POLITECNICO DI TORINO
Repository ISTITUZIONALE

A quality comparison between professional and crowdsourced data in emergency mapping for potential cooperation of the services

Original
A quality comparison between professional and crowdsourced data in emergency mapping for potential cooperation of the services / Elia, A., Balbo, S., Boccardo, P.. - In: EUROPEAN JOURNAL OF REMOTE SENSING. - ISSN 2279-7254. - ELETTRONICO. - 51:1(2018), pp. 572-586. [10.1080/22797254.2018.1460567]

Availability:
This version is available at: 11583/2715555 since: 2020-09-16T11:42:15Z

Publisher:
Taylor and Francis Ltd.

Published
DOI:10.1080/22797254.2018.1460567

Terms of use:

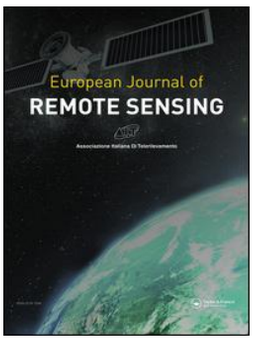
This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
Taylor and Francis postprint/Author's Accepted Manuscript

This is an Accepted Manuscript of an article published by Taylor & Francis in EUROPEAN JOURNAL OF REMOTE SENSING on 2018, available at http://wwww.tandfonline.com/10.1080/22797254.2018.1460567

(Article begins on next page)

26 November 2026



European Journal of Remote Sensing

ISSN: (Print) 2279-7254 (Online) Journal homepage: http://www.tandfonline.com/loi/tejr20

A quality comparison between professional and crowdsourced data in emergency mapping for potential cooperation of the services

Agata Elia, Simone Balbo & Piero Boccardo

To cite this article: Agata Elia, Simone Balbo & Piero Boccardo (2018) A quality comparison between professional and crowdsourced data in emergency mapping for potential cooperation of the services, European Journal of Remote Sensing, 51:1, 572-586, DOI: 10.1080/22797254.2018.1460567

To link to this article: https://doi.org/10.1080/22797254.2018.1460567

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

Published online: 21 May 2018.

Submit your article to this journal

Article views: 143

View Crossmark data

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=tejr20

EUROPEAN JOURNAL OF REMOTE SENSING, 2018
VOL. 51, NO. 1, 572–586
https://doi.org/10.1080/22797254.2018.1460567

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

A quality comparison between professional and crowdsourced data in emergency mapping for potential cooperation of the services

Agata Elia[a], Simone Balbo[b] and Piero Boccardo[c]

[a]Politecnico di Torino, Torino, Italy; [b]ITHACA, Torino, Italy; [c]Politecnico di Torino, DIST, Torino, Italy

ABSTRACT

A protocol for assessing the quality of digital geospatial data is applied to samples of volunteered geographic information (created by Humanitarian OpenStreetMap Team) and professional mappers (Copernicus EMS-rapid mapping). The application on pre-event data shows that a large percentage of them is very similar in terms of quality and is, therefore, potentially interchangeable; post-event data reveal a more divergent behaviour. The results gathered from the comparative analysis and a look at the temporal trends of response of volunteers and professionals justify a framework of interaction of respective activities, which seems to be possible under strong relationship between professionals and volunteers, built upon common operational standards and guidelines.

ARTICLE HISTORY
Received 1 February 2017
Revised 23 March 2018
Accepted 30 March 2018

KEYWORDS
SEM; crowdsourcing; emergency; data quality

Introduction

In recent years, emergency responders have increasingly used geographic information systems (GIS) and earth observation (EO) to help in assessing the impact of a disaster and planning and coordinating the emergency response by means of geospatial data. In this context, satellite-based emergency mapping (SEM) is defined as the creation of geo-information products and spatial analyses dedicated at providing situational awareness and immediate crisis information based on the extraction of reference (pre-event) and crisis (post-event) geographic information from remotely sensed imagery (IWG-SEM, 2014).

"Digital humanitarians" (Meier, 2015) are involved in the effort of creating information, starting from their first widely recognized contributions in the aftermaths of Haiti's earthquake in 2010. Among them, mappers work through visual interpretation of satellite or aerial images, extracting geo-referenced data and making it available to the various stages of crisis response. A branch of SEM is rapid mapping, which supports the immediate response phase by providing on-demand geospatial information in support of the first emergency management activities.

Nowadays, two main groups of contributors have the potential to provide immediate information in the aftermaths of a disaster. One is the "crowd," a wide group of volunteers who perform the remote mapping activity with or without professional background. Based on open-source platforms, volunteers start interpreting satellite images and tracing information in the form of geospatial data. Volunteers contribute to the creation of volunteered geographic information (VGI) based on the principles of open and crowdsourced data and participatory mapping. Creators of VGI are the Humanitarian OpenStreetMap Team (HOT) and Tomnod, among others. The other group of contributors is constituted by professional mappers, who will be defined in this article as mappers who are trained and who are officially designated to work on producing professional emergency mapping products under specific national or international programmes. Representatives of professional mappers are the Copernicus Emergency Management Service (EMS), the United Nations Operational Satellite Applications Programme (UNOSAT) or the National Geospatial-Intelligence Agency (NGA), and others.

The challenges of quality assurance of emergency data and of coordination between all the actors involved in order to avoid duplication of products and reduce efforts in the after-disaster deployment have been pointed out several times in the SEM context (Harvard Humanitarian Initiative, 2011; IWG-SEM, 2015). Quality becomes a highly critical issue in emergency RM, especially in the case of voluntary work, which sees the contribution of volunteers with or without professional background. The two main issues encountered in the use of crowdsourced data are the volume of the information produced and its reliability in terms of data quality. In addition, volunteers and professional mappers are often called to work in producing the same type of information on overlapping areas of interest (AOIs). In the light of previous consideration, the following questions may arise: what quality and type of data are expected to be generated by volunteer mappers with

CONTACT Agata Elia agata.elia1991@gmail.com Politecnico di Torino, C.so Duca degli Abruzzi 24 – 10129, Torino

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

various backgrounds in GIS, remote sensing and the emergency management context? Is it possible to plan and implement a cooperation and interaction of the activities performed by volunteer and professional mappers?

In this study, we provide the results of a comparison of activities and products of the community of volunteer mappers and professional mappers, respectively. The main aim is to identify the strengths of the two initiatives by comparing their production steps and results, with the final goal of proposing a draft framework of integration.

Data quality

The quality and relevance of products of SEM are strictly related to the quality of geospatial digital data used as an input. Geospatial data have four main components of quality upon which data scientists commonly agree: accuracy, precision, consistency and completeness (Veregin, 1999). In addition to the conventional and previously stated components, three more dimensions of quality worth being stated are validity, timeliness and accessibility (Ivánová, 2007; PennState, 2015). The definition of data quality, though, must go together with the definition of geospatial data itself. An object represented by geospatial data is described through the combination of three components: spatial, temporal and thematic. It is critical to evaluate the quality of each of these components.

Materials and methods

Copernicus EMS-RM and HOT

In order to compare the quality of data produced by volunteer and professional mappers, the OpenStreetMap (OSM) community led by the Humanitarian OSM Team and the Copernicus EMS-RM (Copernicus EMS-RM) are chosen as representatives.

Copernicus is a European Union (EU) programme aimed at developing information services based on satellite EO and *in situ* data, with the objective of monitoring and forecasting the state of the environment on land and sea and in the atmosphere. Copernicus is a user-driven programme, and the information services provided are freely and openly accessible to its users, mostly public authorities.

The Copernicus EMS, one of the six Copernicus core services, provides maps and analyses based on satellite imagery in response to a wide variety of disaster types. It thereby supports crisis managers, Civil Protection authorities and humanitarian aid actors dealing with natural disasters, man-made emergency situations and humanitarian crises, as well as those involved in recovery, disaster risk reduction and preparedness activities. The EMS-RM module provides rapid service delivery, and it is available all day for every day of the year (Copernicus EMS, 2015b); the typical response time varies on the basis of the products in a range between 3 and 12 h.

The main products are reference, delineation and grading maps. Reference maps illustrate an updated picture of the area using data prior to the event but as close to the event as possible. Reference maps include features that can be useful in the specific crisis management task, such as selected topographic features on the affected area, in particular exposed assets. Delineation maps describe the post-event situation of the area, providing an assessment of the event extent and evolution. Grading maps provide an assessment of the damage grade, including the extent, magnitude or damage grades specific to each emergency type. They may also provide relevant and up-to-date information that is specific to affected population and assets (e.g. settlements, transport networks, industry and utilities) (Copernicus EMS, 2015b).

The Copernicus EMS-RM is activated by authorized users, which can be national focal points (e.g. the Presidency of the Council of Ministers – Italian Department of Civil protection for Italy) or EU services (e.g. the EU's Emergency Response Coordination Centre) (Copernicus EMS, 2015b). The authorized users request the AOI, the content of the map and the type of analysis; the request is evaluated by the European Response Coordination Centre, and upon approval it is managed by the service provider, which is the commercial entity in charge of the EMS-RM production and provision of geospatial information that actually performs the analysis and produces the maps. The analysis is usually performed by comparing data that describe the pre-event and the post-event situation. In particular, post-event data are usually satellite imagery coming from several providers, public and commercial; the satellite imagery is provided through the European Space Agency, which is entrusted with the responsibility to coordinate access to satellite data, following a Delegation Agreement with the European Commission (European Commission, 2015).

Even though the products of Copernicus EMS-RM are provided to the public in full and open access (with some exceptions due to sensitive locations or events), specific access conditions apply for satellite imagery other than Sentinel data (European Commission, 2015). These have the effect that only the service provider has free and direct access to this imagery, while public organizations (e.g. the authorized user) have to sign a specific licence agreement first. Instead, third parties and the public cannot get free access to the original satellite data.

OSM is a collaborative project aimed at creating a free and editable map of the world. Counting a

community of more than 3 million users in 2017 (OpenStreetMap, 2017), OSM is based on the principles of data openness and free sharing plus civic participation.

The HOT is a global community that works to apply these principles towards humanitarian response and economic development. When a disaster occurs and humanitarian response is needed, HOT coordinates OSM volunteers, contacting responding organization to determine their needs in terms of what to map and where. HOT members investigate the availability of satellite imagery covering the required AOI; pertinent partners are contacted to provide compatible imagery that is then provided to the online community of OSM volunteers. The HOT Tasking Manager is the major tool through which the HOT community coordinates volunteers in emergency mapping by setting up mapping projects; the Tasking Manager is based on the idea of dividing a mapping task into smaller ones in order to complete them rapidly (learnOSM, 2015). Mappers work on tracing from the imagery, focusing on recognizable objects that are useful for humanitarian response. This information is then entered into the database of OSM and can be freely exported into different suitable formats (learnOSM, 2015).

Data generated by these two services in the aftermath of the earthquake that hit Nepal in April 2015 are taken into account for the analysis. In Nepal, the HOT/OSM effort on 28 April saw 2182 digital volunteers (Wired, 2015), tracing more than 14,700 km^2 high-resolution satellite imagery. The activity performed by Copernicus EMS-RM covered more than 7600 km^2 high-resolution satellite imagery (Copernicus EMS, 2015a).

The ELF protocol

In order to technically assess the quality of digital geospatial data, it is necessary to identify an adequate technical methodology and protocol. For this purpose, the quality comparison protocol developed and proposed by the European Location Framework (ELF) is selected for this study. The ELF protocol was developed in order to constitute a proper comparison protocol for evaluating from both a qualitative and quantitative perspective the benefits derived by the use of particular datasets distributed through the ELF and its functionalities in emergency mapping activities (Tonolo, Perez, & Steffenino, 2014).

The ELF protocol for data quality assessment is based on the definition of indicators; each of them is designed for the evaluation of a specific aspect of geospatial data quality. When an indicator is applied to a dataset, a numeric value represents the quality with respect to that specific indicator (Tonolo et al., 2014).

This protocol allows an in-depth analysis of data quality, while short time constraints in emergency mapping activities often do not allow for structured and comprehensive quality control before the delivery.

In detail, the aspects taken into consideration by the mentioned indicators are listed below.

a. Update: it is the date (and the time) associated with measuring a physical quantity. In the SEM case, it is the acquisition date of the satellite imagery used for generating the dataset; the update improves as the time between the acquisition date and the event of interest decreases. The risk to be minimized is constituted by objects on the ground changing independently from the disaster of interest.
b. Geometric accuracy: it refers to the displacement between the objects represented in the dataset to be evaluated and ground truth.
c. Thematic accuracy: it refers to the availability of attributes that describe the objects represented in the dataset. A higher availability of attributes relevant in the context of emergency response indicates a higher quality dataset.
d. Completeness: it reflects the percentage of missing objects (omission errors) and incorrectly produced objects (commission errors) in the dataset to be evaluated (Tonolo et al., 2014).

Methodology

The methodology is based on the application of the protocol described in the previous section: first to datasets generated by volunteers, then to datasets generated by professional mappers; the aim is to detect and report differences in terms of data quality between the two sources. A detailed description of the protocol application methodology is shown in the following section.

The main limitation of the application of the protocol is that no ground truth dataset is available; therefore, when applicable, available high-resolution satellite imagery was used in replacement of ground truth, while in other cases the datasets produced by the two sources were compared among each other. As a result, generally speaking, the considerations derived from this work are mostly related to a relative quality level between the two sources, rather than an absolute quality level.

The application of the protocol is possible on data that respect a set of terms:

a. Data must be thematically comparable; this means that compared data must belong to the same category.
b. Data must be compared on the same AOI, in terms of limiting the comparison only to those areas that are taken into consideration in the mapping process by both services.

c. Data must be representative and significant: representative in terms of evaluating the difference between the two datasets on an adequately variable context, for instance comparing data produced on a highly urbanized context but also data produced on sparsely inhabited area, and significant in terms of the statistic relevance of the amount of data that are compared.

In the light of these requirements, the set of data on which the protocol is applied comprises two pre-event datasets and two post-event datasets: road network and buildings as pre-event data, internally displaced people (IDP) camps and damaged infrastructures (buildings) as post-event data. These features were identified as mapping targets by both services and the extension of the overlap of AOI on which Copernicus EMS-RM and HOT/OSM worked contains a number of data that are considered to be significant and representative in terms of spatial distribution. For instance, area interested by the mapping of the road network features by both services is above 7500 km^2, covering highly urbanized areas as well as sparsely inhabited areas. In addition, this choice allows us to verify the existence of differences in pre- and post-event data mapping activities and behaviour.

The process of application of the protocol can be summarized in the following steps, also illustrated in Figure 1.

a. Data download: the first phase is gathering the necessary data. OSM datasets are mostly collected through the OSM API (application programming interface), through queries that allow filtering and downloading, starting from the whole OSM database. Copernicus EMS-RM datasets are available for download on the Copernicus EMS website, in the section related to the activation for the earthquake in Nepal.
b. Data filtering, cleaning and selecting: the sample datasets are obtained from the previous step by extracting only the theme of interest; this process allows us to reduce the information noise caused by the inclusion of records that are not of interest for this study (e.g. when considering the IDP camps from Copernicus dataset, they have to be extracted from the "crisis_information_p" dataset, which contains much more post-event features in addition to IDP camps).
c. Data filtering by overlapping AOIs: datasets have to be reduced to only those digitized on common AOIs between Copernicus EMS-RM and HOT/OSM.
d. Protocol customization and application: each dataset requires a specific understanding of the correct implementation of the protocol, e.g. point and line features require the implementation of different spatial analyses due to their different geometric nature.
e. Calculation of values for the indicators: the type of values associated with each indicator varies with the indicators, e.g. percentages or numerical values.
f. Analysis of the results: the two sources are compared on a single indicator value at any time.

The following paragraphs briefly unfold the methodology pursued for the calculation of values for each indicator.

Update evaluation

The update indicator is evaluated by retrieving the date of acquisition of satellite imagery used for digitization of features. For post-event features, we refer to the date of post-event imagery. The results are frequency distributions of source dates within the features, as shown in Figure 2.

Geometric evaluation

The geometric accuracy indicator is evaluated based on a three-step workflow, as illustrated in Figure 3.

Data are pre-processed. Then, taking one dataset at any time (road network, IDP camps, buildings and damaged infrastructures), a feature matching is performed in order to identify features in the two data sources (volunteers and professionals) that represent the same object on the field and to match them. Finally, the geometric accuracy is evaluated only on the matches within the two sources.

For the geometric evaluation of data, a buffer approach is adopted in order to evaluate the critical displacement between the two sources, which is the displacement that does not fall within the declared tolerance of final cartographic products taken as reference. In this application, professional-based data are taken as reference and they are buffered using a buffer size equal to the most conservative CE90 (circular error at 90% confidence) geometric accuracy value reported in the Copernicus EMS-RM maps covering the area of analysis. Then, the length (for linear features) or the number of features (for point features) in the other data source, in this case HOT/OSM datasets, which is not included in the defined buffer area, is considered as evidence of spatial displacement between the sources and could be



Figure 1. Standard workflow for the application of the protocol.

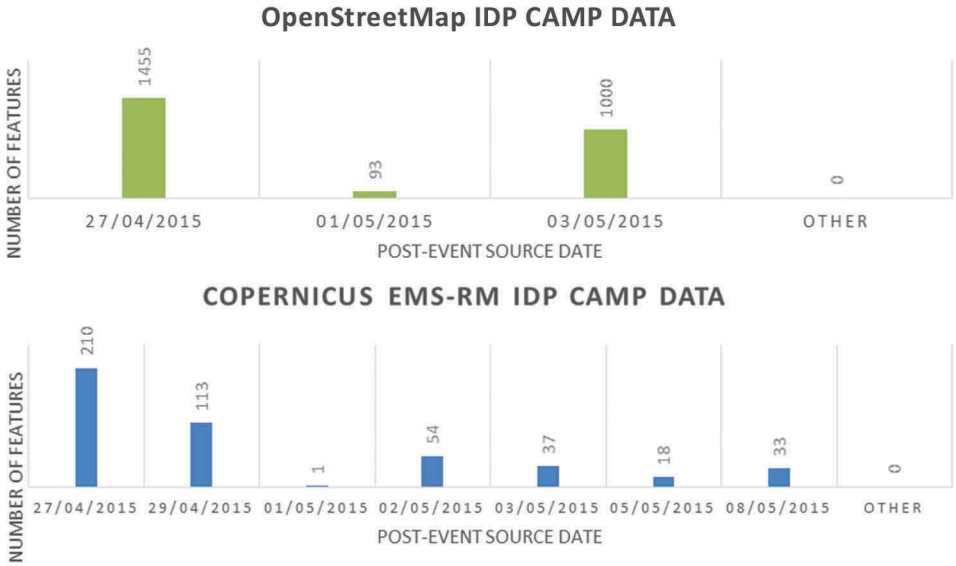


Figure 2. Distribution of date of acquisition of imagery used by OSM volunteers and Copernicus EMS-RM professionals for digitizing IDP camps.

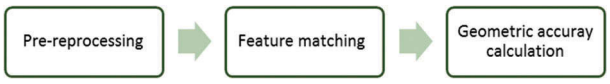


Figure 3. Workflow for the geometric evaluation of data.

considered as critical. Consequently, for line features, the length of the OSM objects that are included in the buffer around corresponding Copernicus EMS-RM dataset is meaningful of geometrically overlapping features.

With this approach, only the geometric accuracy of OSM data is evaluated compared to Copernicus EMS-RM standards; this limitation is due to the lack of a declared geometric accuracy value in OSM data.

A critical aspect in the geometric evaluation is the potential difference in source satellite imagery (e.g. satellite sensors) and geo-referencing process of the two compared sources (e.g. orthorectification). A geometric shift between images used by professionals and volunteers generates a consequent displacement between the sources, which is not necessarily a sign of geometric inaccuracy.

Thematic evaluation

The thematic evaluation of the two datasets is based on the comparison of the number of attributes relevant in the emergency mapping context available for describing features in Copernicus EMS-RM and in OSM datasets and on the comparison of their actual reported values. It is not possible to evaluate the thematic accuracy of the available attributes, in terms of correctness of the information reported, due to the absence of ground truth data.

Completeness evaluation

For evaluating the completeness of datasets, two different approaches are implemented to obviate the lack of ground-truth data:

a. Visual control using high-resolution satellite imagery: an additional satellite imagery interpretation is considered useful for the protocol application because it is carried out outside the emergency response cycle and, therefore, is a less stressing condition.
b. Alignment of datasets taking into consideration : the availability of other data from other satellite emergency mapping services (e.g. Tomnod, NGA and UNOSAT) (NGA, 2016; Tomnod, 2016; UNITAR, 2016) used to assess the convergence of crowdsourced information from the spatial point of view; geotagged pictures and reports collected by the activities of other crowdsourcing services (e.g. Standby Task Force (SBTF)) (Standby Task Force, 2016a) used to ground-truth the datasets and to create on-ground checkpoints.

Results and discussion

The application of the ELF protocol of data quality comparison allows us to first assess that the wide contribution of the crowd in producing data is directly translated in the availability of datasets with a large amount of digitized features. Table 1 compares the number of features (and the length of features in the case of the road network datasets) found in each dataset produced by the two services (number of features evaluated on shared AOIs). Since many Copernicus EMS-RM maps contain OSM data as a source, Copernicus pre-event data were preliminary filtered using the attribute related to the data source in order to not include in the analysis features digitized based on OSM data itself; this step allowed us to avoid misleading results caused by a potentially wrong source attribution (e.g. source attributed to Copernicus EMS-RM when data come originally from OSM).

On average, the age of satellite imagery provided to OSM contributors is the same as that of the imagery provided to professionals in Copernicus EMS-RM activations. Figure 2 shows the results of the update analysis of imagery used for IDP camp digitization. It is noticed, though, that often pre- and post-event imagery are temporally distant up to 2 years. On average, the thematic evaluation shows a larger number of attributes in volunteered data.

Of course, the amount of data is not a direct synonymous of their quality. Indeed, two different outcomes are obtained for pre-event and post-event datasets, especially regarding the geometric accuracy and the completeness indicators. The following sections list the major outcomes of the evaluation of these indicators on the two subsets of data.

Pre-event data

The application of the protocol on the road network and buildings datasets leads to the following considerations:

a. 71% of OSM road network features, in terms of length for the road network dataset, can be considered geometrically overlapping with Copernicus EMS-RM features.
b. The contribution of volunteer mappers in creating pre-event data is valuable in terms of thematic added values compared to Copernicus EMS-RM professionals. In fact, the thematic evaluation of data revealed that volunteers' data are enriched with more thematic attributes compared to professional data, including information such as road and building name and function, road material, and number of building floors, which may result in being fundamental in emergency response and a definitive added value to the geolocation of the road or building itself. This additional information is actually included for 20% of buildings OSM features and 65% of road network OSM features.

These added values potentially derive from local knowledge of OSM contributors or their possibility to gather information from social media.

Post-event data

The application of the protocol on IDP camps and damaged infrastructure datasets leads to the following considerations:

a. Post-event datasets appear extremely divergent among OSM and Copernicus EMS-RM in terms of quantity of identified features. As shown in Table 1, on the same AOIs, OSM volunteers identified 1265 damaged buildings, while Copernicus EMS-RM identified a total of 625 damaged buildings. OSM mapped a total of 4171 IDP camps, while Copernicus identified 466 features over the same AOIs. This leads us to observe that, from the point of view of the end user, it is not possible to indifferently refer to one or the other dataset without incurring significant differences.
b. Since there is a great discrepancy in the number of features belonging to OSM and Copernicus EMS-RM post-event datasets, a sample control on satellite images is performed. Starting from the damaged infrastructures dataset, by looking at the satellite imagery used for tracing the features, some building blocks appear as affected but some others do not show any clear evidence of damage even if digitized as damaged in OSM. As evidence, Figure 4 illustrates a village in the northern area of Trisuli-Valley in pre-event (Bing 2013) and post-event imagery (Pleiades 2015–05-03, CNES, Airbus DS) (OSM Tasking Manager, 2015a).

   This figure shows the imagery as provided in the OSM open platform (OSM Tasking Manager, 2015a). On this area also Copernicus EMS-RM worked on post-event imagery acquired on the same date (Copernicus EMS, 2015c). This portion of imagery has been digitized as damaged only by OSM contributors, while Copernicus EMS-RM did not assess it as affected.

   During the analysis of the satellite imagery used by volunteers, doubts arise on the correctness of their visual interpretation on portions of the area analysed. A much larger number of damages are found by volunteers with respect to professionals; some of them are not evident on the satellite imagery. This missing evidence brings us to think that a quality control step should be included in the volunteers' mapping standard workflow.
c. Nonetheless, after filtering out the elements with scarce evidence on the images, a relevant number of useful post-event features are found. One example of this is illustrated in Figure 5. In a sub-area of Kathmandu, both services recurred

Table 1. Number of features and length of features (in the case of the road network dataset) and extension of shared AOI of selected datasets, respectively, in OSM and Copernicus EMS-RM.

Dataset	Type of data	Shared AOI extension	HOT/OSM	Copernicus EMS-RM
Road network	Pre-event	7678 km^2	15,552 km	6289 km
Buildings	Pre-event	104 km^2	55,373 features	74,899 features
IDP camps	Post-event	2167 km^2	4171 features	466 features
Damaged buildings	Post-event	2570 km^2	1265 features	625 features

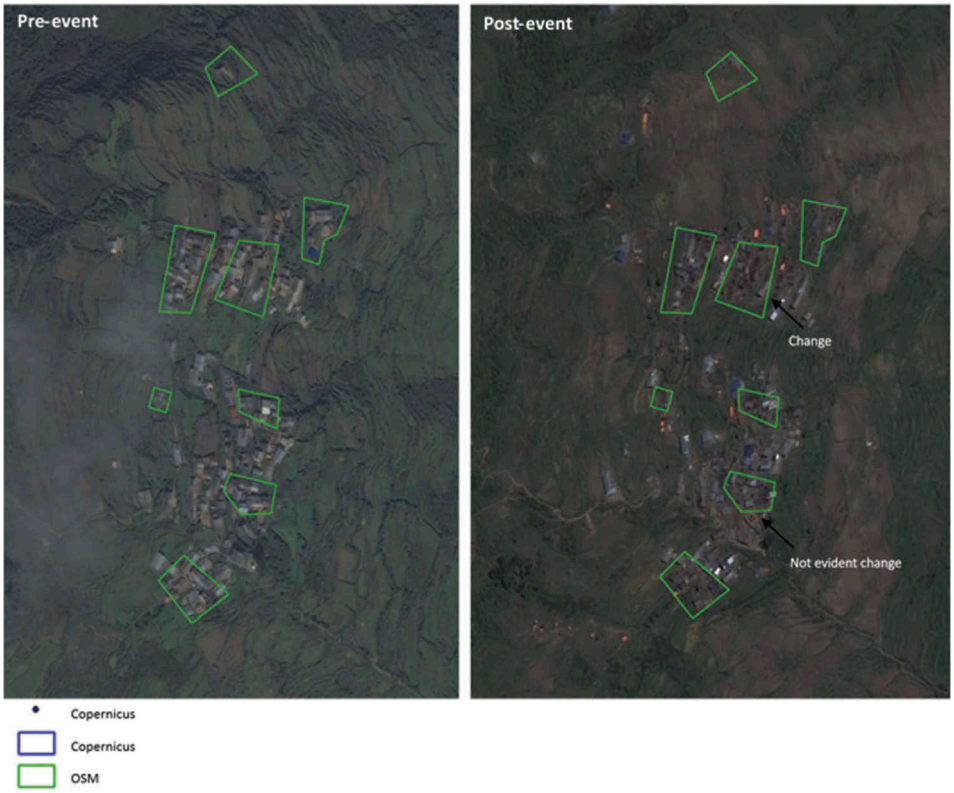


Figure 4. North of Trisuli-Valley. Data from Copernicus EMS-RM (blue) and OSM (green) representing damaged infrastructures on pre- and post-event satellite images.

Source of pre-event image: Bing 2013. Source of post-event image: Pleiades 2015-04-27, CNES, Airbus DS.



Figure 5. Sub-area of Kathmandu. Data from Copernicus EMS-RM (blue) and OSM (green) representing IDP camps on pre- and post-event satellite images.

Source of pre-event image: Pleiades 2014–11-29, CNES, Airbus DS. Source of post-event image: Pleiades 2015–04-27, CNES, Airbus DS.

to post-event imagery acquired on the same date, 27 April 2015, but different sensors (Copernicus EMS, 2015d; OSM Tasking Manager, 2015b). This image illustrates that the same theme is represented differently. Copernicus EMS-RM traces mostly big aggregations of tents, while single or small aggregations of tents are identified by OSM volunteers. In addition, OSM volunteers mapped more features compared to Copernicus EMS-RM, which trace evidences of IDP camps.

d. Overall, the visual control on satellite imagery allows us to hypothesize that volunteer mappers act with an overrepresentation approach compared to the professional mapping behaviour. Volunteers seem to produce datasets with very few omitted features but that include a percentage of commission error; professional products show potentially omitted features, as shown in Figures 4 and 5.

   This is likely due to the fact that professionals must work on clear evidence. Professional mappers have an interpretation handbook in which standard cases are depicted and used as guidelines. In addition, a thematic accuracy equal to 85% is declared on every Copernicus EMS-RM product; in this case, it means that no more than 15% of interpretation errors in detecting damages are present on the map, according to the producers.

e. The quality of satellite imagery, especially of post-event satellite imagery, provided to OSM contributors is sometimes low with respect to the analysis that is required to be carried out, as shown in Figure 4 (OSM Tasking Manager, 2015a) A well-defined building footprint in a pre-event image and an in-shade building in a blurry post-event image may lead to consider the building as collapsed.

   This underlines how the resolution and quality of satellite imagery provided in order to digitize structural damages to buildings is fundamental. This is especially valid in peculiar morphological areas (hilly, mountainous) and a peculiar type of damages occurred, for instance the simple flattening of small houses, hardly recognizable from a satellite image. This makes the damage even harder to detect.

   Detecting post-event features, especially structural damage, still requires field validation or at least a higher resolution of imagery or a different approach, such as drones or low-altitude acquisition. In addition, currently available tools provided in VGI platforms that enhance brightness and contrast of the satellite imagery do not work on the maximum radiometric resolution of the imagery; this aspect could be improved in the effort of helping volunteers.

f. The media effect, hypothesized already by REACH in 2014 (REACH, 2014), probably influenced the action of volunteers in the Nepal activation. In fact, villages that were mentioned in many social media and channels as highly affected by the earthquake present the highest number of digitized features. Furthermore, in some satellite images that do not seem to highlight any damage but that were actually affected, damaged buildings were identified by OSM contributors, probably brought there by information shared by the media or geotagged pictures shared through other crowdsourcing platforms, such as the SBTF (Standby Task Force, 2016b), as shown in Figure 6 (OSM Tasking Manager, 2015c). In this figure, the satellite imagery provided to OSM volunteers does not clearly show any damage; nonetheless, several damaged buildings have been identified by them; this amount of damage was also identified by the SBTF by means of geotagged pictures (Standby Task Force, 2015).

g. The presence of geotagged pictures and videos can help in the process of digitization of post-event features. They can act as ground-truth points to verify the data (Ajmar et al., 2011) and they can be used as event-specific checkpoints to provide visual guidelines for both volunteer and professional mappers. Hence, this practice can be encouraged in order to produce more numerous geotagged photo and video datasets in future events and also be included into the OSM data model with a dedicated tag.

Interaction of the services

The introductory remarks and the results of the data comparison highlighted two main issues: the need for a quality check of crowdsourced data and the duplication of efforts within the existing emergency mapping activities.

An interaction of the activities and data sharing between professional and volunteer mapper communities is seen as a possible resolution to the stated set of problems; in addition, a cooperation can be of mutual support when considering the time constraints in RM.

A cooperation between the two groups of mappers is expected to be beneficial under several aspects, according to the specific needs of the emergency activation, the type of dataset required and its complexity. Some of the proposed gains are listed below:

a. HOT/OSM can contribute in the creation of datasets in order to fasten the work of professionals in RM activities, once their quality is



Figure 6. Village of Arughat in two geotagged pictures collected by the SBTF (REUTERS/Athit Perawongmetha). The damages were identified only by OSM contributors, even though the satellite image provided to OSM volunteers seems not to be helpful at the task.

Source of post-event image: Digital Globe, 2015–04-29.

checked and validated; this contribution would also reduce duplication of work and efforts.

b. Dataset from Copernicus EMS-RM can be enriched by OSM contributors with additional information hardly traceable in the absence of local knowledge or support of other social media sources. For instance, building and road names and functions can be added by OSM volunteers to grading datasets generated by a professional team.
c. HOT may not have immediate access to specific satellite imagery and activate Copernicus EMS-RM. In this scenario, the HOT/OSM community would not have access to the imagery made available to Copernicus EMS-RM for the previously stated licence issues. The same does not apply to the data produced, which are distributed with full and open access.
d. OSM volunteers may not have the adequate skills to respond to some of the mapping project requirements (e.g. evaluation and grading of damaged infrastructures) and activate Copernicus EMS-RM for specific tasks.
e. The contribution of Copernicus EMS-RM may be requested by HOT in order to carry out a quality check on OSM dataset; Copernicus professionals would then get the role of validators in the HOT/OSM workflow.

An efficient interaction of two different and independent services responding to separate workflows requires an in-depth work of coordination and harmonization. Requirements that are expected to be fundamental for the cooperation to work are exposed here:

a. The time of activation and response in mapping activities of the two services have to be comparable in order for the respective activities to fit in mutual original workflows.
b. Framework and guidelines have to guarantee both a successful interaction and a homogenization of the work of volunteers and professionals. As shown by the comparative analysis, it may occur that the two groups represent the same theme (IDP camps) differently (individual units versus bigger group of tents). This can to be avoided through pre-defined standards.
c. As previously discussed, Copernicus EMS-RM and HOT rely on two different mechanisms of provision of the imagery. This aspect can have a negative impact on the temporal coordination of the services, since the actual mapping activities begin when satellite imagery is acquired and received. The interaction of activities is expected to be efficient if both groups will work on imagery that can be considered coher-

ent in terms of both delivery time and acquisition date. Answer to this requirement is not provided in this study.

Temporal trends in the response of volunteers

Time is essential during a humanitarian emergency, especially for those services that must deliver data and information in a rush mode. This stresses the need for evaluating the response time of volunteers; as a preliminary requirement for a potential interaction, response time of volunteers must be of the same order of magnitude of the response time of professionals.

The IDP camp dataset generated in OSM is taken as a sample for the analysis; with dataset it is meant the totality of IDP camp features generated by HOT/OSM activation for Nepal earthquake, at first including also the area not covered by Copernicus EMS-RM activation. In order to assess the temporal trends in volunteers' response, the time lapse of response of volunteer mappers is evaluated. The time lapse of response is evaluated by analysing the time that was necessary to create the feature starting from the moment when the dataset creation task has been published. This allows assessing the actual time of response of OSM volunteers in absolute terms, calculated as the difference between the time of last modification of the single OSM feature and the time of creation of the specific task to which the feature belongs. In order to evaluate the actual speed of data production of volunteers, the only data that were not modified since their first creation are taken under analysis (first version data). Figure 7 shows a distribution chart of response time in the creation of the OSM IDP camp dataset.

It is observed that 9% of all considered features have been created within 24 h from the publication of the respective task. Even though the overall time of response of volunteers is long compared to the response time of professionals, this 9% of the dataset includes approximately 800 features. The numerical contribution of volunteers (230 OSM volunteer users collaborated in the creation of the IDP camp dataset) ensures that even only the 9% of the dataset includes an important number of features.

The evaluation of the time lapse of response is repeated, considering exclusively those areas covered also by Copernicus EMS-RM for Nepal; it is evaluated that OSM volunteers generated approximately 400 features within a time lapse of response of less than 24 h on shared AOI with Copernicus EMS-RM activation. The whole Copernicus EMS-RM activation produced 516 IDP camps features. The graph is shown in Figure 8.

After these evaluations, the amount of data generated by OSM contributors in a time lapse within 0 and 24 h leads us to consider the time of response of the two services (professionals and volunteers) potentially of the same order of magnitude. Consequently, the interaction of the activities can be taken into consideration. To incentivize the overlap in response times of the two groups, volunteers can be provided with smaller areas of priority to be analysed and mapped.

The variability of volunteers' response time with the type of activation and mapped features must be considered. An additional insight could be provided by averaging time of response across multiple events. This analysis is not provided in this study.

Proposition for interaction: framework and guidelines

Establishing a mutually beneficial collaboration framework that includes clear guidance, management and structure could have the potential to mitigate the issues stated in this study, building trust and confidence between all parties involved in the emergency mapping process. One requirement for the interaction of volunteer and professional mapping activities is a comparable time lapse of response of

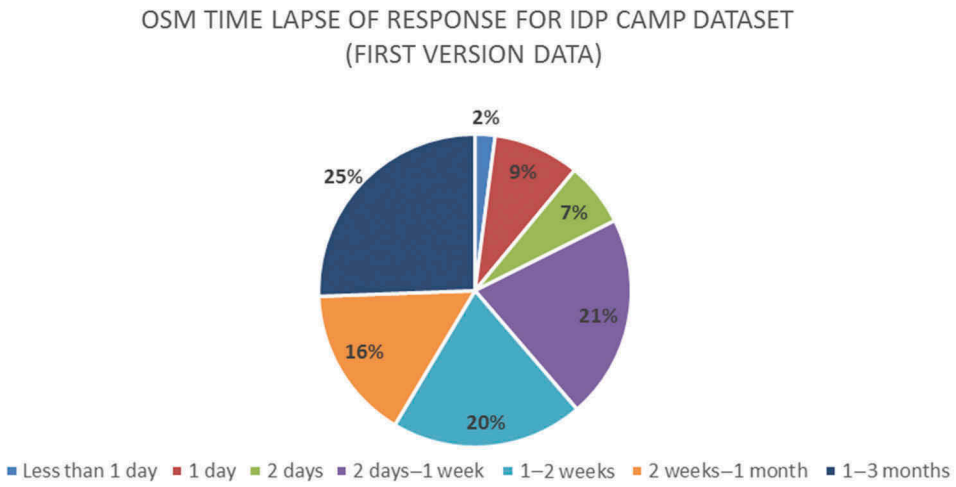


Figure 7. Pie chart showing the shares in time lapse of response of OSM contributors in generating IDP camp dataset. Considered features are exclusively ones whose version is the first.

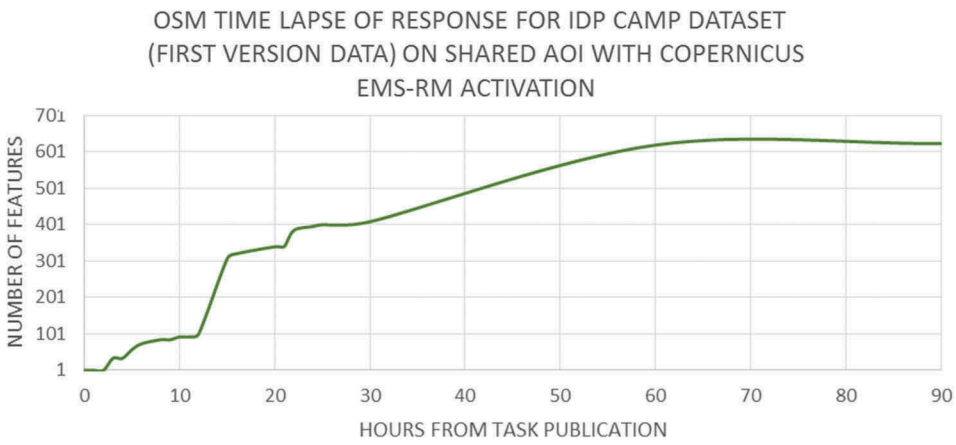


Figure 8. Time lapse of response in hours of OSM contributors on overlapping AOI with Copernicus EMS-RM activation for the production of IDP camp dataset.

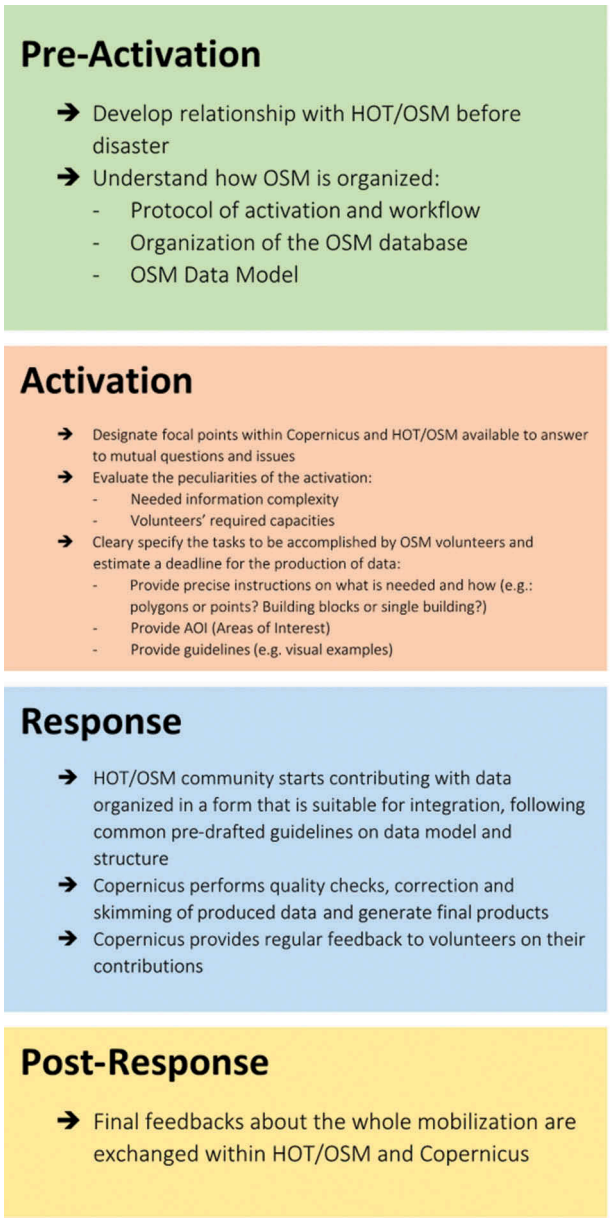


Figure 9. Four stages in the interaction framework between Copernicus EMS-RM and HOT/OSM.

parts involved; once this has been assessed, we propose a general collaboration framework that could be potentially applied in a future activation.

Figures 9 and 10 present the framework proposed for the interaction of the activities in a scenario where the volunteer mapper community is activated by the professional one. The proposed interaction framework is designed with a generalist approach, meaning that the type of products on which professionals and volunteers are called to collaborate in mapping is still left open.

When a disaster strikes, the pre-activation activities listed in Figure 9 must have been already implemented and ready to support the whole activation. Copernicus EMS-RM will receive guidelines on what products and where they are needed from external authorized users, hence the AOIs which Copernicus EMS-RM will be working on. Copernicus EMS-RM will then request the activation of the HOT/OSM community, providing them with the required AOIs and with specific guidelines on what to map and estimated delivery time required.

Another fundamental information exchange is related to the guidelines that must be provided to all parts involved in the activation and potentially previously drafted by the International Working Group on Satellite-based Emergency Mapping. These guidelines must contain common operational standards properly developed in order to standardize the work carried out by volunteers and professionals. Common standards should be referred also to the data model, in order to facilitate the integration of the datasets. This means that metadata and attributes associated with the features must be pre-defined and harmonized within the two datasets. Additional content of the guidelines is visual examples of features, possibly provided on improved satellite imagery (Figure 11) to facilitate the work of volunteers and professionals in identifying pre-event features or post-event features (e.g. damages, IDP camps, etc.) (Humanitarian OpenStreetMap Team, 2015).

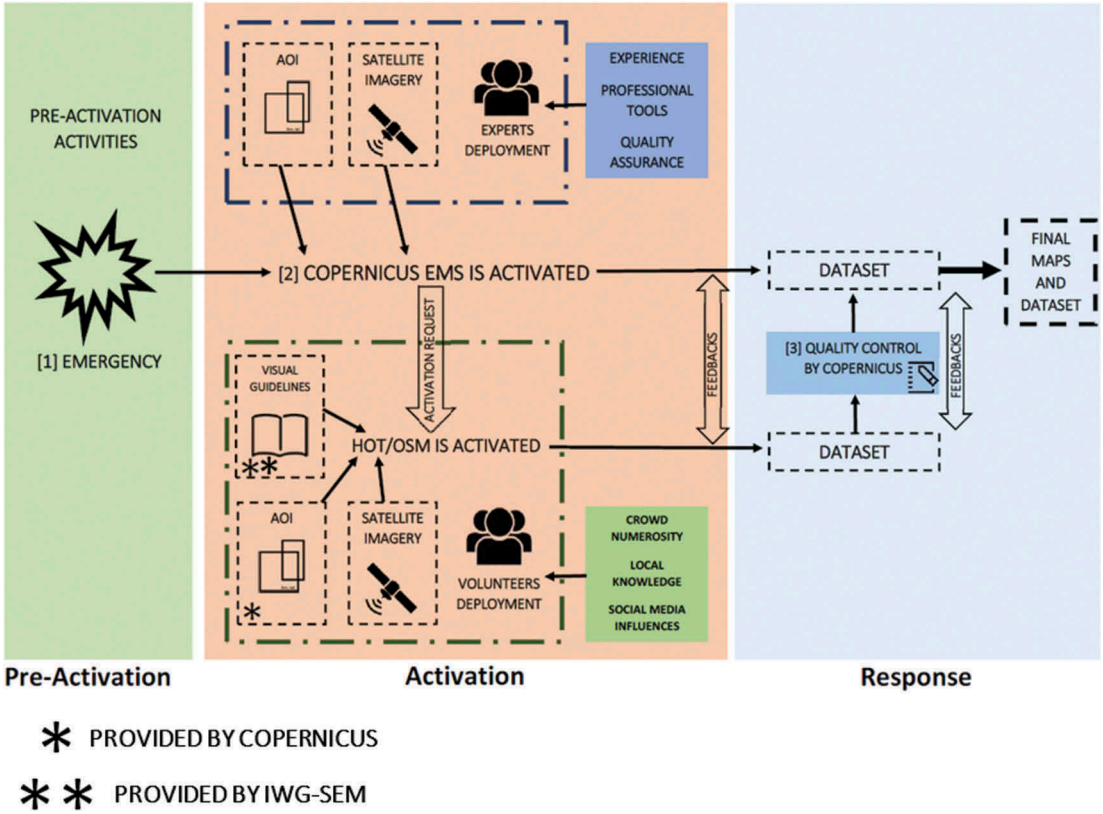


Figure 10. Scheme of the interaction framework between Copernicus EMS-RM and HOT/OSM.

A feedback exchange is suggested for the whole duration of the response. Once HOT/OSM community has contributed with their data, these are delivered to Copernicus EMS-RM that integrates them with the dataset already produced during its activation. Operation of quality checking and skimming of data can be performed to finally deliver the requested product, and a series of feedback about the whole activation are exchanged.

In the case of volunteers activating the professional mapper contribution (e.g. for quality check and validations or for mapping specific features/areas), the working framework is expected to be substantially the same, with some changes in the activation and response phase. In this scenario, the AOIs would be provided by HOT to Copernicus EMS-RM together with guidelines on the requested task.

The presence of built relationship between professionals and volunteers and common operational standards are two fundamental pillars in order to make the potential interaction framework effective and productive.

Conclusions and remarks

Rapid SEM supports the immediate response phase providing on-demand geospatial information in support of emergency management activities. In recent years, it has been possible to identify two main groups of mappers activated in the early response to major disasters: professionals and volunteers. Copernicus EMS-RM and OSM have been considered, respectively, as representatives of the two groups. Despite the widely recognized usefulness of GIS, the main issues of making data reliable and coordinating mapping activities avoiding overlaps in activities still exist.

The application of the ELF protocol for data quality assessment, in the case study of the earthquake in Nepal, allowed noting that OSM geospatial data are valuable. The wide contribution of the crowd in producing data is translated in the availability of datasets with a large number of digitized features. When the same imagery is made available to both groups of mappers, the number of features mapped is larger in OSM.

It was possible to verify that, in a large percentage of the data sample, the geometric accuracy of OSM pre-event data, in this case road network and building datasets, meets the standards of professionals' final products, hence being potentially replaceable. Post-event datasets, instead, revealed a more divergent behaviour between volunteer and professional mappers. We observed a numerical discrepancy within the two datasets and we hypothesized a different mapping behaviour of OSM contributors compared to Copernicus EMS-RM professionals. The latter maps with a conservative approach, since professional products have to respect accuracy thresholds. Volunteers seem to over-represent, producing



Figure 11. Example of improved satellite imagery.

datasets with very few omitted features but including a percentage of commission error. While checking the post-event datasets (damaged buildings) on satellite imagery, in many cases a clear evidence of the objects represented in the datasets could not be found in the images. Consequently, post-event

datasets raise the necessity for a quality assessment. At the same time, OSM post-event datasets (IDP camps) include a relevant number of features that may complement Copernicus EMS-RM datasets once verified, as it was suggested in the proposition for interaction. This divergent behaviour, especially for the remote damage assessment, highlighted the challenge that these features represent and the yet strong necessity for on-the-ground damage assessment and data verification of remotely sensed damage assessments, both in professional and in volunteered products.

Volunteered data were found to have two strengths besides geometry accuracy of pre-event data and number of features created. Volunteered data are enriched by the local knowledge of OSM contributors. Attributes such as road name or buildings' intended use could definitely be an added value to the geolocation of features in the emergency context. Furthermore, the so-called media effect, on the one hand, may catalyse the efforts of volunteers on media-covered area; on the other hand, it allows volunteers to map information acquired through media which are maybe not retrievable exclusively from satellite imagery interpretation.

The presence of geotagged pictures and videos can give good support to the process of detecting especially post-event features such as damages to infrastructures. Geotagged pictures can act as ground-truth representing points to verify the data; they can also assume the role of event-specific visual guidelines and checkpoints to be referred to during the mapping process. The crowdsourcing of geotagged pictures could be encouraged in order to produce more numerous geotagged photo and video datasets in future events; the possibility to include them in the OSM data model can also be investigated.

Another observation retrieved from the application of the protocol is related to the great importance of satellite imagery quality, especially in post-event feature mapping. The quality of the imagery provided to volunteers (e.g. resolution, look angle, contrast, etc.) is strictly related to type and size of objects that can be distinguished on it; the higher the quality, the smaller the damages that can be detected. Volunteer mappers may be helped in their contribution with tools that enhance the quality of satellite imagery by acting directly on the complete radiometric resolution of the imagery, improving contrast, brightness and enhancing details.

In the second section, the analysis of time lapse of response led us to observe a potentially equivalent mapping activity in terms of volume of analysis among professional and volunteer mappers, when taking into account the number of features that the crowd was able to produce in the first 24 h from the data production request, probably given the number of contributors in OSM.

All of these considerations created the basis for drafting a potential cooperation of volunteer and professional mappers, promoting collaboration and reducing the divergence within emergency mapping services; the cooperation of volunteers and professionals is also seen as an answer to the need for quality assurance of volunteered data and the time constraints in the emergency context. Both groups, according to the emergency-specific needs, could potentially trigger the interaction of the services and data. A service interaction would rely on a strong relationship between professionals and volunteers, built upon common operational standards and guidelines.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Piero Boccardo http://orcid.org/0000-0003-4565-7332

References

Ajmar, A., Balbo, S., Boccardo, P., Tonolo, F.G., Piras, M., & Princic, J. (2011). A Low-Cost Mobile Mapping System (LCMMS) for field data acquisition: A potential use to validate aerial/satellite building damage assessment. *International Journal of Digital Earth*, 1–21. doi:10.1080/17538947.2011.638991

Copernicus EMS. (2015a). *EMSR125: Earthquake in Nepal*. Retrieved from http://emergency.copernicus.eu/mapping/list-of-components/EMSR125

Copernicus EMS. (2015b). *Copernicus EMS user guide*. Retrieved from http://emergency.copernicus.eu/mapping/ems/copernicus-ems-user-guide

Copernicus EMS. (2015c). *[EMSR125] Dhunche: Delineation map*. Retrieved from http://emergency.copernicus.eu/mapping/ems-product-component/EMSR125_15DHUNCHE_DELINEATION_OVERVIEW/2

Copernicus EMS. (2015d). *[EMSR125] Kathmandu: Grading Map*. Retrieved from http://emergency.copernicus.eu/mapping/ems-product-component/EMSR125_01KATHMANDU_GRADING_OVERVIEW/4

European Commission. (2015). *Copernicus emergency management service mapping - Manual of operational procedures. Guidelines for EC services, service providers and authorized users. Version 1.1*. Retrieved from http://emergency.copernicus.eu/mapping/sites/default/files/files/EMS_Mapping_Manual_of_Procedures_v1_1.pdf

Harvard Humanitarian Initiative. (2011). *Disaster Relief 2.0: The future of information sharing in humanitarian emergencies*. Washington, DC: UN Foundation & Vodafone Foundation Technology Partnership.

Humanitarian OpenStreetMap Team. (2015). *Nepal Earthquake - Collection guidance*. Retrieved from http://hotosm.github.io/tracing-guides/guide/Nepal.html#about

Ivánová, I. (2007). *Modelling of the data quality in the spatial domain*. Slovak University of Technology in Bratislava, Slovakia. *Slovak Journal of Civil Engineering*, 2007/2 Pages 28 - 34. Retrieved from https://www.svf.stuba.sk/buxus/docs/sjce/2007/2007_2/file1.pdf

IWG-SEM. (2015). *International Working Group on Satellite based Emergency Mapping (IWG SEM) - Minutes of the 2015 spring meeting, held in Bonn on May 28/29, 2015 hosted by UN SPIDER, Germany*. Retrieved from http://www.un-spider.org/sites/default/files/20150528_IWG_SEM_SpringMeeting_Minutes.pdf

IWG-SEM. (2014). *Emergency mapping guidelines*. Retrieved from http://www.un-spider.org/sites/default/files/IWG_SEM_EmergencyMappingGuidelines_A4_v1_March2014.pdf

learnOSM. (2015). *HOT remote response guide*. Retrieved from http://learnosm.org/en/coordination/HOT-Remote-Response-Guide/

Meier, P. (2015). *Digital humanitarians: How big data is changing the face of humanitarian response*. New York, NY: Routledge.

NGA. (2016). *National Geospatial-Intelligence Agency*. Retrieved from https://www.nga.mil/Pages/Default.aspx

OpenStreetMap. (2017). *OpenStreetMap stats report*. Retrieved from http://www.openstreetmap.org/stats/data_stats.html

OSM Tasking Manager. (2015a). *#1044 - Nepal Earthquake, 2015, Severely damaged housing areas and IDP Informal camps, Trisuli Valley - Task #58*. Retrieved from http://tasks.hotosm.org/project/1044#task/58

OSM Tasking Manager. (2015b). *#1008 - Nepal Earthquake, 2015, IDP Camps in Kathmandu*. Retrieved from http://tasks.hotosm.org/project/1008

OSM Tasking Manager. (2015c). *#1024 - Nepal Earthquake, 2015, Severely damaged housing areas and IDP Informal camps, Gorkha - Task #273*. Retrieved from http://tasks.hotosm.org/project/1024#task/273

PennState. (2015). *Geospatial data quality*. Retrieved from https://www.e-education.psu.edu/geog482spring2/c5_p3.html

REACH. (2014). *Groundtruthing open street map building damage assessment – Final assessment report*. Retrieved from https://reliefweb.int/sites/reliefweb.int/files/resources/clay.westrope-28042014-065633-PHL_OSM-Damage-Assessment_Final-Report_To-submit.pdf

Standby Task Force. (2015). *SBTF Nepal EQ Needs*. Retrieved from http://www.arcgis.com/home/webmap/viewer.html?webmap=e887b1aec9d042d58cc8e533b1791d71

Standby Task Force. (2016a). *Standby task force*. Retrieved from http://www.standbytaskforce.org/

Standby Task Force. (2016b). *Nepal*. Retrieved from http://www.standbytaskforce.org/?s=nepal&submit=Go

Tomnod. (2016). *Tomnod*. Retrieved from http://www.tomnod.com/

Tonolo, F.G., Perez, F., & Steffenino, S. (2014). *Emergency mapping - Comparison protocol*. The European Location Framework.

UNITAR. (2016). *UNOSAT*. Retrieved from https://unitar.org/unosat/

Veregin, H. (1999). Data quality parameters. In P.A. Longley, M.F. Goodchild, D.J. Maguire, & D.W. Rhind (Eds.), *Geographical information systems, vol. principles and technical issues* (pp. 177–189). John Wiley & Sons, Inc. Retrieved from https://www.geos.ed.ac.uk/~gisteac/gis_book_abridged/files/ch12.pdf

Wired. (2015). *How Nepal's earthquake was mapped in 48 hours*. Retrieved from http://www.wired.co.uk/news/archive/2015-04/28/mapping-nepal-after-the-earthquake