EARTH OBSERVATION AND SOCIAL MULTIMEDIA DATA FUSION FOR NATURAL HAZARDS AND WATER MANAGEMENT: THE H2020 EOPEN PROJECT PARADIGM

Gialampoukidis I.[1], A. Moumtzidou[1], M. G. Scarpino[2], G. Palumbo[2], S. Vrochidis[1], I. Kompatsiaris[1], F. Zaffanella[3], D. Norbiato[3], M. Ferri[3], & G.Vingione[2]

[1] Information Technologies Institute, Centre for Research and Technology Hellas, Thessaloniki, Greece – e-mail: heliasgj@iti.gr, moumtzid@iti.gr, stefanos@iti.gr, ikom@iti.gr
[2] Serco S.p.A., Rome, Italy – e-mail: Gabriella.Scarpino@serco.com, Giovanna.Palumbo@serco.com, Guido.Vingione@serco.com
[3] Autorità di Bacino Distrettuale delle Alpi Orientali, Venice, Italy – e-mail: francesco.zaffanella@adbve.it, daniele.norbiato@adbve.it, michele.ferri@adbve.it

KEY POINTS

- *Flood detection using satellite data*
- *Data fusion and integration of non-EO data technologies within the H2020-EOPEN project*
- *Future directions in Emergency Management and Early Warning Services that combine EO with non-EO data*

1 INTRODUCTION

Since the adoption of the 2006/60/EU Directive by part of the European Commission the concept of water management has changed; moreover following the adoption of the directive 2007/60/EU about flood risk assessment, water management has become relevant not only in sectors like agriculture, tourism and water safety, but also for the security of citizens. Hazardous flood preparedness and recovery is one aspect of water management which addresses various issues ranging from irrigated agriculture to leisure activities - the contribution of Earth Observation (EO) data in such contexts has been elucidated in the past decades and is at the basis of some current Copernicus Services targeting water governance.

Several H2020 projects develop decision support systems for disaster resilience using social media, weather data and Copernicus[1] data, with a particular focus on flood events and water management. For example, H2020-beAWARE[2] proposes an integrated solution to support forecasting, early warnings, transmission and routing of the emergency data, aggregated analysis of multimodal data and management of the coordination between the first responders and the authorities. Moreover, I-REACT[3] is an innovation project funded by the European Commission. The project aims to integrate emergency management data coming from multiple sources, including that provided by citizens through social media and crowdsourcing to effectively prevent and/or react against natural disasters. The H2020-EOPEN[4] (opEn interOperable Platform for unified access and analysis of Earth observatioN data) project focuses on three use-cases – two of them, namely, "Flood risk assessment and prevention" and "Food security", involve water management. The EOPEN project has already been communicated in ESA BiDS17 (Vingione et al. 2017) and in ESA MWBS2018[5], exchanging ideas and knowledge by its interaction with other H2020-EO-2-2017 projects and similar ESA initiatives. Recently, the user requirements meeting in Vicenza, Veneto area, Italy has already identified many user requirements to be taken into account in the development of the EOPEN platform. The MOSES[6] (Managing crOp water Saving with Enterprise Services) project output is a Platform, based on a variety of technologies and data sources, including EO data, aimed mainly to support irrigation water managers for water procurement, delivery and distribution; moreover, normally irrigation consortia managers face also the management of water outflow on the occurrence of floods. Both, the EOPEN and the

[1] The Copernicus program Copernicus is the European Earth Observation and Monitoring Programme, previously known as GMES (Global Monitoring for Environment and Security) - http://www.copernicus.eu/
[2] http://beaware-project.eu/
[3] http://www.i-react.eu/
[4] http://eopen-project.eu/
[5] http://mwbs2018.esa.int/
[6] http://www.moses-project.eu/

MOSES[7] Projects address water management issues which are connected allowing the exploitation of synergies between such projects - an example is mentioned in the following section.

In this work we present the identification of flooded areas using Earth Observation data by combining a Mahalanobis Distance-based classification for mask creation and a series of morphological post-processing steps for flood mask correction (Section 2). Furthermore, in Section 3, we present EOPEN-related technologies aiming to fuse non-EO data in the case of flood events in an innovative way.

2 FLOOD MONITORING USING EARTH OBSERVATION DATA

In this section, we present an approach (Michail et al., 2018) which identifies flooded areas by considering Earth Observation (EO) satellite images. In particular, it classifies all pixels of a satellite image as a flooded area or not. The method combines a Mahalanobis Distance-based classifier trained for recognizing pixels with water for flood mask creation and morphological post-processing of mask produced for flood mask correction, to separate flood from non-flood areas inside satellite image (see Figure 1).

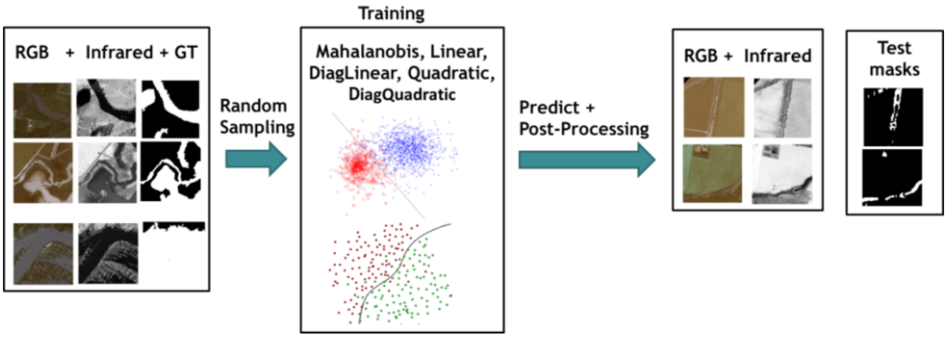

Figure 1. Water pixel-based detection method.

Starting with the classification step, it is performed by considering as input four-dimensional feature vectors that consist of the 4 color channels derived directly from the pixels of the EO images. The channels considered are Red (R), Green (G), Blue (B) and Near-Infrared (NI). During the training step of the classifier, a training set of annotated pixels is produced by considering randomly selected pixels from the EO image training dataset. Apart from the 4-dimensional feature vector a label with values 0 or 1 is considered that signifies water existence. Eventually, training of a classifier is realized by using the pixel feature vectors and their labels using discriminant analysis technique and by evaluating different functions including the linear, diagonal linear, quadratic, diagonal quadratic, and Mahalanobis. The trained classifier is applied on a set of testing images and the output is the binary mask with 1 for water pixels and 0 for non-water pixels.

The step following the flood mask creation is the flood mask correction that consists of the following three steps that aim at removing erroneous areas. The first step is a global filter that check the image as a whole and relies on the assumption that if the percentage of pixels classified as flood is very small, then probably these pixels are misclassified. The second step is a local filter that aims at eliminating small flooded areas (groups of a few pixels, usually up to 10 pixels which are classified as flooded in non-flooded areas), by applying connected-component analysis. Finally, the third aims at eliminating small non-flooded areas inside flooded area while preserving larger areas by applying image dilation and erosion.

Experiments have shown that the accuracy of such methods can be very high, even when the training set is not from the same source of EO data. The aformentioned classifier has been trained on satellite images, collected from PlanetLabs (Planet team, 2017), which have been provided in the context of MediaEval2017 - Multimedia Satellite Task[8]. The training dataset consists of 6000 pixels annotated as positive (water) & 18000 pixels negative (non-water) pixels. The method has been evaluated on a dataset including 1 LandSat7

[7] A list of projects funded under the same call is available at: https://cordis.europa.eu/search/result_en?q=contenttype=%27project%27%20AND%20/project/relations/associations/relatedCall/call/identifier=%27H2020-WATER-2014-two-stage%27

[8] http://www.multimediaeval.org/mediaeval2018

image and 4 LandSat8 images in view to derive the water surface extension of the Al Massira dam located in the Moroccan pilot (irrigated) area of the MOSES Project (Michail et al., 2018). Landsat 7 carries the Landsat Enhanced Thematic Mapper Plus (ETM+) sensor and produces images consisting of seven spectral bands with a spatial resolution of 30 meters for Bands 1-5, and 7, a panchromatic band with resolution 15 meters (Band 8) and a thermal band (Band 6). Landsat 8 carries the Landsat Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS) and produces images consisting of nine spectral bands with a spatial resolution of 30 meters for Bands 1 to 7 and 9, panchromatic band with resolution 15 meters (Band 8), and two thermal bands (bands 10 and 11). Table 1 provides an overview of the bands of the two satellites and the bands that are of our interest and used in our experiments are highlighted in gray.

Table 1. Bands of Landsat 7 and Landsat 8

	Bands										
	Band 1	**Band 2**	**Band 3**	**Band 4**	**Band 5**	**Band 6**	**Band 7**	**Band 8**	**Band 9**	**Band 10**	**Band 11**
Landsat 7	Blue	Green	Red	Near Infrared	Short-wave Infrared	Thermal	Short-wave Infrared	Panchromatic			
Landsat 8	Ultra Blue	Blue	Green	Red	Near Infrared	Short-wave Infrared	Short-wave Infrared	Panchromatic	Cirrus	Thermal Infrared	Thermal Infrared

The measure that was used for measuring the efficiency of the proposed method is the accuracy which was also used during the evaluation of the MediaEval2017 - Multimedia Satellite Task. The accuracy is defined as the ratio of the number of pixels recognized correctly to the total numbers of pixels. Table 2 contains the results of the evaluation of the method by considering: 1) different discriminant functions; 2) different point of evaluation (before and after post-processing); 3) different evaluation region either solely the dam and the surrounding area or the whole image, given that annotation is provided only for dam region.

Table 2. Evaluation of method using various discriminant analysis functions before and after post-processing.

		Linear	**Diagonal Linear**	**Quadratic**	**Diagonal Quadratic**	**Mahalanobis**
Full Image	Mask before post-processing	56.2676	45.8455	50.1789	34.8459	51.1556
	Mask after post-processing	56.1781	45.6994	51.8167	36.9312	53.0569
Dam Region	Mask before post-processing	73.4600	73.8431	87.4349	59.5132	87.7580
	Mask after post-processing	73.2535	73.3361	88.9611	62.2060	**89.0873**

According to the Table 2 we can deduce that in general, post-processing operations improve the mean accuracy. Another conclusion that can be produced is that the mean accuracy of the full image is rather low which is however explained by the fact that the annotation refers to dam area which results to lower accuracy. The final conclusion that can be derived refers to the best performing discriminant analysis method, which seems to be Mahalanobis, though the quadratic functions seems to perform rather well too.

3 EOPEN TECHNOLOGIES FOR FUSING NON-EO DATA

EOPEN project aims to fuse non-EO data in the analysis pipeline and in the following we list some technologies that will further improve the flood monitoring user experience. The EOPEN "Flood risk assessment and prevention" Use Case aims is to provide scalable solutions in given contexts; the Copernicus Emergency Management Service (EMS) is a benchmark for tailoring services in areas related to hazardous events like floods, e.g. the EFAS - The European Flood Awareness System[9]. EOPEN has taken into account such existing technologies, aiming to create additional added-value services, by integrating non-EO data and delivering high-level and real-time knowledge to its end users. The EOPEN technologies for knowledge

[9] http://emergency.copernicus.eu/mapping/ems/early-warning-systems-efas-and-effis

discovery and semantic reasoning are listed as follows:

Change detection: The EOPEN change detection module varies on the domain of the use case scenario. The tool monitors changes in streams of Copernicus data (e.g. EMS, C3S) using statistical and machine learning techniques to monitor changes in climate, floods and crops/weather.

Event detection: The event detection module from social media and online information is based on burst analysis to detect flood events from spatiotemporal information of social media streams.

Concept extraction: The concept extraction module of EOPEN detects concepts from Twitter images to identify whether an image refers to a flood or not, using computer vision techniques.

Similarity fusion: The similarity fusion module provides an advanced search engine that incorporates EO and non-EO data sources and metadata, when searching for past flood events in the database.

Clustering of EO and non-EO data: The parts of EO imagery (e.g. the generated thumbnails, or specific locations), as well as non-EO metadata (concepts, named entities, text), are clustered and representative images will be offered to the end user of EOPEN.

Community detection: EOPEN delivers an animation of the user comminities, using social network analysis to visualise communities and key-players with a particular focus to flood event monitoring.

EOPEN ontology and semantic reasoning support: The EOPEN semantic framework will be based on existing ontologies and models, which will be extended in order to represent all aspects pertinent to EOPEN.

To the best of our knowledge, these technologies are very useful for water authorities in their need to monitor a flood event and to have a holistic view of an area. The development and combination of these technologies, as well as their integration to one unique platform is the main EOPEN technological activity.

CONCLUSIONS AND FUTURE WORK

EOPEN aims to reduce emergency response time by developing an innovative Flood Decision Support System based on EO and non-EO data (meteorological, flood report by citizen, social media data). The current Flood Early Warning System shall be improved by assimilating EO data (snow cover, precipitation, soil moisture). Moreover a real time flood mapping will be implemented to identify sites in high risk flood zones. The presented technique shall be extended in the future by considering synthetic aperture radar (SAR) multimodal data fusion. Flood mapping using satellite images will be beneficial to urban and infrastructure planners, risk managers and disaster response or emergency services during extreme rainfall events.

ACKNOWLEDGEMENTS

This work is supported by the EC-funded projects EOPEN (H2020-776019) and MOSES (H2020-642258).

REFERENCES

G. Vingione, L. Marzell, E. Cadau., I. Gialampoukidis, S. Vrochidis, I. Kompatsiaris, B. Valentin, M. Melcott, L. Gale, W.-K. Lee, S. Woo, M. Gienger, D. Hoppe, C. Kontoes, I. Papoutsis, M. Ferri, F. Baruffi, J. Yoon, H. Yoon, A. Karppinen, A.-M. Harri (2017), The H2020-EO EOPEN project, ESA Conference on Big Data from Space 2017, Toulouse, France.

A. Moumtzidou, S. Andreadis, I. Gialampoukidis, A. Karakostas, S. Vrochidis and I. Kompatsiaris (2018), Flood relevance estimation from visual and textual content in social media streams, WWW 2018 Workshop on Exploitation of Social Media for Emergency Relief and Preparedness (SMERP), 23 April 2018, Lyon, France.

E. Michail, A. Moumtzidou, I. Gialampoukidis, K. Avgerinakis, M. G. Scarpino, S. Vrochidis, G. Vingione, I. Kompatsiaris, K. Labbassi, M. Menenti, F.-E. Elghandour (2018), Testing a Flood Mask Correction Method Of Optical Satellite Imagery Over Irrigated Agricultural Areas, 2nd Mapping Water Bodies from Space Conference (MWBS2018), 27-28 March 2018, ESA-ESRIN Frascati, Italy.

K. Avgerinakis, A. Moumtzidou, S. Andreadis, E. Michail, I. Gialampoukidis, S. Vrochidis, I. Kompatsiaris (2017), Visual and textual analysis of social media and satellite images for flood detection, Multimedia Satellite task of MediaEval 2017.

Planet team (2017), Planet Application Program Interface: In Space for Life on Earth, San Fransisco, CA