Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-560, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 1 November 2016

Hydrology and Earth System Sciences
Discussions

A Multi-sensor Data-driven methodology for all-sky Passive Microwave Inundation Retrieval

Zeinab Takbiri[1,2], Ardeshir M Ebtehaj[1], Efi Foufoula-Georgiou[1,3]

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

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

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 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* database 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 diurnal variability (i.e., morning and evening) of inundation due to localized convective precipitation. At longer time-scales, 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 time scales.

Keywords:
Passive Microwave Inundation Retrievals, Bayesian Inversion, *k*-nearest Neighbors, Deltaic regions, Inverse Problems, Sparse Regularization.

Key points:
Multi-sensor observations improve satellite inundation mapping under cloudy sky.
Retrievals using passive microwave observations capture the diurnal variability of inundation.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-560, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 1 November 2016

Hydrology and Earth System Sciences
Discussions

1 Introduction

Capturing the diurnal spatio-temporal 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 (~0.4–0.8 µm), natural water reflects a small fraction of incident light depending on the water depth and concentration of the suspended and dissolved particulate matter. However, water reflectivity sharply declines and approaches zero in the near infrared bands (~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 and references therein; Frazier and Page, 2000; Smith, 2001; Jain et al., 2005). In the microwave region of the spectrum, the dielectric constant of water (~80) is much higher than the dry soil (~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 been also used through empirical thresholding approaches to distinguish water surfaces from other land 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 (0.5–1.1µm) 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 (0.62–0.67µm) and near infrared (NIR) band 2 (0.84–0.87µm) 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, $NDVI = (NIR - red)/(NIR + red)$ 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, $NDWI = (red - MIR)/(red + MIR)$ introduced by Gao (1996) and later modified to $MNDWI = (green - MIR)/(green + MIR)$ by Xu (2006)—that exploits the mid-infrared (MIR: 1.23–1.25µm) part of the spectrum to improve the mapping. In all thresholding-based 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 inter-annual dynamics of areas covered with water over five deltaic regions.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-560, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 1 November 2016

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 37GHz vertical and horizontal polarization differences (i.e., $Tb_{37v} - Tb_{37h}$), 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 an empirical parameter, called 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 et al. (2010) also used the same method and instrument to map floods for several hundreds of location for the Global Disaster Alert and Coordination System GDACS.

While visible and short-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 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 Special Sensor Microwave Imager/Sounder), only large water bodies can be detected and sub-pixel inundated areas cannot be directly identified (Smith, 1997). Nowadays, there exist several sensors on board different satellites with overlapping in 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 only from passive microwave observations based on a set of paired VNIR and passive microwave training samples. We need to note that the term “inundation” here means regions where water covers the land surface, excluding however permanent water bodies. In particular, as training examples, we use the daily global observations of VNIR data from the Moderate Resolution Imaging Spectroradiometer (MODIS) on

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-560, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 1 November 2016

board Terra (launched in 2000) and Aqua satellites (launched in 2002) and passive microwave data from the Special Sensor Microwave Imager/Sounder (SSMIS) on board 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 database. 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, 1994) that hydrometeors and the atmospheric profile do not significantly affect the low-frequency <60 GHz brightness temperatures. Therefore, the information content of the database 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 sky when only the information in the clear sky database is used, but we expect that this uncertainty will be small since the information of the underlying surfaces in low-frequency channels of the collected database remains almost the same over different atmospheric conditions.

The collected database has a large number of linked pairs of inundation fractions form MODIS data and SSMIS multi-frequency brightness temperature data. For algorithmic development, the database is organized into two fat matrices: the so-called 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 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. Note that, at resolution 12.5 km, we label a pixel as a clear sky condition when less than 50% of the VNIR data at resolution 250m is flagged as missing. We also evaluated the results against the daily and monthly water level data obtained from eleven gages 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* database and the formation of the dictionaries and Section 3 provides detailed information about the retrieval algorithm. Implementation of the method and validation are explained in Section 4. Section 5 presents concluding remarks and directions for future research.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-560, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 1 November 2016

2 Study Area and database

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 the Southeast Asia. This critical region is home to nearly 20 million people, which contains 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 ecologic and ecosystem modeling. The proposed retrieval algorithm is applied to estimate daily inundation fraction at resolution of 12.5 km over some of the lower regions of the Mekong delta in the calendar year 2015 (Fig. 1).

Two sources of information are used to build a database 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, 3 Observations, and no shadow masking) with approximately 250 m spatial resolution (Nigro et al., 2014). 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). 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 AM and Aqua has a descending orbit with the local equatorial crossing time of 1:30 PM.

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 in general based on observations by the SSM/I and SSMIS radiometer on board the DMSP F8, 11, 13 or 17. Since December 2006, F17 satellite has been the only operational satellite form the DMSP series, which carries on board the SSMIS instrument with equatorial crossing times of 05:30–06:30 AM and 17:30–18:30 PM 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 database. 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.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-560, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 1 November 2016

Hydrology and Earth System Sciences Discussions

The first step for building the *a priori* database 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 resolution ~12.5 km, using a nearest neighbor interpolation. The 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 time scales of Tb and MWP values, the Tb values are averaged over a three-day time window to minimize the possible effects of cloud contamination in the VNIR data. Fig. 2 demonstrates schematically the process of producing the explained database.

3 The Retrieval Algorithm

To organize the database in an algebraically tractable manner, let us collect *M* vectors of microwave brightness temperatures $\mathbf{b}_i = \left(Tb_{1i}, Tb_{2i}, ..., Tb_{ni}\right)^T \in \Re^n$ at *n* frequency channels in the column space of an *n*-by-*M* matrix $\mathbf{B} = \left[\mathbf{b}_1 \,|\, \mathbf{b}_2 \,|\, ... \,|\, \mathbf{b}_M\right] \in \Re^{n \times M}$, called brightness temperature dictionary, where $M \gg N$. Analogously, the corresponding inundation fraction values $\left\{f_i\right\}_{i=1}^{M}$ can be collected in the column space of the inundation dictionary $\mathbf{F} = \left[f_1 \,|\, f_2 \,|\, ... \,|\, f_M\right] \in \Re^{1 \times M}$. The algorithm follows two sequential steps: a detection and an estimation step. In the detection step, for each observed vector of brightness temperature $\mathbf{b}_{\text{obs}}$, the algorithm first finds its *K*- neighboring brightness temperatures in $\mathbf{B}$ in the Euclidean sense and stores them in the column space of $\mathbf{B}_s \in \Re^{n \times K}$. Then, knowing the column indices of the neighboring brightness temperatures, it isolates their corresponding inundation fraction values in $\mathbf{F}_s \in \Re^{1 \times K}$. In this step, if more than half of the nearby inundation fraction values in $\mathbf{F}_s$ were non-zero, the algorithm assumes that $\mathbf{b}_{\text{obs}}$ is over an inundated area and attempts to estimate the fraction of inundation in the estimation step.

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

$$\mathbf{b}_{\text{obs}} = \mathbf{B}_s\mathbf{c} + \mathbf{e} \tag{1}$$

where the vector $\mathbf{c} \in \Re^k$ contains a set of representation coefficients to be estimated and $\mathbf{e} \in \Re^n$ is the error vector. Clearly, for an observed vector of brightness temperatures $\mathbf{b}_{\text{obs}}$, the goal is to estimate its unknown inundation fraction value $\hat{f}$. We assume that the two paired dictionaries $\mathbf{B}_s$ and $\mathbf{F}_s$ represent similar manifolds in a geometric sense that their local structures

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-560, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 1 November 2016

Hydrology and Earth System Sciences
Discussions

can be approximated well with the same linear model. This allows us to assume that the representation coefficients in vector $\mathbf{c}$ from Eq. (1) can be used to estimate the inundation fraction $\hat{f}$ as follows:

$$\hat{f} = \mathbf{F}_s \mathbf{c} \tag{2}$$

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

$$\hat{\mathbf{c}} = \underset{\mathbf{c}}{\operatorname{argmin}} \left\{ \left\| \mathbf{W} \left(\mathbf{b}_{\text{obs}} - \mathbf{B}_s \mathbf{c} \right) \right\|_2^2 \right\} \tag{3}$$

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, $K \gg n_c$, making $\mathbf{B}_s$ a rank-deficient matrix and the above problem ill-posed. To make the optimization problem (3) well-posed, we use a mixed $\ell_1 - \ell_2$ norm regularization as follows:

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

which has been successfully deployed for passive microwave precipitation retrievals (Ebtehaj et al., 2015a, 2015b). The non-negativity 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 $\|\mathbf{c}\|_1 = \sum_{i=1}^{K} |c_i|$ and the ℓ_2-norm $\|\mathbf{c}\|_2 = (\sum_{i=1}^{K} |c_i|^2)^{\frac{1}{2}}$. In this mixed regularization, the ℓ_1- norm leverages sparsity in the solution (i.e., forces some of the elements of $\mathbf{c}$ to be zero) while the ℓ_2-norm increases the stability of the solution as the neighboring brightness temperatures in $\mathbf{B}_S$ 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 assign non-zero coefficients to those clusters of nearby brightness temperatures, while the ℓ_2-norm handles the potential correlation between those clustered neighbors and makes the problem sufficiently stable. The proposed algorithm is summarized in a flowchart 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 $\mathbf{B}$ and $\mathbf{F}$ 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).

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-560, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 1 November 2016

Hydrology and Earth System Sciences
Discussions

Since the DMSP satellites have two different equatorial crossing times, here we use two sets of dictionaries for Tb values in the ascending (also called day or morning) and descending (also called night or evening) orbits. From all the available coincident observations we randomly chose 2×10^6 pairs of brightness temperature and inundation fractions in each ascending and descending dictionary. The purpose of stratifying the dictionaries into ascending and descending Tb orbits is to exclude the effects of Tb modulations from the retrieval process caused by the systematic diurnal variation of temperature. In other words, the same inundation fraction has different PMW spectral signature in a daytime versus a night-time 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). Fig. 4(a) presents the systematic difference between the Tbs of the ascending versus descending tracks for various ranges of a pixel's inundated fraction. 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 $\mathbf{F}$. Then for each interval, the average of Tb values is shown. The plot clearly demonstrates that the daytime Tbs are thermally warmer than their night time 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. Fig. 4(b) depicts $\left|\mathrm{Tb}_A - \mathrm{Tb}_D\right|$ where Tb_A and Tb_D stands for Tbs over ascending and descending overpasses, respectively. It can be observed clearly that the $\left|\mathrm{Tb}_A - \mathrm{Tb}_D\right|$ shows the coastlines, the regions with the transient presence and/or absence of water over land.

In the detection step, we found that $K \geq 50$ gives rise to a reasonable detection skill in terms of the probability of hit. 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 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 against both the underlying land surface emissivity and the atmospheric conditions.

For the 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, we empirically found that $\lambda = 0.001$ and $\alpha = 0.1$ provide a reasonable balance

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-560, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 1 November 2016

between sparsity and stability of the solution in Eq. (4). It shall be noted that problem (4) is converted to a constrained quadratic programming problem and solved using an iterative Newton's method with MATLAB optimization Toolbox (see Branch and Grace, 1996).

4 Results, Validation and Discussion

The inundation fractions are estimated during the wet period of the calendar year 2015 from July-to-December when the water levels across the delta begin to rise and eventually recede (see Fig. 7). 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 $P(\hat{f} > 0 \mid MWP > 0)$ and false alarm $P(\hat{f} > 0 \mid MWP = 0)$ of the algorithm outputs. Our analysis indicates that the probability of hit is around 0.92, demonstrating the capability of the algorithm in detecting the inundated areas. However, the probability of false alarm reaches the value of 0.34. The reason for this relatively high rate of false alarm can be due to the MODIS missing data. The MODIS daily data 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.

Fig. 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. Note that here for brevity we only show the results for ascending overpasses, while similar spatial patterns are observed for descending overpasses.

As mentioned before, the inter-annual climatology of the Mekong delta is highly affected by two tropical monsoons that characterize the seasonal patterns of precipitation and river stages and water levels (Delgado, et al., 2012). To better understand whether the results of the retrievals follow the regional climatology, the monthly percentage of the inundated area over the Mekong delta is calculated and shown against the monthly water level data in Fig. 7. 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). 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 approach as the surface water level data are positively correlated with the extent of the inundated surfaces. Fig. 7(a) shows that the temporal seasonal variations in the monthly

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-560, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 1 November 2016

percentage of the inundated surfaces obtained from the proposed model better follow the trend of monthly water elevation data that the standard MWP products. We can see that during the wet months of June-to-November the MWP data report much less total inundated area than the outputs of the proposed retrieval algorithm, while this pattern is reverse 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. For quantitative comparison of the outputs of the algorithm and MWP, the Euclidean distances between the normalized algorithm outputs and MWP with the normalized water level data are calculated. The square root of the Euclidean distance between water level data and the retrieved inundations from ascending and descending orbits are 3.46 and 3.56, respectively, while this distance for MWP and water level data is about 7.89 which is more than twice those distances calculated from the retrieval results. This indicates the superior performance of the proposed inundation fraction retrievals as compared to the MWP products.

Fig. 7(b) supports the assertion that the quality of the MWP products is affected by missing data due to cloudy skies by revealing a strong link between the wet months with a low percentage of inundated areas and high percentage of missing data in MODIS-MWP. The underestimation of the MWP in the dry months perhaps arises from of the general limitation of the empirical Bayesian estimation method regarding the extreme events (see, Petty 2013) and we suspect that it is not just limited to the months of January-to-March but it affects the retrievals at the other months to a lesser extent as well. We also suspect that the dictionaries are under-sampled with respect to the extreme scenarios during the warm months of the year. In other words, the dictionaries may not have an adequate sample, which eventually leads to underestimation. We expect that by improving the representativeness of the database—especially for extreme events in the summer, this shortcoming might be improved.

A closer look of Fig. 7(a) 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 diurnal patterns of the precipitation over the Mekong delta. In effect, 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 daily ascending inundation fractions obtained from MWP and the proposed algorithm with the daily water level data obtained by averaging over the same 11 gages in Fig. 1. Characterization of this dependence is performed using Spearman's rank correlation coefficient. The idea is that 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. These two correlation coefficient values simply

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-560, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 1 November 2016

Hydrology and Earth System Sciences Discussions

summarize the order between the inundation and water level data. To examine the dependence structures across different ranges of inundation and water level quantiles, we also calculated the empirical Copula (see Appendix 1).

The Copula is defined as the joint cumulative distribution of the quantiles of two random variables and contains all the information of their probabilistic dependence either linear or nonlinear. When two random variables are completely dependent, their Copula values are maximum at the diagonal arrays of their quantile matrix (when $i = j$ e.g., at quantiles 0.1 & 0.1 and 0.2 & 0.2, etc.) and almost zero for non-diagonal arrays ($i \neq j$). Therefore, the values of the Copula cumulative density function (CDF) do not change from the quantiles at the diagonal arrays (e.g., 0.1 & 0.1) to the non-diagonal arrays (0.1 & 0.2, 0.1 & 0.3, …) since zero values are added to the Copula CDF. This is the reason for having the L-shaped Copula CDF contours when two variables are completely dependent. Therefore, an "L-shaped" cumulative Copula, which is more skewed toward the origin $(x = 0, y = 0)$, implies a larger inter-quantile dependence between the two variables.

The empirical Copulas, here employed to capture the level of dependency between the daily inundation fractions ($\hat{f}$) and water level values, are calculated and plotted in Fig. 8. In this figure the axes values show the marginal quantiles of each variable and the contours trace the cumulative Copulas at different quantile levels. As the visual inspection of the monthly data also revealed, the empirical Copula of the total daily inundation fraction from the proposed algorithm and water level (*wl*) values shows higher degree of dependence compared to that of MWP and *wl* for medium (< 0.8) quantiles.

5 Conclusions and Future Directions

In this paper, we introduced a methodology to retrieve large-scale inundation form space for almost all-sky conditions to reduce the gaps that exist in using visible-to-near infrared satellite data. The key idea of the proposed method was to explore the links between overlapping daily high-resolution observations in the visible and near infrared range from the MODIS and the lower-resolution passive microwave observations from the Special Sensor Microwave Imager/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 related to 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. Analyses also showed that, at a daily time scale, the outputs of the algorithm exhibit stronger dependence with the water level data than the MWP data.

For further improvement of the proposed method, it is important to enhance the spatial resolution of the estimated inundation fractions by topographic guidance obtained by high-resolution elevation maps from digital elevation models (DEMs, see

Galantowicz, 2002) as ancillary information. Another way to improve the accuracy of estimation is to incorporate passive microwave observations from satellites with better spatial and temporal resolutions which do not necessarily have sun-synchronous orbits (such as the Global Precipitation Measurements (GPM) Microwave Imager (GMI)). However, it is critical to build the paired database independent of the cross-passing time of each orbit to avoid the effect of diurnal changes of land surface emissivity due to the temperature and precipitation variations. One approach to overcome this problem is to take the difference of brightness temperatures from various orbital passes in a day. This difference in Tb value can reveal the trace of flooding surfaces (as Fig. 4(b) also showed) that last less than a day. This is because of the rapid change in the land surface emissivity due to the presence/absence of water.

Acknowledgments: 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 Dr. Dan Slayback from the NASA Goddard Space Flight Center. First author would like to thank Professor Robert Brakenridge for his advice on this research during the AGU Fall Meeting 2015.

Reference

Allison, L. J., Schmugge, T. J. and Byrne, G.: A hydrological analysis of East Australian floods using Nimbus-5 electrically scanning radiometer data. Bulletin of the American Meteorological Society, 60(12), 1414-1427, doi: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. Sp. Res., 16(10), 155–163, doi: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. Meteor., 37(9), 888–911, doi: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, Transbound. floods reducing risks, 2006.

Brakenridge, G. R., Nghiem, S. V., Anderson, E. and Chien, S.: Space-based measurement of river runoff, Eos, Transactions, American Geophysical Union, 86(19), pp.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(4), 1–16, doi:10.1029/2006WR005238, 2007.

Branch, M. A. and Grace, A.: MATLAB: optimization toolbox: user's guide version 1.5., The MathWorks., 1996.

Choudhury, Bhaskar J. "Passive microwave remote sensing contribution to hydrological variables." Land Surface—Atmosphere Interactions for Climate Modeling. Springer Netherlands, 63-84, 1991.

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(1), 48–62, 2011.

De Groeve, T.: Flood monitoring and mapping using passive microwave remote sensing in Namibia, Geomatics Nat. Hazards Risk, 1(1), 19–35, doi: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(5), 1533–1541, doi: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., 1–18, doi:10.1175/JHM-D-15-0164.1, 2015.

Ebtehaj, A. M., Bras, R. L. and Foufoula-Georgiou, E.: Shrunken Locally Linear Embedding for Passive Microwave Retrieval of Precipitation, IEEE Trans. Geosci. Remote Sens., 53(7), 3720–3736, doi:10.1109/TGRS.2014.2382436, 2015.

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, Glob. Planet. Change, 50(1–2), 63–82, doi: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 Trans. Geosci. Remote Sens., GE-24(6), 1008–1013, doi:10.1109/TGRS.1986.289564, 1986.

Frazier, P. S., and Page, K. J.: Water Body Detection and Delineation with Landsat TM Data, Photogramm. Eng. Remote Sens., 66(12), 1461–1467, doi:0099-1112I0OI6612-1461$3.00/0, 2000.

Galantowicz, J. F.: High-resolution flood mapping from low-resolution passive microwave data, IEEE Int. Geosci. Remote Sens. Symp., 3(C), 1499–1502, doi: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(3), 257–266, doi: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. International Journal of Remote Sensing, 10(10), 1673–1686, doi:10.1080/01431168908903998, 1989.

Grody, N. C.: Classification of snow cover and precipitation using the special sensor microwave imager, J. Geophys. Res., 96(D4), 7423, doi:10.1029/91JD00045, 1991.

Guerschman, J. P., Warren, G., Byrne, G., Lymburner, L., Mueller, N. and Dijk, A. Van: 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. Manag., 19(4), 333–347, doi: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(7), 8516–8542, doi: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(8), 2730–2748, doi:10.1080/01431161.2012.750037, 2013

Melack, J. M., Hess, L. L. and Sippel, S.: Remote sensing of lakes and floodplains in the Amazon Basin. Remote Sensing Reviews, 10(1-3), 127-142, doi: 10.1080/02757259409532240, 1994.

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

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(5985), 1517-1520, doi: 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, 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, doi: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(11), 4107–4119, doi: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(21), 4847–4866, doi:10.1080/01431160600675887, 2006.

Petty, G. W.: Dimensionality reduction in Bayesian estimation algorithms, Atmos. Meas. Tech., 6(9), 2267–2276, doi:10.5194/amt-6-2267-2013, 2013.

Prigent, C., Matthews, E., Aires, F. and Rossow, W. B.: Remote sensing of global wetland dynamics with multiple satellite data sets. Geophysical Research Letters, 28(24), pp.4631-4634, doi: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(12), 1993–2000, doi: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(164), 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. Assoc., 10(5), 1060–1081, doi:10.1111/j.1752-1688.1974.tb00625.x, 1974.

Rango, A. and Salmonson, V.: Regional flood mapping from space, Water Resour. Res., 10(3), 1944–7973, doi:10.1029/WR010i003p00473., 1974.

Rouse, J. W., Haas, R.H., Schell, J.A., Deering, D.W., Harlan, J.C.: Monitoring the vernal advancement of retrogradation of natural vegetation, NASA/GSFC, Type III, Final Report, p. 371 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(1), 15003, doi: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(1), 70–76, doi: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(16), 3055–3074, doi:10.1080/014311698214181, 1998 .

Smith, L. C.: Satellite remote sensing of river inundation area, stage, and discharge: A review, Hydrol. Process., 11(10), 1427–1439, doi: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." www. Microimages. Com, 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(5720), 376-380, doi: 10.1126/science.1109454, 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(16), 2623–2634, doi:10.1002/hyp.8020, 2011.

Temimi, M., Leconte, R., Brissette, F. and Chaouch, N.: Flood monitoring over the Mackenzie River Basin using passive microwave data, Remote Sens. Environ., 98(2-3), 344–355, doi:10.1016/j.rse.2005.06.010, 2005.

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(6248), 638-643, doi: 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., (December), 1–9, 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 Sensing Reviews, 11(1-4), 163-194, doi: 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(14), 3025–3033, doi:10.1080/01431160600589179, 2006.

Zou, H. and Hastie, T.: Regularization and variable selection via the elastic net, J. R. Stat. Soc. Ser. B Stat. Methodol., 67(2), 301–320, doi:10.1111/j.1467-9868.2005.00503.x, 2005.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-560, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 1 November 2016

Hydrology and Earth System Sciences
Discussions

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	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 $\mathbf{B}$
$\mathbf{B}$	Brightness temperature dictionary
f	Inundation fraction
$\mathbf{F}$	Inundation dictionary
$\mathbf{b}_{obs}$	Observed vector of brightness temperature
K	Number of nearest neighbors
$\mathbf{B}_s$	Sub-dictionary of $\mathbf{B}$
$\mathbf{F}_s$	Sub-dictionary of $\mathbf{F}$
$\mathbf{c}$	Vector of representation coefficients
$\hat{f}$	Estimated inundation fraction
$\mathbf{W}$	Weight matrix
n	Number of frequency channels
ℓ_1 & ℓ_2	Regularizations norms
λ_1 & λ_2	Regularization parameters

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-560, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 1 November 2016

Hydrology and Earth System Sciences
Discussions

Appendix 1: Copula

Let X_1 and X_2 denote two random variables with marginal cumulative distributions $F_1(x_1) \equiv P\,[X_1 \le x_1]$ and $F_2(x_2) \equiv P\,[X_2 \le x_2]$ with the cumulative joint distribution function $F(x_1, x_2) \equiv P\left[X_1 \le x_1, X_2 \le x_2\right]$. 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:

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

where $C\left(u_1, u_2\right)$, 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 exists, can be calculated by taking the derivative of C and F which results in the following:

$$\begin{aligned} f(x_1, x_2) &= c(u_1, u_2).f(x_1).f(x_2) \\ &= c\left(F(X_1), F(X_2)\right).f(x_1).f(x_2) \end{aligned} \tag{2}$$

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 captures the probabilistic dependence between two random variables x_1 and x_2 by quantifying the strength of the relationship between their corresponding quantiles.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-560, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 1 November 2016

Figures

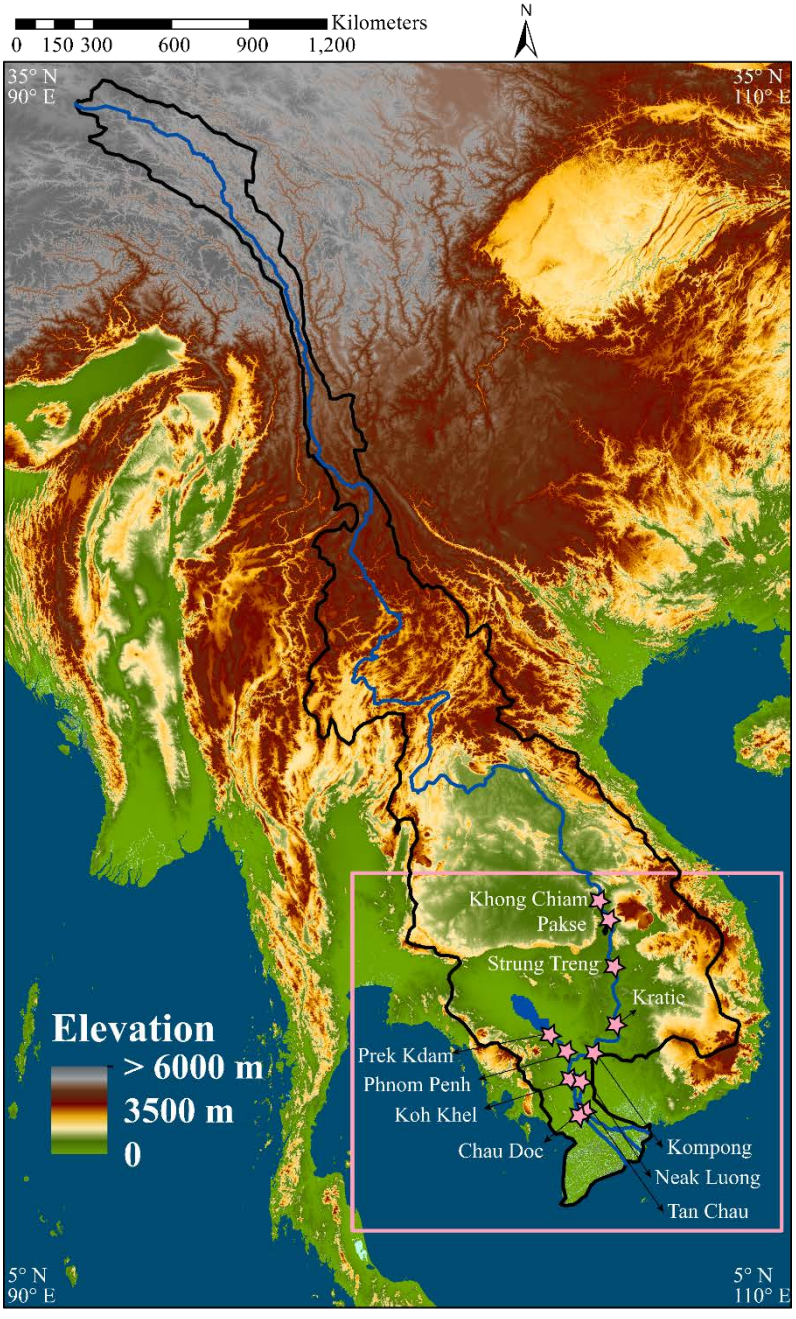

Figure 1. Map of the Mekong river basin and delta. The digital elevation map of the Mekong River Basin with the Mekong River (blue line), the boundary of the basin (black line), and the study area delineated by a pink rectangle. The Mekong River Basin covers an area of about 795,000 km^2. The 11 stations (from Mekong River Commission) that monitor the water level are also marked by pink stars.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-560, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 1 November 2016

Hydrology and Earth System Sciences
Discussions

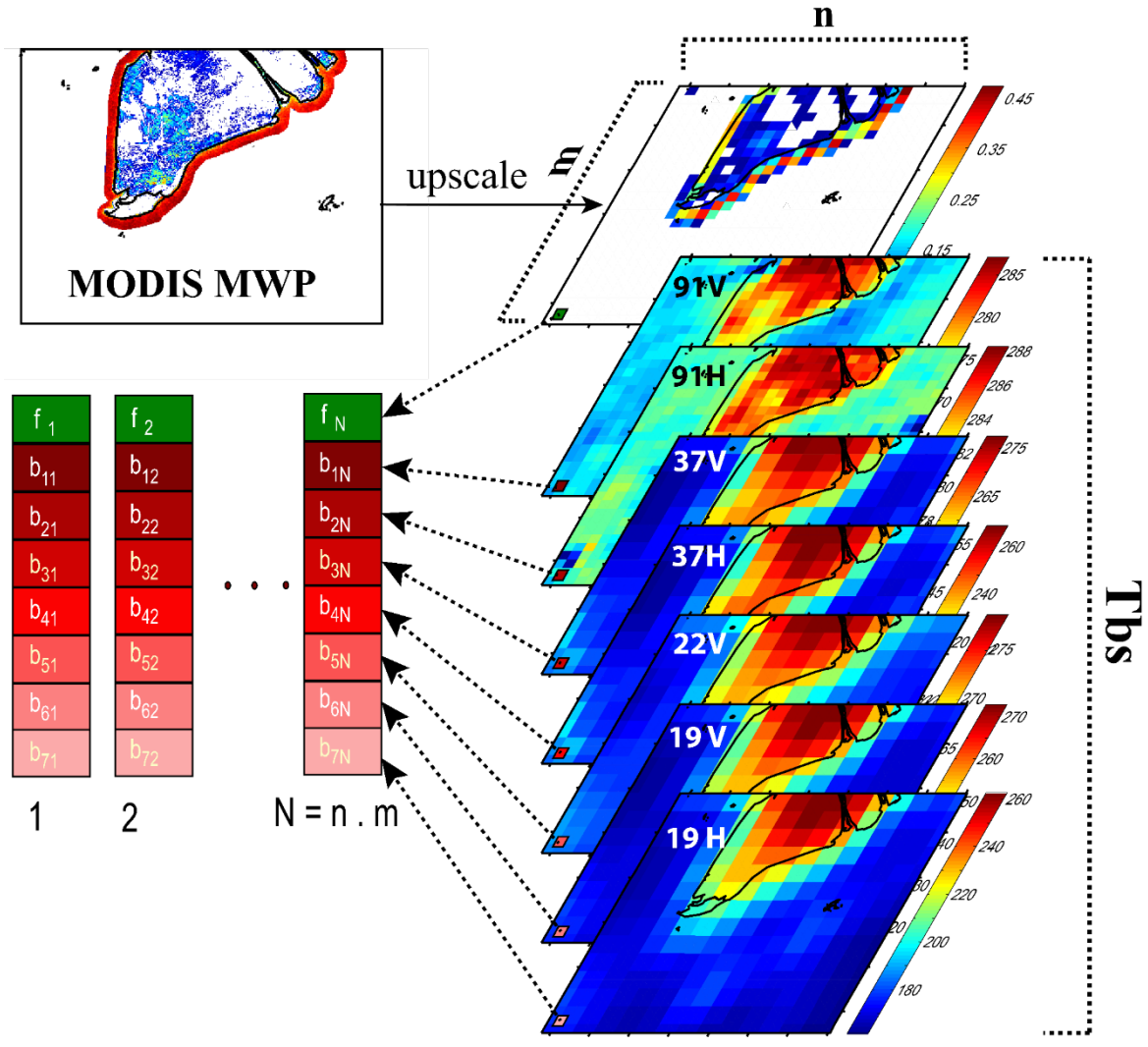

Figure 2. Schematic showing construction steps of the *a priori* database and dictionaries. The top slab is the upscaled MODIS-MWP and the other slabs are the brightness temperature data at seven frequency bands. Each vector on the left is created by stacking a pixel-level information of the multi-frequency brightness temperature by the SSMIS radiometer and the corresponding inundation from the MWP product at 12.5 km resolution. This process generates N number of vectors ($N = n \times m$) to form separate dictionaries for ascending and descending orbits.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-560, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 1 November 2016

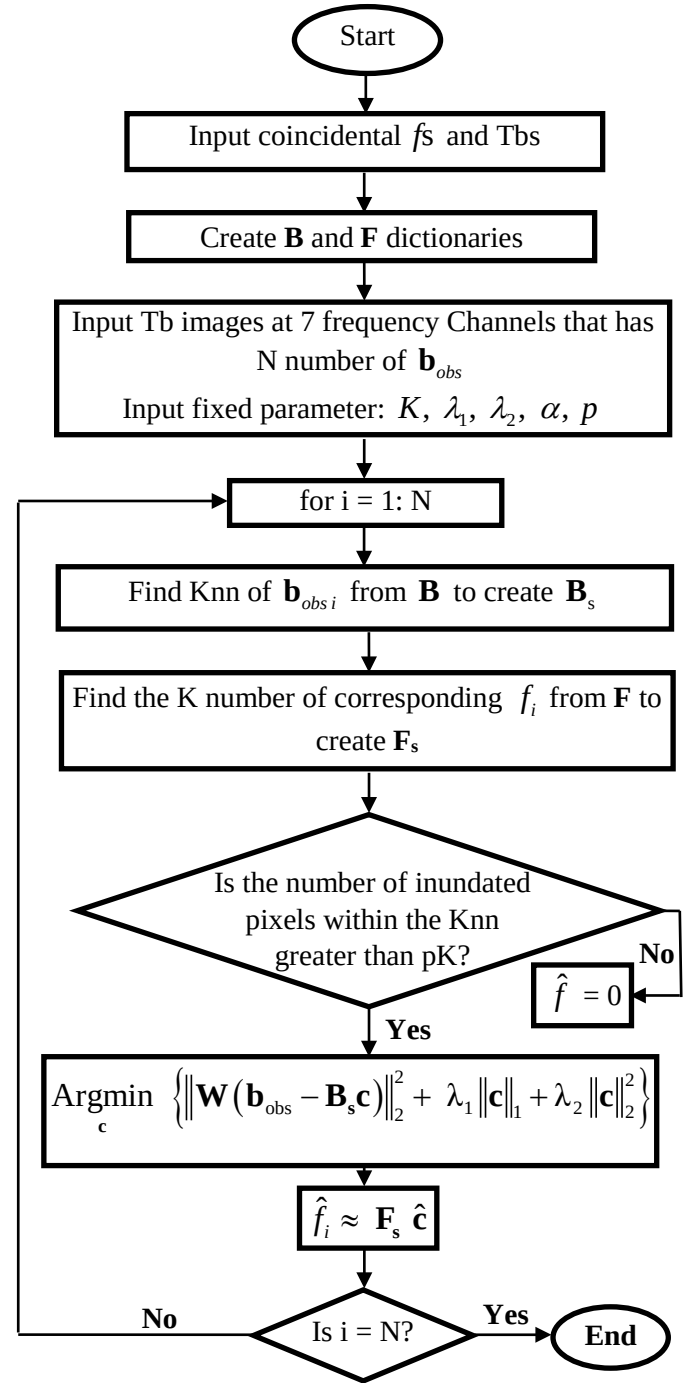

Figure 3. Flowchart of the inundation retrieval algorithm. See text for definitions of the notations and detailed explanation.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-560, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 1 November 2016

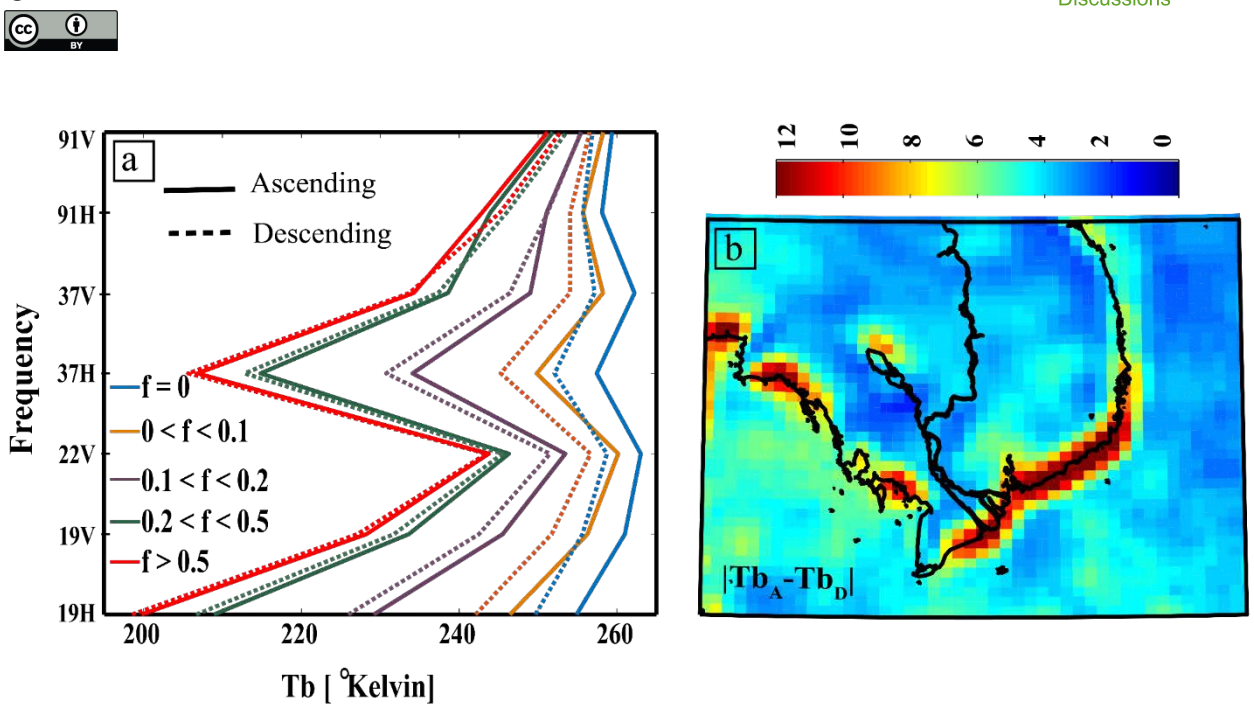

Figure 4. The systematic difference between passive microwave observations from the ascending (solid lines) and descending orbits (broken lines). (a) The averaged ascending and descending brightness temperatures over five different inundation intervals. (b) July-to-Dec daily average of absolute difference between the ascending (Tb_A) and descending (Tb_D) brightness temperatures at 19 GHz for the vertical polarization. The values of $|Tb_A - Tb_D|$ mapped in 4(b) captures the regions with the sub-daily transient presence and absence of water over land (coastlines). The significant systematic difference between Tbs from ascending and descending orbits over those regions caused by changes of land surface emissivity due to sharp sub-daily variations of temperature because of the appearance of water over land.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-560, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 1 November 2016

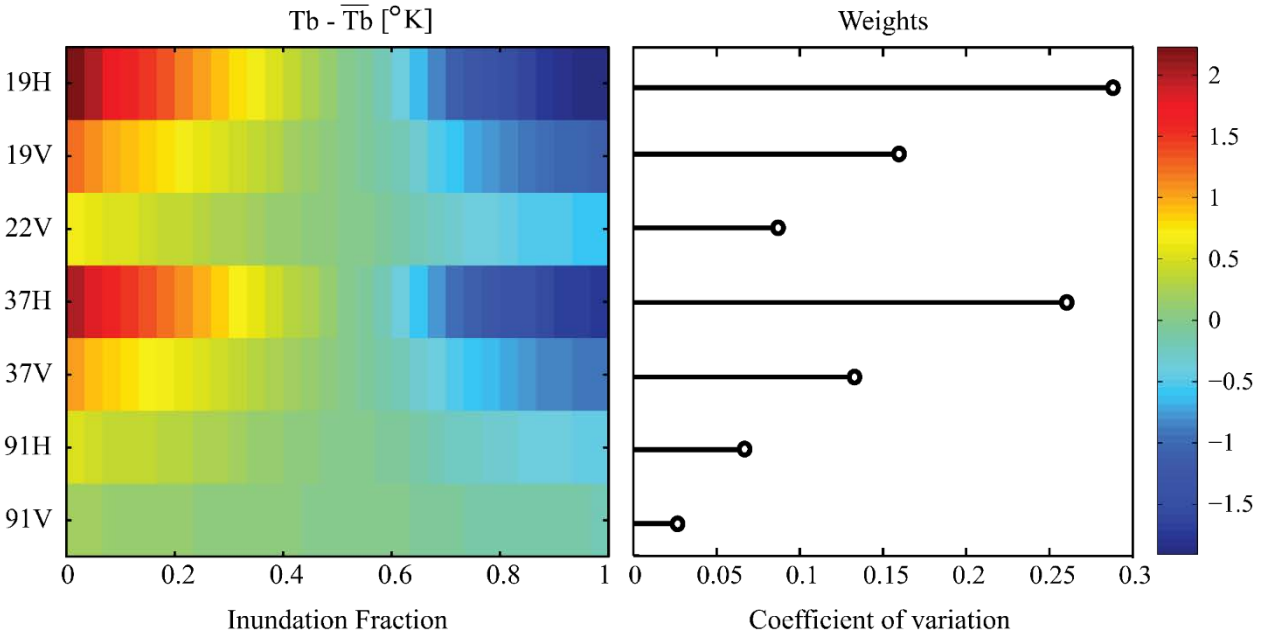

Figure 5. The normalized coefficients of variation (right panel) of the brightness temperatures (Tb) (left panel) averaged over the entire database for different intervals of inundation fractions. Here, $\overline{Tb}$ denotes the average of brightness temperatures over the inundation fractions. The 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 frequency channels to the inundation fraction and are given higher weights.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-560, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 1 November 2016

Hydrology and Earth System Sciences
Discussions

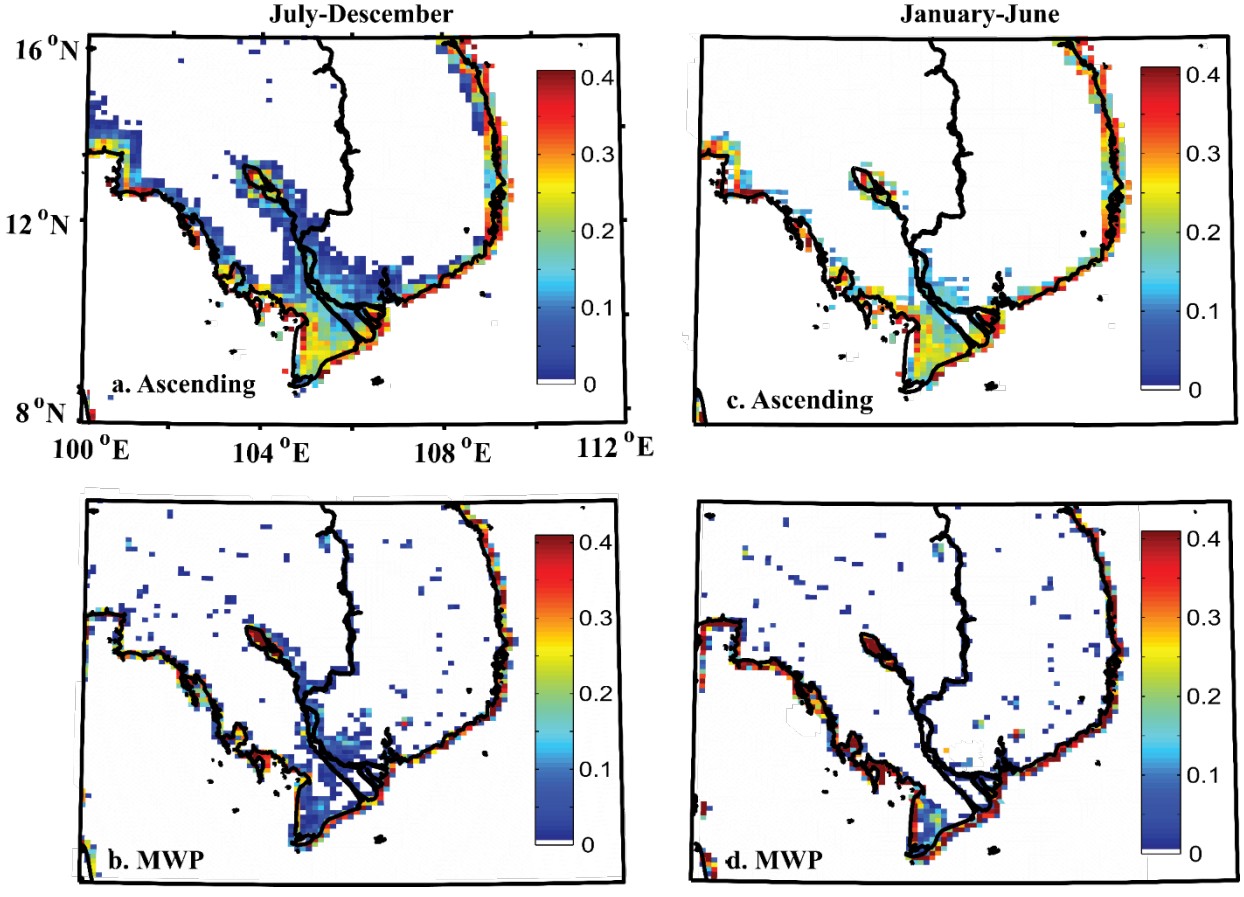

Figure 6. Inundated map of the Mekong delta in the wet (July-December) and dry (January-June) seasons. Based on the water level data collected from 11 stations in the Mekong delta (Fig. 1), two time periods of wet and dry seasons in calendar year 2015 are selected to show the inundated areas at resolution of 12.5 km averaged over all days in: July-to-Dec when the water level is relatively high (left column) and Jan-to-June when the water level is relatively low (right column). 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. Further quantitative comparison is shown in Fig. 7 and 8.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-560, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 1 November 2016

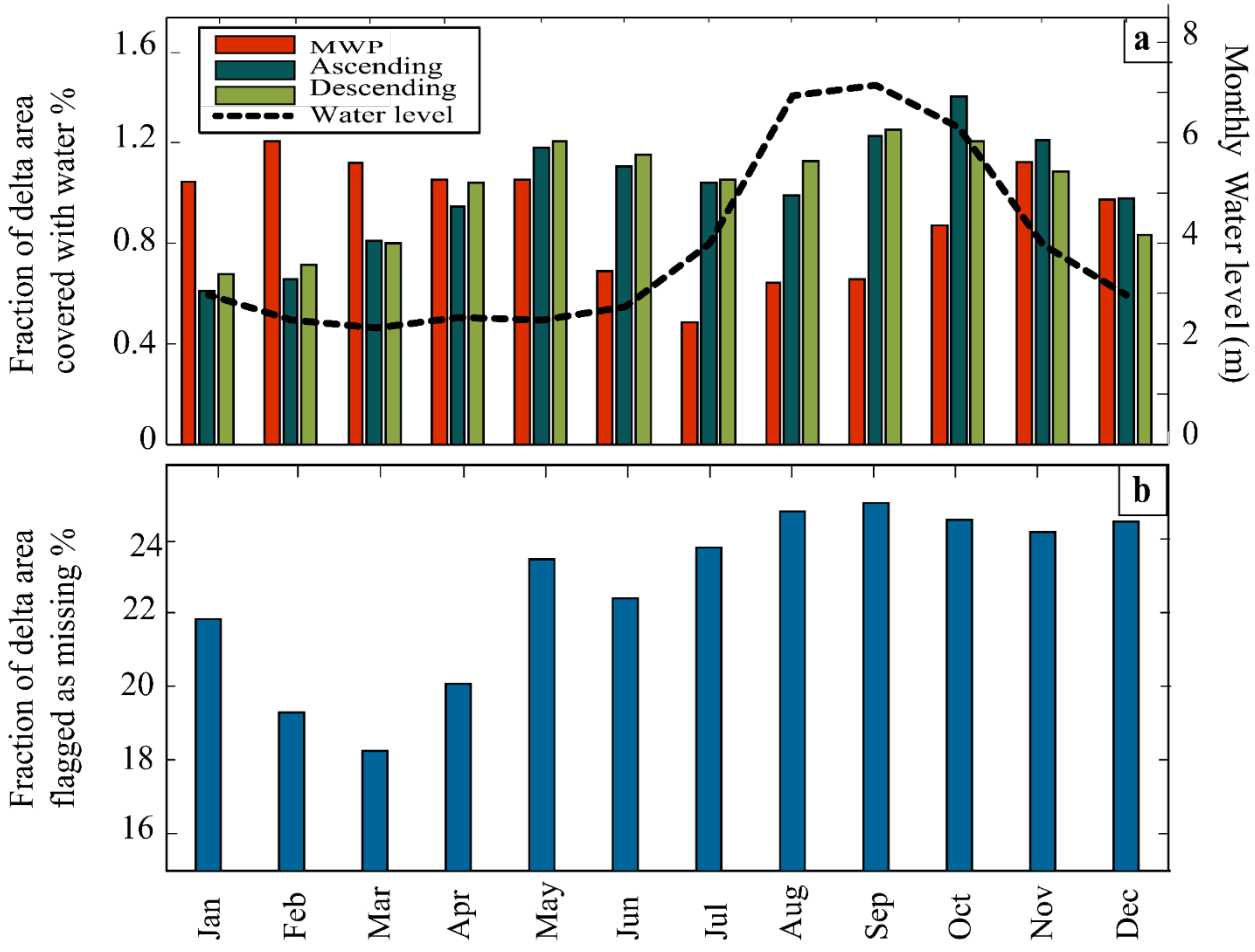

Figure 7. The total monthly inundated area of the Mekong delta calculated from the proposed retrieval algorithm and MODIS Near Real-Time (NRT) Water Product (MWP) data illustrated against the monthly water level data to evaluate their consistency with the climatology of the region. (a) Comparison of the total inundated surface of the Mekong delta from MWP products and from the retrieval algorithm from ascending and descending dictionaries, separately. From visual inspection, it is obvious that the retrieval algorithm can better follow variations of the water levels compared to MWP. More inundation over months of the dry season is reported by MWP products compared to the wet season which contradicts the causality between rivers' stages and more land under water (b) The total fraction of flagged land surface areas that are labeled as missing in MWP product because of atmospheric contaminations. The larger error of MWP products during the wet months is attributed to larger percentage of missing values.

Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-560, 2016
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Published: 1 November 2016

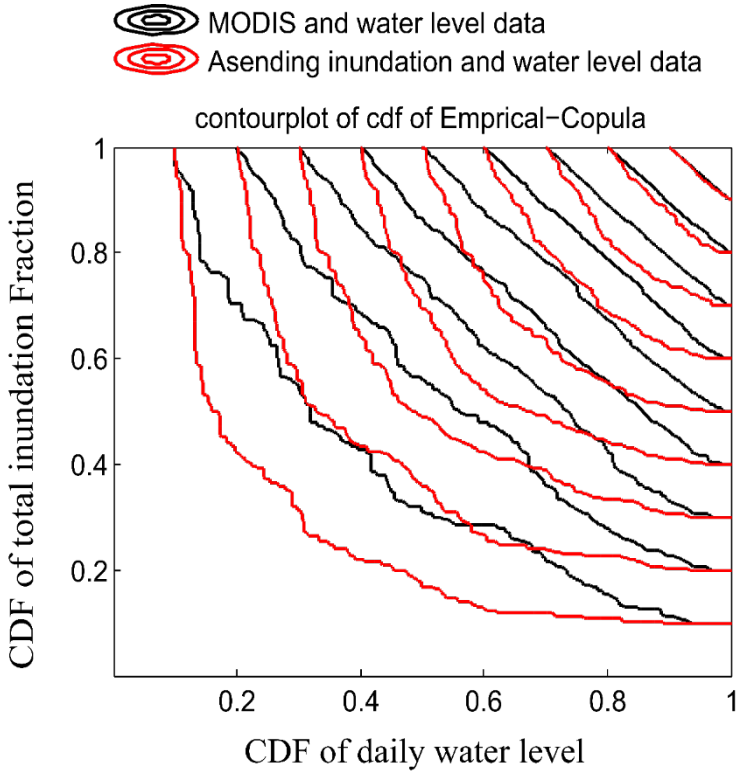

Figure 8. The empirical Copula of the average daily water level versus total daily inundated area from the proposed retrieval algorithm (red curves) and MODIS-MWP data (black curves) to capture the dependence between the marginal quantiles between the inundation retrievals and ground-based water levels. The plots denote that the dependence between the results of our algorithm and water levels is stronger than the MWP product.