



Towards operational SAR-based flood mapping using neuro-fuzzy texture-based approaches



Antara Dasgupta^{a,b,c,*}, Stefania Grimaldi^c, R.A.A.J. Ramsankaran^b, Valentijn R.N. Pauwels^c, Jeffrey P. Walker^c

^a IITB-Monash Research Academy, Mumbai 400076, Maharashtra, India

^b Hydro-Remote Sensing Applications (H-RSA) Group, Department of Civil Engineering, Indian Institute of Technology Bombay, Mumbai 400076, Maharashtra, India

^c Water Group, Department of Civil Engineering, Monash University, Melbourne, VIC 3800, Australia

ARTICLE INFO

Keywords:

Flood mapping
SAR
Texture optimization
GLCM
Neuro-fuzzy classification
ANFIS
Flood extent
COSMO-SkyMed

ABSTRACT

Synthetic Aperture Radar (SAR) data are currently the most reliable resource for flood monitoring, though still subject to various uncertainties, which can be objectively represented with probabilistic flood maps. Moreover, the growing number of SAR satellites has increased the likelihood of observing a flood event from space through at least a single SAR image, but generalized methods for flood classification independent of sensor characteristics need to be developed, to fully utilize these images for disaster management. Consequently, a neuro-fuzzy flood mapping technique is proposed for texture-enhanced single SAR images. Accordingly, any SAR image is first processed to generate second-order statistical textures, which are subsequently optimized using a dimensionality reduction technique. The flood and non-flood classes are then modelled within a fuzzy inference system using Gaussian curves. Parameterization is achieved by training a neural network on the image through user-defined polygons. The results of the optimized texture-based neuro-fuzzy classification were compared against the performance of the SAR image alone and that of SAR enhanced with randomly selected texture features. This approach was tested for a COSMO-SkyMed SAR image at two validation sites, for which high resolution aerial photographs were available. An overall accuracy assessment using reliability diagrams demonstrated a reduction of 54.2% in the Weighted Root Mean Squared Error (WRMSE) values compared to the stand-alone use of SAR. WRMSE values estimated for the proposed method varied from 0.027 to 0.196. A fuzzy validation exercise was also proposed to account for the uncertainty in manual flood identification from aerial photography, resulting in fuzzy spatial similarity values ranging from 0.67 to 0.92, with higher values representing better performance. Results suggest that the proposed approach has demonstrated potential to improve operational SAR-based flood mapping.

1. Introduction

Floods are widely accepted as the most ubiquitous of all natural disasters and an alarming increase in their frequency has been evident for the last few decades (Schumann et al., 2009b). The global socio-economic impacts of flooding are likely to increase as a result of climate change impacts and population growth (CRED and UNISDR, 2015). As satellite data provide a cost-effective, near real-time solution for operational flood mapping, it becomes imperative to exploit its full potential for flood management (Giustarini et al., 2016). Furthermore, satellite-derived flood extent maps can improve flood forecasting skill by allowing more accurate hydraulic model calibration and validation (Grimaldi et al., 2016; Wood et al., 2016), through direct assimilation of flood extent (Hostache et al., 2015; Lai et al., 2014) or spatially

distributed water levels derived using digital elevation models (DEM) (García-Pintado et al., 2013, 2014; Giustarini et al., 2011; Hostache et al., 2009, 2010; Lai and Monnier, 2009; Mason et al., 2012; Matgen et al., 2010).

Synthetic Aperture Radar (SAR) data have proven to be the most useful for the spatial characterization of floods, due to their all-weather, all-day imaging capabilities (Smith, 1997). Inundated pixels often appear dark on SAR images, as specular reflection reflects the radar signal away from the sensor, resulting in low recorded backscatter (Hostache et al., 2009). This usually results in a high land-water contrast and so several flood mapping approaches utilize this characteristic, including but not limited to; radiometric thresholding (Hostache et al., 2006), automatic thresholding (Chini et al., 2017; Twele et al., 2016), region-growing (Boni et al., 2016), object oriented classification (Pradhan

* Corresponding author at: IITB-Monash Research Academy, Mumbai 400076, Maharashtra, India.
E-mail address: antara.dasgupta@monash.edu (A. Dasgupta).

et al., 2016), pixel based supervised classification (Voorrips et al., 2014), and change detection (Giustarini et al., 2013; Long et al., 2014).

Most of the aforementioned approaches rely on a clear separation between land and water pixels achieved at the classification boundary, which is improbable in practice due to an overlap in the class distributions. As shown by O'Grady et al. (2014), even for images exhibiting significant class separability, > 2% of pixels will be misclassified even if the central data value is accurately identified for binarization. Furthermore, SAR images are affected by speckle noise, which causes random backscatter variations within homogeneous image features, making SAR-based classification significantly more challenging (Giustarini et al., 2015). Moreover, the SAR imaging geometry at the time of acquisition, especially the local incidence angle, substantially contributes to backscatter variability (O'Grady et al., 2013, 2014). Factors like submerged vegetation, wind or rain, which roughen open water surfaces and thus alter the backscattering behaviour, may contribute to under detection. Conversely, dark or smooth urban surfaces such as asphalt and concrete, which generate low backscatter similar to water, may lead to over detection (Martinis et al., 2015).

Given that the sources of uncertainty are numerous, stakeholders stand to benefit from a clear representation of these on the resulting flood maps. Probabilistic flood maps provide a unique opportunity for an objective characterization of the various uncertainties associated with SAR-based flood mapping. Multi-algorithm map ensembles, which indicate the possibility of inundation at each pixel, have been proposed to account for the subjectivity in the choice of an appropriate classification algorithm (Schumann et al., 2009b; Schumann and Di Baldassarre, 2010). Such maps have also proved useful for fuzzy model calibration, as the information content of SAR imagery could be extracted while accounting for observational uncertainties (Di Baldassarre et al., 2009). However, the number of ensemble members and the specific algorithms chosen for this exercise could still be subjective (Giustarini et al., 2016).

Similar to the ensemble mapping technique, Schumann et al. (2008) proposed the use of multiple equally plausible thresholds for the land-water interface. Building on this, the merit of acknowledging the uncertainty in the SAR thresholding for flood model calibration was demonstrated (Schumann et al., 2014), with the threshold varied across the whole range of plausible backscatter values to ensure objectivity and to optimize information extraction from the SAR image. Each model simulated binary flood map (water and no water) resulting from a particular parameter set was then compared to all the binary SAR-based maps generated. Although this approach is theoretically promising for model calibration, it remains a computationally intensive exercise which may be unsuitable for operational applications.

More recent studies have used the Bayesian principles of conditional probability, where each pixel is assigned a flood probability based on its backscatter value (Giustarini et al., 2016). Here the probability distributions for flood and non-flood classes were first estimated from the empirical histogram of SAR backscatter values and parameterized as a mixture of two Gaussian functions using the Levenberg–Marquardt algorithm (Marquardt, 1963). The reliability statistic used in the study exhibited a keen sensitivity to the prior probabilities assumed, though authors showed that using a prior value of 0.5 was mostly acceptable. Subsequent research on this method used a time-series of SAR data to parameterize the probability distributions (Schlaffer et al., 2017). The results showed that the varying imaging geometries, the incidence angle in particular, had a large impact on class separability. This implies that reliance on SAR backscatter alone is unable to account for uncertainties contributed by wind and rain conditions, mixed land covers, or water lookalike surfaces. However, these studies established that a reasonable probabilistic definition of flooding at each pixel was possible by modelling the backscatter distribution using some non-linear regression technique, if complementary information was available or priors were accurately estimated.

Ideally, the inclusion of ancillary datasets within this Bayesian framework could eliminate one or more sources of errors in SAR-based flood extraction (D'Addabbo et al., 2016). Integration with a detailed land cover map for example, allows differentiation between water and water lookalike regions in the SAR image (Pierdicca et al., 2008). Several approaches have been proposed for the integration of these separate information layers, from Bayesian networks (Refice et al., 2014) to fuzzy inference systems (Pulvirenti et al., 2013, 2014). A key limitation of such approaches is the assumption that suitable supporting datasets are available for the area of interest, which is often inaccurate especially for developing regions. Moreover, the present cohort of fuzzy rule-based approaches utilizes theoretical electromagnetic backscattering models for parameterization. Given that these are wavelength specific, they typically limit transferability of fuzzy approaches across the range of SAR satellites. Therefore, this study introduced a texture-based image enhancement approach to improve single image flood mapping, which can incorporate the spatial autocorrelation amongst pixel values to minimize the impact of sensor parameters.

Since texture can be derived from the SAR image, it also reduces the dependence on ancillary or complementary datasets. However, state-of-the-art texture based mapping approaches also struggle with the subjectivity in selecting application appropriate texture features, suitable window sizes, and optimal direction for identifying the feature of interest. These challenges currently significantly limit the use of texture in SAR based flood mapping (Di Baldassarre et al., 2011). Consequently, a SAR texture optimization technique is proposed in this paper to improve the utilization of texture in single image flood mapping and address these open research questions.

The optimized texture bands were considered alongside the SAR intensity image, within a neuro-fuzzy classifier to generate a fuzzy flood map. Gaussian membership functions were chosen to represent the backscatter distribution of each class, based on the image histogram as in the probabilistic mapping approaches (Giustarini et al., 2016; Schlaffer et al., 2017). However, using the neural network for a data driven parameter estimation of these membership functions removes the need for identification of suitable prior probability distributions. Training the classifier on the image to be processed, offers the additional advantage of accounting for image specific backscatter variability, caused by the reference incidence angle or wind effects.

Given a filtered SAR image, the ideal window size for texture estimation is first determined through semivariogram analysis. This is followed by an estimation of omnidirectional Grey Level Co-occurrence Matrices (GLCM) from which texture features were derived. An independent component analysis was then used to condense the maximum possible information into minimum bands, which were then added to the SAR image prior to classification. The class distributions were modelled as Gaussian functions within a fuzzy inference system, and parameterized using training data from the image itself.

The resulting maps were evaluated using aerial photographs through reliability diagrams, as well as a fuzzy validation exercise novel to flood mapping literature. The fuzzy map comparison accounts for the uncertainties in manual shoreline extraction for validation data as well. The classification performance of the SAR image with added optimized texture bands was compared against a SAR image without any texture addition and a SAR image with some randomly selected texture features added. Finally, a land-use specific analysis was conducted to assess the spatial variability of classifier performance, to facilitate an area appropriate choice of classifiers for flood mapping.

2. Study area and data

2.1. Study area

The Clarence Catchment of New South Wales, Australia, which spans an area of 22,700 km² was selected as the study site to test the mapping approach. Fig. 1 illustrates the geographic location of the

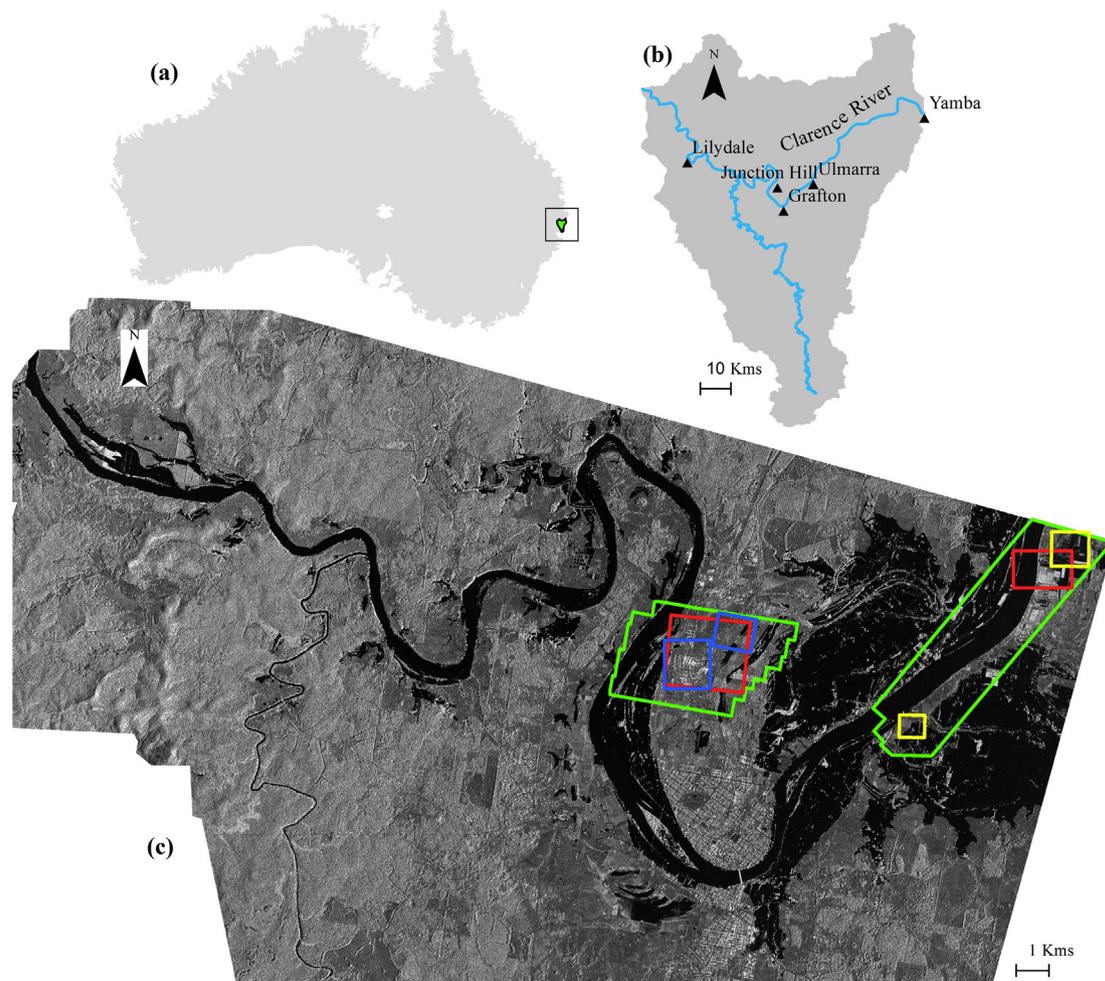


Fig. 1. Map displaying the location of Clarence catchment (a), and the main drainages and towns (b). The COSMO-SkyMed SAR image acquired on 12th Jan, 2011 is also shown (c), with the green polygons indicating the aerial photo coverage used for validation. The example subsets used in Fig. 6 are depicted in red, while those used in Figs. 9 and 14 are shown in blue and yellow respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

study area and the position of the validation targets, where high resolution airborne imagery were available. The land cover of the region is primarily dominated by grassland vegetation and agriculture. The Middle Clarence Catchment near Grafton is characterized by gentle slopes, although the Upper part has a highly undulating topography. The flood event under consideration occurred from January 10 to 17, 2011.

2.2. Data summary

The proposed approach was tested using a COSMO-SkyMed (CSK) X-band HH-Polarized image, acquired by the CSK-3 satellite in Stripmap HIMAGE mode at 3 m resolution on 12 January 2011 at 18:03 h (AEDT). The CSK Level 1D Georeferenced Terrain Corrected (GTC) product delivered as an 8-bit image of digital numbers was used in this study. The calibration process for the GTC product corrects for local incidence angle impacts using a DEM, by normalising the backscatter to a reference incidence angle which was 40° in this case (Italian Space Agency, 2009). The domain comprised of 74,056,858 pixels each having an area of 9 m^2 , bringing the total tile coverage to approximately 666.5 km^2 .

Fig. 2 shows the temporal position of the acquisition on the hydrograph, generated using hydrometric information available from the New South Wales Manly Hydraulics Laboratory. The high-resolution aerial photographs of Junction Hill and Ulmarra, were captured on

January 12, 2011, between 16:41 and 17:17 h and 17:17 to 17:39 h, respectively. The images had a spatial resolution of 10 cm, and were provided for this study by the NSW-LPI. Land cover information was extracted from the National Dynamic Land Cover Dataset distributed by Geoscience Australia at 250 m spatial resolution. Readers are referred to Lymburner et al. (2011) for a detailed description of this dataset.

3. Methodology

An overview of the proposed flood mapping approach is illustrated in Fig. 3. The reasons for choosing each processing step and its subsequent implementation are discussed at length in the following sections.

3.1. SAR preprocessing

The COSMO images were preprocessed using the Gamma Maximum-A-Posteriori (GMAP) filter which suppresses speckle noise while preserving edges and image texture, a property conducive for flood detection (Senthilnath et al., 2013). A window size of 3×3 was used as higher resolution SAR images are more susceptible to speckle noise, due to backscatter interference from neighboring pixels adding to sub-pixel interference. The GMAP filtered image is then used for texture analysis, and hereafter referred to as the SAR image.

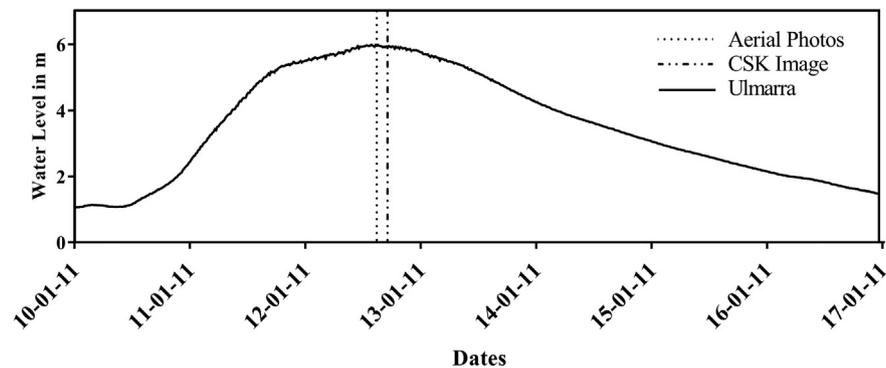


Fig. 2. Flood hydrograph recorded at the validation site Ulmarra (shown in Fig. 1), with the temporal positions of available aerial photos and radar remote sensing from COSMO-SkyMed data represented.

3.2. Texture analysis

Texture based approaches have the advantage that a single SAR image of the event can be used for their derivation and their extraction can be automated. Image texture can be defined as a measure of the transitional probabilities of pixel values, which can facilitate object identification in SAR data (Haack and Bechdol, 1999; He and Wang, 1991). As SAR images are rich in texture, it has often been utilized for flood identification (Pradhan et al., 2014; Schumann et al., 2009b; Senthilnath et al., 2013). However, most of the studies which utilize textural properties of SAR, have unfortunately failed to quantify their contribution to the overall improvement in flood mapping accuracy. Consequently, this study aims to explicitly assess the role of texture in flood delineation and quantify the maximum improvement possible through its inclusion.

Objectively selecting appropriate texture features for a particular case study is critical, as this can significantly impact the subsequent flood classification (Di Baldassarre et al., 2011; Schumann et al., 2009a, 2012; Schumann and Di Baldassarre, 2010). Moreover, SAR-based land cover mapping studies have demonstrated the merit of texture optimization methods to solve this problem (Balaguer et al., 2010; Berberoglu et al., 2000, 2007; Carr, 1996; Carr and De Miranda, 1998; De-yong et al., 2008; Franklin et al., 1996; Haack and Bechdol, 2000). However, the impact of optimized texture still needs to be investigated in the context of SAR-based flood mapping as the implementation of texture based methods is fairly empirical (Amitrano et al., 2018; Ouled Sghaier et al., 2018). As texture is direction and scale dependent, these parameter choices can also influence notably the classification, and so must be explicitly considered in any texture based mapping approach (Di Baldassarre et al., 2011; Franklin et al., 1996).

Statistical texture estimation approaches were chosen for this analysis as they utilize non-deterministic properties, governing the distribution of pixel value pairs. Second-order image statistics are also useful for SAR-based flood extent mapping, as the range of spatial autocorrelation for speckle noise is limited to the image resolution in this space (Ulaby et al., 1986). Grey Level Co-occurrence Matrices (GLCM), which can be interpreted as joint grey level probability density distributions or 2-D image histograms, were used for this study due to their low sensitivity to image contrast (Kuplich et al., 2005). It then follows that the GLCM-based texture features which were subsequently derived and used to enhance flood identification in this study, were also relatively insensitive to the land-water backscatter contrast.

This implies that the effect of wind related surface roughening, which usually hampers accurate flood mapping from SAR by significantly increasing the backscatter of open water surfaces, can largely be mitigated. Adding contrast insensitive texture features as additional information layers reduces the dependence on backscatter, and therefore minimizes the impact of wind-induced backscatter variations. As the patterns in the backscatter rather than the backscatter itself were

analysed, the probability of correctly classifying a wind roughened flood pixel was increased. This is a clear improvement over backscatter contrast dependent flood classification techniques such as histogram thresholding, which are completely unable to identify inundated pixels under windy conditions as the class distributions are no longer separable. Moreover, the use of a fuzzy mapping technique also allows for a clear expression of the uncertainty in flood detection; especially in the overlap between the two class distributions.

Each entry in the $n \times n$ GLCM indicates the number of co-occurrences of pixel value pairs at a specific lag distance in a given direction, where n is the number of grey levels in the image. For example, the 45° GLCM with one pixel lag for a binary image would record the number of times each combination of grey level pairs ($[0,0]$, $[0,1]$, $[1,0]$, $[1,1]$) appears in the image separated in the specified direction by one pixel distance. The optimum window size for GLCM calculation was estimated as the range of sensible intra-class variance through semivariogram analysis (Balaguer et al., 2010). Omnidirectional semivariogram curves were generated for both the flood and non-flood classes, by taking homogeneous subsets of size 400×400 . This step can be automated in the future as the semivariograms for different sensors and resolutions can be precomputed and used as a look up table, based on archived satellite data. According to this scenario, when a new satellite image becomes available the algorithm would select an appropriate window size for texture estimation based on the specific sensor characteristics.

GLCM matrices obtained in the previous step were used to estimate the second-order textures proposed by Haralick et al. (1973). Mean, variance, homogeneity, contrast, dissimilarity, entropy, angular second moment, and correlation were the co-occurrence measures retained for further optimization as they were least correlated. Direction-invariant texture information was obtained prior to optimization, by averaging the texture values in all eight directions. An Independent Component Transform (ICT) was used, to optimize the textural information and reduce the dimensionality.

As ICT assumes the errors to be of unit variance (white noise), the noise adjusted Principal Component Transform was used to whiten the noise and decorrelate it from the signal (Chica-Olmo and Abarca-Hernández, 2000). From the noise whitened data, ICT extracts the direction with the least-Gaussian distribution, and removes the data explained by this variable. A cost function implying non-Gaussianity, such as skewness or kurtosis, is iteratively maximized until the remaining dataset can be explained by statistically independent variables. The first three independent components which contained approximately 99% of the GLCM texture information, characterized by significant Eigen values (Fig. 4), were added to the SAR image by layer stacking.

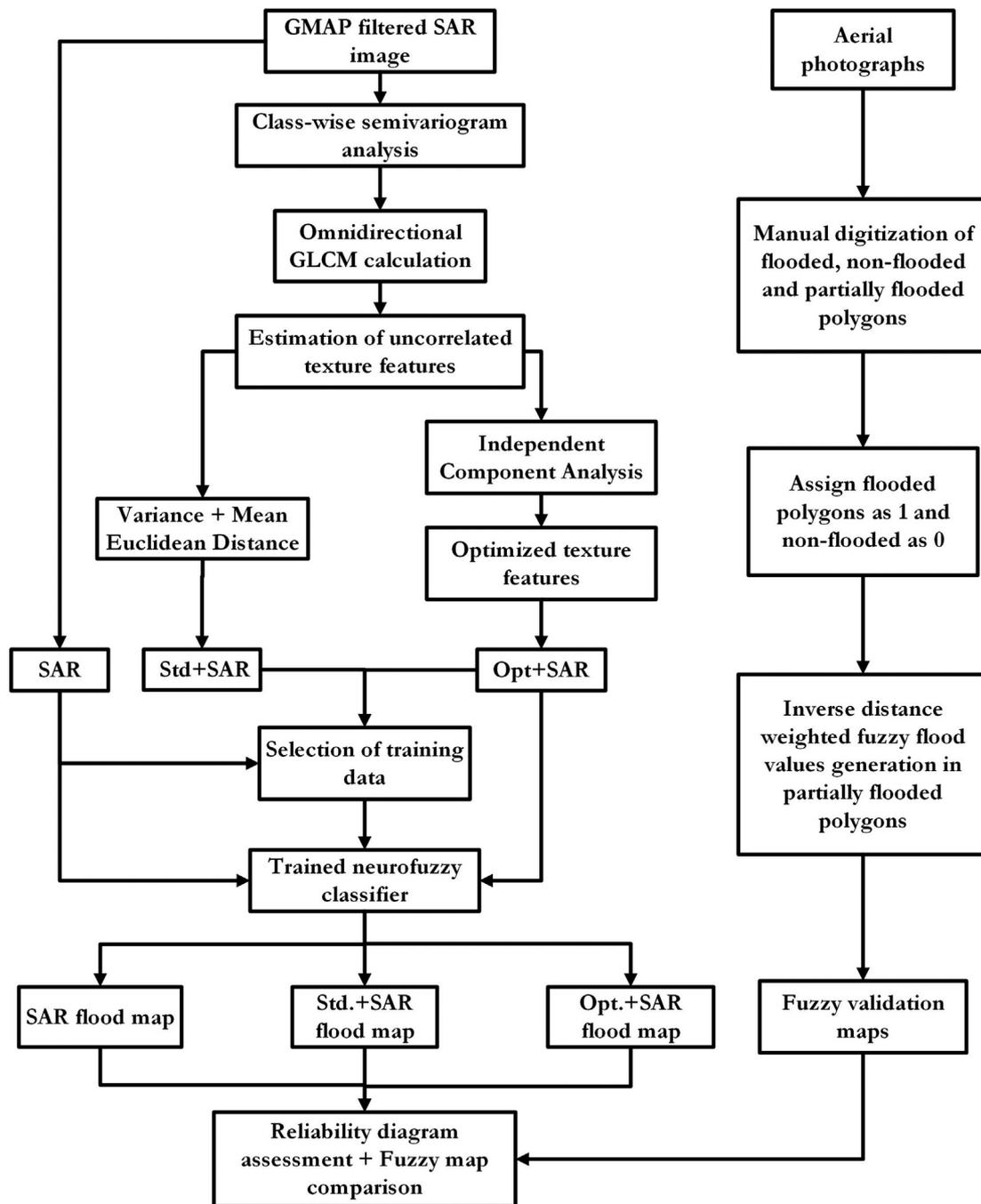


Fig. 3. Schematic of the overall classification framework used to generate the fuzzy flood maps.

3.3. The adaptive neuro-fuzzy inference system (ANFIS) classifier

3.3.1. Fuzzy membership function definition

Fuzzy set theory is a probabilistic adaptation of the classical notion of crisp sets which provides an elegant solution to objectively dealing with the ambiguity of SAR-based flood mapping (Pulvirenti et al., 2011). An element of fuzzy set is its representation of the degree of membership to a particular category or class, characterized by a pre-defined function with values ranging from 0 to 1. In this study, the Takagi-Sugeno type fuzzy inference system was implemented for the classification, as it can handle non-linearities in the data distribution (Takagi and Sugeno, 1985). Gaussian membership functions were chosen for both classes, as the histograms of flooded SAR images can be modelled as a mixture of two normal distributions (Giustarini et al.,

2016). Studies have shown that the bimodality assumption fails if the observed flooded area is not significant compared to the tile size (Chini et al., 2017). Therefore, the image was first subset to extract the area of interest - including the flooded area, flood plains and nearby regions - through visual interpretation. This results in a comparable division of flooded and not flooded pixels, causing the image histogram to exhibit a clear bimodality.

3.3.2. Function parameterization

Theoretical electromagnetic backscattering models have traditionally been used to define fuzzy membership functions and parameters (Pulvirenti et al., 2013). However, such approaches require detailed soil, vegetation, and land cover maps, to accurately estimate the expected backscattering behaviour based on theoretical models, which

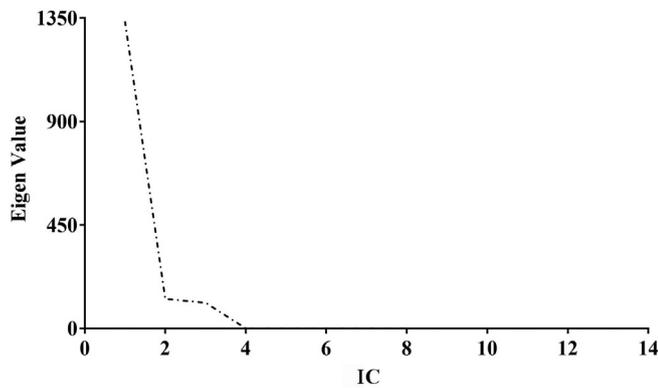


Fig. 4. Eigen values of the independent components obtained after analysing the texture bands.

are often unavailable. Moreover, as these models are wavelength specific, the parameterization is not applicable to data from other sensors or even other areas with a different distribution of scatterers on the ground. Furthermore, theoretical approaches are often unable to capture all practical considerations, for example, seasonal abscission in deciduous vegetation. The use of data driven models like artificial neural networks (ANN), which can learn data characteristics without prior process knowledge, is proposed to counteract this problem. Moreover, the variability in backscatter caused by wind and incidence angle effects - which alter the mean and standard deviation parameters of the class distribution - were implicitly accounted for in this approach, as they were estimated from the same image which needs to be classified.

The ANN was trained using polygons selected from the image, which were consistent throughout the analysis. The hybrid learning algorithm was used, which identifies parameters by iteratively minimizing errors using the gradient descent technique (Jang, 1993). The number of epochs used for training was 100 for each case and the model error was estimated using cross-validation. In order to individually quantify the impact of optimized texture on classification performance, the classifier was tested with the following inputs:

1. The speckle filtered SAR image or the control without any addition of texture, called SAR hereafter.
2. The filtered SAR image enhanced with some arbitrarily selected common texture features - variance and mean Euclidean distance in this case - hereafter called Std + SAR.
3. The filtered SAR image with the optimized rather than arbitrary texture bands added, referred to as Opt + SAR.

3.3.3. Training approach

The classifier was tested with three different sets of training data to ensure repeatability of results and to evaluate the sensitivity. The three training datasets are shown in Fig. 5 and the strategies for each training set described below:

- Set 1: Large area polygons were drawn for each class, encompassing the backscatter variability of the target classes, and spread out across the image. An equal number of polygons were chosen for the flood and non-flood classes.
- Set 2: Smaller area polygons, each comprising of a nearly uniform subset of backscatter values representing one of the signatures, were selected for the training. Polygon selection was restricted to the flooded area and floodplains, where maximum classification accuracy was desired. More non-flood polygons were chosen than flood, to ensure sufficient representation of the entire spectrum of backscatter variability for training.
- Set 3: Same as in training Set 2, except that the size and number of

polygons was further reduced to ensure that the classifier did not overfit the training data and to maximize computational efficiency.

As results of training Set 1 yielding large errors during the classifier assessment phase, it was not tested subsequently for the flood mapping or to assess the classification accuracy. However, a description of the training set has been provided here to highlight the problems with this set and inform readers about these limitations.

3.4. Validation strategies

In order to ensure the reliability of this analysis, results were validated using two different methods. First, a fuzzy set approach was used to validate the SAR-based fuzzy maps against a fuzzy validation target. Second, reliability diagrams were used to assess the flood maps, being the most commonly used validation technique for probabilistic maps.

3.4.1. Fuzzy set validation approach

Shoreline extraction from aerial photography can be quite ambiguous in densely vegetated and built up environments (Giustarini et al., 2013). The accuracy of the derived shoreline may vary from 10 to 100 m, depending largely on the skills of the photo interpreter (Mason et al., 2010). In this particular case, the riparian vegetation at the edge of the storage areas in the flood plain made it rather challenging to “see” the underlying water edge. The illumination differences in the separate flight lines, combined with atmospheric effects, severely affected the clarity of boundaries in the area, as seen in Fig. 6. Studies have proposed marking ambiguous boundary regions as “no data” to remove the associated uncertainty (Giustarini et al., 2016). However, it is more appropriate to use a performance measure which is capable of handling uncertainties, rather than discarding potentially valuable information.

A number of fuzzy performance measures capable of handling uncertainties in validation data have been developed for model calibration, especially those which use remote sensing data as targets (Pappenberger et al., 2007b). However, the utility of such metrics for the validation of SAR-based flood maps has not yet been tested. In order to facilitate this, the fuzzy map comparison method proposed by Hagen (2003) has been adopted in this study for the assessment of probabilistic flood maps for the first time. The proposed approach takes into account locational as well as categorical uncertainty in both the input as well as the validation data (Hagen-Zanker et al., 2005; Wealands et al., 2005). The resulting map comprises of pixel-wise similarity values ranging from 0 to 1, indicative of the local goodness of fit. As this evaluation approach requires an explicit representation of uncertainties in both datasets, a fuzzy flood map for validation was prepared as follows to reflect the errors of manual flood delineation:

1. The aerial photographs were manually digitized into three classes - clearly flooded, clearly non-flooded, and possibly flooded.
2. The flooded polygons were assigned a value of 1 and non-flooded polygons, a value of 0.
3. The partially flooded polygons were filled with intermediate values, interpolated using an inverse distance weighting (IDW) algorithm with an exponent of 2, to fuzzify the validation data.

3.4.2. Fuzzy similarity statistics

In order to characterize the similarity between the SAR-based flood map and the one derived from aerial photos, several statistics were calculated which could highlight the different aspects of classifier performance. First, deterministic grid-based statistics, such as Root Mean Squared Errors (RMSE) and Mean Absolute Errors (MAE), were calculated to assess classifier performance. This was followed by the map comparison method proposed by Hagen (2003), which allows the comparison of fuzzy maps while explicitly accounting for spatial and

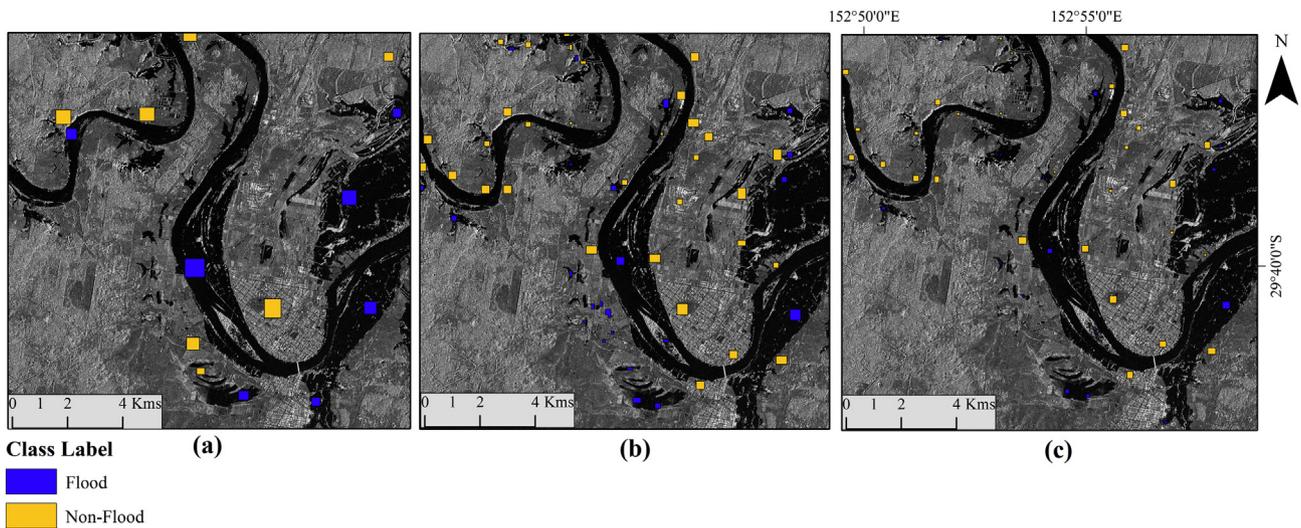


Fig. 5. The actual training polygons selected to train the neuro-fuzzy classifier are shown, where (a), (b), and (c), correspond to training sets 1, 2, and 3, respectively. The difference in the three trainings is primarily the size of individual polygons and their corresponding locations as shown by the coloured squares. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

categorical uncertainties in both datasets.

For the two fuzzy vectors; F_{SAR} was obtained as the SAR-derived flood extent while F_{VAL} represents the validation data. The similarity measure S can therefore be computed for two pixels at the same location in both maps using:

$$S(F_{SAR}, F_{VAL}) = \left[\min(F_{SAR_{flood}}, F_{VAL_{flood}})_{min}, \min(F_{SAR_{non-flood}}, F_{VAL_{non-flood}})_{min} \right]_{max} \tag{1}$$

where $F_{SAR_{flood}}$ and $F_{SAR_{non-flood}}$ denote the fuzzy values of the flood class and the non-flood class memberships at a given pixel in the SAR-based map, while $F_{VAL_{flood}}$ and $F_{VAL_{non-flood}}$ refer to corresponding values in the validation map. S is the maximum value within the set obtained by evaluating the minimum of the two fuzzy vectors F_{SAR} and F_{VAL} on a

pixel by pixel basis (Pappenberger et al., 2007a), which can also be interpreted as the maximum grade of membership to the intersection of the fuzzy sets (Zadeh, 1965). S takes values from 0 to 1 based on the degree of similarity, such that 0 was assigned to cells that were completely different and 1 was assigned to those which were identical. In order to account for fuzziness in location, the influence of neighborhood cells was also considered. The contribution of each neighborhood cell can be calculated using a distance decay function chosen based on the nature and magnitude of uncertainties, and the desired tolerance for spatial error (Hagen, 2003).

In this study, the number and impact of neighbors included in the analysis were selected based on the semivariogram assessment results. The neighborhood influence was estimated based on the 3D exponential decay function with a halving distance of two pixels. This essentially

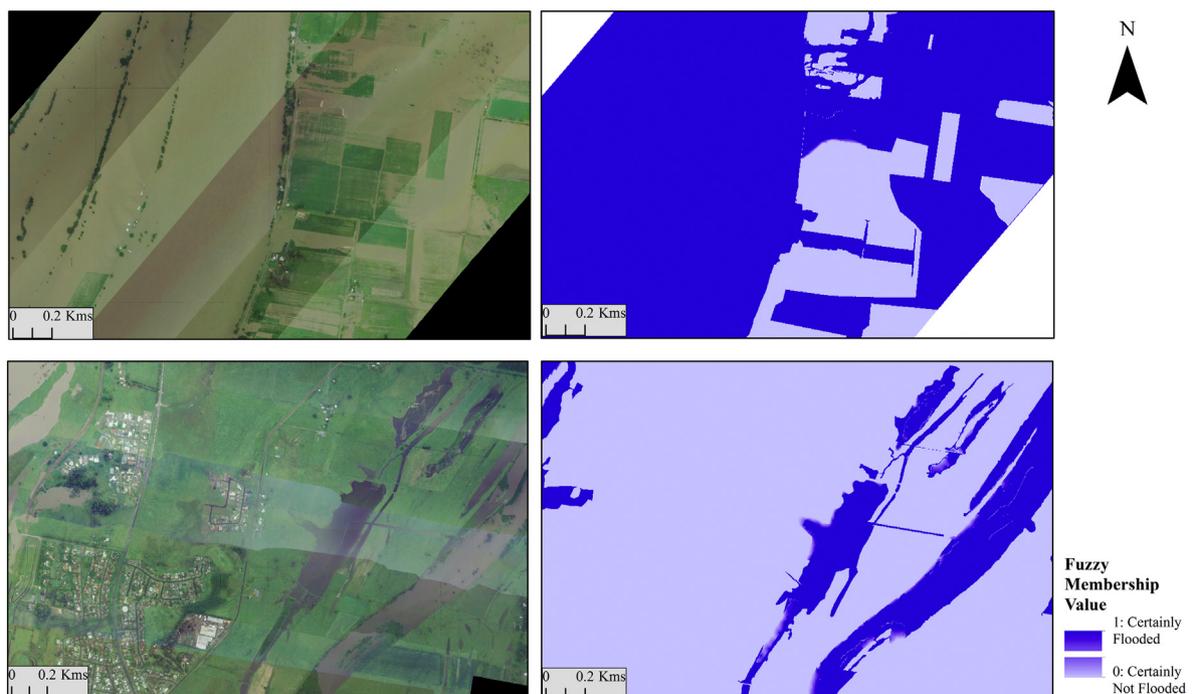


Fig. 6. Example subsets of the true colour aerial photographs (left) shown along with the corresponding manually fuzzified flood maps (right). Locations are shown in Fig. 1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

means that the influence of the neighborhood was halved every two pixels. The two way similarity between the fuzzy distance weighted neighborhood contributions ($F_{SAR_{nbh}}, F_{VAL_{nbh}}$) and the fuzzy membership values for the central cells ($F_{SAR_{cc}}, F_{VAL_{cc}}$) were calculated using the expression:

$$S(F_{SAR}, F_{VAL}) = |S(F_{SAR_{nbh}}, F_{VAL_{cc}}), S(F_{SAR_{cc}}, F_{VAL_{nbh}})|_{min} \tag{2}$$

For further information about the calculation of the neighborhood sets, the reader is referred to (Hagen-Zanker et al., 2005). Last, the fuzzy kappa K_{fuzzy} statistic is proposed as an overall performance measure, calculated using:

$$K_{fuzzy} = \frac{(S_{obs} - S_{exp})}{(1 - S_{exp})}, \tag{3}$$

where S_{obs} and S_{exp} are the observed and expected percentages of fuzzy agreement. Here the observed percentage of fuzzy agreement refers to the spatial similarity between the SAR-based maps and validation data computed using Eq. (2). Conversely, the expected agreement is a function of the number of classes and the image histogram calculated theoretically (Hagen, 2003). K_{fuzzy} only differs from the traditional Cohen's kappa popularly used for map comparison, in the calculation of the expected percentage of agreement S_{exp} (Hagen-Zanker, 2006).

The fuzzy Kappa statistic quantifies the improvement between the compared maps relative to a randomly generated categorical map with an identical histogram (Wealands et al., 2005). S_{exp} is estimated as the probability that a wet pixel observed in the validation data appears within a certain range of pixel distances or neighborhood in the corresponding SAR-based flood maps. This means that a flooded pixel in the validation data, which may have shifted in the SAR-derived map due a variety of factors such as sensor orientation or geo-location errors, is considered a match if it fell within reasonable bounds of uncertainty. S_{exp} can be viewed as a measure of the chance agreement that the two maps in consideration may exhibit, based on chosen neighborhood sizes and the number of classes in the dataset.

Neighborhood rings are defined as the set of cells located at an equal distance from the central cell. This implies that the chance agreement needs to be computed for each central cell and all possible neighborhood rings. Ideally this should be done for each pixel individually and for an infinite zone of influence, as edge pixels may have a different nature and number of possible neighborhood rings. However, on increasing the search radius for neighborhood identification from 10 cells to 500, the difference in values of S_{exp} was found to be insignificant, as the function used to model the influence decays rather rapidly. Furthermore, using a search radius of 500 increases the possible number of permutations and combinations, thereby drastically increasing the computational time. Finally, a radius of 10 pixels was used to calculate the expected agreement between the SAR-derived and aerial photo based fuzzy flood maps.

3.4.3. Reliability diagram assessment

The final performance assessment used reliability diagrams, being

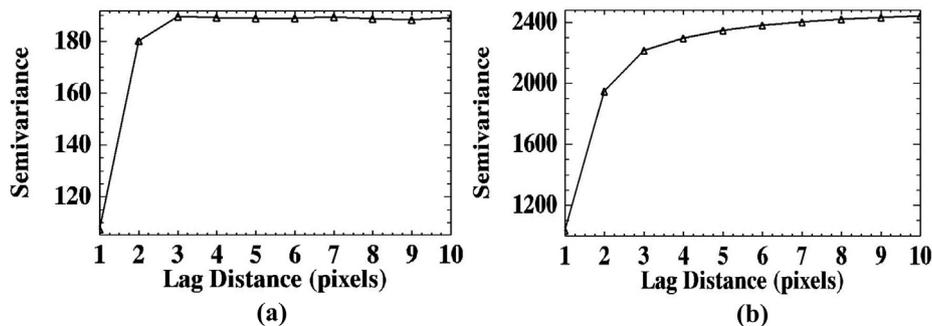


Fig. 7. Semivariograms showing spatial autocorrelation among backscatter values for the different classes in a COSMO-SkyMed 3 m image.

the currently accepted evaluation method for probabilistic maps in literature. As discussed previously, the limitation of this technique is that the validation data need to be binary. However, this analysis was conducted to facilitate the understanding of error characteristics with respect to state of the art mapping techniques and to understand the specific contributions of under or over predictions. A threshold of 0.5 was chosen for defuzzification, to reflect the maximum uncertainty in the resulting deterministic flood map (Schlafter et al., 2017). As the validation map was digitized on an aerial photo with a spatial resolution of 10 cm, the choice of this threshold has limited impact on the outcome of the analysis conducted at 3 m.

The agreement between the fuzzy membership values predicted by the neurofuzzy classifier, and the observed proportion of flooded pixels in the validation data can be characterized by a reliability diagram (Horritt, 2006). The fuzzy membership values were binned into intervals of 0.1 and the ratio of wet cells to total number of cells in each bin, was plotted against the bin means. The ideal classifier assigns fuzzy membership values identical to the proportion of observed wet pixels in the validation data, with deviation from the 1:1 line representative of the classification error. As the distribution of pixels across the bins is non-uniform, a weighted RMSE (WRMSE) was calculated to objectively represent the uncertainty. The WRMSE is calculated by assigning weights to the bin errors, based on the bin population as in (Giustarini et al., 2016):

$$WRMSE = \sqrt{\frac{\sum_{i=1}^N (F_v - F_s)^2 n_i}{\sum_{i=1}^N n_i}}, \tag{4}$$

where n_i is the pixel count of each bin, F_v is the observed proportion of wet cells in the validation map, F_s is the fuzzy membership value predicted based on SAR analysis, and N is the total number of pixels in the validation domain. Reliability diagrams were also used to assess classifier performance for each land cover (LC) class in the study area through reliability diagrams. Such diagnostic analyses may lead to insights on when the proposed approach can be expected to perform well.

4. Results and discussion

4.1. Window size selection

The choice of an appropriate window size is an essential step for texture estimation, to avoid inadvertently interpreting noise as a meaningful pattern. Therefore, the range of spatial autocorrelation observed through the flood and non-flood semivariograms was used as the window size for texture analysis. The semivariogram plots were estimated by visually selecting homogeneous subsets of flood and non-flood classes, to ensure that the observed range reflects only the intra-class variance. The non-flood subsets were located clearly outside and away from the floodplain area, such that purely dry land pixels could be isolated.

Fig. 7 shows the semivariogram plots obtained for the flood and

non-flood classes. The range of both semivariograms is consistent at 3 pixels as expected, as high-resolution data are more prone to noise and thus exhibit low spatial autocorrelation. The non-flood class exhibits significantly higher values of variance due to the diversity of surface scattering characteristics encountered on land.

4.2. Neuro-fuzzy classifier - training, testing, and validation

The results for the neuro-fuzzy mapping are presented in the following manner. First, the general training procedure and results of the training are discussed, followed by an evaluation of the result maps obtained for each validation site. As the classifier uses a data driven estimation of the model parameters, it needs to be trained and subsequently subjected to rigorous testing. It is important to note that for statistical models, two types of errors need to be estimated:

1. Model selection error - which characterizes the prediction error of the selected model in the context of other available models. This analysis allows to select the one most suited to the data distribution;
2. Model assessment error - which evaluates the ability of the selected model to correctly classify new or previously “unseen” data points (Hastie et al., 2009).

The appropriate way to assess both, if sufficient data points are available, is to randomly divide them into three parts: a training set, a validation set, and a testing set. Based on the signal-to-noise ratio expected of a high resolution SAR image and the complexity of the Gaussian model, the split was chosen as 70% training, 15% validation and 15% testing (James et al., 2000). The training set includes the majority of data points by convention, and is used to fit the data driven model or distribution. Ideally, all distributions which are able to sufficiently explain the data characteristics should be examined. A separate dataset, hereafter referred to as the validation set, is used to compute the prediction error for all the competing models or the model selection error. The model with the lowest value of validation error is chosen for further analysis. Since in this case the Gaussian model was already selected based on the histogram bimodality, the validation error was used to select a suitable parameterization of the distribution.

Once a distribution has been trained and validated as the best fit model, another previously unseen batch of data, called the test set, is used for assessment. The test error provides a measure of the generalization capabilities of the chosen model and the ability to correctly classify new data points. If the validation set is reused for this evaluation, the true test error will be substantially underestimated. Ideally, multiple equally plausible model structures should be tested to assign the test error. From the histograms of the SAR image and the optimized texture features presented in Fig. 8, it is apparent that the distributions were exhibiting slight deviations from Gaussianity. However, investigating different distributions with better fits was considered outside the scope of the present investigation, primarily because histograms of flooded SAR images are usually known to consist of a mixture of two Gaussian class distributions, which is also evident from the backscatter histogram shown in Fig. 8(a) (Chini et al., 2017). As the distributions of the texture classes were also nearly symmetrical and choosing a case-specific distribution function would limit transferability of the method, the Gaussian assumption was maintained throughout the analysis.

The error values followed the expected pattern of lower training set errors and larger test and validation set errors as summarized in Table 1. Training Set 3 led to a more generalized classifier as errors were nearly consistent across training, validation, and test sets. As elaborated earlier, minimizing the sample size can reduce overfitting which in turn can improve classifier performance, by reducing the bias-variance trade-off (James et al., 2000). This is in contrast to Set 2, where the validation errors were noticeably higher for both the texture based methods. The larger magnitude of test and validation errors can

be due to overfitting to the training data in Set 2, which adversely affects model generalization. Generalization can be defined as the ability of a classifier to correctly identify previously “unseen” data points, a desirable quality for any classification problem (Hastie et al., 2009). The pixel-wise comparison of deterministic statistics was carried out for both training Set 2 and Set 3, however, the reliability diagrams and fuzzy statistics were only calculated for Set 3. Set 3 was chosen for the more detailed analysis as the classifier achieved a better generalization and was expected to perform better.

Note that the training, testing, and validation datasets defined in this section are subsets of the training data selected to “train” the neurofuzzy classifier, and the meanings of these terms are specific to the data driven modelling part of this study. After the model selection and training process was conducted, the trained model was used to generate fuzzy flood maps from SAR. These were validated against manually derived flood maps from aerial photography, hereafter referred to as the validation data.

4.3. Fuzzy flood maps - accuracy assessment

The fuzzy flood maps obtained post classification were assessed using two approaches - using the fuzzy map comparison and reliability assessment. For each validation site, the fuzzy flood maps, difference maps, spatial similarity maps and reliability diagrams were generated through the procedures detailed in Sub-section 3.4.

4.3.1. Validation site 1: Junction Hill

The pixel-wise assessment results from the two training datasets were found to be somewhat inconsistent, as evident from Table 2. The proposed texture optimization approach succeeded in reducing the RMSE slightly (~2%) in Set 2. However, Set 3 exhibited a slightly higher RMSE, though increase of 10^{-3} can be considered to be negligible. It can easily be observed from the fuzzy flood maps illustrated in Fig. 9 that the optimized texture approach reduced the misdetection of linear and smooth urban features, like road networks, as flooded pixels in the urban land-use dominated (top row) subsets.

The proposed approach also reduced the fuzziness in the sparsely vegetated region which has varying backscatter but relatively homogenous texture, as seen in the second row of Fig. 9 where local improvements are visible. However, the pixel-wise assessment was unable to capture this improvement as both the maps had very different uncertainty characteristics. A deterministic differencing approach is inappropriate for the assessment of continuous random variables as an exact match is nearly impossible. Furthermore, the difference maps depicted in Fig. 10 highlight the disparity between the error characteristics in the validation data and the SAR-based maps. In fact, the SAR-based maps provide more realistic estimates of uncertainty at each pixel, as these are calculated objectively through the ANFIS classifier. Conversely, given that the chief contributor to the uncertainty in manually digitized flood maps is the skill of the analyst (Mason et al., 2010), quantifying this value objectively is significantly more challenging.

The fuzzy set evaluation approach results in two sets of maps; one where only the fuzziness of the membership value or class assignment is considered and one which additionally includes both fuzziness of value as well as the fuzziness of location. The fuzzy similarity set presented in Fig. 11 shows a clear improvement but the transition zone uncertainty is spuriously inflated. As the validation data are manually digitized and fuzzified, the nature of uncertainties in the transition zone may be very different from that of SAR-based approaches. In contrast, the spatial similarity index illustrated in Fig. 12 reflects the improvements offered at the land-water boundary much better. A significant reduction in uncertainties is noticeable across the domain.

The overall performance (average spatial similarity index) values were 0.899, 0.905 and 0.915 for SAR, Std + SAR, and Opt + SAR, respectively (Table 3). The fuzzy kappa statistic, which corrects for the

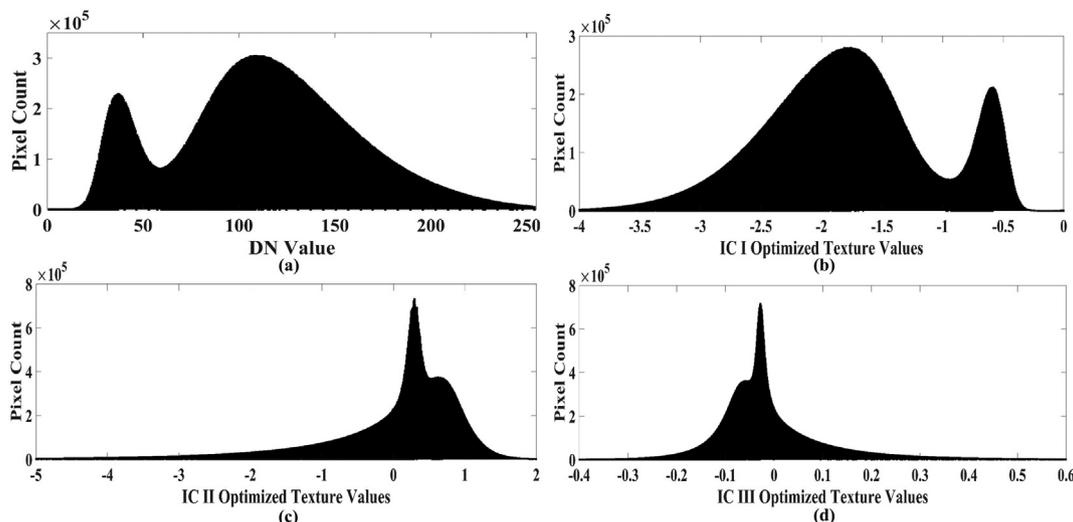


Fig. 8. Histograms depicting the bimodality in the distribution of pixel values for (a) the filtered SAR image in digital numbers, (b) Independent Component (IC) I texture values, (c) IC II texture values, and (d) IC III texture values.

Table 1
Mean Absolute Error statistics for classification model selection and predictive capability assessment.

Input	Training Set 2			Training Set 3		
	Training	Validation	Testing	Training	Validation	Testing
SAR	0.184	0.183	0.186	0.217	0.217	0.216
Std + SAR	0.155	0.178	0.159	0.174	0.175	0.175
Opt + SAR	0.121	0.136	0.125	0.117	0.12	0.12

Table 2
Root Mean Squared Error statistics based on the pixel-wise deterministic difference operation.

	Training Set 2		Training Set 3	
	Junction Hill	Ulmarra	Junction Hill	Ulmarra
SAR	0.242	0.243	0.233	0.261
Std + SAR	0.239	0.246	0.241	0.257
Opt + SAR	0.222	0.24	0.236	0.24

expected percentage of agreement and can be more informative, is identical to average similarity in this case due to low expected agreement.

The K_{fuzzy} can provide a measure of the overall improvement, which is categorically required for many applications. As the expected value for similarity is directly related to the number of classes and the intra-class pixel distribution, the values obtained for this statistic can also be insightful (Hagen, 2003).

Expected similarity can be understood as the probability of chance agreement between the two maps having identical image histograms. In the context of fuzzy spatial similarity, this value of possible chance agreement is computed for all the cells in the neighborhood rings under consideration. One of the possible reasons for the low expected similarity values obtained might be the distance decay function chosen here, which was the 3-D exponential function with a halving distance of two pixels, reducing the neighborhood influence drastically. This observation is in keeping with the expectation, as the nature of uncertainties in the validation data is very different from the SAR-based maps. Since the validation map is manually digitized and fuzzified, error variation is expected to differ substantially from the fuzzy flood maps from SAR, which objectively represent backscatter and classification uncertainty.

The spatial auto-correlation in high-resolution SAR data are also highly localized due to increased speckle noise, as already established through the variogram analysis. However, it is expected that K_{fuzzy} can help with assessing probabilistic flood maps and may add value to the average spatial similarity. The expected similarity S_{exp} would obviously vary with the spatial resolution of the maps, choice of the distance decay function, and the number of classes considered in the analysis. In cases like this where S_{exp} is nearly negligible, the average spatial similarity statistic S_{obs} may suffice as a test statistic for map comparison, as the K_{fuzzy} doesn't add any new information.

The reliability diagram for Junction Hill (Fig. 13) better reflects the improvement offered by Opt + SAR, which consistently gave predictions very close to the 1:1 line. Both the texture based approaches correctly classified nearly all the pixels for the last few bins, containing flood membership values ranging from 0.8–1. These are the certainly flooded pixels of the study area and the ones the algorithm primarily seeks to correctly identify. The addition of optimized texture, especially after optimization, seems to be conducive to this cause. Furthermore, omission errors seem to dominate over commission errors, for most of the uncertain bins. A closer examination reveals that the under-prediction increases with the uncertainty, i.e. the Opt + SAR assigned lesser pixels to bins 4, 5, and 6 than observed in the validation data. The SAR image alone seemed to over predict in bin 9 and under predict in bin 10, both of which contained pixels almost certainly flooded, indicating a contribution of noise in the training. Consequently a reduction in the signal to noise ratio led to ambiguity in classification even within homogeneous image objects such as flood patches.

Texture estimation highlighted patterns and increased the separability between signal and noise. This reduced the uncertainty in feature extraction, offering some advantages for flood assessment. The change in the distribution of bin-sizes on the application of the Opt + SAR approach, evident from the subplot included in Fig. 13, corroborates this interpretation. The WRMSE values based on the deviation from the 1:1 line show a relative improvement of 54.2% over the use of SAR image without texture, and 52.6% over the use of standard textures when using the optimized texture approach. This clearly indicates the importance of choosing appropriate texture features, as the addition of arbitrary texture features reflects no significant improvement.

4.3.2. Validation site 2: Ulmarra

At Ulmarra, the pixel based statistics show a reduction in RMSE for both training Set 2 and 3. However, the magnitude of this reduction was

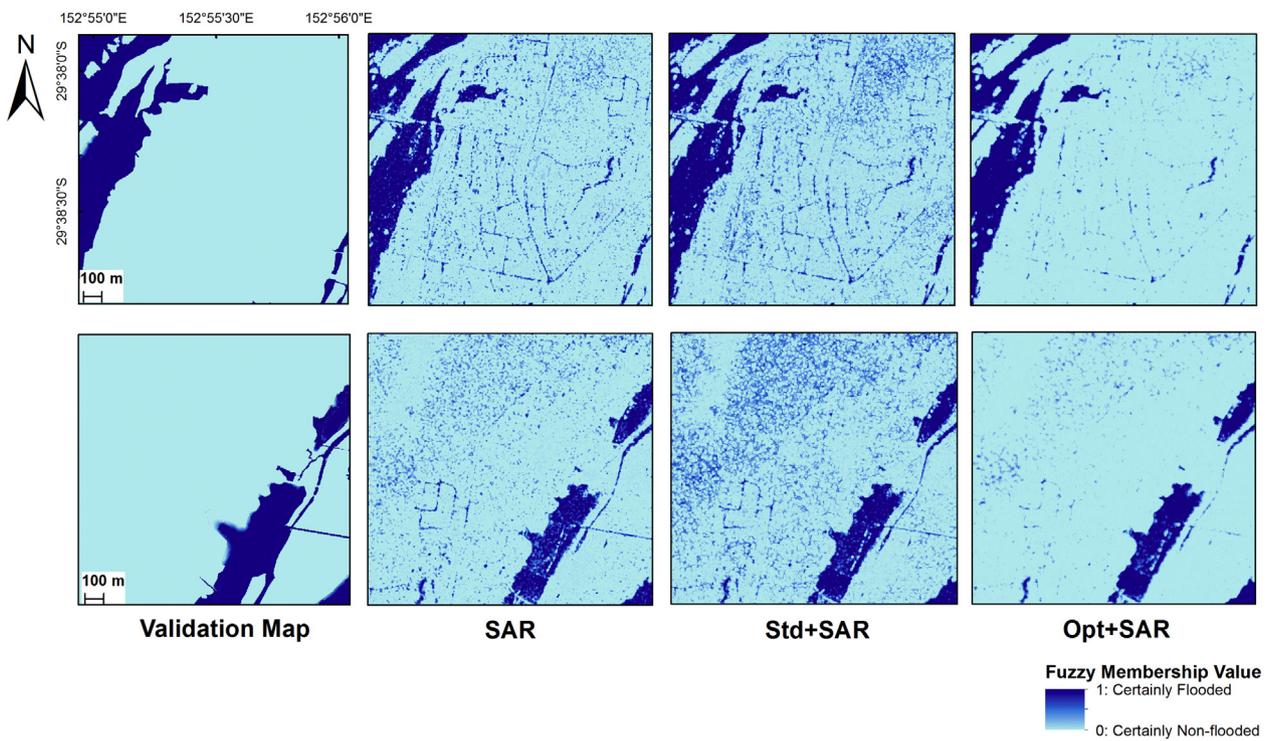


Fig. 9. Validation maps generated from aerial photographs are displayed in the first column, followed by flood maps derived by processing the following inputs through the ANFIS classifier; SAR alone in column two (SAR), arbitrarily selected textures with SAR in column three (Std + SAR), and optimized textures with SAR in column four (Opt + SAR). Areas depicting maximum reductions in uncertainty for the Junction Hill test site were chosen for illustration. The locations of the subsets used here are shown in Fig. 1.

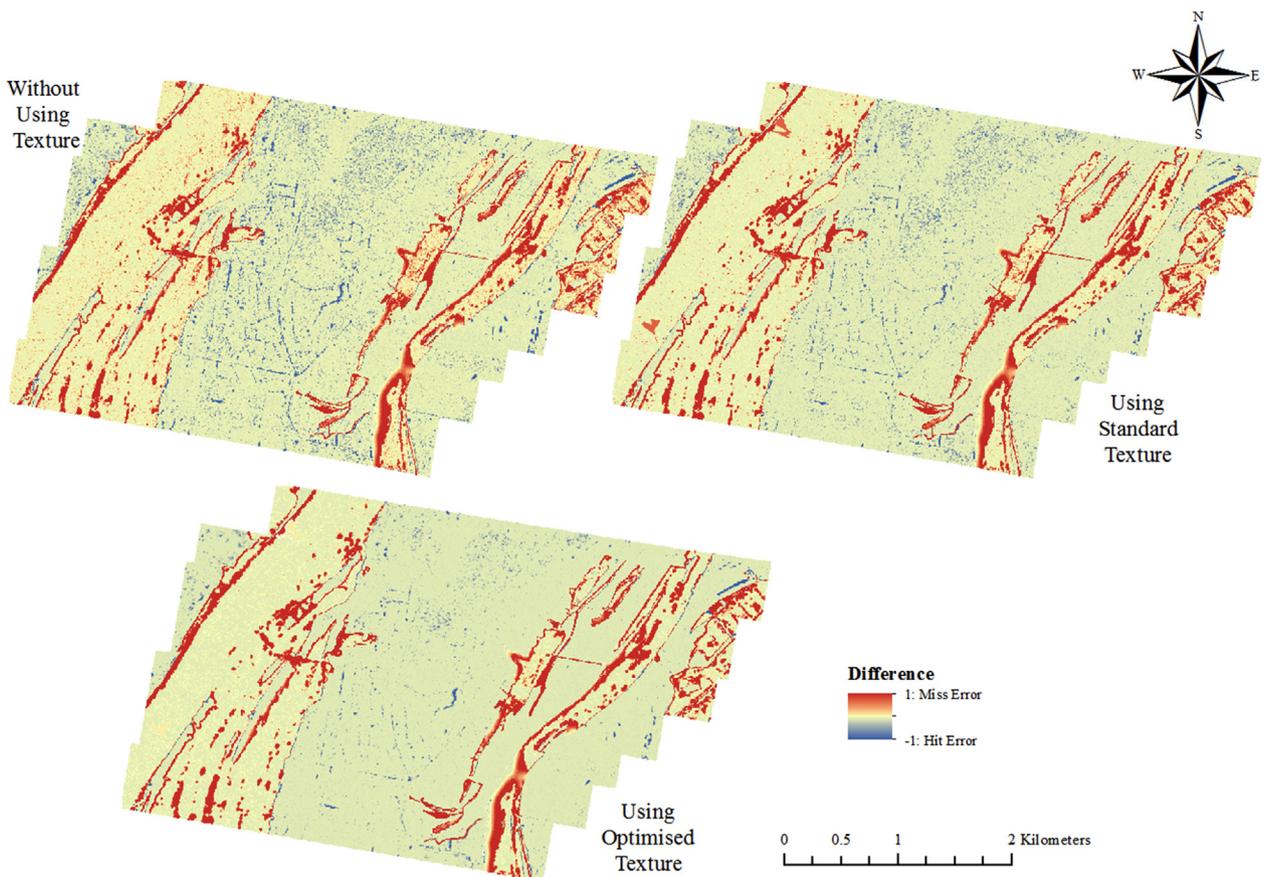


Fig. 10. Pixel-wise difference maps generated by subtracting the SAR-based flood maps, from the fuzzy validation map digitized from aerial photos, for entire Junction Hill region.

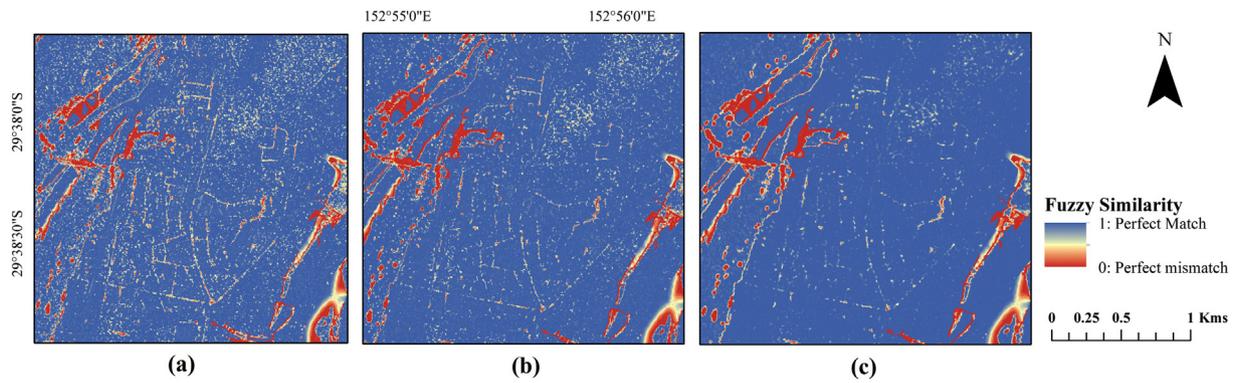


Fig. 11. Fuzzy similarity maps for the central cell comparison between the Junction Hill validation map and the SAR-based fuzzy maps, where (a) SAR represents the use of SAR alone as a classification input, (b) Std + SAR represents arbitrarily selected textures with SAR, and (c) Opt + SAR represents optimized textures with SAR.

larger in Set 3. According to the flood maps in Fig. 14, this validation site had a larger proportion of flooded pixels. Since a significant percentage of the image is covered with a homogeneous land cover type, the impact of noise is more prominent. The optimized approach performed better in Set 3, due to the better model generalization achieved. The generalized model exhibited lower noise sensitivity. The Opt + SAR flood maps showed a noticeable reduction in fuzziness within the homogeneous regions and near the transition zone, in agreement with the findings from Junction Hill.

While the pixel-wise RMSE values reduced with the addition of texture, it is important to consider the failure of the deterministic differencing approach to capture improvements in mapping for the Junction Hill site. The problem of non-contiguous vegetation patches surrounded by flood pixels is observable even in Ulmarra. The deterministic difference maps for Ulmarra have not been included here for brevity as the inferences from the analysis were similar to those from Junction Hill.

Additionally, some portion of the domain had agricultural fields with standing crops, highlighted in the top row of Fig. 14, making the underlying flood water very difficult to detect. A large portion of this area had been classified as non-flooded by all the SAR based approaches. As texture is a backscatter-derived property, the approach is unable to bring significant improvements in this area. Since the experiment was conducted using low-wavelength X-band data, which experiences severe scattering due to emergent vegetation, it was rather challenging to demonstrate the full potential of the proposed approach in this region. However, the RMSE still showed an improvement overall as the texture optimization seemed to reduce noise sensitivity. Since this particular test site was populated with homogeneous flood patches, where the effects of noise dominate, the overall errors were decreased.

The fuzzy similarity maps for Ulmarra (Fig. 15) show large regions of uncertainty in the flood transition zone. Although the Opt + SAR

Table 3
Summary of fuzzy statistics for the two validation sites.

Input	Junction Hill			Ulmarra		
	Observed similarity (S_{obs})	Expected similarity (S_{exp})	Fuzzy kappa (K_{fuzzy})	Observed similarity (S_{obs})	Expected similarity (S_{exp})	Fuzzy kappa (K_{fuzzy})
SAR	0.899	$3.341e^{-12}$	0.899	0.670	$4.843e^{-10}$	0.670
Std + SAR	0.905	$3.168e^{-12}$	0.905	0.689	$3.799e^{-11}$	0.689
Opt + SAR	0.915	$1.342e^{-12}$	0.915	0.713	$3.827e^{-10}$	0.713

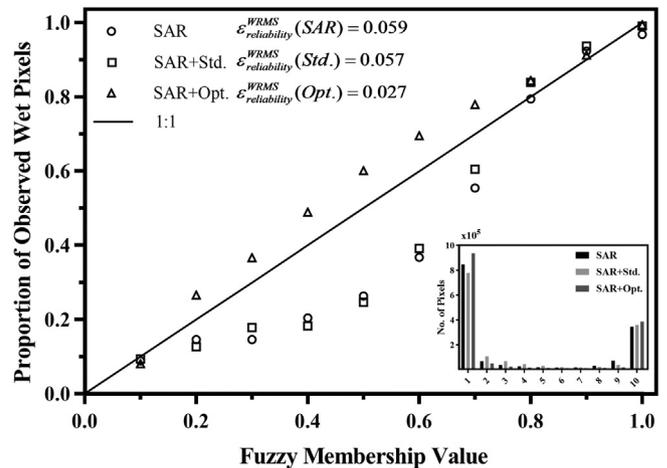


Fig. 13. Reliability diagram for the Junction Hill area with Weighted Root Mean Squared Error values and bin sizes represented in a sub-plot.

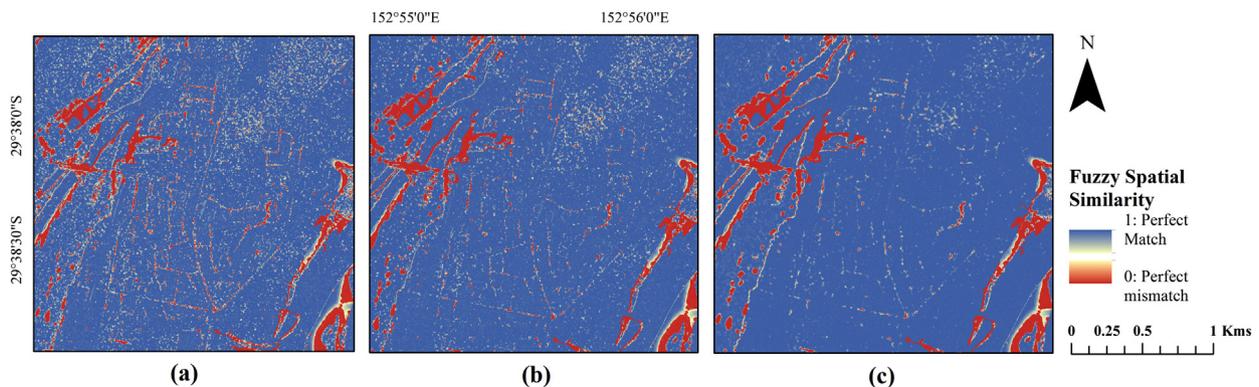


Fig. 12. As for Fig. 11 except with neighborhood context included for the Junction Hill site with (a) SAR, (b) Std + SAR, and (c) Opt + SAR.

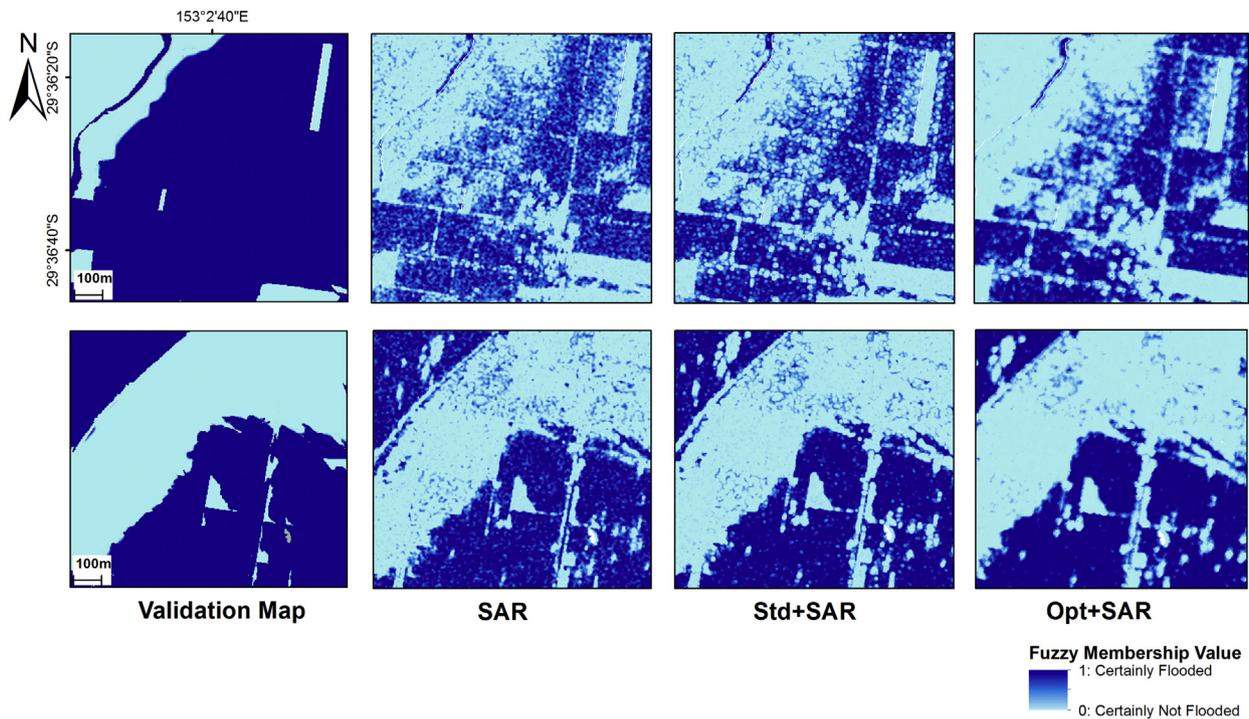


Fig. 14. As for Fig. 9 but for the Ulmarra test site, with locations of the chosen subsets highlighted in Fig. 1.

map shows some improvements compared to the SAR and Std + SAR, it is noteworthy that the inclusion of texture itself offers limited enhancement in classification performance. This is evident from Fig. 14, which shows a notable reduction in uncertainties in the Opt + SAR compared to the Std + SAR approach. This implies that in areas where texture based approaches can be expected to work, such as distinguishing between water and water look-alike surfaces, the Opt + SAR technique will definitely provide better outputs than arbitrarily selected texture features.

Spatial similarity maps provide a suitable assessment measure for probabilistic maps, as a clearer reduction in the transition zone uncertainty is visible in Fig. 16. S_{obs} values were 0.67, 0.69, and 0.71 for SAR, Std + SAR, and Opt + SAR, respectively, as shown in Table 3. The K_{fuzzy} values were identical to S_{obs} , as in the case of Junction Hill, as values of S_{exp} were low for the same reasons. However, the K_{fuzzy} and S_{obs} both showed improvements when using the proposed technique at both test sites. This corroborates the visual observations from the fuzzy maps (Figs. 9 and 14), generated from each input image. The fuzzy map comparison technique used here can also be thought as something which mimics human interpretation of errors. The central cell comparison with the additional consideration of contextual neighborhood

influence is a powerful tool, which should further be imbibed for the assessment of probabilistic maps.

At Ulmarra the reliability diagram statistics are inconsistent with the results obtained from the fuzzy set analysis as illustrated in Fig. 16. It is important to observe that WRMSE increases with the addition of any texture in this case. One of the possible reasons for this could be the large number of certainly flooded pixels classified as not flooded in the SAR based maps. This large discrepancy can be explained by the presence of emergent vegetation in the agricultural region. The deviation from the 1:1 line is large (-1), and as bin 1 has a large population size it consequently has a higher weightage, which amplifies the error statistic. Further, the WRMSE increases when adding optimized texture, as a large number of pixels shift from the uncertain bins to the certainly non flooded and flooded bins, increasing the weight associated with these errors. On closer examination, the reliability diagram reveals a notable shift from all the intermediate bins, containing the uncertain flood values towards the certain bins (1 and 10). This highlights that the optimization approach caused a shift in what backscatter value combinations are classified as uncertain, pushing mixed pixels or those with emergent vegetation towards the flood class. This observation is identical to the findings at Junction Hill, where the optimized texture

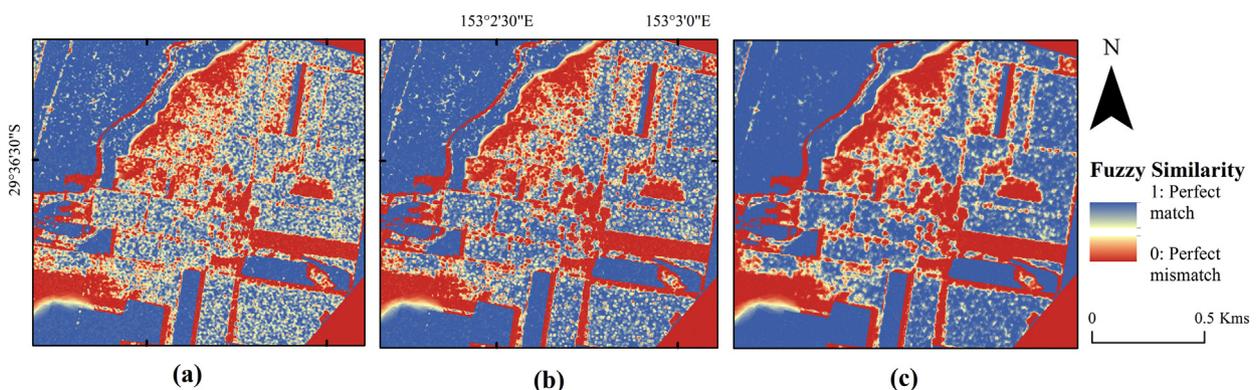


Fig. 15. As for Fig. 11 but for the Ulmarra region, where the agreement with validation data is shown for (a) SAR, (b) Std + SAR, and (c) Opt + SAR.

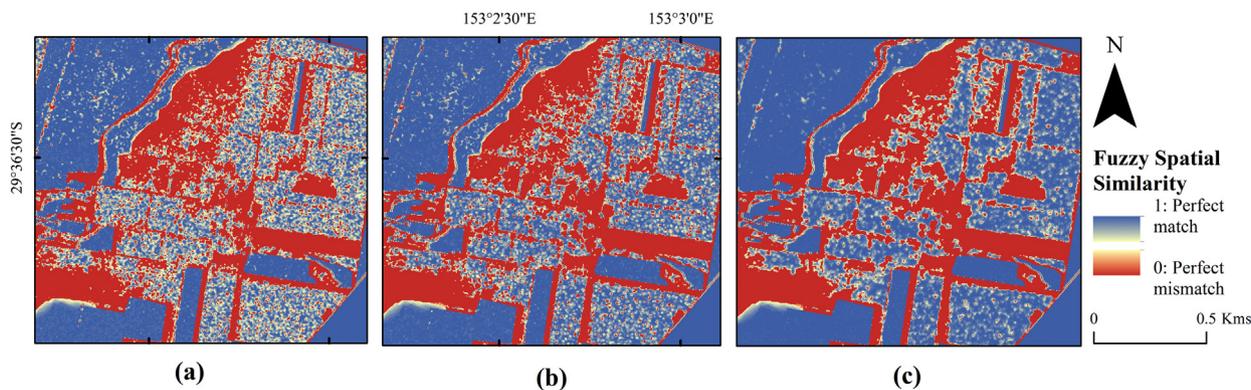


Fig. 16. As for Fig. 12 but for the Ulmarra test site with similarities shown as (a) SAR, (b) Std + SAR, and (c) Opt + SAR.

addition resulted in clear error reductions.

Giving due consideration to the visibly flooded agricultural region (top row, Fig. 11), which is only partially flooded in the SAR-based maps, allows better understanding of the reduction in uncertainty through texture optimization. In the SAR-based maps the region consists of pixels mostly from bin 1 and 2, i.e. closer to the certainly non-flooded end on the fuzzy membership spectrum. However, as optimized texture enhances feature extraction and reduces noise effects, these patches were pushed towards the certainly non-flooded bins. This is consistent with the notable increase in the pixel count of bin 1 for Opt + SAR when compared to other inputs. Many of these pixels from bin 1 and 2, however, belong to the aforementioned error hotspots and were flooded in the validation data. Therefore, the proportion of wet pixels was increased in the validation data for the almost certainly non-flooded classes, especially in bin 2. As more pixels wrongly became “surer” of their non-flooded status, i.e. moving from uncertain bins towards bin 1 and 2, the error margin increased thus amplifying the overall classification error. It is clear from the above analysis that the use of reliability diagrams alone is insufficient for the assessment of fuzzy maps. The need to reconvert the continuous, fuzzy, SAR-based maps into discrete categorical maps with specified bin sizes as well as the need for a binary validation dataset for assessment, limits the applicability of reliability diagrams for an objective evaluation of probabilistic flood maps.

When comparing continuous spatial fields as in the fuzzy flood maps generated here, tolerance for locational or categorical errors is desirable (Pappenberger et al., 2007a,b). Locational tolerance accounts for slight pixel shifts without denoting them in complete disagreement, while categorical tolerance allows to identify the higher similarity between “slightly wrong” values which are common in hydrological spatial fields (Wealands et al., 2005). As the fuzzy spatial similarity statistic accounts for both category and location fuzziness, the improvement brought about by Opt + SAR could be evaluated without the impact of spatial mismatches caused by the notably different nature of the uncertainties in the SAR and the validation data. In this case, addition of the fuzzy statistic to the accuracy assessment approach, proved to be rather informative.

The fuzzy map comparison added value to the reliability diagram analysis, with the local spatial improvement demonstrated through the maps and captured by the overall similarity statistic (Hagen-Zanker et al., 2005). Even though the values of spatial agreement at Ulmarra were lower than those achieved at Junction Hill, which was expected due to the error hotspots critically discussed previously, a clear improvement was evident. All values were > 0.6 which has been considered a satisfactory measure of fit, given the expected uncertainties in shoreline locations (Pappenberger et al., 2007a,b). By accounting for locational and attribute uncertainties in the computation of the local matching, chance agreement and image registration problems were also accounted for (Power et al., 2001). Furthermore, fuzzy map comparison

provides a unique opportunity to assess the spatial characteristics of the classification errors, which can help to better diagnose their underlying cause. Correctly identifying the reasons and nature of uncertainties represents the first step in designing appropriate post-processing strategies, or even formulating better SAR-based flood mapping techniques for the future (Power et al., 2001).

4.4. Land cover based performance analysis

An investigative analysis of the classifier performance within different land cover classes was undertaken, to interpret the classifier response when exposed to pixels with varied surface characteristics. This can help to decide whether or not to use the proposed method based on the dominant land-use class in the region. In order to quantify the above, land cover maps of 250 m resolution were subset for the validation sites. The land cover maps were used to extract the underlying classes of the binned flood maps, which were subsequently overlaid on the binary validation data, to calculate the observed proportion of wet cells in each bin. Further, reliability diagrams were constructed individually for each land cover class and WRMSE values plotted for all the different approaches. Notably, the spatial average values of WRMSE in Figs. 13 and 17 differ from the WRMSE values obtained in Fig. 18, due to the calculation of case individual values for each land-use class. Based on the pixel count in each class, the values were found to be significantly different from the global spatial average. Although, if a weighted average based on the percentage distribution of the land-use classes was considered as shown in the pie diagrams of Fig. 18, the statistics were in agreement.

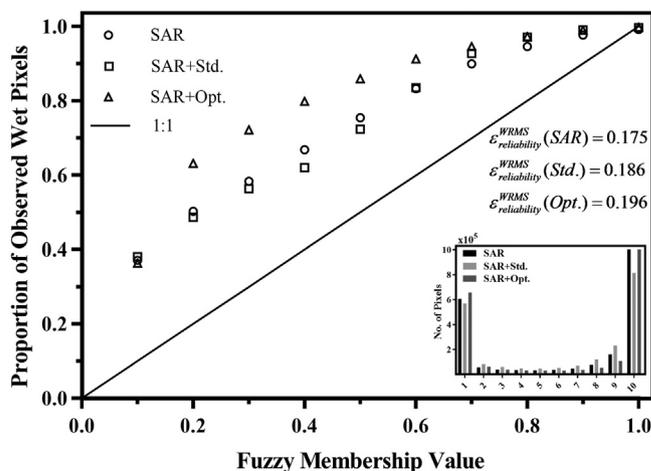


Fig. 17. As for Fig. 13 but for Ulmarra.

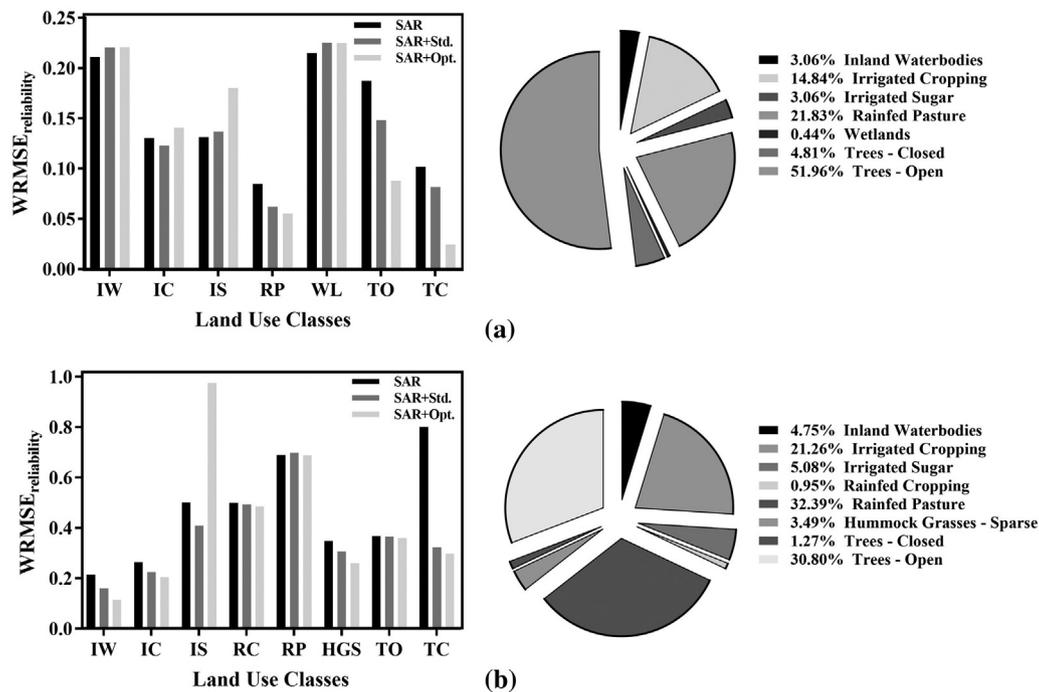


Fig. 18. Distribution of land cover classes and the corresponding Weighted Root Mean Squared Error values obtained for each of the SAR-based flood mapping techniques at (a) Junction Hill and (b) Ulmarra.

4.4.1. Validation site 1: Junction Hill

The Junction Hill area is dominated by the “Trees - Open (TO)” land cover class, followed by the “Rain-fed Pasture (RP)” and “Irrigated Cropping (IC)” classes. According to Fig. 15(a), WRMSE values had a significant reduction in the RP and TO classes, which might be the reason for the algorithm performing well in this region. The image region displaying a visible reduction in uncertainty through texture optimization, observable in the fuzzy flood maps illustrated in Fig. 9, is predominantly covered with the TO land cover class.

In the IC class however, classification performance was relatively poor due to the altered backscatter characteristics caused by emergent vegetation. The stems cause double bounce scattering of microwaves at the irrigated water surface making flood identification from SAR images in these particular land covers challenging. While the other classes may not contribute much to the overall RMSE values they provide crucial insight into classifier behaviour. The “Irrigated Sugar (IS)” shows a sharp increase in WRMSE with the proposed approach for the same reason as IC.

The water dominated classes like “Inland Water bodies (IW)” and “Wetlands (WL)” also showed a slight degradation in performance on texture addition. The contribution of the water based classes is considered limited in this context due to the land cover composition and the small magnitude of the increase. However, this suggests that the slight increase in WRMSE for the Ulmarra region on texture addition may have been caused by the majority of pixels being flooded with 77.6% wet cells. As the Junction Hill region has only 36.7% coverage of flooded area in the aerial photo, the improvement is more noticeable in the mixed land cover classes. This implies that if permanent water can be masked prior to classification, the results may be further improved.

4.4.2. Validation site 2: Ulmarra

Major classes in this region are TO, RP, and IC as evident from Fig. 15(b), none of which seemed to benefit especially from the random addition of texture. IC seems to show a slight improvement over the poor performance at site 1, but as the crop species are unspecified, the difference in RMSE values can be assumed to be a function of plant morphological properties. TO and RP showed almost no change in

WRMSE values for all three approaches. As most of the pixels at this site are inundated, the approach may not be able to demonstrate substantial improvement due to scattering of X-band from emergent vegetation. Conversely, the proposed algorithm was more effective in these land cover classes at Junction Hill, as the false alarm rate was reduced. Other classes in the region include “Hummock Grasses - Sparse (HGS)”, “Rainfed Cropping (RC)”, TC, and IS, of which all except IS exhibited a slight reduction in RMSE for the proposed approach, with TC showing the maximum decrease. IS showed an improvement in classification accuracy with texture addition but a degradation after texture optimization. However, the percentage coverage of the minority classes is insufficient to cause a quantifiable impact on the overall RMSE.

5. Conclusions and future scope

5.1. Conclusions

While texture is often utilized in SAR based flood mapping approaches it has seldom been optimized within this context. An approach to optimize omnidirectional GLCM-based image features, derived using the range of the spatial autocorrelation as the window size, has been proposed through an Independent Component Transform. The optimized texture bands were added to the SAR image pre-classification using a Gaussian Neurofuzzy classifier, which resulted in reduced uncertainties. The classification performance was evaluated against the addition of arbitrarily selected texture features, and without any texture addition to SAR. Results indicate that the texture optimization approach was able to extract the most useful texture information, showing drastic error reductions over the other approaches tested.

The proposed approach was tested at two validation sites situated in the Clarence catchment, Australia - Junction Hill and Ulmarra - for which aerial photographs were available close to the time of acquisition of one of the SAR images. Accuracy was evaluated using a fuzzy set approach and through reliability diagrams which are based on pixel count ratios. Pixels for individual land cover classes were subsequently extracted and the classification performance within each class quantified. This led to the following findings:

1. Texture based image enhancements represent a viable approach to improve flood mapping from single SAR images, leading to a reduction in uncertainties particularly in water lookalike regions e.g. smooth urban surfaces like tarmac. For example, the addition of optimized texture ensured that linear urban features like roads were not wrongly detected as flooded in the Junction Hill region.
2. The addition of optimized texture also led to decreased classifier sensitivity due to speckle noise, corroborated by a notable shift in pixels from uncertain to more certain bin values in the reliability diagrams. Interestingly, the study also demonstrated that the use of randomly selected texture features could actually degrade flood mapping accuracy, highlighting the importance of the proposed texture optimization.
3. The neurofuzzy classifier was not only able to characterize backscatter uncertainties through the use of fuzzy membership functions but the model selection and testing phase also explicitly defined the classification uncertainty.
4. A fuzzy set validation approach was introduced for the first time to assess SAR-based flood maps. While these methods cannot replace more traditional approaches like reliability diagrams which provide a detailed assessment of over and under prediction, they can complement the evaluation strategy to make the validation more robust. Moreover, the fuzzy set approach presents a unique opportunity to represent the uncertainties in the validation data and measure the performance while remaining cognizant of these errors.

5.2. Future scope

The proposed texture optimization and neurofuzzy classification approach improved flood detection capabilities when only a single SAR image of the event is available. This assumption of data availability is typically true in most operational cases. Although the use of a supervised classification technique requires the use of manually selected training data at the moment, this too can be automated by using a combination of image segmentation, thresholding and region growing, by extending methods proposed by Matgen et al. (2011). Once the semivariogram ranges have been precomputed for available SAR sensors, and the training process has been automated, the proposed approach could be operationally implemented. Moreover, if additional data are available, they can be utilized to further refine the resulting flood maps through some of the following methods:

1. In the case of some texture bands, the nature of fuzzy membership functions may deviate from the Gaussian distribution, and thus the sensitivity of this textured approach to the use of different fuzzy membership functions needs to be investigated.
2. Topographic information from the DEM can be added along with the texture bands, within the neurofuzzy classifier, to further increase the information content.
3. Contextual information about neighboring pixel classes can be expected to further enhance the classification outputs, and thus misclassifications or fuzziness within homogeneous image segments can be further reduced.
4. The potential of the fuzzy kappa statistic needs to be further demonstrated through testing different neighborhood influence functions, different resolution SAR images, and different algorithms.

This operational fuzzy flood mapping tool is available on GitHub as it can be useful to flood managers globally (see https://github.com/catantics13/Neurofuzzy_floodmapping_SAR).

Acknowledgments

This study has been conducted within the framework of the project “Improving flood forecast skill using remote sensing data”, funded by the Bushfire and Natural Hazards CRC of Australia. We would also like

to acknowledge E-Geos who provided the SAR imagery and their timely support during pre-processing. Additionally, we would also like to extend our gratitude towards Geoscience Australia and the Clarence Valley Council, for sharing field and ancillary data. Finally, we would like to thank the three anonymous reviewers of this manuscript, whose comments helped immensely to improve the quality of the manuscript.

References

- Amitrano, D., Di Martino, G., Iodice, A., Riccio, D., Ruello, G., 2018. Unsupervised rapid flood mapping using Sentinel-1 GRD SAR images. *IEEE Trans. Geosci. Remote Sens.* 1–10. <http://dx.doi.org/10.1109/TGRS.2018.2797536>.
- Balaguer, A., Ruiz, L.A., Hermosilla, T., Recio, J.A., 2010. Definition of a comprehensive set of texture semivariogram features and their evaluation for object-oriented image classification. *Comput. Geosci.* 36, 231–240. <http://dx.doi.org/10.1016/j.cageo.2009.05.003>.
- Berberoglu, S., Lloyd, C.D., Atkinson, P.M., Curran, P.J., 2000. The integration of spectral and textural information using neural networks for land cover mapping in the Mediterranean. *Comput. Geosci.* 26, 385–396. [http://dx.doi.org/10.1016/S0098-3004\(99\)00119-3](http://dx.doi.org/10.1016/S0098-3004(99)00119-3).
- Berberoglu, S., Curran, P.J., Lloyd, C.D., Atkinson, P.M., 2007. Texture classification of Mediterranean land cover. *Int. J. Appl. Earth Obs. Geoinf.* 9, 322–334. <http://dx.doi.org/10.1016/j.jag.2006.11.004>.
- Boni, G., Ferraris, L., Pulvirenti, L., Squicciarino, G., Pierdicca, N., Candela, L., Pisani, A.R., Zoffoli, S., Onori, R., Proietti, C., Pagliara, P., 2016. A prototype system for flood monitoring based on flood forecast combined with COSMO-SkyMed and Sentinel-1 data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 9, 2794–2805. <http://dx.doi.org/10.1109/JSTARS.2016.2514402>.
- Carr, J.R., 1996. Spectral and textural classification of single and multiple band digital images. *Comput. Geosci.* 22, 849–865. [http://dx.doi.org/10.1016/S0098-3004\(96\)00025-8](http://dx.doi.org/10.1016/S0098-3004(96)00025-8).
- Carr, J.R., De Miranda, F.P., 1998. The semivariogram in comparison to the co-occurrence matrix for classification of image texture. *IEEE Trans. Geosci. Remote Sens.* 36, 1945–1952. <http://dx.doi.org/10.1109/36.729366>.
- Chica-Olmo, M., Abarca-Hernández, F., 2000. Computing geostatistical image texture for remotely sensed data classification. *Comput. Geosci.* 26, 373–383. [http://dx.doi.org/10.1016/S0098-3004\(99\)00118-1](http://dx.doi.org/10.1016/S0098-3004(99)00118-1).
- Chini, M., Member, S., Hostache, R., Giustarini, L., Matgen, P., 2017. A hierarchical split-based approach for parametric thresholding of SAR images: flood inundation as a test case. *IEEE Trans. Geosci. Remote Sens.* 1–14.
- CRED, UNISDR, 2015. The Human Costs of Weather Related Disasters. The United Nations Office for Disaster Risk Reduction (UNISDR) <http://dx.doi.org/10.1017/CBO9781107415324.004>.
- 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. <http://dx.doi.org/10.1109/TGRS.2016.2520487>.
- De-yong, H., Xiao-juan, L., Wen-ji, Z., Hui-li, G., 2008. Texture analysis and its application for single-band SAR thematic information extraction. In: *IGARSS 2008 - 2008 IEEE Int. Geosci. Remote Sens. Symp. II-935-II-938*, <http://dx.doi.org/10.1109/IGARSS.2008.4779149>.
- 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. <http://dx.doi.org/10.1016/j.jhydrol.2009.01.020>.
- Di Baldassarre, G., Schumann, G., Brandimarte, L., Bates, P., 2011. Timely low resolution SAR imagery to support floodplain modelling: a case study review. *Surv. Geophys.* 32, 255–269. <http://dx.doi.org/10.1007/s10712-011-9111-9>.
- Franklin, S.E., Wulder, M.A., Lavigne, M.B., 1996. Automated derivation of geographic window sizes for use in remote sensing digital image texture analysis. *Comput. Geosci.* 22, 665–673. [http://dx.doi.org/10.1016/0098-3004\(96\)00009-X](http://dx.doi.org/10.1016/0098-3004(96)00009-X).
- García-Pintado, J., Neal, J.C., Mason, D.C., Dance, S.L., Bates, P.D., 2013. Scheduling satellite-based SAR acquisition for sequential assimilation of water level observations into flood modelling. *J. Hydrol.* 495, 252–266. <http://dx.doi.org/10.1016/j.jhydrol.2013.03.050>.
- García-Pintado, J., Mason, D.C., Dance, S.L., 2014. Moderation of Ensemble Covariances for Data Assimilation of Satellite-based Water Level Observations Into Flood Modelling. *vj. 16*, pp. 11618.
- Giustarini, L., Matgen, P., Hostache, R., Montanari, M., Plaza, D., Pauwels, V.R.N., De Lannoy, G.J.M., De Keyser, R., Pfister, L., Hoffmann, L., Savenije, H.H.G., 2011. Assimilating SAR-derived water level data into a hydraulic model: a case study. *Hydrol. Earth Syst. Sci.* 15, 2349–2365. <http://dx.doi.org/10.5194/hess-15-2349-2011>.
- Giustarini, L., Hostache, R., Matgen, P., Schumann, G.J., 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. <http://dx.doi.org/10.1109/TGRS.2012.2210901>.
- Giustarini, L., Vernieuwe, H., Verwaeren, J., Chini, M., Hostache, R., Matgen, P., Verhoest, N.E.C.E.C., de Baets, B., 2015. Accounting for image uncertainty in SAR-based flood mapping. *Int. J. Appl. Earth Obs. Geoinf.* 34, 70–77. <http://dx.doi.org/10.1016/j.jag.2014.06.017>.
- 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, 6958–6969.
- 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. <http://dx.doi.org/10.1007/s10712-016-9378-y>.

- Haack, B., Bechdol, M., 1999. Multisensor remote sensing data for land use/cover mapping. *Comput. Environ. Urban. Syst.* 23, 53–69. [http://dx.doi.org/10.1016/S0198-9715\(99\)00003-4](http://dx.doi.org/10.1016/S0198-9715(99)00003-4).
- Haack, B., Bechdol, M., 2000. Integrating multisensor data and RADAR texture measures for land cover mapping. *Comput. Geosci.* 26, 411–421. [http://dx.doi.org/10.1016/S0098-3004\(99\)00121-1](http://dx.doi.org/10.1016/S0098-3004(99)00121-1).
- Hagen, A., 2003. Fuzzy set approach to assessing similarity of categorical maps. *Int. J. Geogr. Inf. Sci.* 17, 235–249. <http://dx.doi.org/10.1080/13658810210157822>.
- Hagen-Zanker, A., 2006. Map comparison methods that simultaneously address overlap and structure. *J. Geogr. Syst.* 8, 165–185. <http://dx.doi.org/10.1007/s10109-006-0024-y>.
- Hagen-Zanker, A., Straatman, B., Uljee, I., 2005. Further developments of a fuzzy set map comparison approach. *Int. J. Geogr. Inf. Sci.* 19, 769–785. <http://dx.doi.org/10.1080/13658810500072137>.
- Haralick, R., Shanmugan, K., Dinstein, I., 1973. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* <http://dx.doi.org/10.1109/TSMC.1973.4309314>.
- Hastie, T., Tibshirani, R., Friedman, J., 2009. The elements of statistical learning. *Elements 1*, 337–387. <http://dx.doi.org/10.1007/b94608>.
- He, D.-C., Wang, L., 1991. Texture features based on texture spectrum. *Pattern Recogn.* 24, 391–399. [http://dx.doi.org/10.1016/0031-3203\(91\)90052-7](http://dx.doi.org/10.1016/0031-3203(91)90052-7).
- Horritt, M.S., 2006. A methodology for the validation of uncertain flood inundation models. *J. Hydrol.* 326, 153–165. <http://dx.doi.org/10.1016/j.jhydrol.2005.10.027>.
- Hostache, R., Schumann, G., Matgen, P., Puech, C., Hoffmann, L., Pfister, L., 2006. 3D flood information from SAR as a mean for reducing uncertainties in flood inundation modelling. In: XXX ISPRS Congress, Commission VII, pp. 217–222.
- Hostache, R., Matgen, P., Schumann, G., Puech, C., Hoffmann, L., Pfister, L., 2009. Water level estimation and reduction of hydraulic model calibration uncertainties using satellite SAR images of floods. *IEEE Trans. Geosci. Remote Sens.* 1–10.
- Hostache, R., Lai, X., Monnier, J., Puech, C., 2010. Assimilation of spatially distributed water levels into a shallow-water flood model. Part II: Use of a remote sensing image of Mosel River. *J. Hydrol.* 390, 257–268. <http://dx.doi.org/10.1016/j.jhydrol.2010.07.003>.
- Hostache, R., Corato, G., Chini, M., Wood, M., Giustarini, L., 2015. A New Approach for Improving Flood Model Predictions Based on the Sequential Assimilation of SAR-derived Flood Extent Maps. vol. 17, pp. 10121.
- Italian Space Agency, 2009. COSMO-SkyMed SAR Products Handbook. pp. 105.
- James, G., Witten, D., Hastie, T., Tibshirani, R., 2000. An introduction to statistical learning. *Curr. Med. Chem.* <http://dx.doi.org/10.1007/978-1-4614-7138-7>.
- Jang, J.S.R., 1993. ANFIS: adaptive-network-based fuzzy inference system. *IEEE Trans. Syst. Man Cybern.* 23, 665–685. <http://dx.doi.org/10.1109/21.256541>.
- Kuplich, T.M., Curran, P.J., Atkinson, P.M., 2005. Relating SAR image texture to the biomass of regenerating tropical forests. *Int. J. Remote Sens.* 26, 4829–4854. <http://dx.doi.org/10.1080/01431160500239107>.
- Lai, X., Monnier, J., 2009. Assimilation of spatially distributed water levels into a shallow-water flood model. Part I: mathematical method and test case. *J. Hydrol.* 377, 1–11. <http://dx.doi.org/10.1016/j.jhydrol.2009.07.058>.
- Lai, X., Liang, Q., Yesou, H., Daillet, S., 2014. Variational assimilation of remotely sensed flood extents using a 2-D flood model. *Hydrol. Earth Syst. Sci.* 18, 4325–4339. <http://dx.doi.org/10.5194/hess-18-4325-2014>.
- Long, S., Fatoyinbo, T.E., Policelli, F., 2014. Flood extent mapping for Namibia using change detection and thresholding with SAR. *Environ. Res. Lett.* 9, 35002–35009. <http://dx.doi.org/10.1088/1748-9326/9/3/035002>.
- Lymburner, L., Tan, P., Mueller, N., Thackway, R., Lewis, A., Thankappan, M., Randall, L., Islam, A., Senarath, U., 2011. The National Dynamic Land Cover Dataset. Geoscience Australia, Symonston, Australia.
- Marquardt, D.W., 1963. An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Ind. Appl. Math.* 11, 431–441.
- Martinis, S., Kuenzer, C., Twele, A., 2015. Flood studies using synthetic aperture radar data. In: *Remote Sens. Water Resour. Disasters, Urban Stud.* pp. 145–173. <http://dx.doi.org/10.1201/b19321-10>.
- Mason, D.C., Schumann, G.J.P., Bates, P.D., 2010. Data utilization in flood inundation modelling. In: *Flood Risk Science and Management*, pp. 209–233. <http://dx.doi.org/10.1002/9781144432484.ch11>.
- Mason, D.C., Schumann, G.J.P., Neal, J.C., Garcia-Pintado, J., Bates, P.D., 2012. Automatic near real-time selection of flood water levels from high resolution Synthetic Aperture Radar images for assimilation into hydraulic models: a case study. *Remote Sens. Environ.* 124, 705–716. <http://dx.doi.org/10.1016/j.rse.2012.06.017>.
- Matgen, P., Montanari, M., Hostache, R., Pfister, L., Hoffmann, L., Plaza, D., Pauwels, V.R.N., De Lannoy, G.J.M., De Keyser, R., Savenije, H.H.G., 2010. Towards the sequential assimilation of SAR-derived water stages into hydraulic models using the Particle Filter: proof of concept. *Hydrol. Earth Syst. Sci.* 14, 1773–1785. <http://dx.doi.org/10.5194/hess-14-1773-2010>.
- 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. *Phys. Chem. Earth* 36, 241–252. <http://dx.doi.org/10.1016/j.pce.2010.12.009>.
- O'Grady, D., Leblanc, M., Gillieson, D., 2013. Relationship of local incidence angle with satellite radar backscatter for different surface conditions. *Int. J. Appl. Earth Obs. Geoinf.* 24, 42–53. <http://dx.doi.org/10.1016/j.jag.2013.02.005>.
- O'Grady, D., Leblanc, M., Bass, A., 2014. The use of radar satellite data from multiple incidence angles improves surface water mapping. *Remote Sens. Environ.* 140, 652–664. <http://dx.doi.org/10.1016/j.rse.2013.10.006>.
- Ouled Sghaier, M., Hammami, I., Foucher, S., Lepage, R., 2018. Flood extent mapping from time-series SAR images based on texture analysis and data fusion. *Remote Sens.* 10, 237. <http://dx.doi.org/10.3390/rs10020237>.
- Pappenberger, F., Beven, K., Frodsham, K., Romanovicz, R., Matgen, P., 2007a. Grasping the unavoidable subjectivity in calibration of flood inundation models: a vulnerability weighted approach. *J. Hydrol.* 333, 275–287. <http://dx.doi.org/10.1016/j.jhydrol.2006.08.017>.
- Pappenberger, F., Frodsham, K., Beven, K.J., Romanovicz, R., Matgen, P., 2007b. Fuzzy set approach to calibrating distributed flood inundation models using remote sensing observations. *Hydrol. Earth Syst. Sci.* 11, 739–752. <http://dx.doi.org/10.5194/hess-11-739-2007>.
- 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, 4151–4164. <http://dx.doi.org/10.3390/s8074151>.
- Power, C., Simms, A., White, R., 2001. Hierarchical fuzzy pattern matching for the regional comparison of land use maps. *Int. J. Geogr. Inf. Sci.* 15, 77–100. <http://dx.doi.org/10.1080/136588100750058715>.
- Pradhan, B., Hagemann, U., Shafapour Tehrani, M., Prechtel, N., 2014. An easy to use ArcMap based texture analysis program for extraction of flooded areas from TerraSAR-X satellite image. *Comput. Geosci.* 63, 34–43. <http://dx.doi.org/10.1016/j.cageo.2013.10.011>.
- Pradhan, B., Tehrani, M.S., Jebur, M.N., 2016. A new semiautomated detection mapping of flood extent from TerraSAR-X satellite image using rule-based classification and Taguchi optimization techniques. *IEEE Trans. Geosci. Remote Sens.* 54, 4331–4342. <http://dx.doi.org/10.1109/TGRS.2016.2539957>.
- Pulvirenti, L., Pierdicca, N., Chini, M., Guerriero, L., 2011. An algorithm for operational flood mapping from Synthetic Aperture Radar (SAR) data using fuzzy logic. *Nat. Hazards Earth Syst. Sci.* 11, 529–540. <http://dx.doi.org/10.5194/nhess-11-529-2011>.
- 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 J. Sel. Top. Appl. Earth Obs. Remote Sens.* 6, 1807–1816. <http://dx.doi.org/10.1109/JSTARS.2012.2219509>.
- Pulvirenti, L., Marzano, F.S., Pierdicca, N., Mori, S., Chini, M., 2014. Discrimination of water surfaces, heavy rainfall, and wet snow using COSMO-SkyMed observations of severe weather events. *IEEE Trans. Geosci. Remote Sens.* 52, 858–869. <http://dx.doi.org/10.1109/TGRS.2013.2244606>.
- Refice, A., Capolongo, D., Pasquariello, G., Daaddabbo, A., Bovenga, F., Nutricato, R., Lovergine, F.P., Pietranera, L., 2014. SAR and InSAR for flood monitoring: examples with COSMO-SkyMed data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 7, 2711–2722. <http://dx.doi.org/10.1109/JSTARS.2014.2305165>.
- 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. <http://dx.doi.org/10.1016/j.jag.2016.12.003>.
- Schumann, G.J.P., Di Baldassarre, G., 2010. The direct use of radar satellites for event-specific flood risk mapping. *Remote Sens. Lett.* 1, 75–84. <http://dx.doi.org/10.1080/0143116093486685>.
- Schumann, G., Cutler, M., Black, A., Matgen, P., Pfister, L., Hoffmann, L., Pappenberger, F., 2008. Evaluating uncertain flood inundation predictions with uncertain remotely sensed water stages. *Int. J. River Basin Manag.* 5124, 37–41. <http://dx.doi.org/10.1080/15715124.2008.9635347>.
- Schumann, G., Bates, P.D., Horritt, M.S., Matgen, P., Pappenberger, F., 2009a. Progress in integration of remote sensing-derived flood extent and stage data and hydraulic models. *Rev. Geophys.* 47, 1–20. <http://dx.doi.org/10.1029/2008RG000274>.
- Schumann, G., Di Baldassarre, G., Bates, P.D., 2009b. The utility of spaceborne radar to render flood inundation maps based on multialgorithm ensembles. *IEEE Trans. Geosci. Remote Sens.* 47, 2801–2807. <http://dx.doi.org/10.1109/TGRS.2009.2017937>.
- Schumann, G.J.P., Bates, P.D., Di Baldassarre, G., Mason, D.C., 2012. The use of radar imagery in riverine flood inundation studies. In: *Fluvial Remote Sensing for Science and Management*, pp. 115–140. <http://dx.doi.org/10.1002/9781119940791.ch6>.
- Schumann, G.J.P., Vernieuwe, H., De Baets, B., Verhoest, N.E.C., 2014. ROC-based calibration of flood inundation models. *Hydrol. Process.* 28, 5495–5502. <http://dx.doi.org/10.1002/hyp.10019>.
- Senthilnath, J., Shenoy, H.V., Rajendra, R., Omkar, S.N., Mani, V., Diwakar, P.G., 2013. Integration of speckle de-noising and image segmentation using Synthetic Aperture Radar image for flood extent extraction. *J. Earth Syst. Sci.* 122, 559–572. <http://dx.doi.org/10.1007/s12040-013-0305-z>.
- Smith, L.C., 1997. Satellite remote sensing of river inundation area, stage, and discharge: a review. *Hydrol. Process.* 11, 1427–1439. [http://dx.doi.org/10.1002/\(sici\)1099-1085\(199708\)11:10<1427::aid-hyp473>3.0.co;2-s](http://dx.doi.org/10.1002/(sici)1099-1085(199708)11:10<1427::aid-hyp473>3.0.co;2-s).
- Takagi, T., Sugeno, M., 1985. Fuzzy identification of systems and its applications to modeling and control. In: *Syst. Man Cybern.* IEEE Trans. SMC-15, pp. 116–132. <http://dx.doi.org/10.1109/TSMC.1985.6313399>.
- 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. <http://dx.doi.org/10.1080/01431161.2016.1192304>.
- Ulaby, F., Kouyate, F., Brisco, B., Williams, T.H., 1986. Textural information in SAR images. *IEEE Trans. Geosci. Remote Sens.* GE-24, 235–245. <http://dx.doi.org/10.1109/TGRS.1986.289643>.
- Voormansik, K., Praks, J., Antropov, O., Jagomagi, J., Zalite, K., 2014. Flood mapping with TerraSAR-X in forested regions in Estonia. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 7, 562–577. <http://dx.doi.org/10.1109/JSTARS.2013.2283340>.
- Wealands, S.R., Grayson, R.B., Walker, J.P., 2005. Quantitative comparison of spatial fields for hydrological model assessment - some promising approaches. *Adv. Water Resour.* 28, 15–32. <http://dx.doi.org/10.1016/j.advwatres.2004.10.001>.
- Wood, M., Hostache, R., Neal, J., Wagener, T., Giustarini, L., Chini, M., Corato, G., Matgen, P., Bates, P., 2016. Calibration of channel depth and friction parameters in the LISFLOOD-FP hydraulic model using medium-resolution SAR data and identification techniques. *Hydrol. Earth Syst. Sci.* 20, 4983–4997. <http://dx.doi.org/10.5194/hess-20-4983-2016>.
- Zadeh, L.A., 1965. Fuzzy sets. *Inf. Control.* 8, 338–353. [http://dx.doi.org/10.1016/S0019-9958\(65\)90241-X](http://dx.doi.org/10.1016/S0019-9958(65)90241-X).