Remote Sensing of Environment 237 (2020) 111582

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Flood mapping under vegetation using single SAR acquisitions

S. Grimaldi[a,*], J. Xu[a], Y. Li[a,b], V.R.N. Pauwels[a], J.P. Walker[a]

[a] Department of Civil Engineering, Monash University, Clayton, Victoria, Australia
[b] WMAwater, Newtown, Victoria, Australia

ARTICLE INFO

Edited by Menghua Wang

Keywords:
SAR
Inundation extent
Flooded vegetation
Fuzzy logic

ABSTRACT

Synthetic Aperture Radar (SAR) enables 24-hour, all-weather flood monitoring. However, accurate detection of inundated areas can be hindered by the extremely complicated electromagnetic interaction phenomena between microwave pulses, and horizontal and vertical targets. This manuscript focuses on the problem of inundation mapping in areas with emerging vegetation, where spatial and seasonal heterogeneity makes the systematic distinction between dry and flooded backscatter response even more difficult. In this context, image interpretation algorithms have mostly used detailed field data and reference image(s) to implement electromagnetic models or change detection techniques. However, field data are rare, and despite the increasing availability of SAR acquisitions, adequate reference image(s) might not be readily available, especially for fine resolution acquisitions. To by-pass this problem, this study presents an algorithm for automatic flood mapping in areas with emerging vegetation when only single SAR acquisitions and common ancillary data are available. First, probability binning is used for statistical analysis of the backscatter response of wet and dry vegetation for different land cover types. This analysis is then complemented with information on land use, morphology and context within a fuzzy logic approach. The algorithm was applied to three fine resolution images (one ALOS-PALSAR and two COSMO-SkyMed) acquired during the January 2011 flood in the Condamine-Balonne catchment (Australia). Flood extent layers derived from optical images were used as validation data, demonstrating that the proposed algorithm had an overall accuracy higher than 80% for all case studies. Notwithstanding the difficulty to fully discriminate between dry and flooded vegetation backscatter heterogeneity using a single SAR image, this paper provides an automatic, data parsimonious algorithm for the detection of floods under vegetation.

1. Introduction

Floods are the most frequent, disastrous and widespread natural hazards of the world (CRED and UNISDR, 2015). Thanks to their synoptic view of the flooded area, Earth observations from space can effectively support emergency management (Ajmar et al., 2017) and enable more comprehensive ways of constraining flood forecast hydraulic models than gauged data (Grimaldi et al., 2016; Schumann and Domeneghetti, 2016). In particular, satellite-borne Synthetic Aperture Radar (SAR) systems, because of their 24 hours and all-weather acquisition capabilities, have become the preferred tool for flood mapping from space (Dasgupta et al., 2018a; Schumann and Moller, 2015). SAR is an active system that emits microwave pulses at an oblique angle towards the target. The amount of microwave energy scattered off an object is mainly a function of its surface roughness, with shape and dielectric properties as secondary factors (Woodhouse, 2005). Rough terrestrial land surfaces reflect the energy in many directions, including back towards the sensor, and therefore appear as high backscatter zones. Conversely, open water has a relatively smooth surface which causes radar radiation to be reflected away from the sensor, resulting in low backscatter (Henderson and Lewis, 2008). This difference in backscatter response generally allows flood mapping (Ulaby et al., 1986). Nevertheless, a number of event-related and catchment-related meteorological and geometric factors can alter the backscatter characteristics causing errors in the detection of the flooded area. For instance, smooth surfaces such as roofs, tarmac and car parks can lead to commission errors (Giustarini et al., 2013), while roughening of the water surface due to rain and wind can lead to omission errors (Zwenzner and Voigt, 2009). Moreover, interpretation of the backscatter response of different targets in urban and vegetated areas in the presence and/or absence of flood water represents the biggest challenge for inundation detection (Pierdicca et al., 2018; Shen et al., 2019). The detrimental impacts of floods on densely populated areas and man-made infrastructures has led to increased research efforts on flood monitoring in urbanized areas (Chini et al., 2019; Li et al., 2019; Lin et al., 2019; Mason et al., 2014; Wang and Tong, 2018). Nevertheless,

* Corresponding author at: Room 115, 18 Alliance Lane, Monash University, 3800 Clayton, VIC, Australia.
E-mail address: Stefania.Grimaldi@monash.edu (S. Grimaldi).

https://doi.org/10.1016/j.rse.2019.111582
Received 17 May 2019; Received in revised form 25 November 2019; Accepted 26 November 2019

flood monitoring in areas with emerging vegetation is essential for comprehensive evaluation of the economic and environmental costs of floods (Dutta et al., 2003; Koks et al., 2019; Molinari et al., 2019). Moreover, adequate understanding of flood dynamics at the large scale is pivotal for emergency and land management in urbanized areas (Falter et al., 2015). The need for methodologies targeting the detection of floods under vegetation has been highlighted by a number of studies (e.g. Brisco et al., 2019; Dasgupta et al., 2018b; Landuyt et al., 2018; Pulvirenti et al., 2011c; Schlaffer et al., 2017; Tsyganskaya et al., 2019; Tsyganskaya et al., 2018b; Tsyganskaya et al., 2016; Twele et al., 2016). Consequently, this study has focused on the investigation of novel numerical techniques to distinguish between dry and flooded vegetation, within the overarching aim of complementing the existing body of literature on the detection of floods in non-obstructed and urbanized areas.

Vegetation backscatter is the result of volume scattering from the canopy, diffuse scattering from the ground, and the double-bounce mechanism due to multiple reflection from the horizontal surface (ground or water) and vertical structures (trunks or stems) (Richards et al., 1987a). When horizontal and vertical surfaces act as a dihedral corner reflector, the double-bounce mechanism becomes the main driver of vegetation backscatter. This effect can be further amplified in flooded conditions when a layer of water covers the ground. In fact, the smooth water surface and high permittivity enhance the coherent specular reflection leading to backscatter in flooded vegetation that is higher than that under non-flooded conditions, as opposed to conditions in which water completely submerges the surface (Engheta and Elachi, 1982; Pierdicca et al., 2018). While optical sensors cannot detect standing water beneath vegetation, this backscatter increase enables SAR the unique opportunity to map floods in areas with protruding vegetation (Schumann and Moller, 2015).

In vegetated areas, however, the electromagnetic interaction phenomena between microwave radiation, horizontal, and vertical surfaces are extremely complex and can mask, diminish or hinder the expected backscatter increase from non-flooded to flooded conditions (Pulvirenti et al., 2013). Backscatter increase is the predominant effect only when the radar signal is able to penetrate the vegetation canopy with enough microwave energy to generate double bounce between horizontal and vertical surfaces (Hess et al., 2003; Richards et al., 1987b; Townsend, 2002a). Consequently, detection of floods is facilitated by low vegetation density, sparse canopy and leaf-off conditions. Moreover, the penetration depth of the SAR signal into vegetation is higher for longer wavelengths, and so use of L-band has been recommended (Henderson and Lewis, 2008; Hess et al., 1990; Richards et al., 1987a). Nevertheless, C-band and X-band data have also been successfully used for the identification of flooded vegetation, especially for sparse forests and leaf-off conditions (Clement et al., 2018; Cohen et al., 2016; Voormansik et al., 2014). Steep (near nadir) incidence angles are generally preferred as they lead to a shorter path of the SAR signal through the canopy, thus enabling the potential for a stronger double bounce mechanism (Lang et al., 2008). Finally, radar signal polarization also impacts the capability of discriminating between flooded and non-flooded vegetation. When available, dual- or quad-polarization SAR data are preferred as they provide more information on the target structural and geometric properties than single-polarization data (Brisco et al., 2013; Plank et al., 2017; Wang and Davis, 1997). When only single polarization data are available, co-polarized signals (HH or VV) are preferred over cross-polarized signals, with HH-polarization leading to more accurate results than VV-polarization (Pierdicca et al., 2013; Townsend, 2002b).

Empirical investigations of radar backscatter increase from non-flooded to flooded conditions reported a wide range of values with up to one order of magnitude difference. For instance, for X-band SAR data, backscatter increase detected in olive groves by Pulvirenti et al. (2013) was 10 times larger than the backscatter increase obtained in dense deciduous forests during the leaf-on phase by Martinis and Rieke (2015). As a consequence of this large variability, the distinction between diffuse vegetation backscattering in dry conditions and the double-bounce phenomena in flooded conditions is a great challenge, and the use of scene-specific thresholds has been recommended (Martinis and Rieke, 2015; Plank et al., 2017; Pulvirenti et al., 2011a). The classification algorithms proposed in the literature broadly rely on change detection techniques and on the use of electromagnetic models. A brief description of the existing techniques is provided below, with a comprehensive review available in Tsyganskaya et al. (2018b).

Change detection methods derive a threshold from the analysis of the backscatter intensity difference between at least two SAR observations, more specifically the image of the flood and a minimum of one acquisition during dry conditions (e.g. Pierdicca et al., 2008). The analysis of backscatter intensity difference can also be complemented with information on interferometric coherence (Canisius et al., 2019; Chaabani et al., 2018; Nico et al., 2000; Pulvirenti et al., 2015). However, in change detection methods, the selection of a dry condition acquisition is a critical step, especially for areas of high backscatter variability such as crops (Hostache et al., 2012; Martinis and Rieke, 2015; Matgen et al., 2011). While consistent acquisition properties have (so far) been required, the hurdle of finding a reference image acquired in the same season of the flooded image can be overcome, at least partially, by producing sets of several multi-year images of non-flooded conditions which contain seasonal information (Cian et al., 2018; Long et al., 2014; Schlaffer et al., 2017). Threshold values of backscatter intensity difference have been derived using supervised approaches based on user-defined training data (Martinis and Rieke, 2015; Plank et al., 2017; Tsyganskaya et al., 2019; Tsyganskaya et al., 2018a), unsupervised approaches such as the application of a Markovian model (Martinis and Twele, 2010), electromagnetic models (Pulvirenti et al., 2011c), or statistical analysis (Tsyganskaya et al., 2016). However, despite the increasing frequency and number of radar acquisitions, an appropriate reference image (or a long time series of images) might not yet be available, especially for high resolution sensors that do not acquire data routinely (Voormansik et al., 2014). Moreover, historical SAR acquisitions have often been used in proof-of-concept studies investigating the use of RS data to improve flood forecast skill (e.g. Hostache et al., 2018). In fact, finding an adequate reference image for selected historical flood events could be an even harder task. Consequently, a flood mapping algorithm for the detection of floods in vegetated areas which makes use of a single SAR acquisition is beneficial to both near-real-time applications and hindcast analysis.

The accuracy of an algorithm for the detection of flooded vegetation using single SAR acquisitions relies on an adequate understanding of the vegetation backscatter response under flooded and non-flooded conditions. Electromagnetic models provide a rigorous theoretical treatment of the scattering of the radiation by vegetation and the underneath horizontal layer in dry and flooded conditions (Bracaglia et al., 1995a, 1995b; Ferrazzoli and Guerriero, 1995; Pulliainen et al., 1994). Nevertheless, application of these models requires a large number of site specific information on the structure of the target (e.g. trunk geometry and permittivity; branch sizes, distribution, and orientation; leaf dimensions and orientation; soil moisture and soil roughness). Pulvirenti et al. (2011c) developed look-up tables for the most common Mediterranean vegetation types. However, in different climatic areas, vegetation parameters are very dissimilar and Cohen et al. (2016), Pulvirenti et al. (2013) and Pulvirenti et al. (2015) tuned the parameters of electromagnetic models using a plethora of detailed ground survey data and optical data. In fact, adequate data for the implementation of electromagnetic models are hardly available and a novel data-parsimonious algorithm for flood mapping in vegetated areas can support a larger number of analysis.

The aim of this study is therefore the development of an algorithm that enables automatic flood mapping in vegetated areas using single SAR acquisitions and commonly available ancillary data. Automatic flood detection is required by near-real-time applications for the purpose of processing several acquisitions over a number of areas, and to provide standardised results. Moreover, despite multi-polarized data being used in continuous wetland monitoring programs (Brisco, 2015; Wohlfart et al., 2018), often only single-polarized or dual polarized data are available for operational flood monitoring (Cohen et al., 2016; Martinis and Rieke, 2015). For this reason, the algorithm proposed in this study has focused on the analysis of single polarized data, specifically HH data, as this is the most frequent acquisition choice in the case of floods (Martinis and Rieke, 2015). This novel approach for the mapping of flooded vegetation uses a statistical technique known as probability binning (Roederer et al., 2001) to analyse the backscatter response of wet and dry vegetation. Thresholds computed using probability binning were then applied within a fuzzy logic approach (Zadeh, 1965) which allows information provided by radar backscatter to be complemented with information on morphology and context. The algorithm was tested on three fine resolution SAR images acquired during the January 2011 flood in the Condamine-Balonne catchment (Queensland, Australia) and the accuracy of the SAR-derived flood extent was evaluated using flood extent layers derived from optical images.

2. Study area and data

2.1. Study area

The study area is the Condamine-Balonne catchment located in Southern Queensland, Australia (Fig. 1a). The drainage area is 136,014 km^2; the main rivers are the Condamine, Balonne, Culgoa, and Maranoa. Two-thirds of the catchment drainage area is nearly flat, with a complex braided river system and numerous creeks joining and breaking away from the main channels. The climate is semi-arid with the area subject to El Niño Southern-Oscillation (ENSO). Accordingly, most of the waterways are ephemeral with extreme flow variability (Arthington and Balcombe, 2011). The Millennium Drought, from the mid-1990s to 2009, was followed by the La Niña Floods (Leblanc et al., 2012) with 50 years average recurrence interval floods occurring in 2010, 2011, and 2012. As a consequence of the large intra and inter-annual hydrological variability, the distribution of the vegetation is driven by landscape position and associated flooding behaviour. The larger rivers are lined by River Red Gum woodland (*Eucalyptus camaldulensis*) which can grow up to 45 m tall, while the smaller rivers and creeks are generally lined by 10 to 20 m tall Coolibah woodlands (*Eucalyptus coolabah*) and river oak (*Casuarina cunninghamiana*). Weeping bottlebrush (*Callistemon viminalis*), which can be up to 10 m tall, is the most common understory. These trees and bushes have sparse canopies, with long and narrow pendulous leaves. Frequently flooded depressions within the major rivers support a range of herbaceous and shrub wetlands species including Cane Grass, Lignum, Golden and Nitre Goosefoot. The grasslands occupy slightly higher and less frequently flooded areas, with extensive areas of the catchment having been cleared for agriculture (Eco Logical Australia, 2016).

2.2. SAR data and optical validation data

The algorithm proposed in this paper was tested on three SAR images acquired during the 2011 flood event. All three images have fine (< 10 m) spatial resolution. Two images were acquired by the X-band wavelength sensor onboard of the COSMO-SkyMed constellation, and one image was acquired by the L-band wavelength sensor onboard of ALOS-PALSAR 1. These SAR images will be hereafter referred to as CSM1, CSM2, and AP, respectively. Airborne (AO) and satellite-borne (SPOT5) optical imagery were used as a validation dataset. These images were digitalised into binary flood extent maps using visual inspection by the Queensland Reconstruction Authority and Geoscience Australia, respectively. Details of the interpretation of the airborne image are available from https://www.data.qld.gov.au/dataset/flood-

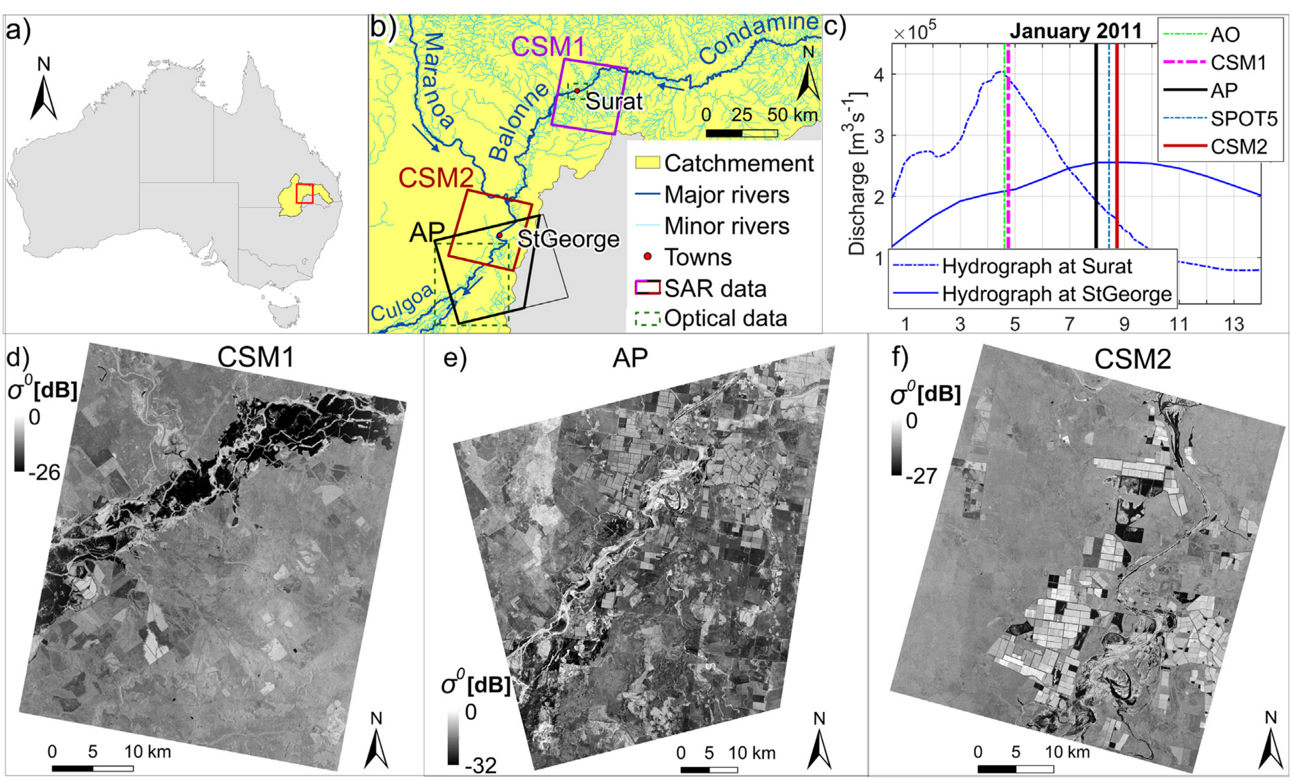

Fig. 1. Location of the Condamine-Balonne catchment (yellow) and of the observed area (red square) (a). Observed area: river network, footprint of the SAR and optical images, major townships (b). 2011 flood event: measured flood hydrographs, acquisition time of the SAR and optical images (c). Backscatter values σ^0 from ALOS PALSAR (d) and COSMO-SkyMed (e, f) acquisitions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
SAR and optical data used for testing the algorithm and validating its results, respectively.

Image acronym	Instrument, sensor (agency)	Date, time [AEST]	Wavelength, polarization	Incidence angle	Pixel resolution [m]
AO	Optical, airborne (Queensland Department of Natural Resources and Mines)	2011/01/04, 14:15–15:12	–	–	2
CSM1	SAR, COSMO-SkyMed2 (Italian Space Agency)	2011/01/04, 18:09	X, HH	40.07–41.89	5
AP	SAR, ALOS PALSAR 1 (Japan Aerospace Exploration Agency)	2011/01/07, 23:10	L, HH	34.3	6.25
SPOT 5	Optical, SPOT5 (Airbus Defense & Space)	2011/01/08, 10:30	–	–	12
CSM2	SAR, COSMO-SkyMed3 (Italian Space Agency)	2011/01/08, 17:27	X, HH	35.92–38.09	5

extent-series. For the satellite image SPOT5, care was taken to include flooded vegetation into the inundation layer (Norman Mueller, personal communication 22/09/2016). In this latter image, areas affected by cloud cover and cloud shadows were excluded from the evaluation dataset.

The sensor, data provider, acquisition time and features of the SAR and optical imagery are detailed in Table 1. Despite a temporal (< 12 h) offset between the SAR and optical sensor overpasses and the well-known difficulty of detecting water under vegetation canopies using optical wavelengths, the airborne and SPOT5-derived flood layers were deemed to provide the most reliable reference for evaluating the accuracy of the presented algorithm. Fig. 1b, c shows the images footprint and acquisition time while Fig. 1d, e, f shows the values of radar backscatter.

2.3. Ancillary datasets

Land cover maps, land use maps, the Water Observation from Space (WOfS) database (Mueller et al., 2016), a Digital Elevation Model (DEM), and extreme flooding assessment maps were the ancillary datasets used for implementation of the interpretation algorithm.

The Dynamic Land Cover Dataset of Australia (Geoscience Australia, 2010) presents land cover information according to the 2007 International Standards Organization (ISO) 19144-2. As detailed in Lymburner et al. (2011), 186 images acquired by the Moderate Resolution Imaging Spectroradiometer (MODIS) from 2000 to 2008 were analysed to produce a 250 m resolution land cover map for the Australian continent. More specifically, 34 land cover classes were used to reflect the structural character of Australian vegetation. These classes can be broadly identified as grasses, shrubs, trees, crops, and bare areas. Shrub and trees were further classified into closed, open, sparse, and scattered according to the following thresholds of canopy cover: > 70%, between 30% and 70%, between 10% and 30%, and < 10%. Urban areas, infrastructure (e.g. roads and water channels), and irrigated areas were identified based on land use data, which are available at the continental scale with 1:2,500,000 resolution (Australian Land Use and Management Classification version 7, ABARES, 2010). Land cover data and extent of the urban and irrigated areas for the footprint of the SAR images are shown in Fig. 2a, b.

The WOfS product is a 25 m resolution gridded dataset indicating areas where surface water has been detected from the analysis of Landsat 5 and 7 acquisitions between 1987 and 2014 (WOfS product version 1.5). The algorithm presented in this paper uses the water filtered confidence layer ($WOfS_{FCL}$), which provides for each pixel the percentage of clear observations for which water was observed. $WOfS_{FCL}$ values higher than 80 identify permanent water bodies (Mueller et al., 2016). Conversely, very low values of $WOfS_{FCL}$ identify areas affected by extreme flood events only. The values for $WOfS_{FCL}$ for the footprint of the SAR images used in this study are shown in Fig. 2c, d.

A Digital Elevation Model (DEM) is a three-dimensional representation of the Earth. The resolution and accuracy of DEMs vary widely depending on the data source and the catchment characteristics (Wang et al., 2018). In this study, the SRTM-derived 30 m resolution DEM-H (Gallant et al., 2011) was used to draw the watersheds and the river network (O'Callaghan and Mark, 1984). These datasets were required for the computation of the Height Above Nearest Drainage (HAND) index and the distance (DIST) index. The HAND index is the height difference between a DEM cell and the nearest cell along the drainage channel while the DIST index is the planar distance along the same drainage path (Nobre et al., 2011; Rennó et al., 2008). These indices were computed using the toolbox developed by Schwanghart and Scherler (2014) in Matlab. Algorithms for the computation of the HAND and the DIST indices have also been developed in Python (Bartos, 2018) and ArcGIS (Samadi, 2018). Moreover, the HAND index can be retrieved from global and continental databases (e.g. Donchyts et al., 2016; Yamazaki et al., 2019).

The Queensland Floodplain Assessment Overlay (Queensland Reconstruction Authority, 2013) represents an estimate of areas potentially at threat of flooding. This layer will be hereafter referred to as the Extreme Flood Boundary (EFB) and was developed utilizing a range of data sources including topographic information and historical flood records. In catchments where information on the Extreme Flood Boundary is not available, a contoured HAND map with a fixed threshold (e.g. 15 m) has been largely employed in RS image analysis to delineate areas at risk of flooding (Huang et al., 2017; Nobre et al., 2016; Twele et al., 2016).

3. Methodology

The flood mapping algorithm proposed here makes use of different information sources; specifically, (1) the backscatter response of the different targets, (2) morphology and land use, and (3) context. These information sources were analysed and combined within a fuzzy logic approach. Fuzzy logic (Zadeh, 1965) has been previously acknowledged as a valuable tool to derive floods beneath vegetation because it enables considering the ambiguities of the radar measurements and the inclusion of ancillary data in the processing algorithm (e.g. Pulvirenti et al., 2013). Within this approach, each pixel is assigned a degree of membership to the flooded area. Degrees of membership are defined by real numbers in the interval [0, 1] and computed using membership functions which allow the comparison of pixel properties with pre-determined parameters of the flooded and non-flooded areas. The workflow and the use of ancillary datasets at each step are shown in Fig. 3. Sections 3.1 to 3.5 present the details of the proposed methodology.

3.1. Pre-processing

The pre-processing of the SAR data requires geometric correction, radiometric calibration, speckle filtering, and the application of exclusion layers. Geometric correction has the purpose to locate the image on the Earth (geocoding) and to correct terrain distortions (orthorectification). Radiometric calibration is conducted by the calculation of the backscattering values (usually indicated with sigma nought, σ^0) for each pixel. More specifically, digital numbers (DNs) of the SAR datasets are converted to backscatter values considering the local incidence angle. While radar intensities can be computed considering both linear and logarithmic (i.e. dB) units, the use of the latter was preferred (Rignot and Van Zyl, 1993).

Speckle is a random noise that occurs when distributed targets are

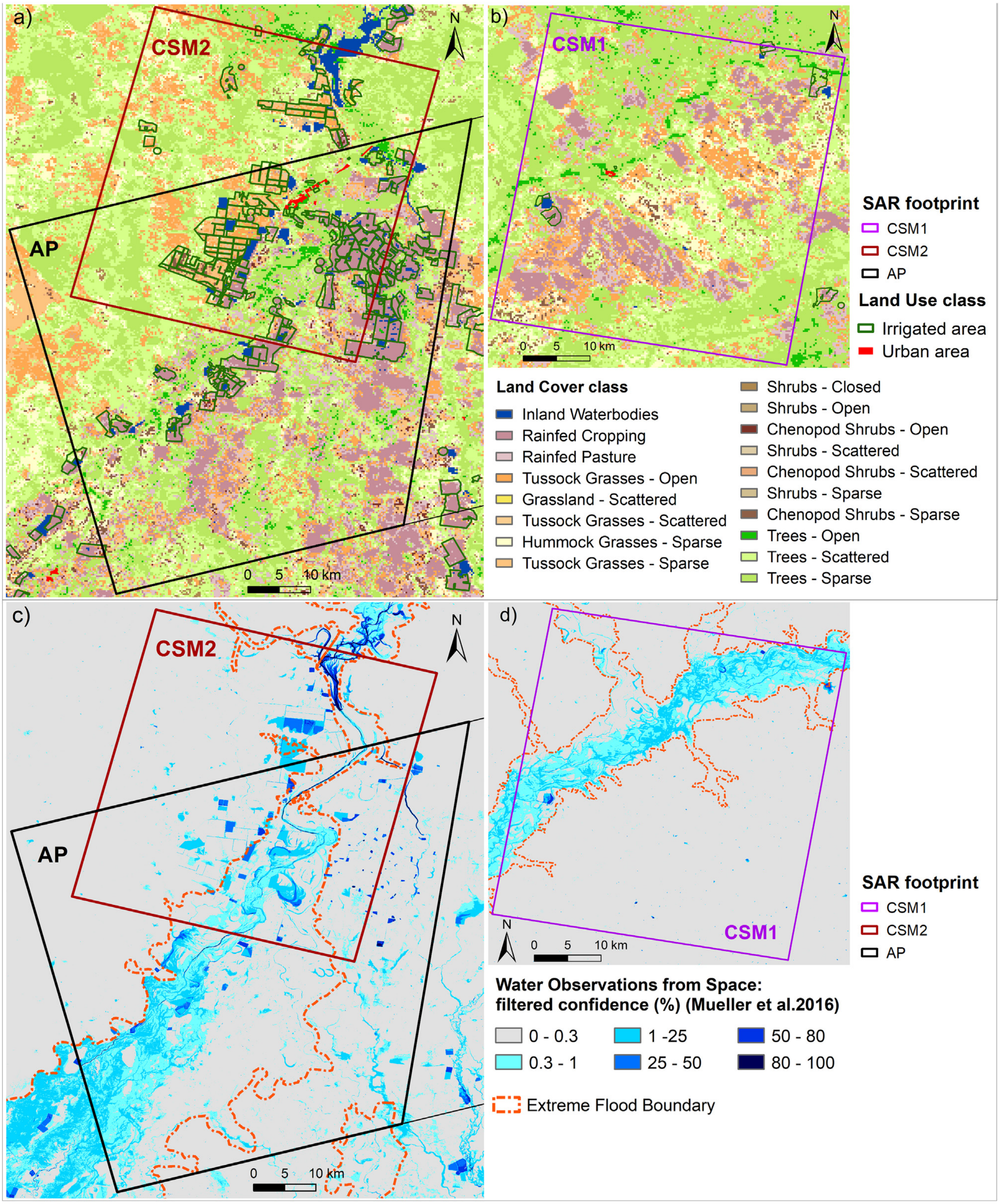

Fig. 2. Land cover data, extent of the irrigated and urban areas for the footprint of AP and CSM2 (a) and CSM1 (b). Water observations from space filtered confidence layer data, Extreme Flood Boundary for the footprint of AP and CSM2 (c) and CSM1 (d).

imaged and the radar intensity represents the contribution of a number of point scatterers (Oliver and Quegan, 2004). A number of techniques exist to diminish this problem; examples include filtering techniques (e.g. Frost et al., 1981; Lee et al., 2009) and bootstrapping methods (Giustarini et al., 2015). This study made use of a 3 × 3 kernel size gamma filter. This technique was selected because it achieves good noise reduction and, at the same time, it preserves details within a limited computational time (e.g. Cohen et al., 2016; Martinis et al., 2009).

The last pre-processing step consists in the application of exclusion layers with the purpose to limit commission errors. More specifically, man-made flat surfaces, such as roads, usually have low backscatter

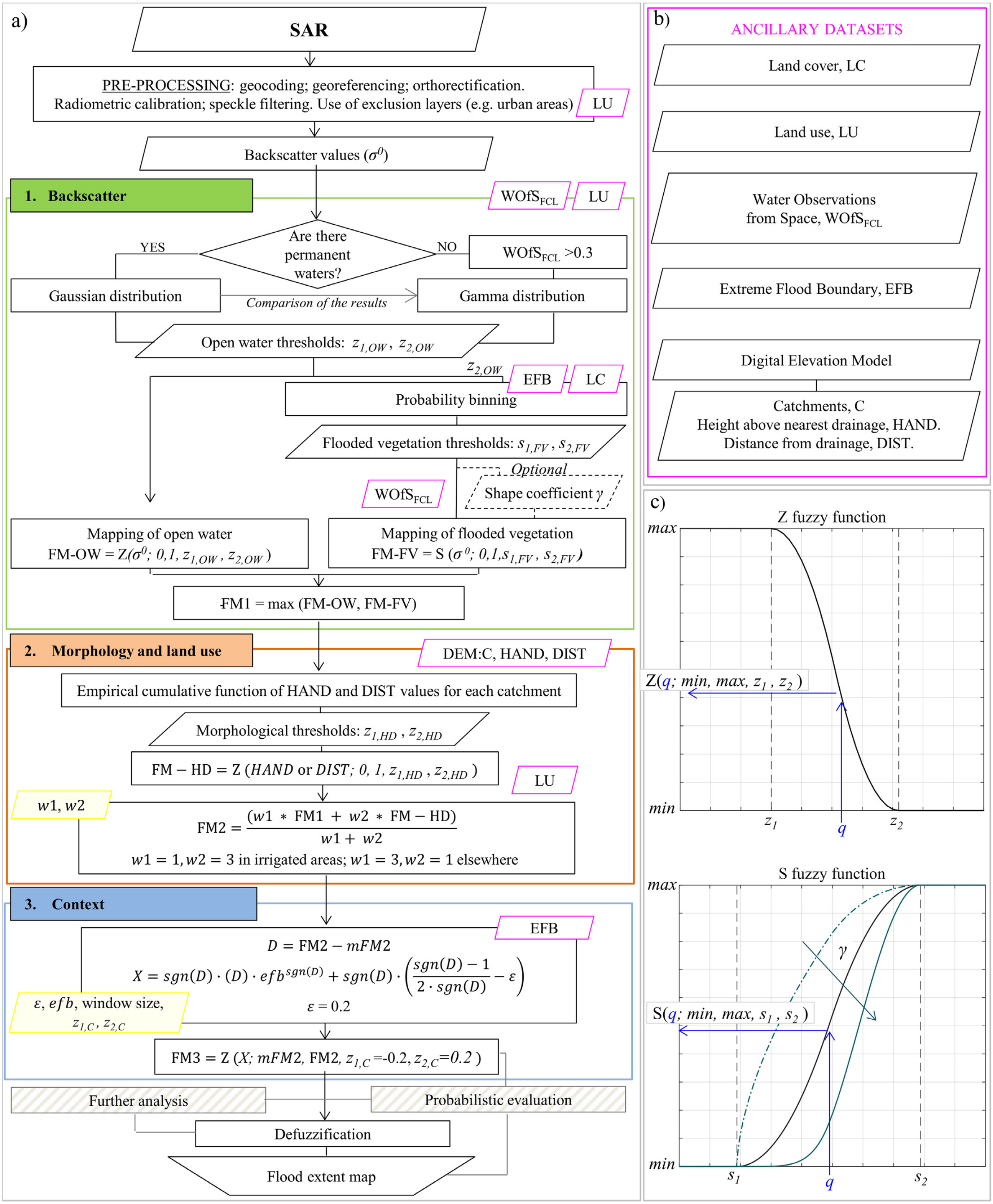

Fig. 3. Workflow of the proposed algorithm (a): parameters that can be changed by the users are in yellow; dashed rectangles feature future improvements to the current algorithm. Ancillary datasets (b). Z and S fuzzy functions; here q is the quantity under investigation (e.g. backscatter value for the computation of FM-OW and FM-FV) (c). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

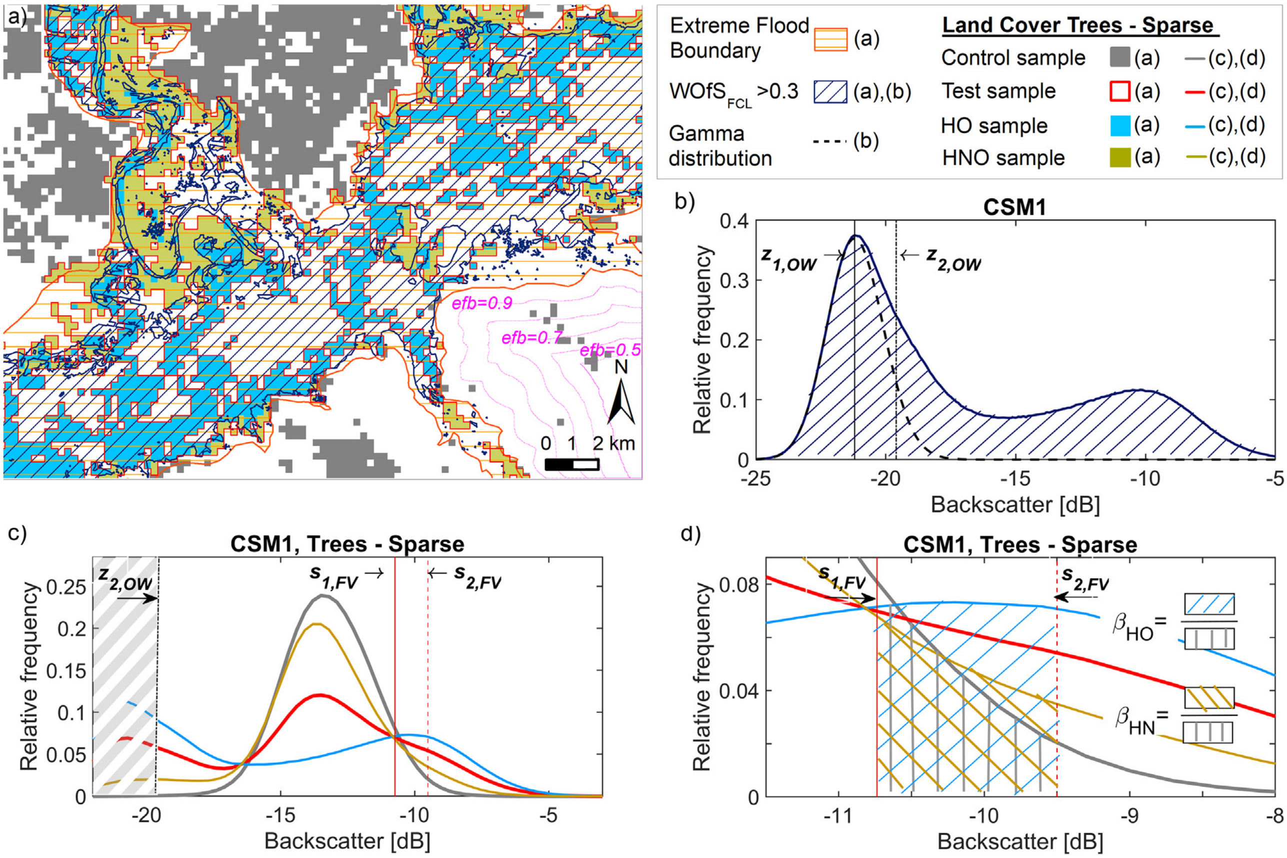

Fig. 4. Schematic of the segmentation of the pixel population according to land cover, EFB, and WOfS. The test case CSM1 and the land cover Trees-Sparse are shown as an example. Spatial distribution of the pixels with WOfS > 0.3 and of the land cover Trees-Sparse with reference to EFB and WOfS (a). Analysis of the backscatter distribution of open water areas (b). Analysis of the backscatter response of the land cover Trees-Sparse for control, test, historically observed (HO), historically non-observed samples (HNO) (c). Computation of β_{HO} and β_{HN} as in Section 3.2.2.3 (d).

response that can be confused with water. Conversely, double bounce backscattering from buildings can be confused with the response of flooded vegetation. Since the purpose of the presented algorithm is the detection of flooded vegetation (Section 1), man-made infrastructures and urbanized areas were identified using land use information and subsequently excluded from the investigated dataset.

3.2. Fuzzy set based on backscatter, FM1

The backscatter distribution of SAR images of areas with extensive flooded vegetation can be interpreted as deriving from the overlapping of the backscatter distribution of open water areas (OW), flooded vegetation (FV), and dry areas (e.g. Tsyganskaya et al., 2018a). In this step of the proposed algorithm, degrees of membership to the open water areas (FM-OW, Section 3.2.1) and to the flooded vegetation (FM-FV, Section 3.2.2) are defined for each pixel based on its backscatter value; the ancillary datasets used are land cover, land use, EFB, and WOfS. To facilitate the understanding of the processing steps, Fig. 4 provides a graphical explanation of the use of the ancillary datasets and of the analysis of the distribution of the backscatter values. It is important to note here that the rapid, yet simple approach for the computation of FM-OW has the purpose to allow the extraction of areas of low backscatter to facilitate the computation of FM-FV. The limitations and potential for improvement of the methodology explained in Section 3.2.1 are discussed in Section 5. FM-OW and FM-FV are combined to compute the membership value FM1 (Section 3.2.3) which provides the core of the SAR-derived inundation layer.

3.2.1. Degree of membership to the open water areas, FM-OW

Smooth water surfaces have a low backscatter and so the standard Z fuzzy function (Pal and Rosenfeld, 1988; Fig. 3c) was selected to compute the membership of a pixel to open water surfaces (FM-OW). In this fuzzy function, the lower the intensity of the pixel, the higher its degree of membership; this choice is consistent with previous studies (e.g. Martinis et al., 2015; Pierdicca et al., 2008; Pulvirenti et al., 2011c). The parameters $z_{1,OW}$ and $z_{2,OW}$ have often been defined after the partitioning of the SAR image with the purpose to segregate areas where backscatter distribution is expected to show a peak in the low range of values (Chini et al., 2017; Shen et al., 2019).

In the proposed application, the images are partitioned using information provided by land cover, land use data, and WOfS. More specifically, if permanent water bodies can be identified using land cover, land use, or $WOfS_{FCL} > 80$ (Mueller et al., 2016), the distribution of the backscatter values sampled from these areas is approximated by the Gaussian probability density function (e.g. Giustarini et al., 2016; Ulaby et al., 1986) according to

$$f(\sigma^0 \mid \mu, s) = \frac{1}{s\sqrt{2\pi}} e^{\frac{-(\sigma^0-\mu)^2}{2s^2}}, \tag{1}$$

where σ^0 is the pixel backscatter value (dB), and μ (dB) and s (dB) are the mean and the standard deviation of the Gaussian density function, and estimated using the Levenberg-Marquardt non-linear least squares algorithm (Marquardt, 1963) as explained in Seber and Wild (2003). The Shapiro-Wilk test (Shapiro and Wilk, 1965) was used to check the possibility of approximating the sample with a Gaussian distribution. The parameters of the Z fuzzy function are then defined as $z_{1,OW} = \mu$ and $z_{2,OW} = \mu + 2s$, meaning that pixels with backscatter values lower than the mean had FM1-OW equal to 1 while pixels with backscatter value higher than twice the standard deviation had an FM1-OW of 0. Backscatter values higher than twice the standard deviation are expected to be due to speckle, wind roughening, or other sources of uncertainty. It is noted that thresholds of one time and three times the standard deviation were also tested (analysis not shown here).

In semi-arid areas with ephemeral rivers, permanent water bodies might not be identified or their backscatter distribution might not be adequately represented by a Gaussian distribution (Fig. 4a). In such a scenario, the algorithm automatically retrieves the backscatter values of pixels with $\mathrm{WOfS_{FCL}} > \alpha$. Inspection of the $\mathrm{WOfS_{FCL}}$ database for a number of Australian catchments demonstrated that a threshold value of $\alpha = 0.3$ allows considering the inundation extent of large and rare flood events. The resulting sample of backscatter values selected from the SAR image is then approximated by a gamma probability density function according to

$$f_{\sigma_m}(\sigma^0 \mid k) = \frac{(\sigma^0 - \sigma_1^0)^{(k-1)}}{\left(\frac{\sigma_m^0 - \sigma_1^0}{k-1}\right)^k \bullet \Gamma(k)} e^{\frac{-(\sigma^0 - \sigma_1^0)(k-1)}{(\sigma_m^0 - \sigma_1^0)}}, \tag{2}$$

where σ^0 is the pixel backscatter value (dB), $\sigma_1{}^0$ (dB) is the minimum backscatter value in the analysed area of the SAR image, $\sigma_m{}^0$ (dB) is the distribution mode, k (–) is the shape parameter, and Γ is the gamma function. Here, similarly to previous studies (e.g. Giustarini et al., 2013; Giustarini et al., 2015; Matgen et al., 2011), the asymmetric and left-skewed gamma distribution, as opposed to the symmetrical Gaussian distribution, was preferred for the retrieval of the OW parameters from an enlarged sample of pixels, which is likely to include the higher backscatter response of dry and flooded vegetation pixels. Consequently, assessment of the parameters $\sigma_m{}^0$ and k is focused on the lower values of the distribution, up to an empirically-derived threshold $\sigma_{th}{}^0$ ($> \sigma_m{}^0$). This threshold represents the backscatter value above which the empirical density function departs from the theoretical gamma distribution. Assessment of $\sigma_m{}^0$, k and $\sigma_{th}{}^0$ is achieved as explained in Matgen et al. (2011), using an iterative approach to minimize the RMSD between the theoretical density function and the empirical density function (Fig. 4b). These values are then used for the computation of FM-OW. More specifically, backscatter values below $\sigma_m{}^0$ are attributed to open water areas, thus $z_{1,OW} = \sigma_m{}^0$, while $\sigma_{th}{}^0$ is used to identify the upper parameter of the Z fuzzy function, thus $z_{2,OW} = \sigma_{th}{}^0$.

3.2.2. Degree of membership to the flooded vegetation, FM-FV

The standard S fuzzy membership function (Pal and Rosenfeld, 1988; Fig. 3c) was selected to compute the degree of membership FM-FV of a pixel to the flooded vegetation. In this function, the higher the intensity of the pixel, the higher its degree of membership. Previous studies defined the parameters of the S function using change detection (e.g. Pierdicca et al., 2008; Tsyganskaya et al., 2016) or electromagnetic models (e.g. Pulvirenti et al., 2011c).

The algorithm implemented in this study adopted the statistical technique called probability binning, proposed by Roederer et al. (2001). Probability binning is a variant of the chi-squared statistic and was formulated to allow automated non-parametric comparisons of highly-overlapping distributions. More specifically, probability binning defines whether a test sample is statistically significantly different from a control sample, i.e. the test sample and the control sample belong to different populations. The description reported below is limited to the steps implemented within the proposed image classification algorithm; a complete demonstration of probability binning can be found in Roederer et al. (2001).

3.2.2.1. Lower threshold value $s_{1,FV}$.

The threshold value $s_{1,FV}$ is computed using probability binning; for this purpose, the first step consists in the identification of the test and control samples. The test sample is composed by n^{test} pixels representing the backscatter response of possibly flooded vegetation and is compared with a control sample of n^{cont} pixels representing the backscatter response of supposedly dry vegetation. The control sample is divided into K bins such that each bin contains the same number of pixels $n_k{}^{cont}$, meaning that a randomly-selected pixel from the control sample has an equal probability of being assigned to any of the bins. Using a higher number of bins resolves distributions with higher confidence. Following a sensitivity analysis (not shown here), a value of K equal to 100 was used for the implementation of the proposed algorithm. The subsequent step counts the number of pixels $n_k{}^{test}$ of the test sample that belong to each bin to allow the computation of a normalized chi-squared $\chi_{PB}{}^2$ value following

$$\chi_{PB}^2 = \sum_{k=1}^{K} \frac{\left(\frac{n_k^{cont}}{n^{cont}} - \frac{n_k^{test}}{n^{test}}\right)^2}{\frac{n_k^{cont}}{n^{cont}} + \frac{n_k^{test}}{n^{test}}}. \tag{3}$$

Roederer et al. (2001) then proposed a $T(\chi)$ metric to quantify whether the control and test samples have the same distribution. This metric is analogous to the t-score as in

$$T(\chi) = \max\left(0, \frac{\chi_{PB}^2 - \overline{\chi_{PB}^2}}{\sigma_{\chi_{PB}^2}}\right), \tag{4}$$

where

$$\overline{\chi_{PB}^2} = \frac{K}{N} \tag{5}$$

$$\sigma_{\chi_{PB}^2} = \frac{\sqrt{K}}{N} \tag{6}$$

$$N = \min(n^{cont}, n^{test}). \tag{7}$$

The value $\overline{\chi_{PB}^2}$ is obtained for comparing equivalent data sets. Consequently, any $\chi_{PB}{}^2$ equal to or less than this value indicates that the two compared data sets have the same distribution. A value $T(\chi) = 0$ implies that the two distributions are indistinguishable; a value $T(\chi) = 1$ means that the two distributions are the same with a probability $p < 0.17$; $T(\chi) > 4$ implies that the two distributions are the same with a $p < 0.01$ (i.e., 99% confidence that the distributions are different). When the latter condition is verified, the quantity Z_k can be computed for each bin to separate the bins in the test dataset that do not belong to the control distribution following

$$Z_k = \frac{\left(\frac{n_k^{cont}}{n^{cont}} - \frac{n_k^{test}}{n^{test}}\right)}{\sqrt{\frac{n_k^{cont}}{(n^{cont})^2} + \frac{n_k^{test}}{(n^{test})^2}}}. \tag{8}$$

Z_k is interpreted as a realization of an approximately Gaussian distribution with zero mean and unit variance. Bins of the test dataset for which absolute values of Z_k scores are larger than 1.96 belong to a different distribution than the control dataset. In this application, bins representative of the double bounce effect are expected to accommodate a larger number of test pixels than the number of control pixels (that is, $n_k{}^{test} \gg n_k{}^{cont}$) in the higher range of backscatter values. Consequently, in the proposed algorithm, the backscatter response of flooded vegetation is given by the union of the bins with Z_k scores lower than -1.96.

The identification of a control sample is pivotal for the implementation of this methodology. Backscatter values of pixels located outside of the EFB are therefore assumed to adequately represent the backscatter response of dry vegetation. Conversely, it is hypothesized that the distribution of backscatter values of pixels inside the EFB results from the overlapping of the response of dry and flooded vegetation and pixels from this area are used to build the test sample (Fig. 4a, c). Since this analysis aims at detecting areas of high backscatter due to double bounce, backscatter values lower than $z_{2,OW}$ (the highest parameter of the Z fuzzy function for the identification of OW, Section 3.2.1) are excluded from both the control and the test sample. The impact of different vegetation structural characteristics on the distribution of radar backscatter is accounted for by using information provided by land cover and land use maps. Cultivated areas subject to irrigation can cause large uncertainty in the interpretation (e.g. Martinis and Rieke, 2015). To limit this problem, irrigated areas are identified using the land use map to define irrigated areas-specific control and test samples. Subsequent to the extraction of irrigated areas, control and test samples are also built for each land cover class. Probability binning is then applied for the assessment of irrigated areas-specific and land cover-specific values of the parameter $s_{1,FV}$ of the S fuzzy membership function (Figs. 3c, 4c). More specifically, $s_{1,FV}$ is given as the lowest value of the union of the bins with Z_k scores lower than -1.96.

3.2.2.2. Higher threshold value, $s_{2,FV}$. The analysis of the right decreasing limb of the relative frequency distribution of the control sample allows the assessment of the upper threshold $s_{2,FV}$. More specifically, having defined m_C as the mode value of the backscatter distribution of the control sample, $s_{2,FV}$ is the backscatter value with relative frequency equal to m_C / 10 (Fig. 4c). This definition was the outcome of a sensitivity analysis (not shown here) according to which such a threshold could account for the noise of radar backscatter while at the same time limiting omission errors.

The Shapiro-Wilk test (Shapiro and Wilk, 1965) was then used to check the possibility of approximating the control sample with a Gaussian distribution (Eq. (1)). The distribution parameters μ and s describing the left portion (up to m_C) of the empirical relative frequency distribution are computed using the Levenberg-Marquardt non-linear fitting algorithm (Marquardt, 1963; Seber and Wild, 2003). A correlation coefficient between the right portion (that is for backscatter values larger than m_C) of the theoretical distribution and of the empirical relative frequency distribution larger than 0.998 is used as an indicator that speckle noise is responsible for backscatter values in the right portion of the control sample. For this reason, backscatter values of the control sample higher than $s_{2,FV}$ are labelled by the interpretation algorithm as non-flooded; a gamma function was also tested but had lower accuracy in representing the left-hand side of the control backscatter empirical distribution. The threshold value 0.998 was selected after extensive testing aiming at maximising the accuracy of the detection of inundated areas (not shown here).

3.2.2.3. Shape coefficient of the S fuzzy function (optional). When information of past inundation extent is available, such as from the WOfS database, the test sample can be split into two sub-samples for the purpose of computing a shape coefficient of the S fuzzy function. The first sub-sample is therefore composed by the pixels which have been previously observed as flooded. This sub-sample will hereafter be referred to as “historically observed” (HO) sub-sample. The remaining pixels are those for which floods have not (yet) been observed; these pixels constitute the “historically non-observed” (HN) sub-sample. In this study, pixels of the test sample having $WOfS_{FCL} > \alpha$ were allocated to the HO sub-sample; remaining pixels were allocated to the HN sub-sample. Consistently with the analysis of the open water areas, α was selected as 0.3. It is here noted that the WOfS database was derived from optical images with 8-days average acquisition frequency, meaning that flooding of pixels of the HN sample could have been missed due to clouds, tree canopies, or low acquisition frequency. The ratios β_{HO} and β_{HN} between the areas under the empirical relative frequency distribution of these two latter samples, and control sample for backscatter values larger than $s_{1,FV}$ and lower than $s_{2,FV}$, is used as a holistic measure of the difference between flooded and dry populations (a schematic is shown in Fig. 4d). This evaluation is reflected in the definition of a γ multiplicative coefficient of the S function as in

$$\gamma = 4{\beta_H}^{-2}, \tag{9}$$

where β_H is β_{HO} or β_{HN} and consequently, γ is the shape coefficient γ_{HO} applied to the analysis of the HO sample or the shape coefficient γ_{HN} applied to the analysis of the HN sample. Only β_H values larger than 1 are considered for the computation of γ. The lower the areas ratio β_H, the more similar the two distributions, and the more difficult the segregation of the flooded population from the dry population. Conversely, the higher the areas ratio β_H the more different the two distributions and thus the easier the segregation of the flooded vegetation from the dry population. As shown in Fig. 3c, the γ multiplicative coefficient modifies the shape of the S function.

3.2.3. FM1

The total inundation extent in any observed area is given by the union of areas of open water and flooded vegetation. Therefore, the operation of fuzzy union, which assigns to each element the largest degree of membership between FM-OW and FM-FV, is used for the computation of the fuzzy membership FM1 as in

$$\mathrm{FM1} = \max(\mathrm{FM-OW}, \mathrm{FM-FV}). \tag{10}$$

FM1 represents the degree of membership of each pixel to the flooded area according to its backscatter value only. In this study, it was hypothesized that the analysis of backscatter intensity yielded a preliminary inundation map that represents the seed region for the final flooded layer, whose retrieval requires the integration of information on morphology, land use, and context.

3.3. Fuzzy set based on morphology and land use, FM2

Inundation layers derived from a pixel-based backscatter intensity analysis are inevitably prone to uncertainties due to the complexity of the interactions between radar signal, highly heterogeneous vegetation structure, and environmental factors. Moreover, agricultural practices such as flood irrigation are a relevant potential cause of commission error. Despite the analysis of the agricultural areas performed in FM-FV, uncertainties are clearly still possible. Therefore, block 2 of the interpretation algorithm (Fig. 3) aims to reduce commission errors due to bright pixels that are not hydraulically connected to the core of the inundated area. For this purpose, a fuzzy membership function FM-HD is derived from the analysis of valley morphology, and subsequently integrated by information on land use to compute the fuzzy membership function FM2.

3.3.1. FM-HD

The HAND (Nobre et al., 2011) and DIST indices allow verifying whether all the pixels identified as flooded by FM1 are hydraulically connected. Rather than using fixed, pre-defined elevation or distance thresholds, the algorithm computes the Empirical Cumulative Distributions of the HAND (ECD-HAND) and DIST (ECD-DIST) values of all the pixels with FM1 $>$ 0.8 within each watershed. It is hypothesized that flooded pixels are located at elevation and distance values gradually increasing from the nearest river network (Fig. 5a). Therefore, the knee points KH (x_{KH}, ECD_{KH}) and KD (x_{KD}, ECD_{KD}), that is the points where the curvature of the empirical cumulative function has a local maximum and the local slope has an abrupt decrease, pinpoint areas with FM1 $>$ 0.8 which are not hydraulically connected to the core of the inundated area. Consequently, pixels at elevation higher than x_{KH} or distance larger than x_{KD} can lead to commission errors. The fuzzy membership value FM-HD is then computed using a Z membership function. Thresholds values $z_{1,HD}$ and $z_{2,HD}$ are defined by imposing a 0.10 buffer around the lowest of ECD_{KH} and ECD_{KD} (Fig. 5b). HAND or DIST values are then used as input to the Z fuzzy membership function (Fig. 3a, c). The use of a 0.10 buffer, rather than a crisp threshold, allows accounting for inaccuracies in the computation of HAND and DIST and for different valley morphologies. Where a knee point was not

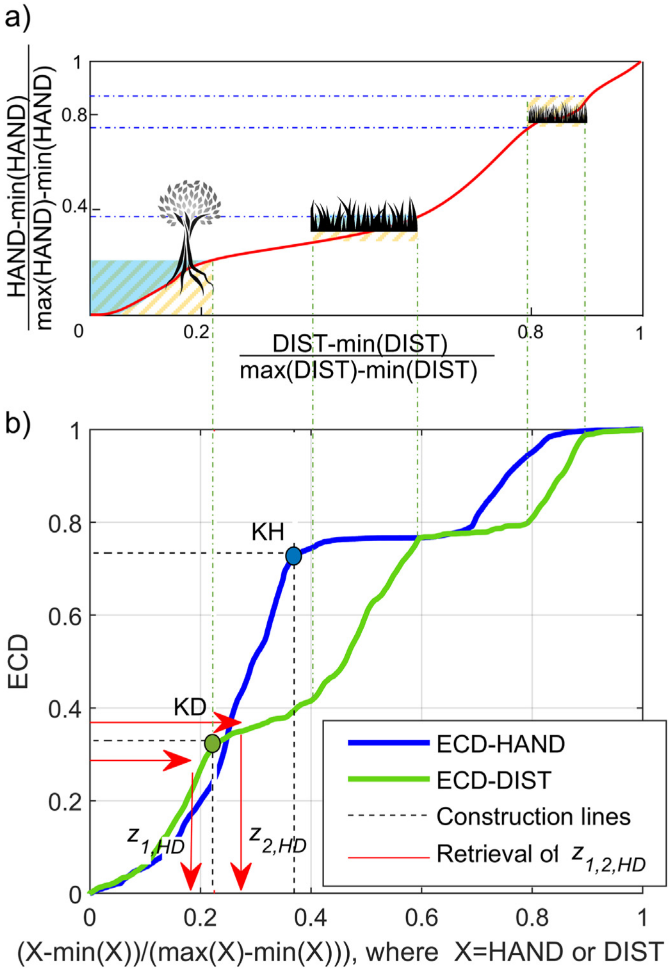

Fig. 5. Schematic of the analysis of HAND and DIST indices: example of the spatial distribution of the indices HAND and DIST for FM1 > 0.8 (a); computation of the thresholds $z_{1,HD}$ and $z_{2,HD}$ (b).

found in either ECD-HAND or ECD-DIST, FM-HD is set equal to 1.

3.3.2. FM2

The weighted average rule (Karmakar and Dooley, 2002; Pulvirenti et al., 2011c) is used to combine the fuzzy sets FM1 and FM-HD into a new fuzzy set FM2 according to

$$\text{FM2} = \frac{(w1 * \text{FM1} + w2 * \text{FM} - \text{HD})}{w1 + \text{w2}}, \tag{11}$$

where weights $w1$ and $w2$ are a function of the land use class. More specifically, FM-HD has the largest weight in irrigated areas ($w1 = 1$, $w2 = 3$). FM1 has the largest weight in the remaining areas ($w1 = 3$, $w2 = 1$). It should be stressed that FM2 is computed only for pixels having FM1 > 0.8. Hence, FM2 depends on the previous analysis, on morphology, on land use.

3.4. Fuzzy set based on context, FM3

Information on context was incorporated into the classification algorithm according to the following considerations: there is a low probability of finding a non-flooded pixel close to flooded ones, especially inside the EFB. Similarly, the probability of the presence of one isolated flooded pixel inside an area of non-flooded ones is low, especially with increasing distance from the EFB.

The fuzzy set FM3 was defined using a Z membership function. First, the difference D between a pixel's membership value FM2 and the average $mFM2$ of the membership values of the neighbouring pixels within a $n \times n$ mobile window is computed as

$$D = \text{FM2} - mFM2. \tag{12}$$

In high resolution images, pixel size smaller than the dimensions of the targets can lead to uncertainties in pixel-based image interpretation (Pulvirenti et al., 2013). As such, considerations on the SAR image pixel size and the expected size of the surface scatterers can provide guidance for the selection of the window size n. A value $n = 11$ was used for all the fine resolution images analysed in this study.

Second, the quantity X is defined as a function of the difference D (Eq. (12)), the position of the pixel with respect to the EFB, and of a parameter ε by

$$X = sgn(D)\bullet(D)\bullet efb^{sgn(D)} + sgn(D)\bullet\left(\frac{sgn(D)-1}{2\bullet sgn(D)} - \varepsilon\right), \tag{13}$$

where efb is a dimensionless index whose value decreases from 1 to 0 with increasing distance from the EFB. For instance, a slow decay rate of −0.1/km, meaning that efb is 1 inside the EFB and 0 at a distance of 10 km, can be used in nearly flat areas (such as the study area presented in this paper; a graphical representation is provided in Fig. 4a). The parameter ε is used to discriminate and consequently remove spurious pixels. After accurate tuning, a value $\varepsilon = 0.2$ was used in the current version of the algorithm. Different values of ε and n can be defined by the operator prior to starting the interpretation algorithm (Fig. 3a). The dimensionless quantity X was then used as input to a Z membership function for the computation of the fuzzy set FM3. The parameters $z_{1,C}$ and $z_{2,C}$ had values of −0.2 and 0.2 respectively, the upper degree of membership was FM2 and the lower degree of membership was $mFM2$ (Fig. 3a, c).

It is noted that the computation of FM3 is conceptually equivalent to the membership functions based on context as defined in Pierdicca et al. (2008) and Pulvirenti et al. (2011c). However, use of the EFB was preferred here over the use of elevation data and could be a viable solution in areas, such as the case study presented here, for which high accuracy Digital Elevation Models are not available.

3.5. Post-processing and evaluation strategy

The final SAR-derived flood extent is obtained through a threshold defuzzification step, which converts the resulting FM3 membership degree into a crisp number. In this study, all the pixels having FM3 larger than the fixed threshold value of 0.5 (e.g. Pierdicca et al., 2008) were allocated to the class water, and the remaining pixels allocated to the class dry. The retrieved flooded area included the permanent water bodies. These areas could be segregated using ancillary datasets such as WOfS_{FCL}, land cover, land use maps, or databases such as OpenStreet Map.

The accuracy of the proposed interpretation algorithm was evaluated by comparing the SAR-derived flood extents with binary flood extent maps retrieved from optical images (Section 2.2, Table 1). Specifically, the classification accuracy was quantified by computing the metrics Overall Accuracy (OA) and the Cohen's kappa coefficient (k). The Producer Accuracy (PA) and User Accuracy (UA) were also computed for both water and dry classes separately. The computation of the accuracy metrics was completed for each main processing step (that is, FM1-OW, FM1, FM2, and FM3) to provide an insight on the relative contribution of backscatter, morphological and context analysis to the accuracy of the SAR-derived inundation layer. It is noted that for the purpose of computing the accuracy metrics, each intermediate fuzzy layer was defuzzified (using a 0.5 threshold). This evaluation was further complemented by an in-depth analysis of the accuracy of the proposed approach for different land cover classes and by the quantification of the contribution of the shape coefficient γ of the S function to the final accuracy of the SAR-derived inundation layer. The former analysis provided insight on the challenges and opportunities for the

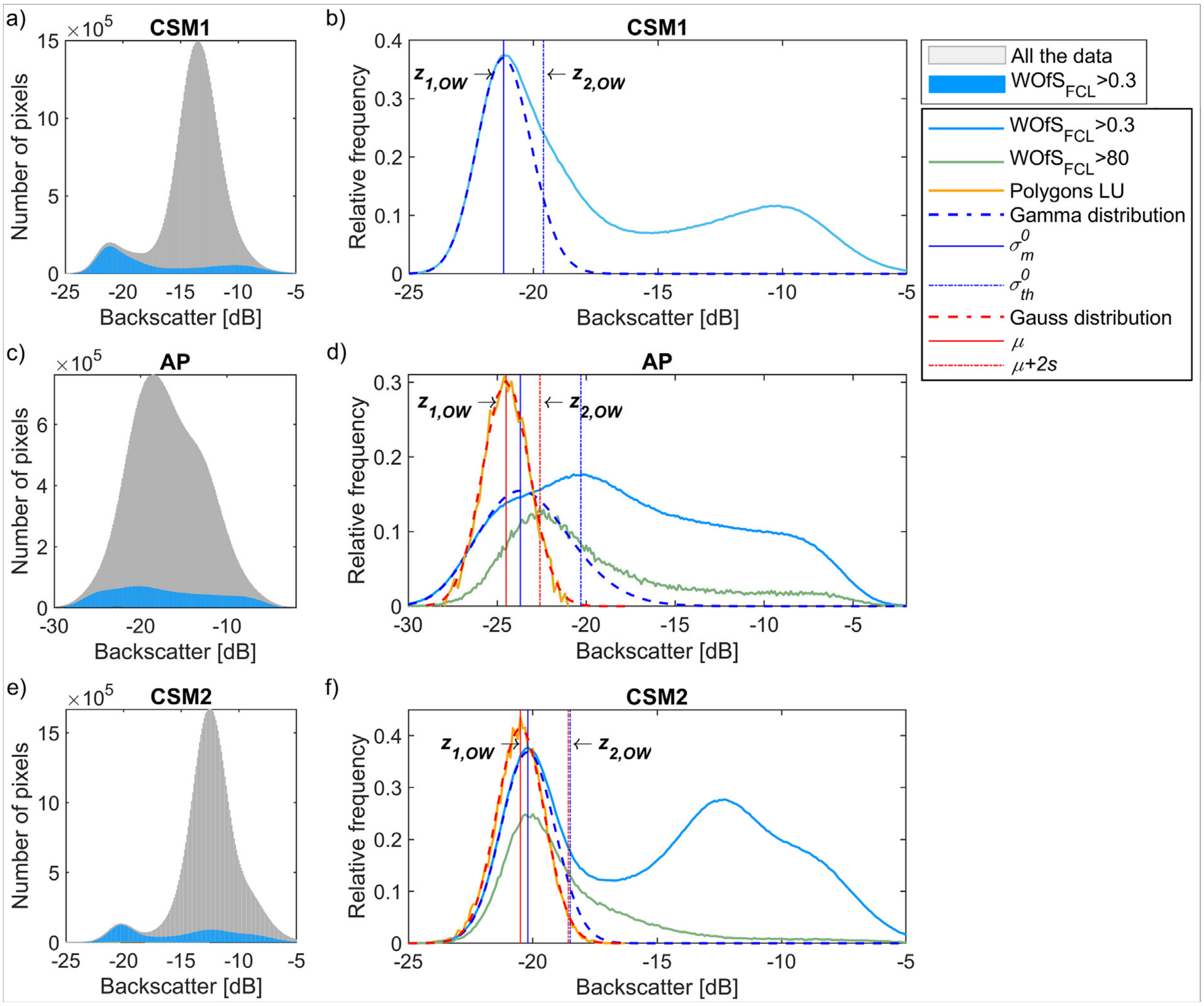

Fig. 6. Distribution of backscatter values of the full image and of pixels having $\text{WOfS}_{\text{FCL}} > 0.3$ for CSM1 (a), AP (c), and CSM2 (e). Analysis of open water areas for CSM1 (b), AP (d), and CSM2 (f). Specifically, distribution of pixels with $\text{WOfS}_{\text{FCL}} > 0.3$ and $\text{WOfS}_{\text{FCL}} > 80$ (pixels with $\text{WOfS}_{\text{FCL}} > 80$ were not found in CSM1). Polygons LU in (d) and (f) are polygons extracted from permanent water areas as identified by the land use (LU) map. Gamma and Gauss functions, computed thresholds $z_{1,OW}$ and $z_{2,OW}$ for CSM1 (b), AP (d), and CSM2 (f).

detection of flooded vegetation with the proposed algorithm. The latter analysis quantified the impact of using historical observations of water on the performance of the interpretation algorithm.

4. Results

4.1. Pre-processing; computation of backscatter threshold values (OW, FV)

Geocoded level 1.5 ALOS PALSAR data were orthorectified using nearest neighbour function (Martinis and Rieke, 2015) and radiometrically calibrated following Lavalle and Wright (2009). COSMO-SkyMed level 1 data were geometrically corrected and radiometrically calibrated by e-GEOS. The cartographic system used for all SAR, optical, and auxiliary data was WGS84, UTM55S.

Fig. 6 shows the thresholds $z_{1,OW}$ and $z_{2,OW}$ for the retrieval of open water areas (Section 3.2.1). More specifically, permanent water areas were not identified within the footprint of CSM1 and the thresholds were derived using a gamma function to approximate the distribution of backscatter values of pixels having $\text{WOfS}_{\text{FCL}} > 0.3$. Conversely, permanent water areas were identified inside the footprints of AP and CSM2 and the thresholds were derived using a Gaussian distribution.

Fig. 7 shows examples of the computation of the thresholds $s_{1,FV}$ and $s_{2,FV}$ for the retrieval of flooded vegetation (Section 3.2.2). Land cover classes having highly different backscatter response were chosen to discuss the results of the analysis. For all the images, land cover classes Trees-Open showed different distributions for the test and control samples; conversely, Tussock Grasses-Open had very similar distributions for the test and control samples. FM-FV was computed only for land cover and land use classes which met the criteria explained in Section 3.2.2.1 and summarised as follows: T (Eq. (4)) higher than 4, a number of bins with Z_k (Eq. (8)) lower than -1.96, and $s_{2,FV}$ larger than $s_{1,FV}$. If one of these requirements was not satisfied (as in Fig. 7g), FM1 depended solely on FM-OW. However, it must be noted that the algorithm does not perform any further evaluation of the results of the probability binning analysis. As such, extremely similar boundary thresholds (as in Fig. 7f) could be used as input to the S fuzzy function. Further testing, featuring the evaluation of the threshold values, could be included in future developments.

4.2. SAR-derived flood extent maps

Figs. 8, 9, and 10 show the results of the computation of the fuzzy membership values FM-OW, FM1, FM2, and FM3 and the performance metrics for CSM1, AP, and CSM2, respectively. In each figure, the red contour is the boundary of the inundation extent as derived from the optical images, whose footprint is shown by the dashed lines.

For each SAR image, increasing values of OA and Cohen's kappa from FM-OW to FM3 demonstrate the contribution of each step of the

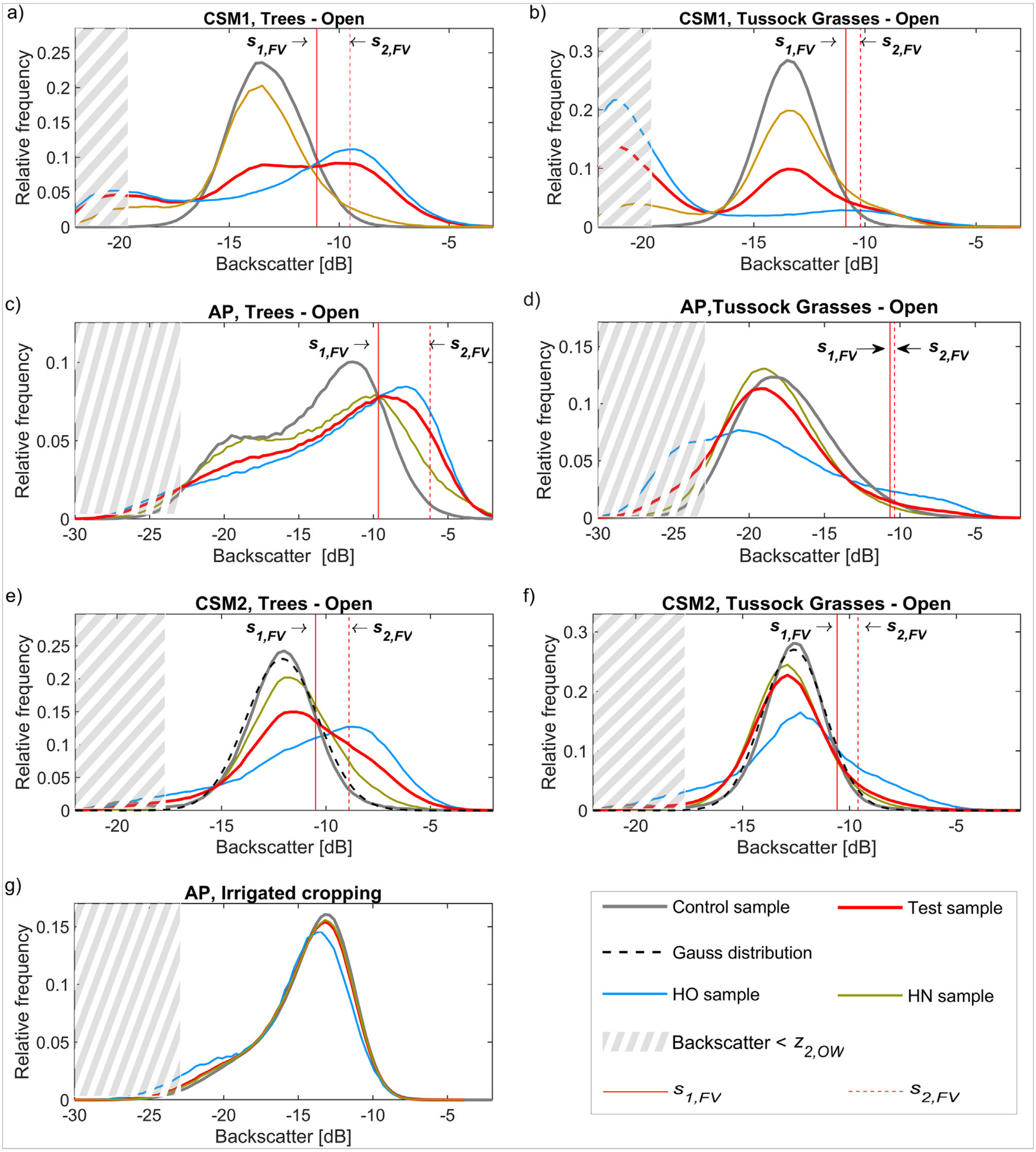

Fig. 7. Examples of empirical distributions of the control sample, test sample, historically observed sub-sample (HO), and historically non-observed (HN) sub-sample, Gaussian approximation of the control sample (where applicable). Computed threshold values $s_{1,FV}$ and $s_{2,FV}$. Land cover Trees-Open for CSM1 (a), AP (c), and CSM2 (e). Land cover Tussock Grasses-Open for CSM1 (b), AP (d), and CSM2 (f). Land Use Irrigated cropping for AP (g).

computational algorithm to the accuracy of the SAR-derived flood extent map. Detection of open water only, that is the use of FM-OW, led to an underestimation of the flooded area with the SAR-derived flood extent maps showing a number of fragmented patches (Figs. 8a, 9a, 10a). Omission errors were highlighted by PA values for the class water (PA_W) as low as 33.3%, 10.1%, and 16.5% for CSM1, AP, and CSM2, respectively. Use of probability binning to detect flooded vegetation allowed to sensibly reduce these errors, and, after computation of FM1, PA_W increased by 62.2%, 75.2%, and 115.1% for CSM1, AP, and CSM2, respectively. The UA for the class dry (UA_D) increased by 14.2%, 2.4%, 5.4% for CSM1, AP, and CSM2, respectively. However, as shown in Figs. 8b, 9b, and 10b, scattered patches of wet areas were returned in the floodplain leading to commission errors. These errors were highlighted by a visual inspection of the areas outside of the footprint of the optical images and were partially quantified by a slight decrease of UA for the class water (UA_w) (−3.6%, −6.7% for CSM1 and AP, respectively) and PA for the class dry (UA_D) (−2.4%, −2.4%, −1.6% for CSM1, AP, and CSM2, respectively).

Commission errors were expected when introducing the analysis of flooded vegetation and FM2 was designed to limit this problem. In these

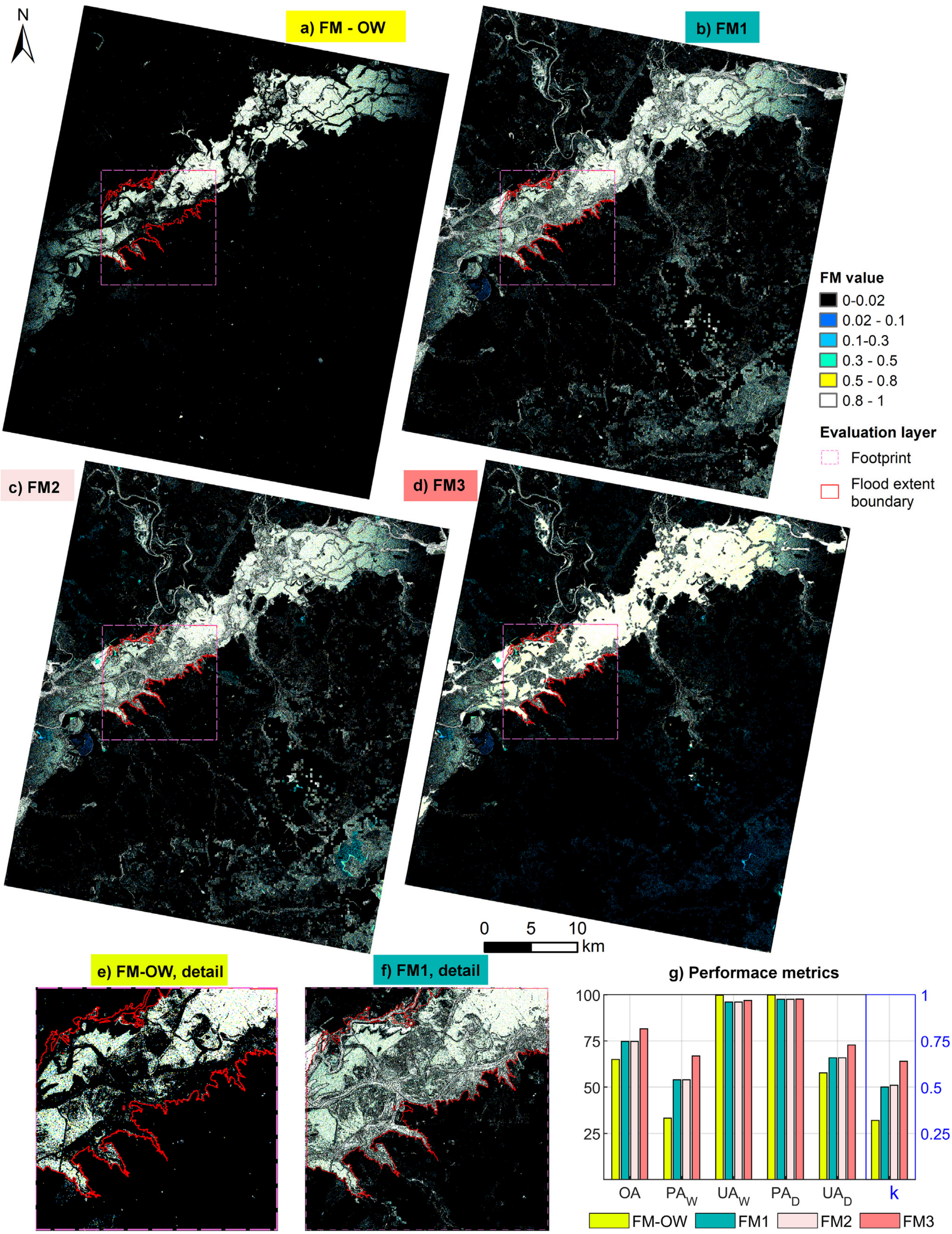

Fig. 8. CSM1: results of the computation of FM-OW (a), FM1 (b), FM2 (c) and FM3 (d). Close-up insets focusing on the footprint of the evaluation data (AO as in Table 1): FM-OW (e), FM1 (f). Performance metrics computed for each processing step (g): Overall Accuracy (OA), Producers' Accuracy for the class water and dry (PA_W, PA_D), Users' Accuracy for the classes water and dry (UA_W, UA_D), Cohen's kappa (k).

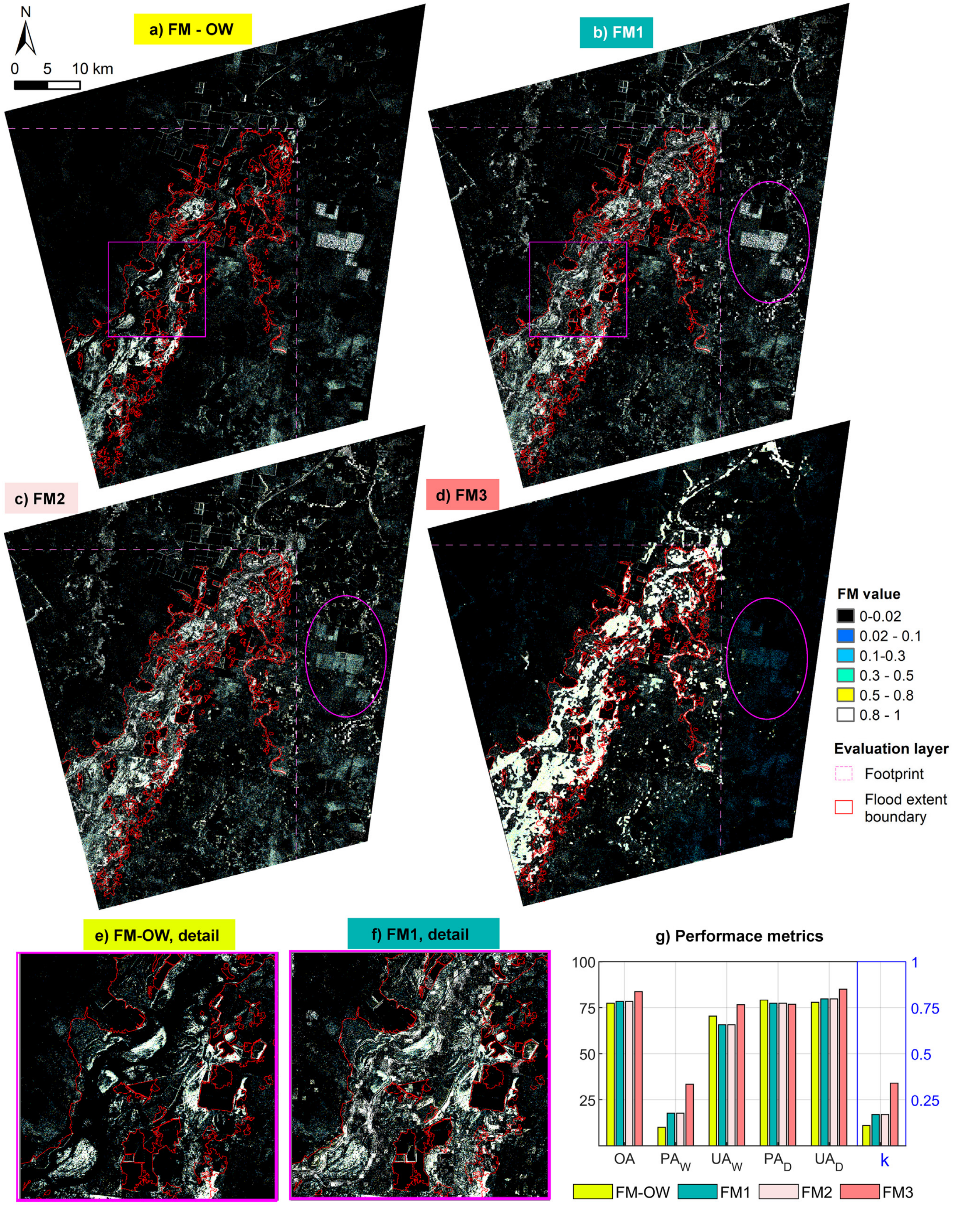

Fig. 9. As per Fig. 8 but for AP. The magenta oval highlights the impact of FM2 for the analysis of irrigated areas. Close-up insets (e, f) feature the area inside the magenta rectangle in (a) and (b). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

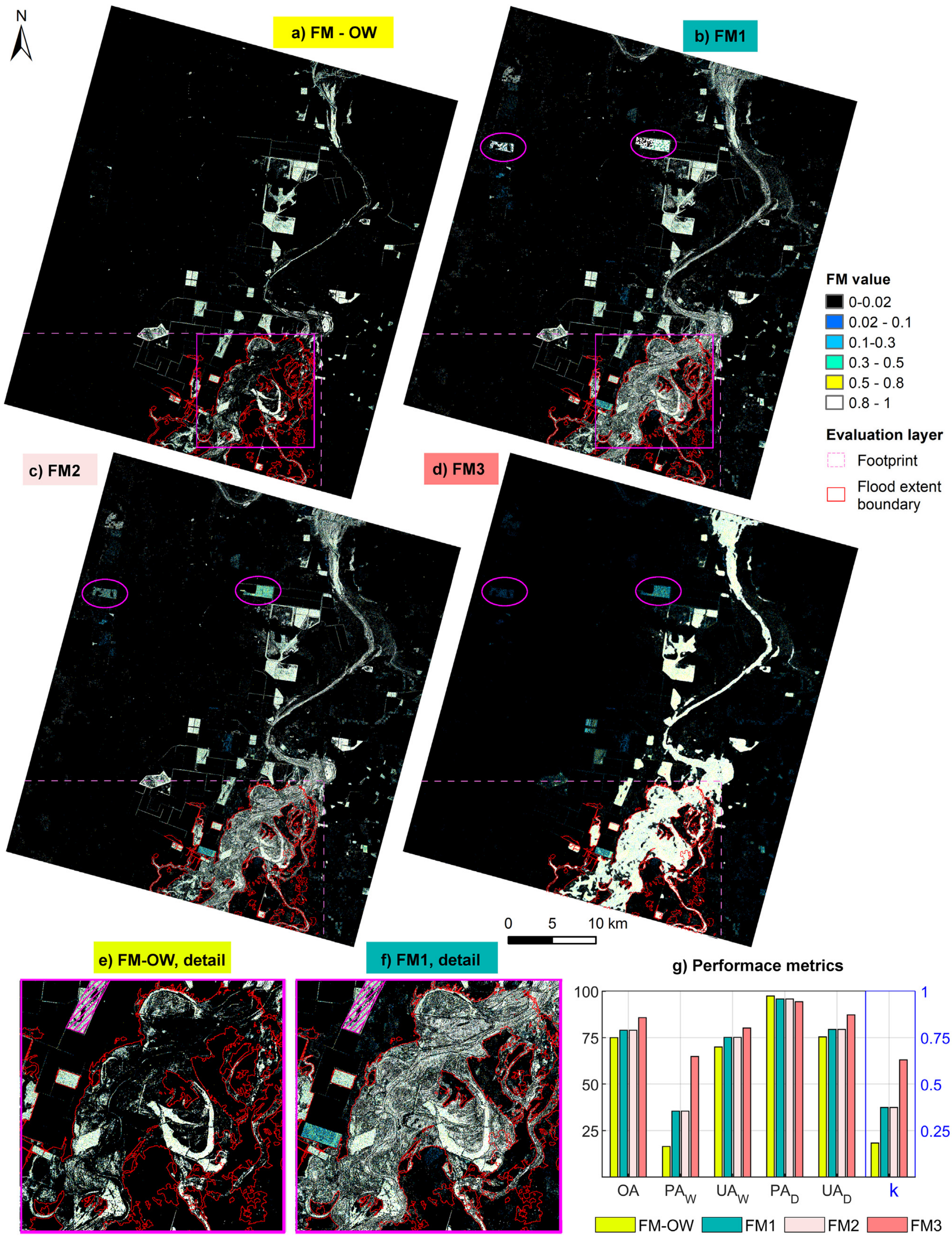

Fig. 10. As per Fig. 8 but for CSM2. The magenta ovals highlight the impact of FM2 for the analysis of irrigated areas. Close-up insets (e, f) feature the area inside the magenta rectangle in (a) and (b); pink dashed area: reservoir. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

case studies, the small footprint of the optical images hindered a comprehensive quantification of the impact of this step of the computational algorithm on the accuracy of the SAR-derived flood extent. Consequently, the values of the performance metrics were similar in FM2 and FM1 for all the images. However, the comparison with Land Use data in Fig. 2a, b highlighted that computation of FM2 diminished the degree of membership to the flooded layer of some irrigated patches, as indicated by the magenta circles in Figs. 9c, 10c. Inclusion of information on context then screened out isolated wet patches (Figs. 8d, 9d, 10d). Upon computation of FM3, PA_W had an increase of 23.9, 88.7, and 82.5% for CSM1, AP, and CSM2, respectively, when compared to FM2. UA increased for both water and dry classes in all the images, with the average increase being 8.5%. PA_D underwent a slight decrease of 0.1%, 0.8% and 1.6% for CSM1, AP, and CSM2, respectively, denoting small false alarms.

These results demonstrated that the fuzzy membership function FM1 was able to identify the seeds of the inundation layer, and that FM2 and FM3 allowed a further refinement of the initial classification. The final classification layer had an OA of 81.5%, 83.7% and 85.7% for CSM1, AP and CSM2, respectively. UA_w, UA_D, and PA_D were larger than 70% for all the images; PA_W had the lowest values being 66.9%, 33.4% and 64.8% for CSM1, AP and CSM2, respectively. The low PA_W of AP led to a low Cohen's kappa of 0.34, while this metric was higher than 0.60 for both CSM1 and CSM2. Despite the encouraging results of the proposed algorithm, further investigations are required to explain the different performances for different images.

Performance metrics were also evaluated for each land cover class. Results were consistent for all images. As an example, Fig. 11 shows results for CSM2. The relative increment of PA_W between FM1 and FM-OW was the highest for the land cover classes Trees-Open (30–70% canopy cover) and Trees-Sparse (10–30% canopy cover). Conversely, the relative increment of PA_W between FM1 and FM-OW was the lowest for land cover classes Tussock Grasses-Sparse and Open, Chenopod Shrub – Sparse, and Hummock Grasses – Sparse. These results can be explained by the more constant structure of the trees (e.g., trunk or branches), as opposed to the irregular morphology of grasses and shrubs. Specifically, the more regular structure of the trees triggered double bounce effects and radar signal increase for all wavelengths. It is noted that canopy cover larger than 70%, i.e. land cover class Trees-Closed, was not detected within these study areas. The analysis of a larger number of case studies is recommended to verify these conclusions.

Finally, the impact of the introduction of the shape coefficient γ of the S fuzzy membership function for the computation of FM1 was tested by running the algorithm with $\gamma = 1$. Impact of using WOfS for the computation of the shape coefficient γ was tangible in CSM1 but negligible in AP and CSM2. In fact, in AP and CSM2, the area ratio β_{HO} had values close to 2 for most of the land cover classes, leading to γ values close to 1. Conversely, in CSM1, β_{HO} and β_{HN} were often close to 4 and 1, respectively, leading to γ values equal to 0.25 and 4. Consequently, omission of the shape coefficient γ for the analysis of CSM1 caused a decrease of the accuracy of the results. More specifically, after computation of the fuzzy set FM1, PA_W, UA_W, OA and Cohen's Kappa were 2.6%, 0.40%, 1.1% and 2.4% lower than the case where the shape coefficient was used. This decrease of accuracy in FM1 led to a further decrease of accuracy in FM2 with a decrease of OA and Cohen's kappa values of 3.5% and 4.5%.

5. Discussion

5.1. Purpose and features of the proposed algorithm

In flooded areas a water layer covers the earth surface. In the presence of emerging vegetation, water surface smoothness and high permittivity generally enhance double-bouncing effects, leading to a radar backscatter higher than that under non-flooded conditions. Such an increase can enable the detection of flooded vegetation using SAR data (Sanyal and Lu, 2004). This work aimed to investigate novel numerical techniques for the mapping of flooded vegetation when data availability or time constraints impede the use of reference images and/or detailed ground information.

The proposed algorithm makes use of probability binning for the analysis of highly overlapping backscatter distributions of dry and flooded vegetation. Differently from other techniques (e.g. mixed distribution analysis; Bioresita et al., 2018), application of probability binning does not require secondary distribution peaks. The use of this methodology allowed correct detection of the initial seed pixels of the flooded area. Identification of these seed pixels is pivotal for the application of further refinements through the use of ancillary data or widely used post-processing methods such as region growing (Giustarini et al., 2013; Li et al., 2018). In fact, as shown in Section 4.2, the use of ancillary datasets within a fuzzy logic approach complemented the results of probability binning analysis, leading to an OA $>$ 80% for all the tested images. The algorithm is automatic and ancillary datasets can be prepared off-line to reduce the computational time. When all the ancillary data were prepared off-line, the computational time to process AP ($\sim 159 * 10^6$ pixels) was 20 min for FM-OW, 23 min for FM-FV, 1 min for FM1, 50 min for FM2, and 14 min for FM3 when using a 3.40 GHz Intel(R) Core™ i7–4770 CPU desktop with 16.0 GB of RAM. However, it must be noted that the current script could be further optimised to limit the computational time.

5.2. Ancillary datasets

Inclusion of ancillary datasets allows for the integration of relevant information on the examined area. However, interpretation accuracy is clearly affected by the quality of these datasets. While only a brief discussion on ancillary data availability and characteristics is presented here, a thorough sensitivity analysis is recommended in a future study through application of the methodology to a large number of catchments. The ancillary datasets used in the proposed algorithm are a DEM, a land cover map, a land use map, and historical observations of flood extents. DEM, land cover maps and land use maps can be derived from satellite and crowd sourced data and are generally available at the global scale (Fritz et al., 2017). In this study, the Dynamic Land Cover Dataset of Australia, despite its coarse resolution (250 m), provided suitable information on the structural characteristics of the vegetation. In fact, the use of this database allowed the comparison of the backscatter response of areas having sufficiently homogeneous vegetation traits. The results showed that the methodology was capable of detecting floods in areas with sparse trees, while it was less effective in non-structured vegetation covers such as grasses and shrubs. The other relevant ancillary dataset used within the proposed algorithm is WOfS (Mueller et al., 2016), which provided information on historical observations of surface water in Australia. A global scale Landsat-derived database of historical water surface observations was published by Pekel et al. (2016). The latter database could enable the application of the proposed algorithm to other continents thus allowing an extensive analysis of the impact of optical-derived historical observations of surface water on the accuracy of the proposed algorithm. This study aimed at providing a first assessment by discussing the impact of the use of WOfS on the SAR-derived flood extent maps. Firstly, WOfS was used for partitioning of the SAR image to facilitate the assessment of the backscatter distribution of open water areas. This step of the algorithm could be replaced by well-established routines, such as automatic tile-based thresholding protocols (e.g. Chini et al., 2017; Martinis et al., 2009; Twele et al., 2016). Second, $WOfS_{FCL}$ allowed comparing the backscatter response of vegetation patches that had been previously observed as flooded with the control sample. Although the improvements shown for one image suggest the potential utility of this comparison, a definite conclusion requires testing other functions and a larger number of case studies.

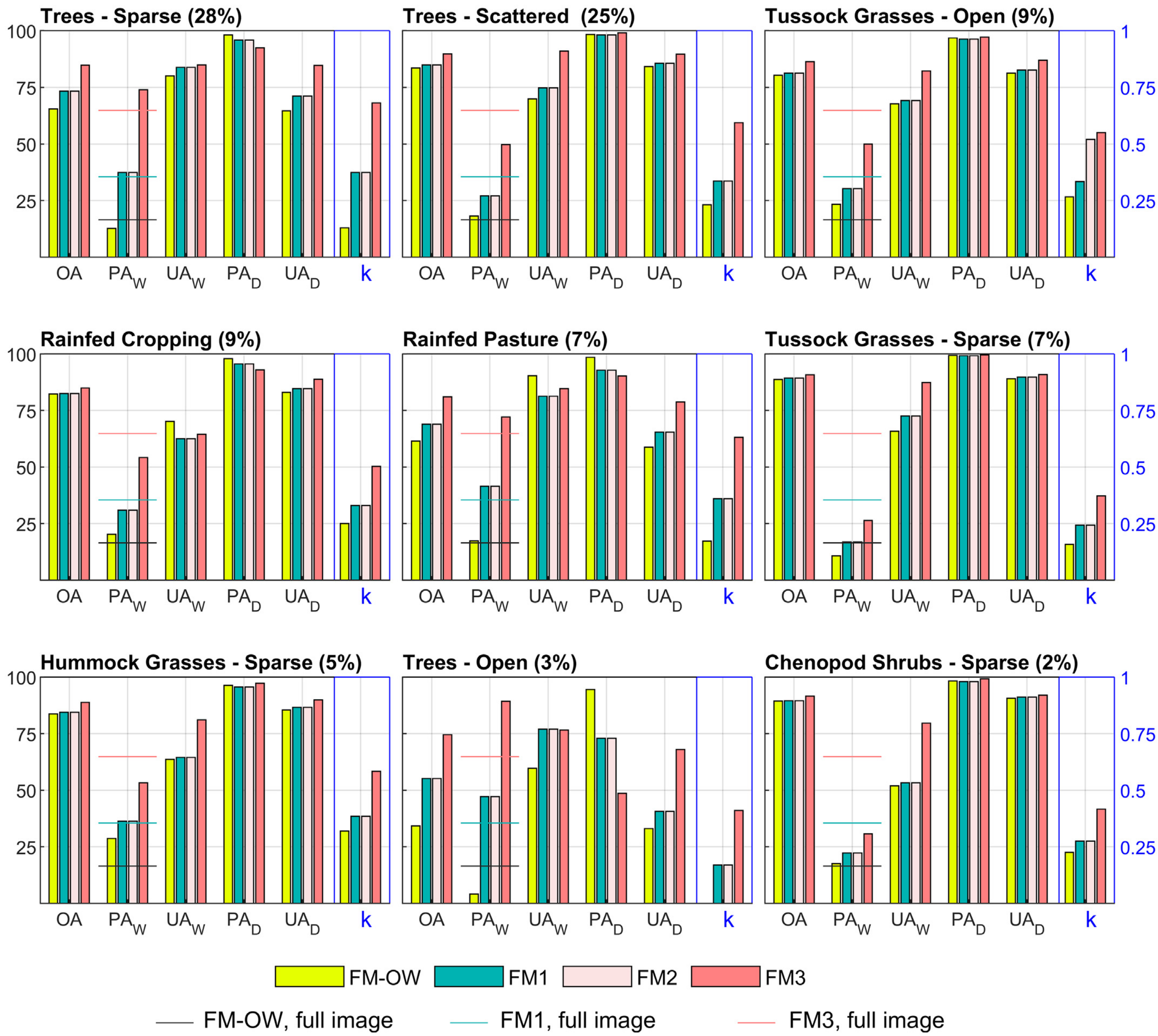

Fig. 11. CSM2: performance metrics computed for the main land cover classes (extent within the footprint of CSM2 indicated as percentage): Overall Accuracy (OA), Producers' Accuracy for the class water and dry (PA_W, PA_D), Users' Accuracy for the classes water and dry (UA_W, UA_D), Cohen's kappa (k). Results forFM2 are reported for completeness, however this processing step targeted land use class irrigated areas and did not impact the analysis of land cover classes. Values of the performance metrics of FM-OW, FM1, and FM3 computed for the full image are reported as reference.

5.3. Limitations and future developments

Notwithstanding the encouraging results shown in this paper, the methodology is affected by limitations that could be overcome by future developments. First, the processing step for the detection of open water areas (FM-OW) is rapid, yet rather simple. The specific scope of the computation of FM-OW was the removal of areas of low backscatter to allow focusing on the discrimination between the backscatter response of dry and flooded vegetation. In fact, the proposal and testing of the numerical analysis steps FM-FV and FM-HD represented the novelty of this study within the overarching aim to complement the existing body of literature on the detection of floods in non-obstructed and urbanized areas. The rapid method for the computation of FM-OW allowed acceptable results in the three test images presented here, where small samples of distinguishable pixels could be found. However, the use of a unique threshold is often not suitable for the whole scene (Cao et al., 2019). This problem is partially highlighted by the results of the analysis of CSM1 (West and East areas of the flood extent layers in Fig. 8a) and is expected to become obvious for images with in-homogeneous or unbalanced backscattering populations. To solve this issue, the first straightforward future development is the integration of the proposed numerical analysis steps with techniques allowing the analysis of smaller tiles, such as hierarchical splitting techniques (e.g. Chini et al., 2017; Martinis and Twele, 2010).

Second, it must be noted that in the three test cases, areas of high backscatter show significant departures from the rather homogenous surroundings. The selection of the test cases was driven by the availability of evaluation data. Nevertheless, the features of the observed areas could impact the algorithm performance. Consequently, extensive testing of the algorithm parameters and several assumptions on a larger number of case studies is required to evaluate the robustness of the methodology. Specifically, application of the algorithm to intensively cultivated areas or areas with significant topography could offer the opportunity for further developments.

In fact, the above suggested further testing would also enable a third area of investigation, that is a sensitivity analysis of the algorithm results to image acquisition properties, pre-processing steps, and features of the land-cover data. Specifically, when SAR data are characterized by a large range of incidence angles, their influence on the backscatter response from flooded vegetation can be significant and has to be considered. Vegetation canopies are intrinsically heterogeneous and radar measurement sensitivity to the spatial arrangement of the scatterers is a big source of uncertainty, especially in high resolution acquisitions (Pulvirenti et al., 2013). Image pre-processing steps and speckle reduction have a smoothing effect that leads to the degradation of the geometric details resulting in information loss, with each spatial resolution characterized by a different compromise between speckle reduction and preservation of geometrical details (Bovolo and Bruzzone, 2005). By segregating areas having the same land cover, the proposed algorithm attempts to analyse samples having similar texture characteristics (with reference to the geometrical features and the phenological phase) and thus enable the detection of backscatter variations due to flooding. However, the use of coarse resolution land cover data (such as the Dynamic Land Cover Map) is likely to impede accurate detection of geometrical details. Hence, future work should investigate the impact of pre-processing steps, the inclusion of a higher resolution land cover dataset and object-based classification (Chini et al., 2011; Pulvirenti et al., 2011b) within each land-cover class.

A fourth area of development could feature a number of additional processing steps. For instance, membership functions formulated in previous studies and focusing on the size and the elevation of the flooded areas, and terrain slope could be included (Gallant and Dowling, 2003; Martinis et al., 2015; Pierdicca et al., 2008; Twele et al., 2016). Moreover, additional ancillary datasets could be used. In non-flooded terrains with high soil moisture, the high dielectric constant causes an increase of backscattering from the ground surface which reduces the contrast between non-flooded and flooded vegetation (Wang et al., 1995). Remote sensing-derived soil moisture information (Gao et al., 2018; Sabaghy et al., 2018; Zhu et al., 2019) could be incorporated in the proposed algorithm to limit this problem. Conversely, in flooded vegetated areas, water depth affects backscatter values (Kasischke et al., 2009). An assessment of water depth could be achieved by overlaying the preliminary SAR-derived flood extent on a DEM (Matgen et al., 2016; Schumann and Di Baldassarre, 2010). Previously investigated relationships between water depth and backscatter values (e.g. Pierdicca et al., 2013) could be incorporated in the interpretation algorithm. Furthermore, such a three-dimensional map of the flooded area could be used to evaluate flow connectivity and verify the connection between irrigated and inundated areas.

Clearly, the inclusion of each suggestion into the classification algorithm should be tested for cost-effectiveness of accuracy compared to required resources and computational time. Nevertheless, SAR-derived flood maps are inevitably affected by inaccuracies and their information content can be considerably increased by including an assessment of the uncertainty of the flood area delineation (D' Addabbo et al., 2016; Di Baldassarre et al., 2009; Giustarini et al., 2016; Schlaffer et al., 2017; Westerhoff et al., 2013). Provision of probabilistic rather than crisp flood extent maps requires evaluation of the statistical significance of each step of the algorithm and would also be the subject of future studies.

Finally, it is important to underline that the proposed algorithm has been tested using high resolution, X and L-band wavelength, HH polarization SAR acquisitions, for which evaluation data were available. However, the algorithm was conceived and designed with the overarching aim of enabling the analysis of SAR acquisitions having different spatial resolution, wavelength, and polarization. While only high-resolution (≤ 10 m) data were used for this study, inundation detection under vegetation with moderate resolution ($\leq 10^2$ m) imagery could be possible in large floodplains with uniform vegetation cover. Extensive testing is strictly required to verify this hypothesis and investigate the trade-off between SAR resolution, vegetation cover properties, land cover data resolution, and accuracy of the methodology. By using both X and L-band acquisitions, this study achieved a first analysis of the flexibility of the methodology for short and long wavelengths at HH polarization. Nevertheless, the reliability of the algorithm for the application to intermediate (e.g. C-band) wavelengths and different polarizations has yet to be investigated. In this framework, the dual polarization (VV and VH) acquisitions made available free of charge by the 10^1 m spatial resolution C-band Sentinel-1 constellation (launched in 2016) provide a relevant dataset for future testing.

6. Conclusions

This paper presented a novel algorithm which makes use of probability binning and fuzzy logic for the mapping of flooded vegetation from single SAR acquisitions and routinely available ancillary datasets. The proposed automatic data parsimonious approach, in which ancillary data can be prepared off-line to limit computational time, can enable near-real time analysis to support emergency management and complement the existing capabilities of flood detection in non-obstructed and urban areas.

The algorithm was tested on three fine resolution images acquired during the 2011 flood event in the Condamine-Balonne catchment (Australia) and its accuracy evaluated using fine resolution optical data. The OA was higher than 80% for all the images; the PA for the class water had an average increase of 84% when adding the analysis of flooded vegetation to the analysis of open water areas only. Possible causes of inaccuracy, the limitations of the proposed methodology, and a number of areas for future developments have been discussed. Despite the impossibility of fully separating backscatter effects caused by flooding and vegetation backscatter spatial heterogeneity using single acquisitions, the results shown in this paper encourage further testing of the proposed methodology and extension of the algorithm to incorporate previous experience and other datasets.

Author contributions

Stefania Grimaldi: Designed the methodology; collected the data; contributed to the implementation of the Matlab code and to the analysis of the results; wrote the paper.

Jin Xu: Contributed to the implementation of the Matlab code and to the analysis of the results.

Yuan Li: Reviewed the methodology, discussed the results and commented on the manuscript.

Valentijn R.N. Pauwels: Acquired the funding and supervised the project. Provided critical feedback and helped shape the research, analysis and manuscript.

Jeffrey P. Walker: Acquired the funding and supervised the project. Provided critical feedback and helped shape the research, analysis and manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Stefania Grimaldi is funded through the Bushfire and Natural Hazards Collaborative Research Centre grant “Improving flood forecast skill using remote sensing data”. The authors would like to acknowledge Geoscience Australia for providing the ALOS PALSAR image, and the SPOT5-derived flood extent layer. COSMO-SkyMed Image Credit Italian Space Agency (ASI). Finally, the authors would like to acknowledge e-GEOS for their timely support during the pre-processing of the COSMO-SkyMed images.

References

ABARES, 2010. In: A.B.o.A.a.R.E.A. Sciences (Ed.), The Australian Land Use and Management (ALUM) Classification Version 7, Detailed. Australian Bureau of Agricultural and Resource Economics and Sciences, Canberra (AUS).
Ajmar, A., Boccardo, P., Broglia, M., Kucera, J., Giulio-Tonolo, F., Wania, A., 2017. Response to flood events: the role of satellite-based emergency mapping and the experience of the Copernicus emergency management service. In: S.M.a.F.B, Molinari, D. (Eds.), Flood Damage Survey and Assessment. John Wiley & Sons, Inc, pp. 213–228.
Arthington, A.H., Balcombe, S.R., 2011. Extreme flow variability and the 'boom and bust' ecology of fish in arid-zone floodplain rivers: a case history with implications for environmental flows, conservation and management. Ecohydrology 4, 708–720.
Bartos, M., 2018. Pysheds: a fast, open-source digital elevation model processing library. In Hatari Labshttps://github.com/mdbartos/pysheds.
Bioresita, F., Puissant, A., Stumpf, A., Malet, J.-P., 2018. A method for automatic and rapid mapping of water surfaces from sentinel-1 imagery. Remote Sens. 10, 217.
Bovolo, F., Bruzzone, L., 2005. A detail-preserving scale-driven approach to change detection in multitemporal SAR images. IEEE Trans. Geosci. Remote Sens. 43, 2963–2972.
Bracaglia, M., Ferrazzoli, P., Guerriero, L., 1995a. A fully polarimetric multiple scattering model for agricultural fields. In: Geoscience and Remote Sensing Symposium, 1995. IGARSS '95. 'Quantitative Remote Sensing for Science and Applications', International. vol. 1332. pp. 1339.
Bracaglia, M., Ferrazzoli, P., Guerriero, L., 1995b. A fully polarimetric multiple scattering model for crops. Remote Sens. Environ. 54, 170–179.
Brisco, B., 2015. Mapping and monitoring surface water and wetlands with synthetic aperture radar. In: Remote Sensing of Wetlands: Applications and Advances, pp. 119–136.
Brisco, B., Schmitt, A., Murnaghan, K., Kaya, S., Roth, A., 2013. SAR polarimetric change detection for flooded vegetation. International Journal of Digital Earth 6, 103–114.
Brisco, B., Shelat, Y., Murnaghan, K., Montgomery, J., Fuss, C., Olthof, I., Hopkinson, C., Deschamps, A., Poncos, V., 2019. Evaluation of C-band SAR for identification of flooded vegetation in emergency response products. Can. J. Remote. Sens. 45, 73–87.
Canisius, F., Brisco, B., Murnaghan, K., Van Der Kooij, M., Keizer, E., 2019. SAR backscatter and InSAR coherence for monitoring wetland extent, flood pulse and vegetation: a study of the Amazon lowland. Remote Sens. 11, 720.
Cao, H., Zhang, H., Wang, C., Zhang, B., 2019. Operational flood detection using sentinel-1 SAR data over large areas. Water 11, 786.
Chaabani, C., Chini, M., Abdelfattah, R., Hostache, R., Chokmani, K., 2018. Flood Mapping in a Complex Environment Using Bistatic TanDEM-X/TerraSAR-X InSAR Coherence.
Chini, M., Pulvirenti, L., Pierdicca, N., Guerriero, L., 2011. Multi-temporal Segmentation of Cosmo-SkyMed SAR Data for Flood Monitoring.
Chini, M., Hostache, R., Giustarini, L., Matgen, P., 2017. A hierarchical split-based approach for parametric thresholding of SAR images: flood inundation as a test case. In: IEEE Transactions on Geoscience and Remote Sensing, pp. 1–14.
Chini, M., Pelich, R., Pulvirenti, L., Pierdicca, N., Hostache, R., Matgen, P., 2019. Sentinel-1 InSAR coherence to detect floodwater in urban areas: Houston and Hurricane Harvey as a test case. Remote Sens. 11, 107.
Cian, F., Marconcini, M., Ceccato, P., 2018. Normalized Difference Flood Index for rapid flood mapping: taking advantage of EO big data. Remote Sens. Environ. 209, 712–730.
Clement, M.A., Kilsby, C.G., Moore, P., 2018. Multi-temporal synthetic aperture radar flood mapping using change detection. Journal of Flood Risk Management 11 (2), 152–168. https://doi.org/10.1111/jfr3.12303.
Cohen, J., Riihimäki, H., Pulliainen, J., Lemmetyinen, J., Heilimo, J., 2016. Implications of boreal forest stand characteristics for X-band SAR flood mapping accuracy. Remote Sens. Environ. 186, 47–63.
CRED, UNISDR, 2015. The Human Cost of Weather-related Disasters 1995–2015. (p. 30).
D' Addabbo, A., Refice, A., Pasquariello, G., Lovergine, F.P., Capolongo, D., Manfreda, S., 2016. A Bayesian network for flood detection combining SAR imagery and ancillary data. IEEE Trans. Geosci. Remote Sens. 54, 3612–3625.
Dasgupta, A., Grimaldi, S., Ramsankaran, R., Pauwels, V., Walker, J.P., Chini, M., Hostache, R., Matgen, P., 2018a. Flood mapping using synthetic aperture radar sensors from local to global scales. In: Global Flood Hazard. American Geophysical Union (AGU), Washington, DC, USA, pp. 55–77.
Dasgupta, A., Grimaldi, S., Ramsankaran, R.A.A.J., Pauwels, V.R.N., Walker, J.P., 2018b. Towards operational SAR-based flood mapping using neuro-fuzzy texture-based approaches. Remote Sens. Environ. 215, 313–329.
Di Baldassarre, G., Schumann, G., Bates, P.D., 2009. A technique for the calibration of hydraulic models using uncertain satellite observations of flood extent. J. Hydrol. 367, 276–282.
Donchyts, G., Schellekens, J., Winsemius, H., Eisemann, E., van de Giesen, N., 2016. A 30 m resolution surface water mask including estimation of positional and thematic differences using Landsat 8, SRTM and OpenStreetMap: a case study in the Murray-Darling Basin, Australia. Remote Sens. 8, 386.
Dutta, D., Herath, S., Musiake, K., 2003. A mathematical model for flood loss estimation. J. Hydrol. 277, 24–49.
Eco Logical Australia, 2016. Vegetation of the Barwon-Darling and Condamine-Balonne Floodplain Systems of New South Wales: Mapping and Survey of Plant Community Types. Prepared for Murray-Darling Basin Authority. pp. 181.
Engheta, N., Elachi, C., 1982. Radar scattering from a diffuse vegetation layer over a smooth surface. IEEE Trans. Geosci. Remote Sens. 212–216.
Falter, D., Schröter, K., Dung, N.V., Vorogushyn, S., Kreibich, H., Hundecha, Y., Apel, H., Merz, B., 2015. Spatially coherent flood risk assessment based on long-term continuous simulation with a coupled model chain. J. Hydrol. 524, 182–193.
Ferrazzoli, P., Guerriero, L., 1995. Radar Sensitivity to Tree Geometry and Woody Volume: A Model Analysis.
Fritz, S., See, L., Perger, C., McCallum, I., Schill, C., Schepaschenko, D., Duerauer, M., Karner, M., Dresel, C., Laso-Bayas, J.-C., 2017. A global dataset of crowdsourced land cover and land use reference data. Scientific data 4, 170075.
Frost, V.S., Stiles, J.A., Shanmugam, K.S., Holtzman, J.C., Smith, S.A., 1981. An adaptive filter for smoothing noisy radar images. Proc. IEEE 69, 133–135.
Gallant, J.C., Dowling, T.I., 2003. A multiresolution index of valley bottom flatness for mapping depositional areas. Water Resour. Res. 39 (n/a-n/a).
Gallant, J.C., Wilson, N., Dowling, T.I., Read, A., Inskeep, C., 2011. In: G. Australia (Ed.), 1 Second SRTM Derived Products User Guide. Commonwealth of Australia (Geoscience Australia).
Gao, Y., Walker, J.P., Ye, N., Panciera, R., Monerris, A., Ryu, D., Rüdiger, C., Jackson, T.J., 2018. Evaluation of the tau–omega model for passive microwave soil moisture retrieval using SMAPEx datasets. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 11, 888–895.
Geoscience Australia, 2010. Dynamic land cover dataset. Bioregional assessment source dataset. In: Giustarini, L., Hostache, R., Kavetski, D., Chini, M., Corato, G., Schlaffer, S., Matgen, P. (Eds.), 2016. Probabilistic Flood Mapping Using Synthetic Aperture Radar Data. IEEE Transactions on Geoscience and Remote Sensing 54. pp. 6958–6969.
Giustarini, L., Hostache, R., Matgen, P., Schumann, G.J.P., Bates, P.D., Mason, D.C., 2013. A change detection approach to flood mapping in urban areas using TerraSAR-X. IEEE Trans. Geosci. Remote Sens. 51, 2417–2430.
Giustarini, L., Vernieuwe, H., Verwaeren, J., Chini, M., Hostache, R., Matgen, P., Verhoest, N.E.C., De Baets, B., 2015. Accounting for image uncertainty in SAR-based flood mapping. Int. J. Appl. Earth Obs. Geoinf. 34, 70–77.
Giustarini, L., Hostache, R., Kavetski, D., Chini, M., Corato, G., Schlaffer, S., Matgen, P., 2016. Probabilistic Flood Mapping Using Synthetic Aperture Radar Data. IEEE Trans. Geosci. Remote Sens. 54 (12), 6958–6969. https://doi.org/10.1109/TGRS.2016.2592951.
Grimaldi, S., Li, Y., Pauwels, V.R.N., Walker, J.P., 2016. Remote sensing-derived water extent and level to constrain hydraulic flood forecasting models: opportunities and challenges. Surv. Geophys. 37, 977–1034.
Henderson, F.M., Lewis, A.J., 2008. Radar detection of wetland ecosystems: a review. Int. J. Remote Sens. 29, 5809–5835.
Hess, L.L., Melack, J.M., Simonett, D.S., 1990. Radar detection of flooding beneath the forest canopy: a review. Int. J. Remote Sens. 11, 1313–1325.
Hess, L.L., Melack, J.M., Novo, E.M., Barbosa, C.C., Gastil, M., 2003. Dual-season mapping of wetland inundation and vegetation for the central Amazon basin. Remote Sens. Environ. 87, 404–428.
Hostache, R., Matgen, P., Wagner, W., 2012. Change detection approaches for flood extent mapping: how to select the most adequate reference image from online archives? Int. J. Appl. Earth Obs. Geoinf. 19, 205–213.
Hostache, R., Chini, M., Giustarini, L., Neal, J., Kavetski, D., Wood, M., Corato, G., Pelich, R.-M., Matgen, P., 2018. Near-real-time assimilation of SAR-derived flood maps for improving flood forecasts. Water Resour. Res. 54, 5516–5535.
Huang, C., Nguyen, B.D., Zhang, S., Cao, S., Wagner, W., 2017. A comparison of terrain indices toward their ability in assisting surface water mapping from sentinel-1 data. ISPRS Int. J. Geo Inf. 6, 140.
Karmakar, G.C., Dooley, L.S., 2002. A generic fuzzy rule based image segmentation algorithm. Pattern Recogn. Lett. 23, 1215–1227.
Kasischke, E.S., Bourgeau-Chavez, L.L., Rober, A.R., Wyatt, K.H., Waddington, J.M., Turetsky, M.R., 2009. Effects of soil moisture and water depth on ERS SAR backscatter measurements from an Alaskan wetland complex. Remote Sens. Environ. 113, 1868–1873.
Koks, E., Thissen, M., Alfieri, L., De Moel, H., Feyen, L., Jongman, B., Aerts, J., 2019. The macroeconomic impacts of future river flooding in Europe. Environ. Res. Lett.(1748-9326) 14 (8) p. 084042, JRC115578.
Landuyt, L., Van Wesemael, A., Schumann, G.J., Hostache, R., Verhoest, N.E.C., Van Coillie, F.M.B., 2018. Flood mapping based on synthetic aperture radar: an assessment of established approaches. IEEE Trans. Geosci. Remote Sens. 1–18.
Lang, M.W., Townsend, P.A., Kasischke, E.S., 2008. Influence of incidence angle on detecting flooded forests using C-HH synthetic aperture radar data. Remote Sens. Environ. 112, 3898–3907.
Lavalle, M., Wright, T., 2009. Absolute Radiometric and Polarimetric Calibration of ALOS PALSAR Products. Document Issue (1), Revision (3).
Leblanc, M., Tweed, S., Van Dijk, A., Timbal, B., 2012. A review of historic and future hydrological changes in the Murray-Darling Basin. Glob. Planet. Chang. 80, 226–246.
Lee, J., Jen-Hung, W., Ainsworth, T.L., Kun-Shan, C., Chen, A.J., 2009. Improved sigma filter for speckle filtering of SAR imagery. Geoscience and Remote Sensing, IEEE Transactions on 47, 202–213.
Li, Y., Martinis, S., Plank, S., Ludwig, R., 2018. An automatic change detection approach for rapid flood mapping in Sentinel-1 SAR data. Int. J. Appl. Earth Obs. Geoinf. 73, 123–135.
Li, Y., Martinis, S., Wieland, M., 2019. Urban flood mapping with an active self-learning convolutional neural network based on TerraSAR-X intensity and interferometric coherence. ISPRS J. Photogramm. Remote Sens. 152, 178–191.
Lin, Y.N., Yun, S.-H., Bhardwaj, A., Hill, E.M., 2019. Urban flood detection with sentinel-1 multi-temporal synthetic aperture radar (SAR) observations in a Bayesian framework: a case study for Hurricane Matthew. Remote Sens. 11, 1778.
Long, S., Fatoyinbo, T.E., Policelli, F., 2014. Flood extent mapping for Namibia using change detection and thresholding with SAR. Environ. Res. Lett. 9, 035002.
Lymburner, L., Tan, P., Mueller, N., Thackway, R., Lewis, A., Thankappan, M., Randall, L., Islam, A., Senarath, U., 2011. Dynamic Land Cover Dataset. Geoscience Australia, Canberra.

Marquardt, D.W., 1963. An algorithm for least-squares estimation of nonlinear parameters. J. Soc. Ind. Appl. Math. 11, 431–441.
Martinis, S., Rieke, C., 2015. Backscatter analysis using multi-temporal and multi-frequency SAR data in the context of flood mapping at River Saale, Germany. Remote Sens. 7, 7732.
Martinis, S., Twele, A., 2010. A hierarchical spatio-temporal Markov model for improved flood mapping using multi-temporal X-band SAR data. Remote Sens. 2, 2240.
Martinis, S., Twele, A., Voigt, S., 2009. Towards operational near real-time flood detection using a split-based automatic thresholding procedure on high resolution TerraSAR-X data. Natural Hazards and Earth System Science 9, 303–314.
Martinis, S., Kersten, J., Twele, A., 2015. A fully automated TerraSAR-X based flood service. ISPRS J. Photogramm. Remote Sens. 104, 203–212.
Mason, D.C., Giustarini, L., Garcia-Pintado, J., Cloke, H.L., 2014. Detection of flooded urban areas in high resolution Synthetic Aperture Radar images using double scattering. Int. J. Appl. Earth Obs. Geoinf. 28, 150–159.
Matgen, P., Hostache, R., Schumann, G., Pfister, L., Hoffmann, L., Savenije, H.H.G., 2011. Towards an automated SAR-based flood monitoring system: lessons learned from two case studies. Physics and Chemistry of the Earth, Parts A/B/C 36, 241–252.
Matgen, P., Giustarini, L., Chini, M., Hostache, R., Wood, M., Schlaffer, S., 2016. Creating a water depth map from SAR flood extent and topography data. In: 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 7635–7638.
Molinari, D., Scorzini, A.R., Gallazzi, A., Ballio, F., 2019. AGRIDE-c, a conceptual model for the estimation of flood damage to crops: development and implementation. Nat. Hazards Earth Syst. Sci. Discuss. 2019, 1–24.
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 25 years of Landsat imagery across Australia. Remote Sens. Environ. 174, 341–352.
Nico, G., Pappalepore, M., Pasquariello, G., Refice, A., Samarelli, S., 2000. Comparison of SAR amplitude vs. coherence flood detection methods - a GIS application. Int. J. Remote Sens. 21, 1619–1631.
Nobre, A.D., Cuartas, L.A., Hodnett, M., Rennó, C.D., Rodrigues, G., Silveira, A., Waterloo, M., Saleska, S., 2011. Height Above the Nearest Drainage – a hydrologically relevant new terrain model. J. Hydrol. 404, 13–29.
Nobre, A.D., Cuartas, L.A., Momo, M.R., Severo, D.L., Pinheiro, A., Nobre, C.A., 2016. HAND contour: a new proxy predictor of inundation extent. Hydrol. Process. 30, 320–333.
O'Callaghan, J.F., Mark, D.M., 1984. The extraction of drainage networks from digital elevation data. Computer Vision, Graphics, and Image Processing 28, 323–344.
Oliver, C., Quegan, S., 2004. Understanding Synthetic Aperture Radar Images. SciTech Publishing, London (UK).
Pal, S.K., Rosenfeld, A., 1988. Image enhancement and thresholding by optimization of fuzzy compactness. Pattern Recogn. Lett. 7, 77–86.
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.
Pierdicca, N., Chini, M., Pulvirenti, L., Macina, F., 2008. Integrating physical and topographic information into a fuzzy scheme to map flooded area by SAR. Sensors 8, 4151.
Pierdicca, N., Pulvirenti, L., Chini, M., Guerriero, L., Candela, L., 2013. Observing floods from space: experience gained from COSMO-SkyMed observations. Acta Astronautica 84, 122–133.
Pierdicca, N., Pulvirenti, L., Chini, M., 2018. Flood mapping in vegetated and urban areas and other challenges: models and methods. In: Refice, A., D'Addabbo, A., Capolongo, D. (Eds.), Flood Monitoring Through Remote Sensing. Springer International Publishing, Cham, pp. 135–179.
Plank, S., Jüssi, M., Martinis, S., Twele, A., 2017. Mapping of flooded vegetation by means of polarimetric Sentinel-1 and ALOS-2/PALSAR-2 imagery. Int. J. Remote Sens. 38, 3831–3850.
Pulliainen, J.T., Heiska, K., Hyyppa, J., Hallikainen, M.T., 1994. Backscattering properties of boreal forests at the C- and X-bands. IEEE Trans. Geosci. Remote Sens. 32, 1041–1050.
Pulvirenti, L., Chini, M., Pierdicca, N., Guerriero, L., 2011a. Combined use of electromagnetic scattering models, fuzzy logic and mathematical morphology for flood mapping using Cosmo-SkyMed data. In: 2011 IEEE International Geoscience and Remote Sensing Symposium, pp. 4154–4156.
Pulvirenti, L., Chini, M., Pierdicca, N., Guerriero, L., Ferrazzoli, P., 2011b. Flood monitoring using multi-temporal COSMO-SkyMed data: image segmentation and signature interpretation. Remote Sens. Environ. 115, 990–1002.
Pulvirenti, L., Pierdicca, N., Chini, M., Guerriero, L., 2011c. An algorithm for operational flood mapping from Synthetic Aperture Radar (SAR) data using fuzzy logic. Nat. Hazards Earth Syst. Sci. 11, 529.
Pulvirenti, L., Pierdicca, N., Chini, M., Guerriero, L., 2013. Monitoring flood evolution in vegetated areas using COSMO-SkyMed data: the Tuscany 2009 case study. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 6, 1807–1816.
Pulvirenti, L., Chini, M., Pierdicca, N., Boni, G., 2015. Use of SAR Data for Detecting Floodwater in Urban and Agricultural Areas: The Role of the Interferometric Coherence.
Queensland Reconstruction Authority, 2013. Queensland Floodplain Assessment Overlay. State of Queensland (Department of Natural Resources and Mines). https://www.data.qld.gov.au/dataset/queensland-floodplain-assessment-overlay.
Rennó, C.D., Nobre, A.D., Cuartas, L.A., Soares, J.V., Hodnett, M.G., Tomasella, J., Waterloo, M.J., 2008. HAND, a new terrain descriptor using SRTM-DEM: mapping terra-firme rainforest environments in Amazonia. Remote Sens. Environ. 112, 3469–3481.
Richards, J.A., Sun, G., Simonett, D.S., 1987a. L-band radar backscatter modeling of forest stands. IEEE Trans. Geosci. Remote Sens. GE-25, 487–498.
Richards, J.A., Woodgate, P.W., Skidmore, A.K., 1987b. An explanation of enhanced radar backscattering from flooded forests. Int. J. Remote Sens. 8, 1093–1100.
Rignot, E.J., Van Zyl, J.J., 1993. Change detection techniques for ERS-1 SAR data. IEEE Trans. Geosci. Remote Sens. 31, 896–906.
Roederer, M., Treister, A., Moore, W., Herzenberg, L.A., 2001. Probability binning comparison: a metric for quantitating univariate distribution differences. Cytometry Part A 45, 37–46.
Sabaghy, S., Walker, J.P., Renzullo, L.J., Jackson, T.J., 2018. Spatially enhanced passive microwave derived soil moisture: capabilities and opportunities. Remote Sens. Environ. 209, 551–580.
Samadi, M., 2018. HANDtool. https://github.com/mahmoodsamadi/HANDTool.
Sanyal, J., Lu, X.X., 2004. Application of remote sensing in flood management with special reference to monsoon Asia: a review. Nat. Hazards 33, 283–301.
Schlaffer, S., Chini, M., Giustarini, L., Matgen, P., 2017. Probabilistic mapping of flood-induced backscatter changes in SAR time series. Int. J. Appl. Earth Obs. Geoinf. 56, 77–87.
Schumann, G., Di Baldassarre, G., 2010. The direct use of radar satellites for event-specific flood risk mapping. Remote Sensing Letters 1, 75–84.
Schumann, G.J.P., Domeneghetti, A., 2016. Exploiting the proliferation of current and future satellite observations of rivers. Hydrol. Process. 30, 2891–2896.
Schumann, G.J.P., Moller, D.K., 2015. Microwave remote sensing of flood inundation. Physics and Chemistry of the Earth, Parts A/B/C 83–84, 84–95.
Schwanghart, W., Scherler, D., 2014. Short communication: TopoToolbox 2 – MATLAB-based software for topographic analysis and modeling in Earth surface sciences. Earth Surf. Dynam. 2, 1–7.
Seber, G.A., Wild, C.J., 2003. Nonlinear Regression. John Wiley & Sons, Hoboken, New Jersey.
Shapiro, S.S., Wilk, M.B., 1965. An analysis of variance test for normality (complete samples). Biometrika 52, 591–611.
Shen, X., Anagnostou, E.N., Allen, G.H., Brakenridge, G.R., Kettner, A.J., 2019. Near-real-time non-obstructed flood inundation mapping using synthetic aperture radar. Remote Sens. Environ. 221, 302–315.
Townsend, P.A., 2002a. Estimating forest structure in wetlands using multitemporal SAR. Remote Sens. Environ. 79, 288–304.
Townsend, P.A., 2002b. Relationships between forest structure and the detection of flood inundation in forested wetlands using C-band SAR. Int. J. Remote Sens. 23, 443–460.
Tsyganskaya, V., Martinis, S., Twele, A., Cao, W., Schmitt, A., Marzahn, P., Ludwig, R., 2016. A fuzzy logic-based approach for the detection of flooded vegetation by means of synthetic aperture radar data. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XLI-B7, 371–378.
Tsyganskaya, V., Martinis, S., Marzahn, P., Ludwig, R., 2018a. Detection of temporary flooded vegetation using sentinel-1 time series data. Remote Sens. 10, 1286.
Tsyganskaya, V., Martinis, S., Marzahn, P., Ludwig, R., 2018b. SAR-based detection of flooded vegetation–a review of characteristics and approaches. Int. J. Remote Sens. 39, 2255–2293.
Tsyganskaya, V., Martinis, S., Marzahn, P., 2019. Flood monitoring in vegetated areas using multitemporal sentinel-1 data: impact of time series features. Water 11, 1938.
Twele, A., Cao, W., Plank, S., Martinis, S., 2016. Sentinel-1-based flood mapping: a fully automated processing chain. Int. J. Remote Sens. 37, 2990–3004.
Ulaby, F.T., Moore, R.K., Fung, A.K., 1986. Microwave Remote Sensing: Active and Passive. Addison-Wesley Publishing Company, Advanced Book Program/World Science Division.
Voormansik, K., Praks, J., Antropov, O., Jagomägi, J., Zalite, K., 2014. Flood mapping with TerraSAR-X in forested regions in Estonia. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 7, 562–577.
Wang, Y., Davis, F.W., 1997. Decomposition of polarimetric synthetic aperture radar backscatter from upland and flooded forests. Int. J. Remote Sens. 18, 1319–1332.
Wang, T., Tong, C., 2018. An analytical model of double scattering for SAR imaging of urban structures. Optik 171, 484–491.
Wang, Y., Hess, L.L., Filoso, S., Melack, J.M., 1995. Understanding the radar backscattering from flooded and nonflooded Amazonian forests: results from canopy backscatter modeling. Remote Sens. Environ. 54, 324–332.
Wang, A., Grimaldi, S., Shaadman, S., Li, Y., Pauwels, V., Walker, J.P., 2018. Evaluation of TanDEM-X and DEM-H digital elevation models over the Condamine-Balonne catchment (Australia). In: Hydrology and Water Resources Symposium (HWRS 2018): Water and Communities. Engineers Australia, Melbourne, pp. 989–1003.
Westerhoff, R.S., Kleuskens, M.P.H., Winsemius, H.C., Huizinga, H.J., Brakenridge, G.R., Bishop, C., 2013. Automated global water mapping based on wide-swath orbital synthetic-aperture radar. Hydrol. Earth Syst. Sci. 17, 651–663.
Wohlfart, C., Winkler, K., Wendleder, A., Roth, A., 2018. TerraSAR-X and wetlands: a review. Remote Sens. 10, 916.
Woodhouse, I.H., 2005. Introduction to Microwave Remote Sensing. CRC Press, Florida (USA).
Yamazaki, D., Ikeshima, D., Sosa, J., Bates, P.D., Allen, G.H., Pavelsky, T.M., 2019. MERIT hydro: a high-resolution global hydrography map based on latest topography dataset. Water Resour. Res. 55, 5053–5073.
Zadeh, L.A., 1965. Fuzzy sets. Inf. Control. 8, 338–353.
Zhu, L., Walker, J.P., Ye, N., Rüdiger, C., 2019. Roughness and vegetation change detection: a pre-processing for soil moisture retrieval from multi-temporal SAR imagery. Remote Sens. Environ. 225, 93–106.
Zwenzner, H., Voigt, S., 2009. Improved estimation of flood parameters by combining space based SAR data with very high resolution digital elevation data. Hydrol. Earth Syst. Sci. 13, 567–576.