

Towards an automated approach to map flooded areas from Sentinel-2 MSI data and soft integration of water spectral features

Alessia Goffi^{a,b}, Daniela Stroppiana^{b,*}, Pietro Alessandro Brivio^b, Gloria Bordogna^b, Mirco Boschetti^b

^a Terraria Srl, via M. Gioia 132, 20125, Milano, Italy

^b Istituto per il Rilevamento Elettromagnetico dell'Ambiente, Sezione Secondaria di Milano, Consiglio Nazionale delle Ricerche, Via Bassini 15, 20133, Milano, Italy



ARTICLE INFO

Keywords:

Multi-source soft integration
Ordered Weighted Averaging
Spectral indices

ABSTRACT

In this work we propose an approach for mapping flooded areas from Sentinel-2 MSI (Multispectral Instrument) data based on soft fuzzy integration of evidence scores derived from both band combinations (i.e. Spectral Indices - SIs) and components of the Hue, Saturation and Value (HSV) colour transformation. Evidence scores are integrated with Ordered Weighted Averaging (OWA) operators, which model user's decision attitude varying smoothly between optimistic and pessimistic approach. Output is a map of global evidence degree showing the plausibility of being flooded for each pixel of the input Sentinel-2 (S2) image. Algorithm set up and validation were carried out with data over three sites in Italy where water surfaces are extracted from stable water bodies (lakes and rivers), natural hazard flooding, and irrigated paddy rice fields. Validation showed more than satisfactory accuracy for the OR-like OWA operators ($F\text{-score} > 0.90$) with performance slightly decreased ($F\text{-score} < 0.75$) over heterogeneous conditions (e.g. rice fields). The algorithm was applied with no changes and/or tuning to independent sites from the Copernicus Emergency Management Service (EMS) activations to simulate operational conditions. Over these sites, the proposed approach achieved greater, more consistent and robust mapping accuracy compared to traditional approaches based on the segmentation of single input features. Moreover, OWA operators offer an appealing way of combining and aggregating multiple information in decision making by modelling uncertainty in decision process.

1. Introduction

Floods are among the most catastrophic natural disasters causing important and/or permanent damages to infrastructures and communication systems, crops and livestock, and property but more relevant, floods provoke loss of human lives. It is estimated that worldwide more than 5000 deaths are caused by flash floods, that are events occurring on small spatial scales with short time scales under conditions of rapid production of surface runoff (Jonkman, 2005). In the future, climate change could exacerbate these phenomena by increasing the frequency of extreme and adverse meteorological events (Kharin et al., 2007). In 2007 the European Commission (EC) adopted the Floods Directive for the assessment and management of flood risk. Remote sensing (RS) is widely recognized as a unique source of data for the implementation of EU directives since it provides synoptic view over large areas, frequent observations and historical archives (Bresciani et al., 2011). Flood mapping could also serve other stakeholders and purposes such as risk management, land use and land management, emergency planning.

RS data have been largely used to map the extent of flooded area worldwide and with instruments operating in the visible, thermal and microwave wavelengths of the Electromagnetic spectrum (ES). Satellite systems represent the most widely used platform for large area mapping and emergency management (Giordan et al., 2018); an operative example is the EC Copernicus Emergency Management Service (Copernicus EMS) (<http://emergency.copernicus.eu>, accessed July 2019).

SAR (Synthetic Aperture Radar) data are a reliable source of information since they provide data in any weather and illumination condition especially during flood events characterised by adverse meteorological conditions. However, if a visible image is available simultaneous and/or close to the peak of the flooding event, retrieving information from optical data is often more straightforward (Schumann, 2015) and optical images have been largely exploited since the early 70's (Robinov, 1978; Rokni et al., 2014; Kumar et al., 2018; Rahman and Di, 2017; Huang et al., 2018).

The European Space Agency's (ESA) Sentinels have started an Earth Observation (EO) new era especially in disaster monitoring and

* Corresponding author.

E-mail addresses: goffi@terraria.com (A. Goffi), stroppiana.d@irea.cnr.it (D. Stroppiana).

emergency management. Sentinel-2 (S2) mission offers advanced spatial resolution and frequency of acquisitions compared to previous satellite missions making available passive optical data, e.g. Landsat. Landsat data, however, still constitute the only historical consistent archive of EO data for long-term flood mapping and monitoring (Díaz-Delgado et al., 2016; Mueller et al., 2016). ESA's Sentinels together with on demand commercial mission data support operational rapid mapping of floods by Copernicus EMS.

Most of the methods proposed in the literature to map standing water with optical data rely on Spectral Indices (SI) or band combinations in colour space, such as the Hue Saturation and Value transformation (Pekel et al., 2014). Water-related indicators are proposed as combinations of reflectance in the shortwave infrared (SWIR, 1250–2500 nm) and the near infrared (NIR, 700–1250 nm) (Gao, 1996) or visible (VIS, 350–700 nm) spectral regions (Li et al., 2014); however, a water index can be found in different combinations of bands as well as with different nomenclature for the same formulation (Boschetti et al., 2014). Mapping is based on the segmentation of a single SI by applying crisp thresholds to single date acquisitions and/or temporal difference; this approach is widely used for the identification of water surfaces because it is fast and simple (Acharya et al., 2018). However, mapping performance for the selected index/threshold varies as a function of study area, water characteristics, environmental and atmospheric conditions: hence, the combination of more indices can lead to better results especially when different conditions coexist in complex scenes (Acharya et al., 2018).

In this study, we propose a method to synergistically combine water evidence scores, i.e., features, building on their complementarity/concurrency to perform a reinforcement of evidence. The approach implements a soft integration of multiple partial evidence scores derived from possibly semantically heterogeneous factors. In this way, we can achieve greater, more consistent and robust mapping accuracy compared to traditional approaches based on segmentation of single water/spectral index. The major long-term objective is to lay the foundation of an automated algorithm for mapping flooded areas requiring less a priori sets and, above all, capable to cope with choices taken under imprecise information, compared to more traditional methods proposed in the literature.

The approach, which was first defined by Bordogna et al. (2007) and Carrara et al. (2008) and applied for the detection of surfaces affected by fires (Stroppiana et al., 2015, 2012), here is extended to map flooded areas from Sentinel-2 imagery by integrating heterogeneous input spectral features: SIs selected from the literature and the H/V components of Hue-Saturation-Value (HSV) colour space that were proposed by Pekel et al. (2014), (2017).

2. Study areas and data

2.1. Study sites

Six sites were selected for algorithm set up and the assessment of mapping products (Fig. 1, Table 1). Three sites were selected in Northern Italy (Cal/Val) for algorithm set up where training and testing pixels were extracted for i) definition of the “standing water” soft constraints, ii) test of different multi-source soft integration operators and iii) validation of algorithm performance. Sites were selected to cover different conditions of standing water in order to capture variable spectral characteristics: flooded area is due to extreme heavy rainfall (Cal/Val_1), river bed (Cal/Val_2) and flooded rice fields (Cal/Val_3). The latter site was selected, although flooding was not due to a natural event, to train and validate the algorithm over heterogeneous conditions of shallow water surface (< 50 cm) mixed with soil patches and vegetation condition (See Supplementary materials S1); moreover, a large dataset of *in situ* observations was available from field surveys.

Given the variability of surface water conditions and characteristics in the Cal/Val sites, we refer to “standing water” rather than flooded

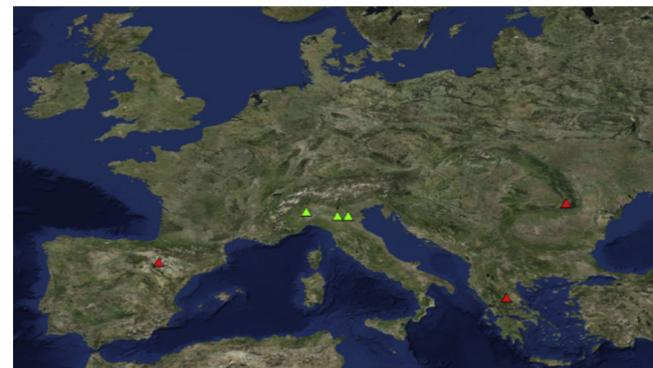


Fig. 1. Location of the calibration sites in Italy (green triangles) and evaluation sites in Greece, Spain and Romania (red triangles). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

areas. But since the primarily objective of the study remains to build an algorithm for automatic mapping of flooded areas, the Evaluation sites were selected among the list of activations of the Copernicus Emergency Management Service (EMS), which provides rapid mapping products to support emergency management activities immediately following a disaster (<http://emergency.copernicus.eu>, accessed July 2019).

Copernicus recent activations were screened to identify flood events with available clear sky post-event Sentinel-2 images close in time to the date of EMS mapping and sites were located in Greece (Farkadona/Thessaly), Spain (Tudela) and Romania (Ozun). Copernicus Emergency's post-event maps, based on COSMO-SkyMed (CSK-Very High Resolution Copernicus Contributing Mission) images were used as external and independent reference data for the evaluation phase (Table 3).

2.2. Sentinel-2 data

The constellation of Sentinel satellites is part of the Copernicus Earth Observation program led by the European Commission and operated by the European Space Agency (<https://earth.esa.int/web/sentinel/home>, accessed July 2019).

S2 mission operates as part of a two-satellite system (A&B) carrying on board the MultiSpectral Instrument (MSI) providing high resolution multispectral optical imagery since June 2015 (A) and March 2017 (B). MSI measures the Earth's reflected radiance in 13 spectral bands from VIS/NIR to SWIR with a spatial resolution ranging from 10 m to 60 m. Its wide field of view (290 km) makes possible a global coverage of the Earth's land surface (between 56 °S and 83 °N) every 10 days with one satellite and every 5 days with two satellites (A&B).

The proposed algorithm builds on S2 data collected for post-event assessment (after flooding occurrence). Dates of S2 images over Cal/Val and Evaluation (Fig. 2) sites are listed in Tables 2 and 3, respectively. Level-2A S2 images were downloaded and pre-processed with sen2r toolbox (Ranghetti and Busetto, 2019), developed in R and released under the GNU General Public License version 3 (GPL-3) freely available on GitHub (<https://ranghetti.github.io/sen2r>, accessed July 2019). The sen2r toolbox accesses image archives through the Copernicus Open Access Hub which provides complete, free and open access to S2 imagery and allows customized pre-processing of multispectral surface reflectance, spectral indices and RGB images. For Cal/Val_1, Cal/Val_2 and Evaluation sites, Level-2A S2 imagery was downloaded as Bottom of Atmosphere (BOA) reflectance, and pre-processing consisted in clipping images to our area of interest and masking clouds: pixels classified as high and medium cloud probability were masked out, while pixels belonging to different classes were retained to avoid masking out water pixels. Within the sen2r toolbox, cloud cover

Table 1

Location/extents of the study sites and characteristics/conditions of the surface water areas.

Site	Location	Lat	Lon	Surface water conditions	Dimension (km ²)
Cal/Val_1	Emilia (IT)	44.968861	10.649674	Flooded areas due to severe rainfall events	2090
Cal/Val_2	Po Valley (IT)	44.992491	11.377019	River in standard conditions	546
Cal/Val_3	Vercelli Province (IT)	45.278927	8.527552	Flooded rice fields	1937
Eval_1	Farkadona (GR) ¹	39.510748	21.968400	EMS flooding events - natural hazards	1354
Eval_2	Tudela (ES) ²	41.918886	1.360486		2722
Eval_3	Ozun (RO) ³	45.849476	25.979288		1049

<https://emergency.copernicus.eu/mapping/list-of-components/EMSR271>.

<https://emergency.copernicus.eu/mapping/list-of-components/EMSR279>.

<https://emergency.copernicus.eu/mapping/list-of-components/EMSR293>.

probability is obtained from the Scene Classification Map (SCL), which is produced by Sen2Cor algorithm and provided with level 2A imagery, and includes the Quality Indicators for cloud probabilities at 20 m resolution. For Cal/Val_3, BOA images were not available in the Copernicus archive and atmospheric correction of Top of Atmosphere Level-1C products was carried out in the sen2r toolbox using Sen2Cor algorithm.

2.3. Reference dataset

The reference dataset is composed of pixels labelled as *water* and *not water* collected over the six sites. Over the three Cal/Val sites, pixels were extracted for algorithm set up (i.e. training) and validation; the *not water* category includes vegetation, agricultural fields, cloud and cloud shadows, urban and industrial areas. Training pixels over Cal/Val_1 and 2 were collected by photo-interpretation of different combination of S2 false colour composites; over Cal/Val 3 *in situ* observations were available from field surveys carried out in the framework of the EU-FP7 ERMES project. For the collection of validation pixels over the same sites, a stratified random sampling scheme was adopted to extract over each site about 1200 pixels, which were independently labelled as *water* and *not water*; uncertain pixels (e.g. presence of clouds and cloud shadows) were discarded leading to a total number of testing pixels lower than 1200 (Table 2). More information is provided in Supplementary materials (S1).

Over the Evaluation sites no training pixels were used, and reference pixels used to assess the accuracy of the mapping products were extracted from Copernicus EMS products.

2.4. Water spectral features

2.4.1. Spectral indices

A set of spectral indices was selected from the literature and computed within the sen2r toolbox (Table 4). The AWEI and AWEI_{SH} indices were proposed by Feyisa et al. (2014) to improve classification accuracy in areas that include shadows and dark surfaces which are often misclassified as water. The NDWI (Normalized Difference Water Index), mNDWI (modified NDWI) and NDFI (Normalized Difference Flooding Index) were chosen among the normalized difference indices based on Boschetti et al. (2014) who identified that indices relying on visible and SWIR reflectance are the best performing for surface water detection. NDWI was proposed first by McFeeters (1996) to identify open water surfaces using GREEN and NIR reflectance. The modified NDWI (mNDWI) was introduced by Xu (2006) to improve water discrimination even in presence of high turbidity and to reduce confusion with built up areas. Finally, SAVI was selected to represent the land/vegetation category.

2.4.2. Hue - Value colour space components

The Hue, Saturation and Value (HSV) colour space transformation was applied to S2 SWIR2, NIR, RED bands to extract the H/V components. Indeed, it has been demonstrated that, in the HSV colour space,

standing water can be effectively estimated by defining a relation between H and V components (Pekel et al., 2016, 2014). The approach, initially developed for MODIS (Pekel et al., 2014) and then adapted to Landsat data (Pekel et al., 2016), is in the process of being extended to S2 (Pekel et al., 2017); it is also the basis for the generation of operational Copernicus water bodies products (<https://land.copernicus.eu/global/content/algorithm-wb-300m-global>, accessed July 2019). Therefore, since it represents the up to date approach for mapping flooded areas, we used H/V components as input features to the algorithm and combined it with SIs relying on the fact that OWA operators can flexibly manage evidence scores derived from heterogeneous inputs.

Standardized colorimetric transformation from RGB (SWIR2, NIR, RED respectively) to HSV colour space (Smith, 1978) is expressed by the following equations:

$$V = \max(R, G, B)$$

$$S = (V - \min(R, G, B))/V$$

$$H = 0 \text{ if } V = \min(R, G, B)$$

$$H = (60^\circ \times (G - B)/(V - \min(R, G, B)) + 360^\circ) \bmod 360^\circ \text{ if } V = R$$

$$H = (60^\circ \times (B - R)/(V - \min(R, G, B)) + 120^\circ) \text{ if } V = G$$

$$H = (60^\circ \times (R - G)/(V - \min(R, G, B)) + 240^\circ) \text{ if } V = B$$

Hue (H), defined as an angle in a range 0°–360°, represents perceived colour spectral composition, i.e. the visual perceptual property corresponding to the categories called red, green, blue and others. Saturation (S) depends on the proportion of reflectance of the dominant wavelength across the spectrum and shows how far a colour is from a grey of equal brightness. Value (V) is defined as colour brightness. In the human recognition process H plays the key role.

3. Methods

The work consisted of two phases: algorithm development (set up and validation) and assessment of the mapping products (Fig. 3). In the first phase, we defined the algorithm scheme and specific elements and we assessed algorithm performance. In the second phase, we quantified the accuracy of flooded area maps obtained by applying the algorithm over independent evaluation sites and without tuning over the specific characteristics of the sites. This phase aims to investigate the exportability of the algorithm when applied to different geographic regions and/or environmental conditions thus simulating operational semi-automated implementation.

3.1. Algorithm set up

The proposed algorithm relies on a multi-criteria approach that aggregates the information brought by multiple input features (SIs and H/V) into a synthetic global evidence degree. Each feature could be used as source for deriving evidence of water/flooded conditions

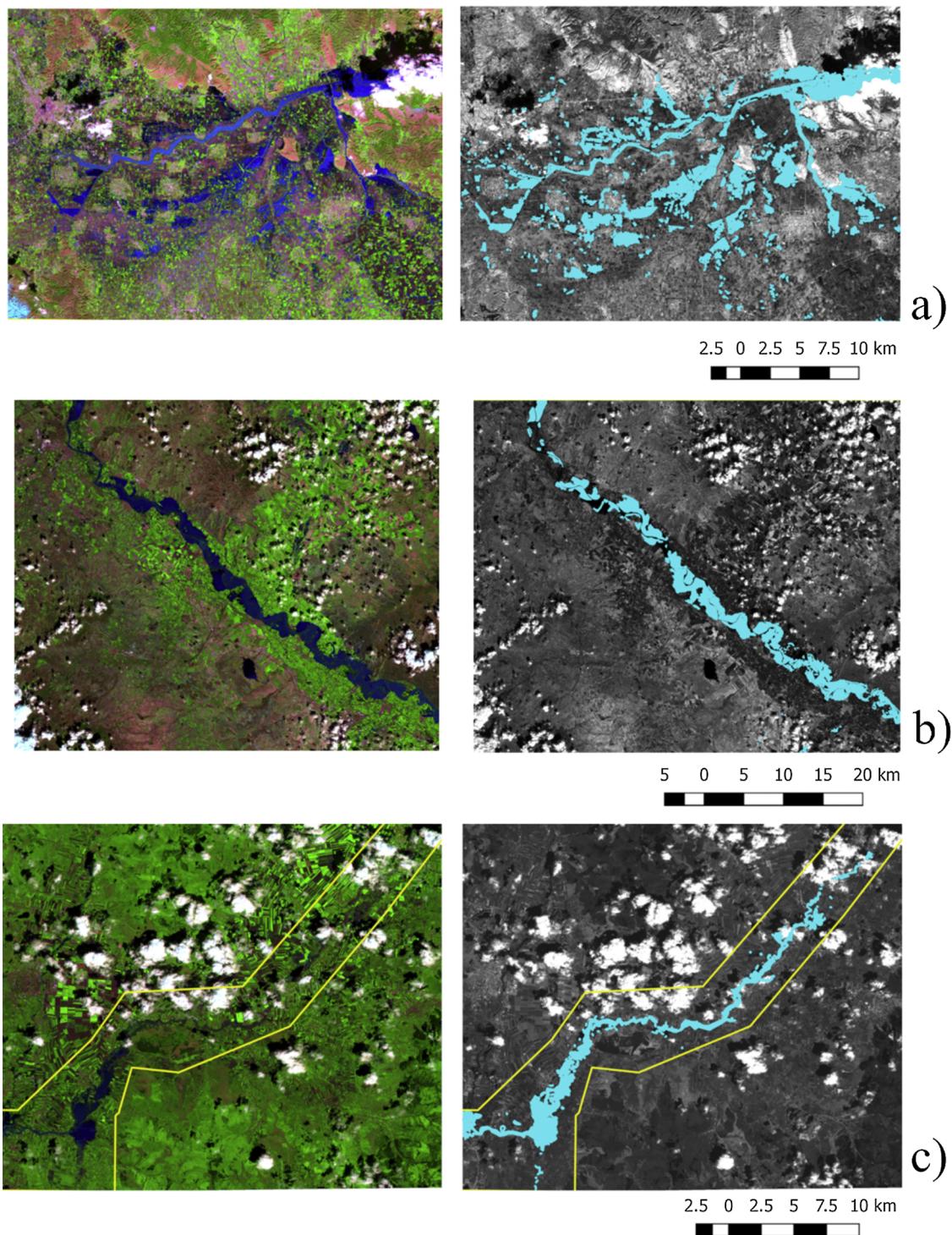


Fig. 2. S2 false colour composites (R = SWIR/band12, G = NIR/band8, B = RED/band4) (left) and grey scale S2 SWIR band 12 (right) over Evaluation sites in Greece (a), Spain (b) and Romania (c). Yellow polygons show the area of interest (AOI) for accuracy assessment. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

Table 2

Number of pixels used for algorithm set up and validation for Cal/Val sites (w/nw stand for water/not water).

Name	Area	Dates	Training (# pixel)	Testing (# pixel) (w/nw)
Cal/Val_1	Emilia (site 1)	S2A 13/12/2017	144689	866 (157/709)
Cal/Val_2	Po valley (site 2)	S2A 14/05/2017	51014	1131 (192/939)
Cal/Val_3	Rice fields (site 3)	S2A 22/04/2016	42015	953 (170/853)

Table 3

Dates of S2 images and EMS mapping for the Evaluation sites. CSK is Cosmo-SkyMed.

Name	Location	Imaging date	Reference date	Resolution reference mapping
Eval_1	Greece	S2B 28/02/2018	CSK 24/02/2018	3 m
Eval_2	Spain	S2A 14/04/2018	CSK 13/04/2018	3 m
Eval_3	Romania	S2A 03/07/2018	CSK 03/07/2018	5 m

(partial evidence of flooding); aggregation reinforces the evidence on one side by exploiting the convergence of partial evidence from multiple possibly redundant sources, and, on the other side, by compensating the inconsistency/conflict of evidence from multiple possibly complementary sources. This allows to strengthen the plausibility of the presence of water and reduces confusion of surfaces with similar characteristics (Stroppiana et al., 2012).

Aggregation is carried out with Ordered Weighted Averaging operators (OWAs): a parameterized family of soft-mean-like aggregation operators. The approach is applied independently to each pixel of the input EO image with the following steps:

- 1 Definition of the soft constraints for each input feature from training data and their application;
- 2 Definition/selection of OWAs to be used for the soft integration;
- 3 Pixel-based soft integration to compute the synthetic global evidence degree;
- 4 Segmentation of the synthetic global evidence for deriving the maps of *standing water/flooded areas* according to pessimistic or optimistic attitude of the decision maker.

In the last step, pessimistic refers to decision maker's attitude that wants to map the maximum extent of the phenomenon to minimise the chance of underestimating. On the contrary, the optimistic attitude wants to minimise the chance of overestimating, reducing false alarms.

3.1.1. Soft constraints on spectral feature

Soft constraints can be defined with different approaches according to the available expertise and training data (Carrara et al., 2008; Stroppiana et al., 2009); in this work, they were defined from training data available over the three Cal/Val sites to compute partial evidence scores of pixels having given spectral feature values. For each feature, distinct soft constraints, i.e. Membership Degree (MD_{SI} , MD_H , and MD_V , respectively), were defined from frequency histograms of training pixels of the three Cal/Val study sites as shown in Fig. 4 (for details see Supplementary Materials). Evidence scores derived from the SIs were computed as $MD_{SI}(SI_p)$, where SI is the spectral index value of pixel p. Evidence scores derived from H/V components were computed by the fuzzy Cartesian product of MD_H and MD_V (i.e., fuzzy relation): $\min(MD_H(H_p), MD_V(V_p))$ where H_p and V_p are the Hue and Value of pixel p.

3.1.2. Soft integration

Partial evidence scores of SIs and the H/V were integrated by applying OWAs: non-linear operators that, given a weighting vector,

$W = [w_1, \dots, w_n]$ so that the weights sum to 1, it applies a weighted average of the arguments rearranged from the highest to the smallest. The key step of OWAs is reordering, meaning that a specific weight w_i is not univocally associated to the specific i_{th} input argument but rather it is associated to the i_{th} position of the reordered arguments (Yager, 1988; Xu and Da, 2003). For example, the weight w_1 will determine the contribution of the greatest input argument to the global evidence degree computed by the OWA which could be the partial evidence score from different features for distinct pixels of the EO image.

The weighting vector W determines the semantics of the integration that can model any decision attitude varying smoothly between an optimistic and a pessimistic one. For example, a weighting vector W of the OWA operator with $w_8 = 1$ considers only the contribution of the minimum argument, i.e. the minimum partial evidence score; hence, in our context, implementing an optimistic attitude means to compute the minimum total flooded area (AND aggregation). Conversely, by setting $w_1 = 1$ the maximum partial evidence will determine the largest flooded area, thus in our context modelling the pessimistic case (OR aggregation). Intermediate cases, in which all or most components of W are not null, model soft "democratic" integrations. An example is the average of the partial evidence degrees, or any operator that exploits only a portion of the input arguments (e.g. "at least three of" which could be defined by a weighting vector $W = [0.5, 0.25, 0.25, 0, 0, 0, 0, 0]$).

In our experiments, we compared the following five OWA weight vectors:

$W_{AND} = [0, 0, 0, 0, 0, 0, 0, 1]$ (completely optimistic global evidence map minimizing watered areas)

$W_{Almost\ AND} = [0, 0, 0, 0, 0, 0, 0.5, 0.5]$ (partially optimistic global evidence map almost minimizing watered areas)

$W_{average} = [0.125, 0.125, 0.125, 0.125, 0.125, 0.125, 0.125, 0.125]$ (fully democratic and neutral global evidence map of watered areas)

$W_{Almost\ OR} = [0.5, 0.5, 0, 0, 0, 0, 0, 0]$ (partially pessimistic global evidence map almost maximizing watered areas)

$W_{OR} = [1, 0, 0, 0, 0, 0, 0, 0]$ (fully pessimistic global evidence map maximizing watered areas)

3.1.3. Algorithm validation

The performance of the algorithm was assessed by comparison with reference dataset for the Cal/Val sites in four major steps:

- 1 Sampling design: stratified random sampling of about 1200 points, i.e. single pixels in the image (10 m × 10 m), with strata for *not water* and *water* classes derived by thresholding AWEI, spectral

Table 4

The spectral indices (SIs).

Index	Category	Reference
AWEI = C1 (GREEN – SWIR1) – (C2 * NIR + C3 * SWIR2)	Water	Feyisa et al., 2014
AWEI _{SH} = BLUE + D1 * GREEN – D2 * (NIR + SWIR1) – D3 * SWIR2	Water	Feyisa et al., 2014
MNDWI = (GREEN – SWIR1) / (GREEN + SWIR1)	Water	Xu, 2006
NDWI = (GREEN – NIR) / (GREEN + NIR)	Water	McFeeters, 1996
NDFI = (RED – SWIR2) / (RED + SWIR2)	Flooding	Boschetti et al., 2014
SAVI = (1 + L) * (NIR – RED) / (NIR + RED + L)	Vegetation	Huete, 1988
WRI = (GREEN + RED) / (NIR + SWIR1)	Water	Shen and Li, 2010

Where $C1 = 4$, $C2 = 0.25$, $C3 = 2.75$, $D1 = 2.5$, $D2 = 1.5$, $D3 = 0.25$, $L = 0.5$ and S2 MSI bands are BLUE = band2 (490 nm), GREEN = band3 (560 nm), RED = band4 (665 nm), NIR = band8 (842 nm), SWIR1 = band11 (1610 nm), SWIR2 = band12 (2190 nm).

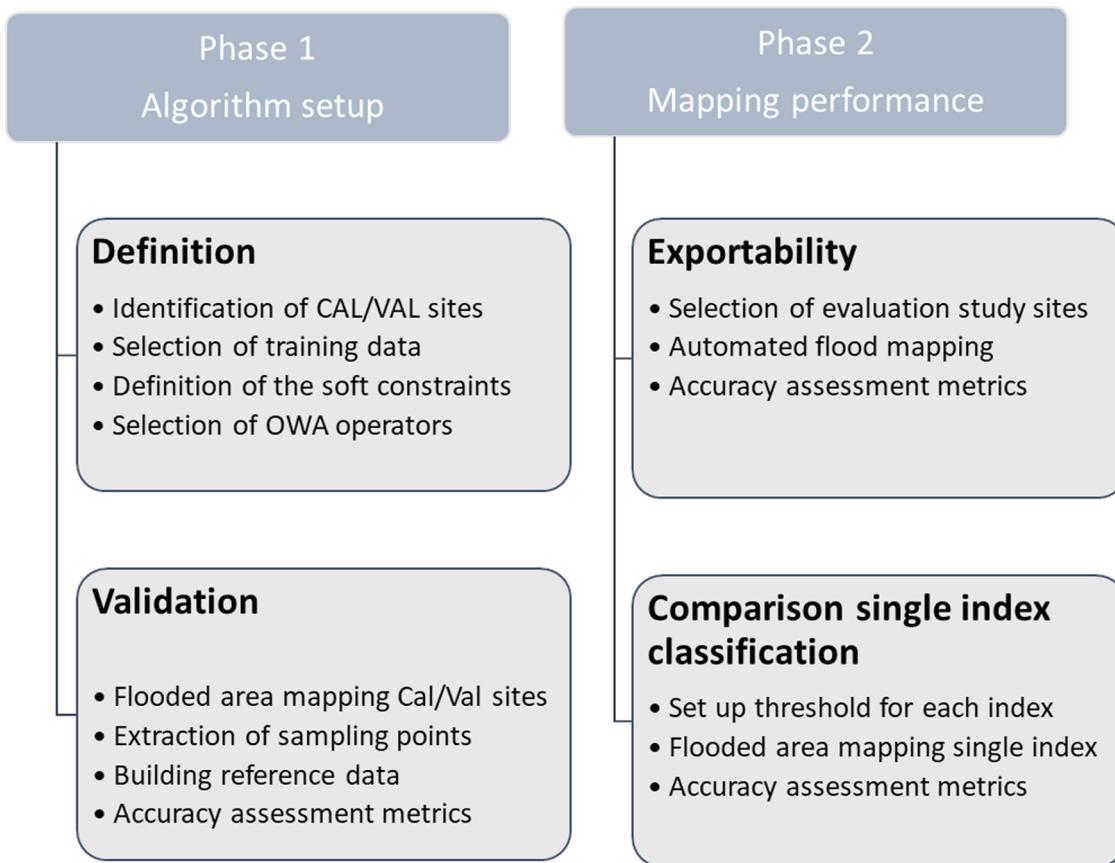


Fig. 3. Algorithm flowchart.

- index that provided the best visual discrimination;
- 2 Supervised labelling: photo-interpretation of the points within each stratum;
 - 3 Confusion matrix: pixel by pixel comparison between reference and maps to compute True Positives (TP), True Negatives (TN), False Positives (FP) and False Negatives (FN);
 - 4 Computation of accuracy metrics from the confusion matrix: commission ($CE = FP/(FP + TP)$) and omission ($OE = FN/(FN + TP)$) errors and F-score according to the following equation:

$$F - score = 2 \times TP / (2 TP + FP + FN)$$

For each OWA operator, CE, OE and F-score were computed adopting a set of thresholds defined in the domain [0, 1] of the integrated global evidence scores with a step of 0.1. Each value specifies the segmentation threshold for identifying *water/flooded* (above the threshold) and *not water/not flooded* (below the threshold) pixels in the synthetic flooded map.

3.2. Assessment of flooding mapping

The proposed algorithm was automatically applied (i.e. with no further tuning and/or changes) to the Evaluation sites to perform a more extensive and robust validation and to provide insights on operational implementation. For each site, we computed accuracy metrics (CE, OE and F-score) through pixel by pixel comparison of algorithm outputs and reference data, i.e. Copernicus EMS perimeters.

Results from the soft integration were also compared with water/flood maps obtained by segmentation of each single spectral index with crisp threshold values proposed in the literature (Table 5). Since threshold values are not optimal for any image and manual adjustment is often necessary (Chen et al., 2013; Hui et al., 2008; Ji et al., 2009),

we also assessed the sensitivity of accuracy metrics to changes of the selected crisp value. For the HSV component, we applied equations proposed in Pekel et al. (2016).

4. Results and discussion

4.1. Partial evidence degrees

The membership functions shown in Fig. 4 were applied pixel by pixel to each single input feature to obtain partial evidence degree of *water/flooding*; examples over zoom areas ($20\text{ km} \times 20\text{ km}$) are depicted in Fig. 5 for the three Cal/Val sites. The higher the degree is (i.e. closer to 1), the more likely the pixel is covered by water and/or flooded. As expected, the degree varies with the considered input feature; SAVI (Fig. 5) provides the lowest partial degree over all types of water surfaces: natural flooding (Cal/Val_1), river (Cal/Val_2) and rice fields (Cal/Val_3). Flooded rice fields have the most variable behaviour/degree in relation to the high rate of different field conditions as observed during *in situ* surveys (See S1) and confirmed by the Fig. 4 where, for most of the indices, the distribution overlaps histograms of the *not water* classes. Yet Cal/Val_3 was selected with the specific purpose to train the algorithm on diverse spectral conditions of standing water. Over this site, H/V and AWEI provide the highest degrees, which, however, might include false detections.

NDFI and mNDWI provide similar results since they are both derived from VIS/SWIR wavelengths, (Boschetti et al., 2014); for example, over Cal/Val_2 the main river is not fully identified (i.e. partial evidence degree < 1). WRI, SAVI and NDWI present very low degree of partial evidence, i.e. not clear identification of the river in the Cal/Val_1 (See S3).

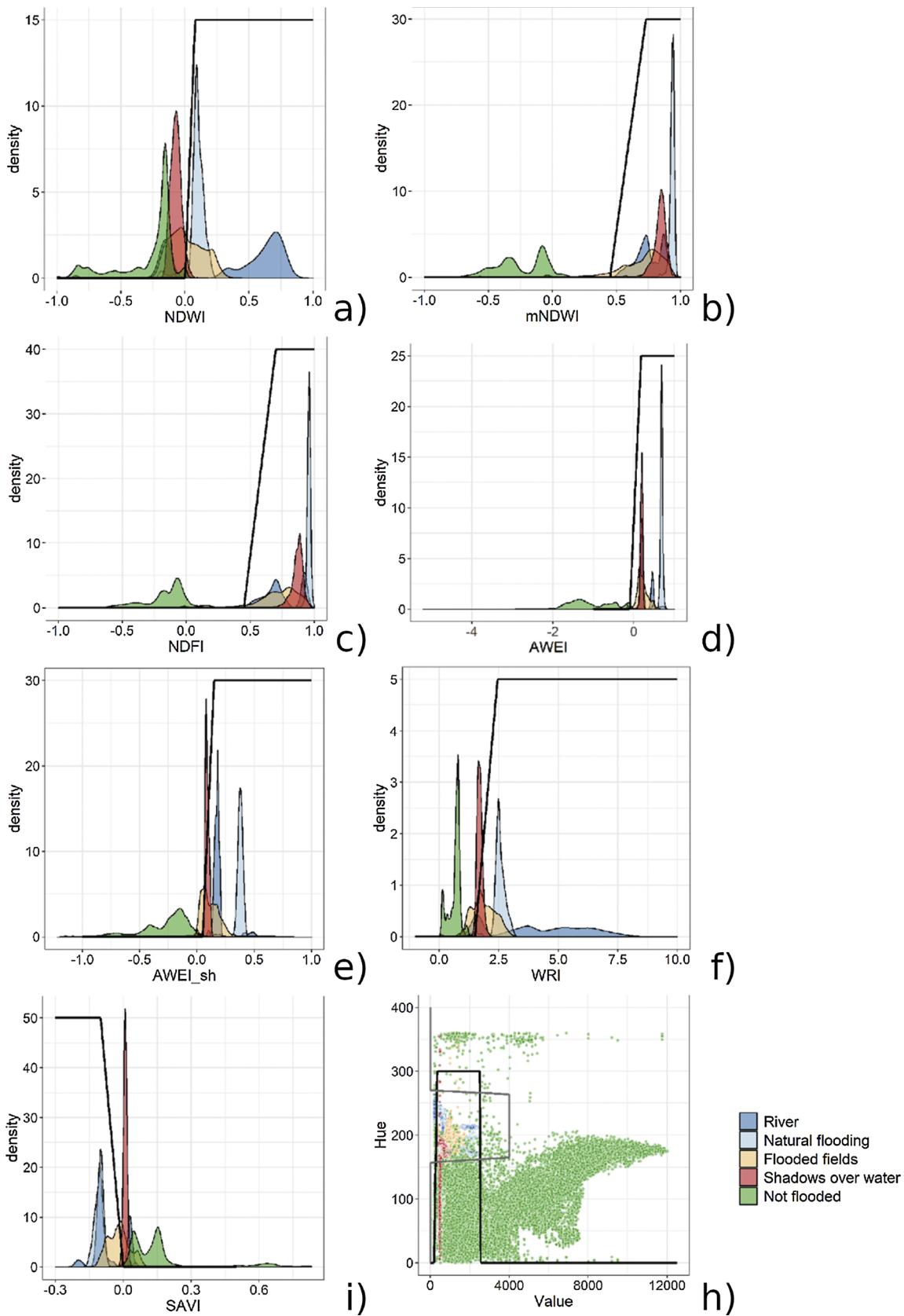


Fig. 4. The soft constraints for the input SIs (a to g) and for Hue (H) and Value (V) components (h) defined on the training pixels over the Cal/Val sites (pooled together). Flooded areas include standing water, rice fields and rivers/lakes; not flooded areas include dry soil, vegetated land, agricultural fields, urban areas, clouds and cloud shadows.

Table 5
Segmentation crisp thresholds on the spectral indices.

Threshold	Reference
AWEI > 0	Feyisa et al., 2014
AWEI _{SH} > 0	Feyisa et al., 2014
mNDWI > 0	Xu, 2006
NDWI > 0	McFeeters, 1996
NDFI > 0.32	Ranghetti et al., 2016
SAVI < -0.25 or SAVI < -0.27	Weinrit et al., 2018
WRI > 1	Acharya et al., 2017

4.2. Global evidence degree

Fig. 6 shows the output maps of soft integrated global evidence degree [0–1] for each Cal/Val site with different OWA compared to false colour RGB images.

The greatest proportion of water surface is obtained with the “OR” ($W = [1, 0, 0, \dots]$) and “Almost_OR” ($W = [0.5, 0.5, 0.0, \dots]$) operators that assign the highest weights to the input features with the highest partial evidence; the opposite for “AND” and “Almost_AND” operators. The “Average” delivers the highest range of variability of the global evidence degree across sites likely due to different water characteristics;

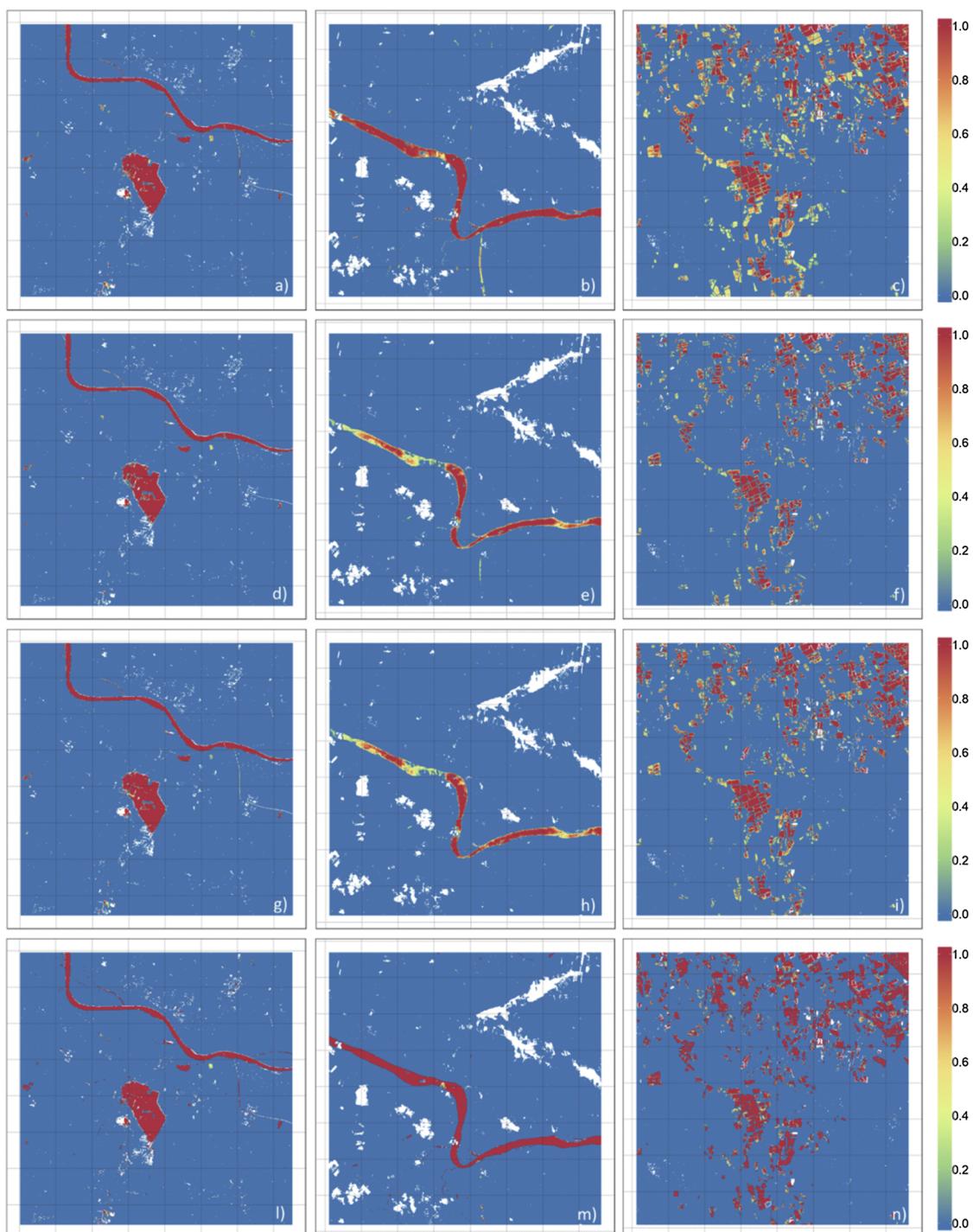


Fig. 5. The partial evidence degrees for AWEI (a, b, c), mNDWI (d, e, f), NDFI (g, h, i), and the H and V components (l, m, n) for zoom areas of 20 km × 20 km of Cal/Val_1 (left), Cal/Val_2 (middle) and Cal/Val_3 (right). Cloud masked areas are white and the degree of partial evidence ranges in [0, 1].

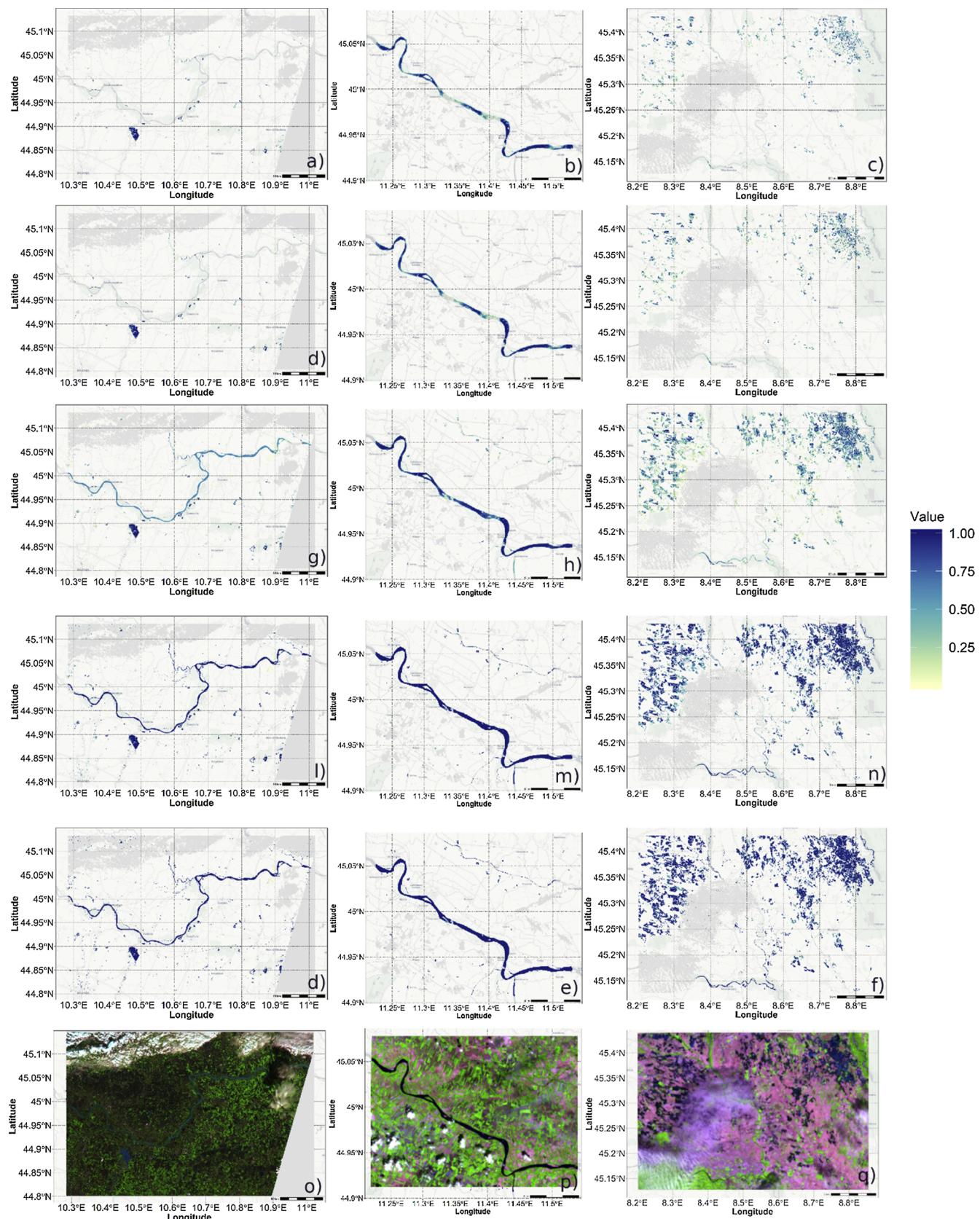


Fig. 6. The global evidence degree over Cal/Val_1 (left), Cal/Val_2 (middle) and Cal/Val_3 (right) obtained with OWA weight vector “AND” (a, b, c), “Almost_AND” (d, e, f), “Average” (g, h, i), “Almost_OR” (l, m, n) and “OR” (o, p, q). In the bottom row (r, s, t) S2 images as false colour composites (R = SWIR/band12, G = NIR/band8, B = RED/band4). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

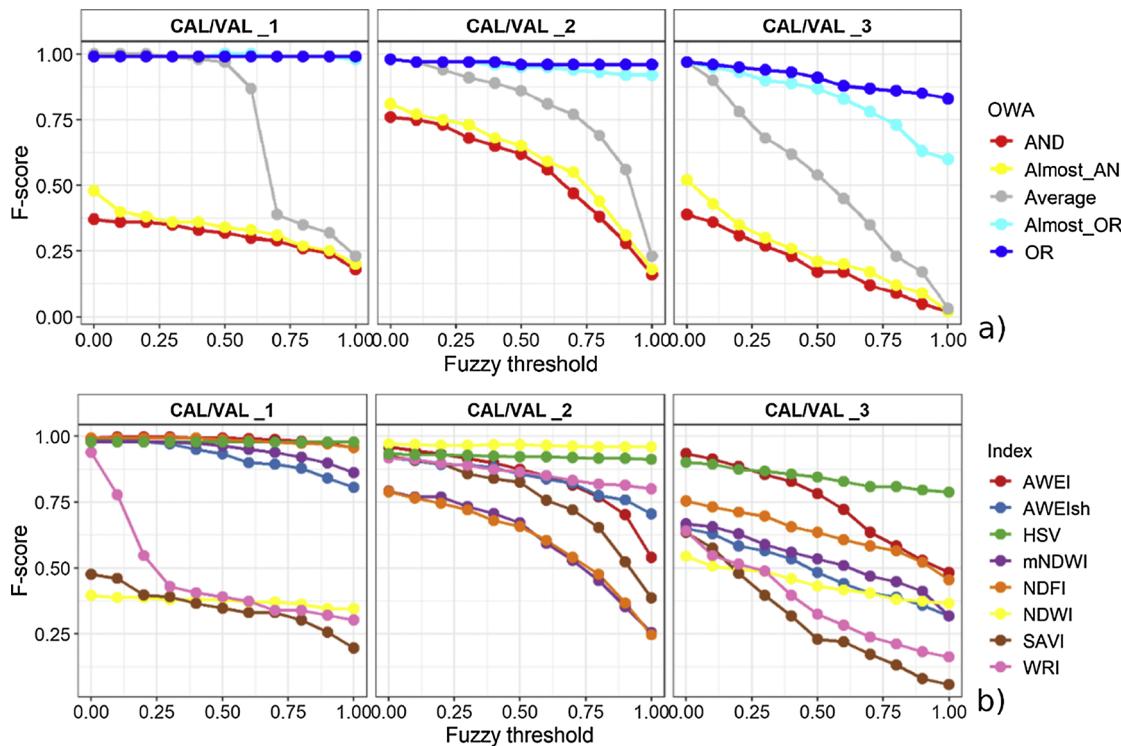


Fig. 7. F-score as a function of the threshold for the three Cal/Val sites for OWA operators (top) and single input features (bottom).

this is particularly evident in the Cal/Val_1 and Cal/Val_3. In the former, degrees close to 1 are assigned to flooded areas while intermediate values are assigned to river pixels. In the latter, rice fields cover the entire range of values [0, 1] probably due to different conditions and timing of agronomic flooding as also highlighted by field photos (See S1).

4.3. Algorithm validation

Over Cal/Val sites, test pixels were used for estimating accuracy of maps obtained by segmenting global as well as partial evidence degrees; segmentation reduces continuous information in [0, 1] to binary two-class maps (*water/no water*). Accuracy metrics, computed for each OWA operator and input feature (SIs and H/V components) for varying thresholds, quantify the improvement brought by integrating redundant/complementary information. F-score is shown in Fig. 7.

Among OWAs, the best performance was obtained with “OR” and “Almost OR” weight vectors (F-score > 95%) suggesting that input features are not only redundant, but instead complementary in characterizing *water/flooded* areas. F-score of the best performing “OR” operator decreases below 90% when threshold > 0.5 due to the increase of False Negative detections over more complex conditions of water surfaces in Cal/Val_3; indeed, rice flooding conditions can be spatially heterogeneous due to difference in water depth, flooding timing and duration, emergence of rice plants (Ranghetti et al., 2018).

For Cal/Val_1, NDFI and H/V have the greatest accuracy with F-score > 90%; comparable performance can be achieved with other indices but only when threshold < 0.5. For Cal/Val_2, NDWI and H/V perform best with F-score is > 90% even for greater thresholds. Accuracy drastically decreases over Cal/Val_3 with F-score below 75% when threshold > 0.1 for all features except AWEI and H/V; even with the best performing features, accuracy is lower than values obtained with “OR” operator confirming that OWAs can dynamically handle variable conditions. Moreover, results of single features are clearly not consistent across sites: the same index does not guarantee robust performance meaning that operational use over large and heterogeneous

areas is straightforward.

4.4. Flooding products

Maps of integrated global evidence degree over Evaluation sites (Greece, Spain and Romania) were obtained with all OWA operators and segmented with variable fuzzy threshold to deliver binary maps of *flooded/not flooded* areas. Pixel by pixel comparison with Copernicus EMS reference data provided accuracy metrics shown in Fig. 8.

The “OR” and “Almost_OR” OWAs show the most consistent performance in the three sites as represented by the greatest F-score and the flat trend line for different thresholds, meaning that “OR-like” operators produced a good balance between CE and OE. On the contrary, “AND” and “Almost AND” perform differently across sites with the lowest commission and the highest omission errors. These operators represent an optimistic attitude showing areas of highest probability of being flooded thus reducing the risk to detect false positive (i.e. false alarms).

Results confirm outcome from validation with the best performing “OR-Like” OWA operators achieving more than satisfactory F-scores (> 75%) yet lower compared to accuracy levels of algorithm validation on Cal/Val sites (F-score ~ 1). Over Evaluation sites, the algorithm was applied automatically with no changes compared to the set-up phase and lower accuracy is due to different environmental/topographic conditions (e.g. narrow river valley or wide flood plains) and water characteristics and permanence (e.g. shallow waters). Moreover, over large areas a decrease of accuracy is expected for the coexistence of complex and highly variable surface conditions (water presence, cloud shadows, built up areas, low-albedo non-water surfaces) (Wang et al., 2018).

Further reasons could be pointed out for the lower performance. Above all, EMS reference polygons are derived from high/very high resolution EO data (pixel size < 5 m) and by supervised classification to achieve highest accuracy while fuzzy flooded areas are derived from S2 with 20-meter spatial resolution with a semi-automatic algorithm. Moreover, EMS maps do not include stable water surfaces (e.g. lakes

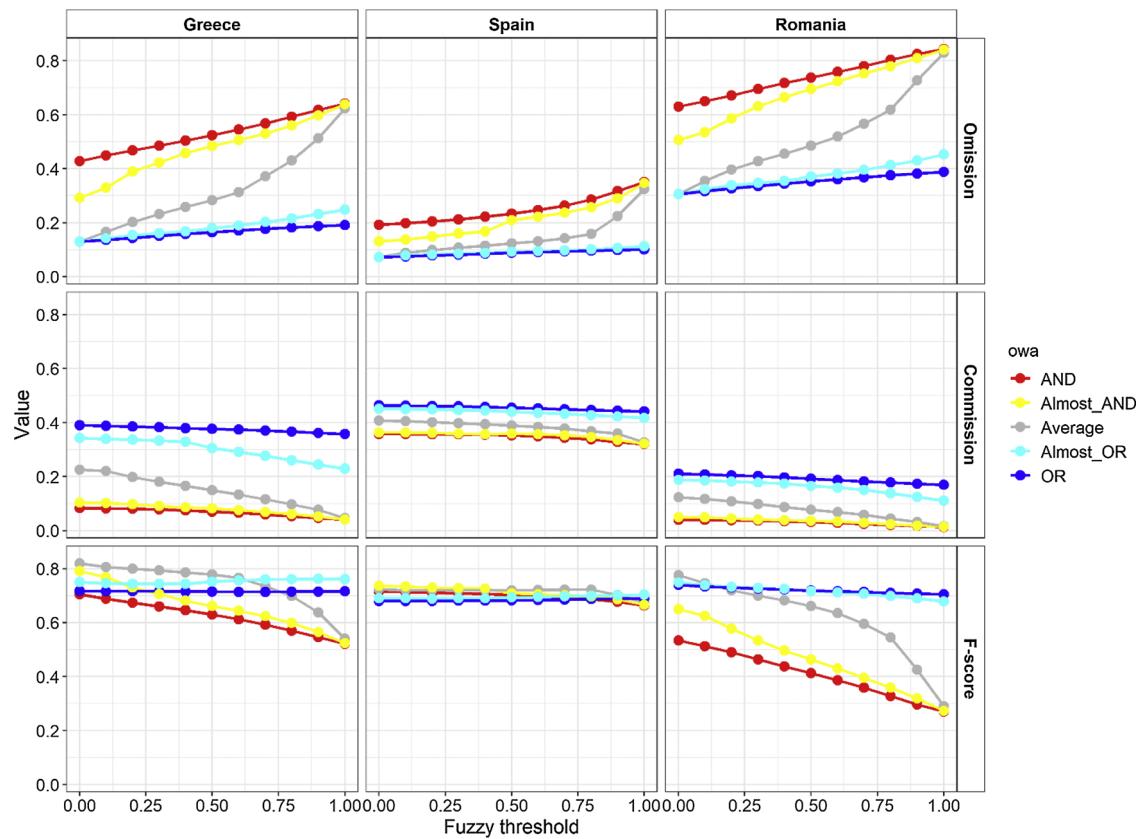


Fig. 8. Accuracy metrics for the Evaluation sites for OWA operators used for the integration of the degree of partial evidences.

and rivers) as flooded areas which are mapped as flooded in the fuzzy maps instead. Finally, (apparent) commission errors could be due to time lag between reference data and S2 imagery acquisition dates (Table 3) that, in fact, is larger over Eval_1 and Eval_2.

To provide an operational-like example, we chose the “Almost_OR” (pessimistic attitude) and “Almost_AND” (optimistic attitude) operators to segment the synthetic global evidence degrees with $f > 0.5$ to obtain maps of *flooded/not flooded* areas for the three Evaluation sites. Fig. 9 shows accuracy maps obtained by per-pixel comparison with reference EMS data, with the following categories: correctly classified *flooded* (blue), correctly classified *not flooded* (grey), commission (orange) and omission (green) errors. The “Almost_AND” operator (left column in Fig. 9) provides different rates of commission/omission (9/48%, 37/21% and 4/69% in panels a, c, and e, respectively) generally picturing the lowest rates of false detections/commission errors (i.e. optimistic attitude) at the expense of a greater omission (green areas) such as for Eval_1 and Eval_3. Over Eval_2, overall accuracy is more stable as measured by F-score (Fig. 8), which is in the range 69–70% and does not significantly change with the OWA and threshold. Over this site, commission errors are greater (32–46%) and mainly due to natural/artificial permanent water surfaces which are not included in EMS mapping. Over Eval_3, F-score is 46% and 71% for “Almost_AND” and “Almost_OR”, respectively due to a significant decrease of the omission error from 69% (Fig. 9e) to 37% (Fig. 9f) and the greater correctly classified flooded areas (blue regions), which can be observed in the lower left portion of Fig. 9f; over this site fuzzy maps show lower performance in the detection of flooded areas along the river course probably due to the lower spatial resolution of the S2 imagery.

Finally, to further analyse results obtained with traditional approaches, we computed accuracy metrics of flooded area maps achieved by segmentation of single input features with thresholds proposed in the literature; results are illustrated in Table 6.

Accuracy metrics are consistent with results from soft integration

with F-score values for Eval_1 (Greece) greater, for most of indices, compared to Eval_2 (Spain) and Eval_3 (Romania). Eval_2 confirms to be the site with highest commission errors for the reasons discussed above.

Although accuracy levels from crisp thresholds are satisfactory, results are not consistent across sites (as already pointed out in Fig. 7). In fact, features ranking based on F-score changes with site, as highlighted in the table by coloured text in the table, thus posing issues for operational implementation over large areas. For example, SAVI is the best for Eval_2 (F-score = 71.1%) and worst for Eval_1 (F-score = 67.3%) and Eval_3 (F-score = 36.1%). More consistent results are obtained with the H/V components with F-score > 70% for all sites although it is not the best performing ranking 4th, 2nd and 5th for Eval_1, Eval_2 and Eval_3, respectively. Note, moreover, that at the time of writing no equations for the HSV approach were specifically calibrated for S2 data and for the comparison we applied equations proposed for Landsat data as in Pekel et al. (2016).

Hence, these results highlight and confirm the need of an approach for combining different inputs since a-priori selection of a single index/feature could not guarantee a stable accuracy across sites.

We further tested the sensitivity of F-score to changes of the crisp threshold around the literature value and found different results according to the investigated SIs (Fig. 10). As suggested by Feyisa et al. (2014), AWEI shows a fairly stable optimal threshold value compared to mNDWI. On the contrary AWEI_{SH} appeared to be sensitive to changes of the threshold value but with different rates over the sites. NDFI and mNDWI, both derived from VIS/SWIR wavebands (Boschetti et al., 2014), show similar trends with a slightly better and flatter accuracy trend of the first. NDWI shows a significant reduction of F-score over Eval_3 (Romania) if threshold > 0; performance comparable to the other indices and sites could be achieved only with threshold set to lower values compared to McFeeters (1996). Similarly, SAVI presents low performance for Eval_3 and likely better results for threshold values

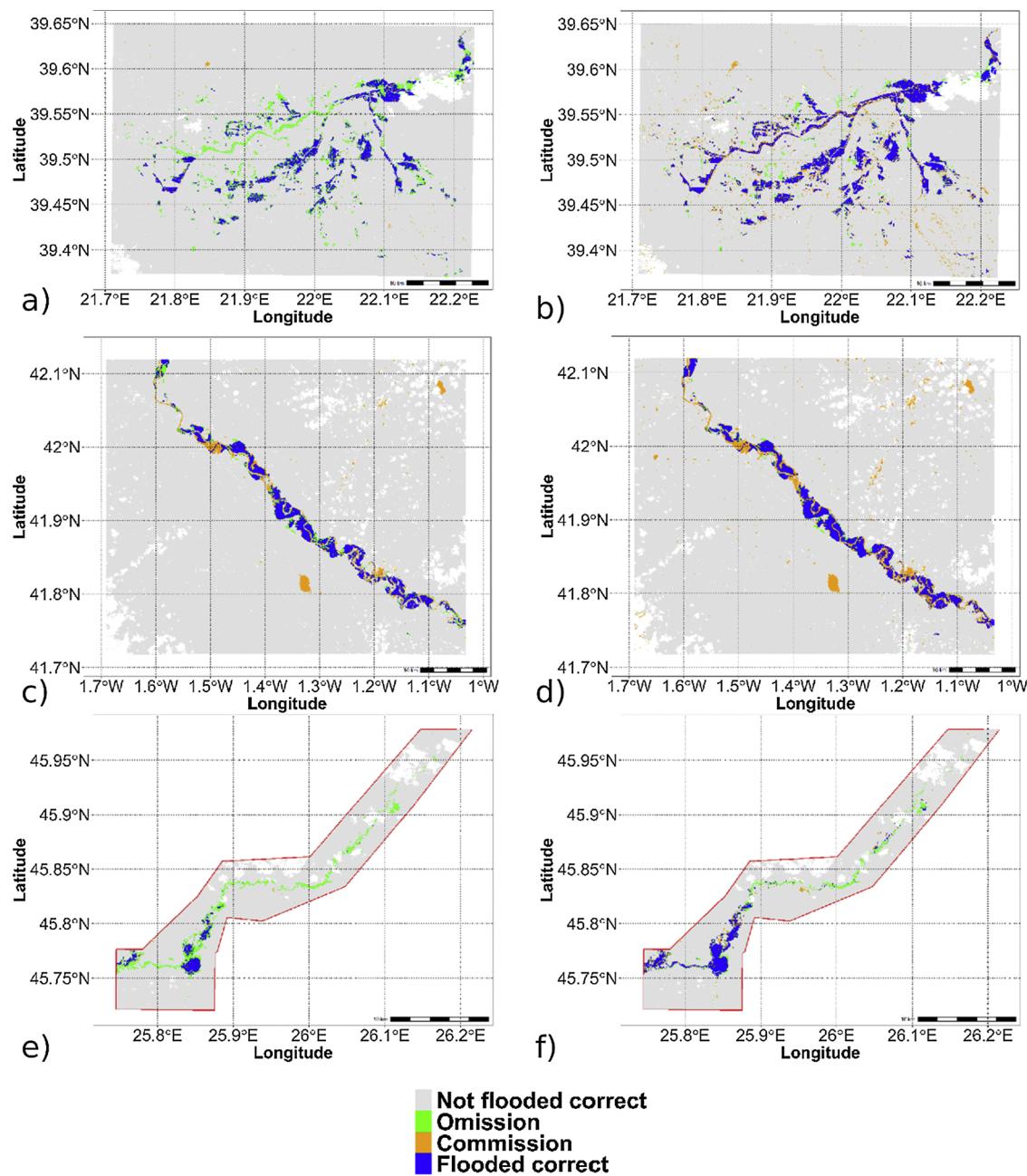


Fig. 9. The accuracy map of flooded areas over the three Evaluation sites (Greece: a, b; Spain: c, d; Romania: e, f) obtained with OWA weight vector “Almost_AND” (left column) and “Almost_OR” (right column) with threshold = 0.5. The map shows correctly classified flooded (blue), correctly classified not flooded (light grey), commission (orange) and omission (green) errors. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

Table 6

Accuracy metrics computed by applying crisp thresholds to the single spectral index for the three Evaluation sites. Best and worst F-score values are highlighted in bold and italics, respectively.

Eval_1-Greece			Eval_2-Spain			Eval_3-Romania		
	OE	CE		OE	CE		OE	CE
AWEI	19.4%	26.0%	77.2%	8.8%	44.3%	69.2%	37.4%	18.5%
AWEI _{SH}	20.5%	28.2%	75.5%	9.8%	44.5%	68.7%	38.5%	18.0%
mNDWI	13.4%	39.0%	71.6%	6.5%	50.1%	65.1%	29.7%	29.3%
NDWI	37.7%	19.1%	70.4%	14.9%	39.7%	70.6%	61.3%	9.4%
WRI	17.1%	38.2%	70.8%	9.1%	46.5%	67.3%	34.5%	22.0%
SAVI	44.7%	14.0%	<i>67.3%</i>	20.4%	35.8%	71.1%	77.6%	7.2%
NDFI	15.2%	35.7%	73.1%	7.3%	46.8%	67.6%	31.7%	25.6%
HSV	20.6%	34.4%	71.9%	10.7%	41.3%	70.8%	39.9%	15.2%

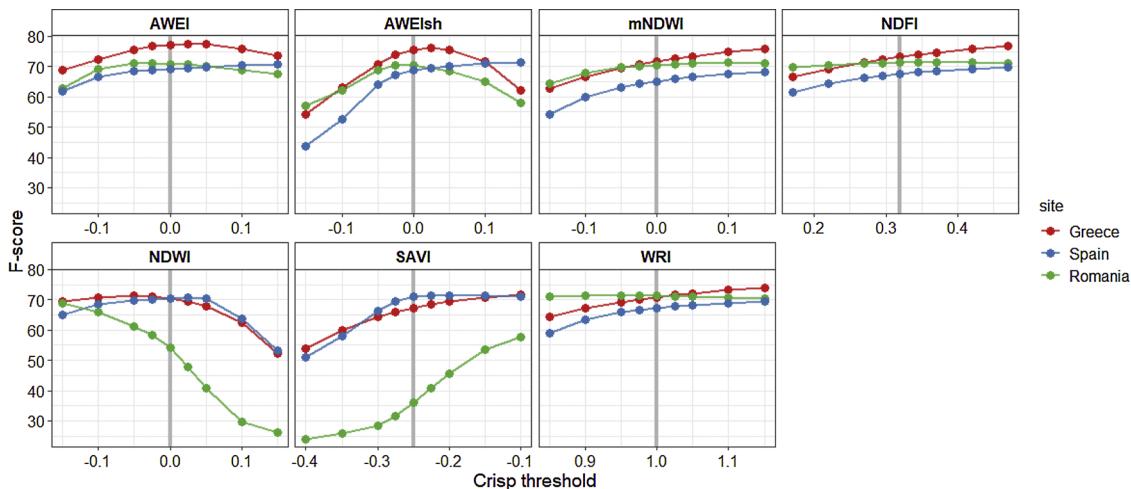


Fig. 10. F-score for the three Evaluation sites (Greece, Spain and Romania respectively Eval_1, Eval_2, Eval_3) and the indices for varying values of the crisp threshold. The range of variability is centred at the value proposed in the literature highlighted by the grey vertical line.

greater than $SAVI > 0$ as proposed in the literature. WRI shows low sensitivity to threshold values greater than the literature one for all sites. Interestingly no change occurs for Eval_3 and WRI appears to be the best performing and more stable index compared to very variable results that it shows for the other sites. In general, these results confirm that threshold selection is a crucial choice to extract water pixels (Yang et al., 2018).

5. Conclusions and perspectives

In this work we developed an algorithm to automatic map flooded areas from multispectral S2 MSI images based on fuzzy set theory. Since the use of multisource data is recognized as the way to achieve improved global and regional water mapping (Huang et al., 2018), we propose an approach to integrate multiple spectral features. Rather than making an a-priori selection of the best (or a few) water indicator(s), we exploited redundancy (convergence of evidence) or complementarity provided by multiple spectral features. Integration is carried out with OWA operators that can flexibly aggregate inputs of different nature: we included traditional spectral indices and the H/V components of the HSV transformation proposed by Pekel et al. (2014), (2016), which represent an up to date approach for water mapping and they are exploited for the generation of Copernicus operational products (<https://land.copernicus.eu/global/products/WB>, accessed July 2019). In fact, OWAs can integrate contributing factors which can be semantically different (spectral indices, H/V, other types of data such as surface elevation gradient, single spectral bands, etc.) by normalizing their distinct contributions by computing evidence score through the application of soft constraints (Bordogna et al., 2007).

The algorithm was trained and validated over three study sites in Italy where water pixels represented a wide range of conditions from shallow to deep water, from clean to turbid and from natural hazard to managed irrigation flooding. Results showed that performance is comparable or better than traditional approaches based on the segmentation of single input features and with more consistent and robust accuracy. By integrating multiple inputs, we limited subjectivity and sensitivity involved in an a-priori selection of the most suitable index and crisp thresholds (Yang et al., 2018).

A key step for the implementation of our approach is the definition of the soft constraints for each input feature, which could be done by integrating expert knowledge (as in this work) as well as in a completely unsupervised manner (Stroppiana et al., 2009) when expert knowledge is not available and/or it requires time/effort to be formalized (Sakamoto et al., 2018). The proposed approach could be

exploited in automated processing for near-real time mapping and monitoring and we showed how, testing over independent and large sites selected from Copernicus EMS operational mapping, it provided satisfactory accuracy. Indeed, over the three Evaluation sites, the algorithm was applied without local tuning over site characteristics: soft constraints defined for the calibration sites in Northern Italy were used without any modification. Over Evaluation sites, lower accuracy was expected due to the complexity of each scene where different surface conditions coexist; however, these accuracy levels could be considered a lower bound of the performance of the algorithm that could be improved by recalibrating the soft constraints and weight vectors.

OWA operators certainly offer an appealing solution for the integration of multiple features and for formalizing distinct expert attitudes when aggregating partial to global evidence (Bordogna et al., 2007). Among the OWA analysed, OR-like operators are more stable providing accuracy levels that are less sensitive to the threshold applied to segment the integrated global evidence degree. On the contrary, AND-like operators show decreasing F-score for increasing thresholds due to the greater rate of omission errors. In fact, AND-like operators implement an optimistic approach by retaining only high probability flooded/water areas. Different OWA operators able to retain only high/low probability water/flood surfaces could be exploited in a framework of multistep contextual classification approaches, such as region growing algorithms (Stroppiana et al., 2012). From a user perspective, OWA operators allow to represent different semantics and to model different attitude of the decision making process smoothly between an optimistic and a pessimistic one; this sort of thematic product could be an advantage in decision support system framework (Ceresi et al., 2018).

In future work, the algorithm will be tested to flexibly integrate feature derived from data acquired by other satellite missions with similar spectral and geometrical characteristics (e.g. Landsat data) as well as other remotely sensed data, such as SAR imagery (e.g. Sentinel-1). We also consider that the soft integrated global evidence degree ranging in [0, 1] could bring additional information on flooding conditions such as the likelihood of being flooded and/or partially flooded pixels; this information could be useful in decision making or as input for modelling. Moreover, the algorithm could be tested with input features taking into account the temporal dimension (pre – post event change detection) to further enhance flooded area mapping accuracy. More sophisticated machine learning approaches to generate flooded area map and relying on automatic training process based on selected high plausible pixels will also be tested.

Acknowledgements

This work has been conducted in the frame of the Fondazione CARIPLO project “STRESS: Strategies, Tools and new data for REsilient Smart Societies” #2016-0766, “Bando Fondazione Rst - Ricerca dedicata al dissesto idrogeologico 2016” and SIMULATOR-ADS project, # 137287, co-funded by Regione Lombardia & FESR “Linea R&S per Aggregazioni”. Authors wish to acknowledge anonymous reviewers for their valuable comments.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi: <https://doi.org/10.1016/j.jag.2019.101951>.

References

- Acharya, T.D., Subedi, A., Yang, I.T., Lee, D.H., 2017. Combining water indices for water and background threshold in Landsat image. *Proceedings* 2, 143. <https://doi.org/10.3390/ecsa-4-04902>.
- Acharya, T.D., Subedi, A., Lee, D.H., 2018. Evaluation of water indices for surface water extraction in a Landsat 8 scene of Nepal. *Sensors* 18, 2580.
- Bordogna, G., Pagani, M., Pasi, G., 2007. A flexible decision support approach to model ill-defined knowledge in GIS. In: Morris, A., Kokhan, S. (Eds.), *Geographic Uncertainty in Environmental Security*. Springer.
- Boschetti, M., Nutini, F., Manfron, G., Brivio, P.A., Nelson, A., 2014. Comparative analysis of normalised difference spectral indices derived from MODIS for detecting surface water in flooded rice cropping systems. *PLoS One* 9. <https://doi.org/10.1371/journal.pone.0088741>.
- Bresciani, M., Stroppiana, D., Odermatt, D., Morabito, G., Giardino, C., 2011. Assessing remotely sensed chlorophyll-a for the implementation of the water framework directive in European perialpine lakes. *Sci. Total Environ.* 409, 3083–3091. <https://doi.org/10.1016/j.scitotenv.2011.05.001>.
- Carrara, P., Bordogna, G., Boschetti, M., Brivio, P.A., Nelson, A., Stroppiana, D., 2008. A flexible multi-source spatial-data fusion system for environmental status assessment at continental scale. *Int. J. Geogr. Inf. Sci.* 22, 781–799. <https://doi.org/10.1080/13658810701703183>.
- Ceresi, A., Goffi, A., Ranghetti, L., Busetto, L., Stroppiana, D., Bordogna, G., Boschetti, M., Brivio, P.A., Pepe, M., Antoninetti, M., Sterlacchini, S., 2018. “A flexible desktop tool for the deployment of periodic downstream services”, IGARSS 2018. *IEEE International Geoscience and Remote Sensing Symposium*.
- Chen, Y., Huang, C., Ticehurst, C., Merrin, L., Thew, P., 2013. An evaluation of MODIS daily and 8-day composite products for floodplain and wetland inundation mapping. *Wetlands* 33, 823–835. <https://doi.org/10.1007/s13157-013-0439-4>.
- Díaz-Delgado, R., Aragónés, D., Afán, I., Bustamante, J., 2016. Long-term monitoring of the flooding regime and hydroperiod of Doñana marshes with landsat time series (1974–2014). *Remote Sens. (Basel)* 8, 1–19. <https://doi.org/10.3390/rs8090775>.
- Feyisa, G.L., Meilby, H., Fensholt, R., Proud, S.R., 2014. Remote sensing of environment automated water extraction index: a new technique for surface water mapping using Landsat imagery. *Remote Sens. Environ.* 140, 23–35. <https://doi.org/10.1016/j.rse.2013.08.029>.
- Gao, B., 1996. NDWI—a normalized difference water index for remote sensing of vegetation/liquid water from space. *Remote Sens. Environ.* 266, 257–266. [https://doi.org/10.1016/S0034-4257\(96\)00067-3](https://doi.org/10.1016/S0034-4257(96)00067-3).
- Giordan, D., Notti, D., Villa, A., Zucca, F., Calò, F., Pepe, A., 2018. Low cost, multiscale and multi-sensor application for flooded area mapping. *Nat. Hazards Earth Syst. Sci.* 1493–1516. <https://doi.org/10.5194/nhess-18-1493-2018>.
- Huang, C., Chen, Y., Zhang, S., Wu, J., 2018. Detecting, extracting, and monitoring surface water from space using optical sensors: a review. *Rev. Geophys.* 56, 333–360. <https://doi.org/10.1029/2018RG000598>.
- Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* 25, 295–309.
- Hui, F., Xu, B., Huang, H., Yu, Q., Gong, P., 2008. Modelling spatial-temporal change of Poyang Lake using multitemporal Landsat imagery. *Int. J. Remote Sens.* 29, 5767–5784. <https://doi.org/10.1080/01431160802060912>.
- Ji, L., Zhang, L., Wylie, B., 2009. Problems of Dynamic NDWI Threshold and Objectives of the Study the NDWI Data Derived from Landsat MSS, TM, and ETM (Jain et al. 75, 1307–1317. <https://doi.org/10.14358/PERS.75.11.1307>.
- Jonkman, S.N., 2005. Global perspectives on loss of human life caused by floods. *Nat. Hazards Dordr. (Dordr)* 34, 151–175. <https://doi.org/10.1007/s11069-004-8891-3>.
- Kharin, V.V., Zwiers, F.W., Zhang, X., Hegerl, G.C., 2007. Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations. *J. Clim.* 20, 1419–1444. <https://doi.org/10.1175/JCLI4066.1>.
- Kumar, R., Singh, R., Gautam, H., Pandey, M.K., 2018. Flood hazard assessment of August 20, 2016 floods in Satna District, Madhya Pradesh, India. *Remote Sens. Appl. Soc. Environ.* 11, 104–118. <https://doi.org/10.1016/j.rse.2018.06.001>.
- Li, M., Zang, S., Zhang, B., Li, S., Wu, C., 2014. A review of remote sensing image classification techniques: the role of Spatio-contextual information. *Eur. J. Remote Sens.* 47, 389–411. <https://doi.org/10.5721/EuJRS20144723>.
- McFeeters, S.K., 1996. The use of the normalized difference water index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* 17, 1425–1432. <https://doi.org/10.1080/01431169608948714>.
- Mueller, N., Lewis, A., Roberts, D., Ring, S., Melrose, R., Sixsmith, J., Lymburner, L., McIntyre, A., Tan, P., Curnow, S., Ip, A., 2016. Water observations from space: mapping surface water from 25years of Landsat imagery across Australia. *Remote Sens. Environ.* 174, 341–352. <https://doi.org/10.1016/j.rse.2015.11.003>.
- Pekel, J.F., Belward, A., Gorelick, N., 2017. *Global Water Surface Dynamics: Toward a Near Real Time Monitoring Using Landsat and Sentinel Data*.
- Pekel, J.F., Cottam, A., Gorelick, N., Belward, A.S., 2016. High-resolution mapping of global surface water and its long-term changes. *Nature* 540, 418–422. <https://doi.org/10.1038/nature20584>.
- Pekel, J.F., Vancutsem, C., Bastin, L., Clerici, M., Vanbogaert, E., Bartholomé, E., Defourny, P., 2014. A near real-time water surface detection method based on HSV transformation of MODIS multi-Spectral time series data. *Remote Sens. Environ.* 140, 704–716. <https://doi.org/10.1016/j.rse.2013.10.008>.
- Rahman, M.S., Di, L., 2017. The state of the art of spaceborne remote sensing in flood management. *Nat. Hazards Dordr. (Dordr)* 85, 1223–1248. <https://doi.org/10.1007/s11069-016-2601-9>.
- Ranghetti, L., Bassano, B., Bogliani, G., Palmonari, A., Formigoni, A., Standardi, L., von Hardenberg, A., 2016. MODIS time series contribution for the estimation of nutritional properties of alpine grassland. *Eur. J. Remote Sens.* 49, 691–718. <https://doi.org/10.5721/EuJRS20164936>.
- Ranghetti, L., Cardarelli, E., Boschetti, M., Busetto, L., Fasola, M., 2018. Assessment of water management changes in the Italian rice paddies from 2000 to 2016 using satellite data: a contribution to agro-ecological studies. *Remote Sens. (Basel)* 10, 416.
- Ranghetti, L., Busetto, L., 2019. Sen2r: An R Toolbox to Find, Download and Preprocess Sentinel-2 Data. R Package Version 1.0.0. <https://doi.org/10.5281/zenodo.1240384>. URL: <http://sen2r.ranghetti.info>.
- Robinove, C.J., 1978. Interpretation of a Landsat image of an unusual flood phenomenon in Australia. *Remote Sens. Environ.* 7, 219–225. [https://doi.org/10.1016/0034-4257\(78\)90033-0](https://doi.org/10.1016/0034-4257(78)90033-0).
- Rokni, K., Ahmad, A., Selamat, A., Hazini, S., 2014. Water feature extraction and change detection using multitemporal landsat imagery. *Remote Sens. (Basel)* 6, 4173–4189. <https://doi.org/10.3390/rs6054173>.
- Schumann, G.J.P., 2015. Preface: remote sensing in flood monitoring and management. *Remote Sens. (Basel)* 7, 17013–17015. <https://doi.org/10.3390/rs71215871>.
- Shen, L., Li, C., 2010. Water body extraction from landsat ETM+ imagery using adaboost algorithm. 2010 18th Int. Conf. Geoinformatics. <https://doi.org/10.1109/GEOINFORMATICS.2010.5567762>. Geoinformatics 2010 1–4.
- Smith, A.R., 1978. Color gamut transform pairs. *Proc. Ofthe 5th Ann. Conf. Comput. Graph. Interact. Tech* 23–25, 12–19.
- Stroppiana, D., Azar, R., Calò, F., Pepe, A., Imperatore, P., Boschetti, M., Silva, J.M.N., Brivio, P.A., Lanari, R., 2015. Integration of optical and SAR data for burned area mapping in Mediterranean regions. *Remote Sens. (Basel)* 7, 1320–1345. <https://doi.org/10.3390/rs70201320>.
- Stroppiana, D., Bordogna, G., Carrara, P., Boschetti, M., Boschetti, L., Brivio, P.A., 2012. A method for extracting burned areas from Landsat TM/ETM+ images by soft aggregation of multiple Spectral Indices and a region growing algorithm. *ISPRS J. Photogramm. Remote Sens.* 69, 88–102. <https://doi.org/10.1016/j.isprsjprs.2012.03.001>.
- Stroppiana, D., Boschetti, M., Brivio, P.A., Bocchi, S., 2009. Estimation of plant nitrogen concentration in paddy rice from field canopy spectra. *Riv. Ital. di Telerilevamento* 41, 45–57.
- Wang, Z., Liu, J., Li, J., Zhang, D.D., 2018. Multi-spectral water index (MuWI): a native 10-m multi-spectral water index for accurate water mapping on Sentinel-2. *Remote Sens. (Basel)* 10, 1643 2018.
- Weintrit, B., Osifaska-Skotak, K., Pilarska, M., 2018. Feasibility study of flood risk monitoring based on optical satellite data. *Tijdschr. Econ. Soc. Geogr.* 22, 172–180. <https://doi.org/10.2478/mgsrd-2018-0011>.
- Xu, H., 2006. 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>.
- Xu, Z.S., Da, Q.L., 2003. An overview of operators for aggregating information. *Int. J. Intell. Syst.* 18, 953–969. <https://doi.org/10.1002/int.10127>.
- Yager, R.R., 1988. On ordered weighted averaging aggregation operators in Multicriteria Decisionmaking. *IEEE Trans. Syst. Man Cybern.* 18, 183–190.
- Yang, X., Qin, Q., Grussenmeyer, P., Koehl, M., 2018. Urban surface water body detection with suppressed built-up noise based on water indices from Sentinel-2 MSI imagery. *Remote Sens. Environ.* 219, 259–270. <https://doi.org/10.1016/j.rse.2018.09.016>.