FloodHub Diachronic Mapping Service in the transboundary Evros river basin

Alexia Tsouni[1], *Anastasis Katsos*[1], *Haris Kontoes*[1]

[1] National Observatory of Athens, Institute of Astronomy, Astrophysics, Space Applications & Remote Sensing, Operational Unit "BEYOND Center of Earth Observation Research and Remote Sensing", Athens, Greece

ABSTRACT

Among the natural disasters, flood had the highest occurrence in 2022 (176 in a total of 387), the second highest death toll (7.954 in a total of 30.704), and the second highest human impact (57.1 in a total of 185) according to the most recent annual report "2022 Disasters in numbers" published by the Centre for Research on the Epidemiology of Disasters (CRED) Université Catholique de Louvain, and the United States Agency for International Development (USAID). Decision makers and civil protection authorities need reliable flood situation awareness in order to improve flood resilience and rapid response during a severe flood. For this purpose, the FloodHub Diachronic Mapping Service has been developed by the Operational Unit "BEYOND Center of Earth Observation and Satellite Remote Sensing", IAASARS, National Observatory of Athens (NOA). It is implemented operationally in the transboundary river basin of Evros (Greece-Bulgaria-Turkey), in the framework of the Balkan Flood Pilot of the Working Group Disasters of the Committee on Earth Observation Satellites (CEOS). Evros is a river basin where severe floods occur almost every year, and it was recently affected by a catastrophic wildfire which burnt around 94,000 hectares in August 2023. The FloodHub Diachronic Mapping Service is based on a fully automated system that searches, downloads and preprocesses Sentinel-1A/B SAR data and then, using machine learning, automatically maps water surfaces in a user-defined area. It uses the Copernicus Dataspace Ecosystem and covers the time period since 2018. It also includes the burnt scar mapping produced by the FireHub Diachronic Mapping Service, which is a fully automated system since 1984 using Sentinel-2 (now) and Landsat (earlier) data. This is very important in order to examine the impact of the wildfires on the floods. The FloodHub Diachronic Mapping Service is a reliable flood mapping tool in line with the requirements for the implementation of the EU Floods Directive 2007/60/EC, the Sendai Framework for Disaster Risk Reduction, the UN SDGs, as well as the GEO's Societal Benefit Areas.

Keywords: natural and man-made disasters, natural hazards, flood, wildfire, diachronic mapping, monitoring, big earth data, resilient society, disasters risk reduction, civil protection, remote sensing, earth observation

1. INTRODUCTION

Among the natural disasters, flood had the highest occurrence in 2022 (176 in a total of 387), the second highest death toll (7.954 in a total of 30.704), and the second highest human impact (57.1 in a total of 185) according to the most recent annual report "2022 Disasters in numbers" [1] published by the Centre for Research on the Epidemiology of Disasters (CRED) Université Catholique de Louvain, and the United States Agency for International Development (USAID) (**Figure 1**).

Tenth International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2024), edited by A. Christofe, S. Michaelides, D. Hadjimitsis, C. Danezis, K. Themistocleous, N. Kyriakides, G. Schreier, Proc. of SPIE Vol. 13212, 132120Y · © 2024 SPIE · 0277-786X · doi: 10.1117/12.3037281

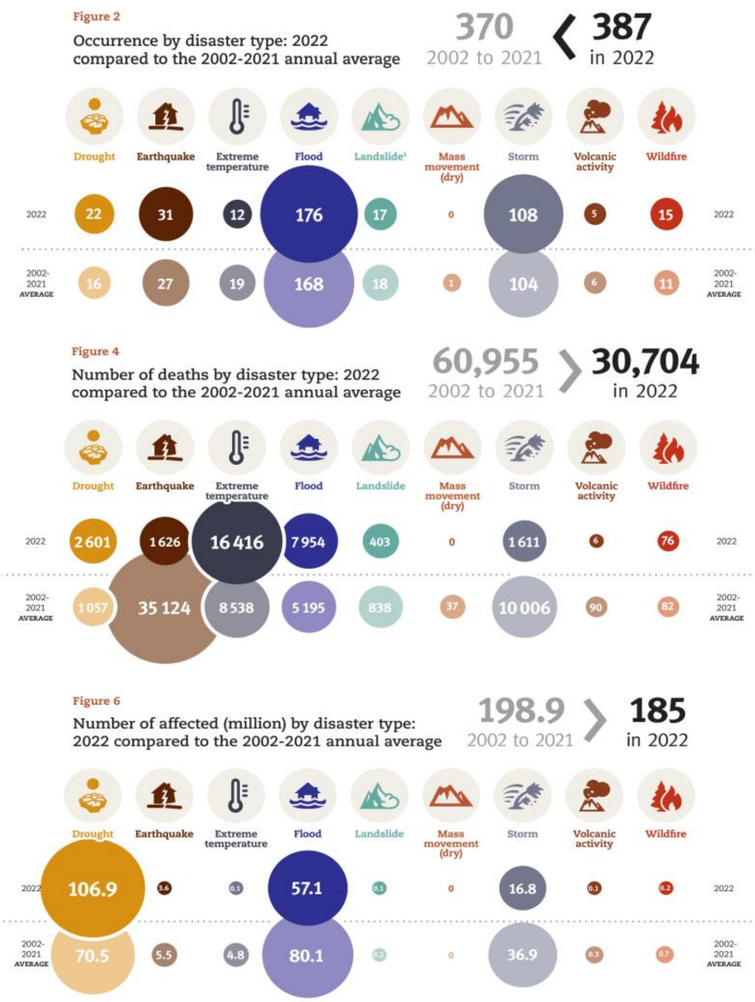

Figure 1: Worldwide disaster statistics according to the most recent annual report "2022 Disasters in numbers"

Decision makers and civil protection authorities need reliable flood situation awareness in order to improve flood resilience and rapid response during a severe flood.

For this purpose, the FloodHub Diachronic Mapping Service has been developed by the Operational Unit "BEYOND Center of Earth Observation and Satellite Remote Sensing", IAASARS, National Observatory of Athens (NOA). It is implemented operationally in the transboundary river basin of Evros, in the framework of the Balkan Flood Pilot of the Working Group Disasters of the Committee on Earth Observation Satellites (CEOS) [2].

2. STUDY AREA

Evros is a transboundary river basin extending in Greece, Bulgaria and Turkey. The total surface of the river basin is 53.075 km^2, with its biggest part in Bulgaria (35.222 km^2), and then Turkey (14.502 km^2), and Greece (3.351 km^2), as presented in **Figure 2**.

Severe floods occur almost every year in the Evros river basin. In addition, it was recently affected by a catastrophic wildfire, which burnt around 94,000 hectares in August 2023, according to the NOA/BEYOND/FireHub Diachronic Mapping Service (**Figure 3**).

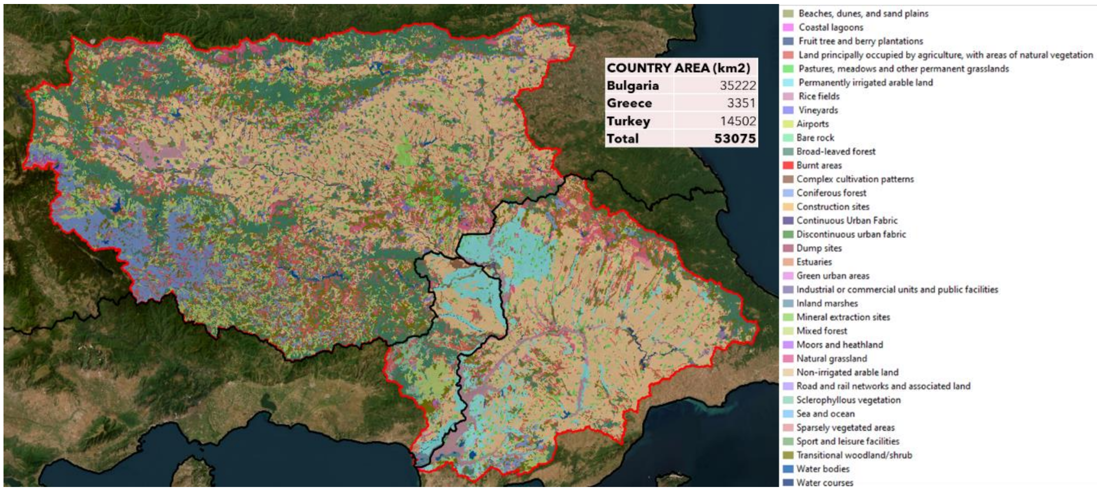

Figure 2: The transboundary river basin of Evros

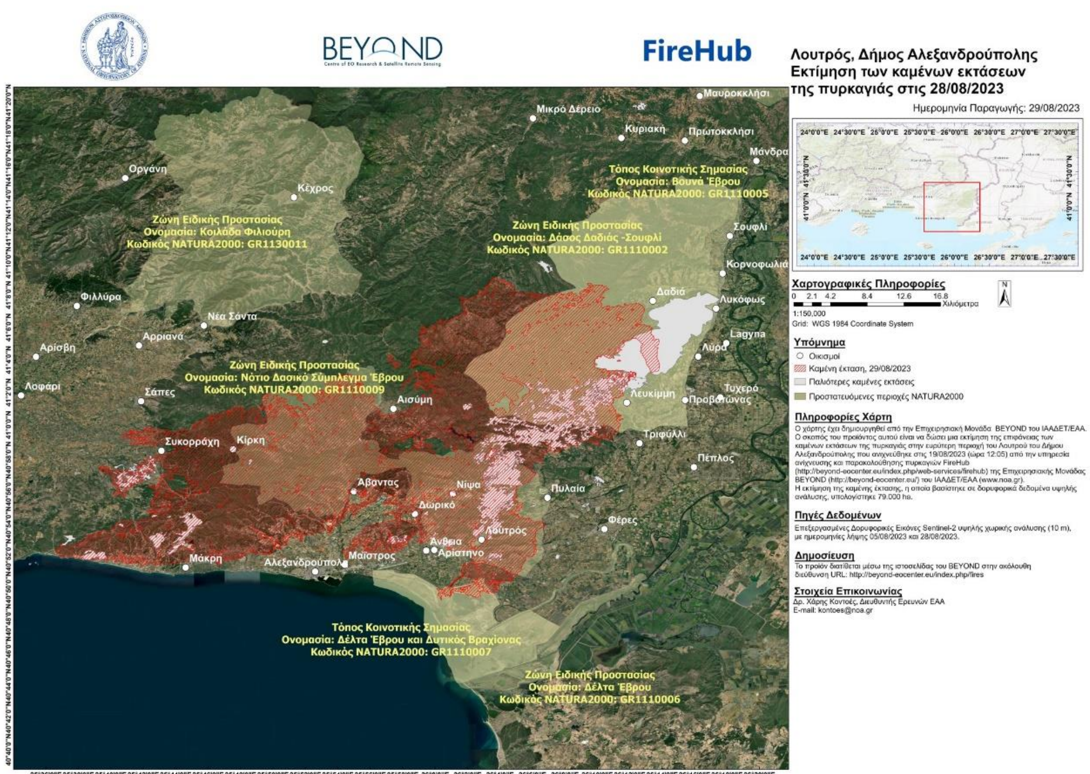

Figure 3: NOA/BEYOND/FireHub burnt scar mapping of the wildfire in Evros river basin in August 2023 (in red) following processing of Sentinel-2 satellite data

3. METHODOLOGY

Figure 4 presents the architecture of the FloodHub Diachronic Mapping Service, which is defined by a five-step solution. These steps span from searching and downloading Sentinel-1 Synthetic Aperture Radar (SAR) data, to training and using a model for identifying floods in the given area of interest.

1. Search
2. Download
3. Pre-Process
4. Ingestion of the Pre-Processed images to the Open Data Cube
5. Flood Mapping

The Sentinel-1 SAR products are pre-processed to get the desired VV and VH polarisations.

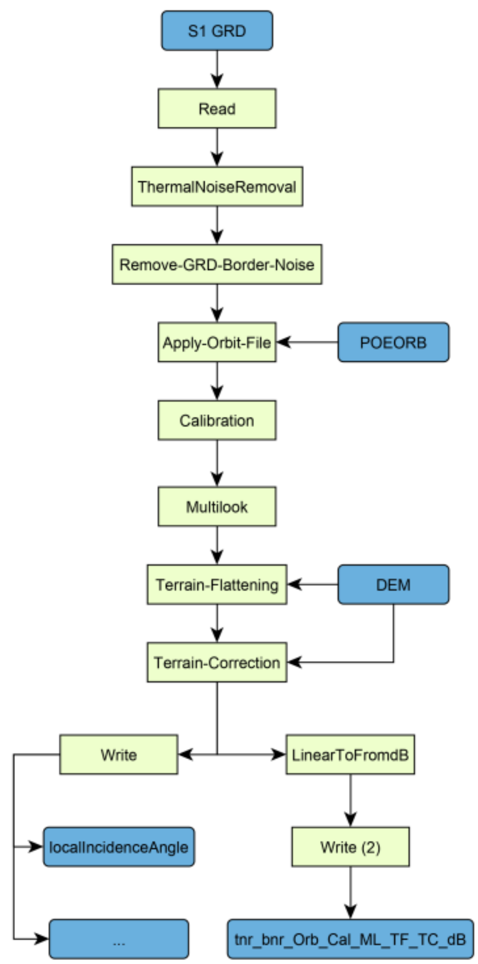

Figure 4: FloodHub Diachronic Mapping Service architecture

1. Search

The service has a configuration file where the desired fields for the query search are given as input, like date, sensing start and stop date, satellite name, area of interest (AOI), pagination, and ordering specified by a field (like date of ingestion).

After the service sends its request, it receives a response (in json format) and inserts the data it needs in the database.

2. Download

After indexing all the desired products in the database, the download process begins. For each satellite product it first searches in CDSE's OData API with the product ID, and inserts its checksum (MD5) into the database. This is a necessary step because in the last step the application uses the OpenSeach API which does not have the checksum of the product as a field.

When the product is downloaded, a routine starts to compare the checksum of the downloaded product and the checksum given by the OData API. If the fields do not match, the downloaded product is deleted and redownloaded.

This continues for each product in the database based on the first step (Search).

3. Pre-Process

After successfully ingesting a product, the Pre-Process step begins. Here, the service interfaces the application ESA SNAP (Sentinel Application Platforms) using the library pyroSAR to perform the pre-processing steps:

- Read of the image
- Border Noise Removal
- Calibration
- Thermal Noise Removal
- Apply Orbit File
- Subset
- Speckle Filtering
- Terrain Flattening
- Range Doppler Terrain Correction
- Conversion to dB

The pre-processing workflow for the Sentinel-1 data involves several crucial steps to prepare the data for analysis. Initially, the raw image is loaded into the processing software.

The first step is border noise removal, which eliminates noise and artefacts from the image edges to ensure cleaner data. Calibration follows, converting raw pixel values to calibrated backscatter values (sigma naught, gamma naught, or beta naught), thus standardising the data. Then the thermal noise removal is performed to reduce noise caused by the sensor's thermal properties.

After that, the area of interest is applied, reducing processing time and computational load. Speckle filtering is applied to mitigate the granular noise inherent in SAR images. Terrain flattening is conducted to correct backscatter variations due to topography. Range Doppler Terrain Correction (RDTC) is then used to correct geometric distortions caused by the side-looking radar geometry and terrain effects, accurately geocoding the image to a map coordinate system.

Finally, the calibrated backscatter values are converted to a logarithmic scale (decibels), enhancing the interpretability of the data by representing the wide range of backscatter intensities in a standardised format.

The service uses the Geographic Reference System EPSG:4326, using a geospatial resolution of 10 meters and the geocoding type of Range-Doppler. This method is based on the Visvalingam-Whyatt algorithm. For each RD images that is successfully ingested and pre-processed, a pair of images is created of type .tiff, of for the backscatter of the channel VV and one backscatter for the channel VH.

4. Ingestion of the Pre-Processed images to the Open Data Cube

After completing the pre-processing step, the pre-processed images are inserted in the Open Data Cube (ODC). For each image a yaml (.yml) file is created, which is important for the insertion of the image to the ODC.

ODC allows cataloguing a big quantity of earth sensing data. It has an existing python API for requesting data in the database. It allows scaling processing of the saved data on a big scale. Open Data Cube also oversees the origin of all the data so it allows qualitative control and update accordingly.

5. Flood Mapping

When selecting a new area for flood mapping, the application must first train a new model specific to that region. Due to the localised nature of flood patterns, the model is area-specific, but the pipeline is optimised to ensure a quick turnaround for the specified area. The algorithm chosen for training the model is the Random Forest, a widely used and robust method suitable for the classification and regression tasks encountered. Training the model is performed in a Jupyter Notebook using Python. From the executed tests Random Forest seems to have the quickest turnaround for this particular architecture.

For creating the ground truth data, libraries such as OpenDataCube, xarray, numpy, and matplotlib are used. The model training utilises numpy, scikit-learn, and joblib libraries. The model requires a variable sample of SAR data from the target area. Once trained on a general area, the model can predict water presence in any AOI within that region.

During the prediction phase, incoming data is evaluated by the model. The output is a flood map of the area, where each pixel is classified as either water (1) or non-water (0). This classification is visually represented as a black and white image, with black pixels indicating non-water areas and white pixels indicating water-covered areas.

4. PRELIMINARY RESULTS

The results of the FloodHub Diachronic Mapping Service are presented on a free and open user-friendly web GIS platform [3] (**Figure 5**), which includes:

- The NOA/BEYOND FloodHub diachronic flood mapping service using Sentinel-1 satellite data (automated processing since 2018)
- The NOA/BEYOND FireHub burnt scar mapping service using Sentinel-2 satellite data (since 1984 – starting with Landsat)
- The NATURA 2000 protected areas covering Europe's most valuable and threatened species and habitats.

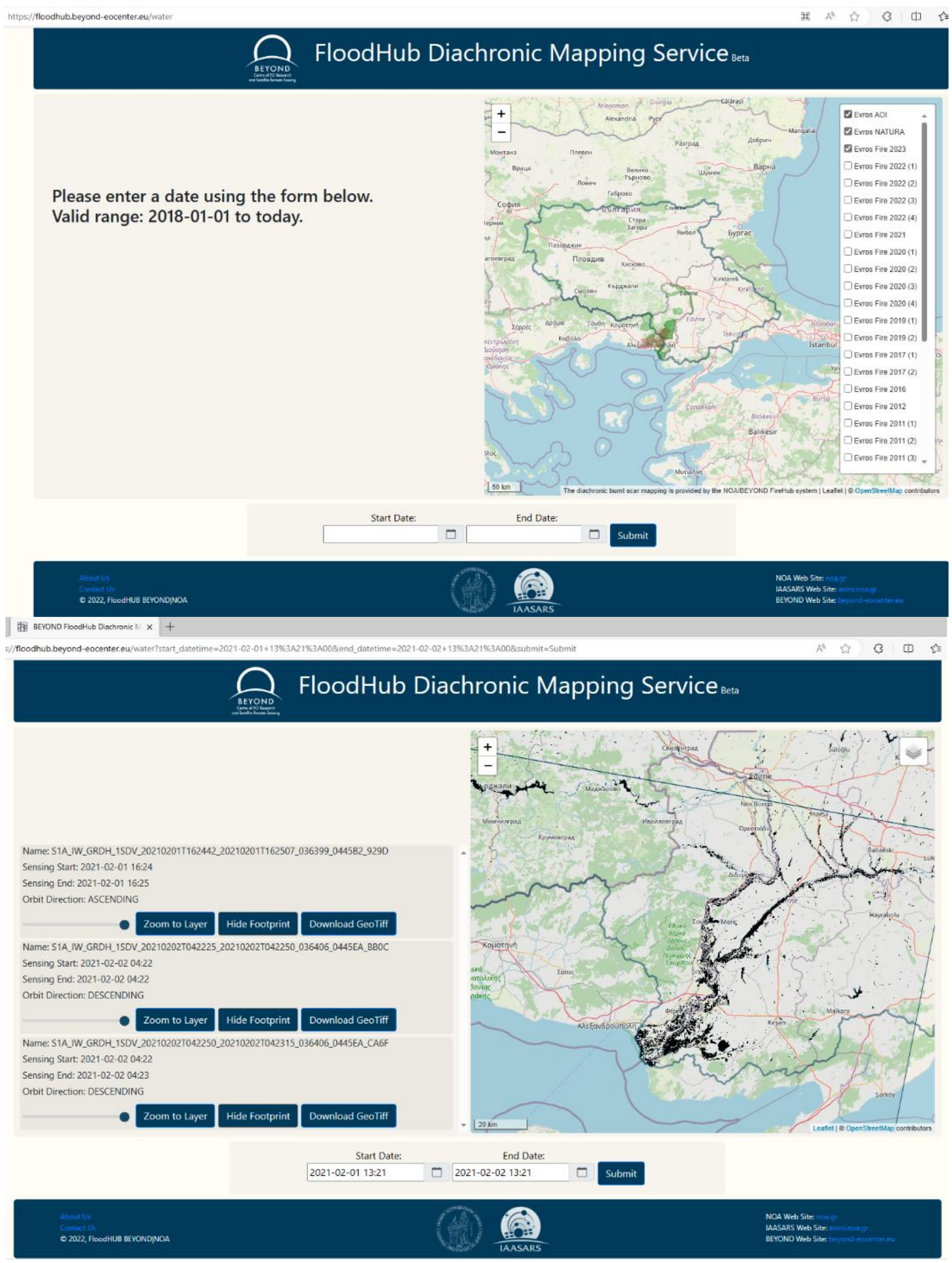

Figure 5: FloodHub Diachronic Mapping Service web GIS platform

Several flood events are detected over the years, some of them severe ones. Furthermore, wildfires are also registered, some of them even affecting the NATURA 2000 protected areas, like the catastrophic wildfire in August 2023.

5. CONCLUSION

The FloodHub Diachronic Mapping Service is based on a fully automated system that searches, downloads and preprocesses Sentinel-1A/B SAR data and then, using machine learning, automatically maps water surfaces in a user-defined area. It uses the Copernicus Dataspace Ecosystem and covers the time period since 2018. It also includes the burnt scar mapping produced by the FireHub Diachronic Mapping Service, which is a fully automated system since 1984 using Sentinel-2 (now) and Landsat (earlier) data. This is very important in order to examine the impact of the wildfires on the floods, which could be a very important future research work.

The FloodHub Diachronic Mapping Service is a reliable flood mapping tool in line with the requirements for the implementation of the EU Floods Directive 2007/60/EC [4], the Sendai Framework for Disaster Risk Reduction [5], the UN SDGs [6], as well as the GEO's Societal Benefit Areas.

REFERENCES

[1] Annual report "2022 Disasters in numbers", Centre for Research on the Epidemiology of Disasters (CRED), Université Catholique de Louvain, and the United States Agency for International Development (USAID), available at: https://www.cred.be/sites/default/files/2022_EMDAT_report.pdf

[2] Working Groups of the Committee on Earth Observation Satellites (CEOS), available at: https://ceos.org/ourwork/workinggroups/disasters/

[3] FloodHub Diachronic Mapping Service, Operational Unit "BEYOND Center of Earth Observation and Satellite Remote Sensing", IAASARS, National Observatory of Athens (NOA), available at: https://floodhub.beyond-eocenter.eu/diachronic

[4] European Union Floods Directive 2007/60/EC, available at: https://eur-lex.europa.eu/eli/dir/2007/60/oj

[5] United Nations (UN) Sendai Framework for Disaster Risk Reduction, available at: https://www.undrr.org/implementing-sendai-framework/what-sendai-framework

[6] United Nations (UN) Sustainable Development Goals (SDGs), available at: https://sdgs.un.org/goals