University of Zurich[UZH]

Perception of landscape along hiking trails in the UNESCO Biosphere Entlebuch, Switzerland

GEO 511 Master's Thesis

Author
Fabian Biland
15-716-749

Supervised by
Prof. Dr. Ross Purves
Vuokko Heikinheimo (vuokko.heikinheimo@syke.fi)

Faculty representative
Prof. Dr. Ross Purves

19.08.2021
Department of Geography, University of Zurich

Perception of landscape along hiking trails in the UNESCO Biosphere Entlebuch, Switzerland

Fabian Biland
15-716-749

Master Thesis
19 August 2021

Prof. Dr. Ross S. Purves
Department of Geography - University of Zurich

Dr. Vuokko Heikinheimo
Department of Geosciences and Geography - University of Helsinki

Abstract

Cultural ecosystem services, a subcategory of ecosystem services, refer to locations where ecosystems offer non-material benefits to people. These benefits are represented by the spiritual, cultural, or recreational importance of a place and are in particular beneficial to mental health and human well-being.

Geotagged pictures, especially from *Flickr*, have become a widely used source of data to study cultural ecosystem services. Their content and location can help to analyze the relationship between humans and landscapes and also provide information about landscape usage or which landscape features contribute to a cultural ecosystem service. In contrast, unstructured text data in combination with social media data is rarely used when studying cultural ecosystem services, especially not to extract their location.

Previous research has introduced two main approaches to extract the location of cultural ecosystem services from social media data. The first approach is the straightforward technique of manually searching for cultural ecosystem services in geotagged image data. Annotating landscape features which appear in the same pictures as the cultural ecosystem services, allow finding possible relations. The second approach is more automatic and detects cultural ecosystem services by locating raster cells that contain many contributors. This technique has not been used to look for relations with landscape features extracted from the pictures of the individual contributors. The relations allow detecting landscape features that are relevant for providing cultural ecosystem services. Additionally, the results of both, the manual and the automatic approach, have never been compared before. Since the techniques have mostly been applied on geotagged image data only, this thesis also aims to fill the research gap on how well they perform when applied to text data.

Consequently, the following two research questions are tackled in this thesis:

- Which technique (manual or automatic) is better suited for text or image data to identify the cultural ecosystem services along hiking trails?

- How can text data help to find relationships between cultural ecosystem services and landscape features?

The UNESCO Biosphere Entlebuch, located in the center of Switzerland, is used as a case study area. It is well-known for recreational activities such as hiking and mountain biking because of its unique landscapes. The manual and automatic approaches are applied on a total of 721 pictures located within 40 meters of a hiking trail segment inside the UNESCO Biosphere Entlebuch. The text data consists of 75 tour descriptions from the cantonal map portal which at least partly lead through the UNESCO Biosphere Entlebuch.

The comparison of the manual and automatic approaches to detect cultural ecosystem services and investigate their relationship to landscape features has revealed that depending on the objective and detail either method can be used. While the manual method is much better suited when comparing different types of CES, the automatic method performs better when an overview of an area and its hotspot is required. Furthermore, the text data lead to different results when compared to picture data only but also reveals new insights when combined with text data. Additionally, text data helps to reduce the influence of pictures when investigating the relationships between cultural ecosystem services and landscape features. This supports the claim that text data should be used more often to study relationships between cultural ecosystem services and landscape features, especially in combination with picture data. Additionally, since this thesis is the first study focusing on a single region, the results can be of interest to local stakeholders. Especially the management of the UNESCO Biosphere Entlebuch and the participating municipalities can use the maps and the results. This can support decision-making and management action in the study site to preserve landscape features contributing to important cultural ecosystem services along hiking trails or grant protection status to them.

Keywords: Landscape Preferences, User-Generated content, UGC, Flickr, Unstructured Text, Hiking, Cultural Ecosystem Services, Landscape Features

Acknowledgements

I would like to sincerely thank my two supervisors, Prof. Dr. Ross S. Purves and Dr. Vuokko Heikinheimo for their guidance throughout this thesis. Their immediate feedback and fruitful ideas were much appreciated!

I am also thankful to Dr. Florian Knauss from the Federal Institute of Technology (ETH) and the UNESCO Biosphere Entlebuch, who, unfortunately, unsuccessfully applied for a Strava Metro data set for this thesis.

A big thank you also goes to the following people:

- Simon Biland, one of my proofreaders, but more importantly, my brother, who always had a sympathetic ear and valuable tips for improving my programming skills during my studies.
- Laura Wysling for proofreading as well as getting my thoughts off the thesis during coffee and lunch breaks.
- Fabienne Christen for testing my annotation rules as well as the continuous exchange of ideas about our theses.
- My parents, who unconditionally supported me not only throughout my education but my entire life.
- My girlfriend and all other friends for the continuing support, fun, and entertainment during the last semesters.

Contents

List of Figures

List of Tables

Acronyms

API	**A**pplication **P**rogramming **I**nterface
CES	**C**ultural **E**cosystem **S**ervices
ELC	**E**uropean **L**andscape **C**onvention
ES	**E**cosystem **S**ervices
GPS	**G**lobal **P**ositioning **S**ystem
KDE	**K**ernel **D**ensity **E**stimation
LF	**L**andscape **F**eatures
MEA	**M**illennium **E**cosystem **A**ssessment
PPGIS	**P**ublic **P**articipatory **G**eographic **I**nformation **S**ystem
SUD	**S**ocial Media **U**ser **D**ays
UBE	**U**NESCO **B**iosphere **E**ntlebuch
UGC	**U**ser **G**enerated **C**ontent
UNESCO	**U**nited **N**ations **E**ducational, **S**cientific and **C**ultural **O**rganization

Introduction

1.1. Motivation

Cultural ecosystem services (CES) represent aspects derived from nature that represent the spiritual or recreational importance of ecosystems (e.g. themed path, traditional customs, summit crosses). These should be conserved and sustainably used (Cardinale et al., 2012; MEA, 2005). In Europe, landscape influenced by humans form a major part of CES due to their cultural value (Agnoletti and Emanueli, 2016) and their positive contribution to the quality of people's life (Díaz et al., 2015). Since cultural ecosystems are not spatially continuous, they can be linked to landscape features (LF) (e.g. rocks, water bodies, grasslands), which are also spatially distinct (Smith and Mark, 2003). Regional or global changes in climate pose a serious threat for CES (MEA, 2005; Palomo, 2017) and protected areas (Foden et al., 2013). Additionally, research has shown that human activities are responsible for the loss of biodiversity and the destruction of ecosystems (Maxwell et al., 2016; Schirpke et al., 2020; van Zanten et al., 2016a). Moreover, ecosystem function degradation (MEA, 2005) and the transformation of ecoregions (Woolmer et al., 2008) have to be better understood in terms of human pressure and their spatial and temporal trends to prevent a collapse of ecosystems (Steffen et al., 2015; Venter et al., 2016). The combination of these threats urges conservation planners and park managers to take appropriate measures to preserve the coexistence of humans, vulnerable species, and their habitats (Hannah et al., 2007). The biospheres of the United Nations Educational, Scientific and Cultural Organizations (UNESCO) are sites where interdisciplinary approaches are tested to create a sustainable setting where humans and nature exist side-by-side (UNESCO, 2020; UNESCO Biosphere Entlebuch, 2021b).

The goal of this thesis is to locate CES in the UNESCO Biosphere Entlebuch (UBE) by comparing two different approaches, a manual approach of identifying CES in photographs and hiking tour descriptions and a more automatic approach of detecting active locations. Then I will identify LF in the two data sets and link them to the appearance of CES along hiking trails in the UBE. The results can help park managers to better protect valuable ecosystem services (ES) from future challenges (MEA, 2005), increase knowledge about factors influencing CES, and improve the understanding of recreational activities and CES.

1.2. Research Questions

This thesis aims to answer the following research questions. The research gaps will be pointed out in Chapter 2.

1. How do the results of locating CES differ between two methods applied on two sources of data?
2. How can different sources of data be used to better investigate the relation between CES and LF along hiking trails?

State of the Art

In this chapter, I aim to show the state of the art of this field of research. First, I demonstrate the importance of landscapes, their features, and discuss the cultural ecosystem services (CES) provided by them. Then, I briefly introduce methods used to extract the locations of CES and show examples of user-generated content (UGC) as well as unstructured text in the field of CES.

2.1. Landscape and its Perception

Landscape is considered useful for studying the relationship between humans and their environment (Plieninger et al., 2015b). The European Landscape Convention (ELC) (2000, p.2) defines "Landscape" as follows: "Landscape means an area, as perceived by people, whose character is the result of the action and interaction of natural and/or human factors." Landscapes of "outstanding universal value" (UNESCO, 1972, p.3) shall be preserved through UNESCO's World Heritage Convention. Because the concept of cultural ecosystem features (Section 2.2) had not been established yet, their definition of cultural heritage mainly focuses on cultural traditions (UNESCO, 1972). The ELC (2000), however, recognized services by landscapes to be a very important part of the quality of life for humans because they represent places of peace and solitude to exercise or follow specific activities (Depellegrin et al., 2012; The Research Box et al., 2009).

Because the definition of landscape in the ELC (2000) includes its perception by people, it can be described by using different types of data. People's perception can either be acquired in the form of pictures or text data (e.g. experience reports or personal stories) (Scott, 2002). Schirpke et al. (2016) used a photo-based survey to model the relationship between spatial landscape patterns and aesthetic values. Higher aesthetic values were found in mountainous areas because more viewpoints could be found along hiking trails. They only took 24 360° pictures into account which limits the studies' outcomes. By using a multiple linear regression model they estimated the aesthetic value of viewpoints along the hiking trails.

There is a Swiss landscape monitoring framework called *LABES* which includes biophysical and cultural indicators of landscapes (Kienast et al., 2015). By using *LABES* Kienast et al. (2015) found out that people living in rural areas perceive their surroundings much more positively when compared to people living in urban areas. They also stated that landscape quality in exceptionally urban or rural areas is very high, but lower in periurban areas. The low quality is explained by the loss of cultural landscapes and valuable green spaces (Kienast et al., 2015).

2.2. Cultural Ecosystem Services (CES)

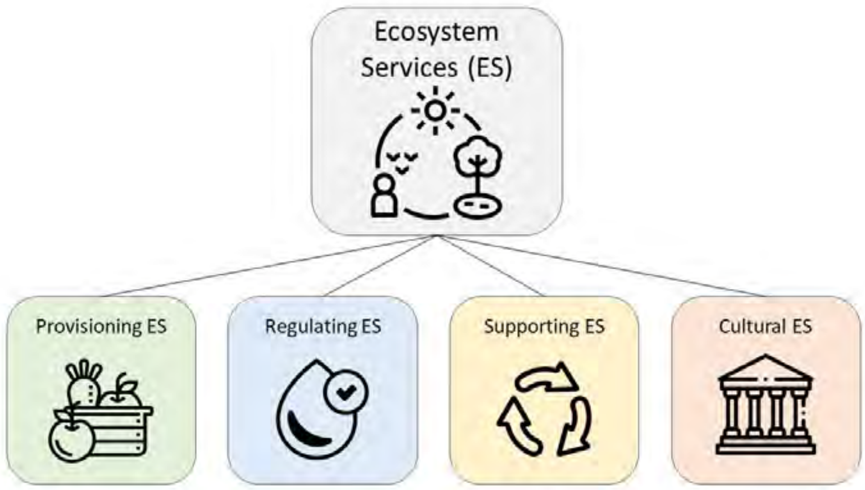


Figure 2.1.: *Overview of the different types of ecosystem services. Icons from Flaticon (2021).*

Ecosystems provide benefits to humans, so-called ecosystem services (ES). Figure 2.1 provides an overview of the different types of services. They include provisioning services such as providing food and other physical resources, regulating services to regulate floods or water quality, supporting services such as soil formation, and cultural services such as spiritual or recreational services (MEA, 2005). CES offer immaterial benefits to people. Specifically, ecosystems can influence cultural diversity, religious and spiritual values, knowledge systems, educational values, aesthetic values, social relations, sense of place, cultural heritage values and recreation, and ecotourism. All these aspects form an important part of culture (MEA, 2005). Church et al. (2014, p.8) offered a different definition of CES: "Environmental spaces [...] within which people interact with the natural environment and the cultural practices (e.g. exercising and playing) that define these interactions and spaces". This definition is much more focused on the environment and activities but still touches on immateriality.

Because CES form a relatively young field of research, the lack of common terminology leads to different interpretations in the different fields of research (Milcu et al., 2013). For example in the MEA (2005), only three types of CES were assessed, while missing out on assessing the other seven (Bieling and Plieninger, 2013). Over the course of time, the term CES represented typologies as for example cultural services (Costanza et al., 1997; MEA, 2005), cultural and amenity services (de Groot et al., 2010) or cultural fulfillment (Wallace, 2007). The common aspect of all terms is their intangibility due to immateriality (Milcu et al., 2013).

Within the subtypes of CES, there is a big diversity of terms to describe a specific subtype and the definitions are vague (Daniel et al., 2012). For example, *sense of place*, *place identity* and *place attachment* all emerged from different fields and try to describe the relation between people and places (Wartmann and Purves, 2018).

The concept of CES also has shortcomings. The concept has been developed by disciplines within natural sciences. When exploring CES, taking into account sociological, anthropological, or psychological disciplines is indispensable (Daniel et al., 2012). The quantification of CES is much more difficult than of the other subtypes of ES. While the intangibility (Milcu et al., 2013) or ambiguity (Daniel et al., 2012) is an often reported problem with CES, other types of ES are mostly dependent on natural attributes instead of human presence (Zoderer et al., 2016) and therefore easier to quantify. Therefore, quantification of CES often relies on mostly economic proxies such as expenditure by tourists or the number of visitors (Bateman et al., 2013; Ghermandi et al., 2009). Each human's cultural perception is different and CES are built up of collective cultural perceptions. Moreover, studying cultural landscape values is very challenging since people most often do not reflect on aesthetic services or immaterial enrichments (CES) provided by the landscape. Additionally, if surveyed, people tend to have problems articulating their thoughts (Stephenson, 2008) because they are not used to speaking about nonmaterial values (Bieling and Plieninger, 2013).

Importance of CES

CES are important for human well-being (MEA, 2005) and can enhance the connection between social and ecological issues (Milcu et al., 2013). They can be used to more deeply integrate people's perspectives and social values (Kumar and Kumar, 2008), shape people's identities, and motivate their actions through cultural attachment (Church et al., 2014). This leads to a stronger integration of social and ecological concepts which is in the favor of the sustainability ideal (Milcu et al., 2013). There are multiple ways in which people benefit from CES. They include recreational activities, aesthetic enjoyment, intellectual development, cultural heritage, artistic and spiritual fulfillment (MEA, 2005). CES are also characterized by the fact that they do not consume any natural resources (MEA, 2005).

The use and importance of ecosystems for recreational and tourism purposes have increased with growing population and income. Especially for developing countries tourism represents the primary economic development strategy, which in turn increases their dependence on ecosystems and their service (MEA, 2005). The Research Box et al. (2009) asked participants to reveal a collection of places which they would access depending on the situation or the type of experience they wanted. Well-known places, such as headlands, high points, or forests, were among the top answers. While some people considered the landscape and its CES as an integral part of their lives, other participants mainly saw it as an opportunity to exercise or as entertainment.

From a research point of view, CES are apparent in many different disciplines. This leads to the fact that their strength is to overcome gaps in different research communities. They also help to understand issues affecting social as well as ecological aspects. By paying more attention to their societal relevance, real-world problems could be addressed much more effectively (MEA, 2005; Milcu et al., 2013).

Threats to CES

The desire to meet the growing demands for food, freshwater and other resources resulted in a degradation of ecosystems and an irreversible loss of biodiversity (Cardinale et al., 2012). In particular urban (MEA, 2005) and periurban areas (Kienast et al., 2015) are affected by the loss of ES. Urban development becomes increasingly important with growing numbers of people living in those areas. Unsustainable urban development degrades CES which reduces the qualities of factors contributing to human well-being. Provisioning and regulating values of ecosystems are as important to local communities as spiritual and cultural services, as they can form important recreational services (e.g. urban parks) or spiritual services (e.g. sacred groves) (Kienast et al., 2015; MEA, 2005). But the degradation process of subtypes of CES is variable. Spiritual and religious values, as well as aesthetic values, show a decrease in quantity and quality, while recreation and ecotourism are rather tied. Mainly because more areas have become accessible but in return, many of them degraded as well. The degradation of CES is risky because the relationship between services and well-being is not linear. A rather small reduction of already scarce CES can have a substantial impact on human well-being. But the loss of specific CES or their threatened attributes could also enhance the appreciation towards them or the remaining services (MEA, 2005). Since we and our well-being are highly dependent on ecosystems and all their services provided, the ultimate goal is to reverse the degradation process of ecosystems while still keeping up with the increasing demands for their services (Cardinale et al., 2012). The conditions of these specific services vary depending on the conservation strategies or policies (MEA, 2005). All ES, whether they are provisioning, regulating, or cultural ecosystems are expected to be found in worse conditions by 2050 in one out of four scenarios when compared to 2005. Because long-term sustainability is endangered, ES need to be protected (MEA, 2005).

The importance of CES also shows vulnerabilities. While regulating and provisioning services can be replaced (e.g. bottled water can replace a polluted well), the loss of CES is irreplaceable (Guo et al., 2010). Even though globally speaking, humanity's dependence on regulating or provisioning services is decreasing, its dependence on CES is increasing. This is the result of improved living standards and more leisure time, as well as urbanization and intensive agricultural farming, which raises the demand for CES, especially in industrialized countries and societies (Guo et al., 2010). Since tourism has increased its contribution to the world's GDP for the 9th consecutive year in 2019 (UN World Travel Organization, 2020), it contributes to the higher demand for CES. CES are attractive to visitors as well as to the local population (Plieninger et al., 2015a).

Despite all these dangers and growing pressures on ecosystems, three out of the four scenarios in the MEA (2005) predict a reduction of the negative consequences if appropriate policies and practices are set in place from an early stage. Actions to prevent ecosystem degradation have shown promising results, but more importantly, they are still overwhelmed by the growing pressures and demands (MEA, 2005). The MEA (2005) also points out that due to the lack of monitoring, assessing the consequences of changes in (cultural) ES is complicated. In addition to that, management decisions often target the economic value of services entering markets, and all other services are neglected or degraded.

2.2.1. Mapping CES

According to the MEA (2005) the consequences of the loss of CES are very difficult to assess. To ensure successful management of all types of ES, understanding their spatial distribution is essential but also challenging, especially for cultural services because of their intangibility (Casalegno et al., 2013; Richards et al., 2018). Mapping CES is a powerful way to explain landscape values to different stakeholders and therefore helps to identify important areas for future park managements (Plieninger et al., 2013). Locating and identifying CES, however, was not among the researcher's top priorities (Ambrose-oji and Pagella, 2012) and was mainly done by mapping all ES and allocating the services to the existing subcategories, with the result, that CES were not the primary target (Bagstad et al., 2013; Bateman et al., 2013). Even now there is still a lack of understanding regarding the spatial distribution of cultural services as well as the nature which underpins these services (Long et al., 2021).

First attempts to map CES have used the numbers of recreational visitors for specific areas (e.g. Hill and Courtney (2006)). But also other measures, such as tourist expenditure (O'Farrell et al., 2011) or the numbers of reported observations of rare species (Raudsepp-Hearne et al., 2010) were used. Even though all these measures showed weak correlations with the spatial distribution of other types of ES, they seemed to be a good proxy for CES (Casalegno et al., 2013). Because quantifying the importance of CES is rather difficult (Bieling and Plieninger, 2013; Gliozzo et al., 2016), crowd-sourced picture and text data from activities form a very valuable source of information to assess the importance of cultural landscape values and eventually map them.

Social media data, especially geotagged pictures are a predestined data source to detect CES and have been increasingly used to extract locations of recreational or aesthetic CES on a regional (Casalegno et al., 2013; Gliozzo et al., 2016; Richards and Friess, 2015; Tenerelli et al., 2016) or continental scale (van Zanten et al., 2016a). Social media data has been used in rural landscapes (e.g. Pinto-Correia et al. (2011); López-Santiago et al. (2014)) and in studies to investigate place attachment (Gliozzo et al., 2016) or landscape characteristics (Tenerelli et al., 2016). They are also said to be able to report cultural heritage values and to be used as a tool to build a shared history (van Dijck, 2011). Consequently, social media data can be very useful to map different subtypes of CES.

Antoniou et al. (2010) were among the first researchers to investigate the geographic distribution of geolocated pictures. They found that users rather take and share pictures of popular places instead of relatively unpopular niches. This leads to the assumption that there is a correlation between the popularity of a place and the density of photographs (Antoniou et al., 2010; Wood et al., 2013). Furthermore, users tend to show selective behavior when sharing information online (Antoniou et al., 2010). Identifying hot and cold spots is necessary to be able to set up or adjust development strategies that would relieve pressure on wildlife and ecosystems in specific areas (Ghermandi, 2016).

Following Antoniou et al. (2010), multiple researchers have taken up the idea of mapping CES by using UGC-data (Casalegno et al., 2013; Gliozzo et al., 2016; Richards and Friess, 2015; Tenerelli et al., 2016). Most of these studies covered different areas including multiple types of landscape, which made it difficult to compare CES between areas. This is why the field of mapping CES has been divided up into different subfields: urban (Guerrero et al., 2016), non-urban (Bieling, 2014; Gliozzo et al., 2016), agricultural landscape (van Zanten et al., 2016b) or mountainous areas (Zoderer et al., 2016).

Mapping Techniques

The methods of participatory mapping, interdisciplinary approaches, or surveys did not achieve the accuracy planners or sustainable land managers need (Plieninger et al., 2013; Scholte et al., 2018). In addition, the identification of CES often relies on complex research teams which monitor an area over a long period of time, as in Bateman et al. (2013). Heterogeneous research teams are required because both natural and social sciences have to be included.

By more frequently integrating social media data into mapping processes, the data collection becomes less time-consuming. Basic spatial analysis techniques such as the number of pictures (Antoniou et al., 2010; Richards and Friess, 2015; Tenerelli et al., 2016) or the number of contributors (Casalegno et al., 2013; Gliozzo et al., 2016) per raster cell (de Smith et al., 2018) can be used. Because users uploading multiple pictures with similar content from the same location (e.g. plane spotter) would lead to a bias, the number of contributors or social media user days (SUD) is better suited to locate CES (Tenkanen et al., 2017).

Casalegno et al. (2013) were the first to use the density of contributors to extract hotspots of aesthetic value and observed a negative correlation between population density and aesthetic value. They mainly focused on differences between inland and coastal areas and concluded that coastal hotspots with a high number of contributors outnumbered the inland hotspots having a lower number of contributors. Gliozzo et al. (2016) applied the same method on nonurban environments by excluding residential areas and comparing different photo-sharing platforms. Antoniou et al. (2010) mainly focused on urban areas, but used the absolute number of pictures.

More advanced studies applied regression models in combination with social media photographs to detect CES (Tenerelli et al., 2016). Long et al. (2021) have developed a model to predict the most valuable places for recreational activities at 540 recreation sites across Europe by using geo-referenced images in combination with environmental metrics and the population density to estimate the number of visitors per km^2 (Long et al., 2021). Long et al.' (2021) results revealed that population centers have a high influence on these valuable sites. Remote locations showed fewer attractive sites than areas close to population centers. Oteros-Rozas et al. (2018) have contributed to mapping ES by analyzing UGC like previous studies (Brown and Raymond, 2007; Guerrero et al., 2016). They used geo-referenced pictures from two different platforms and found, that *Flickr* was better at representing CES at three locations out of five locations, including Obersimmental in the Swiss Prealps.

Besides only including the location of pictures, the content of pictures was also used to identify CES. By joining the coordinates from geotagged pictures to pictures containing CES, these services could also be mapped. Content classification of photographs was performed by Dorwart et al. (2009), while Richards and Friess (2015) were the first to classify geotagged online photographs. They used 9 categories to detect CES in their pictures (nature appreciation (focus on animal or plant), landscape (wide view), social recreation (people), fishing recreation, history, research, infrastructure, other). Their choice of categories seems rather arbitrary since infrastructure does not ultimately mean that a boardwalk represents a CES. This opens up the question of defining CES again and that many ES can be thought of as CES, even a boardwalk. Just like a hiking trail, boardwalks are anthropological features and can be used for recreational purposes which, according to the definition (MEA, 2005), are a CES.

For picture data, CES have not been mapped by using density estimations of pictures containing CES. This approach and the detection of CES by the number of contributors (e.g. Casalegno et al. (2013), Gliozzo et al. (2016)) have never compared to each other. Consequently, it is unknown how each method contributes to mapping CES and what characteristics are revealed by each method.

All previously mentioned studies focused on photographs derived from social media platforms. Gliozzo et al. (2016) argues that the extraction of the locations of CES from text data can be inaccurate and is technically demanding. *Twitter*, however, has also been used to map places that are important to people and offer benefits (Jenkins et al., 2016). By using geotagged tweets and methods from natural language processing clusters of a collective sense of place could be identified (Jenkins et al., 2016). In addition, personal stories and experiences have a high influence on the detection of CES and are often underrepresented or missed out (Bagstad et al., 2013). But they have been included for example in Bieling (2014), who asked residents to submit short stories. This was very time-consuming, especially since the stories had to touch upon several questions. Surprisingly, 93% of all short stories addressed CES, most of them only focusing on this type of ES. Nevertheless, the effect of integrating text data with social media data to map CES is not known yet. Furthermore, it has not been assessed, whether the methods mentioned in the previous paragraph which were applied to photographs can be applied to text data.

2.3. Landscape Features (LF) and CES

Understanding people's preferences in protected areas (Hausmann et al., 2017a) or the main factors attracting people (Hausmann et al., 2017b) is very important and builds a solid foundation to extract CES. Due to the intangibility of CES, LF or biophysical features can be used as a proxy for the identification of CES and are useful to describe the human-nature relationship (Bieling and Plieninger, 2013; Bieling, 2014; The Research Box et al., 2009). Adding LF to the concept of ES, CES in particular, is important because CES are often tied to the perception and cultural recognition of these LF (Kirchhoff, 2012) and consequently attract people. This means, similarly to CES, LF are important for people's individual and societal well-being. This is why the ELC (2000) urges to protect LF because of their cultural value (Agnoletti and Emanueli, 2016). Additionally, understanding the contribution of LF to CES is important for landscape planning (Bieling et al., 2014; Plieninger et al., 2015a).

The Research Box et al. (2009) were among the first researchers to investigate whether cultural services correlate with specific landscape characteristics or features by using interviews. The study reveals that landscape provides many important services and benefits to humans but also that cultural services are delivered by all kinds of combinations of LF. Their main message is that landscape should be preserved and each service is represented by different landscape aspects and fulfills different needs. Participants in participatory mapping studies have shown that they are willing to relate CES to specific LF (Plieninger et al., 2013) but this type of data collection, as well as interviews (The Research Box et al., 2009), is very time-consuming.

Photographs have served as a landscape analysis tool since the 1960s because they have the potential to capture individual perceptions of landscape (Depellegrin et al., 2012). To get information about LF, Depellegrin et al. (2012) asked experts to annotate landscape classes and landscape attributes seen in pictures. Oteros-Rozas et al. (2018) also investigated the important relationship between LF and CES based on the content of photos. By looking for relations between LF and CES at five different sites in Europe, it was possible to compare the landscape values. The annotation was manually done by the co-authors. However, they did not take into account the exact location of pictures within the different sites. For example, Obersimmental, a Prealpine region in Switzerland, showed high relations between the LF *mountain* and *recreational activities* (e.g. skiing or hiking) is predictable. Also, other relations such as aquatic CES and water bodies are rather obvious but support the methodological approach. The work by Oteros-Rozas et al. (2018) shows that manually extracting CES in combination with LF from crowdsourced images works and is promising. All this calls for a more detailed analysis at one specific site because decision-makers usually don't happen to deal with landscapes that are similar to Lesvos (Greece), Madrid (Spain), and Obersimmental (Switzerland) at the same time. Relations between CES and LF on the scale of one predominantly homogeneous area might reveal new insights.

Research with the content of social media is not only done in rural areas, but also urban spaces. Recently, Wang et al. (2021) analyzed the perception of public green spaces from unstructured text data (user comments) and developed a lexicon that captured LF. They then used machine learning algorithms to detect the satisfaction from these LF and found out that the quality of LF in urban parks is very important to the users as well as recreational facilities, water, and service architecture.

Most studies assessing and identifying CES within landscapes deal with places offering aesthetic values (Casalegno et al., 2013; Depellegrin et al., 2012; Martínez Pastur et al., 2016; Vlami et al., 2021; Zoderer et al., 2016). For example, Vlami et al. (2021) launched an extensive study looking at an entire park network to identify CES based on 22 proxy indicators. Their findings showed serious degradation of aesthetic values because of these developments. The study by Martínez Pastur et al. (2016) tried to relate biophysical characteristics (e.g. water bodies or vegetation types) to aesthetic values of the landscape in Patagonia. Other researchers investigated the aesthetic value of landscape in coastal zones from social media tags since they face strong anthropogenic pressure (Depellegrin et al., 2012). Their results revealed that almost 75% hotspots were close to water bodies, which is not surprising in a study dealing with coastal zones. Zoderer et al. (2016) assessed and showed the aesthetic value of mountain regions by using spatial models in the Alpine region of South Tyrol (Italy) by using a photo-based questionnaire survey. Among their findings was that leisure activities and aesthetic beauty were the top CES reported in the Alpine landscape. Cultural heritage and spiritual services were not as frequently reported. Leisure and aesthetic enjoyment and spiritual services were often associated with landscapes that were used traditionally (e.g. larch meadow, hay meadow, permanent crop) but also intensively used hay meadows and spruce forests. Cultural heritage was often associated with landscapes that experienced strong human influence (e.g. intensively used permanent crop). Traditionally used landscapes were associated with hotspots of CES but were hard to identify on the map due to their size and their scattered distribution. All their results depended on cultural background, gender, environmental engagement, and knowledge about landscapes. It has to be mentioned that calling a place aesthetically valuable or identifying CES, in general, is very context-specific and subjective (Daniel et al., 2012). Other studies have targeted multiple CES (Oteros-Rozas et al., 2018; van Zanten et al., 2016a,b) or did not focus on specific subtypes of CES (Plieninger et al., 2013; Gliozzo et al., 2016).

van Zanten et al. (2016a) tried to extract aesthetic and recreational values from different photo-sharing social media platforms at a continental European scale, finding that mountainous areas are preferred. In another study, they focused on a specific landscape type and investigated agricultural landscapes and their LF' contribution to aesthetic and recreational values (van Zanten et al., 2016b).

Besides the content of a picture, tags offer another research opportunity. The use of tags assigned to pictures by the users and owners (e.g. Depellegrin et al. (2012), van Zanten et al. (2016a)) is questioned by Lee et al. (2019) because interpretation is much more difficult with unrestricted and non-standardized tagging. Additionally, a lot of pictures do not even contain tags, which eventually leads to a loss in information and number of pictures, if not analyzed (Lee et al., 2019). Newer methods have tried to use machine learning to improve content analyses of pictures and text data. For example, Lee et al. (2019) used computational methods to analyze the visual content of social media to identify CES. They used an image annotation engine that generates tags for the patterns detected in the picture. They then clustered the tags to identify hotspots of specific types of tags. But by this generalization, the complexity of the data was reduced in two steps (picture $\rightarrow$ tag $\rightarrow$ tag-group). Nevertheless, the automatic assignment of tags showed very high accuracies for the cluster *landscape aesthetic* (Lee et al., 2019). Google Cloud Vision, an online machine learning algorithm, was used by Richards et al. (2018) to analyze photographs regarding their content to map CES and look for relations with the content. Automatic tagging is very time efficient compared to manual annotation which is a big advantage when characterizing content of pictures (Richards et al., 2018). Gosal and Ziv (2020) used machine learning and social media images to model and map landscape aesthetics in the northern English protected area of the Yorkshire Dales National Park which contains several *Areas of Outstanding Natural Beauty* with limestone sceneries as well as farmland, which is very similar to this thesis' research area. Especially in rural areas, mountainous landforms and vegetation contributed to aesthetic value.

2.4. Research on Hiking

Since hiking is the most popular form of recreational activity in this thesis' research area, only data around and about hiking trails will be included. The hiking trail system is supposed to represent recreational activities in nature. Hiking is not considered a major field of research, but since the numbers of recreational activities are rising, it is expected, that the research's focus will be deflected towards hiking trails or recreational activities in the future (Hill et al., 2014). Additionally, the number of newly published articles, indicates that research on hiking trails is on the rise. Relevant studies have mainly targeted the hiker's motivation and social aspects on long-distance hiking trails (Fondren and Brinkman, 2019) or the place attachment and sense of place (Amerson et al., 2020; Kliot and Collins-Kreiner, 2020).

Amerson et al. (2020) states that the place identity, the emotional connection to a place, increased as people spent more time hiking the Pacific Crest Trail (PCT). This conclusion infers that day hikers possibly do not develop a sense of place while hiking and therefore other CES may be of interest. But this is also because long-distance trails are more suitable to conduct research on and the visitor profile is also different when compared to normal hiking trails. Studies on hiking trails representing the recreational area of a nature reserve including geolocated, UGC are still missing.

2.5. User-Generated Content (UGC)

Traditional ways of collecting data to study the interactions between people and nature such as surveys, interviews, counters, or global positioning system (GPS)-trackers, usually exceed the temporal and monetary resources available for researchers or managers of recreational areas (Waldron et al., 2013; Heikinheimo et al., 2020). This is why new ways of efficiently collecting information had to be developed (Toivonen et al., 2019). UGC fills this gap. At its early stage, UGC data was collected through citizen science and crowd-sourcing. Multiple factors have led to a change in data acquisition methodology (Di Minin et al., 2015):

- Continuous internet connection
- Smartphones and cameras taking high-resolution pictures
- Smartphones and cameras being equipped with GPS

These factors not only enabled a new crowd-sourcing technique, but also a new source of data available in the field of UGC: social media data (Di Minin et al., 2015). Heikinheimo et al. (2020, p.11) define social media data as "data from social networking sites that allow users to share (georeferenced) content online." The networking aspect is very important because it enables users to view personally or publicly shared content from other users and interact with them through likes, comments, or other types of reactions.

2.5.1. UGC Data

Attributes

Due to the possibility of geotagging photographs or tweets with coordinates and assigning locations to them, social media has become very valuable and widely used for geographic analyses (Ghermandi and Sinclair, 2019; Sui and Goodchild, 2011). While some studies only used the location of social media data (e.g. Donahue et al. (2018)), others included additional information, such as likes, posts, users information (Hausmann et al., 2017b), timestamp (Heikinheimo et al., 2017), or the content of photographs (Oteros-Rozas et al., 2018). The content of pictures is an important part of information whereby for example activities can be retrieved from

(Oteros-Rozas et al., 2018). Moreover, sophisticated data about the contributors, their main language, or their origin can be extracted (Di Minin et al., 2015; Heikinheimo et al., 2020; Sinclair et al., 2020). All this information can be used by land managers to, for example, restrict access to vulnerable areas (Lee et al., 2019). Resource limitations led to a neglection of analyzing the content of photos (Lee et al., 2019; Martínez Pastur et al., 2016; Richards and Friess, 2015).

Strengths and Weaknesses

Social media data offers multiple aspects which opened up a new branch and increased opportunities in science (Di Minin et al., 2015; Hausmann et al., 2017a; Heikinheimo et al., 2017; Tenkanen et al., 2017; van Zanten et al., 2016a). The sizes of social media data sets (Lawu et al., 2021) often exceed the size of traditionally collected data by several orders of magnitude (Ghermandi and Sinclair, 2019). Moreover, for poorly accessible areas or rather unknown areas, social media poses an attractive alternative source of data, which led to an increase in studies including user-generated data (Di Minin et al., 2015).

Social media data produced by its users allow researchers to examine and study relationships between nature and human behavior. This relationship is at least partly represented in social media data (Ruths and Pfeffer, 2014) and can be used to study individuals' perception and knowledge of the environment (Huang et al., 2013). Moreover, social media data has been successfully used in studies in protected areas and urban green spaces if they were frequently visited by people (Tenkanen et al., 2017) or they contained a topic which was popular on social media platforms (Toivonen et al., 2019). Social media does not only offer a very reliable proxy for large-scale visitation rates (Heikinheimo et al., 2017; Wood et al., 2013), it can also be used to identify hot spots and preferred places of visitors (Gliozzo et al., 2016; Sessions et al., 2016). Consequently, social media can be used as a reliable alternative to quantitatively and qualitatively analyze human behavior.

Despite all these benefits, social media also has its limitations. Drawing conclusions from user-generated data can be difficult and planners or decision-makers have to be careful when doing so (Dunkel, 2015). Especially conclusions drawn from data from a single social media platform should be either cross-validated with traditional data (Di Minin et al., 2015) (e.g. Donahue et al. (2018), Hausmann et al. (2017a), Wood et al. (2013)), combined with other social media data (Tenkanen et al., 2017), or other crowd-sourced data (Heikinheimo et al., 2020; Levin et al., 2017).

Platform-specific characteristics should be addressed and kept in mind when working with any type of social media. Each social media platform has a user-specific profile. For example, while *Instagram* is more often used by younger generations and captures everyday activities and predominantly people, *Flickr* is used by professional photographers who preferably capture nature in high-quality pictures (Tenkanen et al., 2017). Consequently, there is a need to better include underrepresented groups in analyses with social media (e.g. elderly, rural, and indigenous communities) (Oteros-Rozas et al., 2018).

Figure 2.2 by Heikinheimo et al. (2020) shows the applicability of different UGC data to answer specific questions. The closer a data source is to an interrogative, the better suited it is to answer such questions. Depending on the research question(s) different types of data should be

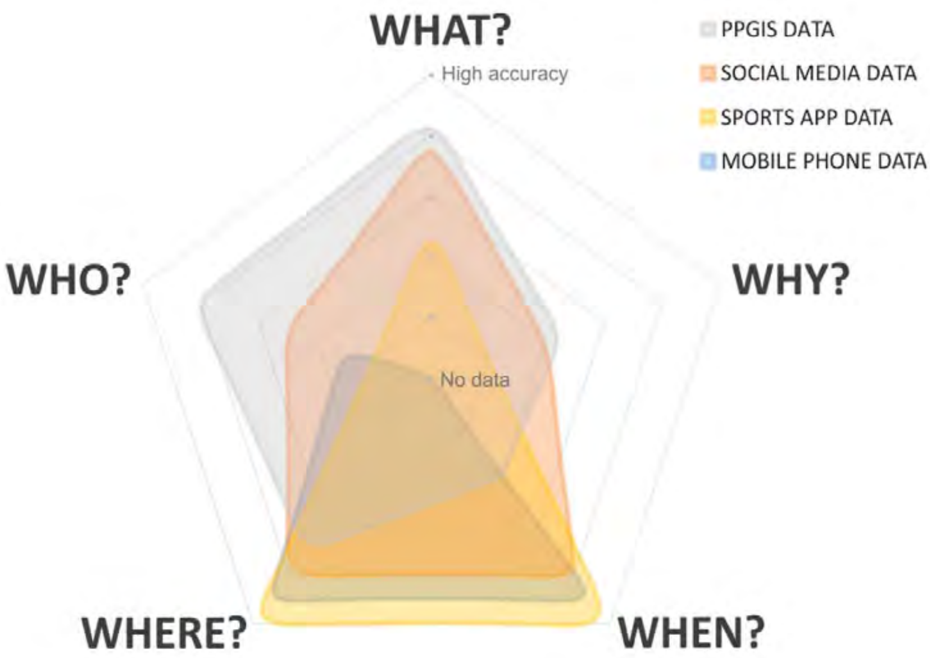


Figure 2.2.: *Comparison of different user-generated data sets to illustrate their applicability to answer the interrogatives where, when, what, why and who (Heikinheimo et al. (2020))*

included.

For example, the density of pictures (Section 2.2.1) does not directly relate to the people's interest in nature or the environment (Richards et al., 2018). The coordinates of photos allow studying the *where?* but also do not provide any information about the reasons *why?* people visited a specific location, which would be interesting and helpful to recreational managers. Therefore the content and the inclusion of additional information is crucial. By analyzing the content of photographs it is possible to understand what aspects of nature are interesting to people and consequently help to answer *why?* or *what?* questions (Richards and Friess, 2015). Not only the content allows answering specific questions, but also the tags of pictures. Ghermandi et al. (2020) used descriptions and tags of pictures as well as their content but finally decided to only use the tags and descriptions because they also revealed aspects that were not visible in the image.

The combination of spatio-temporal approaches with content analysis for social media data has been used previously concerning landscape values (van Zanten et al., 2016a) or human activities (Heikinheimo et al., 2017).

Wu et al. (2017) tried to monitor trail use (*what?/why?*) along linear features (*where?*) in urban areas. Their results were not satisfactory and were mainly restricted by commuting or shopping activities. This should not be a problem in this thesis' area under study, which is predominantly rural. But this study by Wu et al. (2017) is one of the few projects that tried to extract information about preferences along linear features instead of grid squares as in Heikinheimo et al. (2020) or entire national parks as in Hausmann et al. (2017a) or Norman and Pickering (2019). Richards and Friess (2015) found out that footpath routes were among the best explanatory variables for photographs, indicating that the majority of pictures is taken from locations close to a trail, supporting the idea of studying data along linear features. Nevertheless, they also mentioned that locomotion is difficult in a national park in Singapore because of natural or public restrictions. Whether pictures are as close to trails as in Richards and Friess (2015), in areas that are less populated than Singapore, is unclear. Therefore this raises the question of whether content analysis can be performed on pictures to derive CES and extract landscape elements along linear features.

Hence, conducting research on linear features, such as hiking trails, in a non-urban environment including social media to extract CES and LF and answer questions about *where?* and *why?/what?*, will fill a research gap.

Platforms

A lot of studies mentioned have dealt with analyzing the location and/or the content of photographs. But UGC can also originate from other sources (Figure 2.2). Sports data from *Strava* or other fitness applications (Norman et al., 2019; Norman and Pickering, 2019) or mobile phone data (Heikinheimo et al., 2020) have also been used.

Flickr and *Twitter* are the most used social media platforms in environmental research (Ghermandi and Sinclair, 2019). Other widely used platforms include *Panoramio* or *Instagram* (Casalegno et al., 2013; Di Minin et al., 2015; Wood et al., 2013). *Facebook* is the most popular social media platform in the world (Di Minin et al., 2015; Toivonen et al., 2019) but it is not suitable for extensive research because of its limited data access permissions(Tenkanen et al., 2017). In Europe, the Americas, and India, *Instagram* is the most popular platform due to more people owning smartphones with high-quality cameras (Toivonen et al., 2019; van Zanten et al., 2016a). Nevertheless, other photo-sharing platforms, such as *Flickr* and *Panoramio* have more frequently been used in environmental sciences (Casalegno et al., 2013; Gliozzo et al., 2016; van Zanten et al., 2016a; Wood et al., 2013). *Instagram* has been relatively unexplored in science compared to *Flickr* (e.g. Hausmann et al. (2017b), Hausmann et al. (2017a), Heikinheimo et al. (2017), van Zanten et al. (2016a)). Until 2017, 83% of studies have used a single social media platform with *Twitter* and *Flickr* representing the most frequently used platforms. The number of studies including multiple social media platforms has covered 15% of all studies using at least one social media platform in 2017 (Ghermandi and Sinclair, 2019). Tenkanen et al. (2017) and Toivonen et al. (2019) both suggest acquiring data from multiple platforms when working with social media and careful interpretations of results. A combination of multiple platforms hides platform-specific behavior and biases (Tenkanen et al., 2017) and helps to get more robust results a more complete picture on a conservation issue (Toivonen et al., 2019). Furthermore, the combination reduces temporal and spatial mismatch between platforms and different types of data, reducing biases explained in Section 2.5.2 (Di Minin et al., 2015; Heikinheimo et al., 2020; Toivonen et al., 2019). Studies that successfully include social media data from at least two platforms have been published by Ghermandi (2016) (*Flickr* and *Panoramio*) or Tenkanen et al. (2017) (*Instagram*, *Flickr* and *Twitter*). Both studies used the same type of data from different photo-sharing platforms. *Twitter* data is often used as text data as well as as the second source of data (Ghermandi and Sinclair, 2019; Jenkins et al., 2016). If the tweet is not geotagged, natural language processing techniques such as named entity recognition to add coordinates to toponyms are used (Toivonen et al., 2019). The use of geographically untagged text data sources has not been proposed or mentioned. Consequently, its possible application and usefulness is unclear (see Section 2.6). Geographic information retrieval methods on text data are rarely applied or investigated in combination with data from social media platforms.

Heikinheimo et al. (2020) followed the suggestion by Tenkanen et al. (2017) and Toivonen et al. (2019) and compared four different data sources (social media data, sports tracking data, mobile phone data, and public participation geographic information system (PPGIS) data). Figure 2.2 shows their results. Because different data sets are suitable to a varying degree to answer questions, they should be combined. The focus of a study and its methodology highly influence which social media platform(s) should be used because all of them offer different sources of data (e.g. pictures, videos, texts, likes) (Di Minin et al., 2015) and are unequally suitable to answer specific questions (Heikinheimo et al., 2020).

Recent privacy actions by the companies owning social media platforms have limited the variety of social media platforms available for certain research areas, especially geographic analyses. Table 2.1 provides an overview of the restrictions for highly used platforms. Most of the implications are the result of improved user privacy claims (McCrow-Young, 2020) which made application programming interfaces (API) inaccessible. Therefore, *Flickr* remains the best source of social media, especially for geotagged picture data (Lawu et al., 2021).

Table 2.1.: *This table provides an overview of the restrictions for some of the most frequently used social media platforms.*

Platform	**Implication**	**Source**
Geograph	Spatially restricted to British Isles	*Website Geograph*
Instagram	API not accessible since June 2020	McCrow-Young (2020)
Panoramio	API not accessible since November 2016	Tenkanen et al. (2017)
Twitter	Removed possibility of geotagging tweets in June 2019	Hu and Wang (2020)

In addition, a recently started campaign by the Schweizer Alpen-Club (2021) introduced the hashtag *#nogeotag* to reduce the number of people sharing their location when wild camping. This would decrease the number of people staying at the same place, limiting the negative impact on the environment. Whether and to what extent this campaign will be successful in reducing wild camping is not foreseeable. But if its introduction was successful and not only applied to wild camping but also other activities, it would have a negative impact on research with social media data, because fewer people would be sharing information.

2.5.2. UGC in Environmental Research

Until 2015, particularly in conservation science or environmental research, UGC was not used as extensively as in other fields (Di Minin et al., 2015). Since then, however, environmental studies have increasingly integrated social media data in their research (Ghermandi and Sinclair, 2019; Toivonen et al., 2019).

Applications of social media data in environmental science can be found in different fields. They range from visitor monitoring in protected areas (e.g. Heikinheimo et al. (2017), Levin et al. (2017), Tenkanen et al. (2017), Wood et al. (2013)) to understanding how (e.g. Heikinheimo et al. (2020), Norman and Pickering (2019), Norman et al. (2019)) and why (e.g. Hausmann et al. (2017b), Hausmann et al. (2017a), Norman and Pickering (2019)) protected areas or urban parks (e.g. Donahue et al. (2018), Wu et al. (2017) are used or to measure the perceived importance of such areas (e.g. Levin et al. (2017). Because social media posts may include information to answer questions such as *what?*, *where?*, and *when?* (Figure 2.2), it is in particular suitable to extract information about leisure time (Heikinheimo et al., 2020). Social media has also been used sentiment analysis towards the environment (Toivonen et al., 2019) or language identification (Heikinheimo et al., 2020).

Among the principal goals of all of these studies was to provide additional information and new insights to local authorities and park managements which should be included in their decision making and planning processes. This is mainly because the recreational use of green spaces has increased in the last decade which in turn exerts additional pressure on the ecosystems (MEA, 2005; Steffen et al., 2015; Venter et al., 2016).

2.6. Unstructured Text

The merging of two different data sets, not only two picture sharing platforms, but also different data types (e.g. text and pictures) has not been done very often (e.g. Tenkanen et al. (2017)) because of different reasons. Heikinheimo et al. (2018) reported that 50% of all geotagged tweets in their study were originally posted on *Instagram*. Consequently, *Twitter* loses a small amount of interest. The fact that *Twitter* removed the function to precisely geotag purely textual tweets (Table 2.1) also does not help its popularity among scientists working with geotagged text data.

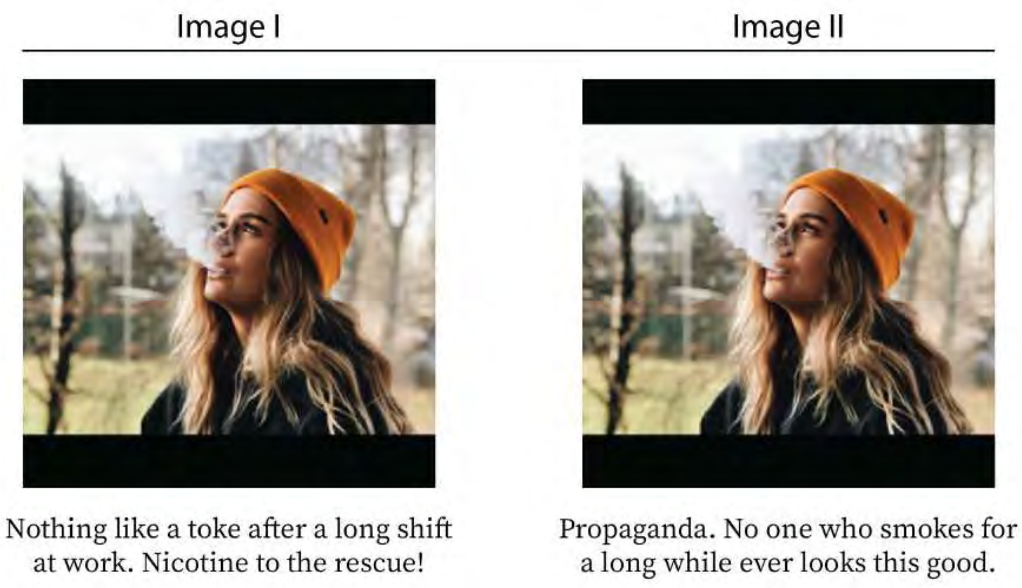


Figure 2.3.: *The image gets a different meaning depending on the caption (Kruk et al., 2019)*

Nevertheless, text data offers advantages, which should not be neglected. Multimodality is one of them and often appears when people try to express themselves through visual and textual content (Ramachandram and Taylor, 2017). The combination of multimodal social media data (text and image) leads to a gain in information because new meanings can be introduced which either source cannot do on its own, inherently increasing the performance of tasks (Kruk et al., 2019) (Figure 2.3). Tenkanen et al. (2017) and Heikinheimo et al. (2020) called for the integration of multiple data types, not only photos but also text or videos from social media platforms. Studies investigating the use of social media in combination with unrelated, independent text data are rare. Consequently, multimodality and geographically untagged and unstructured text data should be more often included in research. Wartmann et al. (2018) have proven that different types of data (free lists, hiking blogs, and *Flickr* tags) can be used to extract different landscape properties and consequently each data source has its characteristics (Figure 2.4). Using the three data sources they highlighted the potential of combining different types of bottom-up data to distinguish landscape types, not relying entirely on experts or single data types. Text data provide insights that other types of data and other methodologies to capture data do not cover. The internet, containing roughly 4 billion pages (de Kunder, Maurice, 2021), offers a big variety of unstructured text data which can be used in research.

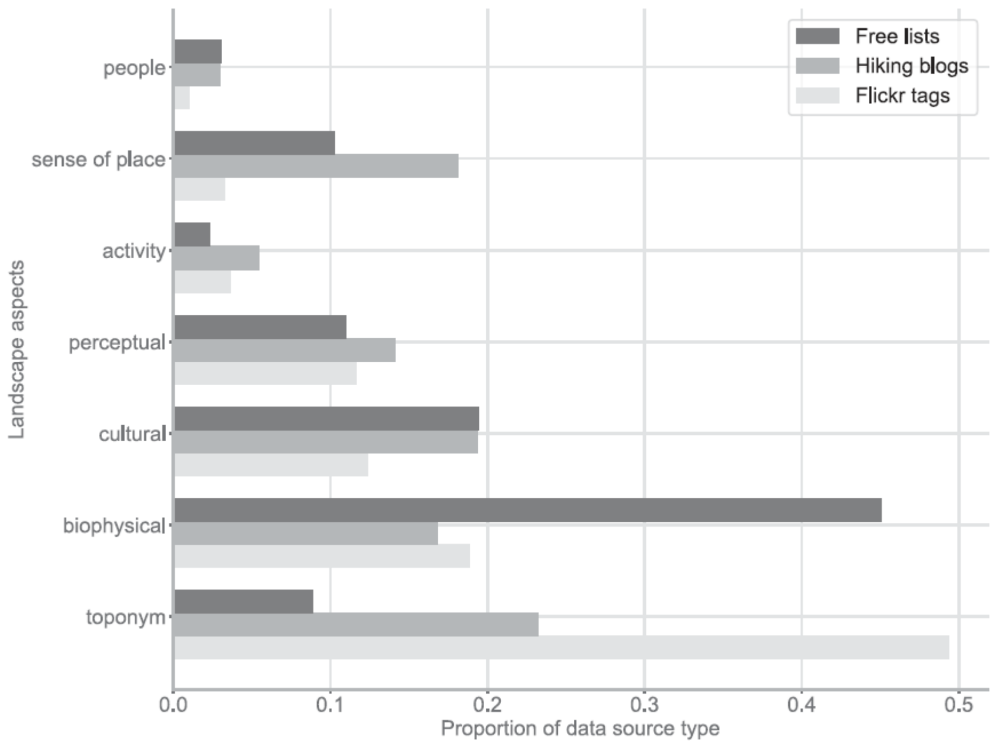


Figure 2.4.: *Comparison of retrieval of different landscape aspects depending on data source (Wartmann et al., 2018).*

The study by Bieling and Plieninger (2013) serves as a perfect example to illustrate the need to include text data. They assessed CES by fieldwalking and listing visible manifestations of CES in the landscape. They only included permanent signs of CES (visible for at least one year) and not, for example, the location of a person taking a picture. However, not all types of CES (e.g. spiritual services) could be assessed to a satisfying degree. Therefore, their method is better suited as a complementary method to more extensive surveys or interviews.

2.6.1. Unstructured Text in Environmental Research

Places become spiritually important by being referred to in literature (Cooper, 2019). This supports the call from Oteros-Rozas et al. (2018), who recommended including other types of data than only pictures in future research to detect CES and LF which have not been detected by the content of photos.

As shown in Figure 2.4, text data revealed promising results when capturing cultural landscape properties, sense of place and also performed well in capturing biophysical elements. Additionally, the proportion of toponyms was also quite low when compared to *Flickr* tags. But since only a subset of all contributors tags pictures, the representativeness of tags is questionable (Guerrero et al., 2016; Lee et al., 2019). Toponyms should still be represented in a decent ratio in text data, to make sure that geolocating annotations is possible (see Section 4.2.3). It is pleasing that roughly 20% of all terms are toponyms and that text data is not overfilled with toponyms. Wartmann et al. (2018) also proposed to extend their methods to less frequently visited places where web content would be decreasingly available, assuming that UGC is unevenly distributed across space (Antoniou et al., 2010; Casalegno et al., 2013).

Derungs and Purves (2014) used a corpus of Swiss alpine landscape descriptions and extracted and georeferenced natural LF to characterize individual regions and compare them to each other. This study was performed on a raster with variable size (between 5km and 40km cell size) and provides good results in characterizing the landscape of a specific raster cell. However, it remains unclear whether manual, instead of automatic, annotation of text data would be feasible on a finer geographic resolution or micro-scale or even on a segment level of linear features. Purves and Derungs (2015) used text and *Flickr* data separately to map vernacular regions. They used kernel density surfaces to map the locations of *Flickr* pictures containing a vernacular place name in their tags (e.g. *Alps* or *Cairngorms*). They also computed χ^2-maps which take the distribution of *Flickr* pictures into account and therefore show whether a vernacular place name is over-or underrepresented when compared to the global distribution of *Flickr* pictures. For text data, they identified unambiguous toponyms and then used Euclidean distances and topographic similarities (derived from digital elevation models) to map them. While this study used two different data sources and compared them, it did not merge them to get one single result from both sources.

Bieling (2014) used text data in the form of short written stories to detect CES. Another study retrieved text from the internet through the use of an API and then extracted descriptions of sights, sounds, and smells (Koblet and Purves, 2020). Lee et al. (2020) combined text mining with opinions of local people to identify CES. They successfully retrieved, mountain climbing and farm or history and culture programs.

Culturomics is a research area where text data is used to detect linguistic and cultural phenomena by analyzing changes in word frequencies (Michel et al., 2011). But it also has the potential to help conservation managers to document changes regarding the interaction between humans and the environment (Ladle et al., 2016). Ladle et al. (2016) present different applications within the field of conservation science where culturomics could be used (e.g. identify conservation flagships or icons, or classify and map CES at macro scale).

2.7. Research Gaps

This thesis aims to fill the following research gaps:

- When working with social media data, research suggests including different sources of data (Heikinheimo et al., 2020; Toivonen et al., 2019). Instead of an additional social media platform, I will use non-geotagged text data, as suggested by Oteros-Rozas et al. (2018) in combination with a social media platform.
- Research has introduced multiple methods to extract the location of CES by looking at the location of pictures (Casalegno et al., 2013; Gliozzo et al., 2016) or their content (Oteros-Rozas et al., 2018). These methods have never been compared to each other. Consequently, the comparison between the two most basic techniques of detecting CES manually through the content of pictures and automatically by counting the number of contributors per raster cell.
- Extracting the LF which contribute to CES has only been done by the manual method (Oteros-Rozas et al., 2018). Applying the automatic method and then extracting LF from the resulting hotspots has not been done before. Additionally, text data has not been used in either method.
- Research on hiking trails is rather scarce. Especially the integration of social media and text data along linear features is missing.

Research Area

3.1. Choice of Research Area

In a first step, all parks of the network *Swiss Parks* were considered as a research area for this project. Currently, the network consists of 19 different parks, which are distributed all over Switzerland from the French-speaking part of Switzerland (e.g. Parc Jura Vaudois or Parc Régional Chasseral) to northern Switzerland (e.g. Regional Nature Park Schaffhausen or Jurapark Aargau) and southeastern Switzerland (e.g. Swiss National Park or Biosfera Val Müstair).

Since it was desirable to include different types of hiking trails (normal hiking trails, mountain hiking trails, and alpine hiking trails) only two parks remained: The Parc naturel régional Gruyère Pays-d'Enhaut and the UNESCO Biosphere Entlebuch (UBE). Additionally, the park should be easily accessible by public transport to visit and get an impression of the study area during the thesis. Moreover, the main language in the Parc naturel régional Gruyère Pays-d'Enhaut is French and the available text data is in French. Therefore, the German-speaking UBE was chosen as the research area.

3.1.1. The Role of a UNESCO Biosphere

> *"Biosphere reserves are 'learning places for sustainable development'."*
> – UNESCO (2020)

This is how the UNESCO (2020) defines a biosphere reserve on their website. It is an area used for hands-on science where new approaches can be tested, applied, and evaluated to further investigate human interactions with nature. Generally speaking, these sites serve as model regions for other biospheres (UNESCO Biosphere Entlebuch, 2021b). There are 714 UNESCO

Biosphere reserves in 129 countries which include terrestrial, marine, and coastal ecosystems. Each site promotes solutions reconciling the conservation of biodiversity with its sustainable use (UNESCO, 2020). The UNESCO (2015) defined 17 sustainable development goals which form the basis of a UNESCO Biosphere. Humans and nature should find an equilibrium between usage and conservation while everybody should profit from the biosphere (UNESCO Biosphere Entlebuch, 2021a). Consequently, because the UNESCO Biosphere's goal, among others, is to conserve its landscape and the CES it is providing, the UBE emerges as a perfect study area.

Each UNESCO Biosphere reserve consists of three different zones where different activities are permitted (see Figure 3.1):

- The core areas are strictly protected and help to conserve landscapes, ecosystems, and genetic variation.
- The buffer zones surround the core areas and may be used used as a research area that is compatible with reasonable ecological practices.
- The transition area is where communities exert human activities in an ecologically sustainable way.

Figure 3.1.: *Illustration showing the different zones and the permitted activities (UNESCO, 2020).*

3.1.2. UNESCO Biosphere Entlebuch (UBE)

With the adoption of the *Rothenturm Initiative* in 1987, moors in Switzerland are placed under protection. In the Entlebuch, however, there was massive resistance against this, as there are over 100 moors in the Entlebuch, which are considered to be obstacles to development. Eventually, the idea of creating a UNESCO Biosphere was brought up and since 2001 the UBE is part of the UNESCO Biospheres. It consists of 7 municipalities (Doppleschwand, Entlebuch, Escholzmatt-Marbach, Flühli, Hasle, Romoos and Schüpfheim) which form the border of the Biosphere and cover an area of $395 km^2$. An overview of the UBE and its hiking trail system can be seen in Figure 3.3. Since 2008, the UBE is also a park of national importance and member of the network *Swiss Parks* (UNESCO Biosphere Entlebuch, 2021b).

The landscape of the UBE is famous for its numerous and spacious moorlands which are among the largest in Switzerland. The karst formations as well as the hilly landscape covered by forest- and river landscapes also characterize the biosphere reserve. Moreover, numerous viewpoints provide a spectacular view of the area (Figure 3.2) (UNESCO Biosphere Entlebuch, 2021c). The moorlands and the karst region of Schrattenfluh and the Napfregion are also part of the Bundesinventar der Landschaften und Naturdenkmäler (Federal Inventory of Landscapes and Natural Monuments). These landscapes are protected because of their contribution to either unique or typical landscapes or the history of landscape (natural monuments) or landscapes which offer great opportunities for recreation (Bundesamt für Umwelt (BAFU), 2017).

(a)

(b)

(c)

Figure 3.2.: *A selection of features which characterize the UBE: (a) Moorland, (b) Viewpoint, (c) Karst formations (all pictures taken by me on 09/10/2020 (a, b) and 25/10/2020 (c)).*

The UNESCO Biosphere Entlebuch's (2021a) slogan "Ein Segen für alle" ("A blessing for everyone") means that their focus lies in benefiting nature and local residents. They defined three main goals to achieve this (UNESCO Biosphere Entlebuch, 2020):

- Conserving diverse nature and culture
- Strengthening a powerful and innovative regional economy
- Leading a road to the future as a learning region and organization

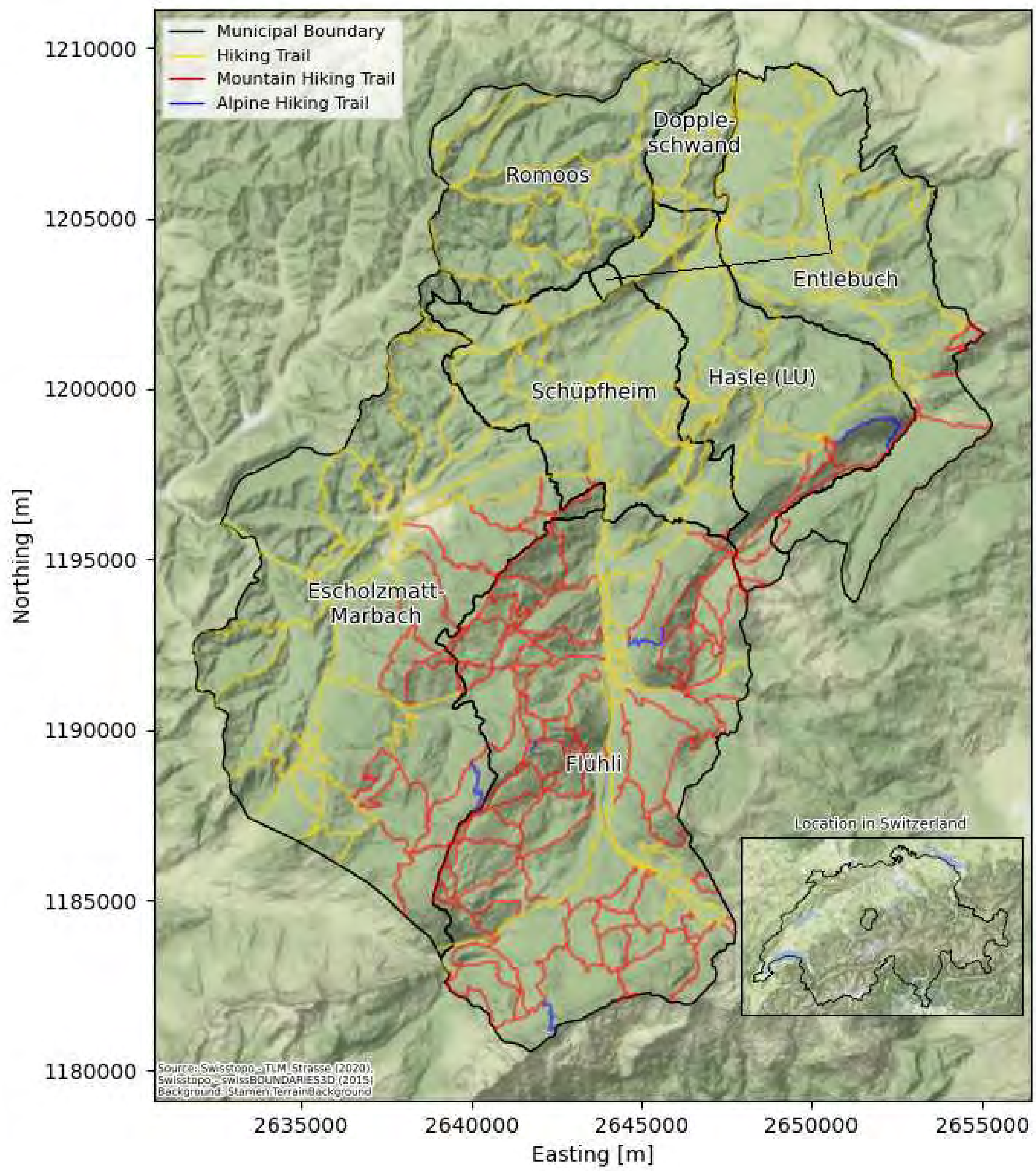


Figure 3.3.: *Overview of the UNESCO Biosphere Entlebuch with its municipalities and the hiking trail network.*

Data and Methods

This chapter introduces the data used for the analysis and underlines the reasons why they were chosen. An overview of all data sets used in this thesis can be found in Appendix A.1. Furthermore, this chapter explains the applied annotation and analysis methods in more detail. Figure 4.1 serves as an overview of the used methods.

4.1. Data

I used two different types of data for this study. On the one hand, there is the text data with tour descriptions in the UBE. On the other hand, I chose *Flickr* as a social media platform that provides photographs as well as metadata.

4.1.1. Tour Descriptions

At the beginning of the thesis, the preferred source of text data was experience reports from hikers or other people following mountaineering activities, such as *hikr.org*. Experience reports are individually written texts about new or already existing tours, which represent the opinions and experiences of users and therefore offer valuable information to other users. Since these texts are about the personal perception and experience and impressions of the viewer, these texts are predestined to study the perception of the landscape of hiking trails as in Bieling and Plieninger (2013) or Derungs and Purves (2016).

Experience reports are available on different platforms: *alltrails.com*, *gps-touren.ch*, *hikr.org*, *komoot.de*. But every single platform comes with its disadvantages. Either the tour reports are too few in number or essential parts were missing (*gps-touren.ch*, *alltrails.com*) or the content is too heavily supplemented with pictures (*komoot.de*) or the geographical search function is

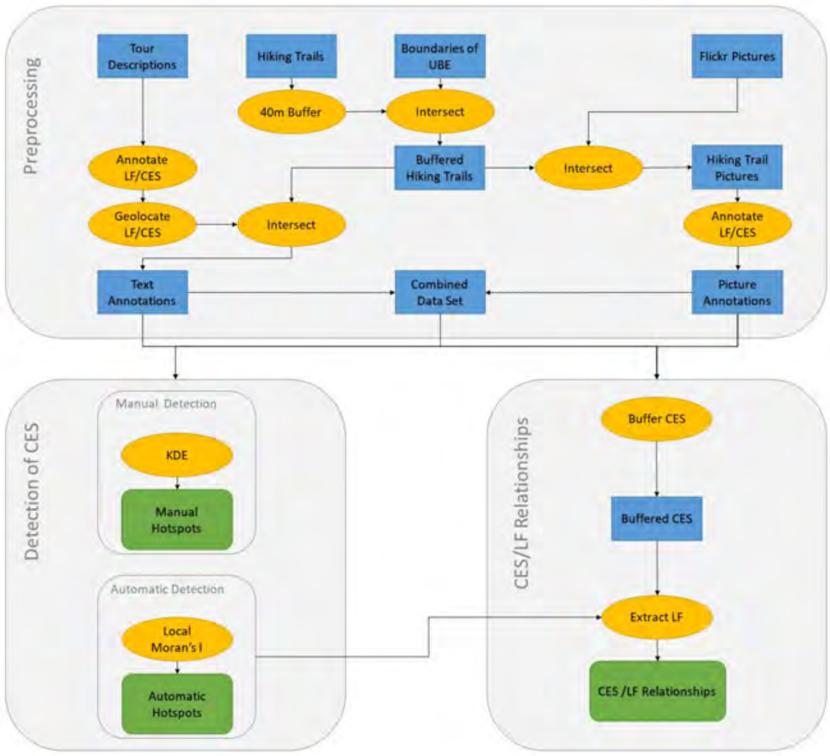


Figure 4.1.: *Overview of the workflow of this thesis (blue: data sets, yellow: actions, green: results).*

not precise enough for the UBE (*hikr.org*).

The choice finally fell on *maps.luzern.com*, the web portal of the canton of Lucerne, whose purpose is not to provide experience reports, but present and recommended activities (e.g. hiking, mountain biking, trail running) descriptively. Moreover, the tours show a high degree of completeness, and the tour descriptions are written in much more detail when compared to the other platforms. Additionally, a GPX file is available for each tour description.

Retrieval of Tour Descriptions

maps.luzern.com contains tour descriptions for various types of activities (e.g. hiking, biking, skating, snowshoeing) in the canton of Luzern. The website's spatial query was applied and returned the tours which at least partly lead through the UBE. For the thematic query, all activities in the following groups of activities were selected:

- *Alles in Wandern* (Everything in Hiking)
- *Alles in Bike und Velo* (Everything in Mountain Biking and Cycling)
- *Alles in Laufen und Skaten* (Everything in Running and Skating)
- *Alles in Bergrouten* (Everything in Mountain Tours)

A total of 83 tour descriptions were returned (date accessed: 09/12/2020).

The GPX files and tour descriptions were downloaded between 09/12/2020 and 29/12/2020. I retrieved the text of the tour descriptions by copying the three parts from the tour page which contained the necessary information and pasting them into a TXT file:

- The main part of the tour description describes the tour in great detail and contains the majority of information. It usually starts with a summary of the tour, followed by a more detailed description.
- *Autorentipps* (recommendations from the author): Possible recommendations are restaurants or specific places along the trail worth visiting. Sometimes taking a detour is required. Not every author/institution added this part when publishing the tour description.
- *Wegbeschreibung* (directions): provide a very short overview of the toponyms and geographical names. Sometimes this part also contains more information than just toponyms (e.g. picnic areas).

Tour descriptions of *Fernwanderwege* (long-distance hiking trails) and multi-day hiking trails which were part of the activity or cover the same path as their individual stages in *Wanderung* (hiking) were excluded from the analysis [1]. They usually were very short tour descriptions and covered multiple individual stages. The individual stages, which were also retrieved, contain much longer descriptions and hence include more information on LF and CES. Shorter distances of the stages also reduce the potential area to geolocate annotations. Additionally, if the long-distance hiking trails were included, some LF and CES would have been annotated multiple times, because they were mentioned in the longer descriptions of the individual stage and the shorter description which covers multiple stages.

This reduction dropped the number to 75 tour descriptions. An overview of the number of tour descriptions per activity can be seen in Table 4.1. Figure 4.2 shows a density map of the GPX files of the tour descriptions.

4.1.2. Picture Data

Flickr emerged as the best widely known source of geotagged social media data for the analysis. Due to data access restrictions and other reasons (Section 2.5), I have not taken into account other popular platforms. *Flickr* showed to be more representative for tourist attractions when compared to *Twitter* (Li et al., 2013) and performed better at capturing CES when compared to *Panoramio* in Obersimmental, a region which is comparable to the UBE (Oteros-Rozas et al., 2018).

Flickr is a social network and web photo platform which is used to manage and share pictures. The photographs and their metadata (owner, coordinates, time, date, and tags) can be accessed by using the application programming interface (API) (*www.flickr.com/services/api/*). Even though social media data have successfully been used in single-year analyses (e.g. Tenkanen et al. (2017)), due to the data scarcity, I decided to aggregate *Flickr* data from multiple years, as it was done by Sessions et al. (2016).

[1] Emmenuferweg, Moorlandschaftspfad UNESCO Biosphaere Entlebuch, Höhenweg Entlebuch Emmental, Steinbock - Trek - Brienzer Rothorn, Im Wanderschuh zum Aelplerrendez-vous

Table 4.1.: *The number of tour descriptions per activity and the total number of tour descriptions used for the analysis.*

Activity	**Count**
Hike	45
Themed Trail	13
Mountain Bike	11
E-Bike	3
Trail Run	1
Mountain Tour	1
Bicycle	1
Total	**75**

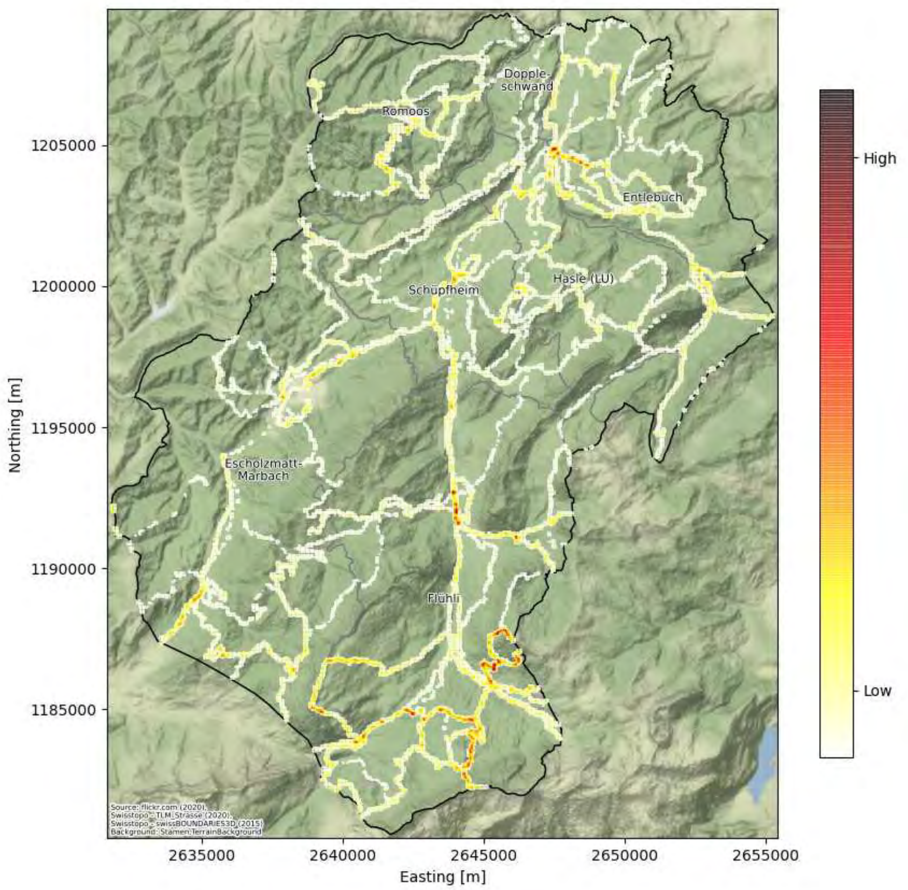


Figure 4.2.: *Density of the GPX points of the tour descriptions used for the analysis in the UBE.*

Figure 4.3.: Image 50211393386 by user 57461617@N07 is used as an example to illustrate the attributes retrieved from Flickr *in Table 4.2.*

Attribute	Example
id	50211393386
owner	57461617@N07
title	BO_08296
datetaken	09.08.2020 09:23
latitude	46.939572
longitude	8.113872
url_c	https://live.staticflickr.com/65535/50211393386_04e4628014_c.jpg

Table 4.2.: Overview of the attributes with the example of the picture in Figure 4.3.

Retrieval of *Flickr* Pictures

I accessed the *Flickr* API through the Python interface *flickrapi*. The *flickr.photos.search* function is used to return the queried data in an XML format, which can be converted to a comma-separated values (CSV) format. The function offers two important arguments, amongst many others. One argument is used to only extract pictures which were geolocated. The other argument allowed me to define a bounding box, which limited the function's spatial operability to the bounding box of the UBE. The *Flickr* API only retrieves pictures that have been declared as *public* by their owners.

The maximum amount of requests and data retrieved by the API is limited to 3'600 per hour. Because the API only returns a random selection of pictures with each access (Brooker et al., 2016), by accessing the Flickr API once, I could only extract a subsample of all data. Consequently, I had to retrieve the pictures multiple times and always add the newly extracted pictures to the existing data set which consisted of previously-stored pictures. This procedure allowed me to include as many pictures as possible in the final data set. Finally, I deleted duplicates that were retrieved by multiple accesses and clipped the point data set to the boundaries of the UBE. A map of all *Flickr* pictures located in the research area can be seen in Figure 4.6a.

I exported the resulting CSV file with the attributes from Table 4.2 where the image in Figure 4.3 is used as an example. This preliminary data set contained information about 2'255 pictures.

Flickr Picture Processing

To extract pictures that were taken along a hiking trail, it was necessary to find a buffer threshold to clip the pictures to the buffer. The hiking trails are part of the TLM_Strasse data set which was obtained from the geodata service of the ETH Zurich (*geovite.ethz.ch*). Wu et al. (2017) used a 50-foot (15.2 meters) buffer and a 200-foot (61 meters) buffer, where the 50-foot buffer did not retrieve enough data for their analysis. Based on this conclusion and the fact that Wu et al. (2017) ran their analysis on trials in the City of Minneapolis, Minnesota, a high-rise area, I decided to decrease the buffer size to 40 meters for the UBE. The distribution of the distances between the location of each picture and its closest hiking trail segment is shown in Figure 4.4 and reveal that the majority (44%) of pictures lie within 40 meters to a hiking trail and contributors which were off the hiking trails or on different trails (e.g. while skiing or snowshoe hiking), were excluded by this procedure.

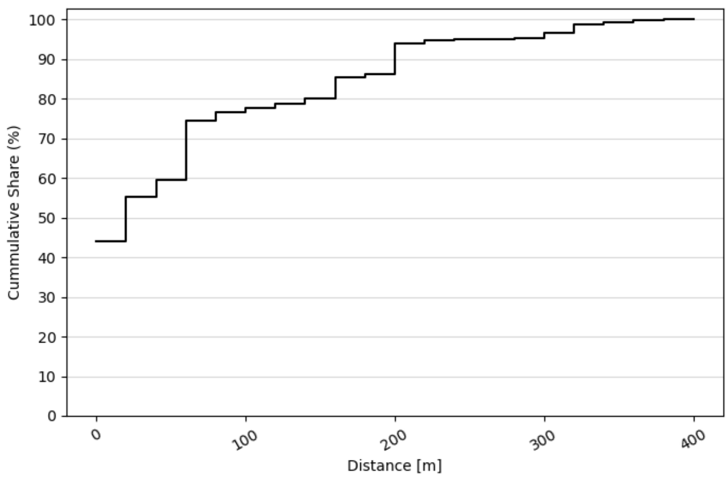


Figure 4.4.: *Histogram of the distance of* Flickr *photographs to its nearest hiking trail segment.*

Since a lot of LF are not visible or identifiable during winter because of snow coverage, I applied a temporal filter. I only included pictures that were taken during the summer season defined by the *Bergbahnen Sörenberg AG*, which starts in May and ends in October.

The final data set used for the annotation procedure and cluster detection contained 729 pictures (Figure 4.6 (b)) from 69 contributors A boxplot of the distribution of the number of pictures among the contributors (without outliers) can be seen in Figure 4.5. The median number of pictures is 2.0. The mean 10.8 (SD: 34.84) is quite large due to far outliers (maximum: 264 pictures).

4.2. Content Analysis

Even though repeatability and extent are limitations for manual content analysis (Heikinheimo et al., 2018), I decided to perform the annotation manually. I did not want to rely on a trained machine learning algorithm to extract LF and CES and time was not a limiting factor. The

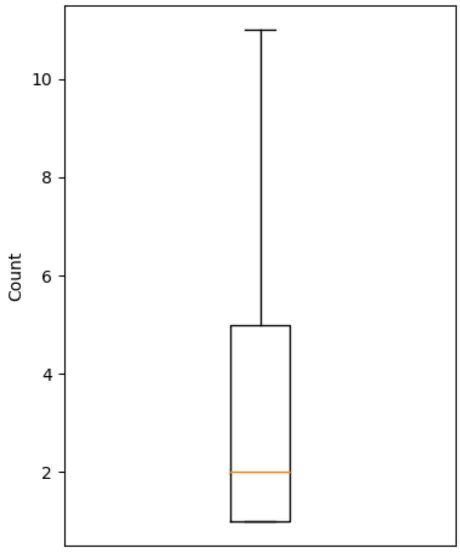


Figure 4.5.: *Number of pictures per contributor without outliers.*

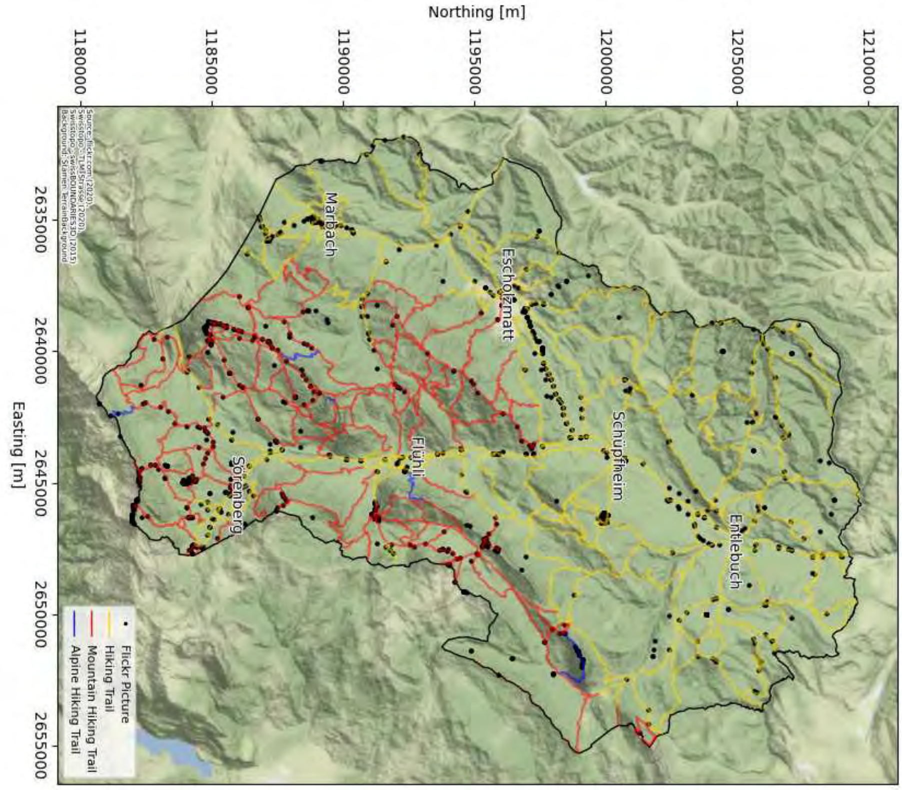
necessary steps were creating lists of LF and CES to be annotated as well as defining rules on how they should be annotated.

4.2.1. Creating the List of LF

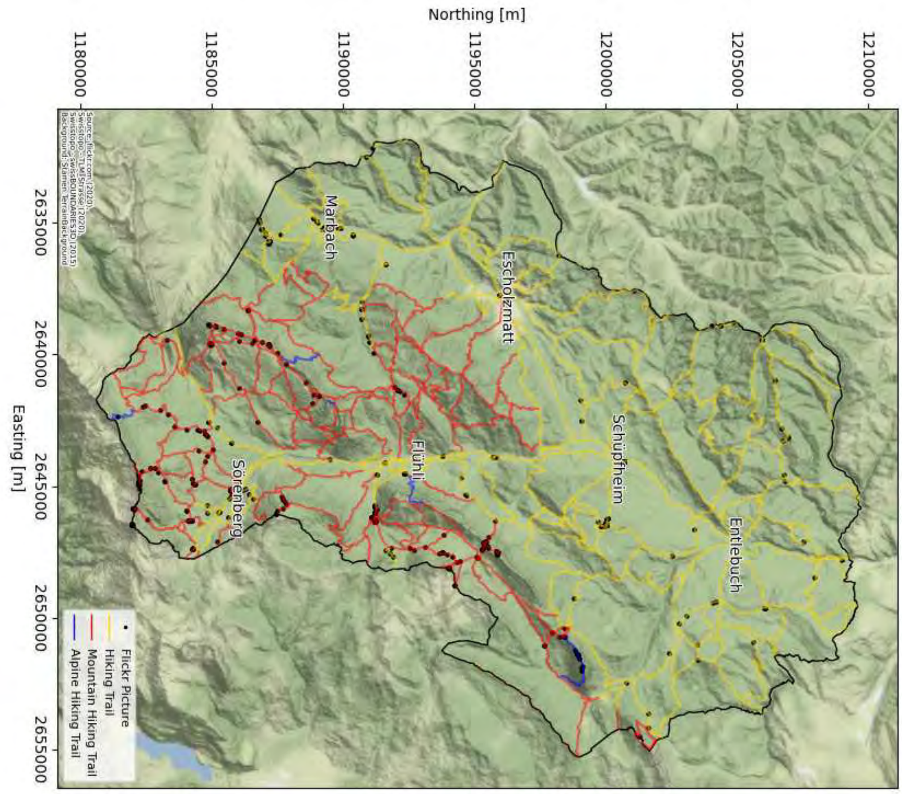
The LF used by Oteros-Rozas et al. (2018) forms the basis for my list of LF. I have adopted many of them, renamed some (e.g. *Wood pasture* to *Forest* and *Mountain* to *Summit*) but also dropped specific LF (e.g. *Coastal/Beach/Dune* or *Sea*) because they would not be appropriate for the properties of the research area. The final list of LF for the annotation process can be seen in Table 4.3.

4.2.2. Creating the List of CES

Oteros-Rozas et al.'s (2018) types of CES were only slightly adjusted by grouping *Recreation / Rec_Terrest / Rec_Aquatic* to *Recreational Facilities*. Additionally, Callau et al. (2019) was particularly helpful to define subtypes of CES (e.g. *Church*, *Viewpoint*, *Signpost* or *Dawn / Sunset*). Moreover, some subtypes cannot be detected in picture data, but only in text data. *Healing Powers* and *Place Attachment* were added to the list of subtypes after having read a couple of tour descriptions. It was important to me to include as many aspects of CES (spiritual, religious, recreational, and educational) as possible since they tend to be left out due to the importance of tourism in local studies (MEA, 2005). A list of the CES and the corresponding subtypes used for the annotation process as well as and their description can be found in Table 4.4.



(a)



(b)

Figure 4.6.: *These maps show the locations of Flickr photographs in the UBE before (a) and after the reduction (b) of the data set.*

Table 4.3.: *Overview of the different types and subtypes of LF.*

Type	Subtype
Natural Landscape	Bedrock
	Flower / Funghi
	Forest
	Grass- and Moorland
	Lake
	River / Creek
	Rock
	Shrub
	Snow / Ice
	Summit
	Tree
	Waterfall
	Wild Animal
Human influenced Landscape	Agriculture
	Infrastructure
	Livestock
	Path / Trail
	Urban

Table 4.4.: *Overview of the different types and subtypes of CES.*

Type	Subtype	Description and Examples
Cultural Heritage	• Identity • Information Board • Information Office • Local History • Tradition • Traditional Architecture	These are sites which represent local history and culture and form a sense of identity. Examples in the UBE are: Farm shops, local cuisine, flags, information boards, traditional houses. Text data which describes the sense of people identifying themselves with e.g. the *biggest moorland in Switzerland* or the *highest peak in the canton of Lucerne* also belong to this type.
Recreational	• Recreational Facilities • Signpost • Viewpoint	These are sites which serve for recreational activities, including signposts providing directions. Examples in the UBE are: Benches, playgrounds, kneipp facilities and places providing a good view.
Social	• Camping • People • Restaurant / Accommodation	These are sites where people are meeting up or can be used as meeting points. Examples in the UBE are: Restaurants, hotels, camping grounds and whenever people are a significant part of the picture
Spiritual	• Dawn / Sunset • Healing Powers (text only) • Place Attachment (text only) • Church • Summit Cross	These are sites which represent a place of increased human attachment. They are also used for religious and spiritual practices, which in some cultures take place in nature. Examples in the UBE are: Summit crosses, churches or chapels and sunrises or sunsets.

4.2.3. Rules for the annotation Process

After I defined all possible LF and CES as explained in Section 4.2.1 and Section 4.2.2, I could set up the rules and procedure of the annotation process which are described in the following section. Each part includes general rules and a more in-depth specification of the annotation of LF and CES for each data type.

Annotating Picture Data

Initially, I had to check all pictures for their validity to decide whether they should be included in the annotation process. The following signs excluded a picture from the annotation process:

- The content of the picture indicates that it was not taken along a hiking trail (e.g. from a plane or indoors).
- The picture was not taken during the summer months (e.g. lots of snow).
- The picture is part of an extensive collection of the same event.
- The content of the picture cannot be annotated (e.g. moon, meteorological phenomena (clouds, fog)) or if the contrast of black and white pictures makes it impossible to recognize LF and CES.

Figure 4.7a shows the annotations in a picture with an example presented in Figure 4.7b. The details of the annotation process are outlined in the following paragraphs.

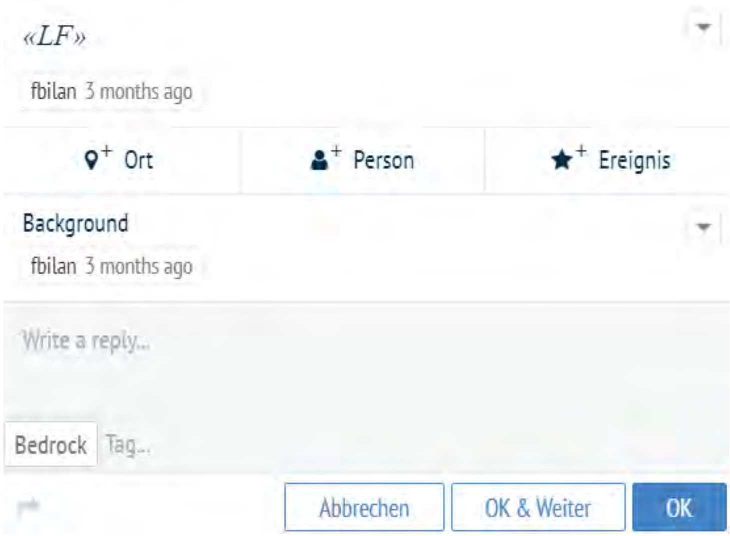
For picture data, I only annotated *subtypes* of LF and CES (see Table 4.3 and Table 4.4). It is possible that the same object is annotated as a LF and a CES (e.g. I annotated a usable bench as *Recreational Facilities* and as *Infrastructure* or a restaurant as *Restaurant / Accommodation* and as *Infrastructure*).

Because I wanted to investigate the relationship between LF and CES, the rules from Oteros-Rozas et al. (2018) formed the basis for my annotation process. However, I also introduced exceptions and adjustments: The main criteria for the annotation of a LF was the presence/absence and coverage of at least 5% of the photograph. This rule is identical to what Oteros-Rozas et al. (2018) used to analyze their *Flickr* data. Multiple parts of LF (e.g. forests) can also be counted as one LF to reach the required coverage.

I decided to add an exception to this coverage rule in favor of subtypes that inherently cover less area in a picture. These subtypes are spatially limited due to their definition. I considered the following subtypes to be affected by this exception:

- River / Creek
- Path / Trail
- Summit
- Tree
- Wild Animal

(a)



(b)

Figure 4.7.: *Illustration of the annotation process in* Recogito. *(a) Image 49969960801 by user 74710161@N02 with its seven annotations (b) example of an annotation with classification (LF) in the top line, the differentiation between fore- and background in the middle section and the corresponding subtype tag (*Bedrock*) at the bottom.*

They still have to form an important part of the picture, should only be annotated if they are located in a central position of the picture, and should be among the reasons the picture was taken. Figure 4.8 shows an example where I annotated *Path / Trail* even though its coverage is below the threshold.

Figure 4.8.: *Image 48228381272 by user 134312632@N06 is an example for the exception where* Path/Trail *was annotated because the network of paths and trails on the plain forms an important part of the picture even though it does not reach the required coverage.*

The annotation process for LF additionally differentiates between foreground and background. For each LF annotation, I added *Foreground* or *Background* as additional information. Since Oteros-Rozas et al. (2018) only looked at the presence or absence and coverage of LF, I expected to gain further insight into the relationship between LF and CES by making this distinction. This means that each LF was annotated at most twice per picture. The differentiation is not based on spatial distance, but recognisability (Edwardes et al., 2007). I used the following definitions of *Foreground* and *Background* (illustration in Figure 4.9):

- Foreground: Details of objects are discernible (e.g. leaves and tree trunks, textures, and treetops can be recognized).
- Background: Details cannot be seen and textures are uniform.

For the annotation of CES, again, only *subtypes* were used. The main criteria for their annotation was solely presence/absence. No specific percentage of coverage was required. But, similar to the exception made for specific LF, these subtypes still have to form an important part of the picture and should be among the reasons why the picture was taken.

A picture may contain multiple CES (e.g. restaurants) but only one per subtype could be annotated. Additionally, no differentiation between *Foreground* and *Background* was necessary for the annotation of CES since the location of the CES in the picture is not relevant.

Annotating Text Data

Because I filtered out the tour descriptions already in Section 4.1.1, no validity check was required and I could use all 75 tour descriptions for the annotation procedure.

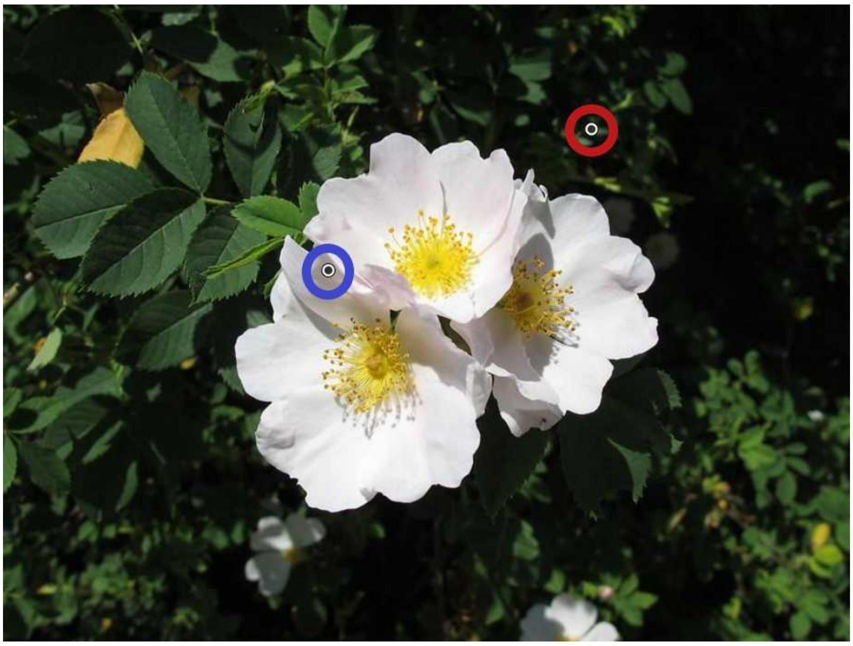


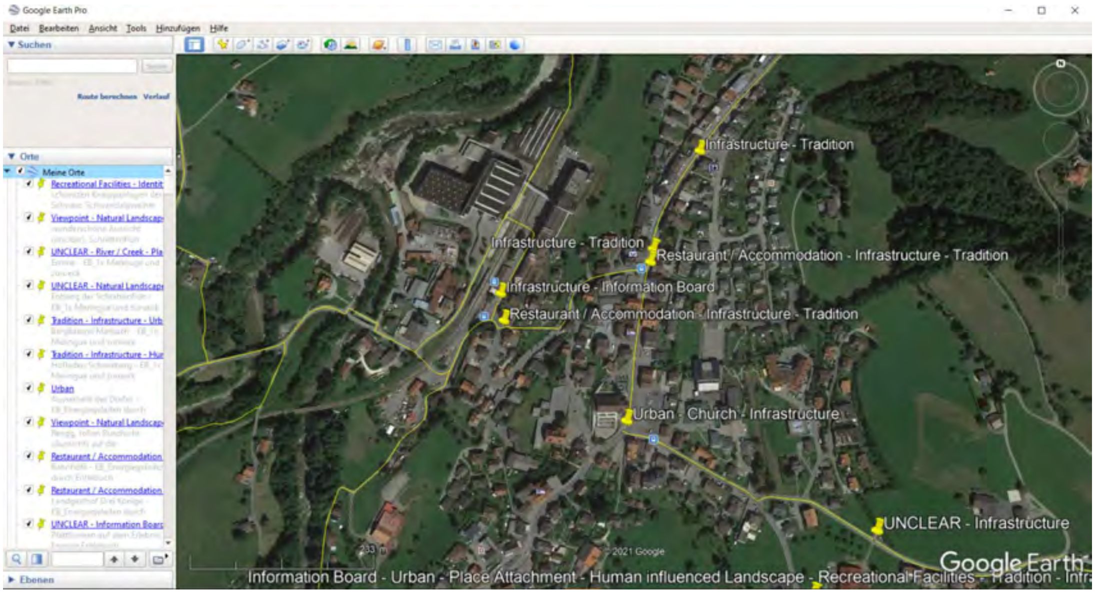
Figure 4.9.: *The two circles in image 35898728460 by user 150752905@N08 highlight the two annotations made in this picture. While the red circle represents an annotation (*Shrub*) in the background, the blue circle represents an annotation in the foreground (*Flower / Funghi*).*

For the tour descriptions, I again used *subtypes* of LF and CES for the annotation process. For specific expressions, the *types* of LF (*Natural Landscape/Human influenced Landscape*) were annotated. This was the case if an expression did not match single but multiple subtypes of a type. By adding the two LF (*types*) to the list of possible annotations, I could annotate more expressions where the annotation of a specific *subtype* was not possible. It was more important to me to include an unspecific expression than ignore it.

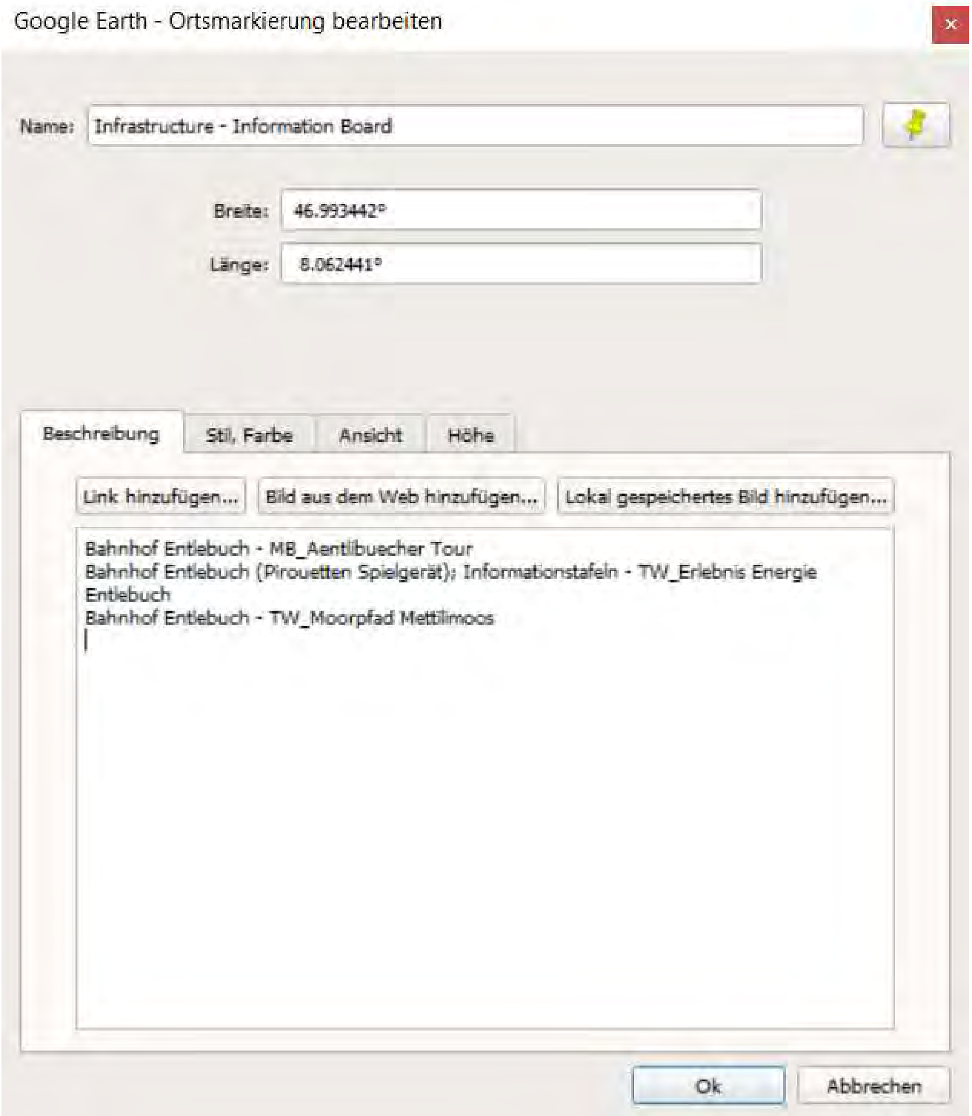
The specialty when working with text data is that words linking text data to their location have to be extracted and pinpointed with GPS coordinates (Derungs and Purves, 2014). I had to assign coordinates to each annotation to be able to use the data for further analysis. Per location, I could only annotate one subtype per tour description. If it was not possible to exactly geolocate an annotation, I could add the tag *unclear* together with coordinates of any part of the tour. Otherwise, the tag *clear* was added. I ignored all annotations outside the boundary of the UBE.

For both, LF and CES, the only criteria for the annotation was whether they were mentioned in the text. It was necessary to match the expressions from the tour descriptions with the available types and subtypes (e.g. "Zaun" (fence) as *Infrastructure* or "das alte Handwerk der Köhlerei" (the old craft of charcoal burning) as *Local History*). For reference, I listed all expressions with their annotated type(s) or subtype(s) in Table 4.5 and Table 4.6. By doing so, I made sure to always assign the same labels. This does not imply that an expression could not be annotated to multiple subtypes (e.g. Speichersee (reservoir lake) was annotated as *Infrastructure* and *Lake*).

Generally, I did not include toponyms unless they represented a LF. For some toponyms, a little research was necessary to identify the toponyms as a village name (e.g. Finsterwald) or LF (e.g. Bühlwald). Since the search function on *map.geo.admin.ch* is not only able to search for town names but also field names, it proofed to be ideal for this task.



(a)



(b)

Figure 4.10.: *Illustration of how the annotation process takes place in Google Earth Pro. (a) municipality of Entlebuch with its annotations. (b) example where multiple tour descriptions mention the same subtypes.*

For geolocating the annotations, search tools from various websites and data were taken into account (e.g. *map.geo.admin.ch*, *SchweizMobil*, *Google Maps*, local websites (e.g. *hiking maps from Zyberliland*) or additional information from the tour page (profile of hiking trail, GPX file)). To prevent geolocating the same annotation of the same object from two different tour descriptions to two different pairs of coordinates, I used *Google Earth Pro* and added each annotation to the map with all previously geolocated annotations. Figure 4.10a illustrates how *Google Earth Pro* looked like with the annotations and Figure 4.10b shows multiple annotations from different tour descriptions which were assigned to the same coordinates. *Google Earth* was used in Cooper and Gregory (2011) to store similar information. If two annotations from the same type or subtype were mentioned and they were located at the same place (e.g. steigen Sie auf der linken Seite mit einer kleinen Leiter über den Zaun (on the left side, climb the ladder and the fence)), only either the ladder or the fence was geolocated and annotated as *Infrastructure*.

4.2.4. Recogito Pelagios

Recogito is a platform to be used as an annotation tool for picture and text data (Pelagios Network, 2021). It includes a download function that returns a well-structured CSV file, which I used for the next steps of the analysis.

Table 4.5 and Table 4.6 show all different expressions from the tour descriptions that were annotated to the types and subtypes of LF and CES.

Table 4.5.: *Overview of the different annotated expressions for LF.*

Type	Expression
Natural Landscape	Alpenkranz, Berg, Berglandschaft, Bergwelt, Brienzer-Rothorn-Kette, Flora, Harder- und Brienzergrat, Hohgant, Hügellandschaft, Hügelwelt, Natur, Pflanze, Pflanzenwelt, Schrattenfluh, urtümlich, Wasserlandschaft, wild
Human influenced Landscape	Alp Brüedere, Alp Imbrig, Alp Schlund, Alpen, alpiger, Alpweid, Alpwiesen, Emmental, Entelbuch, Kräutergarten, Lobenalp, Mittelland, Sonnengarten, Streuesiedlungen, Talboden, Torfstich, Waldemmental, Ziegenalp
Subtype	**Expression**
Bedrock	Bergsturzgebiet, Fels, felsdurchzogen, Felsformation, felsig, Felsvorsprung, Flyschgebiet, Karrenfeld, Karstgebirge, Karstgestein, Karstlandschaft, Nagelfluhflanke, Schrattenfels
Flower / Funghi	blühen, Blumen, blumenreich, blumig, Flockenblumen, Orchideen, Rapunzel, Sonnentau, Waldstorchenschnabel, Wildblume
Forest	Bäumen, Bergföhren, Bergföhrenhochmoorwald, bewaldet, Bockwald, Brameggwald, Bühlwald, Chilewald, Farnwald, Fichtenwald, Föhren und Fichten, Föhrenwäldchen, Forststrasse, Heidelbeerwald, Hochmoorwald, Lichtung, Schachnerwald, Staldigwald, Wald, Waldarena, Waldpartien, Waldpfad, Waldrand, Waldstrasse, Waldstück, Waldweg
Grass- and Moorland	Alpweide, Alpwiese, Bergföhrenhochmoorwald, Flachmoor, grasbewachsen, Graslandschaft, Hochmoor, Hochmoorwald, Lichtung, Moor, Moorboden, Moorgebiet, moorig, Moorlandschaft, Moorpfad, Naturwiese, Schafwiese, Streuewiese, sumpfig, Torfstich, Weide, Weidegebiet, Weidepartie, Wiese, Wiesenhang, Wiesenlandschaft, Wiesenpfad
Lake	Änggenlauenenseeli, Brienzersee, Gewässerbiotop, Schwandalpweiher, Seelein, Speichersee, Teich, Tümpel, Vierwaldstättersee, Weiher
River / Creek	Bach, Bärselbach, Bergbach, Chuterenbach, Eibach, (kleine/Wiss/Wald-) Emme, (kleine / grosse) Entle, Fluss, Flusslandschaft, Goldbach, Ilfis, Kanal, kleine Fontanne, Quelle, Rotbach, Schmelzwasserrinne, Schonbach, Schwefelquelle, Seebenbach, Seelibach, Steiglenbach, Torbach, Wasser, Wasserlandschaft
Rock	Bergsturzgebiet, Geröllhalde
Shrub	farnreiche, Farnwald, Heidelbeersträucher, Heidelbeerwald
Snow / Ice	Schnee, schneebedeckt, weiss gezuckert
Summit	Brienzer Rothorn, Eiger, Finsteraarhorn, First, Fürstein, Gipfel, höchster Punkt des Kantons Luzern, Jungfrau, Mönch, Pilatus, Schibegütsch, Schimbrig, Schreckhorn
Tree	Baum, Bergföhre, Kraftbaum, Lebensbaum, Stamm- und Wurzelreste, Wurzeln
Waterfall	Wasserfall
Wild Animal	Eidechse, Hirsch, Insekt, Libelle, Mooreidechse, Murmeli, Schmetterling, Steinadler, Steinbock, Tier, Vogel
Agriculture	Alp Brüedere, Alp Imbrig, Alp Schlund, Alpen, alpiger, Alpweid, Alpwiesen, Emmental, Entelbuch, Kräutergarten, Lobenalp, Mittelland, Sonnengarten, Streuesiedlungen, Talboden, Torfstich, Waldemmental, Ziegenalp
Infrastructure	Alpbeizli, Alphütte, Alpkäserei, Alpkiosk, Alprestaurant, Alpweidstrasse, Alpwirtschaft, asphaltierte Strasse, Asphaltstrasse, Aussichtsplattform, Bäckerei, Bahnhof, Bahnhöfli, Bahnhofplatz, Bahnübergang, Bahnweg, Bank, Bänkli, Bauernbetrieb, Bauernhaus, Bauernhof, Beizli, Berggasthaus, Bergkäserei, Bergkirche, Bergstation, Biosphärenshop, Birkenhof, Bretzeli-Bahn, Brücke, Brunnen, Burgstelle, Bushaltestelle, Dach, Dorfladen, Erdgasgewinnungsanlage, Erlebnispark, Erlebnis-Restaurant, Feuerstelle, Finishütte, Finsterwaldstrasse, Frutteggstrasse, Fussballfeld, Fussgängerbrücke, Gebäude, geteert, Gipfelrestaurant, Glashütte, Gondelbahn, Grenzmauer, Grenzstein, Grillplatz, Grillstelle, Hängebrücke, Hängesessel, Hauptstrasse, Haus, Hof, Hofarni, Hofladen, Holzbeigen, Holziglus, Hotel, Hürndlihütte, Kanal, Kanalweg, Kappelle, Kirche, Kleinkraftwerk, Kneippanlage, Kohlenmeiler, Köhlerplatz, Kurhaus, Kurve, Laden Würzig, Landgasthof, Landi, Leiter, Lourdes Grotte, Luftseilbahn, Lustenbergstrasse, Moortretbecken, Natursteinmauer, Picknickplatz, Picknickstelle, Pilgerbeizli, Postauto-Haltestelle, Quartierstrasse, Radweg, Rastplatz, Restaurant, Ruheliege, SAC Hütte, Schafmilchbetrieb, Schärhof, Scheune, Schwandalpstrasse, Seelensteg, Sitzgelegenheit, Sonnenterrasse, Speichersee, spielerische Elemente, Spielplatz, Sprung, Stalltüre, Station Sörenberg, Staumauer, Steg, Steinbock-Schaukel, Steinmännchen, Stillaubbrücke, Stolehüttli, Strasse, Talstation, Tische, Torbachbrücke, Torfhüttchen, Tram, Trekkinghof, Trockensteinmauer, Übernachtung, Unterkunft, Wasserbüffelhof, Wasser-Erlebnisse, Wasserspielplatz, Wasserspielwelt, WC, Weidezaun, Welle, Windrad, Wirtschaftsweg, Zaun
Livestock	Kamel, Lama, Schafwiese, Wasserbüffelhof, Ziegenalp
Path / Trail	Abenteuerpfad, Alpstrasse, Bahnweg, Bergwanderweg, Emmenuferweg, Eremitenweg, Flowtrail, Forststrasse, Fruttegweg, Fussweg, Gesundheitspfad, Glasereipfad, Grat, Gratabstieg, Gratrücken, Gratwanderung, Höhenweg, Kanalweg, Karrweg, Kiesweg, Köhlerweg, Krete, Kulturweg, Märchenweg, Moorlandschaftspfad, Moorpfad, Natursträsschen, Naturstrasse, Pfad, Säumerweg, Schlierengrat, Schotterstrasse, Schutzwaldpfad, Singeltrail, Sonnentauweg, Spazierweg, Strasse, Strecke, Themenweg, Trimlegrat, Waldpfad, Waldstrasse, Waldweg, Wanderweg, Weg, Wiesenpfad
Urban	Dorf, Dorfkern, Dorfzentrum, Ferienhäuschen, Ferienhaussiedlung, Gemeinde, Ort, Ortskern, Quartier, Quartierstrasse

Table 4.6.: *Overview of the different annotated expressions for CES.*

Type	Subtype	Expression
Cultural Heritage	Identity	älteste im Kanton Luzern, eindrücklichste, einzige in der Schweiz, erste der Schweiz, geschützte, grösste, höchsten Punkt des Kantons Luzern, längste, nationaler Bedeutung, schönste, Wahrzeichen
	Information Board	Energieplattform, Erlebnisstation, Informationen zu, Informationstafel, Infotafel, Plattform, Posten, Rutschmodell, Schautafel, Station, Tafel
	Information Office	Tourismusbüro
	Local History	1903 gebaut, altes Handwerk, älteste, Dorfgeschichte, ehemalige Burgstelle, historisch, keltisch, Kohlenmeiler, Köhlerplatz, Sage, sagenreich, sagenumworben, uraltes Handwerk, ursprünglich, während den Weltkriegen
	Tradition	Alpkäse, Alpkäserei, Alpkiosk, Älplerzmorge, Bergkäserei, einheimische Produkte, Entlebucher Bier, Hofladen, Kemmeriboden-Meringues, köstliche Spezialitäten, lokale Speisen, regionale Produkte, Schafmilchbetrieb
	Traditional Architecture	alte Fussgängerbrücke aus Holz, bäuerliche Architektur, typischen Entlebucher Bauernhäusern mit ihren ausladenden, weithinunterreichenden Dächern, ursprüngliche Alphütten mit tiefhängenden Schindeldächern
Recreational	Recreational Facilities	Abenteuerpfad, Bänkli, Emmenuferweg, Eremitenweg, Erlebnispark, Feuerstelle, Flowtrail, Fussballfeld, Gesundheitspfad, Glasereipfad, Grillplatz, Grillstelle, Hängesessel, Kneippanlage, Köhlerweg, Kräutergarten, Kulturweg, Märchenweg, Moorlandschaftspfad, Moorpfad, Moortretbecken, Picknickplatz, Picknickstelle, Rastplatz, Schutzwaldpfad, Singeltrail, Sitzgelegenheit, Sonnengarten, Sonnentauweg, Spielplatz, Themenweg, Trekkinghof, Wasserbüffelhof, Wasser-Erlebnis, Wasserspielplatz, Wasserspielwelt
	Signpost	beschildert, Beschilderung, markiert, Markierung, Schild, Schildere, signalisiert, Signet, Wegweiser
	Viewpoint	Ausblicken, Aussicht, Aussichtsberg, Aussichtsplattform, aussichtsreich, Blick, Fernsicht, in die Höhe, Panorama, Rundschau, Rundsicht, Sicht, Tiefblick, Weitsicht
Social	Camping	-
	People	Familie, Kind
	Restaurant / Accommodation	Alpbeizli, Alprestaurant, Alpwirtschaft, Bahnhöfli, Beizli, Berggasthaus, Bergrestaurant, Erlebnis-Restaurant, Gasthof, Gipfelrestaurant, Hotel, Landgasthof, Nachtessen, Pilgerbeizli, SAC Hütte, Übernachtung, Unterkunft, Znüni
Spiritual	Dawn / Sunset	Sonnenaufgang, Sonnenaufgangsfahrt, Sonnenuntergang
	Healing Powers	die Beine stärken, eine pure Wohltat, gestärkt für den Alltag, gesund, Gesundheitspfad, heilend, Heilkräfte, heilsam, Kraft des Wassers, Kraftbaum, Kraftort, Kurhaus, müde Beine entspannen, neugeboren, Seelensteg, wohltuende Wirkung
	Place Attachment	atemberaubend, beliebt, besondere Schönheit, eindrücklich(-ste), einmalig, einzigartig, entzückend, fantastisch, farbenfroh, faszinierend, grandios, grosses Wow, herrlich, idyllisch, imposant, in eine andere Welt versetzt, Kleinod, landschaftlich reizvoll, naturnah, Naturstimmung, oft erwähnt, romantische, Schönheit, schönste, schönsten, schützenwert, spektakulär, unberührt, verzaubern, wildromantisch, wunderbar, wunderschön, wundervoll, zauberhaft
	Church	Alpkapelle, Bergkirche, Eggkapelle, Kapelle, Kirche, Lourdes Grotte, Wallfahrtskirche
	Summit Cross	Beichlenkreuz, Gipfelkreuz

4.3. Inter-Annotator Reliability

The inter-annotator reliability gives an idea about how consistently the set rules have been applied and how much the results differ between two annotators. Errors and ambiguities could occur due to human inaccuracies. However, these should be reduced by setting consistent rules which can easily be applied.

A fellow geography student applied the rules of the annotation process explained in Section 4.2.3 on a sample of 50 pictures and 5 tour descriptions. Since the Cohen's Kappa Index turned out negative I decided to use another measure of similarity. It turned out negative because if I assumed my annotations always as *Actual* or *True* and compare them to the second annotator's annotations *Predicted*, there are no annotations which were *False* by both of us. This means that the overall accuracy (p_0) always turned out smaller than the probability of random agreement (p_e).

I decided to use the Jaccard Index (Equation 4.1) instead, which is a similarity coefficient and takes the ratio between the size of the intersection and the union of two data sets. It returns a value between 0 and 1 whereas a value of 1 represents complete overlap and 0 means no overlap of the two data sets (Jaccard, 1912). The Jaccard-Index can also be used to calculate the similarity between two raster data sets which contain binary values as in Heikinheimo et al. (2020).

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} \tag{4.1}$$

I calculated the Jaccard Index for each picture and each tour description. The Jaccard Indexes were calculated twice for the picture annotations, once including the *Fore-/Background* argument and once without the argument. An example of different annotations of a picture is shown in Table 4.7 and calculated in Equation 4.2. This example does not take into account the *Fore-/Background* argument.

Table 4.7.: *Annotations by a fellow geography student (Annotator A) and me (Annotator B). The overlapping annotations are underlined.*

Annotator A	**Annotator B**
<u>Bedrock</u>	<u>Bedrock</u>
<u>River / Creek</u>	<u>River / Creek</u>
<u>Rock</u>	<u>Rock</u>
Infrastructure	Shrub
Flower / Funghi	Path / Trail

$$J(AnnotatorA, AnnotatorB) = \frac{3}{7} \sim 0.43 \tag{4.2}$$

4.4. Detection of CES

Two different methods are used to map CES. Both approaches are applied to each data source as well as the combination of them. Since the text data also contained annotations that were not along the official hiking trails and the study's focus lies in the perception along hiking trails, I only used annotations that were within 40 meters of a hiking trail, similar to the filter applied to the *Flickr* pictures in Section 4.1.2.

4.4.1. Manual Detection

The manual annotations of *Flickr* pictures and tour descriptions were used for the *manual* detection of CES. Because the pictures already contained coordinates and I manually added coordinates to the tour description annotations, it was possible to map their spatial distribution. This was done by creating a heatmap using kernel density estimation (KDE).

I only used CES annotations which I was able to geolocate precisely in the text data and contained the attribute *clear*. Previous studies have only looked at whether or not a CES is pictured in the photograph. This is not possible for the text data, since the tours span over various distances and cannot be grounded to one point, like the pictures. But to be able to compare the two data sources, I decided to use all CES annotations in the photographs instead of only the pictures which contained a CES. By doing that, I was able to retrieve and map all CES which were annotated in both data sets. The combined data set is made up of the two subsets of the two data sources.

4.4.2. Automatic Detection

For the hotspot analysis, a location-based approach of identifying hotspots of interactions between humans and nature was used (Toivonen et al., 2019; Levin et al., 2017; Tenkanen et al., 2017). To extract locations that have been mentioned (text) or photographed (pictures) by a high number of contributors, I set up a raster which was filled by the number of contributors.

The grid size was chosen according to the estimated visibility in the pictures and tour descriptions. While a grid size of 4km^2 would often include pictures on the other side of a mountain range or a topographical feature, a grid size of 1km^2 represents a good trade-off between the visible distance in pictures and consequently includes the annotated LF and a reasonably grid size to identify significant hotspots. Gliozzo et al. (2016) and Casalegno et al. (2013) also used a cell size of 1km^2 but did not justify their choice.

The Local Moran's I, a local spatial autocorrelation statistic, was used to extract hotspots from the grid with the number of contributors. It detects statistically significant clusters of high and low values but also outliers (Anselin, 1995). I only extracted the hotspots, which were the cells of high numbers of contributors surrounded by other cells of high values of contributors. The outliers and coldspots were not extracted.

4.5. Relations between CES and LF

Oteros-Rozas et al. (2018) looked for relationships between CES and LF by analyzing the LF which appeared in a picture containing a CES. This methodology can be applied to pictures, but not to text data, because the tour descriptions cover different spatial scales and do not capture a moment like pictures. Consequently, looking for LF and CES which appear in the same tour description would be misleading because text data may contain the same or different LF multiple times even though only one is spatially close to a CES and may contribute to its appearance. Therefore, I decided to use all annotations and capture all LF which are within a predefined distance of CES to be able to adequately compare the results of both data sources and their manually annotated CES. While a threshold distance that is too low might not include all relevant LF for the text data, a threshold distance that is too large would include irrelevant LF. The threshold distance also plays an important role in the picture data. By applying the same threshold distance on all annotations of the picture data set, I was able to capture multiple pictures from a location. Because no other study has investigated the relation between CES and LF in text and picture data, I decided to use 20 meters as a threshold distance since this seemed like a valid distance to capture pictures and text annotations from the *same* location. Most of the LF annotations in text data (83%) were within this distance to a CES annotated in the tour descriptions (Figure 4.11). Additionally, 20 meters was a reasonable threshold distance because only the clear annotations were used

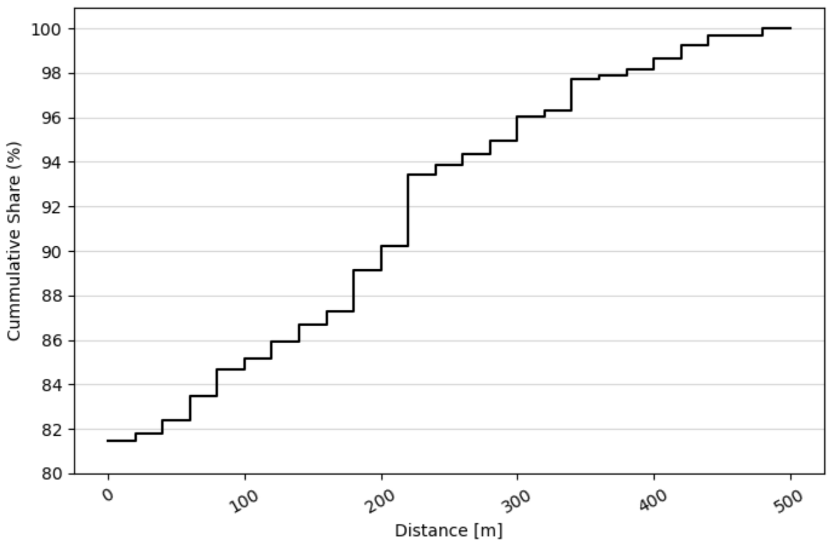


Figure 4.11.: *Cummulative Histogram showing the share of LF based on the distance between each LF and its closest CES annotated in the text data.*

The threshold distance was only applied for the manual detection. For the automatic detection CES, the spatial extents were already defined by the resulting hotspot-polygons. In this case, only the LF within the defined clusters were extracted and analyzed as well as their share among all annotations of the corresponding LF subtype.

Results

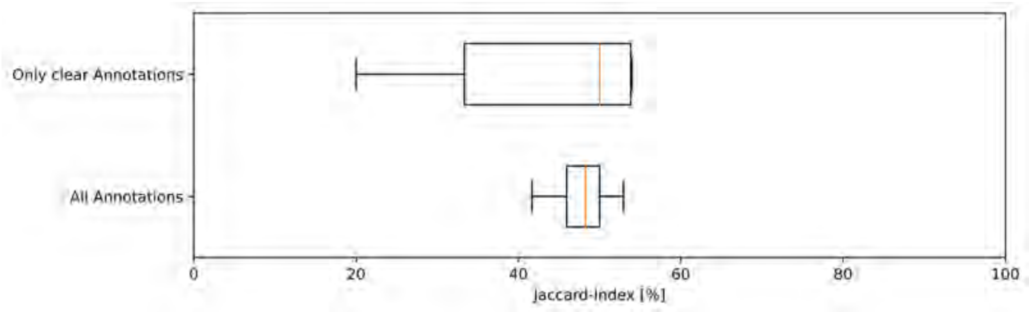
This chapter presents the results of both methods which have been used to detect and map CES as well as their Relations with specific LF. The first section deals with the inter-annotator analysis where I show the results between the second annotator's annotations and my annotations. I also outline what information can be gathered from the results and how they influenced the subsequent sections. Furthermore, I explain the results of the combination of the two data sources and then draw inferences from the individual data sets.

5.1. Inter-Annotator Analysis

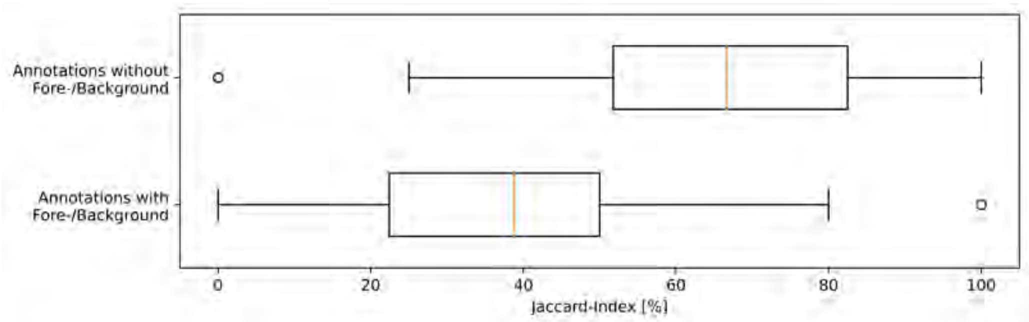
The Jaccard-Index, which calculates the similarity between two annotations, was calculated for each of the 5 annotated tour descriptions and each of the 50 annotated pictures. For the text annotations (Figure 5.1(a)), I differentiated between clear annotations and all annotations. The mean values for the tour descriptions are very close (42.24% for clear annotations and 47.77% for all annotations), meaning that a bit less than half of the annotations overlapped. The standard deviation for the clear annotations (15.06%) is much bigger compared to all annotations (4.26%).

The mean Jaccard-Index for annotations on pictures which did not differentiate between fore- and background is higher (67.04%) than the mean Jaccard-Index for annotations without the fore- and background differentiation (38.36%), while the standard deviations did not differ much (22.70% and 25.05%) (Figure 5.1(b)), especially when compared to the text annotations. But the standard deviations are quite large. The fact that the difference between both mean values and their median values (66.67% and 38.75%) were very small, indicates that the mean values are robust.

The results of this Inter-Annotator analysis influenced the further analysis. Because the Jaccard-Indexes which ignored the *Fore-/Background* attribute performed well, I decided to leave out



(a)



(b)

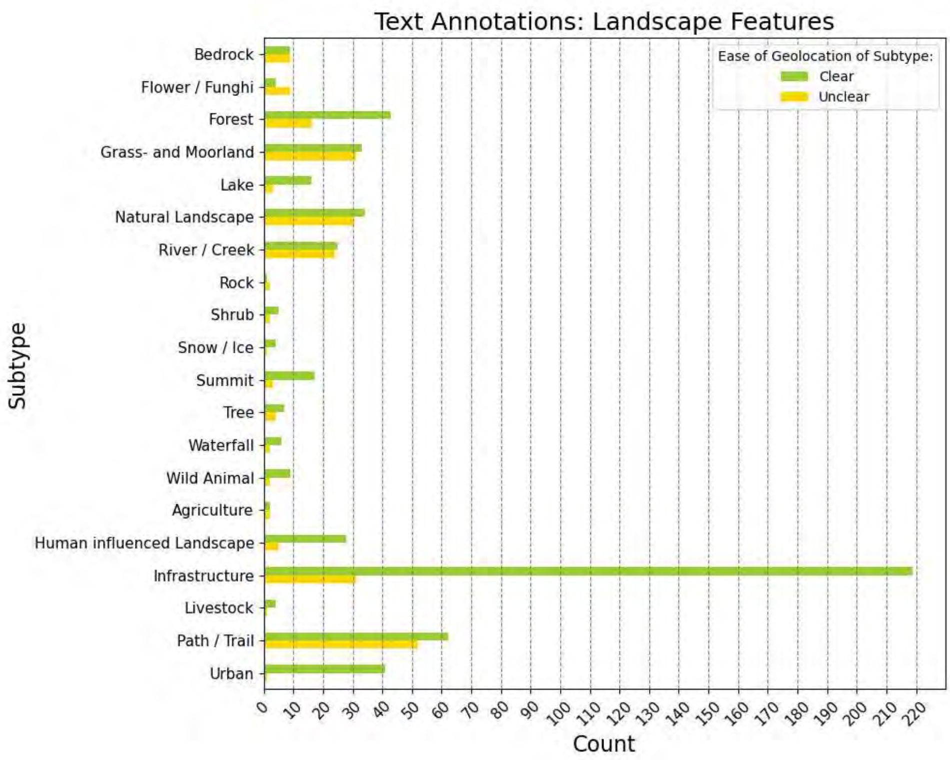
Figure 5.1.: *Jaccard-Index of the annotations for (a) tour descriptions and (b) pictures.*

this attribute when looking for relations. Additionally, this differentiation was not possible for the text data and consequently, a comparison between the two data sets would be difficult. In addition, for the text data, I only used the clearly annotated LF and CES, because the unclear annotations would distort the results through their random locations when automatically detecting CES.

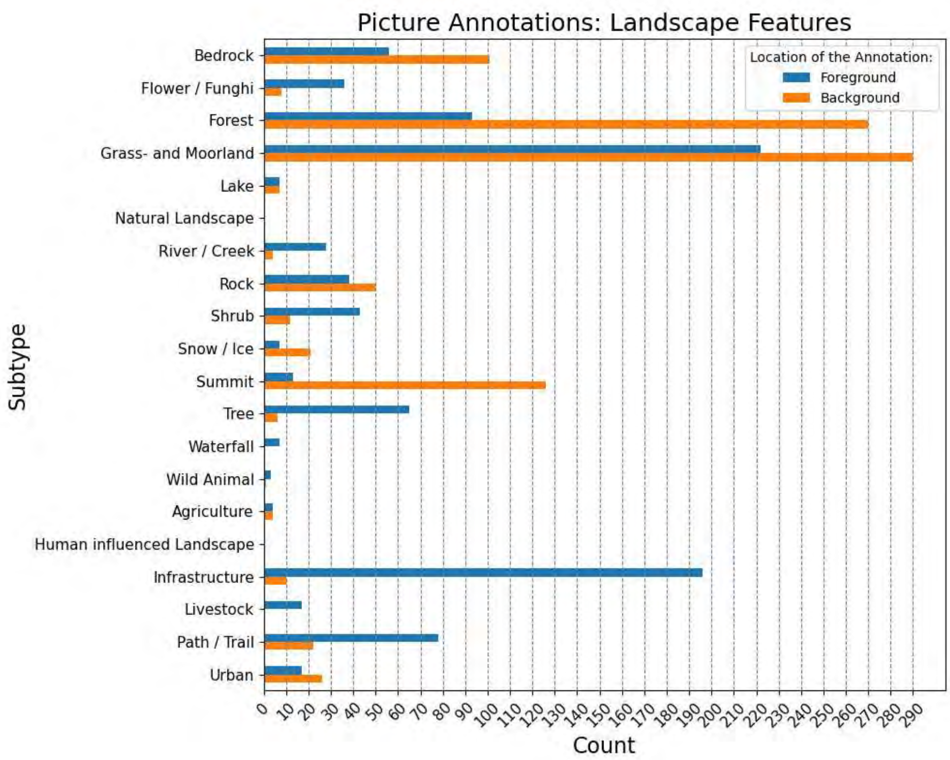
5.2. Annotations

Even though 729 pictures were originally used for the annotation process, only 495 contained at least one annotation. This reduction is because pictures did not meet the requirements (see Section 4.2.3) and mainly due to a horse-riding event. One contributor uploaded 231 pictures of this event. Because the annotations of these pictures would distort the results, I decided to exclude these pictures except one from the annotation process.

Only *clear* text annotations were used for the extraction of CES, the following plots show the bare results of the annotation procedures including unclear labels. The same applies for the *Fore-/Background* differentiation.



(a)



(b)

Figure 5.2.: *Number of LF annotations for (a) tour descriptions and (b) pictures. The LF are ordered alphabetically within their types of LF (Table 4.3).*

5.2.1. Landscape Features

In the text data (Figure 5.2 (a)) 799 LF were annotated. 569 of them were clearly annotated (71%). *Infrastructure* with almost 250 annotations exceeds the second most annotated subtype (*Path / Trail*) by almost 140 annotations and shows a very high proportion of *clear* geolocations (88%). *Forest*, *Grass- and Moorland*, *Urban* as well as *Natural Landscape* showed more than 30 annotations, while *Urban* and *Forest* showed a high share of *clear* annotations (98% and 73%).

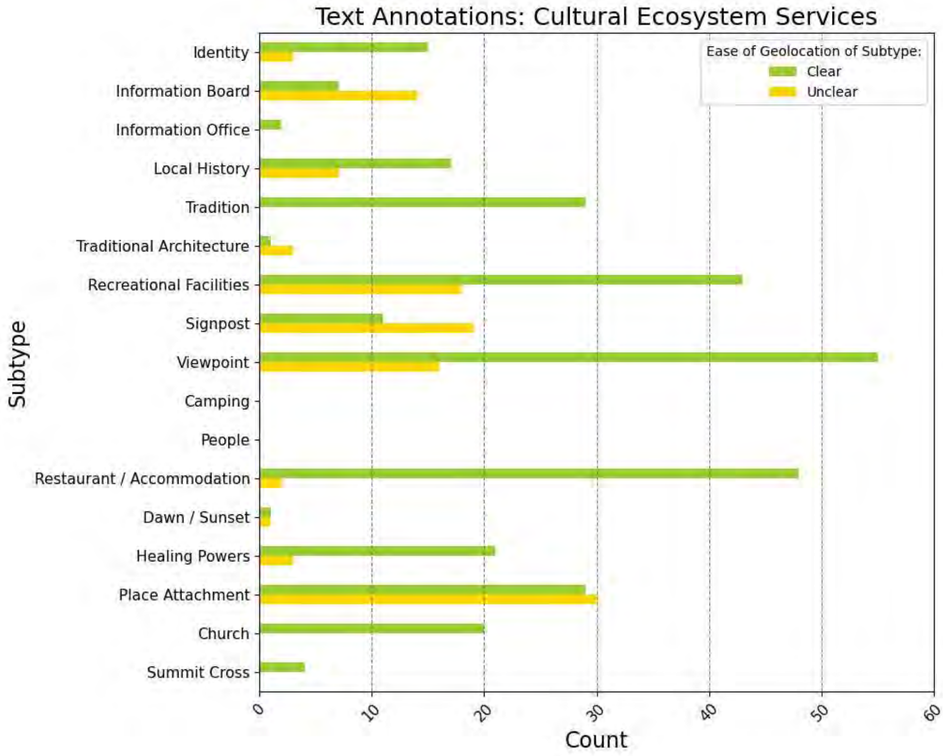
1889 LF annotations were made in the *Flickr* photographs (Figure 5.2 (b)). This number also contains duplicate LF which were annotated in both the fore- and background. By neglecting this attribute, this data set has 1666 annotations for LF. Overall, 49% of all (1889) LF annotations were labeled with *Foreground*. But there are differences among the subtypes. While a lot of subtypes were more often annotated in the *Background* (e.g. *Bedrock*, *Forest*, *Summit*), other subtypes showed a higher share of annotations in the *Foreground* (e.g. *River / Creek*, *Tree*, *Infrastructure*). In terms of total numbers, *Forest* and *Grass- and Moorland* were most often annotated (total of 363 and 512 annotations). *Bedrock* (157), *Summit* (139) and *Infrastructure* (207) also show high numbers of annotations.

When comparing the annotations of LF for the two data sources, the data sets show different patterns. While in the picture data *Forest* and *Grass- and Moorland* stand out, these subtypes do not reach the highest numbers by far in the text. In fact, natural LF (e.g. *Bedrock*, *Forest*, *Grass- and Moorland*) have been more often annotated in the picture data when compared to anthropogenic LF (e.g. *Infrastructure*, *Path / Trail*, *Urban*). The text data reveals the opposite and most of the anthropogenic LF exceed natural LF at least in terms of *clear* annotations, but also to a high degree in terms of total annotations. *Infrastructure* is the only subtype that plays an important role and stands out in both data sources.

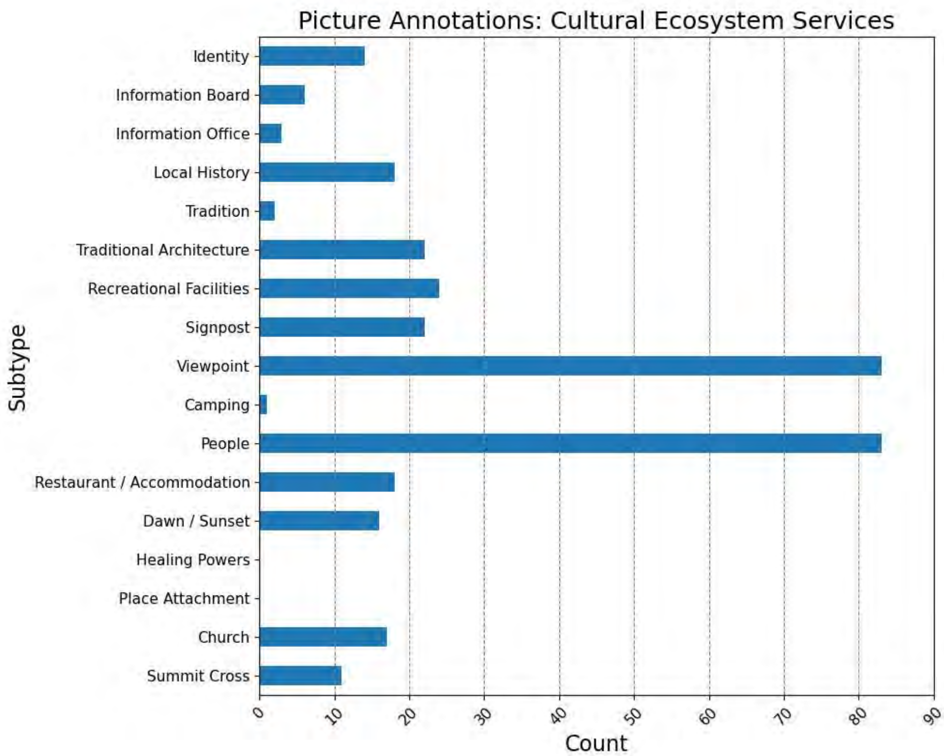
5.2.2. Cultural Ecosystem Services

The text data in Figure 5.3 (a) revealed a total of 419 annotated CES (72% *clear*). Among the clearly geolocated CES, 240 individual CES were found, which means that 57% (240/419) of all CES annotations are mentioned more than once. When looking at the subtypes of CES, *Viewpoint*, *Place Attachment* and *Recreational Facilities* were most often mentioned in the tour descriptions. The share of unclear annotation is small, except for three subtypes. *Information Board*, *Traditional Architecture*, *Signpost* and *Place Attachment* have percentages of more than 50% for *unclear* annotations. In comparison, *Summit Cross*, *Church*, *Information Office* and *Tradition* were never annotated as *unclear*. Other subtypes showed a high share of clearly annotated CES (e.g. *Identity*, *Local History*, *Restaurant / Accommodation*, *Healing Powers*). The subtypes *Camping* and *People* were not found in the tour descriptions.

340 CES were detected in 495 annotated pictures. *Viewpoint* and *People* form the biggest subtypes with more than 80 annotated CES, which exceeds all the other subtypes by more than 50 annotations. Besides the subtypes with only a few annotations (e.g. *Information Board*, *Information Office*, *Tradition*, *Camping*) all the other subtypes do not show any irregularities and are evenly distributed between between 10 and 20 annotations. The subtypes *Healing Powers* and *Place Attachment* were not annotated because they were not identifiable by the content of



(a)



(b)

Figure 5.3.: *Number of CES annotations for (a) tour descriptions and (b) pictures. The CES are ordered alphabetically within their types of CES (Table 4.4).*

pictures.

A comparison between the two data sources shows similarities, but also differences. Both types of data have a high number of annotations for *Viewpoint* and low numbers for *Information Office* and *Camping*. Large differences can be found in subtypes such as *People*, which was not annotated in the text data and showed high numbers in the picture data, or *Tradition* which reached almost 30 annotations in the text data but only 2 in the photographs. *Recreational Facilities* and *Restaurant/Accommodation* are other examples of subtypes, which were more often annotated in the tour descriptions than in the pictures.

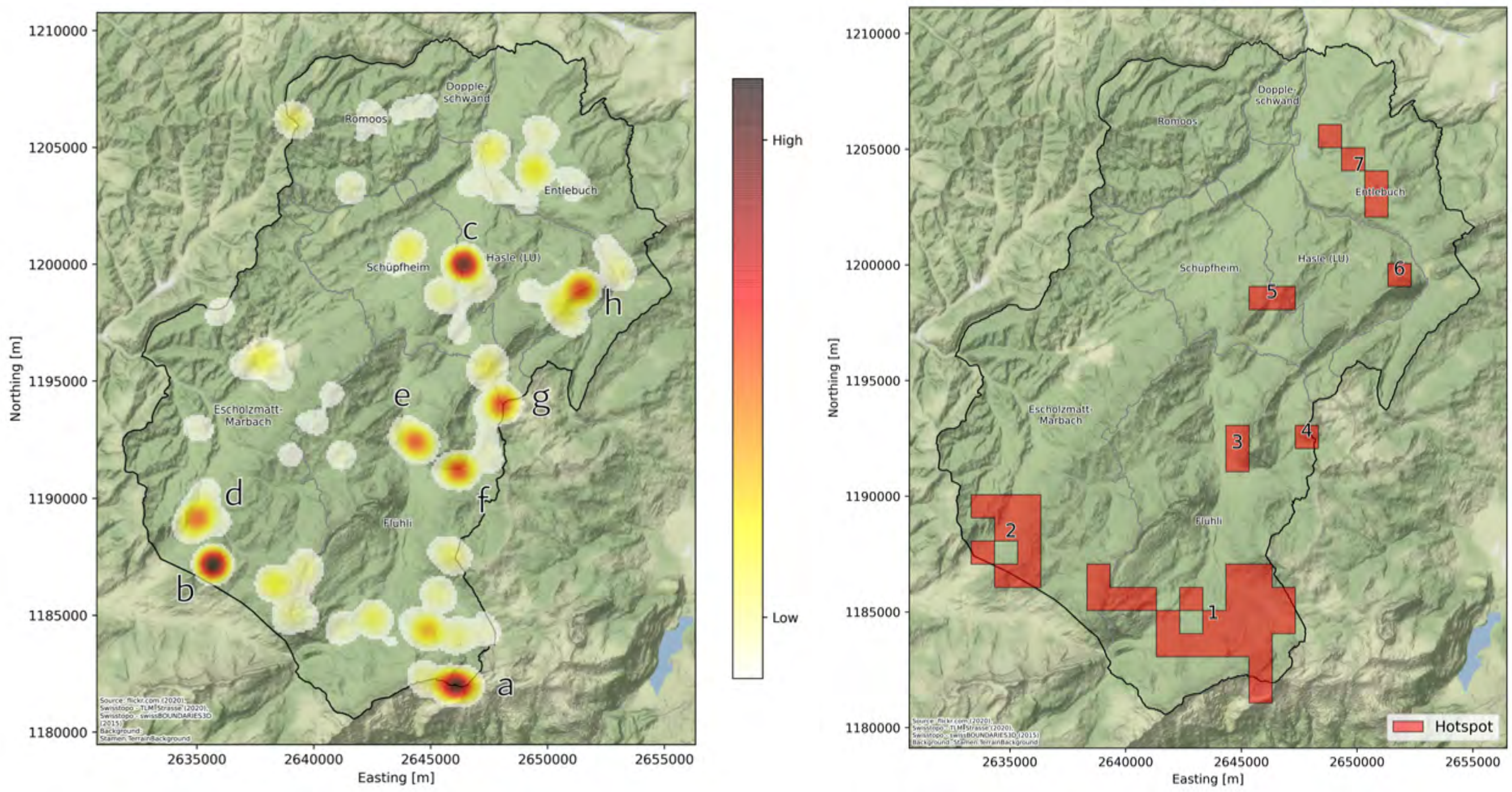
5.3. Detection of CES

The following sections provide maps and background information to the results of the combined data set as well as the two individual data sets. To evaluate the contribution of the two data sources to the combined data set, it is necessary to have a look at the individual results of both the manual and the automatic methods.

5.3.1. Combined Data Set

The manual detection and mapping of CES by using KDE with a combined data set (Figure 5.4 (a)) reveals high densities at specific locations, namely Brienzer Rothorn (a), Marbachegg (b), and Heiligkreuz (c). While the first two sites are most famous for recreational activities and a good view of the surrounding mountains, Heiligkreuz is a pilgrimage destination. Medium densities are shown in the towns of Marbach (d) and Flühli (e). The waterfall in Chessiloch (f) and the summits of Fürstein (g) and Schimbrig (h) also show medium densities. The moorland north of Heiligkreuz and Schimbrig or on the Rossweid, north of Brienzer Rothorn show low densities. The same was found for a lot of regions in the northern, north-western, and western parts of the research area.

The Local Moran's I, of the combination of the *Flickr* and the text data, which was used to extract clusters of high human-nature interaction, revealed 7 clusters for the number of contributors and the number of tour descriptions (Figure 5.4 (b)). Cluster 1 and 2 stand out due to their large extents. Cluster 1 covers the moorland on the Rossweid between Sörenberg, the Brienzer Rothorn and the southern tip of the Schrattenfluh. Cluster 2 reaches from the town of Marbach up to the Marbachegg. The smaller clusters in Flühli (3), Fürstein (4), Heiligkreuz (5), Schimbrig (6) are either villages or summits. Cluster 7 on the Rengg represents a moorland and a small pass road. In the western and north-western parts of the UBE, no clusters were found.



(a) (b)

Figure 5.4.: *KDE of manually extracted CES (a) and automatically extracted clusters by applying a Local Moran's I on the number of contributors and number of tour descriptions per raster cell (b) based on the combined data set.*

5.3.2. Individual Data Sets

Manual Detection

For the text data a total of 274 *clear* annotations were used for the kernel density estimation (Figure 5.5(a)). The highest densities of CES were found in Heiligkreuz (a) and Marbachegg (b). Other mentionable hotspots are located in Flühli (c) and Chessiloch (d) and on the Rengg (e). Overall, the CES seem well distributed: Each municipality and region of the UBE contains CES, even though the densities may vary between locations.

A total of 340 annotations of CES were used for the picture data (Figure 5.5(b)). Brienzer Rothorn (a) is the location that reveals the highest densities in CES in the entire study area for this data type. Other locations with high densities are Marbachegg (b), Heiligkreuz (c) as well as the two summits Fürstein (d) and Schimbrig (e). All other locations of CES for the picture data show much lower densities. Furthermore, the northern, north-western and western part of the research area shows very few signs of CES.

Automatic Detection

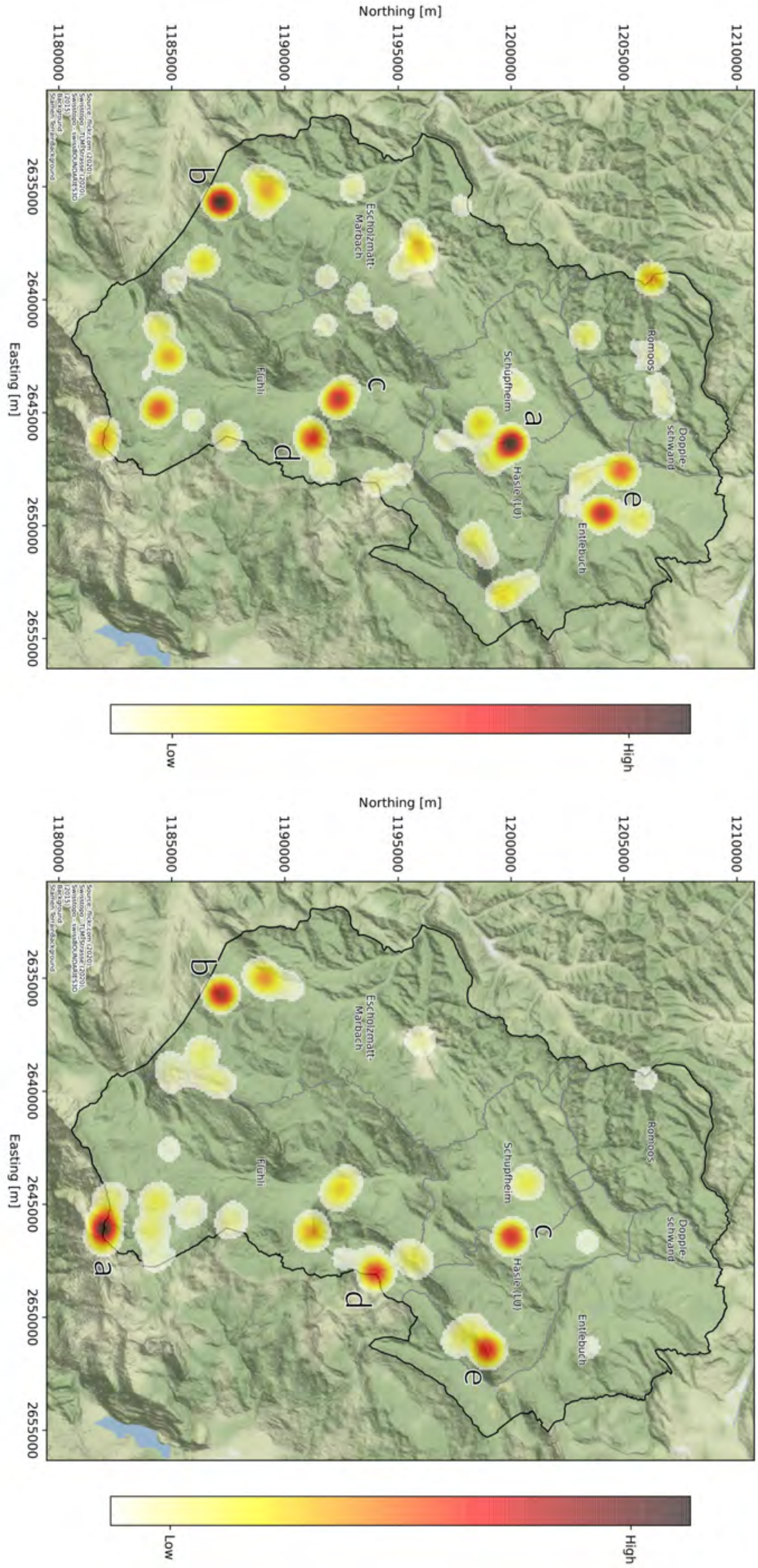
The significant clusters for each type of data based on the number of contributors (pictures) or the number of descriptions (text) were extracted. Figure 5.6 shows that for the text data (a) as well as for the picture data (b), 8 significant clusters were returned.

In the cluster extracted from only the text data, cluster 1 reaches from the Haglere over the town of Sörenberg up to the Rossweid. Cluster 2 is located on the foot of the Schrattenfluh. Cluster 3 covers the area around Marbach and Marbachegg. Clusters 4, 5, and 7 represent the centers of the towns of Escholzmatt (4), Schüpfheim (5), and Entlebuch (7). Cluster 6 is located in the pilgrimage destination of Heiligkreuz and cluster 8 is situated on the Rengg.

The first 3 clusters of the picture data were identified at the same location as the text data but showed different extents. Cluster 1 (Sörenberg and Rossweid) is much larger in the picture data and even reaches the Brienzer Rothorn in the south of the cluster. Cluster 2 spreads over a big part of the Schrattenfluh and not only its foot as cluster 2 in the text data. Cluster 3 in Marbach and Marbachegg is much smaller compared to the text data. Cluster 4 is located in Flühli and clusters 5 and 6 cover the summits of Fürstein (5) and Schimbrig (6). Clusters 7 and 8 are located outside of the town Finsterwald (part of the town Entlebuch) on the Rengg where moorlands are located.

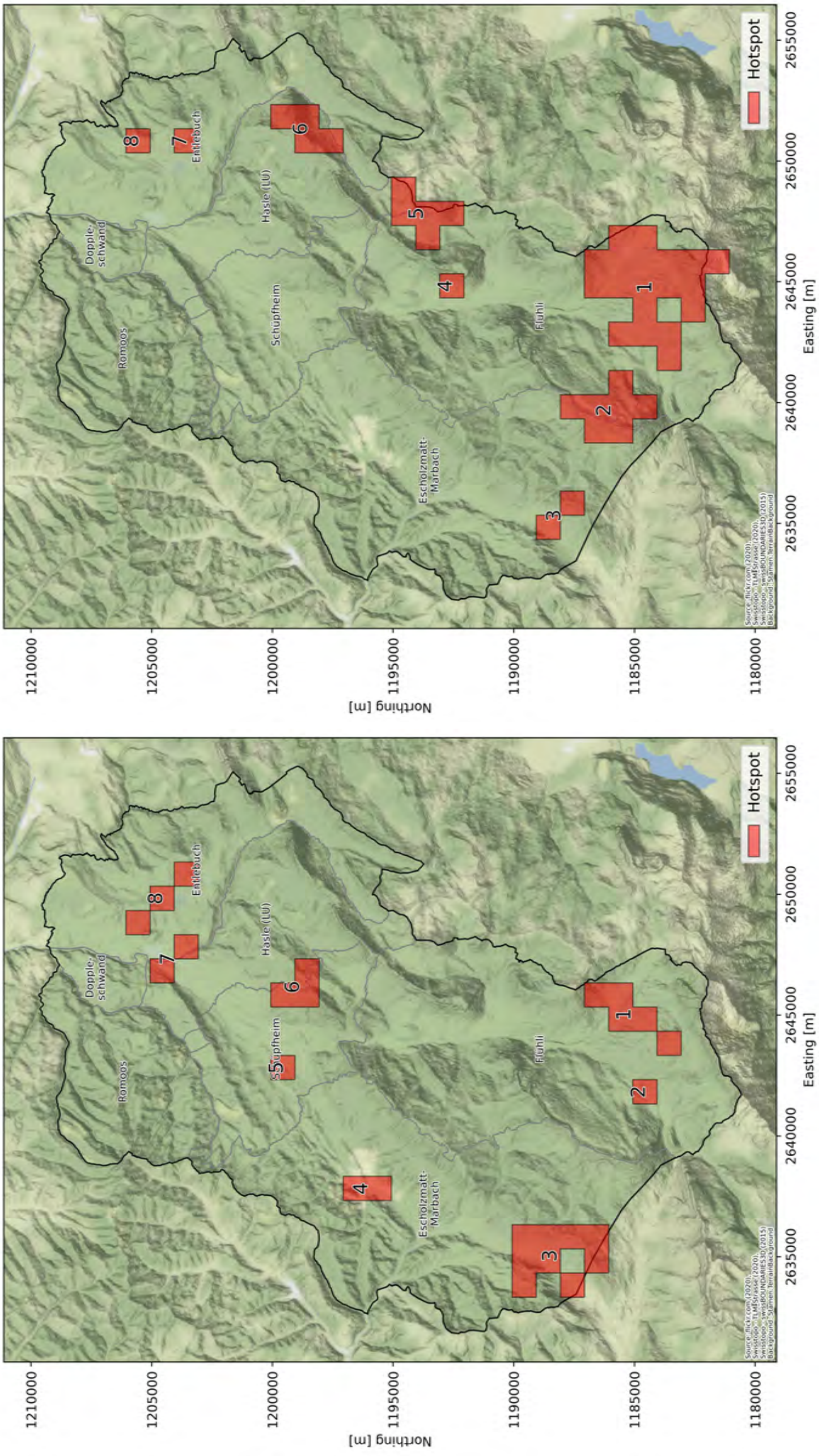
5.3.3. Comparison

Both data sources contribute to the detected hotspots and clusters in both methods. For the kernel density estimations, the combination of the two data sets shows that the densities of hotspots remain high or are even increased if both data sources already show high densities for the same locations (e.g. Marbachegg, Heiligkreuz). In other areas where only one of the two sources showed CES (e.g. Finsterwald (e)) the densities in the combined data set are lowered. Overall, a lot of hotspots are data source dependent, but as soon as the underrepresented data



(a) (b)

Figure 5.5: *KDE of manually annotated CES in (a) the tour descriptions and (b) in pictures.*



(a)

(b)

Figure 5.6.: *Automatically extracted clusters based on (a) the number of tour descriptions and (b) the number of contributors for pictures.*

source shows only small densities, the locations are revealed as hotspots. This is the case for example in Flühli, Chessiloch, and Rengg where the tour descriptions are dominant and the mountain tops of Brienzer Rothorn, Schimbrig, and Fürstein where the picture data is the driving factor. Besides this regulation and amplification effect, each data set's characteristic contributes to the combined result. The CES annotated in the picture data are not very evenly distributed over the research area and enhance the effect of already apparent hotspots, especially if they are already apparent in the text data. The text data contributes by adding CES in regions that have not been captured by the picture data.

For the automatic method similar effects can be discovered. While cluster 3 and 4 in Figure 5.4 (b) are mainly extracted due to the clusters 4 and 5 in the picture data (cluster 4 and 5 in Figure 5.6 (b)), cluster 5 has been revealed in the text data (cluster 6 in Figure 5.6 (a)). Additionally, cluster 7 seems like a union between cluster 8 in the text data and clusters 7 and 8 in the picture data. Cluster 1 is covering the cluster with the number 2 in Figure 5.6 (a). It is also possible to detect the influence of the individual data sets on the extraction of the clusters with the combined data set.

The combination of two data sources cannot only retrieve clusters that are apparent in both data types but also by only one data type. E.g. through the inclusion of different types of data the characteristics of each cluster are changed. Some gain in size because both individual data sets contribute to the combined cluster (e.g. cluster 1 in Figure 5.4 (b)), others lose in size (e.g. cluster 5 in Figure 5.4 (b)) because it is mostly represented by one data source or even disappear because of this effect.

Table 5.1.: *Jaccard-Indexes between the different cluster results.*

Clusters	*Flickr*	Text	Combined
Flickr	1	0.12	0.47
Text	0.12	1	0.43
Combined	0.47	0.43	1

The Jaccard similarity coefficients between different cluster results are displayed in Table 5.1. The coefficient between the two individual data sources (*Flickr* and Text) showed a very low value of 0.12. This low value indicates a very low overlap of the two clusters extracted from the *Flickr* data and the tour descriptions. Both data sources show similar values when compared to the combined data set (0.47 for *Flickr* and 0.43 for text data) which means that both individual data sets contribute to the result of the combined data set.

5.4. Relations between CES and LF

By using the 20 meters buffer around CES to capture LF for specific subtypes of CES, I was able to investigate the relationship between subtypes of CES and LF. The first subsection looks at the results between CES and LF of the combined data set for both approaches to detect CES.

The middle two sections compare CES and clusters and their relations with LF based on the results of each data source.

Because a few subtypes of CES showed very low numbers in Figure 5.3, the focus is on subtypes of CES which appeared in higher numbers to investigate their relations to LF. In addition, primarily subtypes of CES which are not directly linked to a LF were chosen (e.g. *Traditional Architecture* or *Restaurant / Accommodation* are quickly related to *Infrastructure*). Even though only subsets of all results are presented and looked at in more detail, the column *Total* contains the values of all subtypes of CES in Section 5.4.2 and all clusters in Section 5.4.2. The complete tables including all subtypes of CES and all clusters are displayed in Appendix A.2.

The tables for the manually detected CES are colored according to Figure 5.7. The three values have to be explained in more detail:

- **N** is the absolute number of annotations of a LF subtype a CES subtype (e.g. *Identity*) for the manual method was able to capture. For the automatic method, it is the number of annotations of a LF subtype a cluster was able to cover.
- **%LF** is the proportion of N compared to all annotations of a LF subtype. For the automatic method, this value was normalized by the size of the cluster.
- **%CES** is the share of a CES subtype (e.g. *Identity*) that contributed to capturing N.

The results between *Natural Landscape* and *Identity* in Table 5.2 are used for an illustration. The %LF value (28.12) lies in the highest category. Because the %CES value (26.92) is located in the middle category, the bivariate coloration assigns the combination of colors to (high-mid value) to the cell.

For the coloration of the results of the automatic method (Figure 5.8) only the %LF values of the combined data set (Table 5.3) were used. Similar to the manual method, I decided to only analyze a subset of clusters for the automatic method. The goal was to choose a big variety of clusters; clusters that are only apparent in one of the data sources, clusters that were extracted by both types of data, and preferably clusters with different sizes.

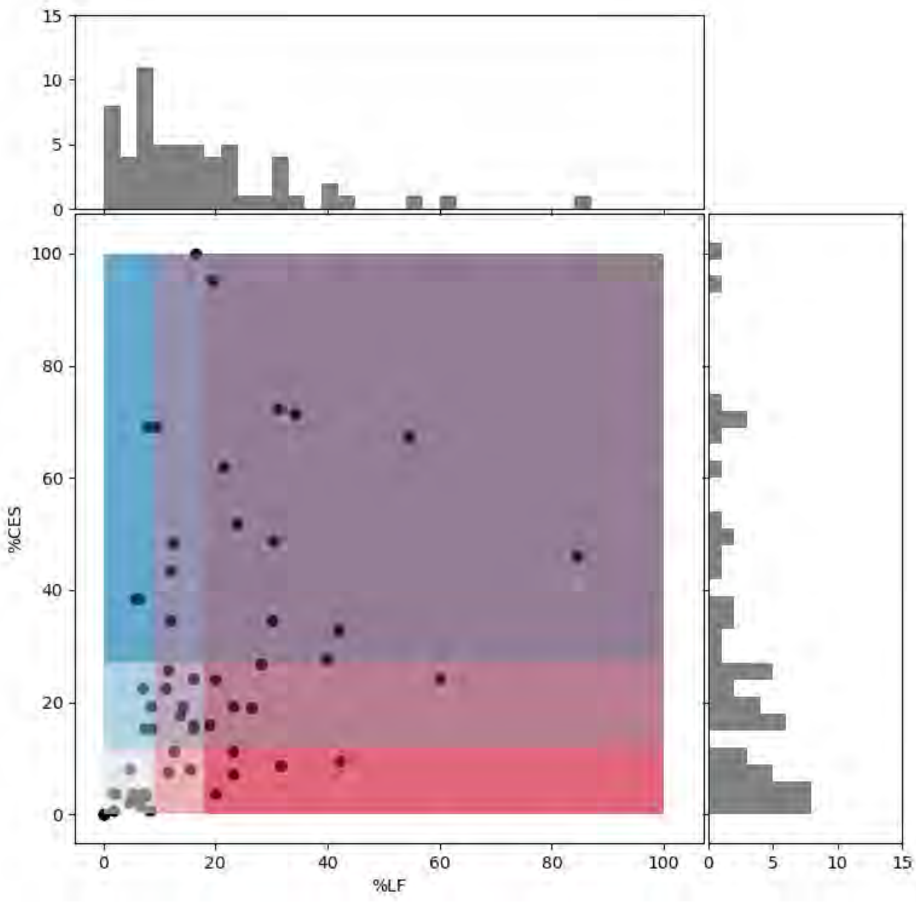


Figure 5.7.: *Coloration of the tables representing the relations between manually detected CES and LF. The data points are the values from the selected subtypes of CES of the combined data set (Table 5.2).*

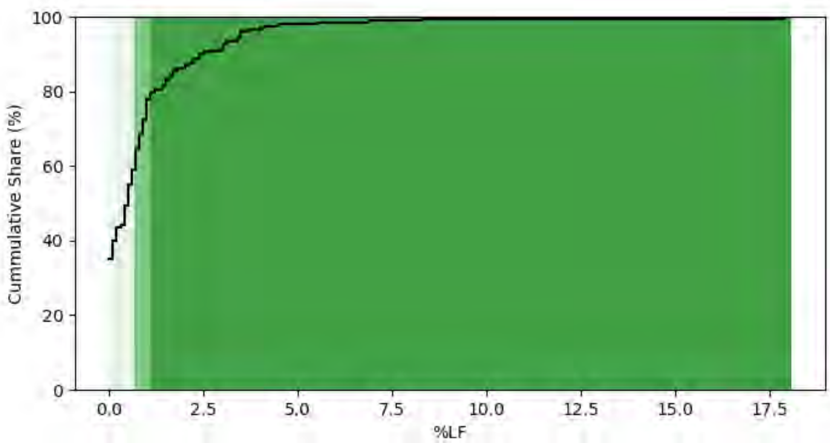


Figure 5.8.: *Coloration of the tables representing the relations between automatically detected CES and LF. The data represent the values all detected clusters of the combined data set (Table 5.3).*

5.4.1. Combined Data Sets

Table 5.2.: *Count of LF within the threshold distance to a CES (N), its proportion of all annotations of a LF subtype (%LF) and the share of CES capturing N (%CES).*

	Identity			Recreational Facilities			Viewpoint			Total		
	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES
Natural Landscape	9	28.12	26.92	6	18.75	16.13	27	84.38	45.99	30	93.75	23.62
Human influenced Landscape	3	11.5	7.7	2	7.7	3.2	6	23.1	7.3	17	65.4	7.0
Bedrock	10	6.3	38.5	11	6.9	22.6	48	30.2	48.9	102	64.2	36.6
Flower / Funghi	1	2.3	3.9	2	4.7	8.1	2	4.7	2.2	12	27.9	5.5
Forest	28	8	69.2	42	12.9	43.6	110	31.3	72.3	208	59.1	59.6
Grass- and Moorland	37	9.3	69.2	49	12.3	48.4	136	34.3	71.5	236	59.5	60.6
Lake	9	30.0	34.6	18	60.0	24.2	12	40.0	27.8	26	86.7	19.7
River / Creek	1	1.8	3.9	7	12.5	11.3	4	7.1	3.7	35	62.5	10.9
Rock	7	8.4	15.4	5	6.0	3.2	22	26.5	19.0	46	55.4	15.0
Shrub	1	1.8	3.9	8	14.0	19.4	1	1.8	0.7	33	57.9	11.1
Snow / Ice	5	16.1	15.4	2	6.5	1.6	13	41.9	32.9	17	54.8	14.3
Summit	13	8.4	19.2	17	11.0	22.6	84	54.6	67.2	100	64.9	35.2
Tree	9	12.0	34.6	12	16.0	24.2	12	16.0	16.1	48	64.0	23.5
Waterfall	0.0	0.0	0.0	0.0	0.0	0.0	1	8.3	0.7	9	75.0	1.8
Wild Animal	3	23.1	19.2	3	23.1	11.3	2	15.4	8.0	6	46.2	7.5
Agriculture	0.0	0.0	0.0	0.0	0.0	0.0	2	20.0	3.7	6	60.0	2.8
Infrastructure	67	16.6	100	79	19.6	95.2	87	21.5	62.0	316	78.2	82.3
Livestock	1	5.3	3.9	8	42.1	9.7	6	31.6	8.8	14	73.7	10.6
Path / Trail	9	5.8	38.5	18	11.5	25.8	37	23.7	51.8	88	56.4	40.7
Urban	6	7.5	15.4	11	13.8	17.7	16	20.0	24.1	48	60.0	19.1

The extraction of LF which are within 20 meters of CES reveals different results depending on the CES subtype (Table 5.2). While *Viewpoint* shows high %LF and high %CES values (high-high) for multiple LF (e.g. *Natural Landscape*, *Forest*, *Lake*, *Infrastructure*), *Recreational Facilities* (*Infrastructure*) and *Identity* (*Lake*) only show high-high values for one LF. Comparing *Identity* to the other CES subtypes, shows that *Identity* is not able to capture high proportions of a LF (low %LF values), but at the same time has medium to high %CES values which means that a high share of *Identity* CES capture the few LF. *Recreational Facilities* shows high %CES numbers for *Forest* (43.55%), *Grass- and Moorland* (48.39%) as well as for *Infrastructure* (95.16%). All subtypes of CES (not only the ones mentioned here) reveal high %LF values, while the %CES values are rather low.

Table 5.3.: *Absolute (N) and relative (%LF) densities of LF for each cluster and km^2 for both data sources.*

	Cluster 1		Cluster 2		Cluster 3		Cluster 5		Total	
	N	%LF	N	%LF	N	%LF	N	%LF	N	%LF
Natural Landscape	0.14	1.7	0.56	6.94	0	0	0	0	0.2	2.44
Human influenced Landscape	0.14	2.27	0.11	1.85	0	0	0	0	0.15	2.44
Bedrock	1.18	0.77	0.78	0.51	1.0	0.65	0	0	0.9	0.59
Flower / Funghi	0.18	0.45	0.44	1.11	1.5	3.75	0	0	0.29	0.73
Forest	2.14	0.67	3.44	1.08	3.0	0.94	0	0	2.2	0.69
Grass- and Moorland	3.45	0.92	4.11	1.09	3.5	0.93	0	0	3.17	0.84
Lake	0.18	0.73	0.44	1.78	4.5	18.0	0	0	0.41	1.66
River / Creek	0.18	0.53	0.22	0.65	0	0	0	0	0.15	0.43
Rock	0.73	0.89	0	0	0	0	0	0	0.44	0.54
Shrub	0.23	0.41	1.0	1.79	2.0	3.57	0	0	0.46	0.83
Snow / Ice	0.14	0.45	0.22	0.74	0	0	0	0	0.15	0.49
Summit	1.09	0.77	1.22	0.86	1.0	0.7	0.5	0.35	1	0.7
Tree	0.55	0.77	1.0	1.41	2.5	3.52	0	0	0.68	0.96
Waterfall	0	0	0	0	0	0	0	0	0	0
Wild Animal	0.05	0.65	0.22	3.17	0	0	0	0	0.07	1.05
Agriculture	0.05	0.57	0.22	2.78	0	0	0	0	0.07	0.91
Infrastructure	3.09	1.14	8.33	3.09	5.5	2.04	1.5	0.56	4.05	1.5
Livestock	0.09	0.53	0	0	0	0	0	0	0.07	0.43
Path / Trail	0.86	0.79	0.89	0.82	0	0	1.0	0.92	0.83	0.76
Urban	0.55	0.94	1.44	2.49	0	0	1.0	1.72	0.71	1.22

In Table 5.3, multiple LF show high values for more than one cluster (e.g. *Natural Landscape* for Cluster 1 and 2, *Flower/Funghi* for Cluster 2 and 3 or *Urban* for cluster 2 and 5. *Infrastructure* shows high values for cluster 1, 2 and 3. Only cluster 2 is able to capture a high share of a specific LF while the other clusters do not perform as well for the same LF, as for example *Grass- and Moorland*, *Wild Animal* or *Agriculture*. Cluster 1 and 3 do not stand out when comparing their values to other clusters. Cluster 5 only captures a few LF and shows a high value for *Urban* (1.72%). All automatically extracted clusters from the combined data set are able to capture high shares of LF for *Natural Landscape*, *Human influenced Landscape*, *Lake*, *Infrastructure* and *Urban*.

5.4.2. Individual Data Sets

Manual Detection

Table 5.4.: *Count of LF within the threshold distance to a CES (N), its proportion of all annotations of a LF subtype (%LF) and the share of CES capturing N (%CES) for text data.*

	Identity			Recreational Facilities			Viewpoint			Place Attachment			Total		
	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES
Natural Landscape	4	12.5	50.0	5	15.6	23.7	27	84.4	63.0	9	28.1	26.9	30	93.8	35.0
Human influenced Landscape	3	11.5	16.7	2	7.7	5.3	6	23.1	16.7	1	3.9	3.9	17	65.4	15.0
Bedrock	2	22.2	33.3	2	22.2	10.5	3	33.3	11.1	3	33.3	11.5	8	88.9	10.2
Flower / Funghi	0.0	0.0	0.0	1	33.3	7.9	0.0	0.0	0.0	2	66.7	7.7	2	66.7	7.3
Forest	1	2.6	25.0	7	18.0	26.3	2	5.1	3.7	6	15.4	30.8	17	43.6	12.8
Grass- and Moorland	2	6.9	25.0	6	20.7	18.4	2	6.9	3.7	12	41.4	34.6	17	58.6	15.3
Lake	6	37.5	58.3	10	62.5	23.7	5	31.3	11.1	7	43.8	19.2	14	87.5	13.1
River / Creek	1	4.2	8.3	6	25.0	15.8	2	8.3	7.4	3	12.5	11.5	14	58.3	9.5
Rock	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Shrub	0.0	0.0	0.0	1	25.0	2.6	0.0	0.0	0.0	3	75.0	11.5	4	100	3.7
Snow / Ice	1	25.0	16.7	0.0	0.0	0.0	4	100	18.5	1	25.0	3.9	4	100	9.9
Summit	3	17.7	16.7	0.0	0.0	0.0	15	88.2	33.3	4	23.5	7.7	15	88.2	15.0
Tree	0.0	0.0	0.0	0.0	0.0	0.0	1	25.0	1.9	0.0	0.0	0.0	2	50.0	1.1
Waterfall	0.0	0.0	0.0	0.0	0.0	0.0	1	20.0	1.9	5	100	11.5	5	100	1.5
Wild Animal	3	33.3	41.7	2	22.2	15.8	1	11.1	3.7	3	33.3	11.5	5	55.6	11.3
Agriculture	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1	50.0	0.4
Infrastructure	28	13.9	100	49	24.3	94.7	37	18.3	64.8	34	16.8	53.9	161	79.7	83.9
Livestock	0.0	0.0	0.0	1	50.0	2.6	1	50.0	5.6	0.0	0.0	0.0	1	50.0	2.9
Path / Trail	1	1.7	8.3	6	10.3	13.2	8	13.8	16.7	7	12.1	34.6	26	44.8	17.9
Urban	2	5.4	8.3	3	8.1	10.5	5	13.5	13.0	2	5.4	7.7	18	48.7	11.7

Table 5.4 shows the results for the text data. Overall, there are a lot of low-low values. *Rock* and *Agriculture* were not found within 20 meters to one of the selected subtypes of CES. Each subtype of CES has several columns with low-low values, but also with high-high values. *Identity* performs well at capturing LF of the subtypes *Bedrock* (22.22%), *Lake* (37.5%) and *Wild Animal* (33.33%). At the same time these subtypes are also able to capture between 33.33% (*Bedrock*) and 58.33% (*Lake)* of all *Identity* CES. On the one hand, *Recreational Facilities* is only good at capturing *Infrastructure* (24.26%) and the other way around (94.74%). All other LF may reach similar values in terms of %LF, but not in %CES, meaning that only a few *Recreational Facilities* capture the LF. *Viewpoint* on the other hand, shows promising results when looking at *Natural Landscape*, *Summit* and *Infrastructure*. *Place Attachment*, similar to *Recreational Facilities*, can capture high shares of specific subtypes of LF, but only a small proportion of CES captures the LF, which results in low %CES values. Only *Grass- and Moorland* is classified as high-high. The column *Total* shows that the results for all LF (except *Rock*) reveal high

%LF values with predominantly low %CES values.

Table 5.5.: *Count of LF within the threshold distance to a CES (N), its proportion of all annotations of a LF subtype (%LF) and the share of CES capturing N (%CES) for picture data.*

	Identity			Recreational Facilities			Viewpoint			People			Total		
	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES
Bedrock	6	4.0	42.9	8	5.3	41.7	45	30.0	65.1	46	30.7	57.8	91	60.7	53.2
Flower / Funghi	1	2.5	7.1	1	2.5	8.3	2	5.0	3.6	3	7.5	4.8	9	22.5	3.8
Forest	20	6.4	92.9	35	11.2	70.8	108	34.5	97.6	78	24.9	77.1	187	59.7	84.4
Grass- and Moorland	28	7.6	92.9	43	11.7	95.8	134	36.4	97.6	96	26.1	74.7	218	59.2	85.9
Lake	2	14.3	14.3	7	50.0	25.0	7	50.0	33.7	8	57.1	12.1	12	85.7	23.2
River / Creek	0.0	0.0	0.0	1	3.1	4.2	2	6.3	1.2	13	40.6	20.5	16	50.0	12.1
Rock	3	3.7	14.3	5	6.1	8.3	22	26.8	27.7	28	34.2	25.3	40	48.8	24.4
Shrub	1	1.9	7.1	7	13.2	45.8	1	1.9	1.2	11	20.8	28.9	24	45.3	16.5
Snow / Ice	2	7.4	7.1	1	3.7	4.2	9	33.3	34.9	4	14.8	4.8	13	48.2	15.3
Summit	6	4.4	21.4	14	10.2	58.3	69	50.4	88.0	25	18.3	36.1	85	62.0	50.9
Tree	9	12.7	64.3	12	16.9	62.5	11	15.5	18.1	20	28.2	41.0	46	64.8	36.5
Waterfall	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4	57.1	6.0	4	57.1	2.0
Wild Animal	0.0	0.0	0.0	0.0	0.0	0.0	1	25.0	1.2	0.0	0.0	0.0	1	25.0	0.6
Agriculture	0.0	0.0	0.0	0.0	0.0	0.0	2	25.0	4.8	1	12.5	1.2	5	62.5	4.4
Infrastructure	31	15.4	100	29	14.4	95.8	50	24.8	56.6	76	37.6	67.5	155	76.7	79.7
Livestock	1	5.9	7.1	7	41.2	20.8	5	29.4	8.4	10	58.8	16.9	13	76.5	15.9
Path / Trail	7	7.1	50.0	12	12.2	45.8	29	29.6	62.7	31	31.6	49.4	62	63.3	51.8
Urban	1	2.3	7.1	8	18.6	29.2	11	25.6	27.7	14	32.6	13.3	30	69.8	20.9

The majority of low-low values can be found in two subtypes of CES (*Identity* and *Recreational Facilities*). The %CES proportions are high for several LF in *Identity*, but this CES subtype does not capture high percentages of LF (%LF), which means that the LF is not very relevant for the CES subtype. The results for *Recreational Facilities* are most often classified as a mid-high and high share of CES capture LF (high %CES value), but not many LF are captured (mid %LF value). 10 LF show high-high results for *Viewpoint* (*Bedrock, Forest, Grass- and Moorland, Lake, Rock, Snow/Ice, Summit, Infrastructure, Path/Trail, Urban*). *People*, a CES which is only found in the picture data, is often represented by 7 LF (*Bedrock*, *Forest*, *Shrub*, *Summit*, *Tree*, *Infrastructure*, *Path/Trail*). Over all subtypes of CES, a lot of LF show high %LF values and mid to high %CES values.

Automatic Detection

The results of the text data (Table 5.6) show that many clusters were only able to partly capture specific LF (e.g. *River/Creek*) or not at all (e.g. *Tree* or *Waterfall*). Cluster 1 (Rossweid) shows high numbers of %LF for many LF with *Human influenced Landscape*, *Shrub* and *Urban* being the most mentionable. Cluster 3 (Marbach and Marbachegg) is the only cluster in this subset to show high values in *Natural Landscape* and *Snow/Ice*. All other high values are either

Table 5.6.: *Absolute (N) and relative (%LF) densities of LF for each cluster and km^2 for text data.*

	Cluster 1		Cluster 3		Cluster 6		Total	
	N	%LF	N	%LF	N	%LF	N	%LF
Natural Landscape	0.2	0.62	0.56	1.74	0	0	0.23	0.72
Human influenced Landscape	0.4	1.54	0.11	0.43	0	0	0.15	0.59
Bedrock	0.2	2.22	0.33	3.7	0	0	0.15	1.71
Flower / Funghi	0.2	1.82	0	0	0.33	3.03	0.08	0.7
Forest	0.8	2.05	0	0	1.0	2.56	0.42	1.08
Grass- and Moorland	0.8	2.76	0	0	0.33	1.15	0.23	0.8
Lake	0.2	1.25	0.44	2.78	0	0	0.19	1.2
River / Creek	0	0	0.22	0.93	0	0	0.19	0.8
Rock	0	0	0	0	0	0	0	0
Shrub	0.6	15.0	0	0	0.33	8.33	0.15	3.85
Snow / Ice	0	0	0.22	5.56	0	0	0.08	1.92
Summit	0	0	0.22	1.31	0.33	1.96	0.12	0.68
Tree	0	0	0	0	0	0	0	0
Waterfall	0	0	0	0	0	0	0	0
Wild Animal	0.2	2.22	0.22	2.47	0	0	0.12	1.28
Agriculture	0	0	0	0	0	0	0	0
Infrastructure	3.6	1.78	3.78	1.87	2.33	1.16	2.88	1.43
Livestock	0	0	0	0	0	0	0	0
Path / Trail	0.2	0.34	0.11	0.19	2.0	3.45	0.65	1.13
Urban	1.2	3.24	0.78	2.1	0.67	1.8	0.65	1.77

shared with Cluster 1 or Cluster 6. Cluster 6 (Heiligkreuz) shows high values for the similar LF as Cluster 1 (e.g. *Flower/Funghi*, *Forest*, *Grass- and Moorland*), while only *Path/Trail* is categorized as high in no other cluster. For all clusters in the text data, a lot of anthropogenic LF show high values (e.g. *Urban*, *Path/Trail*, *Infrastructure*) and only a few natural LF show high values with *Shrub* (3.85%) and *Snow/Ice* (1.92%) revealing the highest %LF values.

Table 5.7.: *Absolute (N) and relative (%LF) densities of LF for each cluster and km^2 for picture data.*

	Cluster 1		Cluster 3		Cluster 4		Total	
	N	%LF	N	%LF	N	%LF	N	%LF
Bedrock	1.79	1.19	2.0	1.33	2.0	1.33	1.98	1.32
Flower / Funghi	0.53	1.32	0	0	1.0	2.5	0.38	0.94
Forest	2.95	0.94	10.5	3.35	3.0	0.96	3.67	1.17
Grass- and Moorland	4.58	1.24	13.0	3.53	5.0	1.36	5.0	1.36
Lake	0.16	1.13	0	0	5.0	35.71	0.28	1.96
River / Creek	0.21	0.66	0	0	0	0	0.1	0.31
Rock	1.21	1.48	0	0	0	0	1.05	1.28
Shrub	0.11	0.2	2.0	3.77	2.0	3.77	0.28	0.52
Snow / Ice	0.32	1.17	0	0	0	0	0.4	1.48
Summit	1.21	0.88	3.5	2.55	2.0	1.46	2.05	1.5
Tree	0.63	0.89	2.0	2.82	5.0	7.04	0.68	0.95
Waterfall	0	0	0	0	0	0	0	0
Wild Animal	0.05	1.32	0	0	0	0	0.05	1.25
Agriculture	0.05	0.66	1.0	12.5	0	0	0.1	1.25
Infrastructure	2.63	1.3	14.0	6.93	4.0	1.98	2.55	1.26
Livestock	0.16	0.93	0	0	0	0	0.17	1.03
Path / Trail	1.21	1.24	1.5	1.53	0	0	1.25	1.28
Urban	0.42	0.98	0.5	1.16	0	0	0.35	0.81

For the picture data, all selected clusters perform well at capturing a high share for specific LF (*Bedrock*, *Grass- and Moorland*, *Infrastructure*). Cluster 1 (Rossweid) shows high values for multiple LF. These values are classified as high, but are still only range from 1.13% for *Lake* to 1.32% *Flower/Funghi* and *Wild Animal*. In comparison, cluster 4 (Flühli) is able to capture 35.71% or all *Lake* annotations and 7.04% of all *Tree* annotations. Cluster 3 (Marbach and Marbachegg) only shows high values, if the cluster was able to capture the LF. The extracted clusters from the picture data performed rather well in capturing high shares of LF, but the range of these values is between 1.17% (*Forest*) and 1.96% (*Lake*).

5.4.3. Comparison

When comparing the two data sources, it stands out that the picture data shows many high-high values for specific CES subtypes (*Viewpoint*, *People*) for the manually extracted CES but also a lot of low-low values (*Identity*, *Recreational Facilities*) when compared to the text data. For the manually annotated CES, the results of the text data are more distributed over the selected CES subtypes where for each subtype, a decent number of relevant LF with high-high values can be extracted. The picture data generally show high %CES values, if a LF was captured. Additionally, often the %LF values also reach a value of at least 40%. The tour descriptions are good at capturing high proportions of LF, while the %CES values often remain low (below 10%). The combination of the two data sets can smooth the variability of the two data sets and combine the attributes and strengths of the individual data sets, which results in fewer high-high values and higher %LF values than in the picture data.

For the automatically extracted clusters, the text data is much more similar to the picture data. Both data sources show a lot of high %LF values compared to the combined data set. The text data shows similar characteristics as the picture data for the manually annotated CES: If a LF was found within the cluster, a high share (high %LF value) is the result.

Discussion

This chapter discusses the results from the previous chapter with the literature. The results show. Afterward, methodological limitations, as well as limitations regarding the data sources, are lined out. The last section provides an outlook for future research.

6.1. Interpretation of Results

6.1.1. Inter-Annotator Analysis

The results of the Jaccard-Indexes for the pictures (without *Fore-/Background*: 67.04%) are better than for the text data (clear annotations: 42.24%). Especially the fact that the Jaccard-Index is lower when including the *Fore-/Background* attribute for the picture data demonstrates, that annotating too many attributes does not always pay off.

Additionally, since humans and not machines did the annotations, the annotators' perceptions build a very crucial aspect in this part of the thesis. Consequently, the estimation of the partly required 5% coverage is error-prone, as well as the perception and interpretation of specific features. This leads to a lower Jaccard-value because I had tested the subtypes by going through a subset of tour descriptions and pictures beforehand. Therefore, I already knew how I would annotate which features, whereas the second annotator did not have such prior knowledge. For example, I knew that I would annotate a Streuwiese (bedding meadow) as *Local History* because it is a historical type of using a meadow, as well as *Grass- and Moorland* because of its natural characteristic.

Oteros-Rozas et al. (2018) worked with researchers with significant knowledge about their study landscapes. They performed cross-checks until a consensus between the annotations was obtained which was not possible in the scope of this study. Additionally, neither I nor the second annotator would call themselves local experts of the research area.

6.1.2. Annotation of CES

The results in Figure 5.3 demonstrate that each data source has its characteristics when it comes to the number of CES annotations. The CES from photographs are more concentrated on *Viewpoint*. This could be expected since *Flickr* is known to provide landscape pictures and since UBE provides a lot of summits, these landscape pictures are classified as *Viewpoint*. Rather surprising is the high number of CES for the subtype *People* because social interactions are more often displayed on platforms such as *Instagram* (Hausmann et al., 2017a; Toivonen et al., 2019). Each data type can better capture specific CES subtypes. Different CES may be better represented in specific data sets (e.g. a sulfur spring is not necessarily identifiable on pictures but in text data) or by specific people (e.g. locals or tourists). While *Tradition* or *Restaurant / Accommodation* is predominantly found in tour descriptions, *People* or *Traditional Architecture* is better represented in the picture data. For example, Richards and Friess (2015) states that social media photographs underrepresent cultural heritage and spiritual values, which can also be seen in Figure 5.3. This agrees with what Casalegno et al. (2013) who found that the representation of different CES is highly dependent on the context as well as the source of data. In addition, cultural values are highly dependent on people's expectations, perceptions, and needs (Daniel et al., 2012).

The tour descriptions revealed more CES (419) than the picture data (341). A cause for this might be that CES represent important locations for directions, but also for recommendations (e.g. "An dieser Stelle geniessen Sie eine herrliche Aussicht [...]" ("At this point, you can enjoy a magnificent view[...]")[1] or a restaurant which serves local food). In Wartmann et al. (2018), text data (hiking blogs), free lists (participants listing words and expressions), and *Flickr* tags were compared concerning their ability to extract landscape aspects. Hiking blogs showed a slightly higher proportion than *Flickr* tags for cultural LF (Figure 2.4). This thesis' results support the findings by Wartmann et al. (2018).

In general, comparisons to previous studies are very difficult because of varying categories of CES and subtypes and different categorizations in each study. Additionally, most studies focus on different spatial scales and multiple types of landscapes are investigated (e.g. Casalegno et al. (2013), van Zanten et al. (2016a), van Zanten et al. (2016b)), which complicates comparisons of landscape preference studies due to differences in research area or context (van Zanten et al., 2016b).

In Bieling and Plieninger (2013) most visible manifestations of CES were assigned to *Recreation* which included several categories which I categorized separately (e.g. hiking trail signs, recreational facilities, benches). But summing up my annotations of these categories would also lead to them being the largest category. Nevertheless, their research focused on a single type of land use, *Streuobstwiesen* (meadows with fruit trees), which is a rather open space, whereas the UBE also consists of mountainous parts as well as many forests. In addition, they observed that some CES can be allocated to different subcategories of CES (e.g. benches could belong to aesthetic (because of the view) and recreational services). Especially that a CES can have a

[1] Hiking Tour: Wiggen-Wachthubel-Marbach

different meaning for different cultures and different people within the same culture has to be mentioned (Byers et al., 2001). This can also be seen with the *Summit Crosses* in the UBE. Some people approach them because of religious reasons, for others, they represent the top of a mountain and hence the highest point of their route. Thus, after locating CES, it is still unclear in what way they are used and perceived.

6.1.3. Annotation of LF

In Richards et al. (2018) the majority of photographs were classified as animals or plants by the Google Cloud Vision. While Figure 5.2 (b) agrees with plants (*Flower/Funghi*, *Forest*, *Grass- and Moorland*) as being a dominant LF extracted from pictures, animals (whether *Wild Animals* or *Livestock*) are not often captured. But this fact is mainly due to the location and characteristics (e.g. attractions and properties) of the research area. Richards et al. (2018) chose nature reserves in Singapore as a case study area. The population is located very close to these reserves and CES, which means that one has to be cautious when comparing my results to theirs because of different motivations. Additionally, the fact that Google Cloud Vision annotated many pictures as animals leads to the assumption that animals may be among the reasons why people visit the nature reserves, while natural LF, such as *Forest*, *Bedrock* and *Summit* are often found in pictures of the UBE.

This shows that the two data sources extract different LF. Tour descriptions are predominantly and more often mentioning anthropogenic LF, such as *Infrastructure*, *Urban* or *Path/Trail*. This difference is caused by the characteristics and the limitations of the text data as tour descriptions instead of tour reports (see Section 6.3).

Text data and *Flickr* tags showed similar values for biophysical landscape aspects (Figure 2.4) in Wartmann et al. (2018). This contradicts the findings of this thesis. While in total 799 LF were annotated in the text data, more than twice as many annotations (1666) (without *Fore-/Background*) were made in the photographs. Even though these numbers represent the total number of LF annotations, Figure 5.2 reveals that there is a big discrepancy between the biophysical LF of the two data sources. It is likely, that a comparison between the content of a picture and *Flickr* tags is responsible for this mismatch. Even though *Flickr* tags have been used in other studies (e.g. Hollenstein and Purves (2010)), I assume that tags are better suited to extract the perception of the landscape instead of specific LF.

6.1.4. Detection of CES

Manual Detection

It is promising that the hotspots detected by the manual method were regulated by the other data set. While some locations with high densities in either source were regularised by the combined data set because the not apparent data source did not show any CES, other locations were enhanced because both data sources contributed to a hotspot. This effect is likely due to similar numbers of CES annotations that have been used for text (274) and picture data (340). It leads to the conclusion that a combination of text and picture data is possible. Additionally, the inclusion of more than one data source allows reaching a better coverage of the entire research area.

When focusing on specific CES subtypes, van Zanten et al. (2016a) compared different social media platforms and their ability to capture landscape values. Their limitation, that for example, estimated recreational values were not possible to be mapped at a local scale, cannot be found by the annotation approach. This supports the use of manual or automatic classifications instead of models which only take into account the locations of pictures instead of their content. Figure A.2 and Figure A.3 show KDE for each CES subtype. By comparing these maps with the KDE of all annotated CES of both data sets (Figure 5.4 (a)), hotspots can be explained. For example, the hotspot on the Marbachegg can be explained by a high amount of *Restaurant / Accommodation* annotations or Heiligkreuz contains a lot of *Healing Powers* and *Church* CES.

(a) *(b)*

Figure 6.1.: *Both pictures are from the same contributor (74710161@N02) with very similar content and two different IDs (49969426688 (a) and 9564860542 (b)).*

Some people contribute far more images than others (Muki Haklay, 2013) to social media platforms which may influence the results. To correct this issue, it would be an option to drop multiple annotations in pictures of the same user within a certain distance. Contributors who uploaded multiple pictures of the same CES have a high influence on the kernel density surfaces and just the visible proof of a CES does not reveal any information about how many people use a CES (Bieling and Plieninger, 2013). Some users may also regularly repost old content (e.g. "throwback") (Toivonen et al., 2019). This was also found in my data when the same contributor uploaded the same picture multiple times or two pictures with the same content but a slightly different angle or with an aesthetic filter which all resulted in different picture-IDs (Figure 6.1). This distorts the results. In contrast, I did not annotate a specific CES more than once even though it was mentioned several times in a tour description. Consequently, further research would need to evaluate and eventually reduce the effect of users who contributed multiple pictures of similar content at more or less the same locations.

Automatic Detection

Picture contributor clusters predominantly cover areas for which the UBE is famous. Cluster 1 and 5 are located in areas of protected moorlands (*Salwide* and part of *Glaubenberg*) and clusters 2 and 6 cover famous summits (*Schrattenfluh* and *Schimbrig*). Especially *Schrattenfluh* is promoted by the UBE because of its karst formations. In contrast, the text data revealed clusters which are located in urban areas of the UBE (clusters 4 (*Escholzmatt*), 5 (*Schüpfheim*) and 6 (*Entlebuch*)). This is similar to Long et al. (2021), where high population densities were often found close to CES.

The extracted text data clusters are characterized by a high amount of *Infrastructure* because of the mentioned restaurants or train stations. This indicates that these clusters might have been found because the hiking descriptions started or ended in those municipalities due to their accessibility by public transport. This needs to be considered during the annotation process of the text data and can easily be ignored during the process.

Table 6.1.: *Jaccard-Indexes between the union of the hotspots of the individual data sets and the hotspots of the of* Flickr *data, text data as well as the combined result.*

Hotspots	*Flickr*	Text	Combined
Union	0.68	0.44	0.64

Similar to the maps of the manually detected CES, both data sources influence the extraction of hotspots of the combined data set. The combined result reveals clusters that have been shown by either data source before but also hides clusters from both data types. The low Jaccard-Index between the two individual data sources (0.12) means that the overlap of these data sets is not very big. This supports the call of including more data sources when extracting CES by Tenkanen et al. (2017) and Oteros-Rozas et al. (2018). Besides the fact that each data source revealed hotspots that are not visible in the other data source, the different sizes and shapes are noteworthy. The Jaccard-scores between the combined data set and the individual data sets (picture (0.47) and tour description (0.43)) data show that neither of the data sources has a disproportional influence on the results of the combined data set, which is very important and desired when combining different types of data. A possible explanation of this outcome could be that the numbers of tour descriptions (75) and *Flickr* contributors (69) are very similar. The methodology of combining annotations from two sources and then extracting hotspots is better suited to combine two different data sets than other methods. For example, a union of the individual hotspots of *Flickr* and text data (Table 6.1) shows Jaccard-Indexes which are different from the results in Table 5.1. *Flickr* shows a very high value when compared to the *Union* data set (0.68), which is even higher than the *Combined* (0.64) and the *Text* (0.44) compared to the *Union*. This indicates that the picture data has a high influence on the union of two individual hotspot data sets and supports the idea of combining two different data sources before extracting hotspots by using the Local Moran's I.

This method of automatically extracting CES stands and falls with the assumption by van Zanten et al. (2016a) that landscape values increase as more people share information about that landscape on social media platforms. With the Local Moran's I, raster cells might be extracted and labeled as clusters, which do not contain many annotations but are surrounded by high values. Additionally, the high-value cells are not classified as cluster cells if their neighbors all have very low values.

Nevertheless, extracting clusters with a Local Moran's I on a raster data set which was built on pictures and geolocated annotations along linear features is questionable. Even though Richards and Friess's (2015) results revealed that social recreation photographs appear more clustered along trails. For other sites, they differed massively depending on the main attractions of these sites (e.g. animals and plants in a nature reserve). Consequently, some nature reserves are more diverse in terms of CES and others that are less diverse. Pictures of wildlife have to be taken opportunistically, which means that unlike landscape pictures from a location with a great view, pictures of wildlife usually are more distributed in space (Richards and Friess, 2015). Cultural heritage and spiritual services did not play a significant role in Richards and Friess (2015) because they do not require physical interaction between an environment and visitors and thus are less often photographed.

Main Messages

Research Question 1: *How do the results of locating CES differ between two methods applied on two sources of data?*

- The results of the manual CES detection can be used for analyses on every local scale since all CES annotated in either data source are mapped.
- The automatic detection of CES is very suitable to gain an overview of areas that are often photographed and/or mentioned in tour descriptions.
- Comparing the two methods to extract CES is difficult. The automatic CES detection returns polygons with different sizes (multiple cells), which may span over several kilometers making it difficult to draw conclusions for landscape planners and conservationists.
- The results of the combination of both data sources show that both individual data sets contribute to it, while none of the data sources is underrepresented. This is mainly because 75 tour descriptions and 69 *Flickr* contributors seem to have similar influences on the other data set, which would not be the case if the numbers differed much more.

6.1.5. Relations between CES and LF

The advantage of the picture data is that any threshold distance to capture the LF annotations around a CES already includes all annotations of a picture because they share the same coordinates. If a contributor took several pictures from a similar location with varying orientations, a panoramic view was included because of the threshold distance. This builds an advantage when compared to studies where only the content of one single picture is analyzed (e.g. Oteros-Rozas et al. (2018)). The use of 360° panorama pictures was discussed in Schirpke et al. (2016). On the contrary, including too many pictures from the same contributor leads to the overrepresentation of specific LF.

Not onl< the presence should also be looked at when investigating relationships between CES and LF, but also the absence and variability of LF. *Infrastructure*, for example, shows high-high values for *Recreational Facilities* as well as *Viewpoint* (Table 5.2). But also shows mid-high values for *Identity* and *Place Attachment*. Consequently, a lot of CES capture mid to high numbers of *Infrastructure*, which indicates that this LF might be an important part of all these

CES. *Infrastructure* is a very commonly annotated LF (Figure 5.2) for *Recreational Facilities* and *Viewpoint* in both data sets, which explains both high values (%LF and %CES). In contrast, *Identity* was not as often annotated and is not able to capture a high proportion of *Infrastructure* (mid %LF value) which explains the mid-high value for this combination. This is why, it is always important to look at the absolute numbers of annotations and also pay attention to CES with below mid-high, high-mid, and high-high values because the absence of *Infrastructure* might also be appreciated.

Besides involving absolute numbers, it is also important to have a closer look at each subtype of CES. Specific subtypes are also very biased because of their manifestation. Coming back to the example of *Recreational Facilities* and *Infrastructure*, a lot of facilities were annotated with these two subtypes, which results in a strong relationship between the CES subtype and the LF subtype just because *Infrastructure* is part of *Recreational Facilities*. The same effect can be found for *Restaurant / Accommodation*, which would not be called a restaurant without its infrastructure.

van Zanten et al. (2016a) hypothesized that topography, proximity to water bodies, and land cover patterns have a high influence on landscape values. In terms of water bodies, I can agree for some specific CES subtypes, such as *Identity* and *Viewpoint* (Table 5.2) or clusters 2 and 3 (Table 5.3). For example cluster 3 includes the Kneipp facilities of Schwandalp, which was mentioned in several tour descriptions but more often in pictures (Figure 5.5 and Figure 5.6). Plieninger et al. (2013) discovered that water bodies were often mentioned with recreational services, which can also be found in the form of *Lake* in Table 5.2. Other previous findings (Martínez Pastur et al., 2016; Oteros-Rozas et al., 2018) have indicated that social and spiritual CES show high correlations with anthropogenic landscapes. Since the UBE is an area that has been formed and modified by anthropogenic actions, this is an expected result (Appendix A.2).

Comparing the found relations between CES and LF to literature is difficult due to similar reasons as stated in Section 6.1.2. For example, Figueroa-Alfaro and Tang (2017) as well as van Zanten et al. (2016a) looked at the aesthetic values and their dependence on LF. Since aesthetic values are not an included CES subtype in this thesis, the results can not be compared directly. Moreover, local circumstances have a big influence on the results and conclusions are easier to draw from larger areas, which makes it even more difficult to relate to studies of other areas containing different landscapes.

Main Messages

Research Question 2: *How can different sources of data be used to better investigate the relation between CES and LF along hiking trails?*

- Both data sources are able to extract LF for the manually and automatically detected CES.
- Pictures are better suited to reveal high %CES values, the text data performed better at capturing high shares of a LF subtype (%LF).
- The combination of both data sets allows to better combine characteristics and reduce biases of the individual data sets which leads to more robust results.
- The comparison and evaluation of relations between CES and LF are very difficult. Mainly

because different subtypes of CES and LF were chosen due to different characteristics of other case study areas.

- This is the first study to have a closer look at one single.

6.2. Methodological Limitations

The chosen threshold distance to capture all relevant pictures and annotations from the tour descriptions which were close to a hiking trail (40 meters) can be changed. Assumed that all people, who uploaded pictures within this distance, were using hiking trails would be wrong. Especially in residential areas, a smaller buffer threshold could have been used to reduce this effect.

6.2.1. Choice of Types and Subtypes

The selected types and subtypes in the lists of LF (Table 4.3) and CES (Table 4.4) for the annotation process need to be discussed in more detail. Certainly, some elements of the list of LF could be changed. Derungs and Purves (2014) differentiated between *summit (peak)* and *mountain*. Since a summit is a geomorphological feature that catches the attention and lies on top of a mountain but might not always be visible due to the extent of the photograph, this differentiation is important, but the difference is small. I decided to use *Summit* and *Bedrock* instead, because of the same reason, but the *Bedrock* is rather general and includes more than *Mountain*. In the text data, a differentiation between a *summit* and a *mountain* (e.g. Brienzer Rothorn) is not possible, whereas the differentiation between *Bedrock* and *Summit* is much easier.

My choice of subtypes also shows some inconsistencies. For example, I decided not to include major geomorphological features (e.g. trenches, valleys, gullies) except the subtype *Summit*, which lead to the term *hügelige Landschaft* (hilly landscape) being only annotated as *Natural Landscape*. Therefore, it is dangerous to neglect geomorphological elements and only focus on LF, because geomorphology, predominantly the hilly landscape, forms an important part of the UBE (UNESCO Biosphere Entlebuch, 2021c).

Since CES are a very broad field and include many different services, it is difficult to include and categorize all of them. For example, some services are not permanent and have been covered in other studies (e.g. horse droppings as a sign of recreational activity or chairs from a party in Bieling and Plieninger (2013)). Bieling and Plieninger (2013) recorded a very high number of temporary signs of CES during their field walks. Whether temporary signs of CES also belong to CES is questionable because adding temporary features (e.g. footprints, waste) to CES would make the categorization much more difficult and entirely based on assumptions. One could argue that in a picture of *People*, walking on a hiking trail is also temporary, but their identification is much easier.

Moreover, it is easy to miss some CES. Some types of CES could also be the result of very long and complex relationships between humans and nature and could not be retrieved from the data sets used in this study. The symbolic meaning may even be impossible to determine by a very detailed content analysis (Albers and James, 1988; Oteros-Rozas et al., 2018). Interviews or stories (Bieling, 2014) could be used instead to capture these services. Also, different sources of data reveal different services or are in turn not suitable to detect some types of CES. For example, field walks in Bieling and Plieninger (2013) were found to not be suited to reveal spiritual ES because it is difficult to extract inspirational values from field walks. Furthermore, some types of recreational activities and their facilities are better suited to be represented on social media. Hiking can be regarded as a favorite activity to share impressions and experiences on social media, unlike rock climbing (Tenerelli et al., 2016) or surfing (Wood et al., 2013), which would lead to underrepresented activities and sites.

Including a sufficient amount of subtypes of LF and CES is crucial to extract diversified and meaningful conclusions from the results. However, including too many subtypes with only small differences is also not desirable. The number of subtypes also depends on the research area and the focus of the study. For example, van Zanten et al. (2016b) extracted the preference of visual landscapes from visitors and only categorized the pictures into four categories (livestock, diversity of agricultural land use, presence of green linear elements (hedgerows, tree lines), presence of point elements (single trees or groups of trees)), while leaving out other biophysical features (e.g. *Grass- and Moorland*, *Infrastructure*, *Wild Animal*).

6.2.2. Annotation Procedure

Besides the different types and subtypes of LF and CES, the procedure of annotating pictures and text data should also be reflected. Recognizing and interpreting are largely based on the individual reader and viewer. However, including local experts in the analysis could reduce this effect (Oteros-Rozas et al., 2018).

The annotation procedure would also affect the results. It could be that the framework of lists with many subcategories and the different rules for text data (e.g. additional CES and types of LF to annotate) and photographs (e.g. *Fore-/Background*) makes the annotation process too tedious. The complexity and ambiguity of an annotation procedure are rarely discussed in literature. It is unclear whether an extensive procedure, as in this thesis or Oteros-Rozas et al. (2018), leads to better results than a procedure which is kept more simple, as in Richards and Friess (2015) or van Zanten et al. (2016b). Richards and Friess (2015) stated that a simplified framework leads to less variation in classifying photographs, is quicker, easier to understand, and ultimately leads to other, but also high-quality results.

By manually geotagging LF and CES for the text data, it was possible to overcome biases that would have been a problem with other methods (e.g. keyword-based text analyses or gazetteers) (Ghermandi and Sinclair, 2019). Nevertheless, the annotation procedure for text data also revealed difficulties, such as how to handle expressions that could not be clearly assigned to a LF subtype (e.g. *Baumgrenze* (timberline), *Alpwiese* (alpine meadow) or *Emmentaler Hügel* (Emmen valley hills)).

Ghermandi and Sinclair (2019) also point out that manual content analysis of photographs, as in this study or by Oteros-Rozas et al. (2018) and Martínez Pastur et al. (2016), introduces a researcher bias. The number of photographs used for this study is not very different from the number of *Flickr* pictures Oteros-Rozas et al. (2018) used in their study. I used a similar amount of pictures for one research area, as Oteros-Rozas et al. (2018) used for five different research areas. In contrast, other studies have used a higher number of pictures (e.g. Gliozzo et al. (2016), Heikinheimo et al. (2020)) for one research area.

Specific events and local conditions may influence observed patterns (Heikinheimo et al., 2020). This could also be observed in my data set where 230 pictures of a horse-riding event were uploaded. A big limitation is that social media does not reveal which places were not used and text data only mentioned important waypoints of places to visit. The motivation of mentioning places or taking pictures is very different between my two data sources.

I only took into account pictures, which are located inside the UBE. Alternatively, a buffer could be added to the boundaries of the UBE, similar to Väisänen et al. (2021) to also include possible pictures outside the UBE whose content (partly) lies in the UBE. This also has the disadvantage that irrelevant pictures could be included from CES outside of the research area.

Regarding the annotation process for picture data, even though specific LF can easily be identified in separate pictures (e.g. *Waterfall*, *Lake* and *River/Creek*), a distinction in the same picture can be tricky (Figure 6.2). But also specific LF are very difficult to be annotated, as *Viewpoint* which was mainly defined by the personal interpretation of the picture.

(a) *(b)*

Figure 6.2.: *These two pictures (49969936331 (a) and 49969417418 (b)) by user 74710161@N02 help to demonstrate difficulties when annotating LF. In picture (a) it is unclear whether it is a* River/Creek *or* Waterfall *running down the bedrock and in picture (b) it is unclear whether the waterfall ends in a* Lake *or a* River/Creek.

6.3. Data Limitations

6.3.1. Choice of Data and general Limitations

Neither the *Flickr* data nor the text data have originally been generated for research purposes and have passively been contributed (Heikinheimo et al., 2020). The fact that social media data represents unfiltered content unlike citizen science projects or surveys where recollection bias is an often reported problem, is a strong advantage of this data source (Dunkel, 2015; Ghermandi and Sinclair, 2019). This reduces biases found in questionnaires or surveys (e.g. Wartmann and Purves (2018)) where the contributors knew that their data will be used for research purposes. Both sources also have other advantages such as being more voluminous and not being limited in duration or extent (Levin et al., 2017) compared to actively contributed data. Nevertheless, while *Flickr* users are most probably unaware of their contribution to research, the authors of text data know that their texts are read and they write the text intending to promote going on hiking trails and visiting the UBE.

The used data sets have to be representative of the visitors in terms of quantity to be able to draw reliable conclusions. Because it is not investigated in this thesis whether the *Flickr* data and the tour descriptions are correlating with visitor statistics, it is assumed that these data and their derived number of contributors and number of tour descriptions per raster cell are representative. For example, Tenkanen et al. (2017)'s study mainly focused on finding relationships between social media activity and visitor statistics. Based on their conclusion that *Instagram* performs best and outperforms *Flickr* in representing the visitor statistics (Tenkanen et al., 2017), the study would likely have performed differently in terms of representing the numbers of visitors, if *Instagram* data was available. But this is just a hypothesis because Tenkanen et al. (2017) also showed that social media activity exhibits weaker correlations in less frequently visited parks.

Text Data

For the text data, it would have been very valuable if I had found text data about the perception of landscape as in Koblet and Purves (2020), the hiking blogs in Wartmann et al. (2018) or short stories as in Bieling (2014). In this thesis, the purpose and meaning of the tour descriptions from *maps.luzern.com* are clear and forms a substantial bias of this data set. The tour descriptions put more emphasis on the directions of the activities and serve to stay on the right path. This is in contrast to tour reports, which also include personal experiences. However, most experience reports also include elements of tour descriptions.

Moreover, the descriptions are not written by private individuals, as on other platforms, but by local tourism institutions (e.g. *Sörenberg Flühli Tourism* or *UNESCO Biosphere Entlebuch*). This leads to the advantage that most of the tour descriptions are uniformly structured and of similar extents. But this also means that it is unclear whether different people from the tourism institutions contributed to the tour description and published them.

Picture Data

Taking into account information regarding platform-specific and population biases is very important (Ruths and Pfeffer, 2014). The representativeness of specific platforms in national parks can be questioned because their users are not representative of the whole population of visitors due to age and content which is shared (Hausmann et al., 2017b,a; Heikinheimo et al., 2017). Younger people (Heikinheimo et al., 2017) and women (Hausmann et al., 2017a) are more likely to post photographs on social media while visiting national parks (Heikinheimo et al., 2017). But this behavior is platform-specific. For example, *Instagram* users are younger than other social media users (Heikinheimo et al., 2017). Therefore, *Instagram* data is better suited to represent opinions and activities which are more popular among this demographic group (e.g. eco-tourism) (Ghermandi and Sinclair, 2019). *Instagram* users are also more likely to share social aspects (Hausmann et al., 2017a; Toivonen et al., 2019). Correcting thematic and demographic biases is very difficult because researchers usually do not have access to the necessary information (Oteros-Rozas et al., 2018). The same problem applies to *Flickr* users, who focus mainly on biodiversity- and nature-related content (Di Minin et al., 2015; Hausmann et al., 2017a). This is why *Flickr* has been more often used in environmental sciences (Tenkanen et al., 2017), which was also suggested by Ghermandi and Sinclair (2019). They stated that social media platforms should be selected depending on the type of research question. Therefore, due to its credibility in past studies and the problems of other platforms (Section 4.1.2), the choice of *Flickr* is justified. Additionally, this platform is said to capture culturally relevant events (van Zanten et al., 2016a) and represent content from users with diverse motivations (Nov et al., 2010). A limitation for *Flickr* is that a relatively small amount of very active users produce a very big amount of data (Li et al., 2013).

Besides thematic and demographic biases, social media has other limitations. Every photograph transfers a perception or meaning. Pictures are seen as a good source to study the main motivation why they were taken (Richards and Friess, 2015; Zoderer et al., 2016). Each message, which the contributors are trying to convey is also a bias that needs to be taken into account (see Section 2.6). People have different motivations to take pictures, which may depend on the level of education or affinity towards specific topics (van Zanten et al., 2016b) and the people's perception of services provided by the landscape may be influenced by socio-demographic factors as well as culture (López-Santiago et al., 2014; Plieninger et al., 2013; Zoderer et al., 2016). Some people take pictures to record negative experiences, others to record positive attributes of the environment (Dorwart et al., 2009). Reinecke and Trepte (2014) found that social media users tend to share positive and likable content. This causes a positivity bias because negative experiences are neglected (Reinecke and Trepte, 2014) and not everything that has been experienced is posted on social media (Di Minin et al., 2015). Social media is rather used as a tool to entertain than to document (Toivonen et al., 2019). But research has shown that social media represents preferences and activities of visitors in national parks (Hausmann et al., 2017a; Heikinheimo et al., 2017), which dampens this limitation's effect. Additionally, social media users who are traveling are more likely to share their experiences and expressions than those staying at home or living in the area (Becken et al., 2017). Nevertheless, the study by Ghermandi et al. (2020) has shown that also local people share data on social media, revealing that locals show a higher appreciation of aesthetic values than tourists. This is why cultural background and the home location of contributors influence what is photographed (Frey, 2020).

Furthermore, some objects are more charismatic or popular than others (Tenkanen et al., 2017) or easier to be photographed. For example, photographing animals often requires specialized equipment (Ghermandi and Sinclair, 2019). This explains the shortcomings of some subtypes (e.g. *Funghi /Flower*, *Shrub*) of LF and CES because the chance that they are pictured is much smaller because of their size when compared to *Grass- and Moorland*, which is very ubiquitous, especially in landscape pictures. If these smaller subtypes are pictured, they often are part of the *foreground* (Figure 5.2(b)), which underlines this statement. Additionally, some users wait until there are no people on the picture or take it at an angle where something specific is in the focus, which disturbs the perception of the landscape. Gliozzo et al. (2016) assumed that multiple decisions have to be taken for a picture to land on social media (going to a place, taking a picture, share it). Some users take these decisions much quicker and more often than others. With newer technologies and internet connection in remote areas, the decisions to take and upload a picture to a platform might be taken much quicker (Gliozzo et al., 2016).

The mismatch between a picture's location and its content is an often reported problem when studying geotagged data in landscape studies. This can be because of the distance between the photographer and what is being photographed (Gliozzo et al., 2016; Guerrero et al., 2016). Especially in mountainous terrains, there is a big mismatch (Dunkel, 2015; Oteros-Rozas et al., 2018) because of topography, which is only apparent in the picture data. Viewsheds from the pictures' location could be used to partly overcome this problem (Yoshimura and Hiura, 2017), where only the orientation would have to be defined.

6.3.2. Spatial Limitations

In the tour descriptions, different types of activities cover different spatial distances and consequently contain different densities of annotations. For example, a themed trail is rather short in length, but because its description is much more detailed due to limited distance, it produces a higher density of annotations when compared to a bike tour (Figure 6.3).

Flickr, as well as other social media platforms, have spatial limitations. First of all, not each post on social media platforms is geotagged (Heikinheimo et al., 2017) because geotagging is only a voluntary functionality when uploading photographs to *Flickr* and can be changed afterward (Toivonen et al., 2019). Spatial accuracy may vary depending on GPS accuracy, platform, and user input (place names in text) (Heikinheimo et al., 2020) and is expected to be lower in remote regions (Heikinheimo et al., 2017), but still high enough because of GPS equipped camera devices (Toivonen et al., 2019). Despite this, Huang et al. (2013) mentioned that the accuracy of the actual geolocation of a picture might be very poor due to missing mechanisms of social media platforms which would ensure quality. In addition, geographic biases due to mobile phone coverage need to be addressed (Di Minin et al., 2015). The mobile coverage by Swisscom (2021), one of the major telecommunication providers in Switzerland, shows some areas without signal in the research area.

It should not be neglected that the use of social media platforms is highly dependent on the location and the purpose because some platforms are used in areas where others are not apparent (Gliozzo et al., 2016). Social media data is much more reliable in areas with large populations or famous tourist attractions, unlike remote locations with rather small attractions (Lawu et al., 2021). Since *Flickr* is often used in Central and Western Europe (van Zanten et al., 2016a), this

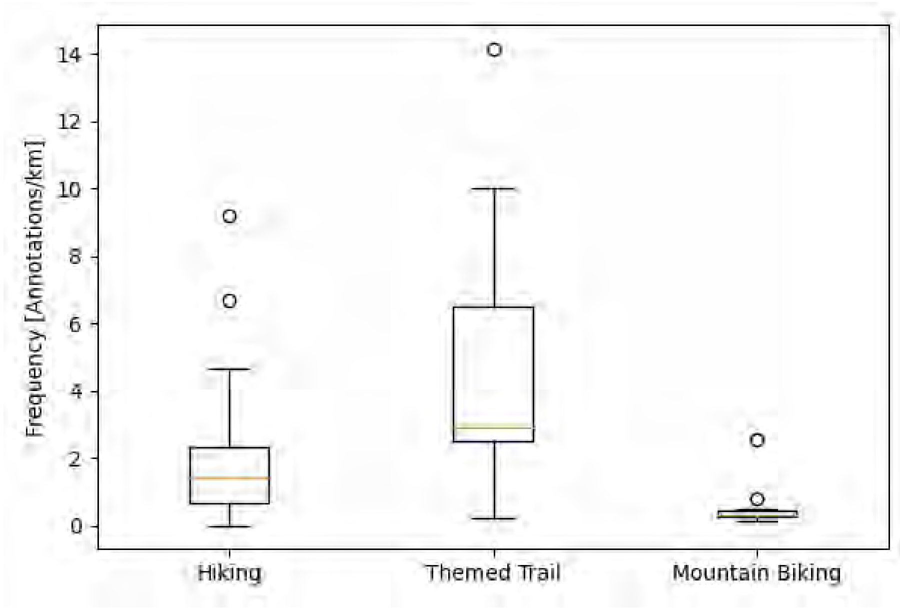


Figure 6.3.: *Annotations per km for the three most common types of activities (see Table 4.1).*

should not be a limitation despite the rural location of the UBE.

6.3.3. Temporal Limitations

Due to the aggregation of *Flickr* data from multiple years, it would also have been insightful to use data from a platform that offers enough data for a single year analysis, because visitation patterns might get hidden (Tenkanen et al., 2017). For example, a CES could be regarded as important in the results, because of one specific year while in all the other years the CES would not have been detected. Using data from multiple years also has an advantage. Photographs from people who did not upload their pictures immediately are also included (Heikinheimo et al., 2020).

An additional bias is created by the fact that only pictures between May and October were used for the analysis. Hiking trails may also be used in other months and might have lead to different results since activities can be expected to take place at lower elevations than during summer. This is why the temporal limitation also leads to a spatial limitation. This has the effect that the results of this thesis can only be regarded as representative for the months between May and October, whereas during winter months other CES might be detected.

6.4. Outlook

6.4.1. Themantic Outlook

This thesis tests two different approaches to locate CES and look for relationships between them and LF. Because there is a big variety of CES, it would be beneficial to look at the perception of landscape from a broader view, for example by assessing the aesthetical value. Especially since the quantity and quality of aesthetical values of natural landscapes have decreased (MEA, 2005).

A limitation is the representativeness of this thesis since no socio-economic data exists and the goal of this project was not to study individual users (Guerrero et al., 2016). If socio-demographic factors and the origin of contributors, participants, or visitors are not assessed, conclusions drawn from studies may be biased towards specific groups and have influences on management strategies. For example, people above 56 years of age more often appreciate cultural heritage values (Zoderer et al., 2016). This is why a differentiation of the perception of landscape between different groups would be beneficial. Additionally, even though Knaus (2012) showed that the majority of visitors do not live within the UBE (Figure 6.4), it is still unclear whether tourists or local people upload pictures to social media. Therefore, it is also questionable whether the detected CES are important to visitors or residents, or both. It is important that CES are not only created and maintained for tourists (e.g. hiking trails) but also other activities and CES which address residents which in turn create a deeper connection to preservable places (Bieling, 2014; Ghermandi et al., 2020).

Figure 6.4.: *Percentages and origin of visitors of the UBE aggregated to the first two zip code numbers from Knaus (2012).*

Most studies have not paid attention to different perceptions of different groups of people in a landscape, like Oteros-Rozas et al. (2018) who assume that the connection between LF and CES are culturally universal. In contrast, van Zanten et al. (2016a) mentions an important point, why cross-cultural comparisons regarding the perception of landscape across the globe are inevitable: Landscape values are treated the same way and the same processes transform the landscapes (e.g. urbanization, population growth, land-use changes, and climate change (Schirpke et al., 2020; van Zanten et al., 2016a), agricultural intensification and abandonment), but since different cultures perceive landscape values differently, a cross-cultural analysis of landscape values is necessary (van Zanten et al., 2016a). Long et al. (2021) agrees by stating that different communities have multiple expectations of something to be aesthetically pleasing or attractive.

A person-based approach on social media data could help classify users into different groups (residents and tourists, single-time users, and multiple-times users, or day-trippers and overnight-stay-tourists (Girardin et al., 2008; Schirpke et al., 2018; Toivonen et al., 2019; Väisänen et al., 2021; van Zanten et al., 2016b). The short stories in Bieling (2014) are an excellent example of extracting CES from residents. But often it is unclear who the participants or contributors really are. As it is only clear in a few papers, who the target audience is, as, tourists in Hausmann et al. (2017a) and Zoderer et al. (2016). To obtain a better understanding of social media data, identifying the home location of users has been done with very high accuracies (Ghermandi, 2018; Frey, 2020; Sinclair et al., 2020). Consequently, it would be interesting to investigate whether the perception of landscape differs by the provenance of each contributor.

6.4.2. Methodological Outlook

Future work could tackle the manually done aspects of this thesis. Especially the annotation process and the geolocation of annotations for text data could be automated. If a task is automated, repetition is much easier and analyses can be run multiple times and detection of landscape changes can be monitored (e.g. Koblet and Purves (2020)).

Manually analyzing the content of social media can be very tedious and laborious. Automatic data classification methods have been developed (Schwartz and Ungar, 2015). A recent study by Väisänen et al. (2021) explored and evaluated three computer-vision methods which analyzed social media photographs to extract information on the interaction between humans and nature. Still, they only classified pictures into four typical regions. Identifying individual LF can be done by using the online machine learning algorithm Google Cloud Vision, as in Richards et al. (2018). Lee et al. (2019) and Frey (2020).

Automated toponym recognition and resolution could save a lot of manual work and time (Amitay et al., 2004). However, in semi-formal texts in the German language where vernacular and idiosyncratic spellings of place names are very common (e.g. Schrattenfluh or Schrattenflue) (Augenstein et al., 2017), automated toponym recognition and resolution is very challenging (Wartmann et al., 2018). Maybe free lists should also be taken into account similarly to Wartmann et al. (2018) since they also activate the memory retrieval process which allows including people's feelings and meanings regarding the landscape (Wartmann et al., 2015).

Based on the approach by Cooper and Gregory (2011), future work could think about extracting LF with the help of land use data and the toponyms which are apparent in hiking descriptions. Even though locating toponyms is very successful (Cooper and Gregory, 2011; Derungs and Purves, 2014; Purves and Derungs, 2015) and the extraction of LF has been done previously (Derungs and Purves, 2014), identifying the LF's exact geolocation would pose a big obstacle since sufficient accuracy is required when geolocating LF along hiking trails.

.

Conclusion

In this thesis, I applied two methods to extract and map CES. The results were used to investigate the relations of CES with LF. Previous research has not focused on an in-depth analysis of one study area and has proposed using different types of data to not rely on biases represented by single data sets. Therefore, I applied these methods to two different sources of data, picture data and unstructured text data, as well as their combination in the UBE. The first method which used manually annotated CES was compared to the automatic method of extracting hotspots by applying a Local Moran's I on the number of contributors for picture data and the number of tour descriptions. Then, I captured all LF which were within 20 meters of each CES subtype as well as within each hotspot which resulted from the automatic CES detection. This approach was applied on both individual data sets as well as the combination of both data sources.

The results showed that the automatic detection of CES is better suited for large-scale areas and may help decision-makers to get an overview of culturally important areas. The manually annotated data helped to identify CES on lower spatial scales, but are strongly influenced by contributors who take several pictures at similar locations.

For both aspects, the detection of CES and the relations between them and LF, the combination of text data and picture data introduced new aspects and contributed to new findings. The text data revealed CES and hotspots which were not detected by the picture data and were more evenly distributed in the research area, which could also be seen by the combined data set. In addition, including text data influenced the relations found between CES and LF by reducing the effect of the picture data and contributing aspects which were neglected by photographs.

Consequently, this thesis contributes to the following aspects:

- Two different methods to detect and map CES were compared.
- Text data was manually geolocated and used to extract CES.
- The combination of two different data sets, text and picture data were also used to detect

and map CES.

- The individual text and picture data as well as the combined data set was used to investigate the relations between CES and LF in one single research area.

Besides its promising results, drawing conclusions from social media and text data should be done carefully. The data does reflect neither the opinion nor the behavior of the complete population, which is why further research is needed to look at socio-demographic aspects of the data sources to create representative results.

Bibliography

Agnoletti, M. and Emanueli, F. (2016). Biocultural Diversity and Landscape in Europe: Framing the Issue. In Agnoletti, M. and Emanueli, F., editors, *Biocultural Diversity in Europe. Environmental History*. Springer.

Albers, P. C. and James, W. R. (1988). Travel photography. A methodological approach. *Annals of Tourism Research*, 15:134–158.

Ambrose-oji, B. and Pagella, T. (2012). Spatial Analysis and Prioritisation of Cultural Ecosystem Services : A Review of Methods. Technical Report June.

Amerson, K., Rose, J., Lepp, A., and Dustin, D. (2020). Time on the trail, smartphone use, and place attachment among Pacific Crest Trail thru-hikers. *Journal of Leisure Research*, 51(3):308–324.

Amitay, E., Har'El, N., Sivan, R., and Soffer, A. (2004). Web-a-where. In *27th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 273–280.

Anselin, L. (1995). Local Indicators of Spatial Association - LISA. 27(2).

Antoniou, V., Morley, J., and Haklay, M. (2010). Web 2.0 Geotagged Photos: Assessing the Spatial Dimension of the Phenomenon. *Geomatica*, 64(1):99–110.

Augenstein, I., Derczynski, L., and Bontcheva, K. (2017). Generalisation in named entity recognition: A quantitative analysis. *Computer Speech and Language*, 44:61–83.

Bagstad, K. J., Johnson, G. W., Voigt, B., and Villa, F. (2013). Spatial dynamics of ecosystem service flows: A comprehensive approach to quantifying actual services. *Ecosystem Services*, 4:117–125.

Bateman, I. J., Harwood, A. R., Mace, G. M., Watson, R. T., Abson, D. J., Andrews, B., Binner, A., Crowe, A., Day, B. H., Dugdale, S., Fezzi, C., Foden, J., Hadley, D., Haines-Young, R., Hulme, M., Kontoleon, A., Lovett, A. A., Munday, P., Pascual, U., Paterson, J., Perino, G., Sen, A., Siriwardena, G., Van Soest, D., and Termansen, M. (2013). Bringing ecosystem services into economic decision-making: Land use in the United Kingdom. *Science*, 341(6141):45–50.

Becken, S., Stantic, B., Chen, J., Alaei, A. R., and Connolly, R. M. (2017). Monitoring the environment and human sentiment on the Great Barrier Reef: Assessing the potential of collective sensing. *Journal of Environmental Management*, 203:87–97.

Bieling, C. (2014). Cultural ecosystem services as revealed through short stories from residents of the Swabian Alb (Germany). *Ecosystem Services*, 8:207–215.

Bieling, C. and Plieninger, T. (2013). Recording Manifestations of Cultural Ecosystem Services in the Landscape. *Landscape Research*, 38(5):649–667.

Bieling, C., Plieninger, T., Pirker, H., and Vogl, C. R. (2014). Linkages between landscapes and human well-being: An empirical exploration with short interviews. *Ecological Economics*, 105:19–30.

Brooker, P., Barnett, J., Cribbin, T., and Sharma, S. (2016). Have We Even Solved the First 'Big Data Challenge?' Practical Issues Concerning Data Collection and Visual Representation for Social Media Analytics. In *Digital Methods for Social Science: An Interdisciplinary Guide to Research Innovation*, pages 34–50. Palgrave Macmillan UK, London.

Brown, G. and Raymond, C. (2007). The relationship between place attachment and landscape values: Toward mapping place attachment. *Applied Geography*, 27(2):89–111.

Bundesamt für Umwelt (BAFU) (2017). Das Bundesinventar der Landschaften und Naturdenkmäler (BLN). https://www.bafu.admin.ch/bafu/de/home/themen/landschaft/fachinformationen/landschaftsqualitaet-erhalten-und-entwickeln/landschaften-von-nationaler-bedeutung/bundesinventar-der-landschaften-und-naturdenkmaeler-von-national.html. Last accessed on May 18, 2021.

Byers, B. A., Cunliffe, R. N., and Hudak, A. T. (2001). Linking the conservation of culture and nature: A case study of sacred forests in Zimbabwe. *Human Ecology*, 29(2):187–218.

Callau, A. , Albert, M. Y. P., Rota, J. J., and Giné, D. S. (2019). Landscape characterization using photographs from crowdsourced platforms: Content analysis of social media photographs. *Open Geosciences*, 11(1):558–571.

Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., Narwani, A., MacE, G. M., Tilman, D., Wardle, D. A., Kinzig, A. P., Daily, G. C., Loreau, M., Grace, J. B., Larigauderie, A., Srivastava, D. S., and Naeem, S. (2012). Biodiversity loss and its impact on humanity. *Nature*, 486:59–67.

Casalegno, S., Inger, R., DeSilvey, C., and Gaston, K. J. (2013). Spatial Covariance between Aesthetic Value & Other Ecosystem Services. *PLoS ONE*, 8(6).

Church, A., Haines-Young, R., Mourato, S., Tratalos, J., Stapleton, L., Willis, C., Coates, P., Gibbons, S., Leyshon, C., Potschin, M., Ravenscroft, N., Sanchis-Guarner, R., Winter, M., and Kenter, J. (2014). UK National Ecosystem Assessment Follow-on. Work package report 5: Cultural ecosystem services and indicators.

Cooper, D. (2019). Digital Re-Enchantment: Place Writing, the Smartphone and Social Media. *Literary Geographies*, 5(1):90–107.

Cooper, D. and Gregory, I. N. (2011). Mapping the English Lake District: A literary GIS. *Transactions of the Institute of British Geographers*, 36(1):89–108.

Costanza, R., D'Arge, R., de Groot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., O'Neill, R. V., Paruelo, J., Raskin, R. G., Sutton, P., and van den Belt, M. (1997). The value of the world's ecosystem services and natural capital. *Nature*, 387(6630):253–260.

Daniel, T. C., Muhar, A., Arnberger, A., Aznar, O., Boyd, J. W., Chan, K. M., Costanza, R., Elmqvist, T., Flint, C. G., Gobster, P. H., Grêt-Regamey, A., Lave, R., Muhar, S., Penker, M., Ribe, R. G., Schauppenlehner, T., Sikor, T., Soloviy, I., Spierenburg, M., Taczanowska, K., Tam, J., and Von Der Dunk, A. (2012). Contributions of cultural services to the ecosystem services agenda. *Proceedings of the National Academy of Sciences of the United States of America*, 109(23):8812–8819.

de Groot, R. S., Alkemade, R., Braat, L., Hein, L., and Willemen, L. (2010). Challenges in integrating the concept of ecosystem services and values in landscape planning, management and decision making. *Ecological Complexity*, 7(3):260–272.

de Kunder, Maurice (2021). The size of the World Wide Web. https://www.worldwidewebsize.com/. Last accessed on May 13, 2021.

de Smith, M. J., Goodchild, M. F., and Longley, P. A. (2018). *Geospatial Analysis: A Comprehensive Guide to Principles, Techniques and Software Tools*. Number 2. Troubador, Kibworth Beauchamp, UK, 6th edition.

Depellegrin, D., Blažauskas, N., and Vigl, L. E. (2012). Aesthetic value characterization of landscapes in coastal zones. *Ocean: Past, Present and Future - 2012 IEEE/OES Baltic International Symposium, BALTIC 2012*.

Derungs, C. and Purves, R. S. (2014). From text to landscape: Locating, identifying and mapping the use of landscape features in a Swiss Alpine corpus. *International Journal of Geographical Information Science*, 28(6):1272–1293.

Derungs, C. and Purves, R. S. (2016). Characterising landscape variation through spatial folksonomies. *Applied Geography*, 75:60–70.

Di Minin, E., Tenkanen, H., and Toivonen, T. (2015). Prospects and challenges for social media data in conservation science. *Frontiers in Environmental Science*, 3.

Díaz, S., Demissew, S., Carabias, J., Joly, C., Lonsdale, M., Ash, N., Larigauderie, A., Adhikari, J. R., Arico, S., Báldi, A., Bartuska, A., Baste, I. A., Bilgin, A., Brondizio, E., Chan, K. M., Figueroa, V. E., Duraiappah, A., Fischer, M., Hill, R., Koetz, T., Leadley, P., Lyver, P., Mace, G. M., Martin-Lopez, B., Okumura, M., Pacheco, D., Pascual, U., Pérez, E. S., Reyers, B., Roth, E., Saito, O., Scholes, R. J., Sharma, N., Tallis, H., Thaman, R., Watson, R., Yahara, T., Hamid, Z. A., Akosim, C., Al-Hafedh, Y., Allahverdiyev, R., Amankwah, E., Asah, T. S., Asfaw, Z., Bartus, G., Brooks, A. L., Caillaux, J., Dalle, G., Darnaedi, D., Driver, A., Erpul, G., Escobar-Eyzaguirre, P., Failler, P., Fouda, A. M. M., Fu, B., Gundimeda, H., Hashimoto, S., Homer, F., Lavorel, S., Lichtenstein, G., Mala, W. A., Mandivenyi, W., Matczak, P., Mbizvo, C., Mehrdadi, M., Metzger, J. P., Mikissa, J. B., Moller, H., Mooney, H. A., Mumby, P., Nagendra, H., Nesshover, C., Oteng-Yeboah, A. A., Pataki, G., Roué, M., Rubis, J., Schultz, M., Smith, P., Sumaila, R., Takeuchi, K., Thomas, S., Verma, M., Yeo-Chang, Y., and Zlatanova, D. (2015). The IPBES Conceptual Framework - connecting nature and people. *Current Opinion in Environmental Sustainability*, 14:1–16.

Donahue, M. L., Keeler, B. L., Wood, S. A., Fisher, D. M., Hamstead, Z. A., and McPhearson, T. (2018). Using social media to understand drivers of urban park visitation in the Twin Cities, MN. *Landscape and Urban Planning*, 175:1–10.

Dorwart, C. E., Moore, R. L., and Leung, Y. F. (2009). Visitors' perceptions of a trail environment and effects on experiences: A model for nature-based recreation experiences. *Leisure Sciences*, 32(1):33–54.

Dunkel, A. (2015). Visualizing the perceived environment using crowdsourced photo geodata. *Landscape and Urban Planning*, 142:173–186.

Edwardes, A. J., Purves, R. S., Bircher, S., and Matyas, C. (2007). Concept ontology experimental report. Technical report.

ELC (2000). European Landscape Convention. In *European Treaty Serices*, number 176, page 26.

Figueroa-Alfaro, R. W. and Tang, Z. (2017). Evaluating the aesthetic value of cultural ecosystem services by mapping geo-tagged photographs from social media data on Panoramio and Flickr. *Journal of Environmental Planning and Management*, 60(2):266–281.

Flaticon (2021). Icons. https://www.flaticon.com/. Last accessed on August 10, 2021.

Foden, W. B., Butchart, S. H., Stuart, S. N., Vié, J. C., Akçakaya, H. R., Angulo, A., DeVantier, L. M., Gutsche, A., Turak, E., Cao, L., Donner, S. D., Katariya, V., Bernard, R., Holland, R. A., Hughes, A. F., O'Hanlon, S. E., Garnett, S. T., Şekercioğlu, H., and Mace, G. M. (2013). Identifying the World's Most Climate Change Vulnerable Species: A Systematic Trait-Based Assessment of all Birds, Amphibians and Corals. *PLoS ONE*, 8(6).

Fondren, K. M. and Brinkman, R. (2019). A Comparison of Hiking Communities on the Appalachian and Pacific Crest Trails. *Leisure Sciences*.

Frey, T. (2020). The Cultural Eye. Exploring Cultural Differences in Landscape Preference based on-Flickr Images. *MSc Thesis, Department of Geography, University of Zurich*.

Ghermandi, A. (2016). Analysis of intensity and spatial patterns of public use in natural treatment systems using geotagged photos from social media. *Water Research*, 105:297–304.

Ghermandi, A. (2018). Integrating social media analysis and revealed preference methods to value the recreation services of ecologically engineered wetlands. *Ecosystem Services*, 31:351–357.

Ghermandi, A., Camacho-Valdez, V., and Trejo-Espinosa, H. (2020). Social media-based analysis of cultural ecosystem services and heritage tourism in a coastal region of Mexico. *Tourism Management*, 77.

Ghermandi, A., Nunes, P. A. L. D., Portela, R., Rao, N., and Teelucksingh, S. S. (2009). Recreational, Cultural and Aesthetic Services from Estuarine and Coastal Ecosystems. *Fondazione Eni Enrico Mattei*.

Ghermandi, A. and Sinclair, M. (2019). Passive crowdsourcing of social media in environmental research: A systematic map. *Global Environmental Change*, 55:36–47.

Girardin, F., Blat, J., Calabrese, F., Dal Fiore, F., and Ratti, C. (2008). Digital footprinting: Uncovering tourists with user-generated content. *IEEE Pervasive Computing*, 7(4):36–44.

Gliozzo, G., Pettorelli, N., and Muki Haklay, M. (2016). Using crowdsourced imagery to detect cultural ecosystem services: A case study in South Wales, UK. *Ecology and Society*, 21(3).

Gosal, A. S. and Ziv, G. (2020). Landscape aesthetics: Spatial modelling and mapping using social media images and machine learning. *Ecological Indicators*, 117.

Guerrero, P., Møller, M. S., Olafsson, A. S., and Snizek, B. (2016). Revealing cultural ecosystem services through instagram images: The potential of social media volunteered geographic information for urban green infrastructure planning and governance. *Urban Planning*, 1(2):1–17.

Guo, Z., Zhang, L., and Li, Y. (2010). Increased dependence of humans on ecosystem services and biodiversity. *PLoS ONE*, 5(10).

Hannah, L., Midgley, G., Andelman, S., Araújo, M., Hughes, G., Martinez-Meyer, E., Pearson, R., and Williams, P. (2007). Protected area needs in a changing climate. *Frontiers in Ecology and the Environment*, 5(3):131–138.

Hausmann, A., Toivonen, T., Heikinheimo, V., Tenkanen, H., Slotow, R., and Minin, E. D. (2017a). Social media reveal that charismatic species are not the main attractor of ecotourists to sub-Saharan protected areas. *Scientific Reports*, 7(763).

Hausmann, A., Toivonen, T., Slotow, R., Tenkanen, H., Moilanen, A., Heikinheimo, V., and Di Minin, E. (2017b). Social Media Data Can Be Used to Understand Tourists' Preferences for Nature-Based Experiences in Protected Areas. *Conservation Letters*, 11(1):1–10.

Heikinheimo, V., Minin, E. D., Tenkanen, H., Hausmann, A., Erkkonen, J., and Toivonen, T. (2017). User-generated geographic information for visitor monitoring in a national park: A comparison of social media data and visitor survey. *ISPRS International Journal of Geo-Information*, 6(3).

Heikinheimo, V., Tenkanen, H., Bergroth, C., Järv, O., Hiippala, T., and Toivonen, T. (2020). Understanding the use of urban green spaces from user-generated geographic information. *Landscape and Urban Planning*, 201.

Heikinheimo, V., Tenkanen, H., and others (2018). Automated social media content analysis from urban green areas–Case Helsinki. *Spatial Big Data*, pages 40–43.

Hill, E., Gómez, E., Goldenberg, M., and Fellows, S. (2014). Appalachian and Pacific Crest Trail Hikers; a Comparison of benefits and motivations. *Journal of Unconventional Parks, Tourism & Recreation Research*, 5(1):9–17.

Hill, G. W. and Courtney, P. R. (2006). Demand analysis projections for recreational visits to countryside woodlands in Great Britain. *Forestry*, 79(2):185–200.

Hollenstein, L. and Purves, R. S. (2010). Exploring place through user-generated content: Using Flickr tags to describe city cores. *Journal of Spatial Information Science*, 1(2010):21–48.

Hu, Y. and Wang, R. Q. (2020). Understanding the removal of precise geotagging in tweets. *Nature Human Behaviour*.

Huang, H., Gartner, G., and Turdean, T. (2013). Social media data as a source for studying people's perception and knowledge of environments. *Mitteilungen der Osterreichischen Geographischen Gesellschaft*, 155:291–302.

Jaccard, P. (1912). The Distribution of the Flora in the Alpine Zone. *New Phytologist*, 11(2):37–50.

Jenkins, A., Croitoru, A., Crooks, A. T., and Stefanidis, A. (2016). Crowdsourcing a collective sense of place. *PLoS ONE*, 11(4).

Kienast, F., Frick, J., van Strien, M. J., and Hunziker, M. (2015). The Swiss Landscape Monitoring Program - A comprehensive indicator set to measure landscape change. *Ecological Modelling*, 295:136–150.

Kirchhoff, T. (2012). Pivotal cultural values of nature cannot be integrated into the ecosystem services framework. In *Proceedings of the National Academy of Sciences of the United States of America*, volume 109.

Kliot, N. and Collins-Kreiner, N. (2020). Social world, hiking and nation: the Israel National Trail. *Social and Cultural Geography*, 21(6):810–838.

Knaus, F. (2012). Bedeutung, Charakteristiken und wirtschaftliche Auswirkungen des Sommertourismus in der UNESCO Biosphäre Entlebuch. *Interner Bericht ETH Zürich und Biosphärenmanagement UBE*, pages 1–37.

Koblet, O. and Purves, R. S. (2020). From online texts to Landscape Character Assessment: Collecting and analysing first-person landscape perception computationally. *Landscape and Urban Planning*, 197.

Kruk, J., Lubin, J., Sikka, K., Lin, X., Jurafsky, D., and Divakaran, A. (2019). Integrating text and image: Determining multimodal document intent in instagram posts. *EMNLP-IJCNLP 2019 - 2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing, Proceedings of the Conference*, pages 4622–4632.

Kumar, M. and Kumar, P. (2008). Valuation of the ecosystem services: A psycho-cultural perspective. *Ecological Economics*, 64(4):808–819.

Ladle, R. J., Correia, R. A., Do, Y., Joo, G. J., Malhado, A. C., Proulx, R., Roberge, J. M., and Jepson, P. (2016). Conservation culturomics. *Frontiers in Ecology and the Environment*, 14(5):269–275.

Lawu, B. L., Lim, F., Susilo, A., and Surantha, N. (2021). Social media data crowdsourcing as a new stream for environmental planning & monitoring: A review. In *IOP Conference Series: Earth and Environmental Science*, volume 729.

Lee, H., Seo, B., Koellner, T., and Lautenbach, S. (2019). Mapping cultural ecosystem services 2.0 – Potential and shortcomings from unlabeled crowd sourced images. *Ecological Indicators*, 96:505–515.

Lee, J. h., Park, H. j., Kim, I., and Kwon, H. s. (2020). Analysis of cultural ecosystem services using text mining of residents' opinions. *Ecological Indicators*, 115.

Levin, N., Lechner, A. M., and Brown, G. (2017). An evaluation of crowdsourced information for assessing the visitation and perceived importance of protected areas. *Applied Geography*, 79:115–126.

Li, L., Goodchild, M. F., and Xu, B. (2013). Spatial, temporal, and socioeconomic patterns in the use of twitter and flickr. *Cartography and Geographic Information Science*, 40(2):61–77.

Long, P. R., Nogué, S., Benz, D., and Willis, K. J. (2021). Devising a method to remotely model and map the distribution of natural landscapes in Europe with the greatest recreational amenity value (cultural services). *Frontiers of Biogeography*, 13(1).

López-Santiago, C. A., Oteros-Rozas, E., Martín-López, B., Plieninger, T., Martín, E. G., and González, J. A. (2014). Using visual stimuli to explore the social perceptions of ecosystem services in cultural landscapes: The case of transhumance in Mediterranean Spain. *Ecology and Society*, 19(2).

Martínez Pastur, G., Peri, P. L., Lencinas, M. V., García-Llorente, M., and Martín-López, B. (2016). Spatial patterns of cultural ecosystem services provision in Southern Patagonia. *Landscape Ecology*, 31(2):383–399.

Maxwell, S. L., Fuller, R. A., Brooks, T. M., and Watson, J. E. (2016). Biodiversity: The ravages of guns, nets and bulldozers. *Nature*, 536(7615):143–145.

McCrow-Young, A. (2020). Approaching Instagram data: reflections on accessing, archiving and anonymising visual social media. *Communication Research and Practice*.

MEA (2005). Millennium Ecosystem Assessment: Ecosystems and Human Well-being: Synthesis. *Island Press, Washington, DC.*

Michel, J. B., Kui Shen, Y., Presser Aiden, A., Veres, A., Gray, M. K., Pickett, J. P., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J., Pinker, S., Nowak, M. A., and Aiden, E. L. (2011). Quantitative analysis of culture using millions of digitized books. *Science*, 331(6014):176–182.

Milcu, A. I., Hanspach, J., Abson, D., and Fischer, J. (2013). Cultural ecosystem services: A literature review and prospects for future research. *Ecology and Society*, 18(3).

Muki Haklay (2013). Citizen science and volunteered geographic information: overview and typology of participation. In Sui, D., Elwood, S., and Goodchild, M., editors, *Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice*, chapter 7, pages 105–122. Springer Netherlands.

Norman, P. and Pickering, C. M. (2019). Factors influencing park popularity for mountain bikers, walkers and runners as indicated by social media route data. *Journal of Environmental Management*, 249.

Norman, P., Pickering, C. M., and Castley, G. (2019). What can volunteered geographic information tell us about the different ways mountain bikers, runners and walkers use urban reserves? *Landscape and Urban Planning*, 185:180–190.

Nov, O., Naaman, M., and Ye, C. (2010). Analysis of participation in an online photo-sharing community: A multidimensional perspective. *Journal of the American Society for Information Science and Technology*, 61(3):555–566.

O'Farrell, P. J., De Lange, W. J., Le Maitre, D. C., Reyers, B., Blignaut, J. N., Milton, S. J., Atkinson, D., Egoh, B., Maherry, A., Colvin, C., and Cowling, R. M. (2011). The possibilities and pitfalls presented by a pragmatic approach to ecosystem service valuation in an arid biodiversity hotspot. *Journal of Arid Environments*, 75(6):612–623.

Oteros-Rozas, E., Martín-López, B., Fagerholm, N., Bieling, C., and Plieninger, T. (2018). Using social media photos to explore the relation between cultural ecosystem services and landscape features across five European sites. *Ecological Indicators*, 94:74–86.

Palomo, I. (2017). Climate Change Impacts on Ecosystem Services in High Mountain Areas: A Literature Review. *Mountain Research and Development*, 37(2):179–187.

Pelagios Network (2021). Recogito Pelagios. https://recogito.pelagios.org/. Last accessed on Mar 26, 2021.

Pinto-Correia, T., Barroso, F., Surová, D., and Menezes, H. (2011). The fuzziness of Montado landscapes: Progress in assessing user preferences through photo-based surveys. *Agroforestry Systems*, 82(2):209–224.

Plieninger, T., Bieling, C., Fagerholm, N., Byg, A., Hartel, T., Hurley, P., López-Santiago, C. A., Nagabhatla, N., Oteros-Rozas, E., Raymond, C. M., van der Horst, D., and Huntsinger, L. (2015a). The role of cultural ecosystem services in landscape management and planning. *Current Opinion in Environmental Sustainability*, 14:28–33.

Plieninger, T., Dijks, S., Oteros-Rozas, E., and Bieling, C. (2013). Assessing, mapping, and quantifying cultural ecosystem services at community level. *Land Use Policy*, 33:118–129.

Plieninger, T., Kizos, T., Bieling, C., Dû-Blayo, L. L., Budniok, M. A., Bürgi, M., Crumley, C. L., Girod, G., Howard, P., Kolen, J., Kuemmerle, T., Milcinski, G., Palang, H., Trommler, K., and Verburg, P. H. (2015b). Exploring ecosystem-change and society through a landscape lens: Recent progress in european landscape research. *Ecology and Society*, 20(2).

Purves, R. S. and Derungs, C. (2015). From Space to Place: Place-Based Explorations of Text. *International Journal of Humanities and Arts Computing*, 9(1):74–94.

Ramachandram, D. and Taylor, G. W. (2017). Deep Multimodal Learning. *IEEE SIgnal ProcESSIng MagazInE*, (November):96–108.

Raudsepp-Hearne, C., Peterson, G. D., and Bennett, E. M. (2010). Ecosystem service bundles for analyzing tradeoffs in diverse landscapes. *Proceedings of the National Academy of Sciences of the United States of America*, 107(11):5242–5247.

Reinecke, L. and Trepte, S. (2014). Authenticity and well-being on social network sites: A two-wave longitudinal study on the effects of online authenticity and the positivity bias in SNS communication. *Computers in Human Behavior*, 30:95–102.

Richards, D. R. and Friess, D. A. (2015). A rapid indicator of cultural ecosystem service usage at a fine spatial scale: Content analysis of social media photographs. *Ecological Indicators*, 53:187–195.

Richards, D. R., Tunçer, B., and Tunçer, B. (2018). Using image recognition to automate assessment of cultural ecosystem services from social media photographs. *Ecosystem Services*, 31:318–325.

Ruths, D. and Pfeffer, J. (2014). Social media for large studies of behavior. *Science*, 346(6213):1063–1064.

Schirpke, U., Scolozzi, R., Da Re, R., Masiero, M., Pellegrino, D., and Marino, D. (2018). Recreational ecosystem services in protected areas: A survey of visitors to Natura 2000 sites in Italy. *Journal of Outdoor Recreation and Tourism*, 21:39–50.

Schirpke, U., Scolozzi, R., Dean, G., Haller, A., Hieronymus, J., Sarmiento, F. O., Sattler, B., Schleyer, C., Kister, J., and Kov, B. (2020). Cultural ecosystem services in mountain regions : Conceptualising conflicts among users and limitations of use. *Ecosystem Services*, 46.

Schirpke, U., Timmermann, F., Tappeiner, U., and Tasser, E. (2016). Cultural ecosystem services of mountain regions: Modelling the aesthetic value. *Ecological Indicators*, 69:78–90.

Scholte, S. S., Daams, M., Farjon, H., Sijtsma, F. J., van Teeffelen, A. J., and Verburg, P. H. (2018). Mapping recreation as an ecosystem service: Considering scale, interregional differences and the influence of physical attributes. *Landscape and Urban Planning*, 175:149–160.

Schwartz, H. A. and Ungar, L. H. (2015). Data-Driven Content Analysis of Social Media: A Systematic Overview of Automated Methods. *Annals of the American Academy of Political and Social Science*, 659:78–94.

Schweizer Alpen-Club (2021). vanlife: Wild-Campen im Aufwind - SAC lanciert kritische Sensibilisierungskampagne zum Boom. https://www.sac-cas.ch/de/umwelt/vanlife-wild-campen-im-aufwind-33292/. Last accessed on May 27, 2021.

Scott, A. (2002). Assessing public perception of landscape: The LANDMAP experience. *Landscape Research*, 27(3):271–295.

Sessions, C., Wood, S. A., Rabotyagov, S., and Fisher, D. M. (2016). Measuring recreational visitation at U.S. National Parks with crowd-sourced photographs. *Journal of Environmental Management*, 183:703–711.

Sinclair, M., Mayer, M., Woltering, M., and Ghermandi, A. (2020). Using social media to estimate visitor provenance and patterns of recreation in Germany's national parks. *Journal of Environmental Management*, 263.

Smith, B. and Mark, D. M. (2003). Do mountains exist? Towards an ontology of landforms. *Environment and Planning B: Planning and Design*, 30(3):411–427.

Steffen, W., Richardson, K., Rockström, J., Cornell, S. E., Fetzer, I., Bennett, E. M., Biggs, R., Carpenter, S. R., De Vries, W., De Wit, C. A., Folke, C., Gerten, D., Heinke, J., Mace, G. M., Persson, L. M., Ramanathan, V., Reyers, B., and Sörlin, S. (2015). Planetary boundaries: Guiding human development on a changing planet. *Science*, 347(6223).

Stephenson, J. (2008). The Cultural Values Model: An integrated approach to values in landscapes. *Landscape and Urban Planning*, 84(2):127–139.

Sui, D. and Goodchild, M. (2011). The convergence of GIS and social media: Challenges for GIScience. *International Journal of Geographical Information Science*, 25(11):1737–1748.

Swisscom (2021). Mobile Coverage. https://scmplc.begasoft.ch/plcapp/pages/gis/netzabdeckung.jsf?netztyp=lte. Last accessed on July 22, 2021.

Tenerelli, P., Demšar, U., and Luque, S. (2016). Crowdsourcing indicators for cultural ecosystem services: A geographically weighted approach for mountain landscapes. *Ecological Indicators*, 64:237–248.

Tenkanen, H., Di Minin, E., Heikinheimo, V., Hausmann, A., Herbst, M., Kajala, L., and Toivonen, T. (2017). Instagram, Flickr, or Twitter: Assessing the usability of social media data for visitor monitoring in protected areas. *Scientific Reports*, 7.

The Research Box, Land Use Consultants, and Minter, R. (2009). Natural England commissioned Report NECR024: Experiencing Landscapes: capturing the cultural services and experiential qualities of landscape. Technical report.

Toivonen, T., Heikinheimo, V., Fink, C., Hausmann, A., Hiippala, T., Järv, O., Tenkanen, H., and Di Minin, E. (2019). Social media data for conservation science: A methodological overview. *Biological Conservation*, 233:298–315.

UN World Travel Organization (2020). International Tourism Highlights. https://www.e-unwto.org/doi/epdf/10.18111/9789284422456. Last accessed on May 22, 2021.

UNESCO (1972). Convention concerning the protection of the world cultural and natural heritage. Technical report.

UNESCO (2015). Resolutions and Decisions by the General Assembly at its 70th session. https://undocs.org/en/a/70/49(Vol.I). Last accessed on May 17, 2021.

UNESCO (2020). UNESCO Biosphere. https://en.unesco.org/biosphere/about. Last accessed on Nov 22, 2020.

UNESCO Biosphere Entlebuch (2020). Vision - Mission - Leitbild - Ziele. https://www.biosphaere.ch/de/unesco-biosphaere-a-z/biosphaere-entlebuch-was-ist-das/vision-mission-leitbild-ziele/. Last accessed on May 17, 2021.

UNESCO Biosphere Entlebuch (2021a). Biosphäre Entlebuch - Ein Segen für alle. https://www.biosphaere.ch/de/unesco-biosphaere-a-z/biosphaere-entlebuch-was-ist-das/ein-segen-fuer-alle/. Last accessed on May 17, 2021.

UNESCO Biosphere Entlebuch (2021b). Biosphäre Entlebuch - Was ist das? https://www.biosphaere.ch/de/unesco-biosphaere-a-z/biosphaere-entlebuch-was-ist-das/. Last accessed on May 17, 2021.

UNESCO Biosphere Entlebuch (2021c). Natur Landschaft. https://www.biosphaere.ch/de/unesco-biosphaere-a-z/natur-landschaft/. Last accessed on May 24, 2021.

Väisänen, T., Heikinheimo, V., Hiippala, T., and Toivonen, T. (2021). Exploring human–nature interactions in national parks with social media photographs and computer vision. *Conservation Biology*, 35(2):424–436.

van Dijck, J. (2011). Flickr and the culture of connectivity: Sharing views, experiences, memories. *Memory Studies*, 4(4):401–415.

van Zanten, B. T., van Berkel, D. B., Meentemeyer, R. K., Smith, J. W., Tieskens, K. F., and Verburg, P. H. (2016a). Continental-scale quantification of landscape values using social media data. *Proceedings of the National Academy of Sciences of the United States of America*, 113(46):12974–12979.

van Zanten, B. T., Zasada, I., Koetse, M. J., Ungaro, F., Häfner, K., and Verburg, P. H. (2016b). A comparative approach to assess the contribution of landscape features to aesthetic and recreational values in agricultural landscapes. *Ecosystem Services*, 17:87–98.

Venter, O., Sanderson, E. W., Magrach, A., Allan, J. R., Beher, J., Jones, K. R., Possingham, H. P., Laurance, W. F., Wood, P., Fekete, B. M., Levy, M. A., and Watson, J. E. (2016). Sixteen years of change in the global terrestrial human footprint and implications for biodiversity conservation. *Nature Communications*, 7.

Vlami, V., Kokkoris, I. P., Zogaris, S., Kehayias, G., and Dimopoulos, P. (2021). Cultural ecosystem services in the natura 2000 network: Introducing proxy indicators and conflict risk in Greece. *Land*, 10(4).

Waldron, A., Mooers, A. O., Miller, D. C., Nibbelink, N., Redding, D., Kuhn, T. S., Roberts, J. T., and Gittleman, J. L. (2013). Targeting global conservation funding to limit immediate biodiversity declines. *Proceedings of the National Academy of Sciences of the United States of America*, 110(29):12144–12148.

Wallace, K. J. (2007). Classification of ecosystem services: Problems and solutions. *Biological Conservation*, 139:235–246.

Wang, Z., Zhu, Z., Xu, M., and Qureshi, S. (2021). Fine-grained assessment of greenspace satisfaction at regional scale using content analysis of social media and machine learning. *Science of the Total Environment*, 776.

Wartmann, F., Acheson, E., and Purves, R. S. (2018). Describing and comparing landscapes using tags, texts, and free lists: an interdisciplinary approach. *International Journal of Geographical Information Science*, 32(8):1572–1592.

Wartmann, F. M., Egorova, E., Derungs, C., Mark, D. M., and Purves, R. S. (2015). More Than a List: What Outdoor Free Listings of Landscape Categories Reveal About Commonsense Geographic Concepts and Memory Search Strategies. In Fabrikant, S. I., Raubal, M., Bertolotto, M., Davies, C., Freundschuh, S., and Bell, S., editors, *Spatial Information Theory*, pages 224–243, Cham. Springer International Publishing.

Wartmann, F. M. and Purves, R. S. (2018). Investigating sense of place as a cultural ecosystem service in different landscapes through the lens of language. *Landscape and Urban Planning*, 175:169–183.

Wood, S. A., Guerry, A. D., Silver, J. M., and Lacayo, M. (2013). Using social media to quantify nature-based tourism and recreation. *Scientific Reports*, 3.

Woolmer, G., Trombulak, S. C., Ray, J. C., Doran, P. J., Anderson, M. G., Baldwin, R. F., Morgan, A., and Sanderson, E. W. (2008). Rescaling the Human Footprint: A tool for conservation planning at an ecoregional scale. *Landscape and Urban Planning*, 87:42–53.

Wu, X., Wood, S. A., Fisher, D., and Lindsey, G. (2017). Photos, tweets, and trails: Are social media proxies for urban trail use? *Journal of Transport and Land Use*, 10(1):789–804.

Yoshimura, N. and Hiura, T. (2017). Demand and supply of cultural ecosystem services: Use of geotagged photos to map the aesthetic value of landscapes in Hokkaido. *Ecosystem Services*, 24:68–78.

Zoderer, B. M., Tasser, E., Erb, K. H., Lupo Stanghellini, P. S., and Tappeiner, U. (2016). Identifying and mapping the tourists' perception of cultural ecosystem services: A case study from an Alpine region. *Land Use Policy*, 56:251–261.

Appendix

A.1. Metadata

A.1.1. Overview of the Data Sources

Table A.1.: *Data used in this thesis and their sources.*

Data Set	Source
Flickr	Flickr API (*flickr.com/services/api/*) (last accessed: 17/01/2021)
Tour Descriptions	Manual extraction from *maps.luzern.com/* (last accessed: 29/12/2021)
Hiking Trails	swissTLM3D (TLM_Strasse) from *GeoVITe ETH Zurich* (2020)
Boundaries and Municipalities	swissBOUNDARIES3D from Swisstopo (2015)

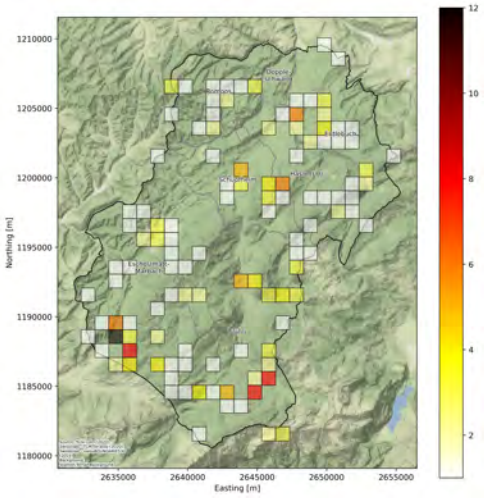
A.1.2. Software

The code for this thesis can be found on github.com. Table A.2 shows the software and libraries used in this thesis.

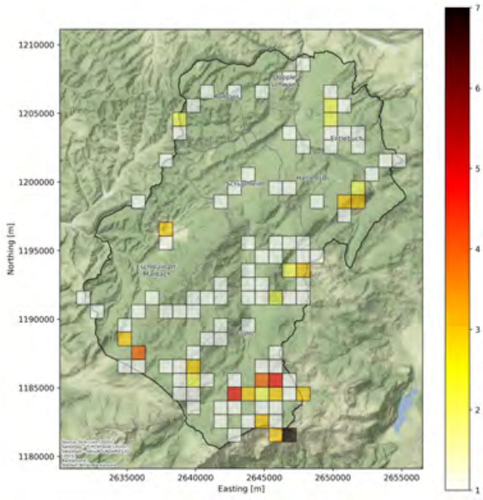
Table A.2.: *Python environment and libraries used in this thesis.*

Program	**Version**	**URL**
Python	3.9.5	python.org/
PyCharm	2020.3.3	jetbrains.com/pycharm/
External Library	**Version**	**URL**
contextily	1.1.0	github.com/geopandas/contextily/
esda	2.3.6	github.com/pysal/esda/
flickrapi	2.4.0	pypi.org/project/flickrapi/
geopandas	0.9.0	geopandas.org/
gpxpy	1.4.2	pypi.org/project/gpxpy/
libpysal	4.4.0	pysal.org/libpysal/
matplotlib	3.3.4	matplotlib.org/
numpy	1.20.3	numpy.org/
pandas	1.2.4	pandas.pydata.org/
splot	1.1.3	pypi.org/project/splot/
statistics	3.4	docs.python.org/3/library/statistics
urllib	1.26.4	docs.python.org/3.6/library/urllib

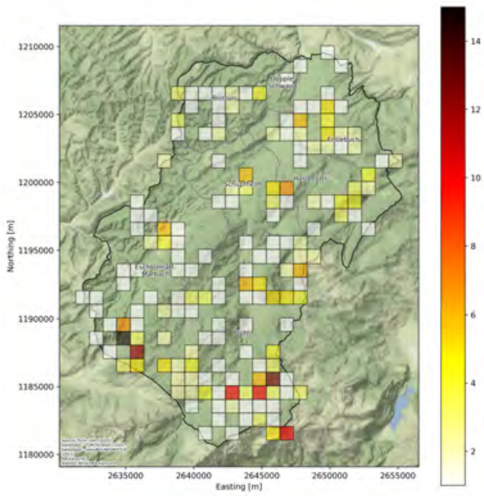
A.2. Additional and Complete Results



(a)



(b)



(c)

Figure A.1.: *Number of contributors per raster cell for (a) text data, (b) pictures and the (c) combined data set which were used as input data to extract significant clusters with the Local Moran's I.*

Table A.3.: *Count of LF within the threshold to a CES (N), its proportion of all annotations of a LF subtype (%LF) and the share of CES capturing N (%CES) for the combined data set (1).*

	Identity			Information Board			Information Office			Local History			Tradition			Traditional Architecture			Recreational Facilities			Signpost			Viewpoint		
	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES
Natural Landscape	9	28.1	26.9	2	6.3	36.4	0.0	0.0	0.0	10	31.3	28.6	5	15.6	10.7	0.0	0.0	0.0	6	18.8	16.1	5	15.6	15.6	27	84.4	46.0
Human influenced Landscape	3	11.5	7.7	0.0	0.0	0.0	0.0	0.0	0.0	3	11.5	11.4	7	26.9	32.1	0.0	0.0	0.0	2	7.7	3.2	0.0	0.0	0.0	6	23.1	7.3
Bedrock	10	6.3	38.5	18	11.3	54.6	0.0	0.0	0.0	7	4.4	62.9	0.0	0.0	0.0	2	1.3	13.0	11	6.9	22.6	25	15.7	50.0	48	30.2	48.9
Flower / Funghi	1	2.3	3.9	1	2.3	18.2	0.0	0.0	0.0	1	2.3	8.6	1	2.3	7.1	2	4.7	4.4	2	4.7	8.1	2	4.7	6.3	2	4.7	2.2
Forest	28	8.0	69.2	17	4.8	45.5	3	0.9	100	15	4.3	57.1	4	1.1	14.3	12	3.4	65.2	42	11.9	43.6	51	14.5	71.9	110	31.3	72.3
Grass- and Moorland	37	9.3	69.2	16	4.0	36.4	3	0.8	100	32	8.1	62.9	8	2.0	21.4	18	4.5	73.9	49	12.3	48.4	59	14.9	75.0	136	34.3	71.5
Lake	9	30.0	34.6	3	10.0	18.2	0.0	0.0	0.0	3	10.0	51.4	0.0	0.0	0.0	0.0	0.0	0.0	18	60.0	24.2	8	26.7	12.5	12	40.0	27.7
River / Creek	1	1.8	3.9	5	8.9	36.4	0.0	0.0	0.0	5	8.9	62.9	0.0	0.0	0.0	1	1.8	4.4	7	12.5	11.3	6	10.7	3.1	4	7.1	3.7
Rock	7	8.4	15.4	7	8.4	36.4	0.0	0.0	0.0	12	14.5	51.4	1	1.2	3.6	1	1.2	13.0	5	6.0	3.2	13	15.7	15.6	22	26.5	19.0
Shrub	1	1.8	3.9	2	3.5	18.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	8	14.0	39.1	8	14.0	19.4	5	8.8	3.1	1	1.8	0.7
Snow / Ice	5	16.1	15.4	2	6.5	18.2	0.0	0.0	0.0	4	12.9	8.6	2	6.5	10.7	0.0	0.0	0.0	2	6.5	1.6	5	16.1	12.5	13	41.9	32.9
Summit	13	8.4	19.2	7	4.6	18.2	0.0	0.0	0.0	13	8.4	57.1	5	3.3	10.7	0.0	0.0	0.0	17	11.0	22.6	27	17.5	28.1	84	54.6	67.2
Tree	9	12.0	34.6	0.0	0.0	0.0	0.0	0.0	0.0	1	1.3	2.9	1	1.3	10.7	6	8.0	47.8	12	16.0	24.2	10	13.3	28.1	12	16.0	16.1
Waterfall	0.0	0.0	0.0	4	33.3	18.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1	8.3	0.7
Wild Animal	3	23.1	19.2	2	15.4	36.4	0.0	0.0	0.0	3	23.1	14.3	0.0	0.0	0.0	0.0	0.0	0.0	3	23.1	11.3	2	15.4	6.3	2	15.4	8.0
Agriculture	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1	10.0	21.7	0.0	0.0	0.0	4	40.0	12.5	2	20.0	3.7
Infrastructure	67	16.6	100	21	5.2	100	13	3.2	100	45	11.1	91.4	45	11.1	100	28	6.9	100	79	19.6	95.2	44	10.9	81.3	87	21.5	62.0
Livestock	1	5.3	3.9	0.0	0.0	0.0	0.0	0.0	0.0	3	15.8	57.1	0.0	0.0	0.0	1	5.3	4.4	8	42.1	9.7	2	10.5	6.3	6	31.6	8.8
Path / Trail	9	5.8	38.5	8	5.1	63.6	0.0	0.0	0.0	6	3.9	60.0	3	1.9	14.3	4	2.6	26.1	18	11.5	25.8	16	10.3	46.9	37	23.7	51.8
Urban	6	7.5	15.4	4	5.0	18.2	3	3.8	100	4	5.0	5.7	3	3.8	21.4	5	6.3	39.1	11	13.8	17.7	6	7.5	6.3	16	20.0	24.1

Table A.4.: *Count of LF within the threshold to a CES (N), its proportion of all annotations of a LF subtype (%LF) and the share of CES capturing N (%CES) for the combined data set (2).*

	Camping			People			Restaurant / Accommodation			Dawn / Sunset			Healing Powers			Place Attachment			Church			Summit Cross			Total		
	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES
Natural Landscape	0.0	0.0	0.0	1	3.1	1.2	9	28.1	25.0	3	9.4	52.9	1	3.1	17.7	9	28.1	26.9	2	6.3	8.3	4	12.5	33.3	30	93.8	23.6
Human influenced Landscape	0.0	0.0	0.0	0.0	0.0	0.0	6	23.1	8.3	1	3.9	5.9	1	3.9	17.7	1	3.7	3.7	3	11.5	16.7	0.0	0.0	0.0	17	65.4	7.0
Bedrock	0.0	0.0	0.0	47	29.6	57.8	8	5.0	20.0	10	6.3	58.8	2	1.3	5.9	13	8.2	19.2	4	2.5	11.1	15	9.4	46.7	102	64.2	36.6
Flower / Funghi	0.0	0.0	0.0	3	7.0	4.8	2	4.7	10.0	0.0	0.0	0.0	1	2.3	17.7	2	4.7	7.7	0.0	0.0	0.0	0.0	0.0	0.0	12	27.9	5.5
Forest	3	0.9	100	78	22.2	77.1	21	6.0	40.0	26	7.4	94.1	7	2.0	52.9	22	6.3	38.5	23	6.5	33.3	34	9.7	93.3	208	59.1	59.6
Grass- and Moorland	2	0.5	100	96	24.2	74.7	45	11.3	48.3	30	7.6	100	4	1.0	17.7	27	6.8	42.3	26	6.6	30.6	36	9.1	93.3	236	59.5	60.6
Lake	0.0	0.0	0.0	9	30.0	12.1	5	16.7	10.0	3	10.0	52.9	4	13.3	5.9	8	26.7	23.1	0.0	0.0	0.0	2	6.7	20.0	26	86.7	19.7
River / Creek	0.0	0.0	0.0	13	23.2	20.5	3	5.4	3.3	0.0	0.0	0.0	7	12.5	11.8	3	5.4	11.5	1	1.8	5.6	0.0	0.0	0.0	35	62.5	10.9
Rock	0.0	0.0	0.0	28	33.7	25.3	12	14.5	8.3	1	1.2	5.9	5	6.0	5.9	0.0	0.0	0.0	0.0	0.0	0.0	1	1.2	6.7	46	55.4	15.0
Shrub	0.0	0.0	0.0	11	19.3	28.9	2	3.5	3.3	0.0	0.0	0.0	6	10.5	41.2	3	5.3	11.5	9	15.8	16.7	0.0	0.0	0.0	33	57.9	11.1
Snow / Ice	0.0	0.0	0.0	5	16.1	4.8	5	16.1	16.7	2	6.5	52.9	0.0	0.0	0.0	2	6.5	7.7	0.0	0.0	0.0	1	3.2	6.7	17	54.8	14.3
Summit	1	0.7	100	28	18.2	36.1	13	8.4	20.0	20	13.0	70.6	0.0	0.0	0.0	16	10.4	7.7	0.0	0.0	0.0	28	18.2	93.3	100	64.9	35.2
Tree	1	1.3	100	20	26.7	41.0	9	12.0	26.7	8	10.7	35.3	1	1.3	5.9	0.0	0.0	0.0	17	22.7	27.8	2	2.7	40.0	48	64.0	23.5
Waterfall	0.0	0.0	0.0	4	33.3	6.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5	41.7	11.5	0.0	0.0	0.0	0.0	0.0	0.0	9	75.0	1.8
Wild Animal	0.0	0.0	0.0	1	7.7	1.2	1	7.7	5.0	0.0	0.0	0.0	1	7.7	17.7	3	23.1	11.5	1	7.7	5.6	0.0	0.0	0.0	6	46.2	1.3
Agriculture	0.0	0.0	0.0	1	10.0	1.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2	20.0	5.6	0.0	0.0	0.0	6	60.0	2.8
Infrastructure	1	0.3	100	88	21.8	67.5	110	27.2	100	16	4.0	88.2	19	4.7	88.2	37	9.2	57.7	58	14.4	100	8	2.0	80.0	316	78.2	82.3
Livestock	0.0	0.0	0.0	10	52.6	16.9	3	15.8	3.3	1	5.3	5.9	0.0	0.0	0.0	1	5.3	3.9	0.0	0.0	0.0	1	5.3	33.3	14	73.7	10.6
Path / Trail	0.0	0.0	0.0	31	19.9	49.4	9	5.8	31.7	3	1.9	52.9	3	1.9	41.2	10	6.4	38.5	4	2.6	16.7	6	3.9	53.3	88	56.4	40.7
Urban	0.0	0.0	0.0	17	21.3	15.7	8	10.0	10.0	0.0	0.0	0.0	2	2.5	5.9	2	2.5	7.7	15	18.8	38.9	4	5.0	46.7	48	60.0	19.1

Table A.5.: *Absolute (N) and relative (%LF) densities of LF for each cluster and km^2 for the combined data set.*

	Cluster 1		Cluster 2		Cluster 3		Cluster 4		Cluster 5		Cluster 6		Cluster 7		Total	
	N	%LF	N	%LF	N	%LF	N	%LF	N	%LF	N	%LF	N	%LF	N	%LF
Natural Landscape	0.14	1.7	0.56	6.94	0	0	0	0	0	0	0	0	0	0	0.2	2.44
Human influenced Landscape	0.14	2.27	0.11	1.85	0	0	0	0	0	0	1.0	16.67	0.25	4.17	0.15	2.44
Bedrock	1.18	0.77	0.78	0.51	1.0	0.65	1.0	0.65	0	0	1.0	0.65	0	0	0.9	0.59
Flower / Funghi	0.18	0.45	0.44	1.11	1.5	3.75	0	0	0	0	0	0	0.25	0.62	0.29	0.73
Forest	2.14	0.67	3.44	1.08	3.0	0.94	2.0	0.63	0	0	1.0	0.31	0	0	2.2	0.69
Grass- and Moorland	3.45	0.92	4.11	1.09	3.5	0.93	2.0	0.53	0	0	4.0	1.06	1.0	0.27	3.17	0.84
Lake	0.18	0.73	0.44	1.78	4.5	18.0	0	0	0	0	0	0	0	0	0.41	1.66
River / Creek	0.18	0.53	0.22	0.65	0	0	0	0	0	0	0	0	0	0	0.15	0.43
Rock	0.73	0.89	0	0	0	0	0	0	0	0	2.0	2.44	0	0	0.44	0.54
Shrub	0.23	0.41	1.0	1.79	2.0	3.57	0	0	0	0	1.0	1.79	0	0	0.46	0.83
Snow / Ice	0.14	0.45	0.22	0.74	0	0	0	0	0	0	1.0	3.33	0	0	0.15	0.49
Summit	1.09	0.77	1.22	0.86	1.0	0.7	1.0	0.7	0.5	0.35	2.0	1.41	0	0	1	0.7
Tree	0.55	0.77	1.0	1.41	2.5	3.52	0	0	0	0	1.0	1.41	0.25	0.35	0.68	0.96
Waterfall	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Wild Animal	0.05	0.65	0.22	3.17	0	0	0	0	0	0	0	0	0	0	0.07	1.05
Agriculture	0.05	0.57	0.22	2.78	0	0	0	0	0	0	0	0	0	0	0.07	0.91
Infrastructure	3.09	1.14	8.33	3.09	5.5	2.04	1.0	0.37	1.5	0.56	0	0	2.0	0.74	4.05	1.5
Livestock	0.09	0.53	0	0	0	0	0	0	0	0	1.0	5.88	0	0	0.07	0.43
Path / Trail	0.86	0.79	0.89	0.82	0	0	0	0	1.0	0.92	2.0	1.83	0.75	0.69	0.83	0.76
Urban	0.55	0.94	1.44	2.49	0	0	0	0	1.0	1.72	0	0	0.5	0.86	0.71	1.22

Table A.6.: *Count of LF within the threshold to a CES (N), its proportion of all annotations of a LF subtype (%LF) and the share of CES capturing N (%CES) for text data (1).*

	Identity			Information Board			Information Office			Local History			Tradition			Traditional Architecture			Recreational Facilities			Signpost			Viewpoint		
	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES
Natural Landscape																											
Human influenced Landscape	4	12.5	50.0	1	3.1	40.0	0.0	0.0	0.0	10	31.3	58.8	5	15.6	11.5	0.0	0.0	0.0	5	15.6	23.7	1	3.1	10.0	27	84.4	63.0
Bedrock	2	22.2	33.3	0.0	0.0	0.0	0.0	0.0	0.0	2	22.2	23.5	0.0	0.0	0.0	0.0	0.0	0.0	2	22.2	10.5	0.0	0.0	0.0	3	33.3	11.1
Flower / Funghi	0.0	0.0	0.0	1	33.3	40.0	0.0	0.0	0.0	1	33.3	17.7	1	33.3	7.7	0.0	0.0	0.0	1	33.3	7.9	0.0	0.0	0.0	0.0	0.0	0.0
Forest	1	2.6	25.0	2	5.1	20.0	0.0	0.0	0.0	1	2.6	5.9	0.0	0.0	0.0	0.0	0.0	0.0	7	18.0	26.3	3	7.7	30.0	2	5.1	3.7
Grass- and Moorland	2	6.9	25.0	4	13.8	40.0	0.0	0.0	0.0	4	13.8	17.7	2	6.9	7.7	0.0	0.0	0.0	6	20.7	18.4	2	6.9	20.0	2	6.9	3.7
Lake	6	37.5	58.3	0.0	0.0	0.0	0.0	0.0	0.0	1	6.3	5.9	0.0	0.0	0.0	0.0	0.0	0.0	10	62.5	23.7	4	25.0	10.0	5	31.3	11.1
River / Creek	1	4.2	8.3	0.0	0.0	0.0	0.0	0.0	0.0	3	12.5	29.4	0.0	0.0	0.0	1	4.2	100	6	25.0	15.8	1	4.2	10.0	2	8.3	7.4
Rock	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Shrub	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1	25.0	2.6	0.0	0.0	0.0	0.0	0.0	0.0
Snow / Ice	1	25.0	16.7	0.0	0.0	0.0	0.0	0.0	0.0	3	75.0	11.8	2	50.0	11.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4	100	18.5
Summit	3	17.7	16.7	0.0	0.0	0.0	0.0	0.0	0.0	5	29.4	11.8	2	11.8	11.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	15	88.2	33.3
Tree	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1	25.0	10.0	1	25.0	1.[illegible]
Waterfall	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1	20.0	1.[illegible]
Wild Animal	3	33.3	41.7	1	11.1	40.0	0.0	0.0	0.0	3	33.3	29.4	0.0	0.0	0.0	0.0	0.0	0.0	2	22.2	15.8	0.0	0.0	0.0	1	11.1	3.[illegible]
Agriculture	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1	50.0	10.0	0.0	0.0	0.0
Infrastructure	28	13.9	100	10	5.0	100	10	5.0	100	20	9.9	82.4	37	18.3	100	1	0.5	100	49	24.3	94.7	8	4.0	50.0	37	18.3	64.8
Livestock	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1	50.0	17.7	0.0	0.0	0.0	0.0	0.0	0.0	1	50.0	2.6	0.0	0.0	0.0	1	50.0	5.[illegible]
Path / Trail	1	1.7	8.3	1	1.7	20.0	0.0	0.0	0.0	2	3.5	11.8	0.0	0.0	0.0	0.0	0.0	0.0	6	10.3	13.2	5	8.6	50.0	8	13.8	16.[illegible]
Urban	2	5.4	8.3	0.0	0.0	0.0	3	8.1	100	1	2.7	5.9	3	8.1	23.1	0.0	0.0	0.0	3	8.1	10.5	0.0	0.0	0.0	5	13.5	13.0

Table A.7.: *Count of LF within the threshold to a CES (N), its proportion of all annotations of a LF subtype (%LF) and the share of CES capturing N (%CES) for text data (2).*

	Camping			People			Restaurant / Accommodation			Dawn / Sunset			Healing Powers			Place Attachment			Church			Summit Cross			Total		
	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES
Natural Landscape	0.0	0.0	0.0	0.0	0.0	0.0	9	28.1	33.3	1	3.1	100	1	3.1	17.7	9	28.1	26.9	2	6.3	15.8	2	6.3	75.0	30	93.8	35.0
Human influenced Landscape	0.0	0.0	0.0	0.0	0.0	0.0	6	23.1	9.5	1	3.9	100	1	3.9	17.7	1	3.9	3.9	3	11.5	31.6	0.0	0.0	0.0	17	65.4	15.0
Bedrock	0.0	0.0	0.0	0.0	0.0	0.0	2	22.2	9.5	0.0	0.0	0.0	0.0	0.0	0.0	3	33.3	11.5	3	33.3	15.8	0.0	0.0	0.0	8	88.9	10.2
Flower / Funghi	0.0	0.0	0.0	0.0	0.0	0.0	1	33.3	11.9	0.0	0.0	0.0	1	33.3	17.7	2	66.7	7.7	0.0	0.0	0.0	0.0	0.0	0.0	2	66.7	7.3
Forest	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2	5.1	41.2	6	15.4	30.8	0.0	0.0	0.0	0.0	0.0	0.0	17	43.6	12.8
Grass- and Moorland	0.0	0.0	0.0	0.0	0.0	0.0	3	10.3	19.1	0.0	0.0	0.0	4	13.8	17.7	12	41.4	34.6	0.0	0.0	0.0	1	3.5	25.0	17	58.6	15.3
Lake	0.0	0.0	0.0	0.0	0.0	0.0	3	18.8	9.5	2	12.5	100	4	25.0	5.9	7	43.8	19.2	0.0	0.0	0.0	1	6.3	25.0	14	87.5	13.1
River / Creek	0.0	0.0	0.0	0.0	0.0	0.0	1	4.2	2.4	0.0	0.0	0.0	2	8.3	11.8	3	12.5	11.5	1	4.2	10.5	0.0	0.0	0.0	14	58.3	9.5
Rock	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Shrub	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1	25.0	35.3	3	75.0	11.5	0.0	0.0	0.0	0.0	0.0	0.0	4	100	3.7
Snow / Ice	0.0	0.0	0.0	0.0	0.0	0.0	3	75.0	19.1	1	25.0	100	0.0	0.0	0.0	1	25.0	3.9	0.0	0.0	0.0	0.0	0.0	0.0	4	100	9.9
Summit	0.0	0.0	0.0	0.0	0.0	0.0	6	35.3	23.8	2	11.8	100	0.0	0.0	0.0	4	23.5	7.7	0.0	0.0	0.0	3	17.7	75.0	15	88.2	15.0
Tree	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1	25.0	5.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2	50.0	1.1
Waterfall	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5	100	11.5	0.0	0.0	0.0	0.0	0.0	0.0	5	100	1.5
Wild Animal	0.0	0.0	0.0	0.0	0.0	0.0	1	11.1	7.1	0.0	0.0	0.0	1	11.1	17.7	3	33.3	11.5	1	11.1	10.5	0.0	0.0	0.0	5	55.6	11.3
Agriculture	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1	50.0	0.4
Infrastructure	0.0	0.0	0.0	0.0	0.0	0.0	53	26.2	100	2	1.0	100	17	8.4	88.2	34	16.8	53.9	20	9.9	100	2	1.0	75.0	161	79.7	83.9
Livestock	0.0	0.0	0.0	0.0	0.0	0.0	1	50.0	2.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1	50.0	2.9
Path / Trail	0.0	0.0	0.0	0.0	0.0	0.0	2	3.5	9.5	1	1.7	100	3	5.2	41.2	7	12.1	34.6	3	5.2	21.1	1	1.7	25.0	26	44.8	17.9
Urban	0.0	0.0	0.0	0.0	0.0	0.0	1	2.7	2.4	0.0	0.0	0.0	0.0	0.0	0.0	2	5.4	7.7	7	18.9	31.6	1	2.7	50.0	18	48.7	11.7

Table A.8.: *Count of LF within the threshold to a CES (N), its proportion of all annotations of a LF subtype (%LF) and the share of CES capturing N (%CES) for picture data (1).*

	Identity			Information Board			Information Office			Local History			Tradition			Traditional Architecture			Recreational Facilities			Signpost			Viewpoint		
	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES
Bedrock	6	4.0	42.9	17	11.3	100	0.0	0.0	0.0	2	1.3	100	0.0	0.0	0.0	2	1.3	13.6	8	5.3	41.7	22	14.7	68.2	45	30.0	65.1
Flower / Funghi	1	2.5	7.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2	5.0	4.6	1	2.5	8.3	2	5.0	9.1	2	5.0	3.6
Forest	20	6.4	92.9	15	4.8	66.7	3	1.0	100	2	0.6	94.4	1	0.3	50.0	12	3.8	68.2	35	11.2	70.8	45	14.4	86.4	108	34.5	97.6
Grass- and Moorland	28	7.6	92.9	12	3.3	33.3	3	0.8	100	16	4.4	94.4	3	0.8	50.0	18	4.9	77.2	43	11.7	95.8	57	15.5	100	134	36.4	97.6
Lake	2	14.3	14.3	2	14.3	33.3	0.0	0.0	0.0	1	7.1	94.4	0.0	0.0	0.0	0.0	0.0	0.0	7	50.0	25.0	3	21.4	13.6	7	50.0	33.7
River / Creek	0.0	0.0	0.0	5	15.6	66.7	0.0	0.0	0.0	2	6.3	94.4	0.0	0.0	0.0	0.0	0.0	0.0	1	3.1	4.2	0.0	0.0	0.0	2	6.3	1.2
Rock	3	3.7	14.3	7	8.5	66.7	0.0	0.0	0.0	8	9.8	94.4	1	1.2	50.0	1	1.2	13.6	5	6.1	8.3	8	9.8	18.2	22	26.8	27.7
Shrub	1	1.9	7.1	2	3.8	33.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	8	15.1	40.9	7	13.2	45.8	0.0	0.0	0.0	1	1.9	1.2
Snow / Ice	2	7.4	7.1	1	3.7	33.3	0.0	0.0	0.0	1	3.7	5.6	0.0	0.0	0.0	0.0	0.0	0.0	1	3.7	4.2	4	14.8	18.2	9	33.3	34.9
Summit	6	4.4	21.4	4	2.9	33.3	0.0	0.0	0.0	3	2.2	100	0.0	0.0	0.0	0.0	0.0	0.0	14	10.2	58.3	22	16.1	36.4	69	50.4	88.0
Tree	9	12.7	64.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6	8.5	50.0	12	16.9	62.5	9	12.7	36.4	11	15.5	18.1
Waterfall	0.0	0.0	0.0	4	57.1	33.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Wild Animal	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1	25.0	4.6	1	25.0	1.2
Agriculture	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1	12.5	22.7	0.0	0.0	0.0	3	37.5	13.6	2	25.0	4.8
Infrastructure	31	15.4	100	10	5.0	100	3	1.5	100	20	9.9	100	4	2.0	100	27	13.4	100	29	14.4	95.8	35	17.3	95.5	50	24.8	56.6
Livestock	1	5.9	7.1	0.0	0.0	0.0	0.0	0.0	0.0	2	11.8	94.4	0.0	0.0	0.0	1	5.9	4.6	7	41.2	20.8	2	11.8	9.1	5	29.4	8.4
Path / Trail	7	7.1	50.0	7	7.1	100	0.0	0.0	0.0	1	1.0	94.4	2	2.0	50.0	4	4.1	27.3	12	12.2	45.8	11	11.2	45.5	29	29.6	62.7
Urban	1	2.3	7.1	4	9.3	33.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5	11.6	40.9	8	18.6	29.2	6	14.0	9.1	11	25.6	27.7

Table A.9.: *Count of LF within the threshold to a CES (N), its proportion of all annotations of a LF subtype (%LF) and the share of CES capturing N (%CES) for picture data (2).*

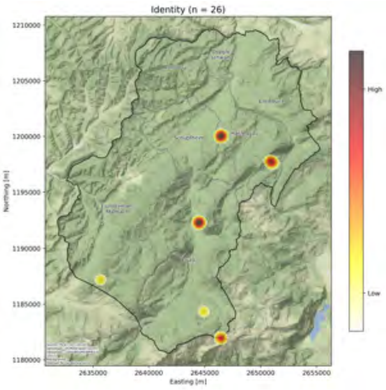
	Camping			People			Restaurant / Accommodation			Dawn / Sunset			Healing Powers			Place Attachment			Church			Summit Cross			Total		
	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES	N	%LF	%CES
Bedrock	0.0	0.0	0.0	46	30.7	57.8	5	3.3	33.3	9	6.0	56.3	0.0	0.0	0.0	0.0	0.0	0.0	1	0.7	598	14	9.3	45.5	91	60.7	53.2
Flower / Funghi	0.0	0.0	0.0	3	7.5	4.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	9	22.5	3.8
Forest	3	1.0	100	78	24.9	77.1	17	5.4	77.8	25	8.0	93.8	0.0	0.0	0.0	0.0	0.0	0.0	23	7.4	70.6	31	9.9	100	187	59.7	84.4
Grass- and Moorland	2	0.5	100	96	26.1	74.7	38	10.3	66.7	29	7.9	100	0.0	0.0	0.0	0.0	0.0	0.0	26	7.1	64.7	32	8.7	100	218	59.2	85.9
Lake	0.0	0.0	0.0	8	57.1	12.1	2	14.3	11.1	1	7.1	50.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1	7.1	9.1	12	85.7	23.2
River / Creek	0.0	0.0	0.0	13	40.6	20.5	2	6.3	5.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	16	50.0	12.1
Rock	0.0	0.0	0.0	28	34.2	25.3	11	13.4	22.2	1	1.2	6.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1	1.2	9.1	40	48.8	24.4
Shrub	0.0	0.0	0.0	11	20.8	28.9	2	3.8	11.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	9	17.0	35.3	0.0	0.0	0.0	24	45.3	16.5
Snow / Ice	0.0	0.0	0.0	4	14.8	4.8	2	7.4	5.6	1	3.7	50.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1	3.7	9.1	13	48.2	15.3
Summit	1	0.7	100	25	18.3	36.1	4	2.9	11.1	15	11.0	68.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	22	16.1	100	85	62.0	50.9
Tree	1	1.4	100	20	28.2	41.0	8	11.3	50.0	8	11.3	37.5	0.0	0.0	0.0	0.0	0.0	0.0	17	23.9	58.8	2	2.8	54.6	46	64.8	36.5
Waterfall	0.0	0.0	0.0	4	57.1	6.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4	57.1	2.1
Wild Animal	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1	25.0	0.6
Agriculture	0.0	0.0	0.0	1	12.5	1.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2	25.0	11.8	0.0	0.0	0.0	5	62.5	4.4
Infrastructure	1	0.5	100	76	37.6	67.5	53	26.2	100	13	6.4	87.5	0.0	0.0	0.0	0.0	0.0	0.0	38	18.8	100	6	3.0	81.8	155	76.7	79.7
Livestock	0.0	0.0	0.0	10	58.8	16.9	2	11.8	5.6	1	5.9	6.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1	5.9	45.5	13	76.5	15.9
Path / Trail	0.0	0.0	0.0	31	31.6	49.4	6	6.1	44.4	2	2.0	50.0	0.0	0.0	0.0	0.0	0.0	0.0	1	1.0	11.8	5	5.1	63.6	62	63.3	51.8
Urban	0.0	0.0	0.0	14	32.6	13.3	7	16.3	22.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	8	18.6	47.1	2	4.7	36.4	30	69.8	20.9

Table A.10.: *Absolute (N) and relative (%LF) densities of LF for each cluster and km^2 for text data.*

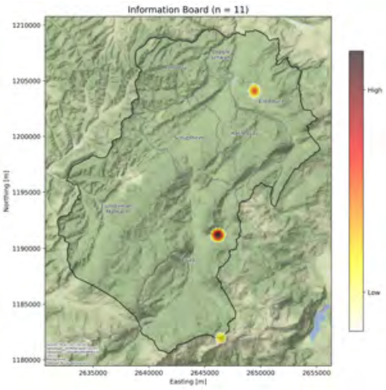
	Cluster 1		Cluster 2		Cluster 3		Cluster 4		Cluster 5		Cluster 6		Cluster 7		Cluster 8		Total	
	N	%LF	N	%LF	N	%LF	N	%LF	N	%LF	N	%LF	N	%LF	N	%LF	N	%LF
Natural Landscape	0.2	0.62	0	0	0.56	1.74	0	0	0	0	0	0	0	0	0	0	0.23	0.72
Human influenced Landscape	0.4	1.54	0	0	0.11	0.43	0.5	1.92	0	0	0	0	0	0	0	0	0.15	0.59
Bedrock	0.2	2.22	0	0	0.33	3.7	0	0	0	0	0	0	0	0	0	0	0.15	1.71
Flower / Funghi	0.2	1.82	0	0	0	0	0	0	0	0	0	0	0	0	0.33	3.03	0.08	0.7
Forest	0.8	2.05	0	0	0	0	0.5	1.28	0	0	1.0	2.56	1.0	2.56	0.33	0.85	0.42	1.08
Grass- and Moorland	0.8	2.76	0	0	0	0	0	0	0	0	0.33	1.15	0	0	0.33	1.15	0.23	0.8
Lake	0.2	1.25	0	0	0.44	2.78	0	0	0	0	0	0	0	0	0	0	0.19	1.2
River / Creek	0	0	0	0	0.22	0.93	0	0	1.0	4.17	0	0	1.0	4.17	0	0	0.19	0.8
Rock	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Shrub	0.6	15.0	0	0	0	0	0	0	0	0	0.33	8.33	0	0	0	0	0.15	3.85
Snow / Ice	0	0	0	0	0.22	5.56	0	0	0	0	0	0	0	0	0	0	0.08	1.92
Summit	0	0	0	0	0.22	1.31	0	0	0	0	0.33	1.96	0	0	0	0	0.12	0.68
Tree	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Waterfall	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Wild Animal	0.2	2.22	0	0	0.22	2.47	0	0	0	0	0	0	0	0	0	0	0.12	1.28
Agriculture	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Infrastructure	3.6	1.78	0	0	3.78	1.87	4.0	1.98	1.0	0.5	2.33	1.16	2.0	0.99	1.0	0.5	2.88	1.43
Livestock	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Path / Trail	0.2	0.34	1.0	1.72	0.11	0.19	1.0	1.72	0	0	2.0	3.45	1.5	2.59	1.0	1.72	0.65	1.13
Urban	1.2	3.24	0	0	0.78	2.1	1.0	2.7	0	0	0.67	1.8	0	0	0	0	0.65	1.77

Table A.11.: *Absolute (N) and relative (%LF) densities of LF for each cluster and km^2 for picture data.*

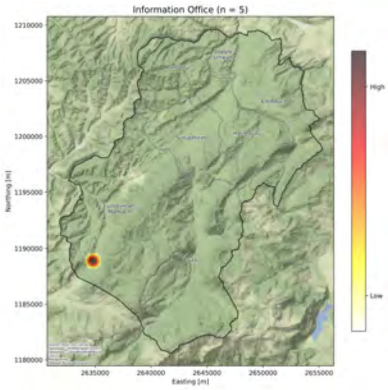
	Cluster 1		Cluster 2		Cluster 3		Cluster 4		Cluster 5		Cluster 6		Cluster 7		Cluster 8		Total	
	N	%LF	N	%LF	N	%LF	N	%LF	N	%LF	N	%LF	N	%LF	N	%LF	N	%LF
Bedrock	1.79	1.19	2.43	1.62	2.0	1.33	2.0	1.33	3.4	2.27	1.25	0.83	0	0	0	0	1.98	1.32
Flower / Funghi	0.53	1.32	0.43	1.07	0	0	1.0	2.5	0.2	0.5	0	0	0	0	0	0	0.38	0.94
Forest	2.95	0.94	1.57	0.5	10.5	3.35	3.0	0.96	5.8	1.85	6.25	2.0	1.0	0.32	1.0	0.32	3.67	1.17
Grass- and Moorland	4.58	1.24	3.29	0.89	13.0	3.53	5.0	1.36	5.6	1.52	7.25	1.97	1.0	0.27	1.0	0.27	5.0	1.36
Lake	0.16	1.13	0	0	0	0	5.0	35.71	0.6	4.29	0	0	0	0	0	0	0.28	1.96
River / Creek	0.21	0.66	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1	0.31
Rock	1.21	1.48	1.57	1.92	0	0	0	0	1.2	1.46	0.5	0.61	0	0	0	0	1.05	1.28
Shrub	0.11	0.2	0	0	2.0	3.77	2.0	3.77	0.4	0.75	0.25	0.47	0	0	0	0	0.28	0.52
Snow / Ice	0.32	1.17	0.86	3.17	0	0	0	0	0.2	0.74	0.75	2.78	0	0	0	0	0.4	1.48
Summit	1.21	0.88	1.71	1.25	3.5	2.55	2.0	1.46	4.4	3.21	4.0	2.92	0	0	0	0	2.05	1.5
Tree	0.63	0.89	0.14	0.2	2.0	2.82	5.0	7.04	0.2	0.28	1.0	1.41	0	0	0	0	0.68	0.95
Waterfall	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Wild Animal	0.05	1.32	0.14	3.57	0	0	0	0	0	0	0	0	0	0	0	0	0.05	1.25
Agriculture	0.05	0.66	0.14	1.79	1.0	12.5	0	0	0	0	0	0	0	0	0	0	0.1	1.25
Infrastructure	2.63	1.3	0.43	0.21	14.0	6.93	4.0	1.98	1.0	0.5	2.5	1.24	1.0	0.5	1.0	0.5	2.55	1.26
Livestock	0.16	0.93	0	0	0	0	0	0	0.2	1.18	0.75	4.41	0	0	0	0	0.17	1.03
Path / Trail	1.21	1.24	1.0	1.02	1.5	1.53	0	0	0.8	0.82	3.25	3.32	0	0	0	0	1.25	1.28
Urban	0.42	0.98	0.14	0.33	0.5	1.16	0	0	0	0	0.5	1.16	1.0	2.33	1.0	2.33	0.35	0.81



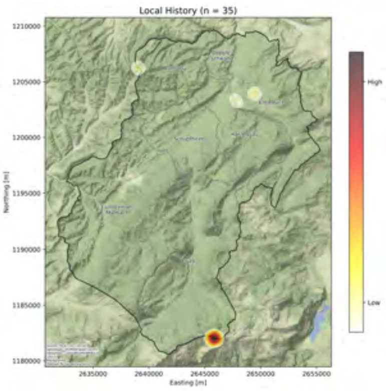
(a)



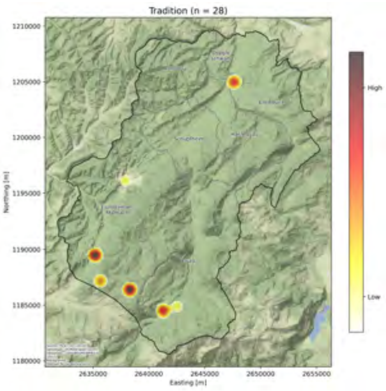
(b)



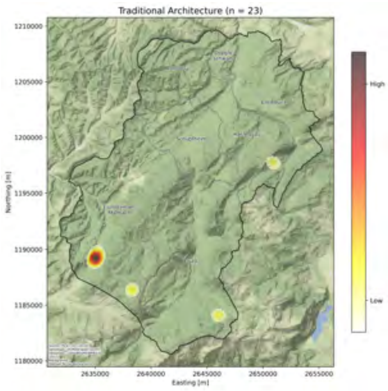
(c)



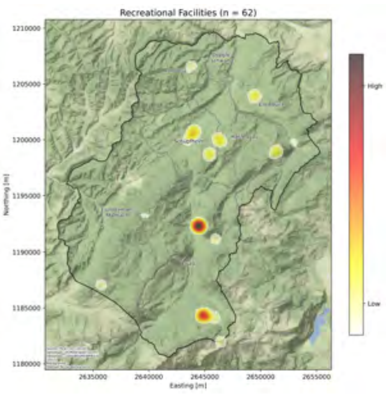
(d)



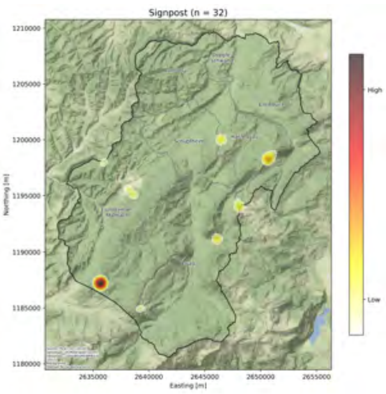
(e)



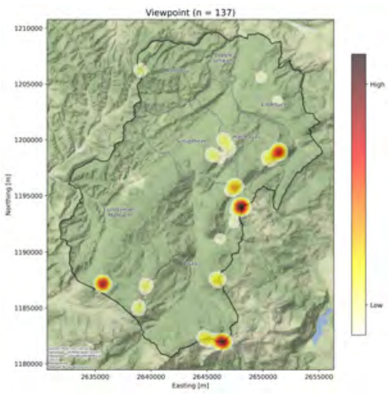
(f)



(g)



(h)



(i)

Figure A.2.: *Kernel density estimations of the combined data set for 9 subtypes: Identity (a), Information Board (b), Information Office (c), Local History (d), Tradition (e), Traditional Architecture (f), Recreational Facilities (g), Signpost (h), Viewpoint (i).*

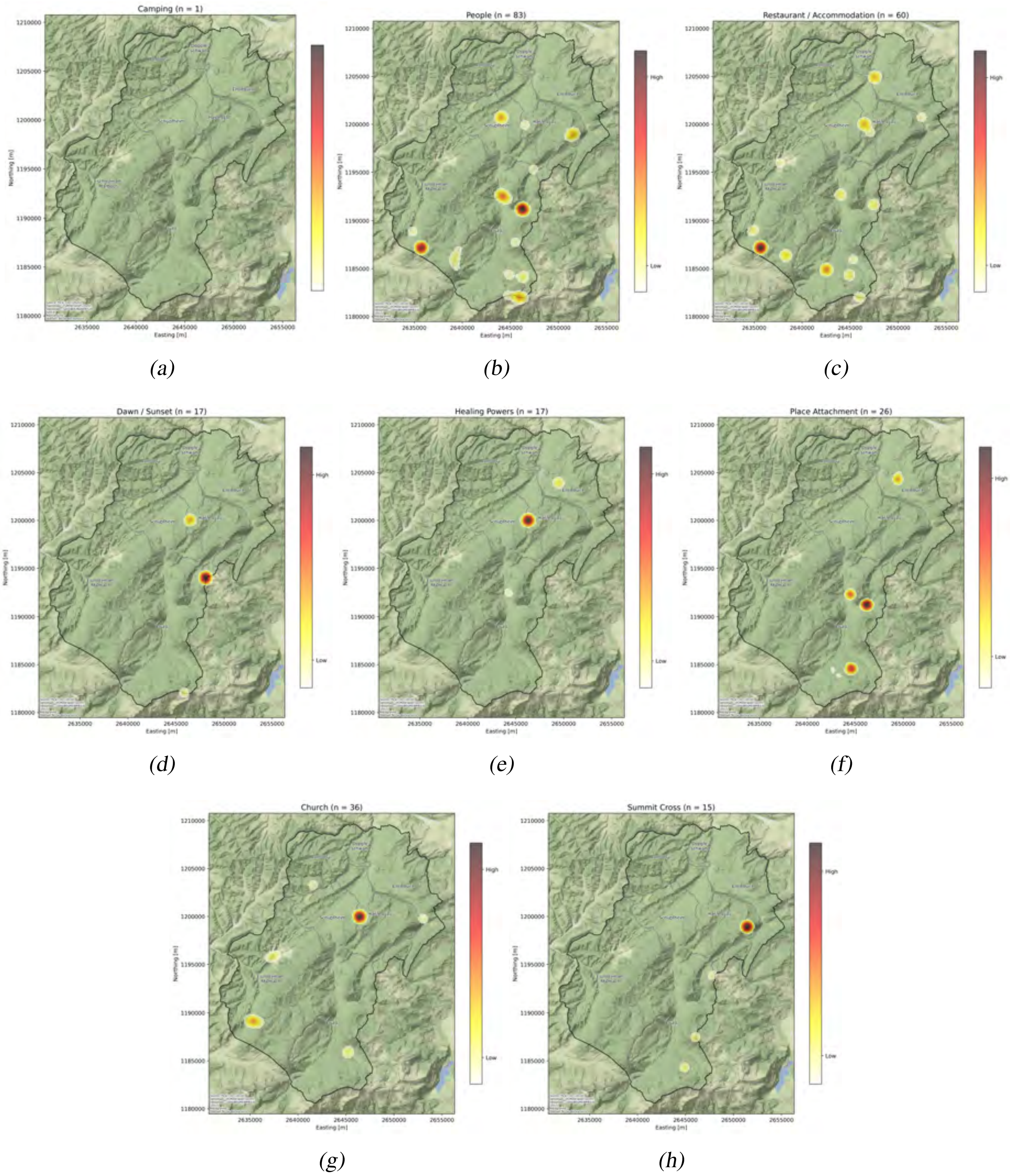


Figure A.3.: *Kernel density estimations of the combined data set for 8 subtypes: Camping (a), People (b), Restaurant / Accommodation (c), Dawn / Sunset (d), Healing Powers (e), Place Attachment (f), Church (g), Summit Cross (h).*

Personal Declaration

I hereby declare that the submitted Thesis is the result of my own, independent work. All external sources are explicitly acknowledged in the Thesis.

F. Biland

Fabian Biland, August 2021