Quality Improvement of Remotely Volunteered Geographic Information via Country-Specific Mapping Instructions

Carolin Klonner
Institute of Geography, Heidelberg Academy of Sciences and Humanities (HAW), Heidelberg, Germany
c.klonner@uni-heidelberg.de

Melanie Eckle
Institute of Geography, Heidelberg, Germany
eckle@uni-heidelberg.de

Tomás Usón
Institute of Geography, Heidelberg Academy of Sciences and Humanities (HAW), Heidelberg, Germany
uson@uni-heidelberg.de

Bernhard Höfle
Institute of Geography, Heidelberg Academy of Sciences and Humanities (HAW), Heidelberg Center for the Environment (HCE), Heidelberg, Germany
hoefle@uni-heidelberg.de

ABSTRACT

Volunteered geographic information can be seen as valuable data for various applications such as within disaster management. OpenStreetMap data, for example, are mainly contributed by remote mappers based on satellite imagery and have increasingly been implemented in response actions to various disasters. Yet, the quality often depends on the local and country-specific knowledge of the mappers, which is required for performing the mapping task. Hence, the question is raised whether there is a possibility to train remote mappers with country-specific mapping instructions in order to improve the quality of OpenStreetMap data. An experiment is conducted with Geography students to evaluate the effect of additional material that is provided in wiki format. Furthermore, a questionnaire is applied to collect participants' socio-demographic information, mapping experience and feedback about the material. This pre-study gives hints for future designs of country-specific mapping instructions as well as the experiment design itself.

Keywords

OpenStreetMap, country-specific mapping instructions, VGI, quality, disaster.

INTRODUCTION

The number of people affected and the damage produced by natural hazards, like floods, are increasing in the last decades (EM-DAT, 2016). This can be mainly attributed to changing climate conditions, the urban expansion in risk areas due to a rapidly-growing world population and the impact of human beings in nature (Ebert et al., 2009). Thus, disaster management plays an important role in dealing with such events. However, oftentimes, official data, e.g., map material for routing or the location of buildings, is not available or out of date. Therefore, in order to react in an efficient and fast way, there is already a high number of cases, in which volunteered geographic information (VGI) is used within natural hazard analysis (Horita et al., 2015; Klonner et al., 2016).

The map project OpenStreetMap (OSM) can be considered as very useful for such applications and since its first crisis setting in 2010 after the earthquake in Haiti, OSM has been applied for many use cases (Soden et al., 2014b). Further, the Humanitarian OSM Team (HOT), which evolved from the Haiti earthquake response, develops

https://doi.org/10.11588/heidok.00037174

mapping tasks in cooperation with humanitarian aid organizations. In addition, projects like Missing Maps[1] bring forward the need of map material for preparedness in order to have current map data already available before a disaster occurs. However, the quality has to be evaluated before data of a collaborative project like OSM, which is collected, edited and shared by volunteers from all over the world, can be applied in disaster management (Barron et al., 2014; Fan et al., 2014). This is especially important for crisis maps because most of the mappers work remotely on the basis of satellite imagery, and therefore, they might neither be familiar with the local conditions nor have country-specific (geographic) knowledge (Eckle et al., 2015).

An experiment conducted by Eckle et al. (2015) tackled this issue and they compared remotely mapped data by Geography students to the results of a local mapper from Kathmandu. They focused on Kathmandu in Nepal because this area is earthquake prone and the local Kathmandu Living Labs team[2] supports the mapping of the area in OSM since official data is very sparse (Soden et al., 2014a). The experiment of Eckle et al. (2015) can be considered as an initial study as it only had eight participants for the remote mapping. They state that their methods should be tested with a larger sample size and that the results showed that errors made by the remote mappers could be minimized by providing material explaining country-specific features. Moreover, See et al. (2013) portrayed in their study that with specific training and individual feedback, volunteers were able to improve more than experts of the domain.

Therefore, the objective of the following study is the quality analysis of remote mapping and the evaluation of the effect of additional material about country-specific features with a larger experiment setting, which is based on the Solomon Design and 72 Geography students as participants. A wiki page provides the additional country-specific instructions for the remote mapping and the experiment concludes with a questionnaire. The analysis of the results is based on correctness and completeness via a comparison of data contributions by the students versus OSM reference data.

STUDY AREA AND DATASETS

In general, during or after a disaster, remote mapping data is contributed to OSM mainly in areas, where no or only little map material is available from the official side. Therefore, there is generally a lack of independent reference data in order to evaluate the mapping quality of the volunteers.

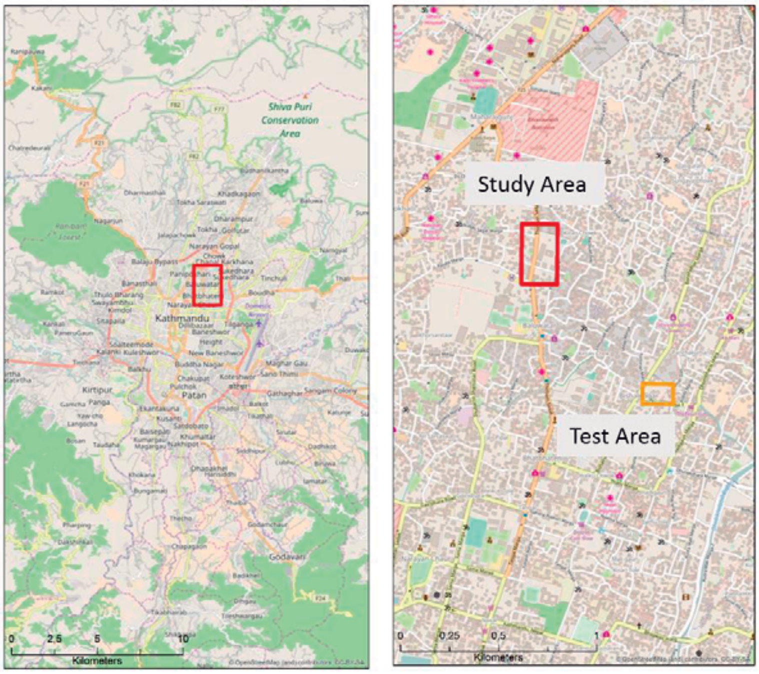

Figure 1. Study area in Kathmandu, Nepal.

The research area of the following quality examination is located in the city centre of Kathmandu, Nepal, (Figure 1) due to the following advantages in comparison to other remotely mapped areas. In 2012, a project aiming at the seismic resilience of the Kathmandu Valley was launched by the World Bank, the Global Facility for Disaster

[1] http://www.missingmaps.org/ (accessed: 15.01.2017)

[2] http://www.kathmandulivinglabs.org/ (accessed: 14.01.2017)

Reduction and Recovery (GFDRR) and the Government of Nepal[3] (Soden et al., 2014a). There was no complete database of schools and health facilities combined with coordinates and information about construction type available for a disaster risk model (ibid.). The collection of these data by locals within the Open Cities Kathmandu project was taken over after the end of the project by the Nepalese NGO Kathmandu Living Lab (ibid.). These efforts led to a rich OSM database, which is constantly updated. Moreover, during and after the earthquake in Nepal 2015, OSM was updated by thousands of remote and local mappers, including many experienced ones. This suggests high data quality and the OSM data are therefore used as a reference for the evaluation of the experiment data. To avoid any influence due to new satellite imagery, the reference is downloaded from the OSM data base on the day of the experiment (15.12.2015). The students' mapping data is based on the same Bing satellite imagery and for their mapping they use JOSM[4], a Java based open source editor for OSM. In addition, Kathmandu is chosen as research area following the previous experiment conducted by Eckle et al. (2015), in order to have comparable conditions.

METHODOLOGY

Experiment Design

An experiment can be seen as a form of an empirical study, which aims at identifying cause-effect-relations in order to explain social phenomena (Eifler, 2014). Processes are actively evoked by the experiment leader to identify the real cause of a phenomenon and therefore it is important to control factors that could also be causes to avoid alternative explanations (ibid.).

Further, Eifler (2014) states that a control group makes it possible to control the factors which influence the internal validity and that specific controlling techniques can be applied such as eliminating disruptive elements or keeping them constant as well as selecting the group randomly or parallel. In accordance, the Solomon Four Group Design is used as it aims at avoiding the influence of a pre-test, which is the mapping of a testing area in the following study, on the mapping results of the post-test, i.e. the mapping of the study area in the experiment. Since this method allows to exclude influences on the internal validity (ibid), it is chosen for conducting the experiment.

Participants

A Cartography lecture of 90 minutes was chosen as experiment setting and 72 students, mostly in their first semester, were taking part. It was assumed that these students represent to a certain extent the mapping community of new users. Moreover, this setting allows to have a similar set of participants regarding age and education and to have enough participants in order to provide about 20 participants per group. As Geography is a subject usually with quite equal gender distribution, the final group of participants consists of 36 female and 36 male students.

Tasks

All participants were attending the same introductory session with information about the experiment itself and explanations about the tools the students were going to use for the mapping (JOSM editor). Afterwards, they were randomly distributed into four different groups (Figure 2). Due to practical reasons and the setting, the distribution was not completely random in the experiment described here, but it was intended to make it as randomly as possible by handing out the same amount of task sheets for each group to the students. The task sheet included specific instructions about what to map and whether to read an additional wiki page (link was provided on the sheet). Each task on the sheet was assigned with a certain amount of time for fulfilling. The instructors of the experiment also reminded the participants of the single groups about the time. Moreover, the test area as well as the study area outline was provided on an online learning platform for the download into JOSM, which allowed the students to easily identify the area of mapping in the satellite image. After the mapping part, the students had to upload their results to the online platform. The task sheet gave further information about the folders on the online learning platform and the way they should save their data. The last task for all group participants was the questionnaire, for which a link was provided on the sheets.

Half of the participants had to do a pre-test, the test area in the experiment. One part (group 1) had to read the additional material on the wiki page first, before starting mapping the test area. They got a certain amount of time to do that. Group 2 could directly start with the test area. After mapping the test area for 10 minutes, both groups started the post-test, the study area in our experiment. The other half of the participants had no pre-test and again

[3] http://www.opencitiesproject.org/cities/kathmandu/ (accessed: 19.01.2017)

[4] https://josm.openstreetmap.de/ (accessed: 19.01.2017)

one group had additional material (group 3) in contrast to the control group (group 4). These two groups only mapped the post-test (study area).

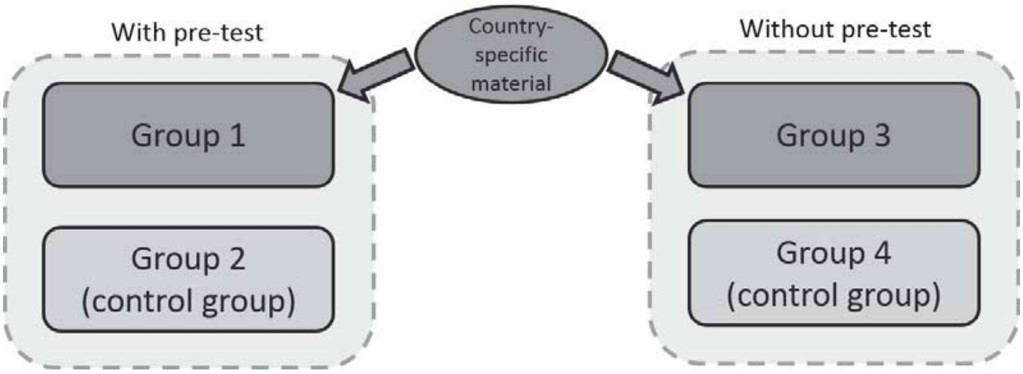

Figure 2. Experiment design based on the Solomon Four Group Design.

Wiki Page

The form of a wiki page is used for the experiment as, especially within educational contexts, this method proved to be a successful way for collaborative knowledge building (Kump et al., 2013). Cultural specific information for remote mappers can be provided and the content is based on experience and local knowledge of many people. Thus, a wiki enables, for example, the large OSM community in general, or the crisis mappers in specific, to contribute and share their local knowledge about a certain area with remote mappers. In urgent disaster cases a specific use case wiki page can be added. Moreover, it is possible, to share the information in different languages. Some of the content of the applied wiki page for the experiment was already in use after the earthquake in Nepal in 2015 and proved to be useful for remote mapping. This shows both the importance and applicability of such methods but also the urgent need to evaluate these tools in order to make statements about the quality of the resulting mapping and the possibilities of improving the portrayal of additional material.

The wiki of the experiment[5] comprises of characteristics of buildings of this region with specific mapping hints. It considers buildings that are irregular in elevation and therefore might appear as a set of houses, a row of houses, which may look like one big house, and complex building structures like multi-polygons (Figure 3). Further, the correct mapping of the actual layout of a building is explained. Focus is set on buildings as they are mostly the features new mappers start with. Moreover, especially for disaster management, the exact outline and the number of buildings is important for population estimation or to identify the type of usage of the building, e.g., elements at risk like schools or hospitals have often a certain shape.

Figure 3. Wiki page with additional information about specific features. Example of multi-polygon creation[6].

[5] https://wiki.openstreetmap.org/wiki/Nepal_remote_mapping_guide_Experiment (accessed: 15.01.2017)

[6] Ibid.

Questionnaire

The experiment consists of a mapping part followed by a questionnaire in order to gain background information about the students such as their age, gender, semester, and experience with geoinformation, remote mapping and OSM. Another issue is their knowledge about Kathmandu. Additionally, information about the mapping material and the applicability is inquired. Furthermore, they are asked about the time for finishing the task and the language issues. Finally, the motivation of the students for further mapping is evaluated. A link for the questionnaire is included in the task sheet.

Analysis

The overall hypothesis of the experiment is that additional country-specific mapping instructions improve the mapping quality of remote mappers. Therefore, the following section presents analysis methods to measure the quality of the data mapped by the students in comparison to the OSM reference data.

Different indicators can be used in order to evaluate the data created by the students. The ISO standards 19157:2013 (International Organization for Standardization) provide a set of quality indicators for geographic information, which are also described in the work of van Oort (2006) and Haklay (2010). The ISO standards define correctness or positional accuracy as "the accuracy of the position of features within a spatial reference system", while completeness is referred to as "the presence and absence of features, their attributes and relationships" (ISO 19157:2013). In the context of this analysis, the correctness indicates the accuracy of the mapped features by the students, whereas completeness reveals excess and missing data in the student dataset (cf. Rutzinger et al., 2009). Therefore, the indicators of correctness and completeness can be used to evaluate the quality of the mapping results of the students.

Klonner et al. (2015) applied two comparative methods for the work with OSM data in urban areas, namely the centroid and the overlap method. They conclude that the overlap method achieves more realistic results in areas with terraced houses and blocks of buildings. The urban area of Kathmandu resembles this kind of building structures, and therefore this method is chosen for the study at hand and applied for the comparison of the students' mapping results to the reference OSM data. Figure 4 shows an example of a mapped building within the student data (orange outline) and the reference data (red outline) as well as the area of their intersection. This intersection of the two building polygons represents the overlapping area (Klonner et al., 2015). Thus, in other words, the overlapping area shows the correctly mapped student data. In the example (Figure 4), the overlapping area (light orange) covers the entire reference building, which indicates that the student even overestimated the building area.

Figure 4. Example of a building in the reference data and mapped by a student. The overlapping area can be used for a comparison.

The overlapping area of the houses of the two datasets can be compared to the overall area of the building polygons mapped by the students and, of the reference buildings. These calculations can be used to assess the quality indicators. The completeness refers to the ratio of the overlapping area of all buildings and the sum of all areas of the buildings in the reference data. Correctness stands for the ratio of the overlapping area of all buildings and the area mapped by the students (cf. Rutzinger et al., 2009).

completeness = overlapping area of all buildings/ area of buildings in the reference data.

correctness = overlapping area of all buildings/ area mapped by the students.

RESULTS

In the following, the results of the mapping by the students during the experiment are evaluated based on the indicators of completeness and correctness. Moreover, the outcomes of the specific tasks such as the mapping of a multi-polygon and a row of houses are presented. The final part gives further information provided by the students within the questionnaire.

Mapping

The calculations of completeness and correctness can be done for the whole study area as well as for the test area. In this way percentages for all student datasets can be evaluated. The following section shows the results of a qualitative analysis due to the available sample size. The final number of student submissions (69) differs from the participant number (72) because three of the student datasets had to be excluded due to invalid data.

Figure 5 shows an example of a student of group 1. The map shows the study area with the reference building area (red outline), building polygons mapped by the student (orange outline) and the overlapping area (light orange area). This student resulted in a mapping completeness of 86% and correctness of 62%.

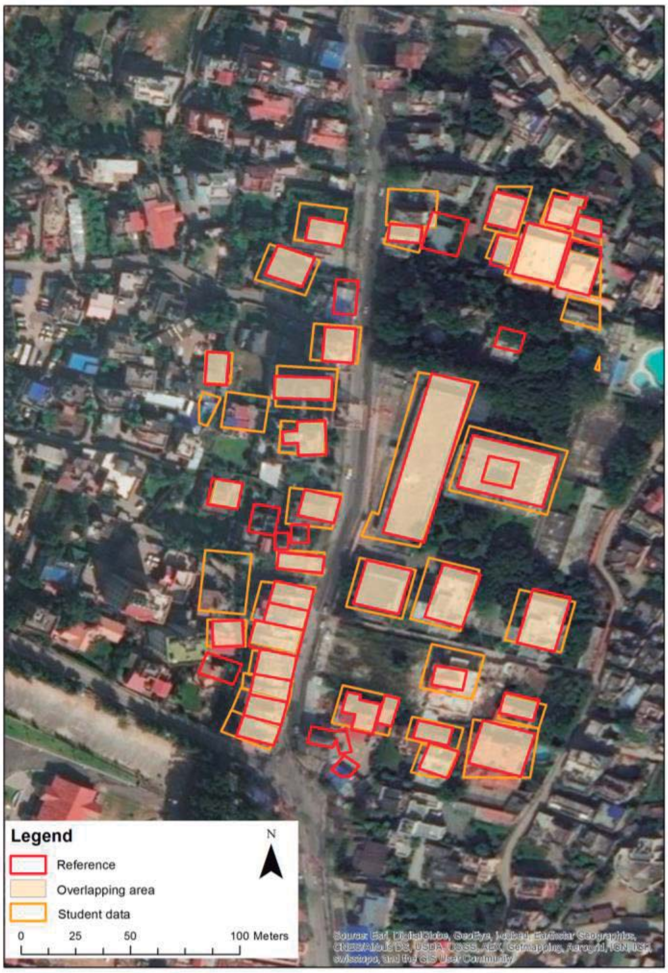

Figure 5. Buildings in the reference data (red) and mapped by a student (orange) as well as the overlapping area (light orange) in the study area of Kathmandu, Nepal.

The sample size of each group ranges from 16 to 18 participants and thus the following statistical analyses are not representative but can give some insights for further experiment designs. First of all, the average correctness values from group 1 and group 3 are compared in order to see whether there is an influence of the test area. A corresponding t-test results in a non-significant outcome ($p = 0.84$). The same comparison for the completeness results in a p-value of 0.53, and is therefore also not significant. So, it can be assumed that there is no influence of the test area, which allows future experiments to have only 2 groups instead of the 4 groups, which enables larger sample sizes of the groups for representative evaluations.

The assessment of the impact of additional information is based on the analysis of the country-specific features portrayed in the wiki and how the students succeeded in their mapping. In the following, focus is set on the mapping of a multi-polygon and a row-of-houses. The results of these analyses can give hints for the improvement of the wiki page for the follow-up experiment with only 2 groups and a larger sample size.

Moreover, the feature of the multi-polygon (Figure 6) can be used as an indicator whether the students really used the additional material. In the experiment of Eckle et al. (2015) it was not identifiable whether the participants read the material or mapped without reading it. This addition of the multi-polygon makes this possible as new mapper usually do not apply such specific features.

Table 1. Multi-polygon

	Group 1	Group 2	Group 3	Group 4
Use of multi-polygon	14 of 18 (78%)	1 of 18 (6%)	9 of 16 (56%)	2 of 17 (12%)

Figure 6. Example of a multi-polygon mapped by a student of group 1 (left) and not mapped by a student of group 2 (right).

The results show that the students of group 1 and group 3 read the material and used the multi-polygon in contrast to only a small number of mapped multi-polygons in the groups without the additional material (Table 1).

Table 2. Row of houses

	Group 1	Group 2	Group 3	Group 4
row of houses as several single houses AND correctly mapped (7±1 and >80% completeness)	3	3	1	0
row of houses as several single houses BUT ≤80% of area of houses are overlapped (completeness) or/and not 7 (±1) polygons	8	13	11	15
row of houses mapped as 1 single building	4	0	2	2
row of houses not mapped	3	2	2	0

For the evaluation of the mapping of the row of houses, different categories have to be applied because there are several ways to map this specific feature (Table 2). Some students did not map the row of houses at all, while others made one single large building. For the students who mapped single buildings a threshold was applied: The row of houses was only counted as mapped correctly when the number of polygons was 7±1 (of the 7 reference polygons) and the completeness of this specific feature was >80% (Figure 7).

Figure 7. Mapping of a row of houses. From left to right: reference data with 7 single polygons, correctly mapped data by the student, several buildings mapped by student but under the threshold of completeness, only one large building mapped for the row of houses.

On the wiki, the instruction for a row of houses also includes the remark that if they are not sure, they should map it as one big building. The comparison of group 1 and 2 shows that the students might have followed this advice as 4 students of group 1 made one big building in comparison to no one in group 2.

The overall results of the mapping of a row of houses of all groups indicate that the correct mapping is still very difficult for beginners. This has to be taken into consideration when using OSM data for disaster management or other applications.

Questionnaire

Only one of the students had been in Kathmandu before the experiment. The results of this student (group 4) are above the average of all groups, which might be a hint to the usefulness of local knowledge.

Moreover, the questionnaire reveals that some students had problems regarding the language. While this is only true for 15%, this still might indicate that future material provided for crisis management by HOT, for example, might even lead to more accurate mapping results if it is translated in different languages, which is technically straightforward in a wiki.

Regarding the time the students had available for the mapping task, the questionnaire shows that 24 students had no time problem while 44 either had to hurry up to finish the task or had not enough time to finish it. 4 did not respond. This time issue is due to the experiment setting because in real mapping situations the students could take their time for mapping.

DISCUSSION AND CONCLUSION

The experiment can be seen as a first step towards the evaluation of the impact of additional country specific material provided for remote mappers. In this paper, only a small analysis is portrayed as the work is still in progress. An extensive analysis of the questionnaire and the mapping of the remaining specific features will also show the different mapping results with respect to the time constraints, to previous knowledge or how the students rate their improvement themselves due to the mapping material.

Additionally, this pre-study gives hints about the experiment design and that the follow-up experiment does not need to use the Solomon Design anymore. The evaluation shows that there is no significant influence of the test area. Thus, only 2 groups are necessary for another mapping experiment allowing for larger sample sizes.

Moreover, regarding the wiki page, this experiment also reveals valuable information for improvement of a future experiment setting as there are still some issues that need to be taken into consideration such as language barriers. Furthermore, on the one hand the mapping might have been influenced by the motivation as, although the experiment was a near-to reality mapping, the students still had a different motivation than normal OSM mappers. On the other hand, the tasks of the experiment did not clearly distinguish whether the overall aim is correct mapping (so no time constraints) or mapping as much as possible (with time pressure). In a follow up testing, the exact indication of the aim of the mapping, whether completeness or correctness is required, should be made clear at the beginning in order to have the same setting for all.

Future work may also include further quality measures in order to face the complexity of crisis mapping. Semantic accuracy, for example, plays an important role especially regarding critical infrastructure. It gives insights into the link between the (geometric) representation of an object and the intended interpretation. Moreover, temporal accuracy can provide information about updates of the dataset regarding changes in the real world (van Oort, 2006; Haklay, 2010).

As remotely contributed VGI is gaining more and more importance in the disaster management sector (Eckle et al., 2016; Horita et al., 2015; Klonner et al., 2016), there need to be tools for enhancing the mapping quality already during the production phase. The study reveals that students use such guides like the description of the mapping of multi-polygons. However, for more complex issues like the outline of a row of houses, there is no clear difference between mappers who used additional material and mappers without specific information. Therefore, it has to be kept in mind that remotely mapped OSM data is very useful for crisis mapping, but that it cannot give all the details local mappers can provide and which might also be required for efficient disaster management. This shows that it is important to select the mapper type (remote or local) according to the task, which has to be supported. In addition, the management of the input of such heterogeneous groups needs to be taken into consideration in future work. Further, it is also advisable to include local communities to the creation of more specific mapping instructions because it is necessary to find out the needs of local actors or specific use case requirements. In the experiment case, the material was already approved to be useful during the crisis mapping of the Nepal earthquake in 2015.

ACKNOWLEDGEMENTS

This work was supported by the WIN-Kolleg of the Heidelberg Academy of Sciences and Humanities (HAW). The authors would like to express gratitude to the students of the Cartography lecture for participating in the experiment. Special thanks to René Westerholt for his advice on the statistical evaluation.

REFERENCES

Barron, C., Neis, P. and Zipf, A. (2014) A Comprehensive Framework for Intrinsic OpenStreetMap Quality Analysis. Transactions in GIS, 18, 6, 877-895.

Ebert, A., Banzhaf, E. and McPhee, J. (2009) The influence of urban expansion on the flood hazard in Santiago de Chile, Proceedings of the Joint Urban Remote Sensing Event, Shanghai, China.

Eckle, M. and de Albuquerque, J. P. (2015) Quality Assessment of Remote Mapping in OpenStreetMap for Disaster Management Purposes, Proceedings of the 12th International Conference on Information Systems for Crisis Response and Management (ISCRAM), Kristiansand, Norway.

Eckle, M., Herfort, B., de Alberquerque, J. P., Leiner, R., Wolff, R., Jacobs, C. and Zipf, A. (2016) Leveraging OpenStreetMap to support flood risk management in municipalities: A prototype decision support system, Proceedings of the 13th International Conference on Information Systems for Crisis Response and Management (ISCRAM), Rio de Janeiro, Brazil.

Eifler, S. (2014) Experiment, In Baur and Blasius (Eds.), Handbuch Methoden der empirischen Sozialforschung. Wiesbaden: Springer Fachmedien, 195-209.

EM-DAT. (2016) The OFDA/CRED International Disaster Database. Accessed: 14.01.2017, http://www.emdat.be/disaster_trends/index.html.

Fan, H., Zipf, A., Fu, Q. and Neis, P. (2014) Quality assessment for building footprints data on OpenStreetMap.

International Journal of Geographical Information Science, 28, 4, 700-719.

Haklay, M. (2010) How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. Environment and Planning B: Planning and Design, 37, 4, 682-703.

Horita, F. E. A., Albuquerque, J. P. de, Degrossi, L. C., Mendiondo, E. M. and Ueyama, J. (2015) Development of a spatial decision support system for flood risk management in Brazil that combines volunteered geographic information with wireless sensor networks. Computers & Geosciences, 80, 84-94.

International Organization for Standardization. ISO 19157:2013. Geographic Information. Data Quality.

Klonner, C., Barron, C., Neis, P. and Höfle, B. (2015) Updating digital elevation models via change detection and fusion of human and remote sensor data in urban environments. International Journal of Digital Earth, 8, 2, 153-171.

Klonner, C., Marx, S., Usón, T., Porto de Albuquerque, J. and Höfle, B. (2016) Volunteered Geographic Information in Natural Hazard Analysis: A Systematic Literature Review of Current Approaches with a Focus on Preparedness and Mitigation. ISPRS International Journal of Geo-Information, 5, 7, 103.

Kump, B., Moskaliuk, J., Dennerlein, S. and Ley, T. (2013) Tracing knowledge co-evolution in a realistic course setting: A wiki-based field experiment. Computers & Education, 69, 60-70.

Rutzinger, M., Rottensteiner, F. and Pfeifer, N. (2009) A Comparison of Evaluation Techniques for Building Extraction From Airborne Laser Scanning. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2, 1, 11-20.

See, L., Comber, A., Salk, C., Fritz, S., van der Velde, M., Perger, C., Schill, C., McCallum, I., Kraxner, F. and Obersteiner, M. (2013) Comparing the Quality of Crowdsourced Data Contributed by Expert and Non-Experts. PLoS ONE, 8, 7, e69958.

Soden, R., Budhathoki, N. R. and Palen, L. (2014a) Resilience-Building and the Crisis Informatics Agenda: Lessons Learned from Open Cities Kathmandu, Proceedings of the 11th International Conference on Information Systems for Crisis Response and Management (ISCRAM), University Park, Pennsylvania, USA.

Soden, R. and Palen, L. (2014b) From Crowdsourced Mapping to Community Mapping: The Post-Earthquake Work of OpenStreetMap Haiti, Proceedings of the 11th International Conference on the Design of Cooperative Systems (COOP), Nice, France.

van Oort, P. (2006) Spatial Data Quality: From Description to Application. PhD thesis, Wageningen University.