

## ASSESSMENT OF FLOODED AREAS CAUSED BY A DAM BREAK (SARDOBA DAM, UZBEKISTAN)

B. Tavus<sup>1</sup>, S. Kocaman<sup>1\*</sup>, C. Gokceoglu<sup>2</sup>

<sup>1</sup> Dept. of Geomatics Engineering, Hacettepe University, 06800 Beytepe Ankara, Turkey – (beste.tavus, sultankocaman)@hacettepe.edu.tr

<sup>2</sup> Dept. of Geological Engineering, Hacettepe University, 06800 Beytepe Ankara, Turkey - cgokce@hacettepe.edu.tr

Commission III, WG III/6

**KEY WORDS:** Dam Break, Flash Flood, Synthetic Aperture Radar, Optical Sensors, Earth Observation Data, Data Fusion, Disaster Assessment, Sardoba (Uzbekistan)

### ABSTRACT:

Although dams are very useful engineering structures, they can have extremely harmful consequences if they fail. One example of these failures occurred in Sardoba Reservoir (Uzbekistan). On May 1, 2020, a part of an earthfill dam failed along the Sardoba Reservoir, and a large region with settlements and agricultural areas in Uzbekistan and Kazakhstan was flooded. Accurate mapping and monitoring of the flooded areas are crucial for the damage assessment and the mitigation efforts. Satellite Earth Observation datasets can serve for these purposes due to their greater availability with high spatial and temporal resolutions. However, the optical sensors have limitations for data acquisition due to the atmospheric conditions, particularly the cloud cover, which often severely affects the image usability when floods occur. The synthetic aperture radar sensors provide valuable information under all weather conditions, but their interpretation is relatively difficult. Therefore, a data fusion methodology is proposed here for the integrated use of Sentinel-1 and Sentinel-2 datasets using a set of features obtained from both. Four different feature combinations were evaluated using the random forest classifier. The pre-processing steps for the feature extraction are explained in detail and the results are discussed here. The proposed algorithm exhibits very high classification accuracy for the flooded areas and flooded vegetation classes. The method can be employed for the flash flood mapping at regional scale. In addition, the damage assessment especially for agricultural areas in the region is very important for accounting the economic losses and the resilience purposes.

### 1. INTRODUCTION

Dams are one of the major infrastructures, and they are constructed for various purposes such as energy production, flood control, irrigation, industrial and domestic water supply etc. However, the consequences of dam failures can be severe for downstream communities due to sudden flash flood. In addition to losses of human lives, agricultural areas, natural flora and fauna, and various infrastructures can be damaged by floods occur after a dam failure. Project and construction faults, unsuitable site selection, extreme or differential settlements, unexpected heavy rainfalls and landslides, earthquakes and fault movements can be the causes of dam failures. After a dam failure, the rapid determination of flooded areas is extremely important for emergency aid and identification of damaged areas since the floods caused by a dam failure develop very rapidly and affect very large areas. On May 1, 2020, a part of an earthfill dam failed along the Sardoba Reservoir in Uzbekistan, and a large area in Uzbekistan and Kazakhstan was flooded. Six person died and thousands were evacuated as flood water spilled across the region and into neighbouring Kazakhstan (Putz, 2020). In this study, it was aimed to determine the areas affected by the flood after the dam failure by using Sentinel-1 synthetic aperture radar (SAR) and Sentinel-2 optical datasets.

The SAR sensors, which have all weather conditions and all-time provide data, have an important value in flood mapping (Tong et al., 2018). SAR data are processed using different methods in the literature to determine water surfaces that have different reflections from other surfaces (Ouled Sghaier et al., 2018). However, information extraction from SAR data is typically considered to be more difficult than multispectral sensors (Amitrano et al., 2018). Due to various advantages of SAR and optical data, many studies in the literature have employed these two data types in a complementary fashion via data fusion methods (Notti et al., 2018; Vanama et al., 2021; Anusha and Bharathi, 2020; Tavus et al., 2019; Tavus et al., 2020).

Here, a feature-level fusion method employing the Sentinel-1 and Sentinel-2 datasets was proposed to determine flooded areas using Random Forest (RF) supervised classification method. For this purpose, three Sentinel-1 and four Sentinel-2 datasets provided freely by the Copernicus Open Access Center (Copernicus, 2020) were used. Initially, the Sentinel-1 and Sentinel-2 images were geometrically and radiometrically corrected for the feature level fusion with high accuracy. The pre-processing steps include spatial resampling to 10 m and mosaicking of five Sentinel-2 bands, radiometric correction, speckle filtering, and terrain correction of the Sentinel-1 SAR images. Finally, the RF was performed to map the flooded areas

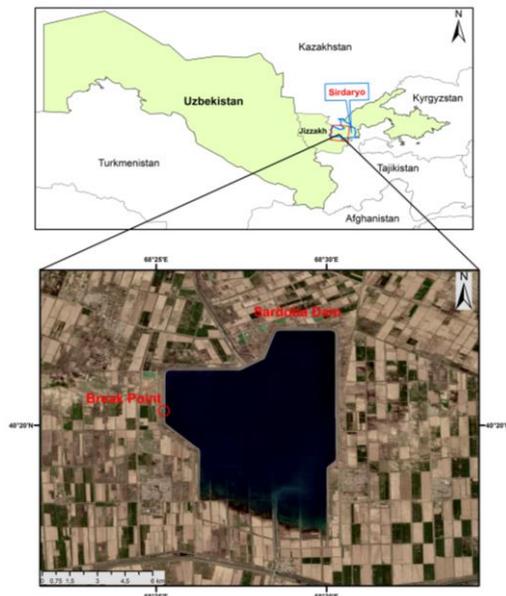
\* Corresponding author

with different features. All processing steps were performed using the SNAP Tool provided by ESA.

## 2. STUDY AREA AND DATASETS

### 2.1 Description of the Event

The Sardoba Dam, constructed between 2010-2017 on the Syr Darya River, has a reservoir with a volume of ~922 million m<sup>3</sup>, which mainly supply for the irrigation of agricultural lands of Sirdaryo and Jizzakh regions (Simonow, 2020). Figure 1 shows the study area location and a zoomed view of the dam on one Sentinel-2 RGB image.



**Figure 1.** Location of the study area on Sentinel-2 image

The dam breach followed seven days of heavy rainfall and high winds in the Sirdaryo and Jizzakh regions, at 5.55 am on May 1 2020, caused a Sardoba dam wall to collapse partially, flooding a large land area (Figure 2). The flooding affected more than 35,000 hectares of land in Uzbekistan and Kazakhstan. Six people died and at least 111,000 were evacuated from the Syr Darya river basin (Simonow, 2020). The agricultural areas were severely damaged as well.



**Figure 2.** Views from the failure and the damages of caused by the Sardoba Dam break (photo credits: Fergana News, 2020)

### 2.2 Datasets

Sentinel-1 is constellation of two SAR satellites, Sentinel-1A and Sentinel-1B, with ~5.7 cm wavelength. The satellites are operated by the Copernicus Program of European Space Agency (ESA); have a revisit time of twelve days for each (six

days for the constellation) (Nagler et al., 2016). The Sentinel-2 optical sensors are twins (2A/2B) and take images with ground sampling distances (GSDs) ranging between 10 m - 60 m (Drusch et al., 2012). The sensors have a total of thirteen spectral channels with diverse wavelengths (e.g. visible, near-infrared (NIR), short-wave-infrared (SWIR)) with a swath-width of 290 km and five day temporal resolution (Gascon et al., 2017).

In this study, Sentinel-1A C-band Interferometric Wide (IW) swath mode and Level 1 ground range detected (GRD) products were used. Each product is dual polarization (VV+VH) at a spatial resolution of 5 m × 20 m and has a swath coverage of 250 km. In addition, Sentinel-2B MSI data with Level-2A from five spectral bands, i.e., blue (B1), green (B2), red (B3), NIR (B8) and SWIR (B11) were employed here. Both datasets were obtained from the ESA Copernicus Programme (Copernicus, 2020). The main properties of the utilized datasets, such as the surface state with respect to the event, the date, and the dataset (DS) ID number are summarized in Table 1.

Sensor	Surface State	Date	DS ID
Sentinel-1A	Pre-event	2020/04/29	DS1
Sentinel-1A	Post-event	2020/05/05	DS2
Sentinel-1A			DS3
Sentinel-2B	Pre-event	2020/04/24	DS4
Sentinel-2B			DS5
Sentinel-2B	Post-event	2020/05/04	DS6
Sentinel-2B			DS7

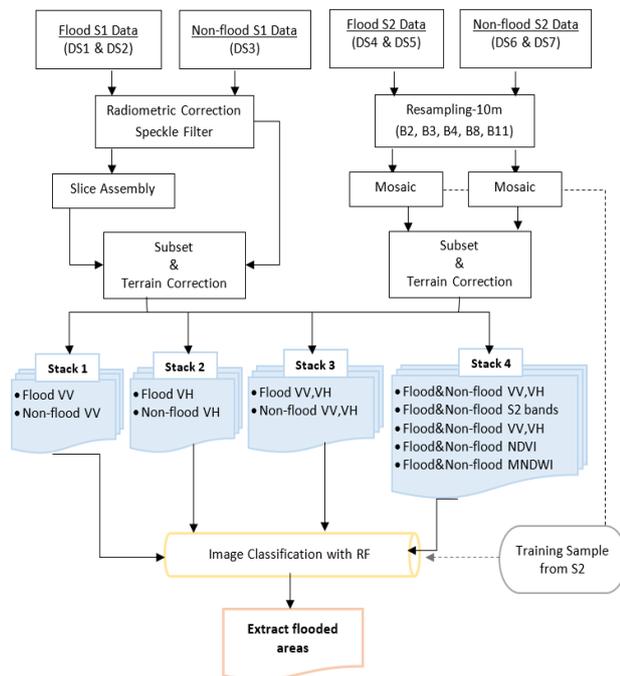
**Table 1.** The main features of the datasets employed here.

The Sentinel-1 and Sentinel-2 data were selected at the most suitable dates to represent the pre- and post-flood conditions. The Sentinel-1A pre-event data has the relative orbit number 71 and ascending pass direction, while the post-event dataset has relative orbit number 152 and descending pass direction. Similarly, Sentinel-2B pre- and post-event datasets have the relative orbit number of 34 and descending pass direction. As can be seen from Table 1, two Sentinel-2 pre-event and two post-event images cover the study area. On the other hand, the post-event Sentinel-1 datasets were employed to ensure the full-coverage of the study area extent. The datasets were merged in the pre-processing steps.

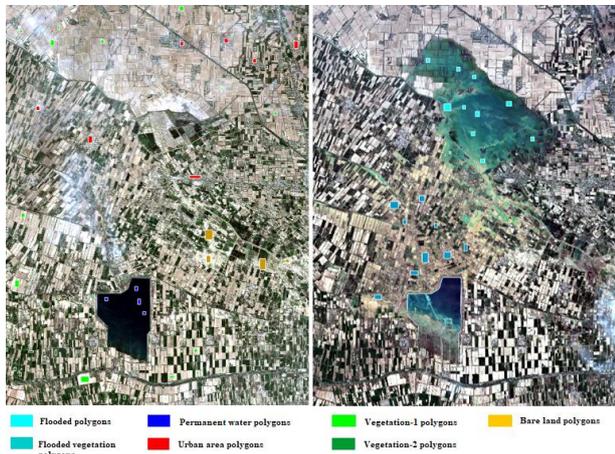
## 3. METHODOLOGY

The methodological workflow of the study can be evaluated in three main stages, i.e., data pre-processing, classification, and evaluation of the results (Figure 3). First, a set of pre-processing algorithms for the Sentinel-1 and Sentinel-2 data was carried out; and the data obtained were stacked in different combinations. In the second stage, each stack was classified using the training samples, which were manually delineated from the Sentinel-2 datasets. Seven land use land cover (LULC) classes were identified (i.e., permanent water, flooded area, flooded vegetation, urban, bare land, and two types of agricultural areas such as vegetation-1, vegetation-2) prior to the selection of the training data. Representative image parts were selected from the pre- and post-event Sentinel-2 mosaics for each class in the form of polygons (Figure 4). Finally, the classification was carried out by using the RF algorithm. The evaluation of the results was performed by visual assessments

and using the overall accuracy values obtained from the RF method.



**Figure 3.** The methodological workflow of the study.



**Figure 4.** Pre- and post-event Sentinel-2 mosaics in true colour combination and the training areas selected for different classes.

The radiometric calibration and the speckle filtering are crucial steps in the pre-processing of SAR data. The calibration process is essential for the quantitative analysis for the multi-sensor and multi-temporal SAR images. As a result, radar backscatter can be represented. The  $\sigma^0$  bands were also produced in the calibration procedure. The variation in radar return within a pixel caused by multiple scattering sources causes a speckle effect in SAR data (Giustarini et al., 2015). The reduction of the speckle effect is important for obtaining accurate results in further processing (Clement et al., 2018; Carreño Conde et al., 2019). For this purpose, SNAP software (SNAP, 2018) offers various filter types with different filter sizes such as Lee, Lee Sigma, Refined Lee, and Gamma Map. At this step, Lee Sigma speckle filter method was used to reduce the speckle, as the Lee

Sigma speckle filter has been found to be useful in many studies (Lee and Pottier, 2009; Jaybhay and Shastri, 2015; Tavus et al., 2018).

The  $\sigma^0$  results were terrain-corrected and orthorectified using the Range Doppler Terrain Correction algorithm. As external topographic dataset, the SRTM (Shuttle Radar Topography Mission) with ca. 30 m spatial resolution was utilized; and the bilinear interpolation resampling technique was performed to obtain the geometrically corrected data.

As shown in Figure 3, the SWIR band (B11) images of Sentinel-2 sensor was also downsampled to the resolution of the visible and the NIR band images (i.e. 10 m). By merging the data of the two acquisitions, mosaics representing the pre- and post-flood conditions were created and then cropped according to the study area extent.

The NDVI (normalized difference vegetation index) and MNDWI (modified normalized difference water index) indices allowing the identification of flooded areas were generated from the pre- and post-event mosaics. These spectral indices provide useful information regarding the effects of floods on the soil even after the water is absorbed (Notti et al., 2018; Mohammadi et al., 2017).

Finally, all features were collected in four different stacks with different feature set combinations. The stacks were designed for the purpose of investigating the success of classification by using different feature types as input; and thus evaluate the use of Sentinel-1 sensor for flood mapping as the sole data source. The stack combinations were as following:

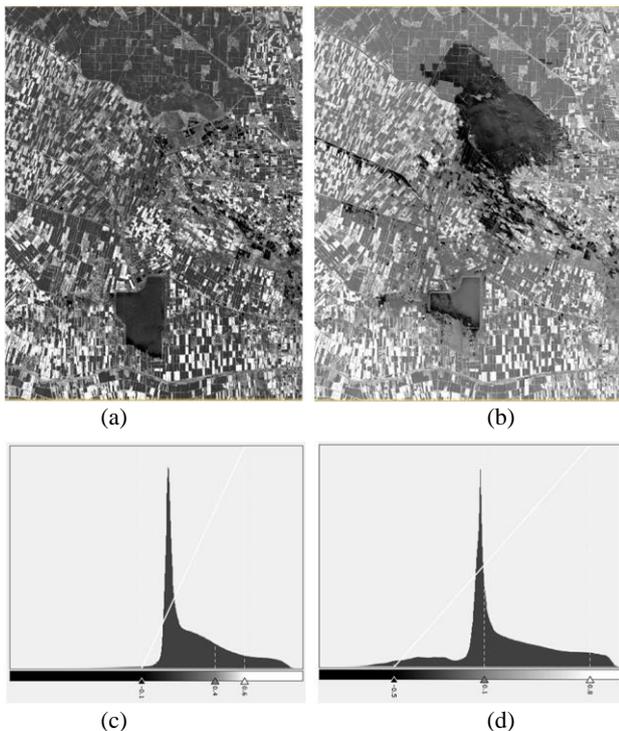
- *Stack 1:* Sentinel-1 flood and non-flood VV bands,
- *Stack 2:* Sentinel-1 flood and non-flood VH bands,
- *Stack 3:* Sentinel-1 flood and non-flood VV and VH bands,
- *Stack 4:* Sentinel-1 flood and non-flood VV and VH bands, Sentinel-2 bands, NDVI and MNDWI indices.

The RF algorithm was employed for the classification of the defined LULC classes with the help of the training data using the SNAP software. The RF belongs to the family of decision tree methods, and was proposed by Breiman (2001). In the RF, the decision tree is created randomly at the training stage; and the trees are evaluated for the best score. A total of 303.640 training samples exists in the area, and 48.000 of them were used for testing. 3-fold cross validation was applied in the SNAP software using 300 as tree size value. The results of the classification were assessed in terms of overall accuracy for all classes.

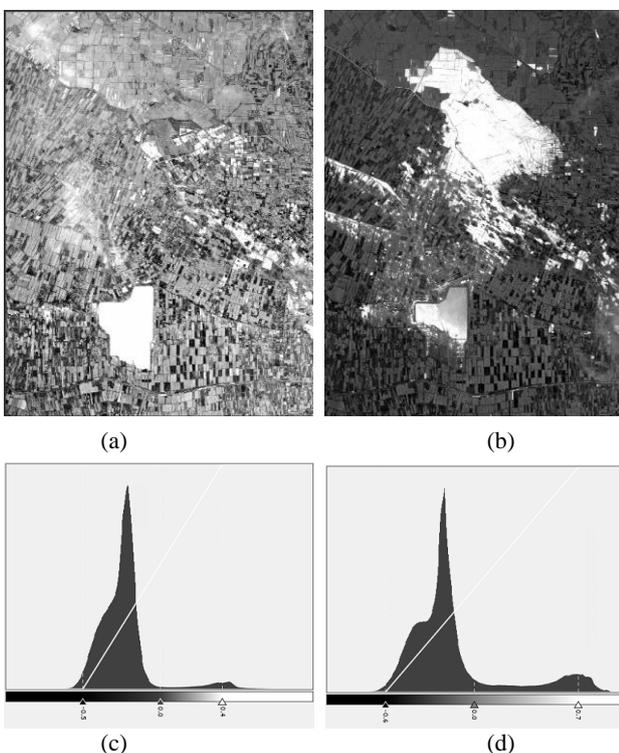
## 4. RESULTS AND DISCUSSIONS

### 4.1 Pre-processing Results

The pre- and post-event NDVI and MNDWI images, and their histograms are presented in Figures 5 and 6, respectively. The index values range between (-1:1) in both. The values were scaled to eight-bit images for visualisation purposes. As can be seen in Figure 5, flooded areas can be recognized in the NDVI feature, and the agricultural areas under the water also have high contrast. Figure 6 shows that the MNDWI reflect distinctive characteristics for the areas covered by water.



**Figure 5.** (a) Pre-event and (b) post-event NDVI indices; (c) pre-event and (d) post-event NDVI histograms.

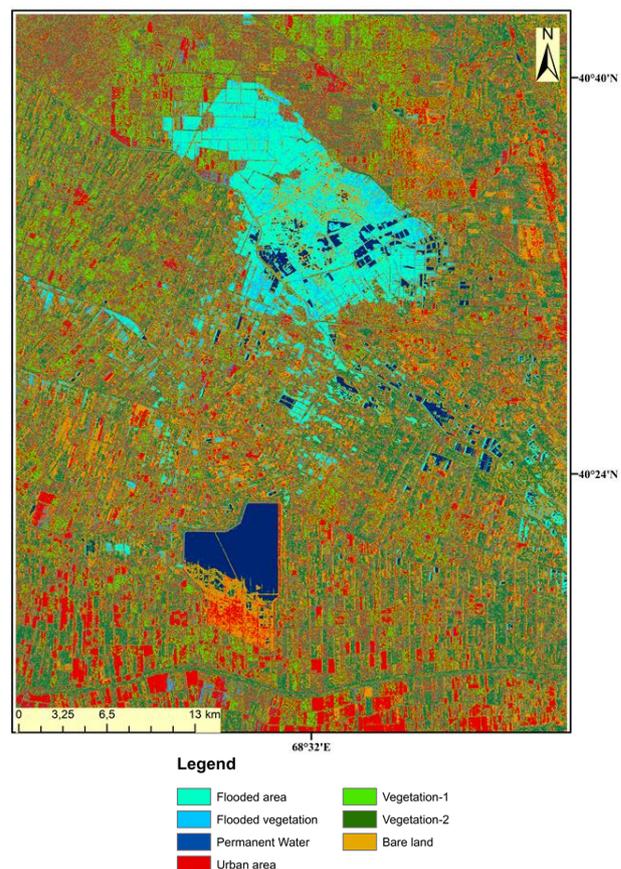


**Figure 6.** (a) Pre-event and (b) post-event MNDWI indices; (c) pre-event and (d) post-event MNDWI histograms.

#### 4.2 Flood Extent Map

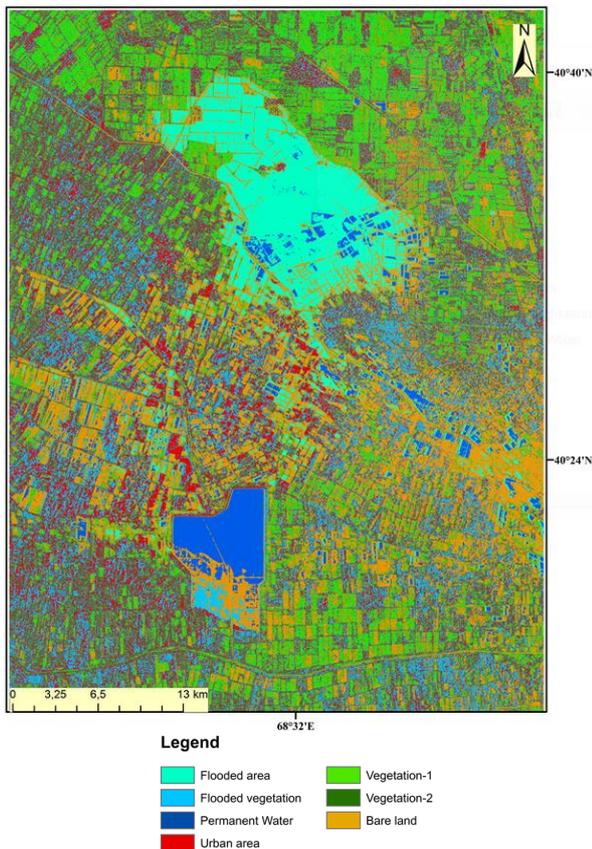
Four different LULC maps obtained from four different stacks (feature sets) using the RF are presented in Figures 7, 8, 9 and 10. The maps include the classes of permanent water, flooded area, flooded vegetation, urban, bare land, and two types of agricultural areas as Vegetation-1, Vegetation-2. The overall

accuracies achieved from the Stacks 1, 2, 3 and 4 are 65%, 62%, 87% and 99%, respectively. The overall accuracy values show that the integration of the Sentinel-1 and Sentinel-2 feature sets provide highest prediction performance for the classification of the defined classes. The use of single polarizations of Sentinel-1 VV (Stack 1, Figure 7) or VH (Stack 2, Figure 8) yielded to lower performances, and therefore is not recommended for flood mapping here. When the Sentinel-1 VV&VH images were used together, although the flooded areas could be separated, the flooded vegetation could still not be extracted (Figure 9). On the other hand, one should consider that the training polygons were delineated on the optical images. Especially the flooded vegetation class is not visible in SAR images. Thus, this class is not separable in SAR data. Here, the first three stacks (SAR polarizations only) were evaluated to analyse the contribution of SAR features to the prediction success. In order to evaluate the prediction performance of SAR-only features accurately, the training data must also be selected from these images, although the ground truth (flooded areas) remain the same (unchanged).

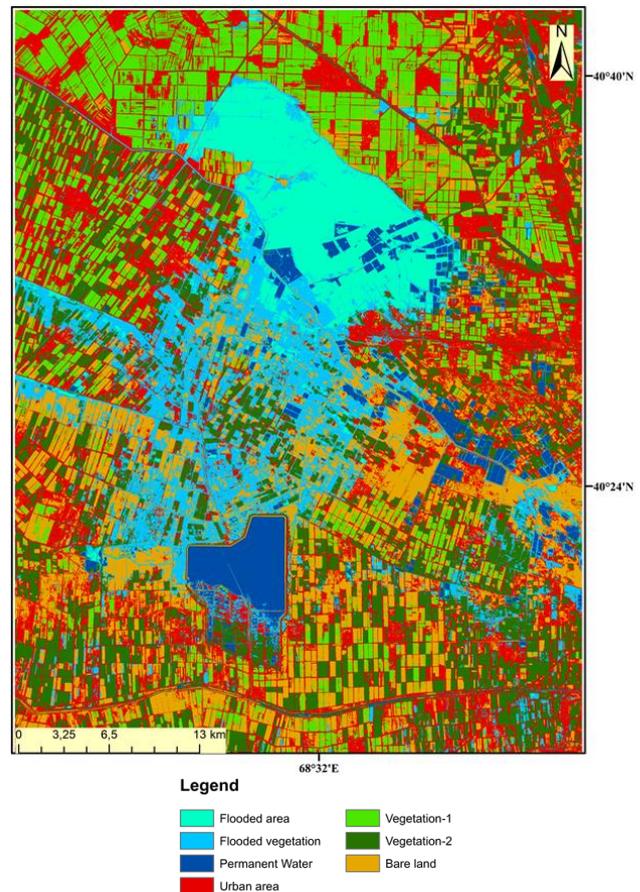


**Figure 7.** Results of the RF classification with the pre- and post-event Sentinel-1 VV images.

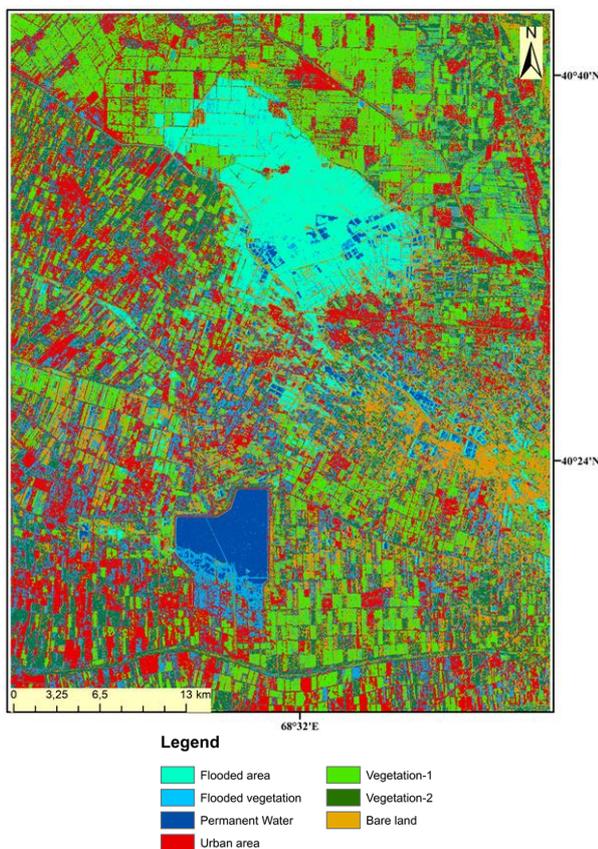
After visual assessments of the flood map given in Figure 10, it was observed that the non-flooded vegetation and settlement areas could be clearly identified. These areas exhibited noisy pattern in Figures 7, 8 and 9. In Figure 9, flooded vegetation could also not be determined since Sentinel-2 data was not employed. These areas were rather muddy with mixed colours in as observed the optical images. These areas are especially visible in the MNDWI index image (Figure 6).



**Figure 8.** Results of the RF classification with the pre- and post-event Sentinel-1 VH image.



**Figure 10.** Results of the RF classification with Stack-4 containing all features extracted from the Sentinel-1 and Sentinel-2 datasets.



**Figure 9.** Results of the RF classification with the pre- and post-event Sentinel-1 VV +VH images.

## 5. CONCLUSIONS AND FUTURE WORK

In this study, a fusion methodology for the integrated use of Sentinel-1 SAR and Sentinel-2 datasets for a flooded region in Sirdaryo Region, Uzbekistan was proposed. A flash flood occurred on May 1<sup>st</sup>, 2020 in the area after the break of a dam constructed on Sardoba Reservoir; and caused losses of lives, evacuations from the settlements, and severe damages in agricultural areas. Accurate mapping of the flooded areas and the flooded vegetation has become thus crucial for the damage assessment and disaster mitigation. Thanks to the frequent data collection schedule of the Sentinel-1 and Sentinel-2 satellites of ESA globally, the maps could be produced in this study by using an ensemble machine learning method (random forest). Seven different LULC types (permanent water, flooded area, flooded vegetation, urban, bare land, and two types of agricultural areas such as vegetation-1, vegetation-2) were identified in the region and the training samples were manually delineated on the Sentinel-2 images and stored as vector data (polygons).

Four different feature sets derived from the Sentinel-1 and Sentinel-2 datasets were evaluated in the proposed methodology. The combinations involved the use of Sentinel-1 VV (Stack 1), VH (Stack-2), VV+VH (Stack 3) polarization images. The last stack was a combination of feature sets obtained from both the Sentinel-1 and Sentinel-2 datasets, including the NDVI and MDNWI indices. The results showed that the integration of both datasets yielded to high prediction

performance for these areas with an overall accuracy of 0.99. Stack 3 (Sentinel-1 VV+VH polarization images) resulted in 0.87 overall accuracy. The future works of the study include the assessment of the potential causes of the dam break; and the development of further evaluations for producing the flood extent maps by using the SAR data only.

#### ACKNOWLEDGEMENTS

This research is part of the Ph.D. thesis work of the first author.

#### REFERENCES

- Anusha, N., Bharathi, B., 2020. Flood detection and flood mapping using multi-temporal synthetic aperture radar and optical data. *The Egyptian Journal of Remote Sensing and Space Science*, 23(2), 207-219. doi.org/10.1016/j.ejrs.2019.01.001
- Amitrano, D., Di Martino, G., Iodice, A., Riccio, D., Ruello, G., 2018. Unsupervised rapid flood mapping using Sentinel-1 GRD SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 56(6), 3290-3299. doi.org/10.1109/TGRS.2018.2797536
- Breiman, L., 2001: *Random forests*. Machine Learning. Springer, 45(1), 5-32.
- Carreño Conde, F., De Mata Muñoz, M., 2019. Flood monitoring based on the study of Sentinel-1 SAR images: The Ebro River case study. *Water*, 11(12), 2454. doi.org/10.3390/w11122454
- Chan, K.L., Qin K., 2017: Biomass burning related pollution and their contributions to the local air quality in Hong Kong. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-2/W7, 29-36. doi.org/10.5194/isprs-archives-XLII-2-W7-29-2017
- Clement, M.A., Kilsby, C.G., Moore, P., 2018. Multi-temporal synthetic aperture radar flood mapping using change detection. *Journal of Flood Risk Management*, 11(2), 152-168. doi.org/10.1111/jfr3.12303
- Copernicus, 2020. Copernicus Open Access Hub. <https://scihub.copernicus.eu/dhus/#/home> (5 July 2020).
- Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Marchese, F., Bargellini, P., 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment*, 120, 25-36.
- Fergana News, 2020. Uzbek foreign ministry criticised for distorting conclusions of Dutch report on Sardoba dam burst. <https://en.fergana.ru/news/118410/> (20 January 2020).
- Gascon, F., Bouzinac, C., Thépaut, O., Jung, M., Francesconi, B., Louis, J., Lonjou, V., Lafrance, B., Masser, S., Gaudel Vacaresse, A., Languille, F., Alhammoud, B., Viallefon, F., Pflug, B., Bieniarz, J., Clerc, S., Pessiot, L., Tremas, T., Cadau, E., De Bonis, R., C., Martimort, P., Fernandez, V., 2017. Copernicus Sentinel-2A calibration and products validation status. *Remote Sensing*, 9(6), 584.
- Giustarini L., Vernieuwe H., Verwaeren J., Chini M., Hostache R., Matgen P., Verhoest N.E.C. De Baets, B., 2015. Accounting for image uncertainty in SAR-based flood mapping. *International Journal of Applied Earth Observation and Geoinformation*, 34, 70-77. doi.org/10.1016/j.jag.2014.06.017
- Jaybhay, J., Shastri, R., 2015. A study of speckle noise reduction Filters. *Signal & Image Processing: An International Journal (SIPIJ)*, 6(3), 71-80. doi.org/10.5121/sipij.2015.6306
- Lee, J.S., Pottier, E., 2009. Polarimetric Radar Imaging: From basics to applications. In *Optical Science and Engineering*; CRC Press: Boca Raton, FL, USA, Volume 142.
- Mohammadi, A.; Costelloe, J.F., Ryu, D., 2017. Application of time series of remotely sensed normalized difference water, vegetation and moisture indices in characterizing flood dynamics of large-scale arid zone floodplains. *Remote Sensing Environment*, 190, 70-82. <https://doi.org/10.1016/j.rse.2016.12.003>
- Nagler, T., Rott, H., Ripper, E., Bippus, G., Hetzenecker, M., 2016: Advancements for snowmelt monitoring by means of sentinel-1 SAR. *Remote Sensing*, 8(4), 348.
- Notti, D., Giordan, D., Caló, F., Pepe, A., Zucca, F., Galve, J.P., 2018. Potential and limitations of open satellite data for flood mapping. *Remote Sensing*, 10 (11), 1673. doi.org/10.3390/rs10111673
- 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 Sensing*, 10(2), 237. doi.org/10.3390/rs10020237
- Putz, C., 2020. Uzbekistan dam collapse was a disaster waiting to happen. <https://www.thethirdpole.net/2020/06/23/uzbekistan-dam-collapse/> (30 January 2021).
- Simonow, E., 2020. Uzbekistan dam collapse was a disaster waiting to happen. <https://www.thethirdpole.net/2020/06/23/uzbekistan-dam-collapse/> (20 January 2021)
- SNAP, 2018. Sentinel Application Platform. <http://step.esa.int/main/toolboxes/snap/> (1 September 2018).
- Tavus, B., Kocaman, S., Gokceoglu, C., Nefeslioglu, H.A., 2018. Considerations on the Use of Sentinel-1 Data in Flood Mapping In Urban Areas: Ankara (Turkey) 2018 Floods. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 42(5), 575-581. <https://doi.org/10.5194/isprs-archives-XLII-5-575-2018>
- Tavus, B., Kocaman, S., Nefeslioglu, H.A., Gokceoglu, C., 2019. Flood Mapping Using Sentinel-1 SAR Data: A Case Study of Ordu 8 August 2018 Flood. Bi-annual Symposium of Turkish Society for Photogrammetry and Remote Sensing, April 25-27, Aksaray, Turkey.
- Tavus, B., Kocaman, S., Nefeslioglu, H.A., Gokceoglu, C., 2020. A fusion approach for flood mapping using Sentinel-1 and Sentinel-2 datasets. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 43, 641-648.

Tong, X., Luo, X., Liu, S., Xie, H., Chao, W., Liu, S., Jiang, Y., 2018. An approach for flood monitoring by the combined use of Landsat 8 optical imagery and COSMO-SkyMed radar imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 136, 144-153. doi.org/10.1016/j.isprsjprs.2017.11.006

Vanama, V. S. K., Rao, Y. S., Bhatt, C. M., 2021. Change detection based flood mapping using multi-temporal Earth Observation satellite images: 2018 flood event of Kerala, India. *European Journal of Remote Sensing*, 54(1), 42-58. doi.org/10.1080/22797254.2020.1867901

Xu, H., 2006. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, 27(14), 3025–3033. doi.org/10.1080/01431160600589179