Master Thesis

submitted within the UNIGIS MSc programme

Interfaculty Department of Geoinformatics - Z_GIS

University of Salzburg

The quality of OpenStreetMap in protected areas

Assessed on the example of road networks in German national parks

by

M. A. Markus Rehn

105419, UNIGIS MSc 2020

A thesis submitted in partial fulfilment of the requirements of

the degree of

Master of Science – MSc

Advisor:

Prof. Dr. Josef Strobl

Helsinki, 22.05.2022

ABSTRACT

Worldwide, hundreds of millions of people with a rising demand for recreation visit protected areas every year. With the advent of GPS technologies, which are available to almost everyone via smartphones, -watches etc., visitors increasingly use mobile routing applications to access these remote landscapes and navigate the terrain.

Most outdoor routing engines are built on data from the OpenStreetMap (OSM) project. Being the most popular example of VGI (Volunteered Geographic Information), OSM initially comes without any quality assurance. With more people trusting on OSM-based applications, the quality of OSM-data becomes more and more important also in protected areas such as national parks.

This thesis aimed to present and utilize an easy-to-apply method and suitable quality indicators to determine the reliability of OSM road networks within protected areas. With the help of the OpenStreetMap History Database and the ohsome API, intrinsic quality evaluation methods were used to estimate feature completeness, attribute completeness and currentness of OSM data in ten national parks in Germany. In a subsequent case study, OSM's fitness-for-use for outdoor routing was evaluated on the example of the Bavarian Forest national park. Here, extrinsic methods and a reference dataset as well as intrinsic measures were used to assess feature completeness, attribute completeness, classification correctness and trustworthiness of the data.

Results indicated that OSM road networks are of high feature completeness in protected areas. At the same time, they show insufficient presence of important attributes for several outdoor activities. Determined by specific legislation on nature conservation that differs from one area to another, access restrictions are incompletely represented within OSM. Furthermore, OSM data in most national parks is maintained by only a small number of contributors. Thus, data quality usually depends on the mapping work of just a few persons.

Overall, results of this study revealed a questionable usability of OSM for routing in protected areas. They also uncovered the opportunities for administrators of national parks to reduce conflicts between recreationalists and the purposes of nature conservation by actively contributing to OSM.

CONTENT

FIGURES

TABLES

1. Introduction

1.1 Background

Spending time outdoors and explore nature while being physically active is one of the most important forms of recreation in modern societies. Worldwide, hundreds of millions of people visit protected areas like national parks and nature reserves every year with a rising demand for recreation. At the same time, new technologies increase the access to and within those large-scale areas which leads to growing disturbance of ecosystems through leisure activities (BALMFORD ET AL. 2015; MONZ ET AL. 2021).

Many of such new technologies and digital services to plan activities and navigate the terrain are based on data from the OpenStreetMap (OSM) project. OSM is one of the most prominent examples of Volunteered Geographic Information (VGI), a term first defined by GOODCHILD (2007) to describe the voluntary collection and contribution of georeferenced information by a large number of private individuals via the internet. Due to data collection mainly by amateurs who use a wide range of tools and methods without any strict standards, VGI projects like OSM initially come without an assurance of quality (GOODCHILD AND LI 2012).

With more people trusting on OSM-based applications to navigate in remote landscapes, the reliability of OSM-data becomes increasingly important in these regions. This particularly applies to the OSM road network in protected areas, including all tracks and paths that can potentially be used to move from A to B. Complete and correct annotation and classification of the data ensures reliable navigation of users and avoids conflicts between the areas' purposes of nature conservation and offering opportunities for nature-based activities.

This thesis is organised as follows: Chapter 1 gives a brief introduction into the OpenStreetMap project and current quality standards for geographic information. After a compact review of quality studies conducted on OSM road networks, research question and objects of the study are being presented. In chapter 2, quality measures and study areas are introduced, followed by a presentation of the analysis' tools and workflows. Results are reported and discussed in chapter 3. Finally, chapter 4 draws conclusions and indicates directions for future research.

Throughout the thesis, the expressions "road" and "way" are used synonymously. The term "road" thus also includes paths and tracks that are usually not described as roads in everyday life.

1.2 OpenStreetMap (OSM)

Founded in 2004 at the University College London, OpenStreetMap's original idea was the creation of a road dataset covering the United Kingdom as an alternative to official map products. Based on the voluntary mapping work of a growing number of contributors, the project expanded globally within a few years, covering not only roads but all kinds of features like Points of Interest (POIs), buildings and land use areas. OSM is licensed under the Open Database License 1.0 (ODbL) which ensures its data is always free to create products from, to share and to adopt as long as the OSM community is attributed as the producer of the data (RAMM ET AL. 2010).

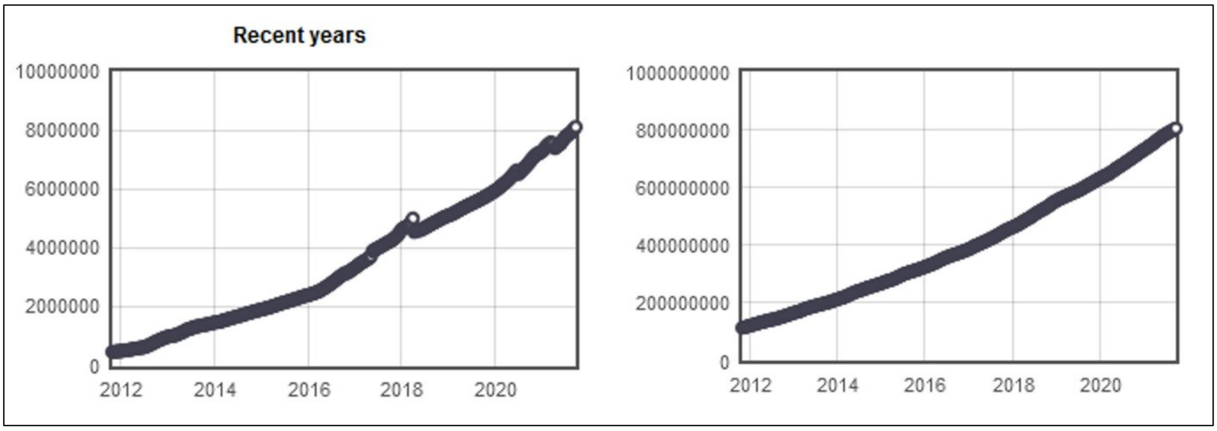


Figure 1: Registered OSM members (left) and ways (right) in the OSM database (NEIS 2021).

Any individual can contribute to OSM (after registering at the project for free) and all OSM data can be edited and deleted by any contributor (all changes to the data are saved and can be revoked). Within the past 18 years, OSM has evolved from a national initiative with a few hundred contributors to one of the greatest and most famous VGI projects counting more than eight million registered members in May 2022 (OSM1). Figure 1 shows that the amount of data in the OSM database grew at a similar pace, making the global OpenStreetMap more than 80% complete by 2017 (BARRINGTON-LEIGH AND MILLARD-BALL 2017).

1.2.1 Data model

The OSM database has a relatively simple XML-based approach. Figure 2 illustrates the conceptual data model which contains three main data types that include a geometric component (RAMM ET AL. 2010):

- **Nodes**: Nodes represent geographic points with a unique latitude and longitude value. They are used to store information on the geographic position and to represent POIs or vertices of polylines and polygons.

- **Ways**: Ways represent either lines (e. g. streets) or polygon boundaries (e. g. buildings). A way is a collection of between 2 and 2,000 nodes that define a polyline. If it represents a polygon, the first and the last node must be identical and form a "closed way".

- **Relation:** A relation is a multi-purpose data structure that represents a logical or geographical relationship between two or more nodes, ways and other relations (e. g. multipolygons or boundaries). It is primarily an ordered list of above-named elements, which are known as the relation's members. As relations can have different meanings, a relation's meaning is usually defined by a "type"-tag (e. g. type=boundary).

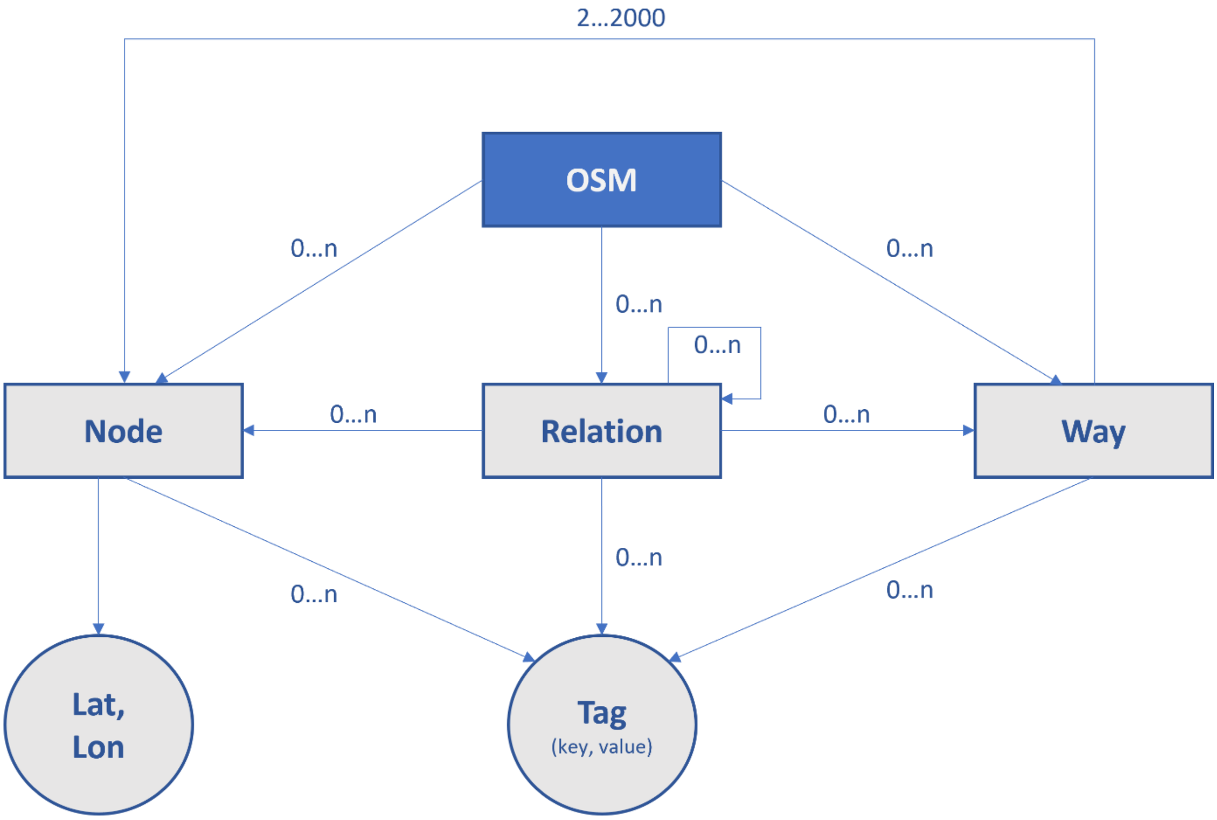


Figure 2: OSM Data Model based on RAMM ET AL.. (2010).

All coordinates in OSM are stored in the WGS 84 (EPSG:4326) coordinate reference system. Every object (node, way or relation) receives a unique ID during its creation process that will not be reassigned even if the object is deleted at a later stage. Because every data type has its own ID space, e. g. a node and a way can have the same ID but are still uniquely defined (RAMM ET AL. 2010). Next to the unique ID, every created object receives a number of predefined attributes (see table 1). Additionally, all tags and a full editing history of every object is stored. Edits are published immediately after the contributor saved all changes. All saved changes done by a contributor in one editing session form a so-called "changeset".

Table 1: Common attributes of all OSM data types (OSM3).

Name	**Value**	**Description**
id	integer	Unique numeric identifier of the object.
user	string	Name of the user who last modified the object. A user can change his/her display name any time.
uid	integer	Numeric identifier of the user who last modified the object. An uid never changes.
timestamp	W3C standard date and time formats	Time of last modification.
visible	"true" or "false"	Defines whether the object is deleted or not in the database.
version	integer	Edit version of the object, starting with 1 when newly created.
changeset	integer	Number of changeset in which the object was created or updated.

1.2.2 Semantic ecosystem

All objects in OSM are attributed by using so-called "Tags". **Tags** describe the meaning of the element (node, way or relation) they are attached to. A tag consists of two free format Unicode string text fields (up to 255 characters), a *key* and a *value*. The key is used to describe a topic, category, or type of feature (e. g. *name* or *highway*) while the value provides detail for the key-specific feature (e. g. *highway=path*). There is no restriction or limit to the usage of tags. They can be attached and modified by any registered contributor and every object can have zero to infinite tags assigned (DAVIDOVIC ET AL. 2016).

Unlike conventional cartographic datasets, the conceptual model of OSM does not follow a top-down, expert-driven approach but allows contributors to create any new tag to describe the objects they are mapping. This lightweight semantic model is responsible for the absence of a "ground truth" in the annotation process of OSM (BALLATORE AND ZIPF 2015). Nevertheless, the OSM community agreed on a list of certain key-value pairs for most commonly used tags. This catalogue called "Map Features" acts as an informal standard and should be followed during the tagging process (OSM8). Since it represents the progressing product of discussions between contributors, such type of reference list has been termed "folksonomy" (BALLATORE AND MOONEY 2015).

The pure existence of this guidance does not guarantee consistent tagging, though. Several studies have shown a high "semantic heterogeneity" and a rather poor compliance of tags with the suggestions of the OSM Wiki, indicating that contributors do not always tag features the way it is recommended by the folksonomy (VANDECASTEELE AND DEVILLERS 2015; DAVIDOVIC ET AL. 2016).

1.2.3 Contributors

Data in OSM is collected and edited by using a variety of methods and tools. Most contributors use a GPS device to map features and add relevant attribute information, but data has also been created by digitizing objects from orthophotos and satellite imagery or by importing official datasets (GIRRES AND TOUYA 2010).

Besides the fact that OSM-contributors are not faced with any restrictions or strict rules during the collecting and annotation process, they also do not necessarily have expert knowledge of surveying (GOODCHILD 2007). Consequently, and in contrast to traditional centralized procedures of gathering geographic data, within OSM the majority of data is collected by amateurs from diverse social backgrounds. Their motivation varies significantly from self-expression over representation of their online identity to simply having fun (BARRON ET AL. 2014). Furthermore, how they contribute to the project also highly depends on their technical equipment, their individual interpretation of data, possible misunderstandings of commonly used key-value-pairs and the willingness of taking the effort to assign attributes. GIRRES AND TOUYA (2010) noted, that socio-demographic factors like high income and low population age result into a higher number of contributions. Most contributions in OSM are isolated without any planned collaboration (MOONEY AND CORCORAN 2012) and a majority of contributors is most active within the first three months after their registration (NEIS AND ZIPF 2012).

Although the huge number of contributors usually quickly corrects and solves errors along with conflicts and makes OSM data the most up-to-date map data available for many regions (OVER ET AL. 2010), it has to be noted that the community is extremely heterogeneous due to an uneven spread of contributors and their diverse experience and contribution efforts (MA ET AL. 2015). NEIS

AND ZIPF (2012) have shown that about 5% of all registered users have produced almost 90% of all changesets. Thus, a minority of OSM users accounts for a majority of contributions and edits.

Mapping practices

Even though contributors are theoretically totally free in how to map, the OSM community tried to agree on a code of conduct that should be considered during the mapping process. According to these guidelines, "verifiability" serves as a main principle for all mapping activities in OSM, meaning data should be verifiable as far as possible. As a result, mappers should only map "what's on the ground", meaning all data must be represented by some real-world object. This also applies to local legislation such as traffic rules that should only be mapped (attributed) if bound to a real object like a traffic sign. Thus, the current on the ground situation is seen as more important than any official source and rules that cannot be recognized in some way on site should not be mapped (OSM4).

1.2.4 OSM-based services and applications

An extensive coverage and the growing size of its database have made OSM "a potential competitor to public and commercial geodata providers" in many domains (NEIS ET AL. 2012). Obviously, a number of benefits arise from OSM's user-generated data approach: It is cost-free, enables the use of potentially up-to-date data at any time and can be shared easily (DORN ET AL. 2015).

Those benefits have led to the development of countless applications built on the OSM database. The list of OSM-based services includes projects from art, history, 3D, public transport, indoor and outdoor routing as well as numerous other fields and services (OSM7).

Outdoor routing and navigation

The availability of large amounts of OSM linear features, combined with an unbroken trend of visiting nature for recreational purposes and the spreading of personal devices carrying a GPS sensor, fostered the development of mobile navigation applications (BALMFORD ET AL. 2015; BERGMAN AND OKSANEN 2016; ROUSELL AND ZIPF 2017). Mainstream car-routing services such as Apple Maps, Bing Maps or Google Maps are all based on proprietary and often inaccessible data. For the development of outdoor routing apps however, OSM has been the data source of choice and trust for many developers. Routing engines of leading applications for outdoor navigation are build either partly or completely on OSM data (see table 2). As a consequence, millions of users worldwide rely on OSM when accessing remote landscapes and navigate the terrain. In the DACH-region (Germany, Austria and Switzerland), OSM-based outdoor routing-apps were downloaded 22.8 million times until January 2022, which represents a market share of over 84%. Additionally, maps and routing engines of many GPS-devices (Garmin, Bosch eBike etc.) are also build on the OSM-database (HALLERMANN 2022).

Table 2: Examples of leading outdoor platforms using OSM data.

Platform	**Registered users**	**OSM-usage**
AllTrails	35 mio (ALLTRAILS 2022a)	Derives trail segments from OSM data and stores them in a separate database. Offers all users an own custom-built map that is based on OSM. Additionally, OpenStreetMap and the OSM-based OpenCycleMap (OCM) are provided as alternative base layers (ALLTRAILS 2022b).
Komoot	>25 mio (KOMOOT GMBH 2022a)	Komoot maps and route planner use map data from OSM. Users are encouraged to correct errors directly in OSM (KOMOOT GMBH 2022b).
Outdooractive	>12 mio (OUTDOORACTIVE GMBH 2022b)	OSM is the only basemap available in the cost-free basic version of the Outdooractive application. Official topographic maps and other professional maps are only accessible for users that have a paid Pro-account (OUTDOORACTIVE GMBH 2022a).

Because not only geometries but also relevant attributes are an important part of route calculation within routing engines (GRASER ET AL. 2014), the reliability of geometries as well as of tags plays a substantial role for correct navigation and thus for avoiding conflicts between nature conservation and recreational outdoor activities.

1.2.5 Road networks in OSM

In OSM, roads are represented by elements of the data type *way*. The key *highway* is used for defining any *way*-element as some kind of road, street or path etc. The associated value to the *highway*-key indicates the type of road and its importance within the road network as a whole (e. g. *highway=motorway*). This practice distinguishes roads by function and importance rather than any physical characteristic and legal classification. Although mappers are not obligated to follow national road classification schemes, a high correlation is typically existing.

Principal tags for the general street network range from *highway=motorway* (most "important" category in the OSM road network) to *highway=residential* (least important). Additionally, there are several categories of link roads (e. g. *highway=motorway_link*), special road types (e. g. *highway=track*) and paths (e. g. *highway=bridleway*).

Naturally, while in urban areas the road network is mainly consisting of roads that belong to a category of the general street network, in rural landscapes like forests and protected areas, more special road types like tracks or paths are to be found (see figure 3).

Tag	Description	Example
highway=track	Minor land-access road that is not considered part of the general-purpose road network. Mostly used for agriculture, forestry, outdoor recreation and similar activities on open land.	
highway=path	A generic path, either multi-use or of unspecified usage, that is open to all non-motorized vehicles and not intended for motorized vehicles unless tagged so separately.	

Figure 3: Paths and tracks in OSM (OSM5).

OSM tags for routing and access restrictions

Routing engines need to include a lot of attribute information in their computations to calculate a correct and reliable route. Although the selection of tags and how they are used by the routing applications is usually confidential and differs between platforms, all of them are dependent on crucial information like access restrictions. Access restrictions in OSM always concern the legal permissions or restrictions and not the practical access (OSM9). To describe a general access restriction that applies to all transport modes, the *access* key is used. Access restrictions that are depending on the transport mode are usually expressed with the respective key (e. g. *motor_vehicle=no*, *bicycle=yes*). As these keys are often not existing in the data, routing engines need to use some default, e. g. the engine assumes that cycling is allowed on all tracks if not tagged otherwise. Even though a general statement on which defaults are used for which road type is not possible and may vary among platforms, this underlines the importance of attributes in OSM.

1.3 Data quality of geographic information

Quality is a crucial component of any (spatial) dataset. It is defined as the "degree to which a set of inherent characteristics fulfils requirements" (ISO19157:2013).

The problem of understanding the quality of geographic data was identified many decades ago. Already in the late 1960s it received attention from surveyors, cartographers and geographers (VAN OORT 2006). With the emergence of Geographic Information Systems (GIS) in the 1980s, the research on quality of spatial data experienced rapid growth which resulted in first efforts to develop standards for the description of spatial data quality in the early 1990s.

1.3.1 Quality standards and elements

The first spatial data quality standard was accepted in the United States in 1992. Later on in 2002, the technical committee 211 (ISO/TC211) of the International Organisation for Standardization (ISO) developed a number of standards concerning spatial data quality (VAN OORT 2006). For quality assessment, the most important ones were *ISO 19113:2002 quality principals* and *ISO 19114:2002 quality evaluation procedures*. The standards were aiming for ensuring that quality measurements of different datasets and executed by various users are comparable. In 2013, after an ongoing revision process, they were revised by the newly published ISO Standard *ISO 19157:2013 Geographic Information: Data Quality* that is still valid at the time of this writing. It is the latest release among other internationally known standards for describing spatial data quality such as the ones published by the International Cartographic Association (ICA), the Federal Geographic Data Committee (FGDC) and the Committee on Standardization (CEN) (FONTE ET AL. 2017).

ISO 19157 describes general principles of geodata quality as well as procedures for the evaluation of quality of digital geographic datasets. As shown in figure 4, it defines a set of data quality elements where every element is a component describing a certain aspect of geographic data quality. All these elements have been structured into six categories (ISO19157:2013):

- **Completeness:** Describes how complete a dataset is. A surplus of data (dataset contains features, that do not exist in reality / in the reference data) is referred to as *Error of Commission*. A lack of data in contrast as *Error of Omission* (dataset misses features, that do exist in reality / in the reference data). Completeness always depends on the purpose of the analysis and can be evaluated for objects (feature completeness) as well as for their relationships and attributes (SEHRA ET AL. 2017).

- **Logical Consistency:** Evaluates the accuracy of the rules and relations established within a dataset. It can be divided into *intra-theme consistency* that expresses the consistency of different objects with other objects of the same theme and *inter-theme consistency* (with objects of other themes). Logical consistency is indicated with percentages (ANTONIOU AND SKOPELITI 2015). ISO19157:2013 distinguishes between four different logical consistency elements:
    - *Conceptual consistency* describes how well data follows its conceptual schema.
    - *Domain consistency* shows if values lie within their domain.
    - *Format consistency* checks if data is stored according to the physical structure of the dataset.
    - *Topological consistency* is existing when data follows predefined topological rules. Typical errors are ways that are not linked in the data but connected in the real world (e. g. street network), dislocation of features that should have the same position or junctions that do not have a node.

- **Positional Accuracy:** Defines the absolute and relative exactness of coordinates within a spatial reference system. It is usually assessed by comparing the position of features to the position of their counterparts in a reference dataset that is considered to represent the "ground truth". Within ISO19157:2013, positional accuracy is determined by three different elements: *Absolute or external accuracy* (difference to true coordinate values), *relative or internal accuracy* (measures the positional difference between features relative to each other in the dataset) and *gridded data position accuracy* (to evaluate position of a grid against its true value in raster data).

- **Temporal Quality:** Describes the historical evolution of the data and refers to the quality of temporal attributes such as date of collection, date of publication, frequency of updates and "up-to-dateness". While the element *accuracy of time measurement* describes the accuracy of time information (closeness of reported time measurements to true time values), *temporal consistency* explains if events are ordered correctly, and *time validity* indicates if time information is valid (e. g. June 31st is an invalid time).

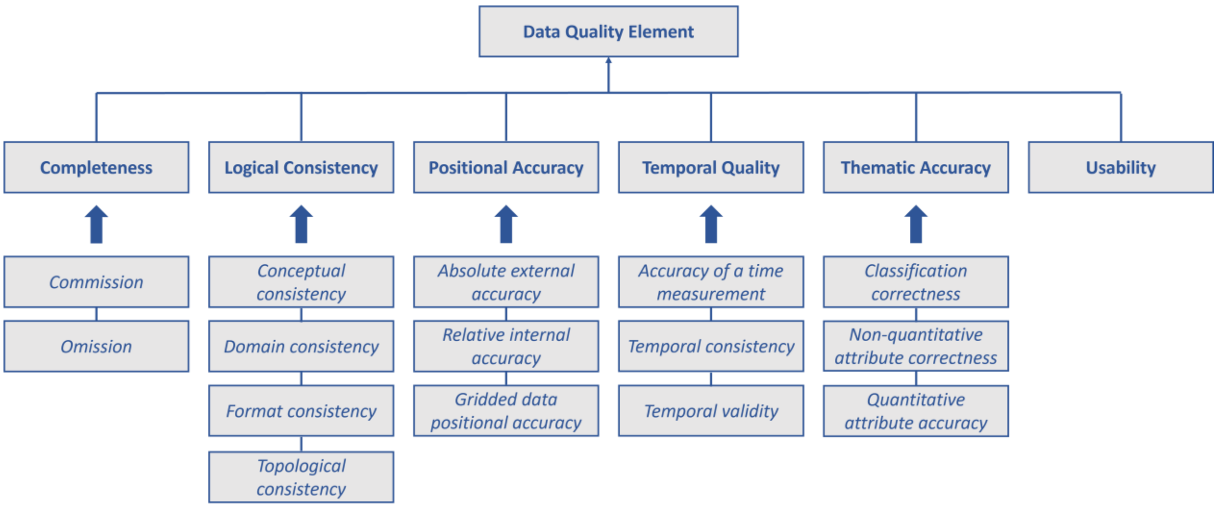


Figure 4: Overview of ISO 19157:2013 data quality elements.

- **Thematic Accuracy:** Describes the correctness of classes, attributes or tags associated with specific geographical objects, e. g. a highway. Often the term "semantic accuracy" is synonymously used. It is traditionally assessed by comparing the information to reference data like satellite imagery or authoritative datasets and subdivided into three elements by ISO19157:2013:
  - *Classification correctness* outlines if objects are classified correctly.
  - *Non-quantitative attribute correctness* describes if the value of non-measurable attributes like "*name*" is correct.
  - *Quantitative attribute accuracy* indicates if the value of measurable (quantitative) attributes like "*area*", "*length*" etc. is correct.

- **Usability:** The above named five categories and their associated quality elements are focused on the quality of a data product from a producer's point of view, also termed the *internal quality* of a dataset. Usability on the other hand, focusses on the user's needs and requirements and is referred to as *external quality*. As a result of a quality evaluation process, it might happen that the internal quality of a dataset is high (as it meets all specifications that were set by its producers), while at the same time the external quality might be poor because the dataset does not meet the requirements of a user. Usability can be evaluated by all above named quality elements or may be characterized by requirements that cannot be described by the other elements. Thus, usability acts as a complementary element by connecting user requirements and data quality measures to check whether the data can be used for a specific application, i. e. usability describes the data's *fitness-for-purpose* (ANTONIOU AND SKOPELITI 2015; FONTE ET AL. 2017).

Non-quantitative quality indicators

As part of the ISO standard *ISO 19115-1:2014 Geographic information – Metadata – Part 1: Fundamentals*, which defines how to describe metadata on geographic information, quality can be further assessed through non-quantitative indicators such as **purpose**, **usage**, and **lineage**. They are used to express a quality overview for the dataset estimated from e. g. knowledge about the source, tool or methods used to collect the data. While purpose describes the dataset's intended usage, usage describes the application(s) in which it has been utilized. The history of a dataset from collection to its form at the time of use is described by lineage (SENARATNE ET AL. 2017).

1.3.2 Data quality evaluation process

Quality assessment can be performed in many ways. Generally, evaluation methods are classified into **direct** and **indirect** approaches. Direct methods evaluate data quality by comparing the data with internal (intrinsic) and/or external (extrinsic) reference information. They comprise the analysis of items within the dataset and can be performed as a full inspection or based on a subset of data. Indirect methods on the other hand are based on external knowledge of the data product. They only estimate data quality based on non-quantitative information on the data such as lineage. Indirectly evaluated data quality may also be solely based on experience. According to ISO19157:2013, direct methods should be used in preference over indirect methods. The actual evaluation process consists of several working steps. Figure 5 illustrates the workflow for assessing data quality as proposed by the ISO standard.

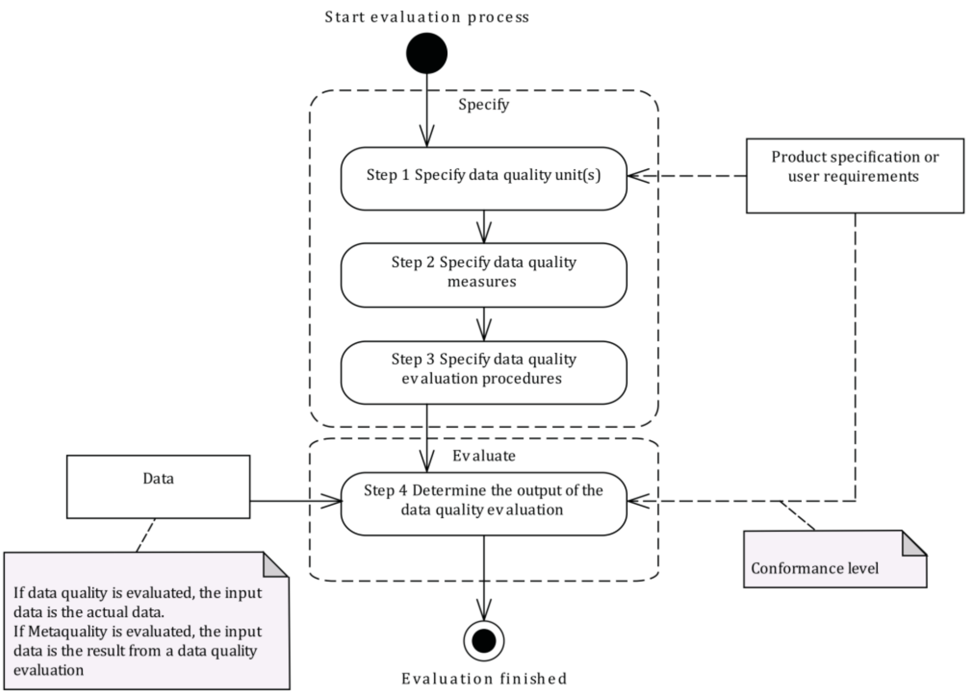


Figure 5: Process workflow of evaluating data quality (ISO19157:2013)

1.4 Related work – quality studies on OSM road networks

Resulting from a heterogeneity of contributors and contributions paired with the absence of quality assurance and strict standards, concerns about the data quality have accompanied the OSM project from its beginning. The more OSM became an essential geodata source for countless areas of application, the more the assessment of its data quality has been of substantial interest to the research community and gave OSM a significant role in Geographic Information Science in the last 15 years (YAN ET AL. 2020).

Quality assessments on OSM data have been carried out for all kinds of purposes and focused on different aspects of comparison. Numerous OSM quality studies of all data types (lines, polygons, points), manifold related thematic layers (e. g. Points of Interest, building footprints, land use or street networks) and for different regions of the world can be found in literature.

As OSM originally started with the idea of creating a free road dataset, many quality studies have been devoted to the assessment of OSM street networks (see appendix A1). The development of new indicators and suitable methods to measure OSM data quality has been another major focus of OSM quality research.

1.4.1 Extrinsic studies and methods

Particularly in the first years of research on OSM, many scientists have followed the traditional and common approach of assessing a dataset's quality by comparing it with a "ground truth" dataset. This approach assumes, that this "official" data, which is mostly provided by public authorities or private companies, comes with an acceptable quality that justifies its usage as reference data representing ground truth. Primarily, the standard quality elements defined by *ISO 19157:2013* have been the subject of extrinsic quality evaluations, with positional accuracy, completeness and thematic accuracy being the ones most investigated (MINGHINI AND FRASSINELLI 2019).

The first systematic quality analysis of an OSM road network was conducted by HAKLAY (2010), who compared the OSM street networks of London and England with the official Ordnance Survey dataset "Meridian 2". He reviewed the quality elements positional accuracy (by performing a buffer comparison method developed by GOODCHILD AND HUNTER in 1997) and completeness (based on a grid-based road length comparison). In his results, the quality and coverage for OSM in England showed heterogeneous patterns with stronger road network concentrations in populated regions (but with a lack of details such as street names) whereas rural areas showed a complete absence of coverage at times. In another study, HAKLAY ET AL. (2010) confirmed the validity of "Linus' Law" (the more contributors, the higher the quality) for the positional accuracy of OSM on the example of the road network in London, UK.

ZIELSTRA AND ZIPF (2010) applied the methodology of HAKLAY (2010) to evaluate the completeness of the OSM road network in Germany. They compared OSM to a commercial TeleAtlas (TomTom) MultiNet dataset by calculating the length differences for the entire country and several cities of different sizes plus within different buffer sizes around city centres. Their findings supported the conclusion of HAKLAY (2010) that OSM datasets are often more complete than authoritative or commercial datasets in densely populated areas while being of poor coverage in rural regions. As they concluded, OSM data for rural areas is not sufficient for the creation of advanced applications such as route planners.

GIRRES AND TOUYA (2010) analysed the quality of the French OpenStreetMap by comparing it with official BD Topo data. They extended the analysis of HAKLAY (2010) to other features such as POIs, waterways and coastlines. Their results showed a similar heterogeneity of OSM for France as previously discovered by HAKLAY (2010) and ZIELSTRA AND ZIPF (2010) for Great Britain and Germany. Their analysis also evaluated a broader range of quality elements such as thematic accuracy and included several different extrinsic methods (see table 3). As one result, they found a high heterogeneity in the classification and attribute correctness of the OSM key *highway* and a notable inaccuracy for the OSM key *name* within the road network. As another result, their study showed that an increase in the number of contributors is strongly related to an increase in positional accuracy and the spatial data volume.

Table 3: Methods used in extrinsic quality assessments of OSM road networks.

COMPLETENESS		
Type	**Method**	**Example Studies**
Unit-based	Length comparison (grid-based)	HAKLAY 2010, LUDWIG ET AL. 2011, ZIELSTRA AND ZIPF 2010, FORGHANI AND DELAVAR 2014
	Number of features comparison	GIRRES AND TOUYA 2010
	Comparison of total length	GIRRES AND TOUYA 2010, KOUKOLETSOS ET AL. 2011, GRASER ET AL. 2014, CAMBOIM ET AL. 2015
	Comparison of density	HOCHMAIR ET AL. 2015, CAMBOIM ET AL. 2015
Feature-based	Geometric matching (line-based)	KOUKOLETSOS ET AL. 2012, ABDOLMAJIDI ET AL. 2015, MOHAMMADI AND MALEK 2015b

	Geometric matching (polygon-based)	Fan et al. 2016
	Attribute matching	Koukoletsos et al. 2012
Attribute completeness	Tag presence	Girres and Touya 2010, Ludwig et al. 2011
	Name completeness	Ludwig et al. 2011, Wang et al. 2013, Graser et al. 2014

LOGICAL CONSISTENCY	
Method	**Example Studies**
Intra-theme consistency (in %)	Girres and Touya 2010
Inter-theme consistency (in %)	Girres and Touya 2010
Spatial similarity in multi-representation	Hashemi and Ali Abbaspour 2015
Calculation of topological errors (Floating Island, Almost Junction, Intersection without junction etc.)	Wu et al. 2021

POSITIONAL ACCURACY	
Method	**Example Studies**
Buffer zone method	Haklay 2010, Jokar Arsanjani et al. 2013, Graser et al. 2014
Average Euclidean distance	Girres and Touya 2010
Hausdorff distance	Girres and Touya 2010
Distance between corresponding intersections of a road network	Antoniou 2011
Minimum bounding geometry (grid-based)	Forghani and Delavar 2014
Directional distribution (grid-based)	Forghani and Delavar 2014
Median centre distance	Forghani and Delavar 2014
Maximum deviation (grid-based)	Brovelli et al. 2016
Vector adjustment model using stereo imagery	Canavosio-Zuzelski et al. 2013
Box-counting dimension difference	Wu et al. 2021
Link accuracy	Wu et al. 2021

TEMPORAL ACCURACY	
Method	**Example Studies**
Calculation of correlation between no. of contributors and mean-capture date	Girres and Touya 2010
Calculation of correlation between no. of contributors and mean version of captured object	Girres and Touya 2010

THEMATIC ACCURACY	
Method	**Example Studies**
Percentage (%) of correct classification	Girres and Touya 2010
Percentage (%) of specific values existing in tags	Girres and Touya 2010, Antoniou 2011
Levenstein distance	Girres and Touya 2010

In a different study for Germany, LUDWIG ET AL. (2011) compared OSM data with a commercial Navteq-dataset. They implemented a feature matching method to find corresponding road objects in both datasets and to perform an object-based comparison of geometric and thematic attributes. Several measures were applied to evaluate completeness and positional accuracy. Overall, their study showed higher deviations in positional accuracy and a lower completeness for rural compared to urban areas. Regarding the attribute completeness in several OSM datasets from German cities, they could show a higher completeness in populated areas compared to uninhabited ones for selected attributes like *name*, *oneway* and *maxspeed*.

While above named quality assessments were based on a comparison of the data itself, MONDZECH AND SESTER (2011) were among the first to assess OSM quality by comparing the fitness for use of the data for a specific purpose and thus focussing on the quality element usability. They compared simulated routes on the OSM network of the German region of Hannover to simulated routes on the official German topographic dataset ATKIS. Concluding that the calculated route length can be a suitable indicator for data quality, they found OSM being of lower quality in the rural study area while showing similar quality to ATKIS in more populated areas.

NEIS ET AL. (2012) analysed the development of the OSM street network in Germany from 2007 to 2011 and its fitness-for-use in car navigation compared to a commercial TomTom dataset. Although they found that OSM was still missing about 9% of data related to car navigation compared to the commercial dataset, they predicted a comparative OSM road network for the end of 2012. For countries like Germany, where the OSM project is well developed, they concluded a comparable quality of OSM to commercial geodata regarding temporal and positional accuracy. They also illustrated the fact that roads in OSM are mapped predominantly in order of their hierarchy. Motorways are usually the first road-types to be mapped completely, followed by municipal roads, streets in residential areas and all other roads such as tracks and smaller paths.

Automated feature-based matching methods, explicitly created for VGI, were presented by KOUKOLETSOS ET AL. (2012) and MOHAMMADI AND MALEK (2015b). In a subsequent case study of KOUKOLETSOS ET AL. (2012), the authors compared OSM data with authoritative datasets from Ordnance Survey for urban and rural areas in the greater London and Newcastle regions to evaluate data completeness. Their results agree with former UK studies of HAKLAY (2010) with OSM proving to be more complete in urban than in rural areas.

Comparing the OSM database with a NavInfo reference dataset for Wuhan in China, WANG ET AL. (2013) showed a better positional accuracy and completeness of OSM data for high-level roads than for the rest of the database, where quality was relatively poor.

GRASER ET AL. (2014) presented a toolbox for street network comparison based on QGIS and performed a comparison between the OSM database for the greater Vienna region and the official Austrian reference graph “GIP”. They calculated key indicators for positional accuracy and completeness and additionally suggested new indicators for turn restrictions and one-way street comparisons to test street network quality for routing. Their results showed that both street networks were of comparable length and positional accuracy was high at least for the analysed motorway geometries. Regarding attribute completeness, OSM showed significantly less complete attribute information for the attributes *name* and *maxspeed*.

In FORGHANI AND DELAVAR (2014), the authors propose a new metric for OSM quality analysis by evaluating the consistency of OSM data with that of a reference map from the municipality of Tehran, Iran. Their approach is based on common heuristic methods that are evaluated for both datasets in separate grids before evaluating the consistency using a fuzzy logic approach. They finally visualise the uncertainty between the two datasets in a grid-based map, verifying that the quality of OSM maps in the study area is fairly good but varies throughout the dataset.

Comparing OSM cycling features with reference data from local planning agencies, HOCHMAIR ET AL. (2015) assessed the completeness of bicycle trail and lane features in the OSM database for several US cities. Bicycle trail and lane objects were filtered based on the OSM keys *highway, bicycle,* and *cycleway*. They found OSM having a higher trail or lane density in only a few cities. Additionally, separately mapped bicycle trails were more complete than bicycle lanes that belonged to a road.

By developing and applying a network-matching algorithm, ABDOLMAJIDI ET AL. (2015) compared the OSM road network of Scania (Sweden) with data from the Swedish National Road Database. Their study revealed that the OSM data had a completeness of 87% in urban and 69% in rural areas.

With the help of an automated procedure based on GRASS GIS, BROVELLI ET AL. (2016) compared the Paris OSM road network against the French official road dataset and calculated measures for positional accuracy and completeness. Results show a high positional accuracy and completeness of the OSM road network in Paris, with OSM having a higher total length of roads than the official dataset.

To tackle the challenging computational problem of comparing two or more spatial datasets against each other, BROVELLI ET AL. (2017) introduced a methodology for the automated comparison of OSM road networks against a complimentary reference dataset. Their method only considers the comparison of geometries but not of their attributes or other information. As a result, it computes quantitative measures for completeness and positional accuracy of an OSM dataset.

WU ET AL. (2021) assessed positional accuracy, logical consistency, and completeness of the OSM road dataset for Allegheny County and Pittsburgh in the United States by calculating geometry-based and topology-based metrics. They introduced a comprehensive framework for quality assessment, consisting of common metrics like road length difference and median centre distance as well as two novel geometry-based quality metrics for VGI linear features: Box-counting dimension difference and Link accuracy. As one result of their analysis, highly complete linear features were clustered in urban areas while logical consistency showed to be higher in rural areas.

Overall, results of extrinsic studies showed that completeness and positional accuracy of OSM datasets have improved over time and are generally high but vary significantly between features and regions. In those studies that comprised both urban and rural regions, OSM data often showed to be very accurate and complete in densely populated areas, while being usually less complete and accurate in rural regions (HAKLAY 2010; ZIELSTRA AND ZIPF 2010; GIRRES AND TOUYA 2010; KOUKOLETSOS ET AL. 2012; ABDOLMAJIDI ET AL. 2015; WU ET AL. 2021). Regarding attribute completeness and thematic accuracy, several studies showed a rather low quality of OSM datasets (GIRRES AND TOUYA 2010; LUDWIG ET AL. 2011; HOCHMAIR ET AL. 2015). It should be noted that all studies mentioned above reflect a specific state of analysed OSM data at the time of the respective analysis. Consequently, results do not represent the current quality of OSM.

To the best of the authors' knowledge, no extrinsic study has been conducted yet explicitly for rural areas. Rural regions have usually been included in studies when analysis took place on a country level or urban areas were compared to suburban areas (see appendix A1).

Questioning authoritative data

Over time, OSM data got more and more accurate in many parts of the world, in some regions even more detailed and complete than authoritative datasets (ANTONIOU AND SKOPELITI 2015). Consequently, the assumption that external data is a valid choice for representing ground truth can be problematic. Authoritative data is often updated infrequently which can create gaps between the map data and conditions in the real world. In fact, a major advantage of VGI is its up-to-dateness. Additionally, the availability of reference data is often limited because of license restrictions, high costs or non-existence. Along with VGI-specific challenges such as heterogeneity of data and contributors, spatial bias, lack of specifications, a dynamic update-process and the inconsistency of contributions, this may result in extrinsic methods being insufficient to assess certain aspects of VGI quality (FONTE ET AL. 2017).

1.4.2 Intrinsic studies and methods

The importance of identifying VGI specific indicators and techniques has been realized at an early stage by FLANAGIN AND METZGER (2008). Addressing the limitations of extrinsic methods, researchers have developed several frameworks (BALLATORE AND ZIPF 2015; REHRL AND GRÖCHENIG 2016; WU ET AL. 2021) and numerous intrinsic methods and tools to evaluate VGI data quality. In general, all intrinsic methods aim at assessing data quality by analysing the evolution of the data itself. Consequently, the focus of intrinsic approaches is to be found in the analysis of the editing history of OSM datasets and corresponding contributors. Whether quality assessments are data-focused, contributor-focused or a combination of both, intrinsic evaluations always depend on the availability of sufficient history data (KEßLER AND GROOT 2013). In the end, intrinsic quality evaluation methods can only be used to approximate data quality. Absolute statements on data quality are merely possible with the help of a ground truth reference dataset (BARRON ET AL. 2014).

Many tools have been developed to automatically analyse VGI specific quality indicators based on OSM Full History files (ROICK ET AL. 2011; BARRON ET AL. 2014; AUER ET AL. 2018; MINGHINI AND FRASSINELLI 2019; RAIFER ET AL. 2019). Other approaches were developed to analyse tagging practices (MOONEY AND CORCORAN 2012; DAVIDOVIC ET AL. 2016; ALMENDROS-JIMENEZ AND BECERRA-TERON 2018; HALL ET AL. 2018), to support these annotation processes (VANDECASTEELE AND DEVILLERS 2013; ALI ET AL. 2014), to estimate specific ISO quality elements like positional accuracy (MOHAMMADI AND MALEK 2015a), completeness (CAMBOIM ET AL. 2015) or thematic accuracy (ALGHANIM ET AL. 2021) or to propose new VGI specific quality indicators like “trustworthiness” (KEßLER AND GROOT 2013; FOGLIARONI ET AL. 2018; SEVERINSEN ET AL. 2019), “aggregated expertise” (MUTTAQIEN ET AL. 2018) or “reliability” (TEIMOORY ET AL. 2021).

In most studies, the intrinsic assessment of OSM road networks served as a case study to test a newly developed method. Only a few intrinsic studies were conducted explicitly on OSM road networks and no intrinsic analysis has specifically analysed OSM road networks in rural regions.

JILANI ET AL. (2014) proposed a machine learning model that learned geometrical and topological characteristics of different street classes to support a correct road type classification. The developed model was applied to two different OSM street network datasets from London and East Sussex in the UK by analysing values associated with the *highway*-tag. Results demonstrated that automated learning of road classes only works for certain road types.

BARRON ET AL. (2014) made a fundamental attempt on quantifying quality in OSM by developing the iOSMAnalyzer tool, which takes the OSM Full History Planet File as input and generates a range of 25 intrinsic quality indicators focusing on the spatial-temporal evolution of the data (see figure 6). For the cities of Madrid, San Francisco and Yaoundé, they exemplarily calculated four intrinsic indicators: Road network completeness, positional accuracy of the dataset, house number completeness and the geometric representation of natural polygons. To estimate the completeness of the road networks, they calculated the development of the road network length over time. While they considered the datasets of Madrid and San Francisco to be close to completion due to a stable length of the network and an active community, in the case of Yaoundé no statements on the road network completeness were possible without the use of a reference dataset.

Figure 6: Overview of the iOSMAnalyzer's intrinsic quality indicators (BARRON ET AL. 2014).

To specifically address semantic aspects of quality, BALLATORE AND ZIPF (2015) developed a framework that allows the measurement of conceptual quality and thus the interpretability of a VGI dataset. The framework identifies six dimensions of conceptual quality: Accuracy, completeness, compliance, consistency, granularity, and richness. In a subsequent case study, they exemplarily analysed the conceptual compliance of OSM road networks from Germany and the UK with the attributes defined in the OSM Wiki website. The analysis was restricted to objects tagged with at least one *highway*-tag and showed an average conceptual compliance of 93%.

SEHRA ET AL. (2017) built an extension to the QGIS Processing Toolbox to assess the quality of OSM line features calculating the following indicators intrinsically: Network length completeness, attribute completeness, semantic accuracy and road network feature completeness. On the example of Punjab (India), above named indicators were calculated. Attribute completeness was evaluated based on the tags *name* and *maxspeed* and resulted in a completeness of only 18.48%.

MOBASHERI ET AL. (2017) assessed the completeness of OSM data regarding sidewalk information to analyse the fitness for use of OSM for routing and navigation of people with limited mobility on the example of Heidelberg, Germany. Through an extrinsic analysis they evaluated the completeness of sidewalk data in an OSM dataset of Heidelberg (Germany) by comparing it to a reference dataset. Afterwards, they intrinsically analysed the completeness at attribute level, i. e. a road segment was complete if it had a sidewalk tag assigned to it and incomplete if it didn't. Results showed that only 17.6% of all sidewalks have been mapped in OSM.

NASIRI ET AL. (2018) aimed at improving the positional accuracy and completeness of the OSM road network by developing a five-stage intrinsic approach based on a Voronoi diagram. For validation, they compared the latest version of OSM data and the result of their new method to an external reference dataset from Teheran, Iran. As a result, the novel approach enhanced positional accuracy from 82.5% to 95.3% and completeness from 86.2% to 97.1%, proving that the quality of road features can be improved with the help of OSM data history.

The ability of contributors to map OSM road features was used as a proxy for data quality by JACOBS AND MITCHELL (2020). On the example of Ottawa-Gatineau in Canada, they bundled contributors into different user cluster groups, assuming each of them has generalizable associations with data reliability. For verification, they extrinsically calculated the overlap between an authoritative reference dataset and the OSM segments last edited by different cluster groups based on the extrinsic buffer zone method (GOODCHILD AND HUNTER 1997). Unsurprisingly, segments that were last edited by the most experienced mappers demonstrated the highest overlap with the reference dataset (76.2%).

With the aim of developing a machine learning approach to better predict road types in OSM, ALGHANIM ET AL. (2021) assessed the classification correctness of the *highway*-tag for all drivable road types in London, UK. They obtained an average classification correctness of 84.12%.

1.4.3 OSM in protected areas

While many studies have shown that populated areas usually receive more contribution to OSM datasets and demonstrate a better data quality than rural regions (see 1.4.1), no research has specifically evaluated the quality of OSM-data in protected areas, yet.

The only study dealing with OSM in protected areas was conducted by HENNIG (2017). Focussing on recreational infrastructure data, the author analysed the possibilities of using OSM in visitor management of Berchtesgaden national park in Germany. The analysis revealed a need to increase the amount of data held in the OSM database and to expand the OSM tagging system to allow the description of different types of nature-based recreational infrastructure. To profit from the benefits of OSM, national park administrations are required to actively contribute to the OSM community of their region. Consequently, personnel that has or builds skills in handling OSM data is needed within protected area management organisations.

Many German national parks have realized the need of dealing with the OSM database for their territory. According to SELTMANN AND ZINK (2021), some even already have employees that are actively involved in the OSM community and aim to enhance digital visitor guidance by improving the quality of the OSM database. A major task lies in the provision of additional information like access restrictions, that is specifically important in protected areas. Primarily through adding new or correcting false tags, relevant information is added to the OSM database.

A major problem authorities face when dealing with OSM is the mismatch between the OSM community's "on-the-ground-rule" and the legislation of a national park. The correct attribution of access restrictions is particularly important for ways that are dedicated to renaturation, where an undisturbed development of nature needs to be ensured. Usually, a way that is not accessible to the public is attributed with the *access*-tag (*access=no, access=private etc.*) and/or with more specific tags like *foot=*, bicycle=** etc. Due to the "on-the-ground"-rule, OSM contributors only mark a way e. g. with *access=no* if the inaccessibility is clearly indicated by a barrier or sign in the real world. As a consequence, park authorities would have to put a sign or barrier on every way where access is not allowed to ensure its correct representation in OSM, which is often impossible due to local conditions such as hard ground or the number of ways that would have to be marked. Deletion of objects in OSM is only allowed when they are not existing in the real world. This does not apply to most ways under renaturation as they are not completely renatured yet (which would mean that they are not visible and existing anymore and thus could be deleted). In agreement with local OSM communities, some national parks have started to use the lifecycle prefixes *abandoned:highway*=* and *disused:highway*=* to mark those ways that are dedicated for renaturation. This way, data is kept in the database and at the same time ways can be marked as "not in use" any more (SELTMANN AND ZINK 2021).

1.5 Research question and objectives

The main intention of this work is to assess the quality of OpenStreetMap road network data specifically in German national parks as a proxy for protected areas. It aims to give an answer to the questions how the quality of OSM road networks within protected areas can be analysed and how reliable OSM is within national parks in Germany.

The following sub-objectives have been defined for the evaluation process:

1. To present an easy-to-apply method and suitable quality indicators for the assessment of OSM road networks in protected areas.
2. To apply the developed method to a quality analysis of OSM datasets in German national parks.
3. To show typical characteristics of OSM datasets and particularly road networks in German national parks.
4. To investigate if and to what extent data quality differs between OSM and an authoritative reference dataset with respect to road networks.
5. To demonstrate the usability of OSM-data in the context of route planning for different outdoor activities in a protected area.

2. Methods

Data can be used for many different purposes. OSM data is often used for routing and navigation as well as for geocoding or producing maps, for example. As interpretations of the data can differ significantly or even be contradictory, the context in which the data is grounded is essential for its use (MOCNIK ET AL. 2017). The context of this OSM quality assessment is defined by the study areas and their characteristics as national parks as well as by the area of application "outdoor routing".

All methods used in this work follow a direct approach as OSM data is compared with internal and/or external reference data (see 1.3.2). At first, the evolution of OSM features as well as the number of active contributors are investigated for all study areas in an exploratory analysis. Both "indicators" proved to provide valuable first insights into the quality of an OSM dataset (NEIS ET AL. 2013, 2012) and help to draw conclusions on the characteristics of OSM in protected areas.

The subsequent, more detailed quality analysis of the road networks goes along with the evaluation process defined by ISO as shown in figure 5. Accordingly, quality units, elements and measures are defined in the following.

2.1 Specification of data quality units

Corresponding to the ISO quality evaluation process, a quality unit is composed by a scope and one to several quality elements. For this study, two data quality units are defined:

Unit 1 – General Road Network Analysis
The first quality unit aims to give an impression of the OSM road network quality within ten land-based German national parks (see table 5) by examining the quality elements feature completeness, currentness (also called "up-to-dateness") and attribute completeness. Feature completeness is chosen as it is the quality element that has been evaluated in almost all OSM road network quality analysis and thus allows comparison to past studies. Currentness is a valuable intrinsic quality indicator as after the initial collection, further maintenance of data is essential for a high quality and up-to-date dataset (VAN EXEL ET AL. 2010). Finally, the assessment of attribute completeness allows to include legislation into the quality analysis that is very relevant in the context of protected areas.

Unit 2 – Case Study "Usability of OSM for outdoor routing in the Bavarian Forest NP"
In this quality unit, the usability of the OSM road network and thus its fitness-for-purpose in the context of outdoor routing is analysed on the example of the Bavarian Forest National Park. The Bavarian Forest is chosen as a study area because its administration is already actively contributing to the local OSM dataset and was willing to provide a reference dataset for the road network. Additionally, in the Bavarian Forest a very specific legislation applies to the national park area which well represents the challenges of dealing with data quality in a protected area.

As shown, completeness of features as well as of attributes in the OSM road network plays a crucial role in routing and navigation applications and has been subject of many OSM quality assessments. Thus, the quality elements feature completeness, attribute completeness and classification correctness (thematic accuracy) are assessed with respect to outdoor routing for the activities cycling, hiking and skiing. Additionally, the activity of the local OSM community is investigated to get a proxy for trustworthiness of the OSM data.

2.2 Specification of quality measures

2.2.1 General Road Network Analysis

To evaluate the absolute quality of ten different national parks, extrinsic methods and the use of reference datasets of all these areas would be required. Due to the unavailability of reference data plus the limited scope of this thesis, an extrinsic approach is not applicable for this part of the study. Thus, intrinsic measures are chosen to assess the quality elements defined for this quality unit and give an approximation to the quality of road networks in the study areas.

Feature Completeness
According to BARRON ET AL. (2014), the **growth of a road dataset** can serve as a proxy to approximate feature completeness. A category of OSM roads can be stated as close to completion, if the monthly increase in length is very small. The development of the OSM road networks is calculated to visualise their growth and draw conclusions on their completeness.

Currentness ("Up-to-dateness")
The visualisation of an **element's last modification** is a possible way to represent its currentness (BARRON ET AL. 2014) and thus is used to infer on the up-to-dateness of the OSM road networks.

Attribute completeness

The evaluation of attribute completeness always requires a precise definition of the attributes under investigation, as datasets can show a very high completeness of some attributes while having a very poor completeness of others. Thus, tags on which completeness should be evaluated are defined first. Attribute completeness is then assessed by calculating the presence of each defined tag in the dataset, meaning how many elements carry the required tag, also called **tag ratio.**

2.2.2 Case Study "Usability of OSM for outdoor routing in the Bavarian Forest NP"

The usability of the OSM road network for outdoor routing is assessed with the help of reference data and extrinsic as well as intrinsic methods. The following quality measures are evaluated:

Feature completeness

Having a reference dataset at hand, a widely used method to measure feature completeness is to perform a length comparison (see table 3). Following this recognized approach, the **length and error rates of excess or missing items** are calculated as indicators for possible errors of commission or omission by comparing the OSM road network with the reference road network.

Attribute completeness

The existence or absence of attributes, also named "tag presence" in several OSM studies, is similarly evaluated like feature completeness. It can be expressed by the **length and error rate of items with excess or missing attributes** in the dataset under investigation. Using the established method also in this study, attributes to be investigated are defined before a comparison of tag presence in both datasets is performed. Some selected tags are evaluated intrinsically by calculating the **tag ratio**.

Classification correctness (Thematic accuracy)

Whether an object has been attributed or classified correctly can only be determined by using a reference dataset that represents ground truth for the classes/attributes to be analysed. By calculating the percentage of incorrectly classified objects, also called **misclassification rate**, a statement on the thematic accuracy can be made. This requires a preceding feature matching which ensures the correct overlay of the dataset under investigation with the reference data. Where feature matching is not possible, visual inspections of the overlay are performed to approximate classification correctness.

Trustworthiness

Assuming that many contributors ensure a better data quality (HAKLAY ET AL. 2010), trustworthiness can be measured with the following parameters: User reputation, versions, users, confirmations, tag corrections and rollback (Keßler and Groot 2013). An all-encompassing analysis of trust as a proxy for data quality would have to take all these parameters into account, which goes far beyond the scope of this thesis. Thus, here trustworthiness is solely analysed in terms of the number of versions of an object. Keßler and Groot (2013) presume, that a revision of an OSM element aims to improve the feature's quality. Consequently, more versions of an OSM element indicate a higher trustworthiness. This is intrinsically evaluated by calculating the **average number of versions per OSM road segment**. Additionally, to better understand the structure of the local OSM community, user activity is analysed by evaluating the **share of edits** of the most active contributors in the study area.

2.2.3 Summary of selected quality measures

The selection of intrinsic measures to evaluate feature completeness, attribute completeness and currentness is based on the framework for intrinsic OSM quality analysis developed by Barron et al. (2014).

Extrinsic measures to assess feature completeness, attribute completeness and classification correctness in the case study are derived from the current ISO-standard on data quality (ISO19157:2013) and based on well-established methods that have been used in past extrinsic OSM quality assessments (see table 3).

Table 4: Quality measures used in the study.

Quality Unit	**Quality Element**	**Quality Measures**	**Approach**	**Reference**
1 – General Road Network Analysis	Feature Completeness	Growth of the OSM road network	Intrinsic	Barron et al. 2014
	Currentness ("Up-to-dateness")	Last modification of elements	Intrinsic	Barron et al. 2014
	Attribute Completeness	Tag Ratio	Intrinsic	Barron et al. 2014
2 – Case Study	Feature Completeness	Length of excess / missing elements	Extrinsic	ISO19157:2013; Girres and Touya 2010; Haklay 2010
		Error rate of excess / missing elements (length)	Extrinsic	ISO19157:2013
		Tag Ratio	Intrinsic	Barron et al. 2014

	Attribute Completeness	Length of elements with excess / missing attributes	Extrinsic	GIRRES AND TOUYA 2010, LUDWIG ET AL. 2011; ISO19157:2013
		Error rate of elements with excess / missing attributes	Extrinsic	ISO19157:2013
	Classification Correctness	Misclassification rate	Extrinsic	ISO19157:2013
		Visible classification errors in the overlay, expressed in maps and textual statements	Extrinsic	
	Trustworthiness	Average number of versions per road segment	Intrinsic	JACOBS AND MITCHELL 2020; KEßLER AND GROOT 2013
		A contributor's share of edits	Intrinsic	

2.3 Study Areas & Data

2.3.1 National Parks in Germany

Protecting a landscape by a National Park status was an idea born in the United States in 1872, when the Yellowstone Area became the first designated National Park worldwide. Europe's first National Parks were declared in Sweden (1909) and Switzerland (1914). In Germany it took until 1970 for the first national park to be opened in the Bavarian Forest. Since 2015 there is a total of 16 national parks in Germany, four of them being of marine character and twelve to be found on the German mainland (see figure 7).

"National parks are large natural or near-natural areas protecting large-scale ecological processes with characteristic species and ecosystems, which also have environmentally and culturally compatible spiritual, scientific, educational, recreational and visitor opportunities." (DUDLEY 2013)

According to the international protected area management categories of the IUCN (International Union for the Conservation of Nature), the primary objective of a national park (= IUCN category II) is to protect natural biodiversity along with its underlying ecological structure, to support environmental processes and to promote education and recreation at the same time (DUDLEY

2013). To meet IUCN-guidelines, at least three quarters of a protected area must be managed according to its main objective. For national parks this means that 75% of their area must correspond to a largely natural state and may not be subject to any use contrary to the protective purpose. Additionally, the area must be large enough to contain one or more complete ecosystems.

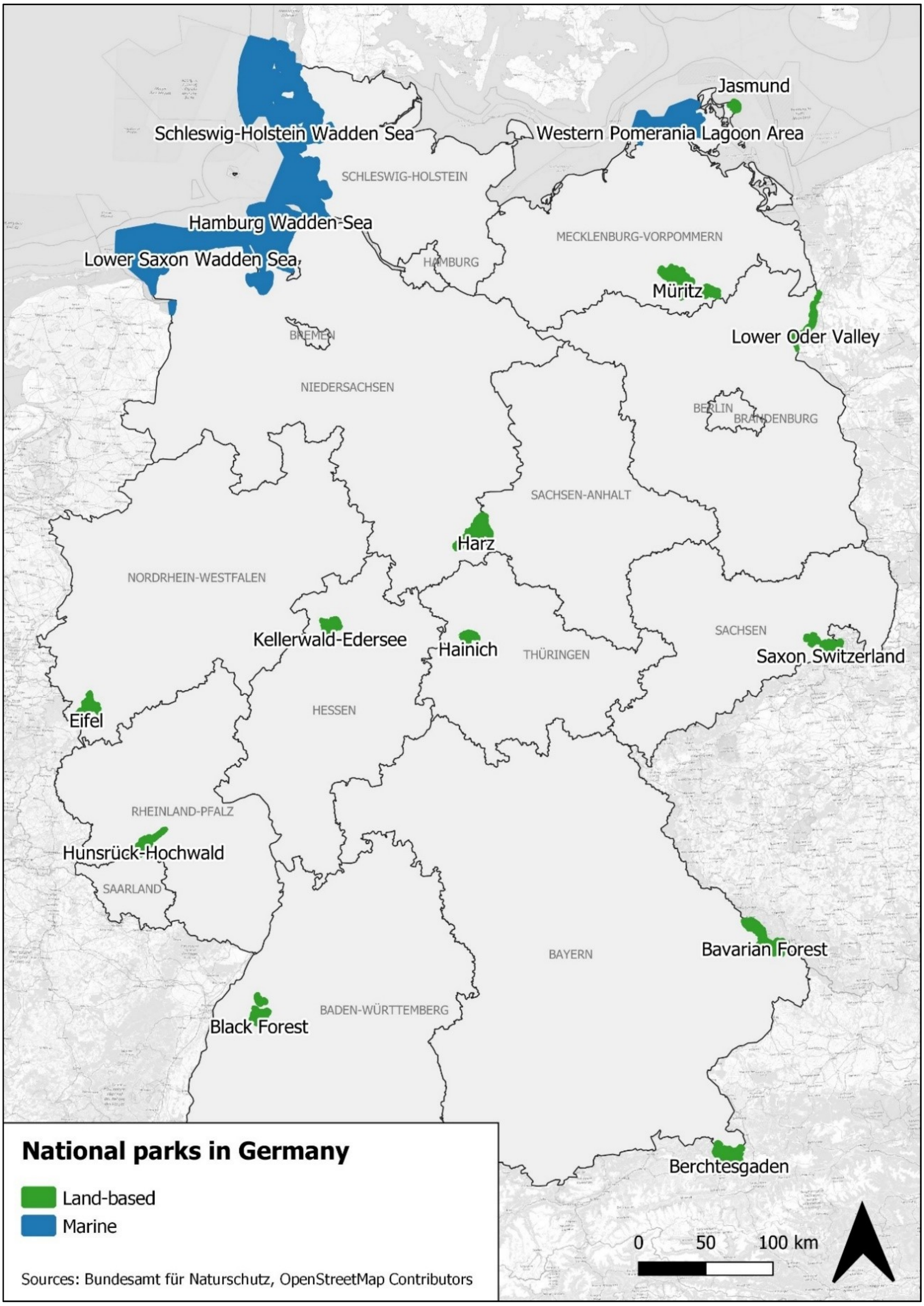


Figure 7: National parks in Germany.

National parks in Germany are designated by the federal states in consultation with the federal government. The legal basis is § 24 of the Federal Nature Conservation Act, which incorporates the IUCN-standards into national law. Most German national parks are currently still so called "developing national parks", meaning they only partially meet the criteria for large-scale, undisturbed natural development. According to the World Commission on Protected Areas (WCPA), in a developing national park transitional measures to support the area's natural development may be implemented up to 30 years after its inauguration. After 30 years at the latest, 75 percent of the national park area is to be left undisturbed for conservation and free of human actions and usage.

Table 5: German National Parks assessed in this study.

National Park	**Year of inauguration**	**Area (in km²)**
Bavarian Forest	1970	243
Berchtesgaden	1978	210
Jasmund	1990	30
Harz	1990 (Saxony-Anhalt) and 1994 (Lower Saxony), fusion in 2006.	247
Saxon Switzerland	1990	93
Müritz	1990	322
Lower Oder Valley	1995	104
Hainich	1997	75
Eifel	2004	108
Kellerwald-Edersee	2004	57

Geodata

As the study focusses on the road network within national parks, marine areas are not relevant in this context. Additionally, the two youngest land-based national parks, Black Forest and Hunsrück-Hochwald are left out of the analysis. Borders of all other ten land-based national parks (see table 5) were received in ETRS89 / UTM zone 32N coordinate reference system (EPSG:25832) from the Web Feature Service "Schutzgebiete" (https://geodienste.bfn.de/ogc/wfs/schutzgebiet) provided by the German Federal Agency For Nature Conservation. For each national park, a separate file containing its borders was reprojected into WGS84 (EPSG:4326) CRS before being exported and saved in geojson-format for later analysis.

2.3.2 Bavarian Forest National Park

The Bavarian Forest National Park ("Nationalpark Bayerischer Wald" in German) is located in south-eastern Germany along the border of Bavaria and the Czech Republic. It was inaugurated in October 1970 as Germany's very first national park and extended in 1997 to its current size of about 242,5 km². Today, it forms the largest contiguous forest reserve in Central Europe together with the adjacent Šumava National Park in the Czech Republic. According to the National Park Authority, about 350 km of signposted hiking trails and more than 200 km designated bike paths are offered for recreational purposes and used by about 1.3 million visitors every year (NATIONALPARKVERWALTUNG BAYERISCHER WALD 2022).

More than 50% of the NP area, primarily in lower altitudes, are freely accessible for pedestrians on all existing trails. The usage of vehicles and bicycles as well as horse riding is limited to public roads and officially signposted biking trails and bridleways in the whole national park area. In particularly sensitive areas, so called "core zones" ("Kernzonen" in German), access for pedestrians is limited to public roads and officially signposted hiking trails between July 15th and November 15th. Additional rules apply to border crossings, border trails and some forest-free areas. All rights of access are regulated in the bylaw "Enactment on the restriction of the right of access in the Bavarian Forest National Park" (REGIERUNG VON NIEDERBAYERN 1997).

In addition to above named core zones, in five wildlife sanctuaries ("Wildschutzgebiete" in German) across the NP area, different wildlife species are specifically under protection (see figure 8). Access to these areas is restricted and regulated by several corresponding bylaws:

- Wildlife sanctuaries "Ahornschachten" (1.56 km²) and "Auwald" (1.63 km²): No access between Nov 15th and Apr 30th (LANDRATSAMT REGEN 1984).
- Wildlife sanctuary "Neuhüttenwiese" (2.19 km²): No access between Dec 1st and May 16th (LANDRATSAMT FREYUNG-GRAFENAU 1990).
- Wildlife sanctuary "Riedlhäng" (3.71 km²): No access between Dec 1st and Mar 30th. Access only on signposted hiking trails between April 01st and May 16th (LANDRATSAMT FREYUNG-GRAFENAU 1990).
- Wildlife sanctuary "Auerwild" (26 km²): Access only on signposted hiking trails and groomed cross-country ski tracks between Jan 1st and Jun 15th (LANDRATSAMT FREYUNG-GRAFENAU 1982).

Seasonal access restrictions apply to all ways within a wildlife sanctuary.

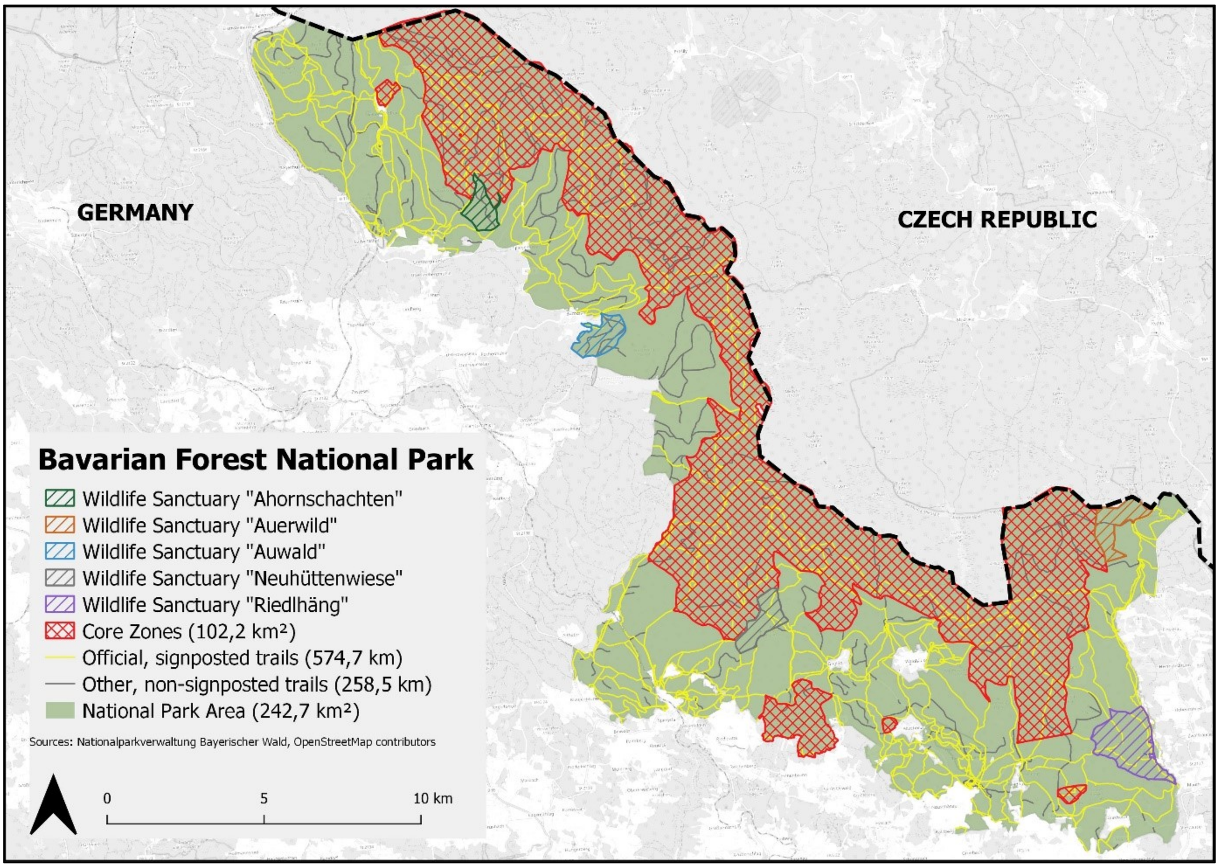


Figure 8: Zones and trails in the Bavarian Forest National Park.

Geodata

The Bavarian Forest National Park administration provided two datasets in Geodatabase-format:

- "Nationalpark.gdb", containing two feature classes that represent the national park's administrative boundary as well as the boundaries of all core zones and wildlife sanctuaries.
- "Wanderwege.gdb", containing two feature classes with all official trails within the national park area. Feature class "Markierte_Wege_DTK25" represents all signposted trails while feature class "Sonstige_Wege_und_Steige" contains all other, non-signposted trails where usage is partly allowed for pedestrians, partly completely prohibited, partly dependent on seasonal restrictions.

The provided Wanderwege.gdb-geodatabase comes with several attribute domains which pre-define the values for different attributes by using codes (see table 6). It comprises 575 km of signposted trails and an additional 258.5 km of other, non-signposted trails (see figure 8).

Table 6: Relevant attributes and values from the provided Wanderwege.gdb-file.

Feature Class "Markierte_Wege_DTK25"		
Attribute	**Value**	**Meaning**
"Wege" (= ways)	WNDWEG	Signposted way for hiking.
	RADWEG	Signposted way for cycling.
	REIWEG	Signposted way for horse riding.
	WRWEG	Signposted way for cycling and hiking.
	RRWEG	Signposted way for cycling and horse riding.
	WRRWEG	Signposted way for cycling, hiking, and horse riding.
"Wintersport" (= winter sports)	LOIPE	Cross country skiing track.
	RODBAH	Sledding track.
	LOIWWW	Cross country skiing track and winter hiking trail.
	WWW	Winter hiking trail.
	LOING	Cross country skiing track currently not prepared.
Feature Class "Sonstige_Wege_und_Steige"		
Attribute	**Value**	**Meaning**
„Renaturierung" (= renaturation)	KR	No renaturation.
	SOM	Shutdown without measures.
	SFG	Shutdown, water streams stay open.
	RZF	Dismantling into footpath.
	RZS	Dismantling into "Schlepperweg".
	VR	Complete renaturation.

The spatial reference of all provided datasets is the "ETRS 1989 UTM Zone 32N" (EPSG:25832) coordinate reference system. According to the national park authority, the datasets were created based on the German Digital Topographic Map 1:25,000 (DTK25). DTK25 is the official German database for topographic information like roads, railways, borders, buildings etc. and has an accuracy of +/- 10 to 20 m. In a few areas of the national park, inaccurate polylines were rectified by individual digitisation based on GPS measurements that have a positional accuracy of ca. 3 m.

2.3.3 OSM Data

For the exploratory as well as the general road network analysis (quality unit 1), OSM full-history data was accessed via the OpenStreetMap History Database and the ohsome API (see 2.4.1).

For the case study analysis in the Bavarian Forest NP (quality unit 2), the file "niederbayern-latest-internal.osm.pbf" was downloaded on Nov 12th, 2021 from Geofabrik GmbH's free download server https://download.geofabrik.de. The osm.pbf data format is the common format for the exchange of raw OpenStreetMap data. Geofabrik's osm.pbf-files contain all data and metadata available in OSM for the respective region at the time of the download. They do not contain information on past edits, though (GEOFABRIK GMBH 2022). The downloaded file covers the whole administrative region of Lower Bavaria (Niederbayern), including the national park, and contains all OSM data that was accumulated up to 11.11.2021 9:21pm.

It must be noted that "internal" OSM-files like the one used in this study contain personal data of OSM contributors. Thus, their usage is governed by EU data protection regulations. Personal information like *user* or *uid* (see table 1) are not allowed to be published anywhere outside the OSM community.

2.4 Tools

There are many tools available that work with OSM data with respect to its quality. Software like Osmose (http://osmose.openstreetmap.fr/de/map) or the OSM Inspector (https://tools.geofabrik.de/osmi/) is mainly used to improve the mapping quality of OSM. A different level of complexity is reached as soon as it comes to the question of fitness-for-purpose of the data. Consequently, tools must be picked that support the answer to the question: "Is OSM data good enough for my specific use case and within my area of interest?".

2.4.1 Ohsome OSM History Analytics Platform

Many studies have shown that analysing historical OSM data can provide meaningful insights into the evolution of OpenStreetMap and performing intrinsic analysis of its data can be a valuable way of quality evaluation (see 1.4.2). Nevertheless, operating on raw OSM full-history data required a certain amount of expert knowledge and could be a complex and time-consuming task in the past, especially when dealing with large areas of interest.

With the aim of making OSM history data more accessible for various user groups and different kinds of data analytics, a team of scientists at the Heidelberg Institute for Geoinformation Technology (HeiGIT) developed the **ohsome History Analytics Platform** (AUER ET AL. 2018). Central component of the ohsome-platform is the **OpenStreetMap History Database** framework (OSHDB) that has been developed at HeiGIT to optimize work with OSM data on a global scale. The OSHDB data model is designed for efficient storage of and access to OSM history data and allows users to investigate data evolution and user contributions in a flexible way (RAIFER ET AL. 2019).

To serve different user groups, the ohsome-platform provides access to the OSHDB on three different abstraction levels as shown in figure 9: At the deepest level, users have direct access to OSHDB-data, containing all individual historical OSM-elements, via Java. The OSHDB data layer provides deep and flexible access into the core of the database and its data structures but requires a good amount of programming knowledge. On a medium abstraction level, the OSHDB-API, also a Java-interface, is available for access to various aggregation functions.

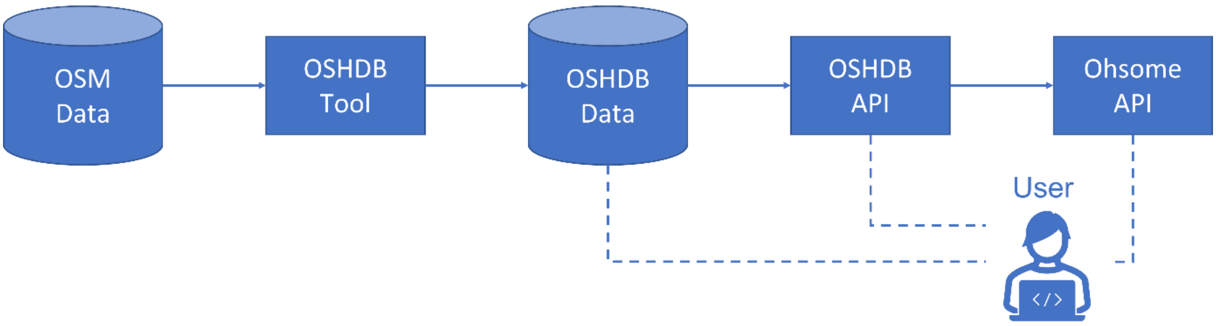


Figure 9: Schematic representation of the ohsome-platform components (AUER ET AL. 2018).

Ohsome API

The highest abstraction level of the ohsome-platform is the ohsome API, a generic web API for analysing OSM history data. It serves as the connection between the OSHDB-database and several potential or existing frontends such as web dashboards like the *OSM History Explorer* (https://hex.ohsome.org) or *ohsome Dashboard* (https://ohsome.org/apps/dashboard/). The ohsome API is designed according to the Representational State Transfer (REST) principle and communicates via HTTP (GET or POST). This way it allows also non-programmers to analyse OSM history data stored in the OSHDB-database. Requests to the API can be send directly by using different programming languages (Python, R) or tools (cURL, Swagger, ohsomeTools) or employing a dashboard-like application (AUER ET AL. 2018).

Figure 10 illustrates the data endpoints (marked in yellow) that can be used to calculate aggregated statistics on the evolution of OSM objects, extract historical OSM data and perform analysis on OSM-contributors and contributions.

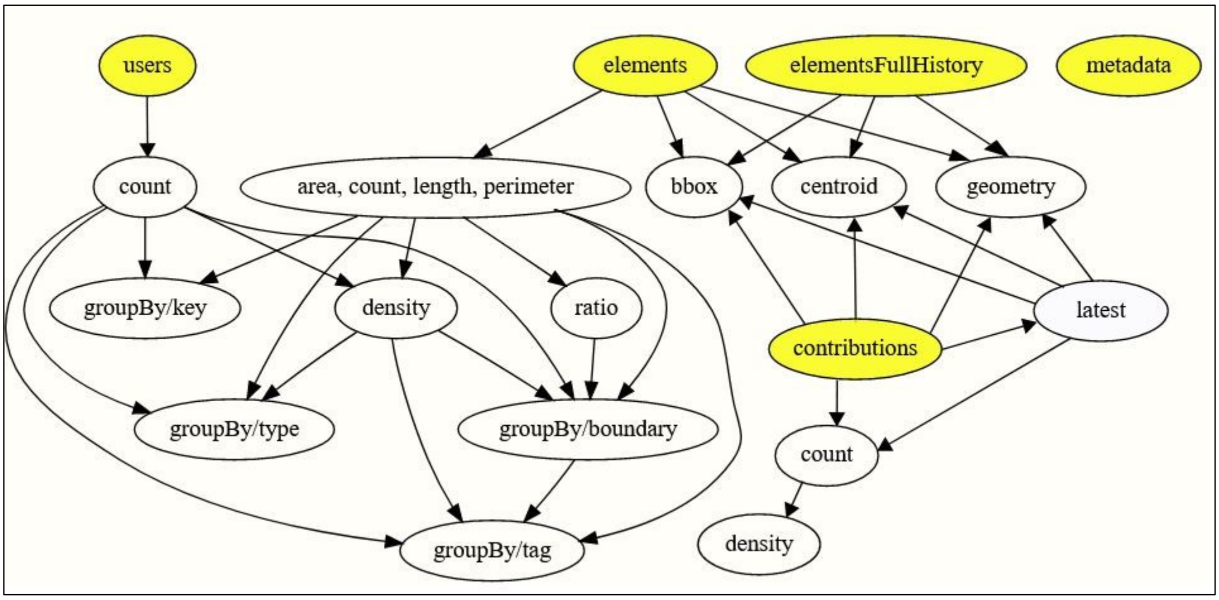


Figure 10: Visualisation of ohsome API Endpoints (HEIGIT GGMBH 2021).

Most of these endpoints provide aggregated data by making use of the *OSHDB snapshot view* and thus return the state of the OSM history data at specific given points in time. This can be used to count elements as well as calculate their length, area, or perimeter. Results are aggregated by given timestamps of the snapshot but can also be grouped by boundaries or attributes. *Contributions views* on the other hand return all modifications to OSM elements within a given time period. Results are aggregated temporally according to intervals and can also be grouped spatially or by attributes. A filter parameter enables users to create custom filter expressions for time, space and attributes. Spatial filters are available as bounding boxes, points with a respective radius or custom polygons. Attribute filters are based on the definition of OSM types (node, way, relation) and/or attributes/tags (AUER ET AL. 2018).

All ohsome-based analysis performed in this thesis made use of the current ohsome API version 1.6.2.

2.4.2 QGIS and QuickOSM Plugin

QGIS is a professional GIS application built on top of Free and Open Source Software. It is an official project of the Open Source Geospatial Foundation (OSGeo) and licensed under GNU General Public License. In this study, QGIS version 3.16 Long Term Release “Hannover” has been used for a variety of analysis and mapping tasks.

QuickOSM

QuickOSM is a plugin for QGIS that enables users to download OSM data right into QGIS. It is based on the Overpass API, a service that allows unfiltered access to OSM raw data and is used by many applications. In QuickOSM, custom queries can be run on an area of interest and local OSM files (.osm or .pbf) can be opened (see documentation at https://docs.3liz.org/QuickOSM/). Plugin version 2.0.1 has been used in this study to open the locally saved "niederbayern-latest-internal.osm.pbf"-file for further analysis of the dataset within QGIS.

2.5 Workflows

The description of workflows in this chapter corresponds to "Step 3 - Specification of the data quality evaluation procedures" in figure 5 as defined by ISO.

2.5.1 Exploratory Analysis

The intrinsically performed exploratory analysis is carried out for the study areas listed in table 5 and based on OSM data stored in the OSHDB-database. The analysis focusses on the OSM data types *node* and *way* only. *Relations* are excluded as they are rarely used compared to nodes and ways (Neis 2021) and almost always consist of one or both of them.

The ohsome API is used to gain statistics about OSM datasets of the study areas. Requests are designed individually and sent directly to the API. This is done by using POST requests through cURL (see figure 11), a command line tool for getting and sending data through URL addresses. cURL is chosen because it doesn't require any programming knowledge and thus meets the objective of developing an easy-to-apply method that can be utilised by anyone. Compared to the other available ohsome-tools that don't require programming skills (see 2.4.1), cURL offers most flexibility in using custom and complex polygons to define the study area.

cURL command:

```
curl --data-urlencode "bpolys@D:\MasterAnalysis\BfN_BayerischerWaldNP.geojson" -d "@D:\MasterAnalysis\06B_parameter.txt"
-X POST https://api.ohsome.org/v1/elements/count -o D:\MasterAnalysis\06B_Result_ELEMENTS_count_WAY_BayerischerWald.csv
```

Parameters defined in parameter-file:

```
format=csv&filter=type:way&showMetadata=yes&time=2008-01-01/2022-01-01/P1M
```

Figure 11: Example of a POST-request to the ohsome API via cURL.

Commands are sent using the command line interpreter of the local operating system (Microsoft Windows 10.0). Each command retrieves the query-parameters from a locally stored txt-file and the study area border from a locally stored geojson-file (see 2.3.1). The response is locally stored in csv-format and thereafter edited in Microsoft Excel for further analysis and visualisation.

The ohsome API endpoints *elements aggregation* and *users aggregation* are used to calculate statistics on the development of the OSM datasets and derive contributor statistics for all study areas plus the German cities of Straubing and Munich for comparison to more densely populated areas.

2.5.2 General Road Network Analysis

All calculations to evaluate the quality measures of quality unit one (general road network analysis) follow the same workflow as in 2.5.1, thus they are based on the OSHDB-database and make use of the ohsome API by sending POST-requests through cURL. Focussing on the OSM road network, only way-elements with a *highway*-tag are assessed.

Feature Completeness

Feature completeness is evaluated by intrinsically analysing the growth of the OSM road networks over time. Using the ohsome API endpoint *elements aggregation*, all road-elements are queried and grouped by road-type for the time period 01-01-2008 to 31-12-2021 and for each national park (see figure 12).

cURL command:

```
curl --data-urlencode "bpolys@D:\MasterAnalysis\BfN_BayerischerWaldNP.geojson" -d "@D:\MasterAnalysis\10_parameter.txt" -X POST
https://api.ohsome.org/v1/elements/length/groupBy/tag -o D:\MasterAnalysis\10_Result_ELEMENTS_length_groupby_tag_BayerischerWald.csv
```

Parameters defined in parameter-file:

```
format=csv&filter=type:way and highway=*&showMetadata=yes&groupByKey=highway&groupByValues=path,track&time=2008-01-01/2022-01-
01/P1M
```

Figure 12: Example of a POST-request used in the analysis of feature completeness.

The results are stored in csv-files that are then edited in Excel to visualise the road network's growth over time in a diagram. For further understanding of the results, additional queries are performed on the keys *abandoned:highway* and *disused:highway*, using the same workflow mentioned above but with different parameters. According to the OSM-Wiki, these keys, which are carrying a so-called lifecycle prefix, are used to map former *highway*-features that are still physically existing (visible in nature) but have been repurposed or are no longer usable (OSM6). These features have usually fallen into disrepair and could only be put back into operation with expensive efforts. Some national park authorities use them to tag paths that are determined for renaturation, thus were a usage is not allowed any more (see 1.4.3).

Currentness

To visualise an element's last modification date, POST-requests via cURL are performed using the ohsome API endpoint *contributions aggregation*, more precisely the sub-endpoint *contributions/latest/count*. This endpoint allows the aggregation of statistics on the latest contribution to all elements in the database, separated by the contribution types *creation*, *deletion*, *geometryChange* and *tagChange*. Statistics are calculated for the same time period as before and for each national park separately (see figure 13). The results are stored in csv-format. In a last step, the number of last modifications within the last 6, 12 and 24 months or older is being calculated within Excel and visualised in a diagram.

cURL command:

```
curl --data-urlencode "bpolys@D:\MasterAnalysis\BfN_BayerischerWaldNP.geojson" -d "@D:\MasterAnalysis\12a_parameter.txt" -X POST
https://api.ohsome.org/v1/contributions/latest/count -o
D:\MasterAnalysis\12a_Result_CONTRIBUTIONS__latest_count_geometrychanges_BayerischerWald.csv
```

Parameters defined in parameter-file:

```
format=csv&filter=type:way and highway=*&contributionType=geometryChange&showMetadata=yes&time=2008-01-01/2022-01-01/P1M
```

Figure 13: Example of a POST-request used in the analysis of currentness.

Attribute completeness

Using a predefined list of tags which characterise completely attributed roads is a valuable approach to assessing attribute completeness without reference data (BARRON ET AL. 2014). Thus, tags on which completeness should be evaluated have to be defined, first. As shown in 1.4, typical road attributes that have been investigated in the past are *name*, *maxspeed* and *oneway*-tags. Apparently, while they are all essential in the context of car navigation, they are not of importance in protected areas, where most highway-elements are expected to be of a road-type that does not provide any of the above-named attributes.

When looking at road networks of protected areas, the most important attributes are related to access rights and restrictions. Conflicts between recreationalists and nature conservation can only be avoided if users of the road network know about the restrictions in place. Therefore, a high attribute completeness in this context can be reached if restrictions are completely and correctly represented in the OSM database. Furthermore, and according to the OSM-Wiki, roads in OSM should ideally always be tagged with a *surface* and/or *tracktype* key to describe their condition (OSM5). This information also improves the quality of navigation. Therefore, attribute completeness within national parks will be assessed on the following attributes:

1) Condition of a way, represented by the keys ***surface*** and ***tracktype***. *Surface* provides information about the material of a highway-element's physical surface (e. g. surface=asphalt). The *tracktype* key measures how well-maintained a track or road is, particularly regarding the surface's firmness. It usually applies to highways of type *track* (*highway=track*) but is often used for non-tracks, too.
2) Usage by pedestrians. Within OSM, this is represented by the key ***foot***, which indicates the legal restriction for pedestrians (e. g. *foot=no* means access on foot is prohibited).
3) Usage by cyclists. Represented by the key ***bicycle*** in OSM, indicating the legal restriction for cyclists on the way it is assigned to (e. g. *bicycle=yes* means the way is accessible to cyclists, but not explicitly designed or signed for their use).

A list of recommended and most-common *key-value* pairs for all above named keys is provided in the OSM-Wiki (OSM2).

Attribute completeness is assessed individually for the four keys *bicycle*, *foot*, *surface* and *tracktype* using the ohsome API endpoint *elements aggregation* (see figure 14). This way, the presence of each tag in the OSM database of each national park can be assessed. After loading the results in Excel, tag ratio is calculated and visualised. For all keys, the tag ratio is calculated within all *highway*-elements.

cURL command:

```
curl --data-urlencode "bpolys@D:\MasterAnalysis\BfN_BayerischerWaldNP.geojson" -d "@D:\MasterAnalysis\18a_parameter.txt" -X POST https://api.ohsome.org/v1/elements/length/groupBy/key -o D:\MasterAnalysis\18a_Result_ELEMENTS_length_groupby_key_bicycle_BayerischerWald.csv
```

Parameters defined in parameter-file:

```
format=csv&filter=type:way and highway=*&groupByKeys=bicycle&showMetadata=yes&time=2008-01-01/2022-01-01/P1M
```

Figure 14: Example of a POST-request used in the analysis of attribute completeness.

2.5.3 Case Study "Usability of OSM for outdoor routing in the Bavarian Forest NP"

Pre-Processing of reference data:
As the reference data comes in gdb-format, the files are opened in ArcGIS Pro 2.9 for pre-processing. The feature class "Markierte_Wege_DTK25" comprises many ways outside the national park and has to be intersected with the feature class "Nationalpark" (incl. a 30 m buffer, see below) to receive a feature class consisting only of the polylines within the national park border. All four feature classes from the reference datasets are then exported in shapefile-format to make them usable in QGIS.

Additionally, to calculate certain quality measures, specific pre-processing steps are necessary and described in the respective workflows below.

Pre-Processing of OSM data:
The downloaded OSM-file "niederbayern-latest-internal.osm.pbf" is opened in QGIS with the help of the QuickOSM-plugin. The plugin allows to query the file for selected geometries and tags, thus only *way*-elements with a *highway*-, *abandoned:highway*- or *disused:highway*-tag are loaded into QGIS. In a next step, the file is intersected with the "Nationalpark"-shapefile to extract the OSM road network within the national park borders (Tool: QGIS Processing Toolbox → Vector overlay → Intersection) and reprojected to EPSG:25832 coordinate system. Visual inspection discloses a geometric offset of up to 30m between the reference and the OSM datasets, which is why a buffer of 30 m is used around the national park borders during the intersection. This way, also OSM-ways that are located right at the national park border are completely included. After pre-processing, 5,954 polylines representing a road or way are still present in the dataset.

Feature Completeness

The length and rate of excess or missing items is calculated by a unit-based *total road network length comparison* that is applied to the OSM and the reference datasets of the study area. The summed length of all road features within the national park border (incl. 30m buffer) is calculated for both datasets in QGIS with the *Sum line length* tool from the processing toolbox.

The same approach is used for a *classified road length comparison* performed for all abandoned and disused roads. In the reference dataset, the status of usage is defined in the attribute "Renaturation" of the file "Sonstige_Wege_und_Steige" (see table 6). Via *Select by attribute*, all ways with the values

“SOM”, “SFG” and “VR” are extracted. From the OSM-dataset, all ways holding an *abandoned:highway*- or *disused:highway*-key are considered.

To identify local differences in feature completeness, an additional *grid-based length comparison* is performed for all roads. First, a grid with 1 km cells is created within the national park borders. For easier calculation, the two reference datasets are merged into one layer or file respectively. Using the *Sum line length*-tool again, the total length of the road network is calculated for each grid cell and separately for the OSM and the reference dataset. In a final step, result layers are merged to calculate and visualise length differences between the two datasets per grid cell. Some regions, where results are particularly conspicuous, are further investigated by visual inspection.

Attribute completeness

Like for feature completeness, the length and rate of items with excess or missing attributes is calculated in a *road network length comparison* (see above). The most relevant attributes for evaluating completeness in protected areas are those on access rights and restrictions. As described in section 2.3.2, in the Bavarian Forest NP different periods of use are regulated by bylaws and dependent on whether a way is located in a core zone or wildlife sanctuary. Thus, the usage of the road network for different activities and modes of transport is highly dependent on seasonal access restrictions in many parts of the study area. Therefore, the presence of related tags for those periods of use and the allowed usage for different activities will be assessed:

1) Usage by cyclists:
    - Represented by the key ***bicycle*** in OSM (see 2.5.2). All ways tagged with *bicycle=yes*, and *bicycle=designated* are considered.
    - In the reference dataset, the accessibility for cyclists is defined in the attribute “Wege” of the file “Markierte_Wege_DTK25” (see table 6). Via the *Select by attributes*-tool, all ways where cycling is allowed (all ways which are attributed as „*RADWEG*“, „*WRWEG*“, „*RRWEG*“ or „*WRRWEG*“) are extracted. On ways from the file “Sonstige_Wege_und_Steige” cycling is generally not allowed.

2) Winter activities:
    - In OSM, ways offering winter sport infrastructure are attributed with the key ***piste:type***. All ways tagged with *piste:type=nordic* and/or *piste:type=hike* are considered.

- In the reference dataset, infrastructure for winter activities is defined in the attribute "Wintersport" of the file "Markierte_Wege_DTK25" (see table 6). All ways that are allowed and prepared to use in the winter (all ways which are attributed as "*LOIPE*", "*LOIWWW*", "*WWW*" or "*LOING*") are extracted. Ways from the file "Sonstige_Wege_und_Steige" do not carry any information on winter activities and thus are not considered.

3) Usage by pedestrians:
   As there are no general access restrictions for pedestrians in the Bavarian Forest NP, the right of access on foot depends on above-named seasonal restrictions and individual access-restrictions.
   - In OSM, such conditional restrictions are attributed with the suffix "***:conditional***". The time period(s) for which the restriction is valid are added in special syntax defined by the OSM-Wiki (Example: *access:conditional=yes @ (Jul15 - Nov15)*). General access restrictions are represented by the key *access=**.
   - In the reference data, information on the period of use can be derived from the boundaries of the core areas and wildlife sanctuaries and respective bylaws that regulate access restrictions.

   Completeness is assessed in detail for the core areas and all wildlife sanctuaries. Due to general seasonal access restrictions in these areas, a road must carry an ***access=no***, ***access:conditional=**** or ***foot:conditional=**** tag to be considered completely attributed. Additionally, roads that are designated for renaturation are evaluated. It is assumed, that they should not be used by any pedestrian within a national park as this would thwart the attempt to enable an undisturbed natural development. In OSM, the status of being dedicated to renaturation is predominantly represented with the lifecycle-prefixes *abandoned:highway*- or *disused:highway*. General access restrictions for all transport modes are expressed with the key *access*. Furthermore, mode-specific keys (e. g. *foot*, *bicycle, motor_vehicle*) allow individual attribution of access restrictions per transport type. To be considered completely attributed in this context, abandoned or disused roads must carry either an ***access***- or a ***foot***-tag.

Calculations are performed analogous to the calculations for feature completeness with the QGIS-tools *Select by attribute* and *Sum line length*.

Classification Correctness

For the evaluation of classification correctness, data from the OSM dataset needs to be overlayed by the reference data to determine if an element was classified correctly or not. Consequently, it is required to match elements from the OSM dataset with corresponding elements from the reference dataset to calculate absolute quality measures like a misclassification rate. Due to a high positional offset between the datasets, different attribution and significant differences in the number of elements present in each dataset, feature matching is not possible with standard tools but would require more complex algorithms to provide a reliable result. Thus, in this context visual inspection of the overlay as the most suitable and easy-to-apply method is performed in QGIS. This way, classification correctness is analysed for the following topics:

1) Cycling-specific access restrictions:
   - In OSM: Represented by the tags *bicycle=yes, bicycle=designated* and *bicycle=no*. No other *bicycle*-tags to be found in the OSM dataset.
   - Reference data: Represented by the values of the attribute "Wege" in the file "Markierte_Wege_DTK25". Values "*RADWEG*", "*RRWEG*", "*WRWEG*" and "*WRRWEG*" correspond to *bicycle=yes* in OSM, the values "*WNDWEG*", "*REIWEG*" and "*NULL*" correspond to *bicycle=no*.

2) Winter activities:
   - OSM: Represented by the tag *piste:type=nordic*. No other *piste:type*-tags to be found in the OSM dataset.
   - Reference data: Represented by the values of the attribute "Wintersport" in the file "Markierte_Wege_DTK25". Values "*LOIPE*", "*LOIWWW*" and "*LOING*" correspond to *piste:type=nordic* in OSM, values "*WWW*", "*RODBAH*" and "*NULL*" correspond to *piste:type=NO VALUE*.

3) Access restrictions for pedestrians:

   Because pedestrians are generally allowed to move on all existing but not only signposted trails (as long as the trail is not designated to renaturation or access is particularly restricted), classification correctness for hiking is determined by the correct representation of (seasonal) access restrictions. This is analysed exemplary for all five wildlife sanctuaries within the national park. Because information on access restrictions is not carried by each element in the reference dataset, a ground truth for general access restrictions is not available. Seasonal restrictions can be derived from the polygon borders (of the wildlife sanctuaries) instead, which makes it easy to match the OSM data with the reference data. Consequently, misclassification rates for seasonal access restrictions are calculated for each wildlife sanctuary additionally to visual inspections.

- In the OSM dataset, general access restrictions are represented by the key ***access***, seasonal access restrictions (for pedestrians) by the keys ***access:conditional*** and ***foot:conditional***.
- Within the reference data, access restrictions generally result from the bylaws (see 2.3.2). Hence, ground truth is represented by the borders of a wildlife sanctuary and the seasonal access restrictions defined in the respective bylaw.

Trustworthiness

Information on the current version of each road segment is saved in the attribute "osm_version" of the "niederbayern-latest-internal.osm.pbf"-file. Because values of this field are saved in string format, they must be copied to a new integer field before the average number of versions per road segment is calculated with the *Basic statistics for fields* tool in QGIS. The same calculation is performed over a 1 km grid to visualise regional differences. The grid, which has been created earlier to assess feature completeness, is joined with the OSM file via *Join attributes by location* tool. This way, the ID field of the grid is added to the OSM file so statistics can be calculated on the field "osm_version" per grid cell with the *Statistics by categories* tool.

Edit shares per contributor must be calculated on the attribute "osm_uid" that holds the unique id of each OSM contributor. The attribute "osm_user" is not practical because a user can change his username but not his id. Calculations are done in Excel after exporting the attribute table of the osm.pbf-file. Due to EU data protection regulations that apply to the personal data of OSM contributors (user and id) saved in the file, the values for the fields *user* and *id* are replaced with synonyms.

3. Results & Discussion

Quality analysis of geographic data often produces quantitative results for the assessed quality elements that can then be aggregated into a quality report. Nevertheless, especially when applying intrinsic methods, this is not always the case. According to ISO19157:2013, "a subjective evaluation of an element can then be expressed with a textual statement as a data quality descriptive result". Therefore, the outcomes of this analysis are presented in a mixture of quantitative results, visualisations and textual evaluations.

3.1 Characteristics of OSM datasets in national parks

The OSM datasets of German national parks have developed differently as figure 15 shows. Most areas saw a moderate and steady growth since 2008. Some steep increases like in the dataset of Eifel NP (see figure 15, nodes in 2016) are likely related to a bulk data import, automated edits (bots) or a collaborative mapping event.

In most national parks like the Bavarian Forest, Berchtesgaden or Jasmund, the number of nodes and ways has remained rather stable over the last years, only a small number of new elements has been added to their OSM database. This indicates an advanced saturation and good feature completeness. In some national parks though, namely Harz, Müritz and to some extent also Hainich, the OSM datasets still see major increases in the number of elements being added. Consequently, OSM can generally not be stated as complete or even close to completion in these areas.

Depending on their size, national parks have between 500 and 3,900 nodes and between 700 and 8,500 ways in their OSM database. Figure 16 shows that the typical density of elements lies between 6 to 35 nodes and 10 to 34 ways per km^2. An outstanding exception among the study areas is Saxon Switzerland NP, which shows a remarkably rich dataset with roughly 10,000 nodes and almost 40,000 ways, making up almost 110 nodes and 418 ways per km^2. For comparison: As an example of a typical mid-sized German town, the OSM-dataset of Straubing (67.6 km^2; 47,600 inhabitants) shows a density of roughly 65 nodes and 300 ways per km^2 (as of 31.12.2021). Munich (310.7 km^2; 1.5 million inhabitants) - as an example for a densely populated metropolis - has about 922 nodes and 1,069 ways per km^2 in its dataset by the end of 2021.

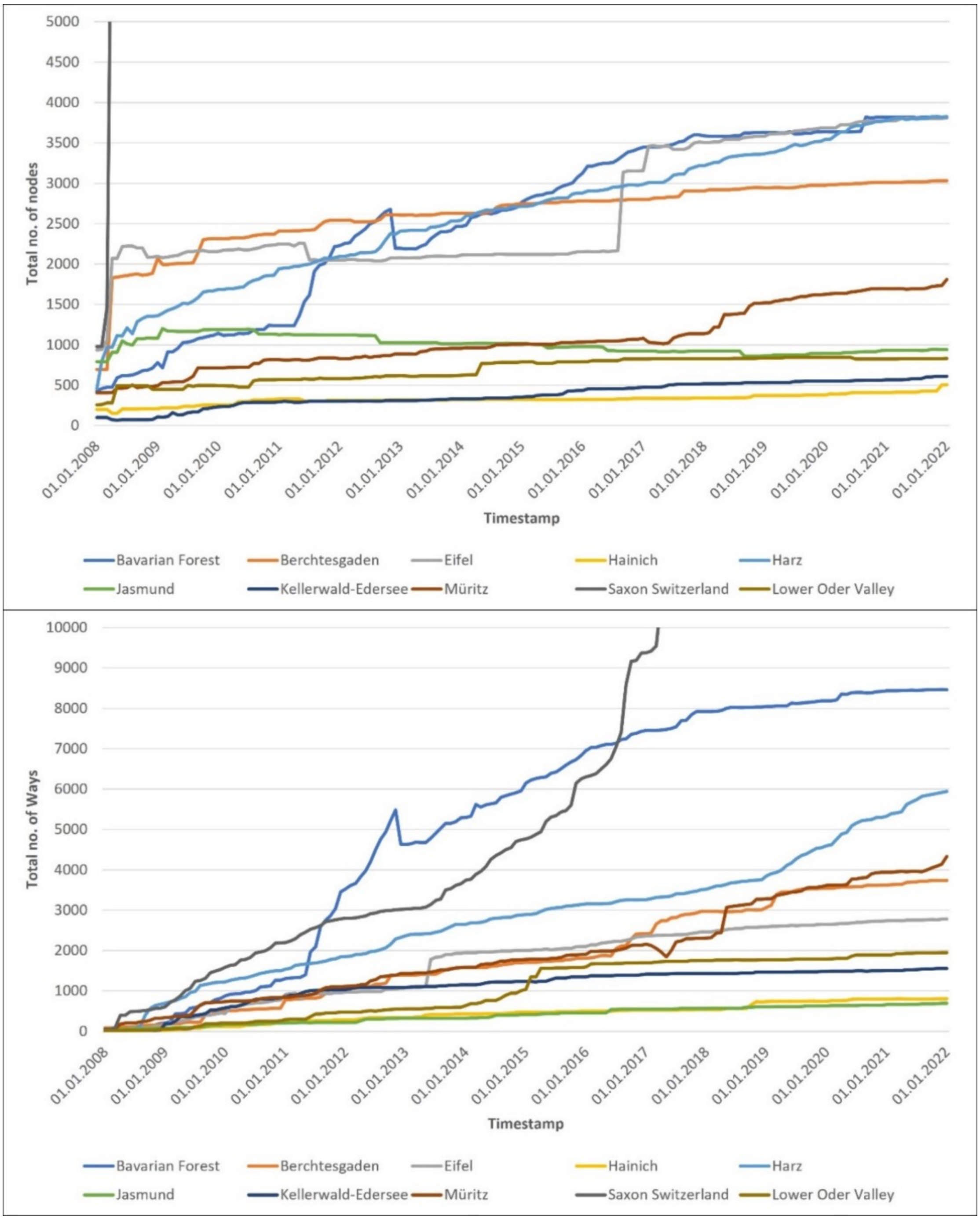


Figure 15: Development of the OSM database in German NP (top = nodes, bottom = ways).

OSM data in protected areas is contributed and maintained by a rather small number of users. Table 7 demonstrates that smaller national parks like Jasmund, Hainich or Kellerwald-Edersee have only around three to four distinct active contributors on average per month. The density of contributors differs widely between only 0.21 per km² in Müritz NP and 1.5 per km² in Saxon Switzerland NP. Comparing those numbers with more populated areas it stands out, that most national parks have a smaller OSM community than a mid-sized city like Straubing (9.26 active contributors on average per month, 1.62 per km²), and all NP have significantly less contributors than a metropolis like Munich (165.4 on average each month, 4.17 per km² in 2021).

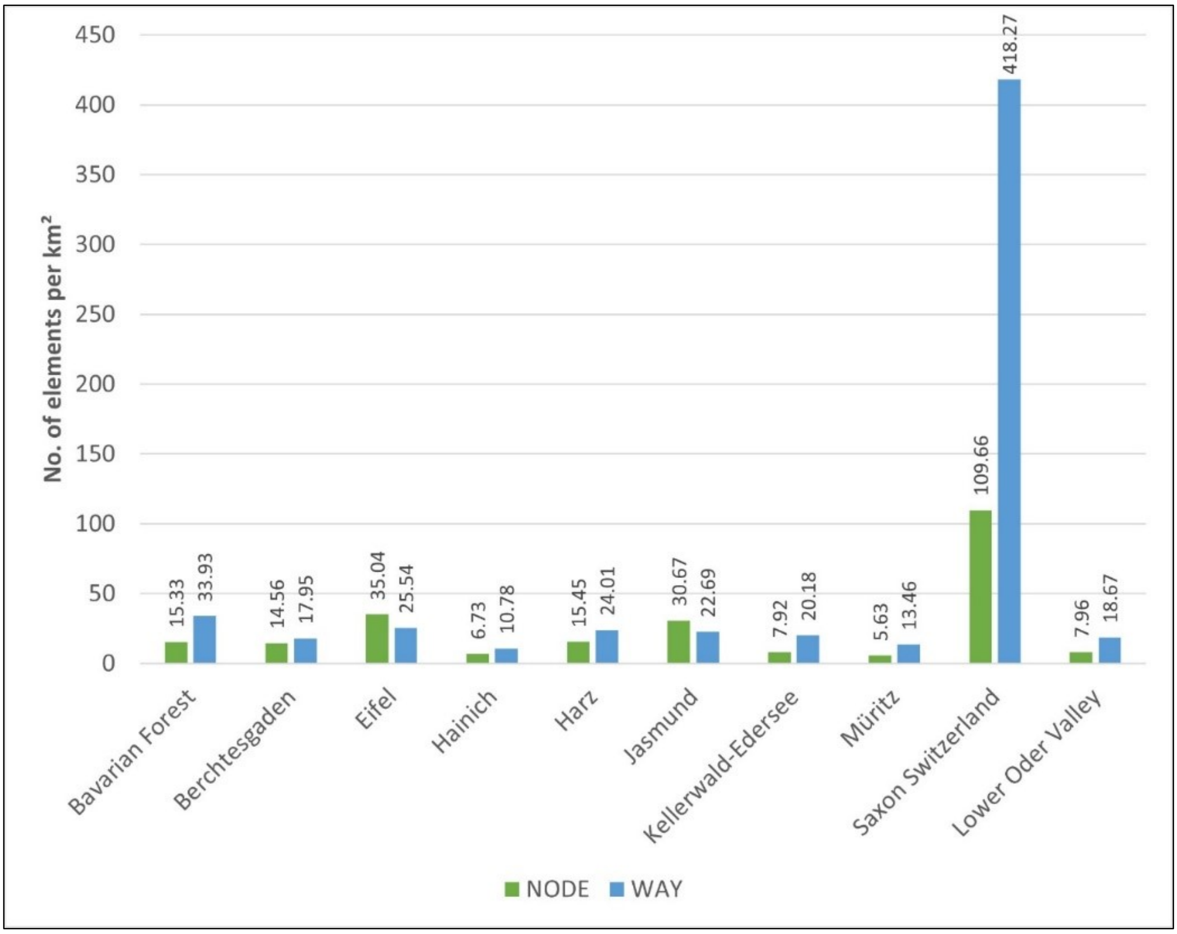
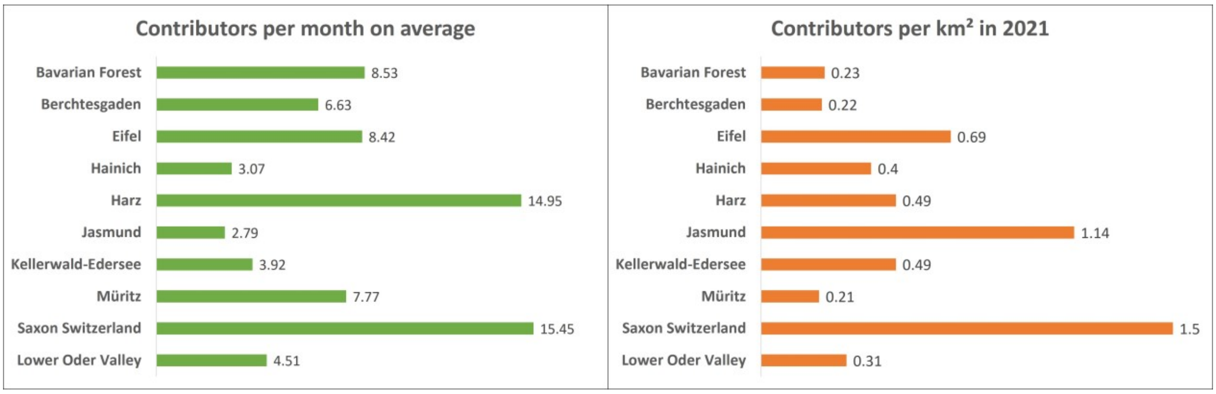


Figure 16: Density of elements in OSM database as of 31.12.2021.

Consequently, in many national parks the quality of the OSM dataset depends on the work and mapping experience of just a few persons. Referring to several investigations that demonstrated a correlation between a high number of active contributors and a stable, up-to-date OSM dataset of good quality (GIRRES AND TOUYA 2010; HAKLAY 2010; NEIS AND ZIPF 2012), this raises doubts about the quality of OSM in respective areas. To draw any conclusions on data quality, the local mappers' experience as well as their approach on how to collect and annotate data needs to be further investigated, though.

Table 7: Distinct active contributors to the OSM database.



3.2 General Road Network Analysis

The typical road density in German national parks lies between 2,000 and 5,000 meters per km² which is naturally less than in populated areas. While in the city of Straubing the road density amounts to almost 12,600 m per km², in Munich it is even close to 28,400 m per km². Again, Saxon Switzerland NP is an exception with a significantly higher road density than in all other national parks (see figure 17).

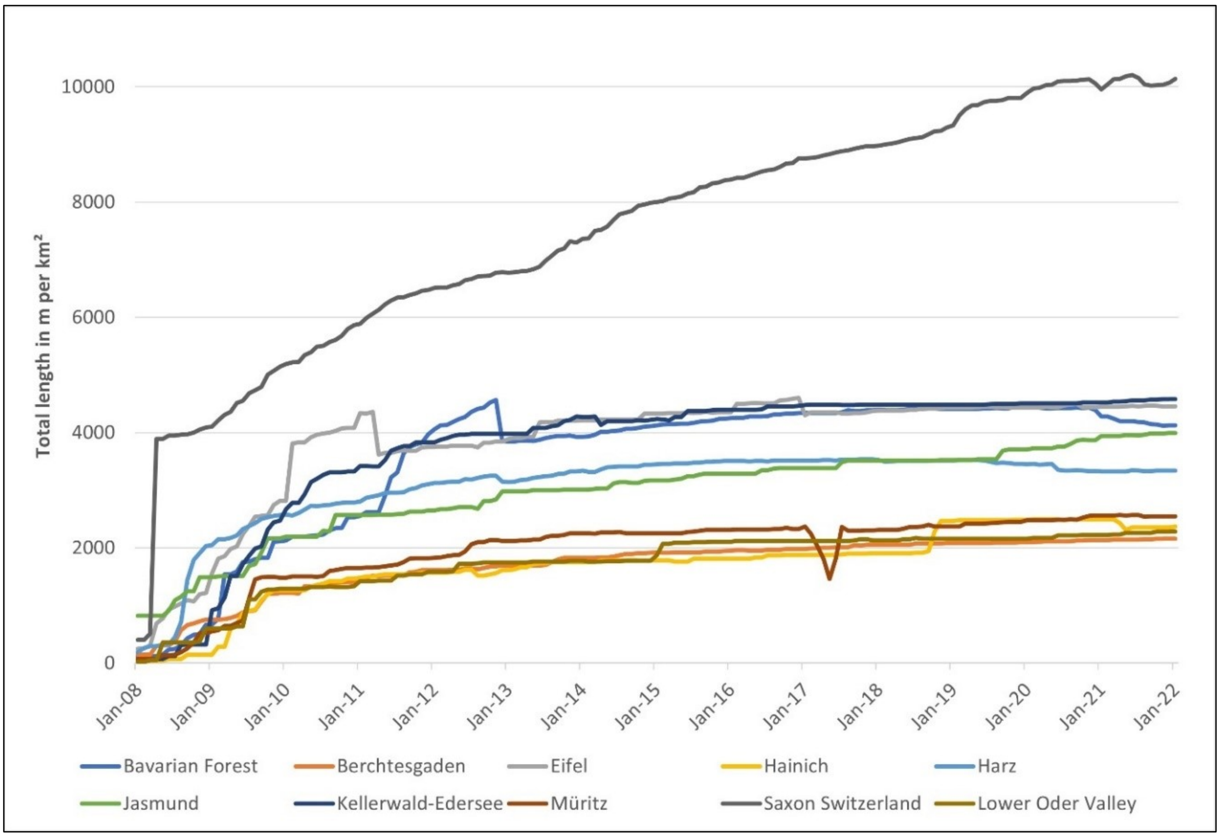
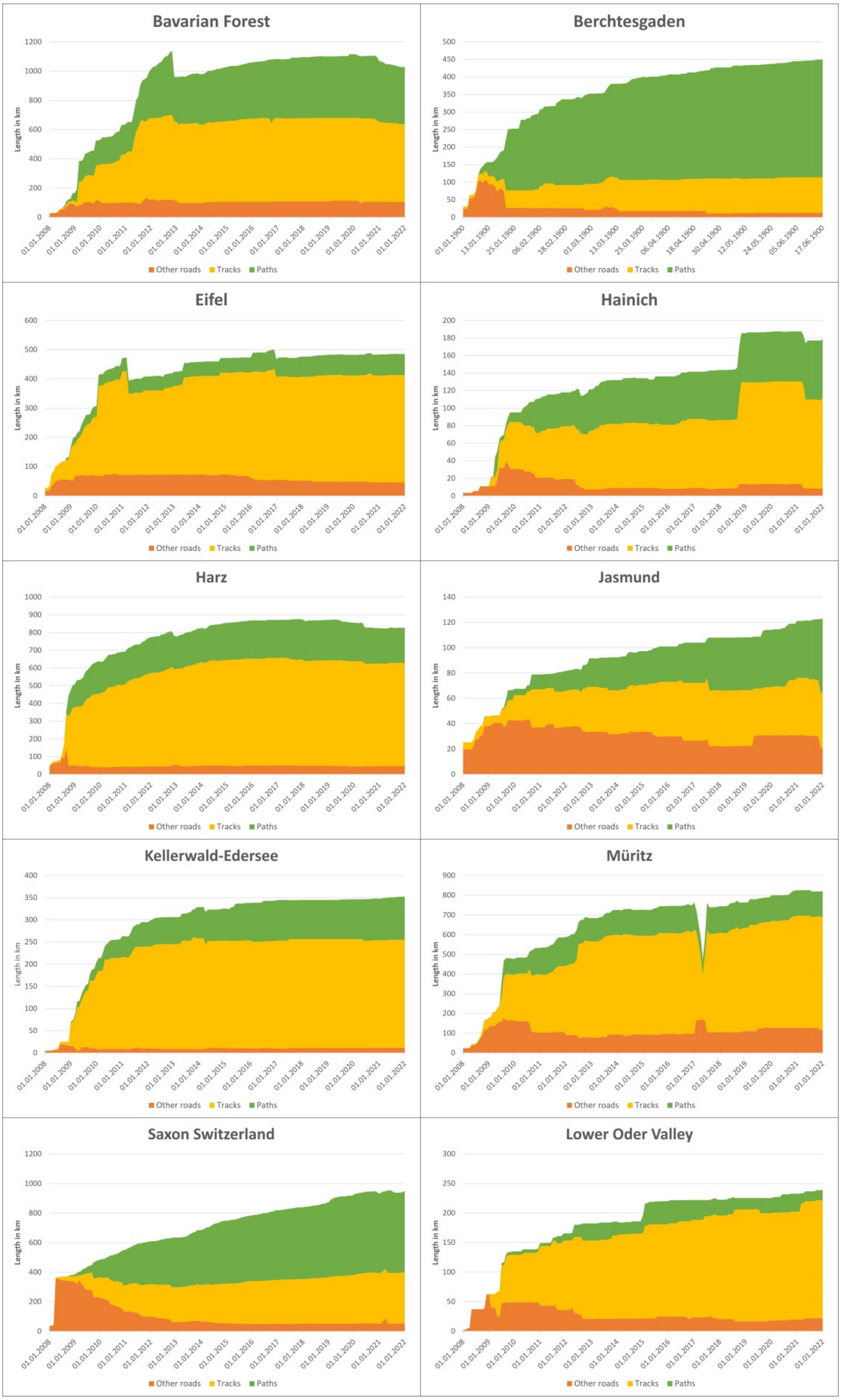


Figure 17: Development of the OSM road density in German national parks.

3.2.1 Feature Completeness

Table 8 shows the development of the total road network length for all analysed national parks. While the results clearly outline differences concerning feature completeness, they also illustrate a special characteristic of road networks in protected areas: A large majority of all roads (between 82.9% and 97.2%) are classified as paths or tracks. As it was to be expected, the road network of national parks is predominantly characterized by ways that are only accessible to the general public without any motorized vehicle.

Table 8: Development of OSM road network (=ways with a highway-tag) in German NP.



To some extent, the total road network length saw an increase or decrease in all analysed areas over the last years as can be seen in figure 18. Nevertheless, specifically in the national parks of Berchtesgaden, Eifel and Kellerwald-Edersee, growth rates have been very small or even close to zero. Minor changes in length are not necessarily caused by new roads but can also be related to a change of value within the highway key. Therefore, these road networks can be acknowledged as "close to completion" according to (BARRON ET AL. 2014).

The road networks of Jasmund, Lower Oder Valley, Müritz and Saxon Switzerland are still seeing a growth of 2-6% per year. This suggests a quite active community of contributors that is stilling actively mapping new road features to bring the road network closer to completion.

Having a look at specific road categories shown in table 8, it stands out that Hainich, Harz and Müritz NP show a high mapping saturation and thus likely a good completeness at least for *paths*, as the length development of all roads tagged with *highway=path* looks stable over the last years.

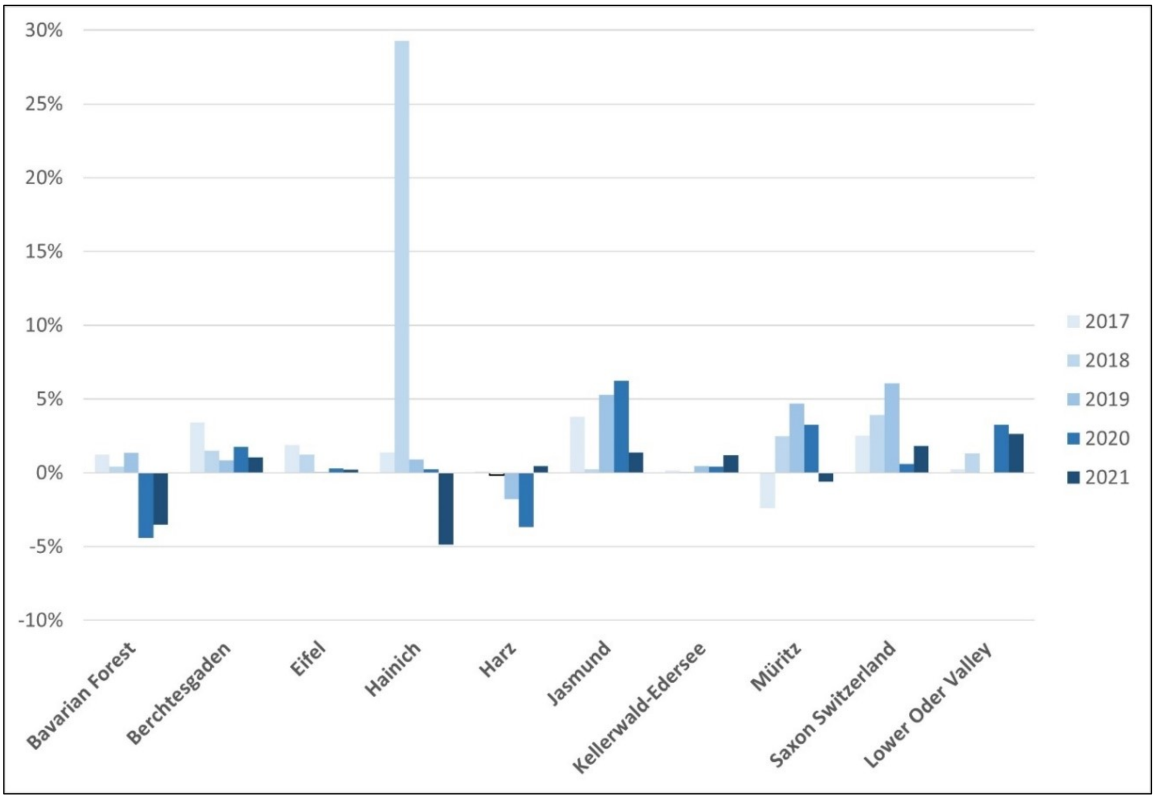
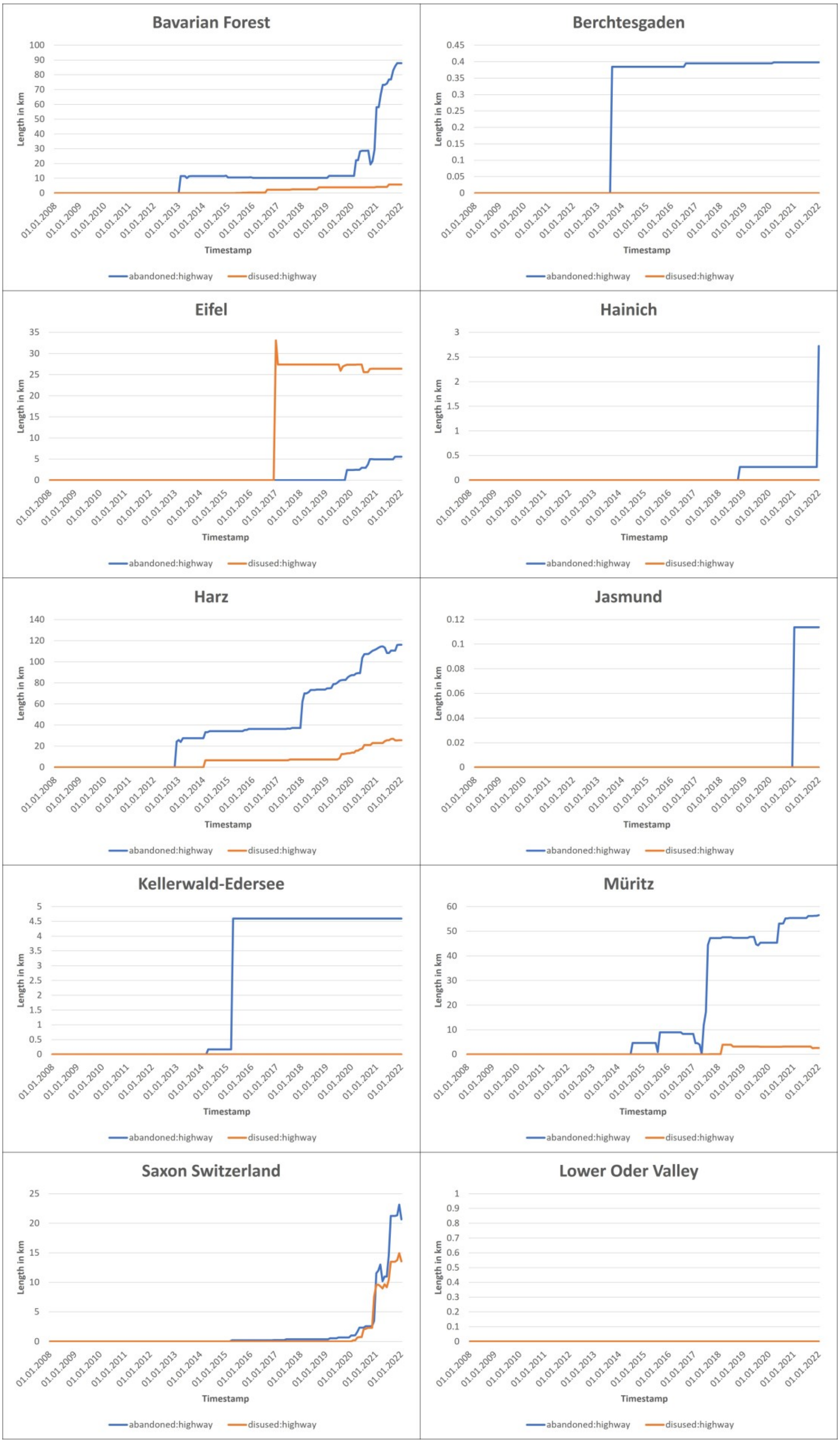


Figure 18: Yearly growth rates of the OSM road network length (all highway-elements).

The road networks of Bavarian Forest, Hainich and Harz national parks show a significant decrease of length in the last year(s). In the Bavarian Forest as well as in Harz national park, this can be explained by a parallel increase of the length of roads categorized as "disused" or "abandoned" (see table 9). Because the increase of lengths of those roads that are not in use any more matches the decrease of length of all roads in use, these road networks overall look rather stable, too. Thus, they can be referred to as possibly close to completion as well.

Table 9: Development of abandoned and disused highways in German national parks



A similar effect can be seen in the dataset of Müritz national park, where the road network length decreased in 2017 when at the same time the length of abandoned highways increased sharply. The Eifel and Saxon Switzerland national parks also show a significant growth of disused and abandoned roads in their datasets, but here this increase did not cause a decrease of length in the overall road network, meaning that the growth rates shown in figure 18 might even have been higher.

Berchtesgaden, Hainich, Jasmund, Kellerwald-Edersee and Lower Oder Valley national park have no or very few roads categorised as abandoned or disused in their datasets. This could be reasoned either in the fact that all existing ways are still in "active" use or mappers haven't attributed them accordingly. Generally, the growing network of abandoned and disused roads in many national parks reflects the development towards more undisturbed parts left for nature conservation and free of human usage. Where national park authorities are actively contributing to the OSM database (e. g. Bavarian Forest and Saxon Switzerland according to SELTMANN AND ZINK 2021), the amount of abandoned or disused highways in the OSM datasets has grown. It is thus unlikely that in those national parks, which do not have any ways marked as abandoned or disused in their OSM datasets, all ways are still in use. Rather, the absence of those lifecycle prefixes indicates that ways under renaturation haven't been attributed accordingly and the OSM datasets of those areas provide insufficient reliability regarding the allowed usage of all ways.

In the absence of above-named lifecycle prefixes, the decrease of the total road network length in Hainich NP in 2021 is probably due to a mass deletion of *track*-elements in the same year. The unusually high increase in 2018 (almost 30%) can likely be explained with a bulk import of road features, most of them being *tracks*, too (see table 8).

All in all, most national parks have a stable road network that is complete to a high extent. Nevertheless, it must be noted that a stable network length can also be caused by an unactive community and absent mapping activities. Thus, an analysis of contributor activity to answer the question "How up to date is the data?" can help to get a better understanding of feature completeness.

High feature completeness within national parks contradicts the results of past studies that concluded a poor completeness of OSM road networks in rural regions and insufficiency for using them for navigation (see 1.4.1). Given the time that has passed since then as well as the general growth of the OSM database and the many OSM-based services that arose in this field of application, a better coverage of OSM in the year 2021 was to be expected. The results of this analysis represent the steady growth of OSM and indicate a good suitability of OSM data for routing engines with respect to a complete road network also in rural areas.

3.2.2 Currentness

Figure 19 shows significant differences in the level of "up-to-dateness" between national parks. The "most up-to-date road network" is to be found in Jasmund NP. With almost one third of all elements being updated within the last six months the road network is unlikely outdated. This also applies to Lower Oder Valley NP and somewhat to Harz NP and Saxon Switzerland NP. In all other national parks, less than one third of all elements in the road network have been edited within the last 12 months. For comparison: In the road network of Munich, 32.8% of all elements were updated within the last year, in Straubing only 18.8%. Apparently, most NP datasets are not less up-to-date than those of more populated regions. A correlation between the local OSM community and the currentness of a dataset is not evident from these results as areas with a higher density of active contributors do not necessarily have a more up-to-date dataset.

In almost all national parks, the majority of recent edits is related to either geometry changes or tag changes (see table 10). This specifically applies to the areas with the highest currentness. Creations were seemingly more important in the first years of all the dataset's development, while deletions in general only account for a comparably small number of edits.

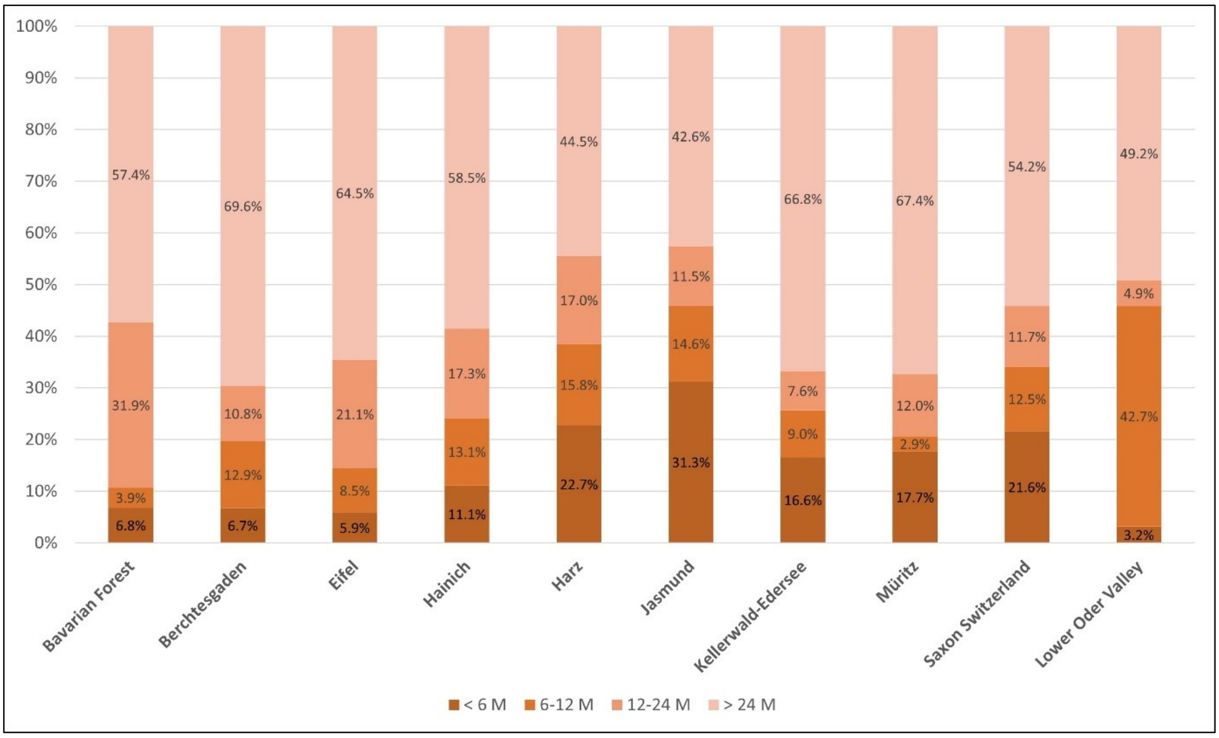
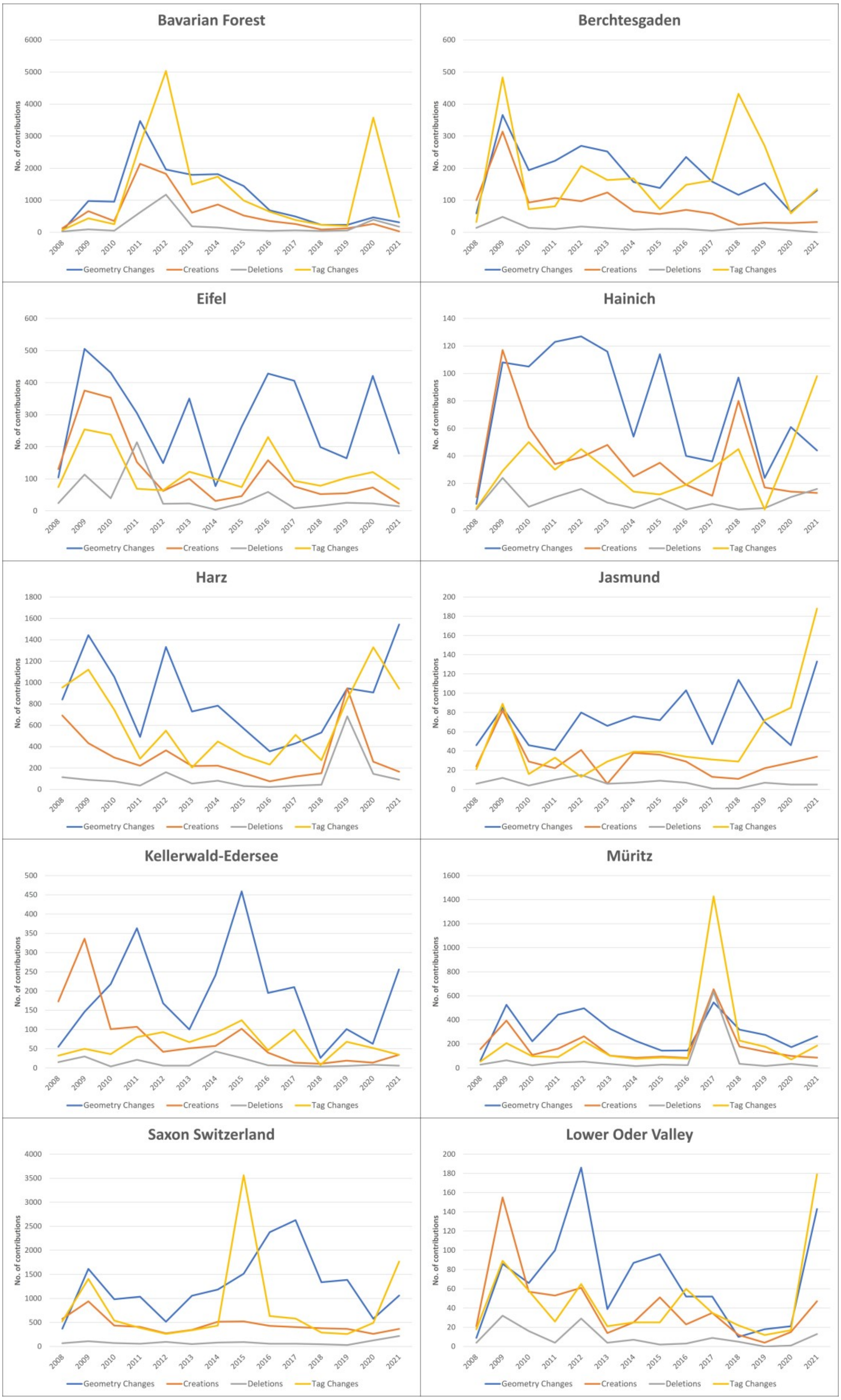


Figure 19: Time of latest contribution to highway-elements as of 31.12.2021.

Table 10: Total Contributions to the OSM road network in German national parks.



Changes of geometry as well as attributes can both have a significant impact on the quality of the dataset. According to BARRON ET AL. (2014), the concept of up-to-dateness assumes that the last editor of an OSM element is responsible for its correctness. This can be problematic for the quality when an element was already completely and accurately mapped in the past and the last edit decreased the quality. Thus, the indicator currentness gives only insights into the activity of contributors and helps identifying possibly outdated data but is does not allow an absolute statement on the reliability of a dataset for a specific purpose. Because currentness describes how well the data represents the current real-world situation, the results are also only valid for a short time. After e. g. a year, the result of an up-to-dateness-measurement is wrong as it does not represent the current situation any more but rather how well it represented the situation one year ago (ISO19157:2013).

To some extent, a high number of recent geometry changes can also be seen as a proxy for a lower positional accuracy in the past as it can be assumed that elements with a high positional accuracy should not see any more changes in geometry. On the other hand, the heterogeneity of OSM is well reflected here, as positional accuracy of OSM data is always dependent on the way data was collected (GPS signal preciseness, aerial imagery etc.). Table 10 shows that a lot of geometry changes are still being performed in most national parks. This could indicate a poor reliability regarding positional accuracy. However, a trustworthy statement on positional accuracy is not possible here as it requires further analysis of geometric accuracy with the help of ground truth reference data.

3.2.3 Attribute Completeness

The completeness of attributes within OSM road networks varies strongly between different areas and attributes. Figure 20 shows that most national parks have a rather high attribute completeness with respect to the description of a road's condition (keys *surface* and *tracktype*). The usage for cyclists and pedestrians is generally attributed a lot less. Only three national parks, namely Bavarian Forest, Berchtesgaden and Saxon Switzerland show a relatively high attribute completeness as roughly 80% of all highway-elements in those areas carry a *bicycle*-tag. Almost the same results are to be found for paths and tracks (*highway=path* and *highway=track*) only. At the same time, *foot*-tags are present a lot less in those same areas. For all other national parks, the tag ratio for *bicycle=** and *foot=** is below 50%, in many areas even below 30%. In general, *foot*-tags are used a lot less than *bicycle*-tags. All in all, the attribute completeness for pedestrian and cyclists' usage of roads is rather poor in most study areas.

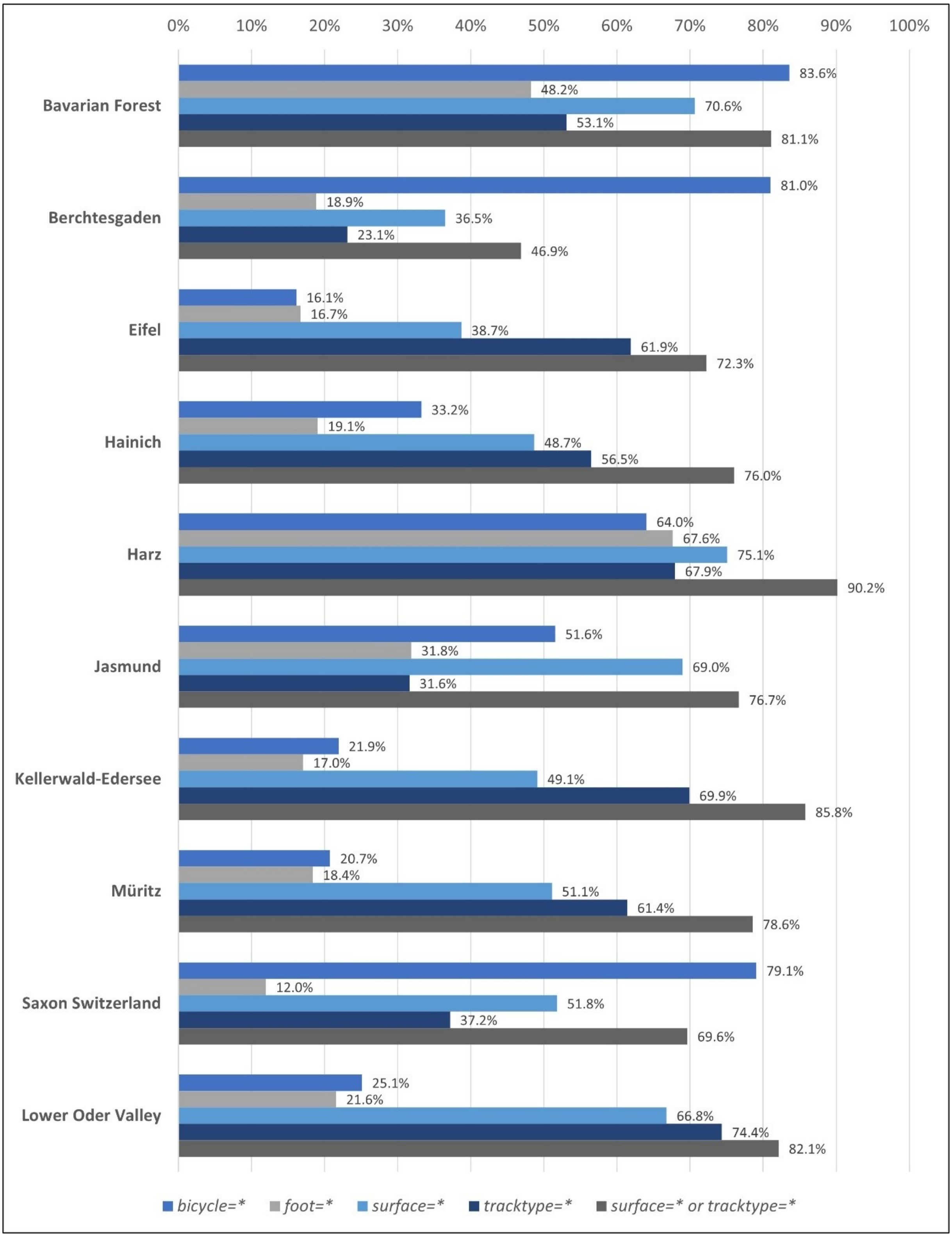


Figure 20: Tag Ratio for selected keys within all highway-elements as of 31.12.2021.

To what extent the results are comparably good or bad is difficult to answer, as no other study has analysed completeness specifically for the above-named attributes. In their fitness-for-use analysis of the OSM dataset for wheelchair routing services, MOBASHERI ET AL. (2017) studied the completeness of sidewalk information in OSM and found a tag ratio between 1.2% and 14.1% within five German cities. In comparison to their results, attribute completeness in this analysis looks comparably good. However, it must be noted that in the context of routing and navigation, every missing attribute can potentially lead to conflicts in the real world and thus a poor attribute completeness is problematic.

As described in 1.2.5, routing engines need to use some default on access restrictions if these restrictions are not tagged on the individual road. The OSM Wiki suggests *bicycle=yes* and *foot=yes* as a default for all highway-elements of the type *path* and *track* if they are not individually tagged otherwise (OSM9). While this is "just" the suggestion by the community and not necessarily common practice of routing engines, it underlines the importance of correct attribution. The absence of a single *bicycle*- and *foot*-tag can cause a user potentially ending up on a way he or she is not allowed to use but guided to by his/her routing app. This is particularly problematic in protected areas where, depending on the area and legislation in place, a lot of access restrictions might apply which can potentially be disregarded (without noticing it). Hence, a low presence of above-named tags indicates a poor attribute quality and questionable reliability of OSM road networks for outdoor routing engines. However, these results only provide a first insight. An in-depth analysis of the individual area including additional attributes (especially tags on general and seasonal access restrictions like *access=** and *access:conditional=**) is necessary to better understand attribute quality and fitness-for-use of the corresponding OSM network. Results of such deeper analysis for the Bavarian Forest NP are presented in the following.

3.3 Case Study "Usability of OSM for outdoor routing in the Bavarian Forest NP"

3.3.1 Feature Completeness

The total OSM road network within the Bavarian Forest national park is 322.2 km (38.27 %) longer than the reference road network provided by the national park administration (see table 11). This represents a very high error of commission. Similar errors are to be found when comparing the length of only the active as well as the disused road network. Obviously, the OSM dataset contains a lot more roads than the sum of both reference datasets. This shows, that the common OSM practice of mapping everything that is existing in the real world does not match with the approach of a national park authority to manage, map and publish only ways that are or have been officially usable for visitors. It also raises doubts on the general usability of the available reference data for extrinsically assessing feature completeness.

Table 11: Results of length comparison.

OSM	**Reference data**	**Error of commission / omission**	
		Length (km)	**Error Rate**
Total length of road network (km)[1]			
1164.05	841.85	+ 322.2	38.27 %
Total length of active road network (km)[2]			
1070.41	781.06	+ 289.35	37.05 %
Total length of disused road network (km)[3]			
94.33	60.79	+ 33.54	55.17 %

[1] OSM: All ways with *highway-*, *abandoned:highway-* or *disused:highway*-keys. Reference dataset: All ways present in the files "Markierte_Wege_DTK25" and "Sonstige_Wege_und_Steige".

[2] OSM: Only ways with a *highway*-key Reference data: Ways with the values *SOM*, *SFG* or *VR* in the attribute "Renaturierung" are not considered.

[3] OSM: Only ways with a *abandoned:highway-* or *disused:highway*-key. Reference data: Only ways with the values "SOM", "SFG" or "VR" in the attribute "Renaturierung".

Looking at the outcomes of the grid-based length comparison shown in figure 21, it stands out that - although in most parts of the national park the OSM road network is longer than the reference network - some regions show opposite results. Other regions, in contrast, show extreme errors of commission. The investigation of some examples discloses that where the reference dataset is longer the road network is usually sparse (see figure 22 B and D). On the other hand, the error of commission is highest where a denser road network is to be found (figure 22 A and C).

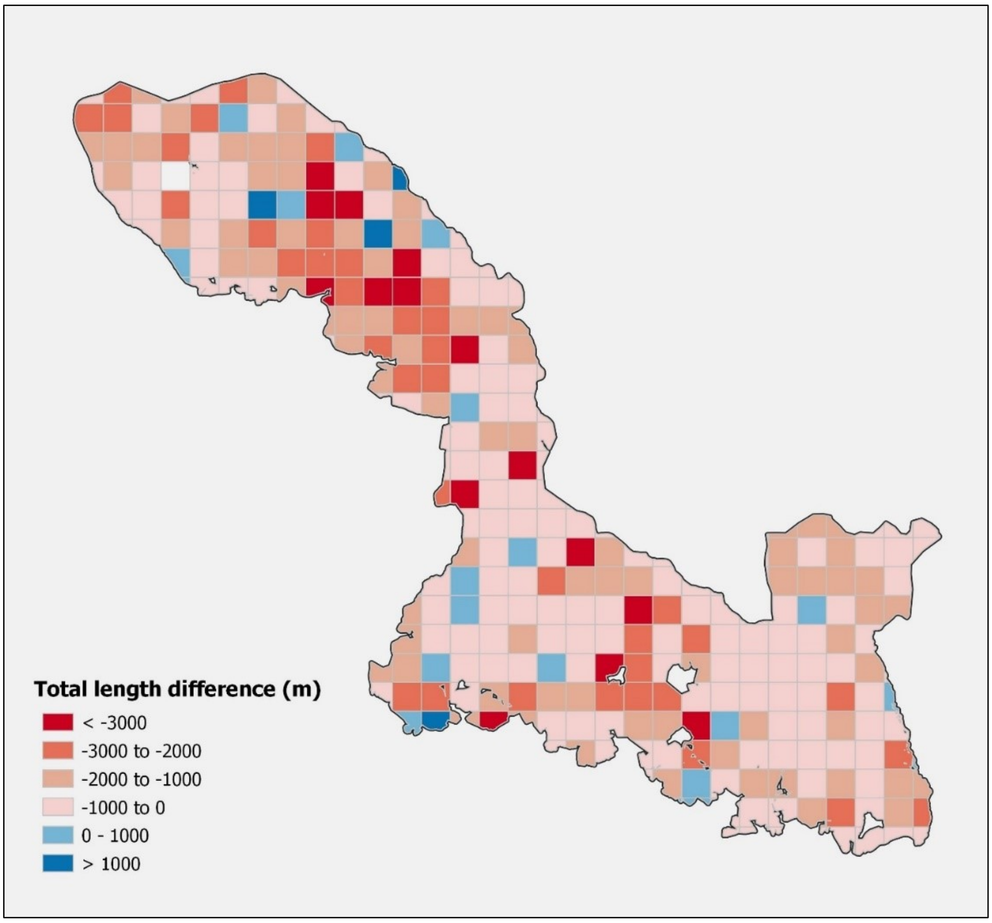


Figure 21: Length difference between OSM and the reference road network.

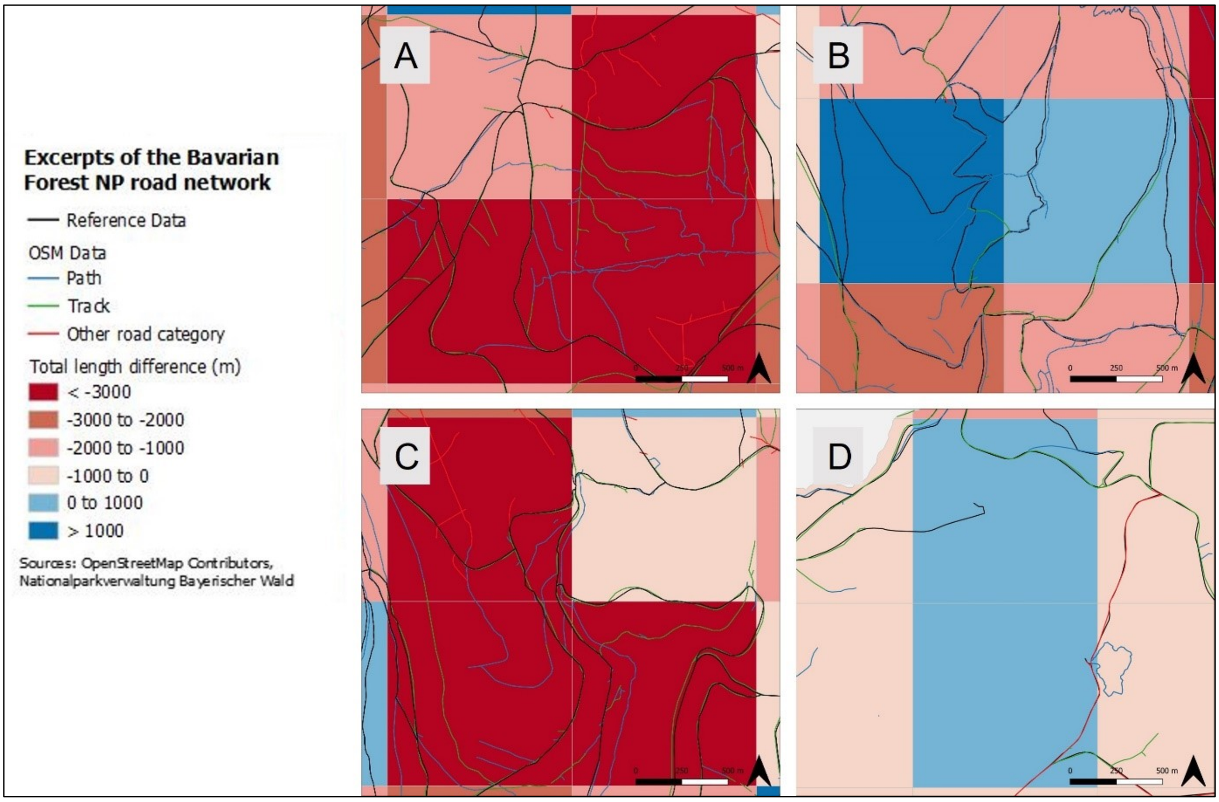


Figure 22: Examples of total length comparison results.

The overlay of both road networks in figure 23 as well as the results from the grid-based length comparison show that the OSM road network in the Bavarian Forest national park can be regarded as close to completion but not fully complete, yet, as there are still some road elements in the reference data that do not exist in OSM. This validates the results of the intrinsic assessment which indicated a rather high road network completeness of the OSM dataset in the Bavarian Forest NP (see table 8).

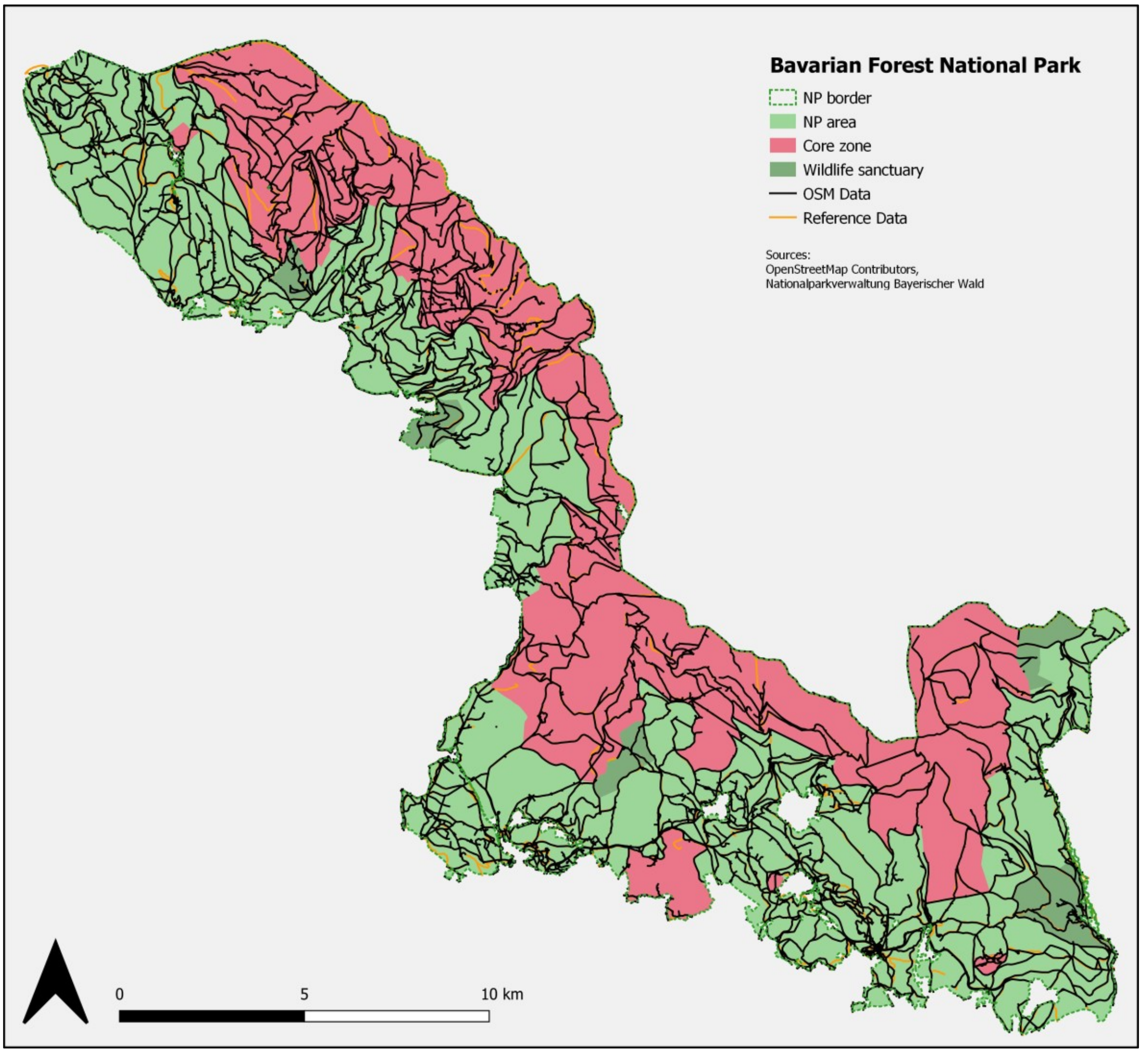


Figure 23: Overlay of reference dataset with OSM (highway=, abandoned:highway=*, disused:highway=*).*

At the same time, the quality of the reference dataset regarding general feature completeness is questionable. It can be assumed that the file containing all officially signposted ways ("Markierte_Wege_DTK25") is more or less complete because it's the basis for all signposting in nature and publication of maps by the NP authority. Thus, it can be regarded as a suitable ground truth dataset for all transport modes that are not allowed on any other ways in the national park. Consequently, the reason for poor feature completeness of the reference data is mainly to be found in an incomplete file "Sonstige_Wege_und_Steige" that comprises all ways which are not officially signposted and thus leaves room for interpretation which ways are relevant to be included.

Overall, OSM proves to be more complete than the reference data and thus better represents ground truth regarding feature completeness. This supports the findings by ANTONIOU AND SKOPELITI (2015) that OSM got more accurate and detailed than authoritative datasets in many parts of the world. In this part of the analysis, the reference dataset was not a valid choice for representing ground truth. Nevertheless, this always depends on the use case as well as the available reference data and cannot be extrapolated to other national parks or protected areas.

3.3.2 Attribute Completeness

In contrast to the high errors of commission in the overall road network, the OSM bicycle network is 2.97% (7.07 km) shorter than the reference network. As cycling is only allowed on officially signposted ways, the suitability of the reference data representing ground truth in this context can be regarded as high. The relatively low error of omission indicates a high attribute completeness of the OSM dataset regarding the bicycle network. Nevertheless, because 16.4% of all roads do not carry a *bicycle*-tag at all according to the intrinsic analysis, there still seems to be a lack of attribution for ways where cycling is not allowed. As shown in 3.2.3, this can be problematic in the context of route planning for cyclists when routing engines take *bicycle=yes* as a default for ways that do not carry a tag that indicates the opposite.

Table 12: Results of extrinsic attribute completeness evaluation.

OSM	**Reference data**	**Error of commission / omission**	
		Length	**Error Rate**
Total length of ways where cycling is allowed[1]			
230.68 km	237.75 km	- 7.07 km	- 2.97 %
Total length of ways prepared for cross country skiing[2]			
38.51 km	96.16 km	- 57.65 km	- 59.95 %
Total length of ways designated for winter hiking[3]			
0.00 km	91.13 km	- 91.13 km	- 100.00 %

[1]OSM: All ways with a tag *bicycle=yes* or *bicycle=designated*. Reference data: All ways from the file "Markierte_Wege_DTK25" where the attribute *Wege* carries one of the values *RADWEG*, *WRWEG*, *RRWEG* or *WRRWEG*.

[2]OSM: All ways carrying a tag *piste:type=nordic*. Reference data: All ways where the attribute *Wintersport* carries one of the values *LOIPE*, *LOIWWW* or *LOING*.

[3]OSM: All ways carrying a tag *piste:type=hike*. Reference data: All ways where the attribute *Wintersport* carries one of the values *WWW* or *LOIWWW*.

For winter sports, where the reference data is a valid ground truth as well, the error of omission is very high and indicates a poor attribute completeness of the OSM dataset regarding cross country skiing tracks and even more winter hiking trails. Consequently, at current state OSM is not a reliable source for planning a cross-country skiing tour in the Bavarian Forest.

Table 13: Attribute completeness for seasonal access restrictions (incl. 30m buffer).

Core zone	
Total length of ways (*highway=**):	370,089 m
Length of ways carrying a general access restriction (*access=no, access=private, foot=no, foot=permit*):	84,192 m
Length of ways that should carry a seasonal access restriction:	285,897 m
Length of ways carrying a seasonal access restriction[1]:	132,732 m
Tag ratio:	**46.43 %**
Wildlife sanctuary "Ahornschachten"	
Total length of ways (*highway=**):	9,868 m
Length of ways carrying a general access restriction (*access=no, foot=no*):	1,996 m
Length of ways that should carry a seasonal access restriction:	7,872 m
Length of ways carrying a seasonal access restriction[1]:	923 m
Tag ratio:	**11.73 %**
Wildlife sanctuary "Auerwild"	
Total length of ways (*highway=**):	3,118 m
Length of ways carrying a general access restriction (*access=no, foot=no*):	19 m
Length of ways that should carry a seasonal access restriction:	3,099 m
Length of ways carrying a seasonal access restriction[1]:	2,976 m
Tag ratio:	**96.03 %**
Wildlife sanctuary "Auwald"	
Total length of ways (*highway=**):	8,746 m
Length of ways carrying a general access restriction (*access=no, foot=no*):	0 m
Length of ways that should carry a seasonal access restriction:	8,746 m
Length of ways carrying a seasonal access restriction[1]:	6,622 m
Tag ratio:	**75.71 %**
Wildlife sanctuary "Neuhüttenwiese"	
Total length of ways (*highway=**):	5,342 m
Length of ways carrying a general access restriction (*access=no, foot=no*):	436 m
Length of ways that should carry a seasonal access restriction:	4,906 m
Length of ways carrying a seasonal access restriction[1]:	3,148 m
Tag ratio:	**64.17 %**
Wildlife sanctuary "Riedlhäng"	
Total length of ways (*highway=**):	20,678 m
Length of ways carrying a general access restriction (*access=no, foot=no*):	0 m
Length of ways that should carry a seasonal access restriction:	20,678 m
Length of ways carrying a seasonal access restriction[1]:	7,326 m
Tag ratio:	**35.43 %**

[1] *access:conditional*=* and/or *foot:conditional*=*

Looking at access restrictions within the core zone and all five wildlife sanctuaries, attribute completeness differs significantly between those areas as table 13 demonstrates. Since seasonal restrictions are valid for the whole area of a wildlife sanctuary and almost all the core zones (see 2.3.2), only a tag ratio close to 100% could indicate a high attribute completeness, which is only the case in one area.

Because the road networks of all wildlife sanctuaries are fairly small and consist of only a few elements plus the analysis represents the status quo of Nov 11th, 2021, changes in the dataset might have already improved the situation in those areas. Results for the core areas on the other hand are likely to still be valid at the time of this writing due to the total length of the road network and the unlikely case of all missing attributes being added in the meantime. Thus, the analysis discloses a lot of room for quality improvement regarding the representation of access restrictions in the OSM dataset.

Assuming that access to an abandoned or disused road is prohibited for any transport mode, the OSM dataset also shows a rather poor attribute completeness regarding access restrictions on abandoned or disused roads, as only 38.5 % carry this information (see table 14).

Table 14: Attribute completeness of abandoned and disused roads.

Tag	**Total length of abandoned/disused[1] roads carrying tag (km)**	**Total length of abandoned/disused[1] roads (km)**	**Tag ratio**
*access=** and/or *foot=**	36.36	94.33	38.55 %
access=no or *access=private*	34.89	94.33	36.99 %
foot=no	3.83	94.33	4.06 %
bicycle=no	14.97	94.33	15.87 %

[1]All roads carrying an abandoned:highway=* or disused:highway=* tag.

All in all, because the evaluation of attribute completeness produces distinct results for every investigated attribute, results are hardly comparable to earlier studies which assessed different attributes like *name* or *maxspeed*. Nevertheless, results of this study point in the same direction already shown by GRASER ET AL. (2014), who found incomplete attribute information related to navigation in OSM road networks, too.

3.3.3 Classification Correctness

Completeness and classification correctness are strongly related. The misclassification of an object to the wrong class will appear in the evaluation of completeness for both classes (one commission and one omission). Thus, some error of omission or commission resulting from the assessment of completeness could be caused by misclassification.

Cycling

The overlay of road networks in figure 24 demonstrates that almost all OSM roads which are attributed with *bicycle=yes* are correctly classified as they all have a counterpart in the reference data.

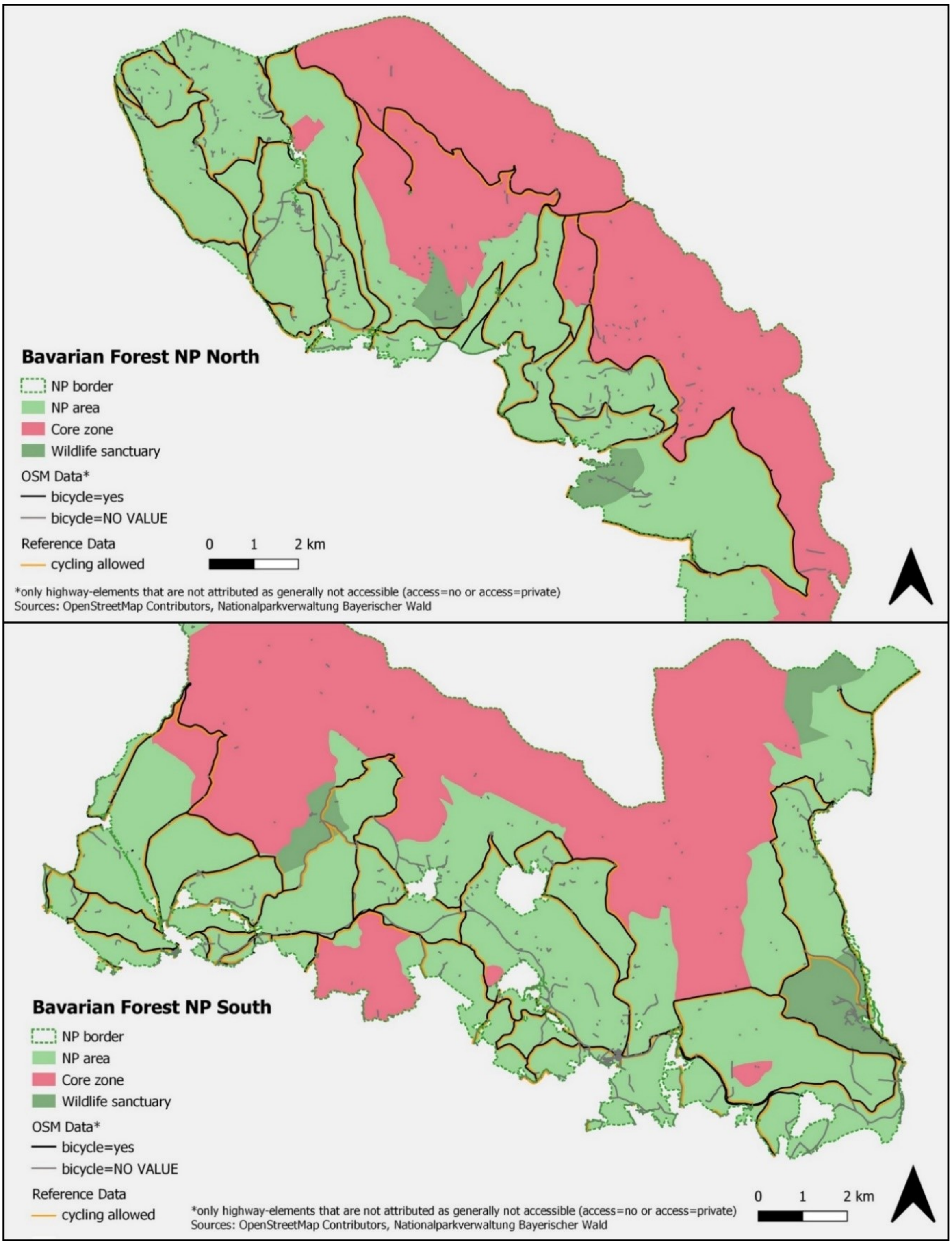
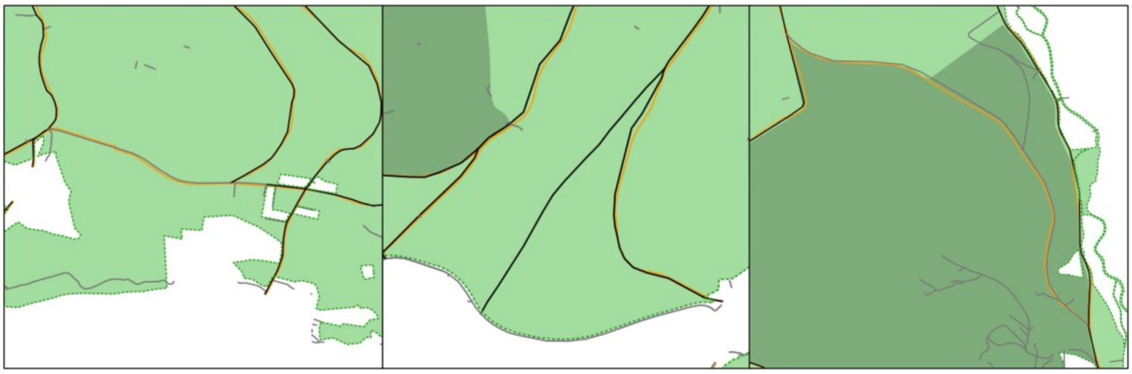


Figure 24: Overlay of ways attributed as "cycling allowed".

Thus, classification correctness of the OSM data regarding *bicycle=yes* is very high and - together with a high attribute completeness for the same tag (see table 12) - makes this information being very accurately represented in OSM. Table 15 summarizes some of the few examples of ways that are usable for cyclists according to the reference dataset but not attributed with a *bicycle*-tag in OSM. The 16.4% of OSM-roads not carrying a bicycle tag at all are mainly shorter paths and tracks all over the national park.

Table 15: Examples of incorrect or missing classification for cycling in OSM.



Cross-country skiing

Despite the overall longer ski track network in the reference dataset and positional offsets visible, classification correctness is at 100% in the northern part of the national park as all ski tracks present in OSM are classified correctly (see figure 25). In the southern part, this is true for most of the OSM elements, with a few exceptions shown in table 16. Although classification correctness of OSM for ski tracks is very high, the low completeness prevents OSM from being a reliable source of information at this point.

Table 16: Classification errors in the cross-country skiing network.

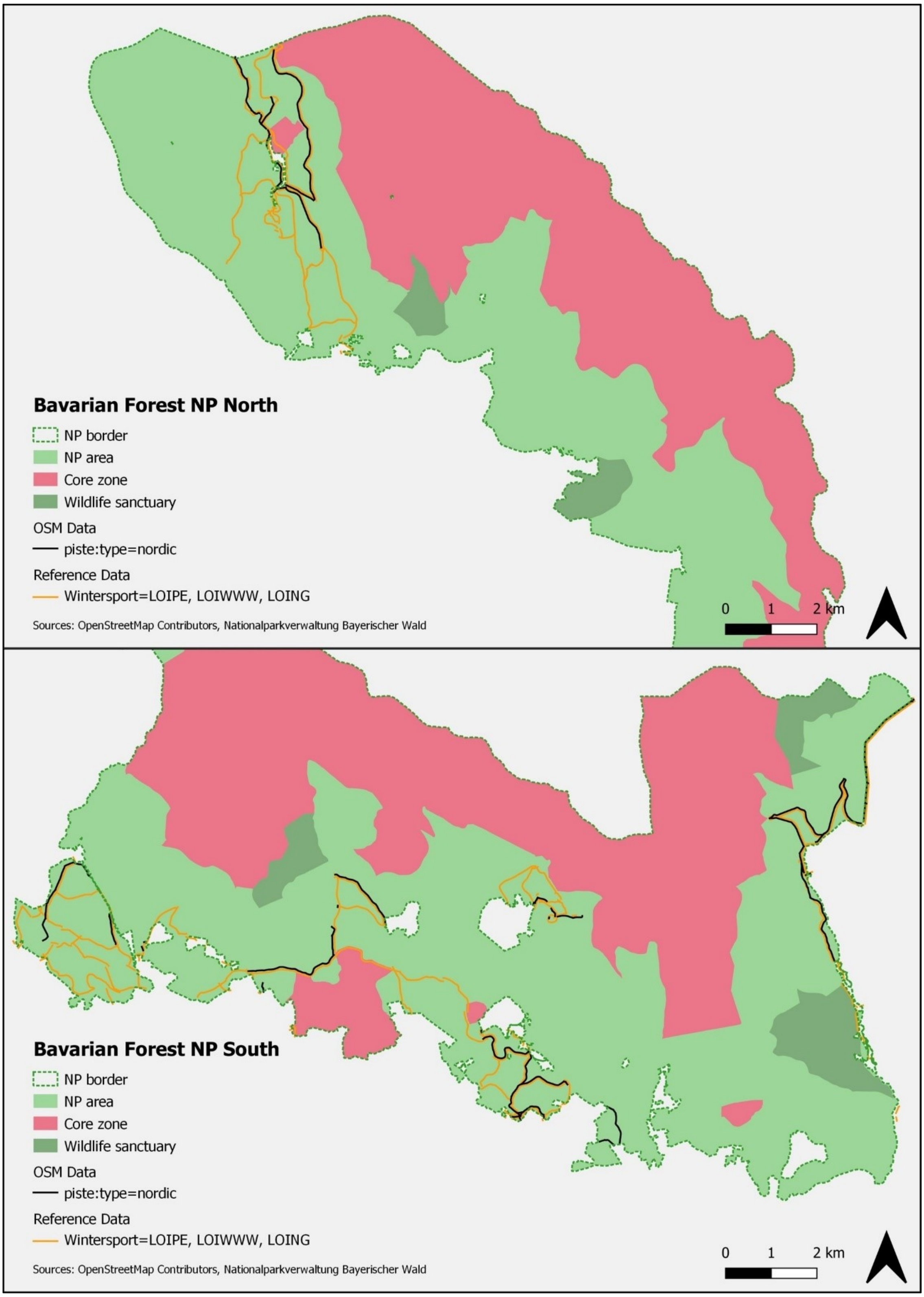


Figure 25: Overlay of ways attributed as cross-country skiing tracks.

Access restrictions for pedestrians

In all five wildlife sanctuaries, the calculated misclassification rate is zero, meaning all existing seasonal restrictions in OSM are 100% correct. At the same time, classification maps (see appendix

A2) demonstrate a low attribute completeness for access restrictions in all areas but the wildlife sanctuary "Auerwild" and thus support the results shown in 3.3.2. Consequently, only for this area the overall attribution quality regarding access restrictions can be stated as "very good" while in all other areas there is still room for improvement.

3.3.4 Trustworthiness

The number of versions of the almost 6.000 road segments in the Bavarian Forest OSM dataset ranges from 1 to 42 as of Nov 11th, 2021. On average, road segments have been edited 4.85 times. It needs to be remarked that in OSM, for *way*-elements only edits that change the referenced nodes (e. g. deleted or added node to the way) and tag changes result in a new version number.

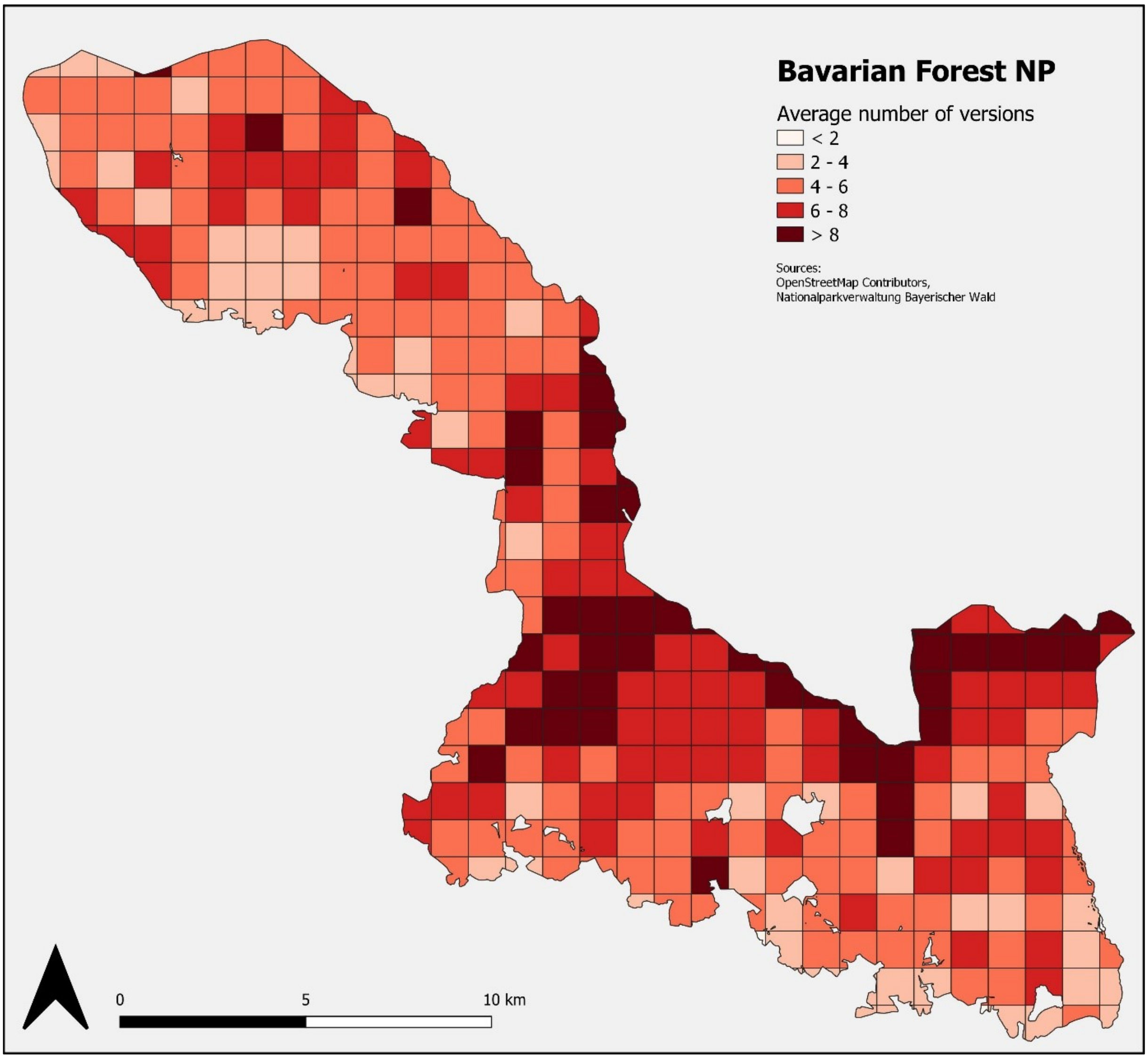


Figure 26: Average no. of versions per road segment over 1km grid cells.

Following the argumentation of KEßLER AND GROOT (2013), OSM data shows the highest trustworthiness in the centre and along the northern border of the southern part of the national park as these areas show the highest number of versions on average (see figure 26). These results must be handled with care though, as they come without any further assessment of other parameters that influence trustworthiness like user reputation or rollbacks. Nevertheless, they give valuable hints towards where better trustworthiness and thus better quality is to be found in the study area. Looking at the different contribution types that were analysed intrinsically (see table 10) it stands out, that most edits are related to tag changes and thus to the completeness and correctness of attributes. Hence, it can be assumed that a high number of versions of a road segment is predominantly related to tag changes and the improvement of its attributes.

In total, the road elements present in the OSM database as of Nov 11th, 2021, have been edited by 158 different users since 2009. In accordance with the results of NEIS AND ZIPF (2012), the minority of users (10%) accounts for more than 90% of all latest edits. As shown in table 17, almost 47% of all road elements (64% / 744.37 km when looking at their length) were edited the latest by the same user (user A).

Table 17: Users who contributed to the OSM road network in the Bavarian Forest NP.

OSM user (Synonym)	**No. of elements last edited**	**Edit Share**
A	2794	46.93%
B	1443	24.24%
C	493	8.28%
D	146	2.45%
E	130	2.18%
F	104	1.75%
Others	844	14.18%

Figure 27 illustrates, that user A has been by far the most active contributor to the road network in the last two years and thus is responsible for the large majority of all performed edits since it started its editing activity in March 2020. Obviously, this user has had a high impact the on the quality of the current dataset and thus influences the fitness-for-use of OSM data significantly. Due to EU data protection regulations, details on user A which are present in the dataset cannot be described in this thesis. However, knowing that the NP administration of the Bavarian Forest started to actively contribute to OSM a while ago (SELTMANN AND ZINK 2021), it is likely that user A is representing this recent activity.

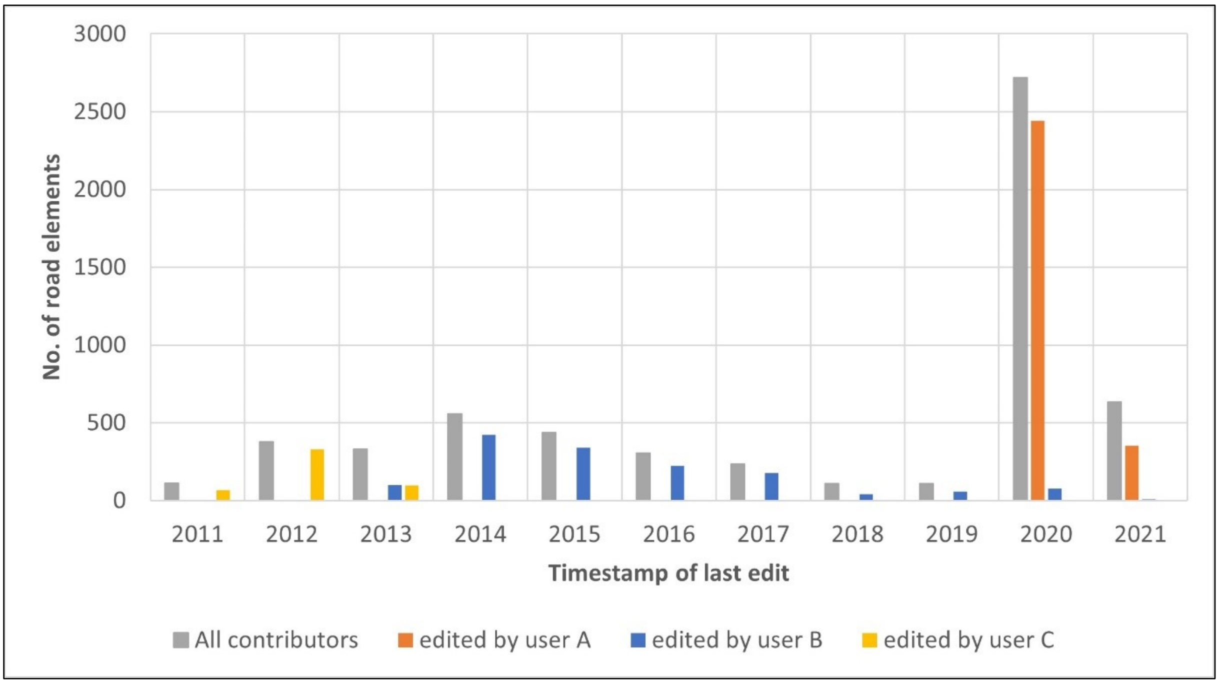


Figure 27: Contributions to the OSM road network of the Bavarian Forest NP (last 10 years).

This study showed earlier that the size of most OSM communities which are actively contributing to the OSM datasets in other national parks is similar to or even lower than in the Bavarian Forest (see table 7). Thus, it can be assumed that the influence of a single, very active contributor has a substantial impact on the quality of the OSM dataset in most national parks. This shows the opportunities for authorities to improve the OSM quality of their territory substantially by actively dealing with OSM and contributing to the database. This way, administrators of protected areas can actively reduce the risk for conflicts between outdoor enthusiasts and the purposes of nature conservation. Assuming an administration's intention would be to improve the data's quality, the fact that an active OSM user account which is responsible for a lot of contributions is run by an official authority would also create additional trust in the dataset.

4. Conclusions

In contrast to urban street networks, OSM road networks in protected areas are predominantly existing of the road type *path* and *track*. Additionally, they are subject to a very specific legislation on nature conservation that differs from one area to another. Usually, they only have a small group of users who are actively contributing to OSM.

In German national parks, OSM road networks are complete to a high extent, although the OSM database still sees some growth in several regions. They proved to almost fully represent roads and ways that are existing in the real world, even small paths and short trails. Regarding attribute completeness, the road networks of national parks leave a lot of room for improvement. This specifically applies to general and specific access restrictions for different recreational activity types like cycling and hiking. Thus, the reliability of the OSM road network for outdoor routing in protected areas is at least questionable and depends a lot on the area and the approach the application provider has on how to deal with missing attribute information.

Germany has a very active OSM community (NEIS 2022). Taking the worldwide growth of OSM road networks into account, it is likely that the results of this work are somewhat transferrable to other protected areas in Germany as well as to national parks in other countries with a similarly active community. Nevertheless, future studies for other countries as well as for other types of protected areas (e. g. nature or biosphere reserves) are necessary to give a more accurate estimation of OSM quality in protected areas on a global scale.

On the example of the Bavarian Forest, it became clear that due to small local OSM communities, the administration of a protected area can have a substantial impact on quality of the OSM dataset in its region if it is actively contributing to OSM. While this work gives some indications for the Bavarian Forest, the actual influence of authorities that are already doing so offers an interesting field for future studies. Additionally, it could be very valuable for interest groups like national park administrations to analyse the contributors to their local OSM datasets in more detail. This has not been done in this thesis due to EU data protections regulations but could easily be performed by authorities for internal purposes with the presented method. A deeper analysis of contributor behaviour (global experience of users, compliance of tags with the OSM Wiki) could also enhance the understanding of reliability of OSM.

Generally, the evaluation methods used in this study proved to be easy-applicable and provide a valuable insight into the quality of OSM road networks from a general interest or an authority's point of view. They can help administrations of protected areas as well as OSM communities to identify data that needs to be corrected or enhanced. The applied methods can be transferred to the analysis of other data types and themes like POIs. Nevertheless, the intrinsic measurements only provide an approximation of data quality as they do not produce any quantitative results. In the context of outdoor routing, high quality reference data and knowledge on the legislation valid in the area under investigation is an essential external part to determine the fitness-for-use of OSM road networks. The external data available for the case study did only partially provide the quality that is needed to accurately do so, as it was less complete than the OSM dataset.

This analysis focussed on the completeness and correctness of attributes, specifically of those relevant for outdoor routing in a protected area. It did not consider other quality elements of which positional accuracy and logical consistency are the most important ones to investigate in future research in the context of outdoor navigation in rural areas.

Moreover, OSM elements of the type *Relation* have not been investigated in this study. Potential studies might find a valuable approach in analysing *routes* in OSM, which are customary or regular lines of passage or travel (cycling routes, hiking routes etc.) that are usually represented by relations.

Furthermore, the approach of this work did not provide the in-depth use-case-specific results that a potential user of OSM like an outdoor routing platform can use to evaluate the usability of data for its specific routing engine. For future studies in this field, it could be beneficial to set up a data product specification that represents the specific requirements to data quality from a platform's point of view before performing the quality analysis.

Finally, it is important to keep in mind that the OSM database is under constant change. The quality analysed in this and any prior or upcoming study does only represent the status quo of the data under investigation. In the meantime, the OSM community might have performed significant edits to the data that influence its quality.

REFERENCES

ABDOLMAJIDI, E.; MANSOURIAN, A.; WILL, J.; HARRIE, L. (2015): Matching authority and VGI road networks using an extended node-based matching algorithm. In *Geo-spatial Information Science* 18 (2-3), pp. 65–80. DOI: 10.1080/10095020.2015.1071065.

ALGHANIM, A.; JILANI, M.; BERTOLOTTO, M.; MCARDLE, G. (2021): Leveraging Road Characteristics and Contributor Behaviour for Assessing Road Type Quality in OSM. In *ISPRS Int. J. Geo-Inf.* 10 (7), p. 436. DOI: 10.3390/ijgi10070436.

ALI, A. L.; SCHMID, F.; AL-SALMAN, R.; KAUPPINEN, T. (2014): Ambiguity and Plausibility: Managing Classification Quality in Volunteered Geographic Information. In Huang, Y., Schneider, M., Gertz, M., Krumm, J., Sankaranarayanan, J. (Eds.): SIGSPATIAL '14: Proceedings of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. Dallas Texas, Nov 04-07, 2014. Association for Computing Machinery. New York, NY, USA, pp. 143–152.

ALLTRAILS (2022a): About AllTrails. Available online at https://www.alltrails.com/about, checked on 4/29/2022.

ALLTRAILS (2022b): OSM Derivative Database - Derivation Methodology. Available online at https://support.alltrails.com/hc/en-us/articles/360019246411-OSM-Derivative-Database-Derivation-Methodology, checked on 4/29/2022.

ALMENDROS-JIMENEZ, J. M.; BECERRA-TERON, A. (2018): Analyzing the Tagging Quality of the Spanish OpenStreetMap. In *ISPRS Int. J. Geo-Inf.* 7 (8). DOI: 10.3390/ijgi7080323.

ANTONIOU, V. (2011): User generated spatial content: an analysis of the phenomenon and its challenges for mapping agenciesDissertation. UNIVERSITY COLLEGE LONDON (UCL), London. Available online at https://discovery.ucl.ac.uk/id/eprint/1318053.

ANTONIOU, V.; SKOPELITI, A. (2015): MEASURES AND INDICATORS OF VGI QUALITY: AN OVERVIEW. In *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* II-3/W5, pp. 345–351. DOI: 10.5194/isprsannals-II-3-W5-345-2015.

AUER, M.; ECKLE, M.; FENDRICH, S. (2018): Ohsome – eine Plattform zur Analyse raumzeitlicher Entwicklungen von OpenStreetMap-Daten für intrinsische Qualitätsbewertungen. In Strobl, J., Zagel, B., Griesebner, G., Blaschke, T. (Eds.): Journal für angewandte Geoinformatik, 4-2018. Berlin: Herbert-Wichmann-Verlag (4-2018).

BALLATORE, A.; MOONEY, P. (2015): Conceptualising the geographic world: the dimensions of negotiation in crowdsourced cartography. In *International Journal of Geographical Information Science* 29 (12), pp. 2310–2327. DOI: 10.1080/13658816.2015.1076825.

BALLATORE, A.; ZIPF, A. (2015): A Conceptual Quality Framework for Volunteered Geographic Information. In Fabrikant, S. I., Raubal, M., Bertolotto, M., Davies, C., Freundschuh, S., Bell, S. (Eds.): Spatial Information Theory, vol. 9368. 12th International Conference on Spatial Information Theory (COSIT). Santa Fe, NM, USA, Oct 12-16, 2015. Cham, Switzerland: Springer International Publishing, pp. 89–107.

BALMFORD, A.; GREEN, J. M. H.; ANDERSON, M.; BERESFORD, J.; HUANG, C.; NAIDOO, R. ET AL. (2015): Walk on the Wild Side: Estimating the Global Magnitude of Visits to Protected Areas. In *PLoS Biol* 13 (2). DOI: 10.1371/journal.pbio.1002074.

BARRINGTON-LEIGH, C.; MILLARD-BALL, A. (2017): The world's user-generated road map is more than 80% complete. In *PLoS ONE* 12 (8). DOI: 10.1371/journal.pone.0180698.

BARRON, C.; NEIS, P.; ZIPF, A. (2014): A Comprehensive Framework for Intrinsic OpenStreetMap Quality Analysis. In *Transactions in GIS* 18 (6), pp. 877–895. DOI: 10.1111/tgis.12073.

BERGMAN, C.; OKSANEN, J. (2016): Conflation of OpenStreetMap and Mobile Sports Tracking Data for Automatic Bicycle Routing. In *Transactions in GIS* 20 (6), pp. 848–868. DOI: 10.1111/tgis.12192.

BROVELLI, M. A.; MINGHINI, M.; MOLINARI, M.; MOONEY, P. (2017): Towards an Automated Comparison of OpenStreetMap with Authoritative Road Datasets. In *Transactions in GIS* 21 (2), pp. 191–206. DOI: 10.1111/tgis.12182.

BROVELLI, M. A.; MINGHINI, M.; MOLINARI, M. E. (2016): AN AUTOMATED GRASS-BASED PROCEDURE TO ASSESS THE GEOMETRICAL ACCURACY OF THE OPENSTREETMAP PARIS ROAD NETWORK. In *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XLI-B7, pp. 919–925. DOI: 10.5194/isprs-archives-XLI-B7-919-2016.

CAMBOIM, S.; BRAVO, J.; SLUTER, C. (2015): An Investigation into the Completeness of, and the Updates to, OpenStreetMap Data in a Heterogeneous Area in Brazil. In *ISPRS Int. J. Geo-Inf.* 4 (3), pp. 1366–1388. DOI: 10.3390/ijgi4031366.

CANAVOSIO-ZUZELSKI, R.; AGOURIS, P.; DOUCETTE, P. (2013): A Photogrammetric Approach for Assessing Positional Accuracy of OpenStreetMap© Roads. In *ISPRS Int. J. Geo-Inf.* 2 (2), pp. 276–301. DOI: 10.3390/ijgi2020276.

CHEHREGHAN, A.; ALI ABBASPOUR, R. (2018): An evaluation of data completeness of VGI through geometric similarity assessment. In *International Journal of Image and Data Fusion* 9 (4), pp. 319–337. DOI: 10.1080/19479832.2018.1504825.

DAVIDOVIC, N.; MOONEY, P.; STOIMENOV, L.; MINGHINI, M. (2016): Tagging in Volunteered Geographic Information: An Analysis of Tagging Practices for Cities and Urban Regions in OpenStreetMap. In *ISPRS Int. J. Geo-Inf.* 5 (12). DOI: 10.3390/ijgi5120232.

DORN, H.; TOERNROS, T.; ZIPF, A. (2015): Quality Evaluation of VGI Using Authoritative Data-A Comparison with Land Use Data in Southern Germany. In *ISPRS Int. J. Geo-Inf.* 4 (3), pp. 1657–1671. DOI: 10.3390/ijgi4031657.

DUDLEY, N. (2013): Guidelines for applying protected area management categories including IUCN WCPA best practice guidance on recognising protected areas and assigning management categories and governance types. Gland: IUCN (Best practice protected area guidelines series, 21). Available online at https://portals.iucn.org/library/efiles/documents/PAG-021.pdf, checked on 4/15/2022.

FAN, H.; YANG, B.; ZIPF, A.; ROUSELL, A. (2016): A polygon-based approach for matching OpenStreetMap road networks with regional transit authority data. In *International Journal of Geographical Information Science* 30 (4), pp. 748–764. DOI: 10.1080/13658816.2015.1100732.

FLANAGIN, A. J.; METZGER, M. J. (2008): The credibility of volunteered geographic information. In *GeoJournal* 72 (3-4), pp. 137–148. DOI: 10.1007/s10708-008-9188-y.

FOGLIARONI, P.; D'ANTONIO, F.; CLEMENTINI, E. (2018): Data trustworthiness and user reputation as indicators of VGI quality. In *Geo-spatial Information Science* 21 (3), pp. 213–233. DOI: 10.1080/10095020.2018.1496556.

FONTE, C. C.; ANTONIOU, V.; BASTIN, L.; ESTIMA, J.; ARSANJANI, J. J.; LASO BAYAS, J.-C. ET AL. (2017): Assessing VGI Data Quality. In Foody, G., See, L., Fritz, S., Mooney, P., Olteanu-Raimond, A.-M., Fonte, C. C., Antoniou, V. (Eds.): Mapping and the Citizen Sensor. London: Ubiquity Press, pp. 137–163.

FORGHANI, M.; DELAVAR, M. (2014): A Quality Study of the OpenStreetMap Dataset for Tehran. In *ISPRS Int. J. Geo-Inf.* 3 (2), pp. 750–763. DOI: 10.3390/ijgi3020750.

GEOFABRIK GMBH (2022): Data Extracts - Technical Details. Available online at https://download.geofabrik.de/technical.html, checked on 5/3/2022.

GIRRES, J.-F.; TOUYA, G. (2010): Quality Assessment of the French OpenStreetMap Dataset. In *Transactions in GIS* 14 (4), pp. 435–459. DOI: 10.1111/j.1467-9671.2010.01203.x.

GOODCHILD, M. F. (2007): Citizens as sensors: the world of volunteered geography. In *GeoJournal* 69 (4), pp. 211–221. DOI: 10.1007/s10708-007-9111-y.

GOODCHILD, M. F.; HUNTER, G. J. (1997): A simple positional accuracy measure for linear features. In *International Journal of Geographical Information Science* 11 (3), pp. 299–306. DOI: 10.1080/136588197242419.

GOODCHILD, M. F.; LI, L. (2012): Assuring the quality of volunteered geographic information. In *Spatial Statistics* 1, pp. 110–120. DOI: 10.1016/j.spasta.2012.03.002.

GRASER, A.; STRAUB, M.; DRAGASCHNIG, M. (2014): Towards an Open Source Analysis Toolbox for Street Network Comparison: Indicators, Tools and Results of a Comparison of OSM and the Official Austrian Reference Graph. In *Transactions in GIS* 18 (4), pp. 510–526. DOI: 10.1111/tgis.12061.

GRÖCHENIG, S.; BRUNAUER, R.; REHRL, K. (2014): 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. Cham, Switzerland: Springer International Publishing (Lecture Notes in Geoinformation and Cartography), pp. 3–18.

HAKLAY, M. (2010): How Good is Volunteered Geographical Information? A Comparative Study of OpenStreetMap and Ordnance Survey Datasets. In *Environ Plann B Plann Des* 37 (4), pp. 682–703. DOI: 10.1068/b35097.

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. In *The Cartographic Journal* 47 (4), pp. 315–322. DOI: 10.1179/000870410X12911304958827.

HALL, A.; THEBAULT-SPIEKER, J.; SEN, S.; HECHT, B.; TERVEEN, L. (2018): Exploring the Relationship Between "Informal Standards" and Contributor Practice in OpenStreetMap. In : Proceedings of the 14th International Symposium on Open Collaboration (OPENSYM'18). Paris, France, Aug 22-24, 2018. Association for Computing Machinery.

HALLERMANN, M. (2022): Wie komoot zum Naturschutz beiträgt und in Besucherlenkung investiert. Available online at https://medium.com/komoot/wie-komoot-zum-naturschutz-beitr%C3%A4gt-und-in-besucherlenkung-investiert-2572230eb4, updated on 3/18/2022, checked on 3/22/2022.

HASHEMI, P.; ALI ABBASPOUR, R. (2015): Assessment of Logical Consistency in OpenStreetMap Based on the Spatial Similarity Concept. In Jokar Arsanjani, J., Zipf, A., Mooney, P. J., Helbich, M. (Eds.): OpenStreetMap in GIScience. Experiences, research, and applications. Cham, Heidelberg: Springer International Publishing (Lecture Notes in Geoinformation and Cartography).

HEIGIT GGMBH (2021): ohsome API Documentation. Available online at https://docs.ohsome.org/ohsome-api/v1/index.html, checked on 3/15/2022.

HENNIG, S. (2017): OpenStreetMap used in protected area management. The example of the recreational infrastructure in Berchtesgaden National Park. In *eco.mont Journal on Protected Mountain Areas Research* 9 (2), pp. 30–41. DOI: 10.1553/eco.mont-9-2s30.

HOCHMAIR, H. H.; ZIELSTRA, D.; NEIS, P. (2015): Assessing the Completeness of Bicycle Trail and Lane Features in OpenStreetMap for the United States. In *Transactions in GIS* 19 (1), pp. 63–81. DOI: 10.1111/tgis.12081.

INTERNATIONAL ORGANISATION FOR STANDARDIZATION (2013): ISO19157:2013: Geographic information - Data quality.

JACOBS, K. T.; MITCHELL, S. W. (2020): OpenStreetMap quality assessment using unsupervised machine learning methods. In *Transactions in GIS* 24 (5), pp. 1280–1298. DOI: 10.1111/tgis.12680.

JILANI, M.; CORCORAN, P.; BERTOLOTTO, M. (2014): Automated highway tag assessment of OpenStreetMap road networks. In Huang, Y., Schneider, M., Gertz, M., Krumm, J., Sankaranarayanan, J. (Eds.): SIGSPATIAL '14: Proceedings of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. Dallas Texas, Nov 04-07, 2014. Association for Computing Machinery. New York, NY, USA.

JOKAR ARSANJANI, J.; BARRON, C.; BAKILLAH, M.; HELBICH, M. (2013): Assessing the Quality of OpenStreetMap Contributors together with their Contributions. In Vandenbroucke, D., Bucher, B., Crompvoets, J. (Eds.): Proceedings 2013. The 16th AGILE International Conference on Geographic Information Science. Leuven, Belgium.

KEßLER, C.; GROOT, R. T. A. DE (2013): Trust as a Proxy Measure for the Quality of Volunteered Geographic Information in the Case of OpenStreetMap. In Vandenbroucke, D., Bucher, B., Crompvoets, J. (Eds.): Geographic Information Science at the Heart of Europe: Springer International Publishing (Lecture Notes in Geoinformation and Cartography), pp. 21–37.

KOMOOT GMBH (2022a): Change The Way People Explore. Available online at https://www.komoot.de/jobs, checked on 4/29/2022.

KOMOOT GMBH (2022b): Improving the komoot map using OpenStreetMap. Available online at https://support.komoot.com/hc/en-us/articles/360022830972-Improving-the-komoot-map-using-OpenStreetMap, checked on 4/29/2022.

KOUKOLETSOS, T.; HAKLAY, M.; ELLUL, C. (2011): An automated method to assess Data Completeness and Positional Accuracy of OpenStreetMap. In : Proceedings of GeoComputation. 11th International Conference on GeoComputation. London, Jul 20-22, 2011.

KOUKOLETSOS, T.; HAKLAY, M.; ELLUL, C. (2012): Assessing Data Completeness of VGI through an Automated Matching Procedure for Linear Data. In *Transactions in GIS* 16 (4), pp. 477–498. DOI: 10.1111/j.1467-9671.2012.01304.x.

LANDRATSAMT FREYUNG-GRAFENAU (1982): Verordnung des Landratsamtes Freyung-Grafenau über das Wildschutzgebiet für Auerwild im Nationalpark Bayerischer Wald. Available online at https://www.nationalpark-bayerischer-wald.bayern.de/service/wegeservice/doc/vo_wschg_auerwild_ba.pdf, checked on 3/7/2022.

LANDRATSAMT FREYUNG-GRAFENAU (1990): Verordnung des Landratsamtes Freyung-Grafenau über die Wildschutzgebiete "Riedlhäng" und "Neuhüttenwiese" im Nationalpark Bayerischer Wald. Available online at https://www.nationalpark-bayerischer-wald.bayern.de/service/wegeservice/doc/vo_wschg_riedlhaeng_teufelsberg_neuhuettenwiese_ba.pdf, checked on 3/7/2022.

LANDRATSAMT REGEN (1984): Verordnung des Landratsamtes Regen über die Wildschutzgebiete "Auwald" und "Ahornschachten" im Bereich des Bayer. Forstamtes Zwiesel. Available online at https://www.nationalpark-bayerischer-wald.bayern.de/service/wegeservice/doc/vo_wschg_auwald_ahornschachten_ba.pdf, checked on 3/7/2022.

LUDWIG, I.; VOSS, A.; KRAUSE-TRAUDES, M. (2011): A Comparison of the Street Networks of Navteq and OSM in Germany. In Geertman, S. C., Reinhardt, W. P., Toppen, F. J. (Eds.):

Advancing Geoinformation Science for a Changing World: Springer International Publishing (Lecture Notes in Geoinformation and Cartography), pp. 65–84.

MA, D.; SANDBERG, M.; JIANG, B. (2015): Characterizing the Heterogeneity of the OpenStreetMap Data and Community. In *ISPRS Int. J. Geo-Inf.* 4 (2), pp. 535–550. DOI: 10.3390/ijgi4020535.

MINGHINI, M.; FRASSINELLI, F. (2019): OpenStreetMap history for intrinsic quality assessment: Is OSM up-to-date? In *Open geospatial data, softw. stand.* 4 (1). DOI: 10.1186/s40965-019-0067-x.

MOBASHERI, A.; SUN, Y.; LOOS, L.; ALI, A. (2017): Are Crowdsourced Datasets Suitable for Specialized Routing Services? Case Study of OpenStreetMap for Routing of People with Limited Mobility. In *Sustainability* 9 (6), p. 997. DOI: 10.3390/su9060997.

MOCNIK, F.-B.; ZIPF, A.; FAN, H. (2017): The Inevitability of Calibration in VGI Quality Assessment. In : Proceedings of VGI-Analytics. 4th Workshop on Volunteered Geographic Information: Integration, Analysis and Applications (VGI-Analytics). Wageningen, The Netherlands, May 9th, 2017.

MOHAMMADI, N.; MALEK, M. (2015a): Artificial intelligence-based solution to estimate the spatial accuracy of volunteered geographic data. In *Journal of Spatial Science* 60 (1), pp. 119–135. DOI: 10.1080/14498596.2014.927337.

MOHAMMADI, N.; MALEK, M. (2015b): VGI and Reference Data Correspondence Based on Location-Orientation Rotary Descriptor and Segment Matching. In *Transactions in GIS* 19 (4), pp. 619–639. DOI: 10.1111/tgis.12116.

MONDZECH, J.; SESTER, M. (2011): Quality Analysis of OpenStreetMap Data Based on Application Needs. In *Cartographica: The International Journal for Geographic Information and Geovisualization* 46 (2), pp. 115–125. DOI: 10.3138/carto.46.2.115.

MONZ, C. A.; GUTZWILLER, K. J.; HAUSNER, V. H.; BRUNSON, M. W.; BUCKLEY, R.; PICKERING, C. M. (2021): Understanding and managing the interactions of impacts from nature-based recreation and climate change. In *Ambio* 50 (3), pp. 631–643. DOI: 10.1007/s13280-020-01403-y.

MOONEY, P.; CORCORAN, P. (2012): The Annotation Process in OpenStreetMap. In *Transactions in GIS* 16 (4), pp. 561–579. DOI: 10.1111/j.1467-9671.2012.01306.x.

MUTTAQIEN, B. I.; OSTERMANN, F. O.; LEMMENS, R. L. G. (2018): Modeling aggregated expertise of user contributions to assess the credibility of OpenStreetMap features. In *Transactions in GIS* 22 (3), pp. 823–841. DOI: 10.1111/tgis.12454.

NASIRI, A.; ALI ABBASPOUR, R.; CHEHREGHAN, A.; JOKAR ARSANJANI, J. (2018): Improving the Quality of Citizen Contributed Geodata through Their Historical Contributions: The Case of the Road Network in OpenStreetMap. In *ISPRS Int. J. Geo-Inf.* 7 (7), p. 253. DOI: 10.3390/ijgi7070253.

NATIONALPARKVERWALTUNG BAYERISCHER WALD (2022): Profile of Bavarian Forest National Park. Available online at https://www.nationalpark-bayerischer-wald.bayern.de/english/about_us/profile/index.htm, checked on 3/7/2022.

NEIS, P. (2021): OSMstats - Members. Statistics of the free wiki world map. Available online at https://osmstats.neis-one.org/?item=members, checked on 9/30/2021.

NEIS, P. (2022): OSMstats - Countries. Statistics of the free wiki world map. Available online at https://osmstats.neis-one.org/?item=countries, checked on 5/11/2022.

NEIS, P.; ZIELSTRA, D.; ZIPF, A. (2012): The Street Network Evolution of Crowdsourced Maps: OpenStreetMap in Germany 2007–2011. In *Future Internet* 4 (1), pp. 1–21. DOI: 10.3390/fi4010001.

NEIS, P.; ZIELSTRA, D.; ZIPF, A. (2013): Comparison of Volunteered Geographic Information Data Contributions and Community Development for Selected World Regions. In *Future Internet* 5 (2), pp. 282–300. DOI: 10.3390/fi5020282.

NEIS, P.; ZIPF, A. (2012): Analyzing the Contributor Activity of a Volunteered Geographic Information Project — The Case of OpenStreetMap. In *ISPRS Int. J. Geo-Inf.* 1 (2), pp. 146–165. DOI: 10.3390/ijgi1020146.

OSM1: OpenStreetMap stats. Available online at https://www.openstreetmap.org/stats/data_stats.html, checked on 5/2/2022.

OSM2: OpenStreetMap Wiki. Available online at https://wiki.openstreetmap.org/wiki/, checked on 4/4/2022.

OSM3: OpenStreetMap Wiki - Elements. Available online at https://wiki.openstreetmap.org/wiki/Elements, checked on 11/9/2021.

OSM4: OpenStreetMap Wiki - Good Practice. Available online at https://wiki.openstreetmap.org/wiki/Good_practice, checked on 2/16/2022.

OSM5: OpenStreetMap Wiki - Key:highway. Available online at https://wiki.openstreetmap.org/wiki/Key:highway, checked on 4/4/2022.

OSM6: OpenStreetMap Wiki - Lifecycle prefix. Available online at https://wiki.openstreetmap.org/wiki/Lifecycle_prefix, checked on 4/4/2022.

OSM7: OpenStreetMap Wiki - List of OSM-based services. Available online at https://wiki.openstreetmap.org/wiki/List_of_OSM-based_services, checked on 9/30/2021.

OSM8: OpenStreetMap Wiki - Map Features. Available online at https://wiki.openstreetmap.org/wiki/Map_features, checked on 10/12/2021.

OSM9: OpenStreetMap Wiki - OSM tags for routing/Access restrictions. Available online at https://wiki.openstreetmap.org/wiki/OSM_tags_for_routing/Access_restrictions#Germany, checked on 4/7/2022.

OUTDOORACTIVE GMBH (2022a): Discover the world with the best map. Available online at https://www.outdooractive.com/en/map/, checked on 4/29/2022.

OUTDOORACTIVE GMBH (2022b): Our worldwide outdoor platform. Available online at https://corporate.outdooractive.com/en/platform/?noredirect=en_US, checked on 4/29/2022.

OVER, M.; SCHILLING, A.; NEUBAUER, S.; ZIPF, A. (2010): Generating web-based 3D City Models from OpenStreetMap: The current situation in Germany. In *Computers, Environment and Urban Systems* 34 (6), pp. 496–507. DOI: 10.1016/j.compenvurbsys.2010.05.001.

RAIFER, M.; TROILO, R.; KOWATSCH, F.; AUER, M.; LOOS, L.; MARX, S. ET AL. (2019): OSHDB: a framework for spatio-temporal analysis of OpenStreetMap history data. In *Open geospatial data, softw. stand.* 4 (1). DOI: 10.1186/s40965-019-0061-3.

RAMM, F.; TOPF, J.; CHILTON, S. (2010): OpenStreetMap: Using and Enhancing the Free Map of the World: UIT Cambridge.

REGIERUNG VON NIEDERBAYERN (1997): Verordnung über die Einschränkung des Betretungsrechts im Nationalpark Bayerischer Wald, revised 12/18/2020. Available online at https://www.nationalpark-bayerischer-wald.bayern.de/ueber_uns/aufgaben/doc/vo_einschr_betretungsrecht_2020_mit_karte.pdf, checked on 3/7/2022.

REHRL, K.; GRÖCHENIG, S. (2016): A Framework for Data-Centric Analysis of Mapping Activity in the Context of Volunteered Geographic Information. In *ISPRS Int. J. Geo-Inf.* 5 (3), p. 37. DOI: 10.3390/ijgi5030037.

ROICK, O.; HAGENAUER, J.; ZIPF, A. (2011): OSMatrix - Grid based analysis and visualization of OpenStreetMap. Available online at https://www.researchgate.net/publication/220021502_OSMatrix_-_Grid_based_analysis_and_visualization_of_OpenStreetMap.

ROUSELL, A.; ZIPF, A. (2017): Towards a Landmark-Based Pedestrian Navigation Service Using OSM Data. In *ISPRS Int. J. Geo-Inf.* 6 (3), p. 64. DOI: 10.3390/ijgi6030064.

SEHRA, S. S. (2014): Assessing the Topological Consistency of Crowdsourced OpenStreetMap Data. In *Human Computation* 1 (2). DOI: 10.15346/hc.v1i2.13.

SEHRA, S. S.; RAI, H. S.; SINGH, J. (2015): Quality Assessment of Crowdsourced Data against Custom Recorded Map Data. In *Indian Journal of Science and Technology* 8 (33). DOI: 10.17485/ijst/2015/v8i33/79884.

SEHRA, S. S.; SINGH, J.; RAI, H. S. (2017): Assessing OpenStreetMap Data Using Intrinsic Quality Indicators: An Extension to the QGIS Processing Toolbox. In *Future Internet* 9 (2). DOI: 10.3390/fi9020015.

SEHRA, S. S.; SINGH, J.; RAI, H. S.; ANAND, S. S. (2020): Extending Processing Toolbox for assessing the logical consistency of OpenStreetMap data. In *Transactions in GIS* 24 (1), pp. 44–71. DOI: 10.1111/tgis.12587.

SELTMANN, A.; ZINK, J. (2021): OpenStreetMap in Nationalparken - Chancen, Folgen und Herausforderungen. In : FOSSGIS 2021 - Anwenderkonferenz für Freie und Open Source Software für Geoinformationssysteme, Open Data und OpenStreetMap. Rapperswil / Online., Jun 06-09, 2021. FOSSGIS e.V., pp. 52–54.

SENARATNE, H.; MOBASHERI, A.; ALI, A. L.; CAPINERI, C.; HAKLAY, M. (2017): A review of volunteered geographic information quality assessment methods. In *International Journal of Geographical Information Science* 31 (1), pp. 139–167. DOI: 10.1080/13658816.2016.1189556.

SEVERINSEN, J.; ROISTE, M. DE; REITSMA, F.; HARTATO, E. (2019): VGTrust: measuring trust for volunteered geographic information. In *International Journal of Geographical Information Science* 33 (8), pp. 1683–1701. DOI: 10.1080/13658816.2019.1572893.

SIEBRITZ, L.; SITHOLE, G.; ZLATANOVA, S. (2012): Assessment of the homogeneity of volunteered geographic information in South Africa. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives* 39.

TEIMOORY, N.; ALI ABBASPOUR, R.; CHEHREGHAN, A. (2021): Reliability extracted from the history file as an intrinsic indicator for assessing the quality of OpenStreetMap. In *Earth Sci Inform* 14 (3), pp. 1413–1432. DOI: 10.1007/s12145-021-00675-6.

VAN EXEL, M.; DIAS, E.; FRUIJTIER, S. (2010): The impact of crowdsourcing on spatial data quality indicators. Available online at https://www.researchgate.net/publication/267398729_The_impact_of_crowdsourcing_on_spatial_data_quality_indicators.

VAN OORT, P. (2006): Spatial Data Quality: From Description to Application.Dissertation. WAGENINGEN UNIVERSITY.

VANDECASTEELE, A.; DEVILLERS, R. (2013): IMPROVING VOLUNTEERED GEOGRAPHIC DATA QUALITY USING SEMANTIC SIMILARITY MEASUREMENTS. In *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XL-2/W1 (2), pp. 143–148. DOI: 10.5194/isprsarchives-XL-2-W1-143-2013.

VANDECASTEELE, A.; DEVILLERS, R. (2015): Improving Volunteered Geographic Information Quality Using a Tag Recommender System: The Case of OpenStreetMap. In Jokar Arsanjani, J., Zipf, A., Mooney, P. J., Helbich, M. (Eds.): OpenStreetMap in GIScience. Experiences, research, and applications. Cham, Heidelberg: Springer International Publishing (Lecture Notes in Geoinformation and Cartography), pp. 59–80.

WANG, M.; LI, Q.; HU, Q.; ZHOU, M. (2013): QUALITY ANALYSIS OF OPEN STREET MAP DATA. In *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XL-2/W1, pp. 155–158. DOI: 10.5194/isprsarchives-XL-2-W1-155-2013.

WU, H.; LIN, A.; CLARKE, K. C.; SHI, W.; CARDENAS-TRISTAN, A.; TU, Z. (2021): A comprehensive quality assessment framework for linear features from Volunteered Geographic Information. In *International Journal of Geographical Information Science* 35 (9), pp. 1826–1847. DOI: 10.1080/13658816.2020.1832228.

YAN, Y.; FENG, C.-C.; HUANG, W.; FAN, H.; WANG, Y.-C.; ZIPF, A. (2020): Volunteered geographic information research in the first decade: a narrative review of selected journal articles in GIScience. In *International Journal of Geographical Information Science* 34 (9), pp. 1765–1791. DOI: 10.1080/13658816.2020.1730848.

ZACHAROPOULOU, D.; SKOPELITI, A.; NAKOS, B. (2021): Assessment and Visualization of OSM Consistency for European Cities. In *ISPRS Int. J. Geo-Inf.* 10 (6), p. 361. DOI: 10.3390/ijgi10060361.

ZIELSTRA, D.; ZIPF, A. (2010): A Comparative Study of Proprietary Geodata and Volunteered Geographic Information for Germany. In Painho, M., Santos, M. Y., Pundt, H. (Eds.): Proceedings 2010. The 13th AGILE International Conference on Geographic Information Science. Guimarães, Portugal. Association of Geographic Information Laboratories in Europe.

APPENDIX

A 1 – Overview of peer-reviewed studies on the quality of OSM road networks

A 2 – Maps on attribute completeness and classification correctness of access restrictions in wildlife sanctuaries

A 1 - Overview of peer-reviewed studies on the quality of OSM road networks*

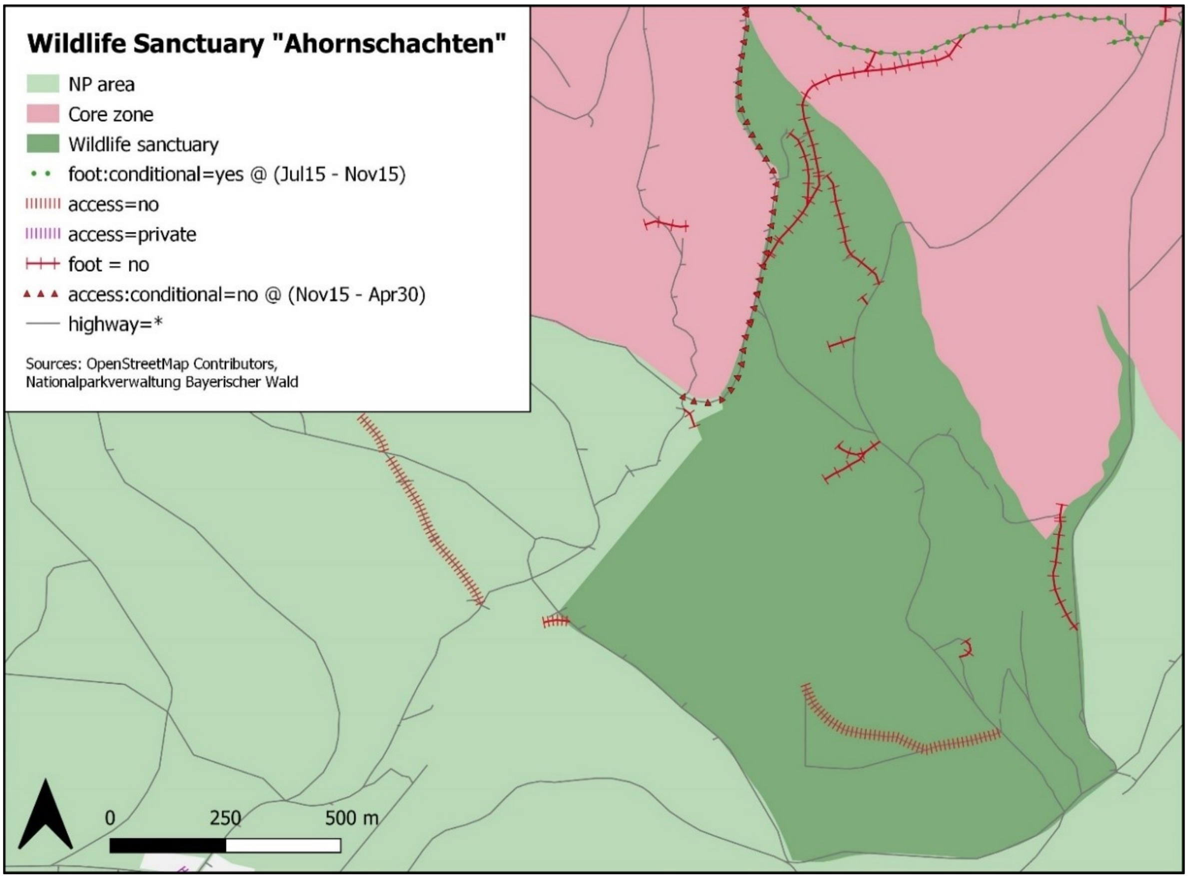
Authors & Year	Study Area(s)			Approach		Assessed quality categories and elements
	Region	Urban	Rural	Extrinsic	Intrinsic	
HAKLAY 2010	London & whole England	x		x		Completeness, Positional Accuracy
GIRRES AND TOUYA 2010	France	x	x	x	x	Completeness, Logical consistency, Positional accuracy, Thematic accuracy, Temporal quality, Usability
MONDZECH AND SESTER 2011	Hannover (Germany)	x	x	x		Usability (for pedestrian navigation)
NEIS ET AL. 2012	Germany	x	x	x		Completeness, Logical consistency, Temporal quality
KOUKOLETSOS ET AL. 2012	London, Newcastle (UK)	x	x	x		Completeness
SIEBRITZ ET AL. 2012	South Africa	x			x	Temporal quality
CANAVOSIO-ZUZELSKI ET AL. 2013	USA	x		x		Positional accuracy
KEßLER AND GROOT 2013	Münster (Germany)	x			x	Trustworthiness
BARRON ET AL. 2014	San Francisco (USA), Madrid (Spain), Yaoundé (Cameroon)	x			x	Completeness, Positional accuracy, Logical consistency, Thematic accuracy, Temporal quality, Usability (several use cases)
FORGHANI AND DELAVAR 2014	Tehran (Iran)	x		x		Completeness, Positional accuracy, Logical consistency
GRASER ET AL. 2014	Vienna (Austria)	x		x		Completeness, Positional accuracy
GRÖCHENIG ET AL. 2014	12 metropolitan areas	x			x	Completeness
SEHRA 2014	Punjab (India)	x		x		Logical consistency (topological consistency)
ABDOLMAJIDI ET AL. 2015	Scania (Sweden)	x	x	x		Completeness
BALLATORE AND ZIPF 2015	Several regions in Germany and UK	x	x		x	Conceptual compliance (measure from a proposed framework to assess conceptual quality / thematic accuracy)

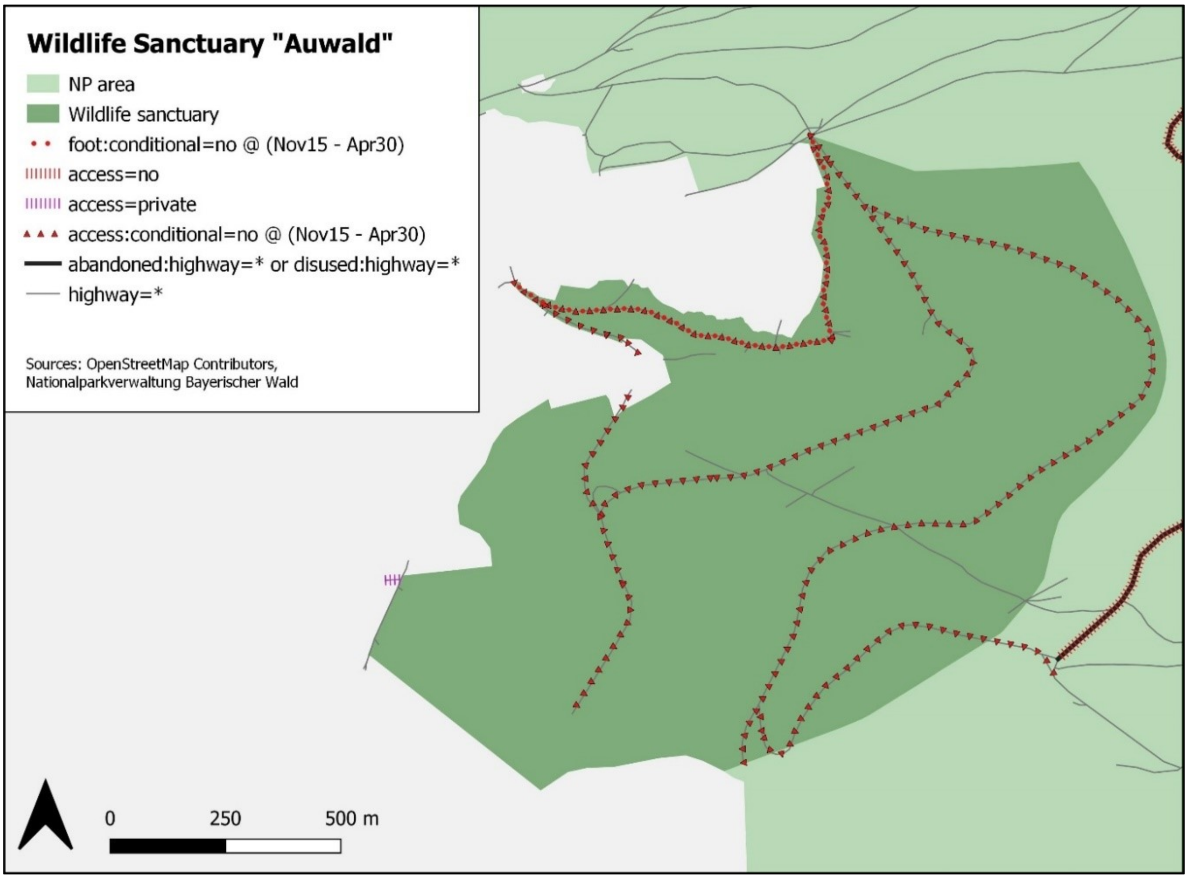
Camboim et al. 2015	Brazil	x	x	x	x	Completeness, Temporal quality
Hochmair et al. 2015	USA (78 urbanized areas)	x		x		Completeness
Mohammadi and Malek 2015a	Tehran (Iran)	x			x	Positional accuracy
Sehra et al. 2015	India			x		Completeness, Positional accuracy, Thematic accuracy
Brovelli et al. 2016	Paris (France)	x		x		Completeness, Positional accuracy
Fan et al. 2016	Heidelberg (Germany), Shanghai (China)	x		x		Completeness, Positional accuracy
Brovelli et al. 2017	Erba (Italy)	x		x		Completeness, Positional accuracy
Mobasheri et al. 2017	Germany (several cities)	x		x	x	Completeness, Usability (routing for people with limited mobility)
Sehra et al. 2017	Punjab (India)	x			x	Completeness, Thematic accuracy, Usability (route navigability)
Almendros-Jimenez and Becerra-Teron 2018	Main cities of Spain	x			x	Tagging quality (Thematic accuracy) based on conceptual quality framework (completeness, compliance, consistence, granularity, richness, trust)
Chehreghan and Ali Abbaspour 2018	Tehran (Iran)	x		x		Completeness
Muttaqien et al. 2018	Jakarta (Indonesia)	x		x	x	Completeness, Logical consistency (topological), Thematic accuracy, Newly developed indicator of "Aggregated Expertise"
Nasiri et al. 2018	Tehran (Iran)	x			x	Completeness, Positional accuracy
Sehra et al. 2020	Punjab (India)	x		x	x	Logical (topological) consistency, Thematic accuracy, Usability (road navigability)
Jacobs and Mitchell 2020	Ottawa-Gatinaeu (Canada)	x			x	Positional accuracy, Experience
Alghanim et al. 2021	London (UK)	x			x	Thematic accuracy, Trustworthiness, Credibility

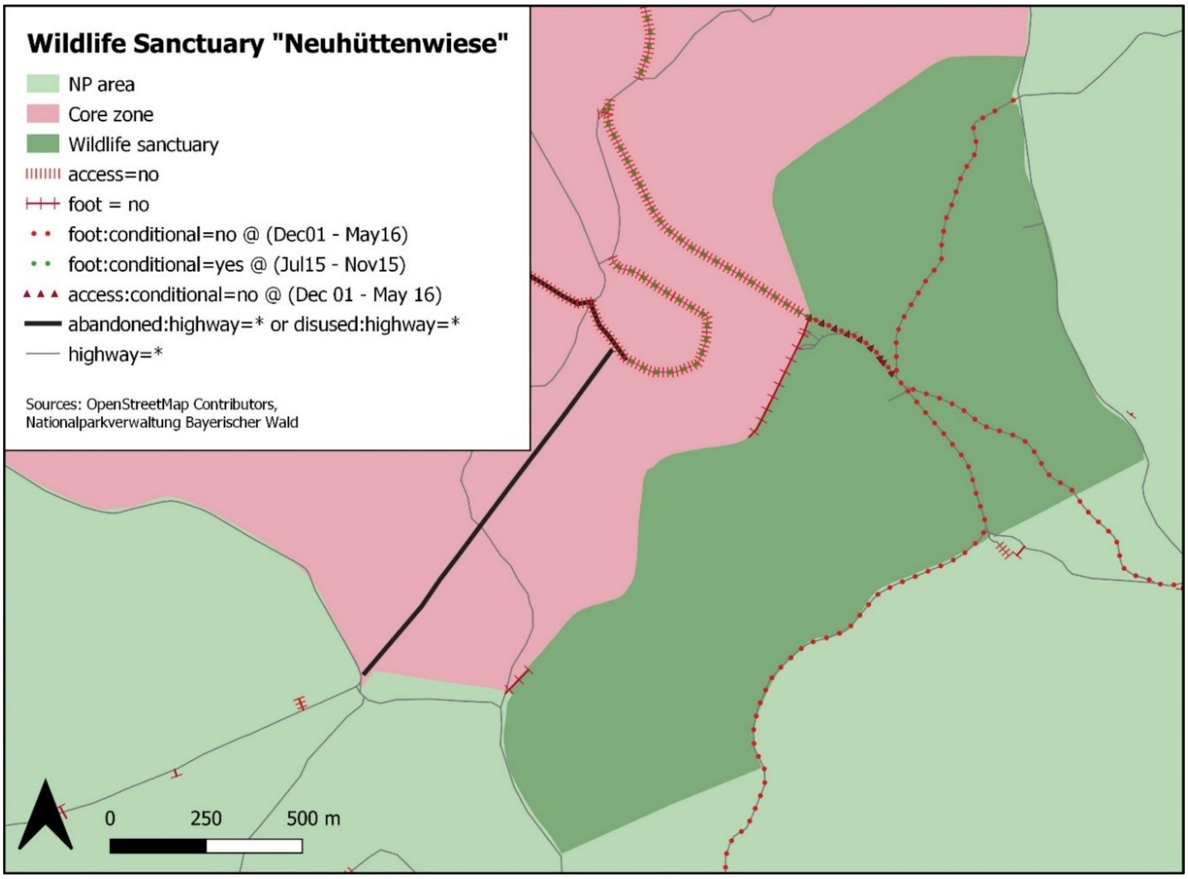
TEIMOORY ET AL. 2021	Tehran (Iran)	x			x	Trustworthiness
WU ET AL. 2021	Allegheny County, Pittsburgh (USA)	x	x	x	x	Completeness, Positional accuracy, Logical consistency
ZACHAROPOULOU ET AL. 2021	Athens, Berlin, Paris, Utrecht, Vienna, Zurich	x			x	Logical consistency

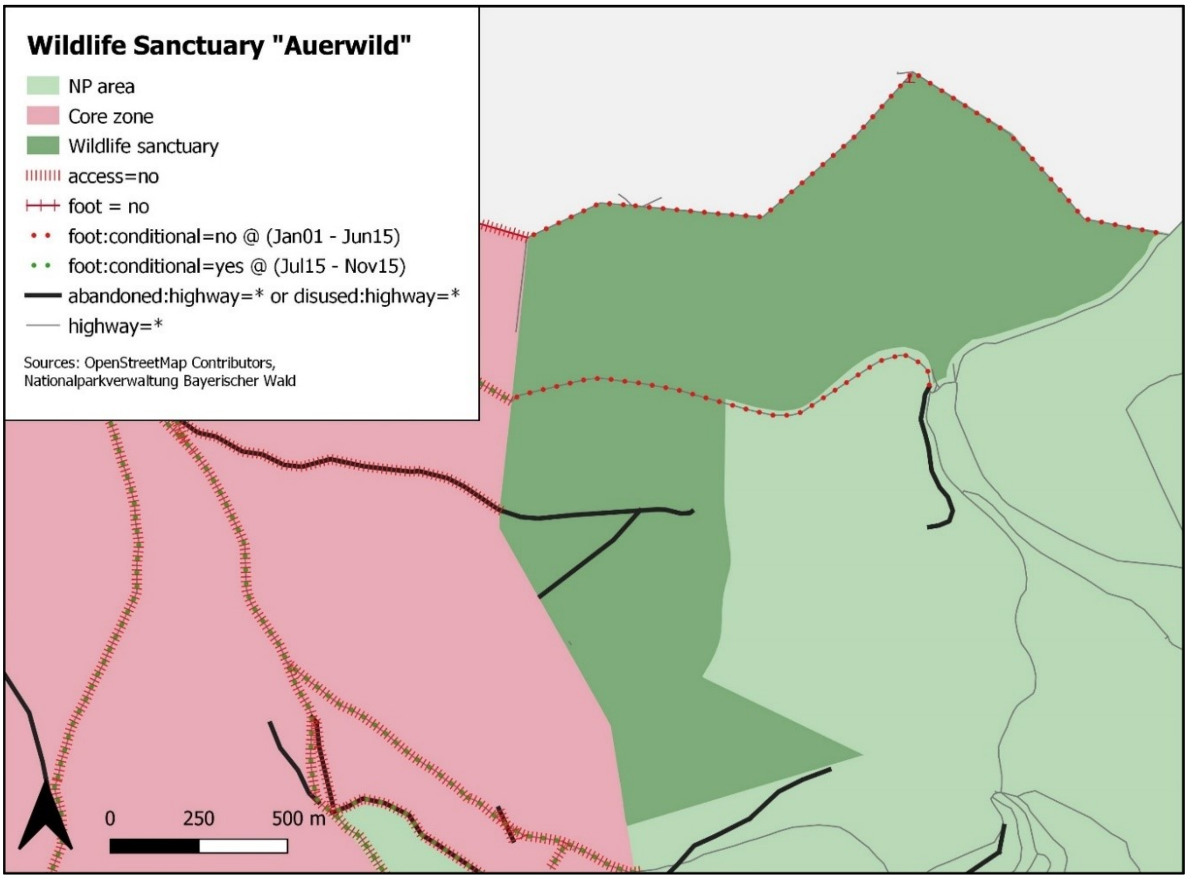
* This overview comprises only studies that were partly or solely assessing OSM road networks / line features (ways) and that were published in scientific journals, sorted by year of publication. A lot more studies have been conducted on other geometry types (POIs, polygons) and/or published in another scientific format (e. g. as conference papers resulting from scientific congresses or symposiums).

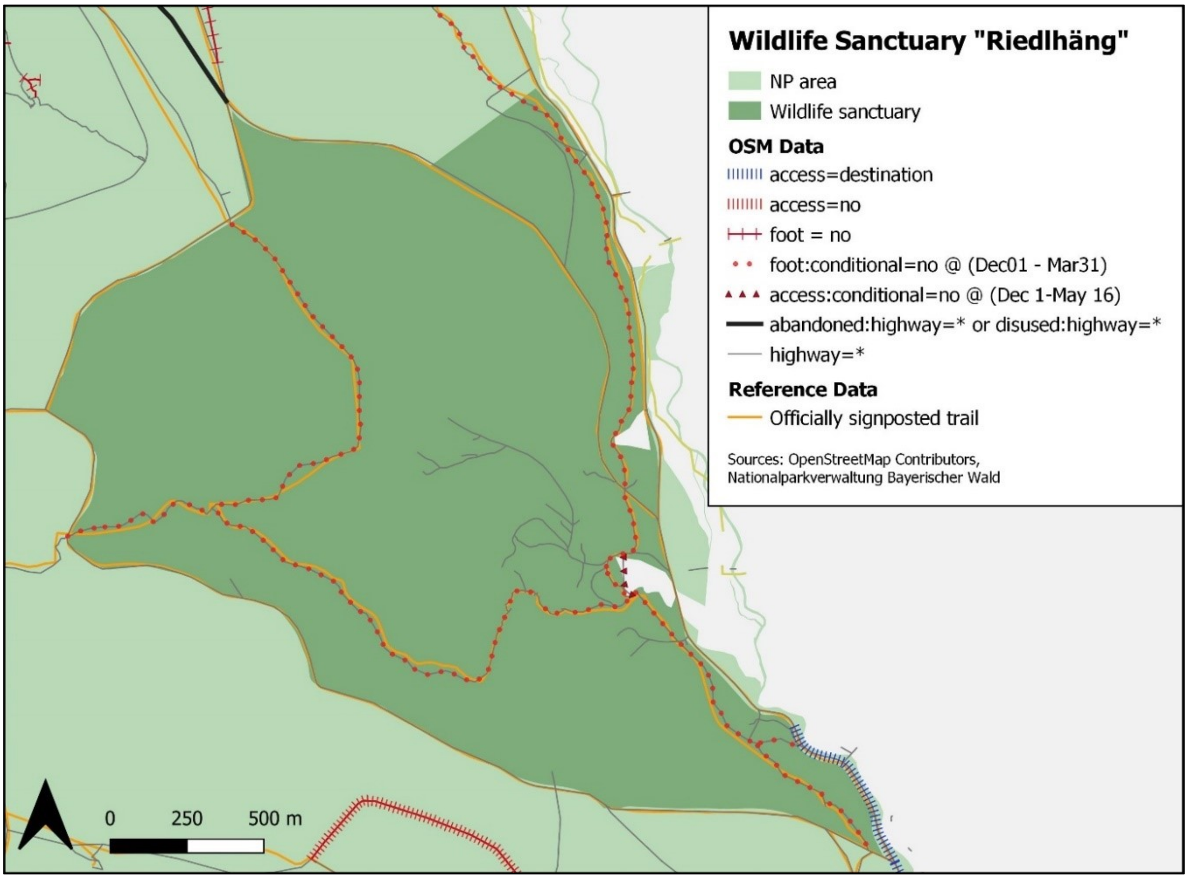
A 2 - Maps on attribute completeness and classification correctness of access restrictions in wildlife sanctuaries






Wildlife Sanctuary "Neuhüttenwiese"
NP area
Core zone
Wildlife sanctuary
access=no
foot = no
foot:conditional=no @ (Dec01 - May16)
foot:conditional=yes @ (Jul15 - Nov15)
access:conditional=no @ (Dec 01 - May 16)
abandoned:highway=* or disused:highway=*
highway=*
Sources: OpenStreetMap Contributors,
Nationalparkverwaltung Bayerischer Wald
0
250
500 m


Wildlife Sanctuary "Auerwild"
NP area
Core zone
Wildlife sanctuary
access=no
foot = no
foot:conditional=no @ (Jan01 - Jun15)
foot:conditional=yes @ (Jul15 - Nov15)
abandoned:highway=* or disused:highway=*
highway=*
Sources: OpenStreetMap Contributors,
Nationalparkverwaltung Bayerischer Wald
0
250
500 m


Wildlife Sanctuary "Riedlhäng"
NP area
Wildlife sanctuary
OSM Data
access=destination
access=no
foot = no
foot:conditional=no @ (Dec01 - Mar31)
access:conditional=no @ (Dec 1-May 16)
abandoned:highway=* or disused:highway=*
highway=*
Reference Data
Officially signposted trail
Sources: OpenStreetMap Contributors,
Nationalparkverwaltung Bayerischer Wald
0
250
500 m