André Keil

OpenStreetMap Land Classification System (OSM LCS)

An Approach to Improve Global Land Cover and Land Use Mapping in OpenStreetMap

Master Thesis
August 2016

Supervisor:
Prof. Dr. Ursula Ripke
(Department for Civil Engineering and Geoinformation)

BEUTH HOCHSCHULE FÜR TECHNIK BERLIN
University of Applied Sciences

ACKNOWLEDGEMENTS

First of all, I would like to express my sincere thanks to Prof. URSULA RIPKE. Whitout her general support and constructive supervision this thesis would have never been finished.

For finally finding a topic for this study, I'm very grateful for the patience and support of Prof. MICHAEL BREUER, the second reviewer of this thesis. There have been several thoughts in very different directions and – also together with Prof. RIPKE – they have all been seriousely discussed.

ABBREVIATIONS

A	AVHRR	Advanced Very High Resolution Radiometer
C	CORINE	Coordination of Information on the Environment
E	ESA	European Space Agency
	EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
	EVI	Enhanced Vegetation Index
F	FAO	Food and Agriculture Organization
G	GLCC	Global Land Cover Characteristics Data Base
I	IGBP	International Geosphere-Biosphere Programme
	ISIC	International Standard Industrial Classification of All Economic Activities
L	LCCS	Land Cover Classification System
	LST	Land Surface Temperature
	LUCAS	Land Use/Cover Area Frame Survey
M	MERIS	Medium Resolution Imaging Spectrometer
	MODIS	Moderate Resolution Imaging Spectroradiometer
	MSS	Multispectral Scanner
N	NASA	National Aeronautics and Space Association
	NACE	Statistical Classification of Economic Activities in the EU
	NOAA	National Oceanic and Atmospheric Administration
O	OSM LCS	OpenStreetMap Land Classification System
P	POES	Polar-orbiting Operational Environmental Satellite
S	SEEA	System of Environmental-Economic Accounts
	SNA	System on National Accounts
	SPOT	Satellite pour l'Observation de la Terre
U	UMD	University of Maryland
	USGS	United States Geological Survey
V	VGI	Volunteered Geographic Information
X	XML	Extensible Markup Language

TABLE OF CONTENTS

Acknowledgements i

Abbreviations ii

Table of Contents iii

Index of Tables v

Index of Figures vi

Summary vii

I Introduction 1
- I.1 Objectives 1
- I.2 Methods 3

II Fundamentals 4
- II.1 OpenStreetMap 4
  - II.1.1 Motivation and historical development 4
  - II.1.2 Basic structure and collaborative community 6
  - II.1.3 From reality to data base 8
  - II.1.4 Using OSM: How, why and why Not? 9
  - II.1.5 Relevance for this study 10
- II.2 Mapping and classifying land cover and land use 11
  - II.2.1 Conceptions 11
  - II.2.2 Origin and development 13
  - II.2.3 Systematizing categorization 15
  - II.2.4 Utilization and challenges 18

III Assessing the Current State 21
- III.1 Land Classification Systems 21
  - III.1.1 FAO Land Cover Classification System (FAO LCCS) 21
  - III.1.2 IGBP-DISCover 28
  - III.1.3 Land Use/Cover Area Frame Statistical Survey (LUCAS) 31
  - III.1.4 Classifying land use 34
- III.2 Land cover and land use in OpenStreetMap 38
  - III.2.1 Classification and codification 39
  - III.2.2 Community perception of tagging land cover/use 44
  - III.2.3 Comparison with other classification systems 49

IV OSM Land Classification System **53**
IV.1 Requirements 53
IV.1.1 Preconditions 53
IV.1.2 General demands 58
IV.1.3 Practical requirements 59
IV.2 Specifications 62
IV.2.1 Objectives 62
IV.2.2 Concepts 62
IV.2.3 General systematics and formative factors 65
IV.3 OSM Land Cover Classification System 73
IV.3.1 Example trees (landcover=trees) 73
IV.3.2 Example water body (landcover=water) 78
IV.3.3 External and internal compatibility 81
IV.4 OSM Land Use Classification System 82
IV.4.1 Example education (landuse=education) 82
IV.4.2 External and internal compatibility 86

V Discussion **87**
V.1 General remarks 87
V.2 Improvements 88
V.3 Limitations 89
V.4 Unsettled issues 91
V.4.1 General mapping issues 91
V.4.2 Mixed properties 91
V.4.3 Integration of complex concepts 92
V.4.4 Rendering 92
V.4.5 Practical application 92

VI Conclusion **94**

VII References **95**
Online Resources 102

Appendix .... **I**

INDEX OF TABLES

Table 1 Motivation for contributing to a VGI project. 7
Table 2 Possible challenges while using land cover information retrieved from remote sensing data. 19
Table 3 Hierarchy and codification of classifiers in the Land Cover Classification System. . 23
Table 4 Overview on three map products using LCCS. 26
Table 5 Land cover classes of selected LCCS based global map products. 27
Table 6 IGBP-DIScover classes and their inherent criteria combinations. 29
Table 7 Overview on three map products using the IGBP's classification system. 30
Table 8 IGBP-DIS and UMD legends in the context of the MODIS Land Cover Type product....................31
Table 9 Classes of 1st and 2nd level used for the LUCAS programme. 32
Table 10 Details on land cover classes applied in the LUCAS programme. 32
Table 11 Details on land use classes used in the LUCAS programme. 33
Table 12 Examples for the 4th level of the LUCAS classification. 34
Table 13 The Structure of the ISIC system. 36
Table 14 Structure of the high-level SNA/ISIC aggregation. 37
Table 15 Land cover and land use classified acc. to SEEA. 38
Table 16 Land cover objects as tagged during OSM mapping activities. 39
Table 17 Land use objects as tagged during OSM mapping activities. 40
Table 18 General structure of primary features in the OSM tagging scheme. 42
Table 19 Land cover and land use as aggregated by RAMM (2015) for Geofabrik. 43
Table 20 Types and examples of namespaces for providing features with additional details. 43
Table 21 List of sources used for reviewing the discussion on land cover/use in the OSM community. 45
Table 22 Approach for classifying land cover in OSM as proposed by user rudolf. 46
Table 23 Comparison of codification and OSM-tagging. 52
Table 24 Mixtures of trees with other vegetation types. 69
Table 25 Tabular layout of the section for vegetated areas in the OpenStreetMap Land Cover Classification System. 74
Table 26 Tabular layout of the section for non-vegetated areas in the OpenStreetMap Land Cover Classification System. 75
Table 27 Preliminary template for a tabular compilation of equivalences between the OSM LCS and the current system. 81
Table 28 Tabular layout of the OpenStreetMap Land Use Classification System (primary and secondary sector). 82
Table 29 Tabular layout of the OpenStreetMap Land Use Classification System (tertiary sector and other land use types). 83
Table 30 Preliminary template for a tabular compilation of equivalences between the current tagging scheme and the proposed OSM LCS. 86
Table 31 Virtual classes for combinations of land cover and land use conditions. 90

Index of Figures

Figure 1 Land cover and land use in the FAO LCCS. 24
Figure 2 Requirements for a new classification system. 54
Figure 3 Schematic illustration of vegetation coverage. 68
Figure 4 Mixture of land cover types. 70

SUMMARY

The OpenStreetMap project was initiated in 2004 and has since then become the most remarkable example in the realm of Volunteered Geographic Information. Geospatial data is collected by means of a crowd-source activities. By this, a free accessible basis for a multitude of purposes has been created: cartographic products and routing services, just to name the most prominent.

In terms of land cover, professional data sets are used differently: Public administration, taxing, environmental protection and parameterizing geo-ecological modelling are only exemplary utilization scenarios that have developed over several decades. Meanwhile, amount, actuality and uncomplicated accessibility of OpenStreetMap data has aroused interest also among professional users. Unfortunately, the use of this data causes well-known problems: Due to substantial differences in terms of source data, mapping procedures and – especially – underlying classification systems, land cover data sets from different providers are very difficult to be combined.

The potentials of OpenStreetMap data as an additional source of reference is still limited: The current way of classifying land cover information causes problems during data extraction. The combination with professional data sets is difficult because they are based on other classification systems. In addition, the OSM data base contains systematic mistakes and inconsistencies. They are caused by shortcomings of the current system, that are known and discussed in the community without having reached a satisfying solution yet.

This study accepts the challenge of finding a way to improve land cover information in OpenStreetMap and to increase compatibility with professional data sets. This is realized by proposing a new classification system, the *OpenStreetMap Land Classification System (OSM LCS)*. It is divided into a land cover and a land use section, that aim on compatibility with the *FAO Land Cover Classification System (LCCS)* and the *International Standard Industrial Classification of all Economic Activities (ISIC),* respectively. The thesis provides a preliminary layout and templates for crucial components of the system, as well as recommendations for the mapping practice.

First of all, these results are meant to initiate a community process and will likely be reworked anyhow. Therefore, another outcome is even more important: Classification systems are characterized by a specific setting of purposes and motivations. These have to be determined in order to set up an appropriate classification system. The study considers classification in general, evaluates the specifications of established classification systems, investigates the background of the OpenStreetMap project and its community, and – of course – analyses the peculiarities of OSM data and their differences to professional classification systems. Herein, especially the tagging scheme has been identified as an essentially limiting factor. In contrast to other systems, the development of classes is strongly affected by the tag format. Upon these and other findings, a set of requirements has been built specifically for the OSM LCS.

In summary, the thesis presents a preliminary version of a new classification system including its conceptual basis. It provides extensive background information that is intended to form a substantial basis for constructive discussions and developments in mapping land cover and land use in OpenStreetMap.

Zusammenfassung

I Introduction

Since its in initiation in 2004 the *OpenStreetMap* project has become the most remarkable example in the realm of *Volunteered Geographic Information (VGI)*. The data is collected worldwide on-ground or by tracing objects on satellite imagery. It is provided under a free content licence and without any intentional constraints in terms of its utilization. It has meanwhile become an essential part for companies that provide geospatial services.

Collection and processing of remote sensing data goes further back in history. Especially since the mid 1980s scopes and complexity of data utilization have increased due to better image resolutions and tremendous progress in computing capacity. The results of automatic land cover investigations were gaining importance. The increasing diversity of utilization scenarios demanded for specifically processed data and higher data quality. Different land cover classification systems have been developed in order to meet these demands.

The biggest challenges that accompany this development until today are the limited amount of reference data and the need for combining differently classified data. Collecting ground truth data or conducting detailed manual investigation on satellites imagery is very resource demanding.

OpenStreetMap provides data mapped in the aforementioned ways. Depending on the general mapping activity in an area, the data can be the most up-to-date available . Consequently, the question arises on how OSM data can be used as a serious alternative or supplement for professionally arranged reference data sets. It is therefore inevitable to understand the classification systematics on both sides. This shall be used as a starting point for addressing the need of improving OSM data quality and expanding the potentials of external utilization.

I.1 Objectives

The final aim of this thesis is to develop a preliminary version of a new *OpenStreetMap Land Classification System (OSM LCS)*. Emphasize will be put on the formulation of the conceptual basis: The analysis of framing conditions and requirements shall support the community during the further developed of the system. In addition, practical recommendations regarding the design of essential components and class specifications will be developed.

In order to reach this goal, information about other classification systems and about classification processes in general have to be provided. This shall lead to a better understanding on how classification actually works and how it is put into practice.

Based on these information an awareness of the potentials of a new classification system can be created. Externally, compatibility with other classification systems can be improved and the role of OSM as a serious alternative in professional applications can be strengthened.

Internally, a new system can provide systematic solutions for misinterpretations, misunderstandings and conflicts that occur during the mapping process.

The knowledge about classification in general, the awareness about the demands of professional data users and the consideration of the situation in the OSM project shall finally provide the basis for a systematic built-up of the OSM LCS.

Research questions

The basic layout of the study follows two major strings that are finally combined: One considers land cover classification in general, and its practical realization by using the examples of established land cover classification systems; another one considers the OpenStreetMap project and their way of mapping and classifying land cover. Thus, every string is subdivided into a level of more general information (chapter II) and a level of rather specific information (chapter III). They are determined by the following questions:

- How can *land cover* and further related concepts be defined? What is *classification* and how is it applied for the case of *land cover*? How important is a separation of *land cover* and *land use*? For what purposes is land cover used?

→ corresponding chapter: II.2 Mapping and classifying land cover and land use

- What are the characteristics of established land cover classification systems? Are there approaches of classifying land use as well?

→ corresponding chapter: III.1 Land Classification Systems

- What is the OpenStreetMap project? What are the characteristics of its organizational structure and its community? How is OSM data created and used?

→ corresponding chapter: II.1 OpenStreetMap

- How is information on land cover currently collected and classified in OSM? What are the shortcommings and problems? How are these topics discussed by the community and which solutions have been proposed so far? What are the differences compared to professional land cover classification systems?

→ corresponding chapter: III.2 Land cover and land use in OpenStreetMap

The information obtained by answering these questions are finally combined for laying the foundation, upon which the new OSM LCS is built. Again, two stacks of questions guide through this part of the thesis (chapter IV):

- What are the major demands the new classification system has to address? How can these demands extracted from the information so far compiled in this study? What aspects have to be especially considered?

→ corresponding chapter: IV.1 Requirements

- How could the new system look like in order to fulfil the developed requirements? Which definitions and concepts should it be based on?

→ corresponding chapters: IV.2 Specifications
IV.3 OSM Land Cover Classification System,
IV.4 OSM Land Use Classification System

Chapter V will finally comment on the process of developing the new system. It will furthermore present improvements, limitations and remaining tasks.

I.2 Methods

This study is solely based on the collection and target-oriented evaluation of technical publications and internet sources, mainly directly related to the OpenStreetMap project. The have been used to collect information and build a pool of arguments based on the questions presented above.

The analysis of the community discussion has been conducted in a qualitative way. The threads have been filtered by using a set of simple search terms. Different opinions and arguments have then been aggregated and evaluated in chapter III.2.2.2 (Perceiving land cover and land use).

In both cases, the general procedure of this thesis aims on extracting a series of information and arguments. They allow a deeper understanding of the current OSM tagging system and the classification of land cover in general. A set of demands and requirements is built upon them in order to be finally considered during the set up of the new OSM LCS. Thus, the chapters III.2 (Land cover and land use in OpenStreetMap) and IV.1 (Requirements) are of particular importance.

II Fundamentals

II.1 OpenStreetMap

OpenStreetMap calls itself a "free, editable map of the world that is being largely built by volunteers from the scratch and released with an open-content license" (OSM_1). By stating this, the main aspects of the project are mentioned: It is a voluntarily organized crowdsource mapping project that acts worldwide (Sehra et al. 2013). It aims on creating, developing and providing *free* and *open* geospatial data on a global scale for any purpose (Mooney & Corcoran 2012a; OSM_2). In order to assure the *openness* of the data, it can be *freely* accessed, used, modified and shared according to the *Open Data Commons Open Database License (ODbL)* and is therefore collected independently from existing data sets that do not comply with this license ($oDef_1$; OSM_3).

This "process of collecting spatial data by individuals, most times on a voluntary basis" is crucial for *volunteered geography* as defined by Goodchild (2007). Compared to other formats of *user-generated content* – as known from blogs, podcasts or wikis – a geographic aspect is added, represented by a pair of coordinates (Neis & Zielstra 2014). The resulting information has been named in numerous ways, e. g. *crowd-sourced geodata* (Heipke 2010; Hudson-Smith et al. 2008), *collaborative geographic information* (Bishr & Mantelas 2008) or *volunteered geographic information (VGI)* (Goodchild 2007). Consequently, the process of "wikification of GIS" (Sui 2008) does not only focus on the information itself but widens the view holistically on the activities and concepts around it (Neis & Zielstra 2014), reflected in terms like *collaborative mapping* (Rouse et al. 2007), *participatory GIS* (Elwood 2006), *public participation GIS* (Sieber 2006) or *web mapping 2.0* (Haklay et al. 2008). In respect of production and utilization of geodata and maps, the "traditional top-down flow of information" (Mooney & Corcoran 2012a) has therefore been inverted into a bottom-up approach (Neis & Zielstra 2014). Following this way, the community of active volunteers made OSM to become one of the largest and well-known VGI-projects (Neis & Zielstra 2014) or "collective mapping endeavours in the history of human civilisation" (Curran et al. 2012).

II.1.1 Motivation and historical development

The registration of the domain *www.openstreetmap.org* on the 9th of August 2004 is still regarded as the official OSM birthday (OSM_4). With an initial focus on the UK, the project's core motivation is the "provision of free and open geographic data for the world" (Gröchenig et al. 2014a; OSM_5). Actually, collection, access and utilization is often controlled by companies or governmental institutions that request payments for acquiring the data or map products accompanied by restrictive rights of use (OSM_6). Thus, geospatial data as provided by OSM is mostly not *free.* But beside the goal of *freely* accessible geodata and map images under an open-content license, the increasing affordability of mobile GPS devices and smart phones became another main driving factor for the OSM project (Sehra et al. 2013).

More than one year after its initiation, about 1 000 members had already registered (December 2005). Three month later, in March 2006, the first draft in the *map feature documentation* was published in the *OSM Wiki* which was launched already more than one year earlier and has since then become the main source of information for the mapping community. Another birthday was celebrated in August 2006: The registration of the *OpenStreetMap Foundation (OSMF),* a non-profit organization as a legally representing and supporting – but not controlling – instance for the OSM project (OSMF$_1$; OSM$_{2, 4 \& 6}$). At this time, 3 000 members had already registered.

Another two milestones have to be mentioned for that year: In November 2006, the *slippy map* was launched, an interactive online map based on OSM-data with images rendered by the *Mapnik* toolkit. Until today, it is the standard visualisation tool for OSM-data on the project's main web page. One month later, *Yahoo!* declared that their satellite imagery can be used as a base layer for remote mapping. Until this time, mapping was conducted basically in the way like Steve Coast started mapping in 2004: By using GPS and mobile computers on the ground, emphasizing on the road network and interesting features in the near vicinity of the contributor (Curran et al. 2012; Neis & Zielstra 2014). The availability of satellite imagery marks the beginning of extensive *armchair mapping,* the process of mapping without surveying the contributed features in the real world (OSM$_{4 \& 7}$). Not only regions on other continents became subject to mapping. From now on, closely located features that were difficult to trace (especially buildings) appeared in the dataset as well. *Microsoft* followed *Yahoo!* in November 2010 and allowed their *Bing* imagery to be legally used for remote mapping. One year later, already 500 000 OSM-members were registered.

In September 2012, a main goal of the OSM project was reached by switching from the *Creative Commons* to the *ODb License* which provided a better alignment to the needs of the OSM project (OSMF$_{2 \& 8}$). At the beginning of the year 2013 the number of registered members exceeded 1 million (OSM$_4$).

The contributions by *Yahoo!* and *Microsoft* also led to a significant increase of mapping activity (Lin 2014). But OSM experienced further important influences: In 2007, huge datasets were imported for the first time, namely in the USA (OSM$_9$) and the Netherlands (OSM$_{10}$). Others followed and gave rise to problems and discussions concerning the impact on community activity and quality maintenance (Zielstra et al. 2013). Further impacts on registration numbers and mapping activities were triggered by humanitarian crisis, namely the earthquake in Haïti in 2010. The event gave start to the *Humanitarian OSM Team (HOT).* Since then, this project has organized mapping activities around various natural or humanitarian disasters with focus on supporting humanitarian aid and economic development (OSM$_{11}$). Especially those events showed significant social impact: Not only private companies like *DigitalGlobe* donated up-to-date satellite imagery for remote mapping of IDP camps or damaged bridges, roads and buildings. Also, governmental institutions changed their attitude and published tax-funded data (Curran et al. 2012).

Especially in well-mapped areas, mapping activities shifted from the completion of street networks towards the contribution of detailed features like buildings and public transport, including adherent information. Furthermore, data usable for indoor navigation, 3D applications and planning touristic routes gains increasing interest (Neis & Zielstra 2014).

II.1.2 Basic structure and collaborative community

Beside the purpose of legally representing the project, the *OSM Foundation* is an important instance for adopting further crucial background tasks: In terms of server administration the foundation cooperates with supportive partners providing computing and bandwidth capacity, mainly the *University College of London (UCL),* the *Imperial College London (ICL)* and private companies like *Bytemark Hosting* (OSM_2; $OSMF_3$). Registered as a limited company, it is allowed to receive donations which are transferred to the OSM project. The foundation also acts as the main organizer of the annual *State of the Map* conference, an important meet up for the OSM community (Neis & Zielstra 2014; $OSMF_3$). Tasks around licensing, public relations, web page and API maintenance, among others, are delegated to various working groups within the foundation ($OSMF_3$).

The complementary part to the visible side of *OpenStreetMap* is a community of meanwhile nearly 3 million (OSM_{12}) registered members, being “diverse, passionate and growing every day” (OSM_2). Research revealed the main aspects of the mentioned diversity:

Firstly in a spatial respect: Since no address is required during the registration process, it is difficult to determine the residence of a registered member. Hence, different approaches were chosen as an indicator, e. g. using a member's first edit, the region of highest activity or the country with the most edited nodes. However, it was concluded that about 75 % of all members reside in Europe, 25 % in North America and Asia, and only a few in Africa, South America and Oceania (BUDHATHOKI 2010; NEIS & ZIPF 2012). Differences in community activity and data coverage reflect this uneven distribution and point at a *digital divide* caused by regional disparities in terms of lacking internet connections, dominance of certain languages on web interfaces and in editing software, as well as higher rates of illiteracy in certain regions (GOODCHILD 2008). A general impact of population density and level of income on data contribution and community efforts is stated by NEIS et al. (2013). According to the conclusions of BUDHATHOKI (2010) and NEIS & ZIPF (2012), the majority of the contributors adds data related to their home region; only a few very active contributors map in two or more different countries. However, the majority of the data was contributed by members living more than 1 000 km away (NEIS et al. 2013).

Secondly, a phenomenon observed by studies on the *Wikipedia* project (JAVANMARDI et al. 2009; WILKINSON & HUBERMAN 2007) has also been identified at the OSM project (BUDHATHOKI 2010; NEIS & ZIPF 2012): A distinct *participation inequality,* quantitatively expressed by the *90-9-1 rule*. Accordingly, 90 % of the members never contribute or do it only once, 9 % contribute on an irregular basis. Only the remaining 1 % accounts for nearly all contributions and gains crucial importance for the entire project (MA et al. 2015; MOONEY & CORCORAN 2012b; NEIS & ZIELSTRA 2014; NIELSEN 2006). Actually, this reflects the difference between *registered members* and *contributors.* Due to high numbers of newly registered members in recent years, the share of active and long term contributors has decreased. Only every 3rd contributor is likely to stay active over several years (NEIS & ZIELSTRA 2014). Contrarily, the number of *users* is nearly impossible to determine because OSM can be utilized in numerous ways without any registration

Thirdly, diversity regarding the members' personal background: More than 97 % of the community are male (BUDHATHOKI 2013; LECHNER 2011; STARK 2010) more the 60 % range in ages between 20 and 40, about 20 % are older (BUDHATHOKI 2013; NEIS & ZIPF 2012). Concerning their educational background, surveys revealed that about 70 % reached at least a college degree (BUDHATHOKI 2013; LECHNER 2011; STEPHENS 2013) and about half of the community works in professions related to geography, geomatics, urban planning or computer and information science (BUDHATHOKI 2010). One could therefore question to call OSM an amateur project (NEIS & ZIELSTRA 2014). However, engagement for OSM still stays voluntary and the mentioned professions do not necessarily imply a mapping expertise.

The last aspect of diversity to be dealt with can be found within the various motivations on which the members build their commitment. According to BUDHATHOKI et al. (2008) contributors in VGI-projects generally share a very strong motivation, fostered by a common identity such as being an *OpenStreetMapper* (LIN 2014). Another unifying aspect is localism, since many contributors' initial motivation is to correct or add familiar features that they found to be wrong or missing in the current state of the map (CURRAN et al. 2012; LIN 2014; MOONEY & CORCORAN 2012a). But apart from that, extrinsic and intrinsic motivations can be quite different (see table 1). Additionally, the social and professional context of a person can influence the individual reason for starting and maintaining the commitment (LIN 2011). Motivation is consequently a very individual factor that influences the selection of objects for mapping and finally – after a period of trail and error – results in an individual or *vernacular* mapping style of each contributor (GERLACH 2010 & 2013).

Table 1 Motivation for contributing to a VGI project.
Sources: BUDHATHOKI (2010) and COLEMAN et al. (2009) summarized by NEIS & ZIELSTRA (2014).

intrinsic	extrinsic
altruism	social rewards
fun, recreation	career
learning, personal enrichment	personal reputation
self-expression, show-off	community/project goals

Although being called a community project, OSM has its non-collective peculiarities. Compared to other internet based social networks or UGC platforms like *Facebook, Youtube* or *Flickr,* OSM doesn't provide sophisticated tools for mutual interconnection that could compete with systems of chat communication, *friends* or *followers* in terms of interactively exchanging experiences or organizing mapping activities. Messages between registered members can be sent like e-mails or as comments to changesets. The activity of members can also be tracked by subscribing to their blog or changeset history via RSS or ATOM feeds. Most of the interaction during the mapping process happens indirectly when contributors edit the same objects. Since most of the mapping work – especially in the case of *armchair mapping* – is done in isolation, collaboration is therefore mostly a result of an accidental meeting at a common virtual geographic feature in the map (MOONEY & CORCORAN 2012b & 2012c).

However, in order to achieve a usable output and as a prerequisite for successful community development in the OSM project (GRÖCHENIG et al. 2014a), methods for communication and cooperation were established. One of the most important tools working groups and sub projects is the *OpenStreetMap Wiki* where relevant information is collected, shared and discussed (NEIS & ZIELSTRA 2014). As a part of it, the *Map Feature Documentation* has become a central platform and hub for getting an overview on how objects should be mapped, edited and *tagged* in order to harmonize with the data base. These contents result from collaborative work and discussions among community members. Despite an increasingly mobile, interconnected, yet individualized, way of mapping, remarkable interaction, cooperation and exchange of experiences takes place in mailing lists, forums, via chat or video meetings or in the real world by organized *mapping parties.* Hence, regular meet ups of enthusiastic mappers to systematically resurveying their neighbourhood still take place, as well as remotely correcting single errors that are randomly chosen world wide, e. g. via *MapRoulette* (CURRAN et al. 2012; LIN 2014).

II.1.3 From reality to data base

Basic prerequisites of adding the digital counterpart of a real object to the OSM data base is the availability of geographic and semantic information – concerning *Where?* (position, course, outline) and *What?* (OSM_{13}). Those can be collected in various ways, roughly dividable into *on site* and *remote data collection*. Gathering information on the ground by using GPS devices, laptop computers, notes and sketches on field papers, as well as video and audio recordings still provide the highest degree of accuracy and detail. It is encouraged and still favoured compared to remote *armchair mapping* (CURRAN et al. 2012; OSM_{7}). Data collected on ground cannot be imported directly into the data base but has to pass an editing process (CURRAN et al. 2012). Contrarily, when mapping remotely by using an *editor* (mainly *iD, JOSM, Potlatch* or *QGIS* (CURRAN et al. 2012; OSM_{15}), collecting and editing happen simultaneously because OSM features are created by directly tracing objects on satellite or airborne imagery. Depending on the image data provided by *Bing, Yahoo!, DigitalGlobe* and others, source and quality of the imagery can strongly vary even within a small region. Difficulties arise from misalignment, lacking rectification, cloud cover or unknown capture date which could actually mean that the mapping basis is out-of-date (OSM_{7}). Available imagery origins from *Landsat, Quickbird, GeoEye* and *WorldView* satellites among others and can provide a reasonable visual resolution (CURRAN et al. 2012).

Either way, for the step of digitizing remotely recognized objects or on site collected data, OSM offers three options for geometric representation: *Nodes* for punctual objects, *ways* for linear and areal objects (the latter one as *closed ways*) or *relations* for objects of different types that share a certain information, e. g. streets of varying types that are used by the course of a certain bus line (NEIS & ZIELSTRA 2014; RAMM et al. 2010). Subsequently, another trinity of basic options is available to be performed on an object's digitized representation: *add, modify* or *delete* (REHRL et al. 2013). Finally, features are usually provided with one or more descriptive information about the object, so called *tags*. They consist of a *key* and a *value*, combined as `key=value`. The former defines a coarse category of an object's attribute, e. g. `place` for more or less populated locations of different size. The latter provides specific details, e. g. `city` for settlements with more than 100 000 inhabitants or of certain regional importance (NEIS & ZIELSTRA 2014; OSM_{16}). Since April 2009, finishing a sequence of creations, modifications or removals causes a *changeset* to be saved containing a maximum of 50 000 *edits*. Information like the member's username, editing software, data source and optional comments are attached. Instead of directly referring to a single *edit*, members can only comment and reply *(discuss)* on the basis of the superior *changeset* (OSM_{14}).

Whereas single features are reviewed and corrected by cautious community members, special working groups take care of the import and conversion of bigger external datasets – another way of contributing to the OSM data base which is critically discussed, treated with caution and generally discouraged (CURRAN et al. 2012; $OSM_{6\ \&\ 17}$). Finally being filled in different ways and from different sources, the data base allows an instant access via its own API (currently version 0.6), read-only APIs like *Overpass* or web mapping frameworks like *Leaflet* or *OpenLayers* ($OSM_{18,\ 19\ \&\ 20}$). Furthermore, dump files are regularly generated and available in different formats via *Planet OSM* back to the beginning in 2004 or e. g. via *Geofabrik* as regional sub-packages. Currently (as of August 28th, 2016), the entire data base has a size of about 51 GB (XML format) containing more than 3.5 billion nodes, more than 365 million ways and nearly 4.5 million relations (OSM_{12}; $plOSM_1$).

II.1.4 Using OSM: How, why and why Not?

The OSM project itself does not favour a certain scenario of utilizing their data. It aims on providing *open* data that shall be used *freely* for any purpose and declares the promotion of “new and interesting uses” of the data ($OSM_{1\,\&\,2}$). Indeed, especially because of the applied license model, developers combine the fully accessible data with the freedom to develop various software applications and GIS-oriented services around it (MAIER 2014). Established companies as well as new start-ups, governmental institutions and non-governmental organizations and of course numerous individuals use OSM data or derived map images for their specific purposes (CURRAN et al. 2012; MAIER 2014; OSM_1). Beside the main utilization for cartographic products (MOONEY & CORCORAN 2012a) as well as for mapping and navigation, the data and its analysis has gained importance for spatial decision making and participatory planning (NEIS & ZIELSTRA 2014). Citizen science (ELWOOD 2010), crisis mapping (ROICK & HEUSER 2012), urban management (SONG & SUN 2010), flood damage estimation (POSER & DRANSCH 2010), wild fire evacuation (PULTAR et al. 2009) or risk/crises/disaster management and response (BONO & GUTIÉRREZ 2011; GOODCHILD & GLENNON 2010; HORITA & ALBUQUERQUE 2013; MANFRÉ et al. 2012; NEIS et al. 2010; OSTERMANN & SPINSANTI 2011) are further areas of usage and scientific interest. Researchers, e.g. working on neighbourhood characteristics, accessibility analysis, spatial proximity studies (MAIER 2014) or process-related modelling in geosciences (MOONEY & CORCORAN 2012a), increasingly appreciate the access to such a huge amount of labelled data (SEHRA et al. 2013). Beside the favourable collection, license and access situation, OSM data is simply the only available, affordable or most actual source for some regions (NEIS & ZIELSTRA 2014).

However, since science *uses* the data rather than *collecting* it (SEHRA et al. 2013), researchers are required to carefully assess the quality of the data according to the intended object and region of research (Gröchenig et al. 2014a). Related studies mainly focus on comparing OSM data with professionally collected or officially provided datasets, often emphasizing on road network evaluation (NEIS & ZIELSTRA 2014). But because of the general data collection and modification concept on the one hand and experiences with other crowd-sourced projects like *Wikipedia* on the other hand, general quality and credibility issues are widely anticipated and often detected. Conclusions ranging from OSM being “spatially rich but semantically poor” (BALLATORE & BERTOLOTTO 2011) and “will never be sufficient as long as it relies on non-expert volunteers of unknown identity” (WELSER et al. 2011) to “probably better [...] no mapping at all, rather than inaccurate” (FAIRBAIRN & AL-BAKRI 2013).

Nevertheless, depending on the desired scenario of utilization, for many projects OSM can be a valuable option for replacing proprietary data (MOONEY & CORCORAN 2012a; ZIELSTRA & ZIPF 2010a & 2010b), e.g. as base map for *Flickr, Foursquare* or *Wikipedia*. For science, general quality issues remain crucial: Credibility and consistency has to be questioned because identity, residence and therefore the trustworthiness of a contributor's local expertise are unknown (WELSER et al. 2011). Due to the influence of the personal background on the individual mapping style, the “database is subject to whims, experimentation and mistakes” (OSM_6). The spatial and semantic heterogeneity of OSM data is caused by differing levels of mapping experience and local knowledge, combined with the tendency of regions with high concentrations of active members showing a more saturated and maintained data coverage (GRÖCHENIG et al. 2014b). Consequently, map data in regions with a low density of community members is much more likely to be contributed from far away. Its credibility and quality is therefore highly influenced by the spatial and temporal resolution of the imagery used during the process of armchair mapping (MOONEY & CORCORAN 2012a; NEIS & ZIELSTRA 2014). Additionally, the heterogeneity of the data is increased by inconsistent metadata, unpredictable modification of features (MOONEY & CORCORAN 2011) and unevenly distributed mapping activity caused by differing popularity of working on certain objects or tasks (NEIS & ZIELSTRA 2014).

However, since the problems mentioned above can be seen as a consequence of the basic concept of OSM, potentials can also be determined out of the same reason. Although a certain degree of incompleteness, inconsistency and heterogeneity exists ubiquitously and is permanently generated, errors can be instantly corrected by every registered member. This causes the map to be highly adaptable to sudden changes in the real world (OSM_6). It provides the requirements of being able to react on intentional or accidental vandalism accompanied with a growing membership (POTTHAST et al. 2008). Also quality in terms of spatial accuracy and data volume benefits from an increasing number of active contributors in a certain area (GIRRES & TOUYA 2010; HAKLAY et al. 2010). The important question whether a tag assigned to a certain feature has been correctly chosen according to the condition of the real object can still only be satisfyingly answered by conducting an on site confirmation. Nevertheless, at least the use of tags themselves is generally in good compliance with the way they are proposed in the OSM wiki, despite the occurrence of spelling errors or misperceptions (MOONEY & CORCORAN 2012a). This might contribute to a general trend of decreasing faultiness as stated by (SEHRA et al. 2013).

II.1.5 Relevance for this study

OSM provides a framework for collecting and managing geospatial data without favouring a certain utilization scenario. Applications and users access the data for their desired benefits and reversely become a source of feedback to the project and to the community on which data is needed and on how information should be stored. Every member is free to introduce a new focus of data collection according to personal interests or needs, e. g. *OpenWheelMap*. Especially for newly introduced features or attributes, heterogeneities are unavoidable due to the above mentioned diversity of personal motivations and arbitrary changes in the dataset. Consequently, novelties and reforms have to be well explained, carefully discussed and convincingly propagated to gain wide acceptance among contributors.

Mapping has become possible at nearly every time and every place because mobile mapping devices are wide spread among the community nowadays (LIN 2014). Thus, a scheme for mapping *land cover* and *land use* – as pursued in this thesis – has to be applicable for on site as well as for remote mapping. It has to consider real world characteristics equally recognizable on the ground and on remote sensing imagery world wide. Further details shall be added by local contributors. However, both remotely mapping contributors and on ground mappers with local knowledge do not necessarily have the technical expertise to correctly assign an attribute provided by a sophisticated tagging scheme; especially if not familiar with the geographic conditions of a region far away.

To promote mapping of *land cover* and *land use* among members of different practice levels and interests, good reasons have to be provided: If not for external utilization, e. g. geospatial modelling or as an alternative source to professionally generated land cover datasets (see chapter III.1), then at least for improving the general usability for the most common purposes of OSM data, namely navigation and production of map images. Additionally, the acknowledged importance of the *Map Feature Documentation* urges the provision of precise explanations required for new *tags* to gain significant acceptance.

II.2 Mapping and classifying land cover and land use

Any land surface is heterogeneous, and the mapping standards to acquire, represent and generalize land characteristics are about as diverse as the land surface itself.

(DI GREGORIO & O'BRIEN 2012, p. 38)

II.2.1 Conceptions

In the first instance, *land cover* and *land use* refer to the term ***land,*** which roughly describes a “delineable area of the earth's terrestrial surface, embracing all attributes [...] immediately above or below this surface” (IDWG/LUP 1994). It is recommended to include inland water and tidal flats as well, since blurry or changing boundaries makes it difficult to reliably delineate wet areas like wetlands, tidelands and swamps (DUHAMEL 2009; OP 2001).

Secondly, *land* can be characterized by ***land objects*** defined by distinguishing properties at a certain position or over a certain extent. Unfortunately, for real geographic objects this identification is often difficult, especially under natural and semi-natural conditions: Whilst it is possible to delineate clearly delimited areas (parking lots, cropland etc.) or objects with physical boundaries (especially buildings), the biophysical properties of *land objects* change seamlessly in ecological transition zones (e. g. from forest over shrubland to grassland). Depending on scale, thin stripes of different crops grown on one parcel can be another challenging example of the horizontal mixture of *land objects*. Further, vertical mixture can occur as well, e. g. agriculturally used forests or storey-bound levels of different human activities in urban areas. Additionally, temporal mixtures of *land objects* can be observed in case of property changes over a certain period of time (DUHAMEL 2009; OP 2001). In contrast to a *land object,* MARTINEZ & MOLLICONE (2012) define ***land key element*** as “a physical component of the land that characterizes one or more land cover classes and/or land use categories”. Hence, they represent a single characteristic part of a *land object,* e. g. a tree as an element for a forest area or a building as an element for a residential area, and contribute to the recognition and characterization of *land objects.*

By using *land* and *land objects* as “physical foundations” (DUHAMEL 2009), the term ***land cover*** describes the directly observable bio-/physical overlay of the Earth's surface (FISHER et al. 2005; VERHEYE 2009), i. e. the material that interacts with electromagnetic radiation including its context, patterns and textures. Thus, accordingly determined *land objects* are described from the view of natural science regarding various quantitative and qualitative aspects (TURNER II 1994), e. g. in a botanical, ecological, biophysical, geological and geomorphological context. Natural surfaces and artificial constructions are means of observation, conducted directly on-site or remotely via airborne and spaceborne imagery (BURLEY 1961; DUHAMEL 2009; OP 2001). Although widely used and scientifically discussed, observation via multispectral analysis and computer-assisted classification (see chapter II.2.2) is actually not the mandatory way for assessing information on *land cover*. In fact, visual true colour evaluation either on high resolution remote sensing imagery, e. g. during the inspection routines in the context of *InVeKoS* (t/a: personal experience), or in reality on-site, e. g. during detailed geoecological biotope mapping (t/a: personal experience), remains important for scientific and official purposes and is fundamental for gathering information for *OpenStreetMap*.

Comparably, defining the term ***land use*** is more complex. On the one hand, it can be equally approached by natural scientists by analysing the “syndromes of human activities” in the context of biodiversity, hydrology or biochemistry (Ellis 2013). Social scientists, on the other hand, have a more general understanding when dealing with the economical, geographical and anthropological aspects of the “human employment of the land” (Turner II 1994). They emphasize on a *functional dimension* by describing how the Earth's surface and its resources are used for retrieving benefits (Verheye 2009). Consequently, a certain *land object* is defined by its socio-economic purpose, e. g. residential, industrial, commercial, agricultural, recreational etc. When the *sequential dimension* is in focus, different stages of changing properties are used for description, e. g. ploughing, seeding, harvesting etc. (OP 2001). Bakker & Velkamp (2008) propose *primary* and *secondary land uses:* The former describes types of *land use* directly influencing *land cover* by aiming on “the production or provision of a certain quantity of a certain commodity”, e. g. agriculture and forestry. Contrarily, the latter neither does directly influence the Earth's surface nor is it necessarily bound to a delimitable area. This is the case for landscape functions or ecosystem services, e. g. water filtering and storage, protection of genetic resources and recreation/leisure/tourism. Turner II et al. (1995) introduce a similar differentiation by distinguishing the “biophysical manipulation” of the land, e. g. irrigation techniques, fertilization systems, grazing patterns, from its underlying intention, e. g. forestry, farming or livestock herding. In addition to the variety of definitions, *land use* is generally more difficult to be observed and often additional information is required (OP 2001) in order to accomplish a successful interpretation of socio-economic human activities (Fisher et al. 2005) or land management practices (Comber 2008a). For gathering this information, methods of social science, like expert interviews and surveys, can be applied (Ellis 2013) or statistics and reports of local authorities can be evaluated (Verburg et al. 2011).

In summary, a *land object* can be most easily characterised by its *material (land cover)* and the *function (land use)* that is directly related to it as a result of human activity (Clawson & Stewart 1965; Duhamel 2009). However, connections between the two aspects are manifold and complex and the one can hardly be inferred from the other based on simple 1 : 1 relations (Comber 2008b; Comber et al. 2008). The common use of *land cover* for characterizing natural environments and *land use* for characterizing urban and agricultural ones (Meinel & Hennersdorf 2002; OP 2001) is another simplification that does not sufficiently account for this complexity. Indeed, a certain cover might be linked with only one single use, as in the case of a cornfield where an agricultural use can be reasonably assumed. And vice versa, where forestry likely indicates the existence of trees. But in most cases a many-to-many relation exist (Comber 2008b; Comber et al. 2008). Consequently, one certain type of *cover* can indicate different kinds of *use:* A grass cover can be used in several ways, e. g. serving as a recreational green space in an urban park, as a runway at a remote airport or as a grazing ground in the country side. Similarly, several types of *land cover* can be characteristic for one single type of *land use:* The surface of a residential area is usually covered by numerous materials, e. g. asphalt, buildings, grass, trees etc. (Duhamel 2009; OP 2001; Turner II 1994). Furthermore, a defined area of a certain *land cover* type may host multiple kinds of *land use* simultaneously (Comber et al. 2008; Fisher et al. 2005), e. g. forests that serve for timber production as well as for recreation and conservation. Regarding a certain time period, the connection of the two can even be a causal one, since a certain type of *land use* can cause a *land cover* to change, and vice versa (LOVELAND 2012). The variety of relationships prevents *land use* to be easily delimited. Not to mention human interests, natural preconditions (VERHEYE 2009) and seasonal fluctuations (COMBER 2008b) that influence the spatial and temporal distribution of how the Earth's surface is covered and used. Additionally, extensive subsurface activities have to be considered: Underground traffic infrastructure or exploitation sites for natural resources like coal mines extend far beyond their observable above-surface counterpart. In order to avoid problems during subsequent analysis, DUHAMEL (2009) proposes to generally restrict the recording of those kinds of *land use* to the extent of their physical impact above ground, e. g. exits of subway stations, headframes or dump sites.

The conceptual differences of *land cover* and *land use* are comprehensibly distinct and often acknowledged (FISHER et al. 2005), yet the practical differentiation during observation and interpretation remains difficult. Both terms are often used synonymously, merged as *land use/cover* or are used to define consistently separated categories that are, nonetheless, arbitrarily mixed with each other within one *classification system* (COMBER et al. 2008; FISHER et al. 2005; LOVELAND 2012). In addition, difficulties arise when both aspects are observed at different spatial and temporal scales by using different instruments because the resulting *land cover* and *land use* types may change independently (DUHAMEL 2009) and lead to inconsistent combinations of the two (VELDKAMP et al. 2001). Not to mention the subsequent treatment with different post-processing and storage procedures.

Considering all these explanations, identifying *land objects* characterized by certain types of *land cover* or *use* is consequently a process of gathering characteristic information in order to enable spatial delimitation and semantic allocation. Necessarily, for this process to succeed, several kinds of information have to be collected, abstracted and aggregated (COMBER 2008a). Hence, characteristic properties and/or *key land elements* are in focus, whereas other minor aspects remain unconsidered. This simplification leads to the term ***categorization.*** Beside being a basic cognitive ability of human beings (Medin & Aguilar 1999) it is intentionally used by scientists as *classification,* a systematic form of *categorization* (JACOB 2004). It is often required to enable simultaneous analysis of multiple objects that – depending on the chosen definitions – share similar or identical attributes and behaviours (COMBER 2008b). For this study, ***classification*** means the process of assigning an object to a certain class including its identification and delimitation based on defined criteria. ***Classification systems*** – or *nomenclatures* (DUHAMEL 2009) – provide these criteria in form of *allocation rules* completed by *class names* and detailed descriptions. This study follows JACOB (2004) by using the term ***category*** for a loosely and individually constructed group of objects or properties that share similar characteristics. They are not precisely specified by an intentionally developed set of *rules,* but a *category* can serve as a conceptual precursor of a more precisely determined ***class***, as finalized in a *classification system* by using *rules* and *names.* The application of a *classification system* on a certain region at certain scale results in a ***legend*** consisting of only those *classes* that occur in the surveyed area (DI GREGORIO & O'BRIEN 2012).

II.2.2 Origin and development

Historically, the of way of describing the land surface has long been characterized by two aspects: Firstly, the influence of individual interests and specifications. The means of observation were determined by the objectives of the institution that ordered a certain assessment. Ecologists, geographers or public administrations did not follow a comprehensive approach but instead concentrated on the distribution of plant species, the character of geomorphological features or the extent of land use patterns and their socio-economic relevance (DEFOURNEY & BONTEMPS 2012). Secondly, recording *land use* has long been more important than mapping *land cover* (COMBER 2008b). From initial activities in parts of the United States during the 1920s and 30s (parts of Michigan and Wisconsin) to later efforts in the 1950s and 60s in Pennsylvania and Massachusetts, as well as in Australia and New Zealand, the focus was mainly on recording *land use.* Among these, seminal improvements have been achieved e. g. by intensively using aerial photos for the first time in order to assess the *Major Land Uses in the United States* (LOVELAND 2012).

The variety of institutional and disciplinary objectives embedded in numerous *classification systems* and map products remain an ongoing challenge for combining and comparing information about the Earth's surface, even when incorporating rather recent datasets (FISHER et al. 2005). In turn, the emphasis on recording of *land use* has shifted significantly towards assessing *land cover,* initiated by the emerging availability of satellite images in the early 1970s. This was accompanied by increasing capabilities of computers and image processing procedures (FISHER et al. 2005). For the first time in history, it became a realistic to completely picture the Earth's surface (GONG et al. 2016). Hence, questions arose about what can actually be measured or seen on images firstly characterized by rather coarse optical and spectral resolutions. The difficulties to reliably relate these results to possible types of *land use* has also lead to the need of discussing the conceptual difference of *land use* and *land cover.* Being in charge of developing the first comprehensive *Land Use and Land Cover Classification System,* especially designed for interpreting remote sensing data, ANDERSON et al. (1976) addressed these issues. However, driven by the interests and demands of the project's initiators (US Geological Survey) to assure the continuation of existing land classifications and to meet the needs of various utilization scenarios, the solution was to simply "[interpret land use by] using land cover as a principal surrogate" (ANDERSON et al. 1976, p. 7). Consequently, this conceptual confusion became constitutive for the developed *classification system,* which finally contained both, *cover* and *use classes.* FISHER et al. (2005) and COMBER et al. (2008) state the importance of the work of ANDERSON et al. (1976) in terms of being the starting point for repeatedly merging or interchanging the two concepts, as well as becoming a reference for numerous *classification systems that* readily accept an arbitrary coexistence of information on *land use* and *land cover* within the same *system* (or even within the same *class*).

The above mentioned relations between *land use* and *land cover,* and furthermore the variety of intentions underlying the generated datasets and constructed *classification systems,* caused the ongoing progress in satellite imagery quality and computer technologies to maintain their crucial importance for the development of global mapping activities during the following decades. Since the beginning in the 1970s, a continuous intensification of research has produced sophisticated image processing and classification algorithms. They incorporated results of fundamental studies on the relation between electromagnetic information and actual surface properties, as well as new statistical approaches, like fuzzy logic and regression/decision trees. (LOVELAND 2012). Meanwhile, remote sensing has become the most common and cost effective source for recording *land cover* in large extents (VERBURG et al. 2011). At the same time, information on *land over* has become a "universal panacea for land inventory [...] adopted by a wide range of disciplines" (DEFOURNEY & BONTEMPS 2012).

Increasing spatial and spectral resolution fostered large scale mapping initiatives, especially following the launch of Landsat 4 (1982), 5 (1984) and SPOT 1 (1986). By the end of the 1990s, the *International Geosphere-Biosphere Programme (IGBP)* presented the first global land cover map providing a resolution of 1 x 1 km (LOVELAND et al. 1999). Studies on data from the *Advanced Very High Resolution Radiometer (AVHRR)*, e. g. TUCKER et al. (1985), not only gave start to the development of global scale mapping (GONG et al. 2016) but also became the origin of incorporating seasonal and annual time series, as well as indices like the NDVI in order to determine *land cover.* Nevertheless, regional initiatives gained fundamental importance as well, especially the *Africover* project realized by the *Food and Agriculture Organization (FAO)*. One of its key elements is the *Land Cover Classification System (FAO LCCS)*. It was the first significant advance in the field of *classification systems* since ANDERSON et. al in 1976 and has meanwhile become a worldwide standard (LOVELAND 2012). It has even been submitted to the ISO committee (DI GREGORIO & O'BRIEN 2012). Similar to *Africover,* the project *CORINE (Coordination of Information on the Environment)* – operated by *European Space Agency (ESA)* – uses Landsat and other high resolution data to provide a regularly updated land cover dataset structured by a *classification system* which is appropriate for European countries.

Since the turn of the millennium, new sensors and projects produce advanced datasets especially on global scale (BAN et al. 2015). The satellites *Terra* and *Aqua* are operated by the *National Aeronautics and Space Agency (NASA)* and carry the sensor *MODIS (Moderate Resolution Imaging Spectrometer).* Together with the sensor *MERIS (Medium Spectral Resolution Imaging Spectrometer)* mounted on *Envisat (Environmental Satellite)* they provide another data source. Based on the former one's data, the NASA provides a land cover product with a resolution of 500 meters, using the IGBP *classification system*. The MERIS data is used by the *GlobCover 2006* project, providing a final 300 meter resolution, classified according to the *FAO LCCS.* Since 2008, the *US Geological Survey (USGS)* provides access to Landsat imagery free of charge and thereby fostering an increase in detail at global scale, as presented by the 30 meter resolution of the project *Finer Resolution Observation and Monitoring of Global Land Cover (FROM-GLC;* GONG et al. 2013*).* In general, the instant access to Landsat imagery caused a shift from baseline mapping to multi-temporal analysis in the context of change detection and seasonal or annual monitoring. Recently, the incorporation of RADAR and LIDAR technology appears as an option to substantially improve the quality of global *land cover* and *land use* data (ESCH et al. 2013; GAMBA & LISINI 2013).

II.2.3 Systematizing categorization

Until today, *classification systems* are developed by different approaches. However, they share the common aspect of being systematic representations of logical processes *(classification)* that guide and conventionalize cognitive processes *(categorization)* (BJELLAND 2004). In case of aiming at accurate identification and delimitation of spatial information and objects, *classification systems* provide criteria, names and relations of distinct *classes.* Additionally, structure and content are shaped by a certain intended purpose. (DI GREGORIO & O'BRIEN 2012; OP 2001).

The sequence of identifying, delineating and labelling objects can be twofold and is important for the layout of a *classification system:*

1. An **a priori approach** provides a framework of predefined *classes* and *rules* that cover all anticipated varieties and combinations of likely occurring objects and properties. The scope of these objects is determined by the purpose of the *classification system* and a certain “'segment' of reality” (OP 2001, p. 22) it refers to. Within this scope, all varieties of objects and properties should be addressed. This is done before the data is collected and can be conducted completely independent from a specific study area. It consequently allows users

and producers of the *classification system* and its subsequent data products to negotiate on a certain level of homogeneity and standardization. Finally, each identifiable object (or pixel in a remote sensing dataset) is delineated and labelled according to its level of conformity with a *class*' definition. (DI GREGORIO & O'BRIEN 2012)

2. On the other hand, an **a posteriori approach** is characterized by developing *classes* after the data has been collected. Therefore, the collection itself follows other approaches (e. g. the Braun-Blanquet method for recording plant species) and is open to every imaginable set of properties or objects. Subsequently, a great range of flexibility is given for categorizing or clustering the collected data. The criteria and rules can be easily adapted to the research question or the unique conditions of a certain study site. The resulting *classification systems* are likely limited in their transferability to other surveys and the classification results are prone to problems during comparison or conflation with other datasets. (DI GREGORIO & O'BRIEN 2012)

In both cases, objects are usually described and successively classified by their properties. Even though, the aggregation can be conducted following a **hierarchical** or a **non-hierarchical system.** One of the most common a priori approaches is a hierarchical **top-down tree** (or *descending tree*). Within the anticipated conditions and relations of all possibly existing objects and properties, some are prioritized according to the purpose of the *classification system*. This means, some of the interrelations are highlighted in terms of likeness, affinity and neighbourhood. They are considered to be more important or significant than others, thus, forming *categories* at a high level in the tree system. To name an example: No matter whether being coniferous or deciduous, areas characterized by numerous densely standing trees of the one or the other type still share the common aspect of being vegetated. Emphasizing this level of likeness, the same relation even links to meadows or shrubland. Emphasizing similarities in canopy cover, life-form and height, the relations between the former and latter two get lost where they persist between the two tree type areas – likely labelled as *forest* or *wood*. Following the different levels of relationships, a system of *categories* and *sub-categories* evolves contains objects aggregated according to successively narrower concepts.

Relations for hierarchically classifying *land cover* and *land use* can be logical (industrial: chemical, iron transformation, car industry, …), ontological (forest > stand > tree), co-ordinational (crop cover: maize and wheat) or temporal (sequence: bare soil → crop). There is a tendency of *land cover* types mainly sharing ontological links, whilst *land use* properties are mostly related logically.

In contrast to this commonly used approach, e. g. by the IGBP and by the ESA for the CORINE project, **combinatory systems** work with basic characteristic properties, so called ***classifiers,*** used for describing objects. In order to capture all possible objects a wide range of *classifiers* has to be defined in advance. They are of equal importance and their mutual relations are not hierarchically weighted. Hence, objects uniquely characterised by certain *classifiers* are firstly described non-hierarchically. According to the assessed combination of *classifiers,* they can be subsequently allocated to a certain *class* in the context of a hierarchical a priori system. On the other hand, they might be used for an a posteriori **clustering** resulting in a hierarchical or non-hierarchical grouping of objects. This kind of classification will only succeed if a wide range of detailed *classifiers* is equally assessable and has been recorded for many objects.

Finally, different approaches and systems can be combined: The FAO's *Land Cover Classification System* provides a structure where objects are described by general *classifiers* of higher ranks and specific *classifiers* of lower ranks. Based on the higher ranking *classifiers,* the relatively rigid *classes* of the hierarchy levels 1 to 3 are formed. The lower levels of the classification system allow more flexibility for users according to the special conditions they face in their area of interest. (Di Gregorio & O'Brien 2012; OP 2001).

Regardless of the *classification system's* structure or development, they are supposed to share some common characteristics or even “key principles” (Di Gregorio & O'Brien 2012; Duhamel 2009; Jansen 2006; OP 2001):

1. A *classification system* should strive for **completeness.** Every object within the scope determined by the system (spatially, temporally, thematically) has to be considered during the design of the *classes.* This way, an allocation of every *object* can be assured. In case of systems dealing with geographical objects, no gaps should be left after finishing a comprehensive data collection or mapping campaign.

2. *Classes* should assure **mutual exclusiveness,** i. e. *classes* should not overlap. Every object should consequently match the definition of only one *class.* Objects with mixed properties have to be excluded. Alternatively, appropriate allocation rules have to be provided or *classes* have to be adjusted. In addition, class names should be unique.

3. Important requirements for realizing completeness and exclusiveness are fulfilled if the basic layout of the *classes* follows **consistent classificatory principles.** This assures transparency regarding the aspects used for delimitation. Accordingly, **consistent naming rules** should be provided to reflect the core principles of the system and to serve as guidelines when new *classes* have to be added. Example: Red squares  and green circles  can be allocated in a system composed of the *classes* “angled”, “blue” and “circles”. Conflicts arise if blue cubes  and yellow ellipses  have to be matched. Although none of the *classes* accurately fits for the ellipses, one could come to a decision through the process of elimination: “Circles” could be the right *class* for objects being neither angled nor blue. Another case for the blue cubes that fit two classes, “blue” and “angled”. The new objects can only be effectively included by creating new classes. However, this fails because classificatory principles and naming rules are not consistent and clear. Instead, the *classes* are obviously defined by different aspects (form and colour) but nevertheless arranged at the same classification level. Additionally, the classes are named by the parallel use of nouns (indicating objects like “circles”) and adjectives (indicating properties like “angled”). Instead, naming rules should be consistent at least at one level. Furthermore, a hierarchical structure of the *classes* according to common aspects is worth to be considered, e. g. “angled”/ “round” on the top level, “rectangular”/ “circular”/ “elliptic” on a second and “cube”/ “circle”/ “ellipse” on a third level, finally followed by subdivisions according to the colours. Depending on the initial objective of the classification, the delimitation by colour could also be on the top or at any other level. However, rules are clearer this way so that further objects can be classified easily, e. g. a brown ellipsoid  at the second level (“round/elliptic”) or in a new *class* one level below (“round/elliptic/ellipsoid”).

4. Class names represent concepts or categories that need to be specified by **definitions.** They preferably provide a context to reconstruct the membership in a superordinate *class.* Further details clarify the differences to neighbouring *classes* at the same level or to similar *classes* elsewhere in the *classification system*. Additional **explanations** should elaborate boundary problems and list objects to be in- or excluded. For latter ones, hints to alternative *classes* should be provided. Presenting prototypes as a basis for comparison can support the given explanations.

5. Detailed criteria, preferably quantifiable, are provided by **identification rules.** They represent the crucial criteria for discriminating objects and allow their allocation within the boundaries of specific *classes.* They should provide further support on how to deal with mixed properties or how to address constellations of “whole and parts”. Beside measurable proper-

ties, like percentage of canopy cover, *land objects* can also be identified by *land key elements* (as defined in above chapter II.2.1) that should be explicitly described and listed.

6. Terms, definitions and explanations have to be carefully chosen in order to create **unambiguousness.** Conceptions not only have to be precisely used during the development of a *classification system*, they also have to be precisely conveyed through attentive formulations in order to avoid doubts and misinterpretations to arise later on. Unambiguousness is therefore another prerequisite for reaching completeness and exclusiveness.

7. **Spatial and temporal consistency** should assure applicability without geographical, seasonal or sequential limitations as wells as compatibility and comparability in temporal and geographical respect. Additionally, temporally consistent classification systems do not consider previous or future states of observed objects (e. g. a class for land designated for future undetermined construction projects). Instead, they are designed for dealing with the state of an object at the moment of observation.

8. *Classification systems* should work **independent from scale and modalities of data collection.** For all *land objects,* the allocation to *classes* has to be guaranteed, irregardless of the influence these factors have on their identification and delimitation. If detailed information is unassessable, the objects should still be allocatable to more general *classes.*

9. In case of objects that have not been considered by the *classification system* so far, **rules of integration** should provide guidance for adjusting existing or adding new *classes.* In this respect, the above mentioned *naming rules* and *classificatory principles* gain further importance. Because objects already classified by the system might be affected, the influence has to be assessed and updates have to be initiated according to the new rules or terminology.

10. For identifying objects not covered by the system, an **index of objects** is recommended. All objects considered and contained by the system are listed and can be found easily. Additionally, indices provide guidance during classification or verification. For convenient abbreviation and identification of classes and objects, combinations of numbers and letters are often used as codes aside class names and definitions.

11. In order to make analysis, evaluation and comparison beneficial for various disciplines and stakeholders, **compatibility with existing classification systems** of global importance should be provided. This consolidates the acceptance of a system and enables the integration into international information systems.

The above mentioned criteria are idealistic and unlikely to be found constitutive for one single *classification system.* However, shortcomings do not necessarily indicate an unheeding development of a system but have to be seen in the light of inevitable compromises because of differing interests and/or lack of resources (see previous chapter).

II.2.4 Utilization and challenges

Considering the processes of global climate change, population growth, urbanization and their consequences, *land* as a resource has become an increasingly valuable asset for generating income, wealth and power (VERHEYE 2009). Creating a basis for taxation has been an interest of governmental authorities ever since. In order to assess the manifold utilization potentials of the Earth's surface, numerous scenarios of using land cover information have created a persisting demand for actual data with variations in spatial and temporal resolution as well as in thematic focus and detail (MEINEL & HENNERSDORF 2002). Users with a wide range of scientific, administrative and economic backgrounds benefit from the availability of information and maps that play a crucial role in modelling, policy making, management and

Table 2 Possible challenges while using land cover information retrieved from remote sensing data. Sources: BAN et al. 2015; DEFOURNEY & BONTEMPS 2012; GONG et al. 2016; MEINEL & HENNERSDORF 2002; VERBURG et al. 2011.

properties & dynamics of the Earth' surface	data collection and procession	human interests and behaviour	definitions and classification systems
mixed land cover → creates dependencies on resolution and minimal mapping units transition zones → difficult delimitation temporal changes → creates dependencies on temporal resolution and observation period huge diversity → difficult generalization and aggregation <u>spectral distinctiveness</u> difficult separation of cropland and natural vegetation difficult delimitation of settlements and urban areas errors due to fluctuating reflectance, e. g. ice vs. water	<u>stage of survey</u> lack of general information on used data base (survey statistics, remote sensing data, collection methods) on-site confirmation only relevant for small areas; however, labour intensive <u>stage of preparing data and classification process</u> unknown pre-processing procedures (e. g. atmospheric correction) lack of appropriate training areas lack of appropriate validation data limitations in computer performance and data storage capacities different projections, resolutions, angles of view the larger the area covered by one data set, the bigger the inconsistencies inclosed <u>stage of classification process</u> spectral confusion between different types of land cover systematic delimitation in favour of distinctive objects → varying level of detail and accuracy lack of additional information if delimitation of land use is intended lack of key land elements small size of important key land elements → low spectral dominance in a pixel leads to wrong classification (dependent on resolution) unreliable reflectance information, e. g. high latitudes or misinterpretation of water and shadow cloud cover	demand for information on land use → generally hard to observe <u>analyst & interpreter</u> the bigger the area, the more knowledge is required errors in evaluating singular sample sites influence subsequent classification results at large scale conflict of increasing accuracy vs. decreasing objectivity subjectivity: different results between different analysts at different times lack of international cooperation <u>general intention</u> no information about the initial user interests, e. g. regarding scale, discipline and application scenario short-term funding by different agencies with individual objectives	design and criteria dependent on intended application → improper classification rules for differing analyses or landscape conditions different definitions of classes with similar or identical names misperception of similarly named classes due to imprecise or lacking description different perceptions of land cover and land use lacking separation of land cover and land use within one classification system

planning (DEFOURNEY & BONTEMPS 2012; MARTINEZ & MOLLICONE 2012; MEINEL & HENNERSDORF 2002).

The urge for adjusting and parameterizing surface process models has emerged from the need to understand global environmental and climate change. Consequently, global land cover information derived from several generations and types of spaceborne instruments have become an important variable in studies on soil erosion, biodiversity and carbon cycles. They provide fundamental knowledge for initiating, evaluating and improving activities on sustainable development, food security, atmospheric quality or environmental protection (BAN et al. 2015; GONG et al. 2016). The relation of certain land management practices and green house gas emissions, like increasing methane production in paddy fields or slash-and-burn activities during the set up of plantations, may serve as a practical example for illustrating the importance of land surface information in monitoring factors of global climate change (VERBURG et al. 2011). Planning and managing conservation areas or assessing business potentials for companies that provide energy, logistic or communication services are further fields of using land cover data (MEINEL & HENNERSDORF 2002).

However, using the great variety of datasets inevitably demands considering inherent inconsistencies, fundamental differences and shortcomings. Especially the need for applications in social and economic contexts and for middle- and long-term monitoring requires consistent data with high spatial and temporal resolution. Because the currently available data does not yet fully meet these requirements, it remains unavoidable to use data from different mapping projects (GONG et al. 2016) that likely originate from different data sources (e. g. sensors), being adapted to different regions and being recorded at different times. Table 2 summarizes the manifold challenges in applying land cover data by a rough categorization regarding their origin: 1. Properties and dynamics of the Earth's surface; 2. methodological and technical shortcomings during processing and application of the data; 3. human interests and behaviour; 4. definitions and classification systems.

III Assessing the Current State

The aim of this study is to develop a new system of classifying *land cover* and *land use* in OpenStreetMap. The understanding of how established *classification systems* are built and applied is crucial for formulating the demands on the new system. Especially when striving for a maximum of transferability and compatibility, the understanding of other *classification systems* is indispensable.

Although active OpenStreetMap members aren't equally distributed over the planet, the evolved tagging scheme has to be applicable on global scale. It has to be aligned to the requirements of remote and local mapping activities worldwide and it has to create the prerequisites for reaching some degree of consistency in the mapping results. Hence, the first part of this chapter will focus on *classification systems* designed for global application, namely the *Land Cover Classification System (LCCS),* developed by the FAO, and the classification system designed by the *International Geosphere-Biosphere Programme (IGBP).* Both are used in several land cover products of global scope, e. g. *MODIS Land Cover, GlobCover* or *Global Land Cover.*

Because of the difficulties in adequately integrating the two concepts of *land cover* and *land use*, another focus will be set on how these *classification systems* deal with this conflict. Therefore, the classification system used by the programme *LUCAS (Land Use/Cover Area Statistical Survey)* is introduced as an example for explicitly separating *land cover* and *land use* in the layout of *classes.* Finally, further possibilities for approaching the classification of *land use* are evaluated by introducing two systems that classify economic activities on a global scale.

The second chapter emphasizes on the OpenStreetMap project. After a general view on the basic structure of its system of keys and values, the categorization of features related to *land cover/use* will be evaluated. In order to formulate the demands on a new *classification system*, there will be an assessment on how the community perceives these two concepts, as well as on how efforts to separate them are discussed. Comparing the current situation of the OSM system with other professional *classification systems* shall reveal challenges to be expected and compatibility issues to be considered.

III.1 Land Classification Systems

III.1.1 FAO Land Cover Classification System (FAO LCCS)

Since the turn of the millennium, development of global land cover products remarkably accelerated. However, FAO and UNEP dealt with the problems of *semantic interoperability* between different *classification systems* already in 1996. New concepts and approaches were discussed in order to create a standardized multi-purpose system, usable for any land cover condition independent from collection method and scale. The *International Africover Working Group on Classification and Legend* presented first proposals in 1996 and 1997. These have further been developed in cooperation with other international initiatives, e. g. the IGBP and the *Land Use Land Cover Change (LUCC) Core Project.* After being exemplarily applied and tested in the *Africover Project* (1997–1999), the first version of the LCCS was published in 2000. Being amended in version 2 in 2005 (DI GREGORIO 2005), including a software framework for improving its usability, version 3 has meanwhile been released in 2011. It is widely accepted as a reference system, especially for its applicability in translating between different *classification systems* in order to compare and interrelate them. Particularly version 3 addressed the problem of building a scheme for bridging differences between varying *classification systems,* while at the same time being one in itself. Therefore, a new emphasize was placed by developing a *Land Cover Meta-Language (LCML),* now already accepted as an ISO standard (ISO 19144-2), acting as a “common reference structure for the comparison and integration of data for any generic land cover classification system, [without intending] to replace those classification systems”. (DI GREGORIO & O'BRIEN 2012)

III.1.1.1 Main features

Recalling the different concepts on building *classification systems* introduced in chapter II.2.3, the LCCS follows an a priori approach. Objects are characterised by differently ranking *classifiers* that had been designed in advanced. To assure an easy mapping process, objects are first of all designated according to three basic properties in a *dichotomous phase: Presence of vegetation, edaphic condition* and *artificiality of cover.* The combinations of these three *classifiers* are arranged in form of a hierarchical top-down tree finally providing the eight basic *classes* of the LCCS:

- cultivated and managed terrestrial areas
- natural and semi-natural terrestrial vegetation
- cultivated aquatic or regularly flooded areas
- natural and semi-natural aquatic or regularly flooded vegetation
- artificial surfaces and associated areas
- bare areas
- artificial waterbodies, snow and ice
- natural waterbodies, snow and ice

However, the hierarchical structure in this first stage of classification is not important. It rather represents how the three basic classifiers are systematically combined according to their different relations and levels of detail, in order to create the eight basic *land cover classes.* Furthermore, it introduces the fundamental outline of the subsequent *modular-hierarchical phase.*

Only after successfully assigning an object to one of the eight basic *classes,* further details can be added. They are represented as hierarchically arranged *land cover classifiers* that can be attached according to the desired level of detail, or the quality and amount of information available. For every basic *class,* a certain set of *classifiers* is available that has to be applied hierarchically. An example: For *cultivated and managed terrestrial areas* information on *crop combinations* cannot be added until the *classifiers* dealing with *life form* and *spatial aspects* are assigned. The *classifiers* are explicitly described and – if possible – quantified, e. g. varieties of vegetation cover and height. All specifications of a *classifier* are codified by a distinct letter-number-combination. Consequently, a *class* characterized by a certain combination of properties can be codified by a unique boolean formula composed of a sequence of *classifier* codes. A paddy rice field would logically codified as A4C1D4*,* where A4 represents *graminoids* as the *life form of the main crop,* C1 means *single crop* and D4 codifies the *surface irrigation* (see table 3).

Table 3 Hierarchy and codification of classifiers in the *Land Cover Classification System* (excerpt; for details see appendix table A2). Source: DI GREGORIO 2005.

1	**primarily vegetated A**							
2	**terrestrial A1**							
3	**cultivated & managed A11**							
I	**life form of the main crop**				**spatial aspect: size**		**spatial aspect: distribution**	
	trees	A1	herbaceous	A3	large- to medium-sized field	B1	continuous	B5
	broadleaved	A7	graminoids	A4	large-sized field	B3	scattered clustered	B6
	needle-leaved	A8	non-graminoids	A5	medium-sized field	B4	scattered isolated	B7
	evergreen	A9	urban vegetated areas	A6	small-sized field	B2		
	deciduous	A10	parks	A11				
	shrubs	A2	parkland	A12				
	broadleaved	A7	lawns	A13				
	needle-leaved	A8						
	evergreen	A9						
	deciduous	A10						
II	**crop combination**							
	single crop	C1	multiple crops	C2				
			one additional crop	C3				
			trees	C5	trees	C13		
			shrubs	C6	shrubs	C14		
			herbaceous terrestrial	C7	gramanoids	C15		
			herbaceous aquatic	C8	non-gramanoids	C16		
			simultaneous	C17	simultaneous	C17		
			overlapping	C18	overlapping	C18		
			sequential	C19	sequential	C19		
III	**cultural practices: water supply**				**cultural practices: cultivation time factor**			
	rainfed	D1			shifting cultivation	D7		
	post-flooding	D2			fallow system	D8		
	irrigated	D3			permanent cultivation	D9		
	surface irrigation	D4						
	sprinkler irrigation	D5						
	drip irrigation	D6						

Additionally, further details can be attached by using *attributes: Environmental attributes* describe climatic, geomorphologic and pedologic conditions among others; *specific technical attributes* refer to methodological aspects during data assessment. Similar to the availability of *classifiers,* every basic *class* provides a specific set of *attributes.* However, they are not hierarchically ordered and can be freely combined and attached after at least one of the basic *land cover classifiers* had been assigned (for detail appendix table A2). Furthermore, the system allows the parallel combination of *classifiers* and *attributes* in order to express spatial and temporal mixtures of properties.

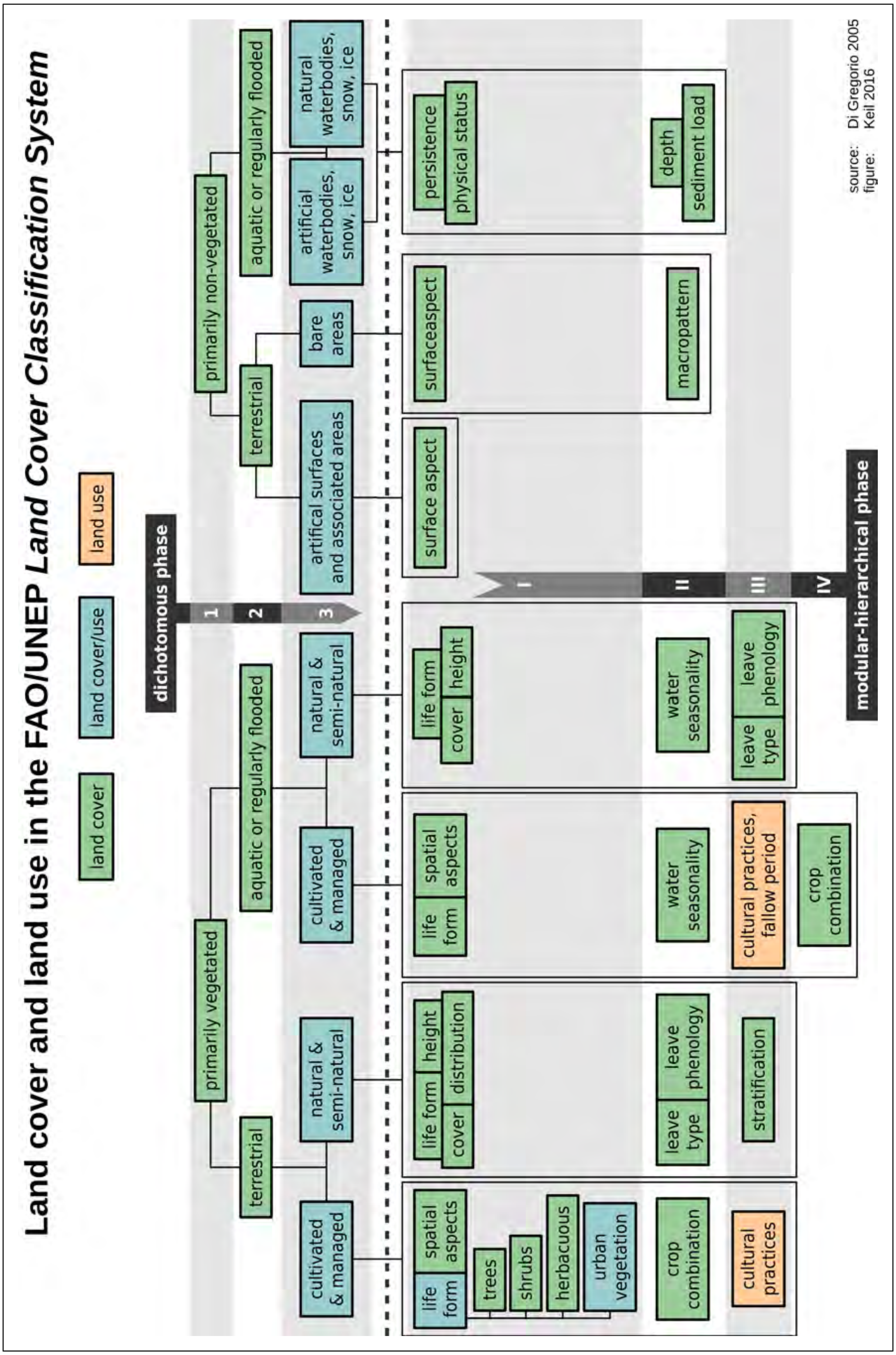


Figure 1 Land cover and land use in the FAO LCCS.

Beside providing a *classification system* and a language for translating and combining different *systems,* the LCCS is used as a basis for map products containing only a part of the initial *classes* or specially designed *classes. Classifiers* and *classes* are chosen, combined and arranged according to special regional conditions and technical objectives. The resulting *legends* might then loosely follow the LCCS hierarchy or are entirely independent from the original structure. Nevertheless, these kinds of *legends* based on LCCS/LMCL remain compatible because their *classes* follow the definitions of the LCCS classifiers.

The *legends* of three important global map products – *GlobCover, Global Land Cover* and *Climate Change Initiative Land Cover* – are introduced below.

III.1.1.2 Assessing land use

The difference between the concepts of *land use* and *land cover* has been explained in chapter II.2, including references to the problems related to their conflation, their synonymous use and their arbitrary integration within one *classification system.* As critically acclaimed by COMBER et al. (2008), the LCCS is also characterized by combining information on both aspects at different positions of the hierarchical structure.

Already during the initial *dichotomous phase,* the classifier *artificiality of cover* is used and does actually contain basic information on the state of land utilization (see figure 1). Further information on *land use* is codified in *classifiers* of the *modular-hierarchical phase.* The *classifier life form* can be specified by the existence of *trees, shrubs* or *herbaceous,* but also by *urban vegetated areas,* namely *parks, parkland* or *lawns.* The use of these specifications depends on how land is used, either *cultivated* or *managed.* Information on *cultural practices,* for example *shifting cultivation* or the type of irrigation, are equally ranked with more *land cover* like information on *stratification* or *leave type.*

Although the manual of the LCCS introduces the two concepts and emphasizes their distinctiveness (DI GREGORIO 2005), the *classification system* does not consistently follow this perception. The design of classes and codes does neither clearly separate nor indicate *land cover* and *land use.* Instead, they are equally used next to each other at one hierarchy level or as specifications within one *classifier.*

III.1.1.3 Map products based on LCCS

Initially, the FAO used the *Africover* project to apply the LCCS in its function as a *classification system.* Due to its modular character and its clearly defined classifiers, the LCCS was successively used by other initiatives and projects as a basis for individually designed *classification systems* of global and regional scope. Three of them shall be introduced in short to exemplify the adoption of the LCCS on a global scale (see table 4).

Global Land Cover 2000 (GLC2000) is a map product provided by the European Union's *Joint Research Center.* It aims on creating a global land cover product that meets the demands disciplines and actors dealing with land cover change. Compared to the data used for earth system modelling, other legends and class definitions have to be arranged for evaluating measures on sustainable development, biodiversity protection, forest conservation and restoration. In order to provide “global consistency and regional flexibility” (BARTHOLOMÉ & BELWARD 2005) the GLC2000 legend follows a bottom-up approach: In cooperation with 30 local expert groups regionally optimized *classes* were developed by using the LCCS' *classifiers*.

Table 4 Overview on three map products using LCCS.

	Global Land Cover 2000 (GLC2000) Bartholomé & Belward 2005; Congalton et al. 2014	**GlobCover 2009** Congalton et al. 2014; Tsendbazar et al. 2015	**Climate Change Initiative Land Cover (CCI-LC)** Defourney et al. 2016; Tsendbazar et al. 2015
producer	EU Joint Research Center	European Space Agency	European Space Agency
instrument	VEGETATION on SPOT (Spot Image)	MERIS on ENVISAT (ESA)	MERIS on ENVISAT VEGETATION on SPOT
primary data	4 spectral bands, NDVI	13 spectral bands, NDVI composits	13 spectral bands (MERIS), NDVI (SPOT)
additional data	elevation (ETOPO5)	elevation (GETASSE30)	
temporal resolution; observation period	daily images; Nov. 1999 – Dec. 2000	annual; Jan. – Dec. 2009	7-day to 5-year; 2003–2012
spatial resolution	1 km	300 m	300 m
classes	up to 44 regional, 22 global	22	22 in level 1, 16 in level 2
sequence	region by region, then globally aggregated	global	global
method	unsupervised	mainly unsupervised spatio-temporal clustering; expert-based labelling	unsupervised spatio-temporal clustering; machine learning classification
validation	statistical sampling	statistical sampling	statistical sampling
accuracy	68.6 %	67.5 %	74,7%

Subsequently, 22 *classes* on the global scale were aggregated (see table 5) by removing *classifiers* that represent detailed information. Although regional legends show limited compatibility with other map products, but they provide details and are easy to be applied under unique local conditions. On the other hand, the design of the global classes was driven by the aim of providing a dataset that is compatible with other global land cover map products, knowing that its usability at local and regional level might be limited.

With *GlobCover 2005,* a successor of GLC2000 was presented by the *European Space Agency.* In cooperation with the *European Environment Agency (EEA),* FAO and IGBP, it was designed to provide a dataset of higher spatial resolution that was used for the programme *Global Observation of Forest and Land Cover Dynamics (GOFC-GOLD),* as well as for studies on land cover and land use dynamics (as conducted by the IGBP; Arino et al. 2007). After the initial success, *GlobCover 2009* provided an updated map product based on MERIS imagery collected during 2009. The production follows two steps: During *pre-processing,* MERIS images are used to create a reflectance mosaic (bi-monthly and annually), including the different steps of geometric and atmospheric corrections. This data is subsequently used in the *classification module* for producing the annual land cover product (Bontemps et al. 2011; Congalton et al. 2014). Beside providing the product itself the project emphasizes on studying the various steps of the creation and validation process, showing and improving the capability of automated map production based on large imagery input (Arino et al. 2010).

The land cover product of the *Climate Change Initiative (CCI-LC)* has been specially created by the ESA to meet the requirements of the *United Nations Framework Convention on Climate Change (UNFCCC)*. So-called *essential climate variables* and their spatial and temporal patterns are assessed by the *Global Climate Observing System (GCOS)* and the *Climate Modelling Community* (CMC). Hence, working with time series is a significant difference to the above mentioned GLC2000 and GlobCover 2009 data. Therefore, MERIS imagery with varying resolutions (7-day to 5-year and 300 to 1 000 m) from 2003 to 2012 were taken as primary input, creating one 10-year land cover dataset serving as a baseline. SPOT-VGT data

Table 5 *Land cover classes* of selected LCCS based global map products.

Global Land Cover 2000 (GLC2000) Bontemps et al. 2011	**GlobCover 2009** Bartholomé & Belward 2005	**Climate Change Initiative Land Cover (CCI-LC)** Defourney et al. 2016
tree cover, broadleaved, evergreen	post-flooding or irrigated croplands (or aquatic)	cropland, rainfed
tree cover, broadleaved, deciduous, closed	rainfed croplands	cropland, irrigated or post-flooding
tree cover, broadleaved, deciduous, open	mosaic: cropland (50–70 %)/ vegetation (20–50 %)	mosaic: cropland (> 50 %)/ natural vegetation (tree, shrub, herbaceous cover) (< 50 %)
tree cover, needle-leaved, evergreen	mosaic: vegetation (50–70 %)/ cropland (20–50 %)	mosaic: natural vegetation (tree, shrub, herbaceous cover) (> 50 %)/ cropland (< 50 %)
tree cover, needle-leaved, deciduous	closed to open (> 15 %) broadleaved evergreen or semi-deciduous forest (> 5 m)	tree cover, broadleaved, evergreen, closed to open (> 15 %)
tree cover, mixed leave type	closed (> 40 %) broadleaved deciduous forest (> 5 m)	tree cover, broadleaved, deciduous, closed to open (> 15 %)
tree cover, regularly flooded, fresh and brackish water	open (15–40 %) broadleaved deciduous forest /woodland (> 5 m)	tree cover, needle-leaved, evergreen, closed to open (> 15 %)
tree cover, regularly flooded, saline water	closed (> 40 %) needle-leaved evergreen forest (> 5 m)	tree cover, needle-leaved, deciduous, closed to open (> 15 %)
mosaic: tree cover/ other natural vegetation	open (15–40 %) needle-leaved deciduous or evergreen forest (> 5 m)	tree cover, mixed leaf type (broadleaved and needle-leaved)
tree cover, burnt	closed to open (> 15 %) mixed broadleaved and needle-leaved forest (> 5 m)	mosaic: tree and shrub (> 50 %)/ herbaceous cover (< 50 %)
shrub cover, closed – open, evergreen	mosaic: forest or shrubland (50–70 %)/ grassland (20–50 %)	mosaic: herbaceous cover (> 50 %)/ tree and shrub (< 50 %)
shrub cover, closed – open, deciduous	mosaic: grassland (50–70 %)/ forest or shrubland (20–50 %)	shrubland
herbaceous cover, closed – open	closed to open (> 15 %) (broadleaved or needle-leaved, evergreen or deciduous) shrubland (< 5 m)	grassland
sparse herbaceous or sparse shrub cover	closed to open (> 15 %) herbaceous vegetation (grassland, savannas or lichens/mosses)	lichens and mosses
regularly flooded shrub and/or herbaceous cover	sparse vegetation (< 15 %)	sparse vegetation (tree, shrub, herbaceous cover) (< 50 %)
cultivated and managed areas	closed to open (> 15 %) broadleaved forest regularly flooded (semi-permanently or temporarily, fresh or brackish water	tree cover, flooded, fresh or brackish water
mosaic: cropland/ tree cover/ other natural vegetation	closed (> 40%) broadleaved forest or shrubland permanently flooded, saline or brackish water	tree cover, flooded, saline water
mosaic: cropland/ shrub or tree cover	closed to open (> 15 %) grassland or woody vegetation on regularly flooded or waterlogged soil, fresh, brackish or saline water	shrub or herbaceous cover, flooded, fresh/saline/brackish water
bare areas	artificial surfaces and associated areas (urban areas > 50 %)	urban areas
water bodies	bare areas	bare areas
snow and ice	water bodies	water bodies
artificial surfaces and associated areas	permanent snow and ice	permanent snow and ice

as secondary input data was subsequently used to calculate past and future steps of the time series. The product finally contains three maps presenting data for 2000, 2005 and 2010. The legend used for classification has two levels: One containing information available or likely to be assessable worldwide; another one, “level 2”, is meant to carry more accurate and specific information available at regional or local scale. However, due to the LCCS classifiers used for the *classes*' definitions the resulting legend remains compatible with other *classification systems,* especially GLC2000 and GlobCover 2009 (DEFOURNEY et al. 2016).

III.1.2 IGBP-DISCover

In the early 1990s, the different projects associated to the *International Geosphere-Biosphere Programme (IGBP)* required information for global environmental modelling that available datasets were not able to provide at that time (TOWNSHEND et al. 1992). A part of the programme, responsible for *Data and Information Systems (DIS),* initiated the development of a global land cover classification product at 1 km resolution, based on AVHRR data. Finally produced by the *US Geological Survey,* the purpose of the dataset was to quickly provide classified Earth surface information with a high degree of worldwide accuracy and consistency, so it could be used by the different IGBP initiatives (FRA 2000). Additionally, the *DISCover* data with its inherent *classification system* was made to become a part of the *Global Land Cover Characteristics database (GLCC),* managed by the USGS and the *University of Nebraska-Lincoln (UNL).*

III.1.2.1 Main features

The legend with 17 *classes* is closely related to the *DISCover* product and has been developed mainly for being used in this context. However, one of the objectives comprises the compatibility with other *classification systems* commonly used for environmental modelling at that time. Furthermore, the *classes* were supposed to be “exhaustive” (so that every part of the Earth can be captured) and “exclusive” (so that *classes* do not overlap). They are also designed to be applied on imagery of higher resolution, during on-site survey and for representing mosaics of different land cover types (LOVELAND et al. 2000). Like the LCCS, the *classification system* used for the *DISCover* product follows an a priori approach. In 1995, the *IGBP Land Cover Working Group* defined a series of basic criteria to be used for class definition: ground biomass (perennial vs. annual), leaf longevity (evergreen or deciduous) and leaf type (broad or needle). The criteria were specified, combined and arranged to *classes* (see table 6). In addition, they were modified to assure compatibility with other *classification systems*. The simplicity and transparency of the criteria, as well as the worldwide consistency, might have been one of the driving factors for the IGBP-DISCover to gain importance as basic data and to be adopted and implemented in further map products (see chapter III.1.2.2 and table 7).

Prior to the allocation of the *land cover classes,* the imagery had been masked (exclusion of water bodies, barren land as well as snow, ice and urban areas) and segmented according to greenness intensities, calculated by using monthly NDVI composites *(preliminary greenness classification)*. Although being regarded as corresponded to a certain degree of seasonal homogeneity, some greenness clusters actually referred to varying land cover conditions. Hence, further separation of greenness clusters was conducted manually by interpreters using high-resolution imagery (e. g. Landsat) and other ancillary information (digital and analogue maps and atlases about land cover, land use, soil, vegetation etc.) (HANSEN et al. 2000a; LOVELAND et al. 2000).

For combining the information contained in the seven land cover layers of the GLCC data base, *seasonal land cover regions* act as minimum mapping units and as spatio-temporal references. They are characterised by similar plant communities, activities and patterns

Table 6 *IGBP-DIScover classes* and their inherent criteria combinations. Source: BELWARD 1996

class	vegetation structure	leaf longevity	leaf type	woody share	woody height
natural vegetation					
evergreen needle-leaved forest	woody	> 1 a	needle-leaf	> 60 %	
evergreen broadleaved forest			broadleaf		
deciduous needle-leaved forest		< 1 a	needle-leaf		> 2 m
deciduous broadleaved forest			broadleaf		
mixed forest		$^{>}/_{<}$ 1 a	needle- and broadleaf		
closed shrubland			needle- or broadleaf		< 2 m
open shrubland				10–60 %	
woody savanna	woody/non-woody		grass, needle- or broadleaf	30–60 %	> 2 m
savanna				10–30 %	
grassland	non-woody		grass	< 10 %	< 2 m
permanent wetlands	woody/non-woody		grass, needle- or broadleaf	0–100 %	$^{>}/_{<}$ 2 m
developed and mosaic land					
cropland	non-woody	< 1 a	broadleaf or grass	< 10 %	< 2 m
urban and built-up	–	–	–	–	–
mosaic: cropland/ natural vegetation	woody/non-woody	$^{>}/_{<}$ 1 a	grass, needle- or broadleaf	< 60 %	$^{>}/_{<}$ 2 m
non-vegetated land					
snow and ice	–	–	–	–	–
barren land	–	–	–	–	–
water bodies	–	–	–	–	–

(CONGALTON et al. 2014). Consequently, the greenness clusters represent certain land cover conditions and were re-arranged into seasonal land cover regions during a phase of *post-classification refinement.* Worldwide, 961 were described in detail by a set of continent-specific attributes. Finally, the heterogeneity of these seasonal land cover regions had to be reduced in order to be usable for modelling and to enable comparisons between regions and other *classification systems.* By aggregating them into different land cover datasets, e. g. into the *Olson Global Ecoregions* (OLSON 1994), the *USGS Land Use/ Land Cover System* (ANDERSON et al. 1976) and of course the *IGPB-DISCover classification system,* the layers of the GLCC were created.

In contrast to the FAO LCCS, the IGBP scheme does not provide criteria for describing *land use.* It can thus be seen as a pure *land cover classification system. Cropland* and *urban/built-up* are the only classes indicating human activity. However, they are basically defined by a set of *land cover* attributes (see table 6). Actually, due to the coarse resolution and the characteristics of the used AVHRR data, details on urban land use are impossible to deduce (LOVELAND et al. 2000).

III.1.2.2 Map products using the IGBP-DIS classification system

Beside being used for the *DISCover* product, the *classification system* developed by the *Land Cover Working Group* (*LCWG*; associated to IGBP) has been adapted for other global land cover maps. Noteworthy are the implementation as one of five layers contained in the *MODIS Land Cover Type* product (MOD12Q1) and the modification for the dataset produced by the *University of Maryland,* also contained in MOD12Q1 (see tables 7 and 8).

Similar to the initial objectives for the *IGBP-DISCover* product, the research group at the *University of Maryland* aimed on providing a global layer with land surface information usable for environmental modelling. However, the AVHRR sensor provides further spectral information that has not been considered by the IGBP. In addition, another classification approach was chosen, based on the earlier success of a AVHRR based global land cover dataset with 8 km resolution, providing accuracy levels between 80–90 % (DE FRIES et al. 1998).

The main steps during the production of the *IGBP-DISCover* product were the clustering of NDVI-based greenness values by an unsupervised classification procedure, followed by manual corrections based on high-resolution imagery. Contrarily, for the land cover dataset produced by the *University of Maryland (UMD)* 150 training sites have been chosen and interpreted by using Landsat MSS imagery of higher resolution and ancillary information. Regarding the AVHRR data, 41 metrics were calculated by using all five spectral bands, in addition to the NDVI composites used for *IGBP-DISCover*. The information provided by the training sites and the calculated AVHRR metrics were finally analysed and restructured statistically in order to create a decision tree. Consequently, links between spectral signatures and surface criteria had been created and were arranged in hierarchical order, leading to the IGBP-based classes at the final level. Pixels could now be assigned to a certain land cover class according to the constellation of their spectral information with an average accuracy of 65 % (HANSEN et al. 2000a & b).

Table 7 Overview on three map products using the IGBP's *classification system.*

	IGBP-DISCover	**MODIS**	**UMD**
producer	US Geological Survey	Bosten University	University of Maryland
instrument	AVHRR on POES (NOAA, EUMETSAT)	MODIS on Terra (NASA)	AVHRR on POES (NOAA, EUMETSAT)
primary data	maximum NDVI composites	12 32-day composites of 7 bands, EVI, LST	41 multi-temporal metrics from spectral bands 1–5, NDVI
additional data	elevation, ecoregions, soil types, vegetation cover		Landsat MSS
temporal resolution; observation period	monthly composites; Apr. 1992 – Mar. 1993	monthly composites	monthly composites; Apr. 1992 – Mar. 1993
spatial resolution	1 km	500 m	1 km
classes	17	17	14
sequence	continental	global	global
method	unsupervised clustering, post-classification refinement	supervised classification tree	supervised classification tree
validation	statistical sampling		–
accuracy	66.9 %		65 %
references	CONGALTON et al. 2014; FRA 2000; LOVELAND et al. 2000	FRIEDL et al. 2010	HANSEN et al. 2000a & b

The creation of the *MODIS Land Cover Type* product follows a similar procedure. The statistic processing of the data differs only in detail and is considered in this study. For the data of *Collection 5,* nearly 1 900 training sites all over the world have been interpreted by using imagery from Landsat 7, Geocover 2000 and GoogleEarth. Most of them covering sizes around 4 km² (4 × 4 MODIS pixel). The spectral information used for assigning pixels to land cover classes are derived from all seven MODIS bands. Among them, other MODIS products are used as input variables, e. g. the *Land Surface Temperature (LST; MOD11)* as well as the *BDRF/Albedo* product *(MOD43)* for calculating the *Enhanced Vegetation Index (EVI).* The data is collected in 8-day intervals and used for the calculation of 135 variables, finally being aggregated to 12 32-day composites on an annual basis. They are subsequently used in a supervised classification tree in order to classify every MODIS pixel according the IGBP scheme. By using the inherent criteria of the IGBP scheme, the four other layers of the final product – one of them the UMD layer – are calculated based on the IGBP layer.

Table 8 IGBP-DIS and UMD legends in the context of the *MODIS Land Cover Type* product.

MODIS Land Cover Type (aggregated)	**IGBP-DISCover**	**UMD**
forests	evergreen needle-leaved forest	evergreen needle-leaved forest
	deciduous needle-leaved forest	deciduous needle-leaved forest
	evergreen broadleaved forest	evergreen broadleaved forest
	deciduous broadleaved forest	deciduous broadleaved forest
	mixed forest	mixed forest
woodlands	woody savanna	woodland
	savanna	wooded grassland/ shrubland
grasses/cereals	grassland	grassland
shrublands	closed shrubland	closed shrubland
	open shrubland	open shrubland
croplands and mosaics	cropland	cropland
	mosaic: cropland/ natural vegetation	–
seasonally or permanently inundated	permanent wetland	–
unvegetated	urban and built-up	urban and built-up
	barren/ sparsely vegetated	barren/ sparsely vegetated
	water bodies	water bodies
	permanent snow and ice	–

sources: Friedl et al. 2010; Hansen et al. 2000b; McCallum et al. 2006

III.1.3 Land Use/Cover Area Frame Statistical Survey (LUCAS)

Although being restricted to the scope of Europe, the *classification system* used for the LUCAS programme shall be introduced in this study because of two reasons: Firstly, it is of particular interest due to its strict separation of *land cover* and *land use* – beginning from the data collection until the statistical aggregation and dataset provision. Secondly, in contrast to the aforementioned *classification systems* and their applications, it is designed to be used on ground, i. e. the classification process is not applied on remotely collected raw data as a more or less supervised automatic classification procedure. Instead, classification is directly conducted in the moment of the survey, on-site or on high-resolution RS imagery. Classes and instructions are designed accordingly and are likely to provide valuable information for an *OSM Land Classification System.*

Initially intended to provide data on crop estimations for the European Commission in 2001, the programme became a crucial geo-statistic reference for EU policy makers. Since 2006, when the scope was extended from agriculture to *land cover* and *land use* in general, the survey is conducted every three years; the most actual data available has been collected in 2015. The increase of information variety has also affected the fields of utilization by the EU: LUCAS data has become important for strategies on soil conservation, environmental protection and actions on facing climate change ($\text{LUCAS}_{1\,\&\,2}$). It has furthermore become an essential part of the production and validation process for the *CORINE Land Cover* data base (Büttner & Eiselt 2013).

Table 9 *Classes* of 1st and 2nd level used for the LUCAS programme. Source: EUROSTAT 2015a & b.

Land Cover			
A	artificial land	A10	roofed built-up areas
		A20	non built-up areas
		A30	other artificial areas
B	cropland	B10	cereals
		B20	root crops
		B30	non-permanent industrial crops
		B40	dry pulses, vegetables and flowers
		B50	fodder crops
		B70	permanent crops: fruit trees
		B80	other permanent crops
C	woodland	C10	broadleaved woodland
		C20	coniferous woodland
		C30	mixed woodland
D	shrubland	D10	with sparse tree cover
		D20	without tree cover
E	grassland	E10	with sparse tree/shrub cover
		E20	without tree/shrub cover
		E30	spontaneously re-vegetated
F	bare land, lichens/mossens	F10	rocks and stones
		F20	sand
		F30	lichens and mossens
		F40	other bare soil
G	water areas	G10	inland water bodies
		G20	inland running water
		G30	transitional water bodies
		G50	glaciers, permanent snow
H	wetland	H10	inland wetlands
		H20	coastal wetlands

Land Use			
U100	primary sector	U110	agriculture
		U120	forestry
		U130	aquaculture and fishing
		U140	mining and quarrying
		U150	other
U200	secondary sector	U210	energy production
		U220	industry and manufacturing
U300	tertiary sector, transport, utilities & residential	U310	transport, communication networks, storage, protection works
		U320	water and waste treatment
		U330	construction
		U340	commerce, financial, professional and information services
		U350	community services
		U360	recreation, leisure, sport
		U370	residential
U400	unused and abandoned areas	U410	abandoned areas
		U420	semi-natural and natural areas not in use

Table 10 Details on *land cover classes* applied in the LUCAS programme (excerpt).
Source: EUROSTAT 2015a & b.

C	**woodland**		tree canopy cover > 10 %; woody hedges, palm trees
C10	**broadleaved woodland**		> 75 % broadleaved
C20	**coniferous woodland**		> 75 % coniferous
	C21	**spruce dominated**	> 75 % coniferous
	C22	**pine dominated**	> 75 % coniferous
	C23	**other coniferous woodland**	spruce or pine < 75 %
C30	**mixed woodland**		broadleaved/coniferous > 25 % each
	C31	**spruce dominated**	> 75 % of the coniferous trees
	C32	**pine dominated**	> 75 % of the coniferous trees
	C33	**other coniferous woodland**	< 75 % of the coniferous trees are pine or spruce
related *land use:*			• agricultural use: grazing (U111) • forestry (U120) • connected to commercial areas (U341) • connected to community services (U350) • specifically managed for recreational purposes (U36x) • in residential areas (urban parks) (U370) • not used (U420)

Compared to the other *classification systems,* the LUCAS system shows some significant differences: Firstly, attributes are assigned to points, not to delineated areas. The location of the points are defined by the crossings of a 2 km-grid, resulting in more than one million points uniformly and regularly distributed all over the EU. Secondly, data collection is conducted manually and split into two phases: During the first phase, all points are assigned to one of eight basic *classes* (see table 9) by using high-resolution imagery, without considering *land use*. For the second phase, a part of the points is randomly chosen by considering their accessibility, former on-site investigations, the size of the administrative area and other weighting factors (MARTINO et al. 2009). This way, more the 270 000 points were selected for detailed on-site investigation conducted during the 2015 data collection cycle (LUCAS$_3$). Only these points are assigned to lower-level *classes* and attributed with further details (see tables 9 and 11). During investigation, points are regarded as circular areas of 3 meters in diameter. If located in heterogeneous conditions, an *extended window of observation* is used (diameter of 40 meters), e. g. in case of wood- and shrublands (EUROSTAT 2015a & b). Nevertheless, the information is stored as punctual data.

The *classes* are defined a priori, both for *land cover* and *land use.* They are arranged in a hierarchical structure, codified by a combination of letters – *A* to *H* for the eight basic *land cover classes, U* for *land use classes* – and numbers for subordinate *classes,* respectively. Unless information is unassessable, the most detailed level is requested to be recorded for the visited points. At this stage, a certain set of *land use types* is available for each *land cover class* (see table 10; EUROSTAT 2015a). In order to avoid allocation mistakes, the guidance documents (EUROSTAT 2015a & b) explicitly name *includes* and *excludes* for each *class,* e. g. for *greenhouses* (A13):

<u>includes</u>	• un-/used greenhouses • temporarily uncovered greenhouses	<u>excludes</u>	crops covered by fabric without raised structure

Table 11 Details on *land use classes* used in the LUCAS programme (excerpt; for full layout see appendix, table A3). Sources: EUROSTAT 2015a & b.

U200	**secondary sector**		
	U210	**energy production** plants for production and generation of electric power, steam, hot water based on fossil fuels, nuclear energy, …	
	U220	**industry and manufacturing**	
		U221	**manufacturing food, beverages and tobacco products** light end product industry for processing material of the primary sector; manufacturing and retail sale (e. g. bakeries)
		U222	**manufacturing of textile products** raw industry (e. g. weaving and dying) and light end product industry (e. g. apparel production)
		U223	**coal, oil and metal processing** mainly raw industry (e. g. smelting metals, refining petrol)
		U224	**production of non-metal mineral goods** e. g. production of glas, ceramic, shaped stones
		U225	**chemical and allied industries and manufacturing** chemical transformation of organic and inorganic raw material (e. g. production of medicine, rubber, plastic and botanical products)
		U226	**machinery and equipment** production of computers, consumer electronics, optical media, general-purpose and special machinery, motor vehicles, furniture etc.
		U227	**wood based products** raw industry processing timber (cork, straw, plaiting; sawing, planning, shaping, laminating)
		U228	**printing and reproduction** light end product industry related to publication (printing newspaper, books; book binding) and media reproduction (video, software etc.)

In some *classes,* a 4th level exists for allocating features not mentioned or outlined in sufficient detail by higher levels (examples are presented in table 12). Beside the allocation into *classes* provided by the basic classification structure (as in tables 9 to 11), additional information are recorded and attached as non-hierarchical attributes, e. g. details on the share of a certain *land cover* (split in steps ranging from *< 5 %* to *≥ 90 %*), on tree height (below or above 5 meters), on special status (e. g. *protected*) or on water management (e. g. type and source of irrigation). Furthermore, the LUCAS framework is designed to save information on secondary *land cover/use* (abbreviated *LC2* and *LU2*), similarly attributed like the primaries. The way additional information is managed reminds on the FAO LCCS where important hierarchically arranged attributes are complemented by optional ones. (EUROSTAT 2015a & b)

Table 12 Examples for the 4th level of the LUCAS classification. Source: EUROSTAT 2015a.

Land Cover		***Land Use***	
B00	cropland	U100	primary sector
B10	cereals	U110	agriculture
B19	other cereals	U111	agriculture
B19a	sorghum	A01.1	growing of non-perennial crops
B19b	buckwheat	A01.2	growing of perennial crops
B19c	millet	A01.3	plat propagation
B19d	canary grasses	A01.4	animal production
B19f	quinoa	A01.5	mixed farming
B19g	fonio	A01.6	support and post-harvest activities
B19h	others not specified elsewhere		

III.1.4 Classifying land use

Similar to the classification of *land cover,* categorizing *land use* depends on the objectives of the intended *classification system* and the concepts contained. Major *classification systems* of global scope, for instance, differentiate between an economic activity and the respective productive output. Accordingly, the *International Standard Industrial Classification of All Economic Activities (ISIC)* is used by the UN for recording economic activities. Contrarily, the *Central Product Classification (CPC)* emphasizes on products and services as an output of economic activities. Both are used as reference systems for national or regional systems in order to maintain compatibility, e. g. the *Statistical Classification of Economic Activities in the European Community (NACE)* and its counterpart, or the *Statistical Classification of Products by Activity in the European Economic Community (CPA).*

However, *land use* related information derived from or related to observable *land cover* conditions, primarily refer to information on the activity taken place. A classification according to products does therefore not seem appropriate. Activities are more likely to be observable by visible *key land elements,* i. e. characteristic textures or features. To finally name and classify an activity according to its actual productive output, it likely needs detailed ancillary information that might be difficult to access, especially in the case of remote mapping. Furthermore, an observed activity might last for a longer period but changes its outcome over time, This likely creates inaccuracies and temporal heterogeneities in a project like *OpenStreetMap* that is characterized by varying spatial and temporal update cycles. Hence, for the purpose of this study, product-based *classification systems* will not be considered.

Instead, another conceptual differentiation might be helpful in approaching *land use* mapping and categorization: JANSEN (2006) seeks for a way of harmonizing different *land use classification systems* and suggests to consider the parameters *function* and *activity.* The former one refers to the economic purpose “[...] and is commonly used for sectoral land use descriptions (e. g. agriculture, forestry, fisheries, etc.)” The latter one refers to an actual operation characterized by processing certain inputs in order to create output products. Consequently, *functions* allow different *land cover/use* types to be grouped “that do not possess the same set of observable characteristics but serve the same purpose”. Although different *land cover/use* types can also be dedicated to one certain *activity,* the “level of data collection increases notably from the 'function' to the 'activity' concept. The use of the 'function' parameter as first level parameter is proposed as a pragmatic choice”. Additionally, information on *function* is easier to be assessed. In comparison, collecting details on *activities* is more difficult and requires ancillary information.

III.1.4.1 *International Standard Industrial Classification of all Economic Activities (ISIC)*

The resolution concerning ISIC was already initiated in 1948 by the *UN Economic and Social Council* following a recommendation of the *Statistical Commission.* In order to create a systematic basis for making economic statistics internationally comparable, it was recommended to the member states to implement ISIC as a national standard. Meanwhile, it is used by the *UN Industrial Development Organisation (UNIDO),* the *International Labour Organisation (ILO)* and the FAO, among others, as a major global classification scheme. (UN-ESA 2008)

During the passing decades, world economic structures have changed and new types of industries emerged. Thus, reviews became necessary and led to amendments and adjustments. After revisions in 1958, 1968 and 1990 (updated as revision 3.1 in 2004), the recent revision 4 was issued in 2008. While increasing details in lower levels of the classification hierarchy and integration of new activities, e. g. information technologies, the main characteristics remain unchanged. (UN-ESA 2008)

The objects of classification are establishments or enterprises that are assigned to a certain class of the system according to the economic activity they are mainly engaged in, disregarding spatial aspects. However, since the design of the classes strives for being exhaustive and mutually exclusive (UN-ESA 2008), the ISIC provides a well maintained and structured list of economic activities, possibly occurring on delineable areas worldwide.

The hierarchy consists of four levels of increasing detail, namely *sections, divisions, groups* and *classes,* including a respective codification system (see table 13). In general, the subdivisions of every level are based on the “input of goods, services and factors of production; the process and technology of production; the characteristics of outputs; and the use to which the outputs are put.” (UN-ESA 2008). *Classes* of higher hierarchy levels were developed according to the activities' output and its utilization; the lower levels with more detailed information are mainly characterised by the processes and technologies of productions. To some extent, this reflects the above mentioned approach of JANSEN (2006) that differentiates between detailed *activities* on lower levels and *purposes* for the crouping at higher levels.

In practice, the ISIC is used for monitoring, analysing and evaluating economic performances. The serious efforts to perpetuate continuity between the revisions allows the data to be analysed regarding changes over time. It has become an important basis for decision- and

Table 13 The Structure of the ISIC system. Levels 2–4 are exemplarily presented for section A. Source: UN-ESA 2008.

A	**agriculture, forestry and fishing**			
	A01	crop and animal production, hunting and related service activities		
	A02	forestry and logging		
	A03	fishing and aquaculture		
		A031	fishing	
			A0311	marine fishing
			A0312	freshwater fishing
		A032	aquaculture	
			A0321	marine fishing
			A0312	freshwater fishing
B	**mining and quarrying** (contains 5 divisions, 10 groups and 14 classes)			
C	**manufacturing** (contains 24 divisions, 69 groups and 137 classes)			
D	**electricity, gas, steam and air conditioning supply** (contains 1 division, 3 groups and 3 classes)			
E	**water supply; sewerage, waste management and remediation activities** (contains 4 divisions, 6 groups and 8 classes)			
F	**construction** (contains 3 divisions, 8 groups and 11 classes)			
G	**wholesale and retail trade; repair of motor vehicles and motorcycles** (contains 3 divisions, 20 groups and 43 classes)			
H	**transportation and storage** (contains 5 divisions, 15 groups and 20 classes)			
I	**accommodation and food service activities** (contains 2 divisions, 6 groups and 7 classes)			
J	**information and communication** (contains 6 divisions, 12 groups and 23 classes)			
K	**financial and insurance activities** (contains 3 divisions, 10 groups and 18 classes)			
L	**real estate activities** (contains 1 division, 2 groups and 2 classes)			
M	**professional, scientific and technical activities** (contains 7 divisions, 14 groups and 14 classes)			
N	**administrative and support service activities** (contains 6 divisions, 6 groups and 26 classes)			
O	**public administration and defence; compulsory social activities** (contains 1 division, 3 groups and 7 classes)			
P	**education** (contains 1 division, 5 groups and 8 classes)			
Q	**human health and social work activities** (contains 3 divisions, 9 groups and 9 classes)			
R	**arts, entertainment and recreation** (contains 4 divisions, 4 groups and 10 classes)			
S	**other service activities** (contains 3 divisions, 5 groups and 17 classes)			
T	**activities of households as employers, undifferentiated goods- and service-producing activities of households for own use** (contains 2 divisions, 3 groups and 3 classes)			
U	**activities of extraterritorial organizations and bodies** (contains 1 division, 1 group and 1 class)			

policy-makers as well as for administrative entities that aim at tax collection, business licensing and other purposes (UN-ESA 2008). Similar to the above mentioned *classification systems* specialized on *land cover,* the ISIC is designed for being adapted and rearranged for special objectives or regional/national peculiarities. NACE in Europe is such an example which was used as a basis for the *land use* classification of the LUCAS programme later on (as indicated by the deviating codification pattern on the 4th level, presented in table 12). The adaption is often accomplished by selecting particular classes and/or regrouping at higher levels. For instance, the *high-level SNA/ISIC aggregation* aims on a clear differentiation of the production stages: from production and processing of raw materials to the production of goods, the provision of services as well as market and non-marked activities. The *System of National Accounts (SNA)* therefore reduced the 21 ISIC sections to only 10, as presented in table 14. (UN-ESA 2008).

Table 14 Structure of the *high-level SNA/ISIC aggregation*. Source: UN-ESA 2008.

1	**agriculture, forestry and fishing** (contains ISIC-section A)
2	**manufacturing, mining, quarrying and other industrial activities** (contains ISIC-sections B–E)
3	**construction** (contains ISIC-section F)
4	**wholesale and retail trade, transportation and storage, accommodation and food service activities** (contains ISIC-sections G–I)
5	**information and communication** (contains ISIC-section J)
6	**financial and insurance activities** (contains ISIC-section K)
7	**real estate activities** (contains ISIC-section L)
8	**professional, scientific, technical, administrative and support service activities** (contains ISIC-sections M–N)
9	**public administration and defence, education, human health and social work activities** (contains ISIC-sections O–Q)
10	**other service activities** (contains ISIC-sections R–U)

III.1.4.2 *System of Environmental-Economic Accounts (SEEA)*

Compared to the ISIC system, spatial aspects are much more considered in the SEEA, developed by the *UN Committee of Experts on Environmental Economic Accounting (UN-CEEA).* It aims on evaluating land as an asset regarding its economic importance for society and business. *Environmental assets* are therefore regarded as bio-physical components of the Earth that provide inputs for production or can be used for consumption and accumulation. For the system land areas, inland water and coastal zones are considered. *Land* as a “unique environmental asset that delineates the space in which economic activities and environmental processes take place” is not only characterized by information about topography, general segmentation patterns and the occurrence of different resources, e. g. mineral, energy, soil and timber resources. *Land cover* and *land use* are regarded as the most important additional aspects.

While the classification of *land cover* is based on the definitions and criteria of the FAO LCCS, the SEEA follows an own approach to categorize *land use,* distinguishing between those typical for terrestrial and aquatic conditions (see table 15). Although regarding *land use* as “both (i) the activities undertaken and (ii) the institutional arrangements put in place” (UN-CEEA 2011, p. 217) the classes are “not defined on the basis of economic activity but rather on consideration of the general purpose and role of the user of the area” (UNCEEA 2011, p. 218). Activities often spatially coincide with the scope of purpose. However, an area generally used for *forestry* may be bigger than the area where actual forestry operations take place (cleaning, clearcutting etc.) In the case of multiple *land uses* (spatially and temporally), the principle of primary or dominant use has to be applied; if possible, subdivision into smaller areas of uniform *land use* is recommended.

Table 15 *Land cover* and *land use* classified acc. to SEEA. Source: EUROSTAT 2015a.

Land Cover	***Land Use***		
artificial surfaces (incl. urban and associated areas)	**land**		
herbaceous crop		agriculture	
woody crops			land under temporary crops
multiple or layered crops			land under temporary meadows and pastures
grassland			land with temporary fallow
tree covered area			land under permanent crops
mangroves			land under permanent meadows and pastures
shrub covered area			land under protective cover
shrub and/or herbaceous vegetation, aquatic or regularly flooded		forestry	
sparsely natural vegetation			forest land
terrestrial barren land			other wood land
permanent snow and glaciers		used for aquaculture	
inland water bodies			land used for hatcheries
coastal water bodies and inter-tidal areas			managed grow-out sites on land
		use of built-up and related areas	
			mining and quarrying
			construction
			manufacturing
			technical infrastructure
			transport and storage
			commercial, financial and public service
			recreational facilities
			residential
		used for maintenance and restoration of environmental functions	
		used otherwise	
		not used	
	inland water	**internal water**	**exclusive economic zone**
	used for aquaculture or holding facilities		
	used for maintenance and restoration of environmental functions		
	used otherwise		
	not used		

III.2 Land cover and land use in OpenStreetMap

"OpenStreetMap represents physical features on the ground (e. g., roads and buildings) using tags to its basic data structures (its nodes, ways and relations)."

(OSM21)

The general way of recording objects in the real world and making them to become a feature in the OSM data base is done by adding or modifying and has been described in chapter II.1.3. Objects related to *land cover* and *land use* are treated the same way. According

Table 16 *Land cover* objects as tagged during OSM mapping activities. Sources: OSM[25 & 26].

object	feature tagging
bare earth	
bog	`natural=wetland + wetland=bog`
grass	`landuse=grass`, `landuse=meadow`, `leisure=park`, `natural=grassland`, `surface=grass`
gravel	`surface=gravel`
mangrove	`natural=wetland + wetland=mangrove`
marsh	`natural=wetland + wetland=marsh`
mud	`natural=mud`
reed bed	`natural=wetland + wetland=reedbed`
salt marsh	`natural=wetland + wetland=saltmarsh`
sand	`natural=sand`
scrub	`natural=scrub`
sea	defined by boundary: `natural=coastline`
swamp	`natural=wetland + wetland=swamp`
tidal flat	`natural=wetland + wetland=tidalflat`
trees	`landuse=forest`, `natural=wood`
water	`landuse=reservoir`, `natural=water`, `waterway=riverbank`

to the respective information in the *map feature documentation* (OSM[21]), the majority is stipulated for being mapped as *areas* (closed ways), less as *ways* and only a very few as *points* (e. g. a single tree by using the tag `natural=tree`). Choosing a certain geometry type for representing an object also depends on the mapping scale which is influenced by the way of data collection (on-ground or remotely) and the available data (observable on-ground details or satellite imagery's accuracy, resolution and actuality). Generally, OSM provides no distinct guidance concerning scale and *minimum mapping unit.* Its open approach allows users to add features of any degree of detail as long as they're real, current, spatially referable (by a certain position or extent expressed by coordinates) and verifiable by others (OSM[22 & 23]). A blurry separation is still made by the term *micromapping,* referring to the recording of objects of very small extent, like single trees, hedges, benches, lampposts, kerbs, entrances etc.

III.2.1 Classification and codification

OSM as an open and dynamic system has ever since been prone to changes on which objects are recorded and how their digital representations are stored in the data base. Consequently, the current outline of *land cover* and *land use* related features is the result of successive practical experiences and negotiations in the community and has not been designed at one point in time. Although OSM cares about objects covering the Earth's surface, *land cover* as a directly observable bio-/physical overlay of the Earth's surface – as defined by FISHER et al. (2005) and VERHEYE (2009) – has not been used to create a central starting point for a systematic declaration of which objects to be mapped and which nomenclature to be followed. Since the start of the OSM project, *features* referring to *land cover* and *land use* have successively been introduced. They have spread over several categories represented by a series of different keys, namely `amenity`, `landuse`, `leisure`, `natural`, `surface`, `tourism`, `water` and others. They have not been systematically developed following an a priori or an a posteriori approach. The history of pages in the OSM-Wiki, e. g. for the key `landuse` (OSM[24]), reveals that some tags have been developed in advance to provide attributes that can be attached to fea-

Table 17 *Land use* objects as tagged during OSM mapping activities. Sources: OSM[25 & 26].

agriculture	
allotments	landuse=allotments
farm land	landuse=farmland, ~=farmyard
forestry	landuse=forest (high tree density; primarily for timber)
grazing	landuse=meadow (for grazing and production of hay or silage; others: surface=grass)
greenhouses	landuse=greenhouse_horticulture
meadow	landuse=meadow
orchard	landuse=orchard
plant nursery	landuse=plant_nursery
vineyard	landuse=vineyard
built environment	
brownfield	landuse=brownfield
cemetery	landuse=cemetery
college	amenity=college
commercial	landuse=commercial
construction	landuse=construction
garages	landuse=garages
greenfield	landuse=greenfield
highway	landuse=highway (proposed)
hospital	amenity=hospital
industrial	landuse=industrial
railway	landuse=railway
recreation ground	landuse=recreation_ground
residential	landuse=residential
retail	landuse=retail
school	amenity=school
university	amenity=university
built environment/ natural	
beach	leisure=beach_resort, natural=beach
cemetery	landuse=cemetery
leisure	
nature reserve	leisure=nature_reserve, boundary=protected area (under discussion)
park	leisure=park (if small sized then part of landuse=residential)
village green	landuse=village_green
military	
military	generally landuse=military; optionally military=naval_base, ~=range, ~=danger_area, ~=baracks, ~=airfield (aeroway=aerodrome + aerodrome=military)
natural	
salt pond	landuse=salt_pond
others	
aerodrome (airport)	aeroway=aerodrome (optionally: landuse=military)
airport	aeroway=aerodrome (optionally: landuse=military)
landfill	landuse=landfill
quarry	landuse=quarry
reservoir	landuse=reservoir, man_made=reservoir_covered
waste water treatment	man_made=wastewater_plant

tures. Others are defined as a reaction of encountering objects or properties not yet represented by tags. Consequently, the classification of OSM features, including those related to *land cover* and *land use,* can be described as following a successive a priori approach that results in a mostly non-hierarchical structure.

Although not precisely reflected in the tagging scheme itself, the OSM-Wiki shows that the community is aware of the concepts of *land cover* and *land use* in the sense of the aforementioned definitions. For both cases, *feature pages* exist where respective real world objects are listed, including the related tag(s) to be used for attribution (see tables 16 and 17). Despite the efforts of the community to separate the two concepts, another approach has been chosen by Ramm (2015): For the OSM data provided in *Layered GIS Format* (e. g. ESRI Shape) by Geofabrik, *land cover* and *land use* have been merged into a *landuse* layer in order to simplify the information for general-use mapping (see table 19).

However, collecting data and developing tags does not systematically follow a certain aspect, like land cover or use. Instead, thematic preferences of single members or community projects are incorporated and affect the appearance of the current OSM tagging scheme. The historically grown system is rather focused on single objects, described by tags that allow only a simply structured hierarchy of coarse categories *(keys)* and detailed information *(values)* attached to them. This is reflected by the structure of the main *map feature* list, containing the most popular tags ($\mathrm{OSM}_{21\ \&\ 27}$): The majority is used to describe *primary features,* coarsely categorized by 26 keys. Some are thematically subdivided which is nevertheless not reflected by the tagging scheme (see table 18), so that the same *key* is used for different subcategories as well.

Besides, another popular six keys are listed for attaching additional information, e. g. address data (house number, street name), sources of information (the provider of the background image or the date of the on-ground survey) and others like leaf type, road condition, wheelchair accessibility etc. Regardless of this compilation of the most commonly used tags, the data provided by the *taginfo* platform shows nearly 5 300 tags to be represented by OSM-Wiki pages in 41 languages (amongst 5 093 in English and 1 518 in German), comprising 238 distinct keys (excl. those containing a colon) and more than 2 100 values. The actual number of unique tags used in the OSM data base exceeds 23 000 (as of June 12th, 2016). Attaching numerous of these tags to one feature is the major way for combining information of different thematic domains as well as for adding further details within the same domain (e. g. `natural=wetland + wetland=swamp`). Additionally, specific details on a certain key can also be provided by using *namespaces* (see table 20): *Prefixes* and *suffixes* are attached to the key, separated by colons. This improves detail in two different ways: 1. for grouping closely related features/tags; 2. for adding information as qualifiers to mutually unrelated keys (OSM_{28}).

In summary, the tagging system for OSM features provides a very detailed, thematically unspecific and – in terms of a *classification system* – only rudimentarily structured collection of identifiers, attributes and *classifiers*. Due to differing degrees of detail and accuracy in the instructions on how to use tags for a certain feature, some only meet the definition of a *category,* others the narrower and more specific character of a *class* (see chapter II.2.1). The *map feature documentation* and the OSM-Wiki in general shows efforts to structure the ever growing amount of objects and properties to be mapped. The *keys* and *values* are highly associative and self-explaining. The mappers approach them directly and memorize them during repeated utilization. On the other hand, they are rarely encountered when editing software like JOSM or iD is used. In the latter case, the user interface provides a selection of objects and properties that are thematically structured by menus and dialogues. This causes a disconnection of object and tag on the semantic level. Consequently, during the development of an

OSM LCS it has to be considered how to appropriately integrate a new *classification system:* By using *feature pages* and *user interfaces,* or by directly representing the classification structure on the level of self-explaining *keys* and *values.*

Table 18 General structure of *primary features* in the OSM tagging scheme (excerpt). Source: OSM_{21}.

aerialway
- equipments (aerialway=pylon, ~=station)
- types (e.g. aerialway=cable_car, ~=chair_lift)
- others (aerialway=canopy, ~=goods)

aeroway (e.g. aeroway=aerodrome, ~=gate, ~=runway)

amenity
- education (e.g. amenity=college, ~=kindergarten)
- entertainment, arts & culture (e.g. amenity=cinema)
- financial (e.g. amenity=atm, ~=bureau_de_change)
- healthcare (e.g. amenity=clinic, ~=dentist)
- sustenance (e.g. amenity=bar, ~=fast_food, ~=pub)
- transportation (e.g. amenity=bus_station, ~=fuel)
- others (e.g. amenity=bench, ~=courthouse)

barrier
- access control (e.g. barrier=block, ~=bollard)
- linear (e.g. barrier=ditch, ~=fence, ~=kerb)

boundary (e.g. boundary=administrative)

building
- accommodation (e.g. building=apartments, ~=farm)
- civic/amenity (e.g. building=cathedral, ~=shrine)
- commercial (building=commercial, ~=industrial)
- others (e.g. building=barn, ~=bridge, ~=bunker)

craft (e.g. craft=blacksmith, ~=brewery, ~=sawmill)

emergency
- firefighters (e.g. emergency=fire_hydrant)
- lifeguards (e.g. emergency=lifeguard_tower)
- medical rescue (e.g. emergency=ambulance_station)
- others (e.g. emergency=access_point, ~=siren)

geological (e.g. geological=moraine, ~=outcrop)

highway
- conditional (e.g. cycleway=track, busway=lane)
- lifecycle (highway=proposed, ~=construction)
- link roads (e.g. highway=trunk_link)
- paths (e.g. highway=footway, ~=steps)
- roads (e.g. highway=motorway, ~=residential)
- special road types (e.g. highway=pedestrian)
- other features (e.g. highway=bus_stop, ~=crossing)

historic (e.g. historic=castle, ~=memorial, ~=ruins)

landuse (e.g. landuse=forest, ~=industrial, ~=port)

leisure (e.g. leisure=beach_resort, ~=dog_park)

man_made (e.g. man_made=bunker_silo, ~=dyke)

military (e.g. military=airfield, ~=danger_area)

natural
- landform (e.g. natural=cliff, ~=peak, ~=volcano)
- vegetation and surface (e.g. natural=wood, ~=scrub)
- water (e.g. natural=bay, ~=glacier, ~=wetland)

office (e.g. office=accountant, ~=association)

places
- administratively declared (e.g. place=municipality)
- populated, urban (e.g. place=city, ~=suburb)
- populated, urban/rural (e.g. place=town, ~=village)
- others (e.g. place=continent, ~=island, ~=islet)

power (e.g. power=plant, ~=cable, ~=compensator)

public transport (e.g. public_transport=platform)

railway
- tracks (e.g. railway=funicular, ~=light_rail)
- others (e.g. railway=buffer_stop, ~=crossing)

route (e.g. route=bicycle, ~=bus, ~=hiking)

shop
- art, music, hobbies (e.g. shop=music, ~=trophy)
- clothing, shoes, accessories (e.g. shop=boutique)
- discount, charity (e.g. shop=second_hand)
- DIY, building materials (e.g. shop=florist, ~=paint)
- electronics (e.g. shop=computer, ~=radiotechnics)
- food, beverages (e.g. shop=bakery, ~=convenience)
- furniture, interior (e.g. shop=antiques, ~=carpet)
- general/department store (e.g. shop=mall, ~=kiosk)
- health, beauty (e.g. shop=cosmetics, ~=optician)
- outdoor, sports, vehicles (e.g. shop=car, ~=hunting)
- stationary, gifts (e.g. shop=books, ~=newsagent)
- others (e.g. shop=copyshop, ~=dry_cleaning)

sport (e.g. sport=archery, ~=athletics, ~=boules)

tourism (e.g. tourism=attraction, ~=camp_site)

waterway
- barriers (e.g. waterway=dam, ~=waterfall, ~=weir)
- facilities (waterway=dock, ~=boatyard)
- man made (e.g. waterway=canal, ~=ditch)
- natural (e.g. waterway=river, ~=stream)
- other features (e.g. waterway=turning_point)

Table 19 *Land cover* and *land use* as aggregated by RAMM (2015) for *Geofabrik*.

class	description	tags
allotments	area with small private gardens	`landuse=allotments`
cemetery	cemetery or graveyard	`landuse=cemetery`
commercial	commercial area	`landuse=commercial`
farm	agricultural land (incl. farms and areas where crops are grown)	`landuse=farm, ~=farmland, ~=farmyard`
forest	forst or woodland	`landuse=forest, natural=wood`
grass	area where grass grows	`landuse=grass`
heath	heath areas	`natural=heath`
industrial	industrial area	`landuse=industrial`
meadow	meadow, possibly used for grazing cattle	`landuse=meadow`
military	military land use, usually no access for civilians	`landuse=military`
national_park	national park	`boundary=national_park`
nature_reserve	nature reserve	`leisure=nature_reserve`
orchard	area used for growing fruit-bearing trees	`landuse=orchard`
park	park	`leisure=park`
quarry	quarry	`landuse=quarry`
recreation_ground	open green space for general recreation	`leisure=recreation_ground, landuse=~`
residential	residential area	`landuse=residential`
retail	area mainly used by shops	`landuse=retail`
scrub	area where scrub grows	`landuse=scrub`
vineyard	area used for growing grapes	`landuse=vineyard`

Table 20 Types and examples of *namespaces* for providing features with additional details.

<table>
<tr><th colspan="3">for grouping</th><th colspan="2">as qualifier</th></tr>
<tr><th colspan="3">specifying power generators</th><th>life cycle prefix</th><th>language suffix</th></tr>
<tr><td colspan="3">grouping of additional information concerning generators (excerpt):</td><td>attributing features according to the stage of their life cycle (excerpt)</td><td>specifying names of features in different languages</td></tr>
<tr><th><code>:source</code></th><th><code>:method</code></th><th><code>:type</code></th><td rowspan="10"><code>planned:</code>
• proposed, approved and funded
<code>construction:</code>
• under construction
<code>disused:</code>
• currently not used
• in reasonable state of repair
<code>abandoned:</code>
• visible, fallen into serious disrepair
• restoration only with expensive effort
<code>demolished:</code>
• intentionally destroyed</td><td rowspan="10">English: <code>:en</code>
German: <code>:de</code>
Chinese: <code>:zh</code></td></tr>
<tr><td rowspan="6"><code>nuclear</code></td><td rowspan="3"><code>fission</code></td><td><code>PWR</code></td></tr>
<tr><td><code>GCR</code></td></tr>
<tr><td><code>FBR</code></td></tr>
<tr><td rowspan="3"><code>fusion</code></td><td><code>tokamak</code></td></tr>
<tr><td><code>ICF</code></td></tr>
<tr><td><code>cold-fusion</code></td></tr>
<tr><td rowspan="2"><code>tidal</code></td><td><code>barrage</code></td><td><code>kaplan_turbine</code></td></tr>
<tr><td><code>stream</code></td><td><code>vertical_axis</code></td></tr>
<tr><td colspan="3"></td></tr>
<tr><td colspan="3"><u>e. g.</u> a solar power tower (produces electricity by concentrating sun beams toward a steam generator):
<code>power=generator</code>
<code>generator:source=solar</code>
<code>generator:method=thermal</code>
<code>generator:type=steam_turbine</code>
<code>generator:output=electricity</code></td><td><u>examples</u>

closed pub: <code>disused:amenity=pub</code>

road damaged by land slide years ago, now only passable on foot or bicycle:
<code>abandoned:highway=unclassified</code>
<code>highway=path</code></td><td><u>e. g.</u> Beijing (capital of P. R. of China):
<code>place=city</code>
<code>capital=yes</code>
<code>name=北京市</code>
<code>name:en=Beijing</code>
<code>name:fr=Pékin</code>
<code>name:zh=北京市</code>
<code>name:zh_pinyin=Běijīng shì</code></td></tr>
<tr><td colspan="3">OSM[29]</td><td>OSM[30 & 31]</td><td>OSM[32]</td></tr>
</table>

III.2.2 Community perception of tagging land cover/use

III.2.2.1 Discussing and changing the tagging scheme

Questions concerning perception and use of tags are discussed in various communication channels of the community. When related to a specific mapping situation *OSM Forum* (OSM_{33}), *OSM Help* (OSM_{34}) and *OSM Mailing Lists* (OSM_{35}) are frequently consulted. Even very detailed and well documented tags are a simplification of the reality and do not necessarily fit to every special situation. In order to understand the tags meaning and scope of application the users contact the community for seeking exchange and advice. As a consequence, such a process of seeking solutions for special mapping cases might end up in realizing that none of the existing tags meets the requirements and a new one has to be introduced.

When starting to use a new tag it is recommended to describe its intended utilization in the OSM-Wiki. Otherwise, users might choose or invent another tag although encountering a similar mapping problem, or the undocumented tag might be used cases it was not invented for. In the long run, inconsistencies in the data base would be created ($OSM_{36\ \&\ 37}$). Only the publication of a new tag allow the community to discuss and improve it.

Especially when introducing tags that are of general interest or difficult to develop, it is recommended to consult the community by creating a proposal which can be discussed and finally democratically decided on. *Proposal pages* in the OSM-Wiki act as the basis for this process, They present necessary information and providing references on the current state of discussion and development. They're set up by following a *proposal template* which contains chapters that are designed for answering questions about the need for introducing a new tag. Similarly, circumstances under which it should be applied, and the criteria of the object are discribed. Furthermore, it has to be outlined for what type of geometry the tag is supposed to be used. Affected tags and wiki pages, as well as suggestions for the visual representation by renderers can be presented as well (OSM_{38}).

Once the status is officially changed from *draft* to *proposed* and a *rfcStateDate* is set *(request for comments),* the discussion is opened on a new wiki page that solely related to this proposal. Here, recommendations are collected and discussed providing input for revising the proposal. A start and an end date mark the duration of the subsequent *voting* status (usually two weeks) during which the proposal is not altered and the community members can express consent or refusal, both preferably including a short explanatory statement. Proposals are considered to be successful when 74 % of at least 10 votes are in favour of it. However, the voting period can be prolonged and even after the status is set to *rejected* the proposal can be reworked and put back to vote. Moreover, it is important to state that even in case of a majority voting in favour of a new tag existing features carrying the old one will not get automatically updated and shall not be altered manually by using the vote as a justification. Users are neither obliged to use the new tag nor has the new tag to be immediately incorporated into rendering algorithms or editing software (OSM_{38}).

III.2.2.2 Perceiving land cover and land use

The tagging of features and properties related to land cover and land use has been introduced above in chapter III.2.1. It has been demonstrated, that related features are located in the domains of a series of different keys. According to the history records provided by the wiki software, the *feature pages* of both *land cover* and *use* ($OSM_{25 \& 26}$) date back to December 2011 and were created by the same user *PeterIto,* alias PETER MILLER, CEO and co-founder of *ITO,* a UK-based company providing spatial transportation and visualisation services based on OSM data. At that time it was already stated that "[…] there have been discussions that land use (the actual use of the land) and land cover (what is actually covering the land) have been mixed between this and other keys." (OSM_{39})

By this, he referred to an even older proposal and its related discussion regarding a new key `landcover`, initiated and officially drafted in November 2010 by user *dieterdreist,* alias MARTIN KOPPENHOEFER (OSM_{40}). He similarly argued that the "feature landcover as such is not new to OSM, but it is currently messed into different other keys (e. g. `landuse`, `natural`, `surface`) hence creating confusion and inconsistencies." (OSM_{41}) First efforts have been made in

Table 21 List of sources used for reviewing the discussion on *land cover/use* in the OSM community. Search terms used: "landcover", "land cover", "landuse" and "land use".

proposals and respective discussions		
draft landcover	03. 04. 2014 – 22. 05. 2014	OSM_{43}
draft landcover woodland	12. 05. 2014 – 22. 08. 2014	OSM_{44}
Proposed features/landcover	16. 11. 2010 – 05. 01. 2016	OSM_{40}
feature pages		
Environment	22. 03. 2008 – 04. 05. 2015	OSM_{45}
Landcover	28. 12. 2011 – 10. 12. 2015	OSM_{26}
Landuse	19. 12. 2011 – 24. 05. 2016	OSM_{25}
Natural	19. 12. 2011 – 24. 05. 2016	OSM_{46}
OpenStreetMap Forum		
Befestigte Flächen mappen	18. 12. 2015 – 21. 01. 2016	OSM_{47}
Das Märchen vom `landuse`	24. 09. 2012	OSM_{48}
`Landuse`	26. 05. 2016	OSM_{49}
`landuse=forest` or `natural=wood`?	02. 12. 2010 – 04. 12. 2010	OSM_{50}
Österreichs Almböden	16. 12. 2015 – 10. 04. 2016	OSM_{51}
OpenStreetMap Help		
How to tag green (vegetated) areas within urban areas?	01. 11. 2011 – 17. 02. 2012	OSM_{52}
Mapping landcover in a park: `landuse` vs. `landcover` vs. `natural`	22. 05. 2016	OSM_{53}
Overlapping land use, cf `residential` and `forest`	21. 11. 2012 – 15. 01. 2016	OSM_{54}
Poplar plantation: how to map?	15. 09. – 09. 10. 2015	OSM_{55}
Should pastures be marked as `farm` or `grass`?	20. 07. 2010 – 22. 02. 2011	OSM_{56}
When should we use `landuse=forest` rather than `natural=wood`?	19. 07. 2010 – 20. 06. 2014	OSM_{57}
OSM mailing list *tagging*		
Defining tag '`natural=wood`'	01. 02. 2016	OSM_{58}
`Landuse=forestRY`?	27./28. 12. 2015	OSM_{59}
`landcover=trees` definition	02. – 20. 08. 2015	OSM_{60}

Table 22 Approach for classifying *land cover* in OSM as proposed by user *rudolf* (OSM[43]).
Tags marked with * are approved, already in use and represented in the OSM-Wiki.

`landcover=water`	open water or permanent ice/snow cover
`water=lake*`	water surrounded by land
`water=river*`	wide river
`water=reservoir*`	artificial lake, used to store water
`water=canal*`	man-made channel for water
`water=ice`	year-long surface cover of ice and/or snow, i. e. glaciers, mountain tops, polar ice
`landcover=wetland`	soil periodically saturated with or covered with water
`wetland=swamp*`	wooded wetlands often flooded for a part of the year; vegetation dominated by trees and tall shrubs
`wetland=marsh*`	wetlands periodically or permanently flooded with water; vegetation typically non-woody, e. g. cattails, rushes, reeds, grasses and sedges
`wetland=bog*`	peat-filled depressions receiving water and nutrients from rainfall; plants partially decomposed; frequently covered in shrubs rooted in the sphagnum moss and peat
`wetland=fen*`	located in areas of groundwater discharge; vegetation typically includes sedges and mosses, along with some grasses, reeds, low shrubs
`landcover=grassland`	dominated by upland grasses
`grassland=natural`	natural/semi-natural grassland
`grassland=savanna`	natural grassland with sparse trees or scrubs
`grassland=tundra`	alpine tundra: habitat above the treeline, characterised by open, low growing vegetation
`grassland=urban`	urban/recreational grassland
`grassland=meadow`	grassland, used for hay or pasture
`landcover=plantation`	used for growing crops or fruits
`plantation=cropland`	used for growing crops
`plantation=orchard`	plantation of fruit trees or shrubs
`plantation=vineyard`	plantation of grape-bearing vines
`plantation=aquatic`	cultivated aquatic or regularly flooded areas
`landcover=woodland`	tree cover
`leaf_type=broadleaved*`	broadleaved vegetation
`leaf_type=needleleaved*`	needle-leaved vegetation
`leaf_type=mixed*`	broadleaved and needle-leaved vegetation
`landcover=shrubland`	woody vegetation with aerial stems, generally less than 6 meters tall, without interlocking
`shrubland=scrub`	uncultivated land covered with bushes or stunted trees
`shrubland=heath`	vegetation dominated by dwarf shrubs of the heather family
`shrubland=tundra`	Arctic tundra: treeless habitat, characterised by open, low growing vegetation
`landcover=bareland`	bare (barren) land, with little or no "green" vegetation
`bareland=sand`	covered with sand
`bareland=bare_soil`	covered earthen material
`bareland=hardpan`	indurated due to chemical or physical processes, e. g. dry lake, salt pan
`bareland=gravel`	covered with loose rock fragments, e. g. scree, gravel beach
`bareland=stones`	covered with stones, e.| g. stone desert
`bareland=bare_rock`	covered with bare rock
`landcover=artificial`	artificial surfaces and associated areas
`artificial=road`	wide roads
`artificial=railroad`	wide railways
`artificial=pipeline`	used for communication lines/ pipelines
`artificial=urban`	urban areas
`artificial=industrial`	industrial areas
`artificial=extraction`	extraction sites

separating "the physical landcover of an area as opposed to the usage […] (landuse)" (ibid.) by proposing new tags, e. g. `landcover=trees` or `~=sand`, partially replacing existing ones, e. g. `natural=sand` and `landuse=grass`. The respective discussion on the proposal (OSM_{42}) started shortly afterwards, showing high activity in November 2010, May 2011 and finally from December 2011 to February 2012 after the creation of the above mentioned *feature pages.* Probably the most substantial contribution to the debate was presented by the user *rudolf* in April/May 2014 (OSM_{43}). His drafted proposal contains approaches on defining *land cover, land use* and surface and presents a two-level classification and tagging scheme for *land cover* features (see table 22), which he states to be inspired by the *classification system* of the USGS, namely the *National Land Cover Database 1992 (NLCD92),* and the FAO's LCCS.

Still, the proposal has never reached the *proposed* state and its wiki page has remained unchanged since May 2014 (OSM_{43}). Although it was mentioned on the discussion page of the 2010's proposal, it didn't revive the debate that had basically come to an end in autumn 2013. The arguments exchanged until then are nevertheless important for gaining an insight on how the community approaches the mapping of *land cover/use* and how ideas on a new tagging scheme are perceived. By additionally referring to selected issues discussed in OSM-Help, OSM-Forum (English and German) and the mailing list *tagging* an attempt is made to outline the current state of discussion (focussing on the years 2015 and 2016; see table 21). Arguments, questions and suggestions can roughly be distinguished into least two different domains, a *semantical* and a *practical,* pointing to conceptual and operational challenges, respectively.

In general, discussions are initiated by specific questions concerning problems encountered during mapping activities (see subject headings in table 21). Some issues are repeatedly discussed, especially concerning the application of the tags `landuse=grass`, `landuse=meadow`, `natural=wood` and `landuse=forest`. The exchange of experiences and suggestions occasionally turns into longer debates about the reasonability of separating land cover and land use and on introducing corresponding new keys and tags. Pro and contra camps often refer to the 2010's proposal and the related discussions in order to emphasize their positions. Despite the numerous occasions, neither have new tags been developed and effectively proposed, nor have the reactions on repeated questions lead to substantial amendments of the existing tags.

Semantical issues

The debate on the semantical level is often focused on explaining the difference between *land cover* and *use*. The definitions developed and advocated are essentially similar to those presented by COMBER, DI GREGORIO, FISHER and other authors mentioned above (see chapter II.2.1); in short, describing "What is there?", opposed to "For what purpose is it used?" By clearer separating these aspects, a more accurate distinction of features is expected. It is believe that possibilities intentional combinations are created and OSM's potentials are better used, so the arguments.

The existing tags are criticized because they don't seem to be semantically clear enough. They appear to be misleading or not free from connotations. Consequently, users complain about a lacking sense for logic and intuition, for example keys that incorporate different semantic domains like landscape features, substances, activities and vegetation cover (`natural=volcano` vs. `natural=water` vs. `natural=scrub`; `landuse=grass` vs. `landuse=peat_cutting`). Other keys and tags contain or produce overlaps (`natural=sand` vs. `natural=beach`;

`natural=water + water=reservoir` vs. `landuse=reservoir`). There are also values that hardly seem to match the key they are assigned to (`landuse=grass`). This "mess" is identified as one cause for newbies repeatedly getting confused when looking for appropriate tagging solutions. Additionally, by the suggestions and comments that they receive on their questions, they are confronted with varying individual interpretations, definitions and disagreements between experienced users, e. g. when to use `natural=wood` or `landuse=forest`, respectively. This is seen to provoke doubtful and incomprehensible choices ("rolling the dice") and is blamed to be an important reason of inconsistent OSM data.

On the other side, for defending the existing tags it is argued that the wiki-pages provide sufficient information for accurately tagging a big variety of features. In addition, the concepts of land cover and land use are not convincingly defined by the supporters of a separation. Questions remain on how to clearly distinguish *land cover* from *surface* and whether *land cover* is characterized only by the objects seen from above (e. g. tree cover under which shrubs and grass my coexist as well) or by those being predominant (like herbaceous plants and mushrooms on a grass dominated area). It is moreover stated that a tag `landcover` will necessarily be misleading because basically every object designated for mapping can somehow be regarded as *land cover.*

<u>Practical issues</u>

As a consequence of formulating clearer definitions and more intuitive keys/values confusions are expected to be reduced, especially for newbies. The supporters are convinced that mapping of objects as there're seen will become easier because further information, e. g. on the way they're used, can firstly remain unknown. Compared to the current way of tagging it is regarded as less problematic to have the opportunity to add and change according information afterwards, e. g. several types of *land use* on an area characterized by one homogeneous *land cover* type or vice versa. The possibility of attaching two separate information parallelly is considered to increase consistency and gain more precise information for the users of the data base.

Contrarily, this kind of "double tagging" is one aspect meeting with criticism. Not only is it seen to create the risk of causing redundancies but it is also believed as not getting accepted by the community: It is argued that most of the users prefer one compound tag instead using several to describe an object. In addition, although users are urged not to "map for the renderer" (OSM_{37}) it is argued, that changes still want to be seen in the map causing users to avoid tags not supported by the renderer. There is still a generally strong conviction, that the existing tags can sufficiently cover the demands of the users and that they offer enough possibilities for dealing with special cases – which would also occur with new tags, anyway. Finally, even for new tags absolute precision in defining them is expected to be out of reach. Uncertainties will remain and continue to be met with arbitrary decisions based on individual interpretations. If incorrect, these cases can be easily reviewed and corrected by the community.

Other arguments emphasize on technical and organizational aspects: Beside the statement that the pursued degree of ontological accuracy has not been established in other parts of the OSM project, it is argued that a new tagging system strictly based on separating *land cover* and *use* is more suitable for the creation of specialized maps but unlikely to succeed in a voluntary project. Another aspect mentioned is the tremendous effort to be expected in terms of adaption: Editing software, renderers and analysing algorithms have to be substantially adjusted, not to mention the community that needs to be convinced. Especially since there are no binding instructions on what tags to use but instead freedom is declared for applying every imaginable tag (OSM37), passionate community members might finally stick to the tags they've used for a long time. Basically, the more profound changes of the tagging system are, so argued, the more difficult it gets to assure backward compatibility, e. g. for conducting comparisons on specific aspects over longer time.

Conclusion

At the bottom line of the positions and discussions outlined above it can be observed that the debate is carried on at two different areas: Most of the arguments in favour of separating *land cover/use* and introducing new keys and tags emphasize semantical aspects. However, the counter-arguments highlight the problems and risks to be expected during implementation and point out their disagreement with insufficiently defined terms. Whereas the pro-camp reacts by anxiously trying to refine and to repeatedly explain the differences and advantages of the concepts, also by presenting practical examples, they have not yet adequately addressed the objections expressed by the contra camp, namely issues concerning adaption, inconsistency and compatibility.

III.2.3 Comparison with other classification systems

III.2.3.1 General remarks

Based on the characteristics elaborated above, comparing OpenStreetMap to other classification systems reveals significant differences. As already mentioned, some *classes* in OSM had been developed in advance and have been adjusted later on, others were successively added afterwards. This process of adjusting and changing *classes* in terms of their definitions, explanations and numbers is dynamic and continues incessantly. In contrast, the developments of the above mentioned systems had a distinct starting point, accompanied by negotations about the objectives of the system. Once they were put into practice only few *classes* were missing. Together with other improvements they were integrated later in periodical revisions, officially published after years or sometime decades. This assures continuity and stability for a long phase and provides consistency to the application of the systems and the utilization of the resulting data. Though suffering from more distinct heterogeneity, but the OpenStreetMap community is able to quickly integrate real world changes and adjust unsuitable class definitions, unfortunately without a consequent and consistent update scenario for the existing data. Simultaneously changing tags of numerous features is basically possible but it is stressed to be carefully discuss with the community since reversions are difficult and may nevertheless cause lasting damages to the data base (OSM61).

The general structure is another aspect of difference: Except the IGBP classification, the above mentioned *classification systems* for *land cover/use* are hierarchically organized, either for provided *classes* at different levels of details (e. g. LUCAS; see chapter III.1.3) or as a decision tree for a guided allocation to a certain class (the dichotomous phase of the FAO LCCS; see chapter III.1.1). Contrarily, the basic classification structure of OpenStreetMap consists of only two levels: broader categories, represented by keys, and more specific sub-categories, represented by values. Further structural arrangements regarding thematic top- or sub-domains are partially provided by feature pages in the OSM-Wiki but are not necessarily reflected on the level of tag names. Other details are not aggregated to new *classes* but solely available as attributive tags to be attached to the *features.* Consequently, compared to the above presented systems developed by FAO etc. OSM is apparently missing a core structure and key principles that allow a monolithic, hierarchical system to develop over various thematic domains and levels of aggregation.

Consequently, degrees of consistency are different. Although not being excluded from criticism (e. g. *land use* and *cover* in the FAO LCCS), the professional *classification systems* presented above can be characterised as being rather consistent. This is assured by developing and documenting classificatory principles, naming rules and elaborated explanations. Since they are results of considering and compromising different objectives and restrictions sporadic contradictions and inconsistencies can nevertheless occur. Contrarily, OSM is characterised by an open approach, also for recording *land cover/use.* Several related concepts (functions, vegetation cover, landscape features etc.) and thematic domains (agriculture, ecology, economy etc.) are mixed in different keys. The scopes of the used tags are not precisely defined and thus give room for interpretation and misunderstandings. Long standing problems repeatedly cause confusions and discussions among newbies and experienced mappers.

Referring to the key principles elaborated in chapter II.2.3 the aspect of being independent from scale is basically fulfilled by the OpenStreetMap approach. Of course, not all tags can be assigned to any object recorded at any scale. But no matter for which extent or degree of detail an appropriate tag is likely to be found representing the observed objects or properties. However, because tags often represent complex concepts some of their characteristics can be assessed at a certain scale while others remain unknown – affecting the validity of the data base in case of assigning those tags. In comparison, the other systems presented above are scale-independent as well. However, detailedly described identification rules are valid for on-ground activities or for mapping high-resolution imagery (see table 6). They are separated from the spectral based classification routines applied for automatically allocating pixels to *classes.* The former ones at a large scale are the basis of calculating and verifying spectral signatures that are used for the classification on a smaller scales. Hence, the different methods enable the entire system to assure classification on all scales: On large scales by directly using the detailed rules, and on small scales by using multispectral signatures. Since the presented *classification systems* of economic activities are not necessarily related to spatial units, independence from scale can hardly be evaluated. However, the hierarchical structure allows bigger areas to be delineated and allocated to a more aggregated class, e. g. “industrial”, while areas of single companies could match more detailed *classes* at levels below, e. g. “metallurgy” or “petrochemistry”.

III.2.3.2 Essential conceptual differences

To understand the reasons for the difference regarding the basic structure is important for the aim of this study. As elaborated in chapters II.2.1 and II.2.3, *categories* are based on rather subjective conceptions that become precisely defined and described in order to develop distinct *classes.* Which criteria and thresholds to use and how to incorporate them, e. g. as classifiers as in LCCS (see table 3) or as part of the definitions as in IGBP (see table 6), are constitutive questions during the set up of a *classification system.* According to the definitions and criteria objects are identified and allocated to a *class,* hence provided with a label that reflects their equivalence to a certain concept, e. g. “woodland” for a *land object* (an area) characterized by a tree dominated canopy cover (as an exemplary definition). However, properties in common with other *classes* might only be implicit or reflected by different levels of aggregation but are not necessarily reflected by the class name. This is not problematic as long as a class name occurs only once in the entire system. Exemplarily, the class name “machinery and equipment” in the LUCAS classification system does neither contain any information about properties in common with objects of other classes nor does it reflect its position in the systems hierarchy. Anyway, because it is unique regarding its name and also regarding the concepts, definitions and criteria it represents it can be unequivocally contextualized as being part of the aggregation “industry and manufacturing” in the superordinate *class* “secondary sector” of the LUCAS land use classification system (see table 11).

All these arrangements can be summarized as being “conceptual”. They decide which concepts are considered by the system, how they are defined, by which criteria objects are identified and delineated, to which *classes* objects are finally allocated and whether they are contextualized according to different degrees of detail or differently weighted commonalities. Subsequently, the results of the classification have to be “technically” managed. This implies the necessity of assigning the classification result to the abstract representation of the object, namely the *feature* in form of a pixel or a certain type of geometry.

Most of the *classification systems* introduced above use combinations of letters and numbers for codification. Usually, they are designed to potentially reflect combinations of classifiers, e. g. in LCCS, or the position of the *class* within the entire system, e. g. “U226” (“machinery and equipment”) in the LUCAS system indicates this *class* to be located at the third level of the “Land Use Classification System” (“U”) as a part of the superordinate classes “industry & manufacturing” (“U220”) and “secondary sector” (“U200”). Although related to specific *classes* they do only refer to their definitional and descriptive content. On the etymological level they are actually independent. They consequently have the advantage of being able to represent different class names as long as the respective *classes* represent the identical concepts, e. g. in the case of translating class names into different languages or adjusting class names without changing definitions or criteria of the respective *class*. In case of LCCS based legends the codes identify the *classes/classifiers* used for aggregation. In summary, codification is not a constitutive part of a *classification system* but an element subsequently added for technically managing classification results.

In this respect, OpenStreetMap is fundamentally different since there is no clear separation of the “conceptual” and the “technical” aspect; of classification and codification, respectively. When searching for where to allocate an observed object, e. g. a group of farmhouses and its associated area, OpenStreetMap will not only lead to a class and its superordinate context: “farmyard” and “landuse”, respectively. These two class names – as value and key – are instantly used for codification, in form of a tag to be attached to the *feature;* `landuse=farmyard` for the given example. Tags can therefore be interpreted as codes that are not etymologically independent from the related class names. In fact, they are an essential component of the OSM classification system and directly reflect the concepts they refer to. Consequently, the development of class names is conversely influenced by the limitations adherent to the tag format.

The origin of the way tags are developed, used and integrated is documented on the history pages of the OSM-Wiki. The wiki-page *Elements* states at the day of its creation in June 2007 that keys and values are used for assigning properties to the basic elements of the map (back then nodes, segments, ways and areas; OSM_{62}). Consequently, tags were apparently not created to technically represent a *class* like a code in a *classification system* but rather for dir-

Table 23 Comparison of codification and OSM-tagging, exemplarily using a land use class of the LUCAS classification system. The presented tags are fictional and only meant to demonstrate the possibilities and challenges of the OSM nomenclature to reproduce concepts and hierarchical context.

class names	code	tag	
		multi-tagging	**namespaces**
land use	U	–	–
secondary sector	U200	`landuse=secondary`	`landuse:econ_sector=secondary`
industry & manufacturing	U220	`landuse=industry` `econ_level=secondary`	`landuse:secondary=industry`

machinery & equipment	U226	`landuse=industry`
		`econ_level=secondary`
		`industry=machinery`
			`landuse:secondary:industry=machinery`

ectly attaching properties to features. This basically corresponds to the concept of classifiers, as used by the FAO LCCS. But beside answering the question of *how* an object is characterised, e. g. by using `surface=asphalt` and `width=10`, tags are nonetheless used to answer the question of *what* an object actually is, e. g. by adding the tag `highway=trunk`.

Since then, the variety of objects and properties recorded has increased which means that numerous concepts and specifications had to be integrated into the tagging system of OSM. In this respect, the representation of further object related details are imaginably easy to build because they contain relatively simple concepts, e. g. the accessibility of an amenity (`wheelchair=*`), the height of a building (`building:levels=*`) or the emptying schedule of post boxes (`collection_time=*`). Comparably, developing tags for representing new objects is much more challenging because they represent complex concepts

Despite the amount and complexity of objects recorded for the OSM project, features are still preferably attributed by one tag in order to answer the *what*-question (OSM_{63}). Changing definitions or criteria for using a certain tag, efforts of re-organising and re-aggregating information differently in order to improve the usability of the large amount of available information – since all those actions may result in changes of sub- or superordinate class names (keys and values), the inevitable adjustment of tags will contrarily prevent substantial changes of the entire system because of three reasons:

1. The basic challenge to cope with the restrictive tag design, basically allowing only two components (key and value). Though namespaces can be alternatively used to increase the possibilities of available expressions, but they do finally not allow an uncomplicated handling of complex multi-level aggregations of objects or properties. In contrast, incrementally adding tags representing different levels of detail might become unhandy during mapping activities in some cases (see table 23).

2. Technical issues, e. g. adjusting editing software and rendering algorithms. Long colon-separated tags might be fundamentally difficult to deal with. Analysing three or four tags before being able to define how to visualize a single object might have an impact on computation requirements.

3. Habitualities of the community members. Because of their self-explaining character, tags facilitate a very direct and precise communication between community members when discussing how to label an observed object. Abstract codes would offer the same but without directly reflecting a distinct meaning (semantical separation). The uncomplicated access to the information of a term directly used on the level of the data base is very attractive for mappers and users. This preferably requires tags to be as succinct as possible, implying the risk of favouring one tag that represents a complex concept over a combination of different tags that provide clear information on simpler concepts.

IV OSM Land Classification System

The main objective of the study is to re-arrange the systematics of collecting and managing land cover/use information in the OpenStreetMap project. The current chapter will define the framing conditions and develop a set of specific requirements by considering the fundamentals, examples and findings compiled in the previous chapters. Subsequently, the chapter presents a preliminary version of a new classification system.

IV.1 Requirements

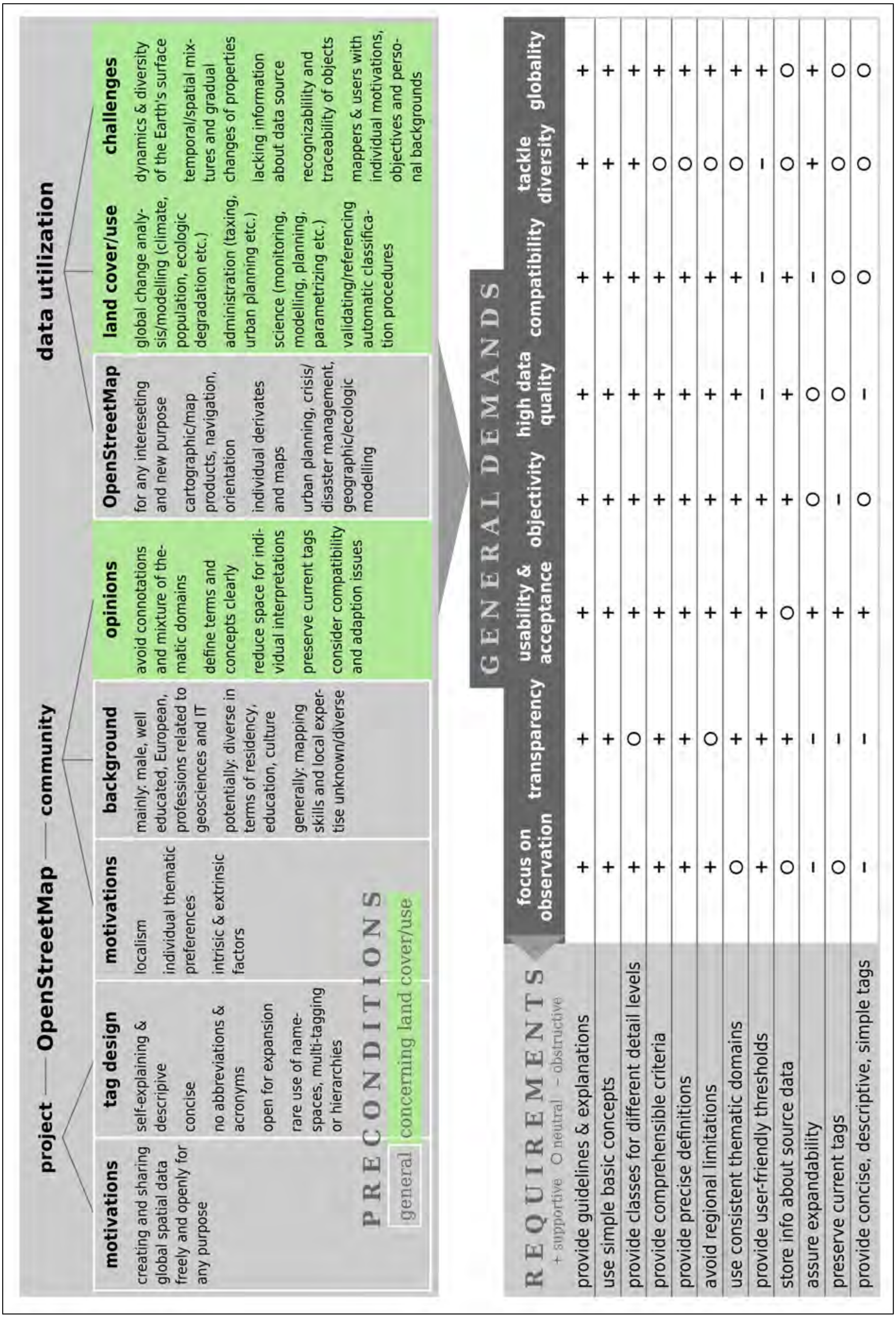
Beside the key principles crucial for *classification system* in general (see chapter II.2.3), guidelines uniquely valid for OpenStreetMap need to be developed. Because inherent purposes and objectives have been pointed out as a source of inconsistencies between different systems they have to be even more insistently considered and expressed in the current case. Consequently, **preconditions** relevant for a new OSM LCS will be presented by summarizing the findings elaborated in the previous chapters. The responses to each of these aspects are used as a basis for formulating **general demands** that are subsequently addressed by a set of **practical requirements.** Figure 2 illustrates this process and how the requirements contribute to meet the demands. This is meant to increase transparency on how the given conditions are actually used as formative factors in designing the *classification system* later on.

IV.1.1 Preconditions

IV.1.1.1 Considering the OSM project

The general motivation of the OpenStreetMap project spreads, of course, beyond the scope of mapping *land use* and *land cover* (see chapter II.1.1). Whatever georeferencable information is collected, it shall finally be provided freely and openly in order to be used for any purpose independent from corporate and governmental limitations. Reducing practical obstacles in collecting and using geospatial information will support this goal.

Consequently, mapping land cover/use objects and extracting respective information from the data base should neither be affected by location or scope of the region of interest nor by the sources and methods of data collection. To serve as an attractive and reliable alternative or complement for other professional datasets, informational and structural foundations need to be transparently laid out in order to gain trust in the objectivity of the data and to increase external compatibility.



	focus on observation	transparency	usability & acceptance	objectivity	high data quality	compatibility	tackle diversity	globality
provide guidelines & explanations	+	+	+	+	+	+	+	+
use simple basic concepts	+	+	+	+	+	+	+	+
provide classes for different detail levels	+	O	+	+	+	+	+	+
provide comprehensible criteria	+	+	+	+	+	+	O	+
provide precise definitions	+	+	+	+	+	+	O	+
avoid regional limitations	+	O	+	+	+	+	O	+
use consistent thematic domains	O	+	+	+	+	+	O	+
provide user-friendly thresholds	+	+	+	+	–	–	–	+
store info about source data	O	+	O	+	+	+	O	O
assure expandability	–	–	+	O	O	–	+	+
preserve current tags	O	–	+	–	O	O	O	O
provide concise, descriptive, simple tags	–	–	+	O	–	O	O	O

Figure 2 Requirements for a new classification system that meet demands derived from case-specific preconditions. Source/layout: the author.

The most important component for realizing the collection, storage, extraction, arrangement and analysis of information captured in the OSM data base is the *tag.* As concluded in chapter III.2.3, the limitations of its format affects the realization of conceptual and definitional changes aiming on improving the mapping of objects and the utilization of OSM data. For the amendments in the realm of land cover/use intended by this study tags will have to inevitably be adjusted, rejected or newly invented. The objects currently described by keys like `landuse` or `natural` are based on subdividing reality into subjective *categories* rather than systematic *classes,* provoking the preferred use of single tags that represent complex but subjectively suitable concepts.

The tag design thus affects objectivity and systematic usability. Contrarily, technical adaptivity and collective acceptance is needed for a classification system to be attractive for mappers and external users. However, in searching for an easy and intuitive but parallelly objective way of tagging general recommendations for the tag design have to remain in consideration: Preserving conciseness and descriptiveness; avoiding abbreviations, acronyms, redefinitions and complex hierarchies; considering the needs and habits of the community.

IV.1.1.2 Considering the OSM community

Beside the general objectives of the entire project, individual motivations of members have to be taken into account (see chapter II.1.2). Intrinsic and extrinsic factors are diverse, but localism and individual interests are uniquely spread among users. The objectivity aspired above is limited by the geographic and thematic preferences of the users. Although profound knowledge on familiar places and certain domains is a valuable source of information for the OSM project, it subjectively influences location and type of objects chosen for mapping.

In order to leverage this source and the motivations of the community members, the classification system should avoid any regional limitations in order to encourage the global mapping of land cover/use. The system should cover local peculiarities or create the opportunity for adding missing classes. Additionally, the system should provide the possibility of adding objects of coarser spatial and informational detail for enabling the correct – but less detailed – mapping of land cover/use data in remote areas world wide. Mapping on a simple level shall encourage members with other interests than land cover/use to also contribute in this domain.

The motivations of members, namely localism and personal thematic interests, are related to the general diversity of backgrounds of OSM community members. Conclusions on data users can be only roughly drawn from the purposes the data is used for (see chapters II.1.4 and II.2.4) because OSM data is freely accessible without the need for registration or for giving any feedback on how the data has been used. In contrast, more detailed findings about the group of registered members were presented in chapter II.1.2, revealing distinct majorities in several aspects: More than 90 % men, up to 70 % at least graduated from college, nearly entirely based in Europe (ca. 75 %) and Northern America & Asia (ca. 25 %). Additionally, professional backgrounds are often located in geosciences, spatial planning or information technologies. However, local and thematic mapping expertise actually remains questionable compared to professional experts engaged for collecting reference and validation data.

Diverse or unknown residences, interests and skills require the classification system to provide clear definitions and explanations in order to synchronize the conceptual diversity among mappers and reduce the scope of subjective interpretation.

Except in case of local peculiarities land cover/use classes should be free from cultural, regional and thematic bias to be applicable by mappers worldwide on objects worldwide. Underlying concepts have to be plausible and reproducible for all mappers, irregardless their professional, educational or cultural background. This implies emphasizing on descriptive observation by following a *WYSIWYM approach: What you see is what you map.*

This actually meets the demands expressed in the community discussions (see chapter III.2.2). Key terms and concepts were asked to be defined transparently, objectively and without connotations and mixtures of different thematic domains. This was meant to foster a logical and intuitive tagging process and to avoid arbitrary classification decisions ("rolling the dice"). However, concerns were expressed in terms of compromising usability and acceptance once tags are changed. Instead, they are supposed to stay simple and self-explaining.

Considering established land cover/use tags, adjusting current and implementing new tags has to be deliberate, transparent and consistent. Backward compatibility and adaption issues have to be taken into account as well, e. g. by compiling transfer catalogues.

IV.1.1.3 Considering data utilization

Apart from mainly creating customized base maps and navigation products OSM data is meanwhile used for administrative purposes like urban planning, disaster management and scientific modelling (see chapter II.1.4). This basically coincides with the utilization scenarios of land cover/use data: The need for respective information is similarly driven in administrative and scientific contexts where monitoring, analysing, modelling and planning of socio-economical and geo-ecological phenomena wants to be facilitated. Actual and reliable data for referencing and parameterizing models is very important. Especially for tasks concerning global processes (climate change, population growth and migration, urban development, degradation of natural resources etc.) data collected on a global scope is needed (see chapters II.2.2 and II.2.4).

The seriousness of the utilization scenarios implies a demand for high data quality. Especially users in professional, administrative or scientific contexts do not necessarily collect data for OSM. Hence, the process of mapping and classifying *land cover/use* has to be transparently documented by using consistent concepts, clear definitions and comprehensible criteria. Because of the diversity of current applications and in order to offer and sustain rich utilization potentials the classification of *land cover/use* information should be as objective as possible and therefore emphasize on observation rather then interpretation. Similar to the diversity within the OSM community the diversity of potential utilization scenarios and users, compatibility and usability has to be increased by incorporating simple basic concepts instead of complex ones that originate in specific technical or scientific domains.

Although data for OpenStreetMap is manually recorded on-ground or derived from mainly true-colour imaginary (see chapter II.1.3) some of the technical challenges faced during the application of automatically classified land cover/use data (see table 2 in chapter II.2.4) have to be considered as well: Different properties and dynamics of the Earth's surface characterise areas in varying degrees or mixed patterns, by gradual changes in transition zones or by temporal changes. Additionally, delineation can be difficult because of high diversity of objects or low spectral distinctiveness.

Following the WYSIWYM approach is important for successfully cope with the diverse reality. Emphasizing observation needs user-friendly thresholds and criteria to be negotiated in order to support multi-purpose applicability and to increase objectivity without running the risk of discouraging people to map. Results of observation will gain attractiveness for subsequent interpretation and analysis if a large variety of objects with different properties and in different degrees of spatial or informational detail can be allocated within the classification system.

Challenges arising from *data collection and processing* occur during the stages of survey, data preparation and classification. Information on the data base (capturing date of the imagery, collection date of on-ground information) might be unclear and require ancillary first hand confirmation. Additionally, varying distinctiveness and visibility on satellite imagery (depending on resolution, spectral similarity or mutual coverage) cause certain objects to be more easily recognized than others and hence to be preferably mapped. This imbalance can most profoundly be compensated by on-ground mapping, fortunately encouraged by the OSM project in contrast to professional data products where detailed on-site confirmation is often problematic due to limited resources.

Transparency about collection and classification procedures will increase credibility and traceability, especially in case of a crowd-sourced project were data is collected by different people at different times from different sources. Problems in recognizability and traceability can be addressed by providing aggregated classes of low detail that support identification and classification of objects on coarse detail levels. The need for confirmation or refinement – e. g. by on-ground mapping – can be indicated by the choice of certain classes, similar to the currently used tags `highway=road` and `building=yes`. This way, the WYSIWYM approach is encouraged and less detailed but correct information can be recorded instead of risking misinterpretation or avoid mapping at all. This will increase objectivity and independence from scales and collection methods.

Human interests and behaviours influence the identification and classification process during the creation of validation and reference data for automatically classified map products. Despite corresponding rules and criteria may exist in those classification systems, different analysts will come to different results when manually examining the same object, especially at different times or on different source data. Additionally, pursuing an increase of accuracy and detail will likely cause a decrease of objectivity because subjective interpretation becomes more readily accepted in order to reach the goal of a completely and detailedly mapped area. These circumstances also have to be presumed when utilizing OSM land cover/use data, given the fact of a highly diverse mapping community and a classification system of inconsistent conceptual structure.

Consequently, following guidelines, definitions and criteria recommended by the classification system will reduce subjectivity. Together with the principle of WYSIWYM honest decisions should be made that result in discussing doubts with the community prior to a delineation and classification or in delegating the judgement to mappers with better knowledge, e. g. local activists. In return, especially doubtful or difficult decisions have to be made comprehensible, e. g. by attaching comments to the changeset or the feature itself.

IV.1.2 General demands

Above, characteristic aspects of the OSM project, its community and the utilization of OSM data and land/cover information in general have been compiled and according responses relevant for the new classification system have been formulated. These preconditions can be addressed by basically eight demands. Some are interrelated or one act as a prerequisite for the other. However, since they represent rather general characteristics then pointing into the practical direction of structural specifications they were treated as an interstage between the initial situation and the final outcome.

The **focus on observation,** above related to the WYSIWYM approach, intends to reduce the space for interpretation mainly during identification and delimitation. It reduces the influence of professional, educational and cultural backgrounds. Emphasizing observation succeeds when information of various degrees of detail – different between mapping on-ground and tracing objects on low resolution satellite imagery – can be allocated in the system. It also requires a high degree of conceptual and structural consistency (concepts, definitions, criteria).

Transparency is relevant for both, mappers and users, because it allows a mutual understanding. Once definitions and criteria are precisely formulated and accessible for every mapper and user existing map features can more readily be assumed as free of subjective interpretation. The trust between mappers and users and the general credibility of the map data will improve. It will also respond to the diversity of the OSM community and the manifold utilization scenarios for which the history of the data has to be traced back to their sources and conceptual origins.

Demanding **usability & acceptance** is one of the basic responses to the preconditions described above. In order to be accepted, classes and tags have to be usable for casual mappers as well as for professional geodata analysts. Usability can comprehensively bridge the different demands of a diverse community and a variety of utilizations. For the case of OSM, the requirements on the tag design will be crucial to meet the demand.

Objectivity is another basic demand. Especially for increasing credibility and data quality the identification and classification of objects needs to be as objective as possible. It is an important factor for harmonizing motivations and objectives among mappers and also between mappers and users.

The utilization of OSM data for many potential purposes, especially for those in the realm of science and administration, demands **high data quality.** It is an important driver for the other demands and aims on aspects like spatial and informational accuracy, consistency and actuality.

The demand for **compatibility** is a response to several aspects: It contains the need for the new system to establish an easy relation to older data and an uncomplicated utilization parallel to the current way of classifying land cover/user in OSM. Furthermore, compatibility should be provided to other professional classification systems in order to enable or simplify comparisons and combined analysis.

Being able to include objects of different spatial scales and different levels of details can also be seen as a kind of compatibility. However, together with the general need to consider the natural diversity of land cover/use objects it is subsumed under the demand of **tackling diversity.** The better diversity is addressed, the more utilization scenarios are addressed as well. Facing the difficulties of mixed or gradually changing properties will increase the quality of the data because subjective interpretation or wrong classification will be reduced. The diverse backgrounds and objectives of mappers and users are further aspects of diversity to be considered by a new classification system.

IV.1.3 Practical requirements

In order to meet the demands elaborated above the new classification system needs to fulfil a set of practical requirements. Some are very similar to the general key principles introduced in chapter II.2.3. However, they originate from the OSM-specific preconditions and will shape the specification formulated in the next chapter. Since the preconditions provoke conflicting demands the requirements introduced below are not necessarily supportive in all respects (see figure 2). Some are not relevant for meeting the demand, others are even obstructive. The final specifications need to harmonize these conflicts as best as possible.

Because being presumably supportive regarding all demands, the new classification system should first of all **provide guidelines and explanations.** Once clear, comprehensible and transparently accessible for mappers and data users, they contribute to a usable, accepted and compatible system that produces objective high quality data on a global scale. Similarly, the use of **simple basic concepts** is a crucial foundation for the design of classes and the articulation of guidelines and explanations. For defining, identifying, delimiting and classifying objects, reduced complexity harmonizes understandings of different peoples and facilitates decisions. In respect of divergent natural or cultural backgrounds worldwide, common ground can be found when taking simple concepts as a starting point. For identifying and describing representative aspects of simple concepts, usually, not much information is needed. For mappers, it reduces the pressure to take decisions without actually having sufficient information and it allows them to map according to the information they can access through observation on satellite imagery and on-ground or through consultation of ancillary sources.

Providing classes for objects at different levels of detail is an important requirement closely related to the incorporation of simple basic concepts. Allowing the identification of informationally or spatially simple objects without facilitating a respective allocation in the system prevents necessary improvements. It would favour interpretation instead of observation, e. g. because of different image resolutions or mutual coverage, because sightings with low informational content can hardly be allocated in system, that does not provide appropriate classes. In contrast, classes at low detail levels can increase the credibility of the data, lower barriers for mapping unknown objects worldwide and increase compatibility because more detailed features can be systematically aggregated. However, in the case of OSM multi-leveled hierarchies can hardly be realized due to the limitations of the tag format. A sophisticated system with several detail levels would affect transparency because it can't be represented by the tagging system.

Providing comprehensible criteria and **precise definitions** are further important requirements to be fulfilled by the aspired classification system. They improve objectivity during the processes of identification, delineation and classification and allow a better mutual understanding between community members regarding the decisions made during mapping. On the other hand, it will be easier for users to interpret and analyse the data. The choice of definitions and the design of criteria is essentially influencing the compatibility with other classification systems. Since this is one important objective of the aspired amendments it is worth to consider the classification systems introduced in chapters III.1 and III.2 as a reference. In order to preserve usability and acceptance the complexity of definitions and criteria has to be limited. Hence, diversity might not be sufficiently represented. However, compared to the current situation better definitions and criteria will likely improve the ability of OSM to address the diversity of land cover/use phenomena.

Since OSM is a global project producing global data by a global community. **regional limitations have to be avoided.** Once implemented, the global universality of the system will improve compatibility and encourage mapping activities independent from cultural and natural circumstances. The demand for transparency is rather influenced by definitions, criteria etc. than by regional limitations. In terms of tackling diversity, reducing regional limitations might get in conflict with the need of considering local peculiarities. A successful solution relies on the expandability of the system (see below) so that a set of universal classes can be enhanced by the community with classes for special local phenomena.

Using consistent thematic domains will cause the general structure of the systems to become clearer and more comprehensible. Being formative for different levels or sections of the system, it is easier for mappers and users to understand the system during mapping or utilization. It will further simplify an expansion of the system because new classes can be designed and located within the system in a more structured manner. However, although mostly supportive, the influence of this aspect on emphasising observation and tackling diversity is rather low.

Whereas consistent thematic domains are important for the general structure of the system the design of **user-friendly thresholds** accompanies the definition of criteria. Especially for a crowdsourced project like OSM the acceptance of innovations depends on their usability. Thresholds in general are important guidelines for decisions that can or have to be made by considering numerical dimensions, e. g. tree height, area proportions etc. Consequently applied, they can improve objectivity and compatibility. However, once thresholds are too many, too complex or too hard to apply (requiring calculations or special equipment) voluntary mappers will likely skip objects or classes decisively defined by thresholds. Given the advantages, not using thresholds is beyond questions. All the more important is their user-friendly design: The required dimensions or proportions have to be easy measurable and split into only few distinct classes in order to allow a quick but accurate decision. On the other hand, rare and coarse thresholds might reduce compatibility with other datasets containing more detailed thresholds. It might affect the data quality and might impair the ability of the system to address details and diversity of reality.

The **availability of information on data sources** mainly addresses demands originating in the utilization of data. Being able to trace back where a feature and its adherent attributes came from allows data users to evaluate accuracy and actuality, especially in case of using the data for referencing or validating. It increases transparency, compatibility and allows verifying objectivity. Although the requirement for source information is derived from the preconditions explained above it won't affect the design and functionality of the classification system. Instead being an integral part of the system, it is rather an additional information recommended to be attached to a feature or a changeset, as currently realized by using the `source`-tag.

In contrast, providing **possibilities for expanding the system** is especially crucial to support its application on a global scale and to face the diversity of reality. Implementing simple basic concepts, preserving a generally simple structure and avoiding regional limitations leads inevitably to shortcomings in terms of addressing local peculiarities or new phenomena (e. g. new types of land use). Hence, the system should allow users to add new classes and tags, respectively, and provide guidelines to follow in order to keep new classes aligned to the rest of the system. This will increase the acceptance because it allows the community to actively participate in improving the system. In contrast, to confront the community with a fixed system not meant to be modified is in opposition to the fundamental values of the OpenStreetMap project and the self-conception of the community. On the other hand, assuring expandability bears the risk of quickly adding a new class although the respective object has been rather interpreted in the sense of a subjective category. Such a new class might not be aligned to the concepts and thematic domains used in the system. This affects the demanded focus on observation and reduces transparency because the added object is defined by individual interpretation. Additionally, increasing amount of new classes, that represent rare and special phenomena, might reduce compatibility with other systems. These problems also affect objectivity and data quality. However, the latter two demands can also have positive influence because by adding new classes objects can be recorded that were impossible to be allocated in the unchanged system.

The remaining requirements consider the OSM tagging system. Because being a constitutive, characteristic and well established component of OpenStreetMap it can't be ignored by a new system. Although rarely supportive regarding most of the presented demands (see figure 2) the limitations of the format and current utilization of the tags have to be considered. The main reason for adapting tags and for following the recommendations is to assure the acceptance of the community and the compatibility within the existing data base. One of the resulting challenges is to **preserve current tags** in order to address habitualities of mappers and users and to minimize adjustments in rendering and analysis algorithms. Parallelly, maintaining existing tags pushes the focus of amendments to the level of definitions, explanations and the classificatory systematics. Given the close relation between the levels of classification and codification (see chapter III.2.3.2), it is difficult to anticipate the effects of definitional or systematic improvements that are not reflected by tags. It will remain unclear, how close mappers stick to new definitions and criteria or simply refer to the perception on which they have decided on so far. Hence, transparency and objectivity will continue to decrease. While internal compatibility will easily be maintained, compatibility to external classification systems will remain difficult because the used tags contain a probably well described but nonetheless complex combination of concepts and criteria that are difficult to match with external classes. Under such conditions it is also hard to address the demands related to natural and global diversity.

Also mainly driven by the need for usability and acceptance, the new system is required to **provide concise, descriptive and simple tags.** Since tags do reversely affect the design of the system, it is a challenge to establish a clear separation of classes by incorporating distinguished definitions and concepts, while knowing that these classes might not be expressed by a usable tag. A simple tag can hardly express the position of a class within a the classification system, not to mention other characteristic details of the feature. Tags are descriptive and self-explaining, if they sufficiently represent a certain informational aspect of the feature to which they had been attached. Reversely, if simplicity limits the expressiveness of tags the only way of assuring transparency is to adjust the set of definitions and criteria, accordingly. Consequently, necessary details and distinguishing aspects are difficult to be integrated into the system if their representation at the codification level complicates the tag structure. Refusing to sharpen class and object definitions in favour of a simple tag format reduces data quality, affects external compatibility and hampers emphasizing observation when visible properties are not represented by the system. However, because of its historically grown structure and format the tag-based OSM system of codifying classes and properties can't be neglected and has to be adopted by a new approach, despite its challenging characteristics and limitations.

IV.2 Specifications

IV.2.1 Objectives

The proposed *classification system* aims on describing the Earth's surface by using classified *land objects.* Their identification and delineation will be accomplished by a global community of voluntary contributors that are characterised by diverse cultural, educational and professional backgrounds. Current and potential utilization scenarios in the realms of administration, environmental planning or geoecological modelling, to only name a few, are similarly diverse and need to be considered in order to support the general goal of the OpenStreetMap project: Providing free geospatial data for any interesting purpose. The new system tries to address these two types of diversity by providing a systematic framework for answering only two fundamental questions about objects observed in the real world: “What is it?” and “What is it for”? (DUHAMEL 2009). Regarding these two aspects of how land is covered and used, improvements compared to the current classification system are pursued: Reducing mistakable classes and misunderstandings between contributors; simplifying and encouraging the mapping of land objects; enhancing attractiveness and usability of OSM data for professional utilization; increasing compatibility with other land classification systems.

Two of the main motivations for amending the current system are parallelly two of the main challenges: Increasing compatibility with other professional classification systems while improving usability and acceptance by the community. The choice and implementation of fundamental components like definitions, explanations, guidelines, criteria & thresholds, as well as the general structure of the system is highly important for addressing this demand (see chapters IV.1.2 and IV.1.3). The underlying concepts and general aspects described below are equally formative for the design and later application of all classes. Especially in terms of applying basic concepts and implementing usable thresholds, the alignment with the FAO LCCS improves the capabilities that allow combination and conflation with other classification systems. Furthermore, the LUCAS scheme is taken for orientation mainly because of the way classes are presented and because of the solutions for practical mapping challenges. But especially for dealing with the separation of land cover and land use within one classification system, the LUCAS scheme is a useful example.

IV.2.2 Concepts

The need for addressing the diversity of contributor's backgrounds and user's interests requires the new system to provide conceptual simplicity. Identifying and classifying an object has to succeed by collecting only a few information that are easy to assess. Concepts depending on special equipment, multiple detailed information and technical expertise are not suitable for a voluntary project.

As a result, the proposed *classification system* is based on rather fundamental concepts. One prominent result of this intention is the separation of land cover and land use. Given the fact of repeated discussions and misunderstandings between community members, summarized in chapter III.2.2.2, this measure had to be seriously considered. The complexity of the relations between the two concepts has been introduced in chapter II.2.1. Facing this, a break down of definitions that require information on both aspects has the potential to simplify and harmonize the classification and mapping process.

For describing and defining classes, the proposed system follows the current practice of the OSM Map Feature Documentation by referring to explanations and definitions provided by the Wikipedia project. If not, the differences will be clearly outlined.

IV.2.2.1 Land

Following the FAO's definition, ***land*** means a “delineable area of the earth's terrestrial surface, embracing all attributes [...] immediately above or below this surface”, including inland water areas, wetlands, tidelands and swamps (see chapter II.2.1).

This definition encompasses floristic, faunistic, climatologic, pedologic and hydrologic characteristics in diverse combinations and fluctuations over space and time. Depending on the objectives or the disciplinary background of a *classification system* certain properties of interest are chosen and used for identifying and classifying *land objects,* the actual means of observation.

IV.2.2.2 Land object

Land objects are spatial entities characterised by a rather homogeneous setting of one or more specific properties that can be investigated and recorded. Depending on the complexity of the underlying concepts multiple properties might be required to be investigated. Information for ecological, geomorphological or socio-economical concepts have to be assessed in special way, to be combined according to a certain method and to be interpreted based on special knowledge or experience. Only afterwards an object can be identified, delimited and classified.

Contrarily, an OpenStreetMap Land Classification System aims on a basic characterization of *land objects* by applying concepts described by a small set of properties that are easy to be assessed and to be interpreted. The conceptual basis follows the aim of classifying *land objects* by describing their visible appearance and the purpose they serve, represented by the concepts of *land cover* and *land use,* respectively. However, it is not the case, the both information are required for mapping a feature. This way, heterogeneous preconditions during contribution (availability of information, personal background of contributors etc.) are considered.

IV.2.2.3 Land cover

Referring to chapter II.2.1 ***land cover*** describes the observable bio-/physical overlay of the Earth's surface. It encompasses natural coverage (e. g. vegetation, inland waterbodies, wet areas, snow/ice), artificial constructions (buildings, sealed ground) and uncovered areas (bare soil, rock). It is limited to phenomena observable on-ground as well as on remotely sensed imagery. Hence, soil types (like Podzol or Chernozem; in German “Bodentyp”), conditions of subsurface water and vegetation (aquifers, roots, mycelia) or artificial structures under ground (traffic tunnels, mines) are not considered by this definition. Because only true colour imagery is usually available for the OSM project the proposed *classification system* will not differentiate between objects that are only distinguishable by using additional spectral information, e. g. different types of snow/ice cover.

The designation of a certain *land cover* type depends on kind, composition and degree of coverage. The most diverse and heterogeneous combinations of these three aspects can be presumed for areas covered by vegetation. Consequently, the new classification system emphasizes on vegetation as a main factor and has to deal with the term *coverage.*

IV.2.2.4 Life form & coverage

Physiological differences of plants play a crucial role in the class definitions of the proposed system. Usually, three main ***life forms*** are separated: *Threes, shrubs* and *herbaceous plants* (short: *herbs*). Trees and shrubs endure over years and decades by forming wooden stems and branches. In contrast, the life cycle of herbaceous plants is much shorter. Furthermore, they reach much lower height due to lacking lignification.

The OSM LCS aspires compatibility with the FAO LCCS. Therefore, life form and coverage follows the perception implemented by the FAO, based on KUECHLER & ZONNEVELD (1988) and Eiten (1968), among others. For the proposed system, *herbaceous plants* accordingly encompass grass, forbs and (non-tree) ferns.

Following EITEN (1968), ***vegetation coverage*** is first of all characterised by the widest and highest level of canopy expansion. For recording *land cover,* this means firstly that the degree of coverage depends on the plant's biggest circumference, e. g. the crown of a tree instead of the circular area occupied by its stem at ground level. Secondly, plants of lower height become irrelevant once they're overgrown and covered by others. In case of seasonal fluctuations, the maximum extend is anticipated in order to estimate the degree of coverage, e. g. at the stage of full leaf development on deciduous trees or shrubs.

This provides consistency between data collected on the ground and data solely retrieved from remote sensing imagery where the uppermost layer dominates the visual impression of a covered area. Consequently, if existent at a certain degree of coverage the higher growing vegetation types are favoured for defining a land cover type (see below). The number of certain species or the area occupied at ground level are not considered. Consequently, grass or low growing bushes become irrelevant for the designation of a certain land cover type once tree canopy reaches a certain degree of closure.

This indicates the need for thresholds, also in order to address the occurrence of mixed properties. These aspects will be elaborated below. However, complains about a lacking separation of the terms *land cover* and *surface* expressed by the community can already be addressed.

IV.2.2.5 Surface

In contrast to *land cover,* the concept of *surface* is already adopted in the OSM project, mainly for characterising the material of roads, paths and sports fields by using the key `surface` (OSM_{64}). Without changing the current perception the scope of use can be extended in favour of the proposed classification system. Since a certain land cover type represents state and characteristics of coverage, *surface* can be used to describe details about the material at ground level. This essential for uncovered or artificial areas but can also be used to provide additional information for covered areas, e. g. indicating an open sand area (`surface=sand`) below a closed palm tree stand (`landcover=trees`) in an oasis.

IV.2.2.6 Land use

In the context of the proposed classification system the concept *land use* is applied for describing the purpose of a *land object.* The question on “What is it for?” is answered by considering human activities that cause an area to serve a set of socio-economical functions and/or that influence land cover characteristics. As stated in chapter II.2.1 *land use* is usually difficult do be determined by observation, especially in case of working with remote sensing imagery. However, the spatial scope of a certain purpose has to be delimitable either by observation or by referring to reliable ancillary data. In case of subsurface activities (transportation, mining etc.), this definition applies only the visible and delineable extend above ground.

IV.2.2.7 Land key element

Although *land objects* can be preliminarily detected as spatial units (due to distinct boundaries, homogeneous appearance etc.), cover and/or function might remain unknown or ambiguous at the first sight. *Land key elements* are observable components serving as typical indicators for a certain *cover* or *function.* Their existence and identification allows a more reliable designation of particular land cover or land use types. Examples are facilities for livestock watering or hay bales that indicate the utilization of grassland.

IV.2.3 General systematics and formative factors

The following chapters are first of all meant to provide inspirations on how the new system could be realized. The study mainly aims on building a framework and presenting recommendations on how it can be filled. In this respect, the aforementioned findings shall provide a comprehensible compilation of aspects worth to be considered in order to assure a maximum of transparency during the formation of the new system. Consequently, the structure and the single class description will remain preliminary and incomplete; also due to the limited scope of this study. However, a foundation shall be provided to be used for further substantial discussions within the community.

IV.2.3.1 Structure & codification

Similar to the LUCAS scheme, the proposed system is divided into two parts, dealing with land cover and land use, respectively. Classes are arranged in different hierarchy levels, expressing different degrees of aggregated information. The classes are defined according to an a priori approach and try to cover all combinations of the chosen properties. Depending on the information and the degree of detail available, the hierarchical tree provides a systematic orientation for mappers to find the appropriate class for the observation they've made.

Realization: *The entrance into the classification system can be provided by a central page in the OSM-Wiki. Here, the separation of land cover and land use can be clarified by* *presenting the according definitions and justifications. Adjacent concepts and characteristics valid for the general structure of the entire system as well as for both subsystems should be provided here. It could also be the place for an index of objects covered by the system, including links to the respective feature pages. Finally, links should lead to the respective subsystems, represented by the keys `landcover` and `landuse`.*

At this level, the specifications of the subsystems for land cover and land use should be provided respectively. Mapping guidelines in general, underlying concepts and naming rules for each hierarchy level should be transparently presented to every contributor and user.

As characteristic for OSM in general (see chapter III.2.3.2), both subsystems try to answer the questions of *What do I see?/What is there?* and *How is it characterized?* In case of land cover, the *What*-question mainly describes the condition of the vegetation cover Existence, life form and coverage are therefore the main aspects used for defining and arranging land cover classes (see below).

Beside the position in the classification tree, every class is described by a *definition*. An *explanation* elaborates the differences to other classes and provides information on the objects addressed. A section already established in the current OSM-Wiki is taken as a standard component for the new system: Under *How to map* advices for identification and delineation are provided, containing observable characteristics, measurable criteria and land key elements that act as class specific indicators. A further section explicitly contains information on *When/ When not to use this tag,* including objects meant to be allocated to this class, so called *includes*. *Excludes* are listed as well, complemented with links to the more appropriate class. The section *Additional information* explains which tags can be attached for expressing further detail on the current object. Similarly, properties already implied by the chosen class are listed as well, so that double tagging can be avoided. This is very important in case of unquestionable 1 : 1-relations between two certain land cover and use types. In order to highlight the connection between the new and the old system, the section *Relations* provides links to current classes of similar content, as well as to classes being replaced by the new one. Reasons for replacing or deprecating current tags are provided in order to clarify the amendments and differences between the old and the new system. Separation and conflation of classes or concepts are important to be made transparent.

Realization: *The various sections used for describing classes are supposed to be compiled at singular Wiki pages, one for each `value` associated to a `landcover`- or `landuse`-key. The page should visualize at which position of the classification system the respective class is located. The structure of every single `value` page follows common OSM practice. Especially in case of land use classes, less adjustments can be expected. In contrast, new pages have to be created for the new land use classes.*

<u>Vegetated areas</u>

For the section dealing with vegetated surfaces, life form, height and coverage are the essential aspects used for subdivision. The resulting classes are arranged in three subordinate hierarchy levels and are represented by newly designed tag values. In contrast to the *What*-questions, some tags for answering *How*-questions are already established in the data base and can be used as they are, e. g. leaf type (`leaf_type=*`) or leaf phenology (`leaf_cycle=*`).

The first level simply represents two basic physiological characteristics: woody and herbaceous plants. Level two allows areas to be designated in a very simple but scale independent way by emphasizing on the predominating life form: *Trees, shrubs, herbaceous* *plants* or *crops.* One level below, the increase in detail concerns the degree of coverage, by separating the ranges of *closed (forests, thickets, meadows, fields)* and *open (woodland, shrubland, grassland and cropland).* The contributors are encouraged to use the last two levels of the system because their degree of detail is similar to the one used for defining the current classes.

Non-vegetated areas

The section for non-vegetated areas contains four major subclasses, namely *water, bare ground, snow & ice* and *artificial.* Because of distinct differences, their further subdivision is irregular, so that this sections consist of two or three subordinate hierarchy levels. Nonetheless, compatibility with the FAO LCCS is assured.

Wetland areas

In contrast to the LCCS and other land classification systems, wetlands are not subsumed in an extra branch or class of the system but are first of all described by the life form settings, too. For example: An area designated with the ecologically termed *bog* can first of all be described according to observable vegetation settings, e. g. by delineating areas dominated by woody or herbaceous plants. Information indicating the wetland character are additionally attached. Following the proposed classification system, a bog can consist of several areal features with the following tags:

feature 1 `landcover=trees + wetland=yes`
feature 2 `landcover=herbs + wetland=yes`

Instead of `yes`, the system allows to add other values if sufficient information is available for classifying the type of wetland: `bog`, `swamp`, `marsh` etc. Most of them do already exist in the current tagging system. Wetland conditions can be expressed for both, vegetated and non-vegetated areas.

Land use

For the land cover subsystem it was necessary to establish the rarely used key `landcover`. In contrast, the key `landuse` is commonly used in combination with numerous values and additional tags (see table 17). Consequently, a lot of tags are already provided at the current state of the system. The proposed classification system presents mainly a new structure and tries to reuse or create connections the current tags. The system emphasizes on terms indicating activities instead of objects. The class names aim on fitting one if the two questions: *The area is used for [noun],* or *The area is used for [adjective] purposes*. The fundamental approach refers to the LUCAS scheme and arranges land use types as they belong to the primary, secondary and tertiary sector. This level is not supposed to be used for mapping. Subsequently, classes are subdivided into a second and a third hierarchy level following thematic aggregations. For the state of this study the choice of classes aspired a balance between established tags and concepts on the one side, and the compliance with the ISIC classification on the other side (see chapter III.1.4.1). Due to the enormous amount and diversity of human activities affecting the Earth's surface, the proposed land use classification only provides a starting point and needs further expansion and rearrangements.

IV.2.3.2 Areal predominance and horizontal mixture

The designation of land cover and land use classes is based on evaluating areal distributions of certain aspects. To assure usability, the design of thresholds and circumscriptions has to support the feasibility of visual estimation, instead of requiring exact quantitative measurements.

As stated above, vegetation coverage refers to the maximum canopy expansion of a plant, disregarding others growing below. First of all, the vegetation type is defined considering life form and height. In order to comply with the FAO LCCS, the following thresholds are used:

- herbs < 0,3m
- shrubs > 0,3 – 3m
- trees > 3m

Anyway, this provides only coarse guidance, so that other factors are often considered, too:

- meadows with herbs taller than 30 cm are still not treated as shrubs;
- shrubs taller than 3 meters remain treated as shrubs;
- high growing reed is treated as shrubs (but never as trees);
- allocation of palm trees and bamboo depends on height, although taxonomically being closer to herbs than to shrubs or trees;
- trees below 3 meters in height are treated as shrubs;
- young shrubs and trees below 30 cm might easily overseen within surrounding herbs and are therefore allocated to the herbs layer.

Once a vegetation type at its maximum canopy expansion covers a distinct majority of an area, it is assumed as predominant and forms a *closed cover.* An *open cover* is given when coverage is less than predominant but more than *sparse,* meaning that at least two different aspects cover the Earth's surface. The FAO LCCS provides the following thresholds:

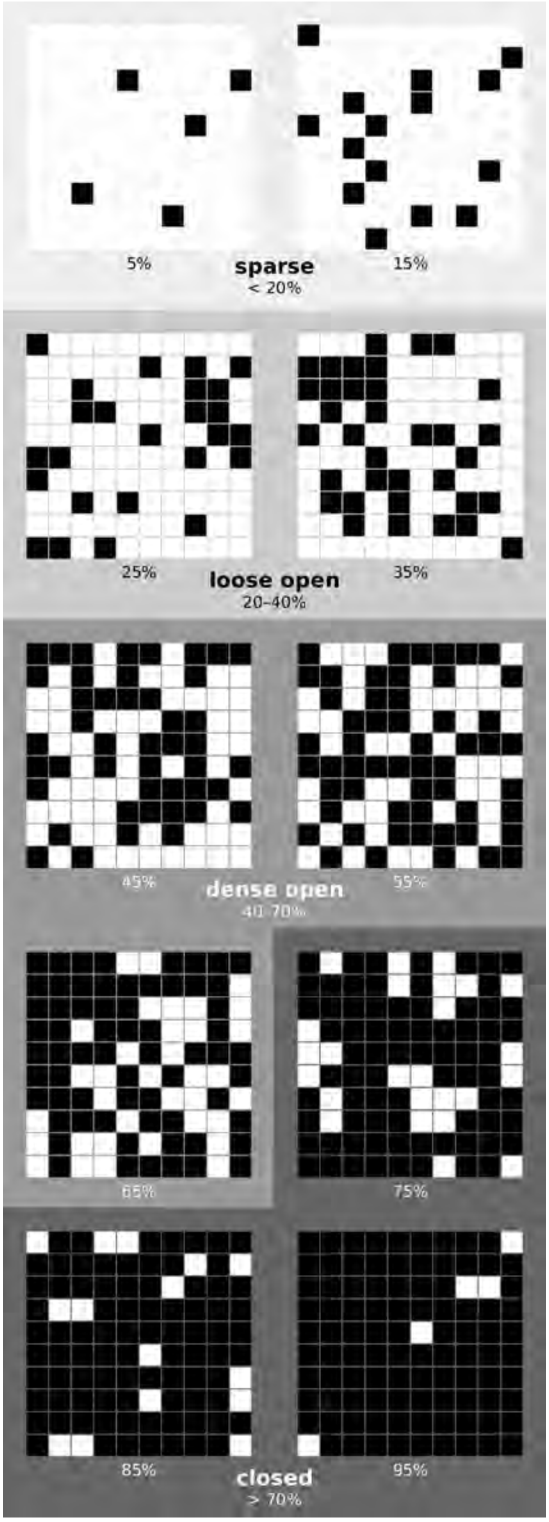


Figure 3 Schematic illustration of vegetation coverage. Thresholds acc. to FAO LCCS; visualization: the author.

- sparse > 10–20 %
- open 10–20 – 60–70 %
- closed > 60–70 %

However, the implementation of the thresholds needs complicated rules to be applied during mapping practice and during the design of mixed classes. Fortunately, the FAO LCCS provides a subdivision for the open cover class, so that the finally implemented thresholds are the following:

- sparse > 5–20 %
- open > 20–40 % (“loose”)
> 40–70 % (“dense”)
- closed > 70 %

The usability of the thresholds during mapping activities shall be improved by the visualisation in figure 3 and the following comparisons:

Trees touching each other's crowns at the outer perimeter cover an area of nearly 80 %. Thus, the chosen threshold of 70 % additionally addresses some further irregularities and gaps in the canopy. When a coverage of 40 % is reached, the distance between the crowns approximately equals their mean radius. Together with the fact, that an area covered to this degree appears closed from far above, the threshold seems appropriate as a pragmatic representation of an areal majority. (DI GREGORIO 2005)

This perception of predominance (coverage > 40 %) is implemented in the first three levels of the classification system. At the lowest level, closed and open cover is distinguished. The latter one actually refers only to a coverage range of 40–70 % (dense open cover). This way, confusion is prevented that is likely to arise when an area has to be labelled with a land

Table 24 Mixtures of trees with other vegetation types. For the corresponding tags and FAO LCCS codes see appendix table A1.

<table>
<tr><th rowspan="2">TREES</th><th rowspan="2">closed</th><th colspan="2">open</th><th rowspan="2">sparse</th></tr>
<tr><th>dense</th><th>loose</th></tr>
<tr><td></td><td>> 70%</td><td>40 – 70%</td><td>20 – 40%</td><td>< 20%</td></tr>
<tr><td>shrubs</td><td></td><td></td><td></td><td></td></tr>
<tr><td>> 70%</td><td>X</td><td>X</td><td>X</td><td>thicket w/ sparse trees</td></tr>
<tr><td>40 – 70%</td><td>X</td><td colspan="2" rowspan="2">acc. to predominance woodland/shrubland or shrubland/woodland</td><td>shrubland /w sparse trees</td></tr>
<tr><td>20 – 40%</td><td>X</td><td>X</td></tr>
<tr><td>< 20%</td><td>forest w/ sparse shrubs</td><td>woodland /w sparse shrubs</td><td>X</td><td>X</td></tr>
<tr><td>herbs</td><td></td><td></td><td></td><td></td></tr>
<tr><td>> 70%</td><td>X</td><td>X</td><td>X</td><td>meadow w/ sparse trees</td></tr>
<tr><td>40 – 70%</td><td>X</td><td colspan="2" rowspan="2">acc. to predominance woodland/grassland or grassland/woodland</td><td>grassland /w sparse trees</td></tr>
<tr><td>20 – 40%</td><td>X</td><td>X</td></tr>
<tr><td>< 20%</td><td>forest w/ sparse herbs</td><td>woodland /w sparse herbs</td><td>X</td><td>X</td></tr>
</table>

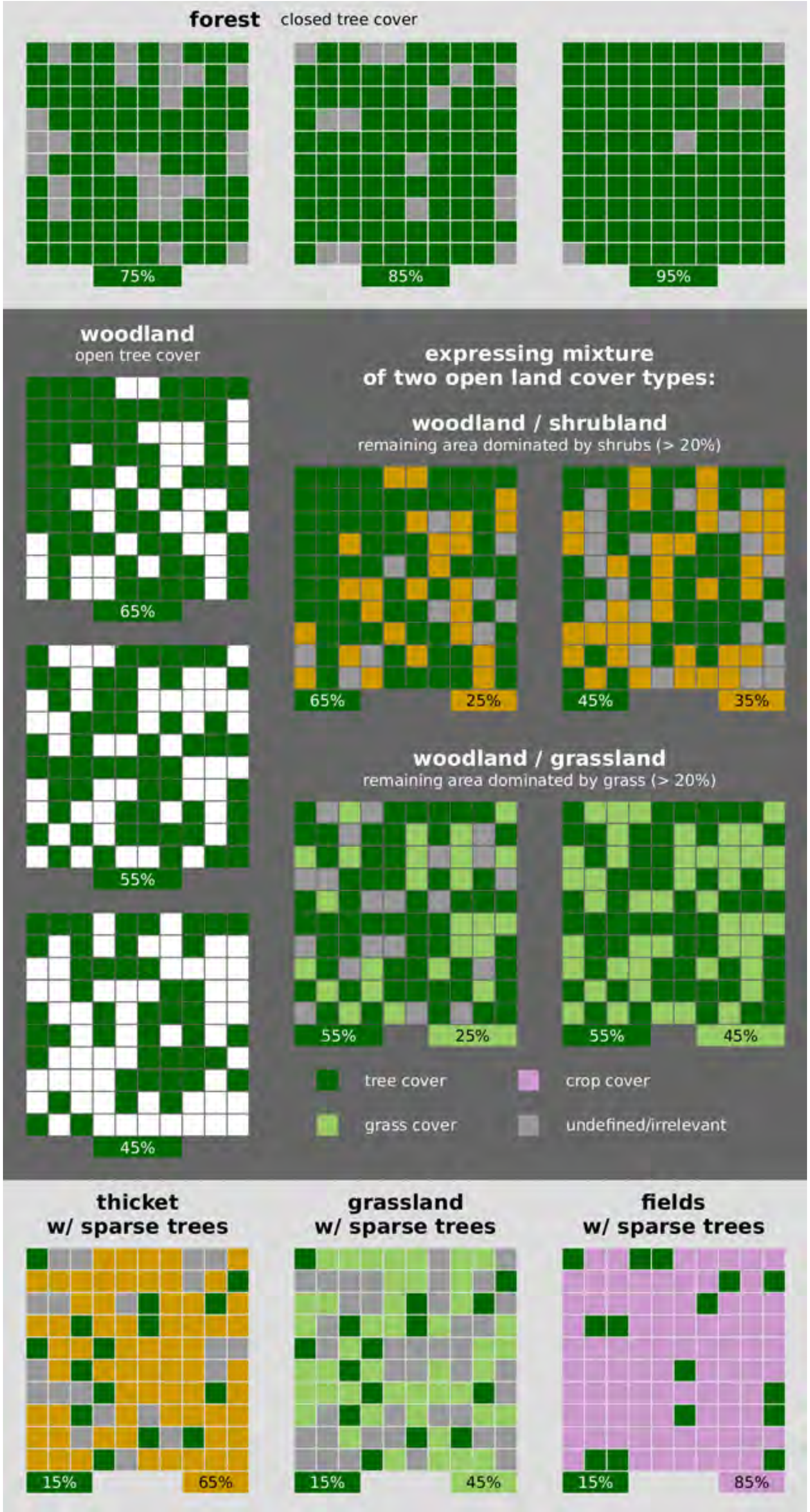


Figure 4 Mixture of land cover types in accordance with table 24. Visualization: the author.

cover type that does not predominantly cover the area. Example: Without the subdivision of dense and loose, an area covered with shrubs by only 25 % and grass by 65 % would have to carry the shrub label. Contrarily, in combination with a tree coverage of 65 % the shrubs become irrelevant and the area is allocated to a tree class. The necessity of parallelly considering coverage and height is likely a source of confusion, misunderstandings and inconsistencies in the data base. It will furthermore complicate the dealing with horizontal mixtures.

The proposed classification system emphasizes the aforementioned perception of predominance (40 % threshold) even in case of an open cover. Consequently, if one major type predominates without forming a closed cover, another type can be identified once it covers the majority of the remaining area (see figure 4). For example, an area predominantly covered by trees (about 55 %) can be solely assigned to the class *woodland*. If the remaining area is dominated by grass (covering 25 % of the entire area) the new system allows to indicate a mixture of wood- and grassland (order indicates major and minor). Tags are accordingly composed in order to complex double tagging (see table 24; for the stated example `landcover=woodland_shrubland`). If sparse shrubs occur as well, this information can be added optionally by using an extra tag (`coverage:shrubs=sparse`). Even in case of attaching this optional information, the compatibility with the FAO LCCS is assured.

Closed and openly vegetated areas are represented by distinct terms at the lowest hierarchy level, e. g. *thicket* or *grassland*. Non-vegetated areas are treated slightly different: The classes defined in this section do always refer to a closed cover. An open cover has to be indicated by expressing a mixture with other land cover types, e. g. rock/vegetation, snow/scree or water/herbs. Consequently, information on the degree of coverage becomes less detailed, but this way the decision on clearly delineating areas without vegetation is easier. It encourages to consider the classes of the former section, once vegetation occurs.

Similar to the designation of land cover types, land use classes are assigned based on areal predominance. The classes defined by the systems introduced in chapter IV.4 are worth to consider. However, the underlying criteria are often related to the monetary outcome of a certain business activity. Since we're working in the realm of geodata collection, the spatial extent of objects remains the major aspect that voluntary mappers are able to estimate. Consequently, according to the criteria used for the ISIC classification, an urban area dominated by multi-story apartment houses with shops at the ground flour would be designated as residential or commercial, depending on the turnovers or profits generated by each economic activity. Because this information is likely to be undisclosed for the eyes of voluntary mappers, the proposed classification system defines land use according to spatial predominance. In case of the mentioned example, a dominance of the residential can observed because more area (not Earth surface) is used for residence than for retail – several flours vs. ground flour, respectively.

IV.2.3.3 Spatial & temporal consistency

The increase in spatial and temporal credibility has been stated above as an important driver of the aspired re-arrangements of the OSM tagging system. The new system prevents the contribution of false information by providing classes of different spatial and informational detail. In general, the more detailed an information is, the more prone it is to heterogeneities and the risk of not being up to date.

In temporal terms, the author argues for emphasizing on mid and long term conditions instead of mapping frequently changing properties. The information about an area to be used for crop production is quite stable. It persists over a longer time period and is less prone to conflicts between satellite imagery and reality. In contrast, recording seasonal changes of crops or special cultivation techniques is very time consuming and probably only conducted by very enthusiastic micro-mappers. In addition, it might turn into a false information once such a change is not instantly updated. Therefore, the classes provided by this system refer to long and mid term conditions. Regularly changing conditions like the seasonality of water conditions can be indicated by adding appropriate attributes (`seasonal=*`, `intermittent=*`).

The proposed system is supposed to by applied to areal objects, represented by areal OSM features. Classes with a more general information content are not necessarily restricted to be used for delineating areas at a small scale. For example, *shrubs* can be used on large areas where no other information is available or intended to be contributed. In this case, it indicates the possibility of further subdivision into parts characterized by a *closed* or *open cover*. At the other end of the scale, this class is assigned to areas that are too small for a reasonable estimation on *open* or a *closed* coverage.

For dealing with linear features, like rivers, a threshold has to be discussed. For remote mapping, the main reasons to be considered are offsets between satellite images and their unknown state of geometric correctness (OSM_{65}). For on-site mapping, the error margin of GPS receivers limits the accuracy. The author consequently suggests, that linear features shall be mapped as areas once their mean widths exceeds the standard GPS error of about 8 meters (USDOD 2008).

A further aspect is the inclusion of *associated areas.* This way, the system follows the FAO LCCS and the LUCAS scheme. Especially in case of land use it simplifies the delineation. As an example: Beside the stone quarry pit in a narrow sense, the definition of the class *mining & quarrying* would also contain the area of adjacent post processing machinery, an office building and a parking lot for costumers. Those are delimited as separate amenities and not covered by a land cover/use system.

Independence from scale and from collection methods also contributes to spatial consistency. The proposed system avoids classes that contain a thematic and regional bias. Not only that allocation is difficult because those classes require multiple information that might be difficult to assess and to interpret. They also cause spatial problems because they might consist of several parts with heterogeneous characteristics or can only be applied under certain regional conditions.

To state some examples:

- The concepts *bog* and *fen* describe landscapes and biotopes. On the first sight, the have a similar appearance, but in fact they're distinguished by very different hydrological conditions that are difficult to asses. Allocation mistakes are likely to occur. The proposed system allows to purely describe the vegetation settings without pushing a mapper to decide between two complex concepts. Anyway, if an interpretation can be surely conducted, bog or fens can still be attributed by an appropriate tag.
- *Tundra* is a concept that describes a biome related to a certain geographic or topographic location (ant-/arctic or alpine). It is characterized by a setting of vegetation types that overlap with other concepts, e. g. *fell* ($OSM_{66\ \&\ 67}$). Additionally, the fringes of those areas are very blurry and make comprehensible spatial delimitation very difficult.

These examples also prove a strong dependence on scale: Heterogeneity of properties can be characteristic for a certain concept. If working on a larger scale, more homogeneous sub-areas can be distinguished. The required heterogeneity gets lost and the concept can not be applied any more. The proposed classification system provides classes that can be applied on various scales. They might express less detail, but their information remains to be right (e. g. classifying grass covered areas only as *vegetated,* although the mapping is conducted on-ground).

IV.2.3.4 Compatibility

The challenge of the new system is to harmonize two aspects: On the one side, for the aspired compatibility with other classification system, class names, tags and the fundamental classification structure had to be inevitably adjusted. On the other side, compatibility to the original OSM tagging scheme wants to be preserved as best as possible. Therefore, the relation of the new classes and their current counterparts need to be evaluated. Tables 27 and 30 are approaches to systematically compile such results. The former describes the relation between new and old tags by considering four aspects:

- "equals": classes of the current system, that are basically identical to the new class;
- "is similar to": classes of the current system, that are similar to the new class;
- "includes": classes of the current system, that fall entirely into the spatial/semantical scope of the new class, but represent only a part of it;
- "is part of": classes of the current system, that have a wider scope than the new class, so that the new class covers only a part of the current class.

Contrarily, the relation should be evaluated in the other direction as well. Table 30 provides a starting point for showing, how current classes/tags could be expressed by using new classes/tags.

IV.3 OSM Land Cover Classification System

The following pages present the land cover section of the OpenStreetMap Land Classification System. The tables 25 and 26 introduce the main classes for vegetated and non-vegetated areas. Chapters IV.3.3 presents templates for a tabular compilation of compatibility isues.

Following the current mapping practice, the Map Feature Documentation will remain the major source of information and guidance, also for the new system. The integration of the system into the OSM Wiki is therefore an important measure for assuring usability and acceptance. In chapters IV.3.1 and IV.3.2 the classes *trees* and *water* are used to exemplify this integration. Hashed underlinings indicate proposed links to other pages of the OSM Wiki.

IV.3.1 Example *trees* (`landcover=trees`)

This class is part of the land cover section in the OSM Land Classification system. It indicates a vegetated land surface and together with the class shrubs it forms the superordinate class woody. It contains two further subdivisions, namely wood and woodland, that represent different degrees of tree coverage.

Definition

Trees are perennial plants, characterized by a major stem and minor branches. They are at least 3 meters tall and can form a crown of various shapes.

Table 25 Tabular layout of the section for vegetated areas in the OpenStreetMap Land Cover Classification System. *Italic tags* are newly proposed or rarely used ([#] number of tagged features; as of August 23rd 2016 via taginfo.openstreetmap.org). Transfer to FAO LCCS does not consider wetland or land use condition (see chapter V.3).

class	tag / short description	FAO LCCS	coverage
vegetated	*landcover=vegetated* predominantly covered by vegetation of any type; further details are unknown, ambiguous or not intended to be contributed; indicates need for refinement	A12 A21	> 40 %
woody	*landcover=woody* predominantly covered by woody vegetation (trees, shrubs and woody crops); further details are unknown, ambiguous or not intended to be contributed; indicates need for refinement	A12 A1A21	> 40 %
trees	*landcover=trees* [9 620] predominantly covered by trees or tree crops; further subdivision difficult, not intended or not reasonable (small areas); indicates need for refinement in case of large areas	A12 A3A21	> 40 %
wood	*landcover=wood* [82] closed tree cover (acc. to degree of crown closure; understory irrelevant)	A12 A3A10	> 70 %
woodland	*landcover=woodland* [3] open tree cover (crown closure fragmentary but still predominant)	A12 A3A12	40–70 %
shrubs	*landcover=shrubs* [15] predominantly covered by shrubs or shrub crops; further subdivision difficult, not intended or not reasonable (small areas); indicates need for refinement in case of large areas	A12 A4A21	> 40 %
thicket	*landcover=thicket* closed shrub cover (acc. to degree of maximum canopy expansion)	A12 A4A10	> 70 %
shrubland	*landcover=shrubland* [1] open shrub cover (canopy closure fragmentary but still predominant)	A12 A4A12	40–70 %
herbaceous	*landcover=herbaceous* predominantly covered by herbaceous vegetation (herbs and crops); further details are unknown, ambiguous or not intended to be contributed; indicates need for refinement	A12 A2A21	> 40 %
herbs	*landcover=herbs* [14] predominantly covered by grass, forbs and ferns; further subdivision difficult, not intended or not reasonable (small areas); indicates need for refinement in case of large areas	A12 A2A21	> 40 %
meadow	*landcover=meadow* [17] areas with a closed grass/forbs/ferns cover	A12 A2A10	> 70 %
grassland	*landcover=grassland* [30] predominantly covered by grass/forbs/ferns; interspersed with other land cover types	A12 A4A12	40–70 %
crops	*landcover=crops* areas predominantly used for/ covered by annual or biennial crops; further subdivision difficult, not intended or not reasonable (small areas); indicates need for refinement in case of large areas	A11 A3	> 40 %
field	*landcover=field* dominated by cultivated crops, incl. temporarily unworked or ploughed areas	A11 A3	> 70 %
cropland	*landcover=cropland* predominantly covered by areas used for crop cultivation; interspersed with other land cover types	A11 A3	40–70 %

Table 26 Tabular layout of the section for non-vegetated areas in the OpenStreetMap Land Cover Classification System. *Italic tags* are newly proposed or rarely used ([#] number of tagged features; as of August 24rd 2016 via taginfo.openstreetmap.org). Transfer to FAO LCCS does not consider wetland or land use condition (see chapter V.3).

class				tag / short description	FAO LCCS
non-vegetated				*landcover=non-vegetated* for areas without vegetation and without any further information or observable indicators for a certain land cover type; indicates the need for further investigation	–
	water			*landcover=water-area* water surface of unknown kind and origin	B27 A1 // B28 A1
		water body		*landcover=water* [322] bodies of standing water; origin and further details unknown	B27 A5 // B28 A5
			lake	*landcover=water* + water=lake natural fresh water bodies	B27 A5
			reservoir	*landcover=water* + water=reservoir artificial water body created by blocking a watercourse	B28 A5
		waterway		*landcover=waterway* linear bodies of flowing water; origin and further details unknown	B27 A4 // B28 A4
			river	*landcover=waterway* + waterway=river water area of a natural river course	B27 A4
			canal	*landcover=waterway* + waterway=canal water area of an artificial waterway that connects other water areas	B28 A4
	ice & snow			*landcover=ice-snow* areas predominantly covered by snow and/or ice most of the time; distinction not possible, not intended or not reasonable (e. g. due to frequent changes)	
		ice		*landcover=ice* areas predominantly by ice most of the time (e. g. glaciers)	B27 A3 // B28 A3
		snow		*landcover=snow* areas predominantly by snow most of the time; excl. snow cover on top of permanent ice bodies	B27 A2 // B28 A2
	bare ground			*landcover=surface* natural areas not covered by vegetation or water, exposing Earth's surface material	B16
		rock		*landcover=surface* + *surface=rock* [95] areas dominated by barren, mainly solid rock surfaces	B16 A7
		sand		*landcover=surface* + surface=sand areas covered by sand	B16 A6
	artificial			*landcover=artificial* areas of artificial origin and appearance	(B15 (A1//A2) // B16 A1
		built-up		*landcover=built-up* areas covered or affected by vertically built structures, like houses, dams, abutments etc.	B15 A1
		non built-up		*landcover=non_built-up* areas with compacted, sealed or shifted surface	B15 A2 // B16 A1
			pavements	*landcover=paved* [2] sealed or artificially compacted areas	B16 A1
			disposal & deposit	*landcover=pile* areas covered by loose intentionally deposited material, e. g. waste dump sites, spoil tips, etc.	B15 A5
			extraction sites	*landcover=pit* open mines and quarries	B15 A6

Explanation

This class is meant to be assigned to areas predominantly covered by trees taller than 3 meters. The degree of coverage refers to the area occupied by the tree crowns; understory vegetation is therefore not considered. Information on hydrological conditions, origin (artificial/natural) or utilization are not required for the application of this class.

When/ When not to use?

Small scale mapping:

- delineating large areas predominantly covered by trees;
- separation into areas with closed and open cover (wood, woodland) not possible or not intended
- expressing necessity/possibility for separation into sub-areas with closed and open cover.

Large scale mapping:

- delineating small areas or groups of trees where further separation in open and close cover is not possible or not reasonable.

Includes

- small tree areas/groups in parks or settlements
- large tree covered areas with heterogeneous coverages (above 40 %)
- mature palm tree and bamboo stands
- tree crop plantations
  recommended: + `landuse=orchardry`
- swamps, mangrove areas or tree covered parts of other wetlands
  recommended (if information on hydrology, species assemblage etc. is available/ reliable): + `wetland=swamp`/ `~=fen`/ `~=mangrove` (instead, + `wetland=yes` if interpretation on wetland type is difficult)
- …

Excludes

- reforestation areas, densely covered with young trees (smaller 3 meter)
  - instead: `landcover=shrubs`
  - recommended: `landcover=thicket`
  - if information available/reliable: + `landuse=forestry`
- tree rows
  instead: `natural=tree_row`
- single trees
  instead: `natural=tree`
- …

How to map

The use of this class is limited to areas. Boundaries are drawn along distinct borders of land cover changes (shorelines, settlements etc.) or – in natural transition zones – along the approximate line where tree cover looses predominance (< 40 %). Subdivision is necessary in case of changing properties that have been additionally recorded and attached to the feature (e. g. tree species or wetland conditions).

As representative geometries, closed ways are the easiest to use to create the feature. Direction of the way is irrelevant. If the area is too big or enclosed land cover types are intended to be delineated, multi-polygon relations should be used.

Additional information

Implies

- `landcover=woody`, `landcover=vegetated`
  both are classes at superordinate hierarchy levels

Does not imply

- absence of other land cover types
- absence or existence of a certain utilization

Can be supplemented by

- `leaf_type=*`
  leaf phenology, e. g. broadleaved, needle-leaved or mixed
- `leaf_cycle=*`
  seasonality of the trees, e. g. deciduous, evergreen or mixed
- `species=*`
  tree species dominant on the area
- `landuse=*`
  purpose of/ activities on the area, e. g. forestry, orchardry, military, leisure etc.
- `wetland=*`
  information on basic conditions (~=`yes`) or a special wetland type, e. g. swamp
- …

Compatibility & relations

OpenStreetMap

equals: –

is similar to:

- `natural=wood`
  mostly used for tree covered areas that are free from forestry operations

includes:

- `landuse=forest`
  mostly used for tree covered areas managed by forestry activities

is part of: –

LCCS

- simple: A12 A3A21
- considering, that the area might be managed/artificial: A12 A3A21 // A11 A1
- under wetland conditions: A24 A3A21
- if used for forestry: A11 A1S10
- if used for orchardry: A11 A1(S1//S2)
- if leaf phenology is mentioned: + D1 (broadleaved); + D2 (leadle-leafed)
- if seasonality is mentioned: + E1 (evergreen); + E2 (deciduous)

IV.3.2 Example *water body* (`landcover=water`)

This class is part of the land cover section in the OSM Land Classification system. It indicates a non-vegetated land surface and together with the class waterway it forms the superordinate class water. It contains further subdivisions, like lake, pond or reservoir, representing different types of standing water bodies.

Definition

Water bodies are areal features, that persist permanently or occur regularly and stay at least for one season. They are fed by precipitation, surface and sub-surface inflow, as well as by discharge from artificial sources. They're characterized by no or only minor currents and can basically be perceived as standing.

Explanation

This class is meant to be assigned to areas predominantly (> 70 %) covered by standing water. Water bodies can be permanent, intermittent or seasonal. Information on origin (artificial/natural) or utilization are not required for the application of this class.

When/ When not to use?

- delineating areas predominantly covered by water
- delineating areas that are regularly covered by water for the duration of a season
- further detailed description not possible or not intended
- delineating water bodies that are known to be standing but without having information about origin or utilization

Includes

- natural lakes
  recommended: + `water=lake`
- reservoirs
  recommended: + `water=reservoir`
- fish ponds
  recommended: + `water=pond` + `landuse=aquaculture`
- …

Excludes

- detention/retention ponds (short and sporadic use in case of heavy weather)
  - instead: `landcover=*` (except ~=`water`) + `landuse=basin`
  - optional: + `wetland=*`
- canals
  instead: `landcover=waterway` + `waterway=canal`
- areas of flowing water, like rivers, streams, ditches etc.
  instead: `landcover=waterway` + `waterway=*`
- riverine vegetation (e. g. reed beds)
  instead: `landcover=shrubs` + `wetland=reedbed`
- …

How to map

The use of this class is limited to areas. Boundaries are drawn along distinct borders (shorelines, reinforcements etc.). Vegetated areas are excluded once coverage reaches open to close conditions (> 40%). In case of water level fluctuations, the lower shoreline should be traced. The area above, up to the seasonal maximum, can either by marked as wetland or as seasonal/intermittent water surface (remains to be discussed).

As features, closed ways are the easiest to use. Direction of the way is irrelevant. If the area is too big or islands are enclosed, multi-polygon relations should be used.

As representative geometries, closed ways are the easiest to use to create the feature. Direction of the way is irrelevant. If the area is too big or islands have to be delineated, multi-polygon relations should be used.

Additional information

Implies

- `landcover=water-area`, `landcover=non-vegetated`
  both are classes at superordinate hierarchy levels

Does not imply

- absence or existence of a certain utilization

Can be supplemented by

- `seasonal=*`
  seasonality, e. g. dry season, wet season etc.
- `depth=*`
  maximum depth of the water body
- `landuse=*`
  purpose of/ activities on the area, e. g. aquaculture, leisure etc.
- …

Compatibility & relations

OpenStreetMap

equals: –
- natural=water

is similar to: –

includes:
- natural=water + water=lake
  natural lakes

- natural=water + water=pond
  small artificial natural water bodies
- natural=water + water=reservoir
  water bodies behind damming structures
- …

is part of: –

LCCS

- A27 A5 // A28 A5
- if depth is indicated: + C1 (> 2 m), +C2 (< 2 m)

IV.3.3 External and internal compatibility

Compatibility to the FAO LCCS is integrated at a simple level in tables 25 and 26. A detailed equivalence table is recommended to be compiled. It should also consider wetland and land use conditions.

Below, a template for compiling equivalence information for internal compatibility is presented. It shows the degree of similarity for the new classes compared with the current.

Table 27 Preliminary template for a tabular compilation of equivalences between the OSM LCS and the current system. For LCS equivalences of current OSM tags see table 30 in the land use section below.

OSM LCS	equals	is similar to	includes	is part of

vegetated	–	–	`landuse=grass`
			`leisure=garden`
			…
				–

woody	–	–	`natural=wood`
			`landuse=orchard`
			…
				–

trees	–	`natural=wood`
		`landuse=forest`
		…
			`natural=wood`
			`landuse=forest`
			…
				`natural=wood`
				`landuse=forest`
				…

wood	–	–	–	`natural=wood`
				`landuse=forest`
				…

woodland	–	–	–	`natural=wood`
				`landuse=forest`
				…

shurbs	…	…	…	…

Table 28 Tabular layout of the OpenStreetMap Land Use Classification System (primary and secondary sector). *Italic tags* are newly proposed or rarely used ([#] number of tagged features; as of August 24rd 2016 via taginfo.openstreetmap.org).

class			tag / short description	ISIC
primary sector				
	farming		*landuse=farming* [44] areas used for crop/animal/mixed farming; further subdivision impossible or not intended	011, 013–015
		crop farming	*landuse=crop_farming* areas used for cultivation of annual or biennial crops (cereals, vegetables, flowers, oilseeds etc.)	011
		animal farming	*landuse=animal_farming* [1] areas used for animal husbandry	014
		mixed farming	*landuse=mixed_farming* areas known to be used for both, crop and animal farming	011, 014
	orchardry		*landuse=orchardry* areas used for cultivation of perennial woody crops (fruits, vine, coffee, rubber, palm oil etc.)	012
	plant propagation		landuse=plant_nursery areas mainly used for open air propagation and sale of woody plants (trees, bushes etc.)	013
	forestry		*landuse=forestry* [6] areas used for growing, logging and managing timber	021, 022, 024
	aquaculture		landuse=aquaculture areas used for cultivating aquatic crops and animals (fish, shrimp, sea weed etc.)	032
	resource extraction		*landuse=extraction* areas used for extraction of mineral resources (ores, stone, clay, coal, oil, gas etc.)	05–09
secondary sector				
	energy production		*landuse=power* [8] areas hosting facilities for/ related to power generation	351–353
	industry & manufacturing		landuse=industrial areas hosting facilities for/ related to industrial processing and production (factories, refineries etc.)	10–33

IV.4 OSM Land Use Classification System

The following pages present the land use section of the OpenStreetMap Land Classification System. Tables 28 and 29 introduce the main classes for the primary, secondary and tertiary sector and present the according ISIC codes. Compatibility issues with the current OSM scheme are presented in chapter IV.4.2 (table 30). Chapter IV.4.1 contains the fundamental outline of the wiki-page for the land use class *education*. Similar to the chapter above: Hashed underlinings indicate proposed links to other pages of the OSM Wiki.

IV.4.1 Example *education* (landuse=education)

This class is part of the land use section in the OSM Land Classification system. As one of the classes provided for community services, it belongs to the major class tertiary sector.

Table 29 Tabular layout of the OpenStreetMap Land Use Classification System (tertiary sector and other land use types). *Italic tags* are newly proposed or rarely used ([#] number of tagged features; as of August 24rd 2016 via taginfo.openstreetmap.org).

class			tag / short description	ISIC
tertiary sector				
	transportation			
		railway transportation	`landuse=railway` areas used for/ associated to railway transportation (tracks, stations, depots etc.)	491
		road transportation	*`landuse=highway`* [3 141] areas used for/ associated to road transportation (roads, median strips, rest areas etc.)	492
		water transportation	*`landuse=shipping`* areas used for/ associated to water transportation (port facilities, watergates, terminals etc.)	50
		air transportation	*`landuse=aeroway`* [6] areas used for/ associated to air transportation (airport ground, runways, terminals etc.)	51
	commerce & business		`landuse=commercial` areas used for hosting activities and services related to trading, business administration etc.	–
		retail & wholesale	*`landuse=trade`* [4] areas used for trading business (shopping centres, malls, markets, wholesale facilities etc.)	45–47
		services	*`landuse=service`* [20] areas used for/ associated to businesses offering services, e. g. office districts	58–68
		hospitality businesses	*`landuse=hospitality`* areas used for/ associated to tourism, accommodation, food and beverage provision	55, 56
	community services			
		health care	*`landuse=health`*care [17] areas used for/ associated to hospitals, nurseries, rehabilitation centres etc.	86, 87
		education	*`landuse=education`* [422] areas used for/ associated to kindergartens, schools, colleges etc.	85
		public services	*`landuse=public`* [300] areas used for/ associated to community welfare, public administration, judiciary, police etc.	–
		religion	`landuse=religious` areas used for/ associated to religious purposes (churches, temples, monasteries etc.)	–
		military	`landuse=military` areas used for/ associated to military activities	–
	sport & recreation		*`landuse=leisure`* [490] areas used for hosting spare time activities (public swimming pools, golf courses, gardens etc.)	–
	dwelling & housing		`landuse=residential` areas used for/ associated to residence	–
miscellaneous				
	abandoned areas		*`landuse=abandoned`* [148]/ `abandoned:landuse=*` areas with signs of former activities; no current maintenance; plans for future utilization unknown	–
	construction sites		*`landuse=construction`* [148]/ `abandoned:landuse=*` areas used for constructing buildings, infrastructure etc.	37–69

Definition

Education comprises every kind of knowledge transfer for basic, leisure or professional purposes. Activities for and with people of every age and conditions are addressed.

Explanation

This class intends to delimit areas that host any kind of education facilities, including. It comprises kindergartens, schools, colleges, universities and training centres. The definition of this class does not separate between state and privately operated institutions.

When/ When not to use?

- delineating areas predominantly used for facilitating and supporting education activities
- delineating areas spatially related to education facilities, including areas used for non-educational purposes (administration, catering, parking etc.)

Includes

- kindergartens (incl. associated areas)
- schools (incl. associated areas)
- college and university campuses (incl. associated areas)
- professional training centres (incl. associated areas)
- …

Excludes

- libraries (incl. associated areas)
  instead: `landuse=leisure` (for the area) + `leisure=library` (for the building)
- …

How to map

The use of this class is limited to areas. Traceable boundaries are most clearly indicated by walls, hedges or fences. A closed way is used delimitation. In case of open boundaries within cities, information might be assessable about ownership or administration of adjacent green space, parking lots etc. If located within the traced areas, those objects are separately mapped (`amenity=parking`, `building=school` etc.)

Additional information

Does not imply

- `landcover=artificial`
  land cover can be diverse (trees/shrubs/grass, water bodies, buildings, paved areas etc.)
- …

Can be supplemented by

- `amenity=*`
  indicating the type of institution (school, university etc.)
- `name=*`
  name of the institution
- `isced:level=*`
  according the *International Standard Classification of Education* (OSM[68])
- …

Compatibility & relations

OpenStreetMap

equals: –

is similar to: –

includes:
- `amenity=school`/ `~=university`/ `~=kindergarten`
- …

is part of: –

ISIC

- 85
- in case of kindergarten and primary schools: 851
- in case of secondary education: 852
- in case of higher education: 853

IV.4.2 External and internal compatibility

Compatibility to the ISIC is integrated in tables 28 and 29. A detailed equivalence table is recommended to be compiled. The ISIC provides subclasses that become relevant if further information is attached to the area.

Below, a template for compiling equivalence information for internal compatibility is presented. In contrast to table 27, it shows how current tags could by expressed in the new system.

Table 30 Preliminary template for a tabular compilation of equivalences between the current tagging scheme and the proposed OSM LCS. For the opposite direction, see table 27 in the land cover section above.

object	current tag	OSM LCS
farmyard	`landuse=farmyard`	`landcover=built-up/ ~=pavement /~=*` `+ landuse=farming/ ~=mixed_farming/ ~=*`
university campus	`amenity=college`	`landcover=trees/ ~=shrubs/ ~=herbs /~=*` `+ landuse=education`
…	…	…

V Discussion

V.1 General remarks

Until formulating definitions, classes and tags of a new OSM Land Classification System, numerous information about the classification of land cover and land use have been systematically collected by this study. The aim was the provision of a basis for arranging a new system in an objective and comprehensible way. It was supposed to support a final definition of classes that allow an objective mapping and classification procedure.

Finally, the objectivity of the process and of the classes remain limited. In fact, objective observation requires measuring *data* that are – separately looked at – of low importance for our mapping endeavours. It is their combination and interpretation that makes them turn into meaningful *information* (Comber et al. 2005). The classification system presented above tries to apply concepts that require only few information input. However, data in its measurable meaning play a minor role and the information needed for allocating an object remains an interpretation that is prone to subjective bias. In addition, the direct dealing with concise tags is well established. The actual information required or transported according to the respective class definition, is therefore at risk to be replaced by the habituated perception of the tag. Consequently, robust definitions of classes and detailed information are important but can not assure an objective classification process in OpenStreetMap.

The misunderstandings between community members on how to apply current land cover tags are a good example of the fundamental problem. The author tried to summarize them in order to extract what might be exacted from a new system. The results are fundamentally important, but remain questionable because the scope of the study did only allow a very shallow investigation. A systematic analysis based on a quantitative approach, e. g. counts of key words or comments, would have produced more solid results; not to mention a dedicated survey among users with carefully developed questionnaires. Beside, the analysed sources (see table 21) are very likely not the only discussions related to *land cover/use* or the introduction of new tags related to these two concepts. The selection was conducted by only using the terms "landcover", "land cover", "landuse" and "land use" in the different platforms.

For identifying and addressing the demands related to a new classification system, a multitude of factors have to be considered (see chapter IV.1). Comparably, the development of a systems like the FAO LCCS, the IGBP or the LUCAS scheme have been accomplished by numerous stakeholders over several years of work. The scope of this study does only allow to develop a starting point and to provide preliminary results. Although completeness is a desirable key principle of a classification system, the state of the OSM LCS presented at the end of this study does not cover all possible properties, objects and features that do exist in reality or are represented by the current OSM tagging scheme. First of all, it had to inevitably emphasize on the most basic and popular objects. This bears the risk of the system to reveal crucial short comings once it is applied or extended.

It has to be generally stated, that the issues mentioned below are not based on practical experience with the new system. At the current state, the proposed OSM LCS has not been tested. In fact, implementing feedback from new and experienced mappers would be of essential importance for improving the system.

V.2 Improvements

Compared with the current system, the proposed OSM LCS allows a more objective mapping. Class definitions require to emphasize on the observation of rather basic aspects. This allows a simple description of small areas, e. g. covered by trees or shrubs, without any further connotations in respect of land use, ecology, climate etc. Similarly, large areas can be described in the way. Currently, the tag `natural=tundra` requires the mapper to consider geographical and climatic conditions. Additionally, the separation from `natural=fell` is difficult. In the new system, the areas are first of all described by `landcover=meadow` or `~=grassland` (depending on the degree of coverage), or more general as `landcover=herbs`.

Also different levels of detail allow areas to be described in a very basic way. A chosen class might represent less information, but at least it is not wrong. Contradictions with later assessments or other data sources are reduced.

A further emphasize has been put on conceptual consistency. Contradictions because of conflicting connotations are reduced. For example, `natural=water` is currently also used for artificial water bodies (supplemented by the according `water=*`-tag). The new system does not contain any connotation: A pond would be tagged as `landcover=water` + `water=pond`.

Consistency has also been improved for land use classes. Here, classes were preferably named according to activities instead of objects (e. g. `landuse=farming` instead of `~=farmland`).

Compared with the current system, the separation of land cover and land use allows a much better representation of their complex relationships. Varying land cover types within an area of homogeneous land use can be expressed – and vice versa. This allows, for instance, to illustrate a sequential development and changes, like *forest → clear cut → reforestation (shrub cover) → forest* without leaving any doubt about the persisting type of land use, forestry. Without the separation, analysing the data leaves room for misinterpretation based on presumed 1:1-relations between land cover and land use (forest → forestry; grass → animal husbandry; etc.)

The separation of the concepts may lead to the separation of features that have so far been undivided and perceived as being of the same type. But actually, this kind of separation is very common for the mapping of roads: One can belong to a certain hierarchy level, e. g. highway=primary, but continuously alter other properties (speed limits, surface material, street names etc.). Whenever a property changes, the feature is split and forms a new feature. Consequently, new procedures in the realm of land cover and land use are not necessarily new for the OpenStreetMap community.

Similarly, the new system arranges major and minor classes as being common practice in the current system. Preferably, current tags are integrated, e. g. `landcover=water` + `water=pond` (instead of `natural=water` + `water=pond`).

V.3 Limitations

The proposed OSM LCS strives for a clean separation of land cover and land use. The results are supposed to be more objective and provide a better compatibility with other classification systems, especially because of the transferability via LCCS codification. Both systems emphasize on the description of vegetation conditions. Nonetheless, for the FAO LCCS the information about an area being natural, cultivated/managed or artificial has to be known right at the start of the classification process. Consequently, the eight major classes of the FAO LCCSA contain both concepts, land cover and land use (see figure 1 and chapter III.1.1.2). In contrast, land cover information in the OSM LCS is intended to be free of these connotations. Allocating an area to the class *wood* does neither indicate natural nor artificial/managed conditions. It remains a pure description of the land cover, until a land use is indicated by a further attribute (e. g. `landuse=forestry`).

Consequently, an FAO LCCS equivalent of an OSM LCS class can only be represented by a combination of two codes. Two slashes are used by the FAO LCCS to indicate such either-or-conditions. For the above mentioned example, A11 A1 // A12 A3A10 would be the according codification. Similarly, the FAO LCCS requires an early decision on wetland conditions. At the current state of the OSM LCS, vegetation can be described without having this information. Consequently, an exact LCCS equivalent would have to consider this as well; again by using to land cover classes, combined with “//” (see also table 26 for exemplary application).

To state a practical example: An area delineated on a satellite image is obviously characterized by an open shrub cover. It has been allocated to the (new) class *shrubland* (`landcover=shrubland`). Information about the state of utilization is not available. The appropriate FAO LCCS code would be: A11 A2 // A12 A4A12 (either “cultivated & managed; shrubs” or “natural/semi-natural, terrestrial; shrubs; coverage 40–70%”). Later on, mapping activities on ground reveal, that the area is covered by blackberry shrubs which are still used. Correctly, the area is allocated to the land use class *orchardry* (`landuse=orchardry`) and supplemented by the information about the crop species (`trees=blackberry`). Once the land use information is available, the according FAO LCCS code would become explicit: A11 A2C1S2 (cultivated & managed; shrubs; single crop; fruit & nuts). The virtual OSM LCS class is named *open shrub orchard,* represented by the tags `landcover=shrubland + landuse=orchardry`.

As a solution, additional classes could be defined that are solely made for the purpose of providing a better compatibility with the FAO LCCS. They represent intentional combinations of land cover and land use information. Because they are not constitutive for the system, their names are not reflected by tags. Instead they are always codified by a `landcover`- and a `landuse`-tag. They could therefore be called *virtual classes*. They can act as a bridge between the new system and established OSM classes, where land cover and land use is often mixed as well.

Table 31 exemplarily compiles some *virtual classes* for a part of the OSM LCS. It can act as a template for systematically illustrating possible combinations of land cover and land use information. It also illustrates the relation between the current system, the new system and the FAO LCCS.

Similar to the lacking unambiguousness in terms of land use, the OSM LCS does also not explicitly distinguish between being or not being a wetland area. Instead, the pure land cover

Table 31 *Virtual classes* for combinations of land cover and land use conditions. In addition, equivalences to FAO LCCS and the current OSM tagging scheme are provided (“OSM”).

land use ▸ **▾ land cover**	**forestry** `landuse=forestry`	**orchadry** `landuse=orchardry`
trees `landcover=trees`	**forest** LCCS: A11 A1S10 OSM: `landuse=forest`	**tree orchard** LCCS: A11 A1(S1//S2) OSM: `landuse=orchard`
wood `landcover=wood`	**closed forest** LCCS: A11 A1S10 OSM: `landuse=forest`	**closed tree orchard** LCCS: A11 A1(S1//S2) OSM: `landuse=orchard`
woodland `landcover=woodland`	**open forest** LCCS: A11 A1S10 OSM: `landuse=forest`	**open tree orchard** LCCS: A11 A1(S1//S2) OSM: `landuse=orchard`
shrubs `landcover=shrubs`	**managed shrubs** LCCS: A11 A2S10 OSM: –	**shrub orchard** LCCS: A11 A2(S1//S2) OSM: `landuse=orchard`
thicket `landcover=thicket`	**managed thicket** LCCS: A11 A2S10 OSM: –	**closed shrub orchard** LCCS: A11 A2(S1//S2) OSM: `landuse=orchard`
shrubland `landcover=shrubland`	**managed woodland** LCCS: A11 A2S10 OSM: –	**open shrub orchard** LCCS: A11 A2(S1//S2) OSM: `landuse=orchard`

information can mean both, either dry or wet. Both cases, land use and wetland conditions, have to be considered during the further development of the system. A distinct tag like `landuse=no` and `wetland=no`, respectively, could be discussed.

But not only compatibility to the FAO LCCS is limited. Maintaining consistency during the design of the new structure and the new classes is limited as well. *Non-vegetated areas* are not consistently subdivided according to land cover – as originally intended. The classes *water body* and *waterway* contain landscape features (*river, canal, lake* etc.). However, they are not directly transferred into a landcover-tag. Instead, the current water-tags are used (e. g. `landcover=water + water=lake`). The classes subsumed under *non built-up* also represent rather complex landscape features. Unlike the water feature, they are represented by distinct landcover tags (e. g. *extraction site:* `landcover=pit`). The land use section is not free from inconsistencies as well: For instance, the classes *road transportation* and *dwelling & housing* are inconsistently codified in two ways: Nouns and adjectives, as well as objects and activities/purposes are used for tag design – `landuse=highway` and `~=residential`, respectively.

Although declared as a key principle, naming rules are unlikely to become or remain consistent during further development of the OSM LCS. The chosen examples show, that neglecting consistency can even have the advantage of increasing compatibility with the current system. For instance, striving for consistency leads to the tag `landuse=orchardry`. It is very likely to be replaced by `landuse=orchard`. Though it indicates an object rather than an activity, but it is an established tag in the current system and has a similar meaning.

V.4 Unsettled issues

V.4.1 General mapping issues

Beside the classification and delineation guidelines for every single class, general mapping challenges remain to be clarified. During the further development of the system, the OSM community is responsible to discuss possible solutions. The tentative application of the system will be essential for this process.

Fundamentally relevant for the new system and for OpenStreetMap in general is the topic of the *minimum mapping unit (MMU),* because it is closely related to the aim of improving spatial consistency. An increase in detail carries the risk, that mapping results become more heterogeneous between different places, mappers and data sources. The author therefore supports a discussion about a general MMU. Guidelines should clarify, under which circumstances patches or strings of trees and grass are separately delineated as *trees* and *herbs,* or whether it is more appropriate to aggregate the area and assign the class *woodland.* Similarly: Should there be a minimum width of linear features, for them to mapped as areas? The official GPS error of about 8 meters was proposed above in chapter IV.2.3.3).

Another topic not settled in OSM does also affect the proposed system: Under which circumstances will land cover/use features be divided by a road, a pipeline or a stream? Is this a question of a minimum width as well? Should the decision be based on canopy cover: As long as the tree crowns are closed, the area remains undivided? Reaching an agreement within the community would be advantageous for the OSM LCS and the data quality in general.

An issue specifically relevant for the practical application of the system is the question of where tags should be assigned to. Should land cover and use information be attached to the same feature? Or shall there be two separate geometries describing the same area? This is an important issue for rendering algorithms that might need both information in order to assign the appropriate signature.

V.4.2 Mixed properties

During the community discussions it was criticized, that a new landcover-key would be misleading because it implies the absence of other species, e. g. in an area labelled `landcover=trees`. Beside actually being depended on the class definition, comments like this point to the issue of expressing vertical mixtures of land cover and – of course – land use types. The reality provides numerous cases in which the representation of likewise conditions appears reasonable: Wildlife crossings, green roofs, residential or commercial buildings that cover on-surface roads or railways, and further more. Not to mention the interest of recording understory vegetation.

The examples of agroforestry (combinations of forestry, crop and animal farming) point to one solution by providing distinct classes the combine the two properties (silvopastoralism, forest gardening etc.). Another solution could the application of namespaces: Similar to the LUCAS scheme, an LCx or LUx could indicate primary/secondary properties or could represent the vertical layering. For example, a silvopastoral areas could be represented by `land-`

`cover=woodland_grassland + landuse:LU1=forestry + landuse:LU2=animal_farming`. At the current state of the of the system, vertical layering has not been considered and the mentioned approach have to be discussed and tested.

Similarly, the representation of horizontal mixtures is also in a very preliminary state. The solution of stringing together different land cover types as proposed in chapter IV.2.3.2 could be one. However, it has to be discussed and tested whether this approach can be universally adapted for all other land cover/use classes.

V.4.3 Integration of complex concepts

Although the proposed system emphasizes to use simple concepts for describing land objects, more complex definitions are not wrong in a general sense. They just contain a more complex combination of properties and therefore need more information as an input. If those are available and the contributor is able to correctly interpret them, why not providing the possibility for contribution. Beside the pure land cover description, landscape elements, geomorphological forms, ecoregions or biotopes could also be part of the OSM data, probably indicated by the key `natural`.

V.4.4 Rendering

Implications or graphical recommendations for the rendering of separated land cover/information have not been provided yet. Requirements of algorithms might have to be considered during the further design of the system or for its practical application (see chapter V.4.1).

A more fundamental question deals with the prioritization and generalization of the displayed information. For rural areas it is probably more important for a map to display objects and conditions that can be seen in reality as well. Rendering land cover would therefore support visible orientation. In contrast, residential, industrial or commercial areas have very different appearances that might not be sufficiently reflected by land cover information (only `landcover=built-up` might have been used). Thus, the illustration of land use might be more usable for larger built-up areas, except enclosed parks or water bodies. The spatial definition of the two rendering priorities is finally a technical problem to be solved.

V.4.5 Practical application

As of the current state of the study, no practical application has been conducted yet. Consequently, further fundamental and minor shortcomings are likely to turn up once the system is applied. At the same time, it will be proved, whether the above mentioned limitations are actual constraints, and whether the presumed improvements do actually evolve.

The practical application of the system will reveal, whether an implementation of simple basic concepts does really lead to a simple mapping and classification process. Given the fact, that land objects have to be assigned to two classes in order to represent their basic properties, the use of current single classes seem to remain more attractive. Therefore, a collection of 1 : 1-relations between might be helpful: For instance, for areas labelled with `amenity=parking` a sealed or compacted surface could be automatically implied, i. e. `land-`

`cover=pavement`. Similarly, `amenity=school` could imply `landuse=education`. This way, extraction of land cover/use information can still be possible without bein explicitly described by the data. An active contribution of the two information could then first of all be restricted to cases without a clear 1 : 1-relation.

A further aspect of successfully applying the new system is the provision of prototypes in form of pictures. The proposals for the wiki pages (see chapters IV.3.1, IV.3.2 and IV.4.1) do not yet contain them. In practice, they are part of the class description and are a very crucial measure of orientation for mappers. Photographs of different landscapes in varying geographical and climatological settings have to be collected and accordingly classified.

VI Conclusion

The aim of this thesis was to create a starting point for a new way of classifying land cover and land use in the OpenStreetMap project. Common classes have been arranged in a new hierarchical structure; templates for wiki-pages and compatibility tables have been developed. For this, there have been three driving factors: Reducing misconceptions, improving data quality and creating compatibility with established classification systems.

The thesis did not fully reach all of these aims. Due to the complexity of the settings to be considered, this had to be accepted somehow. Even the practical recommendation were intended to be preliminary. The state of the OSM LCS at this point is meant as contribution that may revive the discussion on mapping land cover and land in OSM and will hopefully lead to new proposal aiming on establishing a new system. The introduction of relations to the LCCS and ISIC can provide further inspiration.

Beside the practical but preliminary results, the study has reached another aim: In order to develop these preliminary components, the complex constellation of motivations and requirements is now better understood. In this respect, the thesis finally provides valuable information, especially in form of crucial demands and requirements. Disregarding the final structure of a new system, these aspects have to be considered and are now comprehensively formulated.

However, at the current state of the OSM LCS, those requirements are not entirely implemented. And probably, they will never be. This is because a very basic conflict will remain difficult to be solved: Purely focussing on a system that works best for OSM vs. emphasizing on reaching compatibility to other systems, e. g. the LCCS. Although the set of demands and requirements contains efforts of harmonizing these aspects, full compliance with the needs on both sides is unlikely to be reached. Even in case of achieving basic compromises, minor incompatibilities will remain because some classes can't be translated to their full extent of definitional detail.

Despite these difficulties related to a reform of the current classification system, it is a direction worth to head for. Improving external compatibility and internal usability will help to tap OpenStreetMap's full potentials.

VII References

Anderson, J. R., Hardy, E. E., Roach, J. T. & Witmer, R. E. (1976). A land use and land cover classification system for use with remote sensor data. USGS Professional Paper 964.

Arino, O., Leroy, M., Ranera, F., Gross, D., Bicheron, P., Nino, F., Brockmann, C., Defourney, P., Vancutsem, C., Achard, F., Durieux, L., Bourg, L., Latham, J. S., Di Gregorio, A., Witt, R., Herold, M., Sambale, J., Weber, J.-L., Goryl, P. & Houghton, N. (2007). GlobCover: A global land cover services with MERIS. ESA ENVISAT Symposium, Montreux (Switzerland). 6 pp.

Arino, O., Perez, J. R., Kalogirou, V., Defourney, P. & Achard, F. (2010). GlobCover 2009. ESA Living Planet Symposium (Bergen, Norway). 3 pp.

Bakker, A. M. & Veldkamp, A. (2008). Modelling land change: The issue of use and cover in wide-scale applications. *Journal of Land Use Science*; vol. 3, pp. 200–213.

Ballatore, A. & Bertolotto, M. (2011). Semantically enriching VGI in support of implicit feedback analysis. – In: Tanaka, K., Fröhlich, P. & Kim, K.-S. (eds.). Web and Wireless Geographical Information Systems. Berlin; pp. 78–93.

Ban YF., Peng G. & Giri, C. P. (2015). Global land cover mapping using Earth observation satellite data: Recent progresses and challenges. *ISPRS International Journal of Geo-Information;* vol. 103, pp. 1–6.

Bartholomé, E. & Belward, A. S. (2005). GLC2000: A new approach to global land cover mapping from Earth observation data. *International Journal of Remote Sensing;* vol. 26, pp. 1959–1977.

Belward, A. E. (1996). The IGBP-DIS global 1 km land cover data set "DISCover": Proposal and implementation plans. Stockholm; 63 pp. [IGBP Working Paper, No. 13]

Bishr, M. F. & Mantelas, A. (2008). A trust and reputation model for filtering and classification of knowledge about urban growth. *GeoJournal;* vol. 72, pp. 229–237.

Bjelland, T. C. (2004). Classification: Assumptions and Implications for Conceptual Modeling. PhD thesis; University of Bergen; 240 pp.

Bono, F. & Gutiérrez, E. (2011). A network-based analysis of the impact of structural damage on urban accessibility following a disaster: The case of the seismically damaged Port Au Prince and Carrefour urban road networks. *Journal of Transport Geography;* vol. 19, pp. 1443–1455.

Bontemps, S., Defourney, P., van Bogaert, E., Arino, O., Kalogrou, V. & Perez, J. R. (2011). GLOBCOVER 2009: Products Description and Validation Report. 53 pp.

Budhathoki, N., Bruce, B. & Nedovic-Budic, Z. (2008). Reconceptualizing the role of the user of spatial data infrastructure. *GeoJournal;* vol. 72, pp. 149–160.

Budhathoki, N. (2010). Participants' Motivations to Contribute to Geographic Information in an Online Community. PhD Thesis; University of Illinois; 121 pp.
URL: https://www.ideals.illinois.edu/bitstream/handle/2142/16956/1_Budhathoki_Nama.pdf (as of August 28th,2016)

Budhathoki, N. (2013). Motivation for open collaboration: Crowd and community models and the case of OpenStreetMap. *American Behavioral Scientist;* vol. 57, pp. 548–575.

Burley, T. M. (1961). Land use or land utilization? *Professional Geographer;* vol. 13, pp. 18–20.

BÜTTNER, G. & EISELT, B. (2013). LUCAS and CORINE Land Cover.
URL: http://ec.europa.eu/eurostat/documents/205002/274769/LUCAS_UseCase-CLC.pdf (as of May 25th, 2016)

CLAWSON, M. & STEWART, C. L. (1965). Land Use Information: A Critical Survey on U.S. Statistics Including Possibilities for Greater Uniformity. Baltimore; 420 pp.

COLEMAN, D., GEORGIADOU, Y. & LABONTE, Y. (2009). Volunteered geographic information: The nature and motivation of producers. *International Journal of Spatial Data Infrastructures Research;* vol. 4, pp. 332–358.

COMBER, A. J., FISHER, P. F. & WADSWORTH, R. A. (2005). What is land cover? *Environment and Planning B;* vol. 32, pp. 199–209.

COMBER, A. J. (2008a). Land use or land cover? *Journal of Land Use Science;* vol. 3, pp. 199–201.

COMBER, A. J. (2008b). The separation of land cover from land use using data primitives. *Journal of Land Use Science;* vol. 3, pp. 215–229.

COMBER, A. J., WADSWORTH, R. A. & FISHER, P. F. (2008). Using semantics to clarify the conceptual confusion between land cover and land use: The example of 'forest'. *Journal of Land Use Science;* vol. 3, pp. 185–198.

CONGALTON, R. G., GU JY., YADAV, K., THENKABAIL, P. & OZDOGAN, M. (2014). Global land cover mapping: A review and uncertainty analysis. *Remote Sensing;* vol. 6, pp. 12070–12093.

CURRAN, K., CRUMLISH, J. & FISHER, G. (2012). OpenStreetMap. *International Journal of Interactive Communication Systems and Technologies;* vol. 2, pp. 69–78.

DE FRIES, R. S., HANSEN, M. C. TOWNSHEND, J. R. G. & SOHLBERG, R. (1998). Global land cover classifications at 8 km spatial resolution: The use of training data derived from Landsat imagery in decision tree classifiers. *International Journal of Remote Sensing;* vol. 19, pp. 3141–3168.

DEFOURNEY, P. & BONTEMPS, S. (2012). Revisiting land-cover mapping concepts. – In: GIRI, C. P. (ed.). Remote Sensing of Land Use and Land Cover: Principles and Applications. Boca Raton et al.; pp. 49–63.

DEFOURNEY, P., KIRCHES, G., BROCKMANN, K., BOETTCHNER, M., PETERS, M., BONTEMPS, S., LAMARCHE, C., SCHLERF, M. & SANTORO, M. (2016). Land Cover CCI: Product User Guide, Version 2. ed. 2.5; 91 pp.
URL: http://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-PUG-v2.5.pdf (as of August 29th, 2016)

DI GREGORIO, A. (2005). Land Cover Classification System: Classification Concepts and User Manual (Software Version 2). Rome; 190 pp.

DI GREGORIO, A. & O'BRIEN, D. (2012). Overview of land-cover classifications and their interoperability. – In: GIRI, C. P. (ed.). Remote Sensing of Land Use and Land Cover: Principles and Applications. Boca Raton et al.; pp. 33–47.

DUHAMEL, C. (2009). Land use and land cover, including their classification. – In: VERHEYE, W. H. (ed.). Land Cover, Land Use and the Global Change. Oxford; pp. 80–100.

EITEN, G. (1968). Vegetation Forms: A classification of stands of vegetation based on structure, growth form of the components, and vegetative periodicit. Sao Paolo, 67 pp.

ELLIS, E. E. (2006). Land-use and land-cover change. – In: Encyclopedia of Earth.
URL: http://www.eoearth.org/view/article/154143/ (as of April 24th, 2016)

ELWOOD, S. (2006). Citizens as sensors: The world of volunteered geography. *Transactions in GIS;* vol. 10, pp. 693–708.

ELWOOD, S. (2010). Geographic information science: Emerging research on the societal implications of the geospatial web. *Progress in Human Geography;* vol. 34, pp. 349–357.

Esch, T., Marconcini, M., Felbier, A., Roth, T., Heldens, W., Huber, M., Schwinger, M., Taubenböck, H., Müller, A. & Dech, S. (2013). Urban Footprint Processor: Fully automated processing chain generating settlement masks from global data of the TanDEM-X mission. *IEEE Geoscience and Remote Sensing Letters;* vol. 10, pp. 1617–1621.

EUROSTAT (2015a). LUCAS 2015: Technical reference document C3. Classification (Land Cover & Land Use). 93 pp.
URL: http://ec.europa.eu/eurostat/documents/205002/6786255/LUCAS2015-C3-Classification-20150227.pdf/969ca853-e325-48b3-9d59-7e86023b2b27 (as of May 27th, 2016)

EUROSTAT (2015b). LUCAS 2015: Technical reference document C1. Instruction for surveyors. 140 pp.
URL: http://ec.europa.eu/eurostat/documents/205002/6786255/LUCAS2015-C1-Instructions-20150227.pdf/bbc63453-568f-44fc-a149-8ef6b04626d7 (as of May 27th, 2016)

Fairbairn, D. & Al-Bakri, M. (2013). Using geometric properties to evaluate possible integration of authoritative and volunteered geographic information. *ISPRS International Journal of Geo-Information;* vol. 2, pp. 349–370.

Fisher, P., Comber, A.J. & Wadsworth, R.A. (2005). Land use and land cover: Contradiction or complement? – In: Fisher, P. & Unwin, D.J. (eds.). Re-presenting GIS. Chichester; pp. 85–98.

FRA 2000 – Forest Resource Assessment Programme (2000). Forest Cover Mapping & Monitoring with NOAA-AVHRR & other Coarse Spatial Resolution Sensors. Rome; 41 pp.

Friedl, M.A., Sulla-Menashe, D., Tan B., Schneider, A., Ramankutty, N., Sibley, A. & Huang XM. (2010). MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment;* vol. 114, pp. 168–182.

Gamba, P. & Lisini, G. (2013). Fast and efficient urban extent extraction using ASAR Wide Swath Mode data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing;* vol. 6, pp. 2184–2195.

Gerlach, J. (2010). Vernacular mapping, and the ethics of what comes next. *Progress in Human Geography;* vol. 38, pp. 165–168.

Gerlach, J. (2013). Lines, contours and legends: Coordinates for vernacular mapping. *Cartographica;* vol. 45, pp. 22–39.

Girres, J.-F. & Touya, G. (2010). Quality assessment of the French OpenStreetMap dataset. *Transactions in GIS;* vol. 14, pp. 435–459.

Gong P., Wang J., Yu L., Zhao YC., Zhao YY., Liang L. Niu ZG., Huang XM., Fu HH., Liu S., Li CC., Li XY., Fu W., Liu CX., Xu Y., Wang XY., Cheng Q., Hu LY., Yao WB., Zhang H., Zhu P., Zhao ZY., Zhang HY., Zheng YM., Ji LY., Zhang YW., Chen H., Yan A., Guo JH., Yu L., Wang L., Liu XJ., Shi TT., Zhu MH., Chen YL., Yang GW., Tang P., Xu B., Giri, C., Clinton, N., Zhu ZL., Chen J. & Chen J. (2013). Finer resolution observation and monitoring of global land cover: First mapping results with Landsat TM and ETM+ data. *International Journal of Remote Sensing;* vol. 34, pp. 2607–2654.

Gong P., Yu L., Li CC., Wang J., Liang L., Li XC., Li LY. & Bai YQ. (2016). A new research paradigm for global land cover mapping. *Annals of GIS;* vol. 22, pp. 87–102.

Goodchild, M.F. (2007). Citizens as sensors: The world of volunteered geography. *GeoJournal;* vol. 69, pp. 211–221.

Goodchild, M.F. (2008). Assertion and authority: The science of user-generated geographic content. – In: Navratil, G. (2008). Proceedings of the Colloquium for Andrew U. Frank's 60th Birthday. Wien; pp. 1–18.

Goodchild, M.F. & Glennon, J.A. (2010). Crowdsourcing geographic information for disaster response: A research frontier. *International Journal of Digital Earth;* vol. 3, pp. 231–241.

Gröchenig, S., Brunauer, R. & Rehrl, K. (2014a). Digging into the history of VGI data-sets: Results from a worldwide study on OpenStreetMap mapping activity. *Journal of Location Based Services;* vol. 8, pp. 198–210.

Gröchenig, S., Brunauer, R. & Rehrl, K. (2014b). Estimating completeness of VGI datasets by analyzing community activity over time periods. – In: Huerta, J., Schade, S. & Granell, C. (eds.). Connecting a Digital Europe Through Location and Place. pp. 3–18.

Haklay, M., Singleton, A. & Parker, C. (2008). Web mapping 2.0: The neogeography of the GeoWeb. *Geography Compass;* vol. 2, pp. 2011–2039.

Haklay, M., Basiouka, S., Antoniou, V. & Ather, A. (2010). How many volunteers does it take to map an area well? The validity of Linus' Law to Volunteered Geographic Information. *Cartographic Journal;* vol. 47, pp. 315–322.

Hansen, M.C. & Reed, B. (2000a). A comparison of the IGBP DISCover and University of Maryland 1 km global land cover products. *International Journal of Remote Sensing;* vol. 21, pp. 1365–1373.

Hansen, M.C., de Fries, R.S., Townshend, J.R.G. & Sohlberg, R. (2000b). Global land cover classification at 1 km spatial resolution using a classification tree approach. *International Journal of Remote Sensing;* vol. 21, pp. 1331–1364.

Heipke, C. (2010). Crowdsourcing geospatial data. *ISPRS Journal of Photogrammetry and Remote Sensing;* vol. 65, pp. 550–557.

Horita, F. & Albuquerque, J.P. (2013). An approach to support decision-making in disaster management based on volunteer geographic information (VGI) and spatial decision support systems (SDSS). – In: Proceedings of the 10th International Conference on Information Systems for Crisis Response and Management (Baden-Baden, Germany, 12–15 May 2013).

Hudson-Smith, A., Batty, M., Crooks, A. & Milton, R. (2009). Mapping tools for the masses: Web 2.0 and crowdsourcing. *Social Science Computer Review;* vol. 27, pp. 524–538.

Jacob, E.K. (2004). Classification and categorization: A difference that makes a difference. *Library Trends;* vol. 52, pp. 515–540.

Jansen, L.J.M. (2006). Harmonisation of land-use class sets to facilitate compatibility and comparability of data across space and time. *Journal of Land Use Science;* vol. 1, pp. 127–156.

Javanmardi, S., Ganjisaffar,Y., Lopes, C. & Baldi, P. (2009). User contribution and trust in Wikipedia. – In: Proceedings of the 5th International Conference on Collaborative Computing: Networking, Applications and Worksharing (Washington, DC, USA, 11–14 November 2009).

Kuechler, A.W. & Zonneveld, I.S. (eds.; 1988). Vegetation Mapping: Handbook of Vegetation Science. Dordrecht.

LDWG/LUP – Interdepartmental Working Group on Land Use Planning (FAO) Y. (1994). Glossary of Land Use Terms. Rome

Lechner, M. (2011). Nutzungspotentiale crowdsource-erhobener Geodaten auf verschiedenen Skalen. PhD Thesis; University Freiburg.
URL: https://www.freidok.uni-freiburg.de/data/8181 (as of August 28th, 2016)

Lin W. (2014). Revealing the making of OpenStreetMap: A limited account. *The Canadian Geographer;* vol. 59, pp. 69–81.

Lin Y. (2011). A qualitative enquiry into OpenStreetMap making. *New Review of Hypermedia and Multimedia;* vol. 17, pp. 53–71.

Loveland, T.R. (2012). History of land-cover mapping. – In: Giri, C.P. (ed.). Remote Sensing of Land Use and Land Cover: Principles and Applications. Boca Raton et al.; pp. 13–22.

Loveland, T.R., Zhu ZL., Ohlen, D.O., Brown, J.F., Reed, B.C. & Yang LM. (1999). An Analysis of the IGBP global land-cover characterization process. *Photogrammetric Engineering and Remote Sensing*; vol. 65, pp. 1021–1032.

LOVELAND, T. R., REED, B. C., BROWN, J. F., OHLEN, D. O., ZHU Z., YANG L. & MERCHANT, J. W. (2000). Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing*; vol. 21, pp. 1303–1330.

MA D., SANDSBERG, M. & JIANG B. (2015). Characterizing the heterogeneity of the OpenStreetMap data and community. *ISPRS International Journal of Geo-Information;* vol. 4, pp. 535–550.

MAIER, G. (2014). OpenStreetMap, the Wikipedia Map. *Region;* vol. 1, pp. R3–R10.

MANFRÉ, L. A., HIRATA, E., SILVA, J. B., SHINOHARA, E. J., GIANOTTI, M. A., LAROCCA, A. P. C. & QUINTANILHA, J. A. (2012). An analysis of geospatial technologies for risk and natural disaster management. *ISPRS International Journal of Geo-Information;* vol. 1, pp. 166–185.

MARTÍNEZ, S. & MOLLICONE, D. (2012). From land cover to land use: A methodology to assess land use from remote sensing data. *Remote Sensing;* vol. 4, pp. 1024–1045.

MARTINO, L., PALMIERI, A. & GALLEGO, J. (2009). Use of auxiliary information in the sampling strategy of a European area frame agro-environmental survey.
URL: http://ec.europa.eu/eurostat/documents/205002/769457/LUCAS2009_S2-Sampling_20090000.pdf/cb9197df-d621-4436-bd0b-19f8da6b40bc (as of May 30th, 2016)

MCCALLUM, I., OBERSTEINER, M., NILSSON, S. & SHVIDENKO, A. (2006). A spatial comparison of four satellite derived 1 km global land cover datasets. *International Journal of Applied Earth Observation and Geoinformation;* vol. 8, pp. 246–255.

MEDIN, D. L. & AGUILAR, C. (1999). Categorization. – In: WILSON, R. A. & KEIL, F. C. (eds.). The MIT Encyclopedia of the Cognitive Sciences. Cambridge & Massachusetts; pp. 104–106.

MEINEL, G. & HENNERSDORF, J. (2002). Classification Systems of land cover and land use and their challenges for picture processing of remote sensing data: Status of international discussion and program. – In: Proceedings of the 3rd International Symposium Remote Sensing of Urban Areas (Instabul, 11–13 June 2002); pp. 472–479.

MOONEY, P. & CORCORAN, P. (2011). Can volunteered geographic information be a participant in eEnvironment and SDI? – In: Hrebícek, J., Schimak, G. & Denzer, R. (eds.). Environmental Software Systems: Frameworks of eEnvironment. Bosten; pp. 115–122.

MOONEY, P. & CORCORAN, P. (2012a). The annotation process in OpenStreetMap. *Transaction in GIS;* vol. 16, pp. 561–579.

MOONEY, P. & CORCORAN, P. (2012b). Who are the contributors to OpenStreetMap and what do they do? – In: Proceedings of 20th Annual GIS Research UK (Lancaster, UK, 11–13 April 2012).

MOONEY, P. & CORCORAN, P. (2012c). How social is OpenStreetMap? – In: Proceedings of the 15th Association of Geographic Information Laboratories for Europe International Conference on Geographic Information Science (Avignon, France, 24–27 April 2012).

NEIS, P., SINGLER, P. & ZIPF, A. (2010). Collaborative mapping and emergency routing for disaster logistics: Case studies from the Haiti earthquake and the UN portal for Africa. – In: Proceedings of the Geoinformatics Forum (Salzburg, Austria, 6–9 July 2010).

NEIS, P. & ZIPF, A. (2012). Analyzing the contributor activity of a Volunteered Geographic Information project: The case of OpenStreetMap. *ISPRS International Journal for Geo-Information;* vol. 1, pp. 146–165.

NEIS, P., ZIELSTRA, D. & ZIPF, A. (2013). Comparison of Volunteered Geographic Information data contributions and community development for selected world regions. *Future Internet;* vol. 5, pp. 282–300.

NEIS, P. & ZIELSTRA, D. (2014). Recent developments and future trends in volunteered geographic information research: The case of OpenStreetMap. *Future Internet;* vol. 6, pp. 76–106.

Nielsen, J. (2006). Participation Inequality: Encouraging more Users to Contribute. *Nielson Norman Group.*
URL: https://www.nngroup.com/articles/participation-inequality/ (as of August 28th, 2016)

Olson, J. S. (1994). Global Ecosystems Framework: Definitions. Sioux Falls. [USGS EROS Data Center: Internal Report]

OP – Office for Official Publications of the European Communities (2001). Manual of Concepts on Land Cover and Land Use Information Systems. Luxembourg; 106 pp.

Ostermann, F. & Spinsanti, L. A. (2011). Conceptual workflow for automatically assessing the quality of Volunteered Geographic Information for crisis management. – In: Proceedings of the 14th Association of Geographic Information Laboratories for Europe International Conference on Geographic Information Science (Utrecht, The Netherlands, 18–21 April 2011).

Poser, K. & Dransch, D. (2010). Volunteered Geographic Information for disaster management with application to rapid flood damage estimation. *Geomatica;* vol. 64, pp. 89–98.

Potthast, M. Stein, B. & Gerling, R. (2008). Automatic vandalism detection in Wikipedia. – In: Proceedings of the IR Research, 30th European Conference on Advances in Information Retrieval (Glasgow, Scotland, 30 March–3 April 2008).

Pultar, E., Raubal, M., Cova, T. J. & Goodchild, M. F. (2009). Dynamic GIS case studies: Wildfire evacuation and volunteered geographic information. *Transactions in GIS;* vol. 13, pp. 85–104.

Ramm, F. (2015). OpenStreetMap data in layered GIS format.
URL: https://www.geofabrik.de/data/geofabrik-osm-gis-standard-0.6.pdf (as of June 11th, 2016)

Ramm, F., Topf, J. & Chilton, S. (2010). OpenStreetMap: Using and Enhancing the Free Map of the World. Cambridge; – In: Krisp, A. (ed.). Progress in Location-Based Services: Lecture Notes in Geoinformation and Cartography. Berlin; 386 pp.

Rehrl, K., Gröchenig, S., Hochmaier, H. H., Leitinger, S., Steinmann, R. & Wagner, A. (2013). A conceptual model for analyzing contribution patterns in the context of VGI. – In: Krisp, A. (ed.). Progress in Location-Based Services: Lecture Notes in Geoinformation and Cartography. Berlin; pp. 373–388.

Roick, O. & Heuser, S. (2012). Location based social networks: Definition, current state of the art and research agenda. *Transactions in GIS;* vol. 17, pp. 763–784.

Rouse, L. J., Bergeron, S. J. & Harris, T. M. (2007). Participating in the geospatial web: Collaborative mapping, social networks and participatory GIS. – In: Scharl, A. & Tochtermann, K. (eds.). The Geospatial Web: How Geobrowsers, Social Software and the Web 2.0 are Shaping the Network Society. London et al.; pp. 153–158.

Sehra, S. S., Singh, J. & Rai, H. S. (2013). Assessment of OpenStreetMap data: A review. *International Journal of Computer Applications;* vol. 76, pp. 17–20.

Sieber, R. (2006). Public participation geographic information systems: A literature review and framework. *Annals of the American Association of Geographers;* vol. 96, pp. 491–507.

Song W. & Sun G. (2006). The role of mobile Volunteered Geographic Information in urban management. – In: Proceedings of 18th International Conference on Geoinformatics (Peking University, Beijing, China, 18–20 June 2010).

Stark, H.-J. (2010). Umfrage zur Motivation von Freiwilligen im Engagement in Open Geo-Data Projekten. – In: Proceedings of FOSSGIS Anwenderkonferenz für Freie und Open Source Software für Geoinformationssysteme (Osnabrück, Germany, 2–5 March 2010).

Stephens, M. (2013). Gender and the GeoWeb: Divisions in the production of user-generated cartographic information. *GeoJournal;* vol. 78, pp. 981–996.

Sui, D. Z. (2008). The wikification of GIS and its consequences: Or Angelina Jolie's new tattoo and the future of GIS. *Computers Environment and Urban Systems;* vol. 32, pp. 1–5.

TOWNSHEND, J.R.G., CIHLAR, J., JUSTICE, C.O., MALINGREAU, J.-P., RUTTENBERG, S., SADOWSKI, F., SKOLE, D. & TEILLET, P. (1992). Improved Global Data for Land Applications. Stockholm. [IGBP Report No. 20]

TSENDBAZAR, N.-E, DE BRUIN, S., FRITZ, S. & HEROLD, M. (2015). Spatial accuracy assessment and integration of global land cover datasets. *Science;* vol. 7, pp. 15801–15821.

TUCKER, C.J., TOWNSHEND, J.R.G. & GOFF, T.E. (1985). African land-cover classification using satellite data. *Remote Sensing;* vol. 227, pp. 369–375.

TURNER II, B.L. (1994). Global land-use and land-cover change: An overview. – In: MEYER, W.B. & TURNER II, B.L. (eds.). Changes in Land Use and Land Cover: A Global Perspective. Cambridge; pp. 3–10.

TURNER II, B.L., SKOLE, D., SANDERSON, S., FISCHER, G., FRESCO, L. & LEEMANS, R. (1995). Land-Use and Land-Cover Change: Science/Research Plan. Stockholm & Geneva; 132 pp.

UN-ESA – UN DEPARTMENT OF ECONOMIC AND SOCIAL AFFAIRS (2008). International Standard Industrial Classification of All Economic Activities (Rev. 4). New York; 291 pp.

UNCEEA – UN COMMITTEE OF EXPORT ON ENVIRONMENTAL ECONOMIC ACCOUNTING (2011). Revision of the System on Environmental-Economic Accounts (SEEA). 123 pp.

USDOD – US DEPARTMENT OF DEFENSE (2008). Global Positioning System, Standard Positioning System, Performance Standard. 4th ed.; Washington; 160 pp.

VELDKAMP, A., VERBURG, P.H., KOK, K., DE KONING, G.H.J., PRIESS, J. & BERGSMA, A.R. (2001). The need for scale sensitive approaches in spatially explicit land use change modelling. *Environmental Modelling and Assessment;* vol. 6, pp. 111–121.

VERBURG, P.H., NEUMANN, K. & NOL, L. (2011). Challenges in using land use and land cover data for global change studies. *Global Change Biology;* vol. 17, pp. 974–989.

VERHEYE, W.H. (2009). Land cover, land use and the global change. – In: VERHEYE, W.H. (ed.). Land Cover, Land Use and the Global Change. Oxford; pp. 1–45.

WELSER, D.M., COSLEY, D., KOSSINETS, G., LIN, A., DOKSHIN, F., GAY, G. & SMITH, M. (2011). Finding social roles in Wikipedia. – In: Proceedings of the 2011 iConference (Seattle, Washington).

WILKINSON, D.M. & HUBERMAN, B.A. (2007). Assessing the value of cooperation in Wikipedia. *First Monday;* vol. 12.
URL: http://firstmonday.org/article/view/1763/1643 (as of August 28th,2016)

ZIELSTRA, D. & ZIPF, A. (2010a). A comparative study of proprietary geodata and volunteered geographic information for Germany. – In: Proceedings of the Thirteenth AGILE International Conference on Geographic Information Science (Guimarães, Portugal).

ZIELSTRA, D. & ZIPF, A. (2010b). Quantiative studies on the data quality of OpenStreetMap in Germany. – In: FABRIKANT, S.I., REICHENBACHER, T., VAN KREVELD, M. & SCHLIEDER, C. (eds.). Geographic Information Science: Proceedings of GIScience 2010. Berlin; pp. 20–26.

ZIELSTRA, D., HOCHMAIER, H.H. & NEIS, P. (2013). Assessing the Effect of Data Imports on the Completeness of OpenStreetMap: A United States Case Study. *Transactions in GIS;* vol. 17, pp. 315–334.

Online Resources

LUCAS# Land Use/Cover Area Frame Survey

LUCAS1 LUCAS - Land use and land cover survey
URL: http://ec.europa.eu/eurostat/statistics-explained/index.php/LUCAS_-_Land_use_and_land_cover_survey (as of May 26th, 2016)

LUCAS2 What is LUCAS used for?
URL: http://ec.europa.eu/eurostat/web/lucas/publications/use-cases (as of May 26th, 2016)

LUCAS3 LUCAS micro data 2015
URL: http://ec.europa.eu/eurostat/web/lucas/data/primary-data/2015 (as of May 30th, 2016)

oDef# The Open Definition

oDef1 URL: http://opendefinition.org/ (as of April 7th, 2016)

OSM# OpenStreetMap

OSM1 About OpenStreetMap
URL: https://wiki.openstreetmap.org/wiki/about (as of March 15th, 2016)

OSM2 URL: https://www.openstreetmap.org/about (as of March 15th, 2016)

OSM3 Copyright & License
URL: https://www.openstreetmap.org/copyright (as of April 7th, 2016)

OSM4 History of OpenStreetMap
URL: https://wiki.openstreetmap.org/wiki/History_of_OpenStreetMap (as of March 18th, 2016)

OSM5 Legal
URL: https://wiki.openstreetmap.org/wiki/Legal (as of March 15th, 2016)

OSM6 Why OpenStreetMap?
URL: https://wiki.openstreetmap.org/wiki/Why_OpenStreetMap%3F (as of March 15th, 2016)

OSM7 Armchair mapping
URL: https://wiki.openstreetmap.org/wiki/Armchair_mapping (as of April 7th, 2016)

OSM8 Open Database License
URL: https://wiki.openstreetmap.org/wiki/Open_Database_License (as of April 9th, 2016)

OSM9 TIGER
URL: https://wiki.openstreetmap.org/wiki/TIGER (as of April 7th, 2016)

OSM10 AND Data
URL: https://wiki.openstreetmap.org/wiki/AND_Data (as of April 7th, 2016)

OSM11 Humanitarian OSM Team
URL: https://wiki.openstreetmap.org/wiki/Humanitarian_OSM_Team (as of April 7th, 2016)

OSM12 OpenStreetMap stats report
URL: https://www.openstreetmap.org/stats/data_stats.html (as of August 28th, 2016)

OSM13 Pick your data collection technique
URL: https://wiki.openstreetmap.org/wiki/Pick_your_data_collection_technique (as of April 9th, 2016)

OSM14 Changeset
URL: https://wiki.openstreetmap.org/wiki/Changeset (as of April 9th, 2016)

OSM15 Editing
URL: https://wiki.openstreetmap.org/wiki/Editing (as of April 9th, 2016)
OSM16 Tags
URL: https://wiki.openstreetmap.org/wiki/Tags (as of April 9th, 2016)
OSM17 Automated Edits code of conduct
URL: https://wiki.openstreetmap.org/wiki/Automated_Edits/Code_of_Conduct (as of April 5th, 2016)
OSM18 API
URL: https://wiki.openstreetmap.org/wiki/API (as of April 9th, 2016)
OSM19 Overpass API
URL: https://wiki.openstreetmap.org/wiki/Overpass_API (as of April 9th, 2016)
OSM20 Frameworks
URL: https://wiki.openstreetmap.org/wiki/Frameworks (as of April 9th, 2016)
OSM21 Map Features
URL: https://wiki.openstreetmap.org/wiki/Map_Features (as of April 5th, 2016)
OSM22 Features
URL: https://wiki.openstreetmap.org/wiki/Features (as of June 8th, 2016)
OSM23 Verifiability
URL: https://wiki.openstreetmap.org/wiki/Verifiability (as of June 8th, 2016)
OSM24 Revision history of "Key:landuse"
URL: https://wiki.openstreetmap.org/w/index.php?title=Key:landuse&action=history (as of June 11th, 2016)
OSM25 Landuse
URL: https://wiki.openstreetmap.org/wiki/Landuse (as of June 8th, 2016)
OSM26 Landcover
URL: https://wiki.openstreetmap.org/wiki/Landcover (as of June 8th, 2016)
OSM27 Editing Standards and Conventions
URL: https://wiki.openstreetmap.org/wiki/Editing_Standards_and_Conventions (as of June 10th, 2016)
OSM28 Namespace
URL: https://wiki.openstreetmap.org/wiki/Namespace (as of June 11th, 2016)
OSM29 Key:generator:type
URL: https://wiki.openstreetmap.org/wiki/Key:generator:type (as of June 10th, 2016)
OSM30 Lifecycle prefix
URL: https://wiki.openstreetmap.org/wiki/Lifecycle_prefix (as of June 10th, 2016)
OSM31 Key:abandoned:
URL: https://wiki.openstreetmap.org/wiki/Key:abandoned: (as of June 10th, 2016)
OSM32 Multilingual names
URL: https://wiki.openstreetmap.org/wiki/Multilingual_names (as of June 10th, 2016)
OSM33 OpenStreetMap Forum
URL: http://forum.openstreetmap.org/ (as of June 15th, 2016)
OSM34 OpenStreetMap help
URL: https://help.openstreetmap.org/ (as of June 15th, 2016)
OSM35 OSM Mailing Lists
URL: https://lists.openstreetmap.org/listinfo (as of June 15th, 2016)
OSM36 Just Map
URL: https://wiki.openstreetmap.org/wiki/Just_Map (as of June 13th, 2016)

OSM37 Any tags you like
URL: https://wiki.openstreetmap.org/wiki/Any_tags_you_like (as of June 13th, 2016)

OSM38 Proposal process
URL: https://wiki.openstreetmap.org/wiki/Proposal_process (as of June 13th, 2016)

OSM39 Landuse
URL: https://wiki.openstreetmap.org/w/index.php?title=Landuse&oldid=713970 (as of June 14th, 2016)

OSM40 Proposed features/landcover
URL: https://wiki.openstreetmap.org/wiki/Proposed_features/landcover (as of June 14th, 2016)

OSM41 Proposed features/landcover
URL: https://wiki.openstreetmap.org/w/index.php?title=Proposed_features/landcover&oldid=560186 (as of June 14th, 2016)

OSM42 Talk:Proposed features/landcover
URL: https://wiki.openstreetmap.org/wiki/Talk:Proposed_features/landcover (as of June 14th, 2016)

OSM43 User:Rudolf/draft landcover
URL: https://wiki.openstreetmap.org/wiki/User:Rudolf/draft_landcover (as of June 14th, 2016)

OSM44 User:Rudolf/draft landcover woodland
URL: https://wiki.openstreetmap.org/wiki/User:Rudolf/draft_landcover_woodland (as of June 15th, 2016)

OSM45 Environment
URL: https://wiki.openstreetmap.org/wiki/Environment (as of June 15th, 2016)

OSM46 Natural
URL: https://wiki.openstreetmap.org/wiki/Natural (as of June 16th, 2016)

OSM47 Befestigte Flächen mappen
URL: http://forum.openstreetmap.org/viewtopic.php?id=53029 (as of June 15th, 2016)

OSM48 Das Märchen vom landuse
URL: http://forum.openstreetmap.org/viewtopic.php?id=18437 (as of June 15th, 2016)

OSM49 Landuse
URL: http://forum.openstreetmap.org/viewtopic.php?id=54714 (as of June 15th, 2016)

OSM50 landuse=forest or natural=wood?
URL: http://forum.openstreetmap.org/viewtopic.php?id=10204 (as of June 15th, 2016)

OSM51 Österreichs Almböden
URL: http://forum.openstreetmap.org/viewtopic.php?id=53010 (as of June 15th, 2016)

OSM52 How to tag green (vegetated) areas within urban areas?
URL: https://help.openstreetmap.org/questions/8752/how-to-tag-green-vegetated-areas-within-urban-areas (as of June 15th, 2016)

OSM53 Mapping landcover in a park: landuse vs. landcover vs. natural
URL: https://help.openstreetmap.org/questions/49784/mapping-landcover-in-a-park-landuse-vs-landcover-vs-natural (as of June 15th, 2016)

OSM54 Overlapping land use, cf residential and forest
URL: https://help.openstreetmap.org/questions/17874/overlapping-land-use-cf-residential-and-forest (as of June 15th, 2016)

OSM55 Poplar plantation: how to map?
URL: https://help.openstreetmap.org/questions/45251/poplar-plantation-how-to-map (as of June 15th, 2016)

OSM_{56} Should pastures be marked as farm or grass?
URL: https://help.openstreetmap.org/questions/366/should-pastures-be-marked-as-farm-or-grass (as of June 15th, 2016)

OSM_{57} When should we use landuse=forest rather than natural=wood?
URL: https://help.openstreetmap.org/questions/324/when-should-we-use-landuseforest-rather-than-naturalwood (as of June 15th, 2016)

OSM_{58} [Tagging] Defining tag 'natural=wood' ...
URL: https://lists.openstreetmap.org/pipermail/tagging/2016-February/028397.html (as of June 15th, 2016)

OSM_{59} [Tagging] Landuse=forestRY?
URL: https://lists.openstreetmap.org/pipermail/tagging/2015-December/027919.html (as of June 15th, 2016)

OSM_{60} [Tagging] landcover=trees definition
URL: https://lists.openstreetmap.org/pipermail/tagging/2015-August/025848.html (as of June 15th, 2016)

OSM_{61} Automated Edits code of conduct
URL: https://wiki.openstreetmap.org/wiki/Automated_Edits_code_of_conduct (as of June 20th, 2016)

OSM_{62} Elements
URL: https://wiki.openstreetmap.org/w/index.php?title=Elements&oldid=38094 (as of June 20th, 2016)

OSM_{63} User:Joto/How to invent tags
URL: https://wiki.openstreetmap.org/wiki/User:Joto/How_to_invent_tags (as of June 15th, 2016)

OSM_{64} Key:surface
URL: https://wiki.openstreetmap.org/wiki/Key:surface (as of August 11th, 2016)

OSM_{65} Accuracy
URL: https://wiki.openstreetmap.org/wiki/Accuracy (as of August 23th, 2016)

OSM_{66} Tag:natural=fell
URL: https://wiki.openstreetmap.org/wiki/Tag:natural%3Dfell (as of August 23th, 2016)

OSM_{67} Tag:natural=tundra
URL: https://wiki.openstreetmap.org/wiki/Tag:natural%3Dtundra (as of August 23th, 2016)

OSM_{68} Proposed features/ISCED
URL: https://wiki.openstreetmap.org/wiki/Proposed_features/ISCED (as of August 25th, 2016)

$OSMF_{\#}$ OpenStreetMap Foundation

$OSMF_{1}$ OpenStreetMap Foundation
URL: https://wiki.osmfoundation.org/wiki/Main_Page (as of April 7th, 2016)

$OSMF_{2}$ Licence
URL: https://wiki.osmfoundation.org/wiki/License (as of April 9th, 2016)

$OSMF_{3}$ About
URL: https://blog.openstreetmap.org/about/ (as of April 8th, 2016)

$plOSM_{\#}$ Planet OSM

$plOSM_{1}$ Planet OSM
URL: http://planet.openstreetmap.org/ (as of August 29th, 2016)

APPENDIX

Table A1 Tags for horizontal mixture of land cover types, exemplary for trees and shrubs.

TREES	tags acc. to OSM LCS				codes acc. FAO LCCS			
	closed	open		sparse	closed	open		sparse
		dense	loose			dense	loose	
	> 70%	40 - 70%	20 - 40%	< 20	> 70%	40 - 70%	20 - 40%	< 20
shrubs								
> 70%	X	X	X	*=thicket + coverage:trees=sparse	X	X	X	A12 A4A10 // A12 A3A14
40 – 70%	X	*=woodland_shrubland or *=shrubland_woodland		*=shrubland + coverage:trees=sparse	X	A12 A1A21 A12 A3A11 // A12 A4A11 (A12 A3A12 // A12 A4A13 or A12 A4A12 // A12 A3A13)		A12 A4A12 // A12 A3A14
20 – 40%	X			X	X			X
< 20%	*=forest + coverage:shrubs=sparse	*=woodland + coverage:shrubs=sparse	X	X	A12 A3A10 // A12 A4A14	A12 A3A12 // A12 A4A14	X	X
herbs								
> 70%	X	X	X	*=meadow + coverage:trees=sparse	X	X	X	A12 A2A10 // A12 A3A14
40 – 70%	X	*=woodland_grassland or *=grassland_woodland		*=grassland + coverage:trees=sparse	X	A12 A3A11 // A12 A2A11 (A12 A3A12 // A12 A2A13 or A12 A2A12 // A12 A3A13)		A12 A2A12 // A12 A3A14
20 – 40%	X			X	X			X
< 20%	*=forest + coverage:herbs=sparse	*=woodland + coverage:herbs=sparse	X	X	A12 A3A10 // A12 A2A14	A12 A3A12 // A12 A2A14	X	X

Table A2 Detailed listing of the codes used by FAO LCCS. Source: Di Gregorio 2005, pp. 176–179.

PRIMARILY VEGETATED						PRIMARILY NON-VEGETATED	
A11 - CULTIVATED & MANAGED LANDS		**A12 - NAT. & SEMI-NAT. TERRESTRIAL VEG.**		**A24 - NAT. & SEMI-NAT. AQUATIC VEG.**		**B15 - ARTIFICIAL SURFACES AND ASS. AREAS**	
I. A. Life Form of the Main Crop	code	I. A Life Form of the Main Strata	Code	I. A. Life Form of the Main Strata	Code	I. A. Surface Aspect	Code
Trees	A1	**Woody**	A1	**Woody**	A1	**Built Up**	A1
Broadleaved	A7	Trees	A3	Trees	A3	Linear	A3
Needleleaved	A8	Shrubs	A4	Shrubs	A4	Roads	A7
Evergreen	A9	**Herbaceous**	A2	**Herbaceous**	A2	Paved	A8
Deciduous	A10	Forbs	A5	Forbs	A5	Unpaved	A9
Shrubs	A2	Graminoids	A6	Rooted	A8	Railways	A10
Broadleaved	A7	**Lichens/ Mosses**	A7	Free Floating	A9	Comm. Lines/Pipelines	A11
Needleleaved	A8	Lichens	A7	Graminoids	A6	Non-Linear	A4
Evergreen	A9	Mosses	A9	**Lichens/Mosses**	A7	Industrial a/o Other	A12
Deciduous	A10	**A. Cover**		Lichens	A10	High density	A14
Herbaceous	A3	**Closed** (> 70-60%)	A10	Mosses	A11	Medium Density	A15
Graminoids	A4	**Open** (70-60 - 20-10%)	A11	**A. Cover**		Low Density	A16
Non-Graminoids	A5	(70-60 - 40%)	A12	**Closed** (> 70-60%)	A12	Scattered density	A17
Urban Vegetated Area(s)	A6	(40-20 - 10%)	A13	**Open** (70-60 - 20-10%)	A13	Urban Areas	A13
Parks	A11	**Closed to Open** (100 -15%)	A20	**Closed to Open** (100-15%)	A20	High density	A14
Parkland	A12	(100-40%)	A21	(100-40%)	A21	Medium Density	A15
Lawns	A13	**Sparse** (20-10 - 1%)	A14	(70-60 - 40%)	A12	Low Density	A16
B. Spatial Aspect - Size		(<20-10 - 4%)	A15	(40-20 - 10%)	A15	**Non Built Up**	A2
Large-to Medium-Sized Field(s)	B1	Scattered (4-1%)	A16	**Sparse** (20-10 - 1%)	A16	Waste Dump Deposit	A5
Large-Sized Field(s)	B3	**B. Height**		(<20-10 - 4%)	A17	Extraction Sites	A6
Medium-Sized Field(s)	B4	**7-2 m** (for Woody)	B1	Scattered (4-1%)	A18	**A. Built-Up Object**	
Small-Sized Field(s)	B2	**>30-3 m** (for Trees)	B2	**B. Height**		(scroll list with pre-defined objects)	
B. Spatial Aspect - Distribution		>14 m	B5	**7-2 m** (for Woody)	B1		
Continuous	B5	14-7 m	B6	**>30-3 m** (for Trees)	B2		
Scatterred Clustered	B6	7-3	B7	>14 m	B5	**B16 - BARE AREAS**	
Scattered Isolated	B7	**5-0.3 m**	B3	14-7 m	B6	I. **A. Surface aspects**	Code
		5-0.5 m	B14	7-3 m	B7	**Consolidated**	A1
II. **C. Crop Combination**		5-2 m	B8	**5-0.3 m**	B3	Bare Rock a/o Coarse Frgm.	A3
Single Crop	C1	2-0.5 m	B9	5-0.5 m	B14	Bare Rock	A7
Multiple Crop	C2	<0.5 m	B10	5-2 m	B8	Gravel/Stones/Boulders	A8
One Additional Crop	C3						

Trees	C5
Shrubs	C6
Herbaceous Terrestrial	C7
Herbaceous Aquatic	C8
Simultaneous	C17
Overlapping	C18
Sequential	C19
Trees	C13
Shrubs	C14
Graminoids	C15
Non-graminoids	C16
Simulltaneous	C17
Overlapping	C18
Sequential	C19
III. D. Cultural Practices - Water Supply	
Rainfed	D1
Post-flooding	D2
Irrigated	D3
Surface Irrigation	D4
Sprinkler Irrigation	D5
Drip Irrigation	D6
D. Cult. Practices - Cult. Time Factor	
Shifting Cultivcation	D7
Fallow System	D8
Permanent Cultivation	D9
S. Crop Type	
Food Crops	S1
Cereals (& Pseudocereals)	S3
Roots & Tubers	S4
Pulses & Vegetables	S5
Fruit & Nuts	S6
Fodder Crops	S7
Beverages & Stimulants	S8

3-0.03 m	B4
3-0.3 m	B15
3-0.8 m	B11
0.8-0.3 m	B12
0.3-0.03 m	B13
C. Spatial Distribution/Macropattern	
Continuous	C1
Fragmented	C2
Striped	C4
Cellular	C5
Parklike Patches	C3
II. D. Leaf Type	
Broadleaved	D1
Needleleaved	D2
Aphyllous	D3
E. Leaf phenology	
Evergreen	E1
Semi-Evergreen	E4
Deciduous	E2
Semi-Deciduous	E4
Mixed	E3
Mixed (for Forbs/Graminoids)	E5
Annual	E6
Perennial	E7
III. F. Stratification - Second Layer	
Second Layer Absent	F1
Second Layer Present	F2
Woody	F3
Trees	F4
Shrubs	F5
Herbaceous	F4

2-0.5 m	B9
<0.5 m	B10
3 - 0.03 m	B4
3 - 0.3 m	B15
3-0.8 m	B11
0.8-0.3 m	B12
0.3-0.03 m	B13
II. C. Water Seasonality	
More Than Three Months A Year	C1
Persistent for Whole Day	C4
With Daily Variations	C5
Less Than Three Months A Year	C2
Waterlogged	C3
III. D. Leaf Type	
Broadleaved	D1
Needleleaved	D2
Aphyllous	D3
E. Leaf Phenology	
Evergreen	E1
Semi-Evergreen	E3
Deciduous	E2
Semi-Deciduous	E3
Mixed	E4
Mixed (for Forbs/Graminoids)	E5
Annual	E6
Perennial	E7
IV. F. Stratification - Second Layer	
Second Layer Absent	F1
Second Layer Present	F2
Woody	F3
Trees	F4
Shrubs	F5
Herbaceous	F4

Gravel	A14
Stones	A15
Boulders	A16
Hardpans	A4
Ironpan/Laterite	A9
Petrocalcic	A10
Petrogypsic	A11
Unconsolidated	A2
Bare Soil a/o Other Uncon. Mat.	A5
Stony (5 - 40%)	A12
Very Stony (40 - 80%)	A13
Loose and Shifting Sands	A6
Stony (5 - 40%)	A12
Very Stony (40 - 80%)	A13
II. B. Macropattern - Sands	
Dunes	B1
Barchans	B2
Saturated	B5
Unsaturated	B8
Parabolic Dunes	B3
Saturated	B6
Unsaturated	B9
Longitudinal Dunes	B4
Saturated	B7
Unsaturated	B10
Salt Flat	B13
B. Macropattern - Soils	
Gilgai	B11
Termite Mounds	B12

Other	S13
Non-Food Crops	S2
Industrial Crops	S9
Wood/Timber	S10
Other	S14
A23 - CULTIVATED AQUATIC AREAS	
I. A. Life Form of the Main Crop	**Code**
Graminoids	A1
Non-Graminoids	A2
Woody	A3
B. Spatial Aspect - Size	
Large-To Medium-Sized Field(s)	B1
Large-Sized Field(s)	B3
Medium-Sized Field(s)	B4
S mall-Sized Field(s)	B2
B. Spatial Distribution	
Continuous	B5
Scattered Clustered	B6
Scattered Isolated	B7
II. C. Water Seasonality	
Persistent for Whole Day	C1
With Daily Variations	C2
Waterlogged	C3
III. D. Cultural Practices - Fallow period	
Permanent	D1
Relay Intercropping	D2
Sequential	D3
S. CROP TYPE	
Food Crops	S1
Cereals	S3
Fodder Crops	S7
Other	S13
Non-Food Crops	S2

G. Cover - Second Layer	
Closed To Open	F7
Closed (> 70-60%)	F8
Open (70-60 - 20-10%)	F9
Sparse (20-10 - 1%)	F10
H. Height - Second Layer	
7-2 m (for Woody)	G1
>30-3 m	G2
>14 m	G5
14-7 m	G6
7-3 m	G7
5 - 0.3 m	G3
5-2 m	G8
2-0.5 m	G9
< 0.5 m	G10
3 - 0.03 m	G4
3-0.3 m	G11
0.3-0.03 m	G12
F. Stratification - Third Layer	
Third Layer Absent	F1
Third Layer Present	F2
Woody	F3
Trees	F4
Shrubs	F5
Herbaceous	F4
G. Cover - Third Layer	
Closed To Open	F7
Closed (> 70-60%)	F8
Open (70-60 - 20-10%)	F9
Sparse (20-10 - 5%)	F10
H. Height - Third Layer	
7-2 m (for Woody)	G1
>30-3 m	G2
>14 m	G5

G. Cover - Second Layer	
Closed To Open	F7
Closed (> 70-60%)	F8
Open (70-60 - 20-10%)	F9
Sparse (20-10 - 1%)	F10
H. Height - Second Layer	
7 - 2 m (for Woody)	G1
>30 - 3 m	G2
>14 m	G5
14-7 m	G6
7-3 m	G7
5 - 0.3 m	G3
5-2 m	G8
2-0.5 m	G9
< 0.5 m	G10
3 - 0.03 m	G4
3-0.3 m	G11
0.3-0.03 m	G12
T. Floristic Aspect	
Single Plant Species	T1
Dominant Species	T3
Most Frequent Species	T4
Groups of Plant Species	T2
Statistically Derived Groups	T5
Non-Statistically Derived	T6

B27 - ARTIFICIAL WATERBODIES	
I. A. Physical Status	**Code**
Water	A1
Flowing	A4
Standing	A5
Snow	A2
Ice	A3
Moving	A6
Stationary	A7
B. Persistence	
Perennial (> 9 Months)	B1
9-7 (months)	B7
6-4 (months)	B8
3-1 (months)	B9
Non-Perennial (< 9 Months)	B2
Surface Aspect: Bare Rock	B4
Surface Aspect: Bare Soil	B5
Surface Aspect: Sand	B6
Tidal Area	B3
Surface Aspect: Bare Rock	B4
Surface Aspect: Bare Soil	B5
Surface Aspect: Sand	B6
II. C Depth	
Deep to Medium	C1
Shallow	C2
D Sediment Load	
Almost No Sediment	D1
With Sediment	D2
V. SALINITY	
Fresh (<1 000 ppm of TDS)	V1
Slightly Saline	V2
Moderately Saline	V3
Very Brine	V4
Brine	V5

Crops for Biological Filtration	S11
Fibre Crops & Structural Mat.	S12
Other	S14

14-7 m	G6
7-3 m	G7
5 - 0.3 m	G3
5-2 m	G8
2-0.5 m	G9
< 0.5 m	G10
3 - 0.03 m	G4
3-0.3 m	G11
0.3-0.03 m	G12

T. FLORISTIC ASPECT

Single Plant Species	T1
Dominant Species	T3
Most Frequent Species	T4
Groups of Plant Species	T2
Statistically Derived Groups	T5
Non-Statistically Derived	T6

ENVIRONMENTAL ATTRIBUTES FOR EACH

A11: Landform, Lithology, Soil, Climate, Altitude, Erosion and Crop Cover.

A12: Landform, Lithology, Soil, Climate, Altitude and Erosion.

A23: Landform, Lithology, Soil, Climate, Altitude, Erosion and Crop Cover.

A24: Landform, Lithology, Soil, Climate, Altitude and Salinity.

B15: Landform, Climate and Altitude.

B16: Landform, Climate, Altitude, Vegetation and Erosion.

B27/B28: Climate and Altitude.

B28 INLAND WATERBODIES, SNOW & ICE

I. Physical Status	**Code**
Water	A1
Flowing	A4
Standing	A5
Snow	A2
Ice	A3
Moving	A6
Stationary	A7
B. Persistence	
Perennial (> 9 Months)	B1
9-7 (months)	B7
6-4 (months)	B8
3-1 (months)	B9
Non-Perennial (< 9 Months)	B2
Surface Aspect: Bare Rock	B4
Surface Aspect: Bare Soil	B5
Surface Aspect: Sand	B6
Tidal Area	B3
Surface Aspect: Bare Rock	B4
Surface Aspect: Bare Soil	B5
Surface Aspect: Sand	B6
II. C Depth	
Deep to Medium	C1
Shallow	C2
D. Sediment Load	
Almost No Sediment	D1
With Sediment	D2
V. SALINITY	
Fresh (<1 000 ppm of TDS)	V1
Slightly Saline	V2
Moderately Saline	V3
Very Brine	V4
Brine	V5

ENVIRONMENTAL ATTRIBUTES

L. LANDFORM*	Code
Level Land	L1
Sloping Land	L2
Steep Land	L3
Land With Composite Landforms	L4

M. LITHOLOGY*	Code
Igneous rock	M100
Igneous plutonic rock	M110
Granite	M111
Granodiorite	M112
Quartz diorite	M113
Syenite	M114
Monzonite	M115
Diorite	M116
Gabbro	M117
Foidic plutonic rock	M118
Ultramatic plutonic rock	M119
Igneous hypabyssal rock	M120
Aplite	M121
Pegmatite	M122
Porphyry	M123
Dolerite/diabase	M124
Igneous volcanic rock	M130
Rhyolite	M131
Dacite	M132
Trachyte	M133
Latite	M134
Andesite	M135
Basalt	M136
Phonolite	M137
Tephrite	M138
Pyroclastic rock	M140
Ash	M141
Lapilli	M142
Scoria	M143
Tuff	M144
Ignimbrite	M145
Lahar	M146
Agglomerate	M147
Other Igneous rock	M199
Sedimentary rock	M200
Unconsolidated clastic sed. Rock	M210
Clay	M211
Silt	M212
Sand	M213
Gravel	M214
Loess	M215
Loam	M216
Colluvium	M217
Shells	M218
Cons. clastic siliceous sed. Rock	M220
Mudstone	M221
Siltstone	M222
Shale	M223
Quartzararenite	M224
Lithic arenite	M225
Feldspathice arenite/arkose	M226
Graywacke	M227
Conglomerate	M228
Breccia	M229
Calcareous rock	M230
Marl	M231
Calcilutite	M232
Calcarenite	M233
Calcirudite	M234
Algal/reefal limestone	M235
Travertine	M236
Tufa	M237
Dolomite	M238
Evaporite	M240
Gypsum	M241
Halite	M242
Organic rock	M250
Peat	M251
Lignite	M252
Coal	M253
Tar	M254
Residual rock	M260
Laterite	M261
Bauxite	M262
Kaolin	M263
Other Sedimentary rock	M299
Metamorphic rock	M300
Contact metamorphic rock	M310
Hornfels	M311
Spotted slate	M312
Skarn	M313
Cataclastic metamorphic rock	M320
Cataclastic breccia	M321
Mylonite	M322
Regional-metamorphic rock	M330
Slate	M331
Schist	M332
Gneiss	M333
Migmatite	M334
Granulite	M335
Eclogite	M336
Quartzite	M337
Marble	M338
Other metamorphic rock	M399

M. LITHOLOGY - AGE GEOL PARENT MAT	
Quartenary	M400
Holocene	M410
Pleistocene	M420
Late Pleistocene	M421
Middle Pleistocene	M422
Early Pleistocene	M423
Tertiary	M500
Piocene	M510
Miocene	M520
Oligocene	M530
Eocene	M540
Paleocene	M550
Mesozoic	M600
Cretaceous	M610
Jurassic	M620
Triassic	M630
Paleozoic	M700
Permian	M710
Carboniferous	M720
Devonian	M730
Silurian	M740
Ordovician	M750
Cambrian	M760
Precambrian	M800

N. SOIL - SURFACE ASPECT	Code
Bare Rock	N1
Soil Surface	N2
Stony (5 - 40%)	N5
Very Stony (40 - 80%)	N6
Loose and Shifting Sands	N3
Stony (5 - 40%)	N5
With Dunes	N7
Hardpans	N4
Ironpan/Laterite	N8
(petro)Calcic	N9
PetroGypsic	N10
Hardened Plinthite	N11

N. SOIL - SUBSURFACE ASPECT	
FAO's Major Soil Groups*	N12

O. CLIMATE*	Code
Thermal Climate:	
Tropics	O1
Subtropics - Summer rainfall	O2
Subtropics - Winter Rainfall	O3
Temperate Oceanic	O4
Temperate Continental	O5
Boreal Oceanic	O6
Boreal Continental	O7
Polar Arctic	O8
Moisture Determined LGP:	
Hyperarid	O9
Arid	O10
Dry Semi-Arid	O11
Moisture Semi-Arid	O12
Subhumid	O13
Humid	O14
Perhumid	O15

P. ALTITUDE*	Code
< 50 - 300 m	P1
< 50 m	P5
50 - 100 m	P6
100 - 300 m	P7
300 - 1500 m	P2
300 - 600 m	P8
600 - 1000 m	P9
1000 - 1500 m	P10
1500 - 3000 m	P3
1500 - 2000 m	P11
2000 - 2500 m	P12
2500 - 3000 m	P13
3000 - > 5000 m	P4
3000 - 3500 m	P14
3500 - 5000 m	P15
> 5000 m	P16

Q. EROSION	
No Visible Erosion	Q1
Visible Evidence of Erosion	Q2
Water Erosion	Q3
Sheet	Q6
Rill	Q7
Gully	Q8
Wind Erosion	Q4
Mass Movement	Q5

R. WATER QUALITY	
Fresh Water	R1
Brackish Water	R2
Saline Water	R3

U. VEGETATION	
Scattered Vegetation Present	U1
Woody	U2
Herbaceous	U3
Forbs	U5
Graminoids	U6
Lichens/Mosses	U4
Lichens	U7
Mosses	U8

W. CROP COVER/DENSITY	
Permanent Life Forms:	
Closed Cover > (70-60%)	W1
Closed Cover (70-60) - (20-10)%	W2
Sparse Cover <(20-10)%	W3
Temporary Life Forms:	
High Crop Density (> 60%)	W4
Medium Crop Density (60 - 30%)	W5
Low Crop Density (30 - 15%)	W6

Table A3 Complete listing of the LUCAS land use classes (level 1–3). Source: EUROSTAT 2015a

U100			**PRIMARY SECTOR**
	U110		Agriculture
		U111	Agriculture
		U112	fallow land
		U113	kitchen garden
	U120		Forestry
	U130		Aquaculture & Fishing
	U140		Mining and quarrying
	U150		Other primary production
U200			**SECONDARY SECTOR**
	U210		Energy production
	U220		Industry & Manufacturing
		U221	Manufacturing of food, beverages and tobacco products
		U222	Manufacturing of textile products
		U223	Coal, oil and metal processing
		U224	Production of non-metal mineral goods
		U225	Chemical and allied industries and manufacturing
		U226	Machinery and equipment
		U227	Wood based products
		U228	Printing and reproduction
U300			**TERTIARY SECTOR, TRANSPORT, UTILITIES & RESIDENTIAL**
	U310		Transport, communication networks, storage, protection works
		U311	Railway transport
		U312	Road transport
		U313	Water transport
		U314	Air transport
		U315	Transport via pipelines
		U316	Telecommunication
		U317	Logistics and storage
		U318	Protection infrastructures
		U319	Electricity, gas and thermal power distribution
	U320		Water and waste water treatment
		U321	Water supply and treatment
		U322	Waste treatment
	U330		Construction
	U340		Commerce, financial, professional and information services
		U341	Commerce
		U342	Financial, professional and information services
	U350		Community Services
	U360		Recreation, Leisure, Sport
		U361	Amenities, museums, leisure
		U362	Sport
	U370		Residential
U400			**unused and abandoned areas**
	U410		abandoned areas
	U420		semi-natural and natural areas not in use