Gagnon, AS and Quirós, E

Validation of flood risk maps using Open source optical and radar satellite imagery

Validation of flood risk maps using Open source optical and radar satellite imagery

ABSTRACT

Flood risk maps delineate areas potentially at risk of flooding and are thus a crucial tool in flood risk management. In Spain, such maps are provided as Open geospatial data. This paper compares the flood prone areas according to those maps for floods of different return periods with the spatial extent of two floods that severely affected southwestern Spain using Open source optical and radar satellite imagery. Results using the *recall* metric were found to be very good and are the most relevant to emergency preparedness services as this metric focuses on the correctly predicted flooded areas. It was also found that the 500-year flood risk map was the one with the best precision and accuracy for both flood events. The results confirm the accuracy of the flood risk maps based on Open source remote sensing imagery and hence demonstrate the potential of publicly available and freely distributed remote sensing data in their development.

Keywords: Floods, Open geospatial data, remote sensing, risk maps, Spain, validation

1. INTRODUCTION

Flooding is a major issue in many countries, affecting many communities and livelihoods (Olsson, Opondo, Tschakert, Agrawal, & Eriksen, 2014), and, despite major investments in flood protection schemes it remains a major issue in Europe (Kundzewicz et al., 2014). Many parts of Europe have been affected by floods in recent decades, causing loss of life and damages in the order of multiple billions of euros (*Changes in Flood Risk in Europe*, (2012).; Kundzewicz, Pińskwar, & Brakenridge, 2013). In Spain, the economic impacts of floods have increased in recent decades, a trend that some have associated with a changing climate (e.g. Barredo, Saurí, and Llasat (2012)). With the magnitude of floods projected to increase under climate change for the Iberian Peninsula (Roudier et al. (2016); Alfieri, Dottori, Betts, Salamon, and Feyen (2018)), further damage and economic disruption caused by flooding are expected. A national flood risk assessment[1], which was conducted as a result of Spanish Royal Decree 903/2010 July 9 2013 that followed Directive 2007/60/EC 23 October 2007 of the European Parliament on the assessment and management of flood risks, identified more flood prone areas in Spain than any other country of the European Union. Many parts of the autonomous community of Extremadura in Spain, and more specifically the province of Badajoz, given its geographical characteristics, are at risk of flooding, notably the provincial capital city of the same name as the province.

Flood risk maps delineate areas prone to flooding and are thus a crucial tool for flood risk management. For instance, these maps, by identifying areas at risk of flooding during extreme precipitation events, can be used to prioritize adaptive actions and to

[1] http://www.consorsegurosdigital.com/en/numero-03/front-page/flood-risk-management-plans

inform local planning policy. Flood risk maps are produced using a modelled simulation of a flood as a function of precipitation (de Moel, van Alphen, & Aerts, 2009) and are thus limited to precipitation-generated floods such as flash floods and urban floods and do not include the risk of flooding due to accumulated debris in watercourses, for instance. Hence, historical rainfall data and a hydrological model incorporating a Digital Elevation Model (DEM) are the main data type and technique used in the production of flood risk maps, but these can also be supplemented with studies or reports of previous flooding episodes.

Even though there has been increasing use of remote sensing techniques to provide areal estimates of precipitation in data poor regions (Black et al., 2016), and as a result of a decline in the number of gauging stations in many parts of the world (Domeneghetti, Schumann, & Tarpanelli, 2019), the use of remote sensing techniques to map flood prone areas or for validating flood risk maps has to date been limited. Bates (2012) stated that the incorporation of remote sensing data into flood prediction systems could improve forecast accuracy. Accordingly, Di Baldassarre, Schumann, Brandimarte, and Bates (2011) described the potential of freely available remote sensing data to support near real-time modelling of a flood. The potential of using remote sensing and particularly Synthetic Aperture Radar (SAR) images for the monitoring of a flood was also demonstrated by Dewan, Kankam-Yeboah, and Nishigaki (2006) for the city of Dhaka in Bangladesh. However, uncertainties were found when using SAR images in the production of flood extent maps (Di Baldassarre et al., 2011). Although satellite data do not yet appear to be capable of completely substituting in-situ observations of water level, they have the potential to become a valuable source of information in flood risk management (Domeneghetti et al., 2014), such as for validating the reliability of flood risk maps.

2. OBJECTIVES

The aim of this paper is to examine the potential use of remote sensing technology in flood risk mapping. In particular, it validates current flood risk maps by comparing them with the spatial extent of two historical floods as determined using optical remote sensing and radar imagery and compared it with current flood risk maps. Using Open source remote sensing and hydrological data, this paper thus presents the potential for incorporating Open source remote sensing data and their analysis in a Geographical Information System (GIS) environment in the delineation of flood prone areas.

3. STUDY AREA AND DATA

3.1. Study area

The study area is the city of Badajoz, the capital city of the Province of Badajoz in Spain (Figure 1). It is situated close to the border with Portugal, on the left bank of the river Guadiana. The population of the city was 150,543 in 2017.

The first step in the implementation of the EU Floods Directive is a preliminary flood risk assessment. This assessment, identified 82,432,574 m^2 of the catchment of the Guadiana River, which flows through the city of Badajoz, at risk of flooding (Figure 1). However, this study is limited to 33,923,660 m^2 of this area, which covers a 6 km radius around the city of Badajoz and includes the flood risk areas of tributaries of the Guadiana River, i.e., the Rivillas, Gévora and Caja Rivers, which have flood risk areas of 10,194167 m^2, 3,444,777 m^2 and 2,411,576 m^2, respectively. There are 8825 inhabitants living in the study area, and floods have previously affected 2648 of them. The catchment is covered by soils within Group C of the hydrologic soil groups (United States Department of Agriculture Natural Resources Conservation Service, 2007), because water transmission through the soil is restricted and hence these soils have a slow infiltration rate. Figure 1 depicts the spatial extent of the population at risk of flooding.

3.2. Flood Risk maps

Following the principles of Directive 2007/60 on flood risk assessment and management, the Government of Spain launched the National Flood Zone Mapping System (NFZMS) to support river space management, risk prevention, territorial planning and for transparency in administration. This supporting tool includes flood risk maps provided as Open geospatial data (Quirós & Polo, 2018), depicting the risk of flooding according to five probability scenarios:

1) Extremely high probability of flooding associated with a return period of 5 years.

2) Very high probability of flooding associated with a return period of 10 years.

3) High probability of flooding associated with a return period of 50 years.

4) Average probability of flooding associated with a return period of 100 years.

5) Low probability of flooding or scenario of extreme events associated with a return period of 500 years.

The production of flood risk maps is complex, requiring modelling and data from various sources. It requires a DEM of the river basin and the river section, an updated orthophoto (i.e., an aerial photograph geometrically corrected to a uniform scale), both of which at the best resolution available. The production of the flood risk maps, also uses georeferenced historical aerial photos, cartography of the dimensions of the elements or infrastructures located in the study area that may affect flooding, such as bridges, specks and channelling, information on elements located upstream and downstream of the study area that can help define the boundary conditions or simulation edge, such as sea level and reservoir levels, as well as information on land use to determine water losses and maximum flow rates data. The verification of these flood risk maps has been recommended using historical aerial photographs and a reconstruction of historical flood series using surveys and other historical information (Martínez, Moreno, García-Oliva, & Olona, 2011). This is a challenging task and remote sensing, which is not currently used in the development or validation of those maps at government level, potentially presents an opportunity to facilitate this, as well as for gathering the information required for the production of those maps.

3.3. Flood events

The identification of flood events was first based on river discharge time series. Discharge data for the Guadiana River were obtained from an Open database of the *Centro de Estudios y Experimentación de Obras Públicas* (CEDEX), the Centre for

Water Studies or Centre for Studies and Experiments on Public Works in English. Figure 2 illustrates the discharge of the Guadiana River at *Azud de Badajoz* (station 4030), which is situated downstream of Badajoz. The figure shows a number of peak discharge events, which were used to identify potential flood events, as well as the availability of Open source optical and radar satellite images during or near the time occurrence of the peak discharge events.

Two flooding events were selected on the basis of the availability of Open source optical and radar images during or near the occurrence of the peak discharge events, which correspond to the second and third highest peak discharge during the recent recording period. The flood associated with the highest peak discharge during the study period was not selected due to the lack of both optical and SAR images around the timing of the flood. Nonetheless, it was mentioned in the local press that authorities were better at controlled this peak discharge and, for this reason, it did not cause as much damage to the area as the other two selected flooding episodes.

The first flood occurred on November 6, 1997, when the river discharge peaked at the 3089 m^3/s (Figure 2). It was an important flood affecting the area in terms of fatalities, 23 people died, and it caused extensive damage to property (Lorente, Hernández, Queralt, & Ribera, 2008), as it affected a low-income area of the municipality, with low adaptive capacity against events of such magnitude (Olcina, 2016). The second flood occurred on March 8, 2010. Even though the discharge during that event reached 3060 m^3/s and was of a magnitude similar to the first selected flood, it was not as reported and studied as the 1997 flood but still caused the death of a child as well as extensive material damage. The magnitude of the peak discharge during those two floods is just slightly higher than what would be expected for floods with a 10-year return period according to the Spanish map of maximum flows (Álvarez, 2010). Nonetheless, the damage that they caused was extraordinary, particularly for the 1997 flood. The Special Civil Protection Plan for Flood Risks of the Autonomous Community of Extremadura stated that the main cause for the 1997 flood was two tributaries (the Rivillas and Calamón rivers) that reached a discharge of a 500-year return period.

3.4. Remote sensing images

For the flood of 1997, a number of Open source remote sensing images were available. The first remote sensing image that was available subsequent to the occurrence of the flood was an optical image. This image was nearly free of clouds and was taken two days following the flood (Figure 2). The first available Open source image following the 2010 flood event, for its part, is a SAR image, which was taken eight days after the flood.

4. MAPPING FLOODED AREAS

The methodology for mapping the flooded areas differed between the optical and SAR images (figure 3) but all shown operations were implemented by means of the Open source SeNtinel Application Platform (SNAP) and the Open source GIS software QGIS.

4.1. Optical image

For the 1997 flood, optical data were used to delineate the flooded areas. The image was acquired by the Landsat 5 satellite, which was taken only two days following the flood. This Landsat 5 image has seven spectral bands ranging from 0.45 to 2.35 microns and a spatial resolution of 30 m in all bands with the exception of the thermal band.

The processing and analysis of the satellite images using the SNAP software required the Landsat images in the different bands to be combined together. As the Landsat images are georeferenced, layer stacking, which combines the images from the different Landsat bands into a single image, was the only pre-processing step required (Figure 3). The study area was then extracted from the satellite image using a geographical subset.

Dao and Liou (2015) previously demonstrated the use of Normalized Difference Vegetation Index (NDVI) to support the detection of flooded areas using Landsat 8 images. Following their approach, the NDVI was calculated on the subset of the larger satellite image over the study region using the following equation:

$$NDVI = \frac{\mathrm{B4} - \mathrm{B3}}{\mathrm{B4} + \mathrm{B3}}$$

As NDVI compares the reflectance in the red (B3) and near-infrared bands (B4), and that water absorbs the energy in the red band, clear water has a negative NDVI value (-1), its reflectance is nil, and thus appears black in the processed satellite image (Figure 4). The resulting NDVI image was thus stacked as a new band to improve the classification (Figure 3).

After stacking the NDVI image, an unsupervised classification was performed using k-means clustering, in which data in each subset (ideally) share some common trait for the post classification depuration, often proximity according to some defined distance measure. The k-means clustering technique is a simple partitional clustering algorithm that attempts to find k non-overlapping clusters. These clusters are represented by their centroids (mean) (Wu, 2012). Suppose D={x_1, ⋯ , x_n}is the data set to be clustered . K-means can be expressed by the following function, which depends on the proximities of the data points to the cluster centroids:

$$min_{\{m_k\},1\leq k\leq K} \sum_{k=1}^{k} \sum_{x\in C_k} \pi_x dist(x, m_k)$$

where π_x is the weight of x and n_k is the number of data objects assigned to cluster C_k, and $m_k = \sum_{x\in C_k} \frac{\pi_x x}{n_k}$

The flooded and non-flooded areas were then identified as part of the post-classification depuration, as well as the clouds and their corresponding shadows. The latter two elements are important as it is necessary to filter out the clouded areas when it comes to the validation of the flood risk map, as presented in section 5.

4.2. SAR image

The second flooding episode, which took place in 2010, was analysed using a SAR image available eight days following the flood. The only Open source optical image available near the timing of the flood was taken before the flood. For this reason, the delineation of the flooded and non-flooded areas using the procedure described in section 4.1 could not be followed, and an alternative methodology using a SAR image was used for that purpose. The SAR image originates from the European Remote Sensing (ERS)-2 satellite, which operates mainly in a 35-day repeat cycle. ERS-2 images have a spatial resolution of 26 m in range (across track) and between 6 and 30 m in azimuth (along track) and a single VV polarization mode. Di Baldassarre et al. (2011) previously used these SAR images obtained from the ERS satellite to map the extent of a flood with 2 m accuracy.

Figure 3 depicts the processing of the SAR image. First, the image was pre-processed using three steps, consisting of calibration, speckle-filter, and terrain correction:

a) Calibration

Calibration was performed using the Sigma Nought method (σ^0). The radar backscattering coefficient σ^0 was defined by (Laur et al., 2003) and is related to the radar brightness β^0 as follows:

$$\sigma^0 = \beta^0 \cdot \sin\alpha$$

where α is the local incidence angle.

Then, the new Sigma Nought calibrated digital number (DN) of a given pixel was given by:

$$[DN]^2 = C \cdot \beta^0 = C \cdot \frac{\sigma^0}{\sin\alpha}$$

b) Speckle- filter

A Lee sigma speckle was used to reduce the speckle effect. Speckles are produced by random interferences of the de-phased, although coherent, return waves scattered by the elementary scatters within each pixel.

The Lee filter utilize the statistical distribution of the DN values within the moving kernel to estimate the value of the pixel of interest (Mansourpour, Rajabi, & Blais, 2006).

$$DN_{out} = [mean] + K \cdot [DN_{in} - mean]$$

where $K = \frac{\left(\frac{variance+mean^2}{\sigma^2+1}\right)-mean^2}{mean^2 \cdot \sigma^2+\left(\frac{variance+mean^2}{\sigma^2+1}\right)-mean^2}$ and the variance and mean are calculated in a specific window.

c) Terrain correction

Terrain correction was done using the NASA Shuttle Radar Topographic Mission (SRTM) DEM. Images that are not directly at the SAT sensor's Nadir location will have some distortion and terrain corrections intend to correct these.

This terrain correction procedure, consisted of an orthorectification which was based on available orbit state vector information in the metadata or external precise orbit, the radar timing annotations, and the slant to ground range conversion parameters together with the reference DEM.

The procedure was defined by (Small & Schubert, 2008) and consisted of an orthorectification employing "backward geocoding" that takes a slant range and azimuth time as input and uses the same geolocation equations to determine a map geometry. After geolocation, once the range and azimuth indices were determined, a value was extracted from the input image content using the user-selected resampling kernel, and output in the DEM's map geometry.

Once the SAR image was pre-processed, the study area was extracted from the whole scene following the same subset area used for the optical Landsat scene.

Second, as illustrated in Figure 3, two processes were performed in order to improve the SAR image for the classification. On the one hand, a texture analysis was done by means of a Grey Level Co-occurrence Matrix (GLCM). This technique has been widely employed in previous research, e.g. Shanmugan, Narayanan, Frost, Stiles, and Holtzman (1981), Pradhan, Hagemann, Shafapour Tehrany, and Prechtel (2014), Dasgupta, Grimaldi, Ramsankaran, Pauwels, and Walker (2018), and proved to produce meaningful results, notably by improving the lack of spectral information in a SAR single band image (Pulvirenti, Chini, Pierdicca, Guerriero, & Ferrazzoli, 2011).

The texture analysis, defined as correlation texture in Haralick, Shanmugam, and Dinstein (1973), measures the linear dependence of grey levels on those of between neighbouring pixels. There are several texture features can be computed from the GLCM matrix, e.g., mean, variance and correlation. Considering an image as rectangular with N_x resolution cells in the horizontal direction and N_y resolution cells in the vertical direction, the mean, variance, and correlation are computed using the following equations:

$$GLCM\ mean = \sum_{i=2}^{2N_g} i p_{x+y}(i)$$

$$GLCM\ variance = \sum_{i=2}^{2N_g} (i - entropy)^2 p_{x+y}(i)$$

$$GLCM\ correlation = \frac{\sum_i \sum_j (ij)p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$$

where

p(i,j) (i,j)*th* entry in a normalized gray-tone spatial dependence matrix, with which two neighbouring resolution cells separated by distance d occur on the image, one with gray tone i and the other with gray tone j.

$p_x(i)$ (i)th entry in the marginal-probability matrix obtained by summing the rows of p(i,j).

N_g is the number of distinct gray levels in the quantized image.

$$entropy = -\sum_{i=2}^{2N_g} p_{x+y}(i) \cdot log\left\{p_{x+y}(i)\right\}$$

Using the above formulas, the mean, variance and correlation were calculated and stacked to the main bands of the SAR image in order to improve the classification. As an example the obtained GLCM correlation band is presented in Figure 5 to show the improvement in definition in relation to the original band.

Finally, a terrain mask was also included as an additional step to determine areas on the map with steep slopes (Figure 3). This is accomplished using a DEM and has the purpose of separating the flat areas on the map, where flooding is more likely, from the mountainous areas.

The bands resulting from the above procedure were stacked together in order to compose a multiband image for the classification, but, before that, a coregistration with ground control points was necessary in order to increase the geometric accuracy of the image. This step is crucial for images that have been corrected without using orbit files during terrain correction.

In the same way as for the optical image, an unsupervised classification using K-means cluster analysis was performed to extract the flooded pixels from the study area. Li and Wang (2017) demonstrated that this unsupervised clustering applied to a single polarization SAR image is a suitable technique for flood mapping.

Because of the presence of speckle noise in the SAR image data, which was not an issue for the optical data, some post classification depuration was required to achieve greater uniformity of the results. For this reason, a filter based on clumping and sieving was applied to the image to deal with the problem of isolated pixels during the classification. This consisted of clumping adjacent similarly classified pixels together and segmentation to obtain regular results. This sieve filter was defined by (Bangham, Chardaire, Pye, & Ling, 1996) and works at several scales (one per each step of the process) and performs a decomposition by scale. At each stage the filtering operator removes pixels with extreme values of only that scale.

5. VALIDATION OF THE FLOOD RISK MAPS

The above methodological approach on the optical and SAR images resulted in the classification of flooded and non-flooded areas, which was then used for validating the Open access flood risk maps from the government. In order to perform a more precise validation, the hydraulic public domain delimitation was excluded from the study, i.e., the channel areas were excluded from the validation, as the interest was in identifying the areas submerged with water beyond the normal limits of the channel.

The methodology selected for the validation is based on Horritt (2006) for comparing uncertain maps of inundation extent with single observed events. This methodology was selected because extreme flood events are infrequent and, there is often limited availability of satellite images around the timing of a flood.

Different metrics were employed in the validation of the flood risk maps, that is in the comparison of the flood risk maps from government and those drawn on the basis of satellite imagery: measure of fit, accuracy, precision and recall. In all measures, *A* refers to the extent of the flooded area that is correctly predicted by the model, *B* is the area predicted as flooded but that is actually not flooded (i.e., over-prediction), *C* is the flooded area not predicted by the model (i.e., under-prediction) and *D* represents the correctly predicted non flooded areas.

5.1. Measure of fit

The measure of fit, *F*, is defined by Horritt (2006) as:

$$F = \frac{A - B}{A + B + C}$$

The values for this measure can range from -1 to 1. One weakness of this measure is that it ignores the correctly predicted non-flooded areas in its calculation, i.e., *D*.

5.2. Accuracy

This metric represents the ratio of the correctly predicted extent of a flood to the sum of the areas that are correctly predicted and those that are over-predicted. This metric ranges from 0 to 1, with 0 and 1 representing zero and perfect accuracy, respectively:

$$Accuracy = \frac{A}{A + B}$$

As for the *measure of fit*, this metric ignores the correctly predicted non-flooded areas in its calculation, but also the flooded area not predicted by the model.

5.3. Precision

Precision is the only metric that considers all variables in its calculation. As for the accuracy metric, the ratio for this metric ranges from 0 to 1, with 1 representing the best accuracy:

$$Precision = \frac{A + D}{A + C + B + D}$$

5.4. Recall

The measure *recall*, which is also known as the true positive rate and sensitivity is the ratio of the correctly predicted spatial extent of a flood to the sum of the correctly predicted flood extent and the under-predicted areas, the latter representing the areas they were in fact flooded but that the maps did not identify at risk of flooding. This measure also ranges from 0 to 1, with 1 representing perfect agreement between the flood risk map and the area considered as flooded in the satellite image:

$$Recall = \frac{A}{A + C}$$

6. RESULTS AND DISCUSSION

Figures 6 and 7 show the results of the comparison between the areas considered as submerged during the 1997 and 2010 floods, respectively, according to the satellite images and those considered at risk of being flooded in the official flood risk maps for floods of different return periods. Thus, these maps show the correctly predicted flooded areas, the areas that are under predicted and the ones that over predicted on the basis of the flood delineation done using the satellite images in comparison to the flood risk maps.

Figures 8 and 9 show the evolution of each metric ranging from a flood with a 5-year return period to one with a 500-year return period for the 1997 and 2010 floods, respectively.

The results of the validation of the areas delineated as flooded on the satellite images and those at risk of flooding using the *recall* metric were found to be very good, especially in 1997. This metric is considered to be the most relevant to emergency preparedness services as it focuses on the correctly predicted flooded areas. The values obtained using the recall metric for the 1997 flood are excellent, ranging from 0.75 to 0.98, and are significantly higher than the ones obtained for the 2010 flood. Moreover, the values of this metric increase with the return period of a flood, hence it is highest for a flood with a 500-year return period in both cases. The values of the recall metric for the validation of the 1997 flood in Badajoz are slightly higher than the ones obtained by Ouled Sghaier, Hammami, Foucher, and Lepage (2018).

The *precision* metric represents the overall accuracy of the flooded/non-flooded areas using satellite imagery. Unlike the recall statistic, it includes the correctly predicted non-flooded areas and not only the correctly predicted flooded areas. As can be seen in figures 8 and 9, the values of this metric decrease with the return period of the flood, a pattern as is opposite to the one observed for the recall technique. In the case of both floods events, the value of the precision metric obtained during the evaluation of the floods with a return period from 5-year to 50-year are suitable and higher than the ones obtained by Endo, Adriano, Mas, and Koshimura (2018), who examined the areas flooded as a result of a tsunami in Japan.

The results obtained using the *measure of fit* were not successful as negative values were obtained for all the return periods for the two flooding episodes studied. This occurs due to over-prediction of the flooded area. This could be because this statistical measure ignores large areas that remained dry during both floods and that were correctly predicted using the satellite images. This is particularly an issue in the current study, as the correctly predicted non-flooded areas constitute, in some cases, over 80% of the area studied.

The results obtained using the *accuracy* metric were better than those obtained using the *measure of fit*, but not as good as those using the precision and recall techniques. In the same way as the *precision* metric, the values for this metric for the 1997 flood decrease as the return period of the flood increases, while there is no trend for the 2010 flood. The accuracy metric puts an emphasis on the areas that were identified as flooded by the satellite images but that were in fact not at risk of flooding based on the Open source flood risk maps. This over-prediction based on the satellite images reduces the score obtained using this technique. Nonetheless, from an emergency response perspective, under-prediction is more of concern than over-prediction. Horritt (2006) mentioned that a good model is the one that strikes a balance between precision and uncertainty and through, this work, it was found that the flood risk maps with a 500-year return period are those that strike the best balance between *precision* and *accuracy* for both flooding episodes. Moreover, this is the map with the best *recall* values and the floods that would be expected to cause the most severe damage.

The impact of including the NDVI band on the results of the validation of the 1997 flood using the optical satellite imagery is shown in figure 10. As can be seen, there are clear differences in the range of NDVI values between the over-predicted flooded areas versus those that are correctly predicted and under-predicted, as the latter two have lower NDVI values.

A similar pattern is seen for the validation of the 2010 flood risk map. Figure 11 shows that the areas that are over-predicted, i.e., those that were delineated as flooded on the SAR image, but that are not considered at risk of flooding according to the government maps, have higher GLCM correlation band values.

There are differences in the range of NDVI values amongst the three groups between the validation of the 1997 and 2010 floods. For example, there is large variability within all groups, but only the NDVI values for the areas correctly predicted as being flooded include extreme outliers (Figure 10). The situation is different for the validation of the 2010 flood, as there is no extreme outlier for the areas correctly predicted as being flooded, meaning that the classification is more robust.

With regard to the elevation of areas that were correctly predicted and those that were under-predicted during both flood events, figures 12 and 13 show that under-prediction takes place in areas with higher altitudes than the correctly predicted flooded areas. This may be due to the great level of importance that the DEM plays in the flood risk maps provided by the government. Thus, the non-predicted but flooded areas were non-predicted because their elevation was not low enough to be considered as areas potentially at risk.

There are nonetheless differences between the two studied flood events in terms of the altitudes of the correctly predicted flooded areas. During the 2010 flood (figure 13), the flooded areas had further spatial extension that increases the elevation of the correctly predicted, almost matching the altitudes of the under-predicted areas. Attending to the variability in elevation of the predicted areas, it can be observed that for both events there is a high variability and the presence of a large number of extreme outliers, which are always located in high elevations.

Figure 7 previously showed that the under-prediction of the 2010 flood is greater than that of the 1997 flood. Two potential reasons are suggested to explain this pattern. First, the SAR images are more sensitive to soil moisture as the radar signal can penetrate deeper into the vegetation canopy. Second, the optical image for the 1997 flood was not completely free of clouds, and the clouds impeded the validation at the peripheral areas of the main channels.

In terms of validation metrics, in general, the best results were obtained for the 1997-flooding event. This could be caused by the image used for the evaluation of the 1997 flood event, which is closer in time to the occurrence of the flood than the one used for the 2010 flood event, consequently the results are better. However, the soil type of the study area, with slow infiltration rate, in addition to the low topographical gradient of the area results in floodwater not receding quickly. Moreover, the first flood was of bigger magnitude in terms of damage cause than the second and more recent flood and is an additional potential reason for the better results obtained in the evaluation of the 1997 flood episode.

The uncertainty associated with the flood risk maps provided by the government is not known, but inevitably arises to uncertainties due to a potential lack of observations, and uncertainties in the structure and parameters of the models used (Götzinger & Bárdossy, 2008). The small number of Open source satellite imagery in the last decades further increases the difficulty of the validation of floods or any other natural disaster. The ideal situation would be to having ground-truth data and remote sensing images during the flood episode, but as discussed in (Di Baldassarre & Uhlenbrook, 2012) data that are often not available at the right time from the right location.

7. CONCLUSIONS

This paper validated the Open source flood risk maps provided by the Government of Spain by comparing them with the areas flooded during two major floods that affected the city of Badajoz in southwestern Spain. The spatial extent of the 1997 flood was mapped using Open source optical imagery while a SAR image, also Open source, was used for the flood of 2010 due to the lack of availability of optical satellite images soon after the occurrence of that flood. Nonetheless, such a strategy allowed for the use and thus the validation of two types of images. The comparison was performed using maps depicting flood prone areas for floods with a 5-, 10-, 50-, 100- and 500-year return period. For both the 1997 and 2010 floods, the highest values of all the metrics were obtained for the maps depicting the spatial extent of floods with a 500-year return period. The evaluation of the 1997 flood event using the optical image provided better results than that of 2010 using SAR imagery, but this could be due to the longer time interval between the occurrence of the flood and the timing of the satellite image.

Satellite imagery can support the validation of flood risk maps due to their high spatial coverage and can thus complement the field surveys that are undertaken following a flood. The main weakness of satellite images, as this paper demonstrates, is the time lag between the time the satellite passes over the flooded area and the occurrence of a flood. The closer is the satellite image acquired to the occurrence of the flood, the more accurate and reliable will the validation of the flood risk map be. Optical and SAR images both have their advantages and disadvantages in relation to their use in delineating the spatial extent of flooded areas. Optical images have better quality than SAR images, but as they are not acquired from active sensors they cannot penetrate the surface and can only detect areas that are waterlogged. SAR images are more sensitive to soil moisture and are not influenced by cloud coverage, but the noise contained in radar images can sometimes be a problem.

Further research is recommended to determine whether optical images are better at delineating the spatial extent of a flood than SAR images when cloud coverage is limited or if the lower values for the different metrics are due to the longer time interval between the occurrence of the peak discharge and the timing of the satellite image, as the latter was longer for the 2010 flood. Many Open source optical and SAR images have recently become available as a result of the Copernicus Earth observation programme of the European Union. In particular, the inclusion of the Sentinel-1 and Sentinel-2 satellite data with a temporal resolution of five days into the Open source realm will potentially lead to a higher likelihood of the availability of a remote sensing image during or shortly after the occurrence of a flood, which was found to be a limitation for this study, particularly for the 2010 flood, but also to apply the two types of methodologies described as part of this paper (i.e., one using optical images and the other radar images) for the same flood event. The use of both types of satellite imagery to analyse the same floods would allow for a more detailed assessment of the advantages and disadvantages of each type of satellite imagery in validating flood risk maps. Nevertheless, the results of the analyses presented in this paper show the potential of Open geospatial data and their manipulation and processing using Open source remote sensing/GIS software in delineating flood prone areas and could supplement the traditional hydrological modelling method currently use or partly substitute the expensive ground truth observation practiced in many countries for the development of flood risk maps.

ACKNOWLEDGMENTS

This work was funded by mobility grants for teaching and research staff of the University of Extremadura and Technological Centers of the Autonomous Community of Extremadura in foreign centres of higher education and/or research. We also want to thank the Spanish ministry of ecological transition and demographic challenge for all Open source flood risk maps and the hydraulic public domain delimitation.

REFERENCES

Alfieri, Lorenzo, Dottori, Francesco, Betts, Richard, Salamon, Peter, & Feyen, Luc. (2018). Multi-Model Projections of River Flood Risk in Europe under Global Warming. *Climate, 6*(1), 6.

Álvarez, AJ. (2010). Mapa de caudales máximos Memoria Técnica. *Madrid: Centro de Estudios Hidrográficos del CEDEX*.

Bangham, J Andrew, Chardaire, Pierre, Pye, C Jeremy, & Ling, Paul D. (1996). Multiscale nonlinear decomposition: the sieve decomposition theorem. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 18*(5), 529-539.

Barredo, J. I., Saurí, D., & Llasat, M. C. (2012). Assessing trends in insured losses from floods in Spain 1971–2008. *Nat. Hazards Earth Syst. Sci., 12*(5), 1723-1729. doi: 10.5194/nhess-12-1723-2012

Bates, Paul D. (2012). Integrating remote sensing data with flood inundation models: how far have we got? *Hydrological Processes, 26*(16), 2515-2521. doi: 10.1002/hyp.9374

Black, Emily, Tarnavsky, Elena, Maidment, Ross, Greatrex, Helen, Mookerjee, Agrotosh, Quaife, Tristan, & Brown, Matthew. (2016). The Use of Remotely Sensed Rainfall for Managing Drought Risk: A Case Study of Weather Index Insurance in Zambia. *Remote Sensing, 8*(4), 342.

Changes in Flood Risk in Europe. ((2012).). (Z. Kundzewicz Ed.). London: CRC Press.

Dao, Phuong D., & Liou, Yuei-An. (2015). Object-Based Flood Mapping and Affected Rice Field Estimation with Landsat 8 OLI and MODIS Data. *Remote Sensing, 7*(5), 5077-5097.

Dasgupta, Antara, Grimaldi, Stefania, Ramsankaran, R. A. A. J., Pauwels, Valentijn R. N., & Walker, Jeffrey P. (2018). Towards operational SAR-based flood mapping using neuro-fuzzy texture-based approaches. *Remote Sensing of Environment, 215*, 313-329. doi: https://doi.org/10.1016/j.rse.2018.06.019

de Moel, H., van Alphen, J., & Aerts, J. C. J. H. (2009). Flood maps in Europe - methods, availability and use. *Nat. Hazards Earth Syst. Sci., 9*(2), 289-301. doi: 10.5194/nhess-9-289-2009

Dewan, Ashraf M., Kankam-Yeboah, Kwabena, & Nishigaki, Makoto. (2006). Using Synthetic Aperture Radar (SAR) Data for Mapping River Water Flooding in an Urban Landscape: A Case Study of Greater Dhaka, Bangladesh. *Journal of Japan society of hydrology and water resources, 19*(1), 44-54. doi: 10.3178/jjshwr.19.44

Di Baldassarre, Giuliano, Schumann, Guy, Brandimarte, Luigia, & Bates, Paul. (2011). Timely Low Resolution SAR Imagery To Support Floodplain Modelling: a Case Study Review. *Surveys in Geophysics, 32*(3), 255-269. doi: 10.1007/s10712-011-9111-9

Di Baldassarre, Giuliano, & Uhlenbrook, Stefan. (2012). Is the current flood of data enough? A treatise on research needs for the improvement of flood modelling. *Hydrological Processes, 26*(1), 153-158. doi: 10.1002/hyp.8226

Domeneghetti, Alessio, Schumann, Guy J.-P., & Tarpanelli, Angelica. (2019). Preface: Remote Sensing for Flood Mapping and Monitoring of Flood Dynamics. *Remote Sensing, 11*(8), 943.

Domeneghetti, Alessio, Tarpanelli, Angelica, Brocca, Luca, Barbetta, Silvia, Moramarco, Tommaso, Castellarin, Attilio, & Brath, Armando. (2014). The use of remote sensing-derived water surface data for hydraulic model calibration. *Remote Sensing of Environment, 149,* 130-141. doi: https://doi.org/10.1016/j.rse.2014.04.007

Endo, Yukio, Adriano, Bruno, Mas, Erick, & Koshimura, Shunichi. (2018). New Insights into Multiclass Damage Classification of Tsunami-Induced Building Damage from SAR Images. *Remote Sensing, 10*(12), 2059.

Götzinger, Jens, & Bárdossy, András. (2008). Generic error model for calibration and uncertainty estimation of hydrological models. *Water Resources Research, 44*(12). doi: 10.1029/2007wr006691

Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973). Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-3*(6), 610-621. doi: 10.1109/TSMC.1973.4309314

Horritt, M. S. (2006). A methodology for the validation of uncertain flood inundation models. *Journal of Hydrology, 326*(1), 153-165. doi: https://doi.org/10.1016/j.jhydrol.2005.10.027

Kundzewicz, Zbigniew W., Kanae, Shinjiro, Seneviratne, Sonia I., Handmer, John, Nicholls, Neville, Peduzzi, Pascal, . . . Sherstyukov, Boris. (2014). Flood risk and climate change: global and regional perspectives. *Hydrological Sciences Journal, 59*(1), 1-28. doi: 10.1080/02626667.2013.857411

Kundzewicz, Zbigniew W., Pińskwar, Iwona, & Brakenridge, G. Robert. (2013). Large floods in Europe, 1985–2009. *Hydrological Sciences Journal, 58*(1), 1-7. doi: 10.1080/02626667.2012.745082

Laur, H, Bally, P, Meadows, P, Sanchez, J, Schattler, B, Lopinto, E, . . . ERS SAR, Calibration. (2003). Derivation of σ0 in ESA ERS SAR PRI Products, ESA/ESRIN: ES-TN-RS-PM-HL09.

Li, J., & Wang, J. (2017). *Mapping water bodies using SAR imagery – an application over the Spiritwood valley aquifer, Manitoba*. (M103-3/31-2017E-PDF). Canada: Government of Canada.

Lorente, P., Hernández, E., Queralt, S., & Ribera, P. (2008). The flood event that affected Badajoz in November 1997. *Adv. Geosci., 16*, 73-80. doi: 10.5194/adgeo-16-73-2008

Mansourpour, Mostafa, Rajabi, MA, & Blais, JAR. (2006). *Effects and performance of speckle noise reduction filters on active radar and SAR images*. Paper presented at the Proc. ISPRS.

Olcina, Jorge. (2016). Flood policy in Spain: a review for the period 1983-2013. *Disaster Prevention and Management, 25*(1), 41-58. doi: 10.1108/DPM-05-2015-0108

Olsson, Lennart, Opondo, Maggie, Tschakert, Petra, Agrawal, Arun, & Eriksen, Siri EH. (2014). Livelihoods and poverty.

Ouled Sghaier, Moslem, Hammami, Imen, Foucher, Samuel, & Lepage, Richard. (2018). Flood Extent Mapping from Time-Series SAR Images Based on Texture Analysis and Data Fusion. *Remote Sensing, 10*(2), 237.

Pradhan, Biswajeet, Hagemann, Ulrike, Shafapour Tehrany, Mahyat, & Prechtel, Nikolas. (2014). An easy to use ArcMap based texture analysis program for extraction of flooded areas from TerraSAR-X satellite image. *Computers & Geosciences, 63*, 34-43. doi: https://doi.org/10.1016/j.cageo.2013.10.011

Pulvirenti, L., Chini, M., Pierdicca, N., Guerriero, L., & Ferrazzoli, P. (2011). Flood monitoring using multi-temporal COSMO-SkyMed data: Image segmentation and signature interpretation. *Remote Sensing of Environment, 115*(4), 990-1002. doi: https://doi.org/10.1016/j.rse.2010.12.002

Quirós, Elia, & Polo, María Eugenia. (2018). Open resources on geographic information for research and scientific documentation. *Revista española de Documentación Científica, 41*(3), e214. doi: 10.3989/redc.2018.3.1512

Roudier, Philippe, Andersson, Jafet C. M., Donnelly, Chantal, Feyen, Luc, Greuell, Wouter, & Ludwig, Fulco. (2016). Projections of future floods and hydrological droughts in Europe under a +2°C global warming. *Climatic Change, 135*(2), 341-355. doi: 10.1007/s10584-015-1570-4

Shanmugan, K. S., Narayanan, V., Frost, V. S., Stiles, J. A., & Holtzman, J. C. (1981). Textural Features for Radar Image Analysis. *IEEE Transactions on Geoscience and Remote Sensing, GE-19*(3), 153-156. doi: 10.1109/TGRS.1981.350344

Small, David, & Schubert, Adrian. (2008). Guide to ASAR geocoding. *ESA-ESRIN Technical Note RSL-ASAR-GC-AD, 1*, 36.

United States Department of Agriculture Natural Resources Conservation Service. (2007). Chapter 7 Hydrologic Soil Groups *Part 630 Hydrology. National Engineering Handbook* (Vol. 210–VI–NEH, May 2007). United States

Wu, Junjie. (2012). *Advances in K-means Clustering: A Data Mining Thinking*: Springer-Verlag Berlin Heidelberg.

FIGURE LEGENDS

Figure 1. Categorization of the population at risk (a) and flood risk (b) in the study area.

Figure 2. Guadiana River discharge during the period 1995-2014, and availability of Open source optical and radar images.

Figure 3. Schematic representation of the methodological approach.

Figure 4. NDVI band used for the detection of flooded areas.

Figure 5. Texture analysis of the SAR image. Correlation image (left) compared to the original image (right).

Figure 6. Validation of the 1997 flood.

Figure 7. Validation of the 2010 flood.

Figure 8. Accuracy metrics of 1997 flood event compared to flood risk maps.

Figure 9. Accuracy metrics of 2010 flood event compared to flood risk maps.

Figure 10. NDVI values of the areas correctly predicted as flooded, over-predicted, and under-predicted on the optical satellite image according the flood risk maps for the 1997 flood.

Figure 11. GLCM correlation band values of the areas correctly predicted as flooded, over-predicted, and under-predicted on the SAR image according the flood risk maps for the 2010 flood.

Figure 12. Elevation values of the areas correctly predicted as flooded, over-predicted, and under-predicted according the flood risk maps for the 1997 flood.

Figure 13. Elevation values of areas correctly predicted as flooded, over-predicted, and under-predicted according the flood risk maps for the 2010 flood.