Fast and Automatic Data-Driven Thresholding for Inundation Mapping with Sentinel-2 Data

Georgios A. Kordelas, Ioannis Manakos, David Aragonés, Ricardo Díaz-Delgado and Javier Bustamante

This supplementary material supports the main text as follows:

Supplementary 1. Sensitivity Analysis of the Unsupervised Approach.

1. Estimation of $\mathbf{T_{init}}$

There are a few dates of Sentinel-2 (S2) acquisition, where the first deep valley in a Short-Wave Infrared (SWIR) histogram may be spurious and cause an underestimation of the initial SWIR threshold. Histogram smoothing alleviates these spikes, while preserving the same basic histogram shape. Therefore, SWIR and smoothed SWIR histograms are taken into consideration for estimating initial threshold $\mathbf{T_{init}}$. Strong local valleys are evident in both SWIR and smoothed SWIR histograms (see Figure 1). As initial threshold is derived the SWIR value $\mathbf{T_{init}}$, corresponding to a deep valley in the SWIR histogram, which has the closest value to $\mathbf{T_{smooth}}$, which in turn corresponds to the first deep valley in the smoothed SWIR histogram. Figure 1 shows an example using S2 data acquired on 16/07/2016, where the SWIR histogram (blue line) and the smoothed SWIR histogram (red dashed line) are drawn. $\mathbf{T_{init}}$ and $\mathbf{T_{smooth}}$ are equal values, for which deep valleys in SWIR and smoothed SWIR histograms are, respectively, detected. These valleys are plotted with green and yellow color, respectively. $\mathbf{T_1}$ is the SWIR value for which the first deep valley denoted in the SWIR histogram. In the case of not using synergistically information from SWIR and Smoothed SWIR histograms $\mathbf{T_1}$ would be estimated as $\mathbf{T_{init}}$. This threshold would lead to a less accurate initial estimation of the inundated areas.

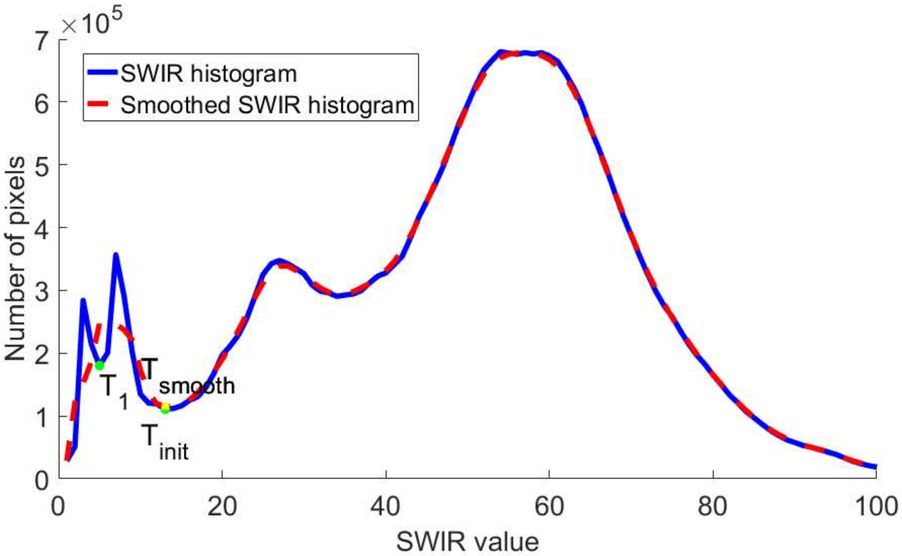
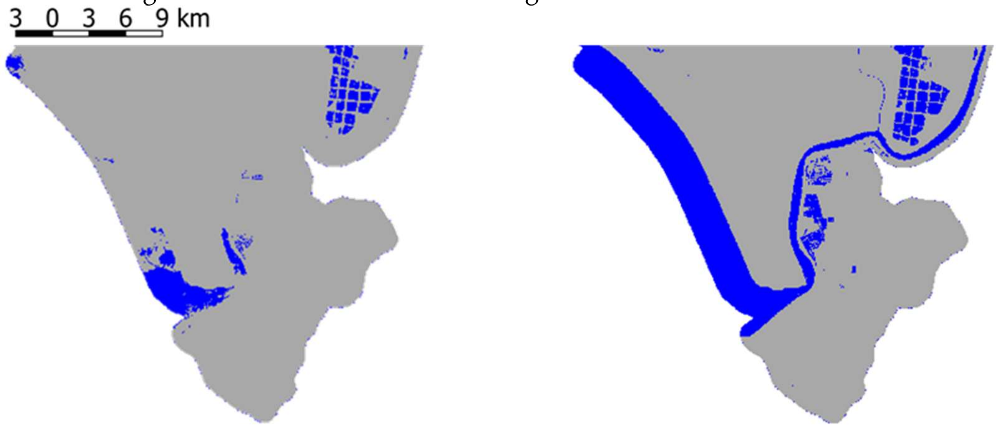


Figure 1. SWIR and smoothed SWIR histograms on 16 July 2016.

Figure 2a and Figure 2b show a zoomed region of the lower part of the inundation map, which is estimated after applying $\mathbf{T_1}$ and $\mathbf{T_{init}}$ to the SWIR band, respectively. Evidently, the inundation map of Figure 2a is less accurate than the one of Figure 2b, since some parts of the marine coastal water and the Guadalquivir river region are not classified as water regions.



(a) (b)

Figure 2. Lower part of the inundation map on 16/07/2016 after applying (a) $\mathbf{T_1}$ and (b) $\mathbf{T_{init}}$ thresholds.

An interesting experiment is to check how the estimation of the final threshold would be affected if $\mathbf{T_1}$ is set as the underestimated initial threshold $\mathbf{T_{init}}$ and then apply the next steps of the unsupervised methodology as described in section 2.4 of the main manuscript. The latter case is the "underestimation case" and in the following is compared to the "actual case" where $\mathbf{T_{init}}$ is estimated synergistically from SWIR and Smoothed SWIR histograms. On 16/07/2016, for the "underestimation case" $\mathbf{T_{init}} = \mathbf{T_1} = 5$, according to Figure 1. After applying the next steps of the unsupervised methodology it is estimated that $\mathbf{T_{final\text{-}1}} = 15.5$. For the "actual case" where $\mathbf{T_{init}} = 13$, according to Figure 1, it is estimated that $\mathbf{T_{final}} = 18$. These results show that for the "underestimation case" the difference $\mathbf{T_{final\text{-}1}} - \mathbf{T_1} = 15.5 - 5 = 10.5$ between underestimated $\mathbf{T_{init}}$ and estimated $\mathbf{T_{final\text{-}1}}$ is greater than the difference $\mathbf{T_{final}} - \mathbf{T_{init}} = 18 - 13 = 5$ between actual $\mathbf{T_{init}}$ and estimated $\mathbf{T_{final}}$. The analysis results on 16/07/2016 are included in the 1st row of **Error! Reference source not found.**. The same analysis is performed for two additional dates: 26/07/2016 and 11/07/2017 and the analysis results are also included in **Error! Reference source not found.**. The results for the three examined dates, indicate that the proposed method is able to cope with $\mathbf{T_{init}}$ underestimation, since the value of the estimated $\mathbf{T_{final\text{-}1}}$, when $\mathbf{T_{init}}$ is underestimated, is close to the value of the estimated $\mathbf{T_{final}}$, when the actual $\mathbf{T_{init}}$ is used.

Table 1. Estimated $\mathbf{T_{final}}$ for the cases of using the "underestimated" $\mathbf{T_{init}}$ ($\mathbf{T_{init}} = \mathbf{T_1}$) and the "actual" $\mathbf{T_{init}}$ as initial threshold and the differences between $\mathbf{T_{final}}$ and the corresponding $\mathbf{T_{init}}$.

	Underestimation Case			**Actual Case**		
Date	$\mathbf{T_1}$	$\mathbf{T_{final\text{-}1}}$	$\mathbf{T_{final\text{-}1}} - \mathbf{T_1}$	$\mathbf{T_{init}}$	$\mathbf{T_{final}}$	$\mathbf{T_{final}} - \mathbf{T_{init}}$
16 July 2016	5	15.5	10.5	13	18	5
26 July 2016	4	13	9	9	14	5
11 July 2017	3	16	13	11	17	6

2. Deviation from the proposed unsupervised approach

The existence of the water-vegetation subclass is reflected in the histogram of the SWIR band of S2 data acquired on 20/08/2017, where two deep valleys can be detected (see Figure 4 in the main manuscript). However, there are other seasons within a year when the water-vegetation subclass cannot be derived, since areas covered concurrently by water and dense vegetation of high greenness are not present. This fact is reflected in the histograms of the SWIR band and the Modified Normalized Vegetation Index (MNDVI) index, suggesting that if there are areas covered with water and vegetation, these are of limited extent, if any, or vegetation has a low density, within the study area. For example, in the histogram of the SWIR band of S2 acquired on 29/12/2015, which is visualized in Figure 3a only one valley is noticed. Additionally, in the histogram of the MNDVI index, which is visualized in Figure 3b, it cannot be detected a deep valley over 0.4. The inundation map generated for this case is visualized in Figure 4.

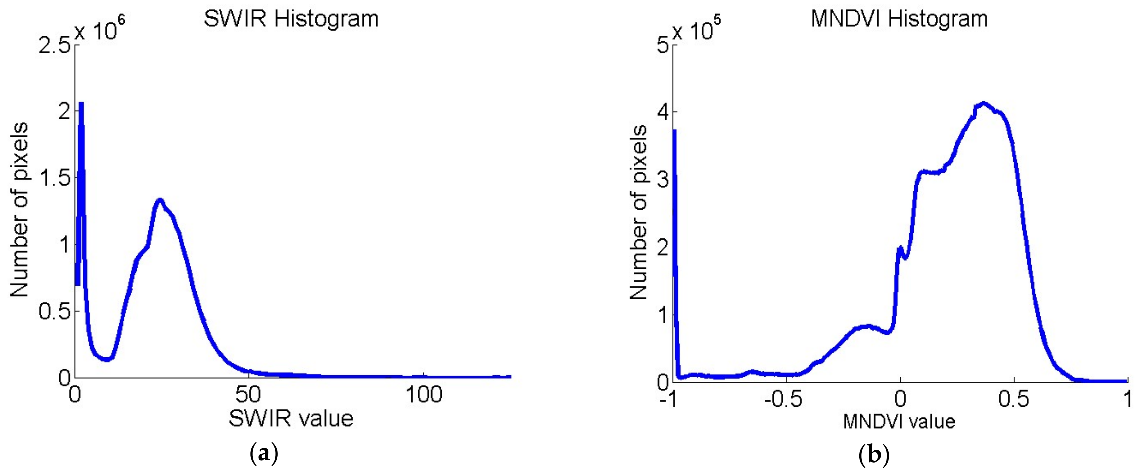


Figure 3. Histograms of (a) SWIR band and (b) MNDVI index on 29 December 2015.

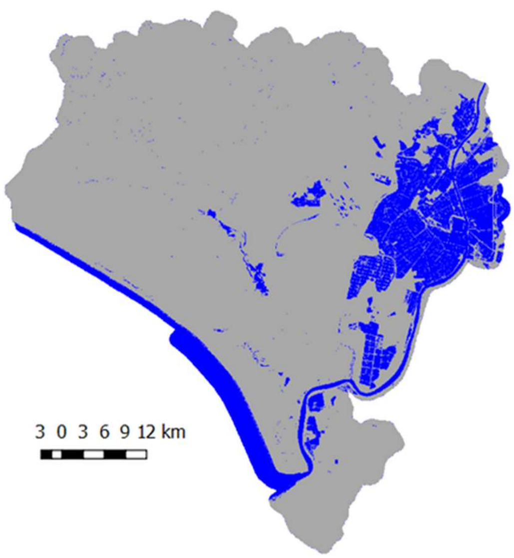


Figure 4. Inundation map on 29 December 2015 containing only open-water subclass pixels.