A WEB APPLICATION FOR THE AUTOMATIC MAPPING OF THE FLOOD EXTENT ON SAR IMAGES

Abdelhakim Benoudjit and Raffaella Guida

Surrey Space Centre, University of Surrey
Guildford GU2 7XH, U.K.
a.benoudjit@surrey.ac.uk

ABSTRACT

Flood is causing devastating damages every year all over the world. One way to improve the readiness of stakeholders (rescue authorities, policy makers, and communities) is by providing flood extent maps promptly after the disaster, preferably in an automated way and with a minimum number of satellite imagery to reduce costs. The web application developed in this paper aims to address this problem by mapping the flood extent automatically from SAR images.

This web application is portable since it runs on the internet browser, and allows to perform the classification of the flooding in an automated fashion. Another strong point is the rapidity of the processing: the whole processing time was around 3 to 5 minutes for a subset of 20 million pixels. The inundation map returned by our algorithm was validated against vector files mapped by the United Nations Institute for Training and Research (UNITAR) for the same flood event.

Regarding the dataset needed in this study, a pair of a pre-flood SAR image and an optical image of the same area were used to build a training dataset of water and non-water classes. The learning phase is immediately followed by the classification of the post-flood SAR image into a binary flood map. The web application described in this paper was built with open-source Python libraries which are backed by large communities (Django, Scikit-learn among others). The flood map was eventually displayed on OpenStreetMap maps provided by Mapbox.

Index Terms— Geographic Information System (GIS), Flood extent mapping, Supervised classification, Synthetic Aperture Radar (SAR)

1. INTRODUCTION

1.1. Introduction

Climate change is causing a hike in the Earth temperature, and manifests itself in the fact that 2016 was the warmest year breaking the record set by the two years before it, while 2017 ranked at the third position overall and as the warmest among years that were not affected by the El Niño meteorological phenomenon [1]. The warming of the Earth increases the quantity of water evaporated, which makes the consequent rainfall more intense and contributes to the risk of flooding. At the same time, the situation is exacerbated since warm air tends to hold more moisture. In fact, a 1°C increase in the atmosphere temperature leads to 7% more water accumulated [2].

Synthetic Aperture Radars (SAR), being active sensors, have the ability to observe the Earth surface regardless of the weather and the time of acquisition. That is why they are more practical than optical satellites for monitoring floods, which occur usually during cloudy conditions. Furthermore, the constellations of Sentinel-1 SAR satellites and Sentinel-2 optical satellites providing revisit-times of 6 and 5 days respectively on the one hand, and the open-data policy of ESA's (European Space Agency) Copernicus Programme on the other, open up new possibilities to deal with Earth observation problems, like the mapping of flooding.

In 2017, more than 60% of the loss of lives caused by natural disasters all over the world were due to inundations [3]. One way satellite imagery can support decision makers and communities in increasing the preparedness, is through flood extent maps. The web application discussed in this paper addresses the same issue tackled by an online service featured in the literature [4], which is to provide response authorities and stakeholders with a map of the extent of flooding promptly following the disaster. However, the flood mapping service proposed in [4] was designed to operate internally on commercial TerraSAR-X images available to the DLR (German Aerospace Center), and therefore it did not take advantage of the freely-available datasets from Sentinel-1 and Sentinel-2, as is the case with the process described in this study.

2. METHODOLOGY

The web application discussed in this paper was developed in Python using the Django Web framework. The chart in Figure 3 illustrates the steps implemented in this web application, which are explained in more detail in the sections below.

2.1. Preprocessing

The Sentinel-2 optical image was atmospherically corrected to eliminate the effect of the atmosphere. This operation was carried out using the Sen2Cor SNAP plugin (Sentinel Application Platform) to calculate the surface reflectance and is necessary before any classification is initiated. During the atmospheric correction, the Top-Of-Atmosphere Level-1C Sentinel-2 product given as input is transformed into a Bottom-Of-Atmosphere Level-2A one.

As for the pre-flood and post-flood SAR images, the preprocessing consists of a calibration to σ^0, a Terrain-Correction with the freely-available Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM), a filtering of the speckle using the Gamma-Map filter, and finally a conversion to Decibels (dB). All of these operations were implemented with the SNAP API for Python.

The SRTM DEM, which is automatically downloaded by SNAP during the terrain-correction, has a resolution of around 30m (1 arc-second). The terrain-correction allows to reduce the inherent geometric distortions in radar imagery, and is required before the collocation of the pre-flood SAR image with the optical one. This DEM was chosen due to not having a free high-resolution DEM, although it has a resolution slightly coarser (30m) compared to the SAR images (20m). Regardless of the resolution of the DEM, the orthorectified SAR images are resampled using the bilinear interpolation to the same pixel spacing as the input Sentinel-1 products (10m).

The speckle noise present mainly in SAR images is caused by signals returning from adjacent and heterogeneous scatterers on the ground, which are summed into the same pixel cell [5]. This type of noise can lead to wrongly-classified pixels, and thus the SAR images have to be speckle-filtered before any processing is performed. The filtering of the speckle was carried out with a 5x5 pixels Gamma-MAP filter, which according to [6] achieves satisfying results in a reasonable time.

2.2. Building the training dataset

Briefly, the training dataset for water and land classes was generated by implementing the following operations with NumPy (a Python library which makes it easy to deal with matrices, e.g. images):

Firstly, the preprocessed pre-flood SAR image and the optical image acquired in dry conditions were collocated. Then, the NDVI was calculated from the red and the near-infrared bands of the Sentinel-2 image by applying the following formula:

$$NDVI = \frac{NIR - Red}{NIR + Red} \qquad (1)$$

The water mask can be subsequently extracted by thresholding the NDVI using a water threshold found in the literature, as shown in the formula below [7]:

$$Class = \begin{cases} Water, & \text{if } NDVI \leq -0.11 \\ Land, & \text{otherwise} \end{cases} \qquad (2)$$

This threshold cannot be directly employed to map the floodwater on an optical image most of the time for lack of cloud-free products during the inundation.

The water threshold in Equation 2 ($NDVI = -0.11$) was validated on three different Sentinel-2 optical images of the same studied region in Myanmar. These images were acquired during the hot season in the same month as the pre-flood SAR image (March), in 2016, 2017, and 2018, with very little or no cloud coverage: 4%, 2%, and 0 %, respectively. It can be appreciated visually from the NDVIs calculated from these three images (Figure 1), that the water from the main river and other wetlands nearby appearing in black, display the lowest values in the NDVI. This is also confirmed by the histograms in Figure 2, where it can be deduced that the values around the smallest mode (on the left of the histogram) represent the water. The NDVI thresholds shown in red on the histograms allow to easily isolate water from the rest of the land cover in the three cases. Moreover, one strength to note after experimenting on the Sentinel-2 image acquired in 2016 is that the clouds and their shadows appearing in dark gray on the NDVI image (Figure 1-a) were not classified as water (Figure 1-b) by this empirical threshold.

The pixels in the preprocessed pre-flood SAR image that do not belong to the water class were filtered out from the latter product, with a simple multiplication with the water mask. The logical negation of the water mask gives the land mask, and the exact same operation is performed again but with the land mask to extract the land pixels from the SAR image:

$$\begin{aligned} TrainingWater &= PreSAR \cdot WaterMask \\ TrainingLand &= PreSAR \cdot \overline{WaterMask} \end{aligned} \qquad (3)$$

2.3. Classification of the post-flood SAR image

The flood was mapped with the Multi-Layer Perceptron (MLP) classifier, which is already implemented in scikit-learn, a machine learning library written in Python. The neural network classifier has only one hidden layer with ten neurons in it.

The flood extent mapping algorithm employed in this study relies on a pixel-based classifier which, as opposed to object-based ones, does not take into account the spatial and textural context. Object-based classification was used in the context of flood mapping from SAR images in multiple studies, particularly in [8]. The main shortcoming of pixel-based classifiers is the noisy nature of the resulting flood map, since each pixel is categorized individually without taking into account the surrounding pixels' classes. However, the object-based classification commonly used in the literature is based
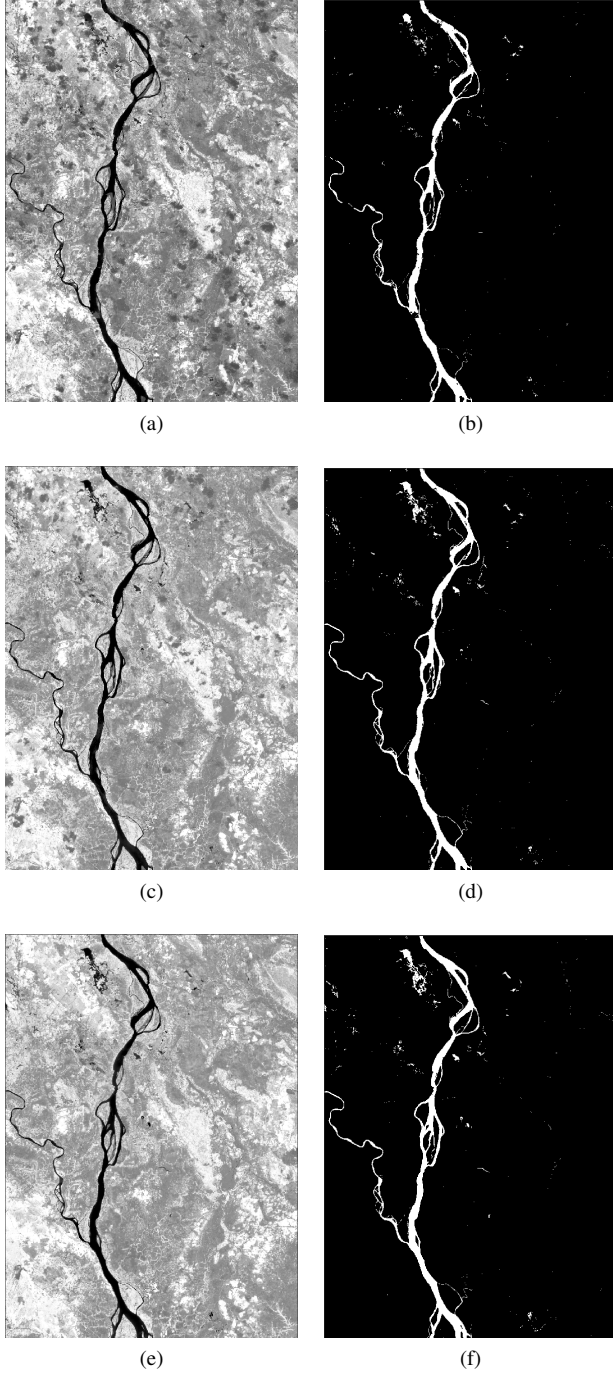
Fig. 1. The Normalized Difference Vegetation Index (NDVI) calculated from Sentinel-2 optical images acquired on the (a) 19 March 2016 (c) 14 March 2017 (e) 04 March 2018 - and the water masks obtained by thresholding the previous NDVIs respectively for the (b) 19 March 2016 (d) 14 March 2017 (f) 04 March 2018.

on the eCognition commercial software, as opposed to the open-source libraries utilized in this paper. More importantly, in the case of the object-based segmentation, it is followed
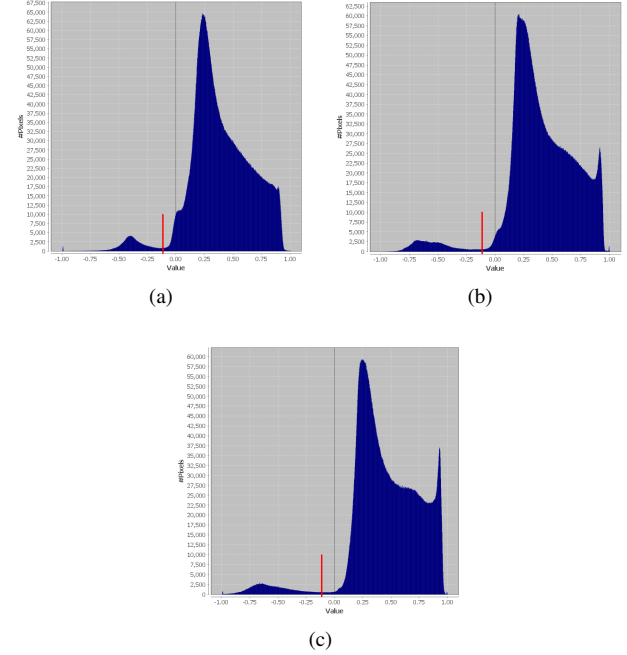

Fig. 2. Histograms of the NDVIs calculated from the Sentinel-2 images acquired on the (a) 19 March 2016 (c) 14 March 2017 (e) 04 March 2018. The vertical red line in the histogram is where the empirical threshold is located (NDVI = -0.11).

by a classification which required the user to step in to set a few scale and thresholds parameters [8], while the web application developed in this paper aims to work in a fully-unsupervised mode which is more adequate for emergency situations.

2.4. Rendering of the flood map

GeoServer is an open-source software that allows to serve maps, and georeferenced rasters and vectors using protocols such as WMS (Web Map Service). The post-flood SAR and the flood map rasters were rendered on Mapbox maps.

3. CASE STUDY AND DATASET

3.1. Case study

The area under study is located in the north of the city of Mawlamyine in southeastern Myanmar (Burma), where the Salween River overflowed its banks on August 2015 during the monsoon season. The flooding occured during the months of July and August 2015, and affected most of the states of Myanmar. Certain states were severely ravaged, and a million of people were forced out of their homes while more than a hundred casualties were reported [9]. The flood extent was mapped on a subset of a Sentinel-1 image taken north of the
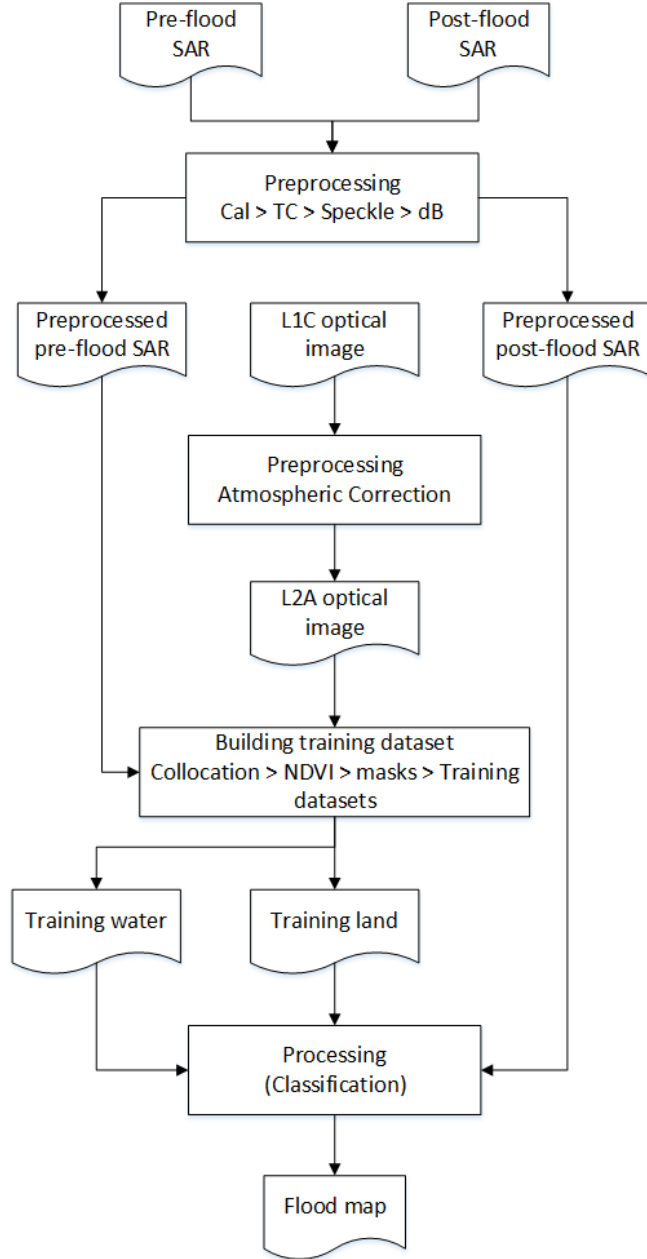
Fig. 3. Flowchart of the flood mapping web application.

mouth of the Salween (Thanlwin) River, where it flows into the Indian Ocean.

3.2. Dataset

No optical image from Sentinel-2 was available at the time of the flooding since Sentinel-2 was not launched till June 2015, and did not deliver its first image on the Sentinel Hub till the end of 2015 after its commissioning phase. A Sentinel-2 image was chosen instead of a Landsat-8 one thanks to its higher resolution (i.e. 10m instead of 30m), which will be

helpful when extracting the water from the NDVI. The optical Sentinel-2 image used in this paper (Figure 4-c) was taken in the same season as the pre-flood SAR image (Hot season), to avoid having water bodies appearing on the optical image when they are dry on the pre-flood SAR image. This is highly likely to happen if the optical image is taken during the Rainy season or the Cool season whereas the pre-flood SAR image is acquired during the Hot season. Moreover, there is a delay of three years in the dates of acquisitions between the pre-flood Sentinel-1 SAR image (March 2015) and the optical Sentinel-2 image (March 2018), because it was the only image virtually cloud-free and in the same month. The presence of clouds over water bodies is normally not a problem for SAR sensors which can see through them. However, because of the way the water mask is extracted in our case, patches of clouds masking water would cause these pixels to appear in the land mask instead of the water mask, and thus be wrongly-categorized with the land class in the training dataset retrieved from the pre-flood SAR image.

The pre-flood (Figure 4-a) and the post-flood (Figure 4-b) SAR images were acquired by Sentinel-1 in the exact same configuration in March 2015 and August 2015, respectively. Both SAR products were captured in the Interferometric Wide Swath (IW) mode with 20m-resolution. The two images are both VV-polarized and were distributed in the multi-looked Ground Range Detected format (GRD).

4. RESULTS

No ground truth was available to validate our results. The flood map (Figure 4-d) was therefore compared to the one obtained from the same Sentinel-1 dataset by the United Nations Institute for Training and Research (UNITAR) [10]. The validation dataset is available as a shapefile vector which was therefore rasterized and clipped with the QGIS software to focus only on the area of interest, prior to the pixel to pixel comparison with the resulting flood map. The commission and omission errors for both the water and land classes as well as the overall accuracy of the classification relatively to the latter ground truth, are shown in Table 1. The flood mapping method utilized allowed to reach a high accuracy in the order of over 90% on the studied Myanmar dataset.

After a visual assessment, it can be easily noted that mountain shadows were misclassified as water on the flood map. This explains the commission error in the flood class (6.97% in Table 1) which was caused by the similar dark backscatter signature of shadow. As for the omission error for the same flood class (6.09% in Table 1), it is possible that it was caused by the missed classification of the water under the vegetation, due to the medium wavelength of the C-band sensor which cannot penetrate deep enough in the canopies.

In addition to the previous assessment metrics, the area affected by the flooding was measured in Km^2 by counting the number of pixels detected as water on both the flood map

	Commission error	Omission error	Overall accuracy
Non-flood	2.92%	3.36%	95.76%
Flood	6.97%	6.09%	

Table 1. The commission and omission errors for the flood and non-flood classes and the overall accuracy of the classification with the MLP classifier on Myanmar 2015 dataset [11].

	Area
Flood map	674.29 Km^2
Ground truth	667.97 Km^2

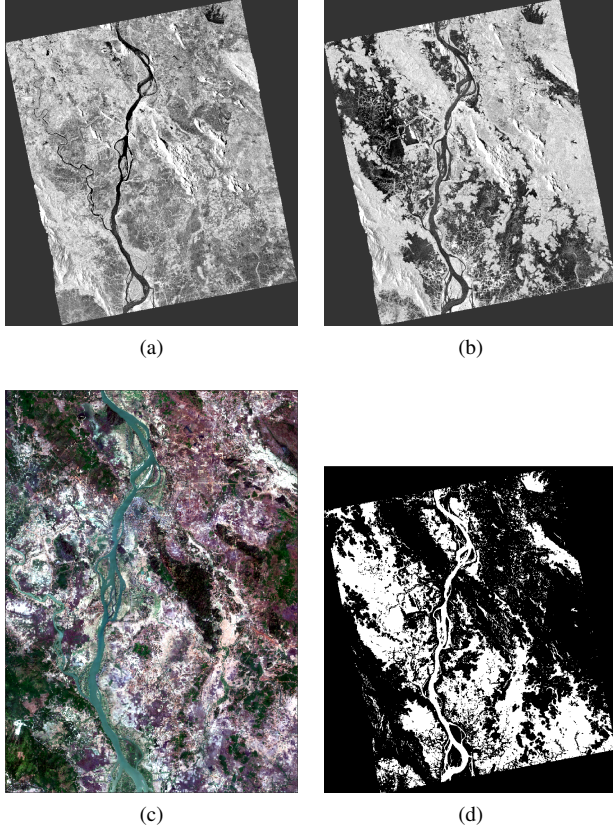**Table 2**. Area that suffered from the flooding in Km^2 on the obtained flood map and on the ground truth.

Fig. 4. Subset of the Salween River on (a) the pre-flood Sentinel-1 SAR image (15 March 2015) (b) the post-flood Sentinel-1 SAR image (06 August 2015) (c) the Sentinel-2 optical image (04 March 2018) (d) the resulting flood map obtained by the MLP classifier.

obtained and the ground truth, and multiplying these two totals each time by the area of a single pixel ($10m \times 10m$ for Sentinel-1 GRD images). The two calculated areas (Table 2) were compared to confirm the overestimation of the flooding expressed by the previous commission error (6.97% in Table 1).

4.1. Results on TerraSAR-X images

In order to test the capabilities of this flood mapping web application on X-band products, the dataset for the inundation in Tewkesbury (south-west of England) in July 2007 was chosen to validate this paper's algorithm. This flood event was captured by TerraSAR-X in the 3m-resolution Stripmap mode, and was extensively studied in the literature about flood mapping from SAR images [8]. This TerraSAR-X image as well as a cloud-free Sentinel-2 captured in the same season but in 2016, went through the same preprocessing as the Myanmar dataset. The classifier used in this case was the Stochastic Gradient Descent (SGD) because its results were assessed visually to be more accurate than those obtained with the

Multi-Layer Perceptron (MLP) used on the previous Myanmar dataset. It should also be noted that to get rid of the noisy nature of the flood map which could create False Positives, a post-processing was performed to remove from the flood map all the polygons having a number of pixels below an arbitrary threshold (1000 pixels).

The post-flood SAR image as well as the resulting flood map are shown in Figure 5. As can be seen from the results reported in Table 3, the precision reached is not too far from the 80% agreement with the ground truth surveyed by the UK Environment Agency using airborne photography on the day of the disaster. One drawback is the high number of missed detections (42.72% in Table 3). However, this can be explained by the fact that the SAR satellite imagery was acquired a few days after the validation dataset was collected, and a potential water withdrawal cannot be excluded.

5. CONCLUSION

The web application developed in this paper was helpful in delineating the extent of the flooding, and managed to reach an accurate classification of the SAR image in a reasonable time (3-5 minutes for the whole process: the preprocessing, the training, and the classification, on a subset of 4000x5000 pixels). Another advantage to note is the fact that the whole process is automatic, which makes it useful to assist response authorities and the communities to be proactive and prepared for potential disasters in the future in flood-prone areas.

Among future improvements, one preprocessing step required to precede the classification is the generation of a mask of shadows from the DEM, which represents areas where the flood cannot be detected on the SAR image, because the presence of water there does not affect the backscatter [12]. Also as a future work, the algorithm developed in [13] to give an
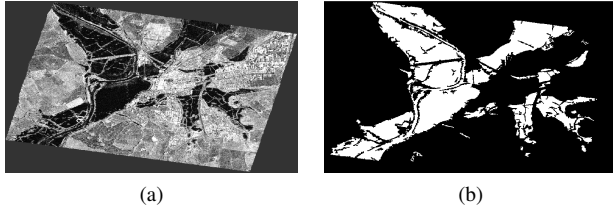
(a) (b)

Fig. 5. (a) The post-flood TerraSAR-X image of the flood in Tewkesbury (25 July 2007) (b) The flood extent map obtained.

	Commission error	Omission error	Overall accuracy
Non-flood	28.27%	0.75%	79.19%
Flood	1.41%	42.72%	

Table 3. The commission and omission errors for the flood and non-flood classes and the overall accuracy of the classification with the SGD classifier on Tewkesbury 2007 dataset [11].

estimation of the depth of the flooding near a building of interest will be integrated in this web application. Finally, the threshold used to extract the water mask from the NDVI will need to be validated against optical images from other regions in the world, where a different land cover might require this threshold to be adjusted accordingly.

6. REFERENCES

[1] Met Office, "2017: warmest year on record without El Niño," https://www.metoffice.gov.uk/news/releases/2018/2017-temperature-announcement, 2017, [Online; accessed 29-Apr-2018].

[2] K. E. Trenberth, "Changes in precipitation with climate change," *Climate Research*, vol. 47, no. 1-2, pp. 123–138, 2011.

[3] AON, "Weather, climate & catastrophe Insight," *2017 Annual Report*, p. 11, 2018.

[4] S. Martinis, J. Kersten, and A. Twele, "A fully automated TerraSAR-X based flood service," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 104, pp. 203–212, 2015.

[5] A. Reddy, *Textbook of remote sensing and geographical information systems*, BS Publications, Hyderabad, 2006.

[6] S. Martinis, A. Twele, and S. Voigt, "Towards operational near real-time flood detection using a split-based automatic thresholding procedure on high resolution TerraSAR-X data," *Natural Hazards and Earth System Science*, vol. 9, no. 2, pp. 303–314, 2009.

[7] P. Nakmuenwai and F. Yamazaki, "Extraction of flooded areas in the 2011 Thailand flood from RADARSAT-2 and ThaiChote images," *International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 3354–3357, 2014.

[8] D. C. Mason, I. J. Davenport, J. C. Neal, G. J.-P. Schumann, and P. D. Bates, "Near Real-Time Flood Detection in Urban and Rural Areas Using High-Resolution Synthetic Aperture Radar Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 8, pp. 3041–3052, 2012.

[9] Government of the Union of Myanmar, "Myanmar - Post-disaster needs assessment of floods and landslides : July - September 2015," *The World Bank*, vol. 1, no. 103631, 2015.

[10] UNITAR, "Flood Waters Over Kayin and Mon State, Myanmar," www.unitar.org/unosat/node/44/2268, 2015, [Online; accessed 23-Apr-2018].

[11] J. R. Jensen, *Introductory Digital Image Processing: A Remote Sensing Perspective*, Pearson, 3rd edition, 2005, p. 507.

[12] L. Giustarini, R. Hostache, P. Matgen, G. J. Schumann, P. D. Bates, and D. C. Mason, "A Change Detection Approach to Flood Mapping in Urban Areas Using TerraSAR-X," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 4, pp. 2417–2430, 2013.

[13] A. Benoudjit and R. Guida, "Semi-automated estimation of the local flood depth on sar images," *2017 IEEE 3rd International Forum on Research and Technologies for Society and Industry (RTSI)*, pp. 1–6, Sept 2017.