



Article The Sketch Map Tool Facilitates the Assessment of OpenStreetMap Data for Participatory Mapping

Carolin Klonner ^{1,*}, Maximilian Hartmann ¹, Rebecca Dischl ¹, Lily Djami ², Liana Anderson ³, Martin Raifer ⁴, Fernanda Lima-Silva ⁵, Lívia Castro Degrossi ⁵, Alexander Zipf ^{1,4}, and João Porto de Albuquerque ⁶

- ¹ Institute of Geography, Heidelberg University, 69120 Heidelberg, Germany; max_hartmann@stud.uni-heidelberg.de (M.H.); rebecca.dischl@stud.uni-heidelberg.de (R.D.); zipf@uni-heidelberg.de (A.Z.)
- ² Institute of Computer Science, Goethe University, 60325 Frankfurt am Main, Germany; lily.djami@stud.uni-frankfurt.de
- ³ National Center for Monitoring and Early Warning of Natural Disasters—CEMADEN, 12247-016 São José dos Campos, Brazil; liana.anderson@cemaden.gov.br
- ⁴ Heidelberg Institute for Geoinformation Technology, Heidelberg University (HeiGIT), 69118 Heidelberg, Germany; martin.raifer@heigit.org
- Getulio Vargas Foundation, Center of Public Administration and Government (FGV CEAPG), 01313-902 São Paulo, Brazil; fernanda.silva@fgv.br (F.L.-S.); livia.degrossi@fgv.br (L.C.D.)
- ⁶ Institute of Global Sustainable Development, University of Warwick, Coventry CV4 7AL, UK; j.porto@warwick.ac.uk
- Correspondence: c.klonner@uni-heidelberg.de

Abstract: A worldwide increase in the number of people and areas affected by disasters has led to more and more approaches that focus on the integration of local knowledge into disaster risk reduction processes. The research at hand shows a method for formalizing this local knowledge via sketch maps in the context of flooding. The Sketch Map Tool enables not only the visualization of this local knowledge and analyses of OpenStreetMap data quality but also the communication of the results of these analyses in an understandable way. Since the tool will be open-source and several analyses are made automatically, the tool also offers a method for local governments in areas where historic data or financial means for flood mitigation are limited. Example analyses for two cities in Brazil show the functionalities of the tool and allow the evaluation of its applicability. Results depict that the fitness-for-purpose analysis of the OpenStreetMap data reveals promising results to identify whether the sketch map approach can be used in a certain area or if citizens might have problems with marking their flood experiences. In this way, an intrinsic quality analysis is incorporated into a participatory mapping approach. Additionally, different paper formats offered for printing enable not only individual mapping but also group mapping. Future work will focus on advancing the automation of all steps of the tool to allow members of local governments without specific technical knowledge to apply the Sketch Map Tool for their own study areas.

Keywords: OpenStreetMap quality evaluation; field papers; field mapping; disaster risk management; flood risk perception; flooding

1. Introduction

International and interdisciplinary research projects with a focus on disaster risk reduction strategies are on the rise due to ever more affected areas and people. This paper focuses on flooding because a lack of adequate financial and technical means to account for flood mitigation leads to insufficient disaster risk reduction in many regions. How can these people prepare themselves better and how can they be supported? A valuable source of information can be used, which is so far rarely applied for disaster risk reduction: The local knowledge and experiences about flooding. Participatory tools can visualize and



Citation: Klonner, C.; Hartmann, M.; Dischl, R.; Djami, L.; Anderson, L.; Raifer, M.; Lima-Silva, F.; Castro Degrossi, L.; Zipf, A.; Porto de Albuquerque, J. The Sketch Map Tool Facilitates the Assessment of OpenStreetMap Data for Participatory Mapping. *ISPRS Int. J. Geo-Inf.* 2021, *10*, 130. http:// doi.org/10.3390/ijgi10030130

Academic Editors: A. Yair Grinberger, Marco Minghini, Peter Mooney, Levente Juhász, Godwin Yeboah and Wolfgang Kainz

Received: 28 December 2020 Accepted: 24 February 2021 Published: 3 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). formalize such local knowledge by people in affected areas mainly in cases where data such as flood maps are sparse and where people are marginalized because they cannot afford the required digital devices or do not have internet access [1]. Sketch maps are a specific way to capture spatial local knowledge about floodings, which can be based on digital devices [2] or on paper maps [3,4]. In contrast to mental maps, sketch maps do not use blank paper but geographic maps and thus, the marking of the participants is guided by the map template [5]. Bustillos et al. [6] base their sketch maps on normal maps, while the method applied by Klonner et al. [4,7,8] offers a step more because it makes the data available in a digital and spatial format based on Field Papers, which offer automatic georeferencing. The advantage of paper-based mapping, i.e., the inclusion of marginalized people, is combined with the advantage of mapping based on digital devices, i.e., the availability of georeferenced data.

The Sketch Map Tool, which will be outlined in Section 2, has the objective to unify all the tools necessary to work with OpenStreetMap (OSM) in projects using participatory mapping approaches. The assessment of the data quality is very important when using user-generated geographic data such as OSM data for participatory mapping. There are already several approaches to assess OSM quality in general [9]. In our setting, in which users are allowed to carry out such studies for every part of the world, it is necessary to provide analyses without the necessity of reference data since they are not available everywhere. We therefore rely on intrinsic analyses. Many approaches concerning this topic can be found in the literature as well (for example [10,11]). We adjust them to the requirements of participatory mapping activities and use additional methods specific to our subject. We focus not only on intrinsic quality analyses but also on how results can be communicated in an understandable and accessible way. This is not only important for disaster managers but also for planners and the general public. Most previous studies regarding quality analyses discussed the analysis itself, and the communication of the results to other stakeholders did not receive enough attention.

A very integral part of carrying out participatory mapping approaches is the generation of sketch maps on which the participants in the study or project can map the information that is supposed to be gathered. Another important aspect is the possibility to digitize these sketch maps after finishing the work in the field. Field Papers [12] enable users to complete both tasks. In the Waterproofing Data Project, we use the functionalities of Field Papers for example to let citizens in the study areas map different levels of flooding and to digitize these markings afterwards. Although the Field Papers tool is very useful and has strongly supported our work, we also experienced its limits. Bustillos et al. [6] and Klonner and Blessing [13] identified group discussions based on maps in the A0 format as a valuable method to gain more information about the context of the flooding. However, the largest format supported by Field Papers is A3, and therefore, the maps had to be enlarged and were therefore pixelated and less convenient to use [13]. This need for sketch maps of higher resolution was one of the reasons for the creation of the Sketch Map Tool. The process of sketch map generation laid out in the Field Papers GitHub repository [14] achieves its purpose by creating an HTML page with an embedded map based on OSM and converting this page into a PDF document. The maximum resolution is therefore limited by the maximum resolution in which the OSM tiles that form the map are available, though the maximum resolution is effectively doubled by using a zoom offset and adjusting the size of the tiles [14]. It is, therefore, necessary to use another approach for higher resolution sketch maps, which will be described in Section 3.2. The second main reason for the development of the Sketch Map Tool was the need for easily accessible quality analyses. A sustainable tool requires the usability for local stakeholders without a researcher, who analyzes the fitness of the OSM data for the use for a data collection based on sketch maps. Results are therefore presented, inter alia, in form of a traffic light, i.e., graded as red, yellow or green, to enable the users to easily identify the usability of the data themselves. We include an automatic quality evaluation that allows the user to check the OSM data of

the required study area and thus, the approach incorporates an intrinsic quality analysis for participatory mapping.

The previously mentioned studies not only showed that sketch maps are a great tool to capture local knowledge but also identified adaptation possibilities of this method to make the tool even more user-friendly and offer a sustainable use by local governments. How can the maps be generated in A0 format without losing resolution? How can a member of a local government without specific technical knowledge conduct a study to formalize local knowledge? Klonner et al. [15] give a first glimpse at a new tool which is based on Field Papers but also includes answers to the previous questions. In this paper, these first steps are taken forward and the Sketch Map Tool with its functionalities is presented. The main two limitations from the Field Papers are overcome by including an intrinsic quality analysis for a participatory method and by offering different paper formats for the printouts. First, the tool is explained in detail and the different components are described. Thereafter, the tool is applied for case studies in Brazil with a focus on river flooding in Rio Branco and urban flooding in M'Boi Mirim, São Paulo. The discussion reflects on the results of these studies and evaluates the tool. As the development of the tool is an ongoing process next steps of advancing the tool are outlined.

2. The Sketch Map Tool

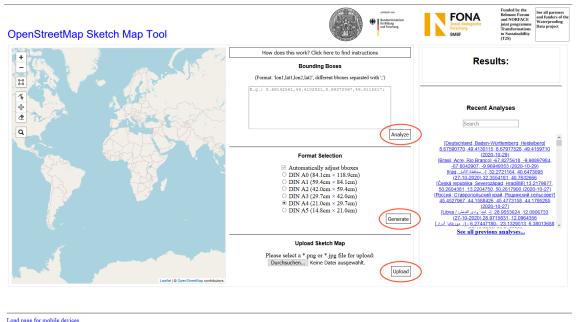
As a whole, the Sketch Map Tool unifies the aforementioned methods: A feature for analyzing the fitness of a selected region's OSM data for usage in sketch maps, a sketch map generation function and a sketch map upload function (see Figure 1). It therefore substantially facilitates the sketch map usage process, especially by pointing out possible difficulties to be aware of during the process and by focusing on the comprehensibility for laypeople in all the steps (see Figure 2). Thus, it is an easy-to-use tool that can be applied by local stakeholders to allow them to gather themselves some information about previous flood events in a specific area. After having successfully completed the fitness analyses for a possible study area and deciding to use the sketch maps for data collection, the users can request the generation of a sketch map in a paper format they prefer, which will then be provided as a PDF document which also contains a QR Code [16] necessary for the automatic georeferencing and little globes, which frame the generated map to detect its boundaries. After applying a printout of this sketch map in the field, e.g., for marking flood risk perception, the users can use the upload functionality, which allows the creation of a GeoTIFF file from a photograph of the map. This file can be used for visualizations or further analyses in a GIS, for example.

3. Materials and Methods

The Sketch Map Tool is mainly written in Python 3 [17], the web-based user interface is written in HTML [18], CSS [19] and JavaScript [20]. The integration between the Python code and the web pages is implemented using Flask [21], which is a framework that enables the development of web apps using Python.

3.1. OpenStreetMap History Database and the Ohsome API

The OpenStreetMap History Database (OSHDB) [22] is a tool developed by the Heidelberg Institute for Geoinformation Technology (HeiGIT) that allows analyses of OSM data with a focus on analyzing the data's evolution. It provides efficient storage and retrieval of OSM data for any available spatio-temporal extent [23]. The Ohsome API, which was developed by the HeiGIT as well, provides a convenient Application Programming Interface (API), to access an instance of the OSHDB and therefore enables the access of OSM data for a chosen spatio-temporal extent [24]. It also enables filtering and aggregation of the data. For example, it is possible to calculate the density of features with a certain tag (e.g., *tourism = hotel*). The analysis feature of the Sketch Map Tool uses the Ohsome API to access the OSM data for performing its quality analyses (see Section 3.4). The fitness of a study region's OSM data for usage in sketch maps is evaluated based on an intrinsic quality analysis, examining both the recent state and the evolution of the OSM data. Users are given a general fitness score, to assess the overall fitness of a region's OSM data as well as detailed reports about each analysis. Furthermore, recommendations on what to consider before and while using OSM-based sketch maps of this region are given.



Cond page for mobile devices This tool accesses OpenStreetMap data, which is partly aggregated, via the <u>obsome AP</u>I by the <u>Heidelberg Institute for Geoinformation Technology (HeiGIT</u>). The data and statistics are based on data by <u>© OpenStreetMap</u> <u>contributors</u>. obsome uses a database that contains <u>ODbL 1.0</u> licensed <u>OSM data</u> and <u>CC-BY-SA 2.0</u> licensed <u>OSM data</u>.

Figure 1. The Sketch Map Tool combines the analysis of the fitness of a selected region's OSM data for usage in sketch maps, a sketch map generation and a sketch map upload function (the buttons to start each of the steps are highlighted in red). The requested study area can either be marked on the map on the left or entered manually using the input field *Bounding Boxes*. For an optimal creation of the sketch map, the user can choose his preferred format. Results of the functionalities can then be obtained on the right side.

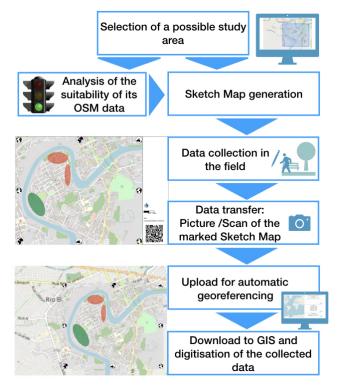


Figure 2. Workflow of a possible field study that is carried out based on the Sketch Map Tool.

Mapnik is an open-source toolkit that provides various functionalities for the creation of maps based on geographic data [25], for example, from OSM. Nik4 is a Python script that uses Mapnik to create image files from OSM data and provides the possibility to conveniently specify parameters such as the zoom level and output resolution [26]. The Sketch Map Tool uses Mapnik via Nik4 to generate maps from OSM data of areas selected by the user (see Figure 3). OSM data can either be accessed from locally stored data of the whole planet or fetched dynamically by requesting data covering the chosen bounding box via an Overpass API instance whose terms of use permit such usage [27]. The maps are rendered in a style that is based on OpenStreetMap Carto [28] and carto-style-hydda [29]. To improve the orientation of participants in sketch map studies, the style was changed to show more points of interest (POIs) in the renderings. The zoom level can be customized in case there is the need for a special level of detail. This has no effect on the scale of the map, i.e., the covered region is not altered. The rendering changes, especially there are more details shown with a higher zoom level.



Figure 3. The sketch map of a study area, here in Rio Branco, is saved in a PDF document which also contains a QR Code necessary for the automatic georeferencing and little globes, which frame the generated map to detect its boundaries when it is uploaded to the Sketch Map Tool.

3.3. Image Processing

OpenCV is an open-source computer vision and machine learning software library and offers many possibilities of analyzing and transforming image files [30]. Binary Robust Invariant Scalable Keypoints (BRISK) is a method for the detection of keypoints, i.e., distinctive regions in images, and their description [31]. It enables the matching of two images using the detected features. As this method proved to work with sufficient precision and in acceptable time in our settings, the Sketch Map Tool uses its implementation in OpenCV to detect sketch maps in photographs that are uploaded for further processing. First, the QR code is read to get information about the bounding box of the map and the date it was created. Using this information, a file of the unmarked map, which had been stored when generating the sketch map, is loaded to be used as a template. With this template, the BRISK implementation in OpenCV is used to detect and extract the actual map from the uploaded photograph in the same perspective as the template. The resulting image is then georeferenced using GDAL [32] with the bounding box information obtained from the QR code, and provided as a GeoTIFF file to the user.

3.4. OpenStreetMap Quality Analyses

The sketch maps used in our project are based on OSM data. Therefore, we need to look at the quality of these data for our use case to be aware of possible problems, which might occur when using the maps. For providing quality analyses of OSM data for any given region of the world, one cannot rely on external reference data. Thus, it is necessary to use intrinsic quality measures. For the selection of such measures, we used both literature research and experiences from the work with sketch maps in the Waterproofing Data Project. Some of the analyses performed by the Sketch Map Tool were already briefly outlined in a contribution to the State of the Map 2019 conference [15]. Each analysis is carried out for the potential study area and the results are combined into a general sketch map fitness score. To provide easily accessible conclusions that can also be used by laypeople and citizen scientists, the quality evaluation is resulting in one of three levels, depending on the fitness for purpose: Green (indicating good fitness), yellow (indicating potential problems to be aware of), and red (indicating that problems when using OSM-based sketch maps of this area are likely). Based on our experience from previous research projects we assign each analysis a weight, reflecting its importance (less important: weight = 0.5, important: weight = 1.0 and very important: weight = 2.0) regarding the usability of the OSM data in sketch maps. The resulting levels for the analyses can be combined by using a weighted average as heuristic. This general score is calculated as follows:

$$generalScore = \left\lfloor \frac{1}{\sum_{i=1}^{6} weight_i} * \sum_{i=1}^{6} [weight_i * score_i] + 0.5 \right\rfloor$$
(1)

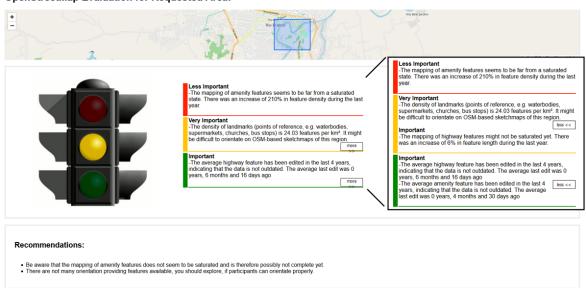
With each score being the result of an analysis as 0 (red), 1 (yellow) or 2 (green) and the weights shown in Table 1.

Table 1. Weighting of the results of the analyses.

i	Analysis	Weight _i
1	Saturation of mapped amenities	0.5 (1.0 when a lack of data is presumed)
2	Saturation of the road network mapping	1.0 (2.0 when a lack of data is presumed)
3	Important sources	0.0 (manual inspection needed)
4	Currentness of the amenity data	1.0
5	Currentness of the road network data	1.0
6	Density of orientation providing features	2.0

All intended user groups, including laypeople and citizen scientists, who want to contribute and gather data for disaster risk reduction, need to easily understand the tool and the results, and therefore, the general score is presented in form of a traffic light set at the calculated level: 0 = red, 1 = yellow, 2 = green.

Besides this general fitness score the results for each analysis are separately reported as well, including conclusions, recommendations and visualizations of the inspected data (Figure 4). We have been working on more quality indicators but decided to focus on those listed below for current analyses. For the other indicators (the duration of *fixme, note,* and similar tags and the positional stability of features) more research on their adequate applicability is required and they are pointed out in the discussion section. The thresholds used in our analyses and the weights are exploratory values, which are based on both literature and own previous research regarding the usage of sketch maps. Our goal in the presented work is the creation of a tool which enables a workflow that integrates the fitness-for-purpose analyses of OSM data in the sketch map usage process. As will be described in the discussion section, it is expected that these parameters are adjusted after the first applications of the tool in the field.



OpenStreetMap Evaluation for Requested Area:

Figure 4. The result of the evaluation of the requested area, here Rio Branco, is visualized using a traffic light to facilitate the classification of the result for laypeople. Next to the traffic light one finds elaborated and assessed information regarding the examined OSM data. The lower part contains recommendations to facilitate a field study in this area.

• Saturation of mapped amenities

As many common POIs, e.g., restaurants, educational and cultural facilities in OSM are tagged with the *amenity* key [33], it is useful to know, if they are likely to be mapped completely or not. In the latter case, some important reference points for the participants to orientate themselves on the sketch maps might be missing. As many categories of amenity features are rendered with special icons (e.g., a coffee cup for a café) [33] they are very valuable when orienting on sketch maps. The analysis of the recent density of orientation providing features, which will be explained below, is related. In the analysis described here, the focus does not lie on the current quantity of orientation providing features, but on the estimated completeness of their mapping. If it is inferred that the mapping of amenity features is close to a state of completion or has been completed, an easier orientation for the participants in a sketch map study is likely, and level green is reported. To obtain an indicator for the mapping completeness of relevant features, we analyze their saturation. When the number of added features falls below a certain threshold, one can infer that the mapping of these features is close to a state of completion or has been completed [34]. Rehrl and Gröchenig [10] use as a threshold the case when the number of created features in a year falls below three percent of the total count of features. We simplify this analysis by inspecting the increase in the density of features per year, as in unsaturated phases one can expect that there are more features to add than to delete, and consequently there is an increase in the density of features above a certain threshold. In our implementation of the analysis, the change is inspected relative to the number of features per $\rm km^2$ in the previous year, instead of relative to the total number of features per $\rm km^2$ in the year in which the new features have been created. As almost complete data can already be expected to provide most of the important landmarks, the three percent threshold of Rehrl and Gröchenig is too strict in our use case, besides the differences in measuring this indicator, which also impede a usage of the same threshold. In our tests 5% for the threshold from yellow to green and 10% for the threshold from red to yellow seemed adequate, when compared to our own estimations of mapping completeness. Like with all parameters of the analyses, it might be necessary to adjust these values as soon as more experiences from practical usage of the tool are available or other use cases are considered. We also included a 25% threshold, which needs

to be surpassed in at least one year, to ensure that there has been a stage of growth at all and not just a lack of data as an explanation for small relative growth. If this threshold is failed, a red level is reported. Generally, we consider this analysis' results as less important (weight = 0.5) compared to the other analyses, as a saturation of the mapping of amenity features does imply an easier orientation, but its impact is expected to be less heavy than that of the density of such orientation providing features. The weight makes the influence of this analysis on the general score smaller than the influence of analyses, which are rated as more important and therefore have higher weights. In case a lack of data is presumed, the importance is considered to be moderate (weight = 1), because a negative impact on the ability of participants to orientate on the sketch maps is expected.

Density of orientation providing features

As described in the previous paragraph, mapped POIs are important for the orientation of participants when using sketch maps. The same significance applies to landmarks, e.g., touristic attractions or natural features. The density of these features can therefore be analyzed for obtaining an indicator of how well participants in sketch map studies can orientate themselves on the maps. We are analyzing the recent combined density of railway stations, shops, places of worship, bus stops, hotels, touristic attractions, pharmacies, hospitals, restaurants, gas stations, schools, universities, colleges, town halls, police stations, fire brigade stations, parks, national parks, mountain peaks, waterways such as rivers, and water bodies such as lakes. Thresholds that seemed adequate in providing a green, yellow or red rating of how many of such features are mapped that can provide better orientation, are 30 features per km² for the threshold from yellow to green and 10 features per km² for the threshold from red to yellow. We derived these values from multiple map extracts of areas in which we could either navigate well, fairly or hardly. The POI densities in these areas have been compared and the thresholds have been chosen in a way in which they reflect our own navigability based on POI density. These exploratory values can be further improved when more data about sufficient POI densities for the various use cases are available. The results are weighted as very important (weight = 2), because it is crucial for the usability of paper-based sketch maps, that participants can orientate themselves properly on them.

Saturation of the road network mapping

The analysis of the saturation of the road network mapping is carried out in a similar way as the analysis of the amenity mapping saturation. In their OSM quality analysis framework "iOSMAnalyzer" [11] Barron, Neis and Zipf analyzed the completion of the mapping of different categories of roads by inspecting their monthly increase in feature length and comparing this increase to the increase in feature length of road categories lower in the hierarchy. When gathering data about disaster risk, there is the need to have reliable data for all road categories. Using sketch maps to capture urban flooding, one must expect that not just roads high in the hierarchy such as motorways, but also especially roads that are lower in the hierarchy such as streets in residential areas are affected. Therefore, we do not distinguish between different road categories when analyzing the increase in feature length. When a possible study area is selected, it is preferable that all the roads that are in this area are mapped completely. As this analysis follows a rationale very similar to that of the amenity mapping saturation analysis, we apply the same way of analyzing, i.e., inspecting the relative yearly growth—now regarding the feature length instead of the density—with exploratory thresholds of 5% and 10% for the detection of saturation and 25% for the detection of growth stages, i.e., the absence of a lack of data. The road network has an integral importance in the use of sketch maps because it is not only an aid for orientation, but also subject to markings in many sketch map applications. This analysis has

therefore weight = 1 when a stage of completion or growth has been detected and weight = 2 when there is a possible lack of data.

Important sources

It is possible to gather knowledge about the precision of mapped features by inspecting the sources of data (e.g., [35]). To enable users to inspect important sources, this analysis reports all sources that have been named in a *source* tag in a share of at least 2% of all mapped features in the chosen area. Although it is not typical to add *source* tags to individual features in the current mapping practice of OSM (they should be included in changesets, which are not inspected here) it was common in the past [36] and therefore many *source* tags remain on single features. These sources and their respective shares are not evaluated automatically but presented to the user so that he can inspect the important sources and consider their trustworthiness and their relevance for the specific study. As the conclusions depend on manual inspections, this analysis is not considered in the general score (weight = 0).

• Currentness of the data

When using sketch maps, it is preferable that all real-world features are mapped in their current state. For example, if a building has been torn down, it should not appear in recent OSM data anymore, and if a new supermarket has opened, it should be visible in OSM. Barron, Neis and Zipf [11] stated the possibility of analyzing the currentness of OSM data by visualizing its latest modification, as the user doing the last change indirectly confirms the correctness of the feature. They point out that this approach is problematic, when features are already mapped correctly and there is no need to change them. To avoid this problem, we inspect the average time passed since the last modification of all features in relevant categories and use thresholds that are long enough so that there is some change to be expected in the real world (e.g., changing opening hours, newly built houses, closed stores). This should cause the average time since the last edit to be below a certain threshold if there is an active community ensuring currentness of the data. The used thresholds are 4 years from yellow to green and 8 years from red to yellow. Two categories of features are checked for their currentness. Because of their explained importance, features tagged with *amenity* = * and *highway* = * are analyzed for their average currentness. The result is reported for both categories. The result's importance is considered to be moderate (weight = 1), as outdated map data might cause confusion among participants in sketch map studies.

3.5. Workflow

The Sketch Map Tool, as shown in Figure 1, is split into three main parts. It comprises a map in the left column, the main functionalities bounding box selection, analysis, sketch map generation and sketch map upload in the central column and the results section in the right column. After the successful completion of the requested task, a link leading to the elaborated analysis, the generated sketch map or the GeoTIFF file will be displayed in the upper right corner. Underneath, the results of previous analyses are listed so that older results can be reviewed again. The tool is supposed to be used as indicated by the workflow depicted in Figure 2.

Selection of possible study areas

To assess the fitness of an intended research area regarding the quality characteristics mentioned in Section 3.4 using the Sketch Map Tool or for the generation of a sketch map, the area of interest can either be marked as a rectangle on the map, or specified directly as a bounding box with the corresponding coordinates. In this case, the values of the bounding box must be entered via minimum longitude, minimum latitude, maximum longitude, maximum latitude, where latitude is a decimal number between

-90.0 and 90.0 and longitude is a decimal number between -180.0 and 180.0. If more than one area needs to be assessed, it is possible to enter or select several areas.

Analysis and Results

Subsequently, the selected areas can be evaluated using the *Analyze* button. The result contains an assessment of the fitness of the OSM data for participatory mapping, which is displayed as a traffic light depending on the quality of the existing data. Additionally, the different criteria which were considered for the outcome are specified in a written form. Concluding recommendations are given to help the user to be aware of possible problems that might occur during a sketch map study and further information which could be helpful to check beforehand. To enable a better understanding of the results as well as to give more details, additional visualizations of the data are provided. These visualizations include for example the plotting of the feature length development for highway features and the feature density development for amenity features. The shares of different POI and landmark categories of these features are provided in form of a pie chart to allow insights on which types of POIs might be used as reference points by sketch map study participants in this area.

• Sketch Map Generation

After successfully evaluating a possible study area's OSM data, having taken notice of possible problems, and deciding for that study area, the next step in the workflow is the generation of printable sketch maps. To generate a map for a sketch map study, the requested area can be selected in the same way as during the analyses by drawing a bounding box on the map or by providing coordinates. The area can, but does not have to be analyzed beforehand. Again, multiple areas can be selected to obtain several maps at one time. The user can choose between different paper formats and has the option to extend the bounding box automatically to generate a map with the optimal scale for the chosen paper format. This means that there are no blank spaces except the area where the QR code and scale are placed. The size of the rendered map is determined by both the paper format and the bounding box selection-and will be automatically chosen to print the map as big as possible on the chosen paper format. As soon as the generation is finished, a PDF file of the acquired sketch map is provided for download, which can be used for a field study. The map can either be printed out or used digitally. In both cases, the findings can be drawn in as needed, but must not cover the QR Code.

Upload

The map can be digitized into the GeoTIFF format using the *Upload* function of the Sketch Map Tool. To do this, the previously created and marked sketch map can be photographed. The maps are designed as such that they can be photographed even under unfavorable illumination conditions and in tilted positions, although the quality is enhanced under favorable conditions and the correct recognition is more likely. After the upload, a GeoTIFF file of the area is provided for download and can be used for further applications. In our work, we use the open-source GIS software QGIS [37] to further process the generated GeoTIFF files.

4. Results

Two areas in Brazil, which were studied within the framework of the Waterproofing Data Project, will be analyzed for demonstration purposes to show how the Sketch Map Tool works. The Waterproofing Data Project (https://www.geog.uni-heidelberg.de/gis/waterproofing.html, accessed on 11 November 2020) aims at improving the resilience of flood-vulnerable communities and at increasing the capacity for local governments and disaster monitoring centers to respond to flood events. Within the project we focus on two flood types, namely urban flooding in M'Boi Mirim, a neighborhood in the southern part of

the city of São Paulo and river flooding in a central region of Rio Branco, the capital of Acre. In 2015, Rio Branco had to face one of its most severe floodings and the city was flooded for about a month. M'Boi Mirim experiences flash floods on a regular basis, which mostly only take a few hours but affect houses and the infrastructure in this densely populated area.

4.1. Case Study—Rio Branco

In the following the analysis of the study region in Rio Branco shall be presented exemplarily. For this purpose, the corresponding region is analyzed by the Sketch Map Tool first. After the successful evaluation by the Sketch Map Tool, the result is displayed as depicted in Figure 4. The elaboration of the OSM data of the requested area is classified in the general categories red, yellow and green according to the indicated fitness for usage in sketch maps (see Section 3.4). As explained previously (ibid.), the quality criteria are then further divided into the subcategories Very Important (weight 2.0), Important (weight 1.0) and Less Important (weight 0.5). Based on these data, the color of the traffic light is chosen, see Equation (1). The traffic light was designed to facilitate the use and comprehension of the tool and its result even for less experienced users. At the right-hand side of the traffic light, more experienced or interested users can read the detailed, written elaboration of the result to make an accurate decision regarding the data situation in the requested study region. In our study region in Rio Branco the categorization of the quality criteria is balanced and consequently the result is yellow.

Following properties are registered by the tool: The high number of newly mapped amenities is classified as problematic. The density has increased by 210% in the last year, which indicates that the mapping of amenity features is currently far from a saturated state. This signals that some important reference points which would offer orientation on the sketch maps might be missing.

Particularly important for orientation is the density of the landmarks, e.g., waterbodies, supermarkets, churches or bus stops. With 24.03 features per km² there is a moderate number of such points in this study region. Participants may, but do not necessarily have to face orientation problems due to a shortage of orientation providing features. This should be checked before a prospective field study. Additionally, the mapping of highway features might not be completed yet. During the last year, the feature length increased by 6%. However, this is already close to the 5% threshold.

Additionally, the fact that the highway features were on average last modified only 6 months ago speaks for a proper mapping of highway features. In our experience, data can be classified as current (level = green) if it was edited by the community on average within the last 4 years. With an average of 4 months and 30 days also the last edit of the amenity features indicates currentness of the data.

Based on these attributes, final recommendations are given that are worth considering during a field study. As already indicated, it should be noted that the amenity features in the study region in Rio Branco are probably not yet completely mapped and that there are not many orientation providing features available. It should be investigated beforehand if this could make orientation more difficult for the participants.

After the region has been positively analyzed for suitability, the sketch map for the study region can be created. Such a map is shown in Figure 3 for the example of Rio Branco. This map can be downloaded and saved as a PDF document. For a field study, one can select the preferred format up to A0 and print it for further usage. After it has been used in a field study, the sketch map can be photographed and then digitized using the upload function. See Figure 5 for an example with the area in Rio Branco.

4.2. Case Study—São Paulo

Additionally to the area in Rio Branco, an urban area in M'Boi Mirim, São Paulo, is now being examined. The results show that while the highway mapping saturation is rated yellow as well, the density of POIs and landmarks is higher and rated green,



which causes the general score to be green as well. Therefore, the traffic light indicates a good data situation (green), see Figure 6.

Figure 5. The sketch map of Figure 3 with an example for possible markings, digitized and georeferenced using the upload function. It is shown with an OSM base layer in QGIS [37].

OpenStreetMap Evaluation for Requested Area:

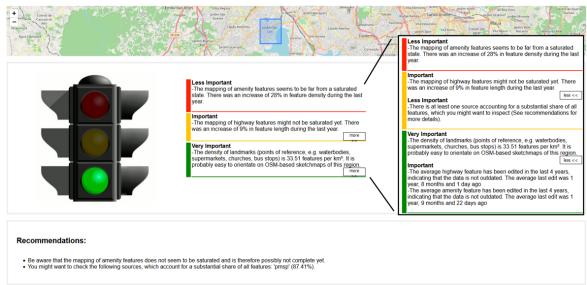


Figure 6. The result of the evaluation of the requested area, here M'Boi Mirim in São Paulo, is visualized using a traffic light. In contrast to the evaluation of Rio Branco, the Sketch Map Tool suggests a good quality of OSM data within this region. Next to the traffic light one finds elaborated and assessed information concerning these OSM data. The lower part contains recommendations to facilitate a field study in this area.

In detail, the highway features are expected to be in a current state (they were last edited on average 1 year and 8 months ago), as well as the amenity features, which were last edited on average 1 year and 10 months ago. The density of landmarks is 33.51 features per km², so that orientation using the sketch map should not be a problem for the participants.

However, the length of the highway features has increased (9%) during the last year, indicating that recently some of them might still be incompletely mapped.

Also, in this study region the mapping of amenity features does not appear to be particularly saturated yet, since there was an increase of 28% during the last year. Nevertheless, it can be seen that the growth is by far not as large as in the previously analyzed region in Rio Branco.

There are two recommendations which should be considered. First, the amenity features might not be mapped completely. Prior to a field study, it should be checked whether this could confuse the participants. Second, a data source was identified in the selected region that accounts for 87.41% of the features. Examination of this source might reveal further knowledge about the precision of mapped features. This source called 'pmsp' guides to the Prefeitura Municipal de São Paulo, i.e., the municipality of São Paulo [38], which only contains data on formal areas, but nothing on slums. This should be considered before carrying out a study in informal areas, which might be even more vulnerable due to less mitigation infrastructure. To counter this lack of data, the Waterproofing Data Project has organized mapping activities.

The generated visualizations of the analyses also enable more in-depth comparisons between different study areas. See Figures 7 and 8 for an example comparing the development of highway feature length in the two discussed areas and the old town of Heidelberg. Although the development in Heidelberg with its decreasing growth suggests a state close to completion, this is not the case for Rio Branco and São Paulo, where the graphs have not flattened yet. A major increase in feature length can be found in all the analyzed areas, suggesting that no lack of data regarding highway features is present. These visualizations also point at the difficulty of this analysis: Both the highway features in Rio Branco and in São Paulo have experienced a phase with very little growth, which would have been interpreted as a state close to mapping completion. Conversely, there is less ambiguity when detecting relative growth above the mentioned thresholds, as this implies that there were some features not yet mapped and probably still are. It also must be considered that from this analysis it cannot be inferred, which kinds of *highway* features are still being added. The effect of incompletely mapped sidewalks (which can have a highway tag [39]), for example, is less problematic than that of incompletely mapped roads in residential areas. Generally, it is preferable that the mapping is close to a state of completion for *highway* features of all kinds, which is why the results of this analysis including the visualizations can still provide useful insights for the user despite the mentioned possible difficulties.

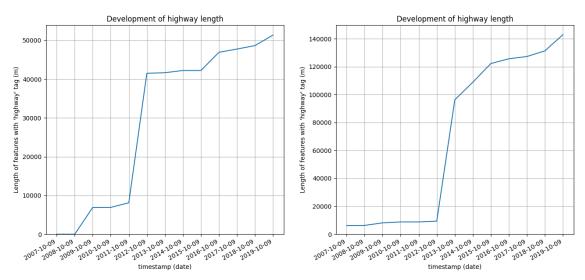


Figure 7. Development of the length of OSM features tagged with *highway* = * in the discussed study areas in Rio Branco (**left**) and São Paulo (**right**).

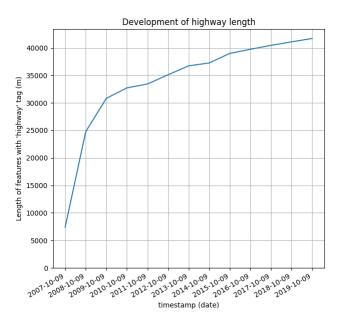


Figure 8. Development of the length of OSM features tagged with *highway* = * in the old town of Heidelberg (for comparison purposes).

5. Discussion

All used thresholds and weights for the analyses are chosen based on either previous research or the ongoing work in the Waterproofing Data Project (see Section 3.4). It is to be expected that broad practical usage of the tool will lead to new insights and potentially the adjustment of these values. The possibility of this adjustment also opens the opportunity to use this tool in other settings and to weight and carry out the analyses as needed for individual use cases. It is also likely that more analyses will be added. The saturation analysis for the road network can be extended to better take into account the difficulties with such analyses explained in Section 4.2, for example, different categories of highway features can be inspected separately, and the development of feature length can be fitted to a sigmoid curve to analyze its shape even better. Another interesting quality indicator for OSM data that could be added to the analyses is the community activity. Seto, Kanasugi and Nishimura [40] analyzed the duration of resolving note tags, which might be connected to the presence or absence of local contributors. With a similar rationale, we have been analyzing the duration of *fixme* and related tags. Further research is needed to gain evidence on the relationship between such tags and community activity. If this relationship is as presumed, it will be useful to add this as an indicator of an active community and thus a presumed higher quality. Another indicator that we have been inspecting and that might be added is the positional stability of features, i.e., the number or extent of changes of their mapped position. A high fluctuation in the positioning of features might lead to the conclusion that there has been some inaccuracy in mapping them at first that might still affect many features. On the other hand, this might also lead to the conclusion that although at first many features might have been positioned inaccurately, the big number of changes indicates that the community is working on correcting such mistakes, which potentially indicates a better quality. Therefore, further evidence is needed in this case as well. Finally, the usability of the tool is essential for the acceptance by different stakeholders and to make it sustainable. The development of the tool is not only based on our own experiences but also on feedback from members of different organizations worldwide who apply OSM Field Papers in their work.

6. Conclusions

The OpenStreetMap Sketch Map Tool proved to be a valuable method to visualize local knowledge and to communicate results of OSM data quality analyses in an easy way

to various stakeholders. In this way, the approach of participatory mapping is supported. The case studies showed how the tool can be applied within the context of disaster risk reduction. Results depict that the method allows evaluation of the fitness of the OSM data for the application of the sketch maps in a specific study area. Local communities can improve these OSM data themselves via mapping activities, for example, which is already done within the Waterproofing Data Project and analyses about their impact are work in progress. Moreover, sustainability will be given by the automatized components and the availability as an open-source tool.

The Sketch Map Tool is work in progress and there are many possible improvements to help the broad group of potential users even better. Different OSM map styles might be used for use case specific renderings of sketch maps, options to separate a study area using a grid tool (Fieldpapers.org [12] already provides this useful option) are planned to be included, and most importantly the further processing of the data will be automatized, i.e., the markings on sketch maps will be detected and evaluated automatically, which has already been partly realized [13]. The result reports of the analyses will be enriched by further explanations and visualizations of the data and their inspection, all integrated into a downloadable PDF file which also serves to make this tool better accessible to laypeople and citizen scientists. The Sketch Map Tool will be tested in the field to evaluate its usability in detail.

Author Contributions: Conceptualization: Carolin Klonner; methodology: Carolin Klonner, Maximilian Hartmann, Lily Djami; software: Carolin Klonner, Maximilian Hartmann, Rebecca Dischl, Lily Djami, Martin Raifer; writing—original draft preparation: Carolin Klonner, Maximilian Hartmann, Rebecca Dischl, Lily Djami, Martin Raifer, Fernanda Lima-Silva, Lívia Castro Degrossi, Liana Anderson, Alexander Zipf, João Porto de Albuquerque; writing—review and editing: Carolin Klonner, Maximilian Hartmann, Rebecca Dischl; visualization, Carolin Klonner, Maximilian Hartmann, Rebecca Dischl; supervision: Carolin Klonner; project administration: Carolin Klonner, Alexander Zipf and João Porto de Albuquerque; funding acquisition: Alexander Zipf and João Porto de Albuquerque. All authors have read and agreed to the published version of the manuscript.

Funding: This article is part of the project T2S Waterproofing Data which is financially supported by the Belmont Forum and NORFACE Joint Research Programme on Transformations to Sustainability (https://www.norface.net/program/transformations-to-sustainability/ accessed on 11 November 2020), co-funded by DLR/BMBF (Federal Ministry of Education and Research) as part of its Social-Ecological Research funding priority, ESRC/Global Challenges Research Fund (ES/S006982/1), FAPESP and the European Commission through Horizon 2020. Responsibility for the content of this publication lies with the authors. Fernanda Lima Silva thanks the São Paulo State Research Support Foundation (FAPESP) for the postdoctoral scholarship (project number 2019/06616-0). Lívia Castro Degrossi gratefully acknowledges the funding by the UK Research and Innovation through the GCRF Global Research Translation Award (EPSRC grant: EP/T015683/1). Liana Anderson thanks the National Council for Scientific and Technological Development (CNPq), Brazil, for the Productivity scholarship (process: 314473/2020-3).

Acknowledgments: We would like to thank the Waterproofing Data, GIScience and HeiGIT teams for their feedback on the tool.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

- Clark-Ginsberg, A. Participatory risk network analysis: A tool for disaster reduction practitioners. *Int. J. Disaster Risk Reduct.* 2017, 21, 430–437. [CrossRef]
- Cheung, W.; Houston, D.; Schubert, J.E.; Basolo, V.; Feldman, D.; Matthew, R.; Sanders, B.F.; Karlin, B.; Goodrich, K.A.; Contreras, S.L.; et al. Integrating resident digital sketch maps with expert knowledge to assess spatial knowledge of flood risk: A case study of participatory mapping in Newport Beach, California. *Appl. Geogr.* 2016, 74, 56–64. [CrossRef]
- O'Neill, E.; Brennan, M.; Brereton, F.; Shahumyan, H. Exploring a spatial statistical approach to quantify flood risk perception using cognitive maps. *Nat. Hazards* 2015, *76*, 1573–1601. [CrossRef]

- 4. Klonner, C.; Usón, T.J.; Marx, S.; Mocnik, F.-B.; Höfle, B. Capturing Flood Risk Perception via Sketch Maps. *ISPRS Int. J. -Geo-Inf.* **2018**, *7*, 359. [CrossRef]
- 5. Boschmann, E.E.; Cubbon, E. Sketch Maps and Qualitative GIS: Using Cartographies of Individual Spatial Narratives in Geographic Research. *Prof. Geogr.* 2014, *66*, 236–248. [CrossRef]
- Bustillos Ardaya, A.; Evers, M.; Ribbe, L. Integrated Participatory Methodologies for Disaster Risk Reduction: Tools to Analyze Complex Systems through Participatory Processes in Brazil. In *Strategies and Tools for a Sustainable Rural Rio de Janeiro*; Nehren, U., Schlüter, S., Raedig, C., Sattler, D., Hissa, H., Eds.; Springer International Publishing: Berlin/Heidelberg, Germany, 2019; pp. 361–376. [CrossRef]
- Klonner, C.; Marx, S.; Usón, T.; Höfle, B. Risk Awareness Maps of Urban Flooding via OSM Field Papers- Case Study Santiago de Chile. In Proceedings of the 13th International Conference on Information Systems for Crisis Response and Management, Rio de Janeiro, Brazil, 22–25 May 2016.
- 8. Klonner, C.; Aeschbach, N.; Usón, T.J.; Höfle, B. Participatory Mapping and Visualization of Flood Risk Perception: An Example from Eberbach, Germany. *Int. J. Disaster Risk Sci.* 2020, 12, 56–71. [CrossRef]
- 9. Degrossi, L.C.; Porto de Albuquerque, J.; Santos Rocha, R.D.; Zipf, A. A taxonomy of quality assessment methods for volunteered and crowdsourced geographic information. *Trans. GIS (TGIS)* **2018**, *22*, 542–560. [CrossRef] [PubMed]
- 10. Rehrl, K.; Gröchenig, S. A Framework for Data-Centric Analysis of Mapping Activity in the Context of Volunteered Geographic Information. *ISPRS Int. J. Geo-Inf.* 2016, *5*, 37. [CrossRef]
- Barron, C.; Neis, P.; Zipf, A. A Comprehensive Framework for Intrinsic OpenStreetMap Quality Analysis. *Trans. GIS* 2014, 18, 877–895. [CrossRef]
- 12. Stamen Design. About Field Papers. Available online: http://fieldpapers.org/about (accessed on 19 October 2020).
- Klonner, C.; Blessing, L. Gathering Local Knowledge for Disaster Risk Reduction: The Use of Sketch Maps in Group Discussions. In Proceedings of the ISCRAM 2019 Conference, Valencia, Spain, 19–22 May 2019; pp. 1397–1398.
- 14. Fitzsimmons, S. Field Papers Printer. Available online: https://github.com/fieldpapers/fp-printer (accessed on 19 October 2020).
- Klonner, C.; Hartmann, M.; Djami, L.; Zipf, A. "Ohsome" OpenStreetMap Data Evaluation: Fitness of Field Papers for Participatory Mapping. In Proceedings of the Academic Track at the State of the Map, Heidelberg, Germany, 21–23 September 2019; Minghini, M., Grinberger, A.Y., Juhász, L., Yeboah, G., Mooney, P., Eds.; Springer: Berlin/Heidelberg, Germany, 2019; pp. 35–36. [CrossRef]
- 16. QR Code Is a Registered Trademark of Denso Wave Incorporated. Available online: https://www.qrcode.com/en/ (accessed on 1 March 2021).
- 17. Python Software Foundation. Python. Available online: https://www.python.org/ (accessed on 31 October 2020).
- 18. WHATWG. HTML Living Standard. Available online: https://html.spec.whatwg.org/ (accessed on 31 October 2020).
- 19. W3C. Cascading Style Sheets Home Page. Available online: https://www.w3.org/Style/CSS/ (accessed on 31 October 2020).
- 20. Ecma International. ECMAScript® 2020 Language Specification. Available online: https://www.ecma-international.org/ecma-262/ (accessed on 31 October 2020).
- 21. Ronacher, A.; Flask Contributors. Flask. Available online: https://palletsprojects.com/p/flask/ (accessed on 31 October 2020).
- 22. Raifer, M.; Troilo, R.; Mocnik, F.; Schott, M. OSHDB—OpenStreetMap History Data Analysis (Version 0.5.10). Zenodo 2020. [CrossRef]
- 23. Raifer, M.; Troilo, R.; Kowatsch, F.; Auer, M.; Loos, L.; Marx, S.; Przybill, K.; Fendrich, S.; Mocnik, F.; Zipf, A. OSHDB: A framework for spatio-temporal analysis of OpenStreetMap history data. *Open Geospat. Data Softw. Stand.* 2019, 4, 3. [CrossRef]
- Auer, M.; Eckle, M.; Fendrich, S.; Kowatsch, F.; Loos, L.; Marx, S.; Raifer, M.; Schott, M.; Troilo, R.; Zipf, A. Ohsome—Eine Plattform zur Analyse raumzeitlicher Entwicklungen von OpenStreetMap-Daten für intrinsische Qualitätsbewertungen. *AGIT J.* 2018, 4, 162–167.
- 25. Pavlenko, A. Mapnik. 2018. Available online: https://mapnik.org/ (accessed on 20 October 2020).
- 26. Zverev, I. Nik4. Available online: https://github.com/Zverik/Nik4 (accessed on 20 October 2020).
- 27. OpenStreetMap Wiki Contributors. Overpass API. OpenStreetMap Wiki. Available online: https://wiki.openstreetmap.org/w/ index.php?title=Overpass_API&oldid=2057486#Public_Overpass_API_instances (accessed on 7 December 2020).
- OpenStreetMap Carto Contributors. OpenStreetMap Carto. N.D. Available online: https://github.com/gravitystorm/openstreetmap-carto (accessed on 31 October 2020).
- 29. Karl Wettin. Carto-Style-Hydda. Available online: https://github.com/karlwettin/carto-style-hydda (accessed on 31 October 2020).
- 30. OpenCV Team. About [OpenCV]. 2020. Available online: https://opencv.org/about/ (accessed on 20 October 2020).
- Leutenegger, S.; Chli, M.; Siegwart, R.Y. BRISK: Binary robust invariant scalable keypoints. In Proceedings of the 2011 International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011; pp. 2548–2555.
- 32. Warmerdam, F.; Rouault, E.; GDAL Contributors. GDAL. Available online: https://gdal.org/index.html (accessed on 31 October 2020).
- 33. OpenStreetMap Wiki Contributors. Key: Amenity. OpenStreetMap Wiki. Available online: https://wiki.openstreetmap.org/w/ index.php?title=Key:amenity&oldid=1989135 (accessed on 7 December 2020).
- 34. Gröchenig, S.; Brunauer, R.; Rehrl, K. Estimating Completeness of VGI Datasets by Analyzing Community Activity over Time Periods. In *Connecting a Digital Europe Through Location and Place*; Lecture Notes in Geoinformation and Cartography; Huerta, J., Schade, S., Granell, C., Eds.; Springer International Publishing: Berlin/Heidelberg, Germany, 2014. [CrossRef]

- 35. Touya, G.; Brando-Escobar, C. Detecting Level of Detail Inconsistencies in VGI Datasets. *Cartogr. Int. J. Geogr. Inf. Geovisualization* **2013**, *48*, 134–143. [CrossRef]
- 36. OpenStreetMap Wiki Contributors. Key: Source. OpenStreetMap Wiki. Available online: https://wiki.openstreetmap.org/w/index.php?title=Key:source&oldid=2061096 (accessed on 7 December 2020).
- Open Source Geospatial Foundation Project. QGIS Geographic Information System. 2020. Available online: https://qgis.org/en/ site/ (accessed on 7 December 2020).
- 38. OpenStreetMap Wiki Contributors. PMSP Buildings Import. OpenStreetMap Wiki. Available online: https://wiki.openstreetmap.org/w/index.php?title=PMSP_Buildings_Import&oldid=1892323 (accessed on 7 December 2020).
- 39. OpenStreetMap Wiki Contributors. Highways. OpenStreetMap Wiki. Available online: https://wiki.openstreetmap.org/w/index.php?title=Highways&oldid=2026021 (accessed on 7 December 2020).
- 40. Seto, T.; Kanasugi, H.; Nishimura, Y. Quality Verification of Volunteered Geographic Information Using OSM Notes Data in a Global Context. *ISPRS Int. J. Geo-Inf.* 2020, *9*, 372. [CrossRef]