Hydrol. Earth Syst. Sci., 21, 2685–2700, 2017
https://doi.org/10.5194/hess-21-2685-2017

Hydrology and Earth System Sciences
Open Access
EGU

A multi-sensor data-driven methodology for all-sky passive microwave inundation retrieval

Zeinab Takbiri[1], **Ardeshir M. Ebtehaj**[1], **and Efi Foufoula-Georgiou**[2,1]

[1]Department of Civil, Environmental and Geo- Engineering and St. Anthony Falls Laboratory, University of Minnesota, Twin Cities, Minneapolis, Minnesota, USA
[2]Department of Civil and Environmental Engineering, University of California, Irvine, California, USA

Correspondence to: Zeinab Takbiri (takbi001@umn.edu)

Received: 26 October 2016 – Discussion started: 1 November 2016
Revised: 14 March 2017 – Accepted: 7 May 2017 – Published: 8 June 2017

Abstract. We present a multi-sensor Bayesian passive microwave retrieval algorithm for flood inundation mapping at high spatial and temporal resolutions. The algorithm takes advantage of observations from multiple sensors in optical, short-infrared, and microwave bands, thereby allowing for detection and mapping of the sub-pixel fraction of inundated areas under almost all-sky conditions. The method relies on a nearest-neighbor search and a modern sparsity-promoting inversion method that make use of an a priori dataset in the form of two joint dictionaries. These dictionaries contain almost overlapping observations by the Special Sensor Microwave Imager and Sounder (SSMIS) on board the Defense Meteorological Satellite Program (DMSP) F17 satellite and the Moderate Resolution Imaging Spectroradiometer (MODIS) on board the Aqua and Terra satellites. Evaluation of the retrieval algorithm over the Mekong Delta shows that it is capable of capturing to a good degree the inundation diurnal variability due to localized convective precipitation. At longer timescales, the results demonstrate consistency with the ground-based water level observations, denoting that the method is properly capturing inundation seasonal patterns in response to regional monsoonal rain. The calculated Euclidean distance, rank-correlation, and also copula quantile analysis demonstrate a good agreement between the outputs of the algorithm and the observed water levels at monthly and daily timescales. The current inundation products are at a resolution of 12.5 km and taken twice per day, but a higher resolution (order of 5 km and every 3 h) can be achieved using the same algorithm with the dictionary populated by the Global Precipitation Mission (GPM) Microwave Imager (GMI) products.

1 Introduction

Capturing the diurnal spatiotemporal dynamics of inundation over coastal regions, deltaic surfaces, and river floodplains requires high-resolution observations in both time and space, which are not available from the typical sparse ground-based sensors. Satellite observations from the visible to the microwave bands of the electromagnetic spectrum have been widely used for mapping floods, estimating surface water storages, river discharge values, and water levels (Smith, 1997). In the visible bands ($\sim$0.4–0.8 µm), natural water reflects a fraction of incident light depending on the water depth and concentration of the optically active components such as suspended and dissolved particulate matter. However, water reflectivity sharply declines and approaches zero in the near-infrared bands ($\sim$0.8–2.5 µm). Thresholding of this sharp gradient is often used to discriminate water bodies from their nearby dry soils and vegetated surfaces (Rango and Anderson, 1974; Smith, 1997, 2001, and references therein; Frazier and Page, 2000; Jain et al., 2005). In the microwave region of the spectrum, the dielectric constant of water ($\sim$80) is much higher than the dry soil ($\sim$4) and thus the inundated areas are substantially less emissive and radiometrically colder than the surrounding soils and vegetation covers. Moreover, emission from smooth water surfaces is more polarized than that from rough soils and vegetated surfaces (Ulaby et al., 1982; Papa et al., 2006; Prigent et al., 2007). This polarization signal has also been used through empirical thresholding approaches to distinguish water surfaces from other land surface types (Allison et al., 1979; Sippel et al., 1994, 1998; Brakenridge et al., 2005, 2007).

Flood mapping from space was first accomplished using visible to near-infrared (VNIR) observations (0.4–1.1 μm), by the multispectral scanner system (MSS) sensors on board Landsat-1 (Rango and Anderson, 1974; Rango and Salmonson, 1974; McGinnis and Rango, 1975). In these pioneering works, flooded areas were mapped where the near-infrared surface reflectance was below a certain threshold as water absorption is strong in this region. More recently, Brakenridge and Anderson (2006) showed that the visible red band 1 (0.62–0.67 μm) and near-infrared (NIR) band 2 (0.84–0.87 μm) from the Moderate Resolution Imaging Spectroradiometer (MODIS) aboard the Terra and Aqua satellites can be used to detect water over land surfaces. They mapped several hundreds of flood events at different sites all over the world by classification of water via thresholding over the NIR band and the normalized difference vegetation index, $\mathrm{NDVI} = (\mathrm{NIR} - \mathrm{red})/(\mathrm{NIR} + \mathrm{red})$ introduced by Rouse et al. (1974). To better discriminate the vegetation from inundated areas in threshold-based methods, Ticehurst et al. (2013) and Guerschman et al. (2011) used a new index – called the normalized difference water index, $\mathrm{NDWI} = (\mathrm{red} - \mathrm{MIR})/(\mathrm{red} + \mathrm{MIR})$ introduced by Gao (1996) and later modified to $\mathrm{MNDWI} = (\mathrm{green} - \mathrm{MIR})/(\mathrm{green} + \mathrm{MIR})$ by Xu (2006). This index exploits the mid-infrared (MIR; 1.23–1.25 μm) part of the spectrum to improve the mapping. In all thresholding methods, the shadows of terrains and clouds are usually miss-classified as inundated areas. Therefore, Kuenzer et al. (2015) used the topography and cloud information data as ancillary variables to obtain improved estimates of the interannual dynamics of areas covered with water over five deltaic regions with high annual cloud cover.

The use of passive microwaves (PMW) to map flooded areas was pioneered by Allison et al. (1979), Giddings and Choudhury (1989), and Choudhury (1991). Allison et al. (1979) used horizontal polarization of brightness temperatures (Tb) at 19.3 GHz, from the electrically scanning microwave radiometer (ESMR) on board the Nimbus-5 satellite, to delineate flooded regions in Australia. Giddings and Choudhury (1989) reported the 37 GHz vertical and horizontal polarization differences (i.e., $\mathrm{Tb}_{37\mathrm{v}} - \mathrm{Tb}_{37\mathrm{h}}$), from the Scanning Multi-frequency Microwave Radiometer (SMMR) on board the Nimbus-7 satellite, as the most responsive channel to identify the seasonal changes in the extent of floodplains over South America. Temimi et al. (2005) used the empirical basin wetness index (BWI) defined by Basist et al. (1998), to obtain real-time water surface fraction (WSF) in the Mackenzie River basin, using multi-frequency information at 19, 37, and 85 GHz. To minimize the contamination effects of atmospheric emission and variations of surface temperatures, Brakenridge et al. (2007) exploited the ratio of Tb values over inundated and dry surfaces at 36 GHz and presented promising results over several river sites all over the globe, using the PMW observations by the Advanced Microwave Scanning Radiometer – Earth Observing System (AMSR-E). De Groeve (2010) also used the same method and instrument to map floods for several hundreds of locations for the Global Disaster Alert and Coordination System.

While visible and shortwave-infrared bands often provide sub-kilometer resolution for inundation mapping, their capability is very limited in a cloudy sky. This limitation is usually very restrictive over prone-to-flooding watersheds and deltas in tropical regions with a high-frequency of heavy precipitation events. For instance, a long-term analysis of Landsat data revealed that due to cloud contamination, only 30 % of overpasses are useful for inundation mapping (Melack et al., 1994). Because of this limitation, most of the related satellite products, including the MODIS inundation products, are available mostly in monthly, seasonal, and/or annual timescales (Ordoyne and Friedl, 2008). However, microwaves can penetrate clouds – and to some extent hydrometeors in frequencies ≤ 37 GHz – to provide water inundation mapping in almost all weather conditions. Unfortunately, due to the coarse resolution of microwave data, e.g., (47×74) km^2 at 19 GHz to (13×16) km^2 at 183 GHz for the SSMIS), only large water bodies can be detected and sub-pixel inundated areas cannot be directly identified (Smith, 1997). Presently, there exist several sensors on board different satellites that overlap in the spatial and time domains that sample land–atmosphere signals at different wavelengths of the electromagnetic spectrum. Therefore, it is imperative to integrate these multi-sensor observations to overcome their individual shortcomings and improve retrievals of land–atmosphere parameters and the extent of flooded areas (Prigent et al., 2001, 2007; Crétaux et al., 2011; Temimi et al., 2011; Schroeder et al., 2010).

In this paper, we develop a method to retrieve sub-pixel inundation fraction ("inundation" referring to regions where water covers the land surface, excluding permanent water bodies) only from passive microwave observations based on a set of paired VNIR and passive microwave training samples. In particular, as training samples, we use global observations of VNIR data from the MODIS on board Terra (launched in 2000) and Aqua satellites (launched in 2002) and passive microwave data from the Special Sensor Microwave Imager and Sounder (SSMIS) on board Defense Meteorological Satellite Program (DMSP) satellites F16–F18. Several years of observations (2000–present) by these two sensors allow us to collect adequate overlapping data to link coarse-scale SSMIS passive microwave data to high-resolution MODIS VNIR data in the form of an organized dataset. Obviously, this collection of almost coincident observations does not contain direct information about surface inundation in a cloudy sky, as the radiative signals in VNIR wavelengths cannot penetrate clouds. However, over land, it is well understood (see Ferraro et al., 1986; Grody, 1991; Wilheit et al., 1994) that hydrometeors and the atmospheric profile do not significantly affect the low-frequency < 60 GHz brightness temperatures. Therefore, the information content of the dataset over low-frequency channels is independent of the atmospheric profile and can be used to a

good degree of accuracy to recover inundated surfaces under cloudy conditions as well. It should be acknowledged that there is an uncertainty for the inundation retrieval under heavy rainy/cloudy skies when only the information in the clear-sky dataset is used. However, we expect that this uncertainty will be small since the information of the underlying surfaces in low-frequency channels of the collected dataset remains almost the same over different atmospheric conditions.

The collected dataset has a large number of linked pairs of inundation fraction data from MODIS data SSMIS multi-frequency brightness temperatures. For algorithmic development, the dataset is organized into two fat matrices: the brightness temperature and inundation dictionaries. For an observed pixel-level brightness temperature, the proposed passive retrieval algorithm uses the nearest-neighbor search to isolate a few vectors in the dictionary of brightness temperatures and their corresponding inundation fraction and then use them to estimate the unknown inundation fraction. The proposed retrieval algorithm is applied to estimate daily inundation fraction at spatial resolution of 12.5 km over the Mekong in 2015. The main motivation for selecting this delta as a case study is that approximately 90 % of the Mekong region is covered by clouds during the rainy season (Leinenkugel et al., 2013), which severely hampers the use of inundation mapping in the VNIR bands. We retrieve the inundation fraction twice per day using the proposed algorithm over the Mekong Delta and compare the results with the flood products of VNIR data during clear skies. We also evaluated the results against the daily and monthly water level data obtained from 11 gauges over the Mekong Delta (Fig. 1) to examine consistency of the retrievals with the regional inundation patterns.

This paper is organized as follows. Section 2 explains the a priori dataset and the formation of the dictionaries and Sect. 3 provides detailed information about the retrieval algorithm. Implementation of the method and validation are explained in Sect. 4. Section 5 presents concluding remarks and directions for future research.

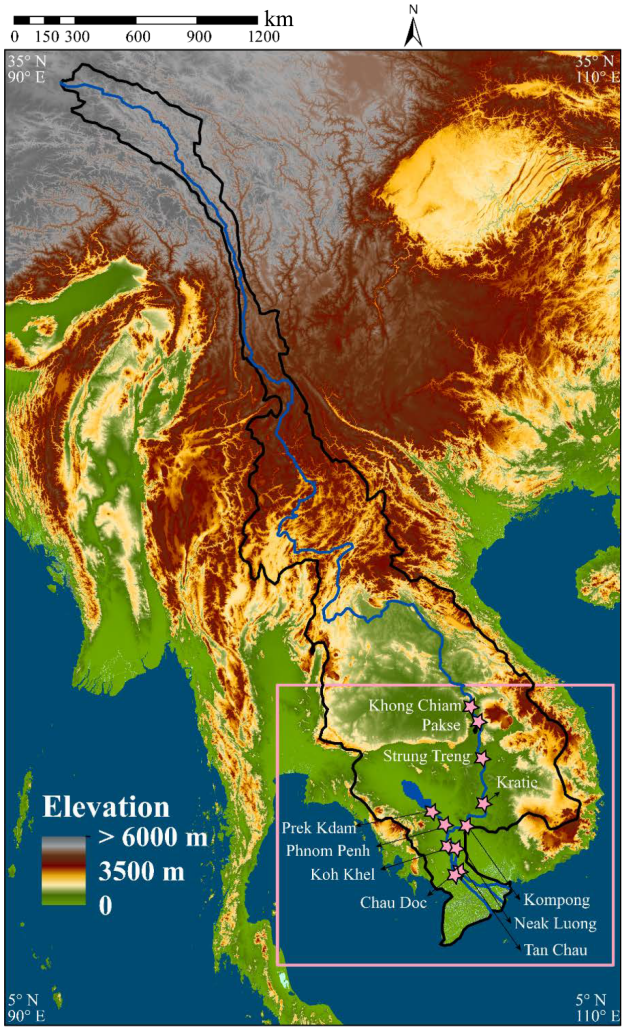

Figure 1. Map and digital elevation of the Mekong River basin (area = 795 000 km^2) and its delta. The study area is delineated by a pink rectangle. The 11 stations (from Mekong River Commission) that monitor the water level are also marked by pink stars.

2 Study area and dataset

The 60 000 km^2 Mekong Delta is in South Vietnam (see Fig. 1) with a tropical monsoon climate system. The delta with its agricultural industry is one of the most important sources of food supply to Southeast Asia. This critical region is home to nearly 20 million people, approximately 22 % of the population of Vietnam, and is one of the most densely populated regions in the world. The area has been exposed to exacerbated erosion due to human activities and increased sea level rise and lowland flood events in the recent decades (e.g., Syvitski et al., 2005; Ericson et al., 2006; Nicholls and Cazenave, 2010; Tessler et al., 2015). Improved quantification of (near) real-time inundation of the Mekong Delta can help (1) to improve flood forecasting by identifying the inundated and thus soil saturated zones and (2) to identify erosional and depositional hotspots that can improve geomorphologic and ecosystem modeling. The proposed retrieval algorithm is applied to estimate sub-daily inundation fraction at resolution of 12.5 km over some of the lower regions of the Mekong Delta in calendar year 2015 (Fig. 1).

Two sources of information are used to build a dataset that connects almost coincident VNIR water inundation data and multi-frequency passive microwave data. The VNIR data consist of the daily NASA standard MODIS near-real-time (NRT) water product (MWP-3D3ON; i.e., 3-Days imagery, three observations, and no shadow masking) with approximately 250 m spatial resolution (Nigro et al., 2014) from both Terra and Aqua satellites. The Terra and Aqua satellites both have a sun-synchronous orbit. They rotate around

the earth in opposite directions: Terra has an ascending orbit with the local equatorial crossing time of 10:30 LT and Aqua has a descending orbit with the local equatorial crossing time of 01:30 p.m. MWP products are binary information of inundation based on the Dartmouth Flood Observatory (DFO) algorithm, which uses a thresholding scheme on MODIS observations at band 1 (0.62–0.67 µm), band 2 (0.84–0.87 µm), and band 7 (2.10–2.15 µm). To minimize the contamination effects of cloud and terrain shadows, we focus on 3-day composite MWP products (3D3ON). Clearly, the use of the 3-day composite MODIS-MWP data can affect daily inundation retrievals; however, in the context of the presented algorithm this is the best choice because, daily MODIS-MWP composites are very uncertain due to the terrain shadows and clouds (Nigro et al., 2014). Typically, there are numerous missing pixels in the daily products, which reduce the sample size dramatically. These errors are significantly reduced in 3-day composite products, as it is less likely that clouds (and their shadows) stay at the same spot during a 3 day period (Nigro et al., 2014).

The microwave data are obtained from the DMSP SSM/I-SSMIS Pathfinder Daily Equal-Area Scalable Earth Grid (EASE-Grid; see Armstrong and Brodzik, 1995) brightness temperatures distributed by the National Snow and Ice Data Center (NSIDC). These datasets are at four central frequencies 19, 22, 37, and 91 GHz. All channels are vertically and horizontally polarized except channel 22 GHz. The effective resolution of the highest frequency channel is ∼ 12.5 km while low-resolution channels are projected onto a grid size of ∼ 25 km. DMSP SSM/I-SSMIS brightness temperature data products are from observations by the SSM/I and SSMIS radiometer on board the DMSP F8, 11, 13, or 17. Since December 2006, the F17 satellite has been the only operational satellite from the DMSP series, which carries on board the SSMIS instrument with equatorial crossing times of 05:30–06:30 a.m. and 17:30–18:30 p.m. for the descending and ascending orbits, respectively. It is important to note that because these satellites revisit every point on Earth at the same local time, repeatedly, the paired MODIS-MWP with DMSP SSMIS data have a fixed diurnal time difference in the entire dataset. Since the MODIS-MWP data are from the combination of Terra and Aqua observations, their time tag is advantageous in the sense that it allows us to enrich the number of samples for the diurnal cycle of inundation dynamics.

The first step for building the a priori dataset is to match the different space–time resolutions of the multi-sensor information. To unify the spatial resolution of the microwave data, the brightness temperatures of the three lower-frequency channels are mapped onto the latitude–longitude grids of the high-frequency channel of 91 GHz with a resolution ∼ 12.5 km, using a nearest-neighbor interpolation. Then the clear-sky MWP data are also upscaled from 250 m to 12.5 km and projected onto the same grids. In the process of upscaling the binary MWP data, we assigned to each upscaled pixel a scalar inundation fraction number f that represents the ratio of the number of inundated sub-pixels to the total number of sub-pixels within a pixel size of 12.5 km. For matching the timescales of Tb and MWP values, the Tb values are averaged over a 3-day time window to minimize the possible effects of cloud contamination in the VNIR data. Figure 2 demonstrates schematically the process of producing the explained dataset.

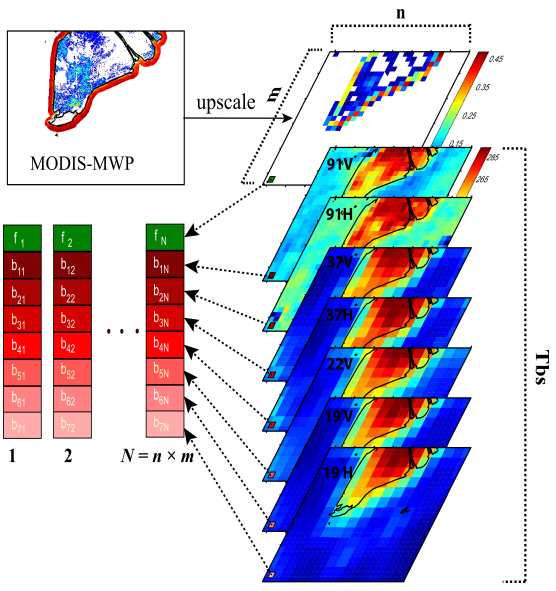

Figure 2. A schematic showing construction steps of the a priori dataset for dictionaries. The top slab is the upscaled MODIS-MWP and the other slabs are the brightness temperature data at seven frequency bands. Each vector on the left is created by stacking a pixel-level information of the multi-frequency brightness temperatures by the SSMIS radiometer and their corresponding inundation fractions from the MWP product at 12.5 km resolution. This process is repeated for each orbit to generate a large number of vectors and form separate dictionaries for ascending and descending orbits using all satellite overpasses in 5 years from 2010 to 2014. $N = n \times m$ is the number of collected vectors for 1 day in a year. The same process is conducted for each day in 5 years (2010–2014) to create the dictionaries with $M = \sum_{i=1}^{5\times365} N_i$ vectors.

3 The retrieval algorithm

The proposed retrieval algorithm uses the link between two available coincidental datasets, passive microwave (SSMIS) and VNIR (MODIS-MWP), to retrieve inundation in the cloudy days. First, the overlapped clear-sky pixels of MODIS-MWP and SSMIS for 5 years (2010–2014) are collected over the study area to create two coincidental dictionaries: the SSMIS dictionary and the MODIS-MWP dictionary. The SSMIS dictionary consists of 8-dimensional vectors of brightness temperature (Tb), where 8 is the number of frequency channels, and the MODIS-MWP dictionary con-

sists of scalar values of inundation fractions for each corresponding pixel in Tb. In other words, the inundation fraction for each Tb in the brightness temperature dictionary is known. The algorithm uses the information embedded in these two dictionaries to estimate the unknown inundation fractions for each Tb observation vector. First, it searches the brightness temperature dictionary to find the K most similar vectors in the Euclidean sense to the Tb observation vector through the K-nearest-neighbors algorithm. Then, for these K-nearest-neighbors, the corresponding known scalar values in the inundation fraction dictionary are picked. If the ratio of the number of inundated vectors in K-nearest neighbors is greater than a threshold (which will be explained later), this pixel is called inundated and the algorithm goes to the estimation step. In the estimation step, the coefficients that can optimally estimate the Tb observation vector based on its K-nearest neighbors are calculated through a least-squares regularization approach. Those coefficients are then used to linearly combine the K known inundation fractions that are associated with the neighboring Tb vectors for calculating the unknown inundation fraction. The above detection and estimation steps are repeated for each orbit at a pixel level of 12.5 km over the study area. The algorithm is mathematically described in what follows.

To organize the dataset in an algebraically tractable manner, M vectors of microwave brightness temperatures $\boldsymbol{b}_i = (\mathrm{Tb}_{1i}, \mathrm{Tb}_{2i}, \ldots, \mathrm{Tb}_{ni})^{\mathrm{T}} \in \Re^n$ at n frequency channels are collected. These vectors form the column space of an n-by-M matrix $\mathbf{B} = [\boldsymbol{b}_1|\boldsymbol{b}_2|\ldots|\boldsymbol{b}_M] \in \Re^{n\times M}$, called a brightness temperature dictionary, where $M \gg n$. Analogously, the corresponding inundation fraction values $\{f_i\}_{i=1}^{M}$ can be collected in the column space of the inundation dictionary $\mathbf{F} = \left[f_1|f_2|\ldots|f_M\right] \in \Re^{1\times M}$. For each vector $\boldsymbol{b}_i$ in the dictionary of brightness temperatures there is an inundation fraction f_i from MODIS-MWP. The collection of these pairs from historical observations forms the two dictionaries $\mathbf{B}$ and $\mathbf{F}$. The algorithm follows two sequential steps: a detection and an estimation step. In the detection step, for each observed vector of brightness temperature $\boldsymbol{b}_{\mathrm{obs}}$, the algorithm first finds its K-neighboring brightness temperatures in $\mathbf{B}$ in the Euclidean sense and stores them in the column space of $\mathbf{B}_{\mathrm{s}} \in \Re^{n\times K}$. Then, knowing the column indices of the neighboring brightness temperatures, it isolates their corresponding inundation fraction values in $\mathbf{F}_{\mathrm{s}} \in \Re^{1\times K}$. In this step, if at least $p \times K$ number of nearby inundation fraction values in $\mathbf{F}_{\mathrm{s}}$ are nonzero, the algorithm assumes that $\boldsymbol{b}_{\mathrm{obs}}$ is over an inundated pixel and attempts to estimate the fraction of inundation in the estimation step. Here, $p \in (0-1)$ is the detection probability parameter. It should be also noted that the K-nearest-neighbor algorithm in this paper does not directly constrain its search to any specific time or location. In other words, for every pixel-level vector of Tb, the K-nearest-neighbors algorithm searches the entire dictionary regardless of any specific time or spatial coherency.

In the estimation step, the method assumes that $\boldsymbol{b}_{\mathrm{obs}}$ can be estimated by a linear combination of a few column vectors of $\mathbf{B}_{\mathrm{s}}$ as follows:

$$\boldsymbol{b}_{\mathrm{obs}} = \mathbf{B}_{\mathrm{s}}\boldsymbol{c} + \boldsymbol{e}, \tag{1}$$

where the vector $\boldsymbol{c} \in \Re^K$ contains a set of representation coefficients to be estimated and $\boldsymbol{e} \in \Re^n$ is the error vector. Clearly, for an observed vector of brightness temperatures $\boldsymbol{b}_{\mathrm{obs}}$, the goal is to estimate its unknown inundation fraction value $\hat{f}$. We assume that the two paired dictionaries $\mathbf{B}_{\mathrm{s}}$ and $\mathbf{F}_{\mathrm{s}}$ represent similar manifolds in a geometric sense that their local structures can be approximated well with the same linear model. This allows us to assume that the representation coefficients in vector $\boldsymbol{c}$ from Eq. (1) can be used to estimate the inundation fraction $\hat{f}$ as follows:

$$\hat{f} = \mathbf{F}_{\mathrm{s}}\boldsymbol{c}. \tag{2}$$

As a result, using a classic-weighted least-squares method, the representation coefficients $\boldsymbol{c}$ can be estimated as

$$\hat{\boldsymbol{c}} = \mathrm{argmin}_c \left\{ \|\mathbf{W}\left(\boldsymbol{b}_{\mathrm{obs}} - \mathbf{B}_{\mathrm{s}}\boldsymbol{c}\|_2^2\right)\right\}, \tag{3}$$

where $\mathbf{W}$ is a weight matrix (to be discussed later in this section) that characterizes the importance of each channel in the retrieval scheme. The number of K-nearest neighbors is often larger than the number of frequency channels, $k \gg n$, making $\mathbf{B}_{\mathrm{s}}$ a rank-deficient matrix and the above problem ill-posed. To make the optimization problem (Eq. 3) well-posed, we use a mixed ℓ_1–ℓ_2-norm regularization as follows:

$$\begin{aligned} &\boldsymbol{c} = \mathrm{Argmin}_c \left\{ \|\mathbf{W}\left(\boldsymbol{b}_{\mathrm{obs}} - \mathbf{B}_{\mathrm{s}}\boldsymbol{c}\right)\|_2^2 + \lambda_1\|\boldsymbol{c}\|_1 + \lambda_2\|\boldsymbol{c}\|_2^2 \right\} \\ &\text{subject to } \boldsymbol{c}0,\ 1^{\mathrm{T}}\boldsymbol{c} = 1, \end{aligned} \tag{4}$$

which has been successfully used for passive microwave precipitation retrievals (Ebtehaj et al., 2015a, b). The nonnegativity of the coefficients assures positivity of the brightness temperatures and the sum-to-one constraint enforces an unbiased estimation. The regularization involves both the ℓ_1-norm $\|\boldsymbol{c}\|_1 = \sum_{i=1}^{K} |c_i|$ and the ℓ_2-norm $\|\boldsymbol{c}\|_2 = (\sum_{i=1}^{K} |c_i|^2)^{\frac{1}{2}}$. The parameters λ_1 and λ_2 in Eq. (4) are regularization parameters that enforce a trade-off between the two regularizations ℓ_1 and ℓ_2. In this mixed regularization, the ℓ_1-norm leverages sparsity in the solution (i.e., forces some of the elements of $\boldsymbol{c}$ to be zero) while the ℓ_2-norm increases the stability of the solution as the neighboring brightness temperatures in $\mathbf{B}_{\mathrm{s}}$ are likely to be highly correlated (see Zou and Hastie, 2005). In effect, due to the use of a mixed regularization, this regularization promotes group sparsity (i.e., some blocks of the representation coefficients are zero) while it keeps the solution sufficiently stable. In other words, it acknowledges the fact that there are a few clusters of nearby brightness temperatures that can properly explain the observation. By enforcing the ℓ_1-norm we select vectors that are parts of clusters of brightness temperatures, while the ℓ_2-norm handles

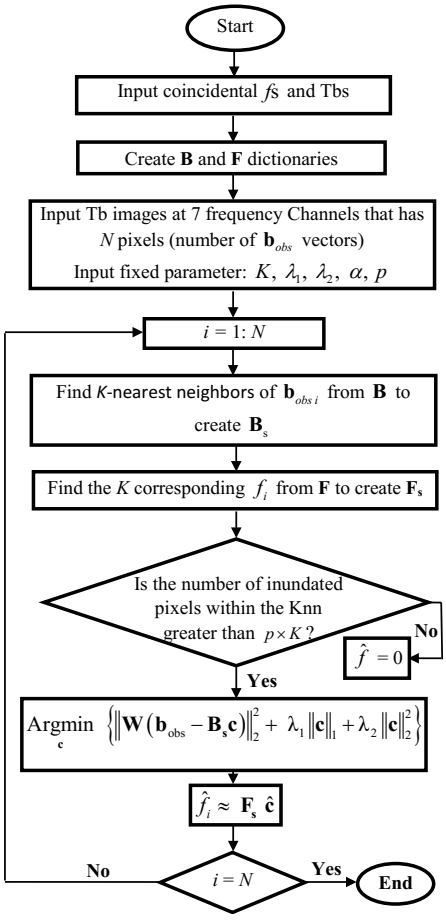

Figure 3. Flowchart of the inundation retrieval algorithm for N pixels in each orbit, where Knn stands for the K-nearest neighbor. See text for definitions of the notations and detailed explanation.

the potential correlation between those clustered neighbors and makes the problem sufficiently stable. The proposed algorithm is summarized in a flowchart shown in Fig. 3.

As previously noted, in the current implementation of the proposed retrieval algorithm, we focus on (almost) coincidental observations of the brightness temperatures and inundation fractions by the SSMIS and MODIS instruments, respectively. The dictionaries **B** and **F** are constructed using 5 years of overlapping data (2010–2014) over the Mekong Delta (latitude: 0–10° N and longitude: 100–110° E) at 12.5 km grid resolution (Fig. 1).

To build the dictionary, only the clear-sky MODIS-MWP products were considered. At resolution 12.5 km, we labeled a pixel as clear-sky when less than 50 % of the VNIR data at resolution 250 m is flagged as non-cloudy. Because the MODIS sensor has a much higher resolution than the footprint of SSMIS and because the number of cloud-free samples over the Mekong are very limited, a threshold above zero is deployed to keep a certain number of partially cloudy pixels and make sure that the dictionary will not be undersampled. For choosing the threshold, we conducted some sensitivity analysis (not shown here) and found a 50 % threshold, as a fair probability choice, results in a minimum of potential biases.

Since the DMSP satellites have two different equatorial crossing times, here, we use two sets of dictionaries for Tb values in the ascending (day or morning) and descending (night or evening) orbits. From all the available coincident observations, we randomly chose 2×10^6 pairs of brightness temperatures and inundation fractions in each ascending and descending dictionary. The purpose of stratifying the dictionaries into ascending and descending orbits is to exclude the effects of Tb modulations from the retrieval process caused by the systematic diurnal variation of surface temperature. In other words, the same inundation fraction has different PMW spectral signature in a daytime versus a nighttime overpass largely due to the diurnal variability of skin temperature, precipitation, and soil moisture (see Mears et al., 2002; Ramage and Isacks, 2003; Norouzi et al., 2012). Figure 4a presents the systematic difference between the Tbs of the ascending versus descending tracks for various ranges of pixel-level inundated fractions. In effect, in this figure, the Tbs in the dictionaries are grouped into five intervals based on their corresponding inundation fraction (from 0 to 1) in **F**. Then for each interval, the average of Tb values is shown. The plot clearly demonstrates that the daytime Tbs are thermally warmer than their nighttime counterparts and this difference begins to shrink when the inundation fraction increases. It is worth noting that the difference between ascending and descending brightness temperatures is larger over the low-frequency channels (≤ 37 GHz) as they respond more to the land surface structural variability than the higher-frequency channels that capture atmospheric signatures. Figure 4b depicts $|\mathrm{Tb_A} - \mathrm{Tb_D}|$ where $\mathrm{Tb_A}$ and $\mathrm{Tb_D}$ stands for ascending and descending overpasses, respectively. It can be observed that high values of $|\mathrm{Tb_A} - \mathrm{Tb_D}|$ depict the coastlines, i.e., regions with the transient presence and/or absence of water over land.

The probability of detection, $p \in (0-1)$, determines if a pixel is inundated or not if the number of inundated vectors in K-nearest neighbors is $\geq p \times K$. We found that the inundation detection with $K \geq 50$ gives a reasonable rate for the probability of hit and false alarms. In other words, the probability of detection does not change significantly for a larger number of nearest neighbors. In the estimation step, to characterize the weight matrix $\mathbf{W} \in \Re^{n \times n}$, we used the coefficients of variation of each channel in response to changes in the inundation fraction (see Fig. 5). In other words, we assume that those channels that exhibit more variability with respect to changes in inundation fraction contain more information about inundation and shall be given more weight in the estimation process. One might ask why it is important to consider the high-frequency channels (e.g., 91 V, H GHz) despite the fact that they show minimal sensitivity to the inundation fraction (Fig. 5) and land surface emissivity compared

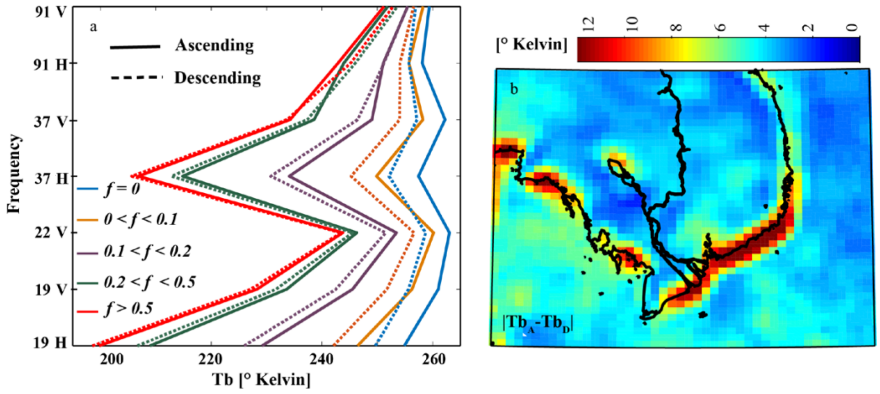

Figure 4. **(a)** The systematic difference between passive microwave observations from the ascending (solid lines) and descending orbits (broken lines) as a function of five different sub-pixel intervals of inundation fractions. **(b)** July to December daily average of absolute differences between the ascending (Tb_A) and descending (Tb_D) brightness temperatures at vertically polarized 19 GHz channel. The values of $|Tb_A - Tb_D|$ mainly capture the coastal regions with significant variability in their surface emissivity values due frequent diurnal tidal effects.

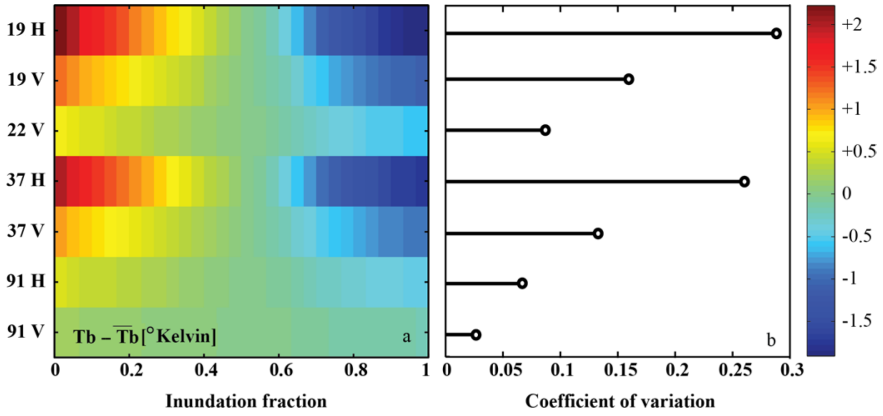

Figure 5. The normalized coefficients of variation (right panel) of the brightness temperatures (Tb) (left panel) averaged over the entire dataset for different intervals of inundation fractions. Here, $\bar{Tb}$ denotes the average of brightness temperatures over the inundation fractions. The coefficients of variation of each channel are used to determine the channel weights for the retrieval algorithm. Channels 19 H GHz and 37 V GHz are the most responsive channels to the variability of inundation fraction and are given higher weights.

to lower-frequency channels. The high-frequency channels mainly capture the information content of the atmospheric profile. Therefore, incorporating them in the proposed retrieval framework allows us to indirectly consider the effect of atmospheric conditions by narrowing down the search for K-nearest neighbors to those Tb candidates that best match both the underlying land surface emissivity and the atmospheric conditions.

For implementation of the algorithm, the regularization parameters are set as $\lambda_1 = \lambda(1-\alpha)$ and $\lambda_2 = \alpha\lambda$, where $\alpha \in (0, 1)$. Here, through cross-validation studies, through cross-validation we empirically found that $\lambda = 0.001$ and $\alpha = 0.1$ provide a reasonable balance between sparsity and stability of the solution in Eq. (4). It should be noted that Eq. (4) is converted to a constrained quadratic programming problem and solved using an iterative Newton's method with MATLAB optimization Toolbox (see Coleman et al., 1999).

4 Results, validation, and discussion

The inundation fractions were estimated during the wet period of calendar year 2015 from July to December when the water levels across the delta begin to rise and eventually recede (see Fig. 6). The wet season of the region is largely characterized by heavy precipitation as a result of the interactions of two monsoons including the Indian monsoon and the East Asia–western North Pacific summer monsoon (Delgado et al., 2012).

To study the performance of the detection step we computed the probability of hit $P(\hat{f} > 0|\mathrm{MWP} > 0)$ and false alarm $P(\hat{f} > 0|\mathrm{MWP} = 0)$ of the algorithm outputs. Our analysis indicates that the probability of hit is around 0.92 for both the dry and wet season, demonstrating the capability of the algorithm in detecting the inundated areas. However, the probability of false alarm is around 0.12 for the

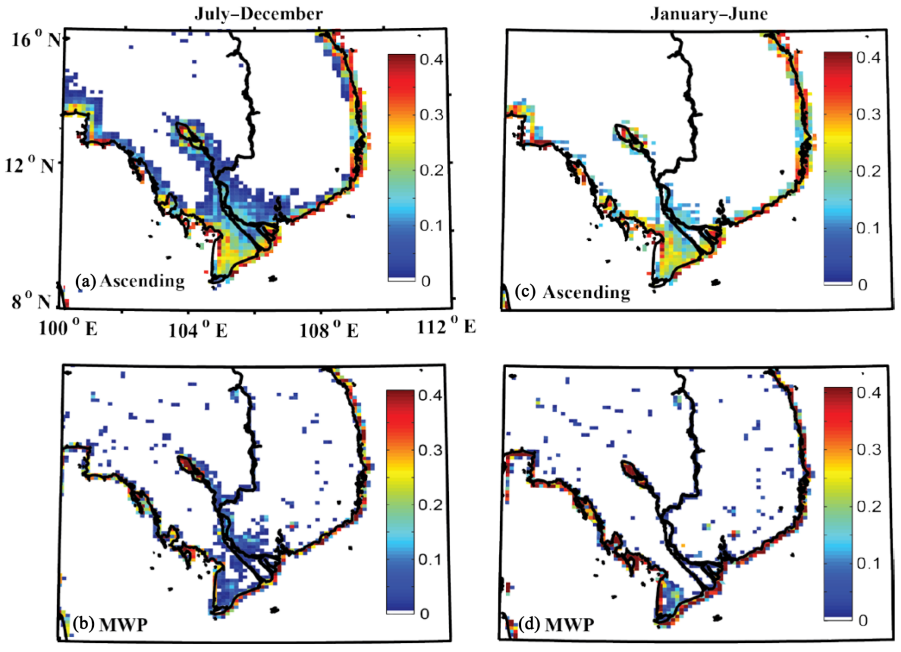

Figure 6. Inundated map of the Mekong Delta in the wet (July–December) and dry (January–June) seasons for the ascending orbits. The results of the proposed retrieval algorithm are presented using the ascending dictionary (top row) against the upscaled MODIS near-real-time (NRT) water product (MWP) data (bottom row). Overall, a good agreement is observed with some overestimation of inundated areas by the proposed algorithm compared to MODIS-MWP data around the river banks.

dry season and reaches the value of 0.34 for the wet season, which might be due to the generalization of the algorithm and MODIS missing data during the wet season. The MODIS daily data, especially in the wet season, contain a large number of missing values due to cloud blockages and frequent heavy rains over the study area. In fact, while we were collecting the overlapping data for constructing the dictionaries, we observed that over 88 % of the MWP products have some missing portion in the 12.5 km resolution. As a result, it is very likely that the MWP data underestimate the actual inundation fraction of regions with prolonged precipitation events.

Figure 6 shows that the algorithm is capable of identifying hotspots of inundation when its outputs are compared with the MODIS-MWP; however, the algorithm slightly overestimates the inundation fractions for some pixels farther from the coastlines, most of which are completely dry in MWP. Here for brevity, we only show the results for ascending overpasses, while similar spatial patterns are observed for descending overpasses. Figure 6 also shows some overestimation of inundation fractions near the riverbanks of major rivers. This might be due to the high soil moisture content (≥ 0.8) during the wet season that increases the dielectric constant of the soil up to 30–50 (Alharthi and Lange, 1987), which is close to the dielectric constant of the water surfaces (75–80). Another reason is the cloud coverage. Since the riverbanks are inundated less frequently than the coastlines, it is possible that these few inundation events were missed by MODIS because of the cloud blockages. There is also some underestimation in the inundation fractions from the proposed algorithm over the hillslopes far away from the riverbanks compared to the MODIS-MWP product. Those sporadically inundated areas, which appear on the MODIS-MWP map (Fig. 6b, c), can be due to the terrain shadows that are misclassified as water. While we cannot directly prove the above assertion within the scope of this paper, the elevation map (Fig. 1) indicates that those hillslopes are very unlikely to get inundated.

Comparison of inundation fractions from MODIS-MWP and the proposed algorithm at daily scale is also challenging. This is because daily MODIS data are often severely corrupted by cloud coverage. On the other hand, under a clear sky, the MODIS-MWP inundation fractions are more precise than the results of the retrieval algorithm. To better illustrate this issue, scatterplots of daily inundation fractions from our retrieval algorithm against those from MODIS-MWP in wet and dry season are displayed in Fig. 7. The scatterplots further demonstrate larger inundation fractions from the retrieval algorithm in July to December (Fig. 7a). In the wet season, there are also a lot of non-zero retrieved inundation values on the y axis that have corresponding zero inundated pixels in MODIS-MWP data. However, in January to June, when there are fewer clouds, the inundation fractions from the proposed algorithm are generally more correlated with the MODIS-MWP data but slightly underestimated (Fig. 7b). This underestimation probably also exists in wet months but

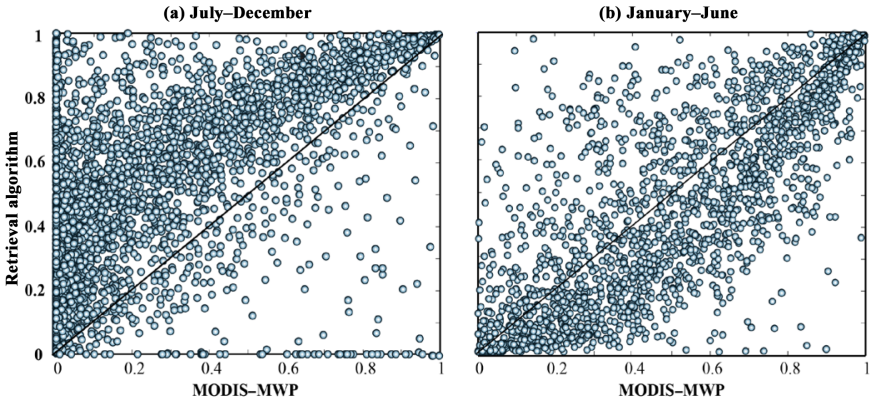

Figure 7. Scatterplots of daily inundation fractions (f) from the retrieval algorithm against those from MODIS-MWP in wet **(a)** and dry seasons **(b)** shown in Fig. 6. The scatterplots demonstrate larger inundation fractions from the retrieval algorithm in July to December **(a)** compared to MODIS-MWP data. However, in January to June, when there are fewer clouds, the inundation fractions from the proposed algorithm are more correlated with the MODIS-MWP data, with only a slight underestimation of their variability.

is masked because of the all-sky retrieval capability of the proposed algorithm in the presence of the clouds and heavy rains. The reason for this underestimation might stem from the choice of 50 % threshold for selecting the clear-sky pixels. In other words, there are a set of brightness temperatures for which the corresponding MODIS data are partially cloudy and potentially underestimate the actual inundation fraction. As a result, it is likely that those pairs will be isolated, used in the retrievals, and eventually lead to some underestimation in passive microwave retrievals.

As previously mentioned, the interannual climatology of the Mekong Delta is highly affected by two tropical monsoons that characterize the seasonal patterns of precipitation, river stages, and water levels (Delgado et al., 2012). To better understand whether the results of our retrieval algorithm follow the regional climatology, the monthly percentage of inundated area over the Mekong Delta is calculated and shown against the monthly water level data in Fig. 8a. The monthly water level data are obtained by averaging over all 11 stations shown in Fig. 1. The specific goal is to compare the monthly variability of the algorithm outputs with the MWP products and investigate whether they are consistent with the regional variations of the surface water level (river stage), which is considered a surrogate for the extent of inundation. It should be acknowledged that this approach is not a direct validation; however, it can provide insight into the performance and climatological consistency of the proposed model as the surface water level data are positively correlated with the extent of the inundated surfaces.

The seasonal variations in the monthly percentage of the total inundated surfaces from the proposed model follow the trend of monthly water level data better than the standard MWP products (Fig. 8a). We can see that during the wet months of June to November, the MWP data report much less inundated area than the outputs of the proposed algorithm, whereas this pattern is reversed during the dry months of January to March. As previously noted, we suspect that the differences in the wet season are due to the large portion of missing data in the MWP products because of the high cloud coverage in the rainy season (Fig. 8b). For quantitative comparison of the outputs of the algorithm with MWP, a Euclidean distance between normalized version of the algorithm outputs and water level data is calculated and compared with its MWP counterpart. The Euclidean distance between water level and the retrieved inundation from ascending and descending orbits is 3.46 and 3.56, respectively, while this distance for MWP and water level data is about 7.89, which is more than twice the distances calculated from the microwave retrieval results. This indicates the superior performance of the proposed inundation fraction retrievals as compared to the MWP products, chiefly because of its all-sky skills during the rain dominant seasons.

When is compared to MODIS-MWP, the inundated area obtained by the retrieval algorithm in the dry months (Fig. 8a) shows some underestimation. One reason for this underestimation is the general limitation of the empirical Bayesian estimation method regarding the extreme events (see Coles and Powell, 1996, and the references therein) and we suspect that it is not just limited to the months of January to March but it affects the retrievals at other months to a lesser extent, as well. This limitation arises by the sample scarcity of large flooding scenarios during the warm months of the year, which probably lead to the underestimation of inundation fractions related to those events by our retrieval algorithm. We expect that by improving the representativeness of the dataset – especially for extreme events in the summer months – this shortcoming can be significantly improved.

A closer look at Fig. 8a also reveals slightly larger inundated surfaces in each month for the ascending (evening overpasses) compared to the descending (morning overpasses) tracks. This small difference between the ascending and descending retrievals can be attributed to the expected

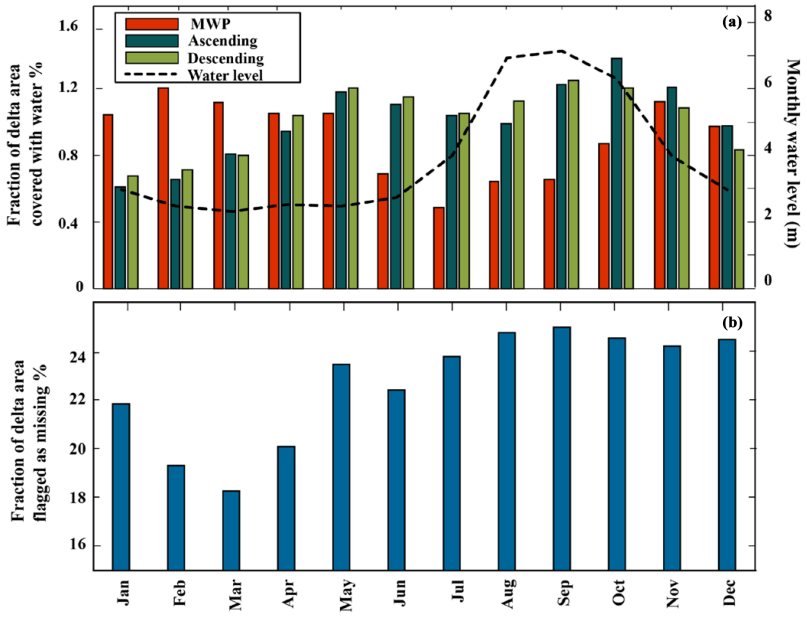

Figure 8. The monthly inundated areas of the Mekong Delta calculated from the proposed retrieval algorithm and MODIS near-real-time (NRT) water product (MWP) data in comparison with ground-based monthly water level data. **(a)** Comparison of the total inundated surface of the Mekong Delta from MWP products and the retrieval algorithm from ascending and descending dictionaries. From visual inspection, it is obvious that the retrieval algorithm can better follow variations of the water levels compared to MWP. More inundation over the dry season is reported by MWP products than the wet season, which contradicts the causality between rivers' stages and the extent of inundated areas. **(b)** The total fraction of land surface areas that are labeled as missing in MWP product because of atmospheric contaminations. The larger deviations of the MWP products from water level data during the wet months might be attributed to the larger percentage of missing values.

diurnal patterns of the precipitation over the Mekong Delta. Indeed, it is well documented (Gupta 2005) that localized convective precipitation events are more likely during the evening, which can increase the extent of the inundated areas. To further assess the proposed algorithm performance at a daily scale, we compare the dependence of the total area of ascending daily inundation fractions of the algorithmic outputs with the average daily water level data, using Spearman's rank-correlation coefficient. The rationale is that a stronger rank correlation of an inundation product with the water level data implies an improved retrieval. The correlation coefficient between the daily water level of the rivers and the total inundated surfaces of the Mekong Delta is equal to 0.22, which drops to -0.38 for the MWP products. To go beyond a rank correlation, we also examined the dependence structures across different ranges of inundation and water level quantiles using an empirical copula (see Appendix A1).

Copulas provide an effective non-parametric way for simple representation of multivariate joint distributions of high-dimensional random variables to describe their dependence structure. When dependence of two random variables increases, their bivariate "L-shaped" cumulative copulas tend more to the origin. In Fig. 9, the axes show the marginal quantiles of daily inundation fractions versus those of water level elevations and the contours trace the cumulative copulas. To characterize the dependence of water level and inundation as a function of topography, we divided the study area into two sub-regions covering the steeper upper parts (above the Phnom Penh gauge in Fig. 1) and the flatter downstream region. The copula analysis for each region was presented separately in Fig. 9. As is evident, the empirical copula of the total daily inundation fraction from the proposed algorithm shows higher degree of dependence to the water level, as compared to MWP, especially for the quantiles with less than 0.8 cumulative probability for both upstream and the downstream regions. However, comparing the downstream (Fig. 9a) and the upstream (Fig. 9b) regions, we see an increased dependency of the retrievals with the water levels in the upstream region. This observation seems to be consistent from a geomorphological point of view, because over a steeper region of the basin the hill slopes are naturally steeper and any small water variability can give rise to significant water extension of inundated areas. However, over fat floodplains water levels and extent of inundations may not be strongly correlated as small changes of water levels may give rise a large extension of flooded surfaces.

5 Conclusions and future directions

In this paper, we introduced a methodology to retrieve inundation from space for almost all-sky conditions to reduce the gaps that exist in using satellite data in visible to mi-

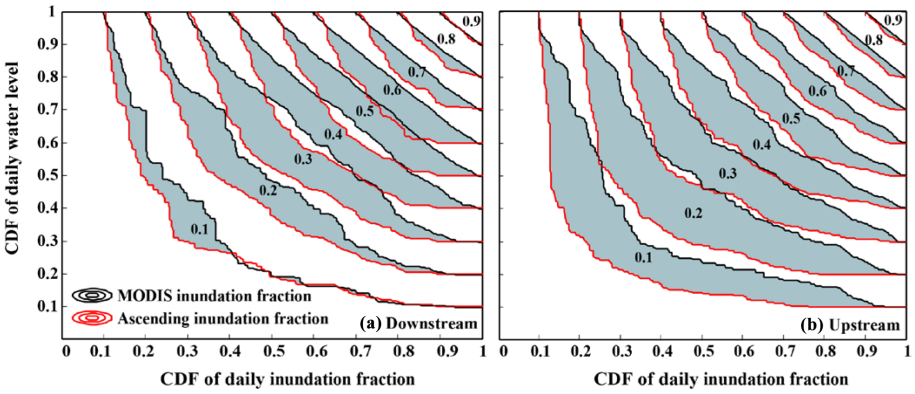

Figure 9. The empirical copula (joint probability distribution of quantiles) of the average daily water level and total daily inundated areas from the proposed retrieval algorithm (red curves) and MODIS-MWP data (black curves) for 2015. These plots indicate that our products have stronger dependence to water levels than the MWP products (more L-shaped curves) for both the downstream **(a)** and upstream **(b)** regions of the Mekong Delta. The shaded areas (which quantify the difference between the degree of dependence of our products and the MWP products to the daily water levels) are larger in the upstream region, indicating an enhanced performance of the proposed algorithm to retrieve inundation fraction where potential inundation areas are better defined due to topography, e.g., around major riverbanks.

crowave bands. The key idea of the proposed method was to explore the links between overlapping daily high-resolution observations in the visible and near-infrared bands from the MODIS and the lower-resolution passive microwave observations from the Special Sensor Microwave Imager and Sounder (SSMIS) sensor. The developed multi-frequency inundation retrieval algorithm uses the K-nearest matching method in conjunction with a sparsity-promoting regularization technique. The proposed method demonstrated promising results in resolving the spatial patterns of inundation, compared with the MODIS-MWP data. Over the months with high cloud coverage, the monthly results are consistent with the seasonal dynamics of water level variation, which is controlled by tropical monsoons in the Mekong Delta. Analysis also showed that, at a daily timescale, the outputs of the algorithm exhibit stronger dependency with the water level data than the MWP data.

There were three major sources of uncertainty in the proposed retrieval model in this paper. The first one related to the use of the 3-day composite MODIS-MWP data (daily products of MODIS-MWP were avoided due to missing values and cloud blockages), which might have introduced some bias in the daily retrievals due to mismatch of timescales. This source of error can be significantly reduced if the MODIS dictionary is populated with more accurate daily products. The second source of error related to the lack of adequate fully clear-sky samples in our dictionary and therefore the need to define a cloud coverage threshold in order to increase the sample size. Using partially cloudy MODIS data was the main reason for some observed underestimation of inundation fractions, especially in the dry months (Figs. 7 and 8), which can be mitigated by increasing the sample size. The last source of error was more related to the general limitation of the Bayesian estimation method regarding the retrieval of extreme events (see Coles and Powell, 1996, and references therein). This limitation is due to scarcity of large floods in the dictionary, which can be treated by adding more scenarios of extreme events to the dataset from different geographic locations.

One of the limitations of the proposed algorithm (because of the spatial resolution of microwave data used in this paper) was its lack of information about the spatial patterns of inundation within the 12.5 km pixels. The spatial pattern of the estimated inundation fractions can be further enhanced by using the guidance of a high-resolution topographic data (see Galantowicz, 2002). The database can also expand to include some high-resolution cloud-free imageries from newly launched satellites, such as Sentinel-2, which can aid in capturing the high-resolution inundation areas. Finally, expanding the dictionary to include data from the passive microwave channels of the new satellites, such as Global Precipitation Mission (GPM) Microwave Imager (GMI), will increase the spatial resolution of the retrievals to approximately 5 km. In this paper, the seasonality and also different land surface classes have not been directly taken into account in the retrieval algorithm. Future research should include the stratification of the dictionary based on different land surface types and time periods (e.g., seasons).

Code availability. MATLAB code available at ftp://ebtehaj.safl.umn.edu/Codes/ShARP_Demo/. The dictionary of overlapped MODIS-MWP and SSMIS and also the resulting database are publicly available upon request (email to takbi001@umn.edu). The data that have been used to create the dictionary are available at https://floodmap.modaps.eosdis.nasa.gov/ (NASA Goddard's Hydrology Laboratory, 2016) and ftp://sidads.colorado.edu/pub/DATASETS/nsidc0032_ease_grid_tbs/global/ for or MODIS-MWP and SSMIS data, respectively.

Appendix A: Copula

Let X_1 and X_2 denote two random variables with marginal cumulative distributions $F_1(x_1) \equiv P[X_1 \leq x_1]$ and $F_2(x_2) \equiv P[X_2 \leq x_2]$ with the cumulative joint distribution function $F(x_1, x_2) \equiv P[X_1 \leq x_1, X_2 \leq x_2]$. According to the Sklar's theorem (Nelsen, 1999), the cumulative joint distribution $F(x_1, x_2)$ of X_1 and X_2 is equal to the cumulative joint distribution function $C(u_1, u_2)$ of the quantiles $u_1 = F_1(x_1)$ and $u_2 = F_2(x_2)$ by

$$\begin{aligned} F(x_1, x_2) &= P[X_1 \leq x_1, X_2 \leq x_2] \\ &= P\left[X_1 \leq F_1^{-1}(u_1), X_2 \leq F_2^{-1}(u_2)\right] \\ &\equiv C[U_1 \leq u_1, U_2 \leq u_2] \\ &= C(u_1, u_2,) \end{aligned} \tag{A1}$$

where $C(u_1, u_2)$, is the cumulative copula with uniform marginal random variables $F_1(x_1)$ and $F_2(x_2)$ on the interval $[0, 1]$. The multivariate density function $f(x_1, x_2)$, if it exists, can be calculated by taking the derivative of C and F, which results in the following:

$$\begin{aligned} f(x_1, x_2) &= c(u_1, u_2) \cdot f(x_1) \cdot f(x_2) \\ &= c(F(X_1), F(X_2)) \cdot f(x_1) \cdot f(x_2). \end{aligned} \tag{A2}$$

It shows the copula density function $c(u_1, u_2)$ separates the joint distribution function $f(x_1, x_2)$ from its marginal probability distribution functions $f(x_1)$ and $f(x_2)$; therefore, it can capture the probabilistic dependence between the two random variables x_1 and x_2 by quantifying the strength of the relationship between their corresponding quantiles.

Appendix B

Table B1. Acronyms and abbreviations.

SSMIS	Special Sensor Microwave Imager and Sounder
SSM/I	Special Sensor Microwave Imager
DMSP	Defense Meteorological Satellite Program
MSS	Multispectral scanner system
VNIR	Visible to near infrared
MODIS	Moderate Resolution Imaging Spectroradiometer
NIR	Near infrared
MIR	Mid-infrared
PMW	Passive microwaves
ESMR	Electrically scanning microwave radiometer
SMMR	Multi-frequency Microwave Radiometer
BWI	Basin wetness index
WSF	Water surface fraction
AMSR-E	Advanced Microwave Scanning Radiometer - Earth Observing System
NRT	Near-real time
NSIDC	National Snow and Ice Data Center
DFO	Dartmouth Flood Observatory
MODIS-MWP	MODIS near-real-time (NRT) water product
CDF	Cumulative probability function
M	Number of vectors of microwave brightness temperatures $\mathbf{B}$
k	Number of nearest neighbors
$\mathbf{B}$	Brightness temperature dictionary
f	Inundation fraction
$\mathbf{F}$	Inundation dictionary
$\boldsymbol{b}_{\text{obs}}$	Observed vector of brightness temperature
K	Number of nearest neighbors
$\mathbf{B}_{\text{s}}$	Sub-dictionary of $\mathbf{B}$
$\mathbf{F}_{\text{s}}$	Sub-dictionary of $\mathbf{F}$
$\boldsymbol{c}$	Vector of representation coefficients
$\hat{f}$	Estimated inundation fraction
$\mathbf{W}$	Weight matrix
n	Number of frequency channels
p	Detection probability $\in (0, 1)$
ℓ_1& ℓ_2	Regularizations norms
λ_1 & λ_2	Regularization parameters

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This work was supported by the NASA Global Precipitation Measurement Program under grants NNX13AG33G and NNX16AO56G. It was also partially supported by NSF under the Belmont Forum DELTAS project (EAR-1342944) and the LIFE project (EAR-1242458). The MODIS-MWP data over the Mekong Delta were kindly provided by Dan Slayback from the NASA Goddard Space Flight Center. The first author would like to thank Professor Robert Brakenridge for his advice on this research during the AGU Fall Meeting 2015.

Edited by: M. McCabe
Reviewed by: two anonymous referees

References

Alharthi, A. and Lange, J.: Soil water saturation: Dielectric determination, Water Resour. Res., 23, 591–595, https://doi.org/10.1029/WR023i004p00591, 1987.

Allison, L. J., Schmugge, T. J., and Byrne, G.: A hydrological analysis of East Australian floods using Nimbus-5 electrically scanning radiometer data, B. Am. Meteorol. Soc., 60, 1414–1427, https://doi.org/10.1175/1520-0477(1979)060<1414:AHAOEA>2.0.CO;2, 1979.

Armstrong, R. L. and Brodzik, M. J.: An earth-gridded SSM/I data set for cryospheric studies and global change monitoring, Adv. Space Res., 16, 155–163, https://doi.org/10.1016/0273-1177(95)00397-W, 1995.

Basist, A., Grody, N. C., Peterson, T. C., and Williams, C. N.: Using the Special Sensor Microwave/Imager to Monitor Land Surface Temperatures, Wetness, and Snow Cover, J. Appl. Meteorol., 37, 888–911, https://doi.org/10.1175/1520-0450(1998)037<0888:UTSSMI>2.0.CO;2, 1998.

Brakenridge, G. R. and Anderson, E.: MODIS-based flood detection mapping and measurement: The potential for operational hydrological applications, in: Transboundary Floods: Reducing Risks Through Flood Management, edited by: Marsalek, J., Stancalie, G., and Balint, G., Nato Science Series: IV: Earth and Environmental Sciences, 72, Springer, Dordrecht, 2006.

Brakenridge, G. R., Nghiem, S. V., Anderson, E., and Chien, S.: Space-based measurement of river runoff, Eos T. Am. Geophys. Un., 86, 185–188, 2005.

Brakenridge, G. R., Nghiem, S. V., Anderson, E., and Mic, R.: Orbital microwave measurement of river discharge and ice status, Water Resour. Res., 43, 1–16, https://doi.org/10.1029/2006WR005238, 2007.

Choudhury, B. J.: Passive microwave remote sensing contribution to hydrological variables, Land Surface–Atmosphere Interactions for Climate Modeling, Springer Netherlands, 63–84, 1991.

Coleman, T. F., Branch, M. A., and Grace, A.: Optimization Toolbox: For Use with MATLAB: User's Guide, Version 2. MathWorks, available at: https://pdfs.semanticscholar.org/35e3/d60a0bdbd53c008fad425e9f8668e1528542.pdf (last access: 5 June 2017), 1999.

Coles, S. G. and Powell, E. A.: Bayesian methods in extreme value modelling: a review and new developments, Int. Stat. Rev., 119–136, https://doi.org/10.2307/1403426, 1996.

Crétaux, J.-F., Bergé-Nguyen, M., Leblanc, M., Abarca Del Rio, R., Delclaux, F., Mognard, N., Lion, C., Pandey, R. K., Tweed, S., Calmant, S., and Maisongrande, P.: Flood mapping inferred from remote sensing data, Int. Water Technol. J., 1, 48–62, 2011.

De Groeve, T.: Flood monitoring and mapping using passive microwave remote sensing in Namibia, Geomatics Nat. Hazards Risk, 1, 19–35, https://doi.org/10.1080/19475701003648085, 2010.

Delgado, J. M., Merz, B., and Apel, H.: A climate-flood link for the lower Mekong River, Hydrol. Earth Syst. Sci., 16, 1533–1541, https://doi.org/10.5194/hess-16-1533-2012, 2012.

Ebtehaj, A. M., Bras, R. L., and Foufoula-Georgiou, E.: On evaluation of ShARP passive rainfall retrievals over snow-covered land surfaces and coastal zones, arXiv Prepr., 17, 1013–1029, https://doi.org/10.1175/JHM-D-15-0164.1, 2015a.

Ebtehaj, A. M., Bras, R. L., and Foufoula-Georgiou, E.: Shrunken Locally Linear Embedding for Passive Microwave Retrieval of Precipitation, IEEE T. Geosci. Remote Sens., 53, 3720–3736, https://doi.org/10.1109/TGRS.2014.2382436, 2015b.

Ericson, J. P., Vörösmarty, C. J., Dingman, S. L., Ward, L. G., and Meybeck, M.: Effective sea-level rise and deltas: Causes of change and human dimension implications, Global Planet. Change, 50, 63–82, https://doi.org/10.1016/j.gloplacha.2005.07.004, 2006.

Ferraro, R., Grody, N., and Kogut, J.: Classification of Geophysical Parameters Using Passive Microwave Satellite Measurements, IEEE T. Geosci. Remote Sens., 24, 1008–1013, https://doi.org/10.1109/TGRS.1986.289564, 1986.

Frazier, P. S. and Page, K. J.: Water Body Detection and Delineation with Landsat TM Data, Photogramm. Eng. Rem. S., 66, 1461–1467, 2000.

Galantowicz, J. F.: High-resolution flood mapping from low-resolution passive microwave data, IEEE Int. Geosci. Remote Sens. Symp., 3, 1499–1502, https://doi.org/10.1109/IGARSS.2002.1026161, 2002.

Gao, B. C.: NDWI – A normalized difference water index for remote sensing of vegetation liquid water from space, Remote Sens. Environ., 58, 257–266, https://doi.org/10.1016/S0034-4257(96)00067-3, 1996.

Giddings, L. and Choudhury, B. J.: Observation of hydrological features with Nimbus-7 37 GHz data, applied to South America, Int. J. Remote Sens., 10, 1673–1686, https://doi.org/10.1080/01431168908903998, 1989.

Grody, N. C.: Classification of snow cover and precipitation using the special sensor microwave imager, J. Geophys. Res., 96, 7423, https://doi.org/10.1029/91JD00045, 1991.

Guerschman, J. P., Warren, G., Byrne, G., Lymburner, L., Mueller, N., and Van Dijk, A.: MODIS-based standing water detection for flood and large reservoir mapping: algorithm development and applications for the Australian continent, CSIRO: Water for a Healthy Country National Research Flagship Report, Canberra, 2011.

Gupta, A.: The physical geography of Southeast Asia, Vol. 4, Oxford University Press on Demand, 2005.

Jain, S. K., Singh, R. D., Jain, M. K., and Lohani, A. K.: Delineation of flood-prone areas using remote sensing techniques, Water

Resour. Manage., 19, 333–347, https://doi.org/10.1007/s11269-005-3281-5, 2005.

Kuenzer, C., Klein, I., Ullmann, T., Georgiou, E., Baumhauer, R., and Dech, S.: Remote Sensing of River Delta Inundation: Exploiting the Potential of Coarse Spatial Resolution, Temporally-Dense MODIS Time Series, Remote Sens., 7, 8516–8542, https://doi.org/10.3390/rs70708516, 2015.

Leinenkugel, P., Kuenzer, C., and Dech, S.: Comparison and enhancement of MODIS cloud mask products for Southeast Asia, Int. J. Remote Sens., 34, 2730–2748, https://doi.org/10.1080/01431161.2012.750037, 2013.

Mears, C. A., Schabel, M. C., Wentz, F. J., Santer, B. D., Govindasamy, B., and Ieee: Correcting the MSU middle tropospheric temperature for diurnal drifts, Igarss 2002 Ieee Int. Geosci. Remote Sens. Symp. 24th Can. Symp. Remote Sensing, Vols. I–Vi, Proc. Remote Sens. Integr. Our View Planet, 1839–1841, 2002.

Melack, J. M., Hess, L. L., and Sippel, S.: Remote sensing of lakes and floodplains in the Amazon Basin, Remote Sens. Rev., 10, 127–142, https://doi.org/10.1080/02757259409532240, 1994.

McGinnis, D. F. and Rango, A.: Earth resources satellite systems for flood monitoring, Geophys. Res. Lett., 2, 132–135, https://doi.org/10.1029/GL002i004p00132, 1975.

NASA Goddard's Hydrology Laboratory: NRT global flood mapping, available at: https://floodmap.modaps.eosdis.nasa.gov/, last access: 5 June 2016.

Nelsen, R. B.: An Introduction to Copulas, Lecture notes in statistics, 39, Springer-Verlag, New York, New York, 1999.

Nicholls, R. J. and Cazenave, A.: Sea-level rise and its impact on coastal zones, Science, 328, 1517–1520, https://doi.org/10.1126/science.1185782, 2010.

Nigro, J., Slayback, D., Policelli, F., and Brakenridge, G. R.: NASA/DFO MODIS Near Real-Time (NRT) Global Flood Mapping Product Evaluation of Flood and Permanent Water Detection, available at: http://oas.gsfc.nasa.gov/floodmap/documents/NASAGlobalNRTEvaluationSummary_v4.pdf (last access: 5 June 2017), 2014.

Norouzi, H., Rossow, W., Temimi, M., Prigent, C., Azarderakhsh, M., Boukabara, S., and Khanbilvardi, R.: Using microwave brightness temperature diurnal cycle to improve emissivity retrievals over land, Remote Sens. Environ., 123, 470–482, https://doi.org/10.1016/j.rse.2012.04.015, 2012.

Ordoyne, C. and Friedl, M. A.: Using MODIS data to characterize seasonal inundation patterns in the Florida Everglades, Remote Sens. Environ., 112, 4107–4119, https://doi.org/10.1016/j.rse.2007.08.027, 2008.

Papa, F., Prigent, C., Rossow, W. B., Legresy, B., and Remy, F.: Inundated wetland dynamics over boreal regions from remote sensing: the use of Topex-Poseidon dual-frequency radar altimeter observations, Int. J. Remote Sens., 27, 4847–4866, https://doi.org/10.1080/01431160600675887, 2006.

Prigent, C., Matthews, E., Aires, F., and Rossow, W. B.: Remote sensing of global wetland dynamics with multiple satellite data sets, Geophys. Res. Lett., 28, 4631–4634, https://doi.org/10.1029/2001GL013263, 2001.

Prigent, C., Papa, F., Aires, F., Rossow, W. B., and Matthews, E.: Global inundation dynamics inferred from multiple satellite observations, 1993–2000, J. Geophys. Res.-Atmos., 112, 1993–2000, https://doi.org/10.1029/2006JD007847, 2007.

Ramage, J. M. and Isacks, B. L.: Interannual variations of snowmelt and refreeze timing on southeast-Alaskan icefields, U.S.A, J. Glaciol., 49, 102–116, doi:10.3189-172756503781830908, 2003.

Rango, A. and Anderson, A. T.: flood hazard studies in the mississippi river basin using remote sensing, J. Am. Water Resour. As., 10, 1060–1081, https://doi.org/10.1111/j.1752-1688.1974.tb00625.x, 1974.

Rango, A. and Salmonson, V.: Regional flood mapping from space, Water Resour. Res., 10, 1944–7973, https://doi.org/10.1029/WR010i003p00473., 1974.

Rouse, J. W., Haas, R. H., Schell, J. A., Deering, D. W., and Harlan, J. C.: Monitoring the vernal advancement of retrogradation of natural vegetation, NASA/GSFC, Type III, Final Report, 371 pp., Greenbelt, MD, 1974.

Schroeder, R., Rawlins, M. A., McDonald, K. C., Podest, E., Zimmermann, R., and Kueppers, M.: Satellite microwave remote sensing of North Eurasian inundation dynamics: development of coarse-resolution products and comparison with high-resolution synthetic aperture radar data, Environ. Res. Lett., 5, 15003, https://doi.org/10.1088/1748-9326/5/1/015003, 2010.

Sippel, S. J., Hamilton, S. K., Melack, J. M., and Choudhury, B. J.: Determination of inundation area in the Amazon River floodplain using the SMMR 37 GHz polarization difference, Remote Sens. Environ., 48, 70–76, https://doi.org/10.1016/0034-4257(94)90115-5, 1994.

Sippel, S. J., Hamilton, S. K., Melack, J. M., and Novo, E. M. M.: Passive microwave observations of inundation area and the area/stage relation in the Amazon River floodplain, Int. J. Remote Sens., 19, 3055–3074, https://doi.org/10.1080/014311698214181, 1998.

Smith, L. C.: Satellite remote sensing of river inundation area, stage, and discharge: A review, Hydrol. Process., 11, 1427–1439, https://doi.org/10.1002/(sici)1099-1085(199708)11:10<1427::aid-hyp473>3.0.co;2-s, 1997.

Smith, R. B.: Introduction to remote sensing of the environment, available at: www.microimages.com (last access: 5 June 2017), 2001.

Syvitski, J. P., Vörösmarty, C. J., Kettner, A. J., and Green, P.: Impact of humans on the flux of terrestrial sediment to the global coastal ocean, Science, 308, 376–380, https://doi.org/10.1126/science.1109454, 2005.

Temimi, M., Leconte, R., Brissette, F., and Chaouch, N.: Flood monitoring over the Mackenzie River Basin using passive microwave data, Remote Sens. Environ., 98, 344–355, https://doi.org/10.1016/j.rse.2005.06.010, 2005.

Temimi, M., Lacava, T., Lakhankar, T., Tramutoli, V., Ghedira, H., Ata, R., and Khanbilvardi, R.: A multi-temporal analysis of AMSR-E data for flood and discharge monitoring during the 2008 flood in Iowa, Hydrol. Process., 25, 2623–2634, https://doi.org/10.1002/hyp.8020, 2011.

Tessler, Z. D., Vörösmarty, C. J., Grossberg, M., Gladkova, I., Aizenman, H., Syvitski, J. P. M., and Foufoula-Georgiou, E.: Profiling risk and sustainability in coastal deltas of the world, Science, 349, 638–643, https://doi.org/10.1126/science.aab3574, 2015.

Ticehurst, C. J., Chen, Y., Karim, F., Dutta, D., and Gouweleeuw, B.: Using MODIS for mapping flood events for use in hydrological and hydrodynamic models: Experiences so far, 20th

Int. Congr. Model. Simul., 1–9, available at: http://www.microimages.com/documentation/Tutorials/introrse.pdf (last access: 5 June 2017), 2013.

Ulaby, F. T., Moore, R. K., and Fung, A. K.: Microwave Remote Sensing: Radar remote sensing and surface scattering and emission theory, Vol. 2, Addison-Wesley Publishing Company, Advanced Book Program/World Science Division, 1982.

Wilheit, T., Adler, R., Avery, S., Barrett, E., Bauer, P., Berg, W., Chang, A., Ferriday, J., Grody, N., Goodman, S., and Kidd, C.: Algorithms for the retrieval of rainfall from passive microwave measurements, Remote Sens. Rev., 11, 163–194, https://doi.org/10.1080/02757259409532264, 1994.

Xu, H.: Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery, Int. J. Remote Sens., 27, 3025–3033, https://doi.org/10.1080/01431160600589179, 2006.

Zou, H. and Hastie, T.: Regularization and variable selection via the elastic net, J. R. Stat. Soc. B, 67, 301–320, https://doi.org/10.1111/j.1467-9868.2005.00503.x, 2005.