Flood hazard assessment and the role of citizen science

\textsuperscript{1}Bocar Sy, \textsuperscript{1}Corine Frischknecht, \textsuperscript{2,3}Hy Dao, \textsuperscript{1,4}David Consuegra, \textsuperscript{3}Gregory Giuliani

\textsuperscript{1} Department of Earth Sciences, Faculty of Science, University of Geneva
\textsuperscript{2} Department of Geography and Environment, Geneva school of Social Sciences, University of Geneva
\textsuperscript{3} Institute for Environmental Sciences, University of Geneva
\textsuperscript{4} Institute of Territorial Engineering, School of Management and Engineering Vaud

Correspondence to: Bocar Sy (Bocar.Sy@unige.ch)

Abstract

Flood hazard assessment is at the core of flood risk management. In order to develop an efficient flood hazard assessment, it is of primary importance to have a well-defined flood scenario encompassing all processes that could occur during an event. Understanding and assessing these processes requires meteorological, topographical and land-use data as well as historical observations. Nowadays, flood delineation is based upon hydrological and hydraulic modelling, ground data collection and remote sensing. Despite the advantages of these tools, they also present some specific limitations, either intrinsic to the approaches or linked with constraints of the local context.

With the rapid advancement of web 2.0 technologies (e.g. Flickr, Wikimapia) and the increase of the use of participatory research, citizen science has the potential to provide valuable and complementary information at all levels of flood risk management and in particular for flood hazard assessment.

After reviewing the capabilities and limitations of the current tools used in flood hazard assessment, this paper demonstrates the role that citizen science can play in providing key information on factors leading to flooding and on flood hazard parameters.

Keywords: flood hazard assessment, citizen science

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/jfr3.12519
Introduction

Natural phenomena of geological, hydrological or meteorological related origins cause substantial loss of lives and economic impacts (UNISDR, 2015a). Between 1997 and 2016, the Emergency Events Database (EM-DAT), an international disaster database maintained by the Centre for Research on the Epidemiology of Disasters (CRED), recorded 5,415 natural disasters worldwide, which claimed 1.02 million lives or almost 51,000 lives on average each year (EM-DAT 2016). The EM-DAT criteria for a disaster are: ten or more people killed, hundred or more people reported affected, a declaration of a state of emergency, a call for international assistance. In terms of floods, the EM-DAT (2016) recorded, for the same period, 2,471 flood disasters, which killed 102,665 people, affected 103 million people and caused more than 470 million USD of economic losses in the world.

Floods are the prevalent type of disasters across the world (Wang et al 2015). According to UNISDR-CRED (2015), floods affect Asia and Africa more than other parts in the world. Different mechanisms are responsible for floods, which, in some cases, can also be influenced by human activities. The major types are: river floods, flash floods, storm surges or groundwater flooding, the latter caused by the rise of water table above surface (Kron 2005; Hughes 2011; Macdonald et al 2012, Von Storch et al 2014; Gascón 2016).

Consequently, there is a need to strengthen capacities in Disaster Risk Reduction (DRR) and Climate Change Adaptation (CCA). In 2015, 187 countries adopted the Sendai Framework for disaster risk reduction for the 2015-2030 period (UNISDR, 2015b), with seven objectives, including understanding disaster risk and improving disaster risk governance.

One key activity that governments through dedicated institutions must carry out in order to reduce the flood effects is to assess and map flood hazards. Both aspects are closely interrelated. Flood hazard assessment predicts flood magnitude and intensities of the various processes involved by analysing flood scenarios including past floods, while flood hazard mapping aims at spatially visualising a potential endangered area following a given flood scenario. Flood hazard assessment and mapping have been carried out in some countries (e.g. USA, Switzerland) for more than 40 years (Burby 2001; Kunz and Hurni 2008). Flood hazard assessment has been mainly used to identify areas at risk and as a tool for land-use regulation and implementation of mitigation strategies (De Moel et al 2009).

Many researchers have explored flood hazard assessment (Kumar 2016; Xiao et al 2017). Although there are different types of assessments, methodologies and purposes, the framework behind them is quite standard, from defining a scenario, acquiring specific basic data and obtaining key flood hazard parameters as sketched in figure 1.

Figure 1. Current framework for flood hazard assessment

Defining scenarios, i.e. understanding the chain of events that can lead to flooding, is not an easy task. Ground measurements, remote sensing analysis and/or hydrological modelling (in this paper hydrological modelling includes both hydrological and hydraulic modelling) are the current tools applied in flood hazard assessment. As such, these classical tools present some specific constraints.

This article is protected by copyright. All rights reserved.
related to their intrinsic characteristics or the context of use. For ground data acquired by instruments, spatial and temporal coverage is a challenge, both in terms of maintenance of a representative network and data processing costs (Houghton-Carr 2014). Remote sensing can compensate for some of these aspects. However other issues, such as weather conditions influencing the quality of satellite images, potential lack of images covering the areas of interest or the processing algorithms used in data mining can impede the spatial coverage of areas of interest (Pulvirenti et al 2011a; Triglav-Čekada and Radovan, 2013). As for hydrological modelling, it is very sensitive to the quality and quantity of input data (Townsend and Walsh 1998) influencing the accuracy of the result.

In an attempt to overcome these shortcomings, several attempts have been made in recent years to apply citizen science in the field of hydrology (e.g. Buytaert et al 2014, Paul et al 2018) or flood assessment, either during flood response or in near real-time mapping of flooding events (Cheong and Cheong 2011; Wan et al 2014; Eilander et al 2016). As defined by Bonney et al (2016), citizen science is a form of collaborative research involving the public into scientific projects. Participatory approaches to research and governance are not new. There is a long tradition of using volunteers to provide information relevant to science, such as Bird count in the United States (Mueller et al 2012), which has been organized annually since 1900 by the Audubon society (Audubon 2017). What is new and different is the widespread diffusion of Internet-based networks and social media. This innovation has been rendered possible thanks to broadband internet, web 2.0 (defined by O'Reilly (2005) as “an architecture of participation” using technologies for data management, display, exchange and interconnection of multiple users), the development of mobile communication, the progress in global positioning in particular with portable GPS receivers, which are inexpensive and often implemented with camera in smartphones or tablets together with other useful sensors (Paul et al 2018). Therefore, data shared by citizens through social media (Twitter, Flickr, and Instagram) has become a useful source of information. Consequently, nowadays citizen scientist engagement can take different ways, depending on the level of implication, from simple data acquisition to active participation, and on the geographical scope as discussed by Haklay (2013) and Paul et al (2018). Various tools are used to involve citizens, which can be classified into two main categories, direct and online. The first category implies observation or mapping activities through active survey (Sy et al 2016), participatory mapping (P-mapping) or participatory Geographic Information System (P-GIS) (Kienberger 2014; Chingombe et al 2015; De Andrade and Szlafsztein 2015). The second category involves citizen’s participation via the use of Web platforms (Dorn et al 2014; See et al 2015) or social networks. It can take the form of crowdsourced data, i.e. online data provided by many volunteers (See et al 2016) or spatial information provided on a voluntary basis and defined as volunteered geographic information (VGI, See et al 2016).

The objective of this paper is to review current practices in flood hazard assessment and provide initial insights on the potential use of citizen science in improving current practices.

This paper is composed of three sections. The first part reviews flood hazard assessment in terms of flood hazard assessment, key parameters and current tools used, their capabilities and limitations. The second section will discuss the potential of citizen science in scenario description, in basic data acquisition and output results. Finally, a brief conclusion will close this paper.
Flood hazard assessment

Flood hazard assessment is a complex task. Current flood hazard assessment uses data coming from fieldwork or ground survey, hydrological modelling and remote sensing. These approaches are employed for scenario description, input data production such as rainfall data, discharge, topographic information or land-use typology, and for output results including water depth, flood extent, flow velocity, inundation duration or a combination of flow velocity and depth of water (De Moel et al 2009; Prinos 2009). Flood hazard assessment can be performed through deterministic or probabilistic approaches (e.g. Di Baldassarre et al 2010). Results are then expressed as a flood probability (or frequency) with a given magnitude or intensity over a specific period of time (e.g. Apel et al 2008; Pappenberger et al 2012; Dottori et al 2016). The reliability of the results depends largely on the data availability and quality, the scale of the analysis and the full understanding of the limits of the approaches used. Consequently, each technique has context-dependent advantages and drawbacks, which are described in table 1.

Table 1. Main strengths and weaknesses of current approaches on flood hazard assessment

Flood scenarios

A proper flood hazard assessment requires defining a flooding scenario. This scenario specifies the set of processes whose combinations trigger flooding (DAEC 2016). It includes three types of processes: basic, associated and aggravating. A basic process (like rainfall) is the main vector of flooding quantified by an intensity and a probability of occurrence. An associated process is a process that accompanies the flood process, like bank overtopping, solid transport or bank erosion for a river, but it does not change the probability of occurrence. The aggravating process refers to incidents that could worsen the flood, like dike failure, jams at bridge crossings or obstructions made by landslides. In some cases, the presence of bridges could be misinterpreted as flood barriers, when in fact they did not stop the water. Since the correlation between aggravating and basic processes is not evident (Consuegra, personal communication), one may have to consider two different scenarios, with different results in terms of flood hazard mapping. Defining a flooding scenario requires an in-depth and accurate knowledge of the area under investigation.

Basic data acquisition

Basic data used as input data are hydrological, topographical and land-use information (WMO 2013). First, the hydrological data used as key input for hydrological modelling are rainfall data and discharge. Furthermore, discharge, velocity and water level recorded in-situ are important because they represent the initial or the boundary condition and they are used for model calibration and validation (Büchele et al 2006). Second, the topographic information is the key information in many hydrological models (Santillan et al 2016; Werren et al 2016), allowing the watershed delineation and the extraction of topography-related parameters such as slope and flow directions (Tarboton et al 2009). Third, the land-use is used as input for hydrological modelling and influences significantly the runoff (Knebl et al 2005). For example, urban areas are characterized by significant imperviousness ratios, thus increasing surface flow and reducing infiltration.
These categories of data are usually obtained from in-situ measurements and remote sensing. In-situ measurements comprise a network of measurement stations and ground survey techniques. Nowadays, various types of instrumental measurements (e.g. rainfall station, metric flow) exist to provide hydrological data. These data are important for characterizing the local site conditions and are used as input parameters in hydrology modelling. However many regions suffer from a declining coverage of hydro-meteorological stations due to financial constraints, reduced capability of acquiring and managing data therefore interrupting long-term observations in some countries (e.g. USA, Canada) (Lins 2008, Mishra and Coulibaly 2009, Hannah et al 2011), which leads to a lack of input data as well as an inadequate hydrological model calibration and validation. To address these issues, discharge or rainfall data can be extrapolated to ungauged areas using regionalisation techniques (Merz and Blöschl 2005). However, these techniques can underestimate or overestimate measured values at local level or induce random errors in the results (Merz and Blöschl 2005).

Ground survey techniques are applied to capture topographic or land-use information at very high resolution. However, as they are time consuming and expensive to put in place, they are only used for small and limited areas (Bates 2004).

As a consequence of these limitations for in-situ and ground survey measurements, remote sensing has become an important tool for basic data production including rainfall data, topographic information and land-use necessary to build flood model and to assess flood hazard. Remote sensing techniques based on precipitation estimates, such as the tropical rainfall measuring mission (TRMM, Huffman et al 2007) and its successor, the global precipitation measurement mission (GPM, Hou et al, 2014) can provide rainfall data with quasi-global coverage (Hong et al 2007, Skofronick-Jackson et al 2017). Such data might be the best source of rainfall data in countries, where there is a lack of dense ground based networks. It provides the most precise available rainfall representation in terms of temporal resolution. Several studies have demonstrated the utility of remote sensing rainfall estimate for hydrological modelling (Hong et al 2007; Nikolopoulos et al 2010). However, satellite precipitation products are associated with random errors that nonlinearily propagate in hydrological modelling (Nikolopoulos 2010) and are inherent on the scale of application and the satellite product resolutions (Nikolopoulos et al 2013).

Remote sensing is also used for topographic information purposes (Marks and Bates 2000; Smith 2002). It can handle large areas in a relatively short amount of time. Actually, the airborne scanning laser altimetry Light Detection and Ranging (LIDAR) has been widely used to acquire data for flood modelling due to its high resolution description of the landscape (Turner et al 2013). However, the cost of data acquisition could be exhibitory, especially in countries lacking of financial resources (Klemas 2014; Yan et al 2015).

Remote sensing particularly based on aerial photographs and high-resolution satellite imagery has been used increasingly for the land-use typology. However, the weather conditions, the lack of images covering the areas of interest and the non-actualization of images in very dynamic settings still represent obstacles.

**Key flood hazard parameters**
Flood hazard assessment aims at characterizing key parameters such as flood extent, water depth, flow velocity, and flood duration. These parameters are of prime concerns for understanding potential impacts on the society. Depending on the final objective of flood hazard assessment, crisis management or risk management, other parameters could be used, including the propagation of the water, water rise, and arrival time of the water to a certain location. As for basic data, key flood hazard parameters can be derived from in-situ measurements, hydrological modelling and remote sensing.

Field measurements are the first line of output data, which can be acquired by instruments or by field experts using flood marks. The main limitation of these field data is the insufficient spatial and temporal coverage, due to a limited number of observations across the area and the lack of repetition of the data acquisition over time (Townsend and Walsh 1998) as well as the financial cost for maintenance and data processing (Houghton-Carr 2014). The situation is worse in developing countries where other factors, such as political and institutional instabilities occurring at national or regional level (Lins, 2008) or the dependency on international aid (Hannah et al 2011) impede the installation and the maintenance of monitoring instruments (see, e.g. Nkwunonwo 2015 and Komi et al 2017).

As ground data are either scarce (e.g. in many countries of Africa, Komi et al 2017), or declining (e.g. in the pan-Arctic region, Lins 2008), and therefore time series incomplete, hydrological modelling is a useful tool to understand main processes involved in flood scenarios and to obtain outputs for flood hazard mapping such as the water depth, the flood extent, the flow velocity and the duration of inundation (Néelz and Pender 2010; Néelz and Pender 2013; Santillan et al 2016). With computational technology progress, modelling can be achieved in 1D (e.g. Brunner 2016), 2D (e.g. Moulinec et al 2011) or 3D (e.g. Prakash et al 2014). However, the more dimensions are involved, the more is needed in terms of quantity, quality of data and computational cost (Hunter et al, 2007).

Therefore, rapid and accurate flood modelling at high spatio-temporal resolutions remains a significant challenge. Accuracy is related to basic data acquisition limitations, the complex and chaotic nature of flooding and uncertainty inherent in hydrological modelling (Freer et al 2011).

Remote sensing, satellite or aircraft-based, has been widely used for flood extent assessment. Since the early satellite image products (e.g Hallberg et al 1973; Currey 1977), remote sensing provides information for areas of poor accessibility or lack of field visits (Smith 1997) and covers large areas. It helps to overcome the lack of spatially distributed data as well as calibrating or validating flood extents obtained by hydrological modelling (Horritt et al 2000; Mason et al 2016). However, there are some limitations, depending on the type of source (e.g. active or passive) and the location on Earth. Issues can raise from incomplete or low spatio-temporal coverage, alteration of the quality of scene or even its cost, especially for recent high-resolution images.

Passive sensors are sensitive to weather conditions and sunlight (Brivio et al 2002). Consequently, cloud cover reduces visibility, because visible and infrared wavelengths (0.4 µm to 1000 µm) do not cross the clouds (Martí-Cardona et al 2010). Similarly, optical images are incapable of detecting flooded areas under vegetation canopies (Hess et al 2003; Wilson et al 2007), which often results in a partial loss of information (Sanyal and Lu 2004) and therefore in an incomplete flooded area evaluation. On the other hand, radar sensors can capture flood extent in cloudy conditions and in both
day and night. However, in urban areas the viewing angle causes difficulties in separating water from other urban features. In addition, Pulvirenti et al (2011b) mentioned that one of the limitations in the use of radar images for flood extent mapping was their low temporal resolution, as most of radar sensors have a repeat orbit cycle of more than one month. With the development of techniques, there are now new sensors with high resolution and improved repeat overpass time in the range 3-4 days (e.g. TerraSAR-X on Europe). However, validation on ground of results obtained with satellite-based results (Pulvirenti et al, 2011b) remains a challenge.

The use of Unmanned Aerial Vehicles (UAV), such as drones, has been attempted in flood extent mapping (Hydro-GIS 2017) and in flow velocity measurement (Tauro et al. 2016). Such vehicles are able to acquire high-resolution data under cloud cover with a cost-effective accessibility to areas that cannot be reached though optical, radar and other aerial images. However, windy conditions, logistics and authorizations required to deploy the drone and the battery life endurance as well as visual line of sight (500 m) can impede data acquisition.

Citizen science and flood hazard assessment

In an ideal world, a flood hazard assessment would be achieved by using a representative dataset, spatially and temporally sound and regularly updated, but as discussed and summarized in table 1, current approaches have some limitations in that respect and many areas in the world are facing one or more of these issues. Therefore, alternatives for flood hazard assessment have been developed. With the rise of mobile devices, web 2.0 technologies (Batita et al 2012) and the increased use of participatory research (Cornwall and Jewkes 1995), citizen science can provide valuable and complementary information in the field of flood hazard assessment.

This section will provide some insights on how citizen science has started contributing in this domain as summarized in Table 2. This table classifies existing works based on their contribution towards basic data acquisition necessary for flood hazard assessment and the resulting key flood hazard parameters as well as defining which citizen science technique has been used.

Table 2. Overview of potential contribution of citizen science in flood hazard assessment

Potential role in flood scenario

Identifying processes that can lead to the occurrence of damage during floods is usually based on field visits in the study area. However, field visits are, in most cases, achieved by specialists coming from outside the concerned zone and some shortcomings can occur, both technical and temporal and spatial knowledge of the area influencing the understanding of the situation. Integrating the on-site knowledge of citizens living in the surroundings can be key to assess the potential floods or understand past events (Tran et al 2009) better. On the one hand, they have a better knowledge about their environment and current issues such as dike maintenance, jams, bridge crossing, which might aggravate the phenomenon. On the other hand, they are eyewitnesses during events, and therefore a valuable source of information on chain of events, which can be retrieved either by calling their explicit memory through interviews, focus groups or participatory mapping or through their photos and high quality video recordings of flood events, which are in most cases with geolocation and time.
Potential contribution in basic data acquisition

The acquisition of basic data for flood hazard assessment faces problems related to spatial and temporal coverage, precision, and update as well as cost for instrumental network installation and maintenance.

Involvement of citizens can help to reduce these issues. For example when dealing with rainfall data, rain gauges can be installed using low cost material (Starkey 2017) and citizens can collect, manage and share measurements across website (Cifelli et al 2005) or social media in order to fill spatial and temporal hydrological network data gap. Social media also can be used to monitor rainfall in real-time. De Andrade et al (2017) for example collected Twitter messages combined with official rainfall data to detect rainfall patterns in real-time. Other promising crowdsourcing approaches for assessing areal rainfall pattern use optical sensors or wipers of moving cars (Rabiei et al 2013, 2016) or the attenuation of radio signals by rain between transmitters and receivers of cellular communication networks (Overeem et al 2013). With regard to discharge, as it is often missed or poorly recorded by ground-based gauges, citizen science can be a valuable option to obtain discharge data that can then be used for model calibration. Indeed, the rise of sharing various types of information on extreme events like floods through social media, in particular videos, provide raw data that can be post-processed using image analysis techniques like Large Scale Particle Image Velocimetry to derive stream discharges as explained by Le Coz et al (2016) and Le Boursicaud et al (2016).

It has been pointed out that flood modelling is highly sensitive to the precision of input data, in particular the topography and land-use. Involving citizens can help improving and updating data, at local and global scale. For example, Shaad et al (2016) developed a digital terrain model using a low-cost alternative to LIDAR remote sensing images. They combined data from unmanned aerial vehicles (UAV) and videos from volunteers. Videos were shot from sites inaccessible by UAVs. Other projects have been using pictures taken by citizens across the world to validate global land cover. For example, Iwao et al (2006) have launched an initiative, where contributors document their surroundings by posting photos and comments in order to improve global land-use map; whereas Dong et al (2012) used geo-referenced field photos to validate existing land-use derived from high-resolution satellite images. Still with respect to land-use, citizen science can help to follow up with the territorial dynamics, especially in cities where the urbanisation rate is very high. In that respect, a VGI initiative, such as OpenStreetMap, is a very promising free source of geo-referenced data over a territory, which can be very useful to verify, complement and update land-use maps obtained from authoritative data or remote sensing techniques. For example, these data were associated with other land-use data sources (LIDAR, remote sensing data) and used as input for hydrological modelling (Dorn et al 2014). At local scale, other methods have also been applied, such as P-mapping in order to complement and update land-use map derived from satellite images as presented in Sy et al (2016).

Potential input for key flood parameters assessment

Depending on the size of the flood, it is often challenging to survey all areas being flooded extensively; therefore citizen science can help identifying flooded areas, before or during events. Several research projects have already involved citizens in that direction. P-mapping and P-GIS have been used in order to map potential or flooded areas (e.g. Peters Guarin 2008; De Andrade and
With regard to crowdsourcing, as social media data can be messages or photos, and can include geolocated information, several authors (see Schnebele and Cervone 2013; Cervone et al 2016; Li et al 2017; Rosser 2017; Smith et al 2015) have demonstrated the usefulness of integrating these data for flood extent assessment, in particular during an event.

Water depth is a crucial parameter to characterize flood hazard, however it is not always surveyed or easily monitored. Therefore, several approaches involving citizens have been attempted to retrieve these data from past events. At local scale, some studies (Singh 2014; Sy et al 2016) have been trying to retrieve this information for past events through P-mapping and P-GIS. In both cases, citizens were asked to indicate the water depth affecting their house during historical flood events by localizing this information. Another approach to retrieve historical data in data-scarce settings mentioned by Michelsen et al (2016) is the use of YouTube videos as the water level could be deduced from objects whose sizes are known or approximately known. Regarding acquisition of water depth during the occurrence of floods, the rise of web 2.0 technologies helps to implement various ways of obtaining data. Some projects focus for example on collecting water level information from citizens by asking them to send information through mobile applications or websites (e.g. Degrossi et al 2014), whereas others (Fohringer et al 2015; Kutija et al 2014) use text messages or photos from social media to derive this information.

Flood extent and water depth are the easiest parameters to obtain, whereas flood velocity is neither easily measured during events, nor well retrieved by modelling (Consuegra, personal communication). Approaches analysing post-event videos posted on social media as done by Le Boursicaud et al (2016) are a very interesting way to obtain in-situ measurements, where official instruments failed or were non-existent.

As showed, various citizen science approaches have already been investigated in order to collect and share data useful for flood hazard assessment. A step forward would be to take advantage of the second generation of environmental virtual observatories (EVO), which are not only virtual platforms for scientists, but aiming at enhancing participation of citizens and decision makers and integrating processes of knowledge co-generation and exchange among all users (Karpouzoglou et al 2016; Paul et al 2018). The architecture of such an EVO would allow incorporating all the different types and sources of data, from monitoring network to citizen data (Karpouzoglou et al 2016).

Some issues to consider

Citizen science is clearly demonstrating a huge potential in the field of flood hazard assessment, from clarifying flood scenarios, improving basic data collection and acquiring real-time data during events. Nevertheless, some aspects should be accounted for, some of them also discussed at length by Buytaert et al (2014).

One of the main concerns is the quality and reliability of citizen science data, being either directly acquired or crowdsourced, as raised by several authors in various domains (e.g. De Longueville et al 2010; Snäll et al 2011; Gollan et al 2012; Lukyanenko et al 2016). Associated to these aspects and a challenge on its own is how to assess and handle uncertainties related to citizen science data (Lahoz et al 2016).
Lewandowski and Specht (2015) investigated the quality of data collected by volunteers through the analysis of peer-reviewed literature and listed recommendations provided by different studies on how data quality could be improved. These recommendations highlighted the use of survey protocols and training for volunteers, statistical approaches to collect and treat the data as well as defining a clear structure for the project. When applied to flood hazard, some researchers (e.g. Sy et al 2016) made provision training prior to data acquisition to ensure that the requested information on flood extent and water level would be surveyed as accurately as possible.

On the side of VGI, its credibility as a reliable source of data has been discussed (Flanagin and Metzger 2008) as it provides a user-generated content. A way to analyse its reliability is to compare results with traditional and trusted providers of data (e.g. offices, research centres, newspapers). For example, De Longueville et al (2010) looked at the potentiality of VGI information derived from Flickr to locate major floods in UK. To ensure reliability, they developed a workflow to obtain reliable data from raw VGI information and they compared the results with medias and scientific databases. Their results show that VGI valuably complement other types of data.

Other aspects to consider when citizen science projects involve contributions through dedicated websites are related to its design and usability. As contributors have different levels of expertise (Newman et al 2010), websites should be designed to ensure the ease of use to be successful (Parr and Scholz 2015), a fact even more so for EVOs as discussed by Karpouzoglou et al 2016.

Another crucial challenge is the commitment of people to the project and over time. Studies demonstrate that engagement and motivation of contributors are necessary elements for success of citizen science projects (e.g. Eveleigh 2014). Understanding the drivers of motivation will provide the background to set up a successful partnership between the different groups (Rotman et al 2012) of the project and factors that can help long term involvement include regular feedback, recognition of the volunteers work, local meetings and concern for their community (Lewandowski and Specht 2015).

People’s motivations in flood hazard assessment seem therefore to be self-evident because it is about their livelihoods, so they could have a personal interest in participating because they can feel their contribution is useful for them and their societies. However, convincing people to be part require some communication skills and the flood hazard assessment project should be designed and presented in a way that all contributors are convinced that their action will be beneficial for them and for the society they belong to. Ensuring people engagement over time on flood hazard project requires understanding the local context, and to work with community leaders representing citizen groups as defined by Bénit-Gbaffou and Katsaura (2014) or well implemented local associations.

**Conclusions**

Evaluating flood risk and implementing mitigation measures to reduce the negative consequences in case of an event requires flood hazard assessments to be as realistic as possible. This paper demonstrates that current flood hazard assessment faces some issues, related to the flood causes, the spatial and temporal scale as well as the accuracy of data, where citizen science could help. At first, defining an accurate flood scenario is a challenge as considering the associated and aggravating processes requires understanding the local context as well as the mechanisms of past events. For this
reason, integrating local knowledge is indispensable, which is often ignored. Another point is the
temporal and spatial coverage of an area under study for both initial and output data. Various tools
and techniques are available that can be used to achieve assessments at local or regional scale, with
more or less high precision. However, availability of human and financial resources and the time
required to monitor and keep all information up to date are issues in most, if not all, contexts. In areas
where land-use evolves rapidly, updating data and consequently hazard and risk maps is crucial if one
wishes to have a proactive risk management aimed at reducing negative consequences.

Citizen science cannot resolve all the issues by itself, but it represents a valuable complementary
approach that can help to improve knowledge in particular in not-well covered areas, but also bring
new perspectives into the process of flood hazard assessment. Moreover, involving citizens leads to
real positive results in line with the Sendai framework for action (UNISDR, 2015b), such as raising
awareness among citizens in understanding hazard and risk and involving them in the process of risk
management.

Acknowledgements

The authors are grateful to the two anonymous reviewers and the associate editor for their comments,
which helped to improve the manuscript.

References

events: Proceedings of the 9th International Conference on Hydroinformatics, Tianhin, China.,1447-
1454.
Apel, H., Merz, B., and Thieken, A. H.(2008) Quantification of uncertainties in flood risk assessments:
count [accessed 31 May 2017].
analysis of disaster related human sensor network observations: Proceedings of the 11th International
ISCRAM conference-University Park, Pennsylvania. USA., 802-806.
2593-2597.
Legitimacy: Unpacking Bourdieu’s Political Capital in Post-Apartheid Johannesburg: International
understanding of science?: Public Understanding of Science., 25 (1), 2-16.
Brivio, P., Colombo, R., Maggi, M., and Tomasoni, R.(2002) Integration of remote sensing data and

This article is protected by copyright. All rights reserved.


This article is protected by copyright. All rights reserved.


Figure 1. Current framework for flood hazard assessment

- **Flood scenario**
  - primary cause of events
  - associated and aggravating processes

- **Basic data acquisition**
  - rainfall, discharge, topography, land use
  - ground-based measurement
  - remote sensing
  - hydrological modelling

- **Flood hazard assessment**
  - time, space
  - flood extent
  - water depth
  - flow velocity
  - duration

This article is protected by copyright. All rights reserved.
Table 1: Main strengths and weaknesses of current approaches on flood hazard assessment

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Key flood hazard parameters</th>
<th>Ground measurement</th>
<th>Hydrological modelling</th>
<th>Remote sensing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Associated and aggravated processes</td>
<td>Duration</td>
<td>- In situ data</td>
<td>- High precision</td>
<td>- High precision</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Spatial coverage</td>
<td>- Time consuming</td>
<td>- High cost</td>
</tr>
<tr>
<td></td>
<td>Water depth</td>
<td>- Real conditions</td>
<td>- High precision</td>
<td>- Real conditions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Punctual data</td>
<td>- Temporal coverage</td>
<td>- Time consuming</td>
</tr>
<tr>
<td></td>
<td>Extent</td>
<td>- Local conditions</td>
<td>- High precision</td>
<td>- Local conditions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Temporal coverage</td>
<td>- Time consuming</td>
<td>- Local conditions</td>
</tr>
<tr>
<td></td>
<td>Topography</td>
<td>- High resolution data acquisition</td>
<td>- Present conditions</td>
<td>- Not applicable</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Spatial coverage</td>
<td>- Temporal coverage</td>
<td>- Time consuming</td>
</tr>
<tr>
<td></td>
<td>Land-use</td>
<td>- Current conditions</td>
<td>- Temporal coverage</td>
<td>- Time consuming</td>
</tr>
<tr>
<td></td>
<td>Discharge</td>
<td>- In situ observations</td>
<td>- Validating data</td>
<td>- Not applicable</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Punctual data</td>
<td>- Temporal coverage</td>
<td>- Time consuming</td>
</tr>
<tr>
<td></td>
<td>Basic data acquisition</td>
<td>- Local data</td>
<td>- Accountability of local weather conditions</td>
<td>- Spatial coverage</td>
</tr>
<tr>
<td></td>
<td>Rainfall</td>
<td>- Instrument precision</td>
<td>- Inaccurate measurement</td>
<td>- Understanding of flood processes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Maintenance and data processing cost</td>
<td>- Not applicable</td>
<td>- Not applicable</td>
</tr>
</tbody>
</table>

This article is protected by copyright. All rights reserved.
Table 2: Overview of potential contribution of citizen science in flood hazard assessment

<table>
<thead>
<tr>
<th>Citizen science techniques</th>
<th>Methodology</th>
<th>Event phase</th>
<th>Objective</th>
<th>Number</th>
<th>Studies areas</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Basic data acquisition</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Rainfall</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Citizen observatory</td>
<td>Local rain monitoring network</td>
<td>During</td>
<td>Modelling</td>
<td>1000 people</td>
<td>Haltwhistle Burn catchment (UK)</td>
<td>Starkey 2017</td>
</tr>
<tr>
<td>Social media</td>
<td>Analysis of Twitter messages on rainfall</td>
<td>During</td>
<td>Monitoring</td>
<td>620 tweets</td>
<td>Sao Paulo (Brazil)</td>
<td>De Andrade 2017</td>
</tr>
<tr>
<td>Web Platform</td>
<td>Measure of rainfall data through simple device and sent on website or phoned</td>
<td>During</td>
<td>Monitoring</td>
<td>-</td>
<td>Colorado (USA)</td>
<td>Cifelli et al 2017</td>
</tr>
<tr>
<td><strong>Discharge</strong></td>
<td>Web Platform</td>
<td>After</td>
<td>Modelling</td>
<td>-</td>
<td>Cordoba (Argentina)</td>
<td>Le Coz et al 2016</td>
</tr>
<tr>
<td><strong>Topography</strong></td>
<td>Citizen observatory</td>
<td>After</td>
<td>Modelling</td>
<td>171 photos</td>
<td>Jakarta (Indonesia)</td>
<td>Shaad et al 2016</td>
</tr>
<tr>
<td><strong>Land-use</strong></td>
<td>P-mapping and P-GIS</td>
<td>After</td>
<td>Hazard mapping</td>
<td>502 people</td>
<td>Yeumbeul Nord (Senegal)</td>
<td>Sy et al 2016</td>
</tr>
<tr>
<td>Web Platform</td>
<td>Visual interpretation of tagged pictures sent through application/web site</td>
<td>Ø</td>
<td>Land-use Mapping validation</td>
<td>749 confluences</td>
<td>Global land-use</td>
<td>Iwao 2006</td>
</tr>
<tr>
<td>Web Platform</td>
<td>Collection and analysis of tagged pictures from Global Geo-Referenced Field photos</td>
<td>Ø</td>
<td>Land-use Mapping validation</td>
<td>2106 photos</td>
<td>Global land-use</td>
<td>Dong et al 2012</td>
</tr>
<tr>
<td>Web Platform</td>
<td>Analysis of vector polygons: buildings and polylines: roads, rivers from Open Street Maps</td>
<td>Before</td>
<td>Flood Modelling</td>
<td>-</td>
<td>Bregenzer Ache (Austria)</td>
<td>Dorn et al 2014</td>
</tr>
<tr>
<td>Web Platform</td>
<td>Visual interpretation of Google Earth and pictures sent through application and website (Geowiki)</td>
<td>Ø</td>
<td>Land-use Mapping validation</td>
<td>65 people</td>
<td>Global land-use</td>
<td>See et al 2015</td>
</tr>
<tr>
<td><strong>Extent</strong></td>
<td>P-mapping and P-GIS</td>
<td>After</td>
<td>Hazard mapping</td>
<td>-</td>
<td>Nega city (Phillipines)</td>
<td>Peters Guarin 2008</td>
</tr>
<tr>
<td><strong>Key flood hazard parameters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Water depth</strong></td>
<td>P-mapping</td>
<td>After</td>
<td>Hazard</td>
<td>300 people</td>
<td>Gorakhput (India)</td>
<td>Singh 2014</td>
</tr>
</tbody>
</table>

This article is protected by copyright. All rights reserved.
<table>
<thead>
<tr>
<th>Method</th>
<th>Activities</th>
<th>Phase</th>
<th>Dataset Size</th>
<th>Location</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>P-mapping</strong></td>
<td>Water depth measurement, handled GPS, mobile GIS, land-use map categories</td>
<td>After</td>
<td>502 people</td>
<td>Yeumbeul Nord (Senegal)</td>
<td>Sy et al 2016</td>
</tr>
<tr>
<td>Citizen observatory</td>
<td>Reading of water level gauges and information sent by text messages</td>
<td>During</td>
<td>4 people</td>
<td>Pijnacker (Netherlands)</td>
<td>Alfonso et al 2010</td>
</tr>
<tr>
<td>Citizen observatory</td>
<td>Water level read from river gauges and given to an official office by hard copy</td>
<td>After</td>
<td>-</td>
<td>Dangila woreda (Ethiopia)</td>
<td>Walker et al 2015</td>
</tr>
<tr>
<td>Citizen observatory</td>
<td>Collection and analysis of texts and pictures from interviews and telephone</td>
<td>During</td>
<td>1700 photos</td>
<td>Elbe river Bassin (USA)</td>
<td>Poser and Dransch 2010</td>
</tr>
<tr>
<td>Social media</td>
<td>Collection and analysis of texts and pictures from Twitter, Instagram</td>
<td>During</td>
<td>18000 tweets, 14000 photos</td>
<td>New York city (USA)</td>
<td>Aulov et al 2014</td>
</tr>
<tr>
<td>Social media</td>
<td>Collection and analysis of texts and pictures (Twitter, Flickr)</td>
<td>During</td>
<td>84 posts</td>
<td>Dresden, river Elbe (Germany)</td>
<td>Fohringer et al 2015</td>
</tr>
<tr>
<td>Social media</td>
<td>Collection and analysis of texts and pictures from Twitter</td>
<td>After</td>
<td>43 tweets</td>
<td>Tyne river (UK)</td>
<td>Smith et al 2015</td>
</tr>
<tr>
<td>Social media</td>
<td>Collection and analysis videos from YouTube</td>
<td>After</td>
<td>16 videos</td>
<td>Dahl Hith (Saudi Arabia)</td>
<td>Michelsen et al 2016</td>
</tr>
<tr>
<td>Social media</td>
<td>Extraction and analysis georeferenced tweets through Twitter Stream API and REST API (dev.Twitter.com)</td>
<td>After</td>
<td>4268 tweets</td>
<td>Carolina (USA)</td>
<td>Li et al 2017</td>
</tr>
<tr>
<td>Web Platform</td>
<td>Citizen’s reading of water level gauges and sent by message texts through internet-based interface</td>
<td>During</td>
<td>68 observers</td>
<td>Watersheds, New York (USA)</td>
<td>Lowry and Fienen 2013</td>
</tr>
<tr>
<td>Web Platform</td>
<td>Citizens reading water level on rulers and information sent through mobile application and website</td>
<td>During</td>
<td>10 observers</td>
<td>Watershed, Sao Carlos (Brazil)</td>
<td>Degrossi et al 2014</td>
</tr>
<tr>
<td>Web Platform</td>
<td>Analysis of a collection of texts and pictures from City Council and website/Newcastle University</td>
<td>After</td>
<td>-</td>
<td>Tyne river (UK)</td>
<td>Kutija et al 2014</td>
</tr>
<tr>
<td>Social media/Web platform</td>
<td>Citizens reading water level gauges and analysis of pictures and videos through Twitter, e-mail, mobile application and website</td>
<td>After</td>
<td>-</td>
<td>Haltwhiste Burn catchment (UK)</td>
<td>Starkey et al 2017</td>
</tr>
<tr>
<td>Social media/Web platform</td>
<td>Collection and analysis of texts and pictures from Twitter, Facebook, Flickr, e-mails and UShahidi platform</td>
<td>During</td>
<td>1000 photos</td>
<td>Brisbane and Bremer rivers (UK)</td>
<td>McDougall 2011; McDougall and Temple-watts 2012</td>
</tr>
<tr>
<td>Social media</td>
<td>Collection and analysis of texts and pictures from Twitter</td>
<td>During</td>
<td>1800 tweets</td>
<td>Tyne and Wear rivers (UK)</td>
<td>Smith et al 2015</td>
</tr>
<tr>
<td>Social media</td>
<td>LSPIV analysis of video collected from YouTube</td>
<td>After</td>
<td>1 video</td>
<td>Saint-Julien torrent (France)</td>
<td>Le Boursicaud et al 2016</td>
</tr>
<tr>
<td>Social media</td>
<td>LSPIV analysis of drone video collected from YouTube</td>
<td>After</td>
<td>1 video</td>
<td>Ardèche river (France)</td>
<td>Le Coz et al 2016</td>
</tr>
</tbody>
</table>

- Not mentioned
Ø Could be applied for flood analyses

This article is protected by copyright. All rights reserved.