Masters
Program
in **Geospatial**
Technologies

Evaluation of a Volunteered Geographical Information Trust Measure in the case of OpenStreetMap

René Theodore Anton de Groot

Dissertation submitted in partial fulfilment of the requirements
for the Degree of *Master of Science in Geospatial Technologies*

UNIVERSIDADE NOVA ISEGI
UNIVERSITAT JAUME·I
WESTFÄLISCHE WILHELMS-UNIVERSITÄT MÜNSTER

Education and Culture DG

Master of Science in Geospatial Technologies

Erasmus Mundus

EVALUATION OF A VOLUNTEERED GEOGRAPHICAL INFORMATION TRUST MEASURE IN THE CASE OF OPENSTREETMAP

Thesis supervised by:

Doctor Carsten Keßler

Professor Doctor Jorge Mateu

Professor Doctor Pedro Cabral

February 2012

UNIVERSIDADE NOVA ISEGI Estatística e Gestão de Informação

Westfälische Wilhelms-Universität Münster

Universitat Jaume I

ACKNOWLEDGEMENTS

I would like to offer my thanks to my supervisor Dr Carsten Keßler and my co-supervisors Prof Dr Jorge Mateu and Prof Dr Pedro Cabral for their willingness to supervise me. Their interest, support and guidance were of great importance for the progress of this work. At times of despair, their positive attitude helped me to pick up the pace and regain confidence.

I would like to thank my parents Jan and Désirée de Groot, whose door is always open, who are always supportive and whose ears are open for all types of news, good or bad.

I would also like to thank the Schenke family in Münster-Hiltrup for trusting me as a good tenant and being very nice and helpful neighbours. Thanks to them an accommodation problem at the start of the thesis semester was solved.

Thanks as well to my all my friends and class mates. Meeting with them every now and then brought the joy that recharged and stimulated me to stay tuned and make the best of life and work.

I would like to thank the European Commission and the organisers of the Erasmus Mundus Master Programme *Geospatial Technologies* for selecting me and making it possible for me to study and gain skills and knowledge in the field of GI Science. Also the experiences of living in Portugal and Germany and especially studying with, working together with, and making new friends with people from all over the world are for me of great value and I am very thankful for that as well.

ABSTRACT

The presence of Volunteered Geographical Information is attracting research because its high availability and diversity make it an interesting source of information. For many organisations it is important that quality of geographical information is of a certain level. Recent developments in studies related to VGI direct towards the estimation of its quality through the notion of trust as a proxy.

For this thesis is investigated which factors have an important influence on trust and a simple approach was used to come up with an indication of trust levels for geographical features. The indicators were selected based on a literature review and on a dataset extracted from the open mapping project OpenStreetMap. Numbers of users, versions and confirmations were counted or calculated and involved as positive indicators, while numbers of various corrections were treated as indicators having a negative influence on the development of trust in information.

Analysis of the dataset and thinking about how to incorporate what in the trust measure showed for example how ideas about time decay could be different. Importance of tags was determined based on a method adopted from documentation studies and applied on the dataset. It allowed for generating lists of tags to be described when publishing information about particular features. This was of importance for assessing information completeness in measuring the quality of the data.

The results of the trust measure have been compared to those if the quality measure and an evaluation of this comparison shows significant signs of support for the hypothesis that VGI data quality can be estimated based on a trust model that incorporates data provenance. On the other hand there is also a significant number of features of which both measures show opposite indications of quality.

Various single assumptions, simplifications and the relatively small size of the dataset restricted the possibilities for obtaining more accurate results. Confirmation or denial of the ideas that resulted from this research can be made by enlarging the dataset and experimenting with different methods. Automating all the data processing would be necessary.

KEYWORDS

Data Provenance

Information Completeness

OpenStreetMap (OSM)

Quality of Geographical Information

Tags

Thematic Accuracy

Time decay

Trust

Volunteered Geographic Information (VGI)

ACRONYMS

AOI	Area of Interest
FOI	Feature of Interest
GI	Geographical Information
IFGI	Institute for Geoinformatics (in Münster, Germany)
ISO	International Organisation for Standardisation
JOSM	Java OpenStreetMap Editor
OSM	OpenStreetMap
PGI	Professional Geographical Information
SDI	Spatial Data Infrastructure(s)
TF-IDF	Term Frequency - Inverse Document Frequency
TF-IFF	Tag Frequnency - Inverse Feature Type Frequency
VGI	Volunteered Geographical Information

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

The carry out of this thesis is an attempt to develop an intuitive method to measure trust in volunteered geographical information (VGI) [10], acquired from the open collaborative mapping service OpenStreetMap[1] (OSM). The method will be evaluated and the outcome of this evaluation helps finding ways to more accurately estimate quality of this type of information.

A literature study was done to become familiar with this emerging subfield within Geographic Information Science as well as to become aware of the various insights and research carried out so far. The research done for this thesis is particularly inspired by the work on trust and provenance modelling done by scientists at the Institute for Geoinformatics (IFGI), University of Muenster, Germany [17]. They propose a model to determine quality of various geographical features from the online mapping project through potential implications that patterns in a geographical feature's historical information could have on trust.

A comparison between the outcomes of a trust measure and a quality measure shows how well the proposed trust measure assesses the reliability of feature information in the online map from OSM. Working with the data and model results led to new insights and recommendations for future appliance of trust measurement for this mapping service and possibly VGI assessment in general.

1.1 Problem definition & rationale

There are two major developments that trigger research on VGI quality assessment. The first one is that various organisations with Spatial Data Infrastructures (SDIs) are coping with time and speed problems; it takes time to collect, filter and prepare geographical data for professional use. At the same time, letting paid professionals deal with data collection is costly because they are expensive to begin with and secondly, the time necessary for collection multiplies the costs of hiring them [11]. The second development is the growing amount of VGI, which is freely available and produced by relatively many people.

[1] http://www.openstreetmap.org

In between these two developments a tension field arose; on one side there is the costly production of professional geographical information (PGI) (guaranteed quality, according to established standards), on the other side the growing availability of VGI (without quality guarantees and with inconsistency in various ways) thanks to initiatives such as OSM.

It would be beneficial to find a reliable method to filter VGI on its quality. Then it could become useful for various organisations and people who need geographical information (GI) in order to make decisions about space in time. An example for which it is even crucial is the one of early warning systems and hazard assessment (Figure 1) [23]. Knowing what happens where helps distributing emergency services and goods to the right places.

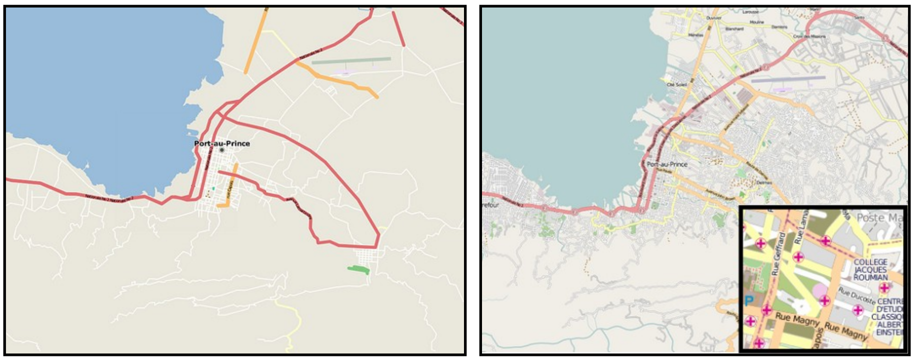


Figure 1: A flurry of activity on OSM after the 2010 Haïti earthquake; 2 days after the earthquake 400 edits had been made, and one day later over 800. Besides infrastructure, the new map provided information about locations of emergency hospitals[2].

One of the most recent research projects in this field is focussed on assessment of VGI quality based on its *provenance* (origin, history) [17]. Based on feature history and producer reputation, users of a feature could potentially be given an indication about the level of trust in the information the feature represents (see section 1.1.3 for a description). For this thesis a practical application has been carried out based on concepts presented in their work, that of others and own insights and assumptions.

[2] http://blog.okfn.org/2010/01/15/open-street-map-community-responds-to-haiti-crisis/

When a feature is being created, changed or assigned a tag in OSM, every instance of one of these processes is saved. Provenance information can be traced back and made explicit. An extended provenance vocabulary allows for recording edits and generating statements about the changes between the instances.

The research in this field is currently at the stage of developing theory and testing assumptions. This is done with the idea of going from coarse to fine, to get useful first insights. An attempt has been made to test theories about trust indication, look at test results and see how well the outcome corresponds with reality. It is a step towards filling the need for evaluation, further refinement and new research directives.

1.2 Aims and Objectives

The main hypothesis is:

Parameters derived from OSM provenance data, including pattern occurrences, determine a feature's trustworthiness. Its trustworthiness is an indicator for data quality and therefore, these two concepts are correlated.

The aim is to evaluate an approach to develop and apply a method to measure trust and to see whether the trust measure result correlates with the dataset's quality by carrying out two parallel assessments; one based on trust indicators, the other one based on higher quality reference data. The outcome of a comparison between the two assessments should confirm or deny the hypothesis, providing new insights about this issue and give directions on further development of VGI quality assessment.

To test the hypothesis, a few questions had to be answered. The main question is *how well the outcome of a trust measure based on proposed indicators corresponds with quality*. Other questions that are important before and after answering the main question are:

1) What are the important factors and parameters to include in VGI trust measure?
2) Which quality aspects can be tested and how?
3) How can the trust measure be concretely implemented?
4) How could the measure's performance be improved?

Achievement of the following objectives was conditional for the process:

1) Getting (the most recent) insights from literature.
2) Defining the area and dataset on which to focus.
3) Collecting parameters of trust from the dataset, keeping in mind what is considered as important according to the acquired insights from literature.
4) Define a model through which a value for trust can be determined and apply it on a selected sample set of features.
5) Investigate which aspects of geographical information quality can be tested.
6) Carry out a field survey to compare (the most recent) OSM feature information with the real situation.
7) Combine for each measure the individual elements to get one value.
8) Compare results of the trust and quality determinations and evaluate how well the trust measure and quality test correspond.

1.3 Research Approach and Methodology

For more clarity, the research approach is represented in a schematic view in Figure 2. OSM history data is available and was analysed to find out which direct and indirect information can be extracted and what variables can thus be involved in a model that allows for derivation of informational trust values. Literature research revealed which variables could be important. It supports the model framework and parameter usage (the words *variable* and *parameter* are used interchangeably).

During the field survey a set of features were checked on their properties by visiting them physically. A sample set was extracted from the data with the intention to have it small enough for a field check within given time and the features with sufficient historical information. All the feature's properties of this sample set were collected and recorded. For comparison, the dataset with selected features (same set as the field sample set) was processed and assigned trust values based on its historical feature information. These values can be seen as proxies for data quality. The same dataset was also assessed against higher quality reference data, acquired through the field survey, knowledge about the topological relations of the features, and information completeness compared with to a list of 'obligatory tags' per feature (inferred from the entire OSM dataset). A comparison between the two test outcomes provided an indication of how well the model's performance is.

An evaluation after the comparison concludes the research and judges whether the proposed approach is satisfyingly contributing to VGI assessment.

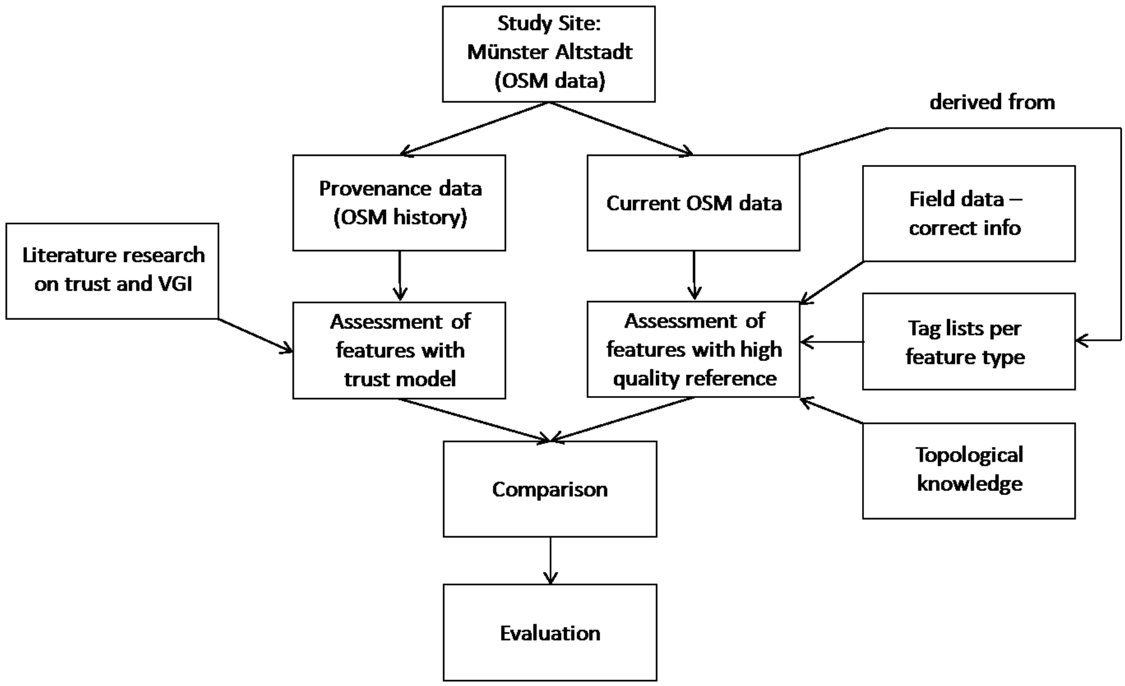


Figure 2: Schematic view of the research approach. Two assessments are carried out: one based on concepts of trust and one with regards to reference data. A comparison and evaluation determines their correspondence.

1.4 Scope

The scope of this project is determined by the idea, the application (OSM) and by time. The trust measure will be evaluated by applying a derivative of current views of how to approach this on the OSM dataset of the Münster Altstadt area and comparing it with results from an assessment based on a three quality elements that are reviewed in professional quality measure standards. OSM supports a limited set of parameters that can be involved in a calculation model and time restricts the field sample size. Because the area focussed on is one with good OSM coverage and relatively long history, the results of this research will only be valid for areas of which feature types and distribution in OSM show similar characteristics.

2 RELATED WORK

A literature study showed what has become clear from research in this field up to now. Most important elements determining trust have been taken into consideration. Claims and ideas about their potential indicative information regarding data quality have been collected and studied to determine their role in assessing VGI quality. Afterwards, a simple model was created for a large part based on the concepts collected for this chapter.

2.1 Volunteered Geographical Information

Volunteered geographical information is a term that was first introduced in 2007 [10]. It is geographic information that is *provided voluntarily by individuals*. Predominantly the Internet is being used as a medium to create, collect and spread VGI.

Traditional GIS sources provide data that is collected and processed based on methods and standards that assure a high quality of information. VGI data origins are mostly unclear. Table 1 (page 18) shows the main differences between the two main types of digital GI [25]. Ideally, professional GIS organisations and professionals in need for data would go for traditional GIS data. However, as stated in the rationale, since cost and time are issues that outgrow the quality guarantee that comes with this type of data, the pressure grows to find alternatives. Alternative data can be extracted from OSM. Another initiative is for example Wikimapia[3], also an online map, privately owned, with satellite imagery and it combines Google Maps with a wiki system (figure 3).

---

[3] http://wikimapia.org

Figure 3: Screenshot of Wikimapia, Münster downtown area. Wikimapia is another open map initiative.

Because of its high availability, it does make sense to look for alternatives in the direction of VGI and its characteristics of abundance and up to date information. On the other hand, there are problems related to data quality, unpredictable data density and inconsistent and uneven distribution. This thesis is focused on predicting data quality for an area for which the information density is high and the distribution of data even.

Table 1: Differences between traditional GIS data and VGI data [23].

Catalog	Traditional GIS data	VGI data
source	Surveying, mapping, Remote sensing Professional	General public, WWW users
procedure	Centralized production,Strictly professional workflow	Volunteered upload, accumulation of data
Advantages	Accuracy, reliable, standardized	richness, comprehensiveness, promptness,
Disadvantages	High cost, limited type	Non-standard, error involved
Product form	DLG, DRG, DOM, DEM, POI, navi data	2D Vector feature: point, polygon, route, track
Application	GIS System, Spatial analysis, Modeling, scientific data visualization	Online browser, dissemination and sharing of geospatial information and knowledge
Consumer	GIS software，Expert	Mapping website, General public
Evaluation indicator	Precise, accuracy, scale, resolution	Authenticity, up-to-date state, data details
Producing platform	RS processing System, Map digitalizer, image vectorizer	VGI Websites with online data editor
Manage mode	Area divide in horizon direction, layers divide in vertical direction、data file and spatial DB	Simple feature management, no uniform organization method
Infrastructure	Satellite/Airborne Sensor, Surveying GPS, professional GIS software, Workstation computer, Skilled GIS expert user	Internet, Web browser, popular GPS unit, online accessible high resolution satellite image, savvy web-mapping user
Technology background	GIS got rapid development, Applications continued to expand	Internet Booming, pervasive Web, the emerging Web 2.0

2.2 OpenStreetMap

OpenStreetMap[4] is a collaborative project that *creates and provides free geographic data such as street maps to anyone who wants them*. The creators state that *most maps one thinks of as free actually have legal or technical restrictions on their use*. This would hold back people from using them in creative, productive, or unexpected ways [24].

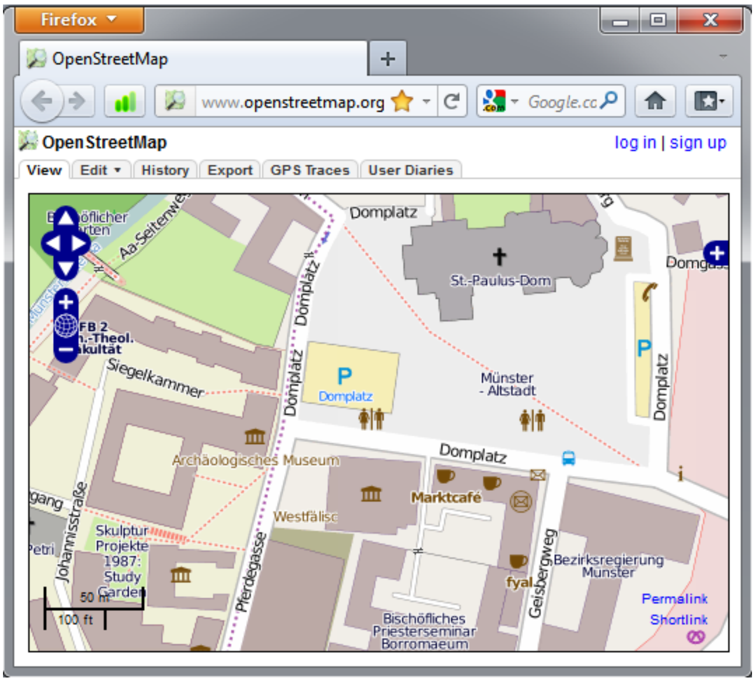


Figure 4: Screenshot of the OSM online map, zoomed in on Münster. The various features are displayed with different types of icons to indicate their nature.

When entering the OSM website, one automatically enters the map interface (Figure 4). The *nominatim* search tool allows for searching places in various ways (name, postal code, street etc.), there are *pan* and *zoom* buttons, there are options to change the base layer, and to add an overlay. The *edit* button shows three options of which two are in-browser editors and one is the Java OpenStreetMap editor (JOSM), one of the OSM editors allowing for offline edits and bulk uploads. The *history* button lets the user view edits (changesets) for the area currently viewed by the map interface. The *export* button allows for exporting the map in a few different formats for further use. *GPS traces* lets users upload and integrate GPS tracks and *user diaries* shows recent user comments about their activities. Lastly, there are various forms of documentation, varying from a help section to guides to licence information when clicking one of the links on the left of the site.

[4] http://www.openstreetmap.org

2.3 Data Quality

Quality has various definitions, depending on the application field. The quality dimensions for making combustion engines are for example very different from those for making a car's frame. The same idea (maybe to lesser extent) also applies to geographical information. Depending on the goal and context the interpretations of *quality* can differ. The International Organisation for Standardisation (ISO) defines GI data quality as *the difference between the dataset and a universe of discourse* [14]. The universe of discourse is the real world view and is defined by a product specification or user requirements; users and producers may have a different universe of discourse and therefore also assess quality differently.

There are many differences that are commonly measured. These are defined in the ISO 19113 standard and showed in Appendix 1. The highlighted data quality elements are those that can be assessed for an OSM dataset. For clear communication in professional GI transfers, these quality elements are measured and the descriptions are given with the product in the form of metadata. The user can then make an informed judgement about the fitness of data for a particular purpose. For OSM data and other forms of VGI metadata is not or very scarcely available. Producers of VGI are mostly contributing information without being aware of data quality elements. Information is corrected or updated by other members of the community based on their ideas of quality.

In spite the efforts of the community, the quality differences between geographical features remain high and extensive metadata is not provided. In an attempt to overcome this problem, various trust and reputation models have been proposed to serve as proxies for data quality [2, 3, 16, and 17].

The important condition for a VGI data quality assessment is to assume that users as a community have a common concept about what high GI data quality generally comprises. Otherwise it is not possible to measure quality with respect to reference data. This assumption can be made because people that are involved in open mapping projects must have an interest in geographic information and it would make sense if these producers and users also share the idea that data quality increases when it is more accurate, complete, consistent and correct.

2.4 Understanding Trust and Related Concepts

The Oxford Student's Dictionary of English has the following definition of *trust*: the belief that somebody is good, honest sincere, etc. and will not try to harm or trick you.[5] This can be seen as a fundamental definition of trust as an important aspect of social interactions. These interactions would not occur when there is no trust between people. In fact, functioning societies rely heavily on trust between members [6, 8, 32] and it makes sense to expect the same in online communities [9]. Trust is also defined (in this particular study area) as *a bet about the future contingent actions of others* [30]. This idea involves mainly interpersonal trust, but informational trust can be linked to that through people-object transitivity of trust [4]. Trust in the context of this work can be seen as the belief that a user has in that an information producer (human sensor) has the intention to be honest and therefore would provide high quality geographical information. And keeping in mind the aspect of GI, there are some important aspects of trust to consider.

One of the spatial aspects of trust is the idea that geographical proximity positively affects trust relationships in communities [5]. Research on friend formation patterns showed that becoming friends with somebody is proportional to the number of people around a user [19]. The reason this might be worth to mention is that becoming friends also requires people to trust each other. In the case of OSM however, it could be less important because the goal of the OSM community is to generate an online map and not to maintain social contacts. Intuitively it also makes sense to assume that geographical proximity positively affects trust in the information about geographical features; when a user is on site he or she is potentially able to publish more accurate information than somebody who is further away from it. This information is however not provided.

Very important is the temporal aspect of trust. It is not something that suddenly exists; it develops over time as a result of continuous interactions. It can however suddenly decrease much, as the consequence of abuse by publishing wrong information [5]. Decay can also be the result of a growing timespan between the last information provision or confirmation and the present. It is obvious that if a piece of information has not been confirmed or updated for a while, a user might wonder whether the information still resembles the current state of the world.

[5] Oxford Student's Dictionary of English, sixth impression 2004

One would most probably search for more up to date information. Therefore, trust in information (informational trust) can decay with time.

Reputation is explained in the dictionary as *the opinion people in general have about what somebody is like (good or bad)*. In [20] reputation is defined as *the subjective perception of trustworthiness inferred from information about the historical behaviour of somebody / something*. From these concepts and from the context of information and trust one can also state that reputation is the knowledge people have about whether a person would provide trustworthy information. This is based on former experiences with the provider. Because everybody can have different experiences, in the end reputation is based on an average of a sequence of information assessments in time, done by multiple users.

Initially, VGI is provided by producers, who are people equipped with positioning tools, cameras or just pen and paper, and their own individual expertise and views on the world that surrounds them. These producers themselves can be trusted or not, in that they provide good information. Users try this information and collectively determine whether this information is useful for them, fits the purpose and therefore is of high quality [2]. The producer builds up a good reputation when he continuously provides high quality information. At the same time, the pieces of information that have been created become trustworthy pieces of information, initially through the producer's reputation (people-object transitivity) and once being published its trustworthiness will also be determined by other factors, related with the spatial and temporal aspects of the information.

Another term that can often be found among VGI related publications is *credibility*, of which the dictionary gives the following definition: *The quality that somebody has that makes people believe or trust him / her*. This sounds very similar to the definition of trust, however, credibility can be more seen as a property of the person who is trusted by the public, while trust is a property of the public themselves [5]. A more clear way of seeing this could be to state that the trust that is put in the trusted person by somebody should be rewarded to him or her with a certain quality coming from this trusted person, in order to keep the balance unchanged. Others state that credibility has two main dimensions: trustworthiness (with aspects of reputation, reliability and trust) and expertise (with aspects of accuracy, authority and competence) [7].

The notion of credibility seems to involve more aspects according to some, but since information about some of these aspects is not provided in OSM, for this thesis the methodology is based on the insight that is more suitable for OSM: *using trust as proxy to determine VGI data quality*.

It turns out that opinions, understanding and definitions vary. Because *trust* is perhaps more evident to most people than information *credibility* or *credible information*, the proposal here is to use the term and notion of *trustworthy information* as being a proxy for high quality information.

2.5 Making Historical Information Explicit and Deriving Patterns

The various interactions that occur with VGI can be investigated by looking at their provenance. *Provenance* is derived from the French verb *provenir* which means *to come from* and it refers to *the chronology of the ownership or location of an historical object*. Another definition is *the details regarding the sources and origins of information* [1].

OSM has every edit recorded and therefore there is a lot of historical object information available. This offers possibilities towards assessing trustworthiness of OSM features. The method to be evaluated is partly inspired by a provenance model that defines various elements such as actors, executions and artefacts. Every element has attributes with information about its provenance [13]. The method is *data-oriented* because the focus is on the origins of specific data items and not on the processes that generate the data [28]. The provenance model has been used with the types and attributes that resemble specifically those from OSM, allowing for making them explicit [17].

Figure 5 gives a high-level overview of this provenance vocabulary. Elements and properties added to the original provenance vocabulary are highlighted in red. Dashed lines represent relationships of which the element pointed to is a sub class of the main element. The class *Edit* is a central one; it links the feature edit with the editor (*User*), feature state (*FeatureState*) and to what is changed (geometry change, tag addition, tag removal, key value change). *Changesets* are collections of edits that were done by one user in one session.

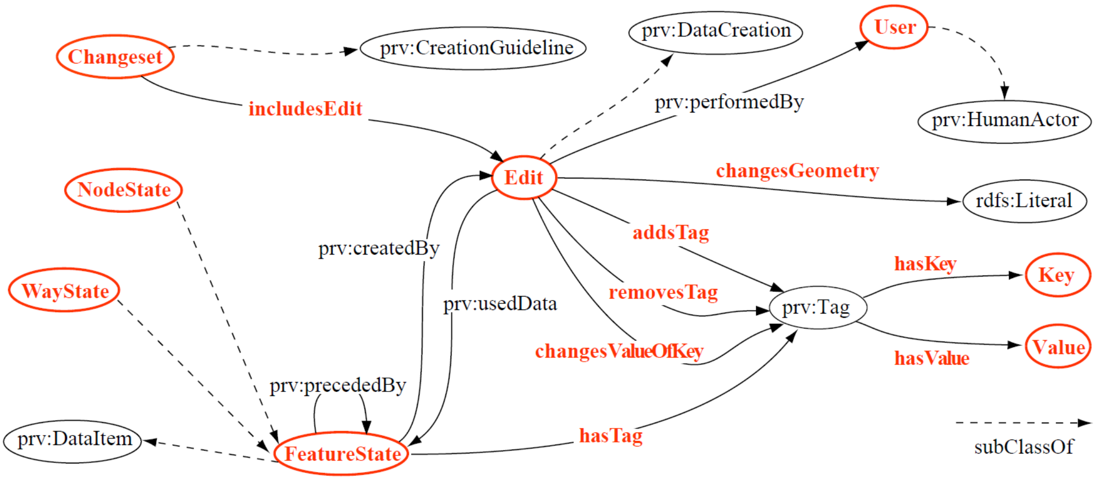


Figure 5: High-level overview of the OSM provenance vocabulary proposed by [17] with edit as a central class through which various actions can occur.

Different versions of a feature are connected through this model (prv: precededBy), which links the edits between two versions of a feature. Based on this model's properties, the recorded instances of features in OSM can be compared and patterns can be derived (Figure 6). These schemes helped in determining the variables to look for in the data.

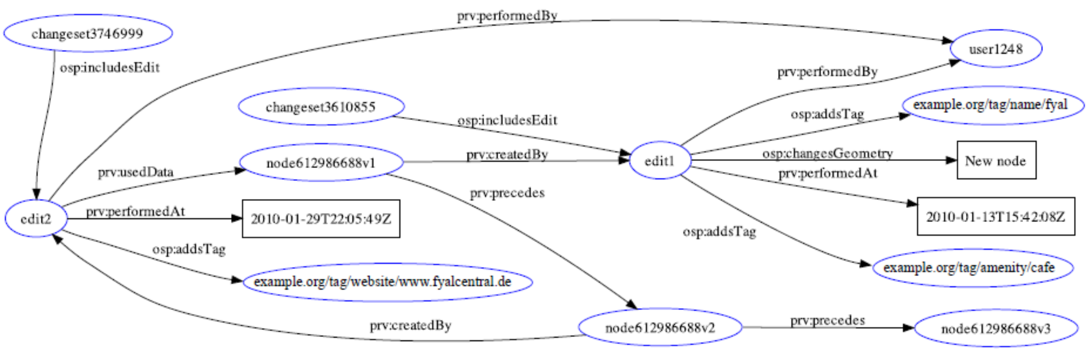


Figure 6: Simplified provenance graph [17]; from one to the next edit.

The following patterns can be derived when looking at the provenance of a feature:

- ✓ Confirmations

  When users check information about a feature without changing it, the more likely it is this information is correct.

- ✓ Corrections

  In case information about a feature is wrong, it is corrected by other users, causing a change in feature state from one to the other instance.

- ✓ Rollbacks

  A rollback occurs when a feature state is reverted to the previous state. Chronologically this means that a feature went from one state to another, after which the 3rd state represents the same information as the 1st state.

- ✓ Self-rollbacks

  This occurs when the same user reverts his / her own edits, mostly because he or she accidentally made a mistake.

2.6 Idea for Assessing Information Trustworthiness

In [17], the proposal is to calculate an overall value of trustworthiness for a feature by assessing user reputation and calculating informational trust. Feature properties being confirmed, added, changed or deleted influence both the information trustworthiness of the feature as well as the reputation of the producer (the previous editor).

When taking feature history and user reputation into account, it is possible to predict quality: in case information from a particular producer turns out to be of good quality several times in history, this producer will build up a good reputation. User reputation is assessed by dividing the number of contributions by the number of corrections and rollbacks (contributions that received negative feedback). To make different users' reputation values comparable, the result of this absolute value should be normalised. Assessing informational trust is done by assigning trust values to every statement based on the feature's provenance. A trust value increases with every confirmation and is also dependant on user reputation. Eventually, an overall trust value for the whole feature is calculated by taking into account the trust values of all statements the feature is based on, while each statement is weighted by the reputation value of the user responsible for the particular statement.

2.7 Uncertainties

A trust and reputation system seems to be a solution worth to try in order to assess VGI quality. But there are various uncertainties that could throw a spanner in that. First of all, there is the definition of quality: the fitness for purpose [14] or the relative value for someone specifically [2].

Since individual users can have different ideas and purposes, it could be hard to determine general quality of OSM information. It could be that people in a community influence each other, which creates bias in the information pool. An example would be a member of the OSM community who has built up a good reputation (indirectly, because in OSM there is no user rating system) and other members therefore do not check anymore whether his / her information is correct. At the same time it could be that this producing member starts providing wrong information because of whichever new intention. Also, information in general could become biased when it is provided under the influence of trends / popularity or group dynamics. Individual qualities of people (human sensors) could then vanish, leading to wrong information coming from a large group of people [18]. A study on these dynamics and interpersonal relations through blogs and / or ratings could reveal why they occur and how they are formed.

As stated earlier, many users of geographical information are also interested in temporal characteristics of spatial objects. Things that are important are among others the temporal span and validity of the data, temporal extent and accuracy. Different spatial objects have different temporal characteristics [29]. It is easily imaginable that a construction site requires its information to be updated a lot more frequently than a downtown church that has been there already for centuries and will be there most probably some centuries more. Tags can change in time; the construction site can be turned into a shopping centre as soon as the construction workers leave the place and the building is officially opened.

OpenStreetMap mentions in its Wiki that there are not any content restrictions on tags.[6] Values however should be verifiable (factual observations). In the wiki is also stated that it would be beneficial when agreements are made about sets of features and corresponding tags. It is imaginable though that because of the freedom in the usage of tags there are semantic problems. Users / producers that are active in a particular area might eventually agree on tags for their local environment. But groups in other areas of the world can have different tags for the same thing. Features that are the same can have different tags due to cultural or linguistic differences.

[6] http://wiki.openstreetmap.org/wiki/Map_Features

2.8 Literature Study Summary and Conclusion

The recent phenomenon of Volunteered Geographic Information concerns large amounts of free data, which makes in interesting when considering rising costs for professional GI data. An example of VGI is data available from OpenStreetMap. Because metadata is not provided, its quality is unknown and it is hard to determine beforehand how well the data fits the purpose. In general, high quality GI is seen as data with high levels of correctness, accuracy and consistency in all ways (as stated also by [14]).

Trust is a concept that comes into play as the way to estimate quality in most scientific publications. Trust and reputation develop over time and can be positively or negatively influenced by space, time and interaction. The idea of how trust and quality are related is as follows: a producer provides geographical feature information (online) and a user decides whether or not to use this information based on the trust he or she has in the information to be of high quality (fits the purpose). Trustworthiness of a feature is based on its history; if repeatedly the information turns out to be good, more trust is gained.

Other things to consider when dealing with VGI are bias (when users influence others), temporal validity of the data and sematic problems with regards to tagging due to personal, cultural or linguistic differences within the community.

Most literature predominantly addresses issues with VGI and provides directives for research. The aim of this work is to provide a method that applies current ideas and generate coarse results. Feature trustworthiness could possibly be determined by taking into account user reputation and informational trust, in turn determined by patterns such as corrections and confirmations in a feature's provenance and other information derived from the data.

The literature review reveals what is considered as important for determining quality based on trust and reputation. But the way the concepts and ideas could be implemented also depended on the dataset that is subject to the assessment. The dataset determines the types and numbers of parameters that could potentially influence information trustworthiness. The next chapter is therefore devoted to data analysis.

3 DATA ANALYSIS

This chapter informs about the raw data, its characteristics and shows filtered and derived information. Derived information is the result of more extensive analysis and findings on how to get desired information. The data were explored, processed and visualised with help of MS Excel and ESRI's ArcGIS software. Interim and final results are stored and available in .xls and .shp formats (available on the CD).

3.1 The Raw Data: Extent, Form and Properties

The raw history file[7] was extracted halfway October and contains every state of every feature recorded within the extent determined by a minimum latitude of 51° 56' 56.6052", maximum of 51° 58' 26.2596", minimum longitude of 7° 27' 2.6382" and maximum of 7° 38' 42.9108". The area of interest (AOI) is meanwhile a relatively small part of this extent (Figure 7).

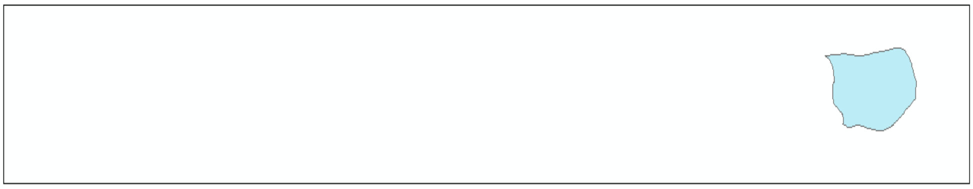


Figure 7: The Münster Altstadt is the area of interest (blue) within the extent (frame) covered by the history file.

The data consists of a large collection of nodes, ways and relations, the elements of OSM (Figure 8). The file was provided by IFGI.


```
<?xml version="1.0" encoding="UTF-8"?>
<osm>
    <node id="21518835" version="1" timestamp="2006-12-10T11:04:03Z" changeset="166491"
    user="SlowRider" uid="3585" visible="true" lat="51.9508583" lon="7.6176480">
  <tag k="created_by" v="JOSM"/>
</node>
<node id="21518835" version="2" timestamp="2008-02-01T20:36:13Z" changeset="7024"
user="SaschaR" uid="8103" visible="true" lat="51.9510336" lon="7.6178164">
  <tag k="created_by" v="JOSM"/>
</node>
<node id="21518835" version="3" timestamp="2008-03-10T21:55:14Z" changeset="300473"
user="Tino" uid="11355" visible="true" lat="51.9510336" lon="7.6178773">
  <tag k="created_by" v="JOSM"/>
</node>
<node id="21518835" version="4" timestamp="2008-07-11T13:31:09Z" changeset="461222"
user="Tino" uid="11355" visible="true" lat="51.9510224" lon="7.6179574">
  <tag k="created_by" v="JOSM"/>
</node>
<node id="21518835" version="5" timestamp="2009-03-22T11:51:04Z" changeset="845019"
user="Gluko" uid="36745" visible="true" lat="51.9510224" lon="7.6179574"/>
<node id="21518836" version="1" timestamp="2006-12-10T11:04:04Z" changeset="166491"
user="SlowRider" uid="3585" visible="true" lat="51.9521713" lon="7.6187092">
  <tag k="created_by" v="JOSM"/>
</node>
```


Figure 8: Screenshot of OSM xml history file. Every edit (version) of a feature is recorded.

[7] CD: ..\HistoryFile\RawFile\muenster-downtown-history.osm

A node is the basic element and has been assigned a latitude and longitude. Optionally an altitude can be assigned as well. A node can represent a standalone feature such as a telephone booth or a place name label. A way is a connection between multiple nodes that represent a linear feature such as a road or power line. Line features like roads can split and a node can therefore belong to multiple ways. Closed ways are used to represent an area (polygon). Relations can be created to group ways and nodes that are geographically connected or adjacent to one another. Nodes and ways can both be part of a relation [24].

3.2 Area of Interest and Data Filtering

Figure 9 shows the part of Muenster with a border that corresponds with the Altstadt (old city) municipal district as defined by Stadt Münster (City of Münster)[8] and the features currently present in OSM for this area. The current OSM data was freely available in shape file format on the website of Geofabrik[9], a consulting company from Karlsruhe involved with the OSM project in various ways. Their server provides an OSM extraction of the Münsterland area.[10] The area of interest was in turn extracted based on the Altstadt border obtained from the 'Gebietsgliederung' (corresponding with the definition of Stadt Münster) of which a shape file was provided by IFGI.

The old city centre is a popular area where people go for shopping, sightseeing and other forms of leisure. Because of the high interest in the area, there is a lot of data available. This way, problems like uneven distribution and density of data are avoided. These are VGI related issues that are outside of the scope of this thesis.

[8] http://www.muenster.de/stadt/stadtplanung/statistik.html
[9] http://www.geofabrik.de
[10] http://download.geofabrik.de/osm/europe/germany/nordrhein-westfalen/muenster.shp.zip

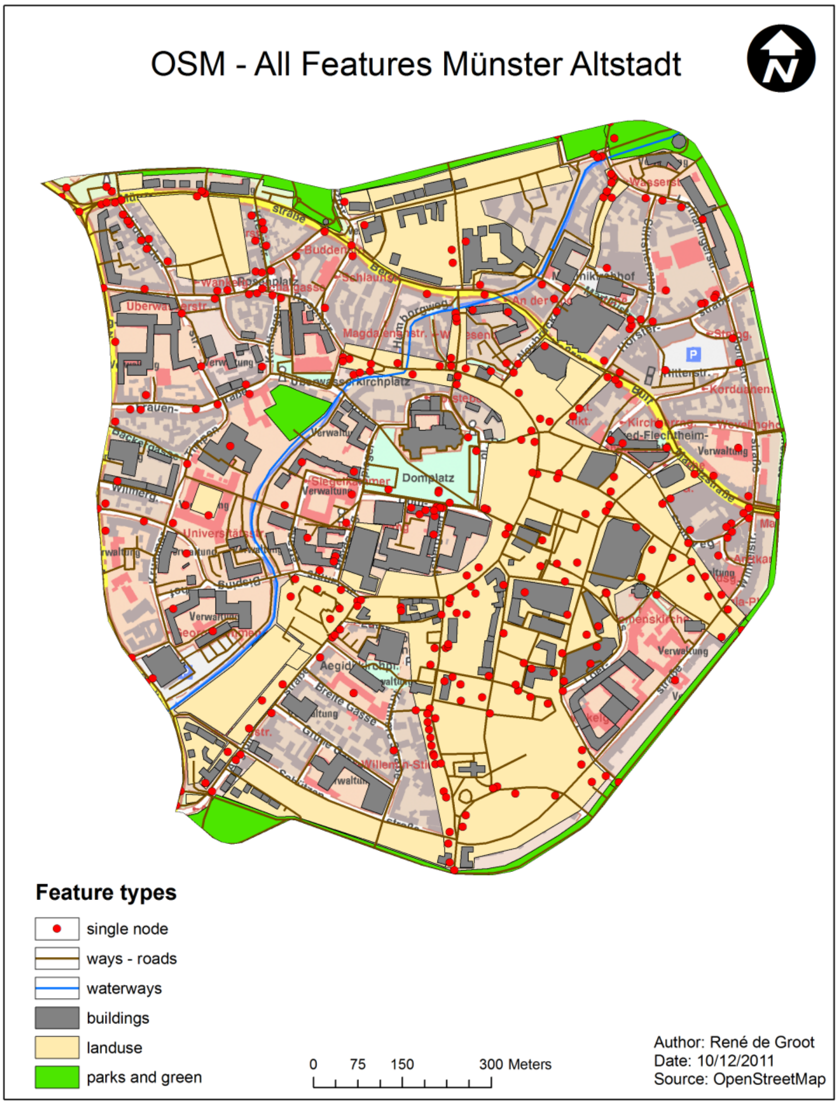


Figure 9: All the features currently present in OpenStreetMap for the Münster Altstadt area.

The features left within this area were further filtered on the number of versions. It was important to find a balance between having enough versions per feature (to have a history long enough to measure trust) and not too many features to deal with for the field survey. It was estimated that applying a filter of 6 or more versions per feature would be the best setting. The remaining features[11] are visualised on the map in Figure 10.

[11] CD: ..\Sheets\HistAltstadt6OrMore.xls

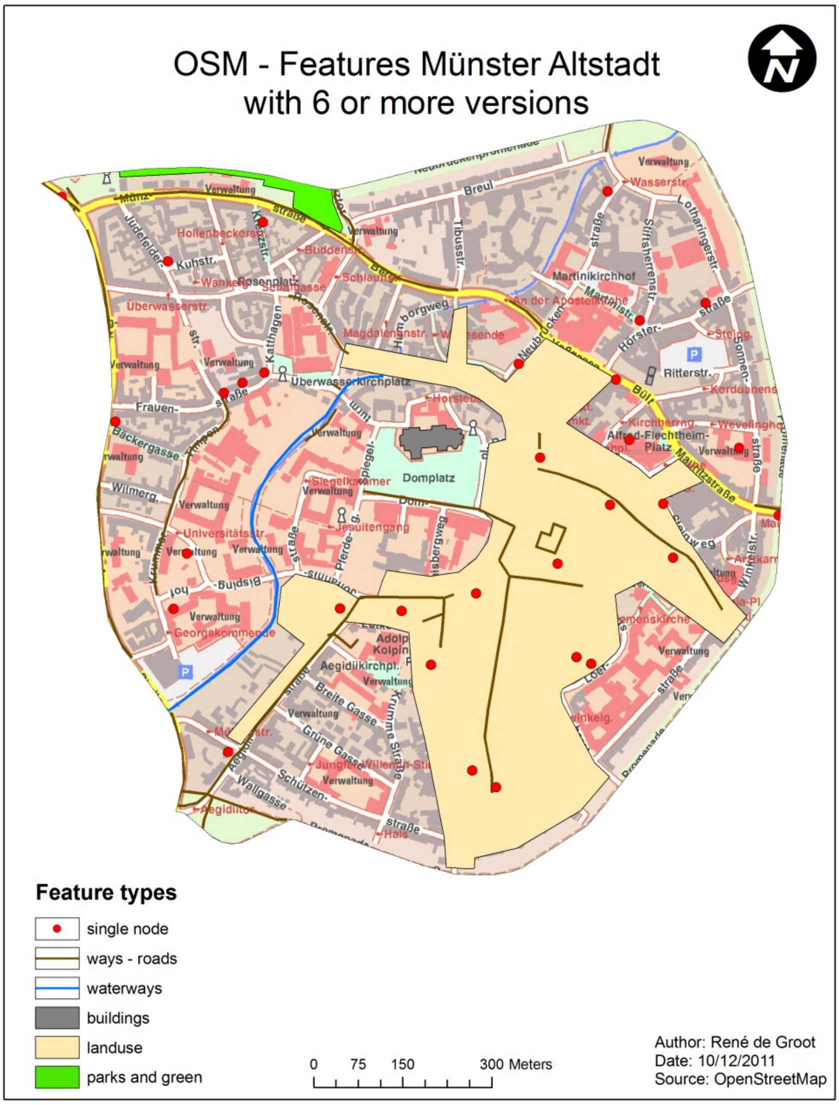


Figure 10: Features remaining after setting a minimum of 6 for the number of versions. These 74 features at the same time offer a suitable sample set for the field survey.

Appendix 2 shows an overview of the method used to come to this result.

3.3 Patterns and Numbers

3.3.1 Counts

Various types of numbers can be extracted from the data (Table 2). These numbers are the result of the interaction within the community and could have an influence on trust. Appendix 3 can be consulted for an overview of the processes used to obtain the results.

Table 2: Factors that could have an influence on trustworthiness.

number of versions
number of indirect confirmations
number of direct confirmations
number of users
number of geometrical or positional corrections
number of tag corrections
number of main tag corrections
number of rollbacks
number of tag additions
number of tag removals
number of tags
number of days since last update

The number of versions is directly coming from the history file, and it represents how many times a feature's information was recorded and thus how many direct interactions occurred regarding the feature of interest. In general, the information in every version is different. The latest version contains the most up-to-date information. The higher the number of versions is, the more lineage data there is to analyse. Therefore this number could have an important influence of trustworthiness. The higher the number of users, the more different users were involved with the feature. It makes sense to assume that collaborative development improves the work (the *many eyes principle*) [31]; the higher this number is the more correct and trustworthy the information should be. Assuming no group trends or popularity, more users should decrease bias and the more users agree, the more likely it is that the information is correct. The number of geometrical corrections shows how many times a feature's position was changed. In the case of point features it involves one coordinate change. In the case of line or polygon (closed line) features it involves correction, addition or removal of one or several points. The number of tag corrections represents how many times a tag has been changed. Hereby also main tag corrections were separately recorded. A main tag correction could be of higher importance because it completely changes the feature type. Rollbacks (complete reversion to the previous state) were very rare among the selected features.

Rollbacks have an important impact, because it brings a feature's information back to the previous state. Tag additions and removals could change the completeness and correctness of information. And a higher number of tags could indicate higher information completeness. Depending on the type of feature however, there could be omission or commission of tags. Time since last update could potentially also affect the information trustworthiness, decaying from the time since the last update. How it affects the information depends on the time duration and the feature type. The result of the count[12] is showed in Appendix 4.

3.3.2 Confirmations

While going through the history of features with more than 6 versions, very occasionally two consecutive lines (representing two consecutive versions) state exactly the same information. When the feature state of a version is the same as the one of the previous version, it could mean a direct confirmation. Here 'could' is used, because other reasons cannot be excluded. Sometimes there is a suspicious repetition of a pattern in user names: several different sets of lines of consecutive versions contain the same user name in the latest version. It almost seems that this user wants to have his / her name stated in the most recent version of features.

For the number of indirect confirmations, an idea of [17] was initially adopted, proposing to use the extents of changesets as implicit confirmations; the feature in consideration could be within these extents. If that is the case, it means the feature has been implicitly confirmed by the user who was responsible for the changeset. However, a closer look at the data revealed that not all of the features of interest are covered by changeset extents (Figure 11).

[12] CD: ..\Sheets\Features&Counts.xls

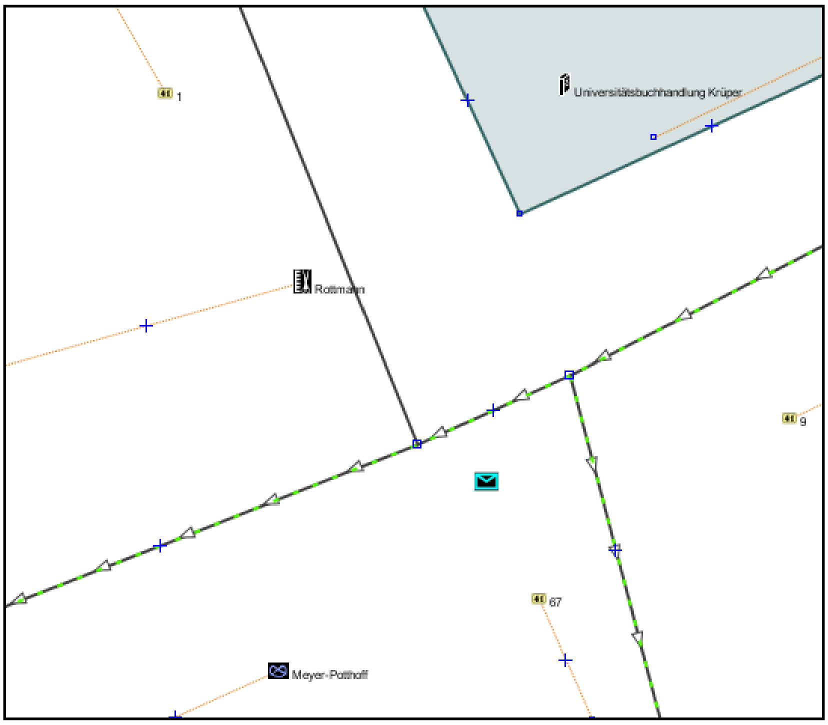


Figure 11: Screenshot (negative) from JOSM of part of the Münster Altstadt. A book shop in the upper right corner has an underlying polygon feature, of which its changesets made after the FOI's most recent timestamp confirm its information. Other features are not covered by changesets belonging to other features.

Therefore the decision was made to involve all the changesets[13] (recorded in the time after the last update of the feature of interest) that fall (partly) within a buffer of 50 metres around the feature of interest (Figure 12).

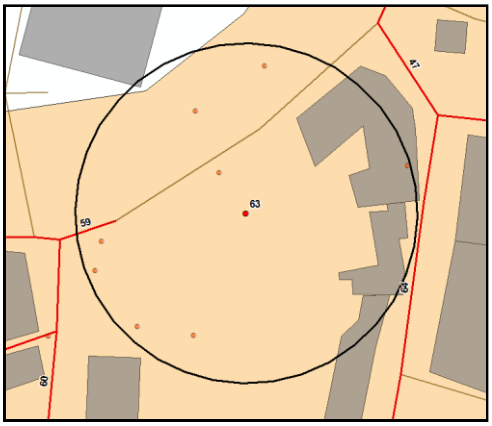


Figure 12: Screenshot from ArcMap, example of a 50 metre buffer around a feature of interest.

[13] CD: ..\Sheets\ChangesetsPerFeature.xls

There is not a clear justification for the 50 metres, but within this distance the assumption can be made that users are aware of the presence of the feature of interest (FOI) and therefore should know something about the correctness of its properties.

After creating the buffers[14], the IDs of the features falling within the buffer[15] were related to the history file and filtered on whether they had a date more recent than the last version of the feature of interest (Appendix 3). The assumption was made that the most recent version of a feature is the most correct one and thus the one for which the implicit confirmations are important. Per FOI these recent changesets of features falling within its 50 metre buffer were counted to get the number of indirect confirmations.

3.3.3 Time Decay

An attempt (Appendix 5) was made to see if the expected differences in time decay among the different feature types are present and how these would be different. For this, the whole history file of the Münster area has been taken into account. The features types representing the features of interest were extracted and all the timespans between two consecutive versions were queried. Timespans between different versions varied from 0 to more than 1000 days.

A general observation is that in most cases the highest number of days is relatively small in frequency. These higher values are randomly distributed in the line of the feature version sequence; long time spans occur between early versions of features as well as between more recent versions of features. One would expect however that as time goes by, the timespans between subsequent versions become longer. This comes with the idea of a feature being edited less and less because after every edit there should be less to be corrected, especially for features like churches and roads, which do usually not change.

[14] CD: ..\ShapeFiles\SelectedFeatures\Buffers
[15] CD: ..\Sheets\IDsBuffer.xls

For example regarding the *highway* feature types there is a common characteristic visible in the data; values between 1 and 100 are much more frequent than other values (Figure 13).

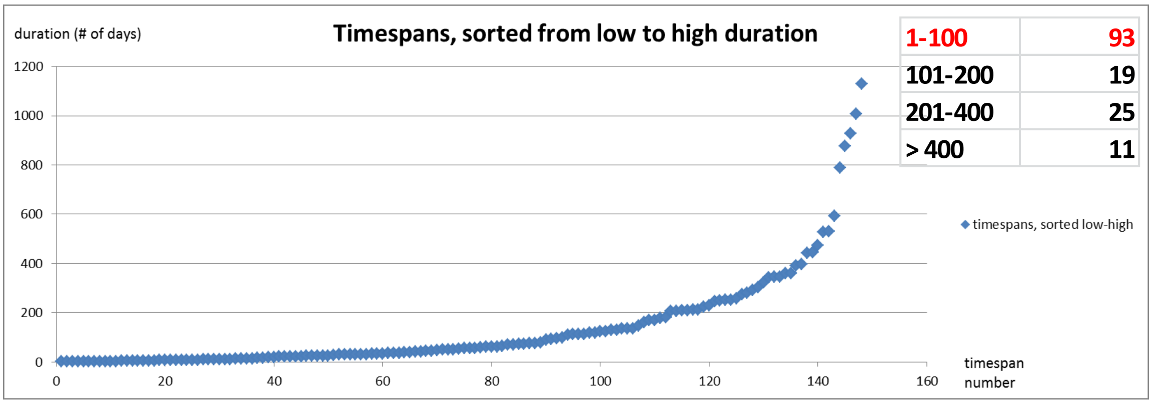


Figure 13: Time decay distribution of highway-cycleway with number of days on y axis

For the other feature types the situation was mostly similar[16]. It is hard to say for some, of which there are not many particular types of these present in the dataset. Based on this exploration there is no evidence of the community holding ideas of specific timespans associated with types of features. For this dataset, it turned out that timespans are randomly distributed in time and across different feature types. However, a more detailed investigation with a larger dataset would be necessary to confirm or deny this idea.

[16] CD: ..\Sheets\TimeSpans.xls

4 TRUST MEASURE

The available dataset is relatively small and this also limits the possibilities for use in the existing approach of [17] that includes user reputation assessment, wherein all the data containing a user's name is needed. The reason for this is based on the assumption that users active in the area of Münster are also active in areas around Münster, in other parts of Germany and maybe even abroad. This does not have to be a bottleneck on the other hand, because patterns in the data provenance can still be counted and incorporated in the trust assessment. That is because these properties eventually have an influence on the overall trustworthiness of a feature anyway. When they initially have a good or bad influence on user reputation, they also have a good or bad influence on feature trustworthiness. Inferred from that the choice was made to apply a direct approach and assume that for example confirmations directly affect information trustworthiness positively and corrections cause a delay in the development of trust in information, indicating instability of the feature.

The trust measure for OSM features that is proposed here has not been implemented so far and with the KISS (Keep It Simple, Stupid!) principle [26] in mind a way to start generating coarse results is by taking available important variables into account with equal importance. The methodology has been suggested based on the properties of the dataset, existing ideas as well as many assumptions and simplifications.

An important uncertainty is there concerning quality. As stated before, quality resembles the fitness for purpose. In the case of OSM, it is very hard to determine the purpose. Every user can have a different purpose. One user could rate a contribution as good, another as bad, depending on whether it serves one's purpose. In spite that, in order to continue with the assessments and evaluation, the assumption can be made that all users have the same general idea about high quality information: as accurate, correct, consistent and complete as possible.

The data analysis helped to see which parameters and patterns are present, which other issues were overlooked and what the data can reveal that is useful for a possible model determination.

4.1 Parameters Involved

The information derived from the data analysis could potentially all tell something about the reliability / trustworthiness of feature information. The number of versions provides more lineage, more historical information and therefore potentially allow for a more accurate determination of trustworthiness. The number of confirmations tells how many times the information of a feature has been confirmed, either directly by a user opening and closing an editing session and not changing the information of the feature, or indirectly by having editing extents overlapping or including the feature in consideration. The number of users is important, because if more people are involved and eventually agree about the information, the higher the chance that the information is correct or at least meets the requirements of the community as a whole. There are cases in which there are over 10 versions for a feature, while only 1 user is responsible for all. There is a high chance of bias here. Corrections decrease the reputation of the previous editor and therefore also negatively influences trust in the feature, as in this case trust in a statement about the incorrect version would be multiplied by a factor representing the decreased reputation of the user who made that statement.

Geometrical and positional corrections have been made. They influence the shape of a feature by moving, deleting or adding nodes, the building stones of a line or polygon, or change the location of a point feature. Tag additions and removals are linked with corrections and show activity and engagement with the feature in history. Since many FOIs are labels, in this case these corrections were not considered as important; labels should be places within the polygon of a building, but not necessarily on a specific spot. It is the information of the label that matters. The number of tags present in the most recent version indicates how much description is published, however it does not tell anything about the content. The information to be described should meet the requirements that are derived from the whole dataset, which is the completeness of information relative to the determined 'obligatory tags' (discussed in section 5.3). These inferred tags will be used as higher quality reference data for which information should be present per feature type. For the time decay no patterns were found that could proof a significant difference in time duration between two subsequent versions.

The parameters taken into account are the ones that are intuitively the most telling and self-contained, backed up by the literature survey and found available in the data. They include the number of:

- ✓ Versions
- ✓ Confirmations
- ✓ Users
- ✓ Tag corrections
- ✓ Rollbacks

A higher number of versions, confirmations and users positively influences trust development, while a higher number of corrections of any kind delays the development of trust because of feature instability (and also indirectly through a downgrading of the reputation of the user who was responsible for publishing the incorrect information).

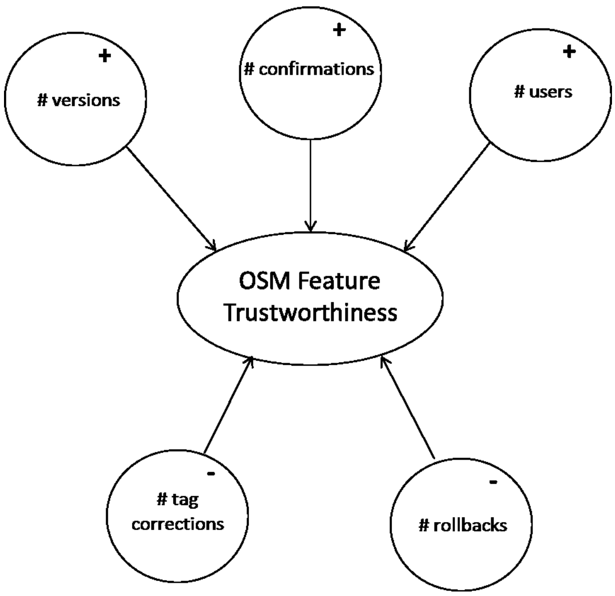


Figure 14: Schematic of the most important influences on feature trustworthiness, their occurrences collected from the provenance of features in the dataset.

4.2 Preparation

Every variable was investigated to check on outliers, which are observations that appear to deviate markedly from other members of the sample in which it occurs [12]. Including outliers may be misleading and have therefore been taken out to avoid the numbers from being undervalued (Figure 15).

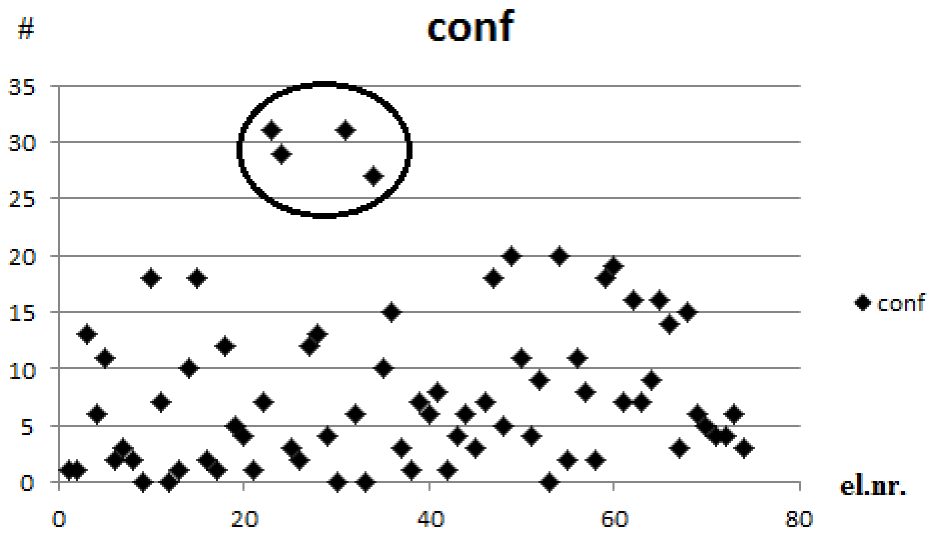


Figure 15: Example of outliers in the case of number of confirmations

The selected variables were prepared[17] and joined to the feature map IDs. All features other than point features were converted to point features and then to raster. This allowed for easier further processing and visualising the outcomes of the trust measure.

The number of geometrical corrections is not taken into account. It turned out there are many features in the selection that are labels. Some of them do also not have an underlying polygon. This could indicate there is less interest from the OSM community in the positional accuracy of a feature.

[17] CD: ..\Sheets\ParametersTrust

4.3 Weighting and Classification

The measure has been kept simple: all variables have the same weight. The sets of numbers have been reclassified based on equal intervals (Figure 16) and for each classification, a map was produced. Equal intervals emphasize the amount of an attribute value relative to other values [21].

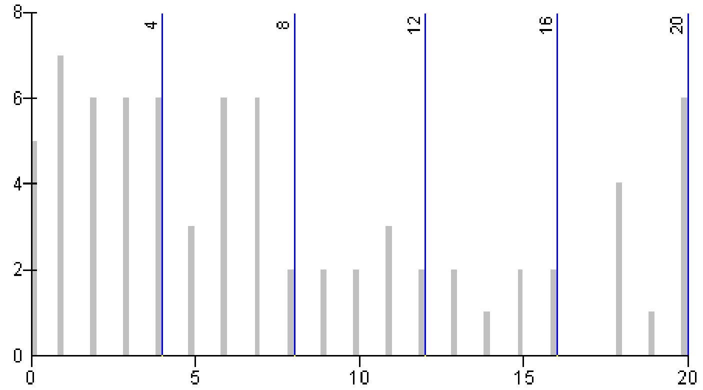


Figure 16: Classification based on equal intervals. All features have been classified into 5 classes equal interval for every parameter count.

4.4 Combined Trust Value

The intermediate map results are shown in Appendix 6.1. The classes for the numbers of versions, confirmations and users represent the values from low to high; a higher class represents higher numbers. The classes for the numbers of tag corrections and rollbacks represent values from high to low; a higher class represents lower numbers. The lower the number of corrections is the less negative influence a feature has from a decreased user reputation and thus the more trustworthy a feature is. All classified values (1-5) were summed up and reclassified, of which the result is shown in Appendix 6.2 and linked to the features in their original form in Figure 17.

Most features are almost equally distributed between the three middle classes (65 of 74), of which class 4 contains the most. Within class 1 there are 8 features and within class 5 only 3. Overall, the trust measure suggests an estimation of the dataset as moderately trustworthy regarding its information quality.

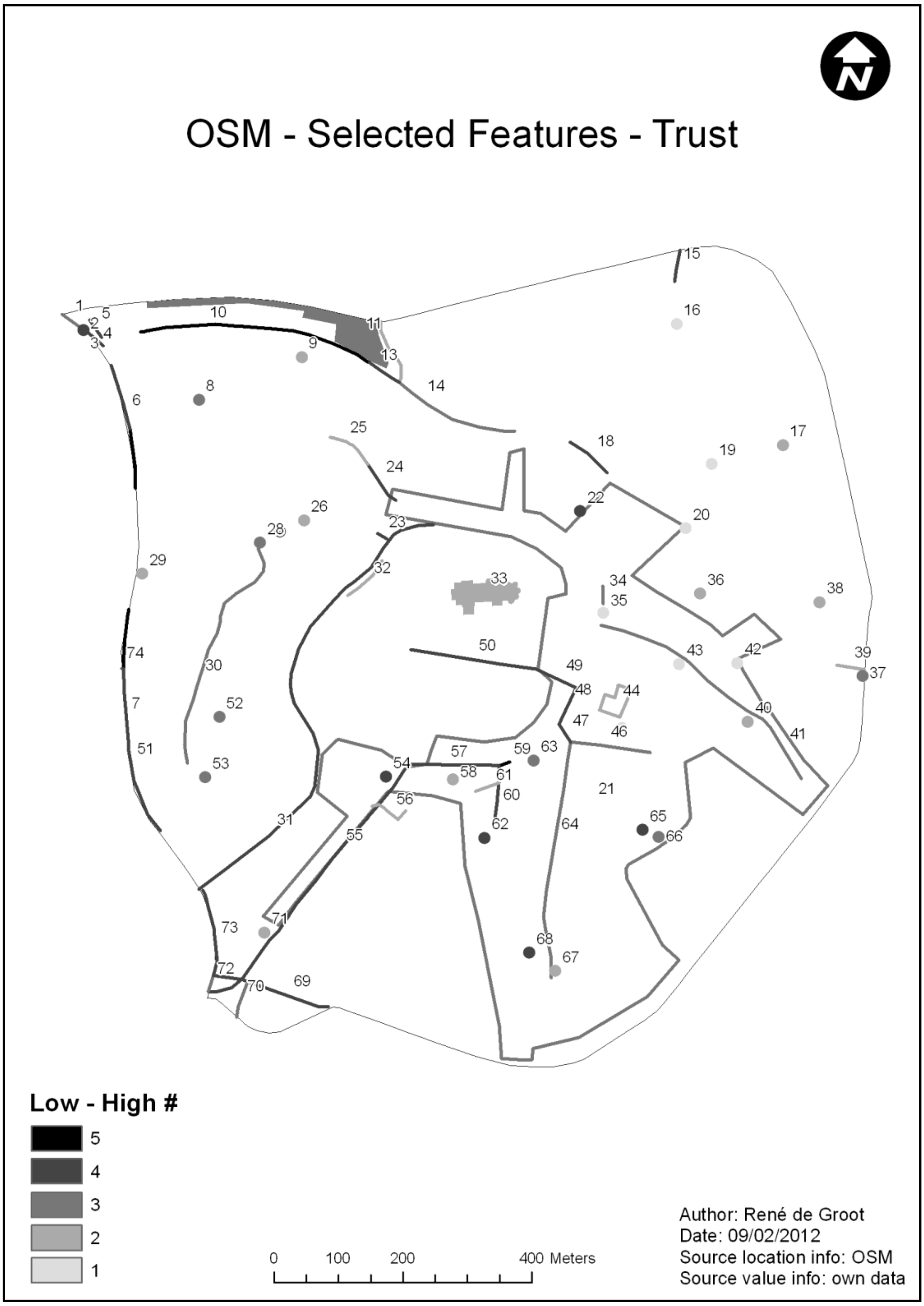


Figure 17: Final map with trust value classes allocated to the features of interest.

5 QUALITY MEASURE

The trust values were compared with the results of a quality assessment based on higher quality reference data, obtained for the quality elements highlighted in Appendix 1.

5.1 Field Survey and Thematic Accuracy

A field survey has been carried out to reveal whether the information given in OSM corresponds with reality as it currently is. The features have been observed, pictures have been recorded[18], and where and when possible people were asked to confirm the denomination of the type of feature in consideration to somewhat suppress any bias. The following components were examined[19] during this survey and are listed here in the order of higher to lesser importance regarding its influence on the quality of the thematic accuracy:

1) The correctness of the main tag: e.g. is this place a restaurant or a café?
2) The correctness of other tags that are described: e.g.$_1$ is the house number stated in OSM the correct number or e.g.$_2$ is the number of lanes of a highway correct?
3) Is there any confusion or doubt about whether the description in OSM represents the feature in the right way:
   - Unclearness about the nature of the feature; lack of description when it is not obvious what the function is (e.g. information - guidepost: street information or historical information?)
   - Doubt about the type within a main feature type (e.g. is it a highway-primary or a highway-secondary?)
   - The feature pointed to could be part of the whole feature instead of the feature itself (e.g. the entrance of parking could be marked as *parking*).

Based on these criteria the features were divided into four classes. Class 1 represents features of which the main tag does not correspond with what has been found in the field. Class 2 is assigned to features of which other tags are incorrect. Class 3 is assigned to features that have a shortcoming as described in the third point. And lastly, class 4 is assigned to those features of which the available information is fully correct. The result of the classification is shown in Appendix 7.

---

[18] CD: ..\FieldWork\Photos74Elements.pdf
[19] CD: ..\Sheets\CheckOSMvsField.xls

5.2 Topological Consistency

From the dataset was learned that many features with a higher number of versions are labels (in the form of point features) and these do not have information about the geometry of the features to which they refer. The accuracy of the position of these labels is not that important, but at least their topological relations should be correct. E.g. for a shop that is positioned on the left side of the street, the label should be on that side as well.

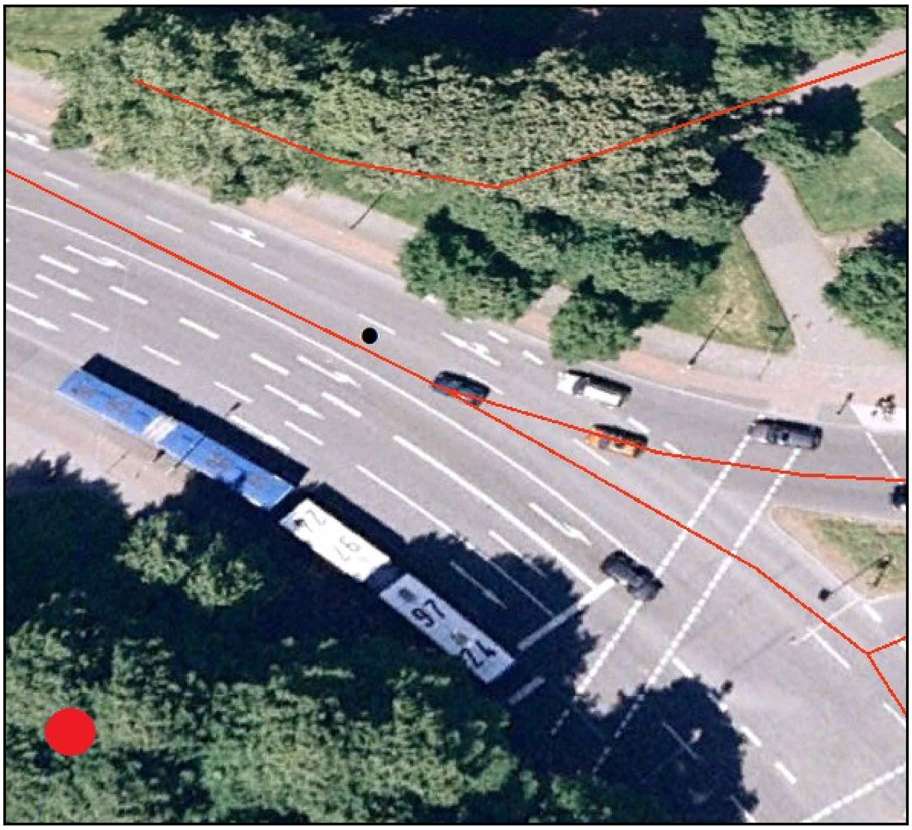


Figure 18: A feature (black dot) is on the wrong side of the street. The red dot is its real position.

To reveal whether the features are topologically correct, a KML was exported and added as a layer to a Google Earth session. Google Earth provided good enough image quality, allowing for the ability to recognise the relation of features relative to their surroundings. The correct location is known from the field survey.

It turned out only one feature, an information panel in the north-western part of the study area, is topologically incorrect (Figure 18).

5.3 Information Completeness

5.3.1 Tag Usage

Members of the OSM community are free to choose any tags they want. Therefore, also regarding this issue the approach was to infer from the data itself which tags ought to be present when retrieving information about a particular type of feature. This allows for assessing whether tags are used correctly and how complete the information about a feature is. Data completeness is one of the quality elements stated in the ISO 19113 report.

For every feature type that occurs in the selected Münster Altstadt dataset the most recent versions of the features were extracted from the raw dataset[20]. An overview of this process is shown in Appendix 8. To determine which tags should be used with a certain feature type the *term frequency - inverse document frequency (tf-idf)* importance measure was adopted from [27]. It is used to evaluate how important a word is to a document in a set of documents. The number of times a term is present in a document is multiplied by a factor that represents the general importance of the same term in the whole set of documents (corpus). This way the individual importance values are normalised with respect to the whole dataset and it becomes clear which terms are document specific.

An *inverse feature type frequency* can serve as a measure of the general importance of a tag by dividing the total number of features in the dataset by the number of features with which this tag is associated. Taking the logarithm of this quotient puts out values that indicate a higher relevance of the tag the closer it is to 0. The equation form is:

$$\mathrm{iff}(t) = \frac{\log |F|}{|\{f : t \in f\}|} \tag{1}$$

whereby $|F|$ is the total number of features in the whole dataset and $|\{f : t \in f\}|$ is the number of features in that set containing tag t.

[20] CD: ..\Sheets\TagExploration

The *tag frequency – inverse feature type frequency* is then determined by:

$$\text{tf-iff}(t, f) = \text{tf}(t, f) \times \text{iff}(f) \qquad (2)$$

whereby *tf* is the tag frequency in the set of features that belong to the same feature type.

The output of this calculation shows the relevance of a tag within a set of features of a certain feature type, considering as well all the features in the whole dataset. Tags that are relatively unique to the feature type in consideration turn up with higher importance values than tags that are also commonly used to describe other feature types.

An issue turned up concerning important tags that are used for more types of features. An example of a very common tag in the whole dataset is *name*. The log function of the general importance measure part (*iff*) really decreases the importance of this tag. Even multiplication by a high tag frequency within the set of features of the same feature type can result in a low value. Going through the results of this calculation (an example is shown in Table 3: values normalised by the highest importance value), in general the tags name and website now fall outside the range of values with a relative high importance, while they can be still important; the name of a feature is often remembered by people and the nature of a name is that many associations are linked to it. Also, a website is these days valued as an important piece of information that could link to more important information about the feature of interest.

When taking as a starting point the idea that the determination of the importance of tags should be really done per feature type, the approach could be to use part of the tf-idf method; in this case only the set of features of the same feature type is taken into account. The approach is now to determine tag importance internally by dividing the total number of features in the feature type set by the number of features annotated with the tag in consideration and taking the logarithm, after which the result is subtracted from 1 (1 – iff). Thus, equation 1 was used again, but now the denominators have a different meaning; F the total number of features within a feature type and $\{f : t \in f\}$ the number of features with a certain tag within a feature type set.

Table 3: Importance values for the amenity - place of worship point feature type

amenity - place of worship							
tag	**#**	**tf**	**iff**	**tf-iff**	**tf-iff-norm**	**i-iff**	**i-iff-norm**
religion	**24**	**0.24**	**1.79**	**0.43**	**1.00**	**0.00**	**1.00**
denomination	**20**	**0.2**	**1.85**	**0.37**	**0.86**	**0.08**	**0.94**
wikipedia	8	0.08	2.04	0.16	0.38	0.48	0.65
website	13	0.13	1.09	0.14	0.33	0.27	0.81
name	24	0.24	0.29	0.07	0.16	0.00	1.00
disused	1	0.01	3.42	0.03	0.08	1.38	0.00
url	1	0.01	3.42	0.03	0.08	1.38	0.00
alt_name	1	0.01	2.22	0.02	0.05	1.38	0.00
building	1	0.01	1.92	0.02	0.04	1.38	0.00
addr:country	1	0.01	1.58	0.02	0.04	1.38	0.00
wheelchair	1	0.01	1.56	0.02	0.04	1.38	0.00
description	1	0.01	1.37	0.01	0.03	1.38	0.00
addr:city	1	0.01	1.31	0.01	0.03	1.38	0.00
addr:postcode	1	0.01	1.18	0.01	0.03	1.38	0.00
addr:housenumber	1	0.01	1.10	0.01	0.03	1.38	0.00
addr:street	1	0.01	1.10	0.01	0.03	1.38	0.00
tags	**100**						
elements	**24**						

It turned out that this method filters out the most important and obvious tags per feature, even if they in frequency are less than half of the features of a particular type. In a feature type like pub the tag *smoking* should be important. It occurred in less than half of all the pub features, but it is still incorporated in the set of ‘obligatory tags’ when determining its importance with the measure of general importance (iff). The results of the internal tag importance determination based on an internal inverse feature type frequency are also shown in the example displayed in Table 3 (i-iff), highlighted by light gray.

For both ways of importance determination the values where normalised and a threshold of 0.5 was used by which a value above indicates that a tag is important enough to fall under the list of ‘obligatory tags’. The underlying idea is that of a pass mark on a scale from 1 to 10 that is in many occasions a mark higher than 5. The results of this investigation for every feature type that exists in the selected dataset of 74 elements can be found on the CD[21].

---

[21] CD: ..\Sheets\TagsPerFeature.xls

5.3.2 Omission

Now that is known which tags should be involved for describing a feature of a particular type, it is possible to assess the completeness of the tag information for the selected features. Completeness is one of the elements involved in data quality assessment. There can be omission (absence of data) and commission (excess data). Omission is determined by dividing the number of tags that are missing by the total number of tags that should be there according to what was inferred from the data (see the example of Table 3). Commission is determined by dividing the number of excess tags by the total number of tags that should be there. An example is shown in Table 4, where the red tag is an obligatory one based on the tf-idf importance measure taking the whole dataset into account. The red AND orange tags are obligatory based on the importance measure applied on just one feature type. The orange shaded tags with black font only are excess data according to the first method and the blue tag is excess data in both cases.

Table 4: Omission of tags

highway - pedestrian							
nr	osm_id	t1	t2	t3	t5	**%o_tf-iff**	**%o_i-iff**
41	5707689	x	x			100	50
64	5967591	x				100	75
44	40812233	x		x		0	50

The fraction of omission was used to measure the completeness of feature information. Omission of red tags was counted as twice as important as omission of orange shaded ones, because the red coloured tags are the ones that are unique to a specific feature. The orange shaded tags are of general importance when restricting the importance measure to the data belonging to one specific feature type.

All other files processed during this investigation are available on the CD[22].

[22] CD: ..\Sheets*tag*

5.4 Combined Information Quality

After preparing the individual quality element tests[23], linking them to the map IDs and classifying them into 5 classes (equal intervals) (Appendix 9.1), the outcomes of the three quality tests have been summed up and the summed values were reclassified again into five classes (Appendix 9.2: Quality Measure). The result linked to the features in their original form is shown in Figure 19.

It turns out many (42 of 74) features fully meet the set quality requirements. Generally, their theme is correctly described and their topological relations with respect to their surroundings are correct. The omission is not for all features that low, but with being consequent in generating the classes and giving each parameter the same weight, many features are of relatively high quality for what is possible to get from OSM. Including the 4th class, 60 out of 74 features are classified into the two highest quality classes.

[23] CD: ..\Sheets\ParametersQuality.xls

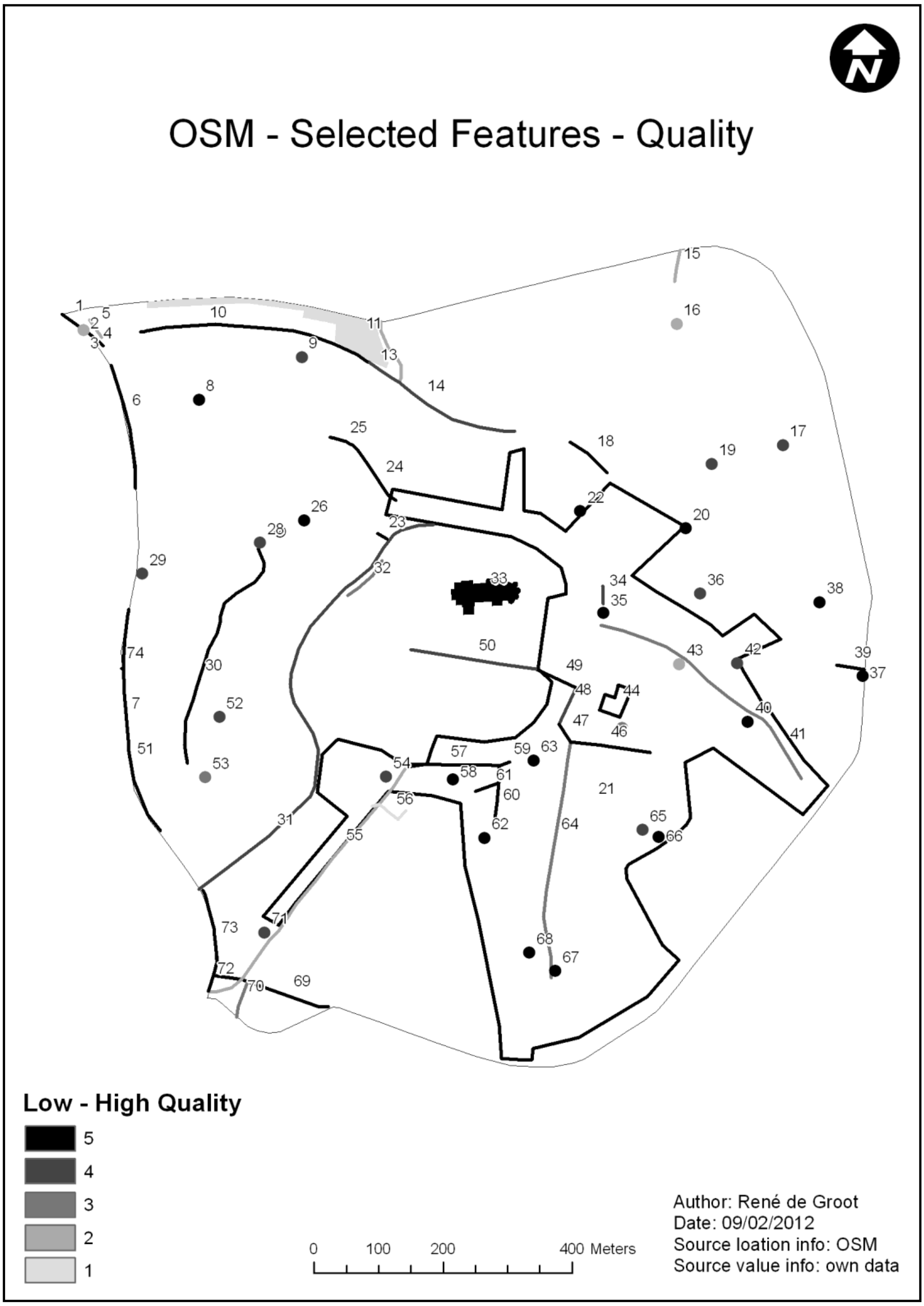


Figure 19: Final map with quality value classes allocated to the features of interest.

6 COMPARISON & RESULTS

6.1 Output

A general observation is that the quality test resulted in a majority of elements assigned to the highest classes, while the trust assessment resulted in a majority of features belonging to middle (class 3) or lesser value classes. The difference in class numbers for features between the two measures varies from -4 to +2 classes (see Figure 20, or Appendix 10.1 for a larger image).

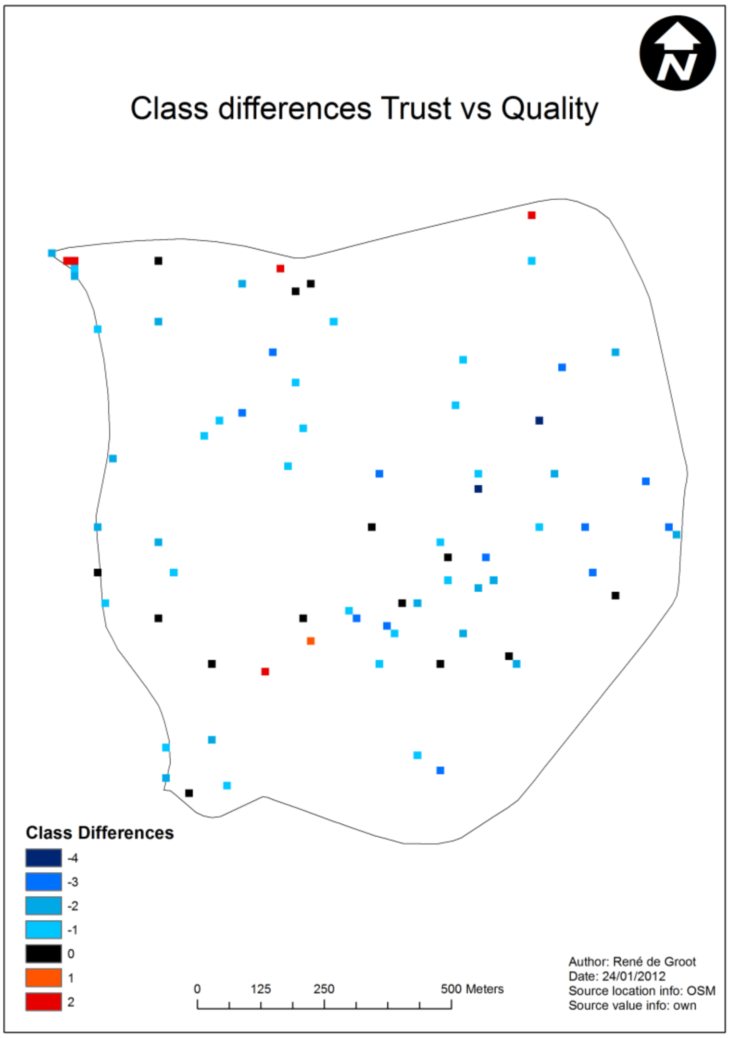


Figure 20: Class differences trust vs. quality (point raster mode). The differences of trust class with regards to quality class vary from -4 (4 classes lower) to +2 (2 classes higher). Most differences are from -2 to -1 (trust is 1 or 2 classes lower than quality).

A look at the table, or at a visualisation of the two measure results, the suggestion is given that for a large part of the dataset there exists similarity in order for both of the measures. There are many cases in which the quality measure 'follows' the measure that estimated quality based on trust in the same direction (see Figure 21, or Appendix 10.2 for a larger image). The basis of this characteristic is a collection of 54 features for which trust measure generally estimates a feature's quality 1 or 2 classes lower than the quality measure (see Figure 22, or Appendix 10.3 for a larger image).

The salient high and low points of both measures revealed that low trust was caused mostly by a low number of confirmations and the presence of tag corrections or rollbacks, while the low quality was caused by mostly a wrong feature type, in one case in combination with high omission. This suggests that confirmations and rollbacks are most important indicators for quality.

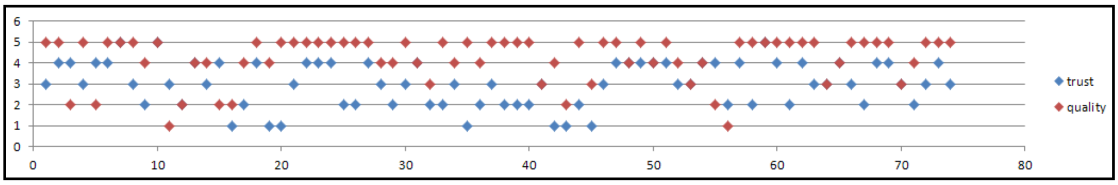


Figure 21: Class value differences between the trust and quality measures; the red dots represent the quality measure results, the blue represent the trust measure results.

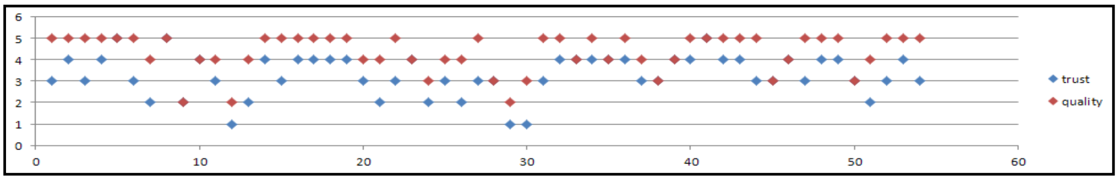


Figure 22: Class value differences between trust and quality measure for differences -2 – 0.

Taking a closer look at the parameter values of the 20 features for which the classes of the two tests show no correspondence but rather opposition, reveals that a low number of versions, a low number of confirmations and the occurrence of tag corrections make these features have a low trust value. In case the differences are in the opposite direction (high quality, but low trust) then it is the theme of the feature that is incorrect. There are only a few of these cases. Most of the big mismatches concern the first observation.

The features with a high negative class difference generally have a higher number of tag corrections and slightly lower numbers of confirmations and users. Less emphasis on the tag corrections would increase the trust value of these features. Trying this out resulted however in more deviation from the trust values as they were. This supports the model as is, but it shows also that one cannot always rely on counts of numbers of variables designated as trust parameters in order to get an indication of the information quality.

6.2 Significance

A statistical measure has been carried out to determine if there is an association between the two classifications. Because the trust and quality measures both consist of a number of reclassifications and the quality measure includes a ranking, the whole set of trust and quality results does not have a clear numerical basis and is rather ordinal. The statistical measure is therefore a non-parametric one and calculates a rank association / correlation coefficient.

6.2.1 Outliers

One definition of outliers has already been described in section 4.2: a datum that appears to deviate markedly from other members of the sample in which it occurs [12]. There are many other definitions and therefore outlier determination is always influenced by subjectivity. Applying various methods to detect outliers again for the final results did not result in removal of most of the 74 feature difference values. However, the 20 features that were noticed because of their suggestion of no association are a minority; they could still be labelled as atypical and not following the characteristic of the majority. These are also properties of outliers. Ultimately, for completeness and room for own interpretation, both the whole dataset and the part of 54 features were taken into account for a rank correlation measure.

6.2.2 Kendall's τ

The Kendall tau rank correlation coefficient [15] measures the association between two measurement results, whereby any pair of observations $\{(x_i, y_i), (x_j,y_j)\}$is *concordant* if $x_i > x_j$ & $y_i > y_j$ OR $x_i < x_j$ & $y_i < y_j$ and *discordant* if $x_i > x_j$ & $y_i < y_j$ OR $x_i < x_j$ and $y_i > y_j$, In case $x_i = x_j$ or $y_i = y_j$ a pair is tied. Kendall's τ is then calculated as follows [22]:

$$\tau = \frac{c - d}{\frac{1}{2}\, n\, (n - 1)} \tag{3}$$

whereby c is the number of concordant pairs and d the number of discordant pairs. *n* is the sample size. If there is a perfect agreement between the two rankings, τ is 1. For the opposite situation τ is -1. If there is no association at all (x and y are independent), τ is expected to be 0 or close to 0.

Calculating τ [33] for the whole dataset of 74 features results in a value of $\tau \approx 0.164$ (Appendix 10.4). This is clearly closer to 0 than to 1 and therefore in this case the null hypothesis that the trust and quality results are not associated cannot be rejected. Calculating τ for the 54 features following the trend of the quality measure, the result is $\tau \approx 0.520$, indicating there is agreement (although not the strongest) and thus association between the two measures. When determining significance, a very low p value of 1.727e-05 suggests a rejection of the null hypothesis and an acceptance of the hypothesis that the trust and quality measure are associated in a linear fashion.

7 CONCLUSION

The information from the literature review and the properties of the dataset determined the trust measure as well as part of the quality measure. The measure developed is coarse and simple, because it is not yet known how important which indicators are. Assumptions were made to open the way for a measure and assessment. These assumptions concerned an understanding of trust and quality, people-object transitivity of trust, equal importance of various indicators of trust and (personal) intuitive notions.

The results of the trust measure strongly suggest that the hypothesis *Patterns and variables derived from VGI feature provenance data can be used to determine a feature's trustworthiness, which in turn indicates its data quality* is likely to contain a kernel of truth, but taking all values into consideration, there is no solid proof. There is also no proof that the hypothesis is wrong. Because there is a great number of 'follow ups', whereby trust measure values decrease or increase when the quality decreases or increases, there is more support for than rejection of the hypothesis.

There is however a group of features, of which the trust and quality values are more or less in opposite classes. Because the quality measure for these features classifies them as high quality, it might not feasible to fully rely on the indicators of trust as presented and proposed to use in trust model. Even when a feature does not have many confirmations, is not looked at by many users, does not have a relatively long history or has multiple corrections, it could be that its information is actually correct. There are various instances of features of which their provenance contains very low numbers of parameters considered as positively influencing trust that are of good quality in the sense its information is correct and complete, as well as there are opposite cases.

Altogether, this could be an indication that it might not be possible to accurately estimate a single feature's data quality based on proposed trust indicators but it might be possible to estimate from a larger set of features whether the majority of these features represent information of high quality. Involving more data and trying out multiple approaches is necessary to acquire results that are more accurate and reliable.

8 LIMITATIONS AND DISCUSSION

The size of the dataset is relatively small. It could not be large due to the field survey as an important restricting factor. Processing a small sample set yields less accurate results than a larger sample set.

For determining the number of confirmations, a buffer of 50 metres was used to select changesets in the surroundings of a feature of interest. The assumption was made that this would be a proper distance about which can be assumed it would be close enough for users to fall within their attention when editing a feature around the feature of interest. The optimal buffer distance could be different. Also all extents in a buffer were treated as equally important. It might be that a larger extent should get a higher weight.

Trust indicators have been involved with equal weights and they were classified the same way, in 5 classes, equal interval. The importance of parameters may however differ. Other methodologies involving multiplication or division would also generate different results.

Tag corrections and rollbacks are assumed to have a negative effect on trust, or at least not contributing to trust. It indicates feature instability and decreases indirectly the reputation of the user who provided the information prior to the correction or rollback, which in turn negatively affects the build-up of the feature trustworthiness. It might also be possible that the corrections and rollbacks positively affect a feature's quality.

The time decay was investigated and no specific timespans were found for particular main types of features. It was therefore not incorporated in the trust measure. The issue here is also that the dataset is most likely too small to confirm that there are no differences between different types.

The assessment of feature information with regards to the outcome of the field study, carried out to measure the thematic accuracy of the features of interest, is based on a personal intuitive notion of how different failures in the information should be classed on a scale from 1 to 4. First of all, the ranking of the different failures might be different if done by somebody else, secondly there might be other failures to be noted and lastly the scaling could be different.

The presence of tags is counted and recorded. A methodology was adopted from the field of documentation in an attempt to acquire lists of tags that should be described for a certain type of feature. This way it would be possible to assess per feature the information completeness by determining the omission of tags. There are different methodologies to determine relevance / importance of tags that might result in a different outcome.

A question that remains is whether the indication of quality of any trust measure would be representing VGI data quality in the same way as the reference quality measure would do; the quality measure based on quality elements involved in assessing data quality might be different from the set of quality elements that are best predicted by the trust measure results. Trial and error in applying various methods and assumptions is a way of getting to know what predicts what.

High positional accuracy was not always available in the examined dataset, because labels and the right descriptions seemed to be more important (based on the numbers of versions and users of single point label features compared to those for polygons and streets). For precision measurement professional organisations in engineering would have to keep collecting high quality data that fits their purpose.

Streets and polygons that are present in OSM could however be examined in a more detailed way to reveal their positional accuracy, which would be a study on its own. This has already been done by various researchers.

What remains certain is that VGI can be very useful. It can show how people describe the world around them and give insight about community semantics. It also reveals which places are interesting according to groups of people and the lack of institutional viewpoints as well as fuzziness in the dataset keeps it rich in types of information and interpretation. Users publishing information about their points of view also brings in privacy and ethical issues. Handling VGI should therefore be done taking into account solutions for these problems, such as making retrieved data anonymous.

9 FUTURE WORK

Important in this field of study is the amount of data. Larger datasets would help to more accurately confirm if measure outputs are plausible. It would allow for the exclusion or acceptance of possible new insights through testing significance of outputs more accurately. Important hereby is also to automate all the data processing. This makes it possible to handle large datasets in a short period of time.

The methodology proposed in this thesis could be applied on a larger dataset to see whether it can be used to determine an indication of the trustworthiness of the VGI dataset as a whole. In case the majority of the trust value follows the quality assessment, the model could be refined and used to determine a general trustworthiness of a dataset. If it turns out otherwise, the model and its underlying ideas have to be thoroughly revised. This could change the model's composition.

Many assumptions were made, but any variation in these assumptions and determining how it changes the output could reveal which ones are closer to the truth. The methodology used to measure trust is coarse and simple and for both measures the individual parameter outcomes can be weighted differently. Besides that there might be many other ways of determining trustworthiness. Testing various approaches from different angles and compare their outcomes could lead to a 'best fit'. The parameters used here could be changed. Also, more could be tested from the angle of user reputation, as initially proposed by [17]. For this work, user reputation has not been directly incorporated in the measure. The assumption was made that parameters influencing user reputation would transfer this influence on to the feature information of the feature of interest. An approach that puts more emphasis on reputation - whereby not every correction has the same effect, but also the user who makes the correction affects trust - could generate different results (if a correction is made by a user with high reputation the net effect is an increase of trust value). Attention should be given though to the issue that not every user is equally active. There should be a minimum for the number of contributions, involving also the up-to-date-ness of the activity etcetera.

The data quality assessment was carried out in a certain way, whereby three quality elements were investigated. More elements could be involved. Furthermore, the way of assessing these quality elements is somewhat subjective (though there is the belief in that its intuitive approach positively influences the objectivity). More people could argue about the ranking of how 'bad' a certain occurrence (see section 5.1) is.

A process that cannot be automated is the field survey. The field survey carried out for this work clearly showed that information of the theme of a feature is not always correctly published and therefore it should always be part of the reference material when evaluating a trust model. To avoid bias in the theme and tag checks, and raise the accuracy and quality of the survey, more people need to be involved to come to a more accurate decision; if a majority of people believes that a particular feature is a 'café' and not a 'pub', 'café' would be the best theme for it. Also larger sample sizes call for more helping hands in the field.

10 REFERENCES

[1] D. Artz and Y. Gil. A survey of trust in computer science and the semantic web. *Web Semantics: Science, Services and Agents on the World Wide Web*, vol. 5, no. 2, pp. 58–71, 2007.

[2] M. Bishr and W. Kuhn. Geospatial information bottom-up: A matter of trust and semantics. *The European information society*, pp. 365–387, 2007.

[3] M. Bishr and L. Mantelas. A trust and reputation model for filtering and classifying knowledge about urban growth. *GeoJournal*, vol. 72, no. 3-4, pp. 229-237, 2008.

[4] M. Bishr and K. Janowicz. Can we trust information? - The case of volunteered geographic information. In A. Devaraju, A. Llaves, P. Maué, and C. Keßler, editors, *Towards Digital Earth: Search, Discover and Share Geospatial Data 2010. Workshop at Future Internet Symposium*, 2010.

[5] M. Bishr. *Trust & Reputation Models for Human Sensor Observations*. PhD thesis, University of Münster, 2011.

[6] K. Cook (e.d.). 2001. *Trust in Society*, New York: Russell Sage Foundation, 2011.

[7] A. J. Flanagin and M. J. Metzger. The credibility of volunteered geographic information. *GeoJournal*, vol. 72, no. 3, pp. 137–148, 2008.

[8] F. Fukuyama. *Trust: The Social Virtues and The Creation of Prosperity*. New York: Free Press, 1996.

[9] J. A. Golbeck. *Computing and applying trust in web-based social networks*. University of Maryland, 2005.

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

[11] R. de Groot. Using citizen-centric data through web 2.0 technology as a solution to the problems of speed and cost in the set-up of a spatial data infrastructure. *Partial fulfilment for the GI Science course at the Higher Institute of Statistics and Information Management*, New University of Lisbon, 2010.

[12] F. Grubbs. Procedures for detecting outlying observations in samples. *Technometrics*, 11(1), 1-21, 1969.

[13] O. Hartig, Provenance information in the web of data. In *Proceedings of the Linked Data on the Web (LDOW)Workshop at the World Wide Web Conference (WWW)*, 2009.

[14] A. Jakobsson and J. Giversen. *Guidelines for Implementing the ISO 19100 Geographic Information Quality Standards in National Mapping and Cadastral Agencies*. EuroGraphics, 2009.

[15] M. Kendall. A new measure of rank correlation. *Biometrika*, 30 (1-2): 81-89, 1938.

[16] C. Keßler, K. Janowicz, and M. Bishr. An agenda for the next generation gazetteer: Geographic information contribution and retrieval. In *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, no. c, pp. 91–100, 2009.

[17] C. Keßler, J. Trame and T. Kauppinen. Provenance and trust in volunteered geographic Information: The case of OpenStreetMap. In *COSIT*, Belfast, ME, USA, 2011.

[18] J. Lanier. Digital Maoism: The Hazards of the New Online Collectivism. In *Edge*, 2006.

[19] D. Liben-Nowell, J. Novak, R. Kumar, P. Raghavan, and A. Tomkins. Geographic routing in social networks. In *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 33, pp. 11623-8, Aug. 2005.

[20] N. Mezetti. A socially inspired reputation model. *Lecture Notes in Computer Science*, pages 191-204, 2004.

[21] A. Mitchell. *The ESRI Guide to GIS Analysis Volume 1: Geographic Patterns & Relationships*. ESRI Press, 2001. ISBN 9781879102064.

[22] R.B. Nelsen. Kendall tau metric. In M. Hazewinkel. *Encyclopedia of Mathematics*, Springer, 2001. ISBN 9781556080104.

[23] Open Knowledge Foundation Blog, *Open Street Map community responds to Haïti crisis*, 2010. URL http://blog.okfn.org/2010/01/15/open-street-map-community-responds-to-haiti-crisis.

[24] OpenStreetMap Community. OpenStreetMap Wiki, 2011. URL http://wiki.openstreetmap.org.

[25] X. Qian, L. Di, D. Li, P. Li, L. Shi, and L. Cai. Data cleaning approaches in Web 2.0 VGI application. In *Proceedings of the 17th International Conference on Geoinformatics*, pp. 1-4, 2009.

[26] B.R. Rich. *Clarence Leonard (Kelly) Johnson, Biographical Memoir.* Washington D.C.: National Academies Press, 1995.

[27] G. Salton, A. Wong, A, & C.S. Yang. A vector space model for automatic indexing. *Communications of the ACM*, 18(11), 1975.

[28] Y. L. Simmhan, B. Plale, and D. Gannon. A survey of data provenance in e-science. *ACM Sigmod Record*, vol. 34, no. 3, pp. 31–36, 2005.

[29] R. Šumrada. Temporal Data and Temporal Reference Systems. In *Proceedings of the International Federation of Surveyors Working Week and 125th Anniversary of FIG Conference*, 2003.

[30] P. Sztompka. *Trust: A Sociological Theory.* Cambridge: Cambridge University Press, 1999.

[31] A. Thomas. Making the most of open content: understanding use. *Joint Information Systems Committee*, 2011. URL http://www.jisc.ac.uk.

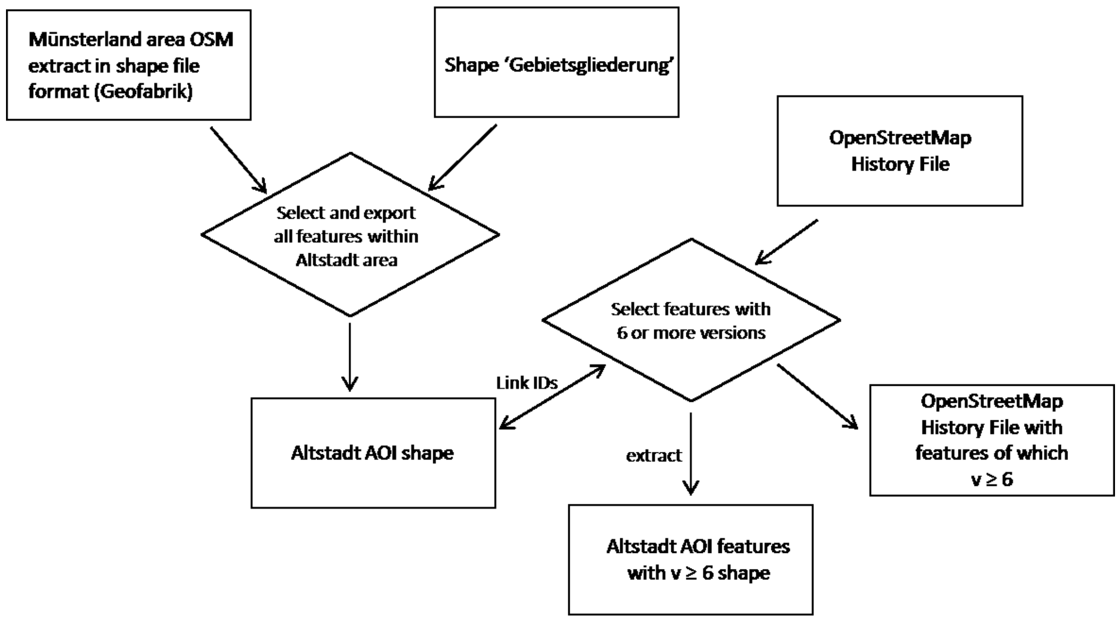
[32] E. Uslaner. *The Moral Foundations of Trust.* Cambridge, UK: Cambridge University Press, 2002.

[33] P. Wessa (2012), Free Statistics Software version 1.1.23-r7, *Office for Research Development and Education*, 2012. URL http://www.wessa.net/rwasp_kendall.wasp

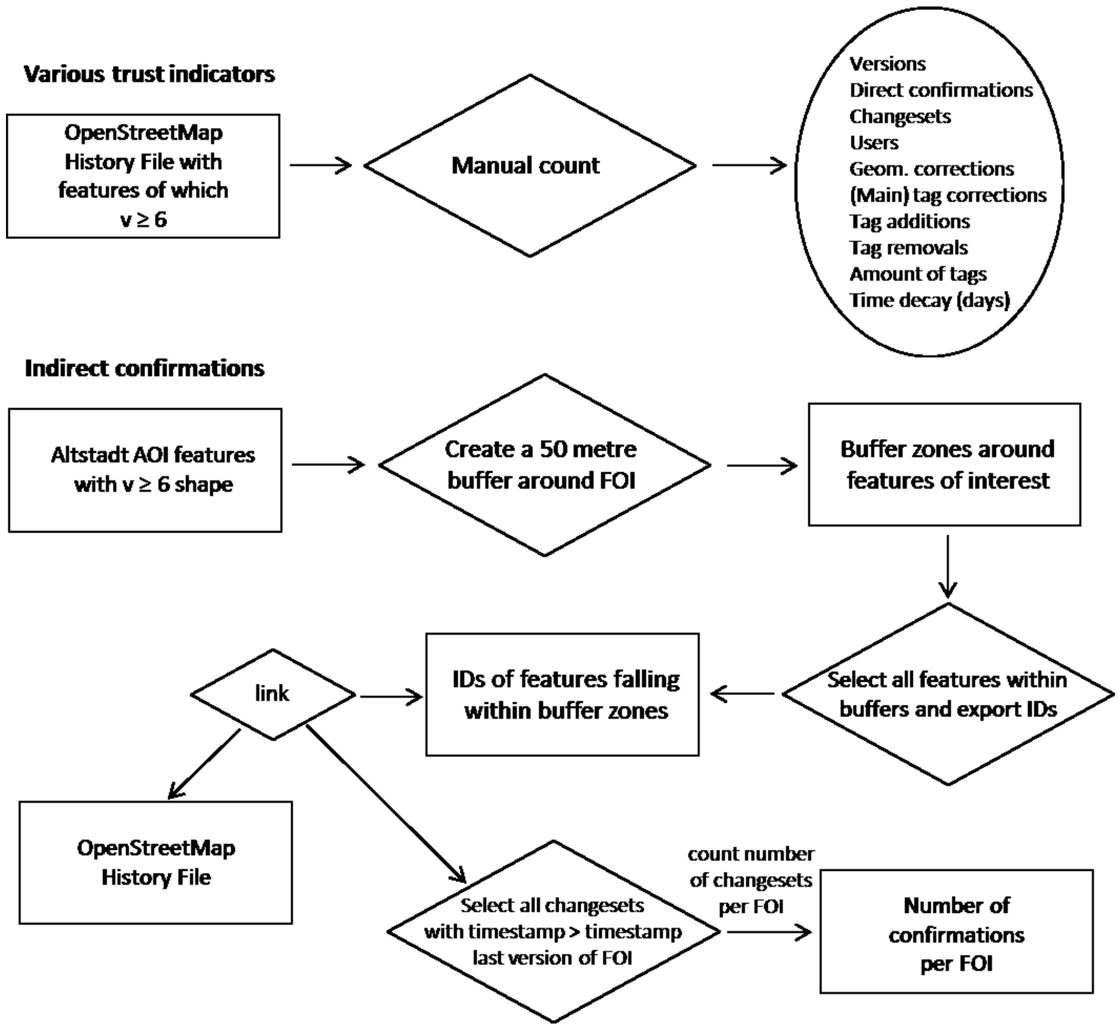
APPENDIX 1: QUALITY ELEMENTS AND SUB-ELEMENTS IN THE ISO 19113 STANDARD

Data quality element Data quality sub-element	Description
Completeness	Presence or absence of features, their attributes and relationships
Commission	Excess data present in a dataset
Omission	Data absent from a dataset
Logical consistency	Degree of adherence to logical rules of data structure, attribution and relationships
Conceptual consistency	Adherence to rules of the conceptual schema
Domain consistency	Adherence of values to the value domains
Format consistency	Degree to which data is stored in accordance with the physical structure of the data set
Topological consistency	Correctness of the explicitly encoded topological characteristics of a dataset
Positional accuracy	Accuracy of the position of features
Absolute or external accuracy	Closeness of reported coordinate values to values accepted as or being true
Relative or internal accuracy	Closeness of the relative positions of features in a dataset to their respective relative positions accepted as or being true
Gridded data position accuracy	Closeness of gridded data position values to values accepted as or being true
Temporal accuracy	Accuracy of the temporal attributes and temporal relationships of features
Accuracy of a time measurement	Correctness of the temporal references of an item (reporting of error in time measurement)
Temporal consistency	Correctness of ordered events or sequences, if reported
Temporal validity	Validity of data with respect to time
Thematic accuracy	Accuracy of quantitative attributes and the correctness of non-quantitative attributes and of the classifications of features and their relationships
Classification correctness	Comparison of the classes assigned to features or their attributes to a universe of discourse (e.g. ground truth or reference data set)
Non-quantitative attribute correctness	Correctness of non-quantitative attributes
Quantitative attribute accuracy	Accuracy of quantitative attributes

APPENDIX 2: EXTRACTION OF FEATURES OF INTEREST



APPENDIX 3: EXTRACTING TRUST PARAMETERS



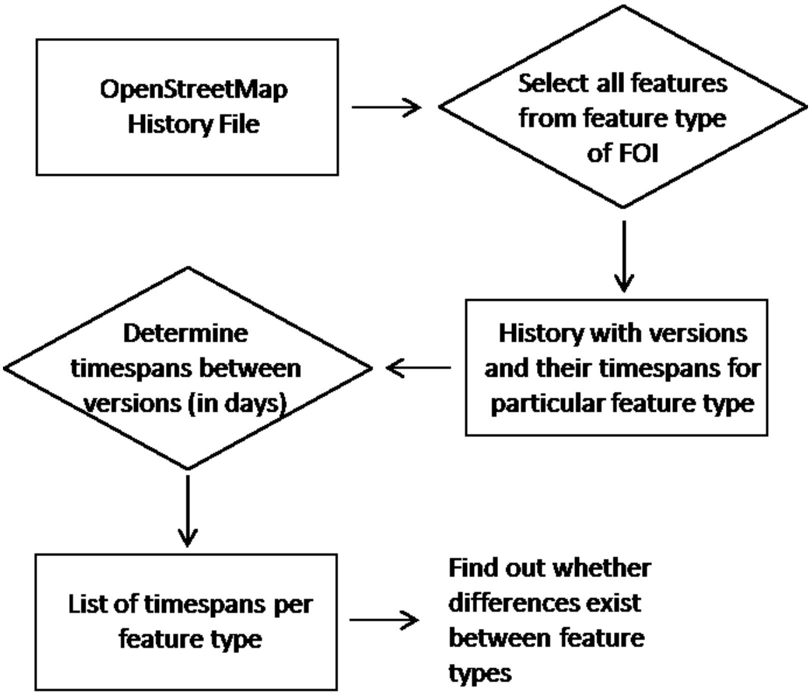
APPENDIX 4: TRUST FACTORS COUNT

mapnr	osm_id	type	v	ic	dc	cs	u	gco	tco	mtco	r	ta	tr	tags	days
1	34428182	hwy primary	6	1	0	6	6	2	0	0	0	3	0	7	15
2	5305703	hwy primary	14	0	1	14	11	6	2	1	0	5	1	6	15
3	395612533	tourism information	6	13	0	6	5	1	0	0	0	3	1	4	305
4	5736755	hwy cycleway	9	6	0	8	5	5	1	1	0	2	1	2	305
5	25631753	hwy cycleway	6	11	0	5	4	4	0	0	0	2	1	2	305
6	32024584	hwy primary	9	1	1	9	7	3	0	0	0	5	0	7	15
7	5967589	hwy primary	14	2	1	14	9	5	1	1	0	5	1	7	15
8	271428115	pub	6	2	0	6	4	1	0	0	0	5	0	11	122
9	271428124	pub	6	0	0	6	4	1	1	0	0	4	0	9	61
10	12342543	hwy tertiary	9	18	0	8	8	4	0	0	0	5	1	6	275
11	10376252	leisure park	7	7	0	6	3	2	0	0	0	3	0	4	76
12	4985978	hwy unclassified	7	0	0	7	6	3	2	2	0	2	0	3	31
13	28745010	hwy tertiary	8	1	0	8	6	3	0	0	0	4	2	3	76
14	5029732	hwy tertiary	11	10	0	11	5	7	2	1	0	3	0	4	61
15	5076410	hwy residential	8	16	2	8	4	6	1	0	0	1	0	3	641
16	301593456	fast food rest.	6	2	0	3	3	1	2	1	0	4	0	6	610
17	318491022	pub	6	0	1	6	4	0	1	0	0	4	0	8	214
18	5029733	hwy tertiary	12	12	0	11	6	6	2	0	0	4	0	5	244
19	314416804	restaurant	7	5	0	5	2	3	2	2	0	2	1	7	214
20	1056653240	bar	6	3	1	5	4	0	3	0	0	2	0	10	92
21	96389136	landuse retail	12	1	0	11	1	11	0	0	0	1	0	3	15
22	273815504	restaurant	6	5	2	6	6	0	0	0	0	4	0	14	122
23	6191699	hwy residential	7	31	0	7	3	3	1	0	0	3	0	8	641
24	5707693	hwy residential	6	29	0	6	3	3	1	0	0	2	0	6	717
25	34073159	hwy residential	6	3	0	6	3	2	1	0	0	2	0	6	76
26	271428135	shop	8	2	0	8	5	1	2	0	0	5	0	7	92
27	273815351	pub	6	12	0	6	5	1	0	0	0	6	0	11	275
28	273815355	shop	6	11	2	6	3	3	1	0	0	3	0	5	244

29	440123814	café	7	3	1	6	5	0	1	0	0	5	0	11	92
30	5967616	hwy residential	7	0	0	8	4	4	0	0	0	2	0	4	76
31	10191758	waterway river	8	31	0	8	5	4	1	0	0	2	1	3	275
32	10191739	hwy footway	6	6	0	6	4	2	2	0	0	2	0	6	275
33	40967526	place of worship	6	0	0	6	2	1	1	0	0	4	0	12	76
34	5707704	hwy residential	6	27	0	6	2	3	1	0	0	1	1	5	885
35	65144348	place of worship	7	9	1	7	6	2	2	0	1	3	6	0	244
36	338961586	library	7	14	1	7	5	1	0	0	1	4	9	0	244
37	36167011	tourism information	7	2	1	5	5	2	0	0	0	1	0	3	305
38	496990981	school	7	0	1	7	3	1	0	0	0	4	0	6	137
39	5029735	hwy tertiary	8	7	0	7	5	3	2	1	0	5	0	6	336
40	1056709820	restaurant	12	6	0	8	3	2	7	0	0	3	2	10	214
41	5707689	hwy pedestrian	9	7	1	9	5	6	1	1	0	1	0	3	92
42	273815562	police station	6	1	0	6	5	1	3	1	0	3	0	8	31
43	609764514	restaurant	6	2	2	6	4	1	3	1	0	3	0	8	15
44	40812233	hwy pedestrian	6	5	1	5	2	3	0	0	0	1	0	4	198
45	499486087	café	6	3	0	6	2	1	3	0	0	3	0	5	198
46	5967602	hwy residential	9	7	0	7	3	3	1	0	0	4	1	10	198
47	5967607	hwy residential	6	18	0	6	3	0	0	0	0	4	2	8	275
48	5707678	hwy residential	10	5	0	9	7	3	1	0	0	4	1	9	198
49	5883198	hwy residential	6	20	0	5	3	0	1	0	0	4	1	9	275
50	6191697	hwy residential	7	11	0	7	6	4	0	0	0	3	0	5	198
51	12687078	hwy primary	11	2	2	11	9	2	1	1	0	5	0	8	15
52	320351374	university	6	8	1	5	3	1	0	0	0	2	1	4	15
53	322903129	restaurant	6	0	0	6	6	1	1	0	0	4	0	7	0
54	616670978	café	7	20	0	6	4	2	1	0	0	3	1	4	244
55	5739833	hwy tertiary	11	1	1	11	7	5	1	1	0	3	0	4	92
56	40819495	hwy service	7	11	0	6	2	4	1	0	0	1	0	3	244
57	34349254	hwy residential	11	8	0	10	6	4	0	0	0	6	0	12	198
58	698066430	restaurant	7	2	0	5	5	1	2	0	0	4	1	8	15

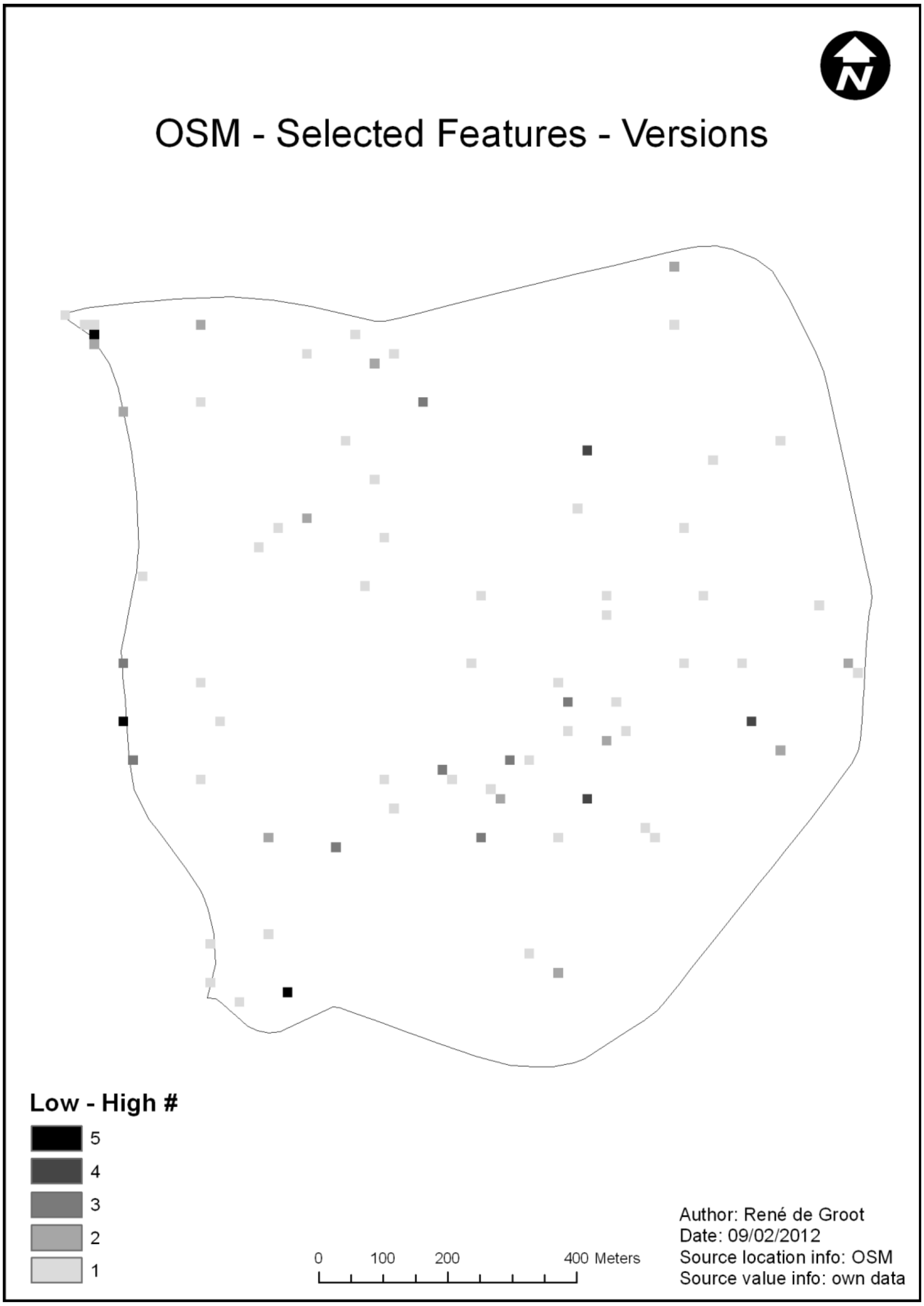
59	5707706	hwy residential	10	18	0	9	6	1	1	0	0	6	3	11	275
60	5707714	hwy residential	9	19	0	9	5	6	2	0	0	6	1	9	275
61	11219405	hwy service	6	6	1	6	3	2	1	1	0	3	0	6	198
62	608543669	restaurant	10	15	1	10	8	2	3	0	0	5	0	15	259
63	523365460	bank	6	6	1	6	4	1	0	0	0	5	0	10	153
64	5967591	hwy pedestrian	7	8	1	7	4	4	1	1	0	1	0	2	214
65	606083346	shop	6	16	0	6	4	4	0	0	0	2	0	4	580
66	496190365	parking	7	14	0	7	1	4	0	0	0	4	0	6	580
67	271428106	fast food rest.	9	3	0	7	6	2	4	0	0	5	0	10	92
68	271406051	place of worship	6	15	0	6	4	2	1	0	0	3	0	6	275
69	5608078	hwy cycleway	15	6	0	14	11	7	2	2	0	6	2	8	31
70	6190121	hwy cycleway	7	5	0	7	5	6	0	0	0	1	0	1	46
71	318497649	café	6	4	0	6	5	1	1	0	0	4	0	5	15
72	14165737	hwy primary	7	4	0	7	6	1	1	1	0	5	1	6	15
73	34428180	hwy primary	7	5	1	7	6	2	0	0	0	2	0	7	15
74	5737990	hwy tertiary	10	2	1	9	5	4	1	1	0	3	0	5	46

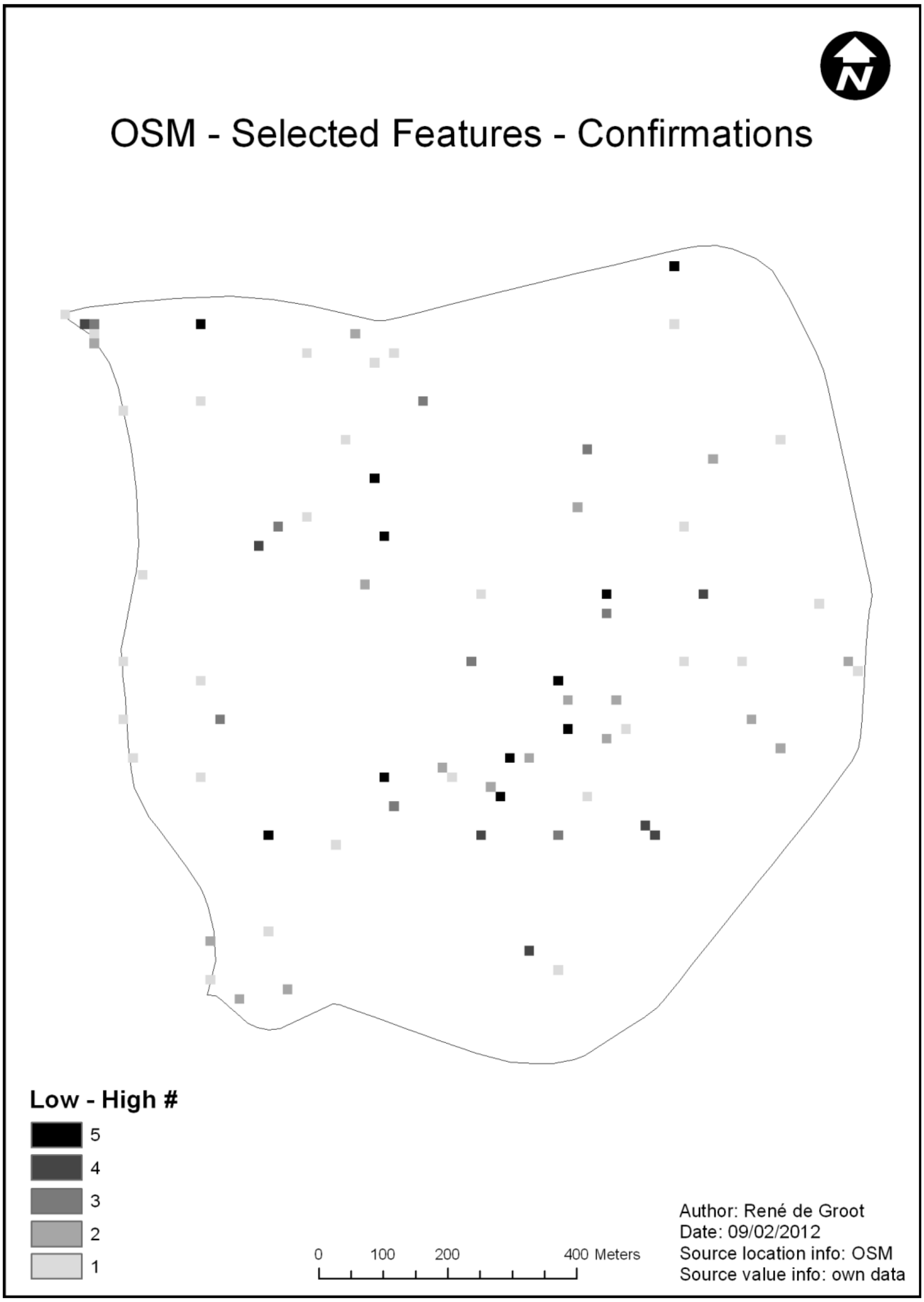
APPENDIX 5: ESTIMATE TIME DECAY

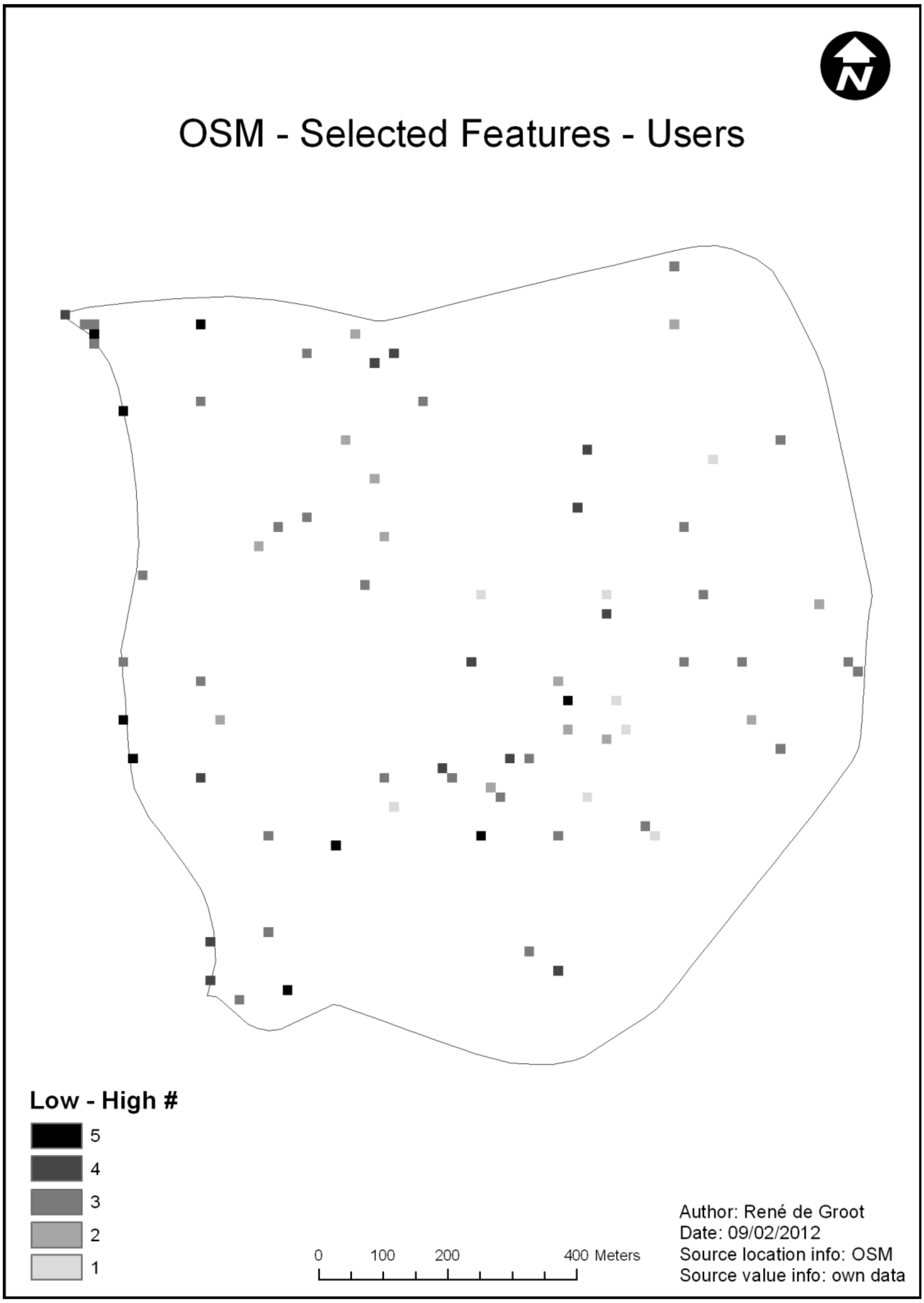


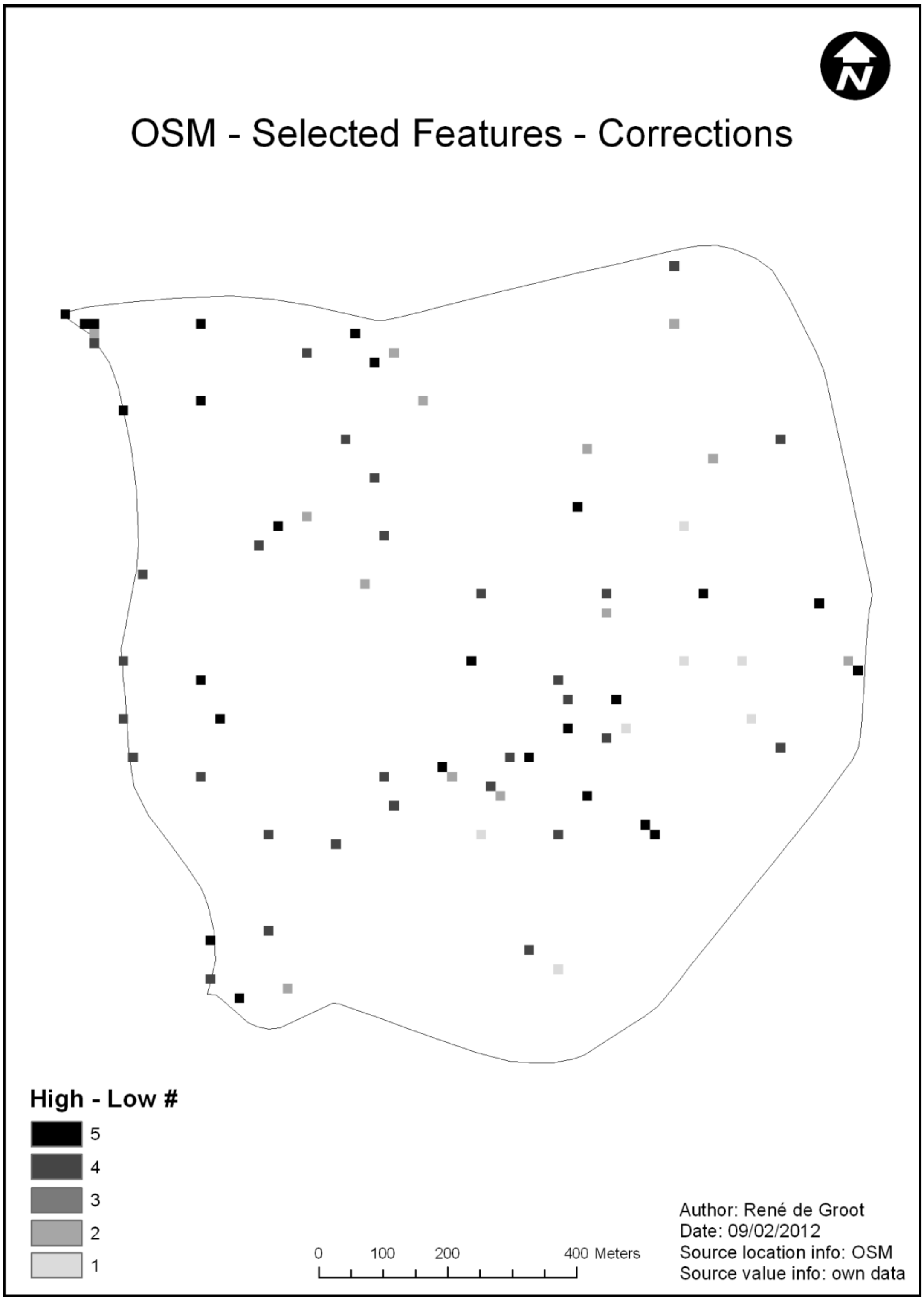
APPENDIX 6: TRUST MEASURE

Appendix 6.1 Trust Parameters Classified and Mapped

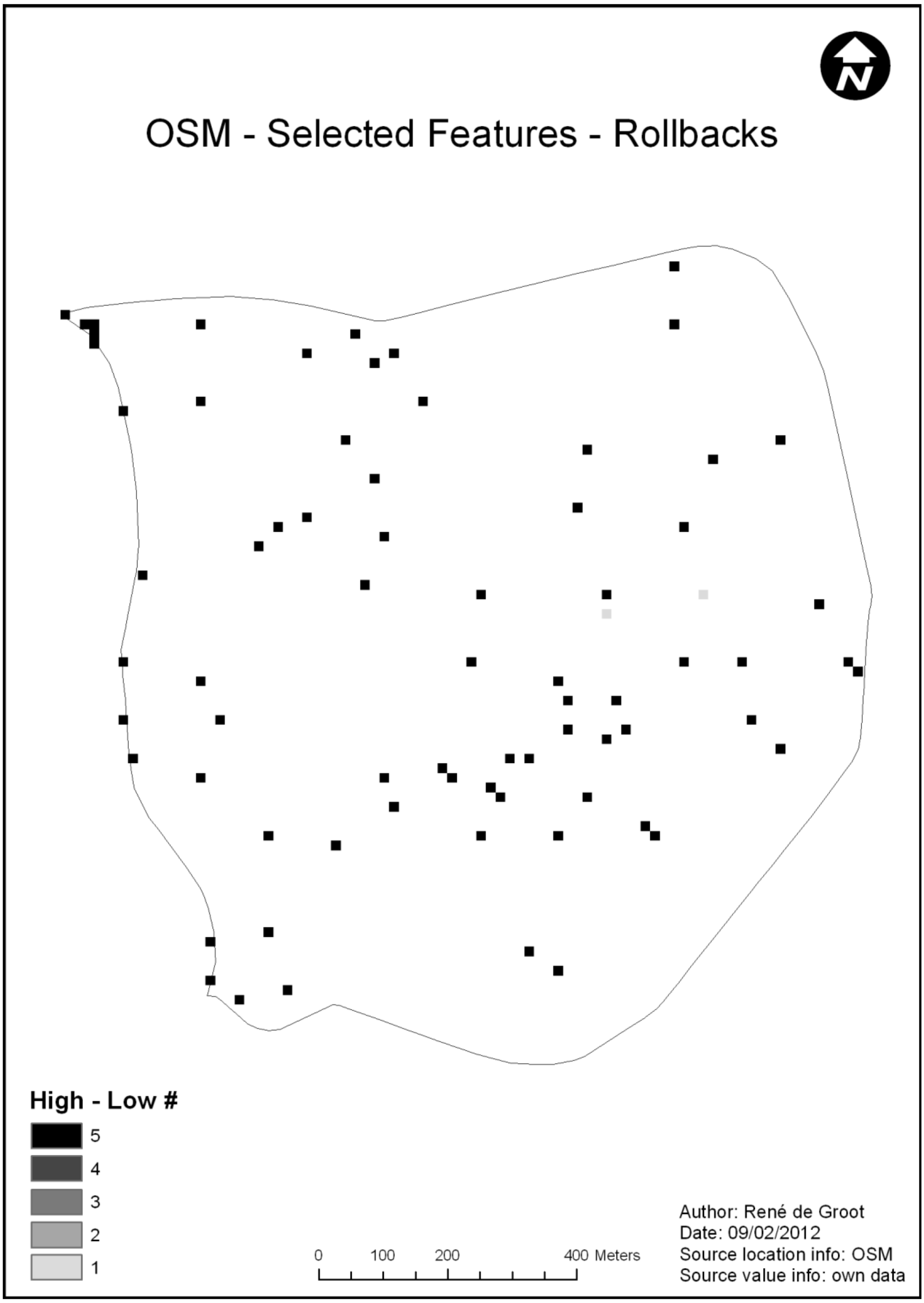



OSM - Selected Features - Confirmations
N
Low - High #
5
4
3
2
1
0
100
200
400 Meters
Author: René de Groot
Date: 09/02/2012
Source location info: OSM
Source value info: own data


OSM - Selected Features - Users
Low - High #
5
4
3
2
1
0
100
200
400 Meters
Author: René de Groot
Date: 09/02/2012
Source location info: OSM
Source value info: own data

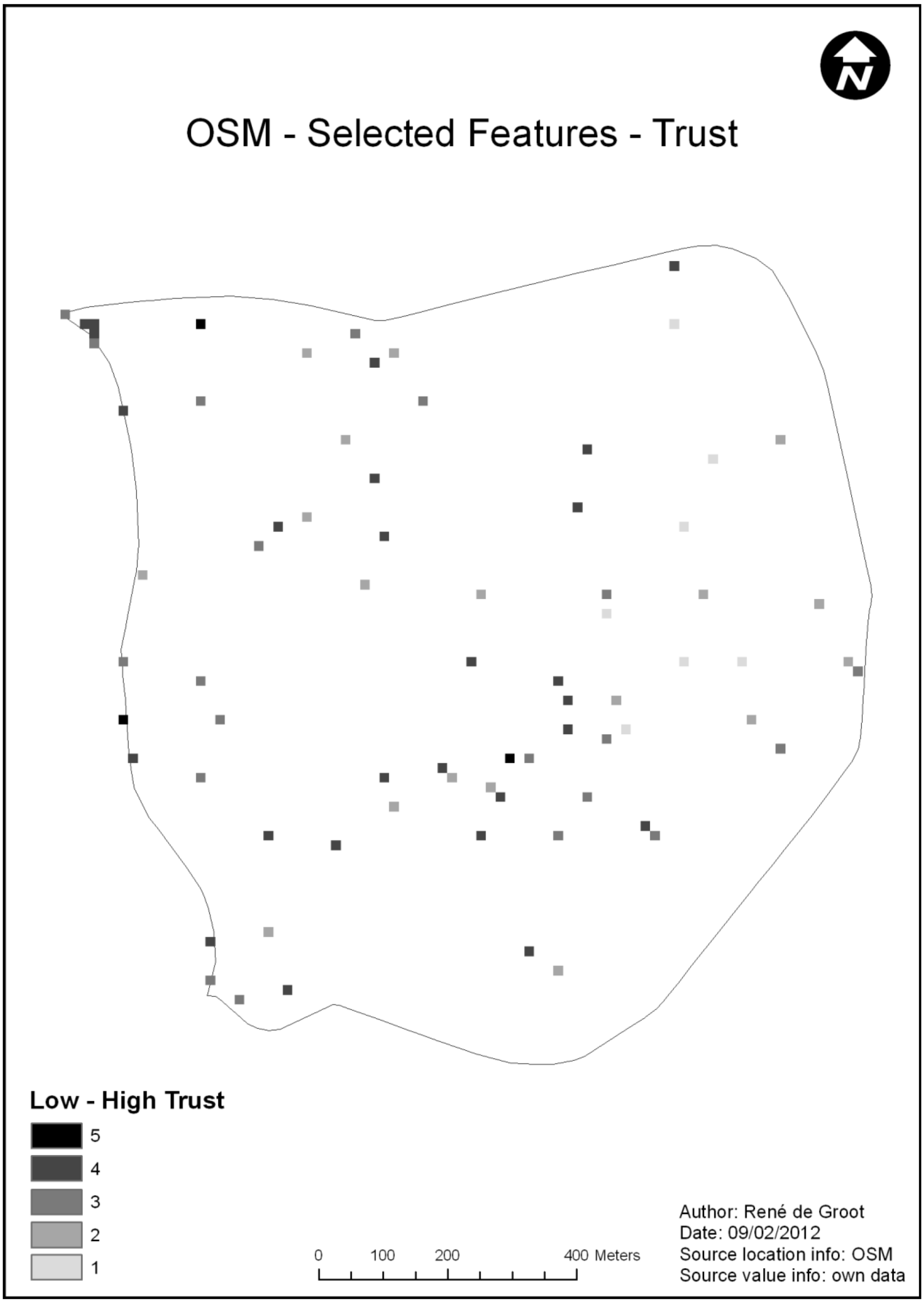


OSM - Selected Features - Corrections
N
High - Low #
5
4
3
2
1
0
100
200
400 Meters
Author: René de Groot
Date: 09/02/2012
Source location info: OSM
Source value info: own data



OSM - Selected Features - Rollbacks
High - Low #
5
4
3
2
1
0
100
200
400 Meters
Author: René de Groot
Date: 09/02/2012
Source location info: OSM
Source value info: own data

Appendix 6.2 Final Trust Map (point raster)

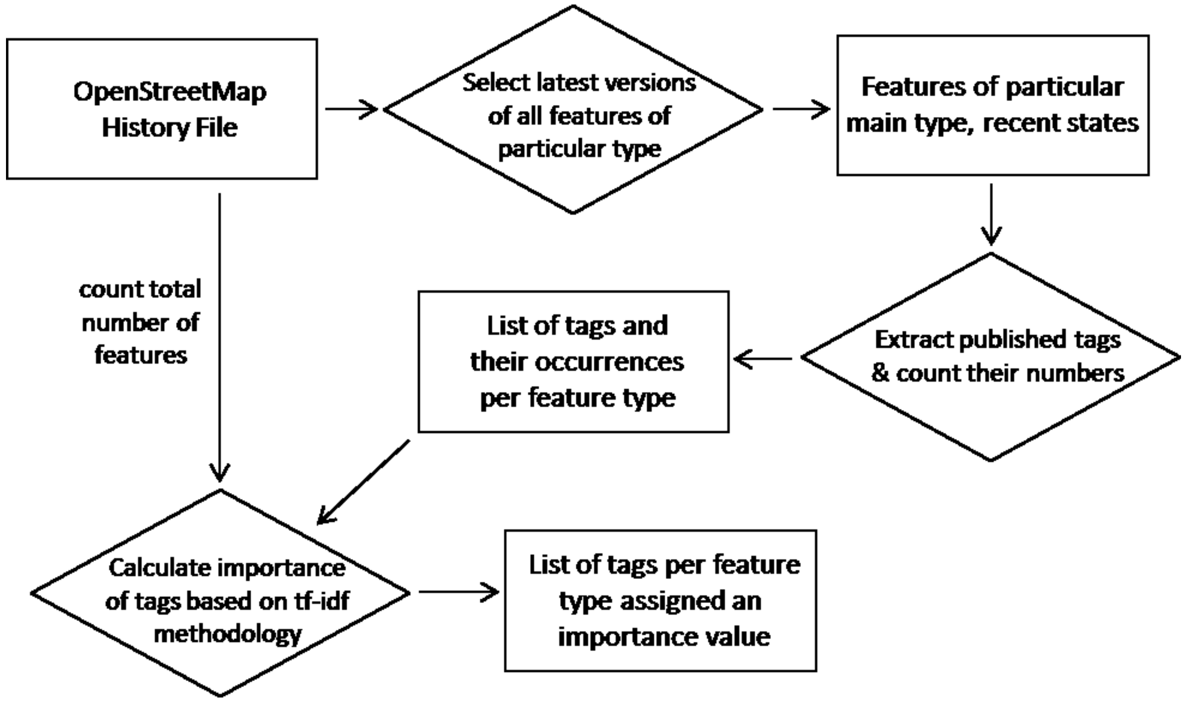


APPENDIX 7: FIELD SURVEY CLASSES

element	type	class
1	highway - primary	4
2	highway - primary	4
3	information - guidepost	3
4	highway - cycleway	4
5	highway - cycleway	1
6	highway - primary	4
7	highway - primary	4
8	amenity - pub	4
9	amenity - pub	4
10	highway - tertiary	3
11	leisure - park (pol)	2
12	highway - unclassified	1
13	highway - tertiary	3
14	highway - tertiary	3
15	highway - residential	1
16	amenity - fast food	2
17	amenity - pub	4
18	highway - tertiary	4
19	amenity - restaurant	4
20	amenity - bar	4
21	landuse - retail	4
22	amenity - restaurant	4
23	highway - residential	4
24	highway - residential	4
25	highway - residential	4
26	shop - books	4
27	amenity - pub	4
28	shop - books	4
29	amenity - café	3
30	highway - residential	4
31	waterway - river	4
32	highway - footway	4
33	amenity - place of worship (poly)	4
34	highway - residential	4
35	amenity - place of worship	4
36	amenity - library	4
37	information - guidepost	3
38	amenity - school	4
39	highway - tertiary	3
40	amenity - restaurant	4
41	highway - pedestrian	4
42	amenity - police	4

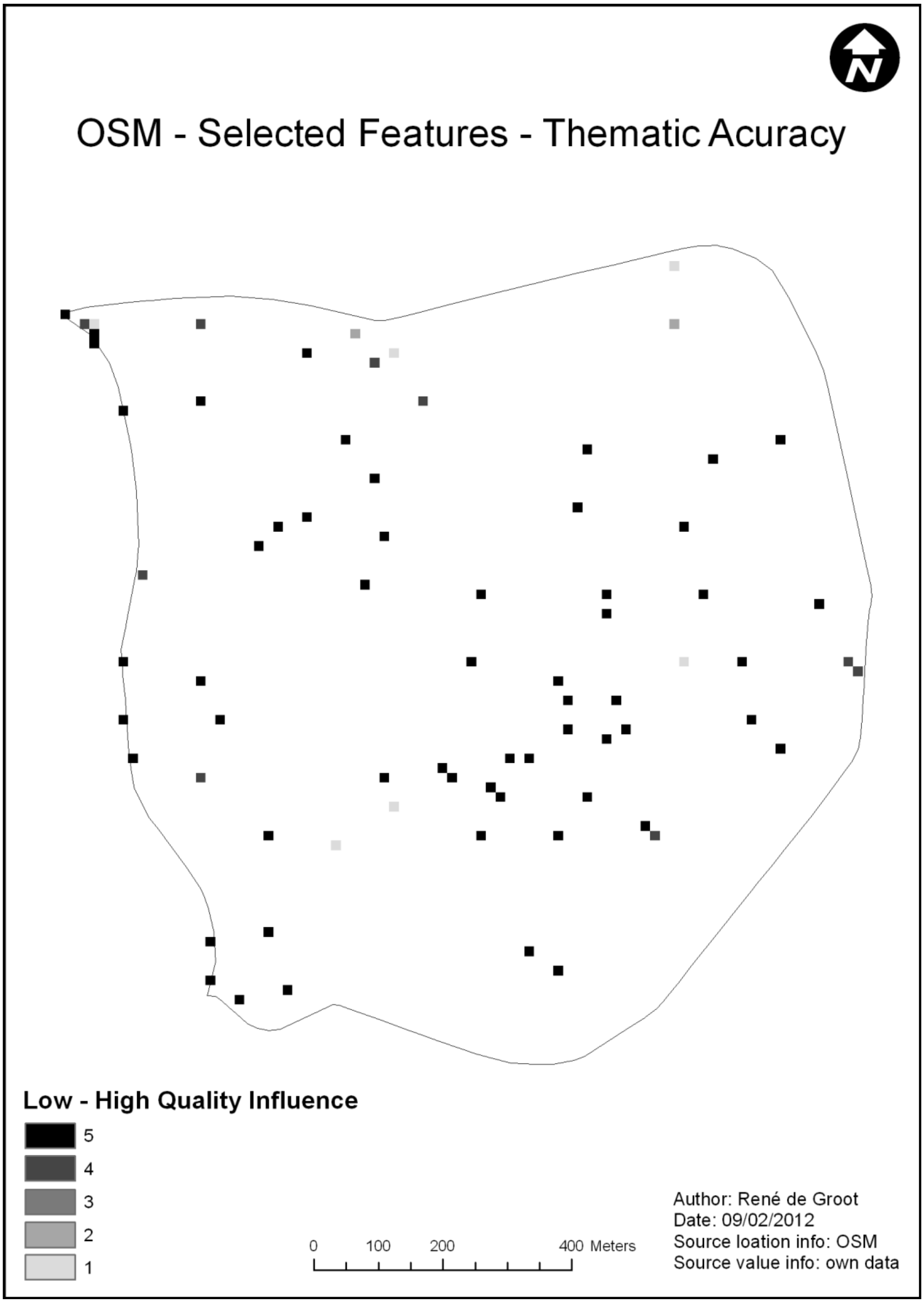
43	amenity - restaurant	1
44	highway - pedestrian	4
45	amenity - café	4
46	highway - residential	4
47	highway - residential	4
48	highway - residential	4
49	highway - residential	4
50	highway - residential	4
51	highway - primary	4
52	amenity - university	4
53	amenity - restaurant	3
54	amenity - café	4
55	highway - tertiary	1
56	highway - service	1
57	highway - residential	4
58	amenity - restaurant	4
59	highway - residential	4
60	highway - residential	4
61	highway - service	4
62	amenity - restaurant	4
63	amenity - bank	4
64	highway - pedestrian	4
65	shop - sports	4
66	amenity - parking	3
67	amenity - fast food	4
68	amenity - place of worship	4
69	highway - cycleway	4
70	highway - cycleway	4
71	amenity - café	4
72	highway - primary	4
73	highway - primary	4
74	highway - tertiary	4

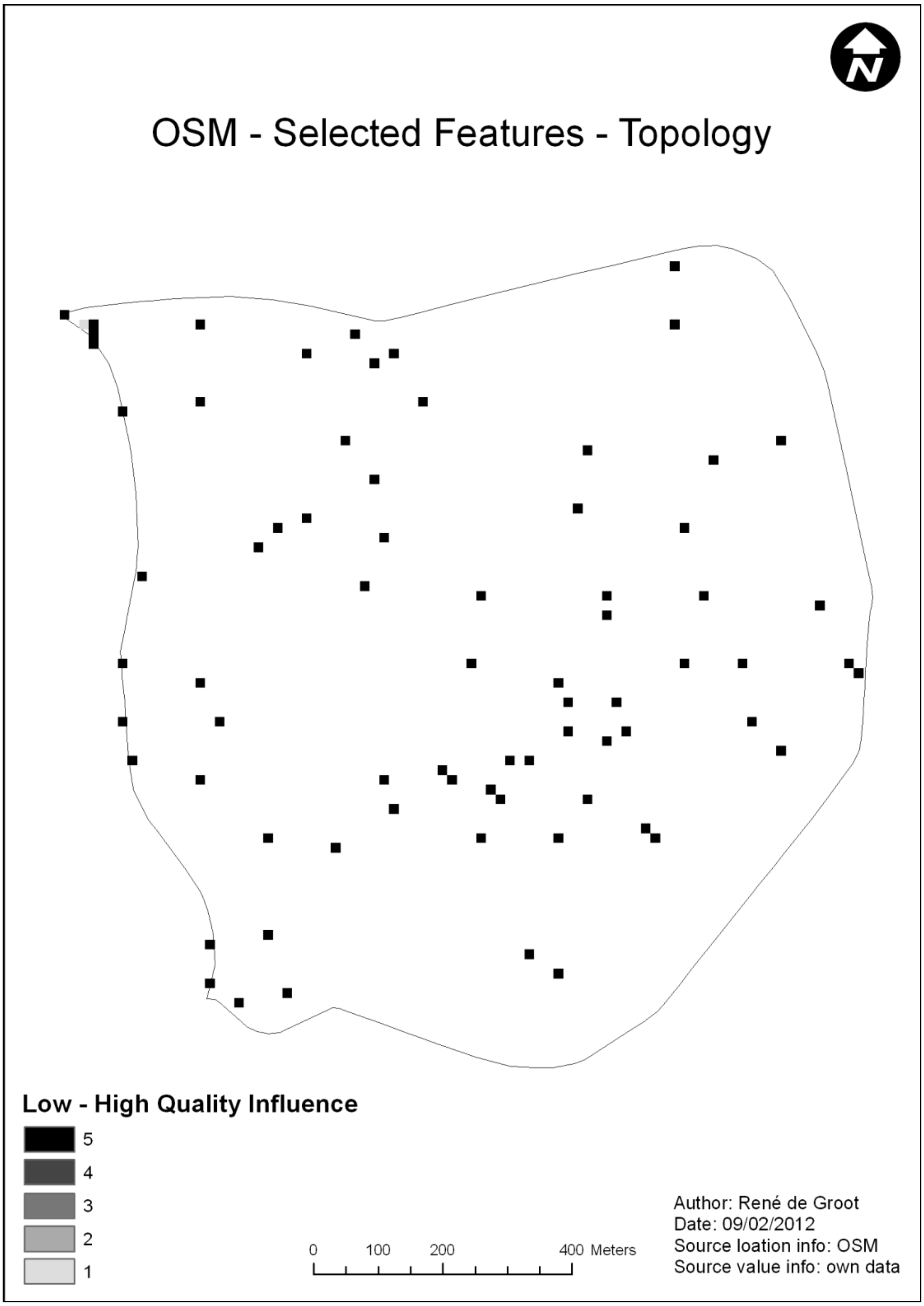
APPENDIX 8: DERIVATION OF 'OBLIGATORY TAGS'

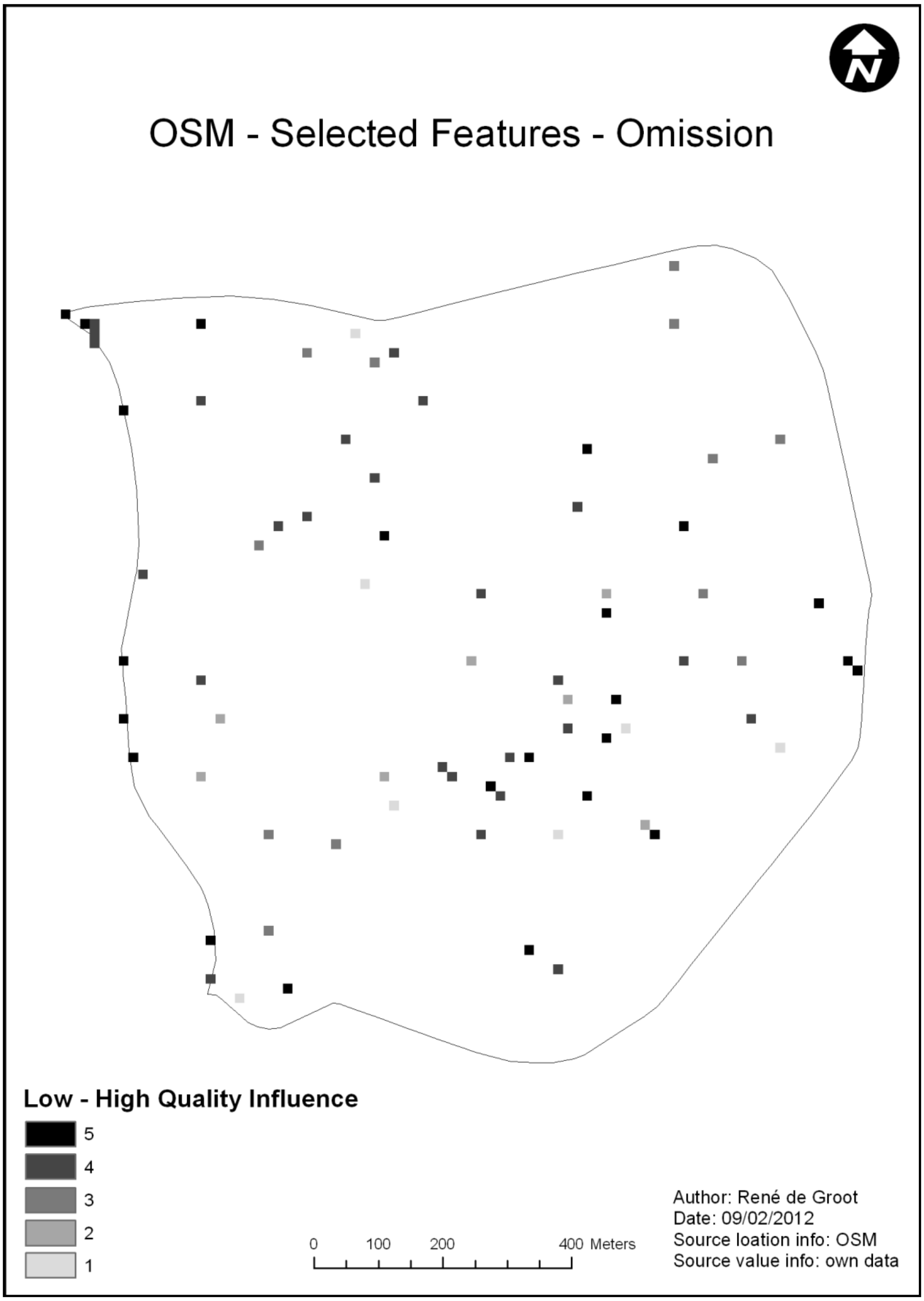


APPENDIX 9: QUALITY MEASURE

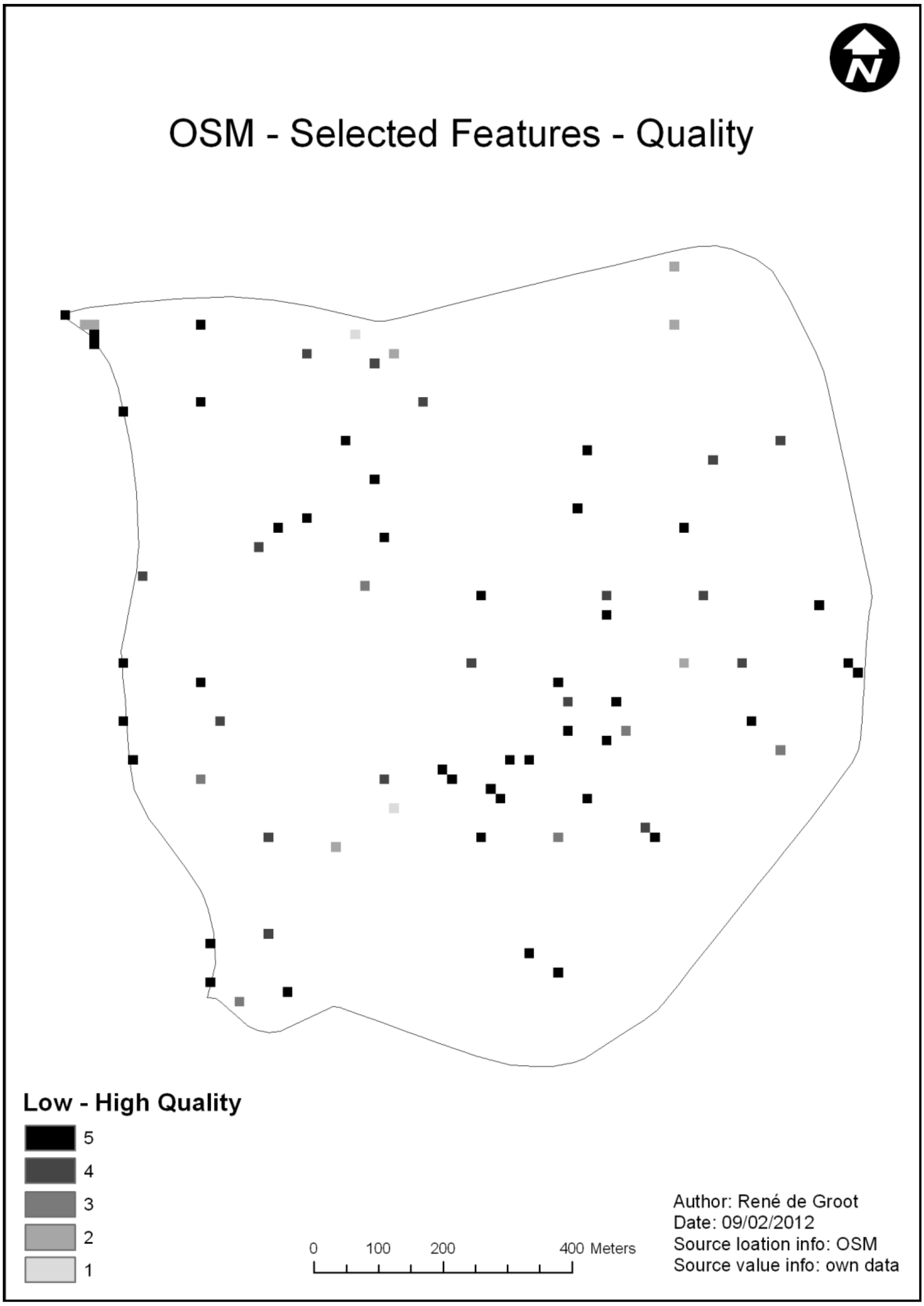
Appendix 9.1: Quality Parameters Classified and Mapped




OSM - Selected Features - Topology
Low - High Quality Influence
5
4
3
2
1
0
100
200
400 Meters
Author: René de Groot
Date: 09/02/2012
Source loation info: OSM
Source value info: own data

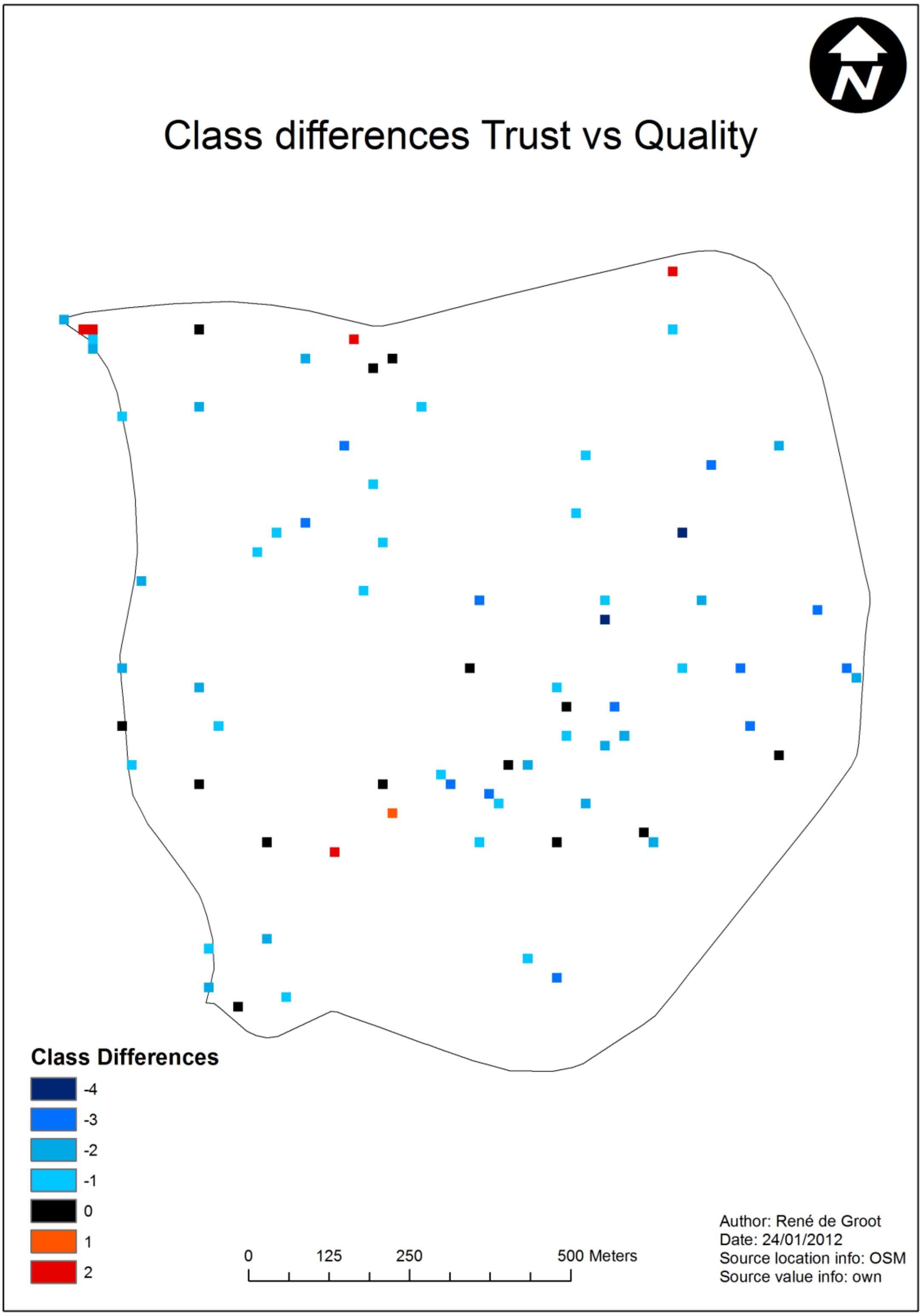

OSM - Selected Features - Omission
Low - High Quality Influence
5
4
3
2
1
0
100
200
400 Meters
Author: René de Groot
Date: 09/02/2012
Source loation info: OSM
Source value info: own data

Appendix 9.2: Final Quality Map (point raster)

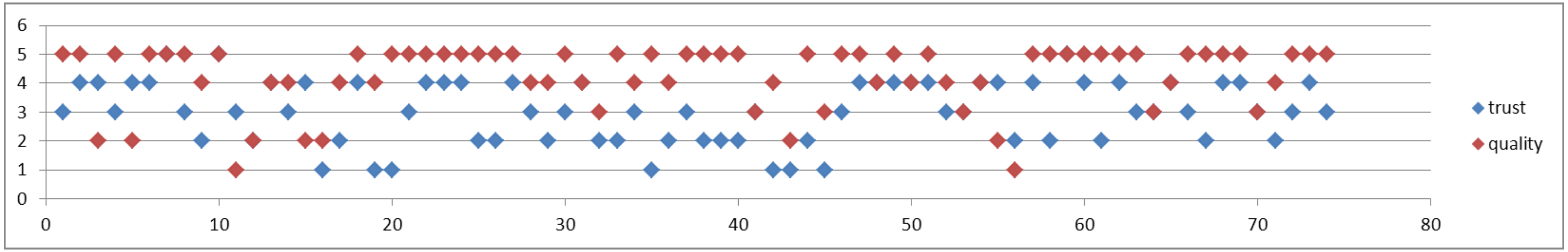


APPENDIX 10: COMPARISON & EVALUATION

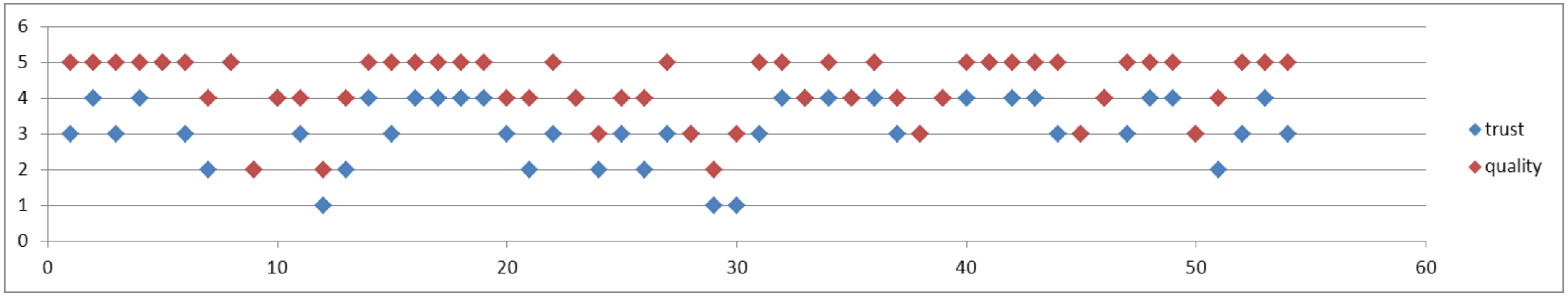
Appendix 10.1: Class differences Trust vs. Quality measures



Appendix 10.2: Distribution of class values for both measures



Appendix 10.3: Distribution of class values for features that have a class difference of -2 – 0



Appendix 10.4: Kendall's τ

All features

Kendall tau Rank Correlation	
Kendall tau	0.164
2-sided p-value	0.104
Score	300
Var(Score)	33828.102
Denominator	1830.0281

54 features

Kendall tau Rank Correlation	
Kendall tau	0.520
2-sided p-value	1.727e-05
Score	486
Var(Score)	12736.348
Denominator	935.279

DECLARATION ON PLAGIARISM

I, René de Groot, declare that the submitted work has been completed by me the undersigned and that I have not used any other than permitted reference sources or materials nor engaged in any plagiarism. I know that plagiarism is wrong. Plagiarism is to use another's work and pretend that it is one's own. All references and other sources used by me have been appropriately acknowledged in the work. I further declare that the work has not been submitted for the purpose of academic examination, either in its original or similar form, anywhere else.

Place: Münster, Germany

Date: 28 February 2012

Signature: ______________________________

Masters
Program
in **Geospatial**
Technologies

Supported by:
Education and Culture
ERASMUS MUNDUS