

An experimental evaluation of a crowdsourcing-based approach for flood risk management

Ranieri de Brito Moreira, Livia Castro Degrossi, and
João Porto de Albuquerque

University of São Paulo,
Institute of Mathematical and Computing Sciences,
São Carlos – SP, Brazil
{ranieribm,degrossi,jporto}@icmc.usp.br

Abstract Volunteered geographic information (VGI) is a potential source of information to complement other sources in flood risk management. However, there is still not enough experimental evidence about the usefulness of VGI in different situations and scenarios. We conducted an experimental evaluation for verifying if VGI, obtained through a crowdsourcing platform, can be useful for the flood risk management context. The experiment occurred in two points of the watershed of São Carlos/SP in Brazil with 15 participants. The results show that volunteered geographic information is, in average, comparable to sensor data. Thus, we can conclude that using crowdsourcing for producing VGI can be a useful source for flood risk management.

Keywords: Experimental Evaluation, Experimental Software Engineering, Volunteered Geographic Information, Flood Risk Management, Disaster Management

1 Introduction

The increase in the number of natural disasters has been a growing concern for national and international organizations, due to their environmental risks and potential hazards [1]. The occurrence of a natural disaster is related to environmental characteristics and social system vulnerability [2]. The prevention of natural disasters and reduction of the social system vulnerability are themes of major concern in a local, national and international level [3].

Disaster management (DM) consists of a set of activities that aim at preventing or reducing social and economic impacts of natural disasters [4, 5]. DM is divided into four main phases: mitigation, preparedness, response and recovery. In each of these phases, different types of actions, policies and information are required. In particular, investments in the first two activities (mitigation and preparedness) are expected to greatly reduce the impact and losses caused by a natural disaster.

According to “The International Disaster Database”¹, flood is the disaster with more occurrences in the world, and in particular in Brazil. Flood risk management (FRM) is a subarea of disaster management that aims to control a flood,

¹ Extracted From: <http://www.emdat.be>

prepare for its occurrence and mitigate its impacts [6]. In particular, an efficient preparation requires the monitoring of risk areas before a disaster [7,8]. Accurate information about the current state of environmental variables is mandatory for enabling the simulation of effects and severity of a disaster [6,8]. This type of simulation helps to reduce the impact of floods, allows the affected population to take preventive measures and enables agencies to develop emergency response actions.

However, effectively managing flood risk depends on the availability of up-to-date and accurate information about environmental variables for improving situational awareness [4]. Such information can be obtained from different sources, including sensors, satellites and other technologies. In addition, another source that can be used is information provided by volunteers, so-called Volunteered Geographic Information (VGI) [9].

The use of “human as sensors” can be a valuable source of information in the context of disaster management, due to the potentially large number of volunteers [4]. Indeed, with the increase on the number of mobile devices with GPS (*Global Positioning System*), along with the interactions enabled by Web 2.0, the creation of geographic information by the general public was facilitated. However, despite the advantages of volunteers’ participation in gathering information, several challenges still need to be faced.

The quality of data generated by volunteers is a major concern. Volunteered information can be created with omissions, exaggerations, and even errors. Another challenge is the lack of structure of volunteered information. This is often regarded as insufficiently structured, documented and validated [10], while information collected with devices such as satellites and sensors has well-defined patterns and structures. Thus, there is a need for verifying whether volunteered information can be useful for the flood risk management context, i.e. whether it is sufficiently structured and accurate so that it can contribute to this application domain.

In this context, Degrossi et al. [11] proposed a crowdsourcing-based approach for gathering useful volunteered information for flood risk management context. The authors performed an experimental evaluation of the approach with 10 participants in one point of the watershed of São Carlos/SP city. Due to the small number of participants and points of the watershed, there is still not enough experimental evidence to support the statement of usefulness of volunteered information in this context. Thus, this paper presents an experimental evaluation for verifying if volunteered information, obtained through a crowdsourcing platform, can be useful for flood risk management context in Brazil.

The remainder of this paper is structured as follows: Section 2 presents the theoretical background for this work. In section 3, the experimental evaluation is explained. Section 4 presents the results. Finally, Section 5 presents the conclusions of this study.

2 Background

In the next sections, we present the concepts and underlying principles related to this work.

2.1 Flood Risk Management

Flood Risk Management is the process of managing a flood risk situation. This aims at controlling a flood, being prepared for it and minimizing its impacts [6]. For this, FRM comprises actions before, during and after a flood. These actions involve early warning and forecasting scenario, contingency plans and restoration [12].

Among natural disasters occurring worldwide, floods are the most frequent, representing 30% of natural disasters [8]. The increase in the number of floods is associated with climate change, being aggravated due to urban sprawl and the phenomenon of rapid urbanization without the availability of essential services [13, 14]. The number of affected people and financial and economic damages increase every year [15].

The preparation phase aims at reducing residual risk through early warning systems and measures that can be taken to minimize flood impacts. For this, the constant monitoring of risks and danger assessment is required. Every dollar invested in flood prevention reduces in U\$ 25 dollars the damage incurred in a natural disaster [12].

Currently, geographic information and related technologies play a fundamental role in all phases of FRM. Natural disasters are typically monitored using different devices such as sensors, satellites, seismometers, among others. However, these devices do not provide information about the impacts of a disaster. Volunteered geographic information can be a valuable source of information about the impacts of natural disasters [16], due to the potentially large number of volunteers who act as “sensors”, noticing important parameters of a disaster in a local environment [4].

2.2 Volunteered Geographic Information

Erst, the creation of geographic information was carried out by official agencies. However, with the increase of interactions made possible by Web 2.0, the use of devices with GPS (Global Positioning System) and the access to broadband Internet, geographic information is being produced by people who have little formal qualification. This type of information is called Volunteered Geographic Information (VGI) [9].

Among the advantages associated with VGI, researchers emphasize its use to enhance, update or complement an existing geospatial data [9]. In different scenarios, volunteered information have better quality than data provided by specialized organization, since in different parts of the world this information is outdated or they were acquired with old and less precise technologies than those currently available to the general public [17].

However, despite the advantages of citizens’ participation in collecting information, there are a lot of challenges to be faced. Data quality is a major concern. Information from many individuals can lead to doubts about its credibility [18]. According to [10], the credibility of VGI can be understood as a subjective concept that describes whether a piece of information can be trusted, considering any possible intentional or unintentional omission or exaggeration error. Moreover, it is not known beforehand how and from where the information will be provided. Another challenge faced refers to the location. Unlike in-situ sensors, people

are in constant movement, so the observations they made need to be located so they become useful [4]. Furthermore, VGI is often regarded as poorly structured, documented and validated [10].

Recent natural disasters have shown that volunteered information can improve situational awareness by providing an overview of the present situation [4]. This fact occurs because VGI offers a great opportunity to raise awareness due to the potentially large number of volunteers that observe important parameters of disaster management in a local environment [4, 9, 19]. Still, despite recent advances in the development of sensors, their observations may not be available due to communication interruptions or even the destruction of the sensor, besides that a sensor is not able to measure certain phenomena such as hail storms [4].

In this scenario, different software platforms, also called crowdsourcing platforms, have been employed for gathering volunteered information and enabling its visualization and analysis. The term crowdsourcing can be understood as a production model where the intelligence and knowledge of volunteers are used to solve problems, create content and develop new technologies. Furthermore, the term refers to a way to organize the work, that involves an information system for coordinating and following up tasks carried out by people [20].

3 Experimental Evaluation

The effectiveness of an approach can be verified through an experimental evaluation conducted in a controlled and well-defined way. For evaluating approach proposed by Degrossi et al. [11], the authors conducted an experimental evaluation for verifying whether volunteered information, obtained through a crowdsourcing platform, was useful for the context of flood risk management. According to Degrossi et al. [11], a volunteered information is considered useful if it can be used in hydrological models or decision support systems. The results indicated that volunteered information can be useful for flood risk management context, since the average of volunteered information about water level was equal to the average of measurements carried out by sensor. However, there is not enough experimental evidence for such statement, since the number of participants in this evaluation was reduced (10 participants). Thus, we saw a need to conduct another experimental evaluation for verifying the usefulness of this type of information.

The experimental evaluation, conducted in this work, aimed at verifying the results obtained in the first experiment. For this, we selected fifteen new participants. The participants were selected from a set of students of a Experimental Software Engineering discipline ministered at the university.

Differently from the first experiment, in this the participants had no experience with the context of flood risk management. The use of volunteers without experience approaches this experimental evaluation of the real scenario of the approach proposed by Degrossi et al. [11], where any volunteer can provide information about environmental variables in the context of flood risk management. In addition, unlike the first experiment, two points of the watershed São Carlos/SP city were selected for conducting the experimental evaluation.

Before the experiment, participants went through a training about the crowdsourcing platform (Flood Citizen Observatory²), the mechanism used to interpret the water level in the riverbed, and how to insert a report in the platform. The Flood Citizen Observatory is a crowdsourcing platform that aims at obtaining volunteered information for flood risk management context, “more specifically about flooded areas and water level in the riverbed” [11]. After training, participants were conducted to the two points of the watershed of São Carlos/SP city for collecting information about water level in the riverbed and insert it in the crowdsourcing platform.

To guide this experimental evaluation, the research question defined by Degrossi et al. [11] was selected. The question aims at determining whether the difference between the average of information provided by participants (volunteers) and the average of data measured by a sensor is statistically significant. As noted by Degrossi et al. [11], it is important to verify this difference because if it is significant “volunteer information may not reflect the real state of the environmental variable observed, resulting in erroneous predictions about the flood risk”. Thus, the research question of this work is:

RQ) Is the difference between volunteered information and sensor data significant?

For this research question, two hypotheses were defined, a null hypothesis and a alternative hypothesis:

Null Hypothesis (H0): the average of volunteered information is equal to the average of sensor data.

$$\mu(volunteer) = \mu(sensor)$$

Alternative Hypothesis (H1): the average of volunteered information is different from the average of sensor data.

$$\mu(volunteer) \neq \mu(sensor)$$

For each experiment variable, this is the average of volunteered information and average of sensor data, metrics were defined for measuring the variable. The average of volunteered information was calculated by the sum of the values observed by each participant at one point and divided by the number of observations. It is noteworthy that, in this context, a volunteer information refers to the value of the water level in the water level ruler. This is an interpretation mechanism laid down on the riverbed that allows a volunteer to determine the water level. On the other hand, the average of sensor data was calculated by the sum of measurements carried out by a sensor during the period of time that the participants performed observations, and divided by the number of measurements.

To analyze the results we used four statistical tests: Shapiro-Wilk Test, Levene’s Test, T Test and Mann-Whitney Test. The Shapiro-Wilk Test [21] was used

² <http://www.agora.icmc.usp.br/enchente>

to check whether the samples had normal distribution. If samples had normal distribution, the Levene's Test [22] was applied for verifying if the samples had equal variance. Then the T Test [23] was performed to statistically compare the sample's average and thus reject or accept the null hypothesis. If samples did not have a normal distribution, the Mann-Whitney Test [23] was performed.

3.1 Instrumentation

The main difference between the first experiment and this evaluation is related to the number and knowledge of participants, and the number of collection points.

Unlike the first experiment, which included the participation of 10 volunteers, 15 new volunteers participated of this experimental evaluation. Thus, the increase in the number of participants improves the confidence in results.

In the first experiment, 50% of participants had at least one month of experience with flood risk management. In this experimental evaluation, the participants selected had no experience with this context. The lack of experience of participants approaches the experiment to the real scenario, since most of the affected population have no experience with the context of flood risk management. Furthermore, the crowdsourcing-based approach, proposed by Degrossi et al. [11], aims at providing mechanisms to help the population in providing useful information for flood risk management context.

Besides the changes related to participants, in this experimental evaluation, information gathering was carried out in two distinct points of the watershed of São Carlos/SP city. It is worth noting that both points are equipped with a water level ruler, that it is used as a mechanism to help participants in better interpreting environmental variable, i.e. water level.

Before the experiment, participants went through a training. The training addressed three points: (i) the mechanisms used to interpret environmental variables; (ii) the crowdsourcing platform used to obtain volunteered information; and (iii) instructions about how to insert a report in the platform.

Degrossi et al. [11] proposed four interpretation mechanism for helping volunteers in better interpreting environmental variables, i.e. water level. The first mechanism is a water level ruler that it is laid down in the riverbed. In this experiment, only this mechanism was used by volunteers to interpret the water level in the riverbed. This mechanism was chosen because it offers a higher accuracy of the water level measure, so it is possible to make a comparison between the measurements carried out by sensors and the observations made by volunteers.

The second mechanism proposed is a dummy with the human form, which has marks that represent the water level. The third mechanism corresponds to a multi-color bands, which represents the hazard index [24]. This index represents the level of danger to the population, this is, the forces exerted on an individual by the water. Finally, the fourth mechanism may be used in points that do not have any of the mechanisms mentioned earlier. This mechanism is based on the popular knowledge, this is, water level is determined based on four marks: low, normal, high and overflowing.

The interpretation mechanisms are used as part of a report inserted in the crowdsourcing platform, which was used for gathering volunteered information. The crowdsourcing platform aims at obtaining useful volunteered information for flood risk management. When inserting a report, a volunteer must provide four

mandatory information: (*i*) report title; (*ii*) report description; (*iii*) category, which represents the interpretation mechanism used in the observation; and (*iv*) the name of the local.

The conduction of the experiment was carried out in three steps. In the first step, participants collected information about the water level in the water level ruler at the first point, called Point A. In the second step, participants also collected information about the water level in the water level ruler, but at the second point, called Point B. Finally, in the third step, participants inserted their observations in the crowdsourcing platform.

It is noteworthy that during the three steps, participants were instructed to not exchange information with each other, avoiding bias in the information gathering. In addition, all steps were followed by the researchers responsible for the experiment.

4 Results and Discussion

The experiment was conducted with 15 participants, with no prior knowledge about flood risk management, in two points of the watershed of the city of São Carlos/SP. The results of this experimental evaluation were based on the four statistical tests, as mentioned in Section 3.

Information inserted by participants in the crowdsourcing platform were used to verify if the difference between the average of volunteer information and the average of sensor data is statistically significant. It is worth mentioning that each point was considered independently for the statistical analysis. In the following sections we present the results for each point.

4.1 Point A

For the statistical analysis related to Point A, 15 volunteered information, inserted by participants in the crowdsourcing platform, and 11 sensor data, which were measured during the experiment, were used (Table 1).

The first statistical test carried out was Shapiro-Wilk Test, in order to determine whether the samples had a normal distribution. In particular, the significance level³ adopted for this test was 5%. The results (Table 1) show that both samples have normal distribution, since the value of Sig. (p-value) is higher than the significance level for both samples.

Table 1. Shapiro-Wilk Test Results (Point A).

Tested Variable	Instrument	Shapiro Wilk Test	
		Number of elements	Sig.
Water Level	Volunteer	15	0,109
	Sensor	11	0,154

According to Shapiro-Wilk Test results, T Test was carried out for verifying if the difference between the average of each sample is statistically significant. Before this test, it is necessary to analyze whether the samples have equal variance.

³ The significance level corresponds the probability to reject the null hypothesis being that true.

Therefore, we performed Levene's Test. The results (Table 2) show that the samples do not have equal variance, since the value of Sig. (p-value) is smaller than the significance level. Thus, at the T Test, we observed the results for samples that do not have equal variance.

Finally, T Test was carried out in order to reject the null hypothesis. The result (Table 2) shows that the null hypothesis can not be refuted, since the value of Sig. (p-value) is higher than the significance level. Te can conclude that there is not statistically significant difference between the average of volunteer information and the average of sensor data, i.e. the average are equal with a confidence level of 95%.

The results indicate that volunteered information is useful for flood risk management context, even with the lack of experience of participants. This can be explained by the fact that participants went through a training before the execution. Thus, we can assert that the training was sufficient for enabling participants to produce useful volunteered information for flood risk management.

It is noteworthy that the water level ruler, used in information gathering, had a slight deviation on its bottom, which can be considered a threat to validity of the study. However, since this is a real scenario, this threat approaches the experiment to the real use proposed in the approach.

Table 2. Results of Levene's Test and T Test (Point A).

Tested Variable	Levene's Test	T Test
	Sig.	Sig.
Water Level	0,001	0,750

4.2 Point B

For the statistical analysis related to Point B, 13 volunteer information was provided by participants, while 21 sensor data were measured during the conduction of the experiment (Table 3).

Similar to Point A, the first statistical test performed was Shapiro-Wilk Test, in order to verify if the samples had normal distribution. The significance level adopted for this test was 5%. However, unlike point A, the results (Table 3) indicated that both samples do not have normal distribution, since the value of Sig. (p-value) is smaller than the significance level for both samples. Thus, for samples that do not have normal distribution, the most appropriate hypothesis test is the Mann-Whitney Test.

Table 3. Shapiro-Wilk Test Results (Point B).

Tested Variable	Instrument	Shapiro Wilk Test	
		Number of elements	Sig.
Water Level	Volunteer	13	0,001
	Sensor	21	0,004

The Mann-Whitney Test, as well as T Test, was carried out for verifying if the difference between the average of samples from Point B is statistically significant. The results (Table 4) show that the null hypothesis can not be rejected, as well

as at Point A, since the value of Sig. (p-value) is higher than the significance level, i.e. the average of both samples at Point B are equal with a confidence level of 95%. Thus, we can assert that volunteer information is useful for context of flood risk management.

Table 4. Results of Mann-Whitney Test (Point B).

Tested Variable	Mann-Whitney Test
	Sig.
Water Level	0,059

As at Point A, the water level ruler also had factors that may threaten the validity of the study. However, such factors approximate the experiment of the real conditions found by volunteers. In the case of Point B, the bottom of the water level ruler was obstructed by vegetation, as shown in Figure 1. This obstruction hindered the determination of the water level by the volunteer, which could cause erroneous values.



Figure 1. Water level ruler at Point B.

5 Conclusion

In this work, an experimental evaluation was carried out for verifying whether volunteered information can be useful for the flood risk management context. In the experiment, volunteers provided information about the water level in the riverbed using a crowdsourcing platform. The results showed that, with the use of interpretation mechanisms and training, volunteers were able to provide information about environmental variable (i.e. water levels) in an accurate way, i.e. they were able to provide useful information to the context of flood risk management. We can make such assertion since the average of observations made by volunteered was statistically equivalent to the average of sensor data.

Furthermore, the use of the Flood Citizen Observatory, as a crowdsourcing platform, was able to mitigate the variability in the structure of volunteered observations, since the platform provides some standards for information provision. Thus, recurring problems, such as the low or lack of structure of volunteered information, can be attenuated. The structure of volunteered information is the first step towards integrating this type of information with data from other sources, such as sensors, rain gauges, among others.

Although the experiment was effective for answering the research question, improvements should be carried out in this context. Volunteers had problems with the mobile application interface, i.e. some volunteers had difficulty in interpreting some report fields at the time to submit an observation. Another challenge of the usefulness of volunteered information is related to VGI quality. In particular, in the crowdsourcing platform, the quality assessment is carried out manually. So, it is necessary to develop an automatic method for the evaluation of the quality of volunteered information.

Acknowledgment

We would like to express our thanks for the financial support provided by CAPES (Coordination for the Improvement of Higher Education Personnel, Brazil). João Porto de Albuquerque is grateful for FAPESP (grant no. FAPESP 2012/18675-1), CAPES (grant no. 12065-13-7) and Heidelberg University (Excellence Initiative II / Action 7) for providing funding for his contribution to this paper.

References

1. UN, U.N.: Humanitarian and disaster relief assistance. Accessed on: 12/02/2013 (2013)
2. Alcántara-Ayala, I.: Geomorphology, natural hazards, vulnerability and prevention of natural disasters in developing countries. *Geomorphology* **47** (2002) 107–124
3. Levy, J.K., Gopalakrishnan, C., Lin, Z.: Advances in decision support systems for flood disaster management: Challenges and opportunities. *International Journal of Water Resources Development* **21**(4) (2005) 593–612
4. Poser, K., Dransch, D.: Volunteered geographic information for disaster management with application to rapid flood damage estimation. *Geomatica* **64**(1) (2010) 89–98
5. Tu, S., Abdelguerf, M.: Web services for geographic information systems. *IEEE Internet Computing* (2006) 13–15
6. Plate, E.J.: Flood risk and flood management. *Journal of Hydrology* **267** (2002) 2–11

7. Kobiyama, M., Mendonça, M., Moreno, D.A., de Oliveira Marcelino, I.P.V., Marcelino, E.V., Gonçalves, E.F., Brazetti, L.L.P., Goerl, R.F., Moller, G.S.F., de Moraes Rudorff, F.: *Prevenção de Desastres Naturais - Conceitos básicos. Organic Trading* (2006)
8. van der Kooij, M.: Flood monitoring and disaster management response. *GEO-connexion International Magazine* (Abril 2013) 26–28
9. Goodchild, M.F.: Citizens as sensors: the world of volunteered geography. *GeoJournal* **69**(4) (Agosto 2007) 211–221
10. Longueville, B.D., Luraschi, G., Smits, P., Peedell, S., Groeve, T.D.: Citizens as sensors for natural hazards: A vgi intergration workflow. *Geomatica* **64** (2010) 41–59
11. Degrossi, L.C., ao Porto de Albuquerque, J., Fava, M.C., Mendiondo, E.M.: Flood citizen observatory: a crowdsourcing-based approach for flood risk management in brazil. In: 26th International Conference on Software Engineering and Knowledge Engineering, Hyatt Regency, Vancouver, Canada (2014)
12. Mendiondo, E.M.: Flood risk management of urban waters in humid tropics: early-warning, protection and rehabilitation. In: Workshop on Integrated Urban Water Management in Humid Tropics, Foz de Iguaçu, Brazil (2005)
13. Munro, R.: Crowdsourcing for pakistan flood relief. Accessed on: 11/27/2013 (August 2010)
14. McDougall, K.: Using volunteered information to map the queensland floods. In: Proceedings of the Surveying & Spatial Sciences Biennial Conference, Wellington, New Zealand (2011)
15. Schade, S., Díaz, L., Ostermann, F., Spinsanti, L., Luraschi, G., Cox, S., nez, M.N., Longueville, B.D.: Citizen-based sensing of crisis events: sensor web enablement for volunteered geographic information. *Applied Geomatics* **5**(1) (2013) 3–18
16. Zook, M., Graham, M., Shelton, T., Gorman, S.: Volunteered geographic information and crowdsourcing disaster relief: A case study of the haitian earthquake. *World Medical & Health Policy* **2** (2010) 7–33
17. Longueville, B.D., Smith, R.S., Luraschi, G.: Omg, from here, i can see the flames! - a use case of mining location based social networks to acquire spatio-temporal data on forest fires. In: Proceedings of the 2009 International Workshop on Location Based Social Networks, New York, USA (2009) 73–80
18. Erskine, M.A., Gregg, D.G.: Utilizing volunteered geographic information to develop a real-time disaster mapping tool: A prototype and research framework. In: International Conference on Information Resources Management, Vienna, Austria (2012)
19. Roche, S., Propeck-Zimmermann, E., Mericskay, B.: Geoweb and crisis management: issues and perspectives of volunteered geographic information. *GeoJournal* **78**(1) (2011) 21–40
20. Miorandi, D., Carreras, I., Gregori, E., Graham, I., Stewart, J.: Neutrality in mobile internet: Towards a crowdsensing-based citizen observatory. In: IEEE International Conference on Communications, Budapest, Hungary (2013)
21. Royston, J.P.: Some techniques for assessing multivariate normality based on the shapiro-wilk w. *Journal of the Royal Statistical Society* **32** (1983) 121–133
22. Olkin, I.: *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*. Stanford University Press (1960)
23. Wohlin, C., Runeson, P., Höst, M., Ohlsson, M.C., Regnell, B., Wesslén, A.: *Experimentation in software engineering: an introduction*. Kluwer Academic Publishers (2000)
24. Rotava, J., Mendiondo, E.M., Souza, V.C.B.: Simulação de instabilidade humana em inundações: Primeiras considerações. In: XX Simpósio Brasileiro de Recursos Hídricos, Bento Gonçalves, Brazil (2013)