria. The lower Niger and Sokoto appear to have significant flooded areas in the MODIS images on 08-16, 08-21, 08-23, and 08-25, which correspond to flooding in the same areas of the AFED maps. By 09-03, an area of flooding appears in the MODIS image and AFED map in the northeast part of the Sokoto. The flooding on the Sokoto has progressed further south on 09-11 in the MODIS images; the AFED results are in good correspondence with this progression.

The time series plots indicate that the microwave algorithm did not detect a flooded fraction signal at Niamey that could be distinguished from false positives. The false positive detection algorithm flagged almost all points in Niamey in August with flooded fraction values greater than zero as false positive flashes. At Say, some 1-pass FF values are above MDFF but the final FF never is. Further downstream on the Niger River (at points Tenda and Kebbi C) more flooding is detected for several reasons. First, the MDFF becomes progressively lower, presumably due to the increase in vegetation cover further south that reduces variability in dry-land *Q* data. Second, the flood plain becomes progressively wider, particularly at Tenda and Kebbi C, as seen in the relative floodability map. Finally, there is likely more water flowing further downriver after other streams have converged with the Niger. The number of false positives at the downstream points is similar to Niamey, suggesting that false positive detection is not the main reason that flooding at Niamey goes undetected. At Tenda, FF exceeds MDFF for a few days in mid August; at Kebbi C, flooded fraction is consistently high from about 08-12 to 10-23.

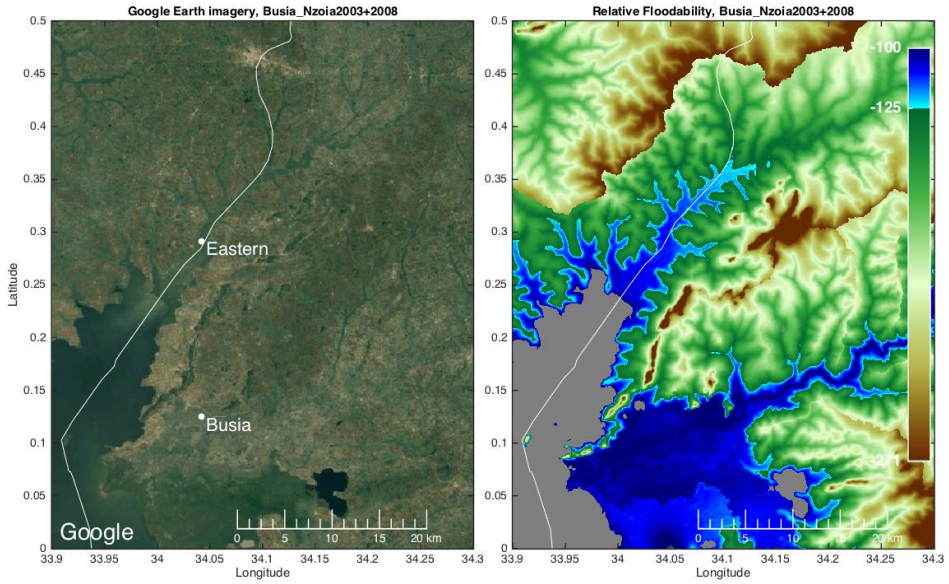
At Kebbi B on the Sokoto River, the flooded fraction is more variable; flooded fraction is above MDFF from mid to late August, falls below MDFF for a time, then increases dramatically around 09-14 and remains above MDFF into November. At Kebbi A, flooding does not begin until early Septmember and lasts through late October. These time series are consistent with the imagery. The Kebbi B area of the Sokoto appears to be flooded early in the event in concert with the flooding on the Niger; the areas upstream of Kebbi B on the Sokoto do not appear to be flooded until the second separate event first appearing on 08-25 arrives. The second peak of flooded fraction beginning on 09-14 at Kebbi B appears tied to the same flood first detected at Kebbi A by the AFED algorithm on 09-05. Note that at Kebbi A and B the dry-land end-member algorithm transitions from the default (black line on plots) to the alternative spatial Q_{dry} (red line) during the September-October flooding. (The transition does not occur at Kebbi C where the wetlands flagging overrides the logic.) The continuity of spatial Q_{dry} before, during, and after the flooding is good evidence that the algorithm accurately estimates true Q_{dry} conditions.

3.5.1.5 *Busia District and the River Nzoia, Kenya, 2003 & 2008*

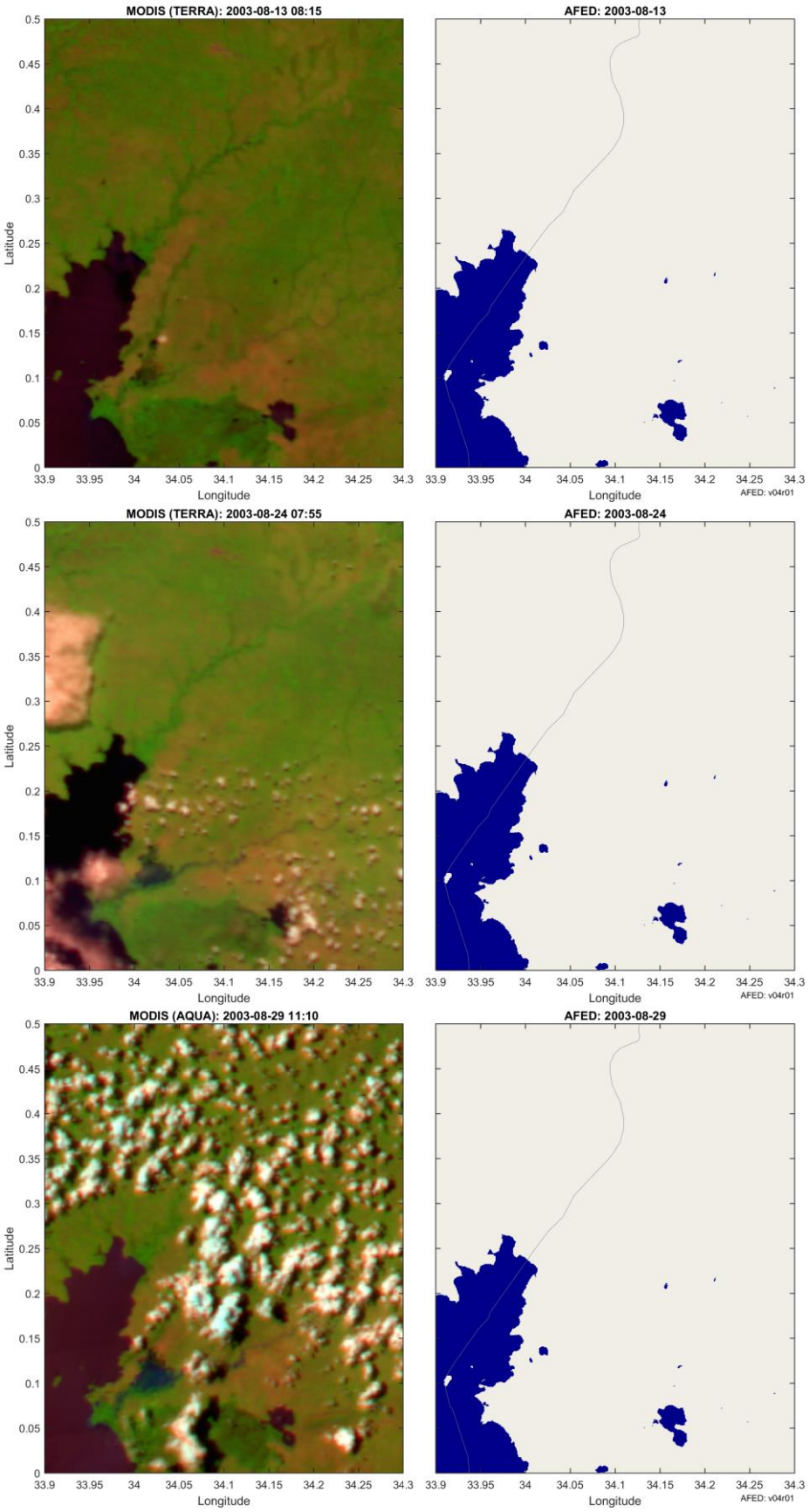
- *2003 event summary (from ARC)*: MODIS observations show flooding near the mouth of the Nzoia River on August 29, 31 and September 5, 7, 12. There were 6500

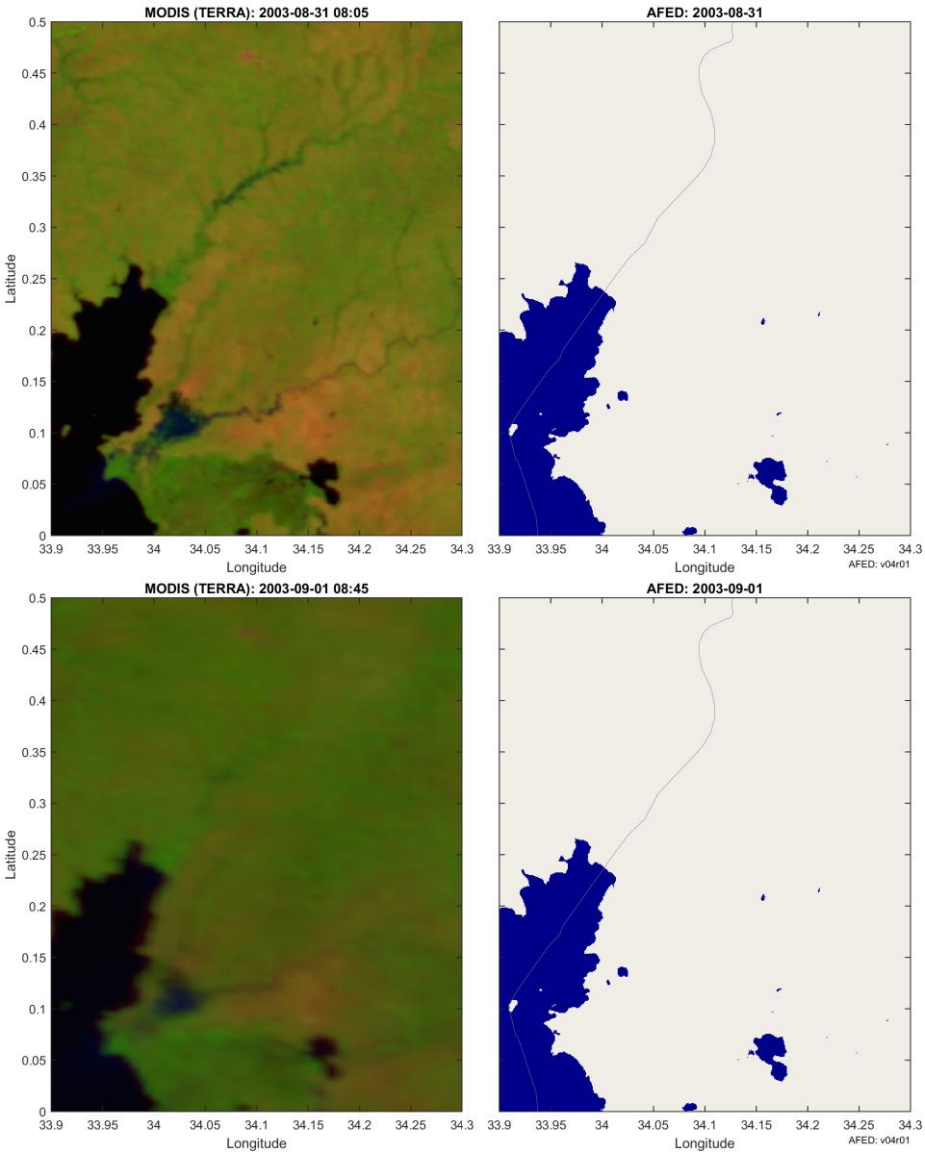
in camps by 4 September (http://www.dartmouth.edu/~floods/Archives/2003sum.htm).

- *2008 event summary*: Like many rivers across the world, the Nzoia River in western Kenya pushes over its banks annually. Dykes usually contain the river, but in November 2008, the river burst through the dykes to flood the low-lying land around it. At least 5,000 people were marooned or evacuated from the banks of the swollen river, said the United Nations Office for the Coordination of Humanitarian Affairs on November 14. (http://earthobservatory.nasa.gov/IOTD/view.php?id=35977; includes image captured by NASA's Advanced Land Imager, ALI, on 13 November 2008.)
- *Location map indicating microwave grid points selected for time series analyses*: Results for a subset of these grid points are given below. See Figure 16 for figure caption.

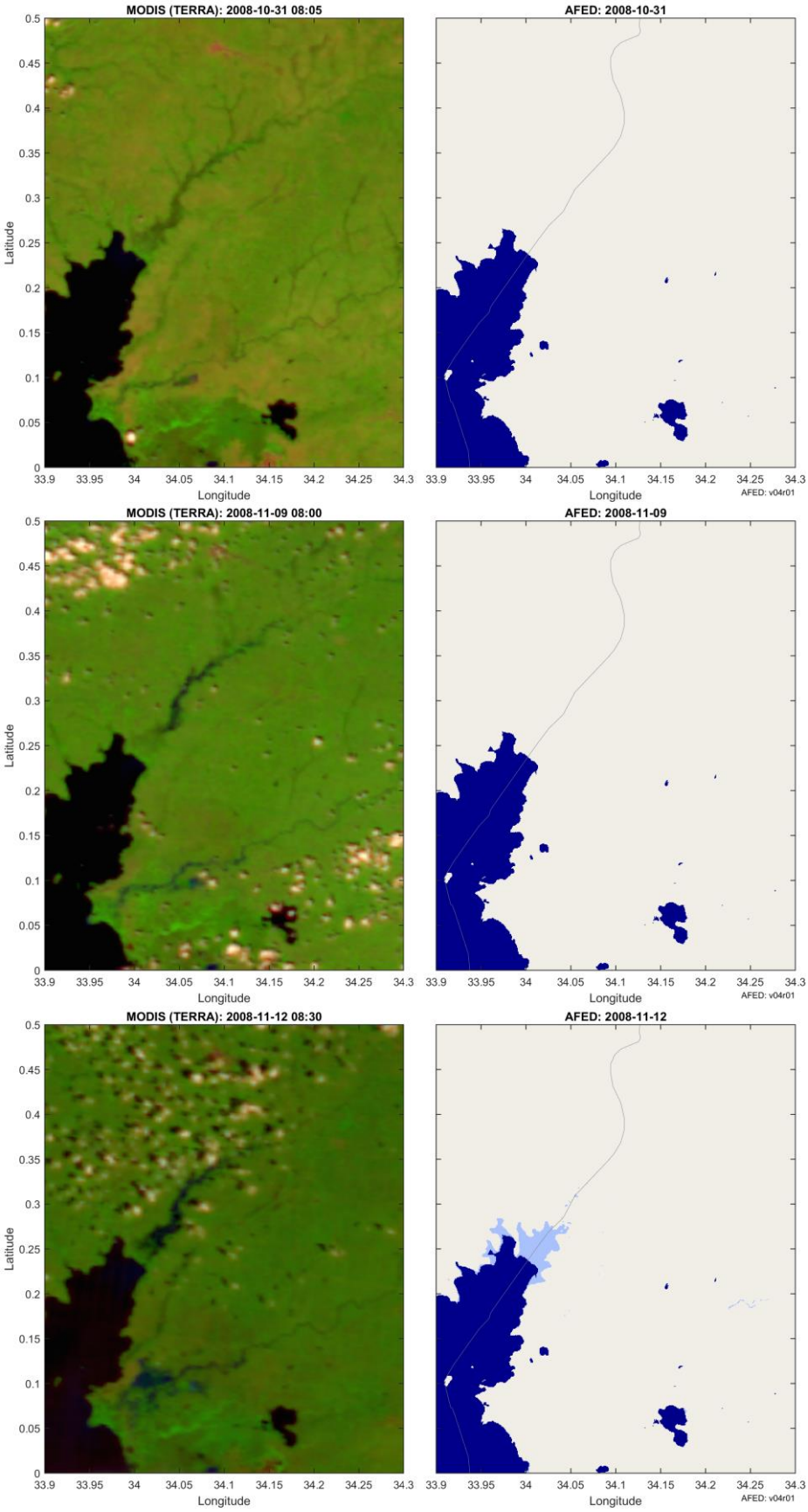


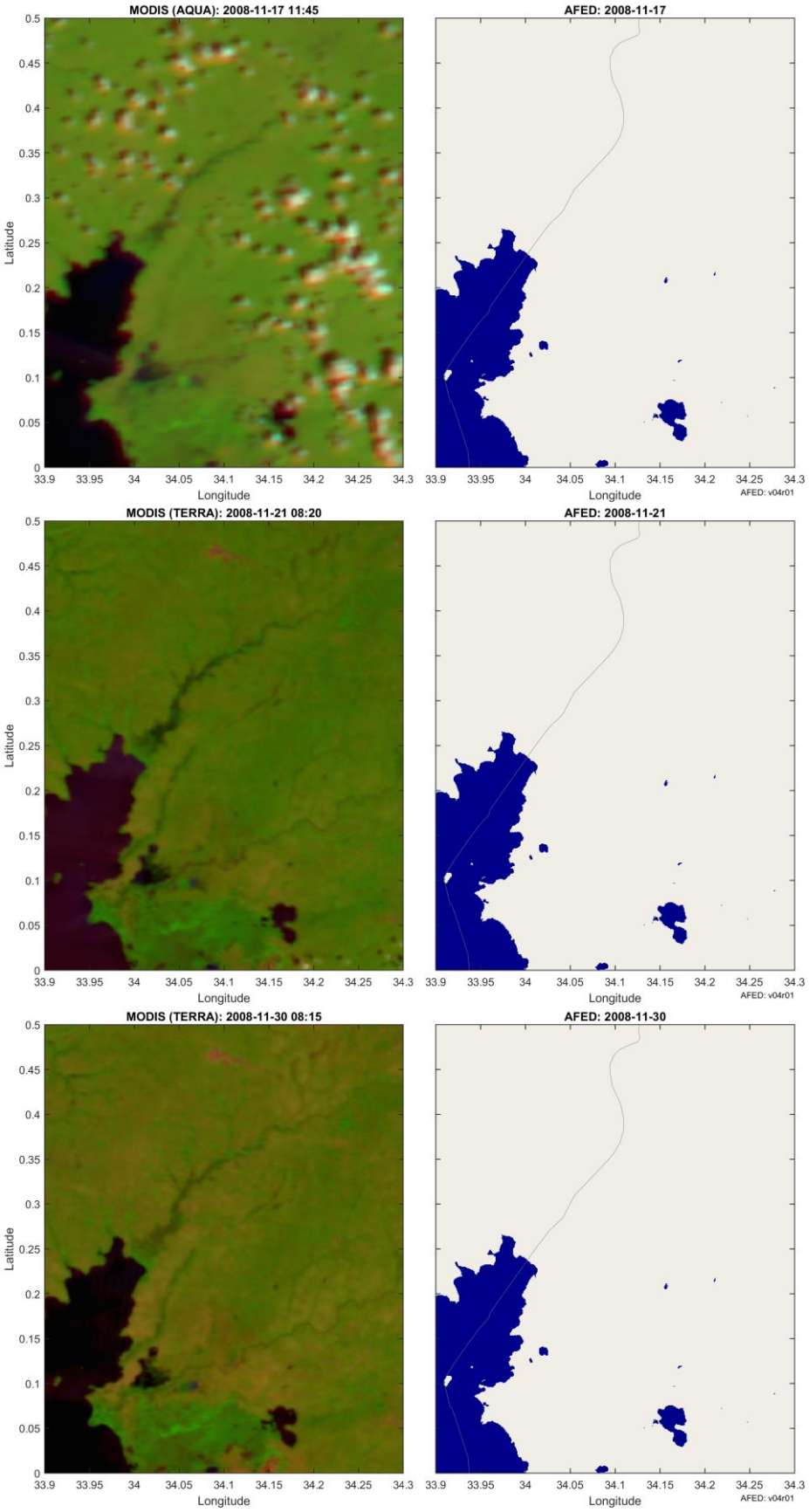
- ***2003*** *side-by-side comparisons of MODIS false-color imagery to AFED*: Images were selected based on proportion of cloud free area and the degree to which the image can be interpreted to infer the timing and location of flooded or non-flooded conditions.



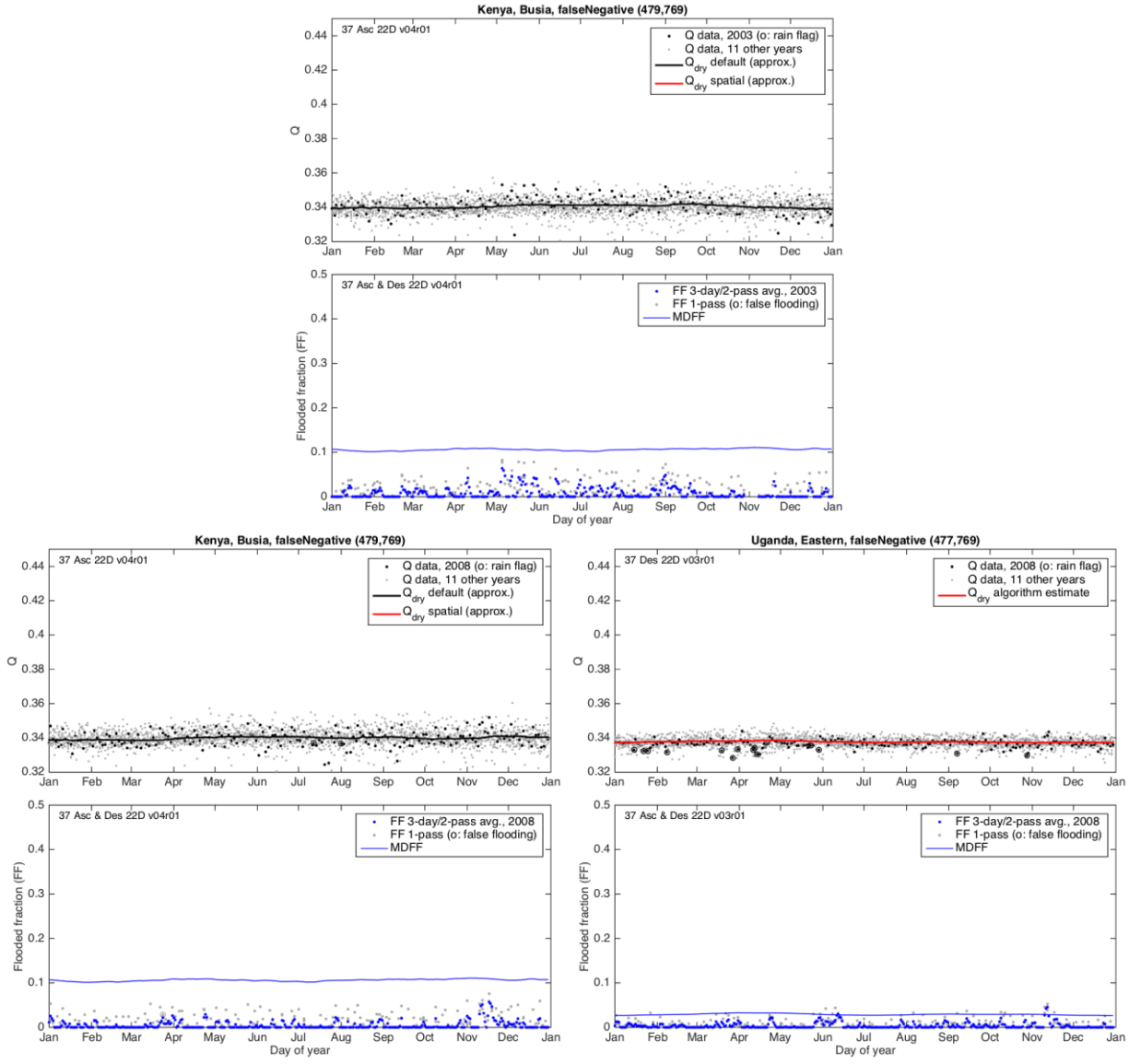


- ***2008*** *side-by-side comparisons of MODIS false-color imagery to AFED [next page]*: Images were selected based on proportion of cloud free area and the degree to which the image can be interpreted to infer the timing and location of flooded or non-flooded conditions.





- *2003 and 2008 time series analyses*: Plots are described in section 3.5.1.1.



Discussion:

The 2003 and 2008 cases are similar in the Nzoia River area. In 2003, the MODIS imagery shows the most significant flooding on 2003-08-24, 08-29, 08-31, and 09-01 while AFED shows flooding no flooding. In 2008, MODIS flooding on the Nzoia can be seen on 2008-11-12, 11-17, and 11-30 while AFED shows flooding no flooding. In 2008 there is flooding on the Kenya-Uganda border that is first apparent in the MODIS image on 2008-11-09, has grown in extent by 11-12, and remains visible through 11-21. AFED depicts flooding in the border area with approximate correspondence to the MODIS imagery only on 11-12.

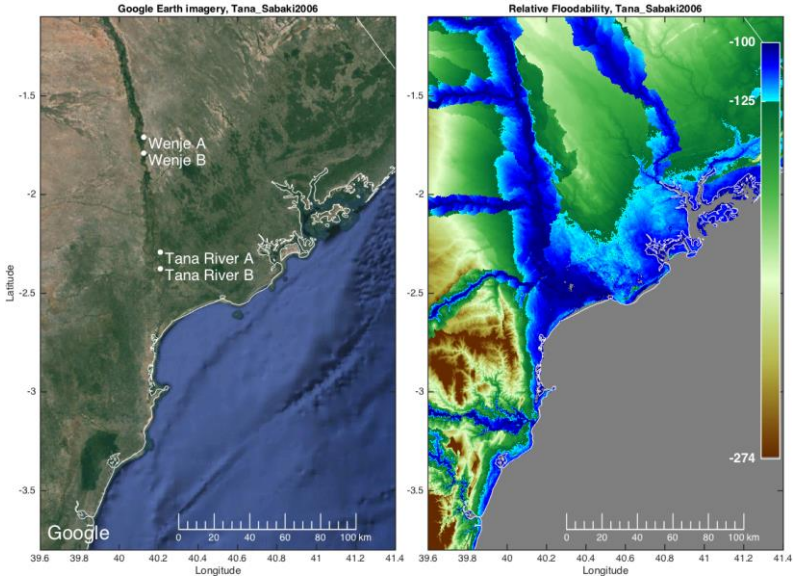
We used Google Earth to estimate that the maximum extent of the flooded area near the mouth of the Nzoia is about 40 km^2, or about 0.067 flooded fraction at 22-km resolution, which is just above the 0.054 MDFF at Busia. A higher-resolution ALI image on 2008-11-13 (not

included) shows that flood coverage is less than 100% in this area, which suggests that the flooded area was never significantly above the MDFF level needed for detection.

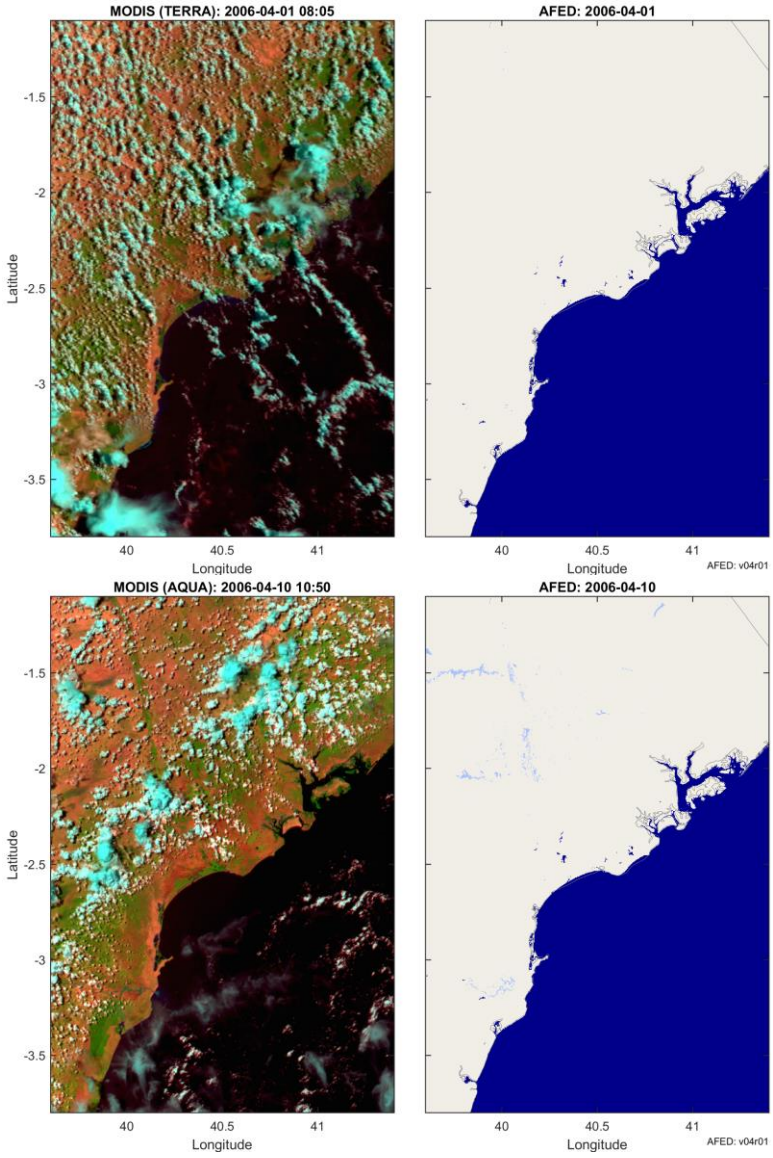
The Busia 2003 and 2008 time series plots are also alike. There is day-to-day noise in *Q* observations that can be attributed to the proximity of the Busia point to the Lake Victoria coastline (about 6 km away). *Q* is generally more variable near coastlines because there are small differences in the shape and location of the microwave footprint with each observation and, as a result, some observations encompass more of the water body than others. Like Busia, the Eastern point is also close to the lake—about 7 km away—but has significantly less day-to-day *Q* variability than Busia because the shape of the coastline means that the lake covers less of the footprints at Eastern than at Busia. MDFF at Eastern is lower than that at Busia as well. For example, the 2008 microwave algorithm flooded fraction peak at Busia is 0.057 (where MDFF is 0.11) but only 0.045 at Eastern (where MDFF is 0.042). Flooding is more easily detected at Eastern than Busia because MDFF is lower at Eastern than it is at Busia.

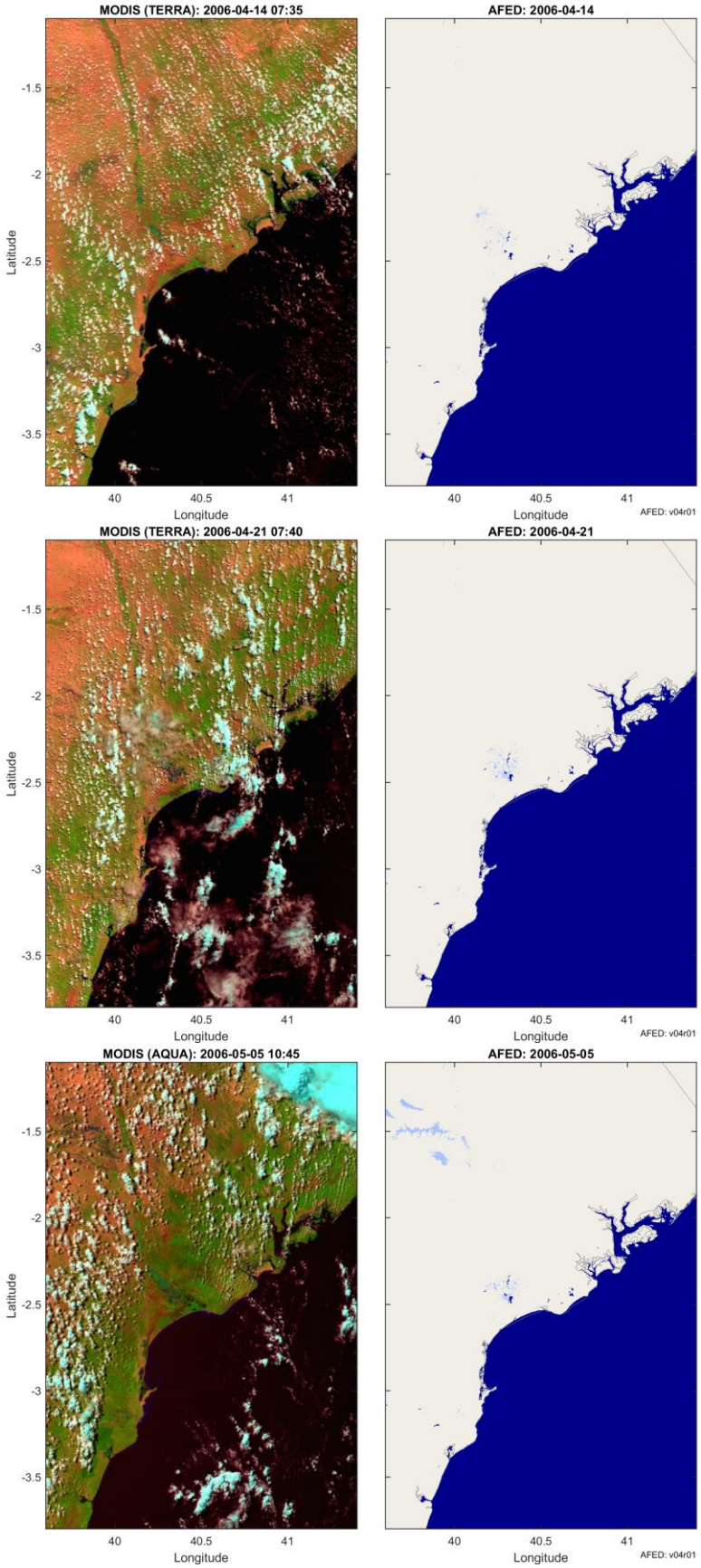
3.5.1.6 *Tana & Sabaki Rivers, Kenya, 2006*

- *Event summary from ARC*: The Dartmouth Flood Observatory groups the following districts of Kenya into one flood event from 2006-04-06 to 2006-05-22: Malindi, Kilifi, Kwale, Nyando, Homa Bay, Migori, Siaya, Rachuonyo, Isiolo, Samburu, Laikipia, Wajir, Garissa, Mandera. The Dartmouth report focuses on flooding on the Tana and Sabaki rivers and an area including only the Tana River, Kilifi, Garissa, and Lamu districts. This analysis also focuses on the Tana and Sabaki river areas.
- *Event information sources*:
  - http://www.dartmouth.edu/~floods/Archives/2006sum.htm
  - http://www.dartmouth.edu/~floods/2006061Kenya.html
- *Location map indicating microwave grid points selected for time series analyses*: Results for a subset of these grid points are given below. See Figure 16 for figure caption.

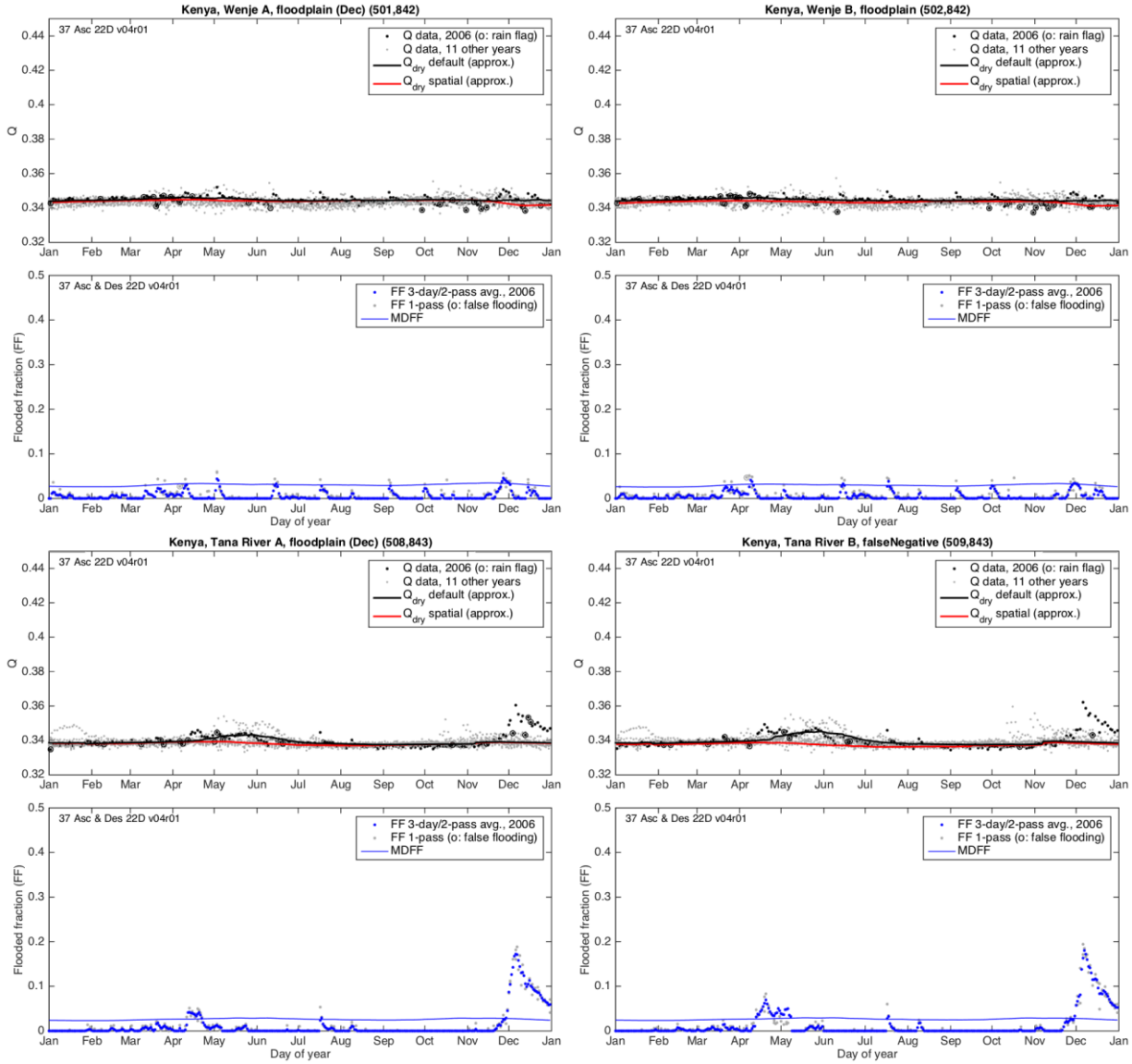


- *Side-by-side comparisons of MODIS false-color imagery to AFED*: Images were selected based on proportion of cloud free area and the degree to which the image can be interpreted to infer the timing and location of flooded or non-flooded conditions.





- *Time series analyses*: Plots are described in section 3.5.1.1.



Discussion:

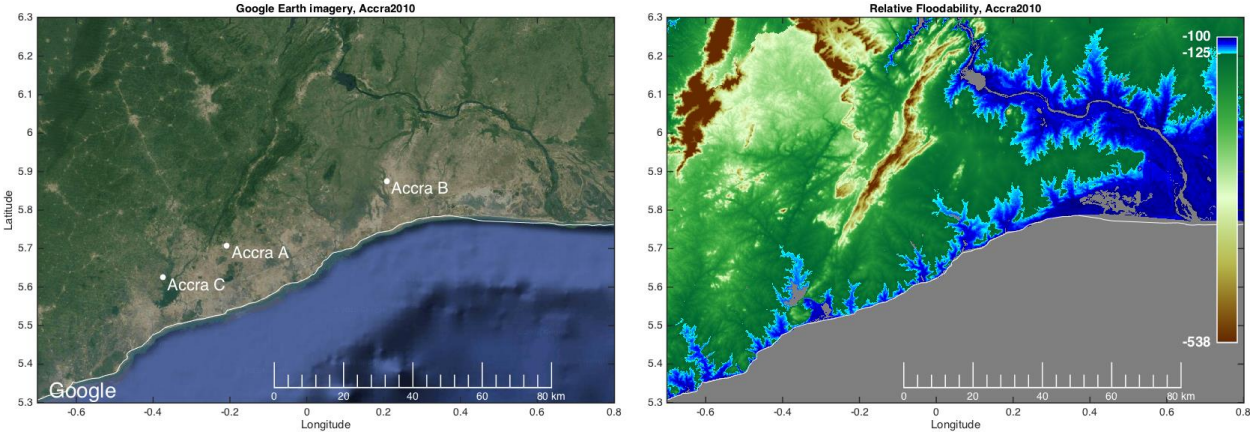
Persistent patching cloudiness make interpretation of the MODIS images difficult for this case. The pre-flood image on 2006-04-01 provides a baseline reference against which subtle changes in the surface in later images may be judged. On 04-10, AFED shows flooding up stream in the Tana basin as well as near the mouth of the Sabaki (between 3.5 and 3° S latitude) that appear to correspond to changes in the MODIS false color image probably indicative of surface water. The MODIS and AFED maps on 04-14 show evidence of water cover closer to the mouth of the Tana, which seems to confirm that flooding had occurred earlier up river. The 04-14 MODIS image also show continued flooding on the Sabaki; AFED shows flooding on the Sabaki 04-10 through 04-13 (not shown). On 04-21, the area of flooding near the Tana mouth has increased in size and there are no other areas of obvious flooding in the MODIS images. By

05-05 an area of widespread flooding is apparent in the northwest corner of the domain in both the MODIS and AFED maps; parts of this area are flooded in the AFED data from 05-02 through 05-08. After 05-08, only the area near the mouth of the Tana River remains flooded.

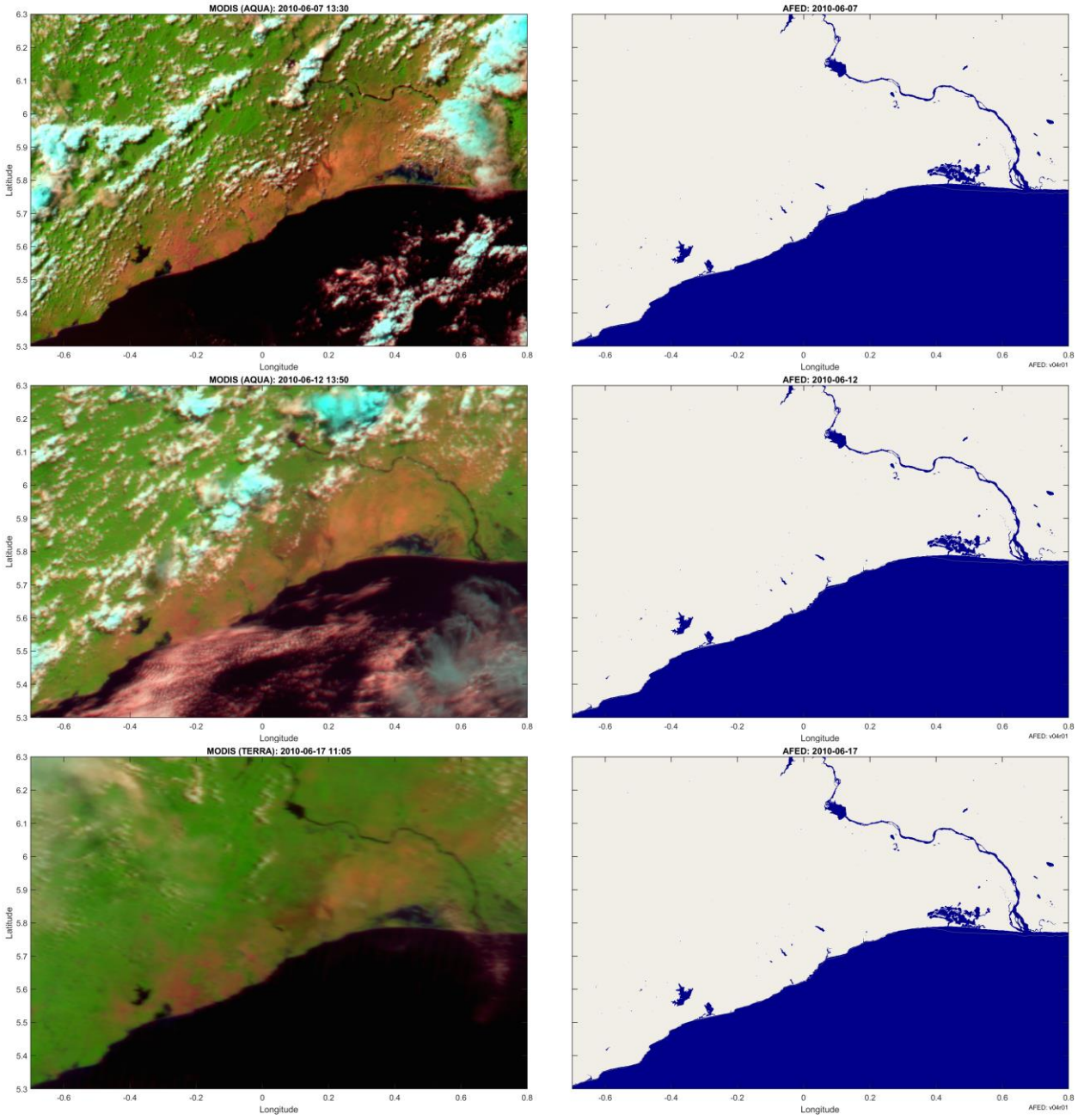
The time series plots show that the flooded fraction values reach peaks of 0.030 (on 04-11) and 0.041 (on 04-09) at points Wenje A and B, respectively, and 0.041 (on 04-20) and 0.069 (on 04-20) at Tana River A and B during the April-May event. At Tana River B, the flooded fraction is elevated (but not always greater than MDFF) from 2006-04-13 to 05-08. The earlier flooding at Wenje A and B is consistent with the large-scale flooding seen in the imagery on 04-10. The later and longer lasting flooding at points Tana River A and B is also consistent with the imagery and generally consistent with flooding in a downriver lowlands area. MDFF is about 0.03 at all time series points during the event; from the time series it can be seen that flooded fraction at Tana River A and B are more highly correlated in time during a large December flood event. During the April event, FF at Tana B is slightly higher than that at Tana A initially, which after 12 days triggers the AFED algorithm to transition to spatial Q_{dry}, which in turn prolongs flooding at Tana B. As noted in the Niamey case, The continuity of spatial Q_{dry} before, during, and after the flooding is good evidence that the spatial Q_{dry} algorithm is an accurate estimate of true Q_{dry} conditions at both Tana A and B.

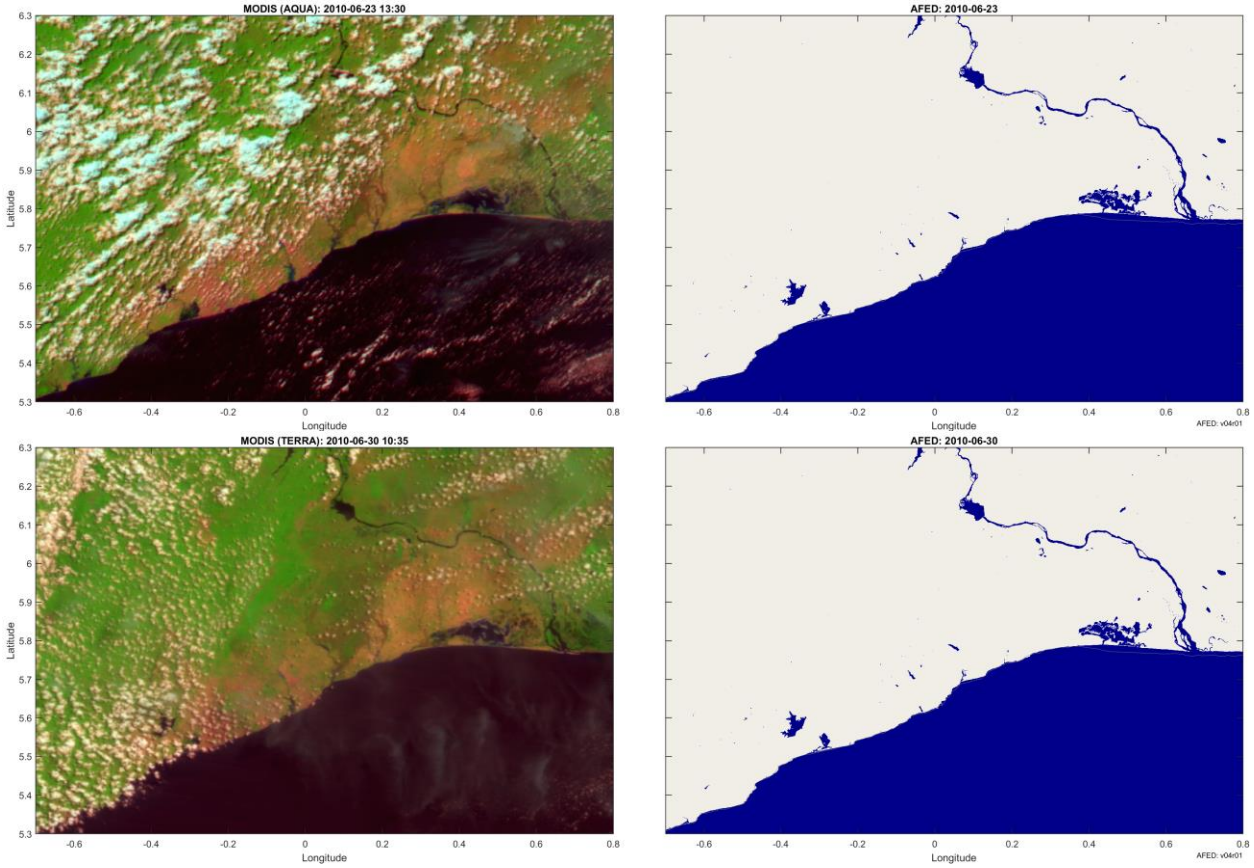
3.5.1.7 *Greater Accra, Ghana, 2010*

- *Event summary from ARC*: Greater Accra, Ghana, 22 June 2010. This is a rainfall related flash flood.
- *Event information sources*:
  - http://www.modernghana.com/news/281376/nations-worst-flood-disaster-death-toll-now-35.html
- *Location map indicating microwave grid points selected for time series analyses*: Results for a subset of these grid points are given below. See Figure 16 for figure caption.

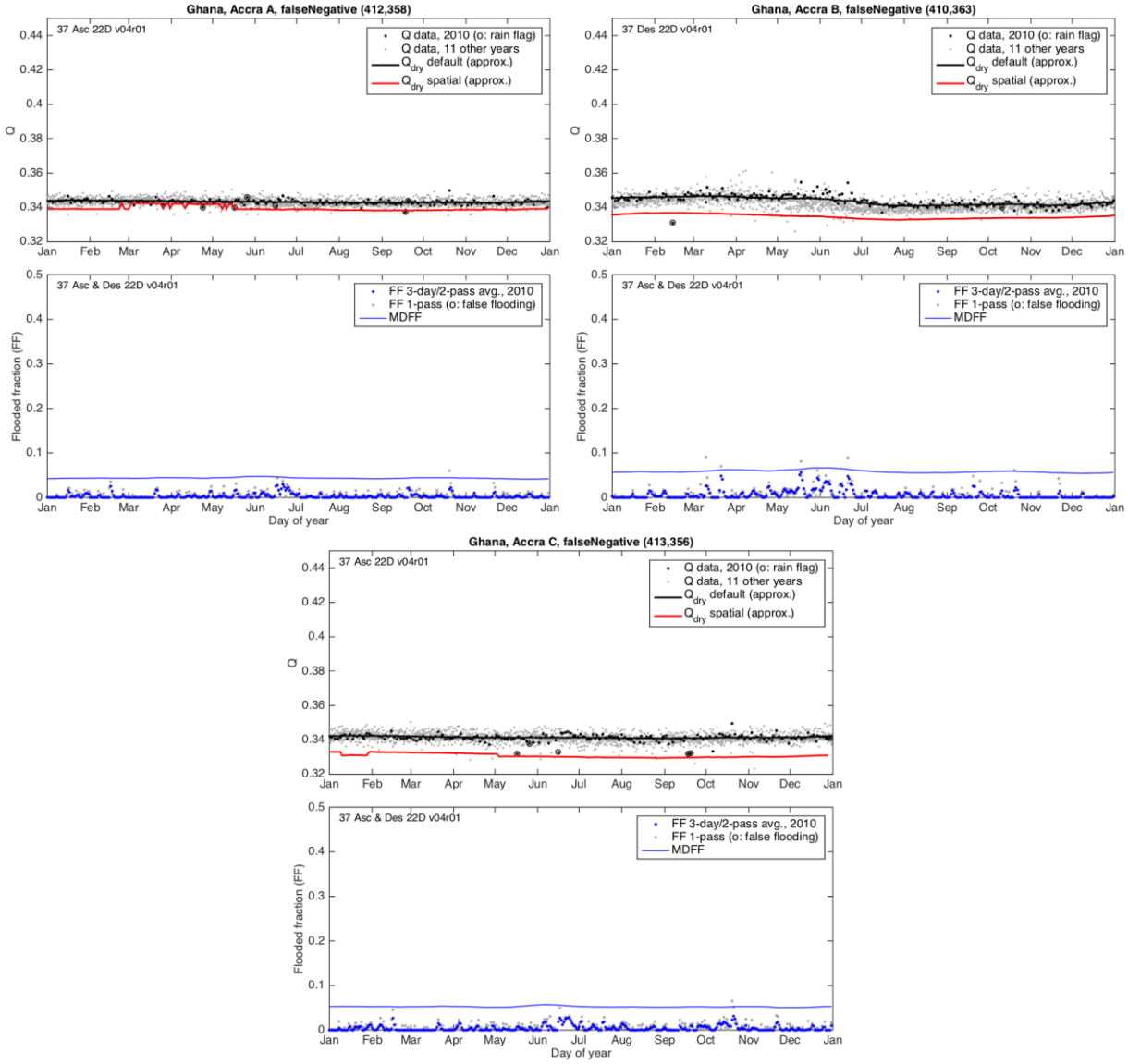


- *Side-by-side comparisons of MODIS false-color imagery to AFED*: Images were selected based on proportion of cloud free area and the degree to which the image can be interpreted to infer the timing and location of flooded or non-flooded conditions.





- *Time series analyses*: Plots are described in section 3.5.1.1.



Discussion:

There is little indication of flooding in MODIS imagery during the June 2010 Greater Accra event and AFED did not depict any flooding in the area during the event. On 2010-06-07 and 06-12 there is no obvious flooding apparent in the MODIS imagery. On 06-16 there is heavy cloud cover (not shown) but on 06-17 there is still no flooding apparent in the MODIS image. Heavy cloud cover occurs again on 06-20 and 06-21. There are hints of flooding in the clear-sky areas of the MODIS image on 06-23, e.g., to the west of time series point Accra B, but no area stands out as flooded in the image with contrast comparable to the Volta River.

The flooded fraction time series plots at points Accra A, B, and C are alike: a small amount of flooding (just above zero flooded fraction) begins around 2010-06-17, peaks around 06-23, and falls back to near zero by about 06-30. The FF peak is about 0.029, 0.048, and 0.028 at points A, B, and C, respectively, with corresponding MDFF values of 0.046, 0.064, and 0.056.

The Africa median MDFF is 0.047 so the day-to-day Q noise at these points is relatively low. The flooded fraction comes close to but does not exceed MDFF at all of the points. The main factor keeping the flood from being represented in AFED appears to be that the flood extent is not large enough relative to the MDFF. If the MDFF algorithm were modified to decrease MDFF, there would likely be an increase in AFED false positives throughout Africa.

3.5.1.8 *Grande Comore, Comoros, April 2012*

- *Event report from ARC and ReliefWeb*: No flooding can be detected in AFED in Comoros. The surface areas of the islands are: Grand Comore 1,148 km^2, Anjouan 290 km^2, and Moheli 424 km^2. We would expect that the second and third island would not show any flooding but for Grand Comore, based on its size, flooding might be possible. A major event happened on [20 April 2012], when a nearby passing cyclone caused extensive flooding. ReliefWeb reports flooding occurred on 20 April: http://reliefweb.int/disaster/fl-2012-000066-com.

Figure 18: Comoros population map from ARC (left) and detail of Grande Comore island from Google Earth (right) showing the microwave data grid points "Comoros N" and "Comoros S" discussed here. The island is about 23 km wide at the widest point orthogonal to the predominantly north-south spine of the island.

Analysis

Our analysis—originally performed for AFED V03R01 and updated for V04R01—indicates several reasons why the AFED algorithm did not detect flooding on Grande Comore (GC) in April 2012 and, furthermore, would not have detected similar flooding on GC during other periods.

At its widest point, GC is only about 23 km wide, which means that all microwave data grid points on the island include significant water cover fraction. We investigated data from two grid points, labeled "Comoros N" and "Comoros S" in Figure 18 (hereafter, CN and CS) in the area where ReliefWeb indicated significant flooding. April 2012 is in the gap period between the end of 22-km AMSR-E coverage in October 2011 and the beginning of AMSR2 coverage in August 2012; AFED during this period is based on 50-km SSM/I data. The 50-km persistent water fractions are 0.71 and 0.75 for CN and CS points, respectively, meaning that the algorithm does

not recognize these—or any other points on the islands—as land where flooded fraction can be derived. Therefore, the algorithm is incapable of detecting flooding during the April 2012 event.

To investigate whether the flood could have been detected if 22-km data had been available during the April 2012 event, we looked at 2015 AMSR2 data at grid points CN and CS. The 22-km persistent water fractions are 0.35 and 0.45 for CN and CS points, respectively, which means that the AFED algorithm recognizes them as land points for flooded fraction derivation. However, the 27-km water fractions are 0.498 and 0.57 for CN and CS, respectively, meaning that the AFED V03R01 rain detection algorithm treats CS as a water point (i.e., water fraction greater than 0.5). For water points, the rain detection algorithm uses a more sensitive detection scheme that can trigger rain false positives at grid cells with partial land cover. Therefore, all AMSR2 data at CS were rain flagged in V0R01. As of AFED V04R00, the rain detection algorithm treats points as land unless they are greater than 0.95 water fraction. As a result, AMSR2 data at CS are no longer incorrectly flagged as rain.

Figure 19 shows 2015 AMSR2 time series results for the Comoros N (CN) and S (CS) grid points. With 27-km persistent water fraction under 0.95 at both points, the land version of the rain-flagging algorithm is in effect and rain flags at CN and CS are uncommon. Q data at CS and CN are alike: there is a high degree of day-to-day and year-to-year Q variability, which is largely attributable to differences in the persistent water fraction falling within the microwave footprint due to satellite data geolocation error and footprint asymmetries. Flooded fraction retrievals are also variable, with some reduction in noise relative to Q due to the 3-day averaging step in the flooded fraction algorithm.

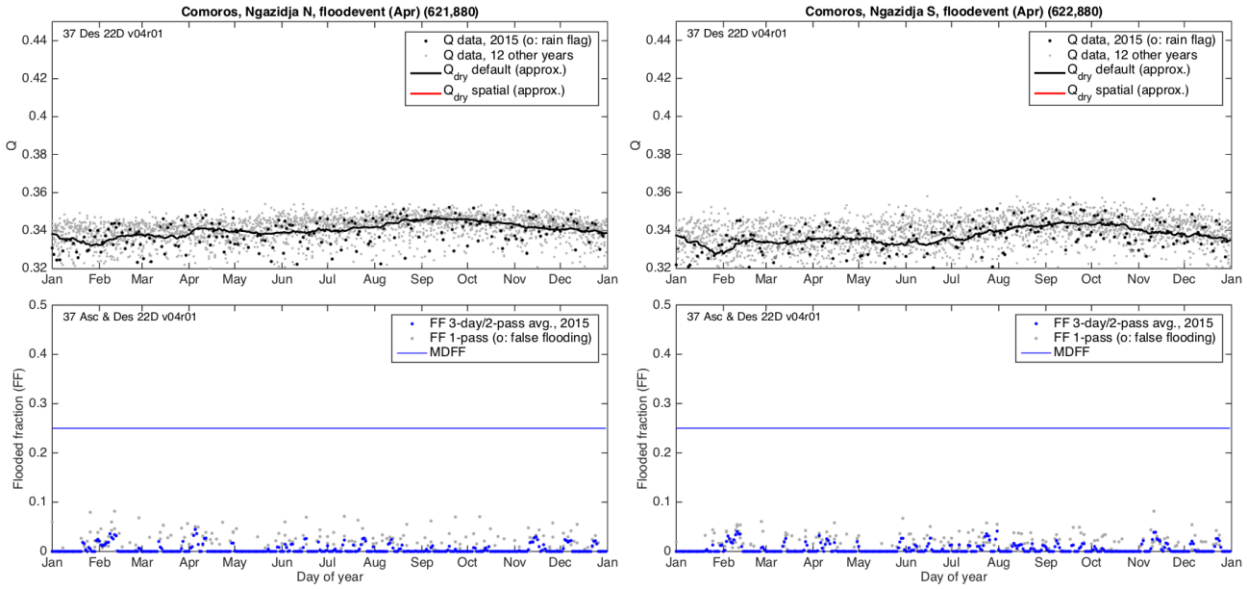


Figure 19: Time series data at the Comoros N (CS, left) and Comoros S grid points. Top: Q data for 2015 indicating that no points are flagged as rain. Bottom: Flooded fraction (FF) results and MDFF.

From these results we can conclude that flood detection is theoretically possible in Comoros but only from 22-km data and only when the flooded fraction exceeds the MDFF of 0.25. However, the steep terrain and lack of large river floodplains on Grande Comore suggests that large extent flooding is unlikely. Furthermore, when floods do occur they are likely to be short-lived flash floods that quickly drain to the ocean and are therefore difficult to detect from daily

or twice daily satellite observations. ReliefWeb indicates that the April 2012 events were flash floods, although no estimate of the duration of extensive flooding is given.

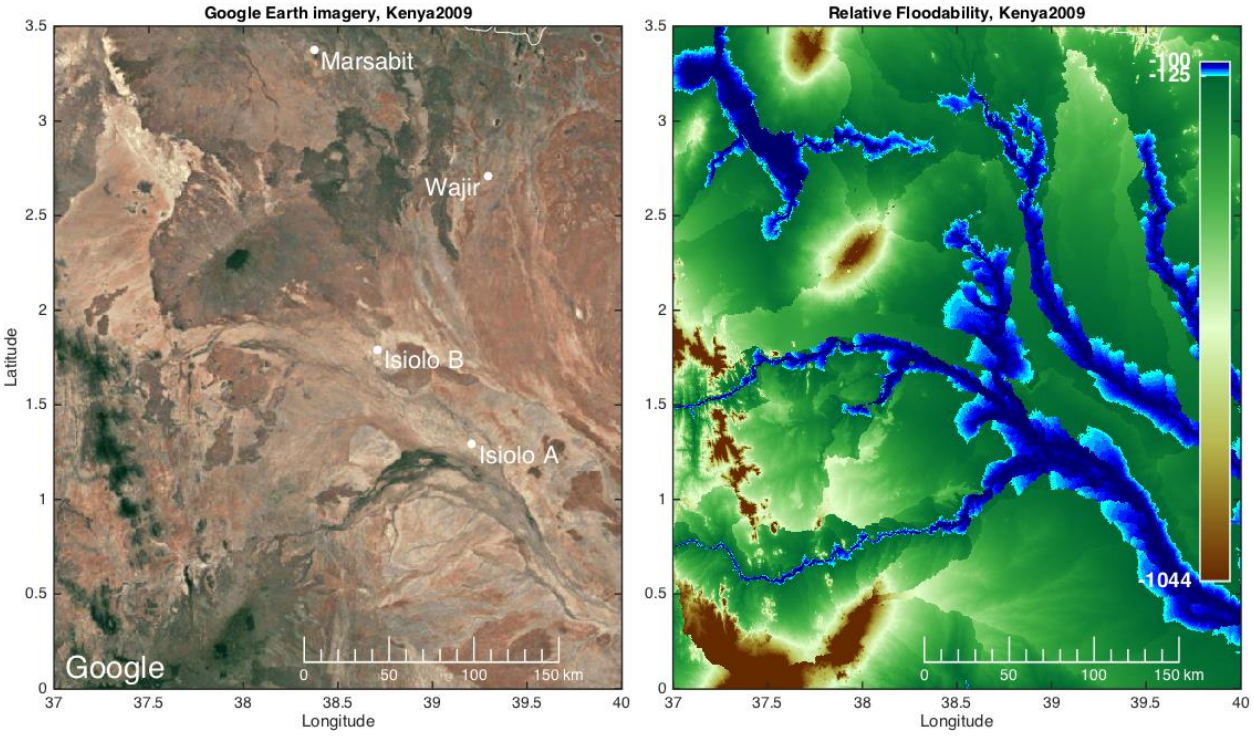
In summary, the April 2012 Comoros flood is not represented in AFED—and floods on Comoros are unlikely be represented at any time—for the following reasons:

1. AFED during April 2012 is based on 50-km SSM/I data. No part of the Comoros islands has 50-km resolution persistent water fraction less than 0.5 in the AFED microwave data grid. The AFED flooded fraction algorithm is not applied to grid points with more than 0.5 persistent water fraction.
2. In some areas the 22-km persistent water fraction is below 0.5 and the algorithm is therefore able to derive flooded fraction. However, even in these areas variability is high due to the effects of nearby water bodies, leading to high MDFF (0.25). The algorithm sets flooded fraction below MDFF to zero.
3. In order to detect flooding above the MDFF level when it occurs, the floods must be extensive, large river floods and must last for several (e.g., 3-5) days. The April 2012 flood was reported to be a flash flood and it does not appear that the steep topography of Grande Comore can generally support these two conditions.

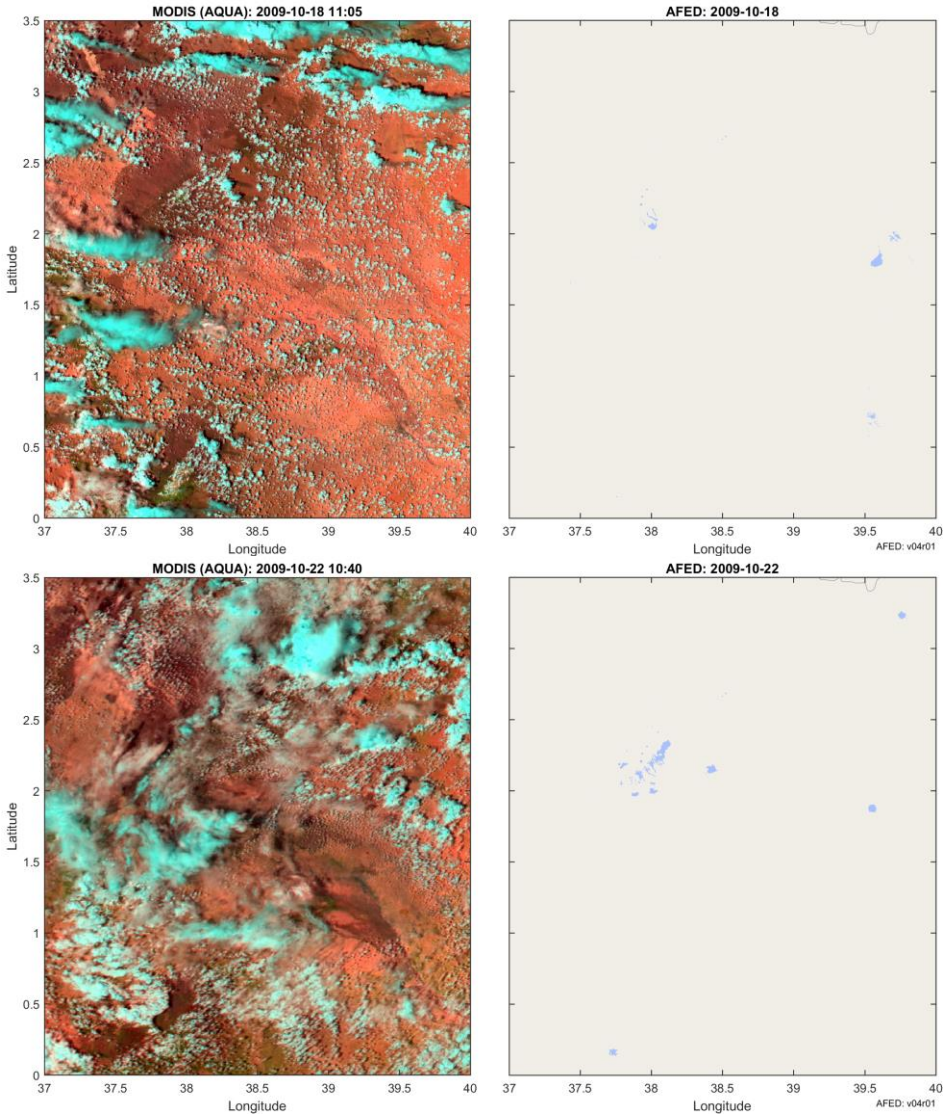
3.5.2 *False positive cases*

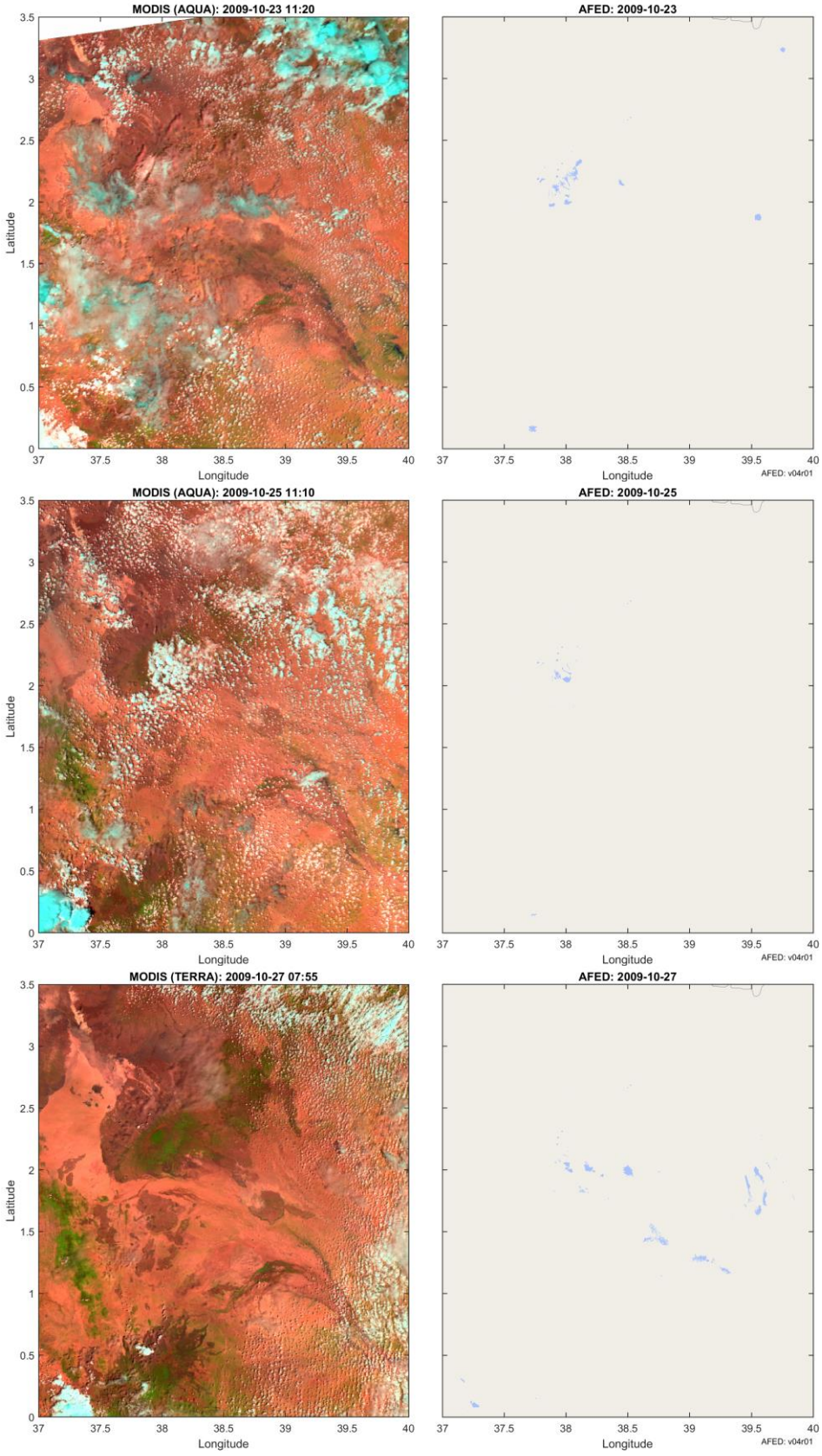
3.5.2.1 *Isiolo region, Kenya, 2009*

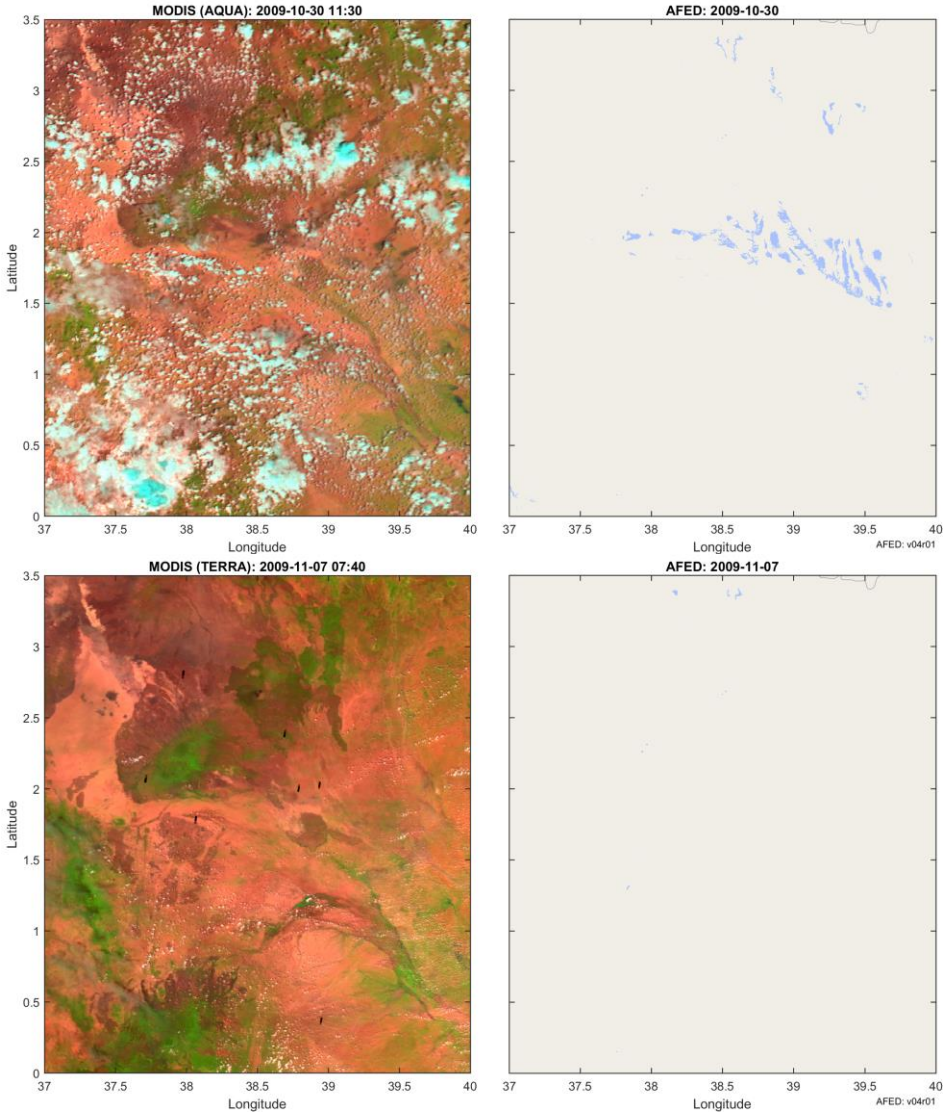
- *Period investigated*: 2009-10-14 to 2009-11-08.
- *Location map indicating microwave grid points selected for time series analyses*: Results for a subset of these grid points are given below. See Figure 16 for figure caption.



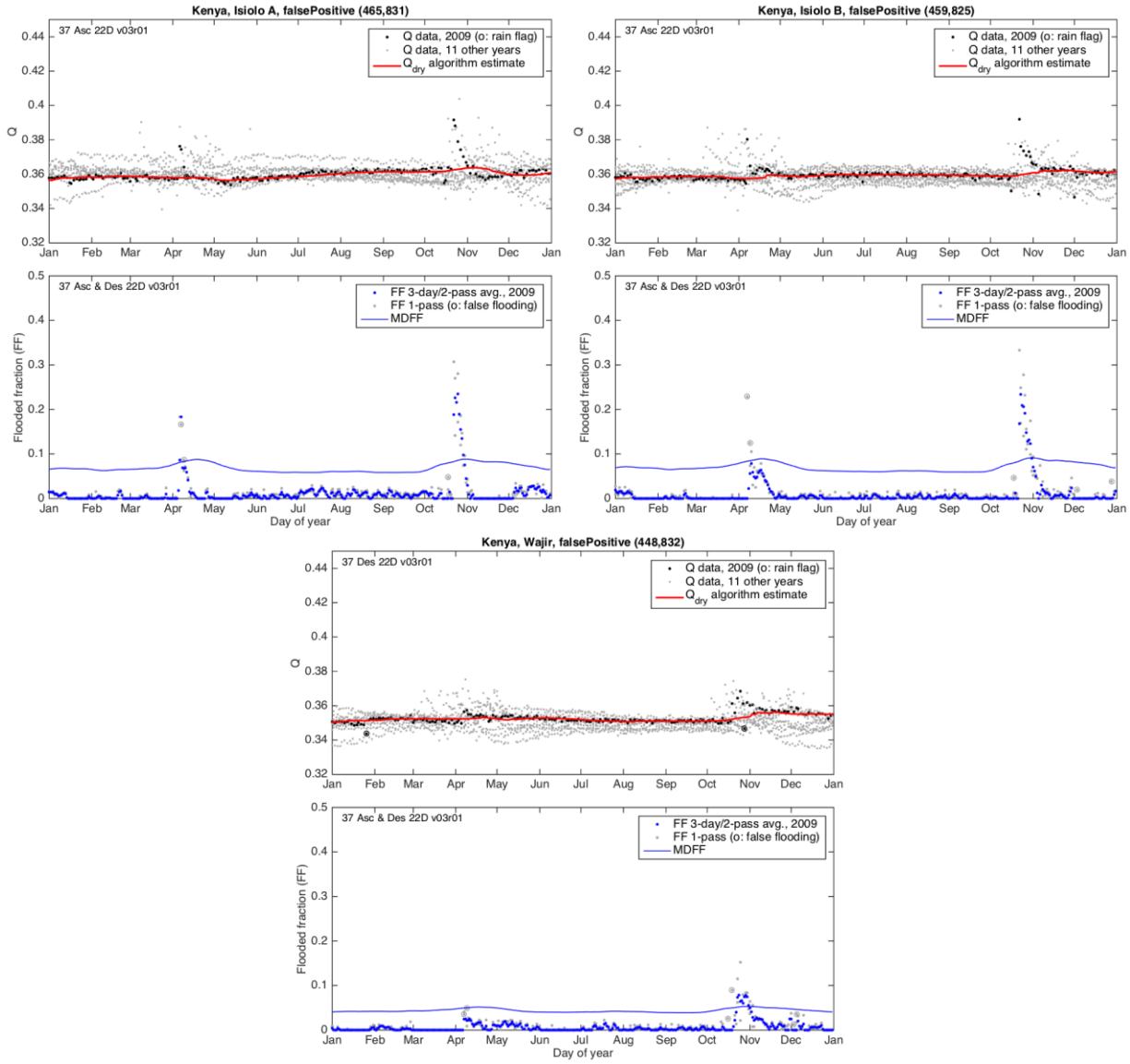
- *Side-by-side comparisons of MODIS false-color imagery to AFED*: Images were selected based on proportion of cloud free area and the degree to which the image can be interpreted to infer the timing and location of flooded or non-flooded conditions.







- *Time series analyses*: Plots are described in section 3.5.1.1.



Discussion:

Some flooding is apparent in the MODIS images although it is difficult to interpret because of persistent patchy cloud cover. The 2009-10-18 image is typical of other prior images and appears to indicate dry conditions compared to later dates. MODIS images on 2009-10-22, 10-23, 10-25, 10-27, and 10-30 show darker surface area with distinct patterns suggestive of flooding along flow paths. Until 10-27, AFED V04R01 has small patches of flooding the most persistent of which is centered near 2.1° N, 39° W. AFED develops a wider area of flooding on 10-27 that reaches peak extent on 10-30 and dissipates by 11-02 but at no point is this area coincident with the dark areas in the MODIS imagery. In the 11-07 case both the MODIS and AFED results indicate little or no obvious flooding. Although some flooding can be corroborated by MODIS imagery, it seems that at a minimum AFED is overestimating the extent of flooding in the 10-27 to 11-01 period even if this is not strictly a false positive.

Time series plots at points Isiola A and B indicate rapid increases in *Q* at the start of the event in October 2009 with flooded fraction increasing to about 0.13 on 10-29. Prior to that time there is very little day-to-day noise in *Q*. MDFF is about 0.09 at both Isiola A and B and the rise in MDFF in October indicates the start of the rainy season when there is more day-to-day and year-to-year *Q* variability. As of V04R01, the AFED algorithm flags a series of FF points as false positives (indicated by gray circles on the FF plot) at Isiola A and B from 10-18 to 10-27 that, which shortens the duration of the event but does not eliminate it entirely.

The time series data at Wajir indicate a smaller but still significant flooded fraction increase during the event, with multiple flash detections prior to FF exceeding MDFF 10-28 to 10-31.

Like the imagery analysis above, the time series data suggest there was wide-scale rain event that resulted in increased soil moisture and perhaps some flooding. A rainstorm can increase surface water—including soil moisture, puddles, and perhaps flooding in streambeds—rapidly and simultaneously over a large area. As surface water is absorbed, evaporates, or collects into low-lying areas, the spatial footprint of water rapidly decreases. The lack of significant vegetation cover in this area helps to amplify the microwave signal of soil moisture. The surface geology may also play a role, e.g., if it infiltration capacity is low. Therefore, the sharply peaked temporal profile of *Q* at different locations simultaneously is consistent with an event dominated by soil moisture and puddles, not river flooding, although it appears from the data that an unusual surface wetting event did occur and probably included some flooding in otherwise dry areas that persisted for days.

Figure 20 illustrates the false positive detection algorithm results during the period from 2009-10-21 to 2009-11-01[16]. Areas flagged as false positives are outlined in red. The algorithm flags a large area of northern Kenya as false positive from 10-21 to the first pass on 10-27. The second pass on 10-27 and the first pass on 10-28 (not shown but similar to 10-27) are not flagged. The second pass on 10-28 and the first pass on 10-29 have no data in this area. In the second pass on 10-29 a large region of high flooded-fraction appears; the algorithm does not flag it as a false positive and it dissipates by 11-01 However, this secondary event is likely a false positive based on its size, duration, quick development, and the absence of a major floodplain in the affected area. Although initial flooded fraction falls quickly, the final flooded fraction algorithm's 3-day averaging approach causes the first effects on AFED to occur on 10-28 and persist through 11-01, as seen in the MODIS-AFED image pairs above.

[16] 19-GHz flooded fraction is used for false positive detection and may differ from the final AFED algorithm flooded fraction results, which are based on 37-GHz data. See the ADD for details.

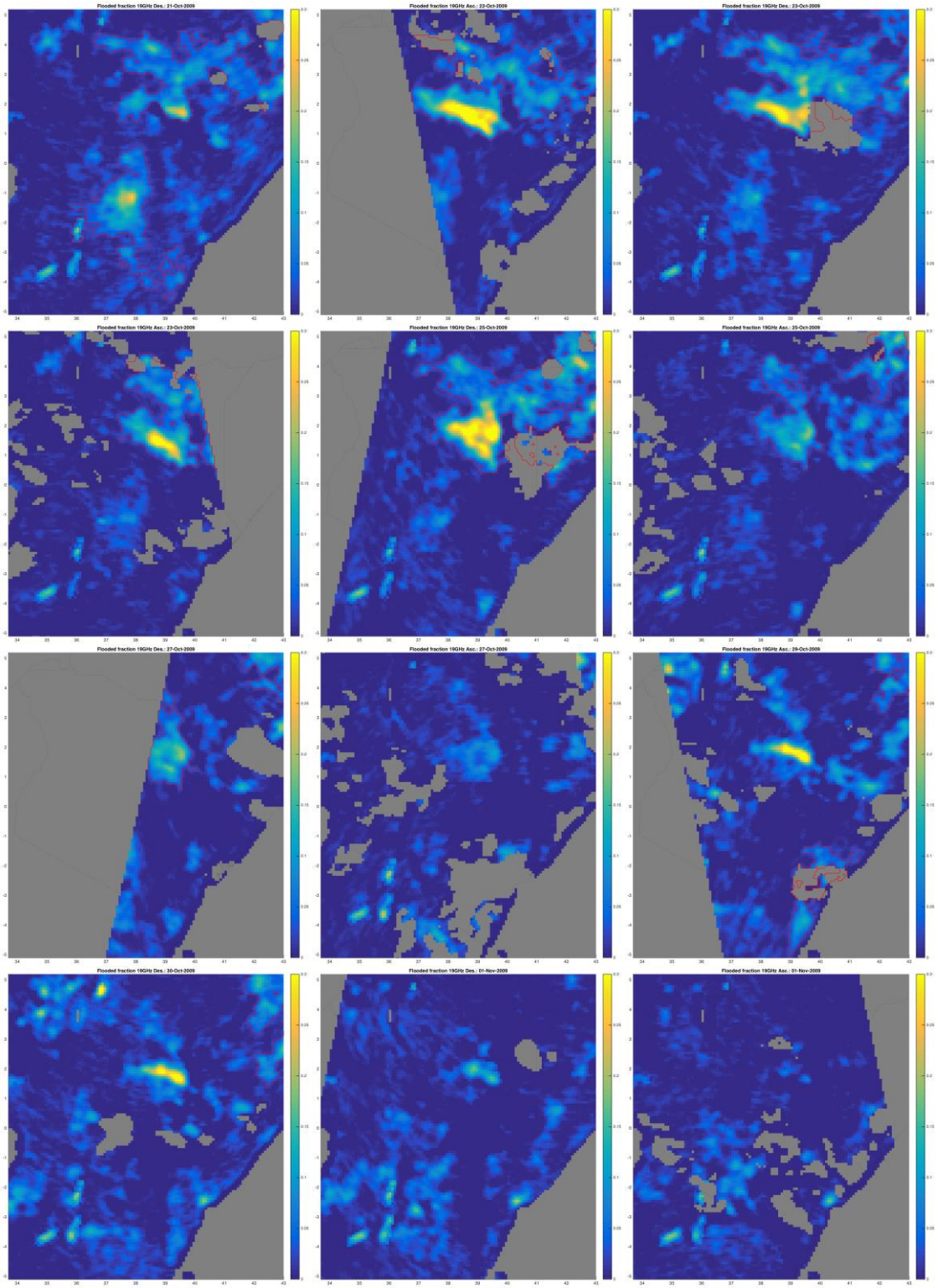
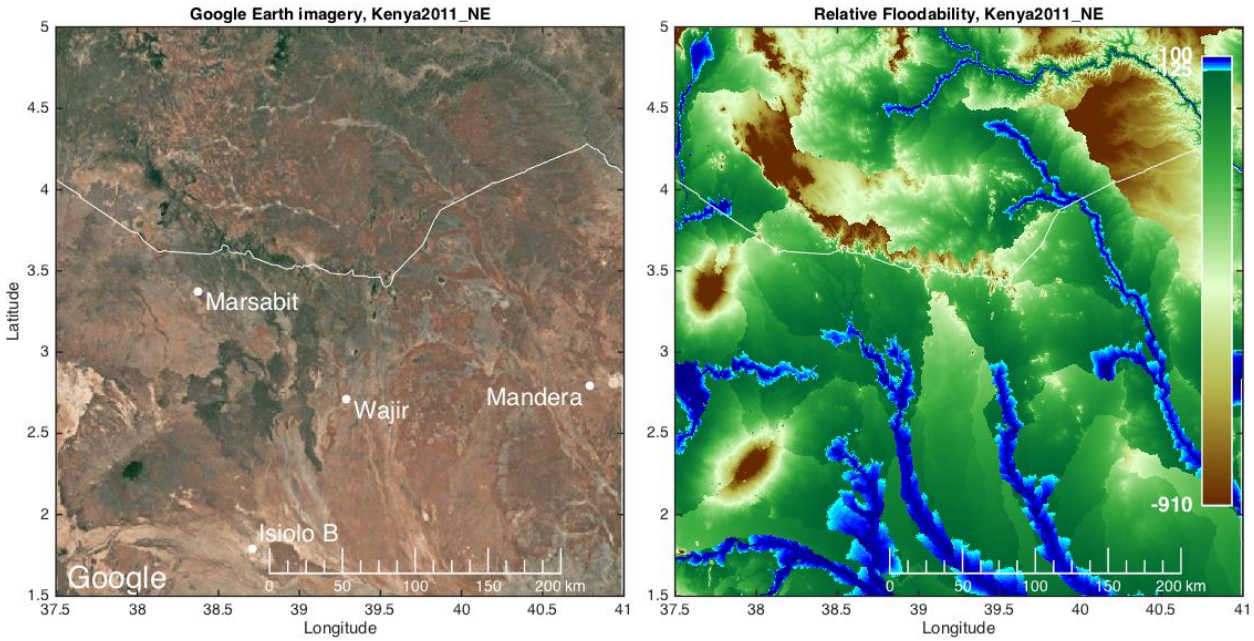


Figure 20: Single pass 19-GHz flooded fraction estimates from the microwave algorithm for selected AMSR-E satellite passes from 2009-10-21 to 2009-11-01. Red outlines indicate areas flagged by the algorithm as false positives. Gray indicates no flooded fraction data either due to the persistent water mask, missing data (i.e., between sensor swaths), or rain flagging.

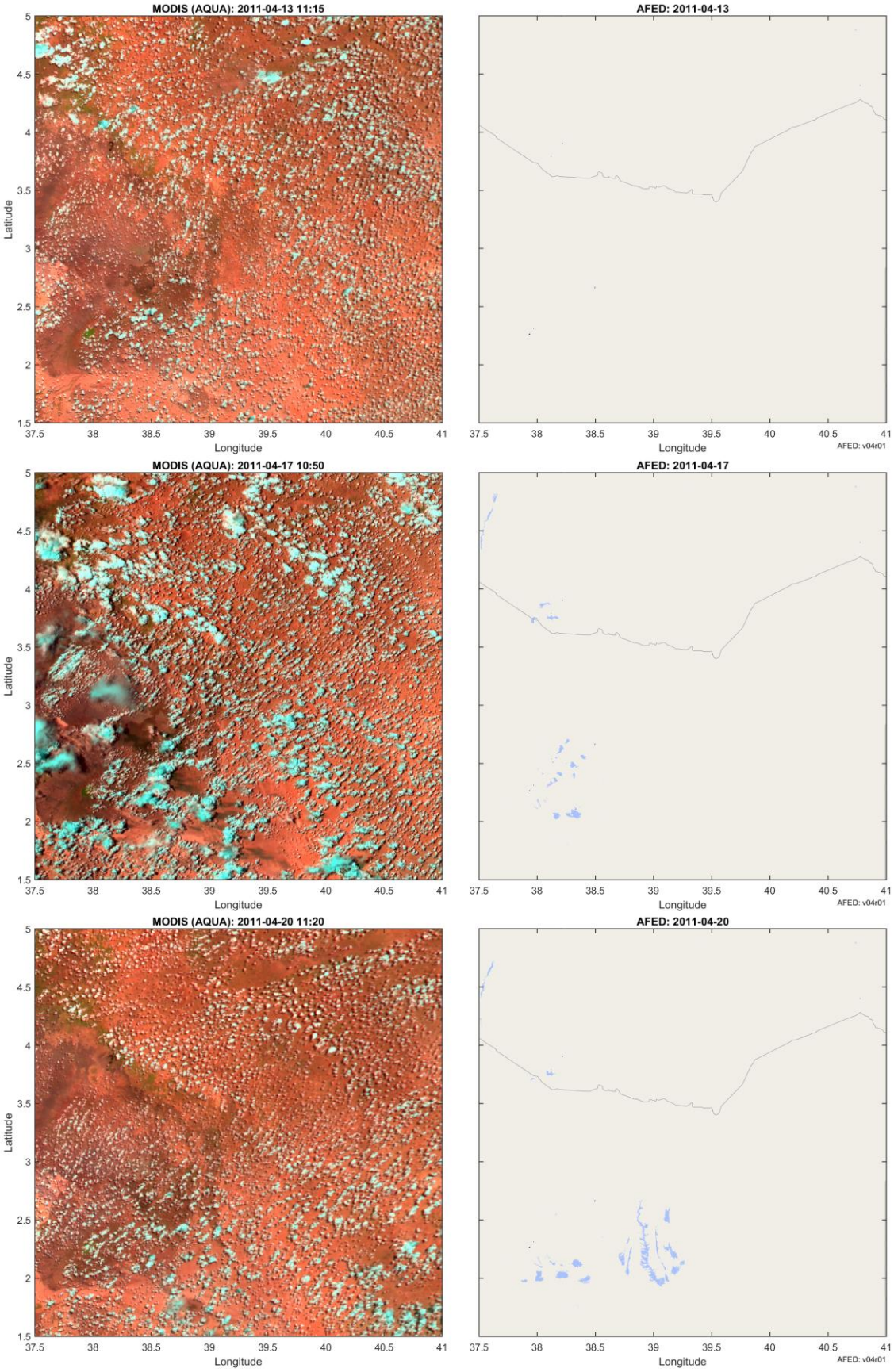
The AFED V04R00 and V04R01 algorithm significantly reduced the size and duration of apparent false flooding in this case relative to V03R01. The algorithm misses the remaining area of concern primarily because the initial appearance of elevated flooded fraction on 10-29 is preceded by two passes in which there is no data in the affected area. The algorithm does not flag false positives unless there is a recent observation against which it can compare the current conditions. Although it would have been unusual, the elapsed time between observations in this and similar cases would have been sufficient for true flooding to develop.

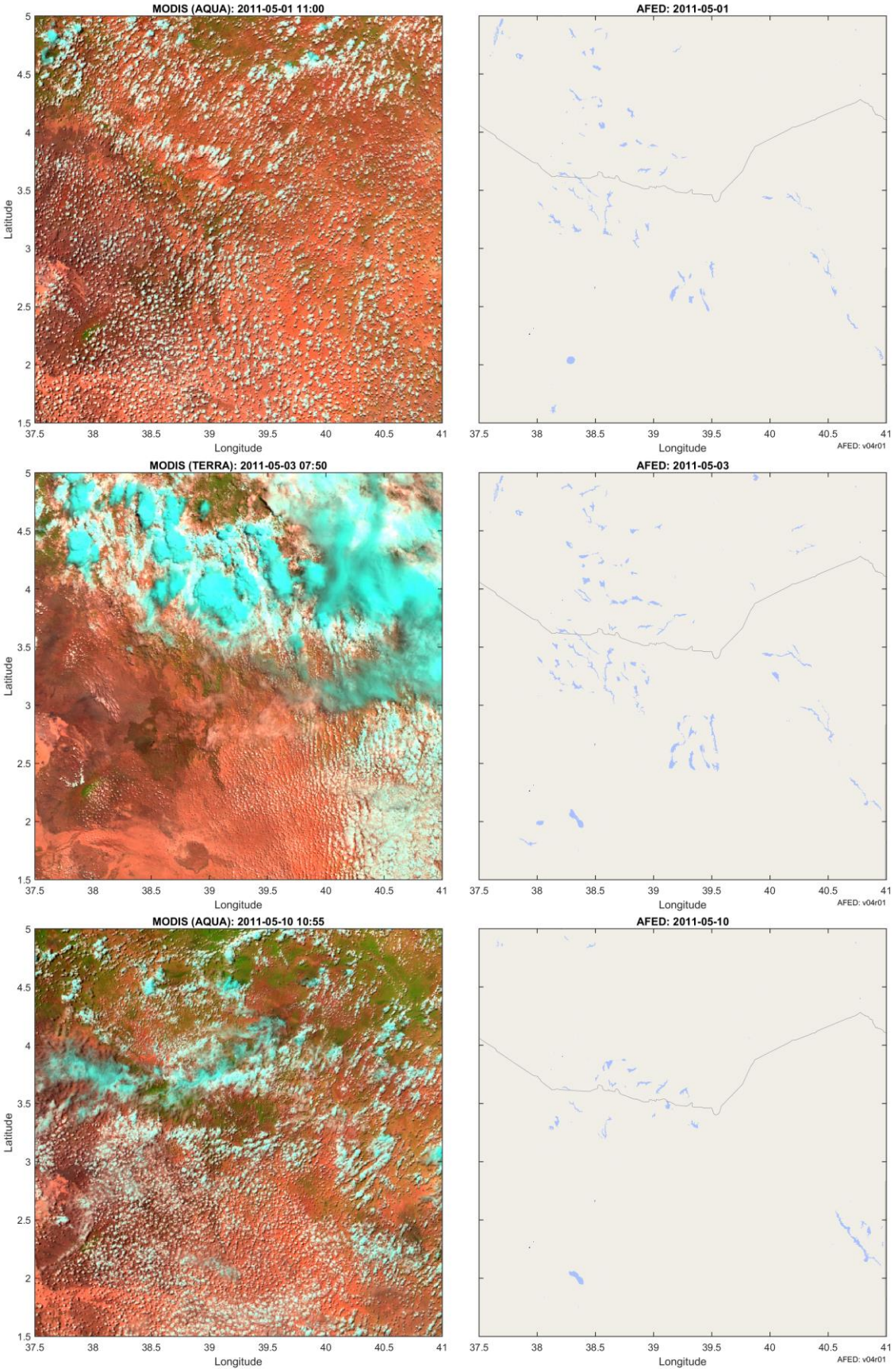
3.5.2.2 *North Eastern region, Kenya, 2011*

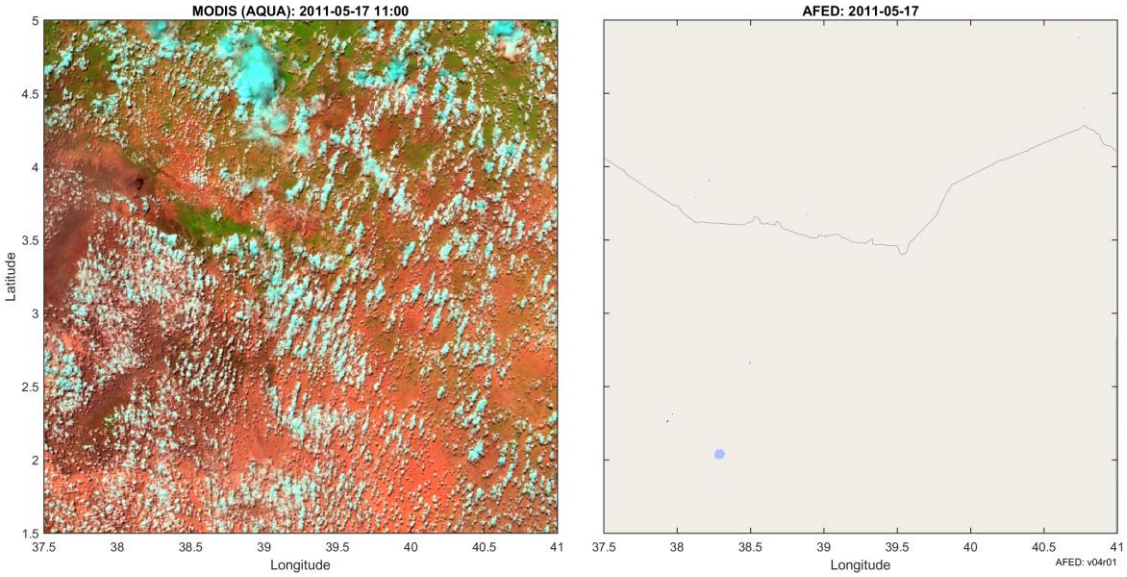
- *Period investigated*: 2011-04-01 to 2011-05-31.
- *Location map indicating microwave grid points selected for time series analyses*: Results for a subset of these grid points are given below. See Figure 16 for figure caption.



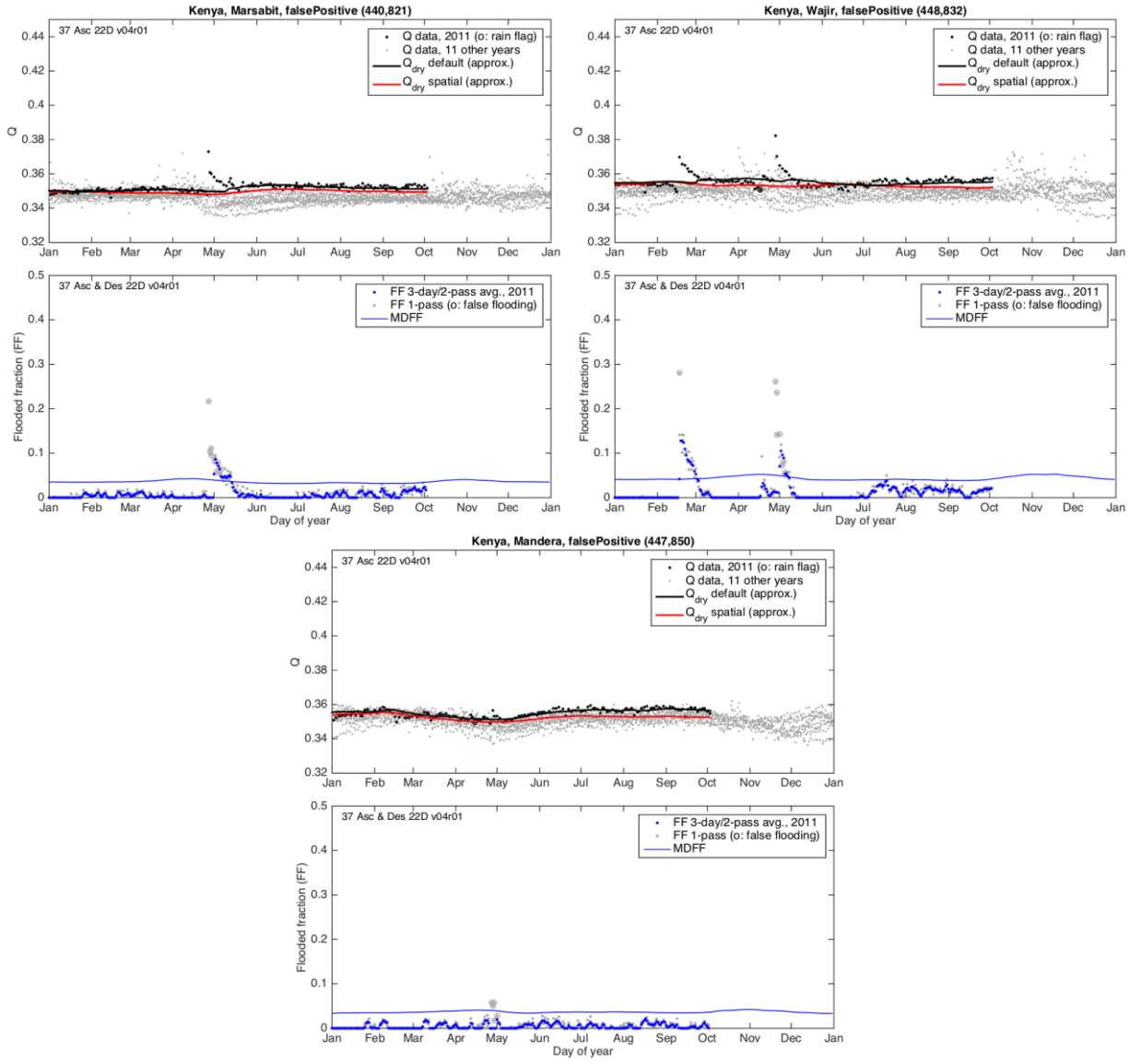
- *Side-by-side comparisons of MODIS false-color imagery to AFED*: Images were selected based on proportion of cloud free area and the degree to which the image can be interpreted to infer the timing and location of flooded or non-flooded conditions.







- *Time series analyses*: Plots are described in section 3.5.1.1. Data ends 3 October 2011 because the plot only shows data from AMSR-E, which concluded operations after that date.



Discussion:

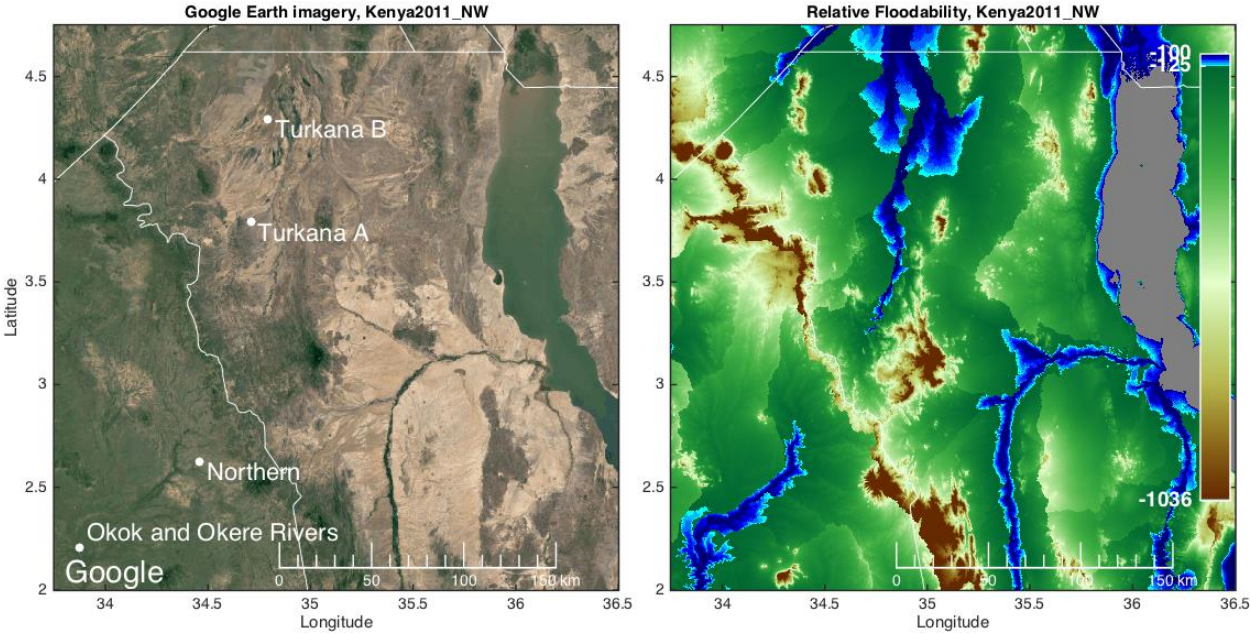
Persistent cloud cover makes this case difficult to assess definitively. The MODIS imagery on 2011-04-13 provides the best view of presumably dry conditions prior to AFED flooding. On 04-17 there is some AFED flooding in the southwestern part of the map where new dark streaks and patches are also apparent in the MODIS image relative to 04-13. The pattern of streaks suggests that passing rainstorms have left behind surface water and that is what the microwave data may be responding to. On 04-20, AFED flooding has spread to the east but it is difficult to discern any change in the imagery in this area relative to 04-13. On 05-01 and 05-03 AFED flooding become more widespread but there continues to be no sign of flooding in the MODIS

imagery. On 05-10, there is some indication in the MODIS imagery to confirm the flooding seen in the southeast part of the AFED map. By 05-17, only a small area of AFED flooding remains. .

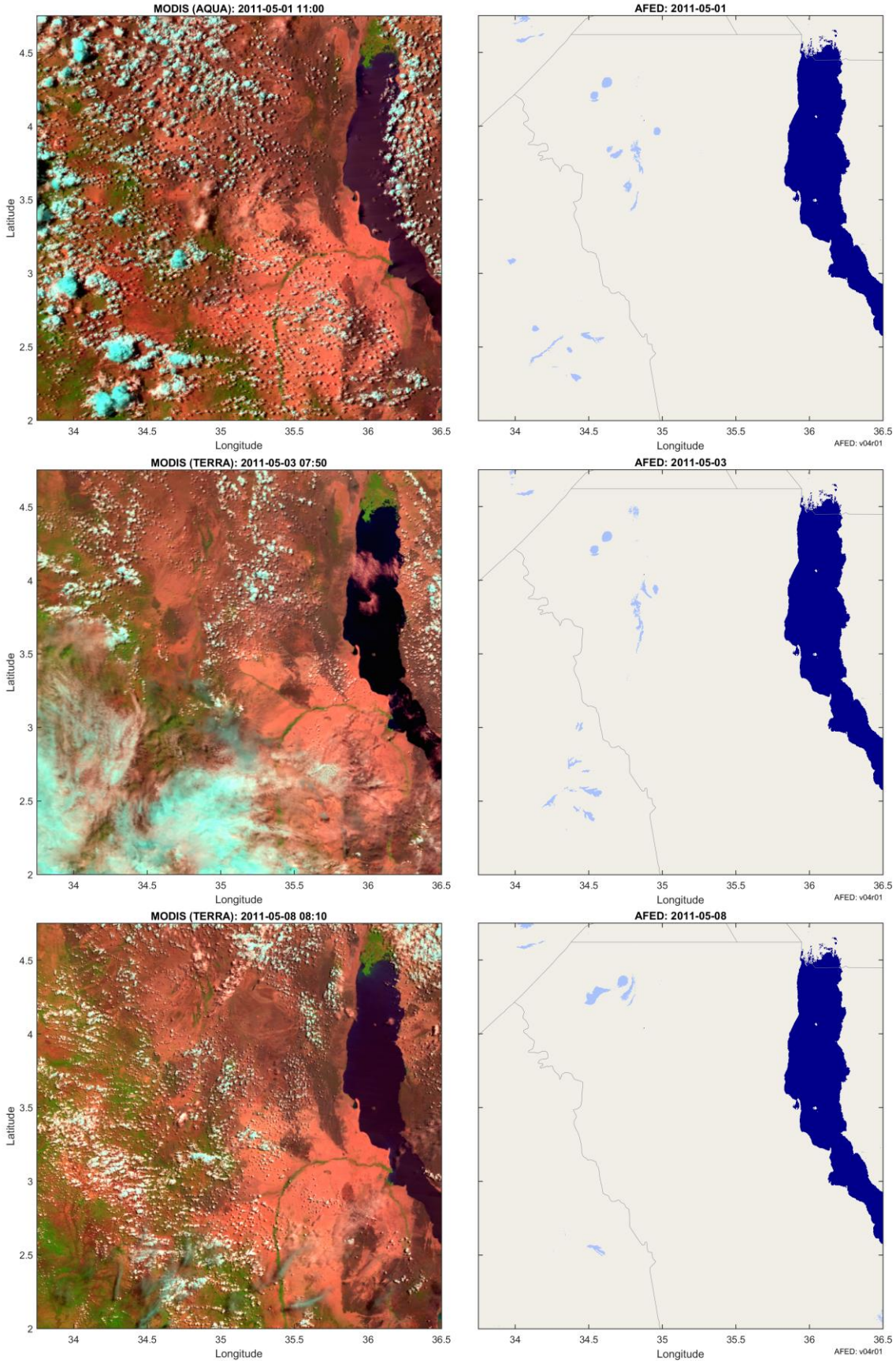
The time series plots from Marsabit and Wajir suggest that event beginning around 2011-04-29 was primarily the result of soil moisture induced by wide-scale rainfall. *Q* data jump rapidly at each point and then fall monotonically over the following days. This pattern is consistent with a rainfall event that quickly spreads surface water over a larger area followed by gradual absorption, evaporation, or collection of the water into low-lying areas. The algorithm flags many points during the study period as false positives but others do not meet the algorithm's criteria for false positive detection. See the Isiolo case for further discussion. Like the Isiolo case, it is likely that there was some flooding of dry streambeds in this area in conjunction with the wider event but there is little evidence in the MODIS imagery to confirm this inference.

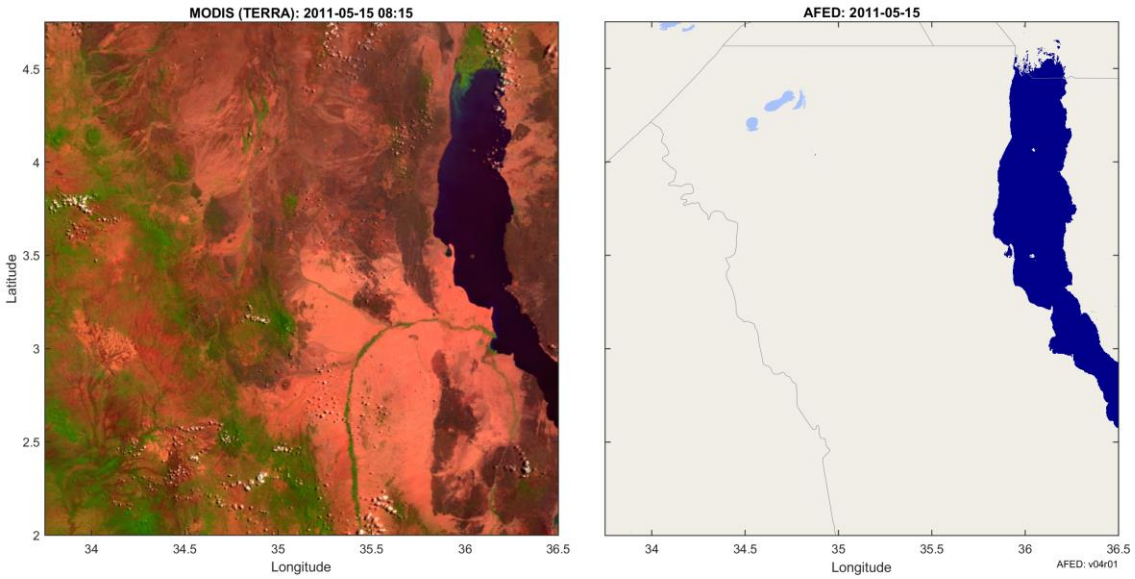
3.5.2.3 *North Western region, Kenya, 2011*

- *Period investigated*: 2011-04-27 to 2011-06-01.
- *Reports*: Some flooding was reported in the region around the end of April including at Lodwar, Lokichoggio, and Kakuma (https://kanere.org/2011/08/21/kakuma-droughts-and-floods/).
- *Location map indicating microwave grid points selected for time series analyses*: Results for a subset of these grid points are given below. See Figure 16 for figure caption.

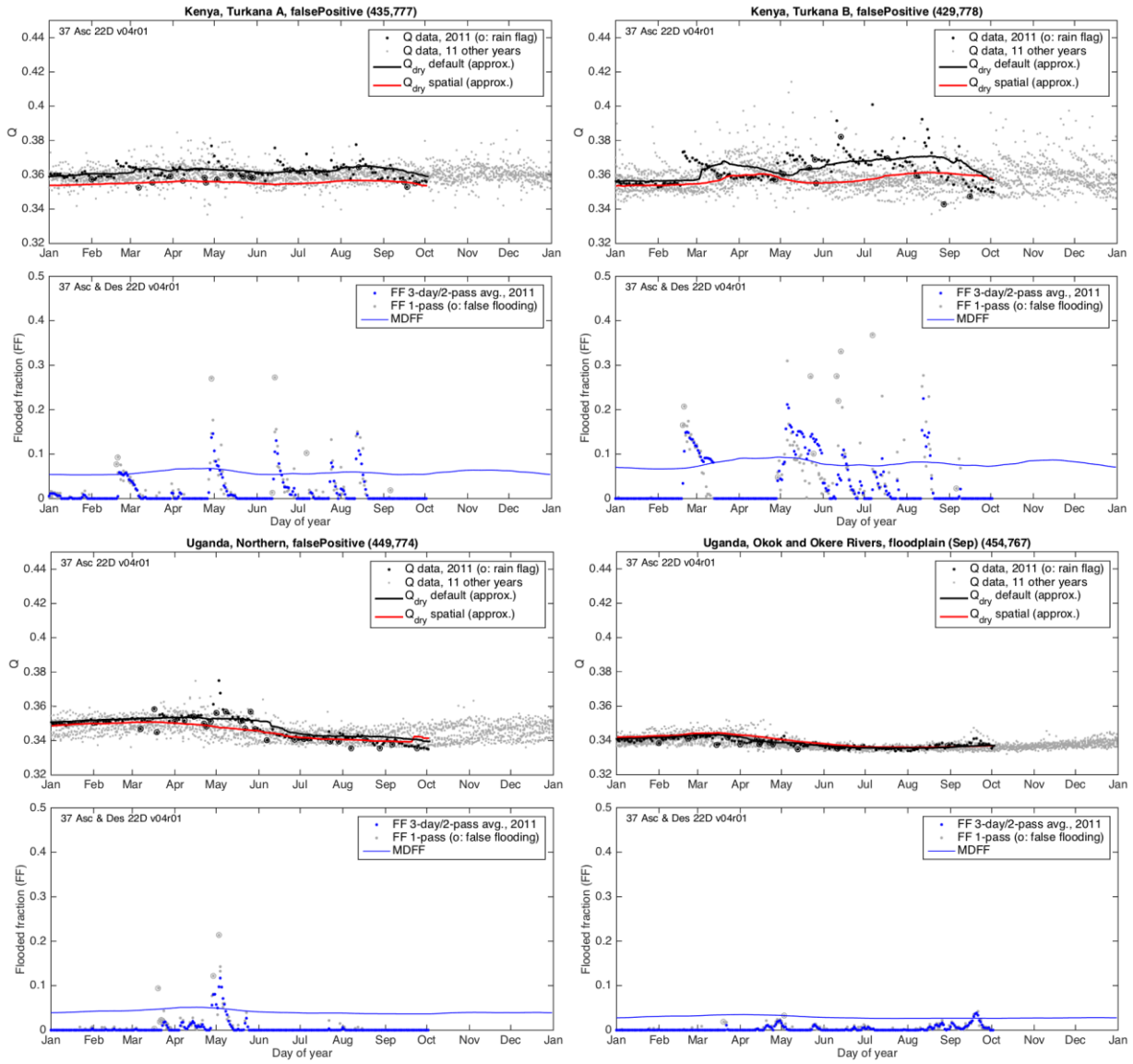


- *Side-by-side comparisons of MODIS false-color imagery to AFED*: Images were selected based on proportion of cloud free area and the degree to which the image can be interpreted to infer the timing and location of flooded or non-flooded conditions.





- *Time series analyses*: Plots are described in section 3.5.1.1. Data ends 3 October 2011 because the plot only shows data from AMSR-E, which concluded operations after that date.



Discussion:

Heavy cloud cover affected most of the MODIS images during the period. The 2011-05-01 image is relatively clear but it is difficult to discern signs of flooding that corroborate the AFED map. However, the report cited above indicates that there was flooding in Kakuma, which is in the area near Turkana A that AFED shows as flooded on 05-01. Cloud cover on 05-03 prevents confirmation of AFED flooding on that date in nearby areas of Uganda around the Northern point.

The time series plots indicate that microwave *Q* data increased rapidly and simultaneously around 04-29 at points Turkana A and B and Northern. As discussed in the Isiolo case, the temporal *Q* profile with rapid onset and gradual decline is attributable to large-scale soil moisture brought about by rainfall and amplified in *Q* by the lack of vegetation cover. This same pattern can be seen several times during the year at Turkana A and B. This evidence and the pattern of dry streams seen in the Google Earth imagery suggest that the region may experience frequent

flash floods. It is likely that this event is not entirely a false positive but is instead a flash flood event for which AFED overestimated total flood extent due to barren conditions.

3.5.2.4 *Coastal flooding, Cote d'Ivoire, Ghana, and Namibia, multiple years*

- *Summary of comments from ARC:* [V03R01] AFED shows recurrent flooding in the coast zone in Côte d'Ivoire and Ghana and there are similar patterns in Namibia. Cumulative flooded pixels (*data footprints* in the Flood Explorer) show the following:
    - West Africa: The [V03R01] data footprint over all years (Figure 21) indicates the recurrent flooding in coastal zones is present at larger scale. There is a lot of flooding in the coastal area in the early years of the SSM/I 50 km resolution data (mainly 1992, less in the period 1993-1997). The coastal zone of Senegal and Guinea (not evaluated here) keeps flooding through the 2002-2015 data.
    - Namibia: The [V03R01] all years footprint shows lots of flooding in the coastal area caused by the first years of 50 km SSM/I data (Figure 22). When each year is examined, there seems to be a spatial pattern. However, from 1996 onward the phenomenon is still visible, but to a much smaller extent.

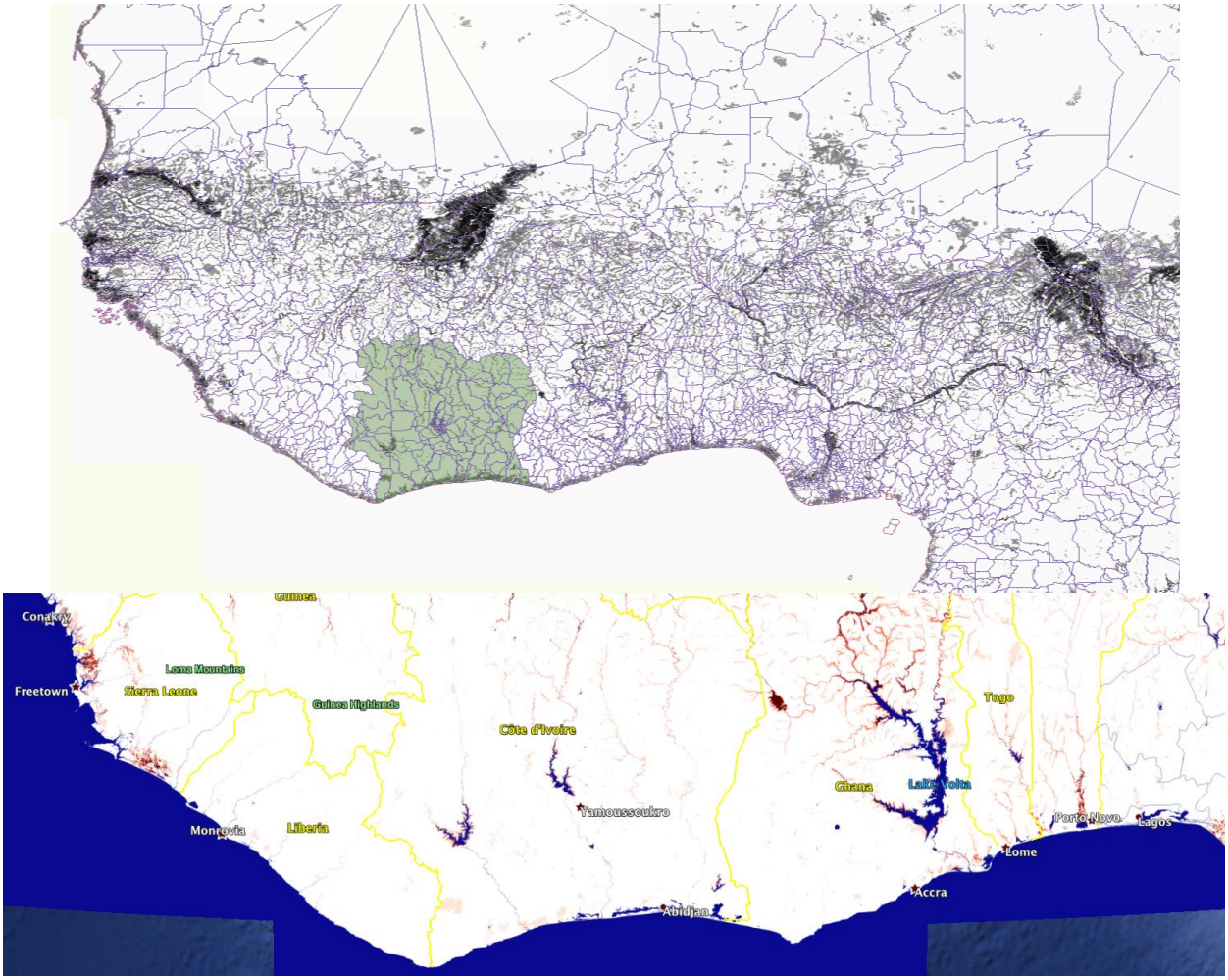


Figure 21: West Africa, all years footprint V03R01 from ARC (top) and V04R01 (bottom).

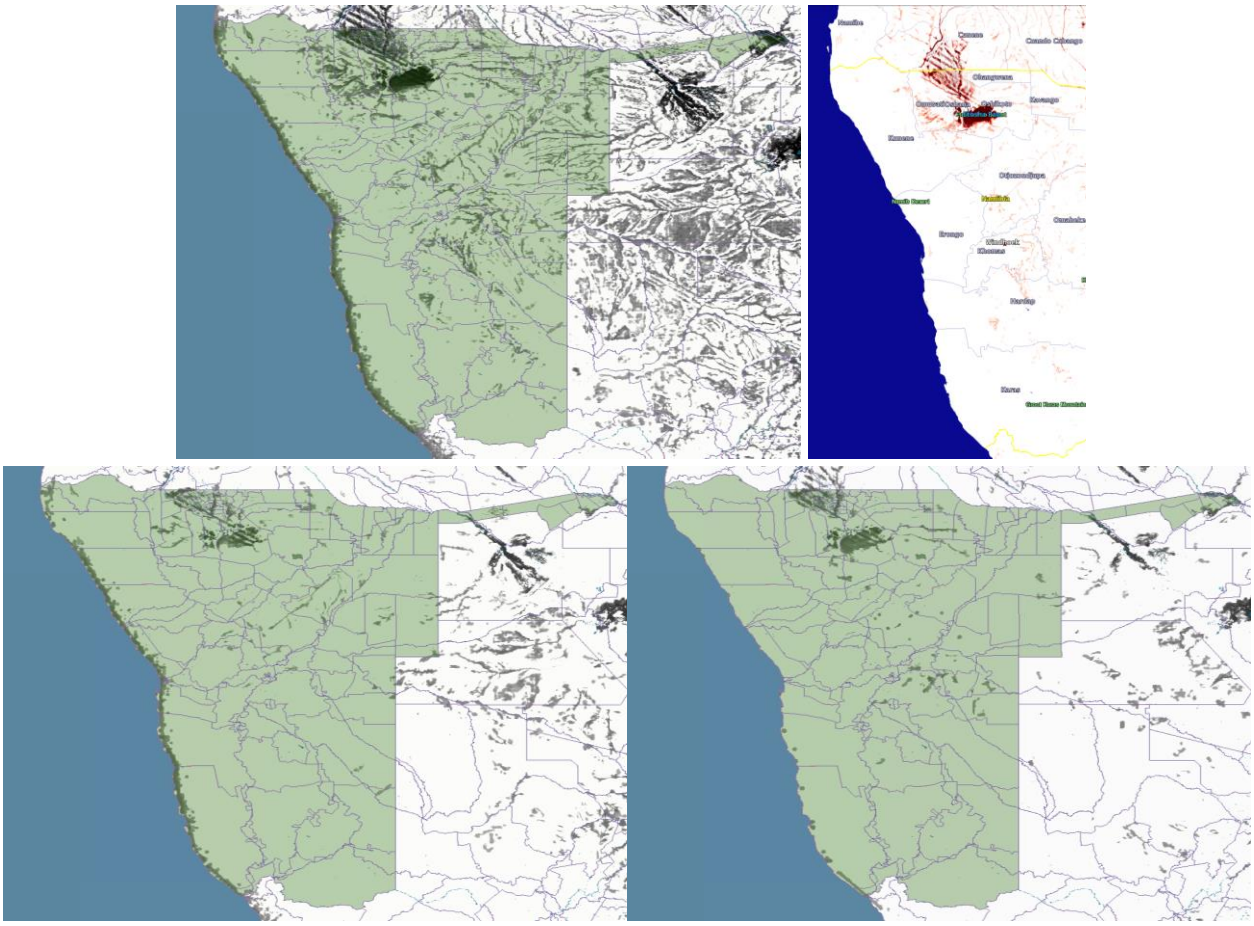


Figure 22: Namibia AFED footprints. Top: all years V03R01 from ARC (left) and V04R01 (right). Bottom: V03R01 1992 (left) and 1996 (right)

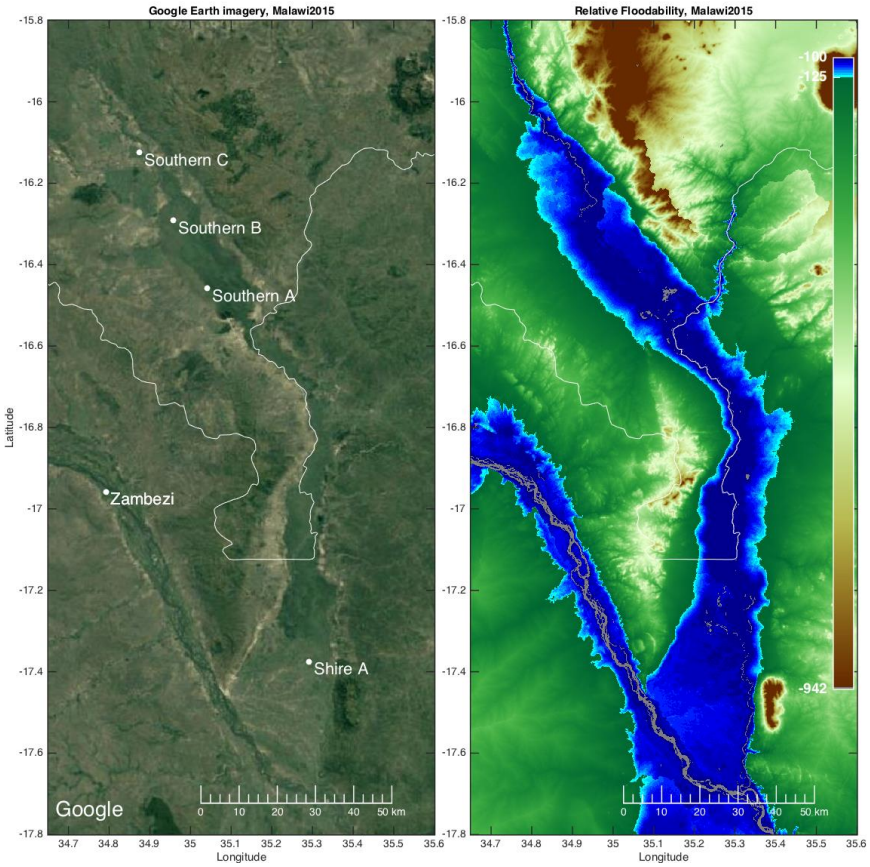
- *Analysis*:
  - For AFED V03R01, we investigated these areas using (1) sets of microwave data and flooded fraction time series plots at coastal grid points for each year 1993-2001 (SSM/I), 2003-2010 (AMSR-E), and 2013-2015 (AMSR2) and (2) total number of days flooded per year and total days flooded less isolated single flooded days and less isolated single days and 2-day pairs.
  - As of AFED V04R01, changes to the AFED algorithm have completely eliminated persistent coastal false positives in these areas. Total V04R01 AFED footprints over all years (Figure 21 and Figure 22) show no pattern of AFED flooding along the coastlines and few apparent false positives along the coast at all. Several types of algorithm changes contributed to the improvements, including: (1) changes to the false positive detection algorithm for spatial constraints near coastlines and dilation of detection regions over rain-flagged areas, (2) changes in MDFF formulation so that spatial smoothing is not applied near coastlines, (3) geolocation error correction, and (4) multi-day flooded fraction algorithm logic that keeps false positives from persisting in the absence of new data.

3.6 Flood duration test results for AFED V0401

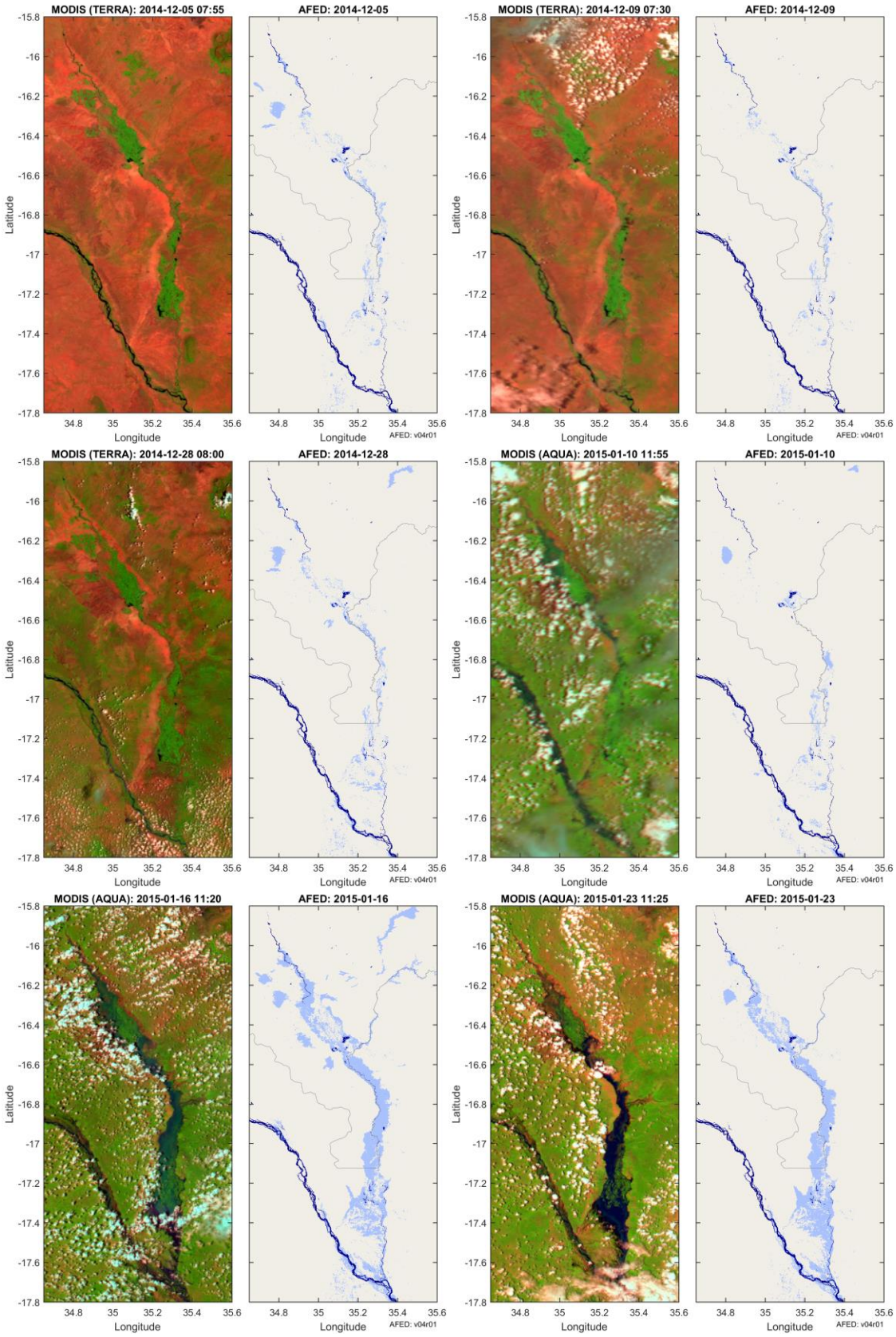
3.6.1 *Shire River, Malawi and Mozambique, 2015*

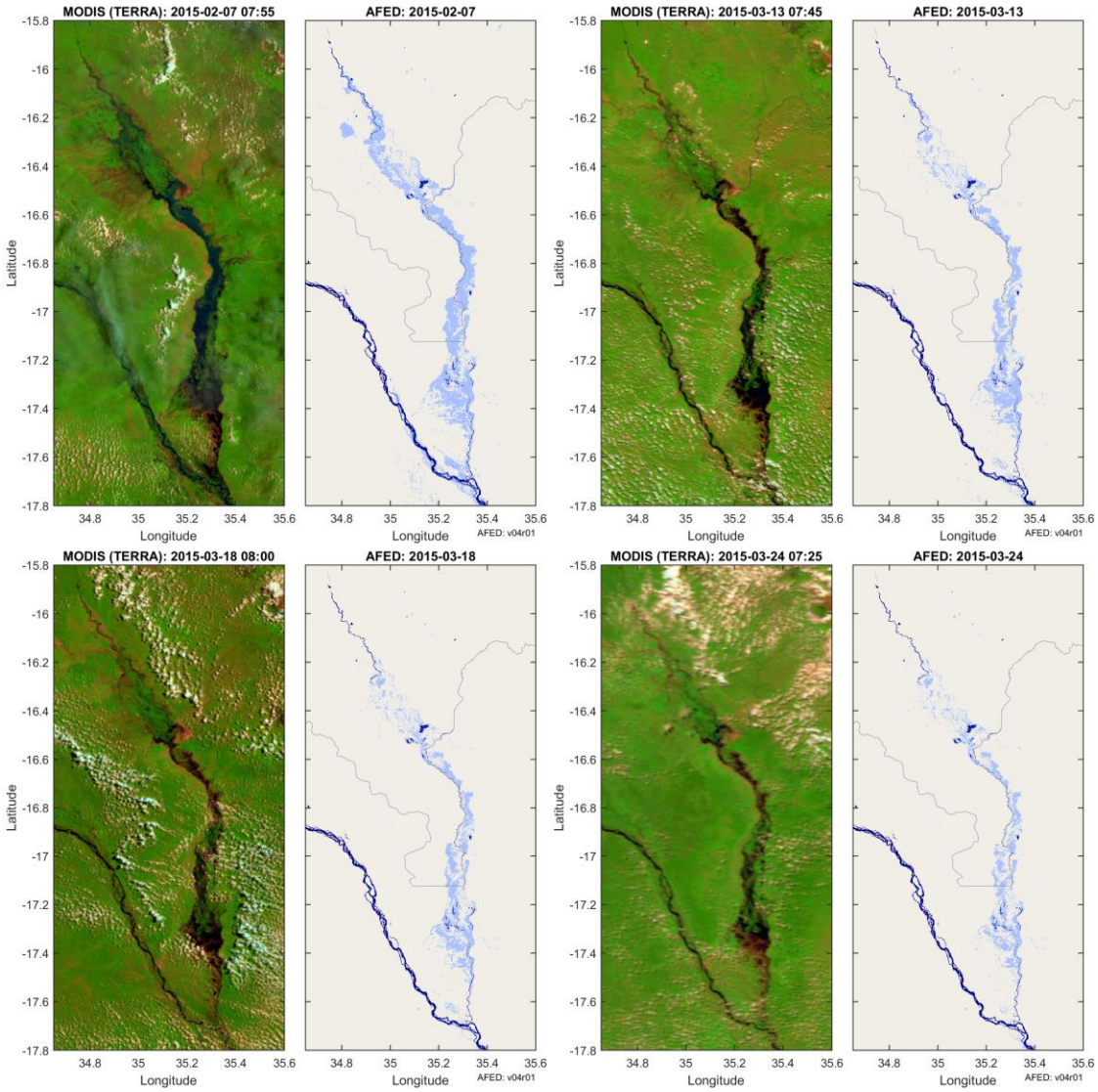
- *Event information*:
  - http://earthobservatory.nasa.gov/NaturalHazards/view.php?id=85192&eocn=image&eoci=related_image: In January 2015, several bouts of heavy rain triggered widespread flooding in southern Africa. The Advanced Land Imager (ALI) on the EO-1 satellite captured an image of this flooding along the Shire River in Malawi and Mozambique on January 26, 2015.
  - https://www.theguardian.com/global-development/2015/feb/10/malawi-floods-devastation-far-worse-than-first-thought [10 Feb. 2015]: Nearly a quarter of a million people, more than originally thought, have been affected by the devastating floods that ripped through Malawi [in January 2015], and with rains still falling, many of the 230,000 who were forced to flee their homes have been unable to return and rebuild their lives, the UN said. Unicef has launched a $9.3m appeal to cover its emergency response for three months. Northern and central regions of neighbouring Mozambique have also been hit by flooding that has affected around 160,000 people and entire communities. Unicef said the flooding had cut off roads, disrupted power supply and destroyed bridges, houses and schools.
  - http://floodlist.com/africa/floods-malawi-facing-worst-food-crisis-10-years [29 Sept. 2015]: The worst floods in living memory struck in Malawi between January and March this year. Around 230,000 people were displaced and 276 reported dead or missing. In a statement last week, the United Nations' World Food Programme (WFP) say the floods washed away food stocks, and ruined fertile land and that the country now faces the worst food crisis in 10 years.

- *Location map indicating microwave grid points selected for time series analyses*: Results for a subset of these grid points are given below. See Figure 16 for figure caption.

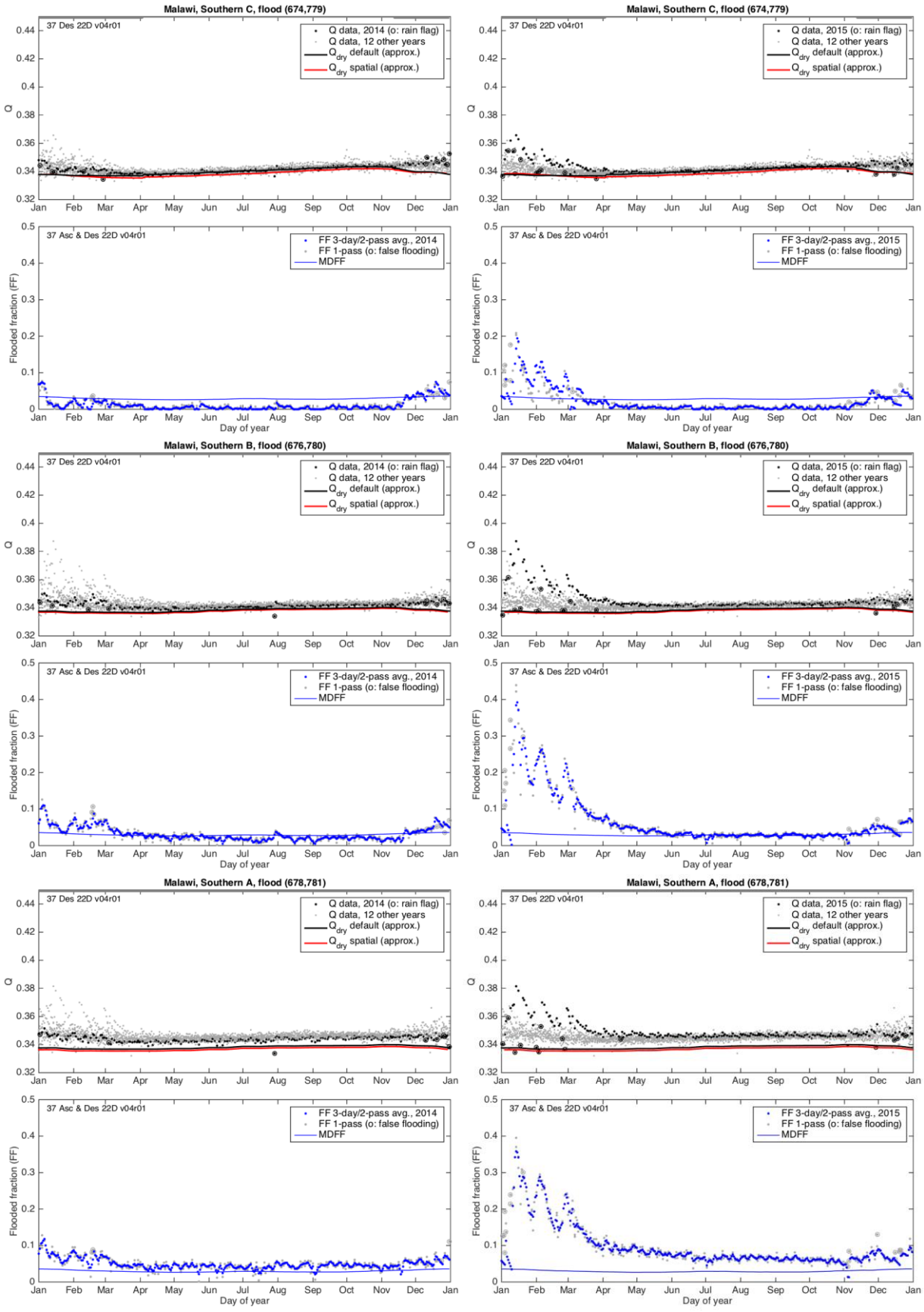


- *Side-by-side comparisons of MODIS false-color imagery to AFED*: Images were selected based on proportion of cloud free area and the degree to which the image can be interpreted to infer the timing and location of flooded or non-flooded conditions.





- *Time series analyses*: [Following page] Plots are described in section 3.5.1.1. Left-hand plots are year 2014 and right-hand plots are year 2015.



Discussion:

The time series plots at points Southern C and Southern B indicate that *Q* begins to increase in mid Dec. 2014; prior to that time *Q* is constant (Southern B) or slowly increasing (Southern C). In mid Nov. and early Dec., Q_{dry} decreases in response to seasonal NDVI changes. (The AFED algorithm flags both points as wetlands, which means that Q_{dry} is estimated from historical average NDVI data. See the ADD for more details.) Flooded fraction, FF, exceeds MDFF for the first time on about **25 Nov. at B and C**, indicating the start of AFED flooding. However, AFED flooding is temporarily disrupted at both points in early January by false flood detection (gray circles on the FF plot), indicating that the period included rain events over large enough areas to trigger the false positive detection algorithm. FF is variable through the rest of Jan. and Feb. but does not finally falls below MDFF until around **15 Mar. at point C and 20 June at point B**, indicating the end of AFED flooding. The slow reduction in FF at point B suggests that this area is influenced by slow-draining wetlands.

Southern A is south of B and C in another area that the AFED algorithm marks as wetlands and FF at point A is at or above MDFF through most of 2014 and 2015. *Q* begins to rise rapidly at point A in early January 2015 but the rise in FF is disrupted by the same false flooding flags at at points B and C. Compared to the 11 other years in the *Q* record, 2014 was a relatively dry year (lower *Q*) at point A and 2015 was wetter (higher *Q*), which is an indication that the Jan. 2015 flood event may have replenished wetlands water levels. The relative floodability map indicates that point C covers a relatively narrow flood plain while B and A are in relatively wide and flat areas conducive to wetlands formation. For this reason, we will focus on analyzing flood duration at point C, which has flood start and end points more distinctive of large river flooding than the wetlands-affected points B and A.

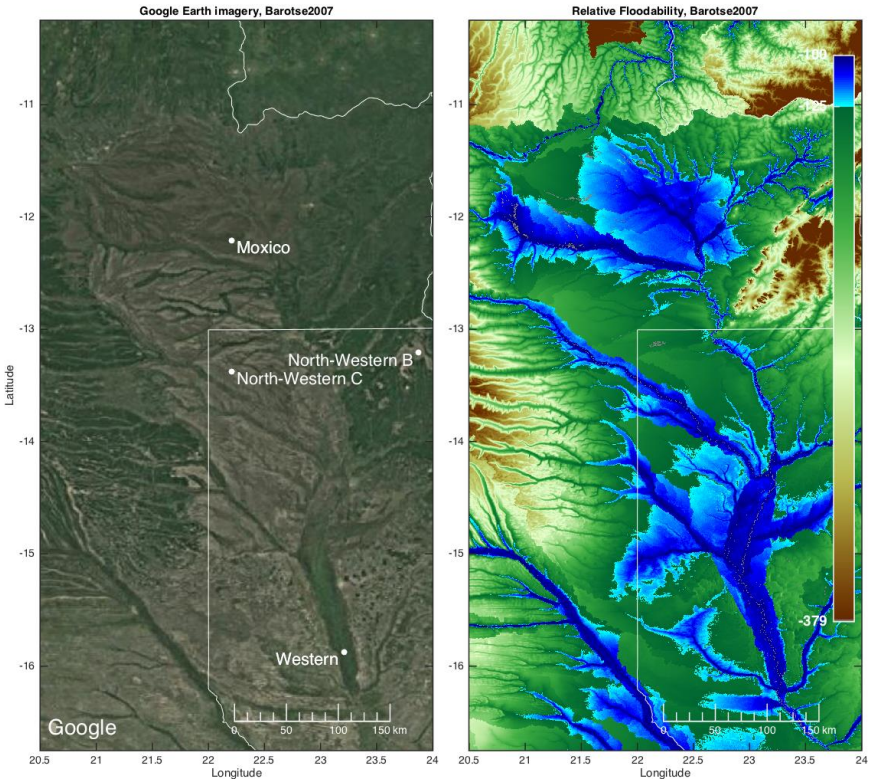
MODIS imagery provides general context for AFED results along the Shire River but the imagery sequence can at best only validate the approximate start, peak, and end points of flooding because of cloudiness and ambiguity in image interpretation:

- **5-9 Dec.**: MODIS imagery shows **patches of water**, some of which are also included in the AFED water mask. AFED indicates the presence of additional water in the wetlands areas near, along, and south of the Malawi-Mozambique border and also patches off the flood plain to the northwest on 5 Dec, which may be false positives or residual water from recent rains.
- **28 Dec.**: **Little to no flooding** in MODIS imagery although rain and false positive "flashes" (indicative of soil moisture) detected in the microwave data time series at points A-C suggest that water has been entering the basin.
- **10 Jan.**: This is the first clear-sky MODIS image of 2015. The image indicates **flooding has begun** in the north end of the mapped area but not the south, confirming the plausibility of AFED point C flooding beginning on Dec. 28.
- **16-23 Jan.**: MODIS indicates **peak flooding** in the northern then southern areas, consistent with AFED results.
- **7 Feb. – 13 Mar.**: **Flooding appears to recede** in the MODIS imagery. AFED flooding decreases significantly over this period at point C.
- **18-24 Mar.**: **Flooding is no longer apparent around point C** in MODIS imagery and AFED flooding in the area has ended. By 24 March most of the floodwater elsewhere has apparently receded except for the lower-lying wetland areas.

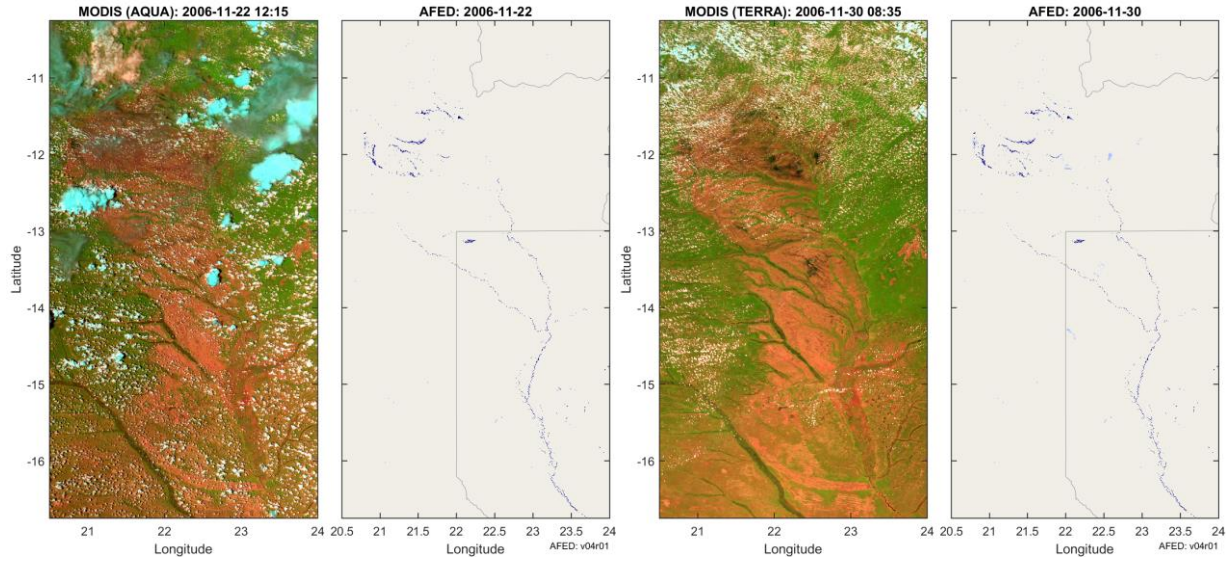
3.6.2 *Upper Zambezi River, Angola and Zambia, 2007*

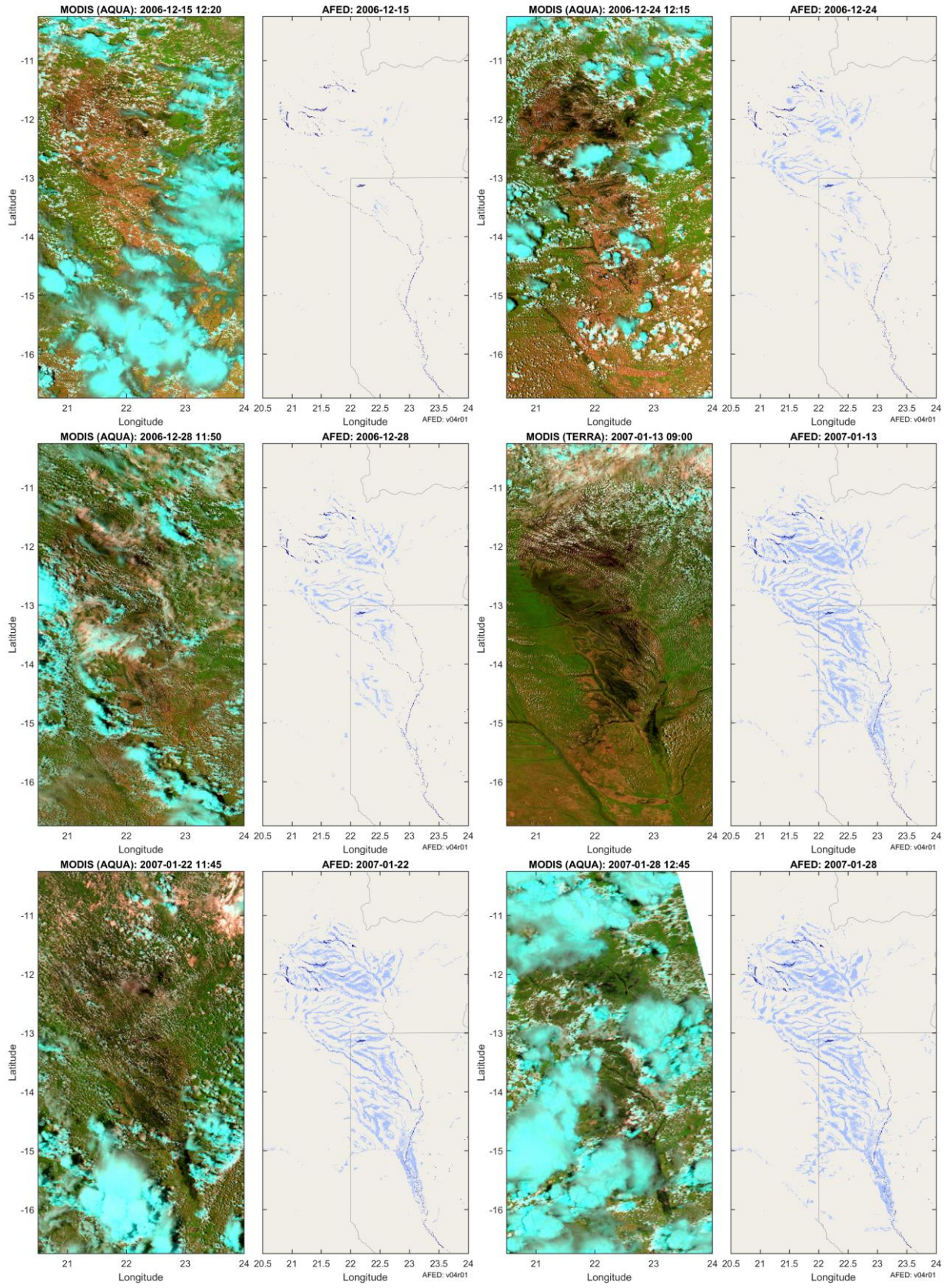
- *Event information*:
  - https://en.wikipedia.org/wiki/Barotse_Floodplain: The **Barotse Floodplain** also known as the **Bulozi Plain**, **Lyondo** or the **Zambezi Floodplain** is one of Africa's great wetlands, on the Zambezi River in the Western Province of Zambia. The peak of the flood occurs on the floodplain about 3 months after the peak of the rainy season in January–February. The flood usually peaks in April, and recedes in May to July, when grasses quickly grow on the exposed plain. At the river's lowest water in November the floodplain still contains about 537 km² of lagoons, swamps and channels.
  - http://www.dartmouth.edu/~floods/wetlands/Barotse.htm: Data from MODIS and Quikscat show seasonal flooding in January-July and November-December 2002.
  - http://visibleearth.nasa.gov/view.php?id=65889 [Flooding in 2003]: In late March 2003, heavy rains in Zambia led to floods that displaced over 10,000 people and destroyed thousands of valuable acres of farmland. The most devastating floods occurred in the southern part of Zambia, and can be seen in the image from March 27, 2003. The loss of cattle and crops in the floods has exacerbated a food shortage due to recent droughts. The food shortage could potentially affect over two million people in the small country. During the dry season there is hardly any standing water in southern and western Zambia.
  - http://earthobservatory.nasa.gov/NaturalHazards/view.php?id=50158 [Flooding in 2011 compared to 2010].

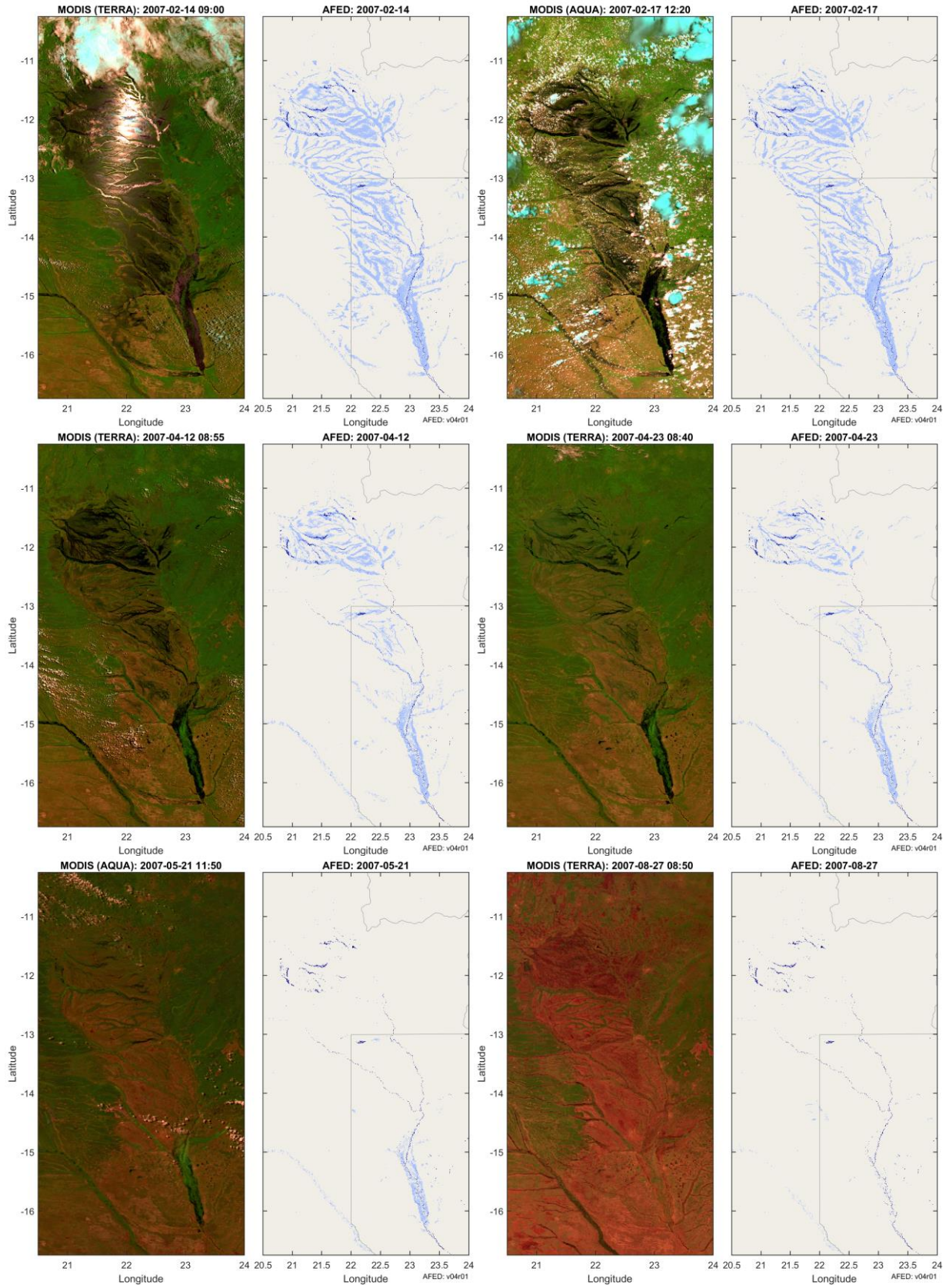
- *Location map indicating microwave grid points selected for time series analyses*: Results for a subset of these grid points are given below. See Figure 16 for figure caption.

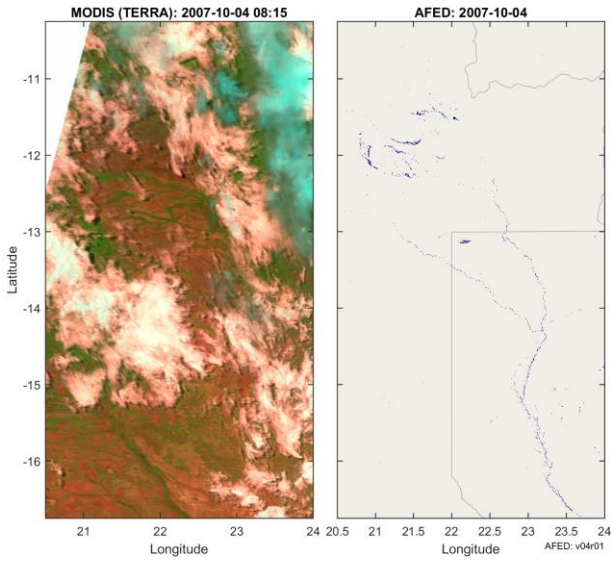


- *Side-by-side comparisons of MODIS false-color imagery to AFED*: Images were selected based on proportion of cloud free area and the degree to which the image can be interpreted to infer the timing and location of flooded or non-flooded conditions.

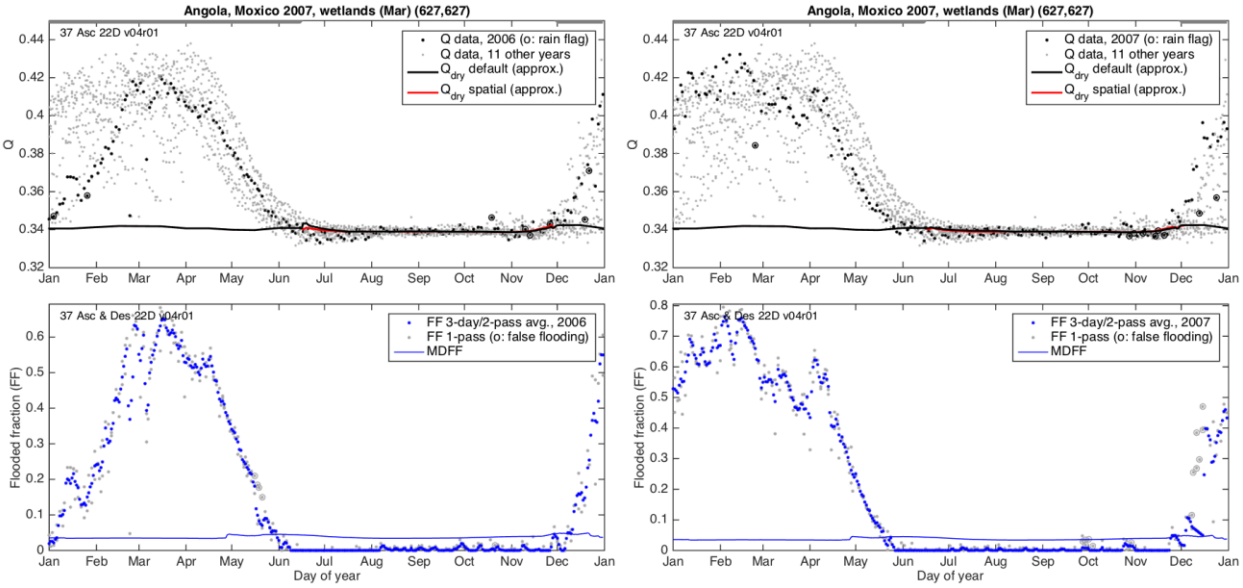


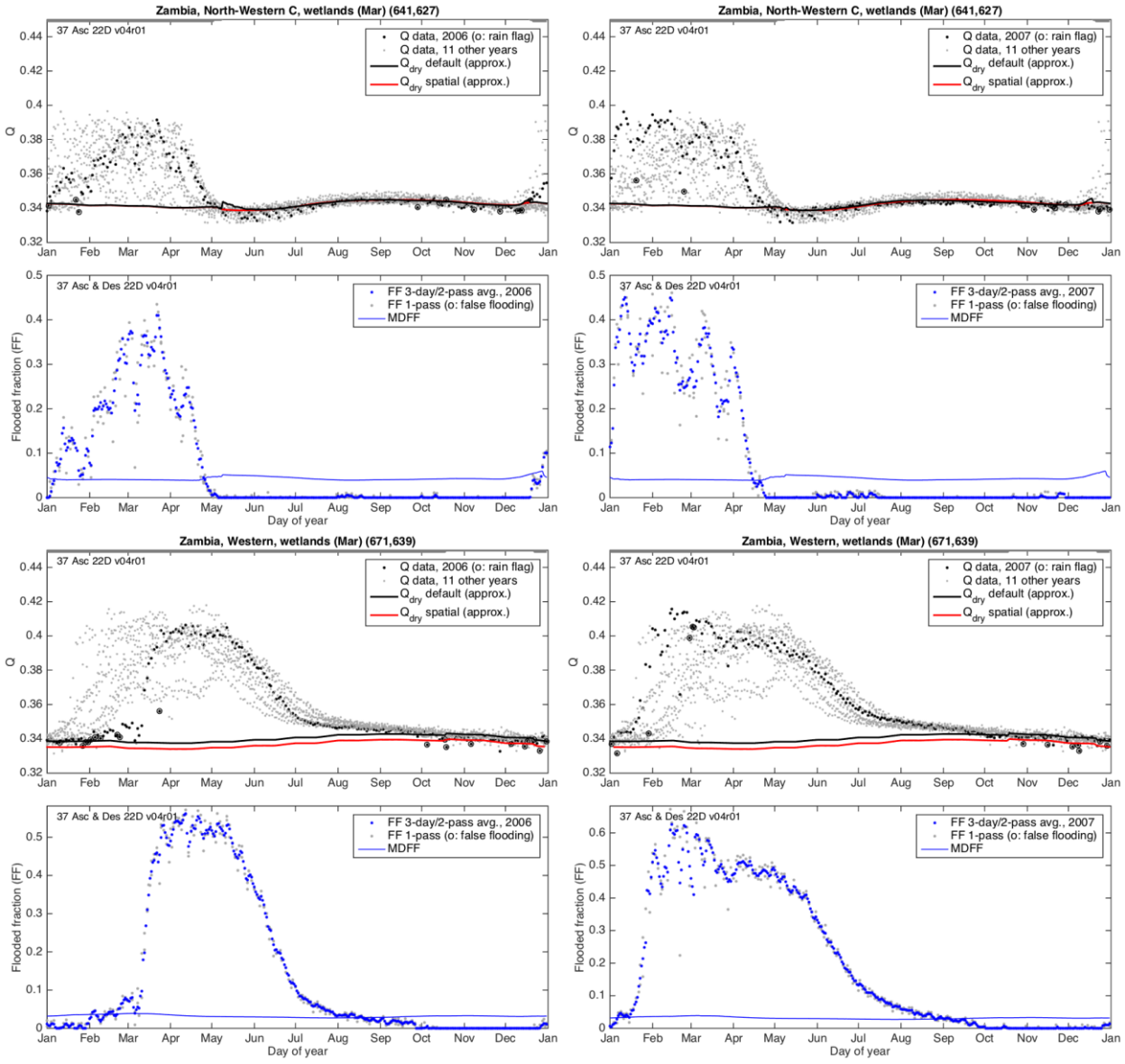






- *Time series analyses*: Plots are described in section 3.5.1.1.





Discussion:

From the time series plots we can infer the following start and end dates for flooding (i.e., FF greater than MDFF):

- Moxico, Angola: Start around 2006-12-07 and end around 2007-05-20.
- North-Western C, Zambia: Start around 2006-12-27 and end around 2007-04-16.
- Western, Zambia: Start around 2007-01-06 and end around 2007-09-01.

Compared to the 11 other years plotted, the *Q* data suggest that 2007 flooded fraction was at or near the maximum value observed at these locations at some point in the flood season.

From the image sequence we observe the following:

- 2006-11-22 (best clear-sky image prior to flood season): Dry conditions throughout the area.
- 2006-11-30 (first clear-sky image with obvious flooding): Apparent flooding around 12N latitude with some AFED flooding in the same areas.

- 2006-12-15 (first confirmation of flooding at Moxico): Widespread flooding can be seen through a clearing around the Moxico area.
- 2006-12-24: MODIS indicates flooding is more widespread in Angola and beginning to become more widespread in Zambia but there is no significant flooding obvious yet at points North-Western C or Western.
- 2006-12-28: AFED flooding at point North-Western B has begun and darkening in the area in the MODIS image provides partial confirmation.
- 2006-01-13: MODIS image confirms flooding at North-Western C
- 2007-01-22: We can infer flooding at Western by noting the continuing darker tone of the image along and around the path of the Zambia River through the Barotse flood plain.
- 2007-01-28: Flooding at point Western has increased substantially and the full width of the Barotse flood plain appears to be flooded.
- 2008-02-14 and 2008-02-17: These MODIS images correspond to near-peak flooded fraction at Moxico, North-Western C, and Western, confirming AFED peak-flood results. The sun glint on 2008-02-14 provides an alternative way of visualizing the extent of flooding that confirms the interpretation of darker areas as flooded in false color images generally.
- 2007-04-12 and 2007-04-23: These images bracket the end of AFED flooding at North-Western C and show a significant reduction in flood extent at that location while Moxico and Western clearly remain flooded.
- 2007-05-21: At the end of AFED flooding at Moxico the MODIS image suggests some residual water cover consistent with flooded fraction levels below the MDFF. Flooding is still apparent at Western.
- 2007-08-27: This image is typical of the images in the period since about 2007-07-01 when it became difficult to distinguish the Barotse Plain from the surrounding terrain. Comparison to later images (e.g., 2007-10-14) shows that some water remains on 2007-08-27 in pools and streambeds that will recede later in the season. This provides indirect confirmation of the 2007-09-01 end date of flooding at Western.

4 Performance summary and recommendations

A summary assessment of overall AFED algorithm performance can be inferred from the individual case studies and broader statistics presented here. During flood events, the algorithm performs best when flooding covers significant parts of more than one 22- (for AMSRX) or 50-km (for SSM/I) microwave data footprint. Flood mapping and false positive performance is also better when there is moderate but not dense vegetation cover that obscures bare soil areas but does not significantly obscure floodwaters.

The ability of the AFED algorithm to detect smaller floods is highly dependent on the microwave sensor resolution. The smallest detectable flood is approximately the product of MDFF and the area within a circle whose diameter is 1.25 times the microwave footprint spatial resolution. Resolution is 50 km from 1992 to 2002 and from October 2011 to August 2012 (i.e., SSM/I sensor period) and 22 km otherwise. Although seasonally varying, MDFF is similar during all years, meaning that resolution is the primary determinant of whether a flood of a given size can be detected in a given season. For example, if MDFF is 0.06 and the flooded fraction resolution is 22 km, the minimum detectable flood size is about 36 km^2 per footprint, which is

equivalent to an 18-km long, 2-km wide floodplain segment. If flooded fraction resolution is 50 km and MDFF remains 0.06, the minimum detectable flood size increases to 184 km^2 per footprint.

In the absence of flooding—which is most of the time—false positives are most prevalent in the marginally vegetated areas of Africa during rainy seasons, including areas of the Sahel, Southern Africa, and the Horn of Africa. As reflected in the MDFF (Appendix A), flood detection is also limited by day-to-day variation in dry-land conditions in these same areas; that is, the size of a flood in, e.g., Burkina Faso in September (when MDFF is highest) must be greater than one in the Congo Basin where MDFF is low year-round.

Notwithstanding resolution differences, the AFED algorithm is consistent from year to year overall. Year to year flooded fraction consistency tests (Appendix A) show that seasonal wetland flooding is consistently represented year over year and areas known to be dry or very rarely flooded (i.e., the Sahara) remain flood-free in AFED. Locally, some patterns of false positives are more prevalent in some years than others, including the false positive cases in Kenya examined here.

Tests indicate that the AFED algorithm is self-consistent with regard to the relationship between flooded fraction derived from microwave data and the high-resolution daily AFED flood depiction derived from the flooded fraction (Appendix A). After downscaling the flooded fraction to produce AFED, the effective flooded fraction of the AFED product itself closely matches the input microwave flooded fractions. This consistency is important for understanding that AFED product performance is directly connected to microwave flooded fraction performance and, therefore, future improvements in flooded fraction, e.g., finer microwave sensor resolution, are critical for making further improvements in AFED.

Tests also indicate that AFED depictions are generally consistent over time during a flood event with some day-to-day variation. Examples here show that by computing an accumulated number-of-days flooded statistic it is possible to produce a flood extent depiction that represents the total extent and duration of flooding with higher confidence than daily depictions alone. Daily depiction of flood extent is needed for ARC to determine the duration of flooding at particular locations. Analyses of flood duration, maximum extent, and other flood characteristics are facilitated by the AFED algorithm's georeferenced representation of flooding on a fixed earth grid. For example, AFED users may easily evaluate the total flooded area within a country or other administrative unit on a daily basis using a predetermined set of grid points. The grid point set for each administrative unit is also consistent between the historical products and near real time flood depictions to be produced in the future. Use of the total number of flooded days per year or overall for each grid point may mitigate potential issues raised by day-to-day uncertainties in the flood depiction by providing an index related to the severity of flooding at a particular location. AFED users may also use this statistic to rule out short-duration events (less than 3 to 5 days) most likely to be false positives.

Algorithm limitations not likely to benefit from algorithm changes

Performance tests and trade studies carried out during AFED algorithm development have shown that some limitations to flood mapping are either unlikely to be ameliorated by further effort or may be addressed only by solutions that degrade performance in some other respect. The following situations are not likely to be improved by any changes within the current algorithm framework:

- Flash and short-duration flood detection: Floods that last less than two or three days in the same area are unlikely to be detectable with current satellite technologies. The false negative case studies of Agadez and Greater Accra provide good examples of this type of event. AFED uses polar-orbiting satellites that pass over areas of Africa at most two times per day and, after excluding rain-affected and false-positive flagged data, coverage at times may be fewer than once per day. Geosynchronous satellites make multiple observations per day but have not historically carried instruments that can make observations through clouds, limiting their usefulness for flood mapping. As discussed below, it may be possible to combine microwave observations from two polar-satellites at times, which may help capture some additional shorter-duration events. Nevertheless, the shortest duration floods are most likely to continue to be missed entirely because of a gap in data coverage or rain during the event. Furthermore, the shortest floods are most likely to be indistinguishable from false-flooding signals that are also short duration.
- Small flood detection: Small floods are difficult to detect with microwave data because microwave sensor spatial resolution is coarse and variability in the natural environment and sensor viewing geometry make it impossible to have perfect information about the area viewed in each observation. This combination of conditions is generally described as a "signal-to-noise" problem and several false negative case studies here provide good examples of this problem (e.g. Ouagadougou, Burkina Faso; Niamey, Niger; Tintane, Mauritania; Busia District and Nzoia River, Kenya; Tana and Sabaki Rivers, Kenya; Grand Comore, Comoros). Small floods provide a smaller signal in the microwave data relative to larger floods because the signal is proportional to the area flooded within a sensor footprint. The algorithm has fundamental uncertainties about the natural environment at the time of the observation because vegetation, soil moisture, atmosphere, and other conditions are continuously changing and the algorithm's estimates of the conditions relevant for measuring flooded fraction are imperfect. These uncertainties cause measurement noise in the flooded fraction algorithm. All remote sensing algorithms are subject to similar signal-to-noise conflicts that limit their measurement accuracy. In the AFED algorithm, small flood detection is more limited when (a) 50-km SSM/I data are used, (b) MDFF is high (generally arid regions during periods when variability is driven by frequent rain but also coastal areas where sensor viewing geometry uncertainty and the presence of persistent water cause day-to-day measurement variability), and (c) urban areas where structures obscure satellite sensors' view of the ground and bare ground and artificial surfaces decrease the difference between dry and flooded microwave data, thereby reducing the net signal of flooding. Other sensor types (e.g., MODIS, Landsat, radar) may be capable of mapping smaller floods but their inability to see through clouds, infrequent observations, or lack of availability historically mean that they do not meet ARC's criteria for flood mapping consistency and continuous coverage of all of Africa.
- Precise flood depiction: The AFED algorithm is limited in its ability to depict floods precisely by uncertainties in microwave flooded fraction estimates and the downscaling process. The Inner Niger Delta and Limpopo validation cases include discussions of downscaling limitations. Even in an ideal situation where flooded fraction and elevation could be known precisely, downscaling would be unable to map floods precisely because it cannot account completely for (a) dynamic hydrological processes that control the distribution of flooding over time, (b) topographic and artificial features in the landscape

like dikes that are not represented at the scale of the elevation model, and (c) neighboring basins with differing degrees of flooding that are too close to be distinguished in microwave data footprints.

- False positives: The AFED algorithm may never be able to entirely distinguish some types of false positives from true flooding. Our analyses (e.g., Isiolo and North Eastern Kenya cases) suggest that many false positives are due to rain events that increase soil moisture over a large area (much greater than a single footprint), mimicking the microwave data signal caused by flooding over a smaller area (less than each footprint but often crossing over many footprints). By implementing a false positive detection scheme, the AFED algorithm strikes a balance between detection of all false positives and detection of smaller, shorter flood events. That is, to detect false positives more completely it is likely that additional true-positive flood events would be incorrectly excluded as well. Therefore, the best approach is to allow some false positives in order to detect as many smaller, shorter duration floods as possible. This trade-off between false positive exclusion and smaller event detection is likely to persist irrespective of the algorithm approach used or the frequency of satellite observations.

Algorithm changes recommended for consideration under a future effort

We recommend that ARC consider the following algorithm changes for investigation in a future effort. Each proposed change would have to be evaluated as a "trade study" in which AFED performance was assessed overall against the baseline algorithm and criteria were established for accepting the alternative. Therefore, in addition to the algorithm work effort estimated for each change, effort should also be devoted to developing validation data and methods needed to support the specific trade studies. The proposals are list roughly in the order of priority based on likelihood and scope of potential impact relative to level of effort required:

1. Estimate flooded land end-member separately from open water: The AFED algorithm currently uses the open water end-member as an estimate for the flooded land end-member. (See the ADD for details.) Test scene results above (see section 3.4 and the discussion of negative bias trends observed in some flooded fraction scatter plots) indicate that the algorithm may underestimate flooded fraction in many situations, which could be attributable in part to over estimation of the flooded land end-member. The proposed algorithm change would compute the flooded-land end-member based on local land cover conditions under the assumption that a proportion of flooding is under-canopy and therefore has a lower end-member value than open water. Roughly two to three months' effort would be needed to implement and test an end-member model and develop higher-resolution validation scenes to more accurately test flooded fraction results in areas with complex canopy structures.
2. Combined use of multiple SSM/I sensors in historical processing: The historical processing system uses only one SSM/I sensor based on a best-available sensor approach. For most of the period from May 1995 through November 2009, at least two SSM/I sensors are available simultaneously from separate platforms. If the AFED algorithm combined the best two SSM/I data sets available each day, AFED performance could be improved throughout Africa in key areas including false positive detection, small flood mapping, and perhaps shorter-duration flood detection simply by having more frequent observations across the two sensors at different times of day. Rainfall, for example, can

be more likely at some times of day than others and having additional daily observations increases the chance that rain-free observations are available. The potential impact of this proposal has significantly decreased with AFED V05R00, which includes 22-km TMI data starting from January 1998. Any improvements from the use of multiple SSM/I sensors would now only affect May 1995 to December 1997. Roughly two to three months' effort would be needed to implement and test algorithm changes to fully exploit the combined data streams.

3. Replacement of the land-water mask: The AER team developed the AFED land-water mask by combining the two best available high-resolution water cover datasets: the Hansen (Hansen et al., 2013) year 2000 mask and the SRTM Waterbody Database (SWBD; Farr et al., 2007; see the ADD for full details.) Both datasets represent a snapshot of water cover in year 2000 and to date we have found no suitable data of similar quality that tracks water cover year-to-year. The mask is important to flooded fraction derivation from microwave data and to flood depiction. In AFED flooded fraction processing, the mask determines how much of the total water detected will be interpreted as flooding; in flood depiction, the mask determines what areas are already water covered and therefore not subject to additional flooding. We have not found significant errors in the mask to date but have noted, for example, that AFED users must account for the flooding of new reservoirs in their applications. New reservoirs or other natural or artificial changes in water distribution over time or a reinterpretation of the definition of persistent water by ARC may necessitate an update to the mask. ARC may also choose to have the mask change over time, e.g., due to the creation or removal of a reservoir or major change in river course. In 2017 AER investigated whether the Global Surface Water 1984-2015 dataset (GSW, https://global-surface-water.appspot.com/) should be adopted for AFED algorithms and recommended that it should not be at that time (notwithstanding the separate use of GSW in the AFED algorithm's relative floodability database). GSW is a more recent depiction of surface water, which makes it incompatible with the current AFED land-water mask, which represents a snapshot of conditions in the year 2000. To merge GSW data with the current mask would double-map rivers that change rapidly over time, e.g. the Zambezi, resulting in an overestimation of total persistent surface water. An advantage of SWBD is that it is a bi-product of SRTM DEM production and is therefore consistent with the MERIT DEM used in AFED processing. The Hansen mask was produced with data contemporaneous with SRTM, notably includes smaller features SWBD excludes, and fills in a number of larger SWBD gaps. Because of these benefits, it makes sense to continue with the current AFED land-water mask. However, we recommend revisiting this question if a (a) new DEM is adopted in the future that no longer has ties to SRTM and/or (b) a land-water dataset become available that accurately tracks year-to-year water cover change. Roughly three to four months effort would be needed to assess the new dataset and implement and test algorithm changes to accommodate a time-varying mask. Alternatively, ARC may choose to maintain the current land-water mask in the AFED algorithm and apply a separate mask over final AFED results to indicate additional areas that should not be interpreted as flooded either by year or over all years.
4. Use of microwave data at full resolution: The AFED algorithm uses the footprint matching and regridding process to normalize microwave data spatial footprints to a common symmetric shape interpolated to a fixed geographic grid. The shapes are

approximately circular-Gaussian weighting functions at 22- and 50-km resolution for AMSRX and SSM/I, respectively. AER developed this novel approach for AFED to push the capabilities for detecting and mapping floods from microwave data by effectively separating temporal and spatial variation in the microwave dataset. An alternative to this approach is to process microwave data through flooded fraction retrieval in the as-measured footprint geometry and downscale flooded fraction to produce an AFED-like representation prior to 3-day averaging and application of MDFF. This approach, which has not been attempted before, would use and build on the current algorithm framework to further push the boundaries of flood mapping from microwave data. The potential benefits would be improved detection of smaller floods throughout Africa (including potentially the case studies with small but sustained flooding, e.g., Ouagadougou, Burkina Faso; Niamey, Niger; Tintane, Mauritania; Busia District and Nzoia River, Kenya; Tana and Sabaki Rivers, Kenya). Roughly four to six months' effort would be needed to implement and test the algorithm change thoroughly.

5. Update NDVI climatology and microwave calibration: The AFED algorithm uses microwave data climatologies calibrated over two periods: 1992-2001 and 2003-2010 plus 2013-2016 for SSM/I and AMSRX, respectively. The algorithm uses the NDVI climatology (11-year average) primarily to provide an index to dry-land end-member seasonal change where the microwave data climatology may be affected by seasonal flooding. The so-called *wetlands signature* areas where NDVI is used are shown in the ADD. The algorithm uses a public NDVI climatology that may be improved in the future by its authors or supplanted by a superior dataset from another group. Although the algorithm discounts the climatologies in regions where there is large year-to-year variability in dry-land conditions, it may improve AFED performance to update the climatologies in the future after a superior NDVI dataset becomes available. Furthermore, in the current algorithm NDVI is used only as a seasonal climatology with no year-over-year change. This is done primarily because NDVI data on a yearly basis is less accurate and may in fact be affected by flooding. However, with a better NDVI dataset it may be possible to model gradual change in addition to the seasonal cycle, e.g., as a 3- to 5-year trend in vegetation amount. The potential benefits would be improved AFED performance over time in areas where there have been shifts in vegetation cover affecting microwave data. Roughly three to four months' effort would be needed to implement and test the proposed changes in the calibration datasets and their use in the AFED algorithm.
6. Combined use of microwave with optical and radar remote sensing: Optical and radar remote sensing can provide more accurate flood extent depiction than the current AFED algorithm based solely on passive microwave observations. These data types present different challenges related to automation and accuracy, including issues with false positives and false negatives. Observations from radar sensors have been less frequent historically than those from microwave sensors and clouds frequently block some or all of optical sensor scenes. The AFED algorithm is well constructed for deriving "best estimate" flood maps with optical and radar sensor inputs. Our approach would be to develop preliminary microwave-only maps of flood likelihood and merge them with optical or radar data when available. A best estimate flood extent algorithm would use flood likelihood to help resolve optical and radar scene ambiguities, continuously fill in areas missed by higher-resolution sensors, and produce flood extent depictions at standard resolution and delivery cadence. The potential benefit would be improved high-

resolution AFED accuracy, reduction in false positives and false negatives, and detection and depiction of smaller floods. An initial research and development effort would require about four to six months to establish feasibility and assess the total effort needed for full implementation in the AFED system or as a separate complementary system.

Table 7 summarizes results for the false positive and false negative case studies listed in Table 2, the Comoros false negative case study, and the coastal zone false positive case studies (Cote d'Ivoire, Ghana, and Namibia). We classify each study by the main reasons there was a false result and the potential for near term improvement with the algorithm changes proposed above.

Table 7: Classification of apparent false negative and false positive cases

Case	Type	Dates	Case Classification	Potential for Near-Future Improvement
Ouagadougou, Burkina Faso	False negative	Early Sept. 2009	Small flood plain size/urban settlement obscuring flood signal	No
Tintane, Maritania	False negative	7 Aug. 2007 - 5 Sept. 2007	Small flood size	No
Agadez, Niger	False negative	Early Sept. 2009	False positive detection algorithm, small flood size, flash flooding	Small chance of improvement spatial-temporal flooded fraction algorithm changes
Niamey, Niger	False negative	10 Aug. 2010	Small flood plain	No
Busia District and the River Nzoia, Kenya	False negative	29 Aug. - 12 Sept. 2003	Coastal interactions, small flood event and small flood plain	No
Busia District and the River Nzoia, Kenya	False negative	13 Nov. 2008	Coastal interactions, small flood event and small flood plain	No
Tana & Sabaki Rivers, Kenya	False negative	April-May 2006	Small flood plain size	No
Greater Accra, Ghana, 2010	False negative	22 Jun. 2010	Flash flood	No
Grande Comore, Comoros	False negative	April 2012	Small flood plain size	No
Isiolo region, Kenya	False positive	October 2009	Flood overestimation due to low vegetation	Small chance of improvement spatial-temporal flooded fraction algorithm changes
North Eastern region, Kenya	False positive	May 2011	Flood overestimation due to low vegetation	Small chance of improvement spatial-temporal flooded fraction algorithm changes
North Western region, Kenya	False positive	May 2011	Flash flood event being overestimated to due low vegetation	Small chance of improvement spatial-temporal flooded fraction algorithm changes

5 References

Carroll, M., Townshend, J., DiMiceli, C., Noojipady, P., Sohlberg, R. 2009. A New Global Raster Water Mask at 250 Meter Resolution. *Int. J. Digital Earth*, 2, 4, 291-308, 2009.

Farr, T.G., P.A Rosen, E. Caro, R. Crippen, R. Duren, S.Hensley, M. Kobrick, M. Paller, E. Rodriguez, L. Roth and Others (2007), The shuttle radar topography mission, *Rev. Geophys.*, 45, 2. doi:10.1029/2005RG000183. Data available on-line from: https://lpdaac.usgs.gov/products/measures_products_table.

Galantowicz, J. F. (2001), *Algorithm Theoretical Basis Document (ATBD) for the Conical-Scanning Microwave Imager/Sounder (CMIS) Environmental Data Records, Volume 11: Vegetation/Surface Type EDR*, AER Document P757-TR-I-ATBD-VST-20010315, AER, Lexington, MA. Available at http://www.aer.com/scienceResearch/mwrs/pubs.html.

Galantowicz, J. F., J. Picton, and B. Root (2018a), ARC Flood Extent Depiction Algorithm Description Document: AFED Version V05R00, *AER document P2181-AFM-ADD-V05R00-R00*, AER, Lexington, MA, USA.

Picton, J., and J. F. Galantowicz (2018b), Task Order 2 : Research Outcomes, *AER document P2181-AFED-DM-TO-02-RO*, AER, Lexington, MA, USA.

Hansen, M. C., P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, S. V. Stehman, S. J. Goetz, T. R. Loveland, A. Kommareddy, A. Egorov, L. Chini, C. O. Justice, and J. R. G. Townshend (2013), High-Resolution Global Maps of 21st-Century Forest Cover Change, Science 342 (15 November): 850–53. Data available on-line from: http://earthenginepartners.appspot.com/science-2013-global-forest.

Lehner, B. and P. Doll (2004). Development and validation of a global database of lakes, reservoirs and wetlands. *Journal of Hydrology*, 296: 1-22.

Pekel, J., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. Nature, 540(7633), 418-422. doi:10.1038/nature20584

Van Leer, K. W., and J. F. Galantowicz (2011), Comparisons of flood affected area derived from MODIS and Landsat imagery. *Tenth Annual AMS Student Conference,* Seattle, WA, American Meteorological Society, S29.

Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., Oloughlin, F., Neal, J. C., Sampson, C. C., Kanae, S., Bates, P. D. (2017). A high-accuracy map of global terrain elevations. *Geophysical Research Letters, 44*(11), 5844-5853. doi:10.1002/2017gl072874

APPENDIX

A. Flooded fraction performance tests

A.1. Minimum detectable flooded fraction

Figure 23 shows maps of the annual minimum, mean, and maximum AMSR-E/AMSR2 22-km and SSM/I 50-km V05R00 MDFF and Table 8 lists the statistics over all land areas and seasons. The areas where MDFF is largest and where it has its largest annual variation are arid and semiarid areas where there is a significant annual rainy season. These areas have a high proportion of exposed ground, which leads to increased day-to-day flooded fraction noise due to atmospheric effects and soil moisture. The lowest MDFF values occur in more vegetated areas (e.g., the Congo Basin) and extremely dry and/or rocky areas (e.g., in the Sahara Desert).

MDFF can be interpreted as a proxy for AFED flood detection ability and flood depiction uncertainty. The smallest detectable flood is approximately the product of MDFF and the area within a circle whose diameter is 1.25 times the microwave footprint spatial resolution[17]. For example, if MDFF is 0.06 and the flooded fraction resolution is 22 km, the minimum detectable flood size is about 36 km^2 per footprint, which is equivalent to a 22-km long, 1.6-km wide floodplain segment. Similarly, for 50-km data detectability begins roughly at 184 km^2 (e.g., 3.7x50 km).

Table 8: V05R00 MDFF overall statistics

	22-km MDFF	50-km MDFF
Mean	0.063	0.070
Median	0.047	0.053
Minimum	0.015	0.016
Maximum	0.025	0.025

[17] The factor of 1.25 accounts for the fact that the microwave footprint spatial weighting function is an approximate circular Gaussian shape with significant weight beyond the half-peak diameter at which the resolution is measured. See the ADD for further discussion.

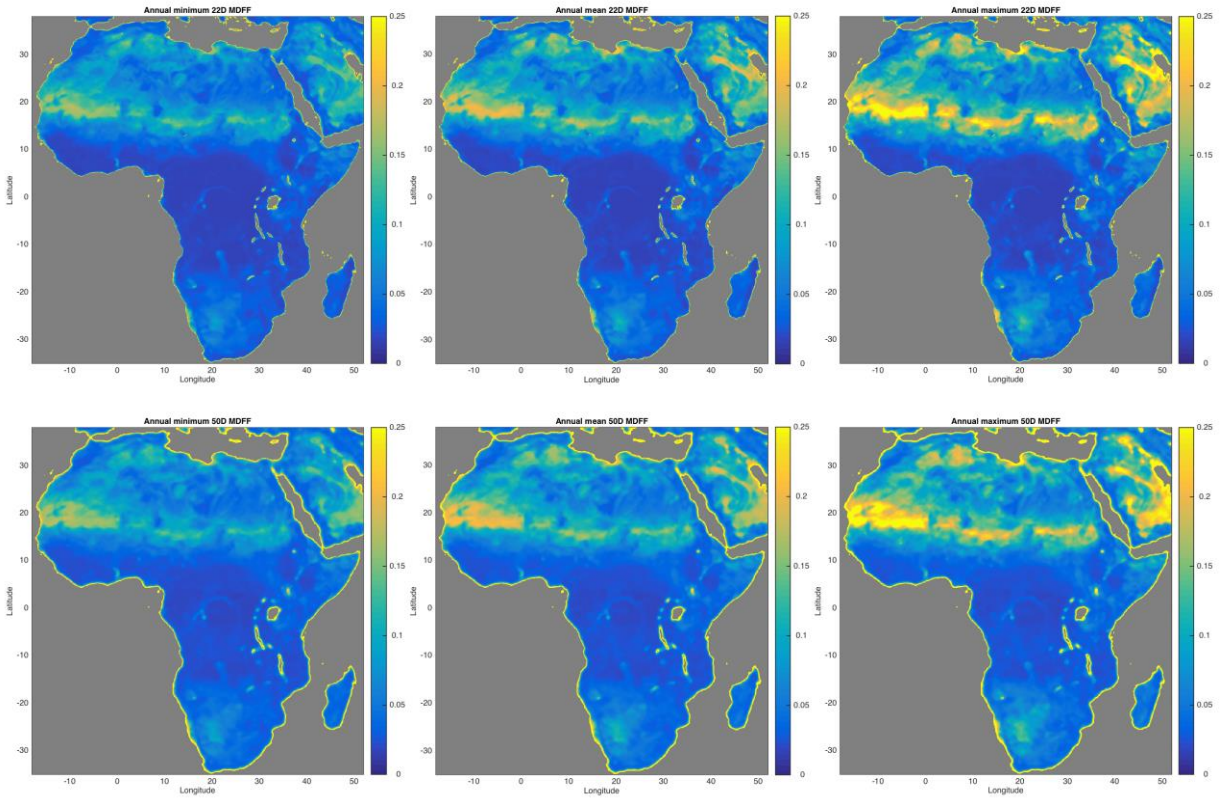


Figure 23: AMSR-E/AMSR2 22-km (top) and SSM/I 50-km (bottom) minimum detectable flooded fraction annual minimum (left), mean (middle), and maximum (right).

A.2. Year to year flooded fraction consistency

The following figures show the number of days flooded per year in flooded fraction data. Days with flooded fraction greater than the minimum detectable flooded fraction are counted. Figure 24 shows 1992-2001 results from V04R01 50-km SSM/I data and Figure 25 shows 1998-2017 V05R00 22-km definitive AFM results, which includes combined data from TMI, AMSR-E, and AMSR2.The pattern of flooded days—which is dominated by Africa's largest wetlands and floodplains—varies locally from year to year but no year-over-year trend of increased flooding or noise is apparent except for the increase in number and spatial detail of detection between the 50-km SSM/I results and the 22-km definitive AFM results.

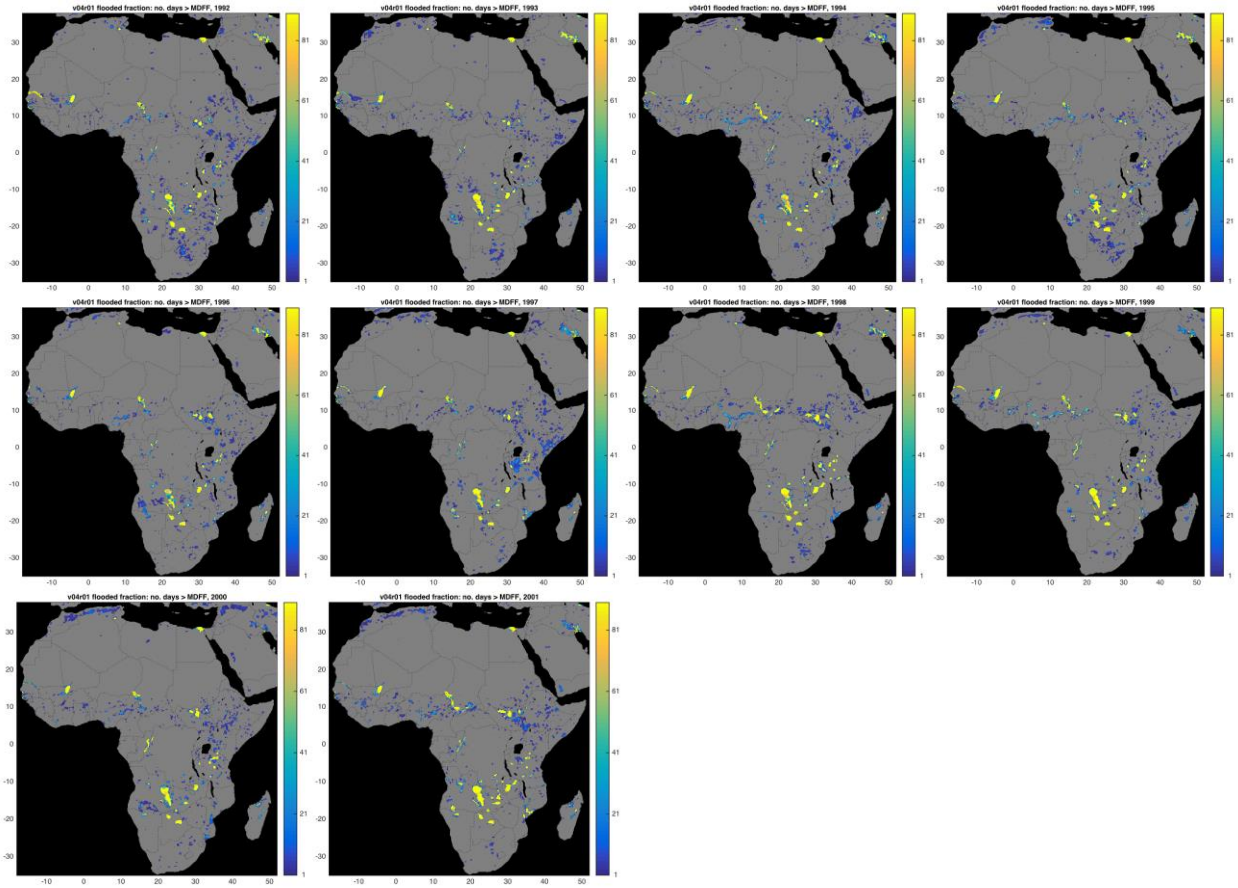


Figure 24: Number of days flooded per year, 1992-2001, counted from V04R01 50-km SSM/I flooded fraction data.

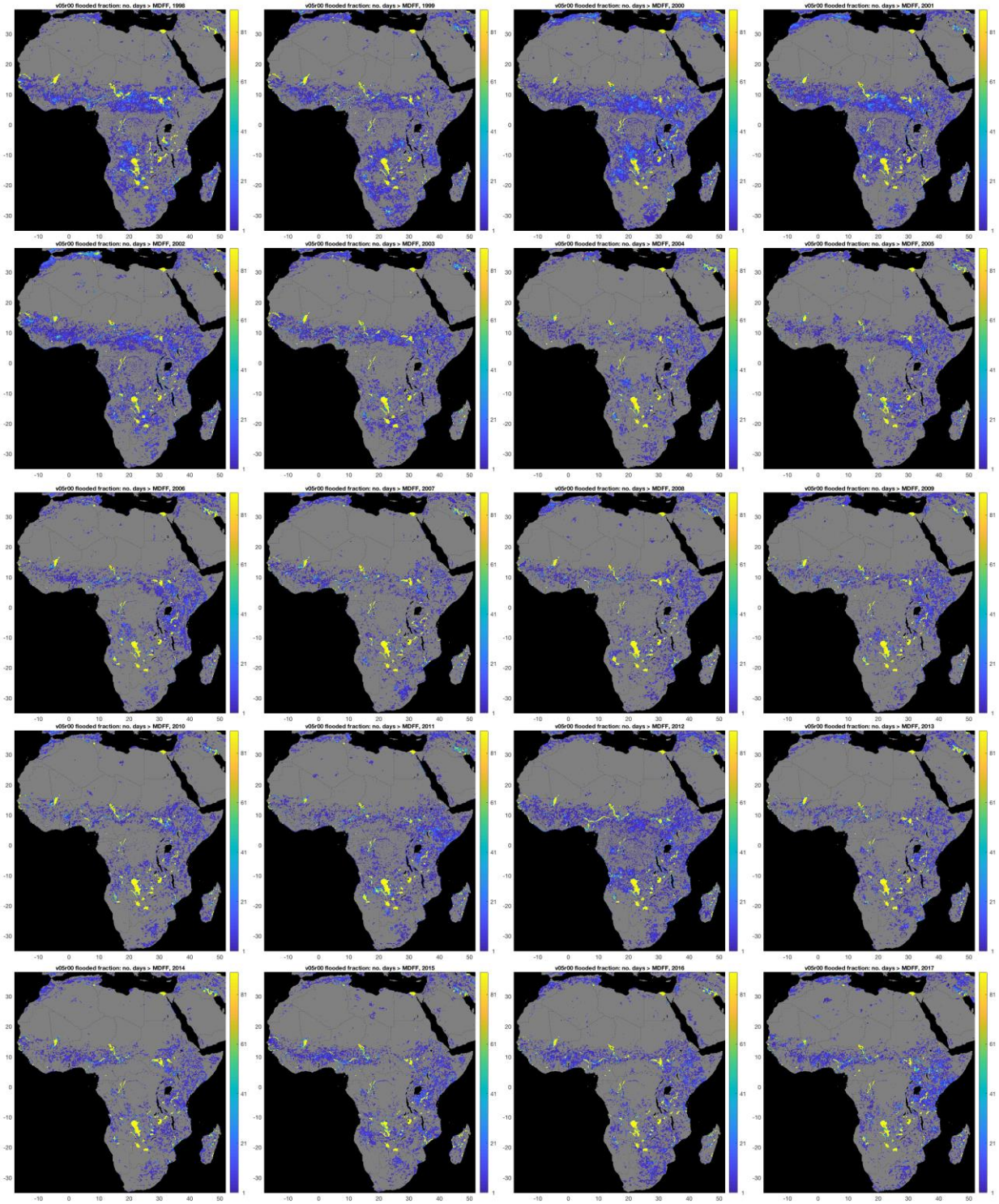


Figure 25: Number of days flooded per year, 1998-2017, counted from V05R00 22-km definitive AFM flooded fraction data.

The maps above indicate many semi-arid and arid areas where flooding occurs for a few days per year. Many of these short-term, large-scale apparent flood events may be attributed to rainfall

that temporarily increases surface water (e.g., soil moisture and ponding) over a large scale but may not represent large-river flooding. Many short-term, larger-scale events are flagged and removed by the AFED “flash” detection algorithm prior to derivation of the final flooded fraction. The AFED algorithm does not remove smaller events to protect against removing actual flood events that are likely to have similar temporal-spatial signatures. AFED users may interpret some of the remaining smaller events as flooding depending on their duration, proximity to other flooded areas, or other criteria. For example, Figure 26 illustrates the prevalence of remaining short-term events by comparing the total number of days flooded in years 2010-20122007-2009 to (a) the total less all isolated single-day data points and (b) the total less all isolated single-day and two-day pair data points. AFED users may significantly reduce false positive rates by eliminating AFED detections not confirmed over two or three consecutive or nearly consecutive days.

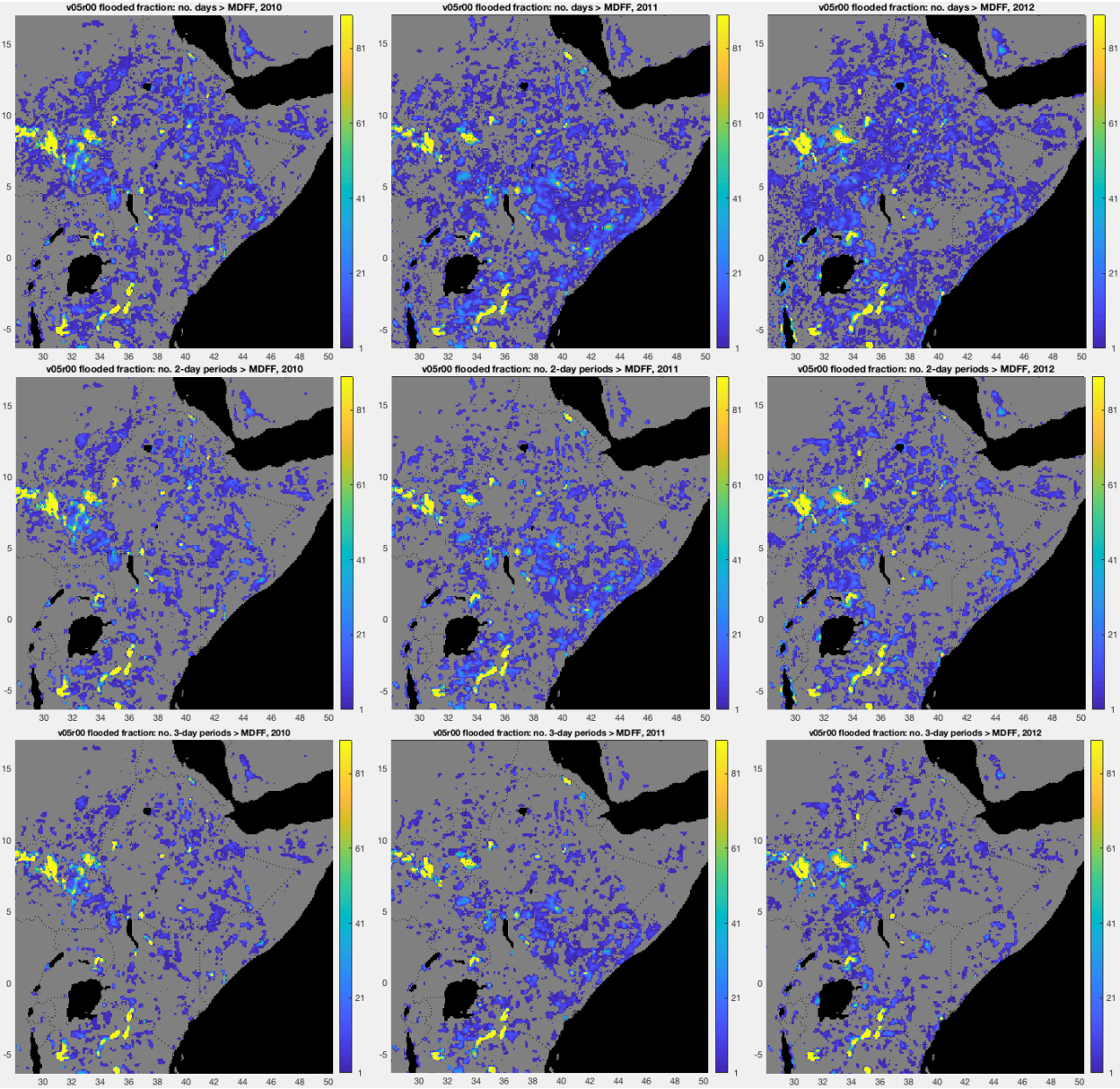


Figure 26: Number of days flooded, year 2010 (left) to 2012 (right), counted from 22-km V05R00 definitive AFMflooded fraction data. Top: total days flooded. Middle: total days flooded less isolated single flooded days. Bottom: total days flooded less isolated single days and 2-day pairs (i.e., only days within 3-day periods).

A.3. Consistency between flooded fraction and AFED

To test the consistency between the input flooded fraction and AFED, we created a series of animations showing daily maps of the two datasets side-by-side every day during flood events. The animations are available separately as GIF or mp4 format files. Figure 27 shows an example of one such map comparison for a flood on the Benue River in Nigeria and Cameroon on 2006-10-2. We observed the following characteristics in these graphics:

- In all cases, AFED flooding is associated with flooded fraction, which confirms that the AFED downscaling process is operating as designed.

- Flooded fraction maps make flooding appear more widespread than the AFED product. This result is consistent with the design of the AFED algorithm. First, AMSR-E flooded fraction is measured on microwave sensor footprints with 22-km half-peak diameter circular Gaussian-like weighing functions. (For scale, the flooded fraction grid cell size in Figure 27 is about 9 km square.) As a result, significant flooded fraction amounts may be measured in footprints off-centered from the flooding. Second, the downscaling process accounts for the footprint weighting function and maps flooding to concentrated areas where relative floodability ranking is highest and water cover would best match the flooded fraction observations.
- At times in all cases the restriction imposed on flooded fraction by the minimum detectable flooded fraction threshold causes apparent spatial or temporal discontinuities in the maps of the flood events.
- At times in some cases (e.g., Senegal 2003 and Shebelle 2006, not shown here) brief flooded fraction "flashes" occur outside the flood plain. We attribute these flashes to rain events that temporarily increase soil moisture and/or surface water (e.g., ponding) over areas typically larger than floodplains.

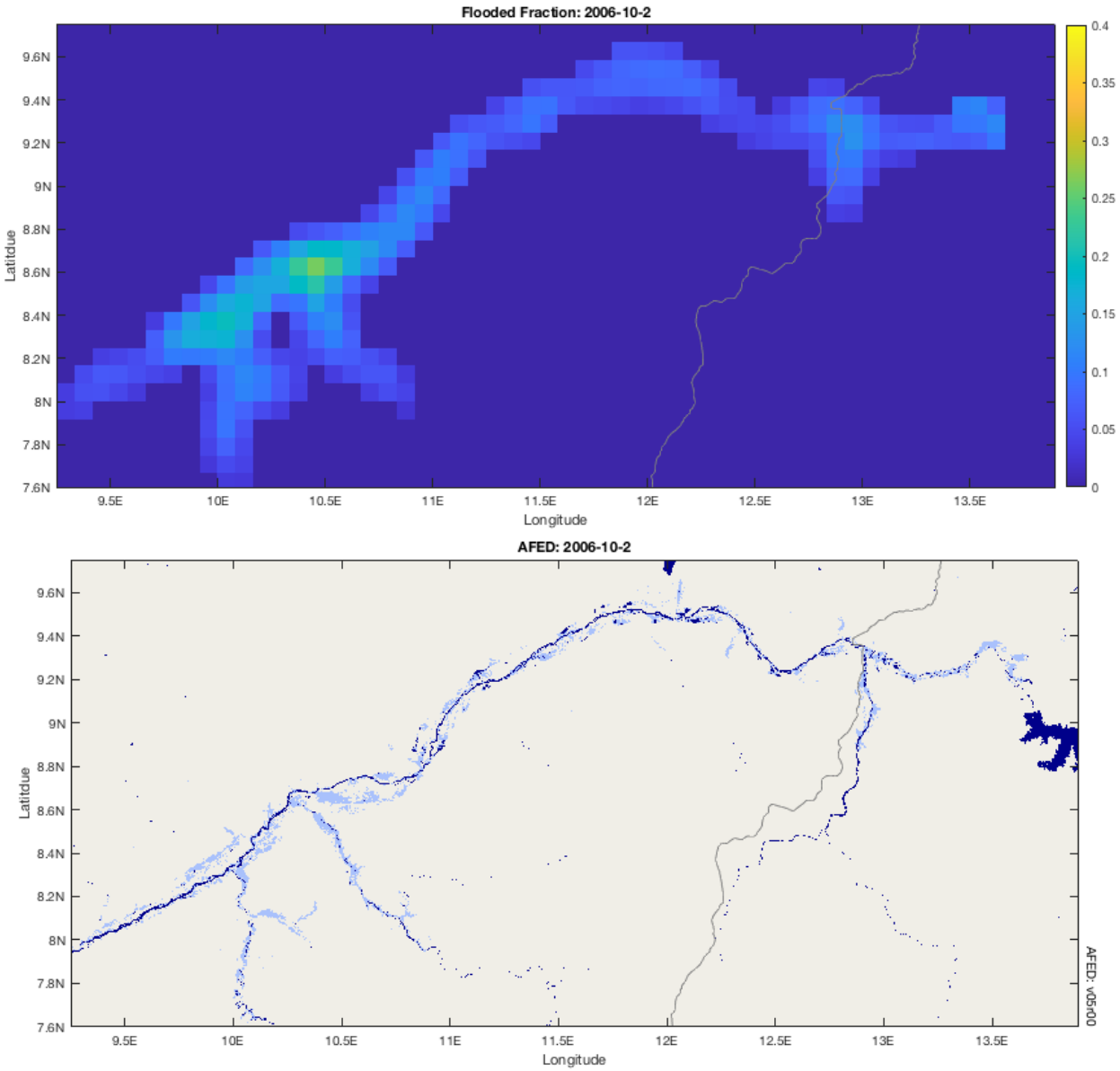


Figure 27: Comparison of flooded fraction (top) and AFED product (bottom) for Benue River, Nigeria and Cameroon, on 2006-10-2. This represents imagery for one day of an animation sequence running from 2006-8-22 to 2006-10-28.

B. Comparison of AFED V04R01 to V04R00

B.1. Introduction

This section summarizes test results demonstrating AFED performance changes for algorithm version V04R01 with respect to the prior version, V04R00.

B.2. Algorithm testing approach

We performed the following types of tests on AFED and the intermediate products from algorithm processing steps prior to AFED downscaling.

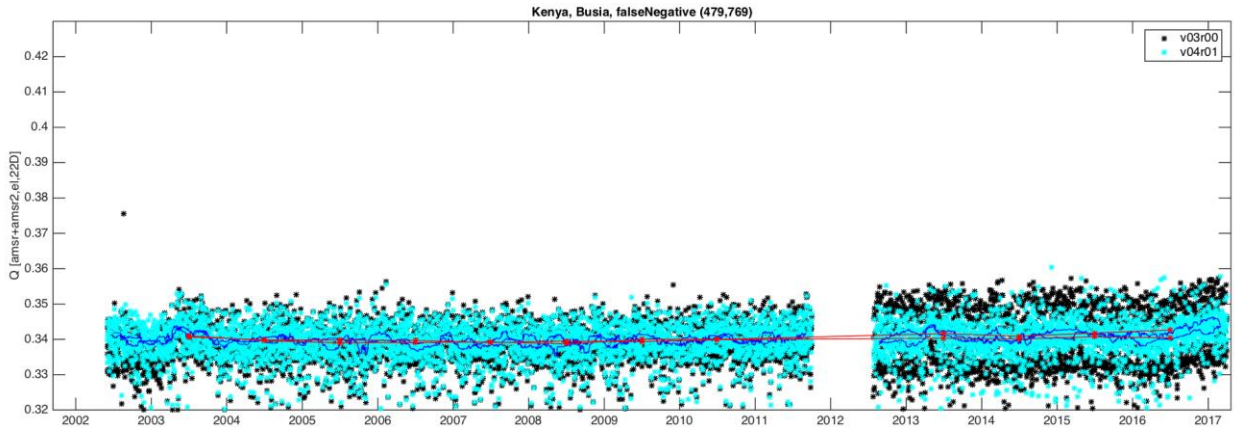
1. **Q time series plots**: Q time series plots combine ascending and descending pass data from AMSR-E and AMSR2 (2002-2017). The data have been atmospherically corrected but rain- and false-positive-flagged data have not been removed. The plots show Q data from V04R01 (cyan markers) plotted over V04R00 (black markers, denoted as V03R00 in the plot legends). The plots include V04R01 ascending and descending 61-day medians (blue lines) and yearly means (red lines and markers), which are included to track year-over-year trends. Look for V04R01 Q to have (a) improved consistency between AMSR-E (2002-2011) and AMSR2 (2012-2017) time periods and (b) reduced day-to-day Q variability.
2. **Flash detection count maps**: Maps counting the number of flash detections per year at the microwave footprint scale. Here we show the difference in detections between the V04R01 and V04R00 algorithm. Look for patterns indicating (a) systematic change in performance or (b) change in flash detection performance for a particular flood event.
3. **MDFF maps and boxplots**: Maps the change in MDFF (minimum detectable flooded fraction) between V04R01 and V04R00. Three maps are shown: (a) maximum annual MDFF increase, (b) maximum annual decrease, and (c) maximum absolute change. Look for systematic regional changes or near features like coastlines. The boxplot shows MDFF change as a function of persistent water fraction. Look for lower MDFF in areas near water bodies to indicate that spatial noise has been reduced.
4. **Flooded fraction time series plots**: Similar to the Q time series, these plots show flooded fraction from the V04R01 algorithm plotted over the V04R00 algorithm flooded fraction prior to application of MDFF. The plots include both versions of MDFF to indicate where flooded fraction would have been clipped in each version. The data span the entire historical AFED period, 1992-2017, and includes 22-km (22D) and 50-km (50D) flooded fraction data (indicated by the legend).
5. **Days flooded fraction exceeds MDFF maps**: Maps counting the total number of days per year that flooded fraction exceeds MDFF (FF>MDFF) less isolated single days and 2-day pairs (i.e., at least three consecutive days must have FF>MDFF to be counted). The maps show the difference between results from the V04R01 and V04R00 algorithm. The counts total the number Look for patterns of systematic increase or decrease in number of days (a) in flood plains, which may indicate change in sensitivity to flooding, or (b) outside flood plains, which may indicate change in false positives.
6. **AFED compared to MODIS imagery:** Three map sets showing (a) false-color MODIS imagery of a flood event, (b) V04R00 AFED, and (c) V04R01 AFED. Look for improved similarity between the overall AFED map and the flood extent apparent in the MODIS imagery.
7. **AFED number of days flooded per year:** Maps of the difference in number of days flooded between V04R01 and V04R00 AFED. Look for significant changes in number of days flooded (>5) coordinated over larger areas (e.g., more than a few pixels), which may indicate significant changes in true flood detection or false positive rates.

B.3. Test results

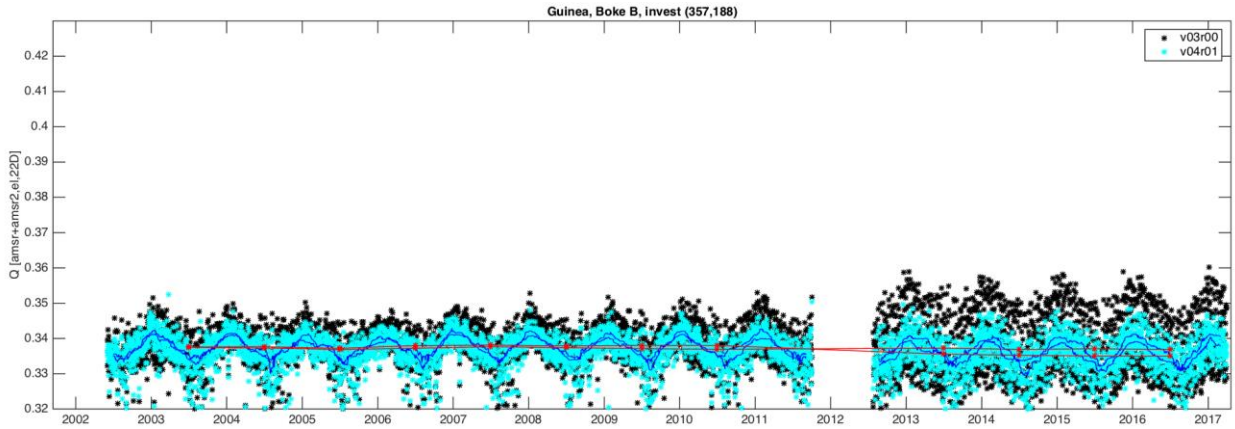
B.3.1. Q time series

The following are Q time series plots for selected locations representing either (a) significant improvement in day-to-day Q variation, (b) a notable lack of improvement, or (c) other features to be discussed. Overall, we made plots for about 391 cases in our database. (Some points appear more than once in the database because of the way the database is used elsewhere.) Of these, only 11 showed no significant improvement in AMSR2 noise; 26 showed significant AMSR2 improvement (making AMSR2 noise more consistent with AMSR-E); and the remainder were not significantly changed. Generally, AMSR-E noise is lower than AMSR2. The V04R01 geolocation correction improved AMSR-E noise in many cases but typically the effect was more significant for AMSR2 data.

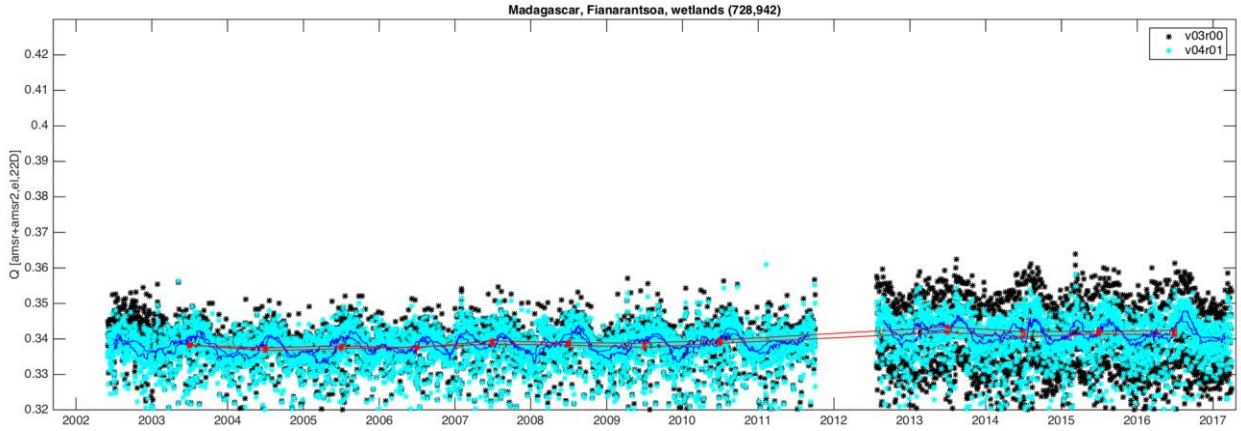
(a) Points with significant day-to-day Q variation improvement



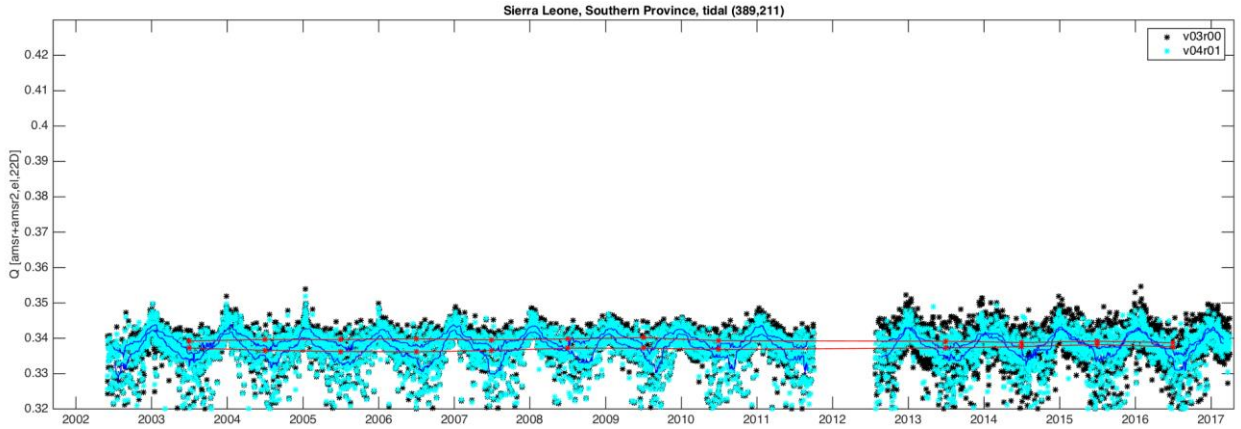
Point near the shore of Lake Victoria included in area of the 2003 and 2008 Busia floods analyzed in the main body of this document. AMSR-E noise is not significantly reduced by the geolocation correction but AMSR2 noise is, which may make it more likely that flooding is detected in this area during the AMSR2 period.



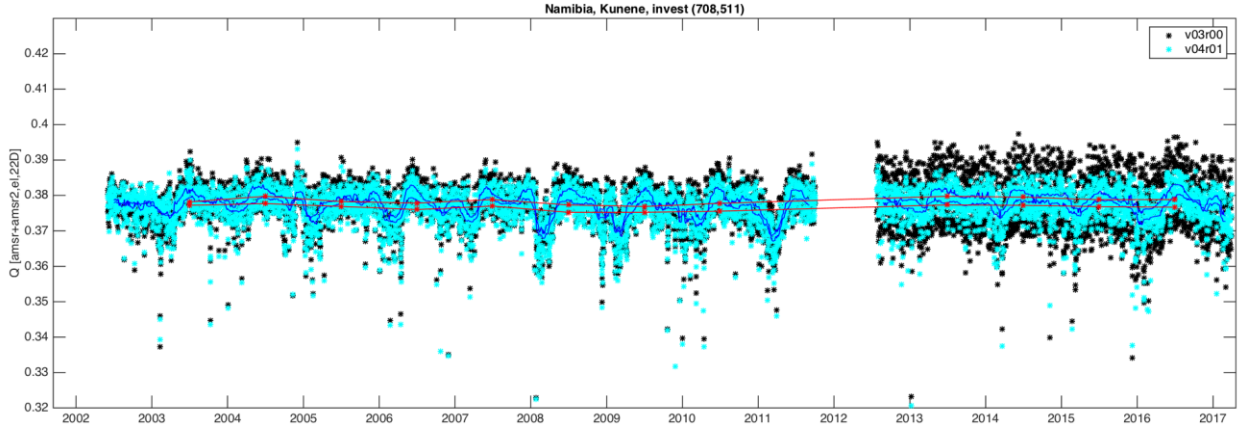
Point near the Guinea coast. Both AMSR-E and AMSR2 noise is significantly reduced.



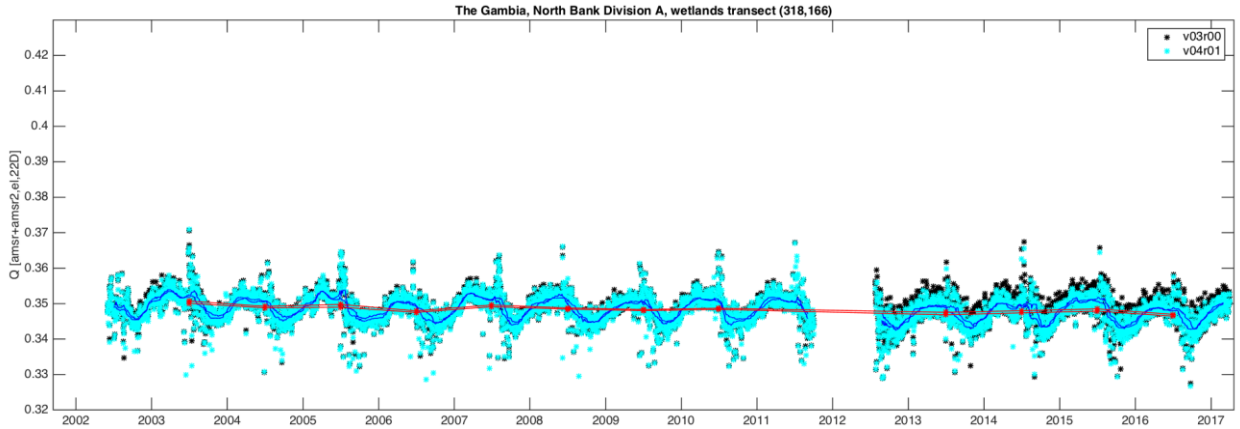
Point near the eastern Madagascar coast near river, wetlands, and possible rice farms. AMSR2 daily noise is significantly reduced, making the seasonal wetlands cycle more clear and consistent with the AMSR-E period.



Point on the Sierra Leone coast near a river and wetlands. AMSR2 noise is reduced more than AMSR-E. This point is very similar to the Madagascar point above but less noisy overall, perhaps because of the land cover or geometry of the coastline.

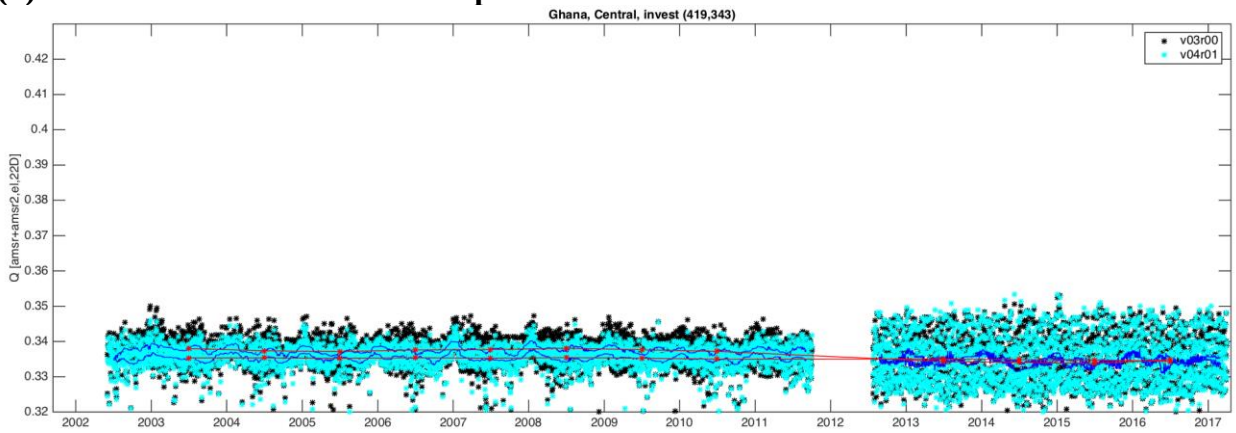


Point on the Namibia coast. AMSR2 noise is significantly reduced in this coastal area.

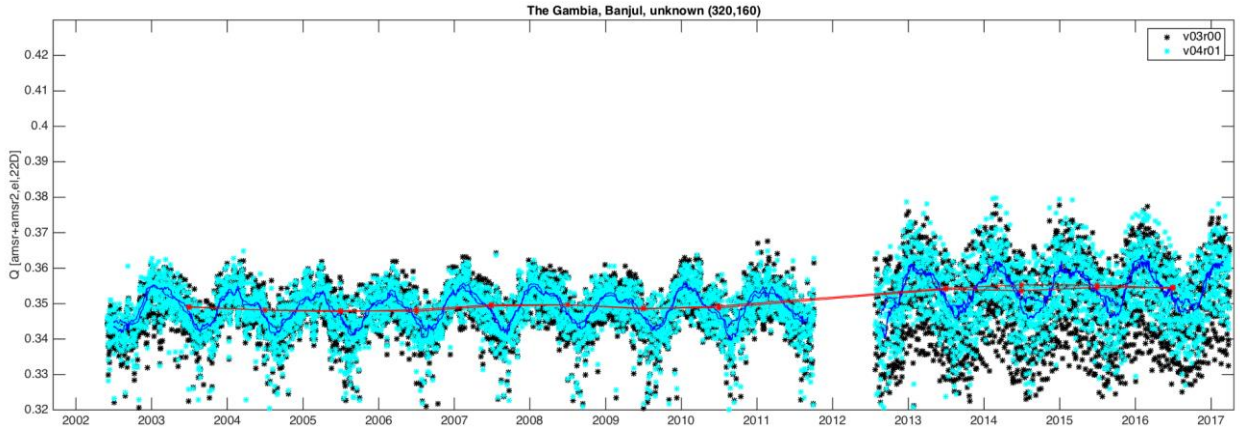


One of a series of points on north bank of The Gambia. AMSR2 noise is marginally reduced but a slight bias has also been corrected, making AMSR2 more consistent with AMSR-E.

(b) Points with a notable lack of improvement

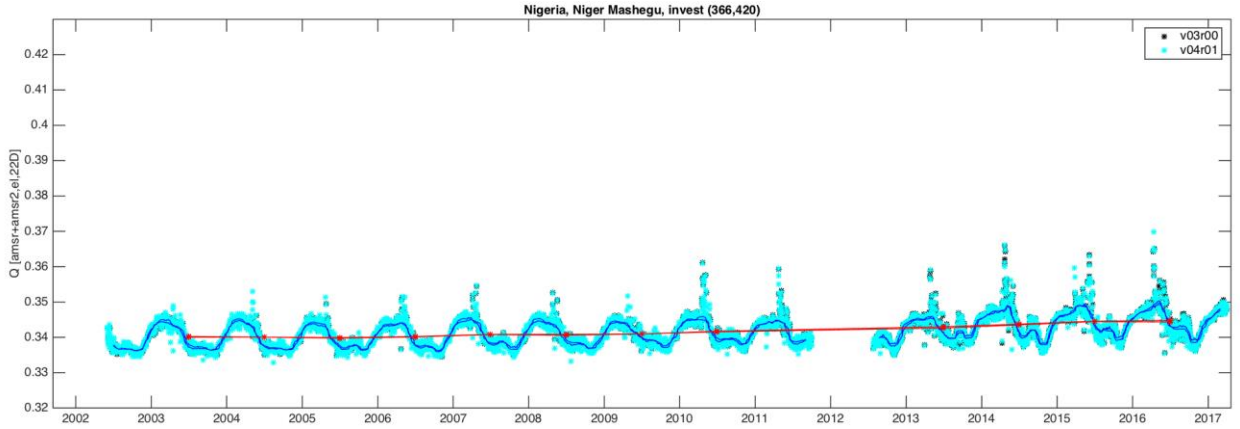


Point 7 km north of the Ghana coast. AMSR-E noise is significantly reduced but AMSR2 is barely affected. It may be that residual noise is worse at particular distances from the coastline and/or particular coastline orientations due to the varying effects of footprint orientation. That is, the resampled microwave footprints are likely to have shape asymmetries that make spatial sampling near a coastline most variable in certain geometries.



Point on the Banjul Peninsula. AMSR2 noise is somewhat reduced but continues to be higher than AMSR-E, perhaps due to the same asymmetries discussed above magnified by the complex coastal geometry around Banjul.

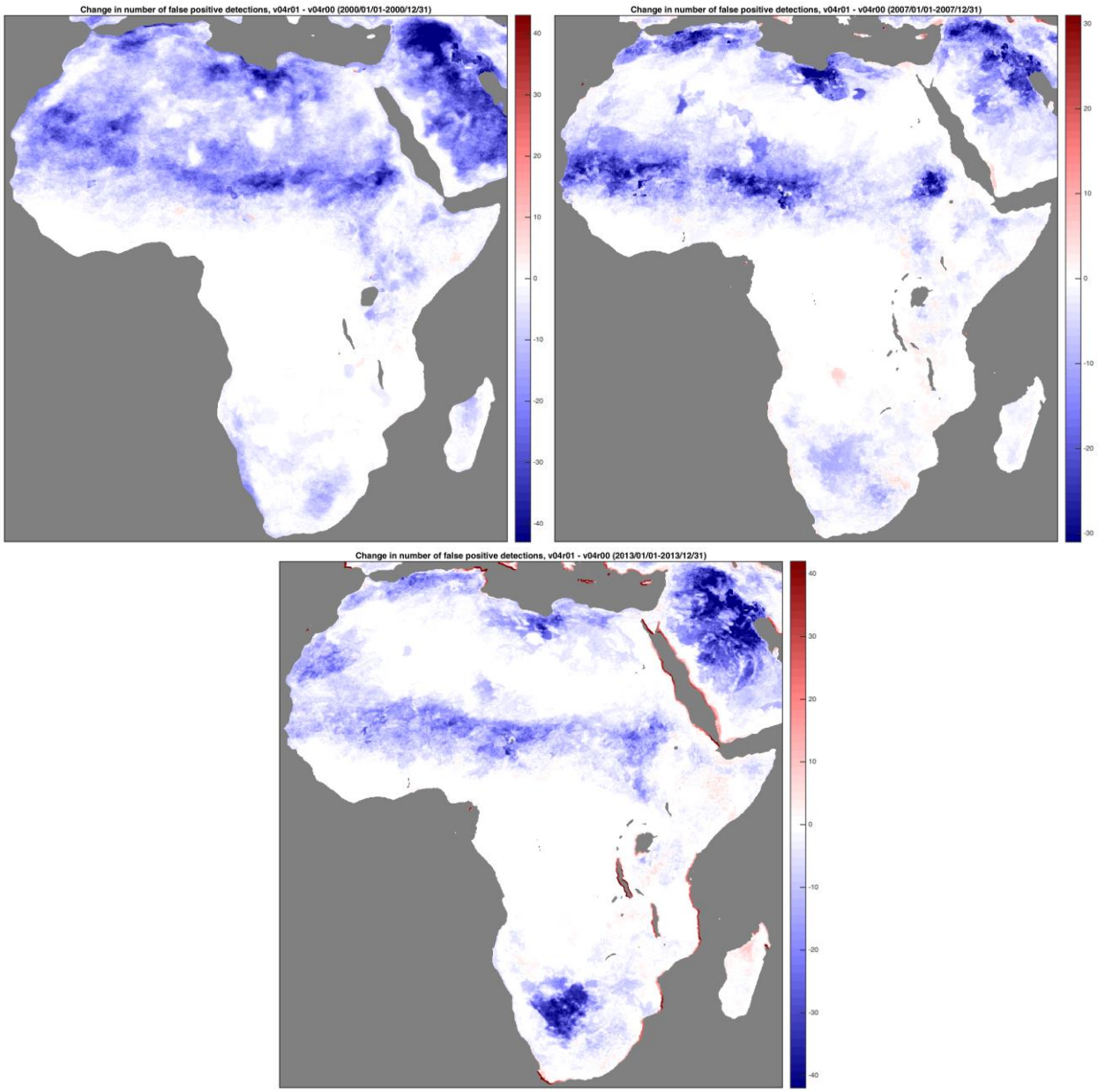
(c) Points with other features of interest



Point in an area where apparent false positives increased during the AMSR2 period and the geolocation correction had negligible effect. Google Earth imagery shows that there has been a gradual deforestation in this area. Our hypothesis is that the increase in exposed or more lightly vegetated ground lead to higher sensitivity of Q to soil moisture and puddling, which we expect is a leading cause of false positives. This plot shows that there is a slight trend in Q year-over-year and that Q outliers in the rainy-season began to be higher and more frequent in starting in 2010.

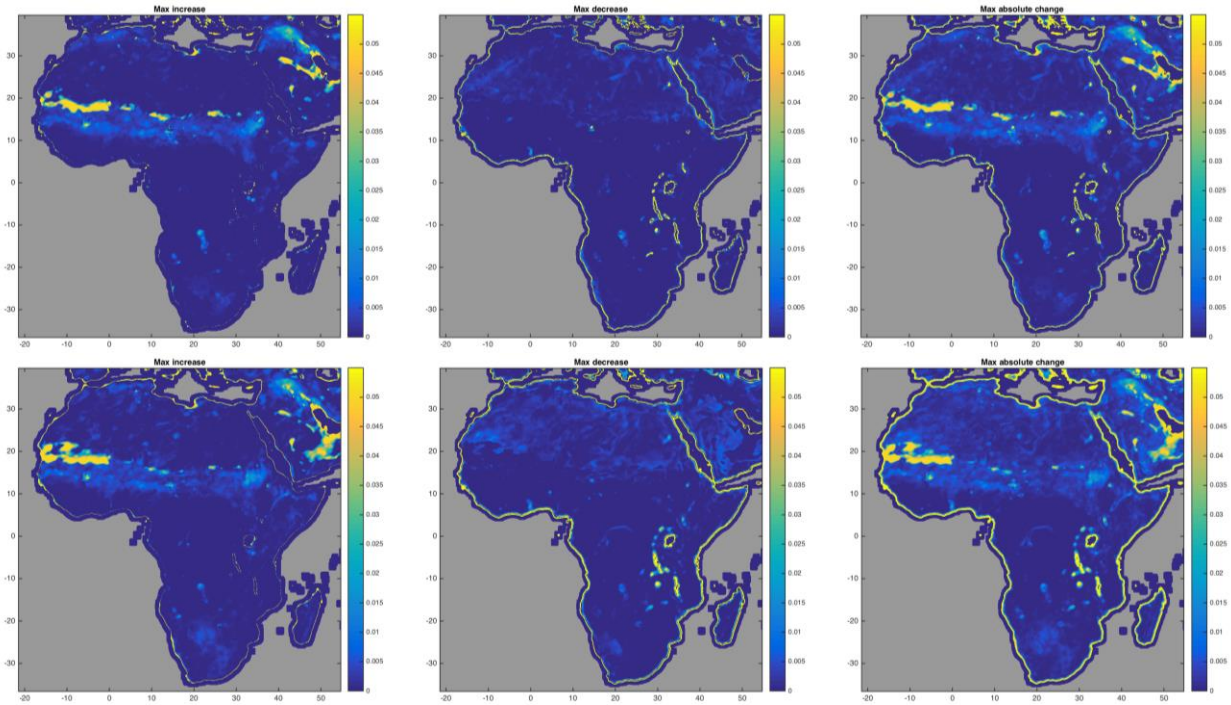
B.3.2. Flash detection count maps

The maps below represent the difference in flash detection counts between V04R01 and V04R00 for 2000 (SSM/I), 2007 (AMSR-E) and 2013 (AMSR2). V04R01 has fewer flash detections overall, which is attributable to algorithm changes that increased Q_{dry} in flash-prone (generally arid) areas which in turn decreased false flooded fraction levels in these areas, resulting in fewer instances triggering the false positive detection thresholds.

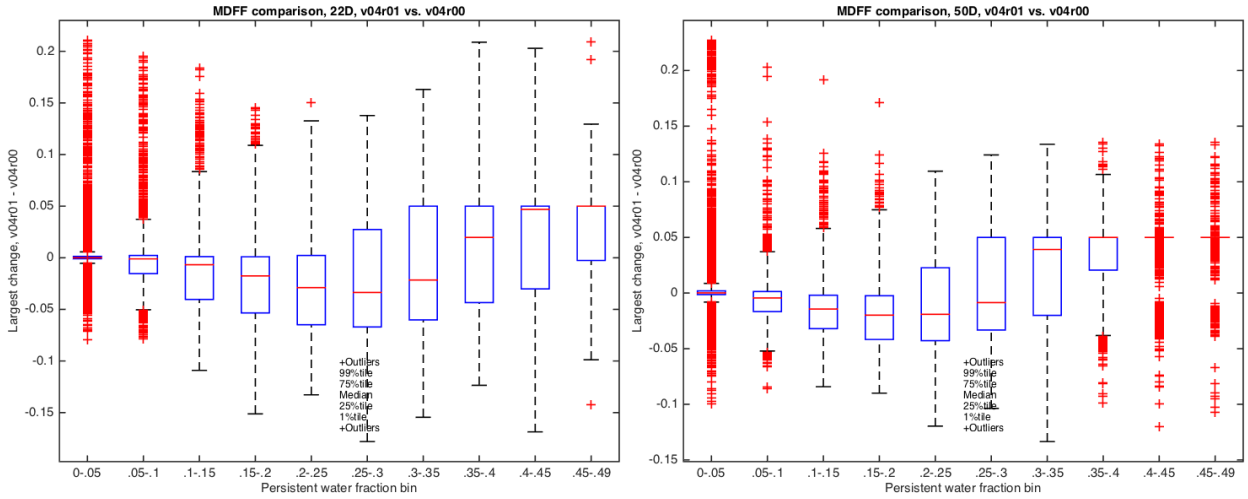


B.3.3. MDFF maps

The maps and boxplots below indicate that the V04R01 geolocation correction algorithm systematically reduces MDFF along coastlines for points with persistent water fraction less than about 0.25. MDFF increases for points above 0.25 water fraction because V04R01 increased the upper limit on MDFF 0.2 to 0.25. The increase in maximum MDFF is also apparent in West Africa in the northern Sahel and southern Sahara.



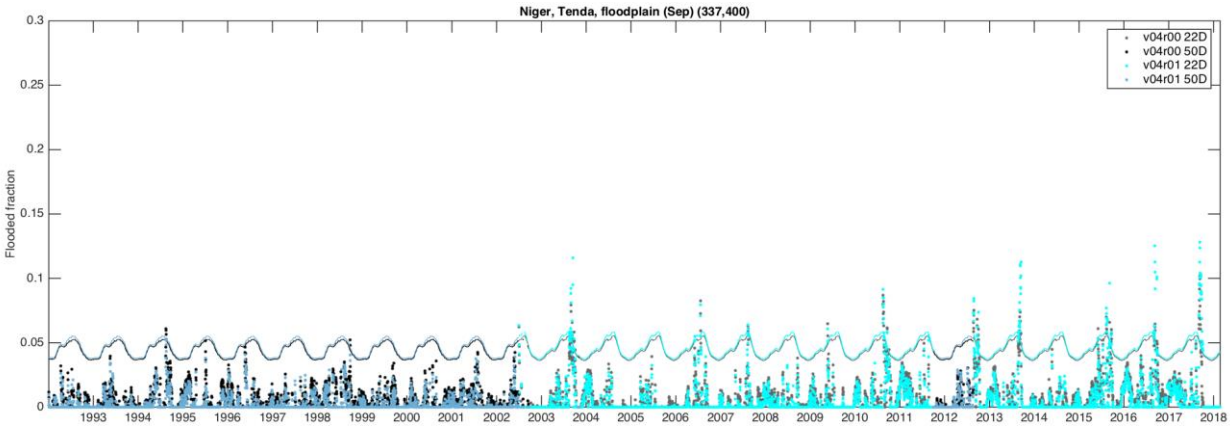
V04R01 vs. V04R00 MDFF maximum increase (left), decrease (middle) and absolute change (right) for 22-km (top) and 50-km (bottom) flooded fraction.



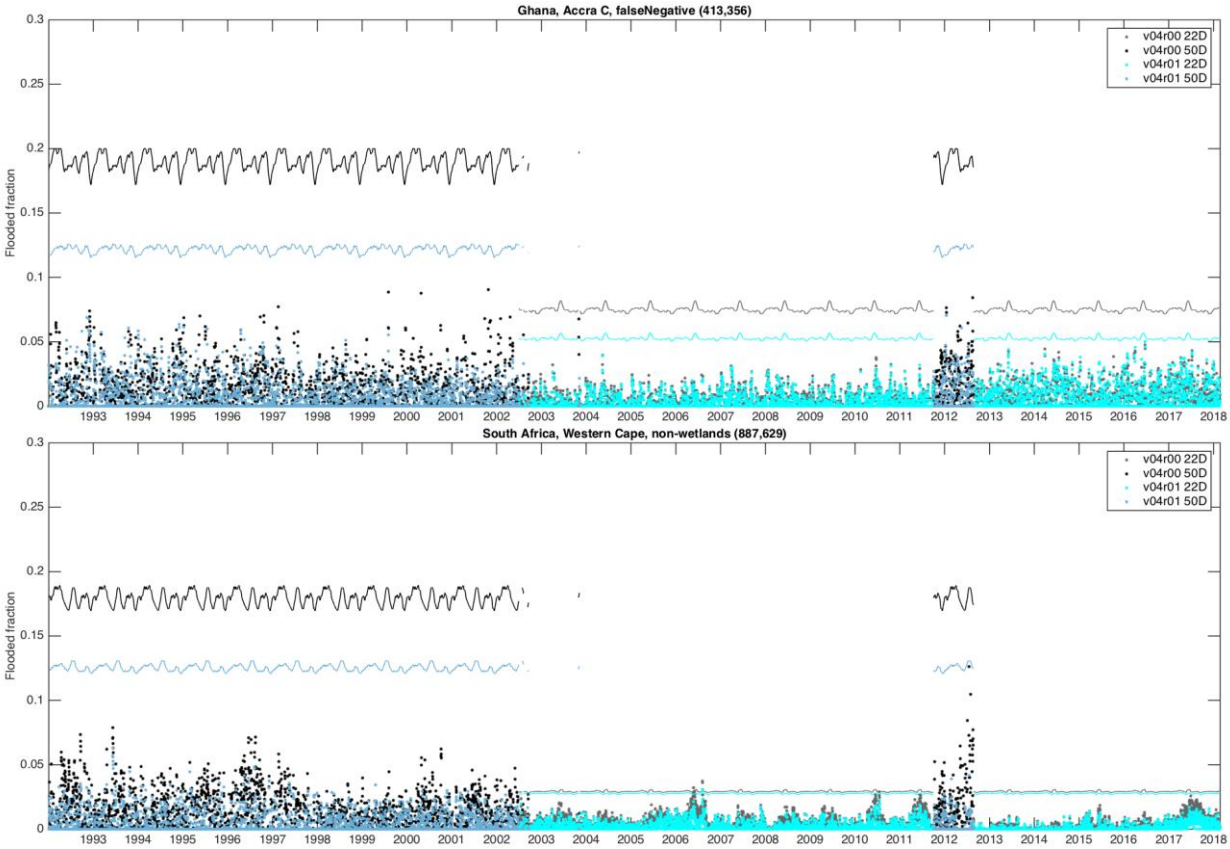
V04R01 vs. V04R00 MDFF change binned by persistent water fraction for 22-km (left) and 50-km (right) flooded fraction.

B.3.4. Flooded fraction time series plots

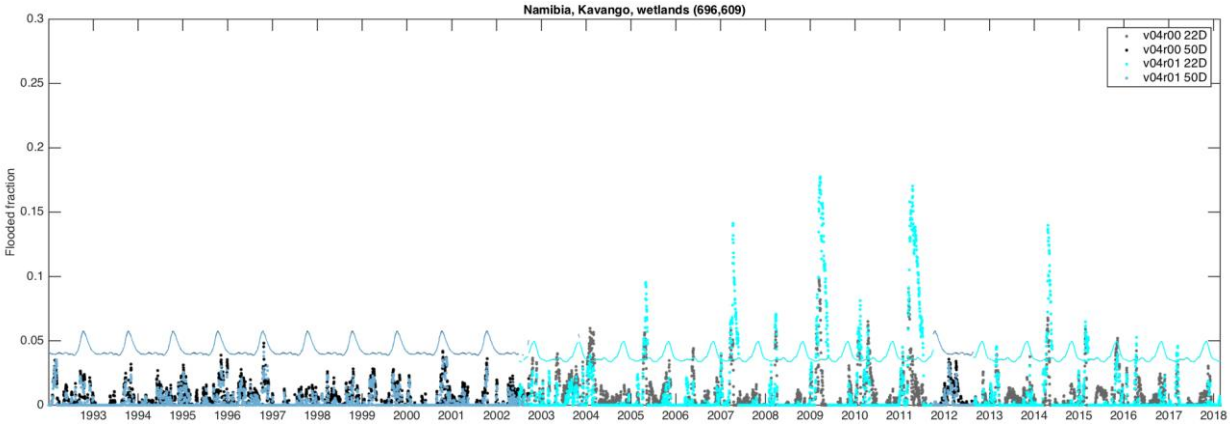
The following are examples from our database of places with either notable changes in flooded fraction time series between V04R01 and V04R00. The points represent specific conditions and are not generally representative of conditions overall.



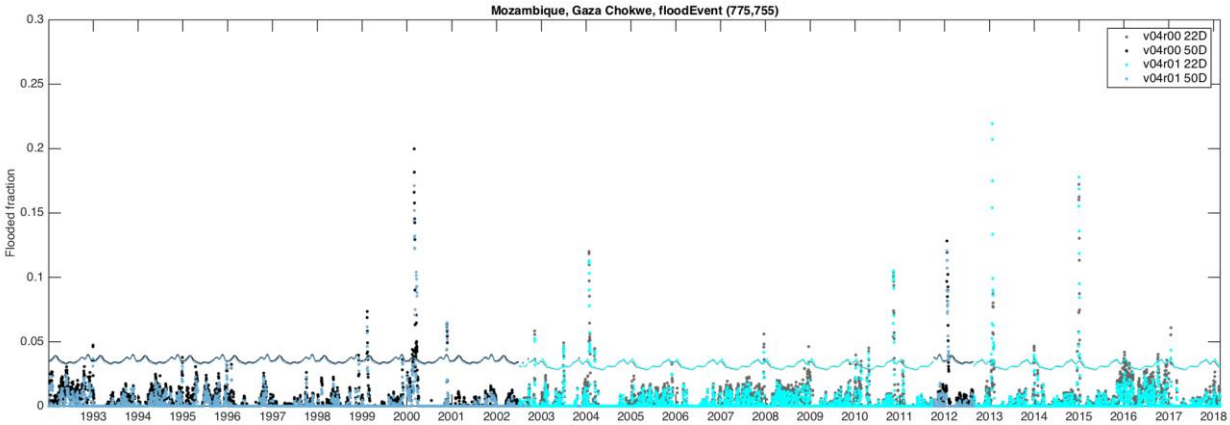
Point on the Niger River near Tenda, Niger where V04R01 enhances apparent true flooding, e.g., in 2003, 2013, 2015, 2016, and 2017.



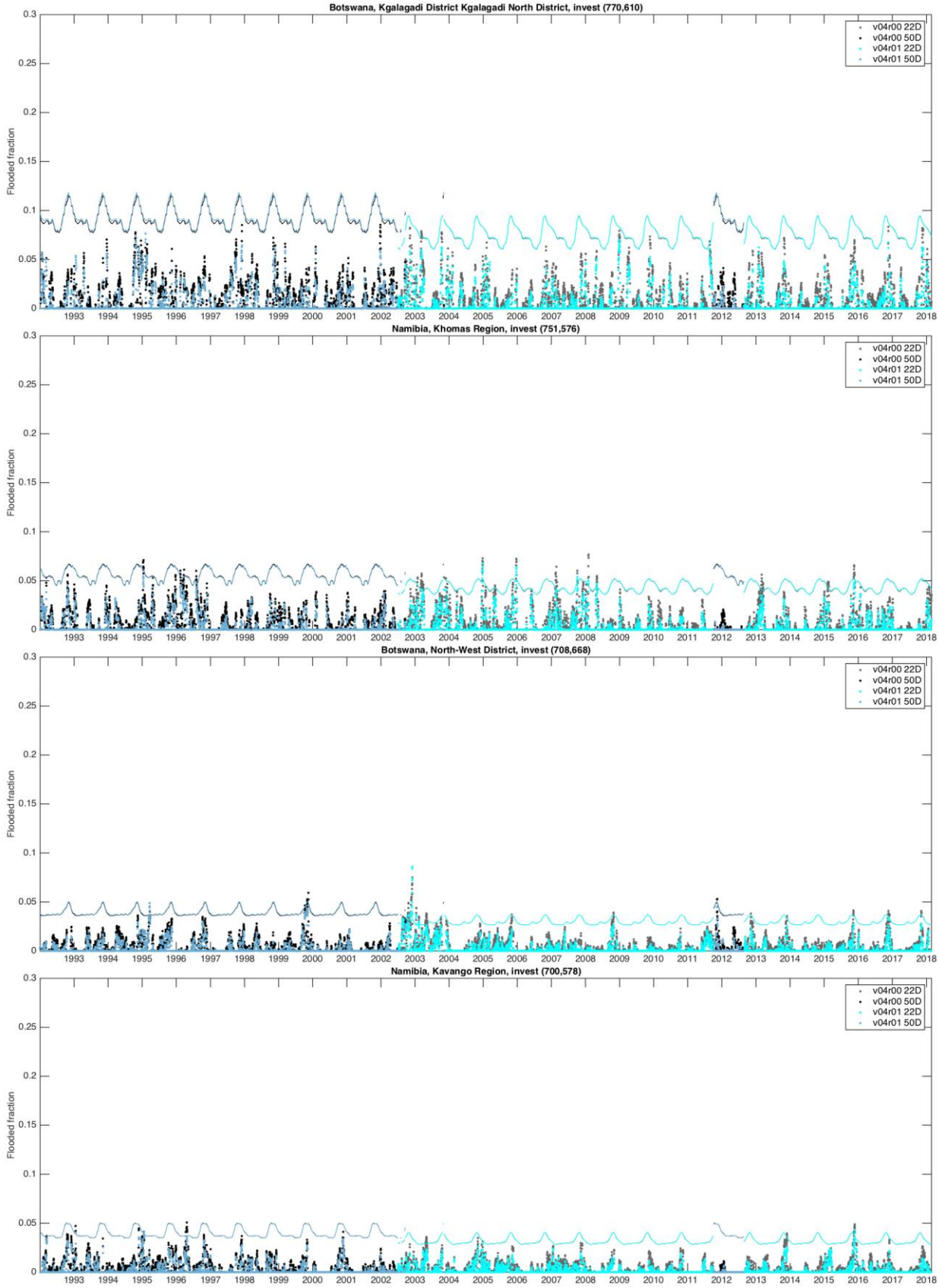
Points near the coast in Ghana (top) and South Africa where V04R01 geolocation correction reduces day-to-day flooded fraction noise and MDFF decreases commensurately.



Point on near Kavango Namibia where the V04R01 flooded fraction temporal-spatial integration algorithm extends flood duration in, e.g., 2005, 2007, 2009, 2011, and 2014.



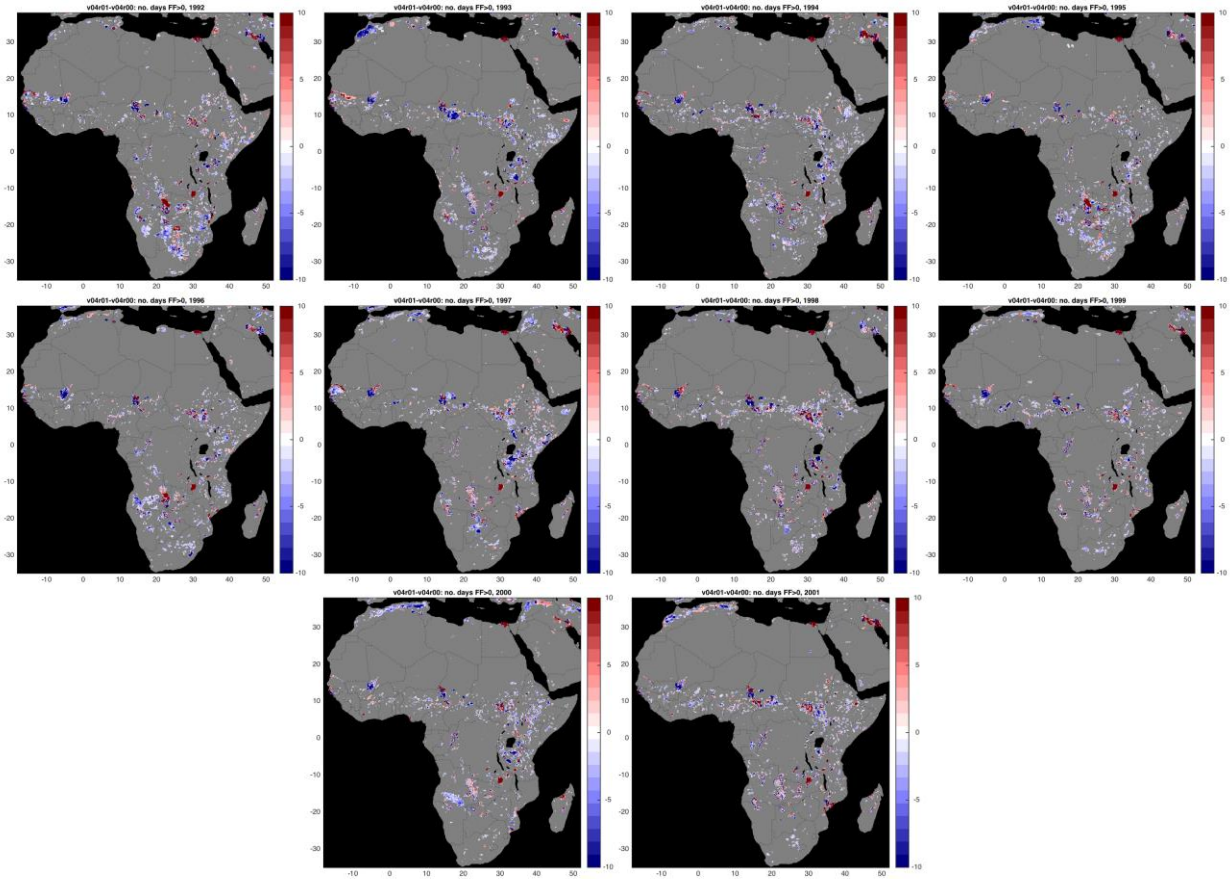
Point near Gaza Chokwe, Mozambique on the Limpopo River where the V04R01 algorithm better captures the peak flooded fraction during 2013 flooding.

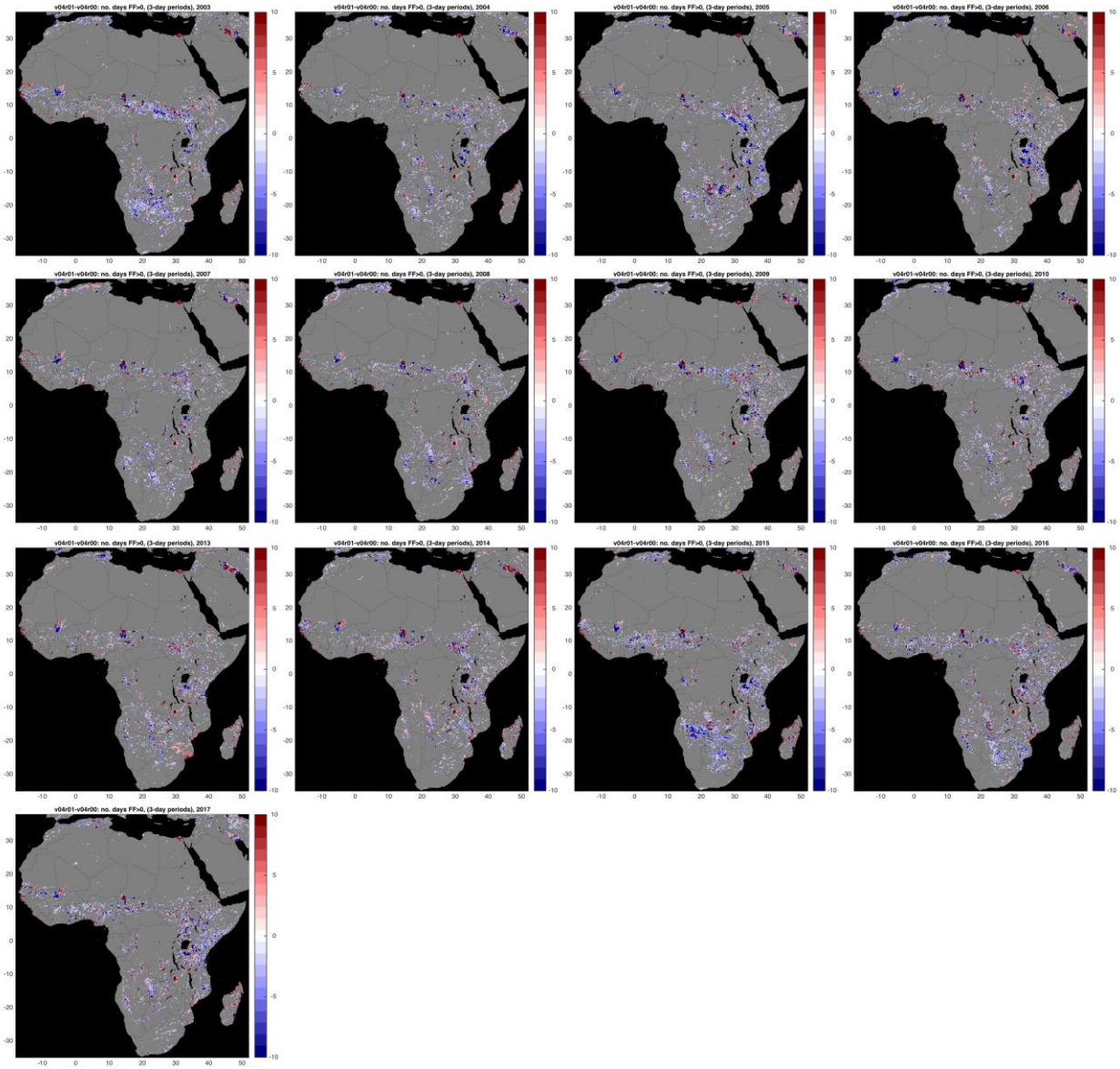


[Previous page] Four points in the Kalahari where the V04R01 algorithm's change in time series Q_{dry} estimation reduced or eliminated false positives by reducing peak flooded fraction values below MDFF.

B.3.5. Days flooded fraction exceeds MDFF maps

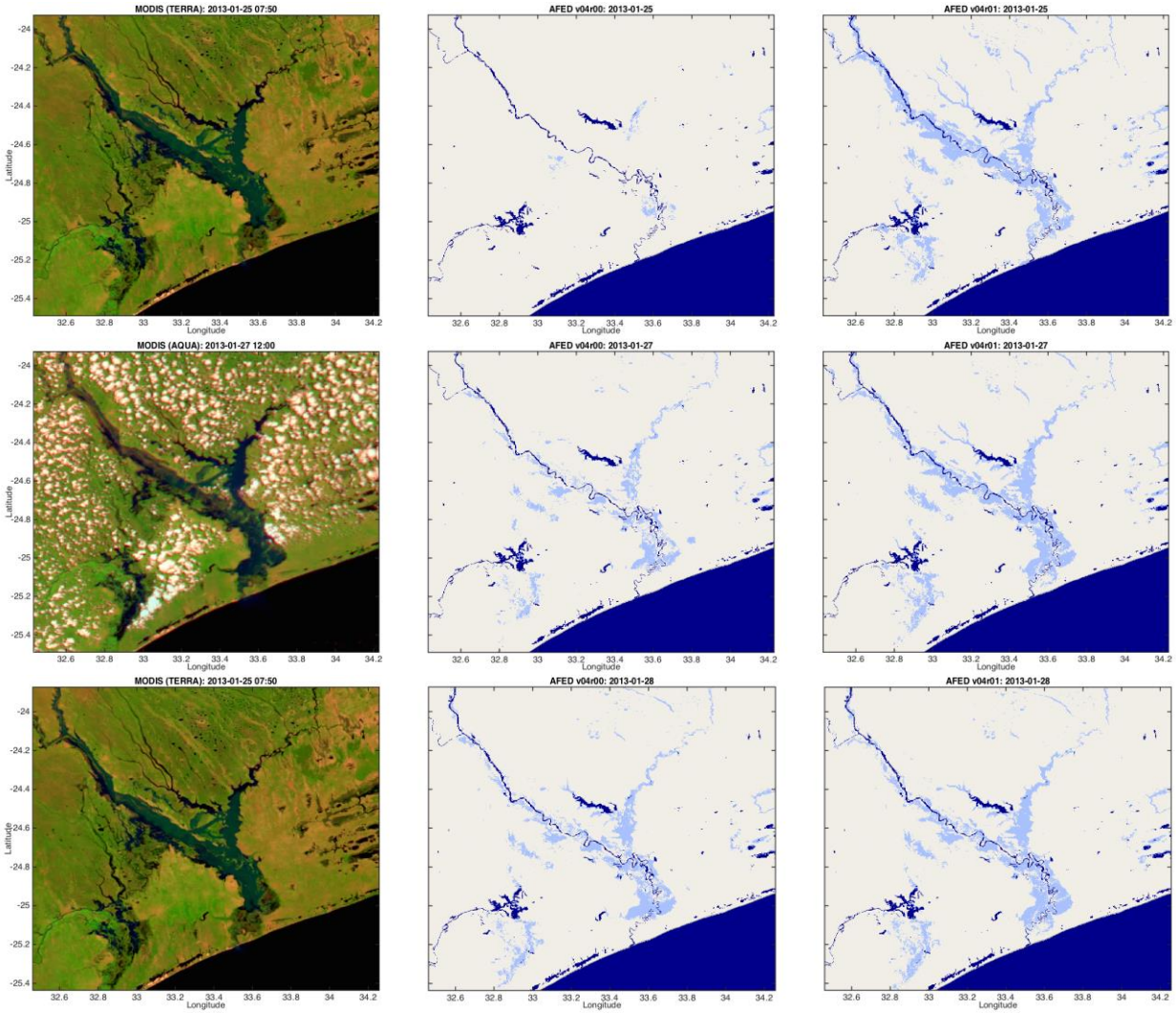
The maps below show the difference in number of days per year with FF>MDFF, V04R01 vs. V04R00. Only days within groups of at least three consecutive days with FF>MDFF are counted. The color scale is set to clip the difference at +/-10 days to highlight the total area with significant changes. There is an overall decrease in the incidence of days with FF>MDFF in all years and for all sensors. Notable decreases occur in the Sahel, Kalahari, The Horn of Africa, Kenya, and Tanzania where false positives have been found to occur. There are also areas with notable increases in number of days flooded, for example: the area of affected by flooding on the Limpopo in 2013, andalong the Senegal River in 2003, 2004, and 2007.



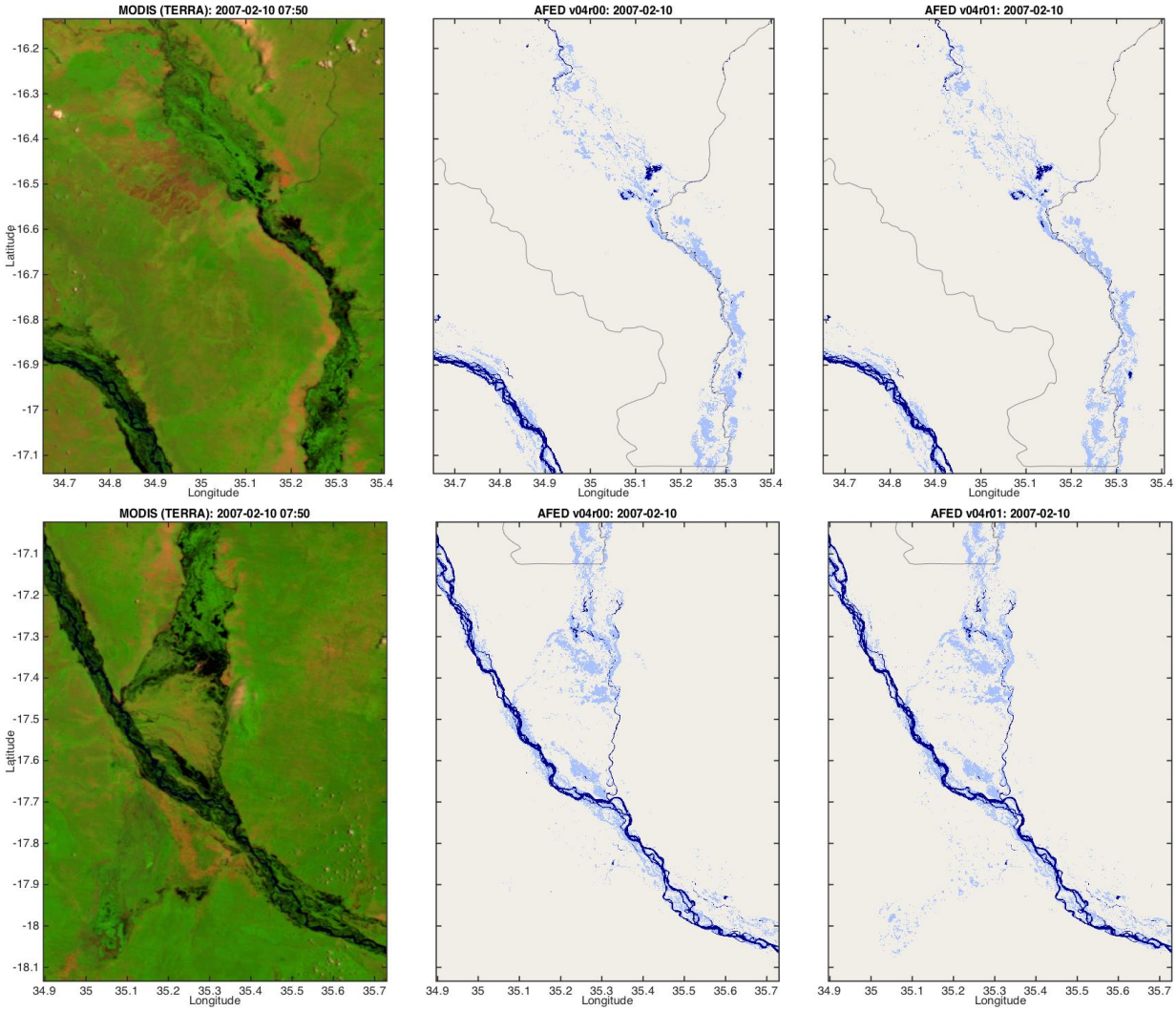


B.3.6. AFED compared to MODIS imagery

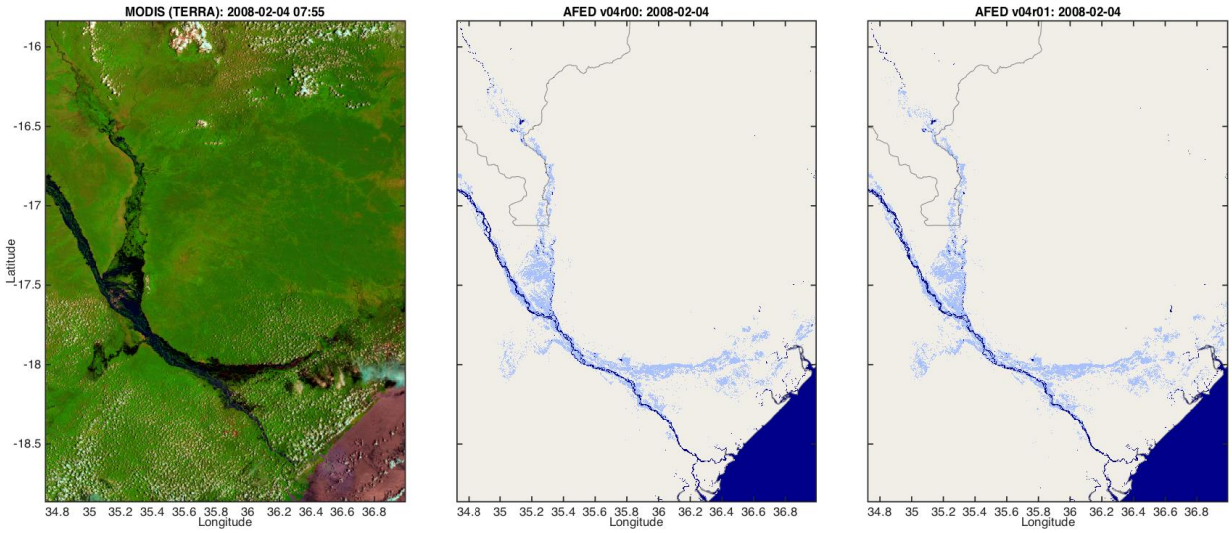
Below are cases studies comparing AFED V04R01 (right plot) to MODIS imagery and AFED V04R00 (center). See comments below each graphic.



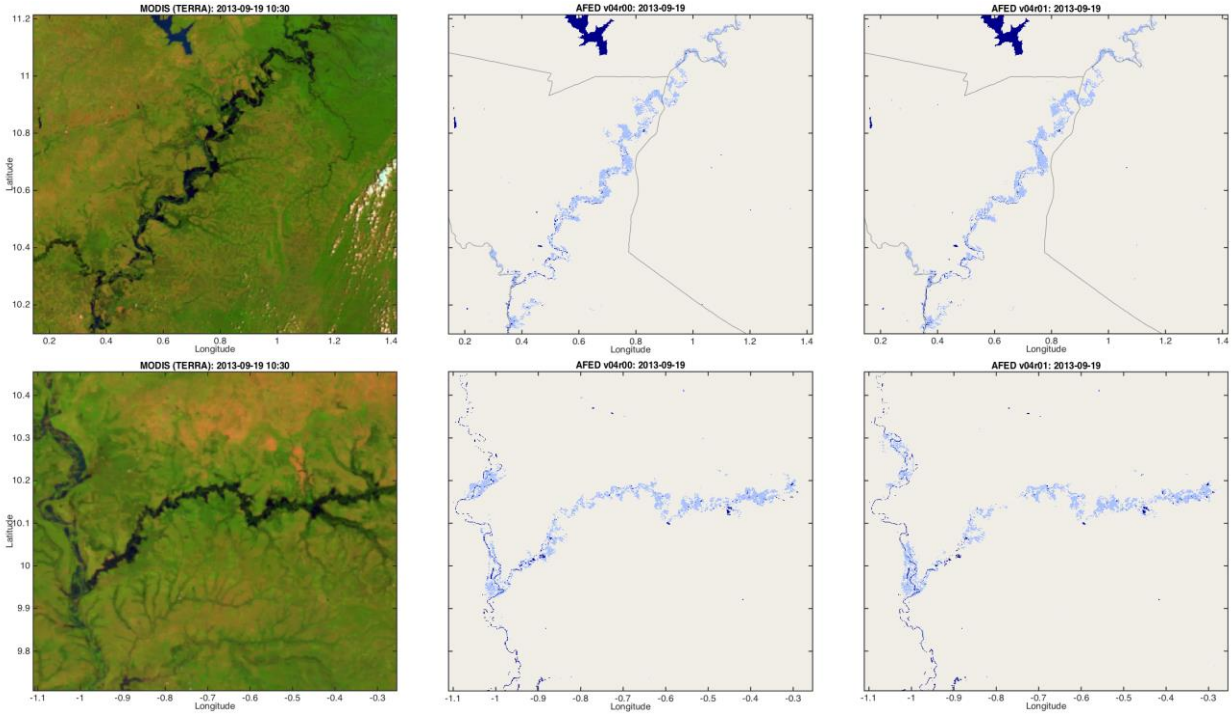
Flooding on the Limpopo River, Mozambique, 2013-01-25 (top), 2013-01-27 (middle), and 2013-01-28 (bottom). V04R01 is a better match to the MODIS imagery on all dates. On 01-25, false positive detections in the V04R00 algorithm prevented flooding from being fully mapped; V04R01 rectifies this issue. False positive detections did not affect either version on 01-27 or 01-28 but V04R01 more accurately maps the full flood extent as seen in the MODIS image, particularly near the coastline, which shows the benefits of better geolocation accuracy and lower MDFF.



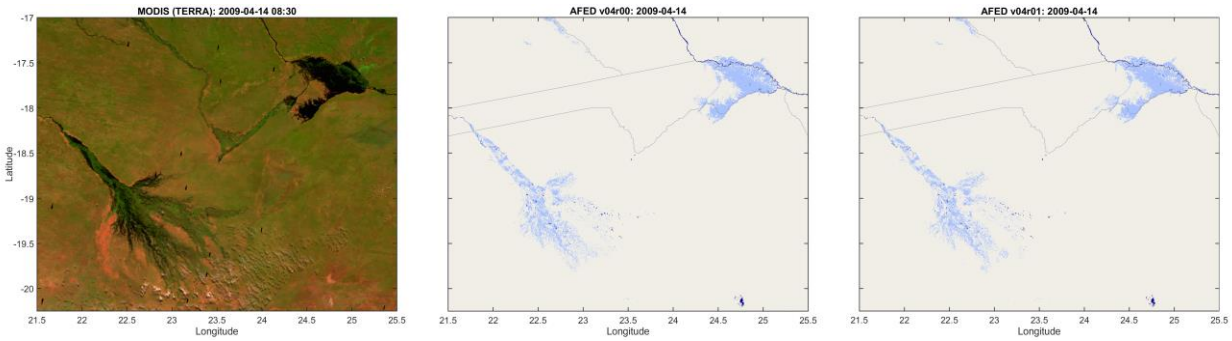
Flooding on the Shire and Zambezi Rivers, Mozambique, 2007. AFED V04R01 (right) is almost indistinguishable from V04R00 (center) except for the area south of the Zambezi where V04R01 detects and maps additional flooding that appears to be corroborated by the MODIS image.



Flooding on the Shire and Zambezi Rivers, Mozambique, 2008. Again, AFED V04R01 (right) is almost indistinguishable from V04R00 (center) except V04R01 depicts flooding closer to the coast due to V04R01 geolocation improvements and lower near-coast MDFF.



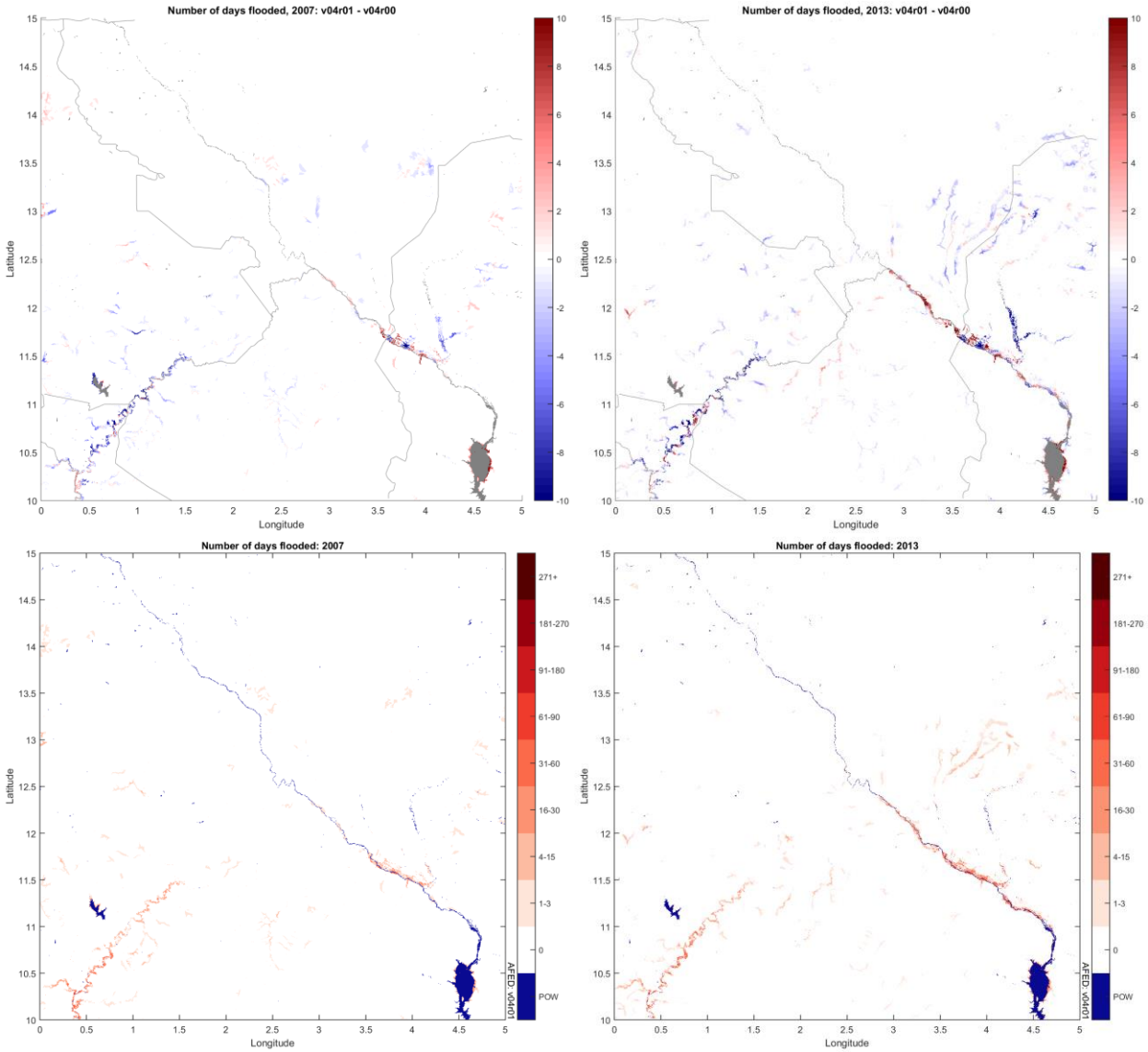
Flooding on the Pendjari/Oti River (top), which crosses the northern part of Togo, and the Volta River system in Ghana (bottom) on 2013-09-19. AFED V04R01 and V04R00 have minor differences but are generally alike and both are in good agreement with the flood extent apparent in the MODIS image.



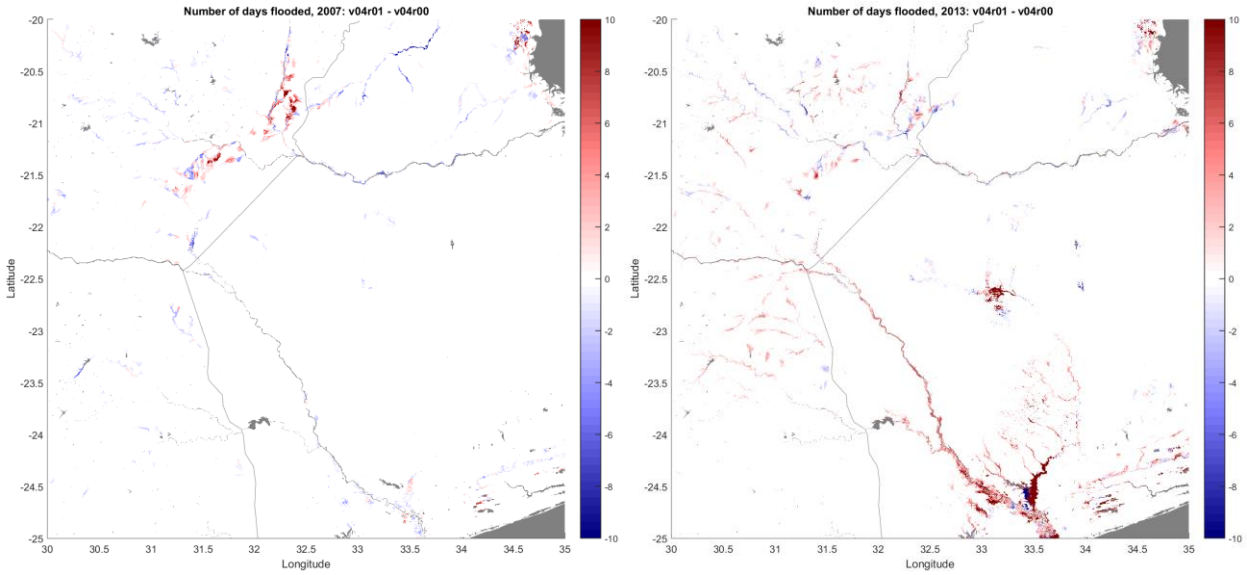
Flooding in the Okavango Delta (lower left) and Caprivi wetlands (upper right) on 2009-04-14. V04R01 and V04R00 have minor differences and both agree well with the MODIS imagery.

B.3.7. AFED number of days flooded per year

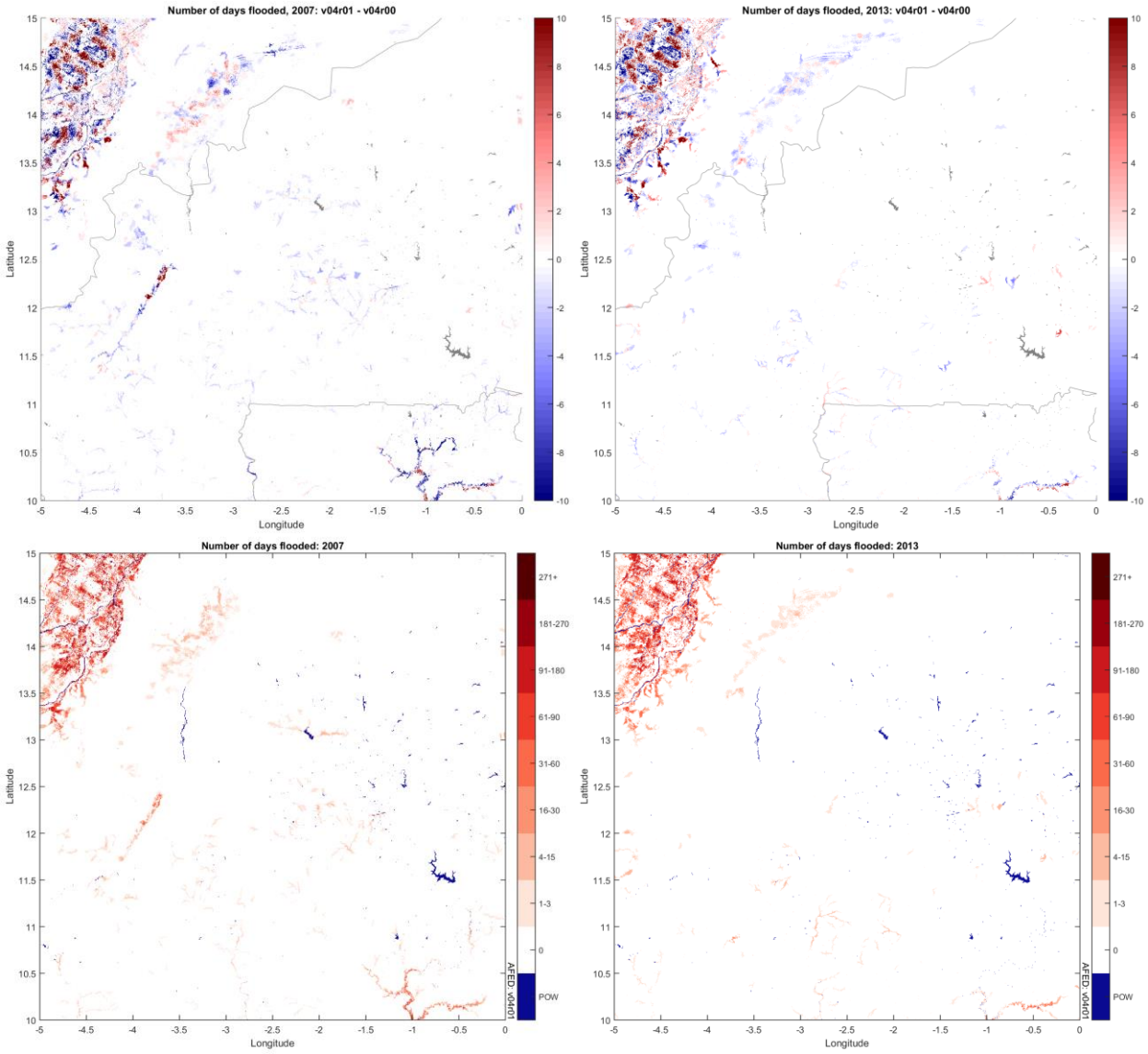
Below is a selected sample of maps comparing the number of days flooded per year, 2007 and 2013, V04R01 vs. V04r00. See comments below each graphic.



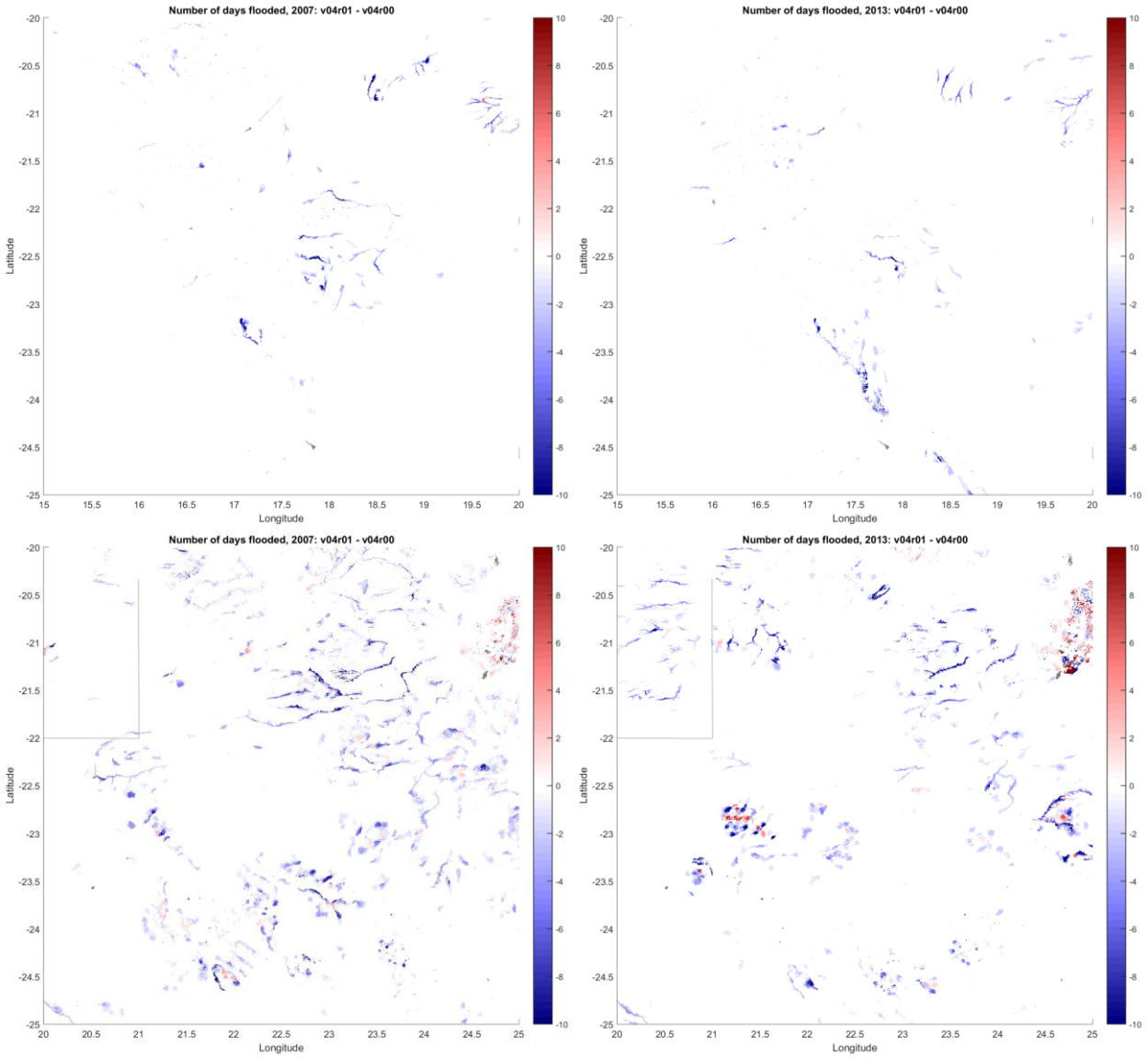
The V04R01-V04R00 difference maps (top) show that V04R01 mostly decreases the number of days flooded on the Pendjari/Oti River (lower left of frame) but there is a mix of increases and decreases on the Niger (right of frame). The difference maps also show that V04R01 reduces flooding in areas outside of the flood plains, which likely indicates a reduction in apparent false positives. The V04R01 number-of-days-flooded maps (bottom) indicate that maximum flooded extent remains well represented despite the reductions in days flooded.



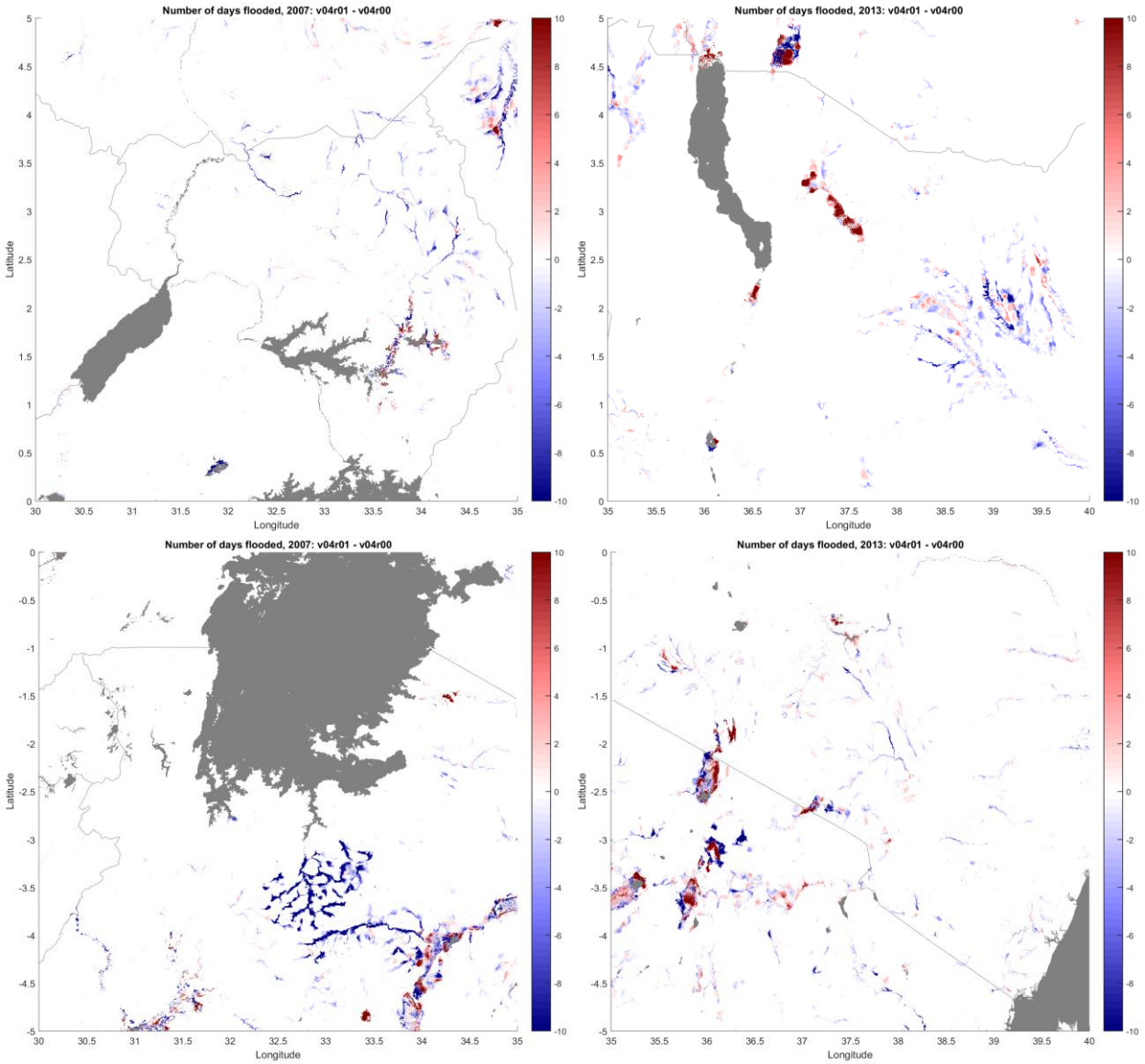
The map on the right shows the additional number of days with AFED-detected flooding in V04R01 during the 2013 Limpopo flood in Mozambique. In 2007 (left) there are relatively minor differences along the Limpopo.



The top left map shows a mix of increases and decreases in flooded days in V04R01 AFED on the Comoe River, Cote d'Ivoire (left center of frame) and in the Volta River system, Ghana (lower right of frame) in 2007. The bottom left map shows that V04R01 captures much of the total flood extent of these events despite the decreased number of days. The maps also indicate a general decrease in apparent false positives in V04R01 (e.g., flooding away from major flood plains).



These four examples show broad decreases in apparent false positives in the Kalahari region in 2007 (left) and 2013 (right).



These four tiles from include parts of Uganda and Kenya (top) and Tanzania and Kenya (bottom) prone to false positives. They show that AFED V04R01 generally decreases the number and extent of apparent false positives in these areas. One areas of significant false positive increase is in the upper right tile at 3° N, 37.5° E, corresponding to the Chalbi Desert. AFED indicates flooding in the area in other years as well. At least part of the area is a saltpan that turns into a shallow lake in the rainy season (http://www.50treasuresofkenya.org/chalbi-desert-huri-hills), so it is possible that AFED's representation is accurate.

C. Comparison of AFED V05R00 to V04R00 and V04R01

C.1. Introduction

This section summarizes flooded fraction test results demonstrating AFED performance changes for algorithm version V05R00 with respect to the prior two versions, V04R00 and

V04R01. Additional comparisons can be made by reference to previous versions of this document.

C.2. Days flooded fraction exceeds MDFF maps

The maps below show the difference in number of days per year with FF>MDFF, V05R00 vs. V04R00 (Figure 28) and V04R01 (Figure 29). Only days within groups of at least three consecutive days with FF>MDFF are counted. The color scale is set to clip the difference at +/- 10 days to highlight the total area with significant changes. Figure 30 shows histogram of the differences for 2017. Overall, V05R00 increases by a relatively small margin the number of flooded days compared to V04R01 but retains the more extensive systematic decrease between V04R00 and V04R01 in flooded days in false-positive prone areas.

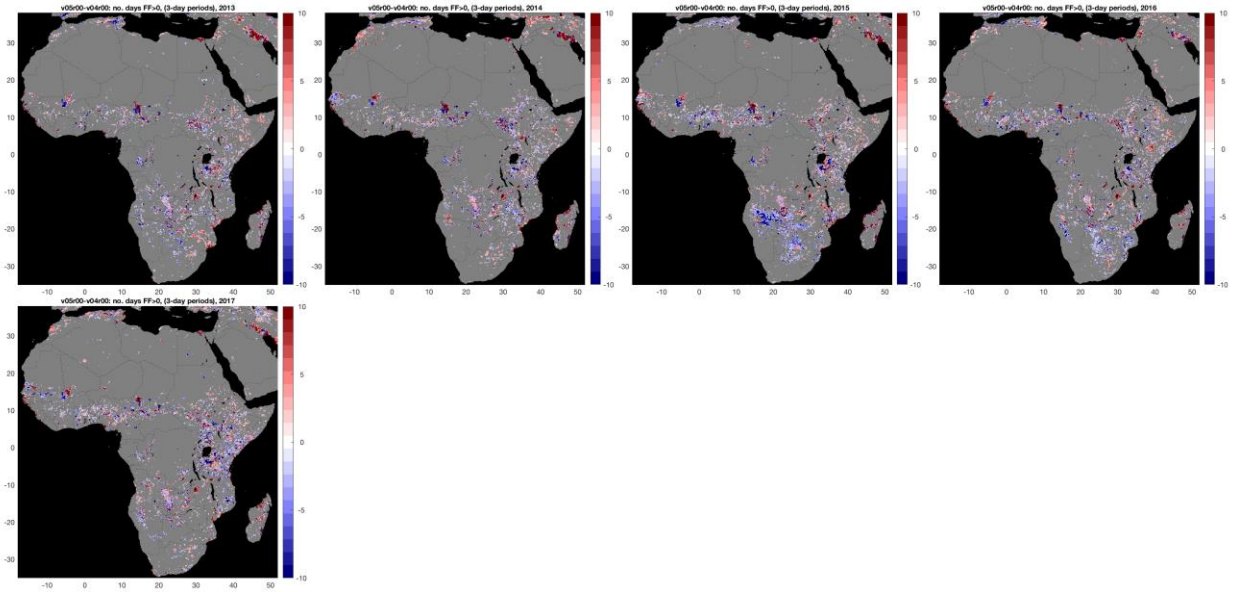


Figure 28: Difference in number of days flooded per year, 2013-2017, V05R00 – V04R00. Includes total days flooded less isolated single days and 2-day pairs (i.e., only days within 3-day periods).

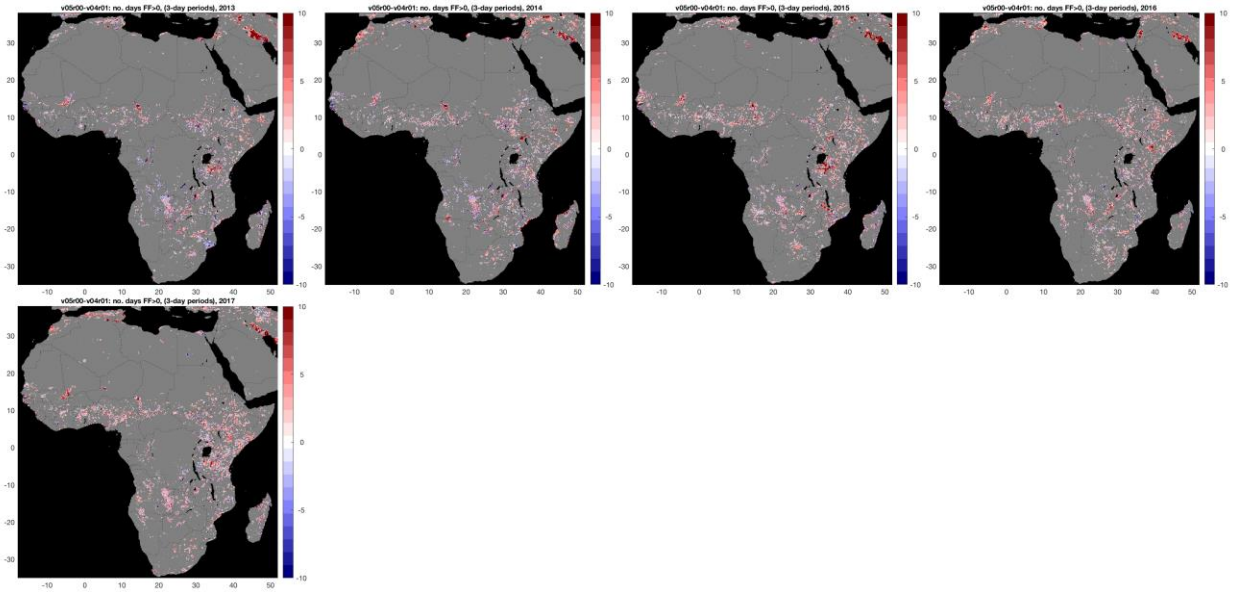


Figure 29: Difference in number of days flooded per year, 2013-2017, V05R00 – V04R01. Includes total days flooded less isolated single days and 2-day pairs (i.e., only days within 3-day periods).

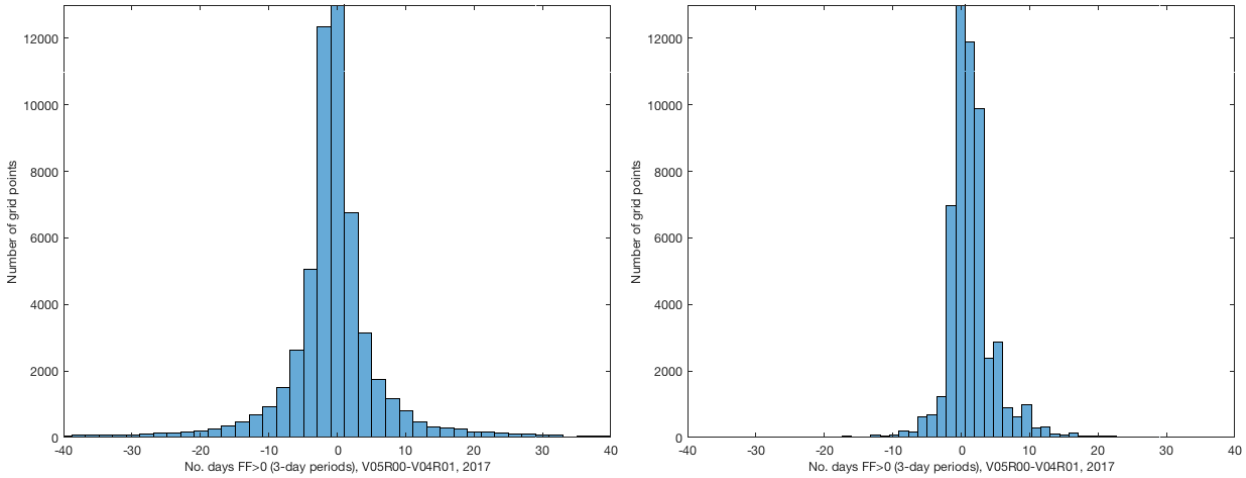


Figure 30: Histogram counts of difference in number of days flooded for 2017 between V05R00 and V04R00 (left) and V04R01 (right). Includes total days flooded less isolated single days and 2-day pairs (i.e., only days within 3-day periods).