

OpenStreetMap Infrastructure Mapping and Its Usage on Flood Impact Assessment Using InaSAFE in Surabaya

Anjar Akrimullah¹, Biondi Sanda Sima¹, Yantisa Akhadi¹

Abstract—*OpenStreetMap has become one of the open data sources used for humanitarian purpose in the last decade. This paper focus on mapping process and flood impact assessment using InaSAFE on exposure data in Surabaya provided by OSM. Key infrastructure mapping by of one of the OSM groups HOT ID in Surabaya has provided various lessons and benefits for Surabaya. These lessons learned are very useful to improve the quality of OSM infrastructure data intrinsic quality and data management, particularly data that related to key and vulnerable infrastructure in the field of disaster risk reduction.*

Keywords—*Administrative boundary, data duality, infrastructure data, OpenStreetMap.*

I. INTRODUCTION

Geographic information systems have been instrumental in various humanitarian activities in the last decade. Due to the internet, geographic information systems have also grown to bring community involvement in geographic data collection processes [1]–[3]. It is also called Volunteered Geographic Information. One of the VGIs that has provided great benefits to humanitarian purposes and various information in the world is OpenStreetMap [1]–[4]. OSM has been one of the best VGI in the world that now still continue to evolve and develop the system therein. One of OSM work methods is to bring together and collect individuals who are interested in providing free data for everyone. These individuals then become a particular group or team that also participates actively in the mapping of the world as well as certain locations considered by the coordinator of OSM priorities [5]. Groups like these have been scattered in different parts of the world carrying humanitarian missions in the mapping process and collaborating with other OSM contributors through various communication channels including mailing lists, discussion forums, and physical meetings [5].

One of the OSM group which is currently doing mapping in Indonesia is Humanitarian OpenStreetMap Team Indonesia (HOT ID). HOT ID mapping process is also done through cooperation with various parties, some of which are Pacific Disaster Center, BNPB, and BPBD. The author himself is one of the members of this team. HOT ID has mapped key and vulnerable infrastructure in the event of a

disaster as the source of humanitarian information and to support many studies in using OSM datasets [4], [6]. This mapping has been done in one of the metropolitan city which is Surabaya in 2016. The data has been collected and very easily accessible to anyone on the OSM website. This information definitely accessible for all stakeholder such as the government, academics, students, to profit and non-profit institutions.

The infrastructure mapping in Surabaya is also one action of preparedness in responding to the increasing of intensity and frequency of disaster events in Indonesia. Surabaya as the second largest city after Jakarta also has considerable exposure to certain disaster hazards, especially floods. The Surabaya city government's current efforts to respond to disaster risk reduction action have reached the establishment of the Surabaya Regional Disaster Management Agency (BPBD Surabaya). Therefore, the government of Surabaya with its BPBD in the future can use the infrastructure data mapped by HOT ID team for the benefit of disaster risk reduction.

In this paper, the authors also apply one of the uses of OSM data in the context of estimating the impact of the disaster hazards in Surabaya that is flood disaster. To be able to estimate the impact, the author uses one of the QGIS plugin InaSAFE which is currently being developed to assist stakeholders in Indonesia in conducting disaster risk reduction processes. InaSAFE on its official website promote that it is free software that produces realistic natural hazard impact scenarios for better planning, preparedness and response activities. It provides a simple but rigorous way to combine data from scientists, local governments and communities to provide insights into the likely impacts of future disaster events. A large scale effective disaster mitigation and preparedness process requires dynamic information exchange between multiple organizations and nations [7]. With this paper case example, it is expected to be one of the applicable models for more complex impact assessment cases for different hazard data.

By looking at the development of nature and human activities in Surabaya, in the future OSM together with volunteers and their groups that may come from Surabaya or not will continue to do voluntary mapping for the updating of existing data. This paper itself is a preliminary initiation for all urban development stakeholders in optimizing the benefits of infrastructure mapping by HOT

¹Anjar Akrimullah, Biondi Sanda Sima, Yantisa Akhadi are with Humanitarian OpenStreetMap Team Indonesia, 12820 Tebet, Tebet Timur, Jakarta Selatan, Indonesia. E-mail: anjar.akrim@hotosm.org; biondi.sima@hotosm.org; yantisa.akhadi@hotosm.org.

ID in 2016 in disaster risk reduction as well as the Surabaya City development planning process itself.

II. DATA AND METHOD

A. Mapping Process, Data, and Quality

In this paper, we use descriptive analysis in exploring mapping process dimension, comparative method in checking data quality, and multi data analysis in impact assessment using InaSAFE. Regarding with the infrastructure mapping process, pertaining to infrastructure mapping in Surabaya, logical consistency is the main issue generally evaluated in terms of spatial data quality by many studies in OSM research subject [1]. It is studied in lesser amount compared to other VGI researches [2] as lack of structured process, crowdsourced system also appears in the absence of logical consistency [8]. HOT ID team then form a structured mapping team regarding with the logical consistency issues. In the process of mapping which HOT ID done in 2016, team used some survey tools such as mobile phone, field paper, village map, map and survey board, motorcycle, ID card, stationery, power-bank, and some administrative requirements. The mapping was done to all sub-districts and sub-districts in Surabaya with details of 31 sub-districts and 154 sub-districts. The HOT ID team consists of 16 data entry specialists (DE) and 4 quality assurance (QA) and 2 supervisors. With the composition of the team and the survey equipment, the city of Surabaya is planned to be completed mapping process for approximately 5 months from October 2016. These team then conducts training for all members and volunteer in Surabaya in order to support in reaching goals. In addition, the process of mapping the infrastructure is also followed by mapping the boundaries of each village RW (community group administrative boundary). RW boundaries collecting is one of HOT ID effort to make an enrichment in administrative data for everyone in order to make a detail spatial unit for better data analysis in various subjects [9]–[11].

The list of survey objects mapped by the team consists of 11 main objects including administrative boundary while another 10 objects are the key infrastructure for disaster risk reduction. Hodgson et al. [12] and Jeon [13] also indicates these 10 objects as the main infrastructure in disaster response. The details of the objects mapped in the table 1 have passed the discussion process with the parties involved in infrastructure mapping activities in Surabaya such as Pacific Disaster Center and BNPB.

TABLE 1.
LIST OF SURVEY OBJECTS

Category	Objects
Finance	Banks
Communications	Communication Towers
Transportation	Airports, Ferry Terminals, Train Stations, Bus Stations, Roads, Railways
Water Supply Systems	Fire Hydrants, Water Towers, Pump House, Reservoirs, Water Gates, Water well

Electrical Power Systems	Generators, Towers, Substation, Power Plant
Fuel Storage	Gas Stations
Public Institutions	Kindergarten, Schools, College, Universities, Places of Worship, Supermarkets, Traditional Marketplace
Health Facilities	Clinics, Hospitals
Emergency Services	Police Stations, Fire Stations
Government Establishments	Governor Office, Townhall, Subdistrict Office, Village Office, Community Group Office, Ministry, Embassies
Administrative Boundaries	City/Regency, Sub-district, Village, Community Group

During the survey process, data entry specialists perform data collection and input directly into the OpenStreetMap under the supervision of quality assurance and supervisor. The inputted data is ensured by the quality assurance that has met the OpenStreetMap standard before the supervisor does random checking on the input result and the quality assurance validation result. Random checking itself is one of the procedures to ensure that OpenStreetMap data is in accordance with OpenStreetMap standards and in accordance with field data.

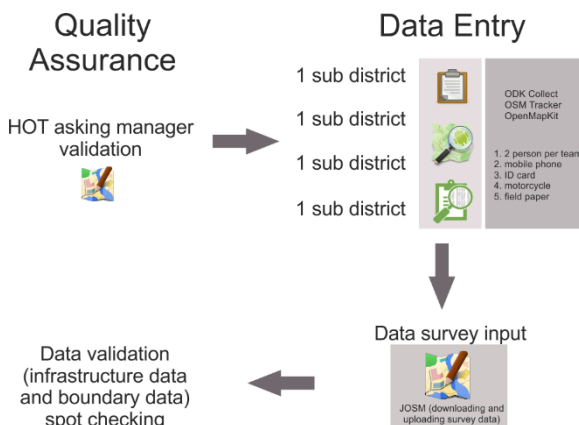


Figure 1. Structure of infrastructure mapping methodology in Surabaya.

In order to add the result of logical consistency, in this paper we also show several cases about logical consistency that we found while downloading several grid on recent OSM data in Surabaya. The data is downloaded three months after the project ended. As described in Hashemi and Abbaspour [2], the quality study in VGI comprises three categories, which are higher quality data source comparison, user activities-based [14], and data history-based. We use second and third categories as part of data intrinsic indicator [15], [16]. We use this method in this paper since the structured mapping process in team is the key differences between HOT ID team and the other contributors that affect data quality [8] in terms of logical consistency. We evaluate the quality by investigate the correctness of polygons [17] However, we also use the first method in checking the quality of administrative boundary data.

B. Impact assessment

The data used in assessing the impact of the flood event in Surabaya is the OSM Surabaya data downloaded in April 2017 using QuickOSM plugin in QGIS. Meanwhile, the data related to flood hazard is obtained from Surabaya Development Planning Agency. This data is a flood hazard mapping based on Musrenbang (government and community discussion in Sub District and Village level) results in 2016. Then, administrative boundaries data is from data collected by the HOT ID team on infrastructure mapping in 2016. The impact assessment is calculated with the help of the InaSAFE plugin in QGIS. By entering the necessary data such as physical object exposure data, administrative boundary data, and flood hazard data, InaSAFE will directly calculate the expected impact on existing physical objects.

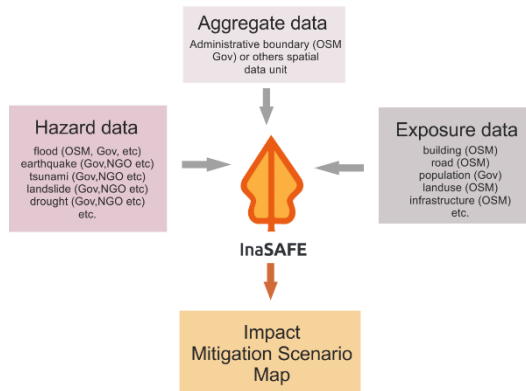


Figure 2. InaSAFE conceptual process in assessing impact.

III. RESULT AND DISCUSSION

A. Mapping Result, Process, and Procedure

As the second largest city after Jakarta [18], Surabaya is also the largest center of activity in the East Java province. Mapping of key infrastructure that took place in Surabaya was completed in approximately 5 months in 2016. This mapping group involved 23 individuals consisting of 16

data entry specialists, 4 quality assurance, 2 supervisors and 1 operational manager. The composition of this team is effective to map key infrastructure in all villages in Surabaya. This mapping process has provided a variety of experiences and lessons that are important to observe as to how infrastructure data should be managed in the field so that the quality and quantity are always in the proper standards for the good planning in the context of urban development planning and disaster risk reduction [12]. As for these lessons, the authors summarize the different sections below.

The growing number of population demands the government to continue in responding the accommodation of public facilities. This process certainly passes the assessment processes on the quality and quantity performance of public facilities that have been there before. To meet government needs related to the data of this public facility, the government usually through village and subdistrict official ordered them to prepare a report in the form of village or sub-district profile. This profile is mostly used by the government of Surabaya as the main source of public facilities planning in Surabaya city besides data sourced from BPS (statistic centre agency) in the form of *Kecamatan Dalam Angka*.

Infrastructure data is a vital data in urban development and regeneration with involvement in data management, analysis and estimation [19], [20]. Both the village profile, the sub-districts in the figures (*Kecamatan Dalam Angka*) and the others government archives that are sourced from various activities related to infrastructure data must have certain deficiencies both in quantity and quality aspects. The reality we found on the ground, the infrastructure data when compared with some secondary sources above, has a significant difference. Factors of data actuality, survey radius to input data errors can be the cause of significant differences between ground data and data contained in secondary documents. Table 2 shows 5 village data comparison from three data sources, village profile, BPS and HOT ID infrastructure mapping in 2016.

TABLE 2.
 INFRASTRUCTURE DATA SOURCE COMPARISON: VILLAGE PROFILE (VP), STATISTIC CENTRE AGENCY (BPS) AND OSM

	Keputih			Menur Pumpungan			Klampis Ngasem			Bangkingan			Menanggal		
	VP	BPS*	OSM	VP	BPS*	OSM	VP	BPS*	OSM	VP	BPS*	OSM	VP	BPS*	OSM
Mosque	11	11	10	9	9	5	8	8	7	5	4	4	4	4	4
Church	0	0	0	4	4	1	4	4	2	1	1	1	3	3	2
Clinic	2	-	2	5	-	6	3	-	2	2	-	1	1	-	1
Hospital	1	-	2	0	-	1	4	-	1	0	-	0	0	-	0
Kindergarten	15	22	9	17	15	8	9	30	12	8	-	3	8	8	5
School	12	-	9	14	-	9	12	-	8	2	-	5	5	5	7
University	3	3	3	4	5	3	4	5	2	0	-	0	0	0	1

*data was compiled in 2016 by Surabaya BPS staff using 2015 database. Access via Surabaya BPS website.

The discussion on the mapping results, processes and procedures in this paper begins by looking at table 2. Table 2 is five examples of village with the completeness of its OSM data in comparison with the extent of data sourced from the government apparatus ie BPS and Village Profile. The OSM data is data mapped by the HOT ID team in 2016 through a survey scheme, data input, and structured quality control. While Village profile data is data obtained from urban village staff when visiting Village office conducted by data entry team. BPS data is data obtained from the official website of BPS that can be downloaded openly by anyone.

More specifically related to the two comparison data, the Village Profile data is usually obtained by the Data Entry team through the village government staff. This data is updated every 3 months by the village staff. The substance includes information related to the general conditions in the Village. In this case, the HOT ID team only focuses on data relating to the amount of infrastructure in the village area of interest. The process of obtaining Village Profile data requires adherence to good administrative procedures. If it does not pass through the procedure, the Village will find it difficult to share data. However, not all urban villages are very adherent to such administrative procedures, under certain conditions, with good communication and negotiation with the Village staff, data relating to the mapping of infrastructure is obtained easily. If the Village staff has shared the expected data, the Data Entry team also needs to do some validation by asking directly with the Village staff about the facility locations recorded in the Village profile. Because of this, there are quite a lot of urban profile files that need to be updated because the conditions in the field with the data on the Village profile are not suitable. For example, if there are 5 primary schools in the Village Profile, then the urban village when questioned will remind that some primary schools have become one, or there are schools being moved or evicted. Then this will also be proven in the field through a survey conducted by the data entry team. This may explain one of the reasons why the Village Profile data that was last obtained was quite different from the mapped OSM data as in some of the cases listed in Table 2. In other cases, sometimes the data entry team found some facilities with slightly larger numbers Compared to the data in the Village Profile. Both of these are enough to reinforce the need for collaboration of data provision among stakeholders to improve data quality for disaster mitigation efforts, economic development, and general development.



Figure 3. OSM vs Village Profile infrastructure data in 5 villages.

The second data source that is the benchmark is the data published by BPS on its official website. Documents containing infrastructure data on the BPS website are only available in a document called "Kecamatan Dalam Angka". Until this website is accessed, the available data is data infrastructure that is processed based on 2015 data, while published by BPS in 2016. In the presentation of the data, not all data presented based on the existing Village in the Sub-District. School facilities and clinics, for example, in the document, information on the number of these facilities is presented in general to the total number of sub-districts, not divided by Village. Thus, the researcher was unable to determine the exact number of facilities required in this study to compare it with the data obtained by the HOT ID team. With regard to the limitations and weaknesses of BPS data, this is one of the things that researchers note as learning the importance of good management and data management. As one of the government's tools, BPS certainly has an important role in providing secondary data that can be used by the community, academia, private and government itself in conducting analyses and studies related to the development process in a region. With these roles, BPS should provide a data source that allows users to sort the necessary data. In this study, the researchers witnessed by themselves how the data obtained from BPS can not meet the researcher's aim in comparing the data with the results of the HOT ID team mapping. This is certainly a lesson learnt where the performance of government tools in data management both analysis and data compilation becomes a very important thing to note because of the quality of human resource constraints and some technical procedures in data management is still a major problem [21]–[23]. More broadly, authoritative data are increasingly out of date in many parts of the world [24], and were acquired using older technologies that were less accurate than those available to the general public today [25].

In relation to the mapping results of the HOT ID team in more depth, besides mapping the tags related to key infrastructure for disaster, the team also mapped out the building tag which is one of the baseline data to estimate physical exposure in an area of a disaster hazard. The total number of buildings mapped in Surabaya is 651,900 based on the calculation of the osm-analytics.org website. Both infrastructure and building data are part of spatial data with the help of computer technology development, both of which are crucial in making effective decisions in disaster mitigation efforts [7], [26].

The mapping of infrastructure in Surabaya City has standard data processing that is adjusted to the basic rules of naming in OSM Indonesia. The term also called "Indonesia Preset" is the labeling nomenclature or tags for various objects mapped in Indonesia. This tags then become the HOT ID team reference to map the various important objects that have been agreed with other relevant stakeholders. These objects are as listed in table 1 for approximately in 5 months, 4461 objects have been successfully mapped to the previous infrastructure of only 658 objects and 4593 km of roads with previously 3879 km.

TABLE 3.
DETAILS OF OBJECT SURVEY RESULTS IN SURABAYA IN 2016.

Infrastructures	Total		
	Before	Ground Survey	Validated
Bank	37	193	306
Communication Tower	0	227	267
Transportation			
Airport	0	0	0
Ferry Terminal	0	1	0
Train Station	4	4	13
Bus Station	3	9	13
Water Supply Systems			
Fire Hydrants	0	2	1
Water Tower	1	2	2
Pump House	1	30	35
Reservoir	0	1	0
Water Gates	0	10	12
Electrical Power Systems			
Power Tower	305	33	441
Substation	0	6	10
Power Plant	0	4	1
Fuel Storage			
Gas Stations	14	73	88
Education Facilities			
Kindergarten	3	464	576
Schools	57	854	956
College	7	32	45
Universities	29	64	46
Place of Worship			
Mosque	13	858	1120

Infrastructures	Total		
	Before	Ground Survey	Validated
Church	1	190	242
Pura	0	15	6
Vihara	0	6	18
Klenteng	0	4	12
Daily Needs and Logistics			
Supermarket	0	25	47
Market Place	10	71	56
Health Facilities			
Clinic	6	206	289
Hospital	31	58	54
Emergency Services			
Police	6	15	36
Fire Station	0	10	11
Government Establishments			
Governor Office	2	13	1
Townhall	5	4	1
Sub-district Office	0	33	30
Village Office	3	139	149
Community Group Office	0	371	412
Government Office	3	279	260
Embassies	0	3	6
Sport Facilities			
Stadium	7	7	3
Sport Centre	8	35	35
Pitch	102	110	480
Total	658	4461	6080

As shown in table 3, prior to mapping by the HOT ID team in 2016, the contribution of data from OSM contributors in Surabaya City is very limited, only 10% of total data has been validated by Quality Assurance. The data contained in the field survey field is the raw data obtained by Data Entry. The data is then validated by Quality Assurance to ensure its validity. So that the final data that has been controlled the quality is on the column validated data. Likewise, with road infrastructure, table 4 shows the differences in the length of each road type before and after the survey.

TABLE 4.
DETAILS OF ROAD LENGTH AFTER SURVEY RESULTS IN SURABAYA IN 2016

Roads	Before (m)	After Validated (m)
Motorway	71,801.50	58,566.25
Trunk	75,931.90	119,797.76
Primary	57,734.71	54,945.93
Secondary	299,571.29	285,178.27
Secondary Link	8,014.79	29,822.77
Tertiary	316,009.58	433,587.24
Tertiary Link	3,591.98	28,856.48
Service	182,142.13	198,910.55
Residential	2,825,591.48	3,101,238.13

Roads	Before (m)	After Validated (m)
Pedestrian	1,201.39	69.49
Path	14,021.50	12,608.69
Living Street	1,238.15	242,099.08
Track	1,134.04	2,076.94
Unclassified	21,566.82	25,253.28

All data either in the form of point objects, polygons or lines, has passed through structured processes and procedures to produce final data uploaded to the OSM website. All that starts from the formulation of coordination system between DE, QA, Mapping Supervisor and Operational Manager. The coordinating system established within the HOT ID team greatly stimulates all team members to continue to provide the best in order to generate open data for everyone. The principle built in this team is the principle of kinship so as to build good communication and will affect the expected results. Every month is always held evaluation related to team performance, constraints and other things to be conveyed related to the achievement of mapping targets.

In the HOT ID team, DE is a team member consisting of 2 members in each group that has a role in field data collection of both infrastructure objects and administrative boundaries. Prior to carrying out data collection in the field, DE will usually conduct an initial briefing to identify survey sites based on secondary data and satellite imagery maps available. When in the field, before collecting infrastructure data, DE is required to collect data on administrative boundaries of RW units at the Village office. Good negotiation and communication skills are required because some Villages have quite strict administrative compliance. If successful negotiation, DE is required to clarify several matters concerning administrative boundaries of RW units, and clarify the location and amount of infrastructure contained in the Village. Then checking the field can be done. In general, the process passed by the DE includes initial briefing, clarification, and discussion with urban village staff and field checks. After all that is done, DE performs the field data input process with the help of JOSM under QA supervision.

QA are some members of the HOT ID team who have a responsibility to ensure the quality of data that DE has inputted in accordance with OSM label nomenclature standards. QA performs active coordination with DE and Mapping Supervisor related to the process of achieving the target. Both QA and QA, each have a history of different experiences in contributing to OSM. If categorized by contributor categorization of Neis & Zipf [27], then almost all team members from the beginning of the mapping are hit-and-run, newbies and casual mappers. Some of the team members, besides Supervisors, have contributed to the level of heavy mappers. Training for DE and QA is one of the key factors in the implementation of the infrastructure mapping process to match the expectation. Through such training, DE and QA can know the minimum standards and targets expected in the mapping. Figure 4 and 5 shows the

activity of more than 1 contributor in one object in Surabaya OSM data. Any object mapped by HOT ID team can be ascertained as validated by QA or by a Supervisor directly.

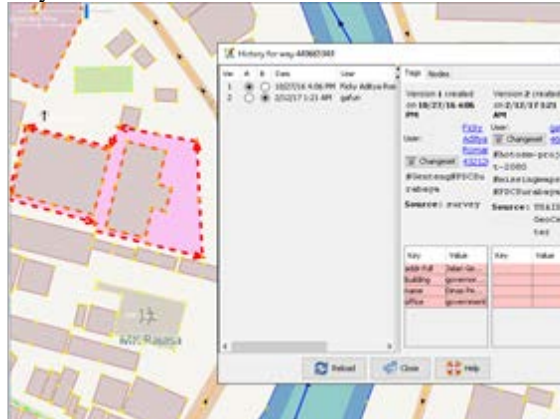


Figure 4. Government office building, input by DE and validated by QA.

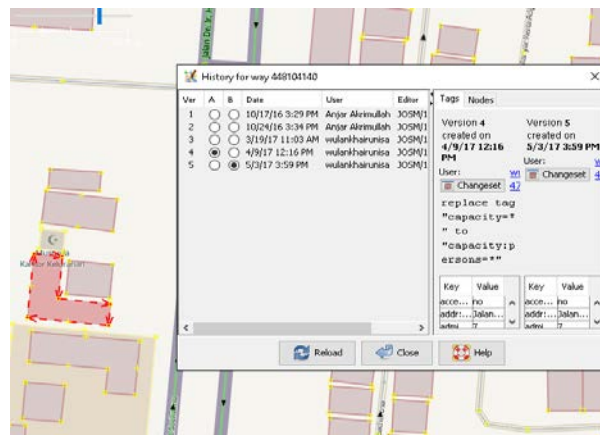


Figure 5. Government office building, input by DE and validated by Supervisor.

In this section, the authors provide an important record of the need for certain OSM-standardized procedures and frameworks standardized by OSM in maintaining the quality of OSM data from contributors. If the Surabaya project is successful because there is direct coordination between the members of the DE, QA and Supervisor teams, then for remote mapping with limited coordination and communication, special quality monitoring procedures are required and more specific that should be aimed at the contributors. At an adequate level of contribution and experience. Although international standards for spatial data quality assessment with ISO TC211 become the main guidance, specific procedures for the assessment of OSM data quality require a certain framework [2], [15], [28]. Further explanation regarding the quality of OSM data in Surabaya itself should be discussed in the another work as it suppose to be a particular research domain [17], [29].

B. Administrative boundary and its inconsistency

In table 1 related to the list of objects surveyed in infrastructure mapping activities in Surabaya, it can be seen that administrative boundaries are also a top priority in the field mapping process. Data related to these administrative

boundaries include the boundaries between urban villages and RW boundaries. When all urban village boundaries and boundaries are collected, it will be seen that the boundaries of the sub-districts and administrative boundaries of Surabaya City are based on reliable sources in Village Offices. The boundary mapping of RW was done by visiting the village office then requesting clarification regarding the initial boundary of the HOT ID team image on the survey map.

This clarification includes questions regarding whether the Village border on the map has been in accordance with the Village boundaries recognized by the Village official. Minimum of this clarification is done together with the secretary of Village and the chief assisted by governmental staff. Once completed with village boundary clarification, the HOT ID team also requests cooperation with the staff to map the RW brick on the map that the HOT ID team carries. The mapping of RW boundaries is quite varied, some village usually delegates this mapping to one of the staff who is considered to have long lived in the village and know with certainty the physical boundaries of each RW in the Village. Through a discussion process with other urban village staff and HOT ID teams, the RW boundary is then depicted with markers over the map until the entire RW boundary is mapped. This process is a fairly short process. However, in some Villages, the urban village staff volunteered to request the HOT ID team to map the RW boundary by tracing directly the physical boundaries in the field. This process occurs if the image quality of the map that the HOT ID team carries is less clearly visible by the Village staff, so the best option is to trace directly to the field and mark the boundary point of reference via smartphone. Figure 6 shows framework of RW Boundaries data collection in Surabaya infrastructure mapping.

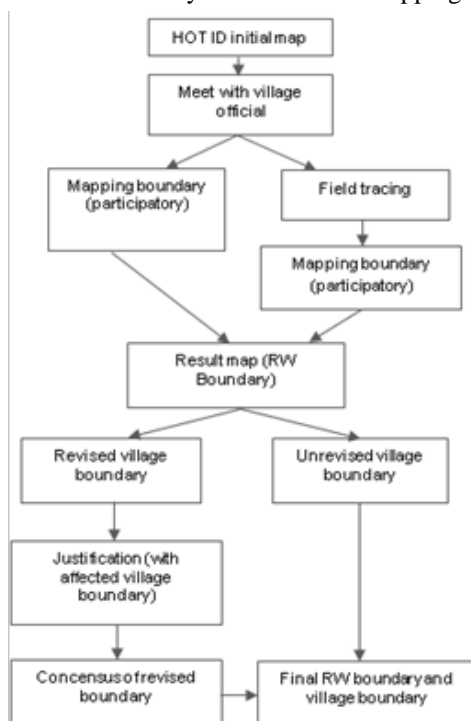


Figure 6. RW boundary mapping method in Surabaya infrastructure map.

When looking at the imbalance between the various sources of administrative boundaries of Surabaya City with a map of survey results timHOT ID, then the question is why these maps have different borders? When was the map created? How is the mapping method? Which map is used by official government agencies in city planning? Which map does the government use to issue building permits at the border area of the City? These questions will not be answered in this paper. The problem of administrative border inconsistency is not new in Indonesia. Confirmation of regional boundaries based on the smallest level of village or sub-district has become mandate of legislation. However, Indonesia's second largest city in the same class of Surabaya still has an inconsistent administrative boundary, what about other areas. Awareness of this if not really in serious follow-up will only become an obstacle in the development of the City. Just imagine, if there is a settlement development at the border of Surabaya by using the basic map of a particular source, but the fact on the ground, the area is partially or even claimed by the district administration or another city nearby. If there is no strict control and control, surely this can be a prolonged problem such as the kinship separator [30]–[32]. At present, it is not clear how the procedures and methods of affirmation of local administrative boundaries in Indonesia. Quite a lot of scientific writings on the affirmation of regional borders in Indonesia with various methods. Still, the official administrative borders issued by the government agency should be used instead of certain individual or group jobs not recognized by the official government body. Similarly, the mapping has been done by the HOT ID team in Surabaya, whether the result is usable or not the administrative boundary is used depends on the official government agency's acknowledgment.

Figure 7 shows the comparison of the completeness side, line geometry and data label of RW administration boundary to data obtained from the official government agency of Surabaya City ie Development Planning Agency. The comparison data has the level of detail of the spatial unit up to the RT limit. While the HOT ID team mapping data is only up to RW level. For label completions, some RW features in comparison data do not have labels. In addition, there is an area with more blank information on comparison data than OSM data. In Figure 8 it shows that for every RW administrative boundary in two village samples, there is a difference between the reference data and the data mapped by the HOT ID team.



Figure 7. RW Boundary in the same extent between government data vs OSM data

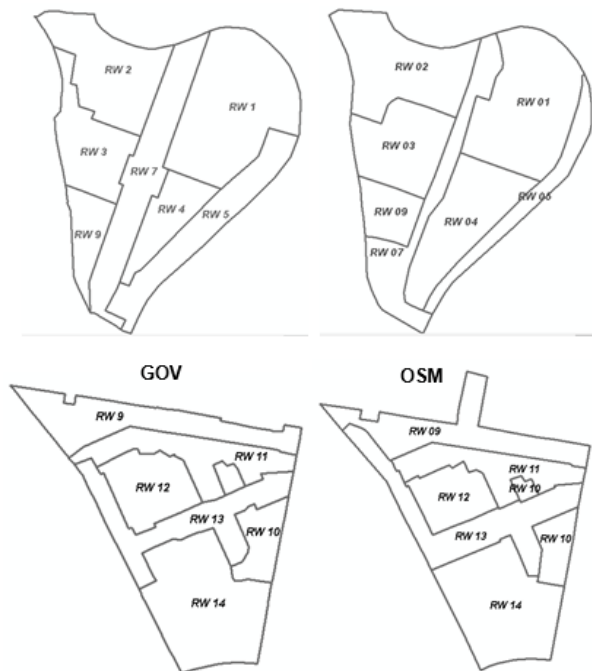


Figure 8. RW Boundary inconsistency between Government data and OSM data in two samples village.

However, all data on both the infrastructure and administrative boundaries that have been mapped in the city of Surabaya by the HOT ID team have passed through data collection and data validation processes that are structured and well organized. If problems arise regarding which administrative data should be used, so the authors, in this case, provide recommendations for users of OSM data as well as other users in general, in order to as much as possible using data that has been recognized by the official government, meanwhile, referring to Goodchild and Lina statements [25] related to actuality of Government data and

other issues, data OSM should be used as complementary data. OSM data is very flexible to use through collaboration with other data sources. Inconsistencies will always be found in every spatial data since the principle of comparing data with one another requires us to accept the flaws and advantages of each comparable data.

C. Logical consistency issues

Logical consistency is highly correlated with positional errors or topological aspects [2], and the validity ranges of values that occur in the data set [33]. There is no indicator to measure it quantitatively [14][28] but in this paper we evaluate the quality by investigate the correctness of polygons [17]. Based on the result of the investigation of the writer as one of Quality Assurance in HOT ID community, the percentage of polygon correctness in OSM data of Surabaya City is 99.70%. Meanwhile some of the errors found in general related to the completeness of the polygon and a mismatch between the geometry of the polygon with the physical appearance of the building in the imagery. Figure 9 and 10 show the data layer that the author downloaded and some examples of the polygon are not quite right.



Figure 9. Samples of correct polygon.

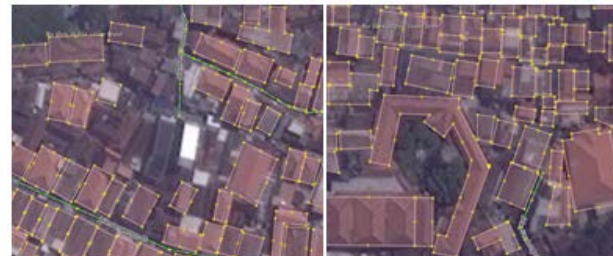


Figure 10. Samples of incomplete building area and incorrect building geometric.

D. Flood impact assessment

This section is one example of using OSM data in the context of disaster mitigation. With the help of the InaSAFE plugin in QGIS, the authors combine several types of data needed to be analyzed automatically to

produce certain information. InaSAFE in this case plays a role in answering the question of how many structures will be affected if a disaster occurs. The data used are data on disaster hazard originating from Surabaya Development Planning Agency, exposure data or building data sourced from OSM, and administrative boundary data sourced from HOT ID team mapping. The author has not yet explored the application of OSM data for use with other analytical purposes in InaSAFE in this paper. For example, analyzing affected population, damaged roads as well as other more in-depth matters as this paper essentially aims to show an example of the overview of OSM data usage in disaster risk reduction contexts.

Furthermore, InaSAFE in principle, requires at least three types of data to be able to assess the impact of a disaster ie, data hazard, exposure, and aggregation data. In the case of Surabaya, the writer used the data hazard of flood disaster because this disaster is one of annual disaster in Surabaya. In addition, the availability of hazard data obtained from the government is also the main reason. Hazard data is the result of government discussion with the community.

After performing the process input and analysis by InaSAFE, from 669,000 buildings, a total of 33,900 buildings affected by flooding. based on high category of inundation, 13,900 buildings were affected by high category, medium 18,900, and 1,200 low. The building of educational facilities is the largest affected with a total of 210 buildings. In addition, educational facilities are affected, mostly in the high flood category. Other public facilities that have significant impacts are facilities of worship and government buildings.

TABLE 5.
ESTIMATE NUMBER OF BUILDINGS.

Hazard Zone	Count
High	13,900
Medium	18,900
Low	1,200
Use Caution	631,000
Total	669,000

TABLE 6.
ESTIMATE NUMBER OF BUILDINGS BY STRUCTURE TYPE.

Structure Type	Affected			Not affected
	High	Medium	Low	
Residential	10	10	0	660
Education	110	90	20	3,900
Health	20	20	0	640
Transport	10	10	0	60
Place of worship	50	40	10	1,400
Government	50	40	10	1,300
Commercial	20	20	10	560
Recreation	10	0	0	10
Public facility	10	10	10	90
Other	13,700	18,800	1,100	622,000
Total	13,900	18,900	1,200	631,000



Figure 10. Structure affected by flood event in Surabaya.

From the results of this analysis, the authors can conclude that the city government can immediately take preventive action for disaster mitigation on public facilities exposed to inundation, specifically educational facilities, worship and government offices. Action scenarios can certainly be an increase in the quality of facility resistance not only functioning for normal conditions but also should be able to work in emergency conditions when disaster strikes [34].

IV. CONCLUSION

Mapping of key infrastructure related to the disaster will continue to be done by HOT ID team in several cities in Indonesia. Surabaya as one of the pioneering City provides various learning in process dimension and quality of OSM data. As an open data provider, OSM with the VGI concept has involved many contributors who actively contribute both remotely and coordinated by a community in a structured way. The mapping process in Surabaya is one form of actively engaging the community, in this case, the HOT ID team in a structured mapping process. Challenges and problems in the mapping process are fundamentally related to a complex system of government administration. This affects the increasing difficulty with access to government data. The authors indicate this as one of the causes of the data of government that is less actual. It has been mentioned in some cases that inconsistencies in both infrastructure and administrative data are among the authors of this paper.

From the mapping results obtained, a structured and coordinated mapping system in Surabaya City has provided an example of evidence of increased logical consistency with such a system. The truth of polygon geometry reaches 99.70%. While other parameters like data completeness is also always a deficiency in all spatial data. Meanwhile, the mapping of administrative boundaries in Surabaya indicates a difference between the borderline version of the

central government and village official. While we all know, in the context of city development, a spatial data processor performs spatial analysis for the benefit of future space utilization plans, administrative boundaries is an important thing needed. With the delineation of administrative boundaries, city planners will know which areas can be utilized for development purposes, including the border areas. If between the central government of the City and its village official there is a misalignment of perceptions about the boundaries, then it is clear there will be land disputes with other regions. Such findings have not been much noticed, whereas land disputes between blocks and land plots of the community are so complicated, not to mention land disputes between different regions of the city administration system.

Apart from some of the above lessons, OSM data wherever it is very open for use in various analyses and specific studies. The utilization of OSM Surabaya data in this paper is intended as an initial initiation for governments, academics and practitioners to broaden and use the data of OSM Surabaya for various analytical interests in various fields. In this paper, the authors provide an overview of field facts that the flood hazard in Surabaya City has a dominant impact on education facilities, place of worship, and government offices. Educational facilities are the most impacted so the authors recommend that these facilities be upgraded to the minimum standards in order to respond to flood disaster events. That is, the government needs to take certain steps to ensure that the education facilities will function normally not only when there is no disaster, but when the government floods are ready with all its efforts to make this facility function normally.

ACKNOWLEDGEMENT

This work and paper publication was financially supported by Humanitarian OpenStreetMap Team Indonesia (HOT ID).

REFERENCES

[1] J. Jokar Arsanjani, A. Zipf, P. Mooney, and M. Helbich, "An Introduction to OpenStreetMap in Geographic Information Science: Experiences, Research, and Applications," in *OpenStreetMap in GIScience: Experiences, Research, and Applications*, J. Jokar Arsanjani, A. Zipf, P. Mooney, and M. Helbich, Eds. Cham: Springer International Publishing, 2015, pp. 1–15.

[2] P. Hashemi and R. Ali Abbaspour, "Assessment of logical consistency in OpenStreetMap based on the spatial similarity concept," *Lect. Notes Geoinf. Cartogr.*, 2015.

[3] M. F. Goodchild, "Citizens as sensors: the world of volunteered geography," *GeoJournal*, vol. 69, no. 4, pp. 211–221, Nov. 2007.

[4] A. Rousell and A. Zipf, "Towards a Landmark-Based Pedestrian Navigation Service Using OSM Data," *ISPRS Int. J. Geo-Information*, vol. 6, no. 3, 2017.

[5] P. Mooney and P. Corcoran, "Analysis of Interaction and Co-editing Patterns amongst OpenStreetMap Contributors," *Trans. GIS*, vol. 18, no. 5, pp. 633–659, Oct. 2014.

[6] L. Palen, R. Soden, T. J. Anderson, and M. Barrenechea, "Success & Scale in a Data-Producing Organization," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*, 2015, pp. 4113–4122.

[7] T. Silva, V. Wuwongse, and H. N. Sharma, "Disaster mitigation and

preparedness using linked open data," *J. Ambient Intell. Humaniz. Comput.*, vol. 4, no. 5, pp. 591–602, Oct. 2013.

[8] M. W. Dobson, "VGI as a Compilation Tool for Navigation Map Databases," in *Crowdsourcing Geographic Knowledge*, Dordrecht: Springer Netherlands, 2013, pp. 307–327.

[9] G. Harper and L. Mayhew, "Applications of Population Counts Based on Administrative Data at Local Level," *Appl. Spat. Anal. Policy*, vol. 5, no. 3, pp. 183–209, Sep. 2012.

[10] M. Riva, P. Apparicio, L. Gauvin, and J.-M. Brodeur, "Establishing the soundness of administrative spatial units for operationalising the active living potential of residential environments: an exemplar for designing optimal zones," *Int. J. Health Geogr.*, vol. 7, no. 1, p. 43, Jul. 2008.

[11] M. Freedman, J. Lane, and M. I. Roemer, "New Approaches to Creating Data for Economic Geographers," *SSRN Electron. J.*, Jun. 2007.

[12] M. E. Hodgson, S. E. Battersby, B. A. Davis, S. Liu, and L. Sulewski, "Geospatial Data Collection/Use in Disaster Response: A United States Nationwide Survey of State Agencies," in *Lecture Notes in Geoinformation and Cartography: Cartography from Pole to Pole*, M. Buchroithner, N. Prechtel, and D. Burghardt, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 407–419.

[13] S.-S. Jeon, "Areas vulnerable to natural disasters and damage estimation of infrastructure in Busan, Korea," *J. Cent. South Univ.*, vol. 21, no. 4, pp. 1499–1507, Apr. 2014.

[14] J. Jokar Arsanjani, C. Barron, M. Bakillah, and M. Helbich, "Assessing the Quality of OpenStreetMap Contributors together with their Contributions," in *Proceeding of 16th AGILE Conference*, 2013.

[15] M. Haklay, S. Basiouka, V. Antoniou, and A. Ather, "How Many Volunteers Does it Take to Map an Area Well? The Validity of Linus' Law to Volunteered Geographic Information," *Cartogr. J.*, vol. 47, no. 4, pp. 3150–322, 2010.

[16] M. Van Exel, E. Dias, and S. Fruijtier, "The impact of crowdsourcing on spatial data quality indicators," in *Proceedings of GIScience 2011*.

[17] A. Vandecasteele and R. Devillers, "Improving Volunteered Geographic Information Quality Using a Tag Recommender System: The Case of OpenStreetMap," in *OpenStreetMap in GIScience: Experiences, Research, and Applications*, J. Jokar Arsanjani, A. Zipf, P. Mooney, and M. Helbich, Eds. Cham: Springer International Publishing, 2015, pp. 59–80.

[18] H. Santosa, "Environmental Management in Surabaya with Reference to National Agenda 21 and The Social Safety Net Programme," *Environment&Urbanization*, vol. 12, no. 2, pp. 175–184, 2000.

[19] Y. Yasuoka, "Monitoring of Urban Infrastructure and Environment by Use of Remote Sensing," in *Spatial Data Infrastructure for Urban Regeneration*, Tokyo: Springer Japan, 2008, pp. 41–51.

[20] Y. Asami, "Urban and Regional Information Infrastructure," in *Spatial Data Infrastructure for Urban Regeneration*, Tokyo: Springer Japan, 2008, pp. 1–14.

[21] M. M. Arif, H. S. Putra, and T. Kurrohman, "Analisa Kesiapan Pemerintah Daerah dalam Penerapan Standar Akuntansi Pemerintah Berbasis Akrua (Studi pada Pemerintah Daerah Kabupaten Bondowoso)," *J. Akunt. Univ. Jember*, vol. 11, no. 2, pp. 92–104, 2013.

[22] O. Y. Fujiyati and Sukadi, "Sistem Informasi Pengolahan Data Kependudukan Di Desa Purwoasri," *J. Speed – Sentra Penelit. Eng. dan Edukasi*, vol. 7, no. 1, pp. 1–8, 2014.

[23] R. Adawiyah, "Analisis dan Perancangan Sistem Informasi Pengolahan Data Penduduk pada Kantor Kelurahan Desa Pamali Kabupaten Bangka," *J. SIFOM*, 2015.

[24] J. E. Estes and D. W. Mooneyhan, "Of Maps and Myths," *Photogramm. Eng. Remote Sensing*, vol. 60, no. 517–524, 1994.

[25] M. F. Goodchild and L. Li, "Assuring the quality of volunteered geographic information," *Spat. Stat.*, vol. 1, pp. 110–120, May 2012.

[26] A. Mansourian, A. Rajabifard, M. J. Valadan Zoej, and I. Williamson, "Using SDI and web-based system to facilitate disaster management," *Comput. Geosci.*, vol. 32, no. 3, pp. 303–315, Apr. 2006.

- [27] P. Neis and A. Zipf, "Analyzing the Contributor Activity of a Volunteered Geographic Information Project — The Case of OpenStreetMap," *ISPRS Int. J. Geo-Information*, vol. 1, no. 2, pp. 146–165, 2012.
- [28] S. S. Sehra, J. Singh, and H. S. Rai, "Assessing OpenStreetMap Data Using Intrinsic Quality Indicators: An Extension to the QGIS Processing Toolbox," *Futur. Internet*, vol. 9, no. 2, 2017.
- [29] G. Touya and A. Reimer, "Inferring the Scale of OpenStreetMap Features," in *OpenStreetMap in GIScience: Experiences, Research, and Applications*, J. Jokar Arsanjani, A. Zipf, P. Mooney, and M. Helbich, Eds. Cham: Springer International Publishing, 2015, pp. 81–99.
- [30] Sumaryo, "Aspek Geospasial dalam Kasus Sengketa Pulau Berhala," in *Prosiding The 1st Conference on Geospatial Information Science and Engineering*, 2012, pp. 247–256.
- [31] Faturochman, "PERANG KECIL: Problem yang Terus Berlangsung," *Bul. Psikol.*, vol. 2, pp. 6–12, 1993.
- [32] B. Riadi and A. Makmuriyanto, "Kajian Percepatan Penetapan dan Penegasan Batas Kecamatan/Distrik, Desa/Kelurahan Secara Kartometri," *Maj. Ilm. Globe*, vol. 16, no. 2, pp. 109–116, 2014.
- [33] M. Caprioli and E. Tarantino, "Standards and Quality in GIS Contexts," in *Proceeding FIG Working Week 2003*, 2003.
- [34] B. Bhattacharya, "Disaster Mitigation of Large Infrastructure Systems," in *Proceedings of the International Symposium on Engineering under Uncertainty: Safety Assessment and Management (ISEUSAM - 2012)*, S. Chakraborty and G. Bhattacharya, Eds. India: Springer India, 2013, pp. 477–482.