Article, Published Version

Popescu, Ioana; Jonoski, Andreja

Citizen's observatories: contributions to the improvement of flood modelling for management

Hydrolink

Citizen's observatories: contributions to the improvement of flood modelling for management

By Ioana Popescu and Andreja Jonoski

Crowdsourcing of environmental data has recently been pro-posed as a possible alternative to augment and enrich available datasets for managing environmental systems. In water-related studies, it could supplement the data available from existing monitoring networks. This article presents experiences from research in collecting and processing crowdsourced data for use in flood modelling studies. The work has been carried out in a past European research project (of the H2020 Research Programme), where data collected by citizens were used to support the development, calibration and validation of hydrodynamic models used for flood analysis. The data were gathered by a dedicated game-like mobile phone app in the form of images and videos that were later post-processed to provide data on e.g. land use/land cover, river geometry, water levels and water velocities.

In water related problems models are important tools that support decision-makers for management in the their water resources. One of the important topics studied through modelling are the river floods and their related risks. Due to the impact that floods have on the environment, either in a positive or negative way, various types of models are used in order to understand and predict their behaviour. For adequate representation of floods in models, data is a critical requirement, as the quality of the developed models depends on data quality[6]. Continuous efforts are being paid for acquiring data, however still many rivers lack the amount and quality of data and existing data access is not always possible[4]. Limitations in data acquisition are due to several reasons: difficulty to reach the location, high cost of complex instrumentation, lack of technical knowledge of the data collection stations, etc. Alternative technologies, such as remote sensing and/or drones can provide complementary data, however these technologies are not available everywhere and they may be expensive. In this context, data collected through crowdsourcing can be relevant for modelling, if robust mechanisms for ensuring data quality are used.

De Sherbinin *et al.*[3] emphasize the critical importance of data as an output of citizen science projects, and their contribution to the collection of scientifically relevant information, due to the advantages of being low-cost and abundant. For flood related projects data is most often collected through citizen observatories (COs), in which participants together with scientists are involved in all stages of a project, from research design to processing of data collection and/or interpretation[2]. This approach, as part of the citizen science field, has often been used in the past decade in diverse projects for different purposes related to floods analysis and management. COs have been explored in the field of flood modelling and management as a monitoring tool, for data assimilation in modelling, and for mapping floods[1]. When used in modelling, all collected data through CO need to be benchmarked and integrated into models, for model set up, calibration, validation, simulation and potentially for forecasting.

In this article we are presenting the experiences in collecting data for flood modelling, through CO, for the particular case of a wetland in the Danube delta in Romania. The work was carried out as part of SCENT research project.

The Sontea Fortuna wetland and CO data collection

The SCENT project was one of the research projects funded by European Union under Horizon 2020 programme, exploring how citizens can be engaged in data collection such that they become the 'eyes' of the policy makers. The project developed a gamified smartphone application dedicated to collect data on land use, water levels and velocities in the form of photos and videos. The main research question was whether such collected data can contribute to the improvement of a flood model. One of the case studies where data was collected is the rural area of the Sontea-Fortuna wetland in Romania. Located in the upstream central part of the Danube Delta, the Sontea-Fortuna wetland is an important lacustrine complex for maintaining the good ecological status of the delta, both from hydrodynamical and morphological point of view (**Fig. 1**). The wetland covers a total surface area of 246.36 km^2. Main canals with total length of 106 km convey water to an inside secondary network of canals, which are 153 km long in total[5]. The canals in the wetland area are interconnected with 11 major lakes. The wetland is only accessible by boats.

Currently the Sontea-Fortuna wetland is well-preserved, however tendencies of increased discharge due to climate change, decreased sediment supply due to upstream interventions (dams and reservoirs) and modified eutrophication rates threaten the maintenance of the ecosystem. Thus, a river management that sustains a proper monitoring network, with real-time capabilities, together with modelling of the flooding patterns, is a necessity.

Flow properties in the studied area, depth and velocity (or depth and discharge), change with time, hence a hydrodynamic model of the Sontea-Fortuna wetland was build using the open source modelling suite developed by the Hydrologic Engineering

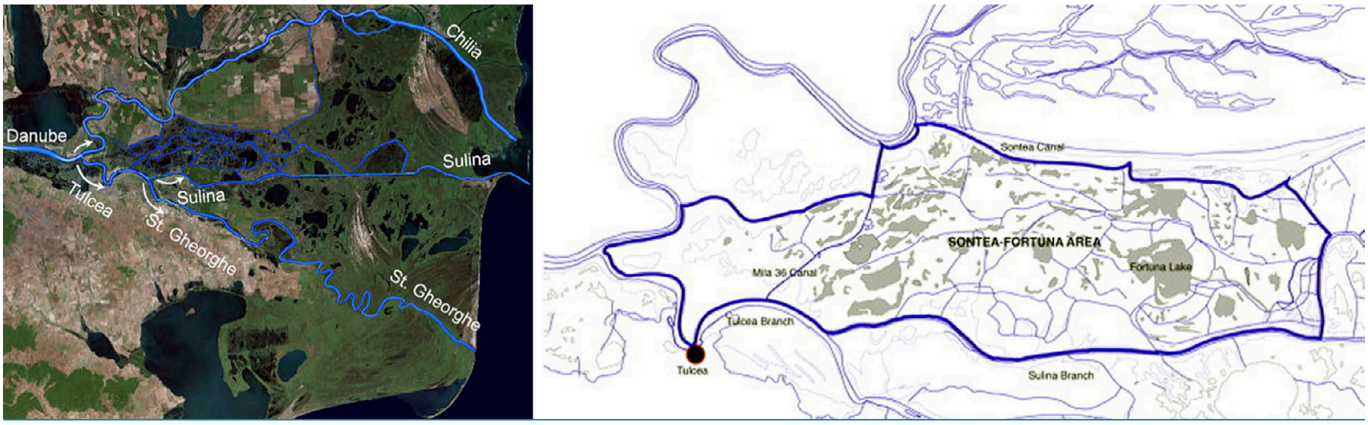

Figure 1 | Main branches of the Danube River (Left) and the Sontea-Fortuna area (Source: Popescu *et al.*, 2015) (Right).

Center (HEC) of the United States Army Corps of Engineers (USACE), the River Analysis System (HEC-RAS). The 1D/2D HEC-RAS model of the Sontea Fortuna area (**Figure 2**) calculates the value of depth and discharge in a spatial grid of computational points, defined by the modeller, over a simulation period of interest. The model build using traditionally available data (obtained from the Danube Delta National Institute) was tested for improvement by adding data collected through COs.

Data collection in the Sontea-Fortuna area entails defining boat routes to follow, when CO campaigns are organised. The approach for defining pathways for data collection, though not specific to the Danube Delta case study, was developed during the SCENT research and its applicability was tested in the Sontea-Fortuna area. The principle of determining the data collection routes take into account the interest of local stakeholders and the characteristics of the study area, such as navigability of canals in terms of minimum water level, boat velocity and maximum available time for a route completion. The generated possible pathways were given scores based on a set of criteria, allowing for their prioritisation and choice. **Figure 3** presents several determined routes, along with the boat position at the location of the measurement. Several possible pathways were determined and followed during data collection campaigns. However, some of these pathways were changed during the data collection due to the need of adjusting the pathway parameters, as it is, for example, the time needed to make an observation. Collected images for measuring water depth and videos (of floating objects for measuring flow velocity, not discussed here), with the phone, contained metadata on location coordinates, date and time. These were pre-processed and made available in the SCENT database. In order to use the crowdsourced data in the flood model of the area, the quality of photos of water depth were first visually analysed and the images with no gauge, very low resolution and no water surface were discarded. **Figure 4** shows images that are bad or good for model usage. The water depth values from filtered good images were manually interpreted.

In order to check the validity of the collected data, during the campaigns, measured data were collected through an Acoustic Doppler Current Profiler (ADCP) instrument, in the same locations where citizens were collecting data. The ADCP data had a series of measurement points along cross-sections of canals.

Modelling results

Three datasets were available to determine the quality of collected data after the processing step: crowdsourced data, measured data and model results. The water depth data of the three datasets were qualitatively compared and classified as good, average and bad matches. It was found that there was a shift in the location of model cross-sections with respect to the measured and crowdsourced data. The measured depths by citizens were aligned to the model cross-sections for analysis and visual comparison with the ADCP data, classifying them as good or bad, depending on how well they matched. The ones

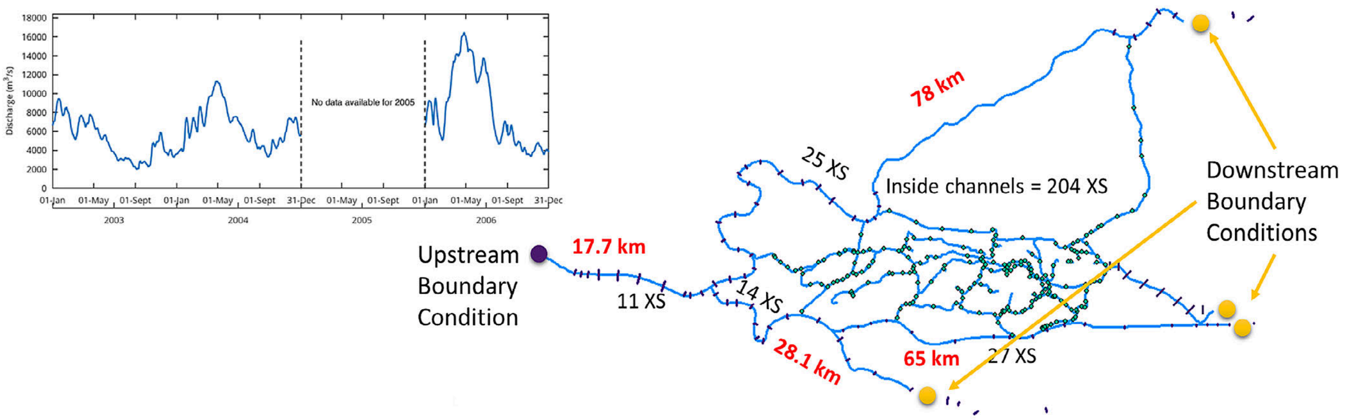

Figure 2 | Schematic of the 1d/2D RAS model of the Sontea-Fortuna area.

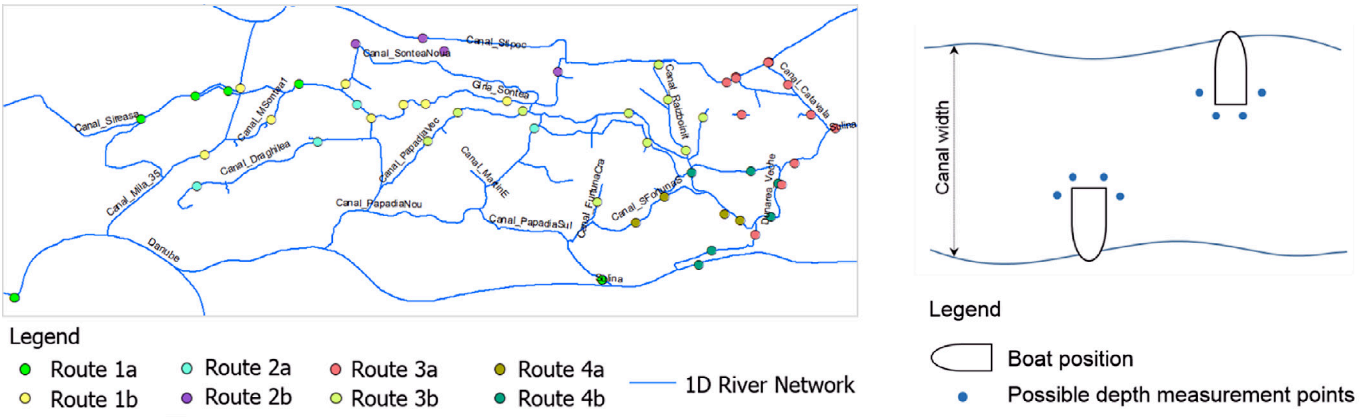

Figure 3 | Example data collection pathways to be followed by boats (Left) and boat position and data collection points (Right).

classified as bad were eliminated. There were two factors which led in crowdsourced data points to be discarded. One was the location of crowdsourced data itself, and the other was the representation of the terrain in the HEC-RAS model. The crowdsourced data locations were derived the GPS of the mobile phones. Sometimes the signal quality was weak resulting in locations potentially to be less accurate. The second factor was that eve though there were locations where, the crowdsourced data had proper spatial position, the spatial locations of crowdsourced data did not match the inundations shown in the model results (**Figure 5**).

In order to understand better how well the model is performing, without the uncertainty of the crowdsourced data, the maximum water depths in a cross-sectional profile and maximum discharge on the day of measurement were compared with the model results at all locations. It was found that the maximum depths in the 1D canals were approximately equal to the measured data. In 2D flow areas, they were close, but not as much as in 1D (**Figure 6**).

The comparisons of the model profiles with the crowdsourced and measured data indicated better matches in the 1D canal networks than the 2D flow areas, because the model was calibrated mostly for 1D networks. The good correlation of crowdsourced water depths with respect to the measured water depths indicate that the crowdsourced data represents valid information and may be used in modelling.

Figure 4 | Collected water levels of bad quality (upper row) and good quality (lower row).

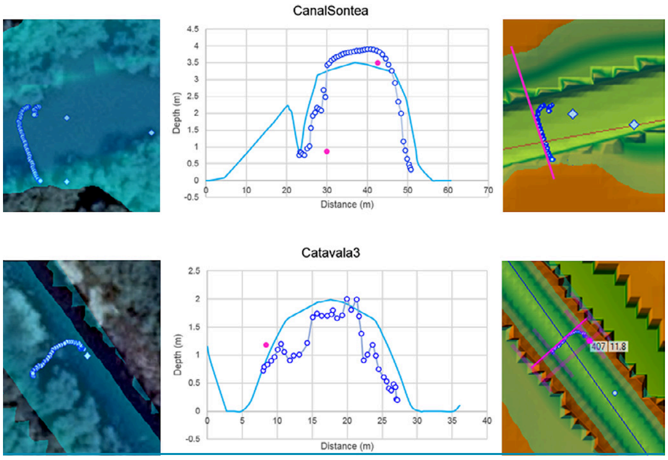

Figure 5 | Example of bad located crowdsourced data as compared with model inundation results.

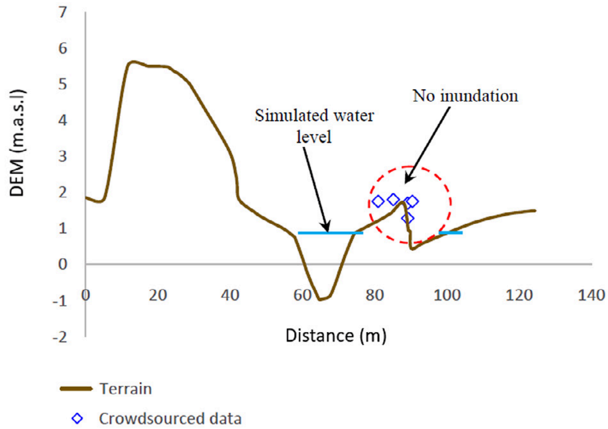

Figure 6 | Water depth comparison between measured, model simulated data and ADCP collected data, on two different canals inside Sontea Fortuna area.

Conclusions

The collected citizen data in SCENT proved that such data is effective as there is enough amount and well distributed data available from crowdsourcing. The comparative analysis carried out on different data sources indicated the value of crowdsourced data, complementing the limited data available from classical sources in the modelling context. Possible usages of crowdsourced data in the modelling of floods are validation datasets, when calibrating the model with the measured data, or vice-versa. Tuning the model using both measured and crowdsourced datasets in combination does improve its performance, and also shows the contribution of crowdsourced data as valid dataset in the modelling process.

Despite being challenging, the implementation of citizen observatories in the Sontea-Fortuna area has demonstrated its usefulness and that it can contribute to the improvement of many aspects of environmental area is better protected and managed. These initial results are promising, however many challenges remain open in crowdsourcing approaches, regarding gathering sufficient amount of data at the right time and on the right locations, as well as the sustainability of applying these methods for collecting data after the main project funding stops. There is a need for better involvement of the citizens in these actions, by having the authorities investing in actions and in continuous organisation of campaigns, in projects that address the concerns of the local communities engaged in COs.

Ioana Popescu
Ioana Popescu is an Associate Professor in Hydroinformatics at IHE Delft Institute for Water Education. She worked in the field of flood modelling over the past 25 years. She has participated in a number of flood related projects in Vietnam, China, Romania and several EU funded projects related to hydraulic and hydrological modelling.

Andreja Jonoski
Andreja Jonoski is an Associate professor in Hydroinformatics at IHE Delft Institute for Water Education. His man research interests are hydrological and hydraulic modelling, groundwater modelling, coupling of simulation models with optimisation models, participatory decision support systems. His past experience is in numerous research projects applying modelling of water systems for decision support.

References

1 | Assumpção, T.H., Popescu, I., Jonoski, A., Solomatine, D.P., 2018. Citizen observations contributing to flood modelling /: opportunities and challenges 1473 1489.

2 | Buytaert, W., Zulkafli, Z., Grainger, S., Acosta, L., Alemie, T.C., Bastiaensen, J., De Bièvre, B., Bhusal, J., Clark, J., Dewulf, A., Foggin, M., Hannah, D.M., Hergarten, C., Isaeva, A., Karpouzoglou, T., Pandeya, B., Paudel, D., Sharma, K., Steenhuis, T., Tilahun, S., Van Hecken, G., Zhumanova, M., 2014. Citizen science in hydrology and water resources: opportunities for knowledge generation, ecosystem service management, and sustainable development. Front. Earth Sci. 2, 1 21. https://doi.org/10.3389/feart.2014.00026.

3 | de Sherbinin, A., Bowser, A., Chuang, T.R., Cooper, C., Danielsen, F., Edmunds, R., Elias, P., Faustman, E., Hultquist, C., Mondardini, R., Popescu, I., Shonowo, A. and Sivakumar, K. (2021) The Critical Importance of Citizen Science Data. Front. Clim. 3:650760. doi: 10.3389/fclim.2021.650760

4 | Hannah, D.M., Demuth, S., van Lanen, H.A.J., Looser, U., Prudhomme, C., Rees, G., Stahl, K., Tallaksen, L.M., 2011. Large-scale river flow archives: Importance, current status and future needs. Hydrol. Process. 25, 1191 1200. https://doi.org/10.1002/hyp.7794

5 | Popescu, I., Cioaca, E., Pan, Q., Jonoski, A., Hanganu, J., 2015. Use of hydrodynamic models for the management of the Danube Delta wetlands: The case study of Sontea-Fortuna ecosystem. Environ. Sci. Policy 46, 48 56. https://doi.org/10.1016/j.envsci.2014.01.012

6 | Schumann, G.J.P., Bates, P.D., Neal, J.C., Andreadis, K.M., 2015. Measuring and Mapping Flood Processes, Hydro-Meteorological Hazards, Risks, and Disasters. Elsevier Inc. https://doi.org/10.1016/B978-0-12-394846-5.00002-3