The Cultural Eye. Exploring Cultural Differences in Landscape Preference based on Flickr Images.

GEO 511 Master's Thesis

Author
Tobias Frey
12-720-157

Supervised by
Dr. Olga Koblet

Faculty representative
Prof. Dr. Ross Purves

28.04.2020
Department of Geography, University of Zurich

«No matter how sophisticated you may be, a large granite mountain cannot be denied – it speaks in silence to the very core of your being» (Adams 1932: 1).

Abstract

Landscapes serve as constant backdrops of human life. Research has shown that the perception and preference of landscapes depends on three major factors: universal landscape ideas, personal experiences and factors common to larger groups. This study focusses on factors common to large groups, more precisely cultural differences and aims to determine to what degree social media imagery (Flickr) can be used as a proxy for landscape preferences. Building on existing work on landscape perception, the main objectives are (1) the geo-referencing of Flickr users based on their online photostream to locate them within a cultural setting and (2) to what extent automatically extracted landscape elements can be used to better understand cultural differences in landscape perception. In this context, culture is defined in simple terms as the country a user is living in.

Based on a review of the literature on landscape perception and preference, culture and user generated content, a semi-automated approach was proposed to geo-reference users. The method clusters image coordinates found from sampling their online photostream and reverse geocoding this information into a derived home country. The study draws on the YFCC100M dataset limited to the following five areas of interest: Jungfrau-Aletsch (Switzerland;), Dolomites (Italy), Geirangerfjord (Norway), Lake District (Great Britain), and Yellowstone National Park (United States of America). Potential landscape elements are extracted from the users' imagery using a machine learning application. While the geo-referencing of Flickr users yielded highly accurate results on a country level, the results obtained by multi-dimensionally scaling the extracted tags do not go beyond possible trends. Although there are slight signs of cultural differences, they remain limited by small sample sizes and the unequal distribution of derived home countries across the five areas of interest. Further research is needed that draws from larger samples and an ideally widened understanding of culture.

Keywords: *Landscape Preferences, Cultural Differences, User-Generated Content, UGC, Volunteered Geographic Information, VGI, Social Media, Flickr, Geographic Information Retrieval, Landscape Elements, Visual Recognition*

Max, this one is for you.

Acknowledgements

Submitting this Master thesis marks the end of my studies at the University of Zürich. I would like to thank everyone who inspired, motivated, and helped me throughout my studies and this thesis. Much of it would have been a lot harder without the support of a number of people. I am particularly grateful for the assistance given by:

Prof. Dr. Ross Purves & Dr. Olga Koblet; Thank you for all the time you spent discussing with me, always asking the perfect questions, and occasionally reminding me to focus on pragmatic solutions.

Lea Bacilieri, Corinne Jörg, Dr. Franziska Frey & Laura Markwalder; Thank you for proofreading at various stages of writing and providing the most valuable feedback.

My dear friends & family; Thank you for your continuous support and getting my mind off this thesis every once in a while.

Tanara, Ben Howard, and the Gardener & the Tree; Thank you for the music that accompanied me through countless hours of programming and writing.

Table of Contents

List of Figures

List of Tables

List of Listings

Acronyms

AOI	Area of Interest
API	Application Programming Interface
UGC	User Generated Content
VGI	Volunteered Geographic Information
CGI	Contributed Geographic Information
ML	Machine Learning
NP	National Park
DHC	Derived Home Country
DHL	Derived Home Location

Chapter 1 | Introduction

Landscapes serve as constant backdrops of human life. But they also have the power to attract millions of visitors a year and access to a scenic environment even promotes the healing of patients (Velarde et al. 2007). There is a wide field of potential benefits of landscapes, from tourism, recreation, and revenue to conservation (Figuerao-Alfaro & Tang 2016). With growing efforts to protect natural and scenic areas and the recognition of the need for spatial planning (Jongman 2002; Sarlöv Herlin 2004), it is crucial to better understand this complex interplay of landscapes and human life. The field of landscape perception and preference studies has a long tradition in theorizing and examining this relationship. While most research has been concerned with user studies, focussing on small glimpses into the relationship of a few cultures and their perception of landscapes, there has been little to no research on a wider cross-cultural scale. While some studies found signs for cultural differences, an equally large number concludes that no significant or relevant differences in landscape preferences were found (Buijs, Elands & Langers 2009; Herzog et al. 2000; Kohsaka & Flitner 2004; Purcell, Peron & Berto 2001; Yu 1994). Due to the diverse methods used and the barely overlapping cultures sampled, it remains difficult to compare and assess their findings. User generated content (UGC) has proven its ability to provide extensive data sources for research concerned with human-environment interaction. In combination with major advances in artificial intelligence, powerful image recognition solutions enable efficient ways to process and analyse imagery in UGC. The inscrutability of such machine learning solutions is often ignored.

1.1 Motivation and Goal

This study focuses on factors common to large groups, with a potential explanatory variable in the form of cultural background (Bourassa 1991: 55–57; Tribot, Deter & Mouquet 2018: 3). More specifically, the project focuses on possible cultural differences in landscape perception. Therefore, the use of Flickr imagery as a potential source in combination with machine learning as a method for finding cultural differences in landscape preferences is evaluated. Within five areas of interest (AOI), extensive samples of users are geo-referenced by deriving their home location. Subsequently, the location then subsequently serves as a cultural proxy to compare and analyse landscape elements automatically derived from the users' image collections. Therein, the quality of the derived landscape tags is of particular importance, to highlight the ambiguity in the unquestioned use of AI based results.

A better understanding of the factors which lead to a positive perception of an area enables improved decision making in the future. Such knowledge could for example be used for future policies in landscape planning and protection with further implications for tourism and health (Fenton 1985; Hausmann et al. 2018; Pan, Lee & Tsai 2014; Velarde, Fry & Tveit 2007). As Sevenant & Antrop (2010: 841) state: «It is important to understand the basis for differences in landscape preference and to fit this into theory. This may help to conduct a landscape policy and planning that seems less haphazard or untrusty and more closely reflects the needs and aspirations of the public, as is aimed at by the European Landscape Convention».

From a methodological perspective, the aim is twofold: first, to improve efficiency of landscape preference studies by proposing a semi-automated workflow based on social media data, and second, to better understand the extent to which Vision API can be used to extract landscape elements.

1.2 Structure

Chapter 2 introduces the theoretical background. Current and past scientific debates relating to the fields of landscape perception as well as potentially varying cultural preferences are discussed. This is crucial to the extraction of measurable units along which people's images are comparable against each other. With Flickr as the main data source, user generated content as an overarching topic is reviewed and introduced as a valuable source for machine learning applications. Thereafter, the relevant scientific research gaps are identified. Building on these, *Chapter 3* summarises and formalises the research gaps found. To close them, two research objectives, namely (1) an approach to geo-reference Flickr users and (2) the exploration of cultural differences in the preferences of landscape elements are defined. *Chapter 4* introduces the main dataset and the five selected areas of interest. The following *Chapter 5* explains the methodology to geo-reference Flickr users, extract landscape elements from their images and combine these two pieces of information to the final analysis of cultural differences in landscape preferences. *Chapter 6* subsequently presents the results. *Chapter 7* revisits the initially formulated research objectives, interprets and contextualises the found results within the scope of the earlier introduced literature, and stresses uncertainties and limitations of the applied methodology. Finally, *Chapter 8* concludes the study and points out future research opportunities.

Chapter 2 | Theoretical Background

This study draws from several fields of research and their concepts, which are introduced in the following. In a first step, the concept of *landscape*, *its perception, and preferences* are introduced. In a second step, there is an overview of *culture*, a prerequisite for the following discussion on culturally influenced landscape preferences. In a third step, as this study is based on Flickr images, *user generated content* (UGC) is discussed with specific information on *Flickr* as a social networking site building on UGC. In a last step and with a focus on the performed data analysis, basic information regarding *machine learning* and more detailed *image recognition* is given. While this chapter serves as a thematic introduction, it further introduces key literature to contextualise and enable an informed discussion of the results.

2.1 Landscape

First a current understanding of the *concept of landscape* is introduced. As human perception is an – if not ‹the› – important factor in the resulting definition, the second part introduces contemporary knowledge and literature on *landscape perception and preferences*. Specific attention is given to *cultural differences in landscape preferences*. Third, the notion of *landscape elements* – as a measurable sub-unit of landscapes – is introduced.

2.1.1 Landscape as a Concept

Landscape as a concept has changed quite drastically over the centuries. The early meaning of landscape could be described as «a composition of man-made spaces» (Jackson 1986: 7); a definition which lacks any associations with aesthetics or emotions as we would expect from its present everyday meaning. It was not until around 300 years ago when these two concepts started to heavily influence the understanding of what landscapes are (Jackson 1986). Therefore, in a modern understanding of landscape, it «is something that is mental as well as physical, subjective as well as objective» (Atha et al. 2019: xxi). This notion accurately reflects the vagueness of the term. Not surprisingly, many definitions for 'landscape' exist and are used in various fields of research (Atha et al. 2019: xx). But since its introduction, the

European Landscape Conventions' definition has become relatively widely adapted. There, landscape is understood as «an area, as perceived by people, whose character is the result of the action and interaction of natural and/or human factors» (Council of Europe 2000: 6). Such a holistic approach incorporates natural, cultural as well as scenic aspects (Antrop 2019: 9) and reflects well the mentioned contradictions of mental and physical as well as subjective and objective.

2.1.2 Landscape Perception and Preferences

The notion that «beauty is in the eye of the beholder» is a frequent generalisation based on the observable diversity of preferences. One could derive from such a statement that taste is a random phenomenon, «as variable as people are» (Kaplan & Kaplan 1989: 40). The same – but equally false – conclusion could be drawn from perception psychology's definition, which states that «different individuals may not affectively appraise the same environment in exactly the same way. Nor would the same individual at different times, nor different populations of individuals who have different backgrounds» (Russel 1988: 127). Fortunately, the (dis-)similarity of preference judgements is an empirically testable hypothesis that has been tested in a vast number of studies and has shown clear trends and influences from external variables (Kaplan 1979). Admittedly, the fact, that certain landscapes are perceived as more attractive than others is hard to deny. However, it is crucial to know the drivers of such differences in preference. Early research on human preferences in scenic quality of landscapes often drew from developmental biology theories. Hunziker (2000) identifies three main theories with this particular focus: (1) the *prospect-refuge theory* by Appleton (1975), (2) the *information-processing theory* developed by Kaplan & Kaplan (1989) and (3) the *savannah hypothesis* going back to Orians (1980, 1986).

(1) Appleton (1975) argues in his *prospect-refuge theory* that people prefer landscapes that meet the basic needs of human life. More specifically, this is protection, orientation as well as access to water and food. Several studies could, at least partially, confirm the hypothesis (Howard 2019: 43).

(2) The *information-processing theory* brought forth by Kaplan & Kaplan (1989) expects that landscapes that can be rapidly processed by humans are preferred more often. The assumption is that faster, easier processing of a landscape supports its navigation and use. To predict landscape preference, they developed the following four dimensions: complexity, mystery, coherence, and legibility. A preferred landscape «offer[s] rich exploratory opportunities (complexity, mystery)» and «must be well structured and easy to understand (coherence,

legibility)» (Arnberger & Eder 2011: 20). Mystery was found to be a strong predictor for landscape preference in many studies (Arnberger & Eder 2011).

(3) The well-known *Savannah hypothesis* is based on the homo sapiens hunter-gatherer past. The assumption is that landscapes that are favourable for such a lifestyle are perceived as more pleasing (Howard 2019: 52). Studies testing this theory have garnered mixed results. While Falk & Balling (1982, 2010; Ulrich 1977) could show a tendency for positive ratings considering such wide-open sceneries, other studies (Han 2007) failed to produce the expected results. But even if some results could show such a preference, the theory seems relatively limited in the types of environments to which it can be applied. Although typical mixed woodland-grassland ecosystems – like the savannah – were reported to be the most positively rated landscapes, they are not the only preferred landscape type, therefore other important factors seem to be excluded from this theory (Hunziker 2000).

Thompson (2019:22) states «that preference is unlikely to be based simply on biological or innate response to the environment». Similar critique is raised by other authors (Menatti & Casado da Rocha 2016; Buller 2009). A potential solution to the explanation of human preferences of scenic quality was forwarded by Bourassa (1991: 55–57). It still builds on universal landscape ideas that influence humans perception, but includes two other parts of influence: personal experience (Gregory 1997: 5) as well as factors common to large groups (i.e. nationality, culture or ethnicity) (Howard 2019: 51). This diversity of factors influencing landscape perception can be found in a quite similar manner in the work of Berque (1995), Cosgrove (1984), and Jackson (1984). An example for personal experiences is the familiarity with landscapes. Various studies have tested the hypothesis wheter people prefer what they know, assuming preference increases with familiarity of a landscape. The results show no clear positive or negative effect, but rather point to more underlying complexity (Nasar, 1988: 326). In conclusion, it can be stated that familiarity with the local natural setting has an effect on perception, but its effect on preference is mixed. As Kaplan & Kaplan (1989: 95) state: «Knowledge of the local scene does not assure increased preference, and the novelty of foreign places may enhance it».

The factors common to large groups shall be the focus of this study, although traces of personal experiences and universal landscape ideas factor into the results as well. One potential way of finding such factors, could be the segmentation of landscapes into a form of smaller units, for example landscape elements. A promising unit which has already been used in several studies related to landscape perception and preference (Seresinhe, Preis & Moat 2017, Conrad 2011) and will be introduced in *2.1.3 Landscape Elements*.

Although the focus of this study will be on visual cues, it is important to note that landscape perception is a multi-sensory experience (Mark et al. 2011: 2; Gibson 1986: 240). This includes the *olfactory* (Pennycook & Otsuji 2015; Quercia et al. 2015), *auditory* (Halonen et al. 2015; Andringa & Lanser 2013) and the much less researched *somatosensory* and *gustatory* senses (Brown 2017). Yet, visual material seems to be particularly effective in evoking related information associated with the information presented. This notion of the strong effects of visual stimuli could be interpreted as a manifestation of «the dominance of the visual mode without the necessity that it refers exclusively to visual information» (Kaplan & Kaplan 1989: 4). But as current research has shown, this is not universally applicable, as cross-cultural variations exist (Hutmacher 2019). Hence the dominance of the visual stimuli might need to be read in a *Western* context.

Another word of caution regards the temporal limitation of aesthetic preferences. There is evidence that – influenced by environmental awareness and education – relatively rapid changes in aesthetic preferences take place. Moreover, this seems also true for «long-term historical and cultural shifts in aesthetic appreciation for particular types of landscape[s] such as mountains» (Jorgensen 2011: 353). It is expected that such influences will be present in the analysed dataset, but only to a limited degree, as 99.9% of the photographs were taken within a 14-year range (2000–2014).

2.1.3 Landscape Elements

To computationally explore differences in landscape preferences, a measurable sub-unit of a landscape is needed. Tversky & Hemenway (1983) have found such **sub-units, or landscape elements** as named in the scope of this study, when they asked participants to describe photographs of landscapes. Rarely, the landscape was referenced as a whole. Therefore, such sub-units – at least in this context of *natural* landscapes – should be called landscape elements (or landscape features). Some examples are various land forms, trees or water (Lothian 2017, Derungs 2014).

Landscape elements are a concept that has been used widely in landscape research (Kaplan & Kaplan 1989; Tversky & Hemenway 1983; Seresinhe, Preis & Moat 2017). Brook (2019: 45) notes that «positive aesthetic value is sensed and appreciated subjectively – it can't be done by a machine – but the qualities appreciated are based on properties of the landform, vegetation, structures, climate and so on and their characteristic way of working together to create a whole». Clearly, this would reject methods that rely solely on landscape elements and ignore their interplay and other factors like climate. Nonetheless, research building on landscape elements has yielded promising results. It seems that at least a part of the appreciation can be explained by the collection of visible features. A complete scope of such landscape elements is most likely impossible to define, not only because of the number of possible categories but also due to its granularity. Potential, but incomplete, collections of such feature sets can be found in a number of studies that have looked at landscape elements and created various taxonomies, either based on theory or experiments (Tversky & Hemenway 1983). One of these approaches are so-called key response themes, which build on *contributors* to landscape characters (Conrad 2011). It is a taxonomy of six classes, consisting of: (A) *rural characteristics* (e.g. fields, unbuilt land, green areas, rubble walls), (B) *natural landforms* (e.g. hills, plateaus, valleys, clay slopes, cliffs, beaches, rocky coasts), (C) *cultural features* (e.g. churches, hilltop villages, prehistoric temples, historic towers, saltpans), (D) *specific locations* (toponyms), (E) *intangible aspects* (e.g. tranquillity, church bell sounds, smells, timelessness) and (F) *visual aesthetic qualities* (e.g. breath-taking views, coastal sceneries, landscape colours, natural beauty). Nearly all of these categories will be present in the data analysed. When looking at other taxonomies (*landscape components* in Lothian 2017; *basic-level scene categories* in Tversky & Hemenway 1983; *(visual) features* in Seresinhe, Preis & Moat 2017, *scene categories* in Zhou et al. 2017, *landscape elements* in Tieskens et al. 2018) it seems like there is a common understanding of the granularity of such elements, although it is never explicitly stated. This is no coincidence and can potentially be explained by a concept called *folksonomy*.

2.1.4 Folksonomy

Folksonomy is a compound word of *folks* and *taxonomy*. It describes collections of user generated, personally tagged information and objects in social (media) environments. A common example are collections of so-called Hashtags on Twitter. Similarly, this study will work with tags, but instead of user-generated tags, the focus will be on machine-generated tags (see *5.4 Extracting Landscape Elements*). Such tags can be used to filter or search the available information or objects (Deng, Chuang & Lemmens 2009), a process to alleviate the semantic gap that is created with otherwise mostly unstructured data (Mousselly-Sergieh et al. 2013). *Folksonomies* do not build on a rigorous structure as it is known from scientific taxonomies. Yet, the ags can normally be identified as either subordinate level, basic-level or superordinate level terms. These three classes follow – at least partly – the natural structure of language, close to our daily use (Jolicoeur 1984). Units – in our case tags – at the subordinate level are highly specific, show a low degree of class inclusion and feature identifiable and detailed gestalts. From a linguistic perspective, they often have the form of polymorphemic composite forms (e.g. compound nouns such as ‹glacial lake›). At the basic-level, units have a medium degree of generality, still feature identifiable gestalts and show a high degree of class inclusion. This is the level of abstraction on which most objects are identified at first (Jolicoeur 1984). At the top level, the superordinate level, units have a high degree of generality and tend to not feature a single identifiable gestalt (Ungerer et al. 1996). An example would be: ‹glacial lake› (subordinate level) – ‹lake› (basic-level) – ‹body of water› (superordinate level). Important in the context of this study is the notion that there is a level – the basic-level – on which people identify most objects in a first instance. This knowledge of folksonomy and various levels of terms will later be useful to discuss the quality of the machine-generated tags and compare them to other sources of tags (i.e. other machine learning algorithms and their tag sets).

Having shown how landscape can be understood and – from a rather naïve perspective – seen as the collection of its landscape elements, the next chapter turns the attention towards *culture*. A working definition is established and *cultural differences in landscape preferences* are discussed.

2.2 Culture

Although culture as a concept is widely used in our daily lives, its definitions are often vague and diverse. Various fields of research have developed their own understanding of culture, while clear definitions of such great and sensitive concepts are difficult to formulate appropriately (Günthner & Linke 2006).

On a very general level a definition as formulated in a dictionary could be expected to be well suited. For example, the definition of culture as «the customary beliefs, social forms, and material traits of a racial, religious, or social group» and «the characteristic features of everyday existence (such as diversions or a way of life) shared by people in a place or time (i.e. popular culture, Southern culture)» (Meriam Webster 2019). But, and this will be shown in the following, this is just one perspective on culture. Therefore, recent development in various fields working with culture as a concept are outlined. As a starting point, the focus lies on the field of geography. Here, since the mid 19th century (*modern era*, Schnädelbach 2004), culture has been seen as the counterpart to nature. While nature was understood as independent from cultural processes, it existed of objectively describable characteristics and laws which would be made accessible through the natural sciences. Therefore, any knowledge generated from such an understanding would be free of social and historical influences. Culture, on the other side, was the sphere of human activity, its creativity and freedom of thought. Such a strict separation was questioned with the introduction of post-structural theories in the 1960s. Present research therefore sees this dichotomy of culture and nature as a societal construct, a product of power and knowledge discourses (Zierhofer 2011, 1080). As it is not within the scope of this thesis to further elaborate on this discourse, this strand of thought is left here and the focus is shifted to other fields of research with more concrete understandings of culture.

The attention should be directed more towards the constitution of ‹culture› and how it is used in linguistics and ethnology. As with most concepts, its definition has changed over time. A broad overview of the understanding of culture can be found in Günthner & Linke (2006). Their introduction to linguistics and cultural analysis gives valuable insights into the relationship between language and culture. An early definition is formulated by Edward B. Tylor by the end of the 19th century:

> *«Culture or civilization taken in its wide ethnographic sense, is that complex whole which includes knowledge, belief, art, morals, law, custom, and any other capabilities and habits acquired by man as a member of society»*
> (Tylor, Edward B. 1871: 1).

Although language is not explicitly named as being part of it, it becomes quite clear that language must be the medium to transfer these capabilities and habits within society (Günthner & Linke 2006). A more recent definition is brought forward in the 1950s by Ward Goodenough. He understands culture no longer just as a collection of characteristics, but as the formation and form of its parts, constituted by the people's perception and interpretation:

> *«A society's culture consists of whatever it is one has to know or believe in order to operate in a manner acceptable to its members […] culture is not a material phenomenon; it does not consist of things, peoples, behavior, or emotions. It is rather an organization of these things. It is the form of things that people have in mind, their models for perceiving, reacting and otherwise interpreting them»*
> (Goodenough 1957: 167).

When focussing explicitly on landscape perception and preference research, the understanding of culture is almost always assumed to be general knowledge (i.e. Arnberger & Eder 2011; Qureshi, Breuste & Jim 2013). Although culture is used as an important factor, it is often not further explained or defined. Therefore, it can only be guessed how culture should be understood. Menatti et al. (2019) offer some insights into their understanding when speaking of «western vs non-western» culture, which gives an idea of the granularity used. For Buijs, Elands & Langers (2009), it is assumed that their understanding of culture is strongly related to ethnicity although this is never clearly stated in the study.

The diversity of definitions of *culture* presented and the implicitness of its use in current studies, show clearly, that *culture* can be understood in such a multitude of ways that a universal definition seems impossible. The only thing that seems to be constant is the notion that culture is embedded in society. If it would now be possible to define a granularity on which societies could be mapped out, the goal of matching location with cultural background could be achieved. But as literature on the granularity of society and on the extent to which culture covers the same spatial extension, is as diverse as the above discusses definitions, a simpler approach is needed.

Keeping in mind how culture is supposed to be used later in the data analysis, culture is understood at the granularity of countries. As shown, the scope of culture should be understood as much wider than the here presented proxy of *national culture*. As shown above, this is a rather obsolete concept, but the main reason for its usage is found in the data that will later be used to perform the analysis. For each Flickr user, the most likely home location will be determined, which serves as a proxy for this person's cultural background. This assumption seems to – at least partly – prove useful, as studies based on national culture have shown interesting results (see Buijs, Elands & Langers 2009 or Yu 1994) as did research based solely on language (Majid et al. 2018). But there is another issue with this rather bold assumption: even if such a thing like national culture exists, to what extent do people's home locations define their cultural belonging? This is a questions this study will not be able to fully answer, but which is further discussed in *7.2 Deriving Home Locations*.

2.2.1 Cultural Differences in Landscape Preferences

The following section focuses on the cultural differences in landscape perception, one of the three main drivers of landscape perception and preference as discussed above. The basic idea is that people «who share a system of though or perhaps a language (or dialect)» (Kaplan & Kaplan 1989: 86) perceive their surroundings, which includes landscapes, similarly and therefore might not share the same, but similar preferences. For the fact that these people live in the same environment, one could expect them to have «greater familiarity with its vegetation patterns and seasonal effects» (Kaplan & Kaplan 1989: 86), which again could influence their perceptions and preferences. Generally, many of these questions have only been looked at in small, highly-biased samples and most results are ambiguous (Kaplan & Kaplan 1989; Hägerhäll 2018).

Several small-scale studies (Kaplan & Herbert 1988; Buijs, Elands & Langers 2009; Herzog et al. 2000; Kohsaka & Flitner 2004; Purcell, Peron & Berto 2001; Yu 1994) have looked at cultural differences in landscape perception and found relatively mixed results. Many of them found much more similarities than discrepancies in the various cultures' perceptions and preferences of landscapes (Ulrich 1983). The set of explanatory variables in most cases covered personal experiences (e.g. education) as well as factors common to large groups (e.g.

nationality). In the case of personal experiences, all studies could identify significant differences, while for the factors common to large groups, the results were less clear.

On the one hand, there are studies that found weak to no cultural differences in landscape perception (Purcell, Peron & Berto 2001; Herzog et al. 2000). Other factors, for example education or experts versus non-experts, were better suited to explain the differences in ratings. On the other hand, the work of Buijs, Elands & Langers (2009) and Yu (1995) showed differing preferences for landscapes with culture as an explanatory variable. But, and this seems to be an important notion, the findings of Yu (1995) show variance in the weight of the cultural factors. The differences could be attributed to landscape types. More precisely, he states that «cultural influence in landscape preference is more likely to appear in specific landscapes that contain certain cultural meaning» (Yu 1995: 108). The study conducted by Kohsaka & Flitner (2004) found differences in the perception and preferences of landscapes dominated by forests. In this very specific landscape type, the preferences of German and Japanese people were compared. As mentioned above, cultural meaning of the landscape – in this case the forests' functionality – was found to explain the differences between the two groups.

An interesting conclusion drawn from such cross-cultural studies comes from Zube (1984), who concludes that «the similarity in cross-cultural evaluation is related to the similarity between the cultures» (Kaplan & Herbert: 379). The greatest differences should therefore be found in cultures very different from each other and their perception of strikingly contrasting landscapes. When applying this to the above discussed studies, this pattern can hardly be denied. All three studies (Buijs, Elands & Langers 2009, Yuu 1995; Kohsaka & Flitner 2004) compared samples of people from quite different cultural backgrounds, while those studies that found just weak differences (Purcell, Peron & Berto 2001; Herzog et al. 2000) worked with samples from Australia and America. Not surprisingly, this is also one of the main explanations brought forward for the weak effects in the study by Herzog et al. (2000).

The wish that «there were more: more research with large enough samples to validate the effects based on the groups included […] as well as more studies to explore differences based on many other comparisons», formulated by Kaplan & Kaplan (1989: 113) sums up the above discussed results. A similar and more recent demand is found in Hägerhäll (2018: 3). Clearly there is a need for studies featuring large samples of people from diverse cultural backgrounds. This could be achieved by the use of user generated content, more specifically the image archives available with social media platforms like Flickr that feature billions of images, of which at least some depict landscapes. Photographs, in the context of visitor employed photography (VEP), have been successfully used for many years in making «the abstract concept of user perception more tangible» (Cherem & Driver 1983: 81) and adding quantitative measurements to natural features. While for VEP the participants were asked to photograph anything they wish within a pre-defined space (e.g. a part of a trail), the photographs available on Flickr were created with many different intentions in mind (Cherem & Driver 1983). More on this in the next section on *user generated content* and a focus on *Flickr*.

2.3 User Generated Content (UGC)

To better understand Flickr as a social image sharing platform, it is essential to first focus on user generated content (UGC) in general and, more specifically, on volunteered geographic information (VGI). UGC and user created content (UCC) are treated as synonyms for the scope of this work. User generated content is understood as defined by the Organization of Economic Cooperation and Development (OECD). Their definition names three requirements for content to be UCC: «(i) content made publicly available over the Internet, (ii) which reflects a certain amount of creative effort, and (iii) which is created outside of professional routines and practices» (Wunsch-Vincent & Vickery 2007: 4). In many cases the users' contribution is voluntary (Bubalo, van Zanten & Verburg 2019) and the quality of the results will most likely vary (Goodchild 2007).

Volunteered/contributed geographic information (VGI/CGI) restriction of UGC to the domain of geographic information. On a basic level, these could be coordinates but of course toponyms can add locational information as well. Several technologies enabled VGI that transformed the web in the late 1990s and allowed a shift from the one-way information download from web pages (located on *servers*) to users (*clients*) towards the upload of information to such web pages and their respective databases. Enabled by this development, the users' role extended and thus anyone was enabled to become a producer and generate content that was in turn accessible to other users. Web 2.0 is the term that is now widely associated with this second generation of web services. Early examples for such web pages are Wikipedia, eBay, and Flickr (Constantinides & Fountain 2008). Data from such services has been used in research focusing on peoples' perceptions and interactions with landscapes (Bubalo, van Zanten & Verburg 2019; Dunkel 2015; Seresinhe, Preis & Moat 2017).

2.3.1 UGC in Landscape Perception Research

As already mentioned, the usage of photographs is not an entirely new approach to landscape perception and preference research (Bubalo, van Zanten & Verburg 2019; Cherem & Driver 1983). In the context of visitor employed photography, this has been done for over 40 years. Building on social media images to look at cross-cultural differences in landscape preferences takes this idea a step further.

It is essential to understand what makes photographs within user generated content a potential and valuable source of information. Therefore, a quick overview of key findings regarding landscape photographs in combination with landscape preferences is given. Photographs are inherently subjective as they summarize «*emotions, opinions, views and values of people in relation* [to] *landscapes and/or ecosystems*» (Bubalo, van Zanten & Verburg 2019: 102). From that perspective, Brooks' quote on the subjectivity of landscape perception and its impossibility to be measured by a machine found in *2.1.3 Landscape Elements* is rendered irrelevant. Because if a photograph is subjective and summarizes a person's values towards landscapes (which is equal to landscape preference), the step to affiliating values to landscapes is already taken. And algorithms (*machines*) can be used to process the images to gain insight into what might have led to the appreciation of the landscape depicted in that specific photograph. The question whether one can derive appreciation from the sole fact that an image is taken in a specific location is discussed in Gliozzo, Pettorelli & Haklay. (2016: para. 1). They state that «*the sequence of decisions and actions taken to share a digital picture of a given place includes the effort to travel to the place, the willingness to take a picture, the decision to geolocate the picture, and the action of sharing it through the Internet. Hence, the social activity of sharing pictures leaves digital proxies of spatial preferences, with people sharing specific photos considering the depicted place not only 'worth visiting' but also 'worth sharing visually'*». Similarly, Seresinhe, Moat & Preis (2018) treat photographs taken by photographers who travel as a proxy for their spatial preferences. While the direct derivation of landscape preferences from representative images created by the study author(s) is criticized in many studies, this problem is not present in this study as the images already hold the subjective information on landscape preference, as they were created by the participants themselves.

Various studies have made us of geo-referenced images, tags, and captions from social media to better understand landscape perception. Images and their captions have been used to build density maps to find spatial patterns of social and cultural landscape values (Chen, Parkins & Sherren 2018), to examine and rediscover the aesthetic attractions in Nebraska (Figueroa-Alfaro & Tang 2016), and assess landscape aesthetics in general (Oteros-Rozas et al. 2018; Casalegno et al. 2013; Depellegrin, Blazauskas & Vigl 2012; Tenerelli, Demsar & Luque 2016; Tenerelli, Püffel & Luque 2017; Seresinhe, Preis & Moat 2017; Tieskens et al. 2018). Data are mainly used with a strong focus on its locational component (Figueroa-Alfaro & Tang 2016; Chen, Parkins & Sherren 2018). The image content is often accessed by manually coding the available visual information (*distance zones* in Tenerelli, Püffel & Luque 2017; *indicators of landscape features* in Ozero-Rozas 2018; *presence of specific objects* in Barry 2014; *landscape elements* in Tieskens et al. 2018) and only in a few studies by fully-automated procedures (*features* in Seresinhe, Preis & Moat 2017). As Seresinhe, Preis & Moat (2017) have shown, such automated approaches are able to process large numbers of images and generate valuable information on their content. Therefore, the next section examines the application of machine learning to UGC.

2.3.2 UGC and Machine Learning

Automated approaches building on artificial intelligence have proven to be well suited for categorising, labelling, and analysing UGC. *Sentiment analysis* (Ye, Zhang & Law 2009; Chaovalit & Zhou 2005), *recommendation systems* (Nguyen, Wistuba & Schmidt-Thieme 2017; Schedl 2019), and *image categorization* (Abdullah, Veltkamp & Wiering 2009) are only a few selected fields of research to which machine learning has been successfully applied. The focus within this study is on object recognition, a computer vision technique with the goal to have the computer understand what an image contains (https://www.mathworks.com/solutions/image-video-processing/object-recognition.html). The basic steps to train and use a machine learning algorithm are the following: (1) Classification and feature extraction to simplify the image to important information which make up a feature vector; (2) Create training and test dataset of labelled image data. (3) To understand how image recognition is performed, it is essential to first understand the basic principle of machine learning. It treats each feature vector as a point in a high dimensional space. Within this space, the algorithm tries to find planes or surfaces that separate the categories found in the labelled data. (4) Each new input feature is transformed to its feature vector which is then categorised based on its relationship to the planes or surfaces. The resulting category is returned (Maruti TechLabs n.d.).

The machine learning application used in the scope of this thesis is the pre-trained Google Vision API. It is a commercial service provided by Google that returns for any image a set of tags which is supposed to reflect the features found in it (Google n.d.). As the Vision API is a commercial product, detailed information on its training data or tag sets are undisclosed (see *5.6 Methodological Limitations*). Despite everything, the high usability, low costs and high reported accuracy (81%, see Perficient Digital Agency 2019) make it attractive for the analysis of large collections of imagery. This is especially the case in the domain of landscape perception and preference research in which such labels can be used as a proxy for image content or ideally even landscape elements. In the case of this study, the Vision API is applied to photographs which were uploaded to the social media platform Flickr.

2.3.3 Flickr

The data used for this study consists of photographs, as well as their metadata including geo-references, extracted from the social networking and image sharing platform Flickr. The site was launched in 2004 and enables its user to upload, sort, bookmark, share, and comment images online. Having started as a side project of two game designers, the site has grown to 100 million registered users, is visited by 60 million people monthly, and features hundreds of millions of images (Flickr n.d. b). The platform itself is a source of user generated content and should be treated as such. Generally, approaches based on social media imagery, especially Flickr, have helped to better understand landscape aesthetics. Therefore, it is expected that such photographs in combination with locational information of its photographers could potentially give insights into cultural differences in landscape perception.

Although the Flickr dataset used is free to use for non-commercial applications, there are still a few legal and privacy concerns that should be considered.

2.3.4 Legal and Privacy Concerns

Working with social media data as a data source can raise legal and privacy issues. A brief overview of the potential concerns about the specific dataset used in this study is given in the following.

Legal concerns

As the analysed dataset, consisting of 100 million Flickr images and videos (more on this in *4.1 Flickr Dataset: YFCC100M*), is entirely compiled of Creative Commons licensed images, no legal issues should arise. Each image is attributed a Creative Commons commercial or non-commercial license. The full dataset, given by its license attribution, can be used freely for academic purposes (Choi, Thomee & Larson 2017).

Privacy concerns

Another issue are privacy and ethical concerns when working with social media data, especially data that includes peoples' locations (Mooney et al. 2017; Da Rugna, Chareyron & Branchet 2012). Since users' profiles as well as their image history will be used to derive a location for that person (which will have to be queried through the Flickr API), special attention should be payed to privacy issues. In the case of the derived home location (DHL) of Flickr photographers, single users are individually analysed and discussed. This is a typical example of data extraction from UGC of which the user is (or was) most likely unaware of when creating the data (Mooney et al. 2017). As locational data alone or linked with other data sources «can potentially expose sensitive private information, such as personal data, living habits […]» (Mooney et al. 2017: 120), such information should be treated accordingly. In those cases where processed and derived information for specific user profiles is shown, their user and image IDs will not be displayed. Fortunately, for most parts of the study, users will be grouped to larger aggregates (i.e. countries) which should sufficiently anonymize the data.

Chapter 3 | Research Gaps and Objectives

The review of the literature in the field of landscape perception and its cultural differences unveiled several research gaps. These are introduced in the following chapter, followed by an overview of the research objectives and the general methodology of this study.

3.1 Research Gaps

On a methodological basis (1), the broad usage of manually coded content analysis, which is the case in most studies, is seen as a research gap. Only a very small number makes use of automated processes and thus can handle larger and potentially less biased samples. This reveals the second research gap, (2), the small sample size found in most landscape preference studies. Many findings are built on less than 50 participants. The chances of biases introduced by such a limited number of people is likely greater than with more participants. Besides these ‹technical› research gaps, there are two more with substantial relevance for the topic itself: (3) the relevance of landscape elements as an explaining variable of landscape preference. Several studies have looked at this, with contradicting results. (4) The a few studies that focussed on cultural differences were based on two more or less distinct cultural groups (e.g. ‹western› versus ‹non-western›, Menatti et al. 2019). Studies comparing a greater number of cultures with each other do – to the knowledge of the author – not exist. The following study tries to close these four gaps, which are in brief summary:

(1) Manually coded content analysis instead of automated processes
(2) Small and therefore potentially biased sample sizes in most studies
(3) Landscape elements as a proxy to landscape preference
(4) A binary set of cultural differences instead of larger and more diverse samples

3.2 Research Objectives and Overall Methodology

The two research gaps focussing on the methodology (manual processing and small sample sizes) are not explicitly referenced as they are inherent in the automated processing of the

Flickr dataset. The first research objective focuses on a pre-processing step enabling the connection of users to cultural backgrounds based on their home location.

RQ.1 **How and with what accuracy can the home location of Flickr users be derived?**
- How can Flickr user's home locations be derived?
- What is the accuracy of such a method?

Based on the derived home location (DHL) each user's images are broken down into the landscape elements which are shown and extractable. A machine learning application is used to automate this task. As there is no information on the training data for the implemented application, the extracted tags are checked for their suitability in the context of landscape elements.

RQ.2 **What, potentially measurable, differences can be found in the preferences of landscape elements between groups of people with similar cultural backgrounds?**
- To what extent can machine generated tags (in the case of Google Vision API) be used for the extraction of landscape elements?
- What are the differences between the tag collection on a derived home country level?

The following chapter introduces the dataset on which the study is based as well as the areas of interest (AOIs) for which the images are examined. This is important, as the nature of these data is what ultimately determines which methods are useful and which are not.

Chapter 4 | Data and Area of Interest

4.1 Flickr Dataset: YFCC100M

The dataset used in this study, namely the Yahoo Flickr Creative Commons 100 Million Dataset (YFCC100M), was created in 2014 as part of the Yahoo Webscope program. It comprises 100 million media items. The data was published in 2014 and consists of 99.2 million images and 0.8 million videos uploaded in a ten-year range from 2004–2014. Close to all media items (99.4%) were created in the period from 2000–2014. The full dataset features, as afore-mentioned, exactly 100 million media files of which 48 469 829 (~48.5%) feature a geo-reference in the form of a coordinate pair (longitude/latitude). Their spatial distribution is shown in *Figure 4.1*. The YFCC100M dataset is entirely compiled of public and Creative Commons licensed media files and therefore it can be used for free and legally under the given license restrictions (Thomee et al. 2016).

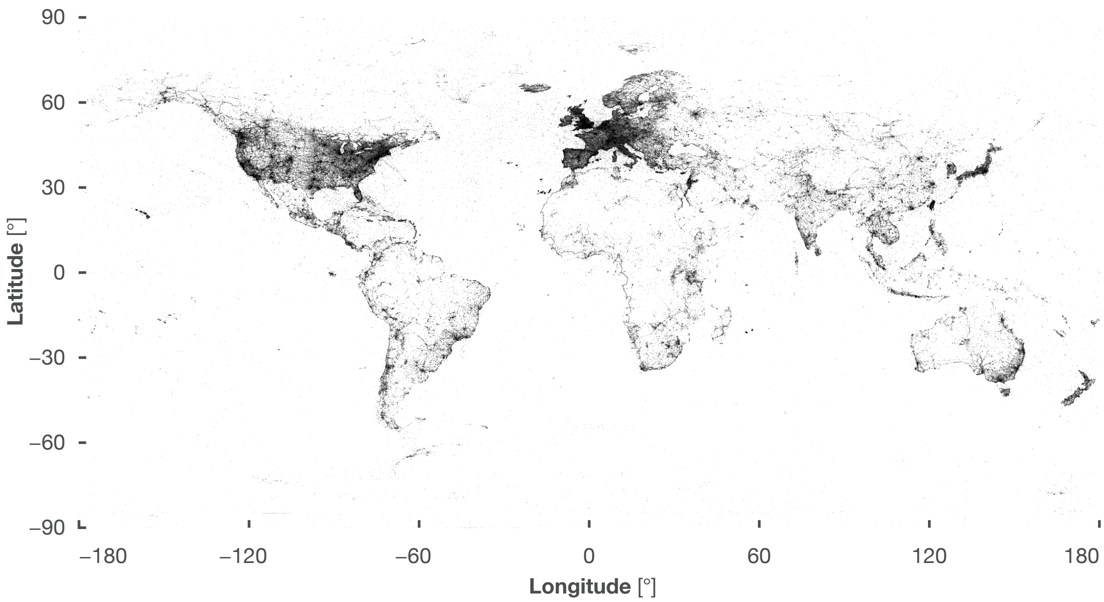


Figure 4.1: Global distribution of the roughly 48 million geo-referenced media items available in the YFCC100M dataset.

To make sure that the media files sampled for the YFCC100M dataset are representative of the full media pool on Flickr, Thomee et al. (2016) created a second sample of 100 million public media files and compared it to the first dataset. For various categories (e.g. cameras or locations), the relative frequencies in both samples were compared and showed an average difference of 0.02% with a standard deviation of 0.1%. Hence, it is assumed that that the YFCC100M dataset is a good representation of the publicly available Flickr image pool. It is important to mention that while the dataset represents well what is publicly available on Flickr, it shows serious distortions on a global level. Koochali et al. (2016) found that when normalizing the image counts for each country by their approximate population, the United States, Japan and most European countries feature the highest ratio values. The authors conclude that «this means, compared to the number of people living in those areas – thus representing their culture – those areas can be expected to provide the most comprehensive and accurate picture of their cultural identity» (Koochali et al. 2016: 37). However, the normalization with population numbers is problematic in two ways. First, it assumes a positive dependency of population and image count, while the more meaningful comparison would be a countries number of photographers normalized by its population. This is rather difficult, as the photographers' location is, if not voluntarily added to the profile information by the user, unknown and would have to be derived. The second issue is less obvious, but equally important. It considers the assumption that if two countries feature the same number of images and a similar population (or even photographer) number, they are equally represented. Such a narrow definition limited to quantity falls short of what should be understood as the most comprehensive and accurate picture of a cultural identity. As it is not the aim of this study to develop a more accurate measurement, it shall be concluded that the dataset is biased on a number of levels. This inherent characteristic of social media data should be kept in mind when interpreting results and draw conclusions.

4.1.1 Data Access

The data used in this study is a mixture of the static source of the YFCC100M dataset and the live information provided when queries are made through the Flickr API. Both ways of access are briefly outlined in the following.

YFCC100M

The YFCC100M dataset can be requested at Yahoo Webscope (Yahoo! Webscope n.d.) and by providing basic information on the planned project, access is granted to the files hosted on a cloud solution provided by Amazon (AWS cloud). Two projects, namely the YFCC100M Browser (Kalkowski et al. 2015) and MMCS (Krell & Li n.d.), have created web based applications to access large parts of the dataset. The core dataset (for which several extensions with additional metadata exist) with 51.8 GB data volume was downloaded and processed locally. As most of the analysis is performed on small subsets of the full dataset, data volumes are kept relatively small. The only exceptions are some basic exploratory data analyses on the full set and the data extraction for each area of interest (see *4.2 Areas of Interest*), which had to be run on a line-by-line basis to account for the massive data size.

Flickr API

For some parts of this study, user profile information is needed. Specifically, these are geo-references of images from people's photostreams and the location stated in their profile information, which is used for validation purposes. Access to this information is possible through the Flickr API or extensions that build on its functionality, such as the Python package *flickrapi*, which was used in the scope of this study. The advantage over the direct use of the Flickr API are several practical features, such as a simplified authentication process. The returned responses have the form of JavaScript Object Notation (JSON), a common exchange format in-between applications (Crockford 2017).

4.1.2 Data Structure

The YFCC100M is available in comma-separated, tabular data (.csv file), in which a single row represents one of the 100 million media files. Each item features 25 attributes, reaching from unique identifiers for each user (User NSID, referenced as *userID*), the devices used to capture the images (Capture Device), a media identifier (Photo/video identifier, referenced as *imageID*) to details on the file extensions. The full list of attributes (for the example image shown in *Figure 4.2*) is available in *Table 4.1*.

Figure 4.2: Image 4786526939 by user 72975926@N00 as an example for media items found in the YFCC100M dataset. Its metadata is found in Table 4.1. Image (http://www.flickr.com/photos/72975926@N00/4786526939/) by «Peter Schaer», licensed under CC BY-NC-SA 2.0.

Table 4.1: YFCC100M data structure with the metadata of image 4786526939 (shown in Figure 4.2) as an example. Several fields are empty, which is not uncommon for this dataset, as there are several user-provided fields which are optional.

#	Attribute	Example value
1	Line number	475699
2	Photo/video identifier	4786526939
3	Photo/video hash	4fdf593e7c18dfb9ad8122be47434c0
4	User NSID	72975926@N00
5	User nickname	Peter Schaer
6	Date taken	2010-07-10 09:47:21.0
7	Date uploaded	2010-07-12 18:14:43.0
8	Capture Device	Canon PowerShot G10
9	Title	Bach
10	Description	tanzboedeli
11	tags (comma-separated)	[no value]
12	Machine tags (comma-separated)	[no value]
13	Longitude	7.888011
14	Latitude	46.527748
15	Accuracy of the longitude and latitude coordinates (1 = world, 16 = street level accuracy)	12
16	Photo/video page URL	http://www.flickr.com/photos/72975926@N00/4786526939/
17	Photo/video download URL	http://farm5.staticflickr.com/4114/4786526939_0ecbc352c2.jpg
18	License name	Attribution-NonCommercial-ShareAlike
19	License URL	License http://creativecommons.org/licenses/by-nc-sa/2.0/
20	Photo/video server identifier	4114
21	Photo/video farm identifier	5
22	Photo/video secret	0ecbc352c2
23	Photo/video secret original	413a046783
24	Extension of the original photo	jpg
25	Photos/video marker (0 = photo, 1 = video)	0

For this study, the following attributes are of most relevance:

– ‹*Photo/video identifier*›: as a unique identifier for each media element, imageID is used synonymously
– ‹*User NSID*›: to identify the author of an image, which can then be used to gather all images of a single photographer, userID is used synonymously
– ‹*Longitude*› & ‹*Latitude*›: to geo-reference the media element, which is relevant to filter images for the various AOIS

– ‹*Date taken*›: knowledge of the date enables temporal analysis and validation

4.1.3 Geotagging

On Flickr, two geo-referencing processes exist (Thomee et al. 2016). The first option is a generic solution where the camera itself (e.g. smartphone) saves the coordinate information – based on GPS or similar positioning techniques – to the images. The stored location within an image's exif data is recognized and used by Flickr. This method is expected to yield relatively consistent results, of course depending on the accuracy of the positioning technique. The second option allows users to semi-automatically geo-reference images by indicating their location on an interactive map.

4.1.4 Demographics

Only very limited information is available on the demographics of Flickr users. Belyi et al. (2017: 1384) find that Flickr is mostly used in ‹developed countries› and usage in China (due to legal regulations), India and most of Africa is relatively low. Similar patterns are found in so-called social network analysis reports (i.e. Ignite Social Media 2012) listing the top eight cities found in Flickr data which are all located either in Europe or North America. Due to the restricted information on sources and details on their methodology, these results should be treated carefully.

4.2 Areas of Interest (AOIs)

In order to reduce the 99.2 million images to a more concise selection of landscape photographs, the study area is not set on a global scale. What follows are descriptions of the five AOIs that give a brief overview of their characteristics and the reasons for including them in the studied sample. Five areas of interest (AOIs) were identified based on the following requirements: The AOI should (1) feature a high proportion of landscape images, (2) attract many people from different cultural backgrounds as well as (3) feature a high number of images in the proposed dataset. Therefore, the project focuses on UNESCO World Natural Heritage Sites, which, due to their famous landscapes often attract many visitors from different countries and potentially different cultures (Adie & Hall 2017, López-Guzmán 2017). Out of the European sites that potentially fulfil the criteria, the following are selected: *Dolomites*, Italy; *Geirangerfjord*, Norway, and *Swiss Alps Jungfrau-Aletsch*, Switzerland. This is a very limited sample when looking at the more than three dozen UNESCO World Natural Heritage Sites in Europe (UNESCO World Heritage Centre 2019). All three have been selected carefully for distinct reasons explained below. Where available, additional information on the distribution of the visitors' countries of residence is given. Additionally, two more AOIs were added, the *Lake District,* UK and the *Yellowstone National Park,* USA. The AOIs locations are shown in *Figure 4.3*.

Swiss Alps Jungfrau-Aletsch, Switzerland (short: Jungfrau-Aletsch)

The Swiss Alps Jungfrau-Aletsch area (example imagery shown in *Figure 4.4*) is selected for the author's extensive local knowledge, which facilitates the interpretation of early study results. Furthermore, the region is known to attract a large number of tourists from Asian countries (Jungfrau Region Tourismus AG 2017: 23), which could facilitate the process of finding groups of various cultures for cross-comparison.

Dolomites, Italy
The Dolomites (shown in *Figure 4.5*) is selected as a second example of alpine landscapes to look at variations between – at least if it comes to typical landscape elements – relatively similar AOI. Compared to the Jungfrau-Aletsch sample, the area is much larger and thus gives an idea on how well the data processing pipeline scales.

Geirangerfjord, Norway
Still similar to the former two AOIs in terms of mountains and rock faces, the *Geirangerfjord* (shown in *Figure 4.6*) shows clear differences. For example, there are fjords as large bodies of water dominating the landscape. They are used by cruise ships to enter the area, which potentially make up for a large part of the tourist flow (Halpern 2007: 9). The confined movement

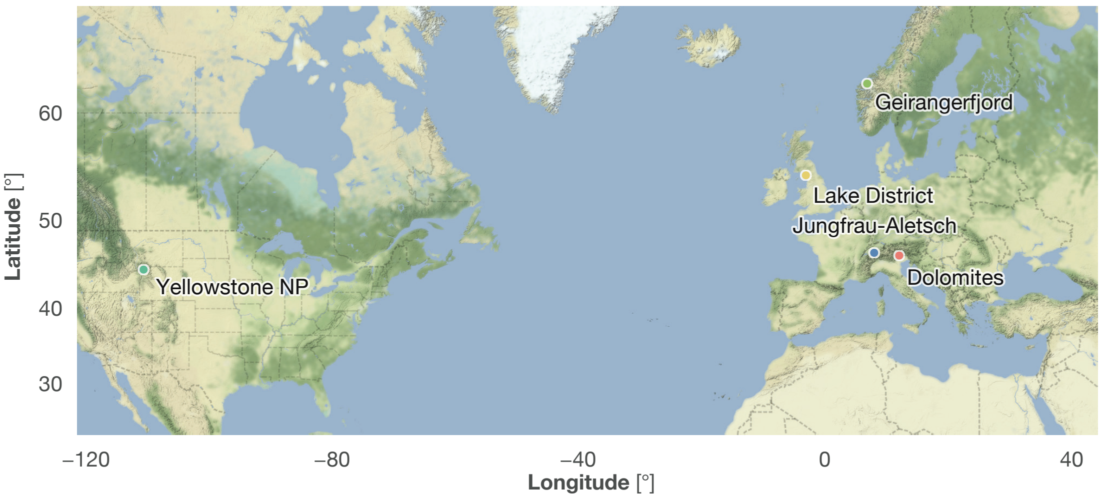


Figure 4.3: Map showing the location of all five AOIs: Jungfrau-Aletsch, Dolomites, Geirangerfjord, Lake District, and Yellowstone NP. (Map tiles by Stamen Design, under CC BY 3.0. Data by OpenStreetMap, under ODbL.)

space given by the relatively similar routes taken by these ships creates an interesting situation. Many users are forced to view the landscape from a more or less defined perspective and it will be interesting to see if this still leads to cultural differences in the landscape elements that are photographed. A study carried out by Halpern (2007) found that 62.3% of all overnights in the area are foreign. Within these group there is 32.2% from Germany, 13.9% from The Netherlands and 7% from Japan to name the nationalities with the largest shares.

To diversify the AOIs in terms of landscape types, the *Lake District, Great Britain* and *Yellowstone National Park, USA* were added to the sampled sites.

Lake District, Great Britain
The main reason to add the Lake District (shown in *Figure 4.7*) to the UNESCO World Natural Heritage Sites that are examined in this study is the profound knowledge of the area by Olga Koblet, supervising this thesis. Apart from that, it is also a mountainous landscape and somehow similar to the already introduced AOIs.

Yellowstone National Park, USA (short: Yellowstone NP)
This AOI was added for its distinctly different landscape to all other AOIs (shown in *Figure 4.8*). It is expected that this is also reflected in the landscape elements extracted from its images. Regarding the visitors' country of residence. a study by Manni et al. (2006) found the

Figure 4.4: Impression of the Jungfrau-Aletsch region. Image (http://www.flickr.com/photos/79463716@N00/170357275/) by «psmorrison», licensed under CC BY-NC-ND 2.0.

Figure 4.5: Impression of the Dolomites. Image (http://www.flickr.com/photos/75495759@N00/102988836/) by «Federico Pelloni», licensed under CC BY-ND 2.0.

following numbers: 90% United States of America, 2.5% Canada, 1.7% The Netherlands and 1% Germany.

Spatial Extents

Each AOI is described by a bounding box, a rectangular selection area marked by a lower left and an upper right coordinate pair. The main advantage of such a simple operationalisation is the reduced complexity (compared to a detailed polygon) in programmatically deciding if an image lies within an AOI or not, as it is done when building the subset for each AOI. Clearly, this is an abstraction of the much more detailed borders of any of the above introduced sites. But, considering the initial requirements for the AOIs, such a generalisation is insignificant. It expands the set of pictures derived for each site, which can, but does not have to, slightly reduce the proportion of photographs that depict landscapes. The goal of a potential focus on landscapes, diverse cultures, and a high image count is still met. For each AOI the bounding box was created manually (using *bboxfinder.com*, see Racicot 2019), based on mixed sources of information (UNEP 2017; UNESCO World Heritage Centre 2019). The selected coordinate pairs for each AOI, later used in *5.2 Data Subsets*, are shown in *Table 4.2*.

Table 4.2: Overview of the bounding boxes used for each AOI. The area covered by the bounding box is given as a measurement of scale.

AOI	Point LL (lower left)		Point UR (upper right)		area
Name	*Longitude*	*Latitude*	*Longitude*	*Latitude*	*km²*
Jungfrau-Aletsch	7.771736° E	46.395081° N	8.164189° E	46.662919° N	896
Dolomites	11.3861° E	45.6611° N	12.6769° E	46.779° N	12 375
Geirangerfjord	6.119° E	61.9858° N	7.5385° E	62.5379° N	4 537
Lake District	–3.5° W	54.25° N	–2.7° W	54.75° N	2 885
Yellowstone NP	–111.127° W	44.1235° N	–109.8364° W	45.1191° N	11 332

Figure 4.6: Impression of the Geirangerfjord. Image (http://www.flickr.com/photos/29059230@N00/12054468456/) by «Vvillamon», licensed under CC BY-SA 2.0.

Figure 4.7: Impression of the Lake District. Image (http://www.flickr.com/photos/32014910@N00/4121849/) by «Sarble», licensed under CC BY-NC-SA 2.0.

Figure 4.8: Impression of the Yellowstone NP. Image (http://www.flickr.com/photos/61111353@N00/2816506901/) by «dominiqs», licensed under CC BY 2.0.

Chapter 5 | Methodology

The overall methodological procedure is visualized in *Figure 5.1*. It can be divided into four main processes: (*5.2*) creating data subsets, (*5.3*) deriving a home location (DHL) and home country (DHC) for each user, (*5.4*) extracting landscape elements from photographs, and (*5.5*) the analysis of cultural differences in landscape preferences. Each of these steps features several smaller sub-processes which are explained in detail in the following sections.

5.1 Software and Scripting

Known for its rich environment of third party packages, Python was used for all data processing steps. This included the data extraction, the geo-reference of users, and the extraction of landscape elements from photographs through the Flickr API. For visualisation and statistical analysis of the results, R Statistics was used. Although Python supports similar functionalities, the author's knowledge in this specific domain is greater in R, which is the reason for this two-fold software strategy. *A.2.1 Software/Scripts* (Appendix) gives an overview of the applied software, including all packages extending the base functionalities.

5.2 Data Subsets

Based on the 100 million available media files ($\text{dataset}_{\text{original}}$), a $\text{subset}_{\text{original}}$ for each AOI is generated. This process builds on the two filter criteria, media type and location (given in longitude/latitude). Each media file within the YFCC100M dataset features a media type marker (0 = photo, 1 = film). By only selecting images that feature a 0, videos are ignored, as the project focusses solely on photographs. Regarding the location, a point in polygon query is used (see *Listing 5.1*) to select only those photographs that lie within the defined AOI (see *Table 4.2*). This approach builds on the available location information of the images (the case for nearly 50% of all media items within YFCC100M, according to Koochali et al. (2016)) and filters out all images that lack such information.

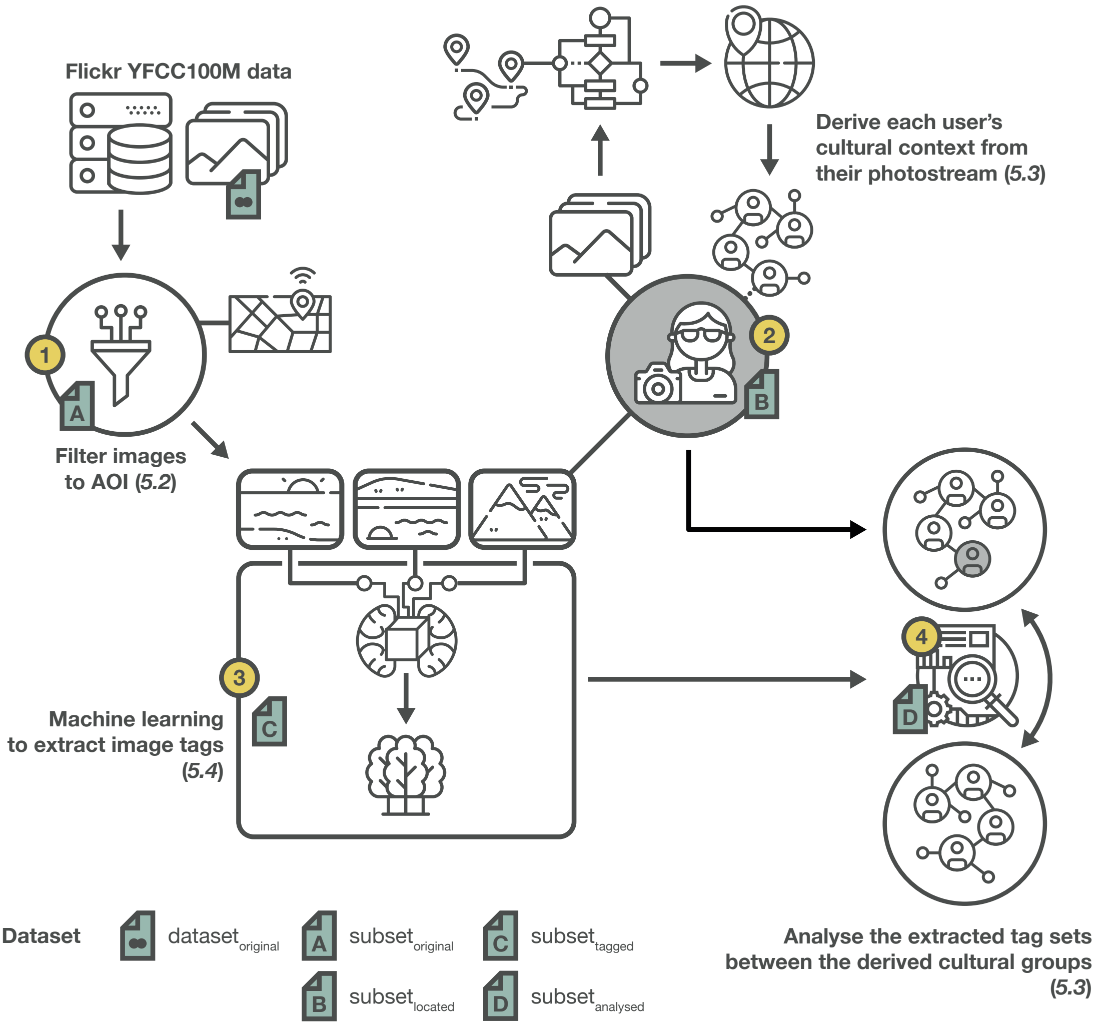


Figure 5.1: Overview of the main processing steps involved: (1) filter, (2) geo-reference user, (3) extract image tags, and (4) final analysis. (Icon sources are listed in A.3 Icons (Appendix)).

Listing 5.1: Python code used to determine if an image is located in the bounding box of the AOI or not. The input is given in the form of coordinates (longitude and latitude) representing the image location and the bounding box (bbox), a nested list in the form of [lower-left corner = [longitude, latitude], upper-right corner = [longitude, latitude]].

```
def within_bbox(lon, lat, bbox):
    is_within = False
    if lon != “” and lat != “”:
        if bbox[0][0] <= float(lon.strip()) <= bbox[1][0]:
            if bbox[0][1] <= float(lat.strip()) <= bbox[1][1]:
                is_within = True
    return is_within
```

It is important to point out that these initial subsets are further filtered based on the results of additional processing steps (as illustrated in *Figure 5.1*). Thus, all users without a DHL are excluded of further analysis and their images are excluded from further processing. In *5.5 Analysis of Cultural Differences in Landscape Preferences* the subset is filtered again, this time to 50 users per DHC and a maximum of 100 images per user for each AOI. Therefore, the following naming will be used from now on to refer to the various versions of datasets:

$\text{dataset}_{\text{original}}$	full YFCC100M dataset, featuring 100 million media items
$\text{subset}_{\text{original}}$	features all geo-referenced photographs within each AOI
$\text{subset}_{\text{located}}$	$\text{subset}_{\text{original}}$ limited to users for whom a home location could be derived
$\text{subset}_{\text{tagged}}$	$\text{subset}_{\text{located}}$ limited to a maximum of 50 users per country, 100 images per user and successfully extracted image tags.
$\text{subset}_{\text{analysed}}$	$\text{subset}_{\text{tagged}}$ limited to countries with a minimum of 10 users

5.2.1 Representativeness of Subsets

With a focus on the distribution of variables like *user counts* or *images per user* across the four levels of subsets, the Nearest Neighbour Index (NNI) is used to report on its spatial properties. The NNI is «the ratio of the actual mean distance between nearest neighbor points in the area to the mean expected distance of random distribution of the same number of points in the same area» (Naqshbandi, Fayaz & Bhat 2016: 16). A resulting index of 0 signifies clustered data, while values around 1 describe randomly distributed data and at the maximum of 2.15 the data is uniformly distributed (Clark & Evans 1954: 447).

5.3 Deriving Home Locations

To be able to look at cultural differences, each image within an AOI needs to be assigned to a cultural background. As discussed in *2.2 Culture*, the derived home country (DHC) of its creator is used as a proxy for this purpose. The basic idea is that by sampling a user's photostream (all available photographs of a user, but in this case limited to the publicly available ones) for geo-referenced images, a set of locations is extracted from which the DHL can be calculated. Based on such a DHL, a coordinate pair of longitude and latitude can then be reverse-geocoded to find the country it lies in. The finding, that users tend to upload an average of 50% of local photographs, created less than 100 Kilometres away from the user's home location, supports the above sketched approach (Hecht & Gergle 2010). An overview of the processing steps is provided in *Figure 5.2*. In order to not confuse the DHL with the user-provided spatial information within their profile, the following terminology is used: derived home location (DHL) or country (DHC) always refers to the derived location in the course of this study. This information is not given by the users, it is calculated based on their photostream. The user

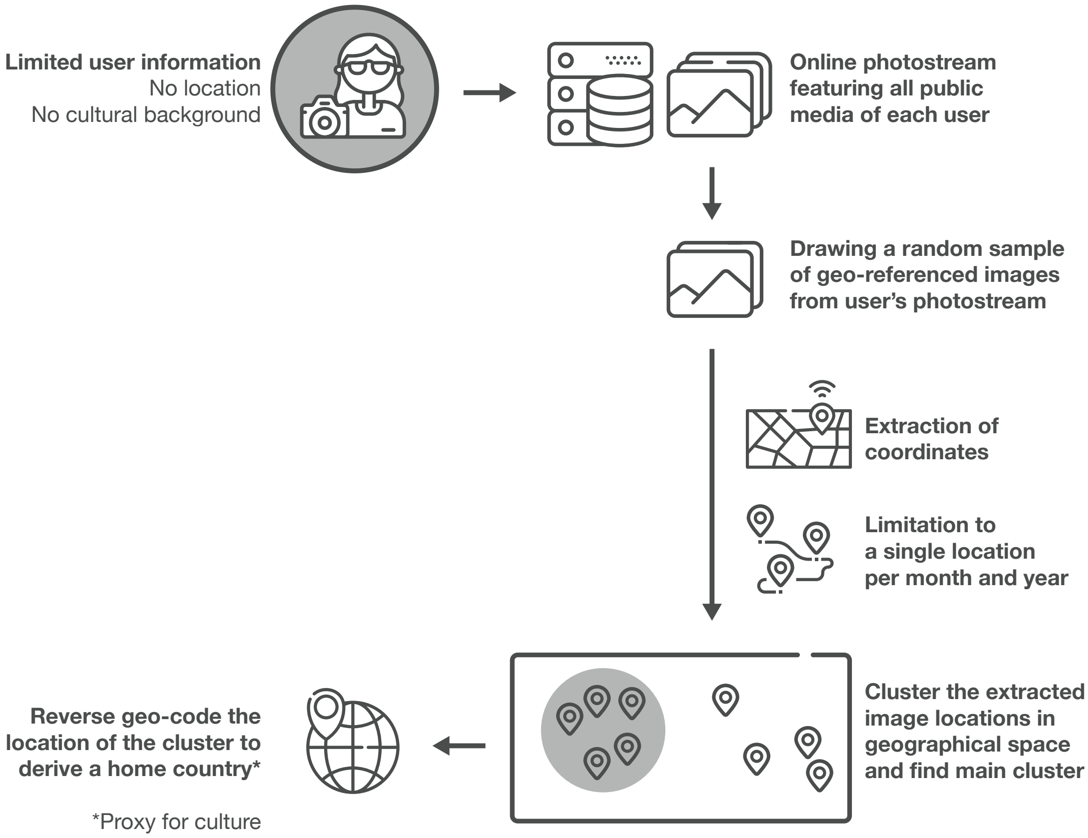


Figure 5.2: Overview on method to derive a user's home location.

location refers to the non-mandatory user-provided information that can be found in some of the user profiles (see *Listing 5.2*, "user location"). This user location is used as ‹ground truth› to test the DHL against.

5.3.1 Extracting Home Locations

A promising approach to derive users home locations is proposed by Popescu & Grefenstette (2010). Their algorithm looks at textual image tags to determine where images were taken for the greatest number of days (or time span if there is a tie). This information is then assumed to be the user's home location. Similar concepts have also been used in studies by Girardin et al. (2008); Da Rugna, Chareyron & Branchet (2012) and Belyi et al. 2017. With limited information on the actual implementations of their approaches, the following algorithm was designed.

Access User Metadata
In an initial step, the users Flickr profile meta information is accessed through the Flickr API. The input is a unique user identification (userID) as used by Flickr. Such a call returns detailed information on the user, including among some others: their real name, date of first upload, their time zone and most importantly for the current task, the count of their online photo library (see *Listing 5.2*). User profiles that are no longer available online are filtered out.

Listing 5.2: API response returning the user information of Flickr user «David G. Hong». The API call was made through the FlickrAPI with flickr.people.getInfo(user_id = "88943918@N00").


```
{"person": {
  "id": "88943918@N00",
  "nsid": "88943918@N00",
  "ispro": 0,
  "can_buy_pro": 0,
  "iconserver": "63",
  "iconfarm": 1,
  "path_alias": "davidhong",
  "has_stats": "1",
  "username": { "_content": "David G. Hong"},
  "realname": { "_content": "David Hong"},
  "location": { "_content": "Sydney, Australia"},
  "timezone": {
     "label": "Canberra, Melbourne, Sydney",
     "offset": "+10:00", "timezone_id": "Australia/Canberra"},
  "description": { "_content": "software engineer / asp.net / c# / javascript / cycling /
        skiing / onsen-ing / starcraft 2 / photography / travelling / food"},
  "photosurl": { "_content": "https://www.flickr.com/photos/davidhong/"},
  "profileurl": { "_content": "https://www.flickr.com/people/davidhong/"},
  "mobileurl": { "_content": "https://m.flickr.com/photostream.gne?id=2460059"},
  "photos": {
     "firstdatetaken": { "_content": "2009-08-15 17:59:10"},
     "firstdate": { "_content": "1303216758"},
     "count": { "_content": 137}
     }
  }, "stat": "ok
}
```

Sampling User Library
Based on their image count, the following processing involves randomly accessing images out of their library and checking these for attached coordinates. Unfortunately, random image extraction does not seem to be a common use case and therefore is not implemented or at least not accessible in the current Flickr API. To be able to draw random samples nonetheless, the circumstance of the Flickr API returning the results structured on pages and the knowledge of the users image count can be (mis-)used.

In a first step, a list containing the index of each image (sequence from 1 – the page numbering start at 1 and not as expected at index 0 – to the count of user images) is created and shuffled. It is used to randomly draw an image index to select an image. In a second step, the users image count is divided by 10 000, the maximum number of pages allowed (*Blog Post*), and the result rounded up. The resulting number is used to define how many images are displayed per page (*per_page* attribute, see *Listing 5.3*). For a user with 25 000 images (divided by 10 000 and rounded up) this would be 3.

(Listing 5.3) per_page = math.ceil(user_image_count ÷ 10000)

In a third step, the randomly drawn index is used to determine the value of the *page* attribute within the API call. By rounding up the result of the randomly drawn index divided by the *per_page* attribute, the *page* attribute is found (see Listing 5.4).

(Listing 5.4) index_page = math.ceil(index ÷ per_page)

For all users with an image count larger than 10 000 this page contains 2 or more images (except for the last page). For these cases the sub-index within the returned results of one page is found by taking the remainder of the division of the randomly drawn image index (corrected by –1, as the ‹within page› index starts at 0) and the *per_page* attribute (see *Listing 5.5*).

(Listing 5.5) index_sub = ((index – 1) % per_page)

Ultimately, the Flickr API is queried using the above discussed attributes along the *extra* attribute set to ‹geo, tags, date_taken› in order to include geo-references, tags used in images, and the creation date of the photograph in the API response. *Listing 5.6* shows an example of such a returned json. What gets returned is the information on two images with all the requested attribute groups in the API call. While the first image does not feature any geo-reference (latitude and longitude are both 0), a location (–33.9/151.1) is found. If in this case the second image is the randomly drawn one, its location as well as the corresponding imageID are saved. Independently of wheter coordinates were found or not, the next sample is drawn, as long neither of the two thresholds – sample size or maximum number of drawn samples – is reached.

This task is repeated for each user until the intended number of 30 coordinate pairs per user (except for Jungfrau-Aletsch with samples of 50 images and a few test cases in Geirangerfjord and Lake District), consisting of longitude and latitude, is collected or more than five times the intended number of samples (30 × 5 = 150) are drawn. This approach works for user profiles with up to five million images, as there is a maximum of 10 000 pages and 500 images per page (*flickr.people.getPublicPhotos* in Flickr n.d. a). As all the studied user profiles feature less images, this seems to be no serious limitation. Because there is a limitation of API calls per

Listing 5.6: API response returning the sampled photostream of Flickr user «David G. Hong». The API call was made through the FlickrAPI with flickr.people.getPublicPhotos(user_id = “88943918@N00”, per_page = 2, page = 3, extras = ‘geo,tags,date_taken’).


```
{‘photos’: {
  ‘page’: 13, ‘pages’: 69, ‘perpage’: 2, ‘total’: ‘137’,
  ‘photo’: [ {
      ‘id’: ‘8289724834’, ‘owner’: ‘88943918@N00’,
      ‘secret’: ‘f70dd27f2b’, ‘server’: ‘8361’, ‘farm’: 9,
      ‘title’: ‘I think Lucy is scared of all the lightening!’, ‘ispublic’: 1, ‘isfriend’: 0,
      ‘isfamily’: 0, ‘datetaken’: ‘2012-11-27 10:14:30’, ‘datetakengranularity’: ‘0’,
      ‘datetakenunknown’: 0, ‘tags’: ‘’,
      ‘latitude’: 0, ‘longitude’: 0, ‘accuracy’: 0, ‘context’: 0
  }, {
      ‘id’: ‘8289724820’, ‘owner’: ‘88943918@N00’,
      ‘secret’: ‘d840e9bf95’, ‘server’: ‘8072’, ‘farm’: 9,
      ‘title’: “Let’s see how you goes Outback!”, ‘ispublic’: 1, ‘isfriend’: 0,
      ‘isfamily’: 0, ‘datetaken’: ‘2012-11-25 07:51:10’, ‘datetakengranularity’: ‘0’,
      ‘datetakenunknown’: 0, ‘tags’: ‘’,
      ‘latitude’: ‘-33.863658’, ‘longitude’: ‘151.088485’, ‘accuracy’: ‘16’, ‘context’: 0,
      ‘place_id’: ‘MVMQ.NxTUL0i.z7KIg’, ‘woeid’: ‘22720667’,
      ‘geo_is_family’: 0, ‘geo_is_friend’: 0, ‘geo_is_contact’: 0, ‘geo_is_public’: 1}
  ]},
  ‘stat’: ‘ok’
}
```

time unit, the script includes timeouts whenever the maximum number of calls is reached. The collected images and, more importantly, their geo-references are now filtered.

Filtering Image Locations
In a next step, the resulting locations of each user are filtered to a maximum of one geo-reference per month and year. This data reduction step is necessary to reduce the bias of over-represented months, as it could for example be the case for vacations. The initial 132 153 images sampled for the 4 585 users are reduced by 50% to 67 842 images (see *Figure 5.3*). Due to the performed pilot study, the users in Jungfrau-Aletsch were processed with a maximum sample size of 50. No significant difference is found in using 40 or 50 samples, as discussed in *6.2.2 Validation of Derived Home Locations.*

Calculating the Home Location
In a last step, the home location is derived from the filtered datasets. A simplified approach based on the median of the found longitude/latitude (*Listing 5.7*) values already yields mostly plausible results.

(Listing 5.7) $lon = median(lon_{image\ 1}, lon_{image\ n})$ and $lat = median(lat_{image\ 1}, lat_{image\ n})$

But there are at least two issues that arise with this approach: (1) *edge effects* and (2) *non-existing locations.* (1) *Edge effects* are expected to be found along the edges of the dataset, which is around –180° and +180° longitude and –90° and +90° latitude. Another example should clarify the issue. Assuming again a one-dimensional vector of longitude values (–160°,–150°, 150°, 160°), the median is calculated to be at 0°. But in reality, we are dealing with a spherical, or in

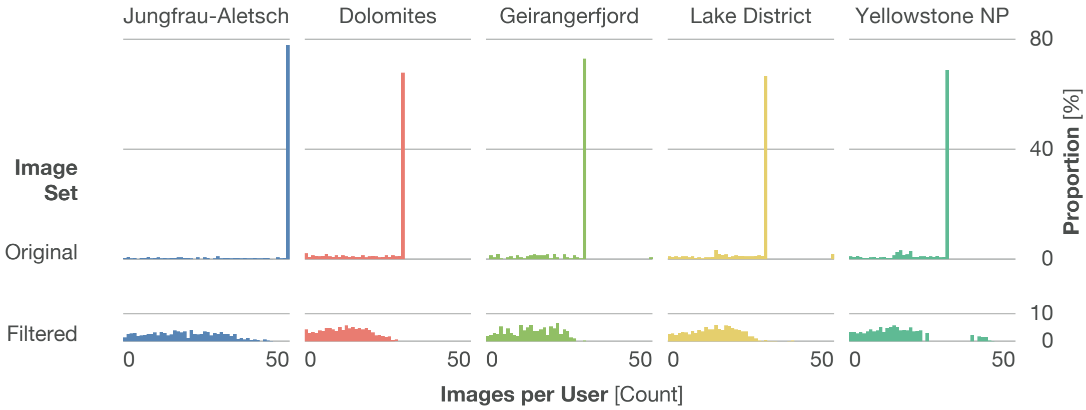


Figure 5.3: Number of sampled images for each user before and after reduction to a single image per month and year.

the simplified example presented here, a circular system where we should expect the value to lie at ± 180°. (2) *Non-existing locations* are generated by breaking the longitude/latitude pairs into two unrelated sets of longitude and latitude coordinates. As for each set the median is derived individually, the resulting longitude/latitude pair is most likely not a combination as seen in the data before. Only with a very small chance a pair results that exactly reflects one of the previous locations.

To solve the problem of (1) *edge effects*, DBScan was used to detect hotspots and select the main cluster (e.g. by the number of locations within each cluster). DBScan is well suited for several reasons: First, it is designed to detect and ignore outliers. This is important as for many users there are clear patterns of clusters, but also many outliers, in their image locations. Second, the number of clusters is not predefined. This gives the flexibility needed to respond to any spatiotemporal pattern found in the image locations. Third, the most important parameter to tune the method is the epsilon distance. This threshold value defines the distance for which points are still part of a cluster or outliers. A value of 200 kilometres was selected. The value is a trade-off between small, compact clusters and many users for which at least one cluster could be found. Fourth, the implementation used integrates the haversine formula (i.e. Alam et al. 2016), which solves the problem of *edge effects*. The disadvantages found are the slightly higher computational power needed for clustering (as compared to k-means) and the characteristic of DBScan to detect clusters of any form (i.e. a long straight line). In the case of home locations, more or less round clusters would have been preferred. Although, a visual inspection did not expose any cases in which this would have been a serious problem.

The second issue discussed above, the (2) *non-existing locations*, is left unsolved, as the increase in accuracy is expected to be minimal. A potential solution could use the centre of a cluster and find its nearest neighbour which would then be labelled as the DHL. But it is unclear if and how it would improve the method. Further examination would be needed.

To get an idea on the certainty of a position a different measure is used for each of the two methods. In the case of the DHL based on the median, the number of image locations used is taken as an absolute measure. The potential range of values is 1 to 50, as at least one image is needed and the maximum number of sampled images per user lies at 50. With more images, the derived location is less sensitive to the influence of a single image. The assumption is that this also increases the quality of the derived location. For the DBScan method another metric was used. It is a relative measure that states the number of images in the ‹home cluster› in relation to all images in a user's sample. Thus, the resulting value is found in a 0–100% range. A value of 100% would mean that all sampled images are found in a single cluster with no outliers.

Figure 5.4 gives an overview of the collected and derived data for two Flickr users. The spatiotemporal dataset is visualised in a space-time cube (Hägerstrand 1970, Kraak 2003). Thereby the horizontal plane represents the earth surface in a Mercator projection and the vertical axis displays time. The sampled images are shown as thin lines (each colour reflects a distinct user), while the DHL is accentuated with a thicker line and labelled with the user location found in their profile. These visualisations are therefore very helpful to quickly validate the quality of the DHL by providing valuable information on the full set of sampled images and their creation date. The creation date is important as it can give information on the timespan a user has been active in a certain area.

5.3.2 Deriving Home Countries

In order to translate the DHL (longitude/latitude) into DHCs, reversed geocoding was used. This was achieved through the geoprocessing library *reverse_geocode* available for Python (Penman 2018). For each coordinate pair inputted, the closest entry in «a database of known placesnames such as cities or villages along with their geographic coordinates» (Ahlers 2013: 74), a so-called gazetteer, is returned. The gazetteer used in *reverse_geocode* features 121 276

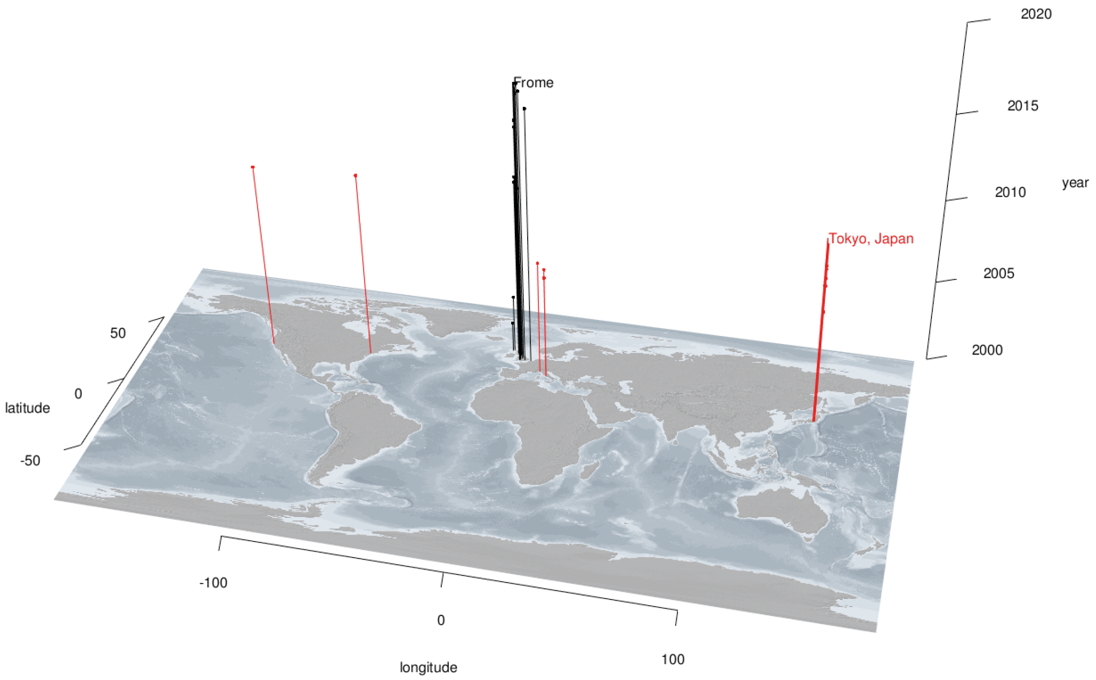


Figure 5.4: Space-time cube showing the sampled images as well as the DHL for two users. The thin lines each represent the geo-reference of a single sampled image. The labelled and slightly thicker line symbolizes the resulting DHL. The label shows the user-provided location in the profile information and were here used to quickly visually validate the results. (Map provided by Natural Earth, naturalearthdata.com.)

entries and is derived from the *Cities1000* subset provided by GeoNames, a popular, freely available gazetteer (Penman 2018, Ahlers 2013). Criteria for a listing in this subset are either a population greater than 1 000 or the seat of a first-, second- or third-order administrative division (GeoNames 2019). For more information on GeoNames and especially uncertainties and anomalies attached to the data, see Ahlers (2013). On one side, the change in spatial scale from a point location to a country is favourable. It potentially reduces some of the uncertainty attached to the DHL. On the other side, when dealing with relatively small countries or DHL lying close to borders, this could have negative effects. Because in such a case the location of a point within a country – and thus a clue for the certainty of its belonging to this very country – is lost.

In the validation process, a comparison on continent scale is interesting for the effect of further aggregation. Therefore, for each of the 150 samples this information is manually added for the given location in the user's profile as well as the DHL through DBScan.

5.3.3 Validating Home Locations and Countries

To validate the DHLs and DHCs, the users' profile information is used. The (spatial) accuracy is measured by comparing the DHLs to the information given in the user profiles. This is manually done for a sample of 150 users. The sampling strategy is designed to consider (a) the five AOIs, (b) the lack of locational information for some users with the DBScan method and (c) the sparse number of users with locational information in their Flickr profile. In a first step, all users' missing locational information in their profiles and for which the DBScan method could not derive a location are filtered out. The resulting dataset is then split up into the five AOIs and for each a random sample of 30 users is drawn. The resulting selection of 150 users is then manually validated. For each user, the locational information from their user profile is evaluated for agreement with the DHL and DHC from both methods on the three spatial granularities: city-, country-, and continent-level.

All images of users for whom no home location can be derived are excluded from the $\text{subset}_{\text{original}}$ of the analysis because they cannot be related to a cultural background. The resulting $\text{subset}_{\text{located}}$ is then inputed to the landscape element extraction described in the following chapter.

5.4 Extracting Landscape Elements

To make landscape preferences measurable and comparable across the samples, the earlier introduced concept of landscape elements is used. In favour of processing time, $\text{subset}_{\text{located}}$ is pre-processed and reduced to a more balanced sample. In the following step, the Vision API is used to extract tags which are then structured and analysed. Two prominent methods, namely co-occurrence and hierarchical clustering, are introduced. In the last step, multidimensional scaling (MDS) is outlined, which is used in the final analysis of cultural differences in landscape preferences.

5.4.1 Pre-processing

As the subsets are relatively unbalanced by image count per user and derived user country, only a selection of the images will be tagged. To limit a single user's weight within a sample, the maximum image count is capped at 100 images per person. This is slightly less than the 95th percentile, as 94.5% of the users found in subset$_{located}$ feature 100 or less images. As seen in *Figure 5.5*, the number of users featuring more than 100 images is small. The share on the full image count of an AOI for the users with more than 100 images is with 55.0% quite large (*Figure 5.6*). The challenge that a small number of users provide large amounts of data is often found in user generated content and was termed participation inequality by Nielsen (2006). The implemented solution to this problem is a simple, but potentially problematic, filtering process as outlined below.

All users providing more than 100 images are limited to a maximum of 100 images, resulting in a sample of their originally provided images. The problem that arises with this straight forward approach is two-fold. On the one hand, it assumes that users with high image counts do not provide useful data and, on the other hand, the overall data volume is drastically limited (Purves 2011). The usefulness of data might be debatable by the fact that most of these users with high image counts generated these within a very short time span. This is clearly visible in *Figure 5.7*, for which the images of all users with more than 100 images are aggregated by

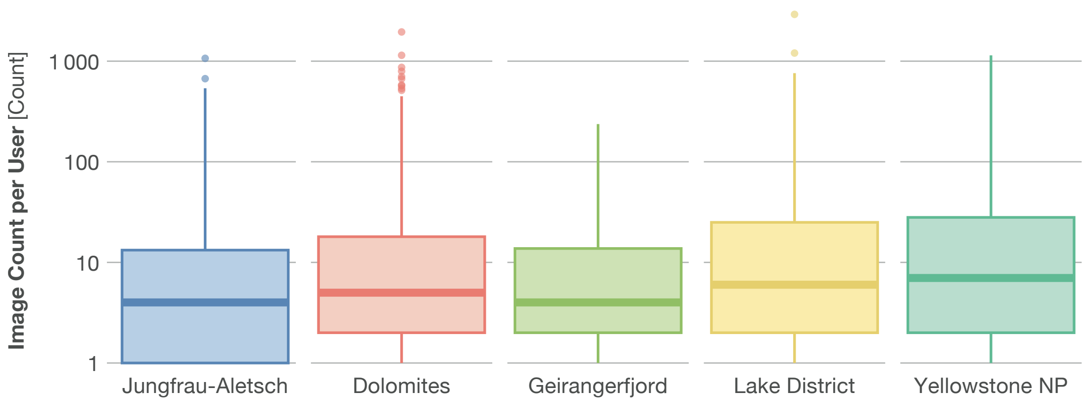


Figure 5.5: Image count per user by AOI. The y-axis is log10 scaled in order to increase readability.

month and year. The resulting sums are visualized as ‹bubbles› and users categorized into the timespans their images were taken in: 1, 2–5 or 6 and more months. It can be seen that for many of these users, a large amount of the images was created in just a few events. Only in the 6+ months group a few users are found that show constant contributions in the form of taken images. The consequences of these findings are not entirely clear, but it supports the idea that these are outliers. The reduced data volume is seen as an acceptable limitation in order to reduce the bias introduced by a small number of users, especially as many of these users created their images in very short timespans. As users with more than 100 images make up for 55.0% of all images, they are not fully excluded from the sample but limited to a maximum of 100 images (see *Figure 5.6*). With the limitation in place, there are 39 135 images (34.9%) in the subset$_{located}$ excluded from further processing.

Having reduced the subset$_{located}$ to a maximum of 100 images per user, it is ready for the next step: the extraction of potential landscape elements.

5.4.2 Google Vision API for Object Recognition

In order to access image contents and thus potential landscape elements, the study makes use of Google's Vision API, a pre-trained machine learning model that – along services like text or image recognition – assigns tags to images. Such tags are meant to describe what is shown in an image (Google Cloud n.d.; Google Cloud 2019), which is a typical implementation of object recognition. The advantage of such a fully automated approach is the precision with which objects are labelled. It is the same for all images, independent of their creator. This is important, as it is expected that not all users tag their images with the same thoroughness

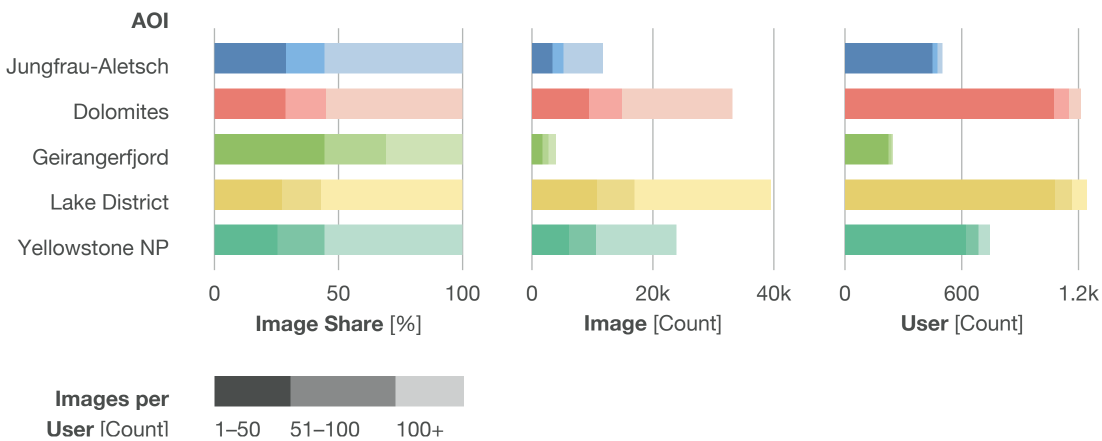


Figure 5.6: Participation inequality in subset$_{located}$. Most users account for 1–50 images. Shown are user counts as well as absolute and relative image counts per AOI.

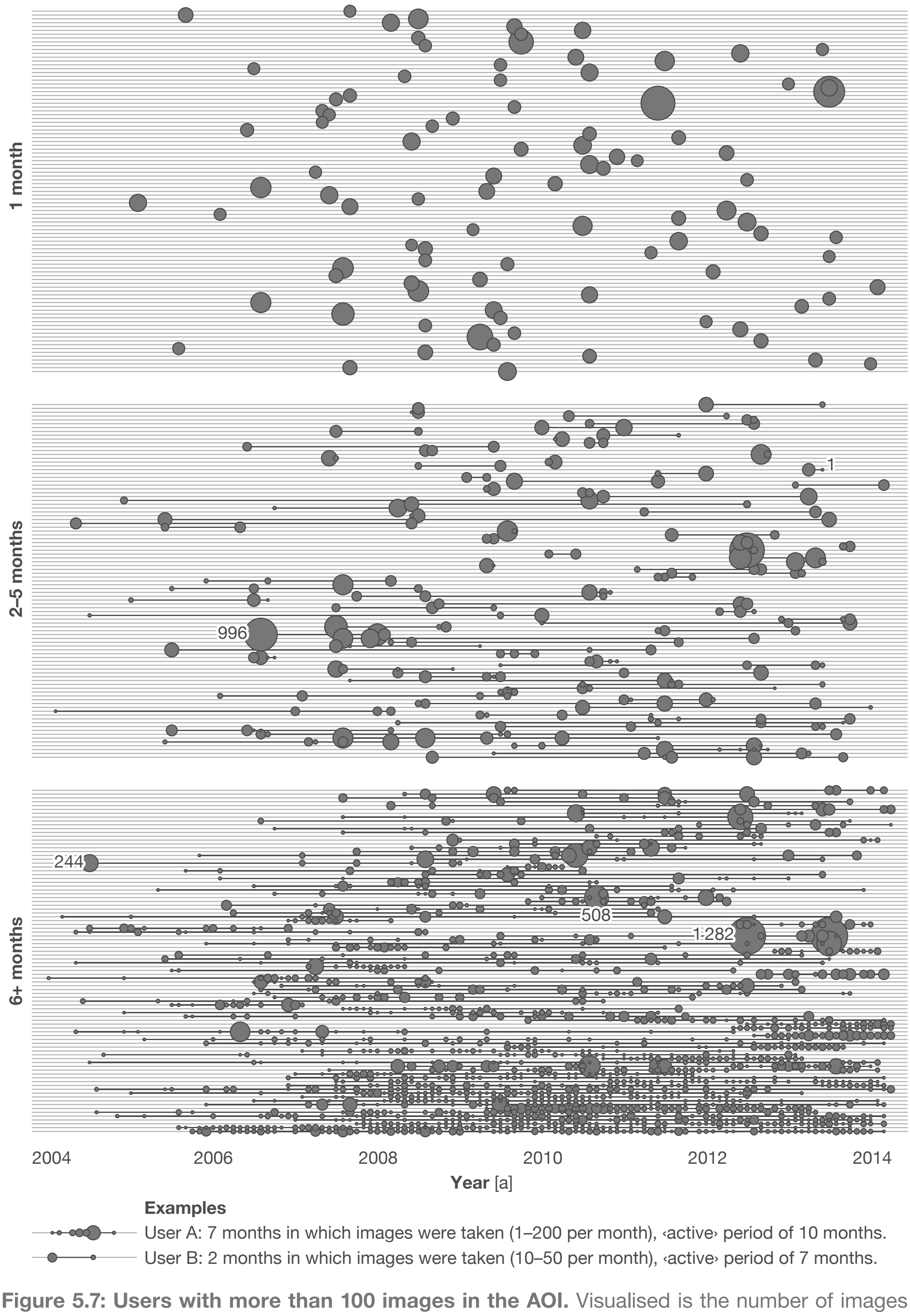


Figure 5.7: Users with more than 100 images in the AOI. Visualised is the number of images taken by each user over the period of 10 years (2004–2014). Around 2% of the images of users with more than 100 images that were created before 2004 or after 2014 are not shown to increase readability. For four out of the total of 282 users, this means that there is no data available in the visualised time span.

nor would they use the same taxonomy. Especially in the context of cultural differences, this could be an issue. Therefore, user generated tags are neglected and the analysis is built solely on these machine-generated tags. While such an approach is attractive for its efficiency and high precision, there are several issues that need to be treated carefully (see *5.6 Methodological Limitations*).

A common API call and its returned response in the common JSON format is depicted in *Listing 5.8*. The respective image is displayed in *Figure 5.8*.

The response features four attributes (Google Cloud 2019): (a) *MID*, which stands for machine-generated identifier; (b) a short *description* of the entity; (c) a *score* which expresses the confidence score (0–1); and lastly (d) topicality, the «relevancy of the Image Content Annotation (ICA) label to the image. It measures how important/central a label is to the overall context of a page» (Google Cloud 2019: para. 5). As of March 2020, there is a known bug which results in the exact same values of (c) and (d). It is unclear which of the two values is correct. The official issue tracker has listed this bug since 2018 (Issue Tracker 2018), but it has not been resolved or answered up to the day of writing [25.04.20]. Therefore, the attributes (c) score and (d) topicality are left unused in this study.

Figure 5.8: Image 24632641 (imageID) located in the AOI Jungfrau-Aletsch. Image (http://www.flickr.com/photos/20375052@N00/24632641/) by «josef.stuefer», licensed under CC BY 2.0.

Listing 5.8: API call and shortened response for image 24632641 (imageID) found in the AOI Jungfrau-Aletsch. Note that the called URL is slightly different from the original Flickr URL as credited in Figure 5.8. It refers to a static copy that was created when the YFCC100M was compiled.

```
CALL {
  “requests”: [ {
          “image”: {
              “source”:{
                  “imageUri”: “http://farm1.staticflickr.com/22/24632641_0b7a039d8e.jpg”
              }
          },
          “features”: [ {
              “type”:”LABEL_DETECTION”, “maxResults”:100
          } ]
      } ]
}

RESPONSE {
  “response”: [ {
      “labelAnnotations”: [ {
                  “mid”: “/m/0c9ph5”,
                  “description”: “Flower”,
                  “score”: 0.9955990314483643, “topicality”: 0.9955990314483643
              }, {
                  “mid”: “/m/04sjm”,
                  “description”: “Flowering plant”,
                  “score”: 0.9854584336280823, “topicality”: 0.9854584336280823
              }, {
                  “mid”: “/m/05s2s”,
                  “description”: “Plant”,
                  “score”: 0.9635068774223328, “topicality”: 0.9635068774223328
              }, […] {
                  “mid”: “/m/0fx65”,
                  “description”: “Gentiana”,
                  “score”: 0.5235731601715088,“topicality”: 0.5235731601715088
              } ]
      } ]
}
```

5.4.3 Tag Structures

Each image that could be tagged with the Vision API features a set of tags of which each tag is described by four attributes (*MID, description, score, topicality*, see *Listing 5.8*). While the score should give an idea on the certainty with which the object is recognized in the image, there is no information available on the relation between returned tags. As there is only very limited information on the tags and their structure (as discussed in *5.6 Methodological Limitations*), a closer look at the returned tags is vital. This is done by calculating the tags co-occurrence and autocorrelation. Based on this, a hierarchical clustering is performed, which partly reveals some of the underlying tag structures.

Co-occurrence of Tags

To better understand the relationship between the various tags, their co-occurrence can be studied. Co-occurrence describes the overlap of two tags (*do they overlap?*) while their correlation is a measures for the relationship's strength (*how much do they overlap?*). This means, if two tags are featured in the same image they are co-occurring (Aaron, Taylor & Chew 2018; Zhang et al. 2012). A so-called co-tag matrix can be generated that consists of counts for every tag on how many times it occurs together with each of the other tags (see *Figure 5.9*). Based on this matrix, their correlation coefficient (e.g. Pearson's r) can be calculated on the asymmetrical co-tag matrix (Deng, Chuang & Lemmens 2009; Mousselly-Sergieh et al. 2013; Leydesdorff & Vaughan 2006). The resulting values can then be used to colour a co-tag matrix or as an input to hierarchical clustering. Studies have used this approach to represent «possible conceptualisations of a place» (Deng, Chuang & Lemmens 2009: 54) or find tag groups (Mousselly-Sergieh et al. 2013).

Hierarchical Clustering of Tags

Various methods for clustering tags according to their similarity, in most cases co-occurence, are discussed in literature (Mousselly-Sergieh et al. 2013; Begelman, Keller & Smadja 2006; Zhang et al. 2012). While many methods result in a two-dimensional representation, several approaches result in three-dimensional representations within a vector space (Shepitsen et al. 2008). As the main reason for clustering within the realm of this study is the comprehensible ordering of tags for various visualisations, a two-dimensional approach is preferred. For its ability to show hierarchies and hence groups as well as their sub-groups, which are derived

	Apple	Pear	Mango
Mango	10	5	–
Pear	60	–	–
Apple	–	–	–

Figure 5.9: An example of an asymmetrical co-tag matrix. A high co-occurrence is found between the tags apple and pear. Apple and mango as well as pear and mango co-occur only ten and five times, respectively.

based on the interpretation of the complete tree, hierarchical clustering is used. It uses the distance between tags, namely their co-occurrence as a measurement of how closely related they are. By cutting the tree at a specific height, 1–n clusters can be generated, each containing n–1 elements.

5.5 Analysis of Cultural Differences in Landscape Preferences

With the analysis of cultural differences in landscape preferences, all of the above introduced processing steps come together. The pre-processed datasets for each AOI are combined with the per image information on the derived user home country as well as the extracted tags through the Google Vision API. To join the datasets together, image and user IDs serve as unique identifiers.

5.5.1 Visual Analysis

To get an overview of the extensive number of tags in combination with various users, DHCs, AOIs, and further grouping variables, an extensive visual analysis is performed based on various visualisations. The focus lies on the detection of irregularities and which of the tags could help to understand cultural differences in landscape preferences. It is expected that terms are unequally distributed over the five AOIs and DHCs.

5.5.2 Multi-dimensional Scaling

Multi-dimensional scaling (MDS) is a method that, based on a table of distances/dissimilarities between objects, outputs a lower dimensional representation of the objects in which their distances are preserved in the best possible manner. The above-mentioned table of dissimilarities in this study is given in the form of the correlation matrix of the co-tag matrix. A resulting representation, ideally two dimensional, could give valuable insights into individual to collective differences. The quality of the MDS is measured in a goodness of fit (GOF) statistic named *stress*. It is based on the differences between the predicted and the initial distances/dissimilarities (Kruskal 1964: 4). For the interpretation of the resulting stress values, *Table 5.1* gives an overview of a suggested verbal evaluation.

As the calculated and later visualised differences in between the objects are relative, their orientation is not fixed and varies with each calculation. Therefore, results can for example be flipped. To avoid such issues, all figures shown in *6.4 Analysis of Cultural Differences in Landscape Preferences* are the result of a single MDS run. All shown aggregations are performed afterwards. All MDS plots include a vertical and a horizontal line that run along the zero values of each dimension. This ‹cross› is thought to simplify interpretation by enabling descriptions in the form of «most users are found on the right» or «no users are found in the third quadrant», see *Figure 5.10*.

5.5.3 Sensitivity Analysis

It is expected, that the sample of DHCs for each AOI will be rather unbalanced, with some countries only featuring a handful of photographers. Therefore, equal samples for each DHC are drawn in order to work with the same number of photographers for each group. The main issue is that in potentially small samples, the weight of a single user can be quite drastic. A sensitivity analysis, looking at the influence of different samples, is performed to validate the MDS results. All results are validated through visual inspection.

5.6 Methodological Limitations

A number of methodological limitations arise, especially with the use of the Vision API. Overall limitations will be discussed in *7.3.2 Uncertainties and Limitations*. But in the case of the Vision API, there are a few restrictions that need to be addressed prior to its usage. These are (a) the linkage of tags and landscape elements, (b) the unknown training set, (c) the inaccessibility of the full tag set, and (d) reproducibility of the results.

(a) Linking tags and landscape elements

A major issue, as already demonstrated in the above example, is the diversity of tag categories provided by the Google Vision API. Clearly, many of these can be easily matched to known classes of landscape features (see *2.1.3 landscape elements*). Nevertheless, there are tags that cannot be directly associated with such features. More details on this unevenness are found

Table 5.1: Verbal evaluation of stress values, the goodness of fit (GOF) statistic used in MDS. (Adapted from Kruskal 1964: 4)

Stress [%]	Goodness of fit
20.0	poor
10.0	fair
5.0	good
2.5	excellent
0.0	perfect

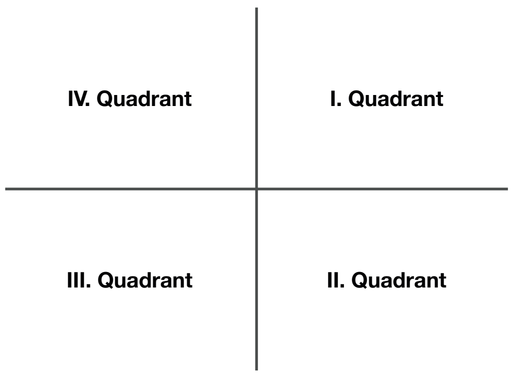


Figure 5.10: Quadrants in MDS plots to simplify descriptions of results.

in *6.3.3 Tag Structures*. The general reasoning is that although certain tags might not be classifiable by the given schemes as present in the literature, this should not disclose these tags from further analysis. The reason for this is that, as focussed on in RQ.2, it might be these tags that show large differences between cultural groups, for which the DHCs serve as a proxy. Therefore, the returned tags will be analysed for their compatibility with a popular classification of landscape elements, but non-landscape elements are not excluded for the final analysis.

(b) Training data unknown

As with nearly every trained machine learning model, it is intended to be exclusively applied to no-training data (Ghojogh & Crowley 2019: 5). In most applications of machine learning, such a situation is easily avoided by knowledge of the training data. The information is especially interesting as it potentially points out the bias introduced by the people who initially labelled the training data, as landscape elements are classified quite differently depending on language and culture (Mark et al. 2011: 5–6). In the case of the Vision API, no information is given on the training set(s) used. Therefore, a technical support request was filed on the 24th of November 2019 concerning the potential inclusion of the YFCC100M dataset in the training dataset. Unfortunately, the provided answer is unable to bring any clarity to this matter. The technical supports only state that they «[…] also have limitations to such information regarding the nature of the dataset where Vision API trains its models and its exact sources» (Google Cloud Platform Support, 2019). The initial suspicion was raised by the appearance of the tag ‹alps› in early tests with images from the Jungfrau-Aletsch region. But similar results – including the tag ‹alps› – were found for images taken within the Jungfrau-Aletsch region from the author's private image archive (that have never been published online) when run through the Vision API.

(c) Full category set unknown

Not only is the training data kept secret, the access to the full set of potentially returned tags is limited as well. In the release-notes it is stated that «label detection, which names objects inside an image, now recognizes more than 10,000 entities» (Google Cloud 2017). Further details are found in the documentation regarding the returned values, more specifically the machine-generated identifier (MID). The MID, «if present, contains a machine-generated identifier (MID) corresponding to the entity's Google Knowledge Graph entry. […]. To inspect MID values, refer to the Google Knowledge Graph API documentation» (Google Cloud 2019). No information was found on the extent of the graph, including an overview of available entities or even just the number of entities.

Chapter 6 | Results

The results are outlined in four sections. First, results of an exploratory data analysis of the full dataset as well as its subset are presented. This information provides basic knowledge of the data the following sections are built on. Second, the DHL and countries are evaluated (*RQ.1*). Third, an overview of the automatically extracted image tags is given. Together with the previously derived country information, they serve as the input for the last section regarding the potential cultural differences (*RQ.2*).

6.1 Representativeness of Subsets

As mentioned in *5.2 Data Subsets*, the study builds on several levels of subsets for each AOI. To better understand the filter and sampling process, several key metrics (i.e. user and image count) are used to describe the data reduction steps taken. While user and image count are available for each sample version, there are a few (i.e. count of derived countries) that can only be stated after a certain level of data processing. These metrics give valuable information on how the data was reduced and how much data was filtered out in which steps. A tabular overview is found in *Tables 6.1* & *6.2*.

The number of users per AOI decreases with each subset as data is filtered out for various reasons. Interestingly, the relative amount is quite different for the five AOIs. While for the Geirangerfjord the user count in the final $subset_{analysed}$ is still to 57.9% of the initial $subset_{original}$, it is reduced to 10.0% in the case of the Lake District. Another relatively high percentage is observed for Jungfrau-Aletsch (42.5%), while Dolomites (20.3%), and Yellowstone NP (12.7%) filtered out more than three out of four users.

6.1.1 User Characteristics

Table 6.1 shows that the number of users is constantly reduced in each subsetting step. To understand why exactly users are filtered out at each step, the metadata generated in the process of locating users and tagging their images is analysed. *Figure 6.1* gives an overview on

Table 6.1: User counts for each subset and AOI. Percentages given in relation to subset$_{original}$.

	Jungfrau-Aletsch		Dolomites		Geirangerfjord		Lake District		Yellowstone NP	
	Count	*%*	*Count*	*%*	*Count*	*%*	*Count*	*%*	*Count*	*%*
subset$_{original}$	638	100.0	1 641	100.0	318	100.0	1 580	100.0	1 019	100.0
subset$_{located}$	500	78.4	1 213	73.9	246	77.4	1 243	78.7	744	73.0
subset$_{tagged}$	367	57.5	428	26.1	230	72.3	249	15.8	217	21.3
subset$_{analysed}$	271	42.5	333	20.3	184	57.9	158	10.0	129	12.7

Table 6.2: Image counts for each subset and AOI. Percentages given in relation to subset$_{original}$.

	Jungfrau-Aletsch		Dolomites		Geirangerfjord		Lake District		Yellowstone NP	
	Count	*%*	*Count*	*%*	*Count*	*%*	*Count*	*%*	*Count*	*%*
subset$_{original}$	14 571	100.0	43 212	100.0	4 901	100.0	44 988	100.0	29 590	100.0
subset$_{located}$	11 703	80.3	33 108	76.6	3 945	80.5	39 530	87.9	23 912	80.8
subset$_{tagged}$	7 057	48.4	9 912	22.9	3 397	69.3	6 485	14.4	6 620	22.4
subset$_{analysed}$	5 566	38.2	8 850	20.5	2 677	54.6	4 859	10.8	4 806	16.2

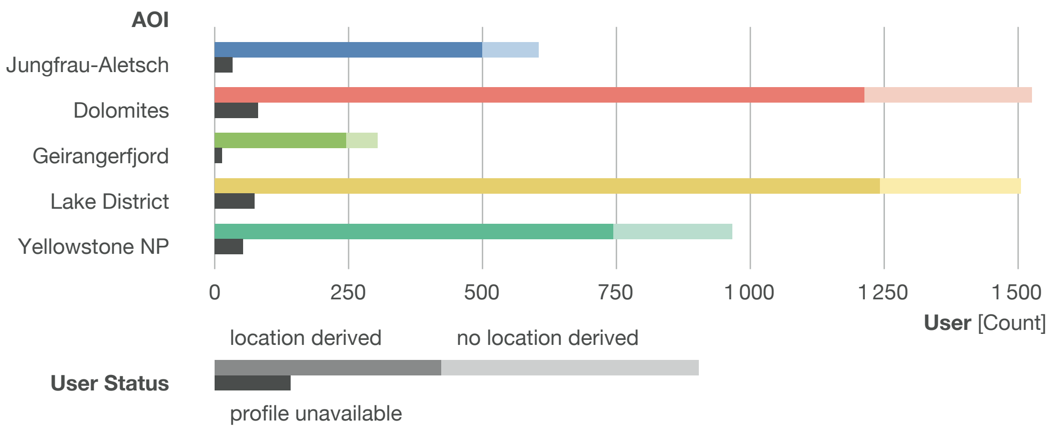


Figure 6.1: Overview of user stratification by availability of online profile and successful extraction of home location by AOI. Only users in the category ‹online info available, location found› are further processed and found in subset$_{located}$.

number of users for which an active online profile was found (which is essential for the user location method as introduced in *5.3 Deriving Home Locations*) combined with the information in which cases a location could be derived based on the sampled images. The percentage of users without an active Flickr profile is very similar over all AOI (~5%). The same is true for the number of users with an active Flickr profile but no DHL (20–30%).

In the extraction of the subset$_{analysed}$ from an AOI's full dataset (subset$_{original}$), the larger proportion of users is filtered out due to errors or limits (see *Figure 6.2*). On one hand, users are filtered out because important data is missing, which is categorized as an error. Limits, on the other hand are understood as deliberate decisions to limit the dataset for methodological reasons. Each class features two types, making this the four filter criteria described below:

error$_1$ ‹*Home location could not be derived*›. The procedure of sampling the user's photostream and finding the main cluster with DBScan did not result in a location. This happens if all sampled images are so far apart from each other that there is not a single cluster and all are labelled as outliers. This was the case for 27.0% (Yellowstone NP), 26.1% (Dolomites), 22.6% (Geirangerfjord), 21.6% (Jungfrau-Aletsch), and 21.3% (Lake District) users within the AOI subset$_{original}$. Additionally, there were 247 users (4.8% of all users in subset$_{original}$) whose online profile did no longer exist and therefore no photostream could be accessed for sampling images.

error$_2$ ‹*None of the images could be tagged*›. This error was only found in a single case within each of the following AOIs: Dolomites, Geirangerfjord, and Yellowstone. This rare error is caused when not a single image of a user can be processed with the Vision API. Often, these are users that have only a single or a very small number of images which are no longer available. Normally, the reason is that the image was removed from Twitter in between the publishing of the YFCC100M dataset and its download to an AWS S3 bucket (see Choi, Thomee & Larson 2017).

limit$_1$ ‹*User limit for country is reached*›. Preliminary data analysis showed that derived user countries are unequally distributed within and between the areas of unit (see *Figure 6.25*). To avoid the samples being dominated by one or two countries and limit the number of images to be processed by the Vision API, the upper limit of users per derived country is set at 50. Hence, if there are for instance 93 users in the subset$_{located}$ associated with the same DHC, 43 will be filtered out in the process of creating subset$_{tagged}$.

limit$_2$ ‹*Minimum number of users per country is not reached*›. There are many countries (73.5% *or* 83 of 113 derived countries) that feature 10 or less users over all AOIs. As the bias introduced by a single user can be very large in such a small sample, these countries are excluded. If limit$_1$ is the upper limit, then this would be the lower limit.

After all filtering steps the final subset$_{analysed}$ holds a total of 1 051 unique users (of which only 24 are found in two AOIs). In the following, the focus lies on the images created by those users and the comparison of the sample with the subset$_{original}$.

6.1.2 Image Characteristics

The final subset$_{analysed}$ holds a total number of 19 168 images. Their temporal distribution reflects – based on a visual comparison – the trends seen in the full dataset (see *Figure 6.3*) relatively well. The absolute numbers of images reflect the quite different initial sample sizes for each AOI. While the samples for the Dolomites, Lake District, and Yellowstone are large, they are quite small for the Jungfrau-Aletsch and especially the Geirangerfjord. To get a clearer picture of the distribution, *Figure 6.4* shows the relative distribution for each AOI over time. Except for a few instances with differences of up to 6 percentage points, the subsets are quite similar. As the temporal aspect is not central to the research objectives this shall suffice.

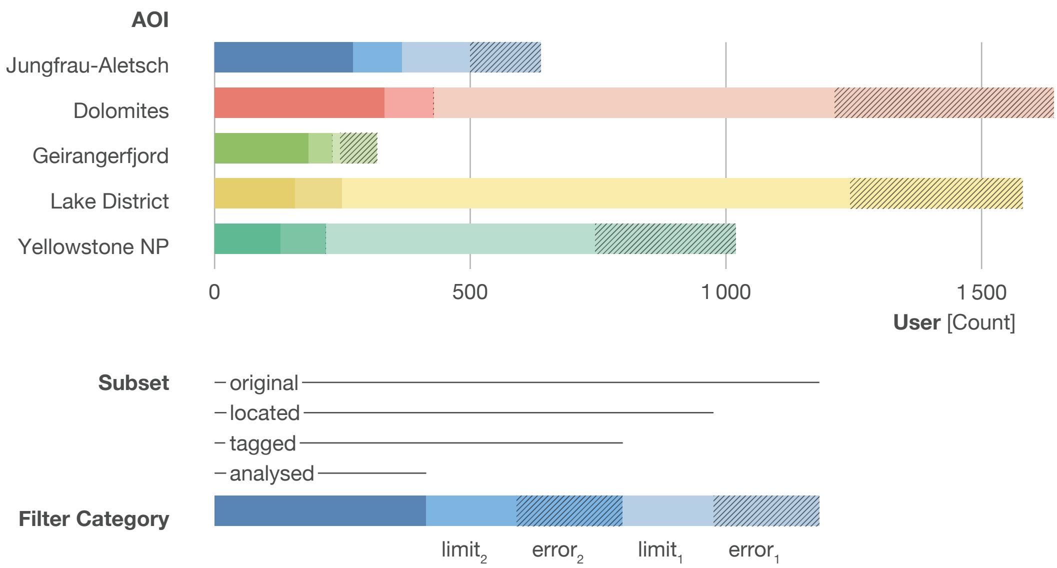


Figure 6.2: Stratification of subsets$_{original}$ by limits and errors. The greatest reduction of data is due to limit$_1$ (max. user / country) accounting for the filtering out of up to 50% of users initially found in an AOI. The lower section shows in which subsets the filter steps take place.

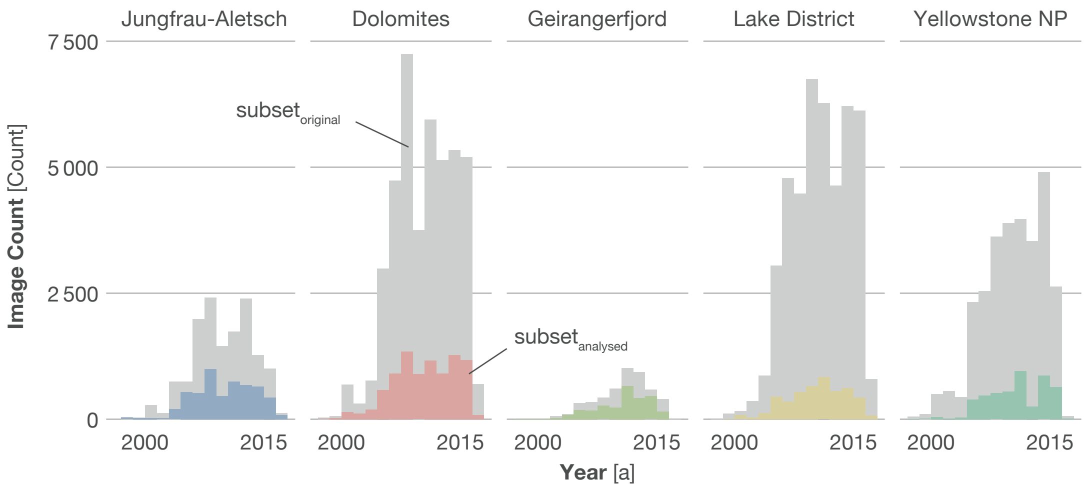


Figure 6.3: Comparison of absolute image counts by year (2000–2014) between subset$_{original}$ and subset$_{analysed}$ by AOI.

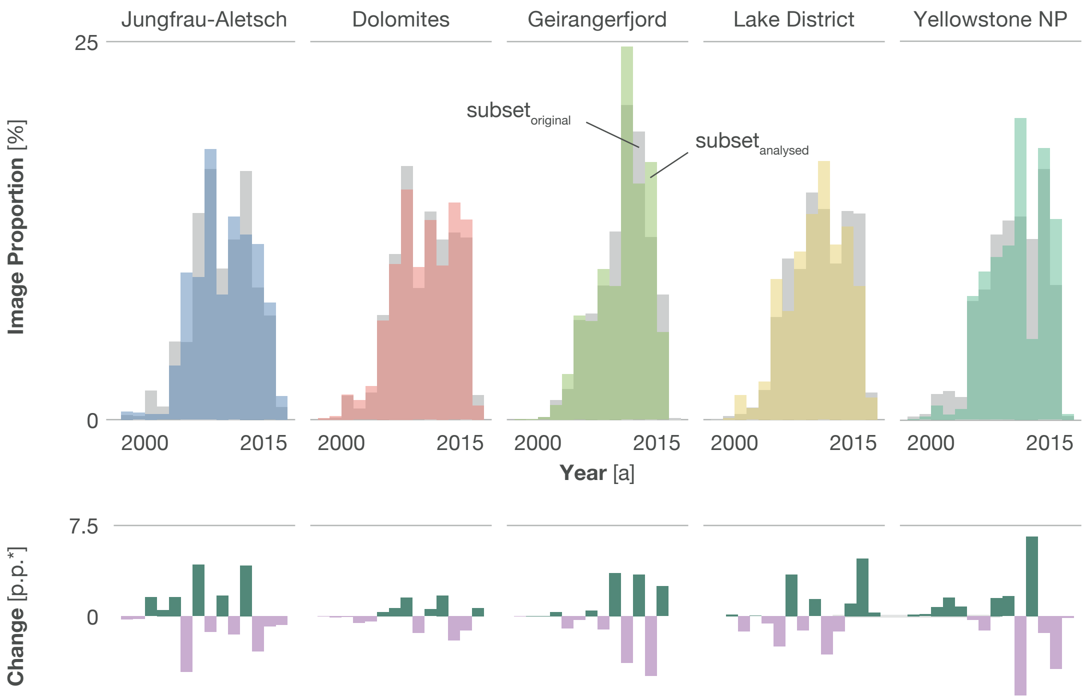


Figure 6.4: Normalized image count by year (2000–2014) and AOI, comparing subset$_{original}$ and subset$_{analysed}$. Shown in the lower section are the differences in percentage points (*) between the two subsets for each year.

In regard to the number of images per user, the distribution is relatively similar over all AOIs as well as the two subsets (see *Figure 6.5*). Two peaks are clearly distinguishable, one at low image counts of around 1–5 and the second one at 100+ images per user. The latter is due to the decision to cap a user's maximum image count at 100 images. While the Dolomites show very high numbers for users with very few images in the subset, Geirangerfjord has four times less users in that category. Once again, the varying sample size for the four AOI is clearly visible. For an easier comparison, and due to the quite unequally distributed user counts between the AOIs, relative values as well as their change are visualised (see *Figure 6.6*). From $\text{subset}_{\text{original}}$ to $\text{subset}_{\text{analysed}}$ the distributions for all AOIs show two main changes: the percentage of users with a minimal number of images gets smaller and the one for users featuring 100+ images increases. The biggest changes are happening from $\text{subset}_{\text{located}}$ to $\text{subset}_{\text{tagged}}$ and to a lesser degree from $\text{subset}_{\text{tagged}}$ to $\text{subset}_{\text{analysed}}$. What causes these changes are the four main filters used in generating the four subsets as visualised in *Figure 6.2*. In the first reduction step, from $\text{subset}_{\text{original}}$ to $\text{subset}_{\text{located}}$, users are filtered out because no locations could be derived. As can be seen, the overall distribution of images per user has only changed less than 1 percentage point for most image counts. In the second reduction step, from which $\text{subset}_{\text{tagged}}$ results, the distributions start to change drastically. The percentage of users with just one image is reduced. Simultaneously, the percentage of users with 100+ images increases. Due to the fact that the number of users with just one or at least a few images is very large in all AOIs, a larger number of users in these will be filtered out due to limit_1 (reduction of the maximum number of users per DHC to 50) or error_2 (none of the users images are no longer available online). The filtering explains the decrease in very low image counts and – due to the nature of relative values – increases the values for the other categories. This change is reflected most notably in the percentage of users featuring 100+ images. In the third reduction step, the same scenario is repeating itself, just that this time the filtering is caused by limit_2 (minimum number of 10 users per DHC in the final $\text{subset}_{\text{analysed}}$).

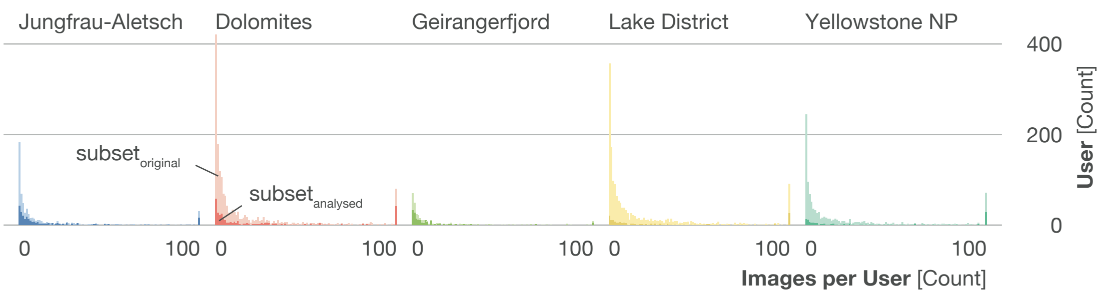


Figure 6.5: User count by image number per user for each AOI in $\text{subset}_{\text{original}}$ and $\text{subset}_{\text{analysed}}$.

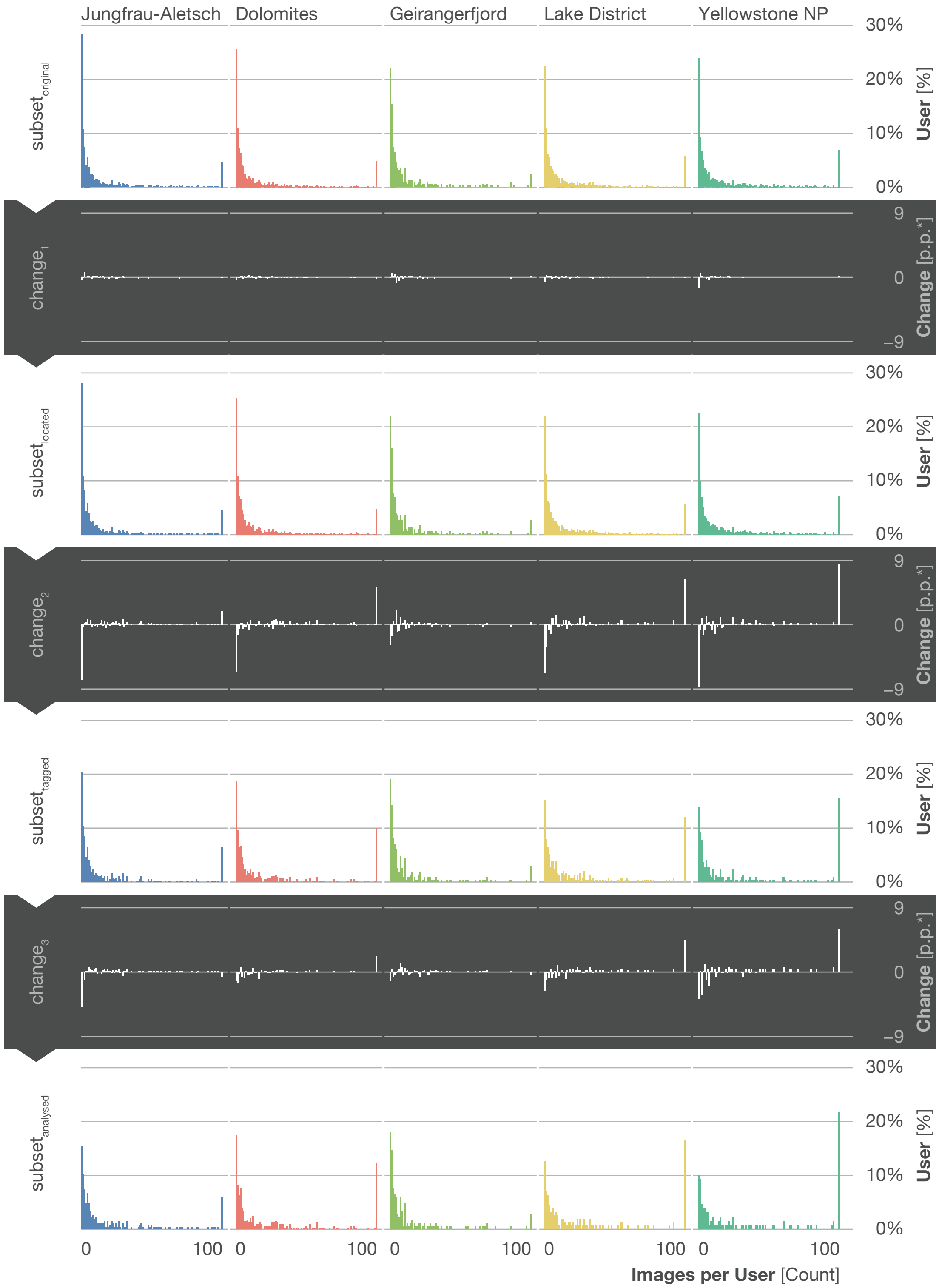


Figure 6.6: Overview of each user's image count taken within the five AOIs as relative values for each of the four subsets as well as the change in percentage points (*) between the subsets. For the completeness of the visualisation, image per user counts greater than 100 in the subset$_{original}$ are capped as well. With each filtering step, the proportion of users featuring only a small number of images is reduced.

Not only should the generated subsets be representative in the matters of users, image counts, and images per user, they should also reflect the spatial distribution of the $subset_{original}$. Therefore, the next chapter's focus lies on the spatial characteristics of the various subsets.

6.1.3 Spatial Characteristics

The spatial distribution of the images is a defining factor for the potential landscape elements. If all images in the proximity of lakes are excluded due to a poor sampling strategy, an unwanted bias would have been introduced. To get a spatial understanding of the various subsets and their sampling, there are three maps created for each AOI (*Figure 6.7a–6.11c*). They show the image density over the three subsets: $subset_{original}$, $subset_{located}$ and $subset_{analysed}$. Overall, the images have very similar distributions over the three subsets. The hotspots, indicating areas with a high density of images, tend to be the same as well. Exceptions are the Dolomites and the Lake District (*Figures 6.8* & *6.10*). In both cases, areas of high density disappeared when reducing the dataset to $subset_{analysed}$. In a more detailed visual analysis of the maps, the following observations can be made:

- *Jungfrau-Aletsch*: the areas of higher density are slightly larger and the images along the train track from Wengen to Kleine Scheidegg are no more present in $subset_{analysed}$ (*Figure 6.7c*).
- *Geirangerfjord*: Around Ålesund there is no visible reduction of images over the three subsets as it can be seen inside the Geirangerfjord (*Figures 6.8a–6.8c*).
- *Yosemite NP*: While for the $subset_{original}$ and $subset_{located}$ (*Figures 6.11a–6.11b*) the ring road leading the visitors through the national park is clearly visible, this is no longer the case for the $subset_{analysed}$ (*Figure 6.11c*).

Following this first visual inspection, the Nearest Neighbour Index (NNI) is used to describe the spatial distribution of the images within the various AOIs and subsets (see *Table 6.3*). The lowest values (NNI: ~0.1) are reported for all subsets of the Yellowstone NP, while the highest values (NNI: ~0.2) are found within the subsets of the Lake District. In general, a trend can be seen towards slightly higher index values with more processing (from $subset_{original}$ to $subset_{analysed}$).

6.2 Deriving Home Locations

In the following chapter, the resulting DHLs and DHCs for the Flickr users in $subset_{located}$ are presented.

6.2.1 Comparison of Methods

In order to geo-reference single users, two methods (Median and DBScan) have been introduced in 5.3.1 *Extracting Home Locations*. Both approaches have been implemented, resulting in relatively similar results. A first overview of the general results of both methods is given for which the data is not separated by AOI (but of course limited to these). Followed by a closer look at the AOIs and more specific results, the two methods are directly compared.

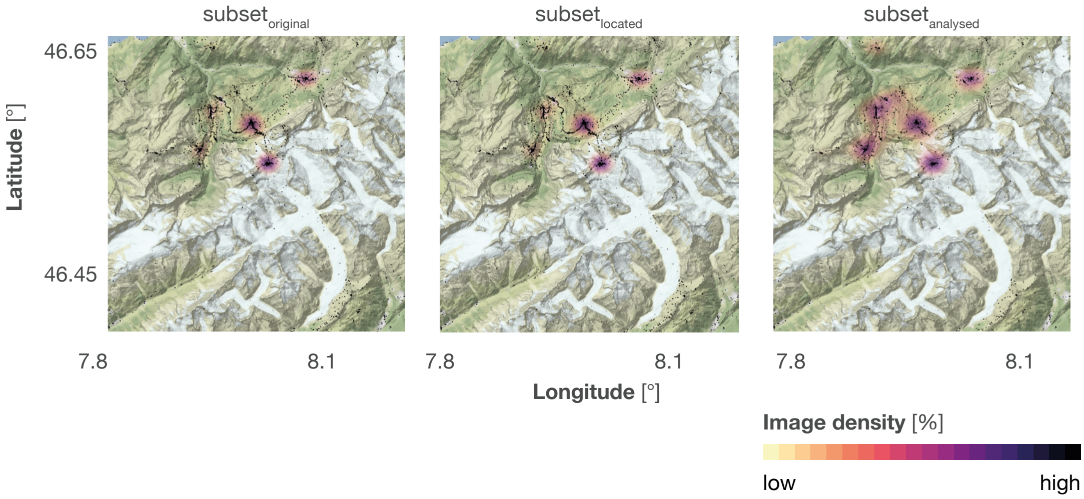
The Median approach resulted in a potential DHL for 91.3% of the users (4 744 of 5 196). With the second approach, using DBScan to cluster the image locations and find the largest cluster, **75.9%** of the users could be geo-referenced (3 946 of 5 196). The reason why for some users no DHL could be derived is found in either their Flickr library, their user profile or the spatial distribution of the sampled image locations. As the number of sampled images per users is limited to 250 and only those featuring coordinates are kept, there is either the chance that users do not feature any images with coordinates at all, the random sampling missed all images with geo-references or that the user account had been deleted by the time of analysis. In the case of DBScan, the method itself potentially reduces the absolute number of geo-referenced users. This can be the case when, based on the given parameters, no cluster could be found and all data points (locations of sampled images from users Flickr library) are classified as outliers.

Granularity of Space

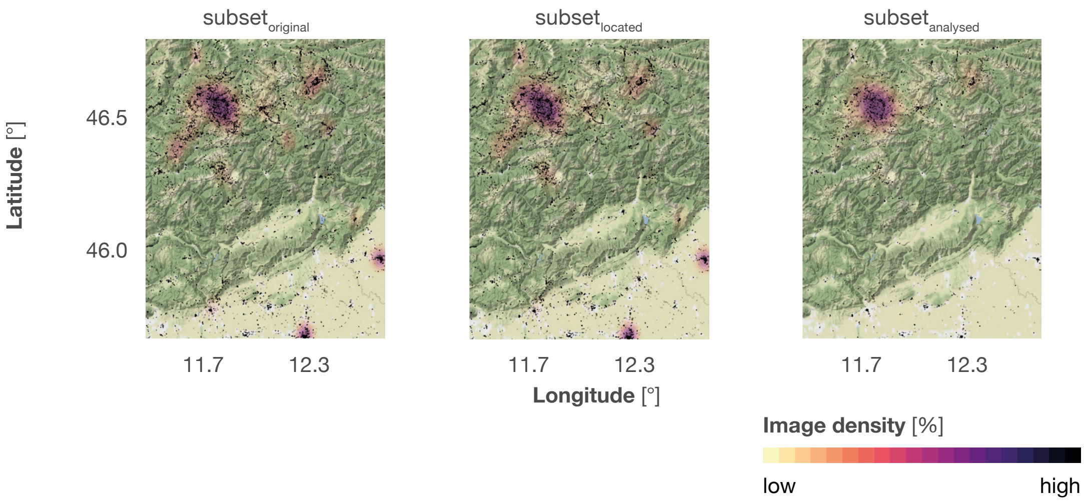
The resulting locational information is analysed on four spatial granularities. On a *coordinate-level*, the focus lies on the derived coordinate pairs. On a *micro-level*, the reverse geocoded city information is further investigated (DHL) and compared to the overlaying country data, the *meso-level* (given by the DHC). The coarsest spatial granularity is found on the *macro-level*, in this case on a continental level.

Table 6.3: Nearest Neighbour Index values for all four subset steps in each AOI. The following reference values are used for interpretation: 0 represents a perfectly clustered distribution, while 1 advises a random distributed and 2.15 corresponds to uniformly distributed data.

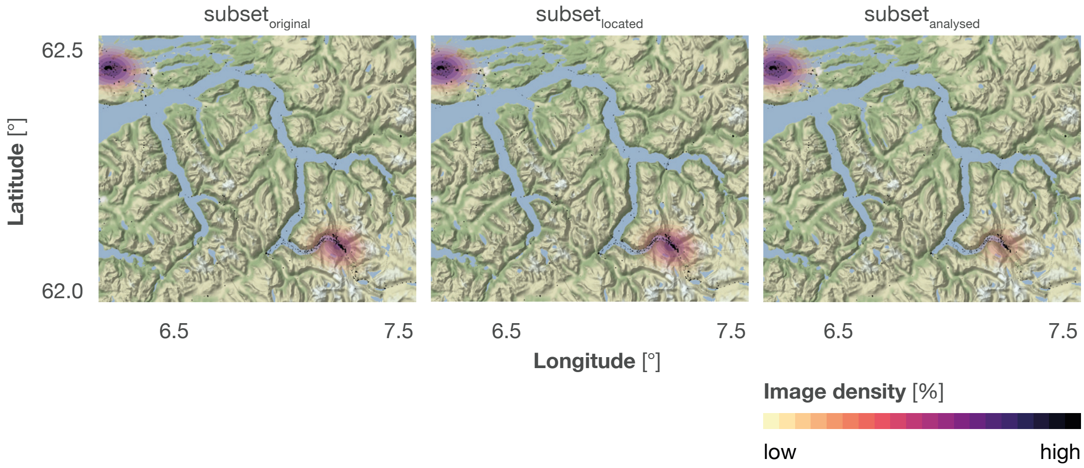
	Nearest Neigbhour Index Values			
AOI	$subset_{original}$	$subset_{located}$	$subset_{tagged}$	$subset_{analysed}$
Jungfrau-Aletsch	0.16	0.16	0.17	0.17
Dolomites	0.18	0.19	0.21	0.21
Geirangerfjord	0.13	0.14	0.14	0.15
Lake District	0.20	0.20	0.22	0.22
Yellowstone NP	0.10	0.10	0.10	0.09



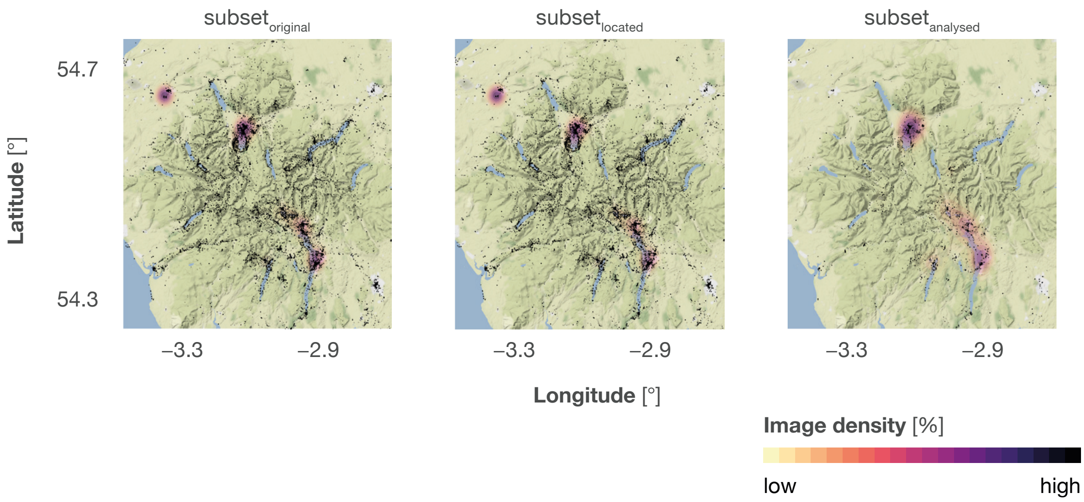
Figures 6.7a–6.7c: Spatial distribution of images in the AOI Jungfrau-Aletsch as found in subset$_{original}$, subset$_{located}$, and subset$_{analysed}$. A number of hotspots are clearly visible. They correspond to the most important stations and villages (e.g. Grindelwald, Kleine Scheidegg, Lauterbrunnen, Wengen), as well as the Jungfraujoch as one of the main tourist attractions in the region. (Map tiles by Stamen Design, under CC BY 3.0. Data by OpenStreetMap, under ODbL.)



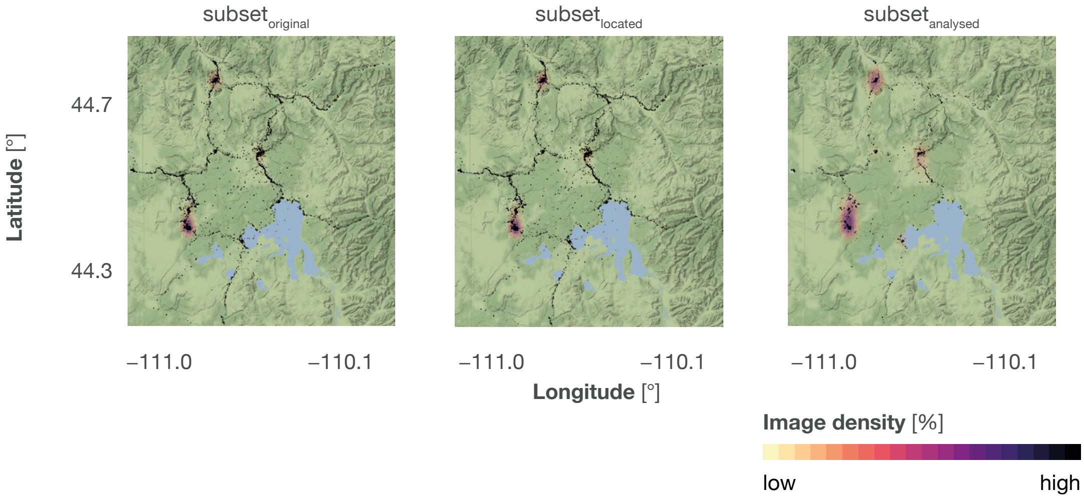
Figures 6.8a–6.8c: Spatial distribution of images in the AOI Dolomites as found in subset$_{original}$, subset$_{located}$, and subset$_{analysed}$. The main hotspot in the north west of the AOI overlaps with the nature park ‹Puez-Geisler›, an area exemplary for the peaks the Dolomites are famous for. (Map tiles by Stamen Design, under CC BY 3.0. Data by OpenStreetMap, under ODbL.)



Figures 6.9a–6.9c: Spatial distribution of images in the AOI Geirangerfjord as found in subset$_{original}$, subset$_{located}$, and subset$_{analysed}$. The north-western hotspot corresponds to Ålesund, the largest city in the area. In the south-eastern part a second hotspot is seen, marking the end of the Geirangerfjord and at the same time the most touristic area in the region. (Map tiles by Stamen Design, under CC BY 3.0. Data by OpenStreetMap, under ODbL.)



Figures 6.10a–6.10c: Spatial distribution of images in the AOI Lake District as found in subset$_{original}$, subset$_{located}$, and subset$_{analysed}$. While there are three hotspots in subset$_{original}$ and subset$_{located}$, the one in Cockermouth (north-east) is missing in the final subset$_{analysed}$. The other two are located at the shores of the two lakes Derwentwater (north) and Windermere (south). (Map tiles by Stamen Design, under CC BY 3.0. Data by OpenStreetMap, under ODbL.)



Figures 6.11a–6.11c: Spatial distribution of images in the AOI Yellowstone NP as found in subset$_{original}$, subset$_{located}$, and subset$_{analysed}$. The two main hotspots correspond to two natural sights: The Mammoth Hot Springs (north-west) and the Old Faithful (south-west). The third area that features a relatively high density of images are the Upper and the Lower Falls along the Yellowstone River. (Map tiles by Stamen Design, under CC BY 3.0. Data by OpenStreetMap, under ODbL.)

The smallest level is given by the derived coordinate pairs from the median coordinates and the calculated median of the DBScan cluster featuring the highest image count. This means that most users – except those where no cluster was found, as discussed above – within the subset$_{located}$ have two potential DHLs that can be compared to each other. Calculating the (haversine) distance between the two DHLs and visualising the sum curve of the resulting distribution gives valuable insights (see *Figure 6.12*). More than 25% of all DHLs – irrespective of their AOI – feature a distance of 0 Kilometres, which means that both methods derived the same DHL. All curves stall at around 75–80%, which is caused by the fact that the method based on DBScan not always yields results (i.e. the case with only outliers and no clusters), resulting in NA values. Therefore, the actual maximum for each curve is slightly different and lies somewhere around 80%. When looking at the shape of the sum curves in the section 0–2 000 Kilometres, there are three clusters: (1) *Relatively high agreement* as seen with the Dolomites and the Lake District. 75% of all DHLs lie no more than 250 Kilometres apart. (2) *Medium agreement* (Jungfrau-Aletsch and Geirangerfjord), which means that the 75% mark is reached at around 750 Kilometres. The last cluster, with (3) *relatively low agreement* consists of Yellowstone NP, for which the third quantile is reached at 2 000 Kilometres. The exact reasons for the described differences remain unclear.

On the micro-, meso-, and macro-level, the focus lies on the accordance of the two methods. Accordance is herein understood as the percentage of users for which both methods returned the exact same city or country. To get their containing administrative unit (*city*, *country*), the DHL, given in longitude and latitude, were reverse geocoded. Their accordance increases with a higher level of aggregation, which could be expected. The larger spatial units simply allow for more error regarding the accuracy of the DHL. *Figure 6.13* shows the accordance as well as the non-matching or missing percentages of all users per AOI for all three levels. The aggregation from micro- to meso-level increases the accordance from around 25% to nearly 75% for all AOIs. The percentage of users for which both methods do not agree on is reduced from around 50% to less than 20%. When aggregating the data on the level of continents, the agreement increases again. Only for a small percentage (around 5%) of users in each AOI the methods do not agree on their home continent.

All AOIs lack around 20% of DHL on each level, as the DBScan method did not identify a minimum of one cluster and therefore no comparison is available for the Median results.

Since the focus lies on the spatial reference unit of countries, some further analysis is performed with a focus on that specific level. As the full set of countries (113 countries) contains a large number with only 1–5 users (i.e. Puerto Rico, Argentina and Reunion), the countries are

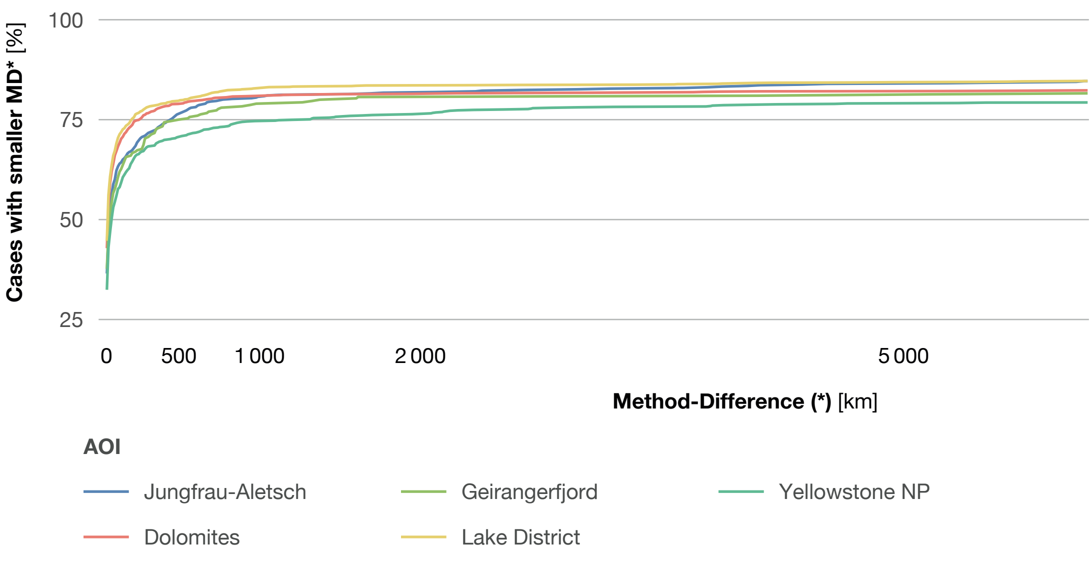


Figure 6.12: Sum curves of calculated distances between the DHLs derived with Median and DBScan approaches for each AOI. For each distance from 0–5 000 kilometres (x-axis) the percentage of DHLs derived with the two methods that lie closer by each other than the arbitrary threshold (e.g. 75% of all DHLs derived with the two methods in Yellowstone NP are less than 2 000 kilometres apart).

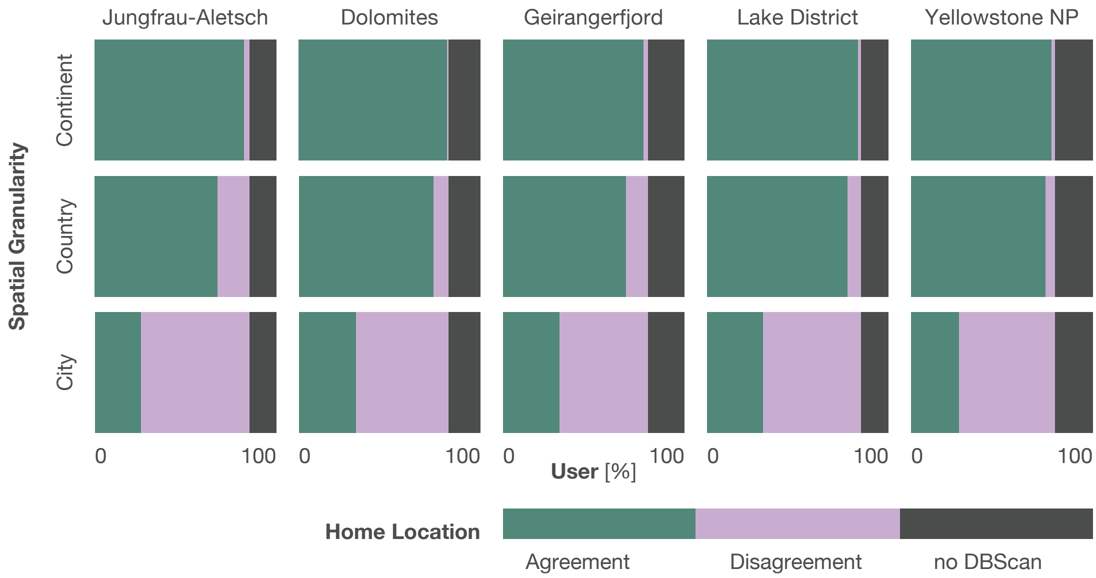


Figure 6.13: Accordance of Median and DBScan approaches on micro- (city), meso- (country), and macro-level (continent).

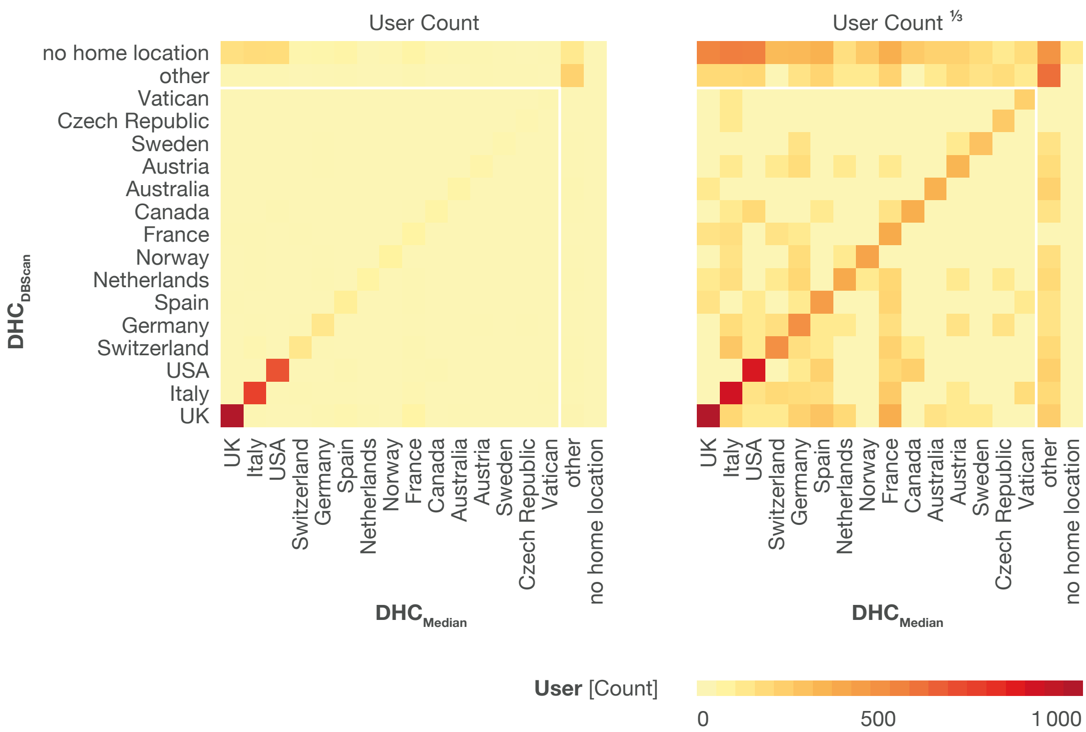
reduced to the top 15 countries by user count and all other countries are classified as ‹other›. Confusion matrices are generated that give some ideas on classification errors and where they appear.

In a first step, the classification results of all AOIs are visualized in a single confusion matrix (*Figure 6.14*). The most obvious feature is the clear diagonal representing all those cases where both methods found the same home country. For several countries, namely the United Kingdom, Italy, and the United States, a smaller proportion of ‹no home locations› was found for the DBScan method. The main problem with the data is the high number of DHLs for the United Kingdom, Italy, and the United States (around 1 000 cases just for the United Kingdom). Countries with smaller counts as well as the small number of wrong classified DHCs are no longer visible. Therefore, the same data is visualised in *Figure 6.15* with major distortion: the square root is taken for all counts. The resulting confusion matrix still shows the diagonal and with it the number of users for which both methods yielded identical home countries. But it now allows a qualitative overview of the misclassifications and with it the following observations:

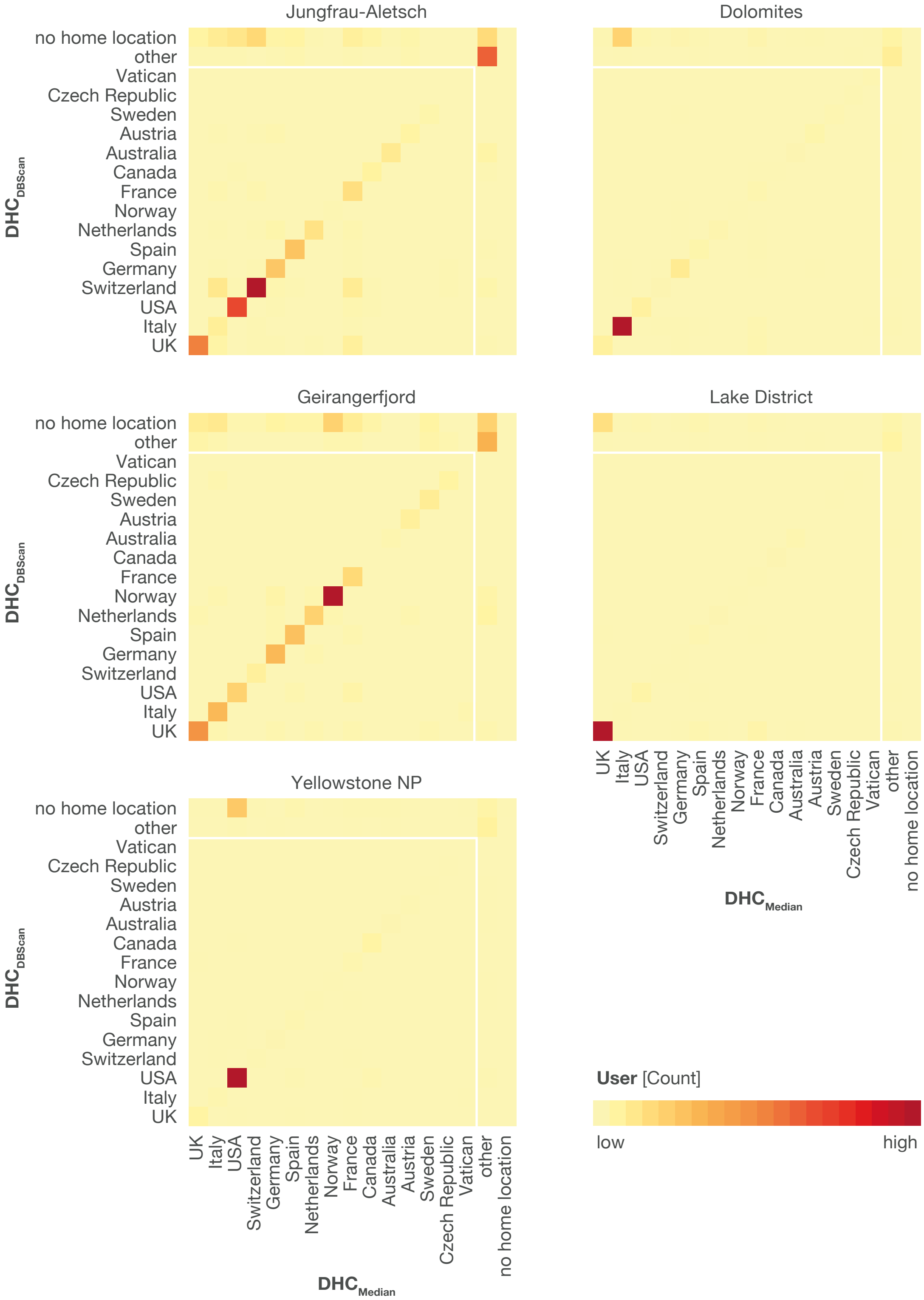
- The main classification error between the two methods are users with DHC found with the Median method that could not be reproduced with the DBScan method and resulted in ‹no home location›.

– The higher the number of users found per country (relatively similar for both methods with slightly less counts for the DBScan method) the more misclassifications are found. Such a trend can be expected, but it is not true for at least one country with high user counts but very few misclassifications: Norway.
– For a great number of cases the methods assigned ‹other› as the DHC. This result has to be treated carefully as it does not mean that both methods resulted in the same DHC (e.g. Swaziland and Lithuania could both be part of the class ‹other›, but clearly this would not be the same classification as suggested by the correlation matrix).

In a second step, there is a confusion matrix calculated for each AOI as shown in *Figures 6.16–6.20*. While there is still a clear tendency to identical results, several differences between the AOI become visible. On one hand, the maximum count for each case corresponds to the country the AOI is embedded in. On the other hand, Geirangerfjord and Jungfrau-Aletsch



Figures 6.14 & 6.15: Confusion matrices of user counts per DHC found by DBScan and Median method with Figure 6.15 showing their third root to highlight lower values. Included are the top 15 countries by user count, while all other countries are classified as ‹other›. The class ‹no home location› consists of all those users for which the DBScan method was unable to determine a DHL (but the Median method did).



Figures 6.16–6.20: Confusion matrices of DHL by DBScan and Median method for each AOI. Not all DHCs are equally represented in each AOI, the local user population (e.g. ‹Switzerland› for the Jungfrau-Aletsch region) is the largest in all five AOIs.

feature a much more balanced set of home countries than the Dolomites, the Lake District and the Yellowstone NP.

6.2.2 Validation of Derived Home Locations

Regardless of the agreement between the two methods on any of the spatial granularities, no information has so far been given on the accuracy of their results. The validation is based on the creation of a structured, semi-random sample and its manual validation with user locations available on their Flickr profiles (see *5.3.3 Validating Home Locations and Countries*). For a small number of users, problems were found in (1) too unspecific user locations in their Flickr profile (e.g. Ireland) to evaluate for agreement on city level, (2) users living in Rome, Italy for whom the derived country was Vatican City, and (3) in one case in an unknown language (i.e. user 10560956@N08). In the specific case with a very small country lying within a large city, two users actually living in Rome with Vatican City as their DHC are rated as correct on all spatial granularity levels. For the spatially higher resolved DHL, the following simplification is made: If either method's DHL lies in close proximity (< 10 km) of the user location, it is treated as correct, even though they do not bear the exact same name.

A sample of three validated users with their DHL within the USA is visualized in *Fig. 6.21* to illustrate the process. All three users live in the eastern parts of the country and their DHL

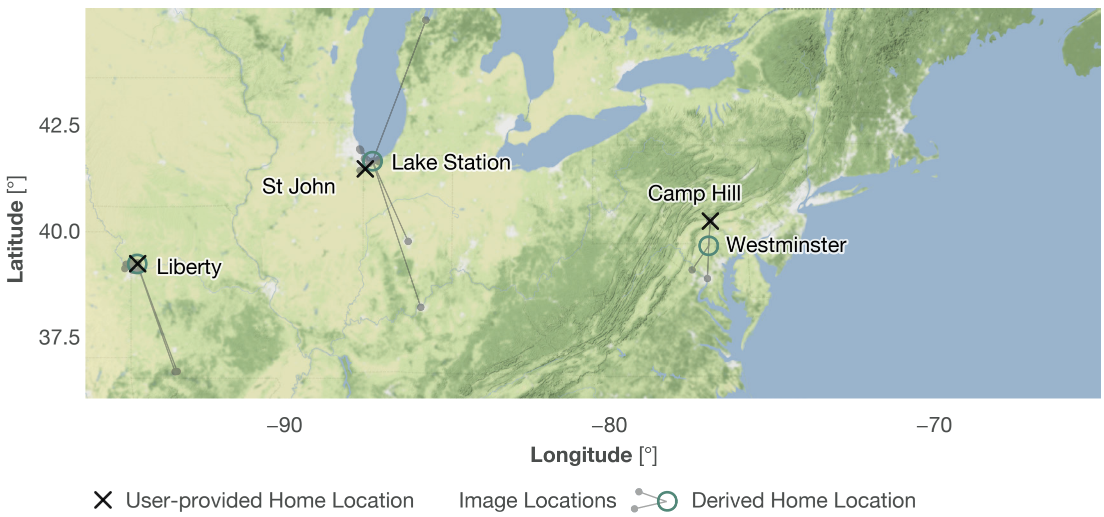


Figure 6.21: Map displaying the image clusters based on which the DHL for three US-American users is derived. The bold label corresponds to the user-provided home location as found in the user profile. The single image locations based on which the location was derived are connected to the DHL, which is marked and labelled as well. (Map tiles by Stamen Design, under CC BY 3.0. Data by OpenStreetMap, under ODbL.)

was calculated on the three clusters displayed. Each cluster consists of the single images the DHL was calculated on, which itself is displayed as a slightly larger point. For each user, the DHL stated in the profile was manually looked up on online (*maps.google.ch*) and added to the map as a reference. The results show a similar pattern as observed in the in-between method agreement: the accuracy is highly dependent on the spatial granularity. For a small scale (city-level) an overall AOI accuracy of 30% (Median) and 50% (DBScan) is found (see *Figure 6.22*). On the next aggregation level, based on countries, the accuracy is around 90% for both methods and on a continent scale rises to nearly 100%. As the number of validated users is limited to 150, it is important to keep in mind that already single cases can introduce a high bias.

Paired two-sided t-tests are performed for each spatial granularity testing for similarity of the accuracy values. The test hypotheses per spatial granularity are: H_0, the mean of the distribution differs not significantly from zero; *and* H_1, the mean of the distribution differs significantly from zero. If the two methods (Median and DBScan) do not differ in the amount of correct locations returned, the test should not be significant at a significance level of $\alpha = 0.05$. The test assumptions (interval/ratio scaled variable, normal distribution of pairwise differences, and random sample from population) are all satisfied, whereby the normal distribution of pairwise differences is neglectable at $n > 30$ (UZH 2018). The results are listed in *Table 6.4*. The reason for the inconsistent degrees of freedom over the three levels is the low spatial detail in the location provided in some user profiles (e.g. Italy, for which no comparison can be made on a city-level). The only significant difference ($p < 0.05$) between the two methods is found on a city-level with DBScan performing 20 percentage points better. Splitting the validation results by AOI (see *Figure 6.23*), the distribution – low percentage of correct locations at city level, high percentages at country level, and nearly 100% at continent level – is

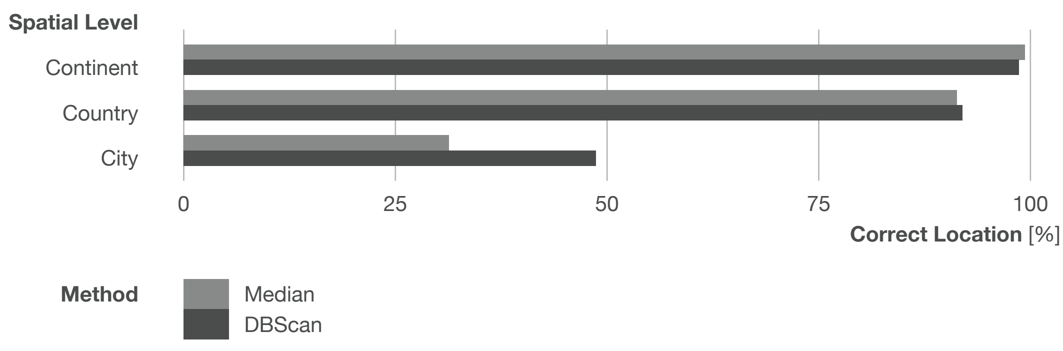


Figure 6.22: Validation of user location for the DBScan and Median method on the three spatial levels city, country, and continent.

Level	p-Value	t-Value	DF*
City	< 0.01	5.19	132
Country	0.76	0.30	149
Continent	0.31	–1.00	148

* DF = Degrees of Freedom. Values vary because of not available values in the dataset.

Table 6.4: Results of paired t-test comparing the validation results of the DBScan and Median method over the three spatial levels: city, country, and continent.

very similar to the overall results. As mentioned earlier, the users in Jungfrau-Aletsch were geo-referenced with a maximum sample of 50 images. For the other AOIs the limitation was set at 30 images to reduce computation time. As the results show, this decision was probably right, as the DHL for users in Jungfrau-Aletsch do not seem to be correct more often than in the other AOIs. The same paired two-sided t-test as performed above is again applied to the data divided by AOI. The general result is the same: for all AOIs except for Jungfrau-Aletsch, a significant ($p < 0.05$) difference was found in the validation results on a city-level (same test hypothesis as used before). The full test results are reported in *Table 6.5*.

As the DBScan approach seems to result in locations with a higher accuracy (more in *7.2.1 Interpretation of Results*), the following analysis will work with that set of DHLs.

AOI	Level	p-Value	t-Value	DF*
Jungfrau-Aletsch	City	0.33	1.00	23
	Country	0.16	1.44	29
	Continent	—	—	28
Dolomites	City	0.01	2.75	24
	Country	1.00	0.00	29
	Continent	0.33	–1.00	29
Geirangerfjord	City	0.01	2.74	25
	Country	0.66	–0.44	29
	Continent	—	—	29
Lake District	City	0.02	2.42	27
	Country	1.00	0.00	29
	Continent	—	—	29
Yellowstone NP	City	0.02	2.54	29
	Country	—	—	29
	Continent	—	—	29

Table 6.5: Comparison of the accuracy of the DBScan and Median methods. For each AOI a t-test was performed for the number of correct cases at each spatial level.

* DF = Degrees of Freedom. Values vary because of not available values in the dataset.

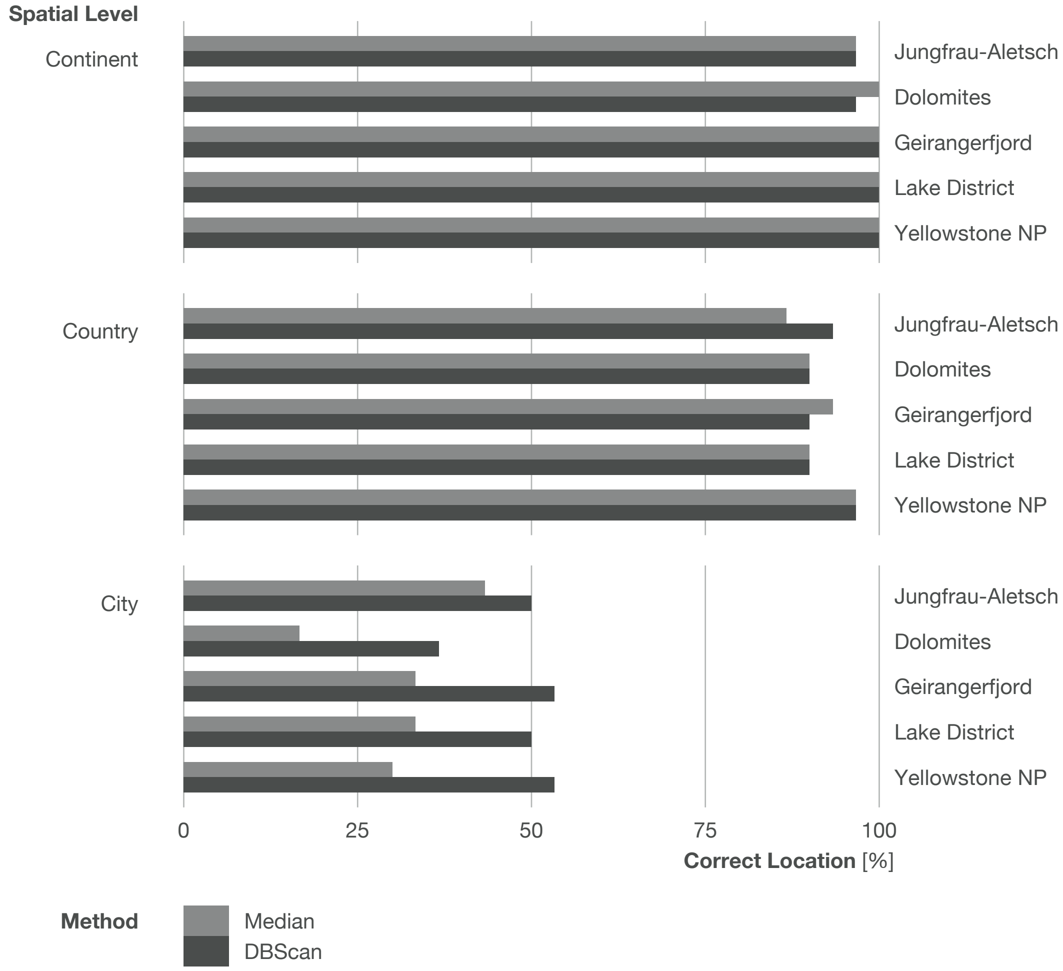


Figure 6.23: Validation of user location for the DBScan and Median method on the three spatial levels city, country, and continent for each AOI. The most variation is found on the city-level with users providing images in the Dolomites yielding the lowest validation results.

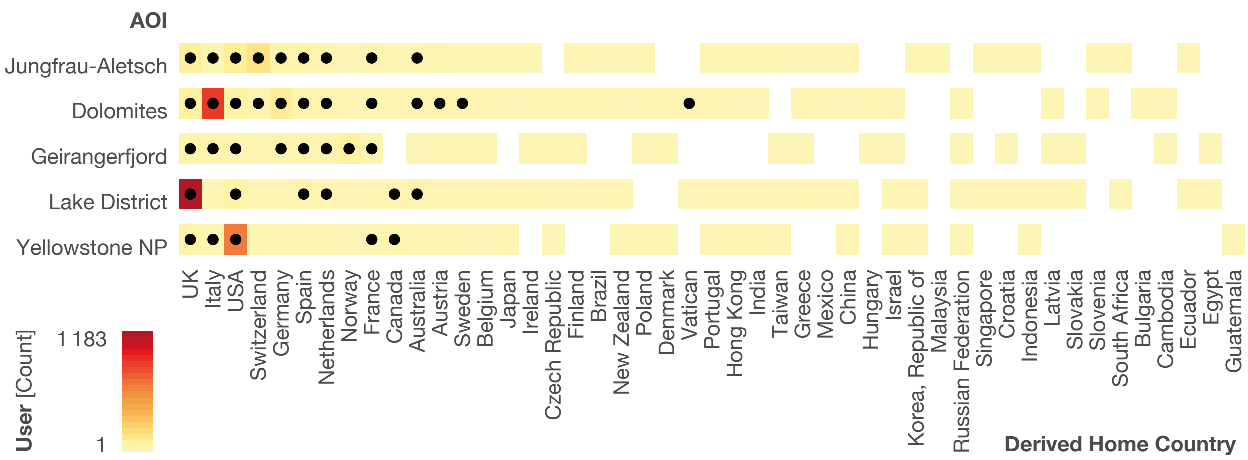


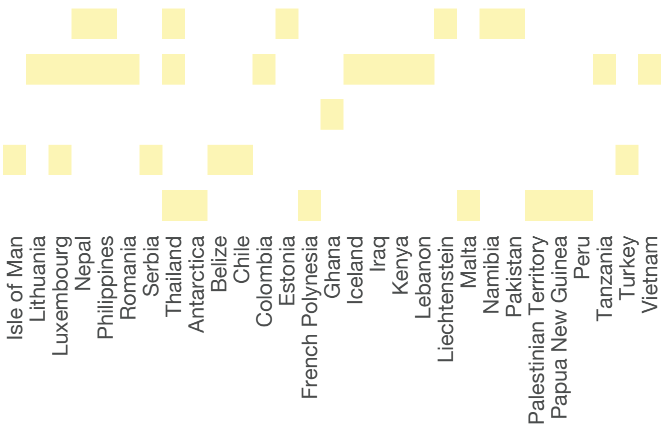
Table 6.6: Derived user home country count by AOI. Percentages are given in relation to $\text{subset}_{\text{located}}$.

	Jungfrau-Aletsch		Dolomites		Geirangerfjord		Lake District		Yellowstone NP	
	Count	*%*	*Count*	*%*	*Count*	*%*	*Count*	*%*	*Count*	*%*
$\text{subset}_{\text{original}}$	–	–	–	–	–	–	–	–	–	–
$\text{subset}_{\text{located}}$	54	100.0	68	100.0	36	100.0	63	100.0	52	100.0
$\text{subset}_{\text{tagged}}$	42	77.8	49	72.0	29	80.6	45	71.4	36	69.2
$\text{subset}_{\text{analysed}}$	9	16.7	12	17.6	8	22.2	6	9.5	5	9.6

6.2.3 Derived Country Characteristics

As DBScan has outperformed the Median approach in the only two significantly different results and as visual inspection suggests it could be the case for more levels if there were larger samples, the following results are always reported based on the home countries derived with the DBScan approach. In the initial $\text{subsets}_{\text{original}}$, around 60 distinct home countries could be localised (except for Geirangerfjord with just 36). With the subsequent filtering procedures applied – of which the minimum users per country filter (limit_2, see *6.1.1 User Characteristics*) has by far the greatest impact – the $\text{subsets}_{\text{tagged}}$ and $\text{subsets}_{\text{analysed}}$ are generated. Over all AOIs, this results in less diverse country sets. They feature only 10–20% of the initial number of countries (see *Table 6.6*).

While in *Figure 6.24* the overview of all users that could be located with the DBScan approach is given, *Figure 6.25* is reduced to users and derived countries available in the final $\text{subset}_{\text{analysed}}$.



◀

Figure 6.24 (incl. left page): User count by DHC and AOI for $\text{subset}_{\text{located}}$. Tiles with a black dot indicate user counts greater than 10, the defined minimal number of users for a DHC to be considered in the final analysis.

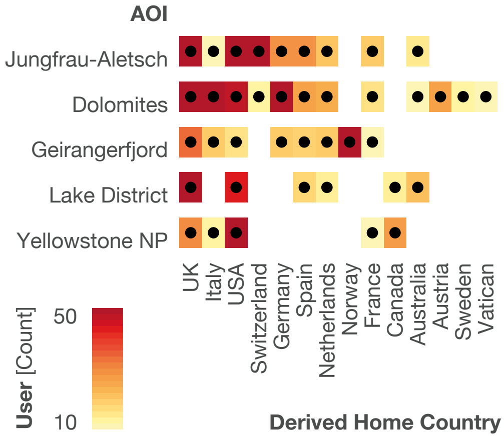


Figure 6.25: User country by DHC and AOI for subset$_{analysed}$. The maximum number of users per DHC and AOI is capped at 50 and all DHCs with less than 10 users in any of the five AOIs are filtered out.

The subsets$_{original}$ feature a diverse selection of derived user countries, but clear hot spots can be seen in the user counts for the country the AOI is lying in. Especially in the case of the Lake District, Dolomites, and Yellowstone NP, this can be seen quite clearly with over 500 users in all three cases. For the Jungfrau-Aletsch Region and the Geirangerfjord, this trend is visible as well but at a much lower amplitude.

The 14 countries within the final analysis are the United Kingdom (UK, 208 users), the United States (USA, 201 users), Germany (96 users), Italy (89 users), Spain (88 users), the Netherlands (74 users), Switzerland (62 users), France (53 users), Norway (50 users), Australia (43 users), Canada (40 users), Austria (26 users), Sweden (12 users), and Vatican City State (10 users). While in the validation process Vatican City State was matched to Italy, it is treated as an independent entity in the following. The main reason is the high level of automation that would be limited by such specific data cleansing steps. While users with DHL in the UK and USA are found in all five AOIs, all other derived countries are found in just 1–4 of the five AOIs. Users from Austria, Sweden, and Vatican City State are only found in the Dolomites. This observation stresses the importance of considering the AOI as a context variable as only a few home countries are found in greater numbers in more than one AOI. Therefore, country specific differences have a greater chance to be highly influenced by the AOI most of its users and their images are found in. An important finding regarding the discussion of RQ.2.

All DHLs come with a certainty measure – the percentage of image locations in the home cluster compared to all image locations (see *5.3.1 Extracting Home Locations*). To see how accurate this measurement is, it can be compared to the validation results. The certainty measurements (see *Figure 6.26*) found in the process of locating each user are of two types: for the Dolomites and the Lake District a high number (~40%) of the locations feature a certainty measure of more than 75%. The other three AOIs – namely Jungfrau-Aletsch, Geirangerfjord, and Yellowstone NP – show a much smaller proportion (~25%) of cases with more than 75% certainty. Note that these values represent the full subset$_{located}$, in comparison to *Figure 6.27* that is limited to the 150 users within the validated subset. *Figure 6.27* compares for the selected users the calculated certainty measurement against the validation results. For each AOI

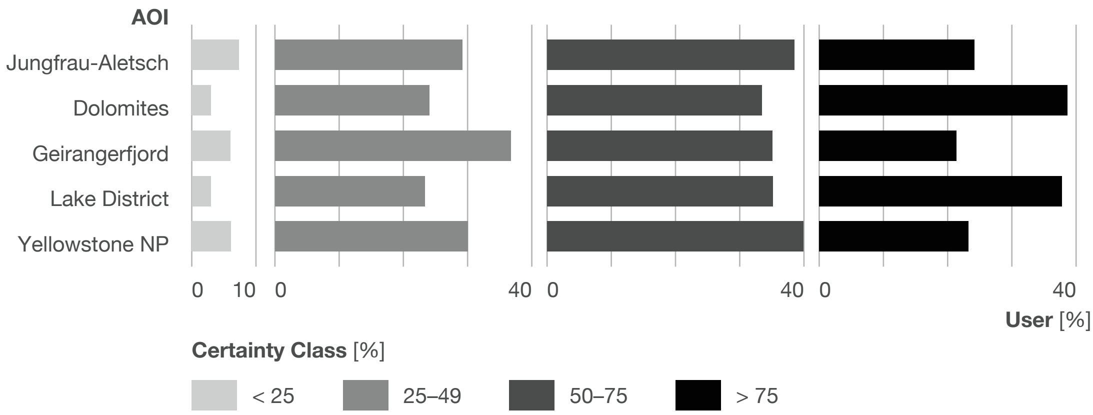


Figure 6.26: Classified certainty values as collected in the process of deriving the home locations shown by AOI.

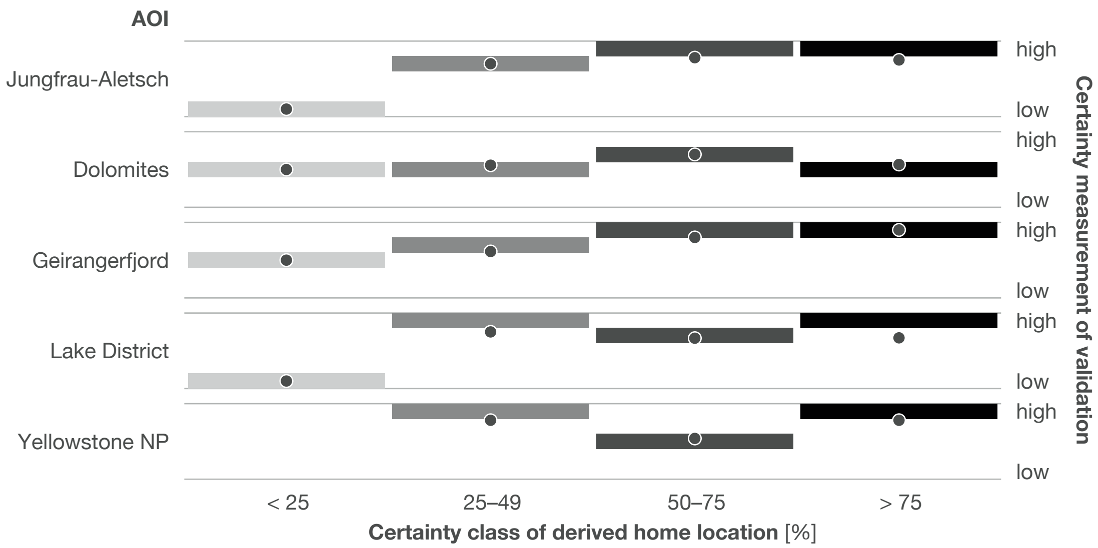


Figure 6.27: Comparison of classified certainty values as collected in the process of deriving home locations compared to the accuracies found in the validation process shown per AOI. The points represent the mean values, while the areas correspond to the median values of the accuracies found in the validation process. The higher the accordance of these values with the certainty classes, the more useful the latter is.

the median as well as the mean values of correct spatial levels (0 if city, country and continent is wrong and a maximum of 3 if all three are correct) are calculated. The higher the value, the better the validation result. Their median and mean per AOI are plotted against the four certainty classes found in the user locating process. The overall trend from an additional visual inspection seems to be a positive correlation, except for two outliers in the 25–49% class of Lake District and Yellowstone NP. Overall, this seems to support the calculated certainty measurement.

6.3 Extracting Landscape Elements

Knowing the DHC of a user allows the matching with what is understood as a cultural background (see *2.2 Culture*). But to know what the preferences of a person are, information is needed on their perception, which in the case of this study is given by the landscape elements found on their images. For this, the subset$_{\text{original}}$ is filtered to those users that could be geo-referenced (subset$_{\text{located}}$) and for each country a maximum of 50 users is randomly extracted. As the image count per users is quite different (see *Figure 5.5*), the image count per user is limited as well. In the first step, an overview is given on tags found in subset$_{\text{tagged}}$ on the level of single images and their tags (*6.3.1 Overview*). These tags can be analysed on several levels of aggregation, from single images or users to groups of users (i.e. their DHC) as well as the five AOIs. Therefore, in the second step, the tags are analysed on three potentially interesting aggregation levels: Image, DHC and AOI (*6.3.2 Tag Characteristics*). In the third and last step, the results on the structure of the tags with a focus on hierarchical clustering, similarity, and linguistic quality are introduced (*6.3.3 Tag Structures*).

6.3.1 Overview

A total of 33 471 images by 1 433 users are tagged by the Vision API, resulting in 689 472 total tags and 5 594 unique tags found in the subset$_{\text{tagged}}$. As the accuracy of the tags generated by the Vision API is unknown, a random sample of three images along with their tags is depicted in *Figures 6.28–6.30*. At a first glance, the resulting tags seem to describe the images quite well. The snowy image (*Figure 6.28*) is tagged with ‹snow›, ‹roof›, ‹house›, and ‹tree›, all of which seem to make sense. If it really depicts a ‹blizzard› and shows ‹freezing› temperatures or ‹frost› is an open question. The animal depicted in *Figure 6.29* is not exactly an Argali (see Morris 2019), but it is most likely a sheep. Tags like ‹mammal›, ‹vertebrate› or ‹livestock› fit this image. Besides the animal, the tags ‹grassland› and ‹pasture› accurately describe the surrounding. For the landscape image shown in *Figure 6.30*, nearly all tags make sense. Maybe ‹highland› does not fit too well and ‹cirque› is quite a guess, as there is no clear

Figure 6.28: Image 86070547 (imageID) located in the Dolomites. Extracted tags: ‹Snow›, ‹Winter›, ‹Freezing›, ‹Tree›, ‹Blizzard›, ‹Roof›, ‹Home›, ‹House›, ‹Architecture›, ‹Plant›, ‹Frost›, ‹Building›, ‹Cottage›. Image (http://www.flickr.com/photos/47271568@N00/86070547/) by «MenguMat», licensed under CC BY-NC-ND 2.0.

Figure 6.29: Image 8534486826 (imageID) located in the Lake District. Extracted tags: ‹Mammal›, ‹Vertebrate›, ‹Sheep›, ‹Livestock›, ‹Goats›, ‹Snout›, ‹Argali›, ‹Pasture›, ‹Horn›, ‹Wildlife›, and ‹Grassland›. Image (http://www.flickr.com/photos/38986305@N06/8534486826/) by «ARG_Flickr», licensed under CC BY 2.0.

sight of an «amphitheatre-shaped basin with precipitous walls, at the head of a glacial valley» (Encyclopaedia Britannica 2020).

A more structured overview is given in *Figure 6.31*, displaying collections of images that depict four of the more frequent tags like ‹mountain river›, ‹wilderness›, ‹wildlife›, and ‹ice›. For each term a random sample of three images was drawn from each AOI based on the $\text{subset}_{\text{tagged}}$. In general, the tags reflect the images well, it seems like the basic idea is right. For ‹wildlife› it is questionable if ‹livestock› is fitting, but at least all except for one image do show some sort of animal. The images tagged ‹wilderness› mostly show landscapes with no visible human impact, except for a trail in the woods, a wooden house on grassland, and a church. ‹Mountain river› shows bodies of water in every case, but sometimes these seem to be lakes rather than rivers. Questionable is what makes a river a mountain river, but this goes beyond the scope of this study. From the four tags, ‹ice› yielded the weakest result. Nearly all images show snow, but only in around half of the cases ice is clearly visible. Still, the very brief overview

Figure 6.30: Image 265769810 (imageID) located in the Jungfrau-Aletsch region. Extracted tags: ‹Mountain›, ‹Ridge›, ‹Highland›, ‹Alps›, ‹Massif›, ‹Summit›, ‹Fell›, ‹Wilderness›, ‹Valley›, ‹Sky›, ‹Moraine›, ‹Cirque›, ‹Hill›, ‹Rock›, ‹Landscape›, ‹Geology›, ‹Terrain›, ‹Photography›, and ‹Glacier›. Image (http://www.flickr.com/photos/61474423@N00/265769810/) by «g.h.vandoorn», licensed under CC BY-NC 2.0.

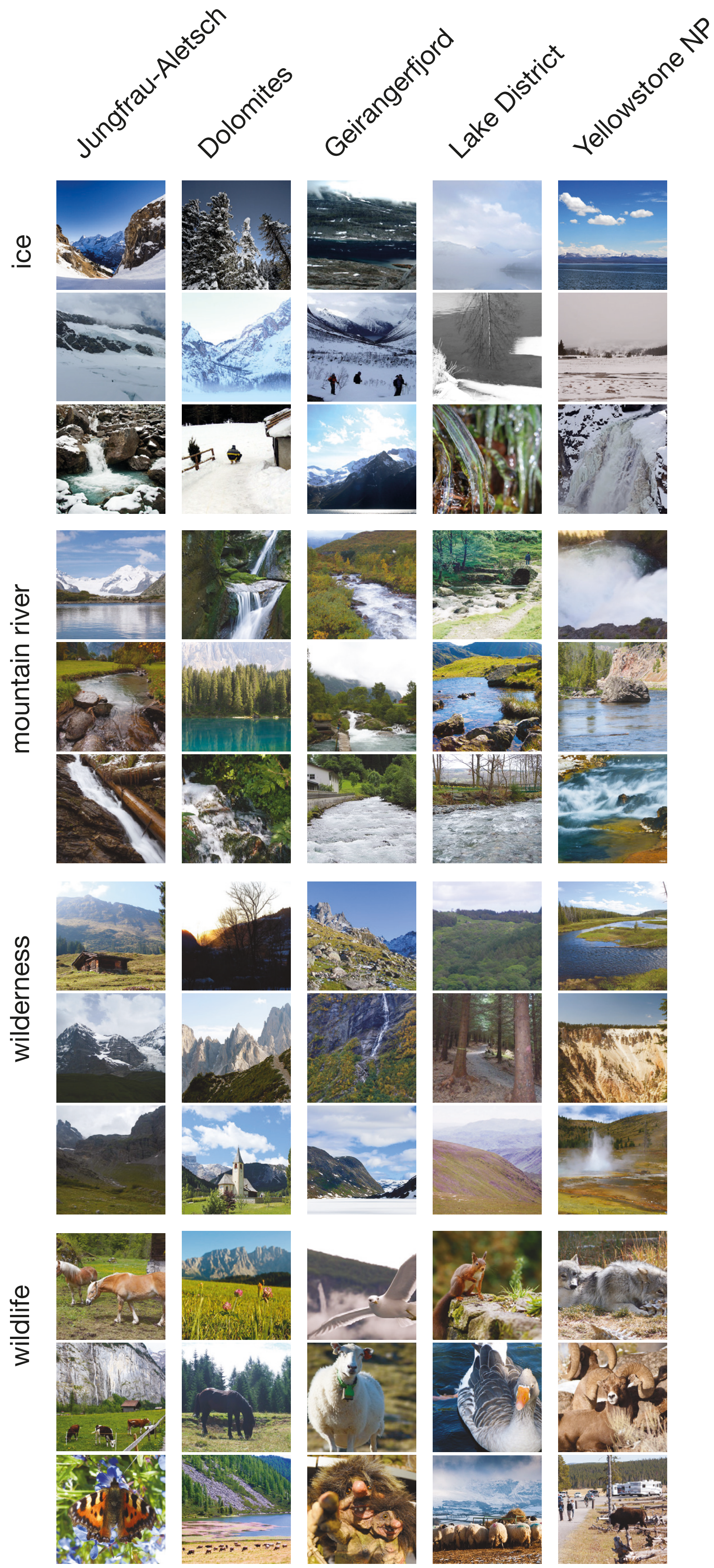


Figure 6.31: Sample images for each AOI and the four tags ‹wildlife›, ‹wilderness›, ‹mountain river›, and ‹ice›. Images credits are found in A.2.2 Sources for Figure 6.31 (Appendix).

gives a promising glimpse into the potential resulting landscape elements used at later stages. An industry study comparing various image recognition services (Perficient Digital Agency 2019) shows promising results. The reported accuracy is 81.7%, and in the comparison with human hand tagged results, especially with a focus on landscapes, the Vision API performs the best out of the four tested products (AWS Rekognition, IBM Watson, Microsoft Azure, and Vision API). As these are results from an industry study, they should be treated carefully. Nevertheless, the study supports the general impression gained from *Figures 6.28–6.30*.

As the existing data is filtered one last time for the final $\text{subset}_{\text{analysed}}$, it seems appropriate to already report on that filtered dataset. Otherwise, the gained insights lose their value, as they do not represent the data used for the final analysis. Therefore, all reported numbers from here on – if not stated differently – are based on the $\text{subset}_{\text{analysed}}$. A total of 26 758 images (–6 713 to $\text{subset}_{\text{tagged}}$) by 1 051 users (–382) are tagged by the Vision API, resulting in 550 996 total tags (–138 476 tags) and 5 254 unique tags (–340) found in the $\text{subset}_{\text{analysed}}$.

The ten tags found most often over all AOIs are landscape (16 725), mountain (15 068), sky (12 890), mountain range (12 402), hill (11 995), hill station (10 663), fell (8 602), mountainous landform (8 405), tree (8 203), and ridge (8 147). Most of the top tags are relatively unspecific and relate to the category of superordinates (i.e. mountain, sky or mountainous landform, see. *6.3.3 Tag Structures*). Besides the generality of the extracted terms, some of them seem to be quite similar as well, for example mountain range, mountain, and mountainous landform. To get a better understanding the results are presented in short chapters denoted to the three main grouping variable (Image, DHC, and AOI). To be able to display the tags in a meaningful way, their sorting is determined by the result of their hierarchical clustering (see *Figure 6.34*).

6.3.2 Tag Characteristics

Image

Each image is annotated with 1–60 tags, while most featured 20–30 tags (see *Figure 6.32*). The pre-defined tag maximum of 100 tags per image for the Vision API is never reached. The distribution varies slightly depending on the AOI. The histogram of the tag count per image taken in the Yellowstone NP shows a peak at around 15–20 tags. For all other AOIs this peak is found at around 30 tags per image.

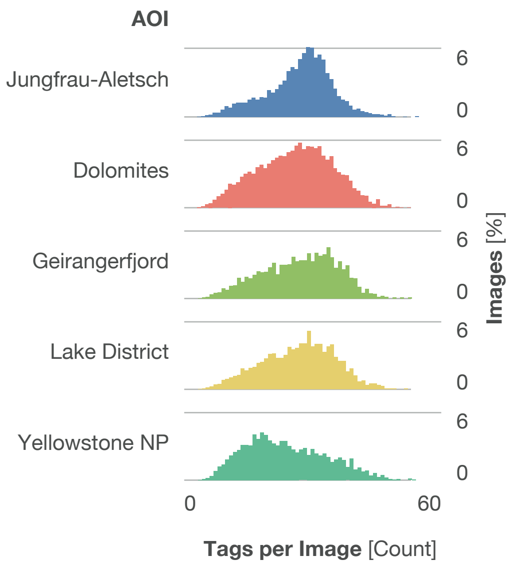


Figure 6.32: Number of extracted tags per image and AOI.

Country

As already mentioned, the distribution of the DHCs in between the AOIs is clearly non-random. In all AOIs the largest portion of users is found to be from the country that embeds the AOI (in the case of Jungfrau-Aletsch and the Dolomites not visible, as the maximum user count was capped at 50). A second observation is that the United Kingdom (208 users) and the United States (201 users) are the only DHCs that are found at least once in all AOIs. Missing in just one AOI are France (53 users), Italy (89 users), the Netherlands (74 users), and Spain (88 users). *Figure 6.33* gives an overview of user, image, and tag count for each DHC stratified by AOI. The imbalance of the sample is clearly visible over all levels: the three variables, the AOIs, and the DHCs. While there are quite big changes in between the distributions of user counts and their image counts, this is not the case for image and tag counts. *Figure 6.34* depicts their nearly perfect linear relationship. Labelled are all DHCs that feature more than 20 000 tags (the exact value of the threshold is arbitrary), while all labels in bold represent the country in which the respective AOI is located. Except for Norway, these are always the countries with the most images (and tags).

Further filtering the subset$_{\text{analysed}}$ to only these countries with sufficient users in each AOI (or similar approaches) would result in further loss of data available for the final analysis. Previous pre-processing steps already reduced the number to just 14 different DHCs. Thus, alternative approaches are needed to overcome the inequality. A potential solution is the inclusion of all users but a standardisation of tag counts by DHC and AOI. By standardizing each tag by all tags found for each DHC in each AOI, the resulting value becomes comparable. In

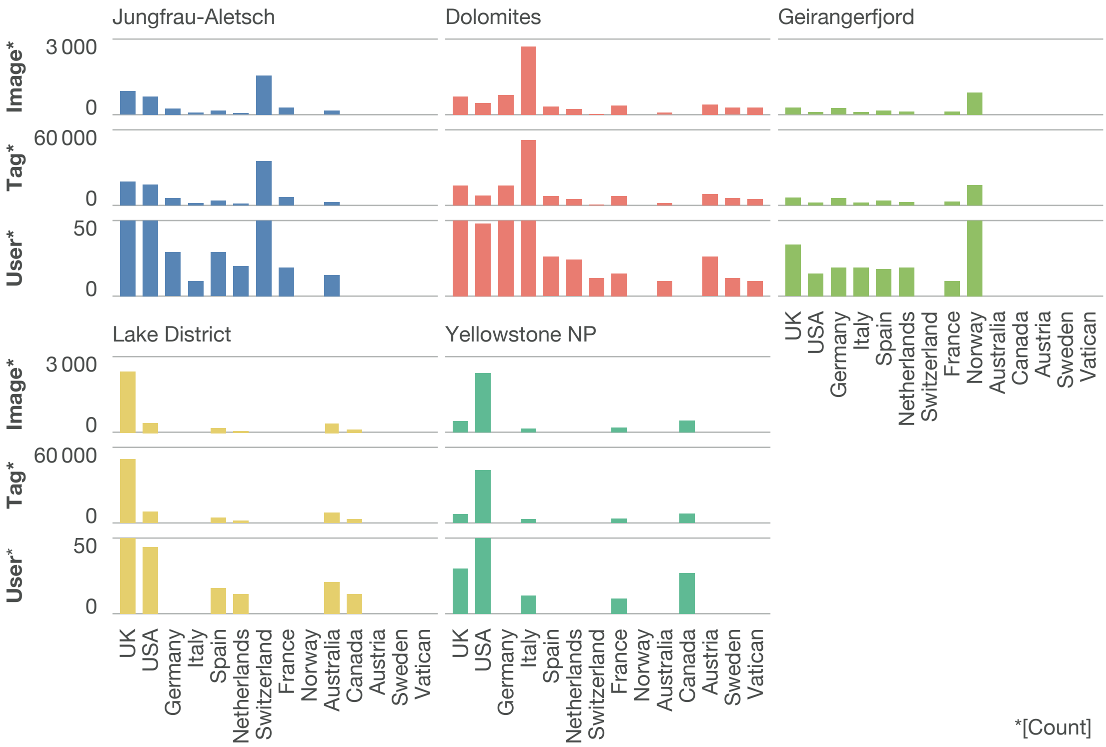


Figure 6.33: Overview of subset$_{analysed}$ comparing user count, image count, and tag count per DHC and AOI.

order to reduce the number of tags for display, a simple measurement is taken: only tags that are available in half of the possible combinations of AOI and DHL are included (*Listing 6.1*).

(Listing 6.1) (5 [*aoi*] × 14 [*derived home locations*]) ÷ 2 = 35

This reduces the number drastically from 5 254 to 157 *relevant* terms, as 43.4% of all tags are only found in one AOI (see *Table 6.7*). In the unexpected case that a tag is found in all DHCs, it would have to be found in at least three AOIs, limiting the potential tags to slightly less than 30%. The circumstance that by far not all tags are found in each DHC explains only 157 tags. Only North American and European countries are left in subset$_{analysed}$ due to the restriction of minimal user counts per DHC. Based on the findings that less similar cultures show greater differences in landscape preferences (see *2.2.1 Cultural Differences in Landscape Perception*), a qualitative look is taken at some DHCs in subset$_{tagged}$, that still holds many of the countries with lower user counts. In order to still get a large enough sample size, all users from the 13 Asian countries available are aggregated. This subset is then compared to the rest of subset$_{tagged}$ that still holds all other countries which are almost completely located in Europe or North America. Visualised in *Figure 6.35* is the relation of the tag counts within the two groups in the form of the ratio of their tag counts. The number of tags found in the Asian countries is divided by the number of tags found in the remaining $subset_{tagged}$. The colour (Asia in blue, all other countries in red) indicates which group features the higher normalized tag count (count per tag divided by the sum of all tags within the group). Only tags that show a large enough difference are visualised (standard deviation of the relative frequency within the two groups larger than 0.03). The threshold of 0.03 is arbitrary and was set at a level that reduced the number of tags to a feasible number. With a share of over 30%, the tags ‹train station›, ‹metro›, and ‹tram› stand out. The whole cluster to their right is devoted to public transport and shows clearly above average values. All these tags are in blue, as the

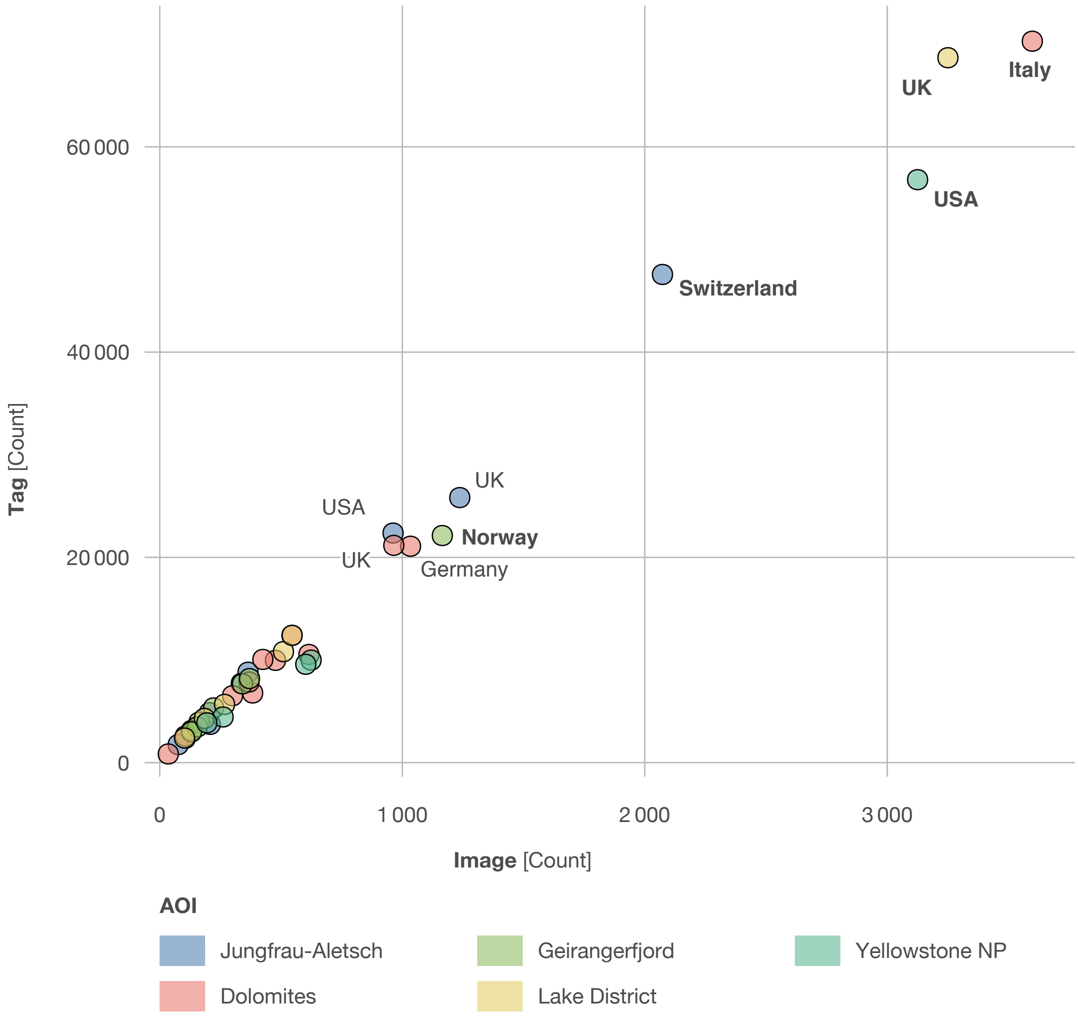


Figure 6.34: Correlation of image counts and tag counts by DHC and AOI. DHCs featuring more than 20 000 tags are labelled, and those representing the country the AOI is lying in are highlighted.

AOIs in which this tag is found	Tags	%
1	2 282	43.4
2	956	18.2
3	674	12.8
4	504	9.6
5	838	15.9

Table 6.7: Absolute and relative count of tags extracted according to the number of AOIs they are found in.

relative frequency within the subset of Asian countries is larger than in the remaining countries. Many of the tags that show very low ratio values are associated with outdoor activities.

As much as these results lend to an optimistic conclusion on the cultural influence in landscape perception if observed on a relatively high aggregation level, these should not be seen as anything more than potential trends. The user numbers in the sample of 13 Asian countries are very small (77 users) and several of these tags are just available in a single AOI. Thus, this brief foray is left here and the results based on the pre-processed subset$_{analysed}$ with a focus on the influence of the different AOIs are presented.

Area of Interest
There are clear trends in the sets of tags that are derived from each AOI. This can be seen clearly in the top section *Frequency of Tag* of *Figure 6.36*. While cluster 3 holds mostly tags extracted from images in the AOI Geirangerfjord and Lake District, cluster 6 is dominated by tags found in images from the Dolomites. These differences – given that the sample is representative and not heavily biased – are likely due to the different nature of the AOIs' landscapes.

So far, the focus has been on tags that are found across DHCs and AOIs. As seen before, this limits the number of used tags drastically. The issue with the full tag set is a phenomenon called the *long-tail*. Besides a small number of tags with high counts over all groups (AOI and DHC), there are many more tags with counts as low as one. For now, these tags are left in the subset, as they might hold valuable information on differences between users of different DHC. Nevertheless, the tag collection shown in visualisations is often reduced to earlier introduced 157 *relevant* tags to increase readability.

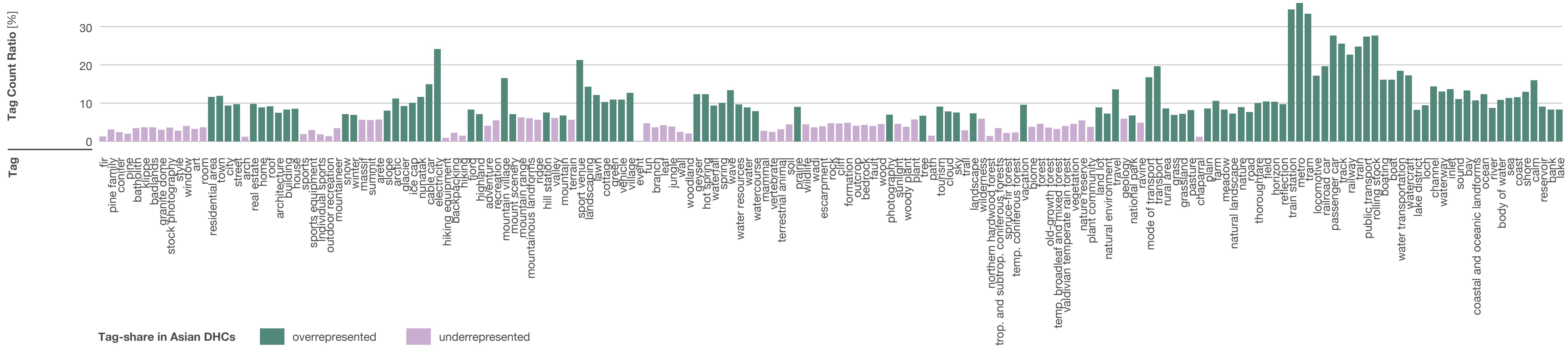


Figure 6.35: Tag counts from Asian countries divided by the tag counts found attributed to the remaining countries for the 157 relevant terms. The values are coloured according to their relationship with the expected value calculated on the assumption that the tags are evenly distributed between users, independent of their DHC.

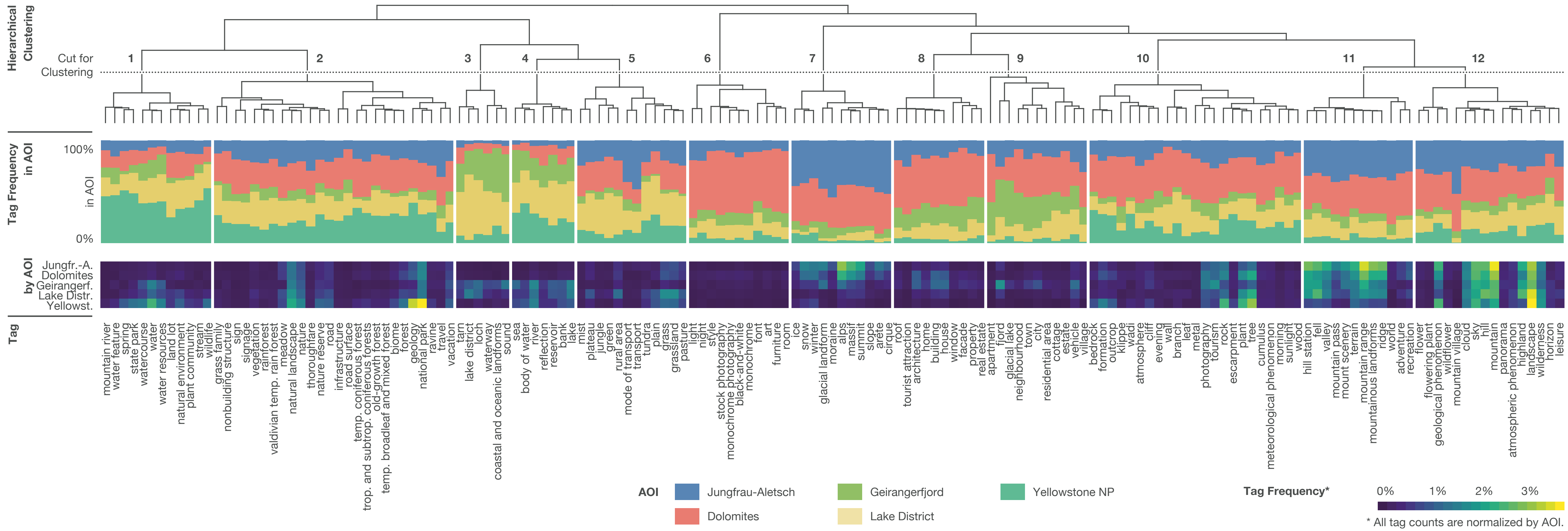


Figure 6.36: Relative distribution of tags within the five AOIs. For each tag their frequency between the five AOIs (top) and within each AOI (bottom) is visualised. Example: Around 50% of all ‹mountain river› (first tag from left) tags extracted from subset$_{analysed}$ are found in Yellowstone NP. Compared to the other tags found in Yellowstone, the frequency is low. This contrasts with the tags ‹geology›, ‹national park› and ‹landscape› that show the highest frequencies within this AOI.

6.3.3 Tag Structures

Hierarchical Clustering

The full tree with its 5 254 tags is too large to display. A smaller version, based on the tag reduction presented earlier, shows 157 tags and the corresponding clustering (see top of *Figure 6.36*). The tree is cut at a height that results in 12 clusters, of which most have a clear thematic focus point. It is observed that there are several variations of clusters that could be named ‹landforms› with no clear thematic distinction. However, for the study it is not important how these clusters could be named, yet it is crucial to see that hierarchical clustering generates thematic clusters of tags.

Similarity

So far, the extracted tags have been looked at with a focus on their overall structure and frequency. In a next step, the intra-tag relationship is analysed. As it already became clear in *Figure 6.36*, there are certain tags that are often extracted in groups from images. Several tags correlate heavily. To understand the co-correlations of tags, the selection of *relevant* tags is further analysed. For each pair of tags, the co-correlation is calculated. The resulting co-correlation matrix is shown as heat maps in *Figures 6.37–6.41*. Although, the co-correlation matrix for all five AOIs shares an overall pattern, they still all have their peculiarities:

- Jungfrau-Aletsch, Dolomites, Geirangerfjord, Lake District; All these AOIs feature a high correlation values in the top right (1). These tags mostly relate to geomorphological features and mountainous landscapes in general. Yellowstone NP features the a similar, but much weaker pattern.
- Geirangerfjord, Lake District: Both AOIs show an area of high correlations towards the lower left corner (2). The tags correlating are mostly associated with forms and bodies of water. However, the two AOIs also have a major difference found in around the centre (3) of the co-correlation matrix. Geirangerfjord shows high correlation values, while the ones for the Lake District are comparably low. The thematic focus in these clusters is mostly on mountainous landforms, alps and winter.
- Yellowstone NP: What is remarkable about this AOI, is the observation, that all co-correlation values in general tend to be lower than in the other AOIs.

Overall, it can be stated that for the 157 *relevant* tags compared there are clear patterns of co-correlation. This knowledge is helpful in two ways. First, it helps to better understand which tags are often found together. With large groups of strongly co-correlating tags over all AOIs, even if the extracted tag collections are vast, their meaningfulness is reduced. Fortunately, the visible clusters are not too large nor are their co-correlations all over 0.5

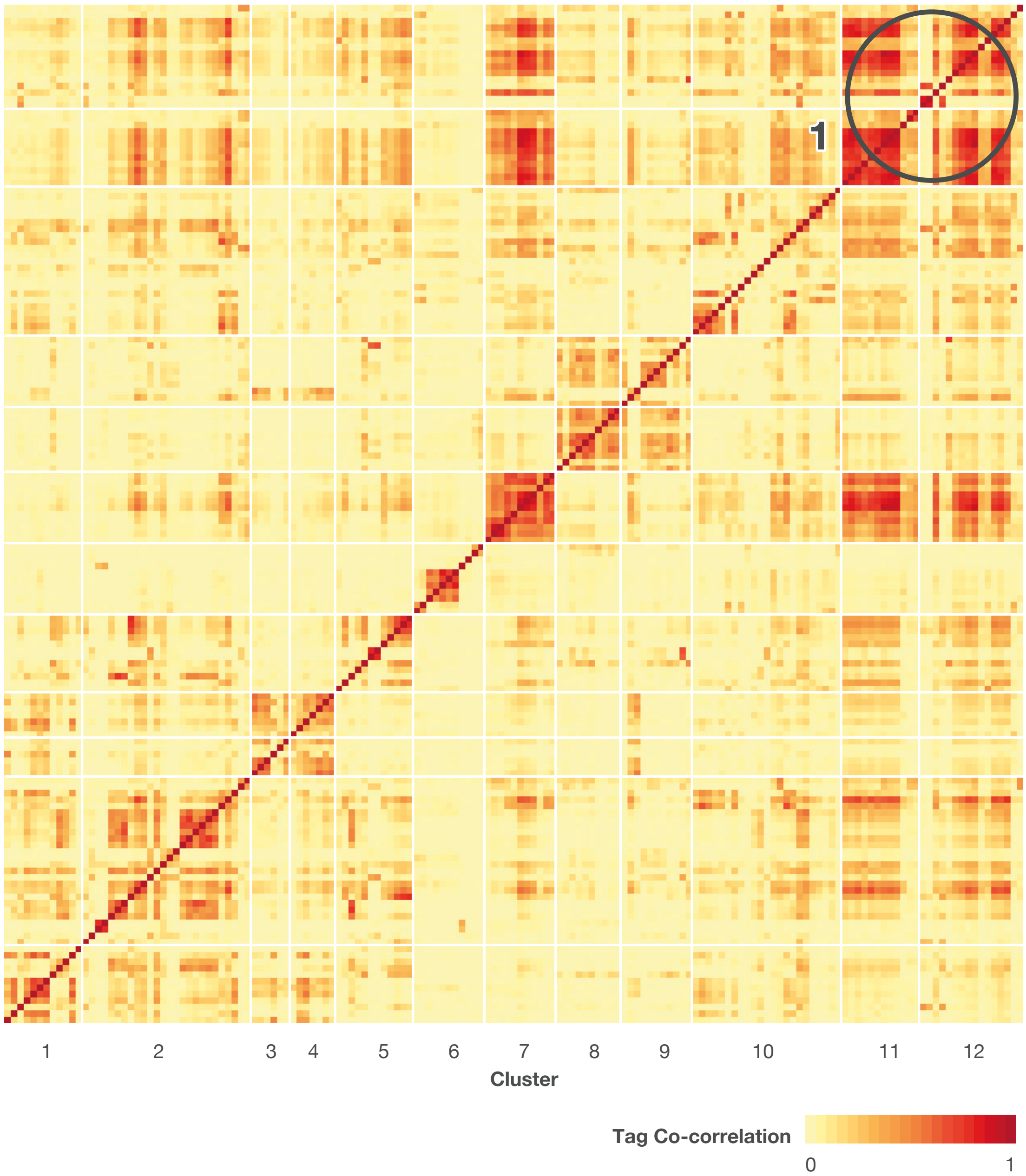


Figure 6.37: Co-correlation matrix with the 157 relevant tags in Jungfrau-Aletsch. The tags are sorted based on the results of the hierarchical clustering and for readabilty reduced to the cluster number. High correlation values correspond to pairs of tags that are often extracted from the same images. The marked area (1) is referenced in 6.3.3 Tag Structures.

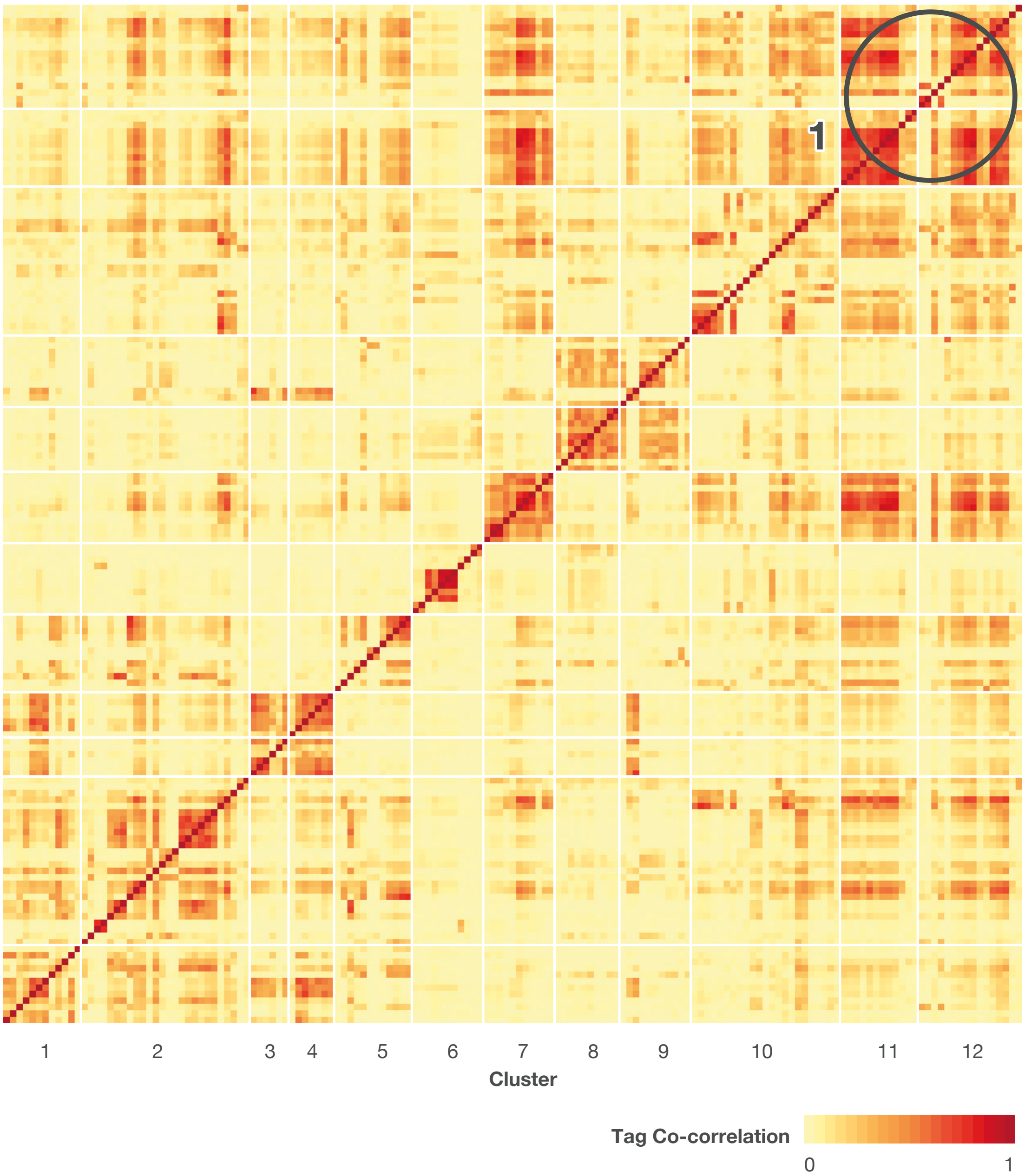


Figure 6.38: Co-correlation matrix with the 157 relevant tags in Dolomites. The tags are sorted based on the results of the hierarchical clustering and for readabilty reduced to the cluster number. High correlation values correspond to pairs of tags that are often extracted from the same images. The marked area (1) is referenced in 6.3.3 Tag Structures.

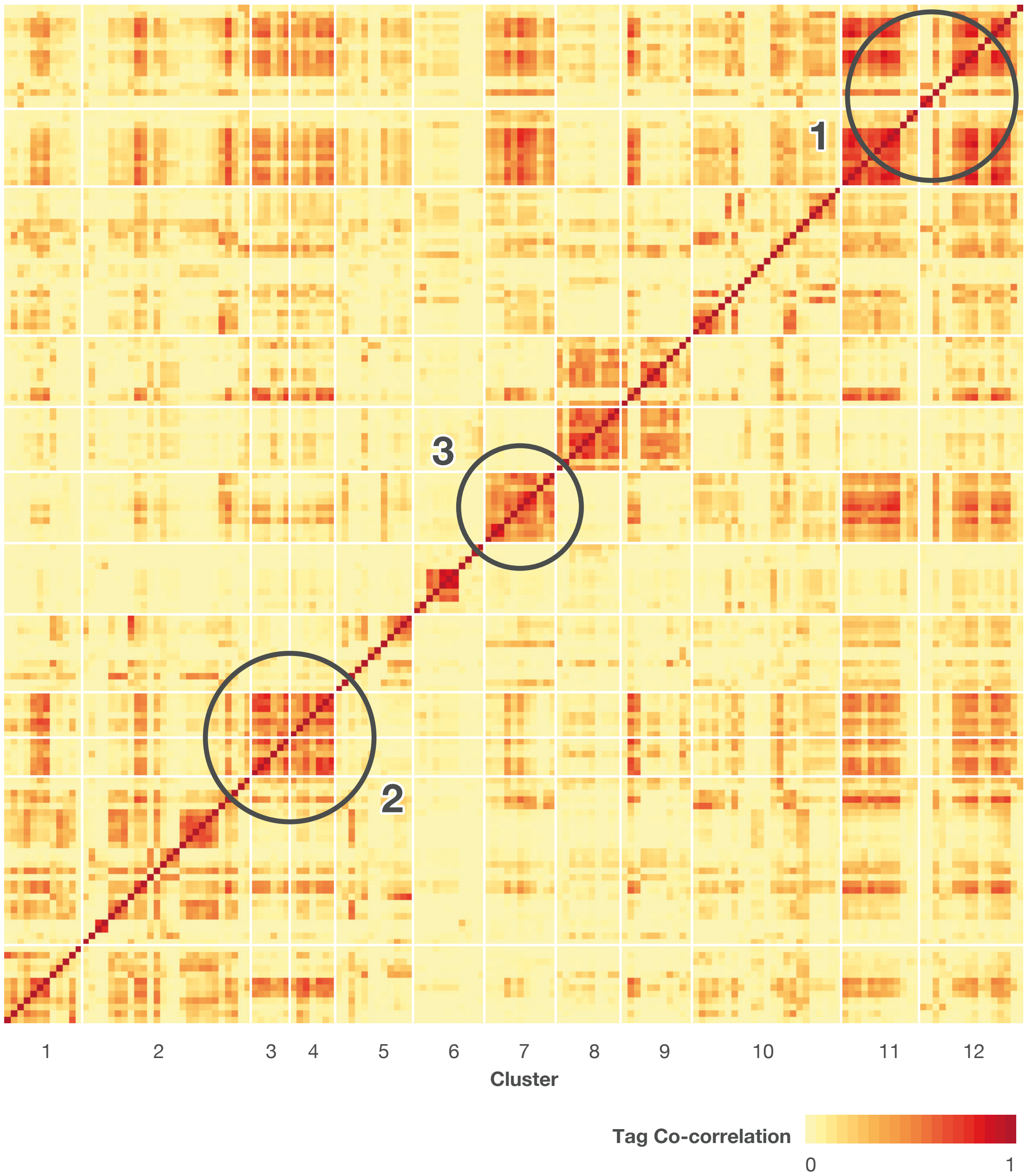


Figure 6.39: Co-correlation matrix with the 157 relevant tags in Lake District. The tags are sorted based on the results of the hierarchical clustering and for readabilty reduced to the cluster number. High correlation values correspond to pairs of tags that are often extracted from the same images. The marked areas (1–3) are referenced in 6.3.3 Tag Structures.

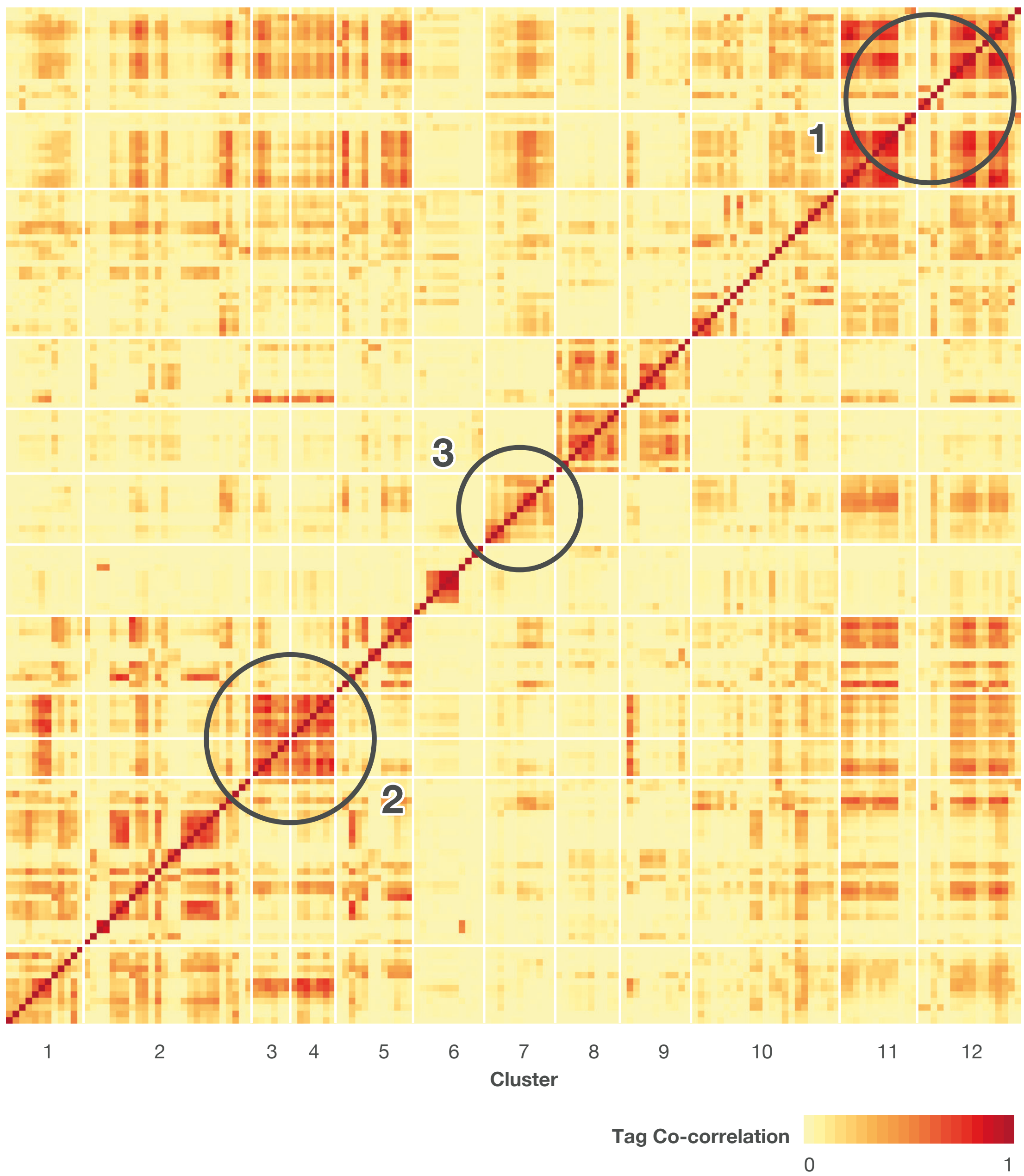


Figure 6.40: Co-correlation matrix with the 157 relevant tags in Lake District. The tags are sorted based on the results of the hierarchical clustering and for readabilty reduced to the cluster number. High correlation values correspond to pairs of tags that are often extracted from the same images. The marked areas (1–3) are referenced in 6.3.3 Tag Structures.

Figure 6.41: Co-correlation matrix with the 157 relevant tags in Yellowstone NP. The tags are sorted based on the results of the hierarchical clustering and for readabilty reduced to the cluster number. High correlation values correspond to pairs of tags that are often extracted from the same images.

(which would mean that in half of the cases these tags were found together). For the cases where high co-correlation values were found, they seem to make a lot of sense, for example in ‹summit› and ‹massif›. Second, it gives an idea on how the five AOIs are similar – or different – to each other. The overall pattern, as already briefly described, gives an idea of the relative composition of tags within an AOI. This helps to find similar (Jungfrau-Aletsch and Dolomites) or dissimilar (Yellowstone NP and Geirangerfjord) AOIs. The similarities found in *Figures 6.37–6.41* do reflect the ones found in *Figure 6.36*.

It can be concluded that an overarching pattern of co-correlating tags is visible, but there are clear variations by AOI, especially between Yellowstone NP and the remaining AOIs. As the co-correlation vary in strength for all tags over the various AOIs, there is no need to aggregate or exclude tags.

Quality

As introduced in *2.1.4 Folksonomy*, tags can be categorized along three levels. This helps to shed some light on the character of the derived tags. The proportion of superordinate, basic, and subordinate level tags gives a hint on how much detail the Vision API is able to extract from images. Ultimately, this is of course also a question of accuracy. However, *Figures 6.28–6.30* have shown that the tendency is towards correct and valuable results, a finding that is also supported by an industry study comparing various image tagging providers (Perficient Digital Agency 2019). *Figure 6.42* shows the distribution of a classification of tag levels found in Seresinhe, Preis & Moat (2017: 6, *Figure 2*) and the 157 *relevant* tags from the Vision API used in this study. In Seresinhe, Preis & Moat, there is a clear differentiation between very few superordinate level tags on the one hand, and an equal number of basic-level categories and subordinate level tags on the other hand. This is not as clear for the tags extracted by the

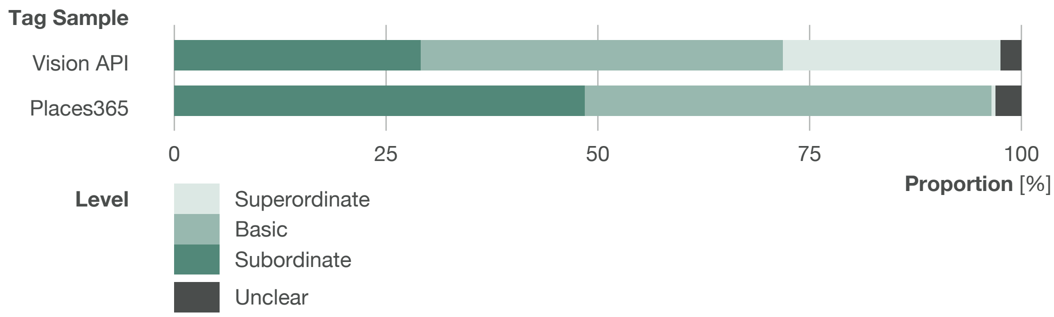


Figure 6.42: Comparison of the distribution along linguistic levels based on the tags found in Figure 2 in Seresinhe, Preis & Moat (2017) and the 157 relevant tags extracted with the Vision API. See Table A.1 (Appendix) for the classified data.

Vision API. Here, the tags are relatively equally distributed over the three levels, with slightly less tags at the superordinate level than in the other two. The main differences between the two classifications is that the Vision API returned many more (25 percentage points) tags at the superordinate level at the expense of subordinate level tags. It is expected that tags at the superordinate level might not be as helpful as subordinate level tags in finding cultural differences in landscape preferences because they are too general to reflect any differences or similarities. Therefore, the quality of tags might be slightly lower than the tags extracted with PlacesCNN as done in Seresinhe, Preis & Moat (2017).

Another interesting view on quality is the amount of landscape elements within the tags returned and their classification along known landscape element classes.

Matching tags with theory

The 157 *relevant* tags in the subset$_{\text{analysed}}$ generally show a high overlay with the introduced six classes as described by Conrad (2011), see *2.1.3 Landscape Elements*. *Table A.2* (Appendix) displays the tags that fit into each of the six categories, namely (A) *rural characteristics*, (B) *natural landforms*, (C) *cultural features*, (D) *specific locations*, (E) *intangible aspects*, and (F) *visual aesthetic qualities*. The results of this classification are summarised in *Figure 6.43*. Most tags are found within the class (B) *natural landforms*. Small numbers are classified as (A) *rural characteristics*, (D) *specific locations*, (E) *intangible aspects*, and (F) *visual aesthetic qualities*. No tags are found in the (C) *cultural features* class. 60% of the 157 tags could be matched with any of the six landscape element classes. This means 40% of the tags could not be matched with any of the six classes proposed by Conrad (2011). Instead of excluding these tags, they are

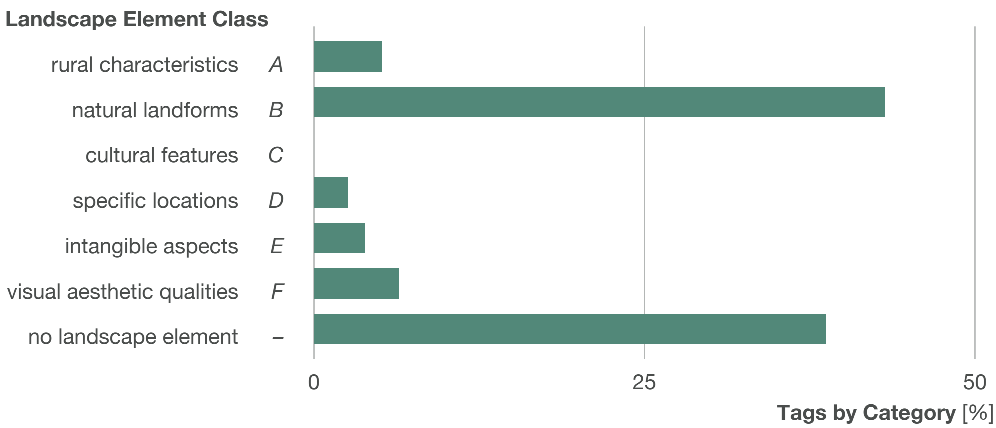


Figure 6.43: Distribution of the 157 relevant tags by landscape element classes as defined by Conrad (2011). See Table A.2 (Appendix) for categorisation of tags.

left in the sample. First, these tags could potentially explain cultural differences as – at least most of them – do describe elements visible in the images. That they are not directly relatable to the classes discussed above is not reason enough to further reduce the subset$_{analysed}$. Second, a reduction by 40% of the 157 tags would result in an even more limited number of tags in a subset that is already filtered by users, AOIs, and image counts. Therefore, subset$_{analysed}$ is not further filtered and ready for the last step of the analysis, as described in the next section, the analysis of cultural differences in landscape preferences.

6.4 Analysis of Cultural Differences in Landscape Preferences

In this last chapter, the information on DHCs and tags is coming together. Although numerous figures already incorporated both variables, it is not until now that the focus is solely on the potential cultural differences in landscape preferences. The dataset on which the analysis is performed is the subset$_{analysed}$ and the limitation to the 157 *relevant* tags as described in *6.3.1 Overview* is in place if not stated differently. In a first step, the tag frequencies along each DHC and AOI are visualised. The previous information on tag frequencies between the AOI and on the clustering result are added for context. In a second step, multidimensional scaling is used to reduce the 157 or up to 5 254 dimensions, each reflecting the count of a unique tag, down to two dimensions to make the data accessible and potential patterns visible. In a third and last step, the MDS results are tested for their sensitivity.

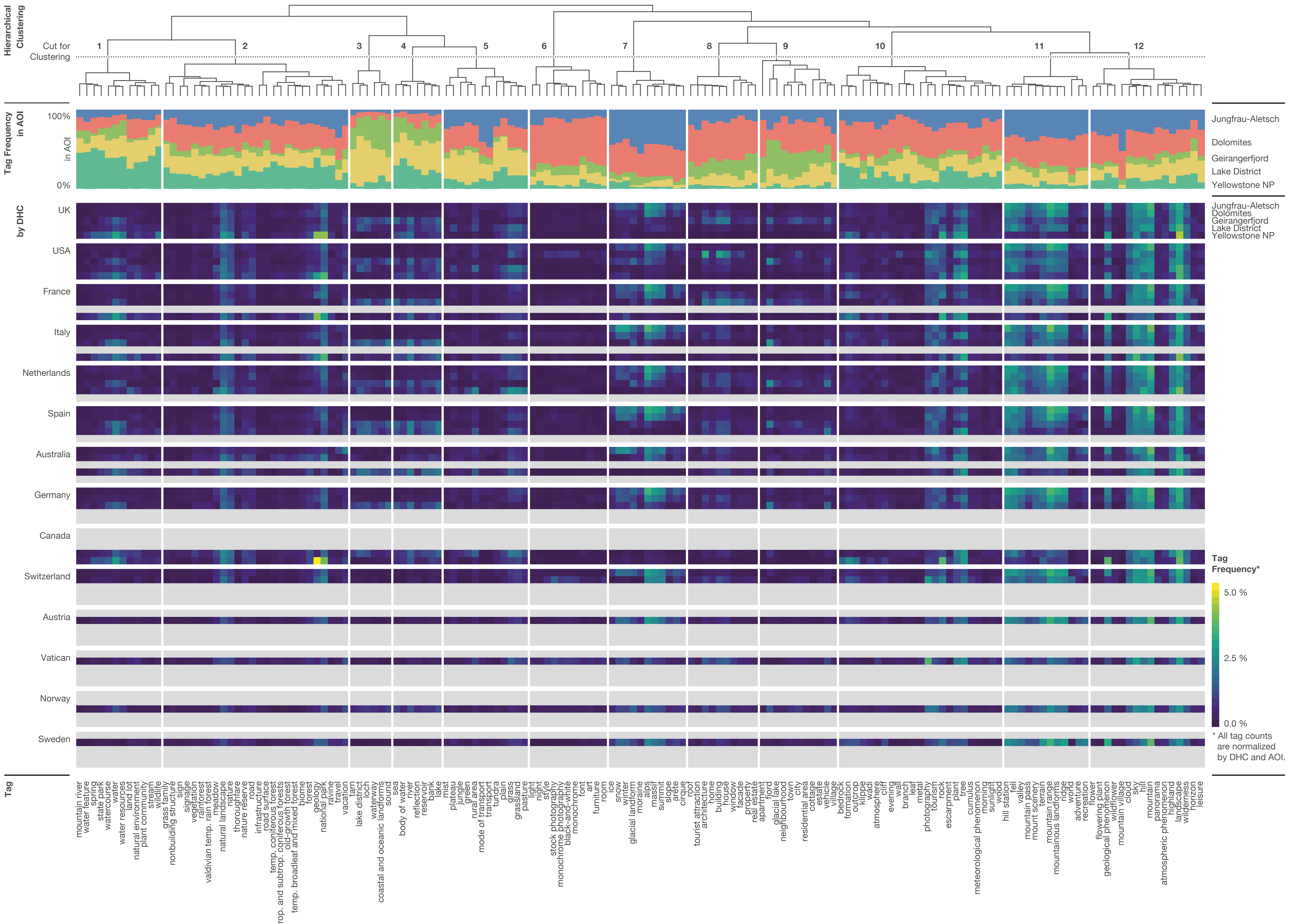
6.4.1 Mapping Cultures

As a starting point and to gain an overview over the 157 tags and their distribution over the 14 DHCs, an extensive heat map is generated based on the normalized frequencies of tags within each DHC and AOI. The resulting visualisation (*Figure 6.44*) gives the following insights:

- Overall, the pattern looks very similar for all DHCs and AOIs. There is only a very small number of irregularities between the DHCs (*potential cultural differences*). An example of such an irregularity is the very high value for the tag ‹geology› for Canada (AOI: Yellowstone NP). The pattern is visible in the other DHCs that include users in Yellowstone NP as well, but not as strongly.

▶

Figure 6.44 (right page): Overview of subset$_{analysed}$ showing the tag frequency normalized by DHC and AOI. Additional information on the clustering as well as the tags' distribution over the AOIs is given in the top section (as already introduced in Figure 6.36).



Hierarchical Clustering
Cut for Clustering
1
2
3
4
5
6
7
8
9
10
11
12
Tag Frequency in AOI
100%
0%
Jungfrau-Aletsch
Dolomites
Geirangerfjord
Lake District
Yellowstone NP
by DHC
UK
USA
France
Italy
Netherlands
Spain
Australia
Germany
Canada
Switzerland
Austria
Vatican
Norway
Sweden
Tag Frequency*
5.0 %
2.5 %
0.0 %
* All tag counts are normalized by DHC and AOI.
Tag
mountain river
water feature
spring
state park
watercourse
water
water resources
land lot
natural environment
plant community
stream
wildlife
grass family
nonbuilding structure
sign
signage
vegetation
rainforest
valdivian temp. rain forest
meadow
natural landscape
nature
thoroughfare
nature reserve
road
infrastructure
road surface
temp. coniferous forest
rop. and subtrop. coniferous forests
old-growth forest
temp. broadleaf and mixed forest
biome
forest
geology
national park
ravine
travel
vacation
tarn
lake district
loch
waterway
coastal and oceanic landforms
sound
sea
body of water
river
reflection
reservoir
bank
lake
mist
plateau
jungle
green
rural area
mode of transport
transport
tundra
plain
grass
grassland
pasture
light
night
style
stock photography
monochrome photography
black-and-white
monochrome
font
art
furniture
room
ice
snow
winter
glacial landform
moraine
alps
massif
summit
slope
arete
cirque
roof
tourist attraction
architecture
home
building
house
window
facade
property
real estate
apartment
fjord
glacial lake
neighbourhood
town
city
residential area
cottage
estate
vehicle
village
bedrock
formation
outcrop
klippe
wadi
atmosphere
cliff
evening
wall
branch
leaf
metal
photography
tourism
rock
escarpment
plant
tree
cumulus
meteorological phenomenon
morning
sunlight
wood
hill station
fell
valley
mountain pass
mount scenery
terrain
mountain range
mountainous landforms
ridge
world
adventure
recreation
flower
flowering plant
geological phenomenon
wildflower
mountain village
cloud
sky
hill
mountain
panorama
atmospheric phenomenon
highland
landscape
wilderness
horizon
leisure

- Although the tags are distributed very similarly in between the DHCs, there are great differences between the tags. The most prominent ones to the right (in respect to *Figure 6.44*) are clusters with a focus on geomorphic features (‹hill›, ‹mountain›, ‹terrain›, ‹mountain range›, ‹mountainous landform›, ‹ridge› but also ‹hill station›, ‹landscape›, and ‹wilderness›). Another, but much smaller cluster, is found in the centre, again oriented towards physical elements of the landscape (‹glacial landform›, ‹massif›, ‹summit›, and ‹alps›).
- There are no DHCs that show a distinctly different pattern.

The main limitation of *Figure 6.44* is its complexity. In the per DHC heat map, values for over 5 800 combinations of DHC, AOI, and tags are visualised. Therefore, multi-dimensional scaling (MDS) is used to reduce the numerous dimensions. In every case the dimensions reflect the count of tags within a specific group (users, DHCs or AOIs). The resulting two dimensions are hard to make sense of and are thus not named. Instead, they are referenced as dimension 1 and dimension 2. While a more characteristic naming of the dimensions would have been helpful for the interpretation of the results, it is not central to the analysis of differences in cultural landscape preferences. As dimension 1 is the axis of greatest variation within the data, it is important to be aware that the same distance along dimension 2 (showing the second largest variation), although it is perceived equally long, must be interpreted as smaller difference.

In a first part, each user is visualised individually. The basic idea is a focus on clusters consisting of users with the same DHC. *Figure 6.45* shows these MDS results in which the AOI functions as the grouping variable for the tags. All users are coloured by their DHC. All five plots show relatively distinct distributions. Regarding the distinctiveness of these differences it is important to note that the scales on both dimensions extend no further than 0.75 units from the centre. With most data points found in an even smaller section, the found differences are expected to be relatively small. Clearly, this is nothing more than a hint at the strength of the found effects than an absolute number. The poor goodness of fit (stress 21%) requires a cautious interpretation of the results, but was expected due to the large number of dimensions reduced to only two (Kruskal 1964: 15–21).

Generally, users found in Jungfrau-Aletsch tend to cluster in the first quadrant (0.25/0.20), for the Dolomites this is at 0.25/0, on the border towards the second quadrant. The user density is much lower for Geirangerfjord and the following two AIOs, due to the limited number of users in the sample (see *Figure 6.2*). A clear hotspot is not present, but most users are found in the second quadrant. For users found in the Lake District, there is a cluster found at around 0.00/−0.25 and for the Yellowstone NP it is found at a similar position, with a strong tendency towards the third quadrant. The visualisation does not indicate any distinct clusters of users with the same DHC.

For various reasons (overview, traceability, minimum frequency, and distribution of tags) the number of tags used has been reduced to the 157 *relevant* tags. Interestingly, the results are very similar if performed on the full set of 5 254 tags as shown in *Figure 6.46*.

In a second step, due to the limitations of *Figures 6.45* & *6.46* to show per DHC clusters, the tags are grouped by DHC. The results shown in *Figure 6.47* stem from the same MDS computation as before. This is important, as MDS results tend to change their orientation with each calculation, resulting in flipped or rotated outputs (see *5.5.2 Multi-dimensional Scaling* for details). The colour now reflects the AOI a user's images are found in. If a user is found in more than one AOI, more than one data point per user are added to the plot. The already known circumstances that the samples per DHC are not of equal sizes and not all AOIs are found in each DHC are clearly visible. For example, all users localised in the Czech Republic took their images in the Dolomites. Similarly, Norway only features users whose images were taken in the Geirangerfjord and the Lake District. But these findings are not why this figure was generated in the first place. The focus lies on patterns, or more precisely, on clusters which are to be expected for each DHC given that their allocated users' images resulted in different tag sets. This does not seem to be the case. Although there are different patterns for the various countries, they are strongly dominated by the AOI a user's images were taken in, as discussed above. To better understand this influence, *Figure 6.48* structures the visualisation further and combines the two previous figures by using both AOI and DHC as grouping variables. The resulting matrix shows the influence of the AOI on the vertical axis and the DHCs on the horizontal axis. Again, the strong patterns for each AOI are visible while there is no visible congruency on the horizontal axis. These observations are all made at the level of single users, and not DHCs. The next section therefore focusses on the aggregated data by DHCs and the thereby resulting tag collections.

In this second part, users are aggregated to their DHCs. The aggregation of users results in single data points per DHC that could suggest that there is no variation within. It is important to be aware of this limitation and special caution is taken to satisfy this shortcoming in the following discussion. *Figure 6.49* shows the results of an MDS based on the tag collections found in each DHC. Each data point is once again coloured by the AOI its images and ultimately tags were extracted from. Moreover, are the points are scaled according to their users' sample sizes. The well-known influence of the AOI is clearly visible here as well. The aggregation removed much of the noise and shows – deceptively – clear patterns in between the

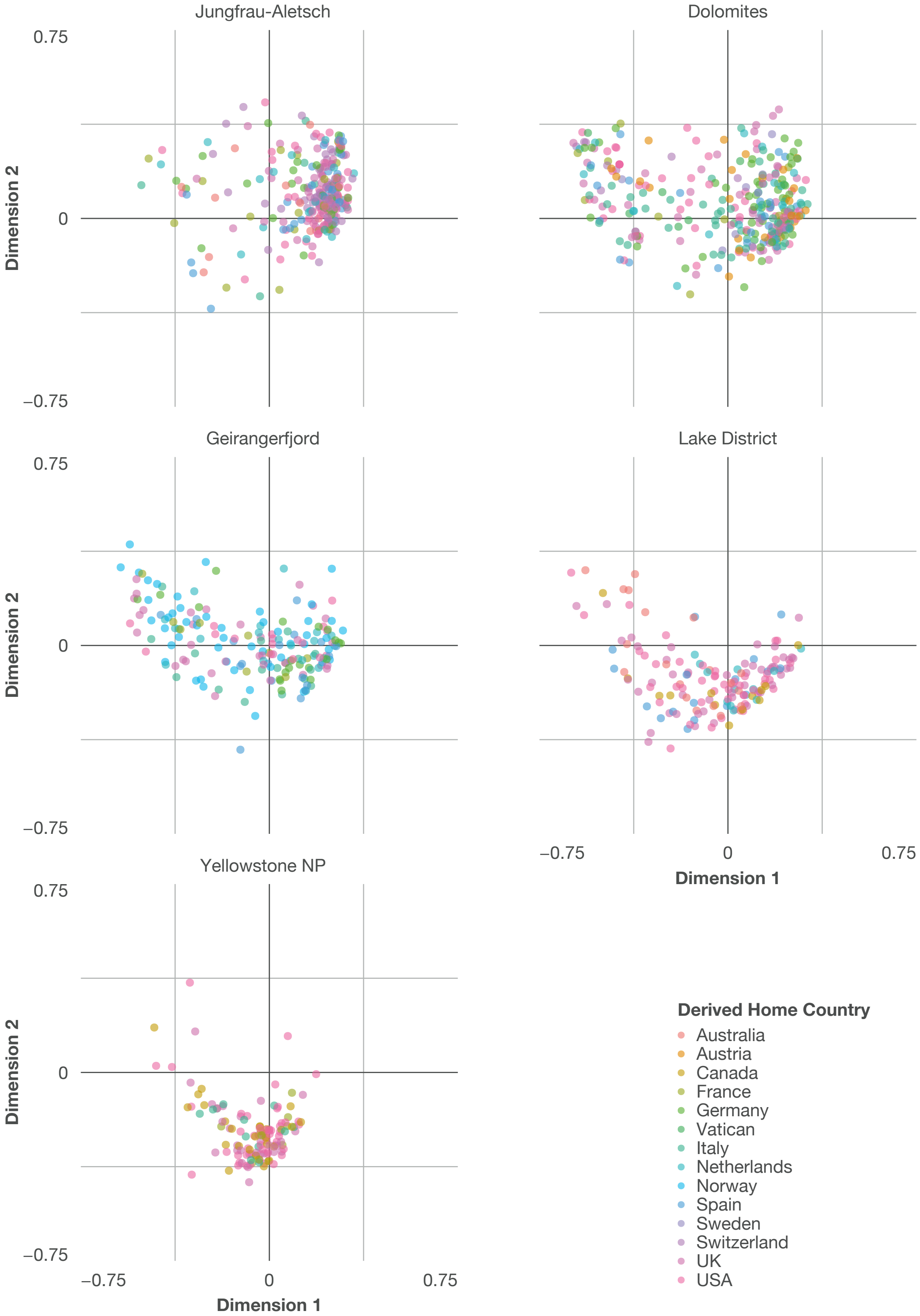


Figure 6.45: Result of MDS of users based on 157 relevant tags, facetted by AOI and coloured by DHC.

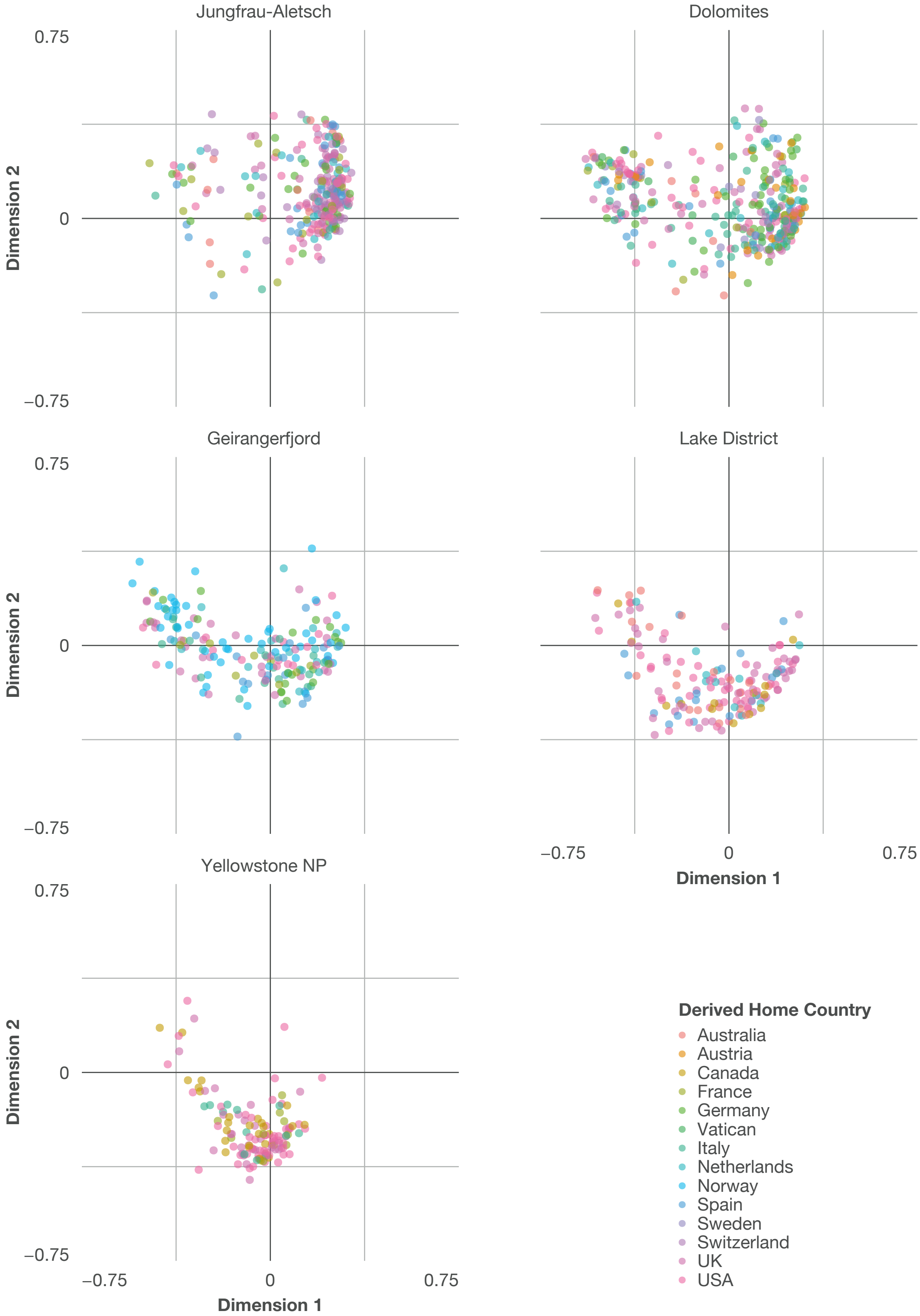
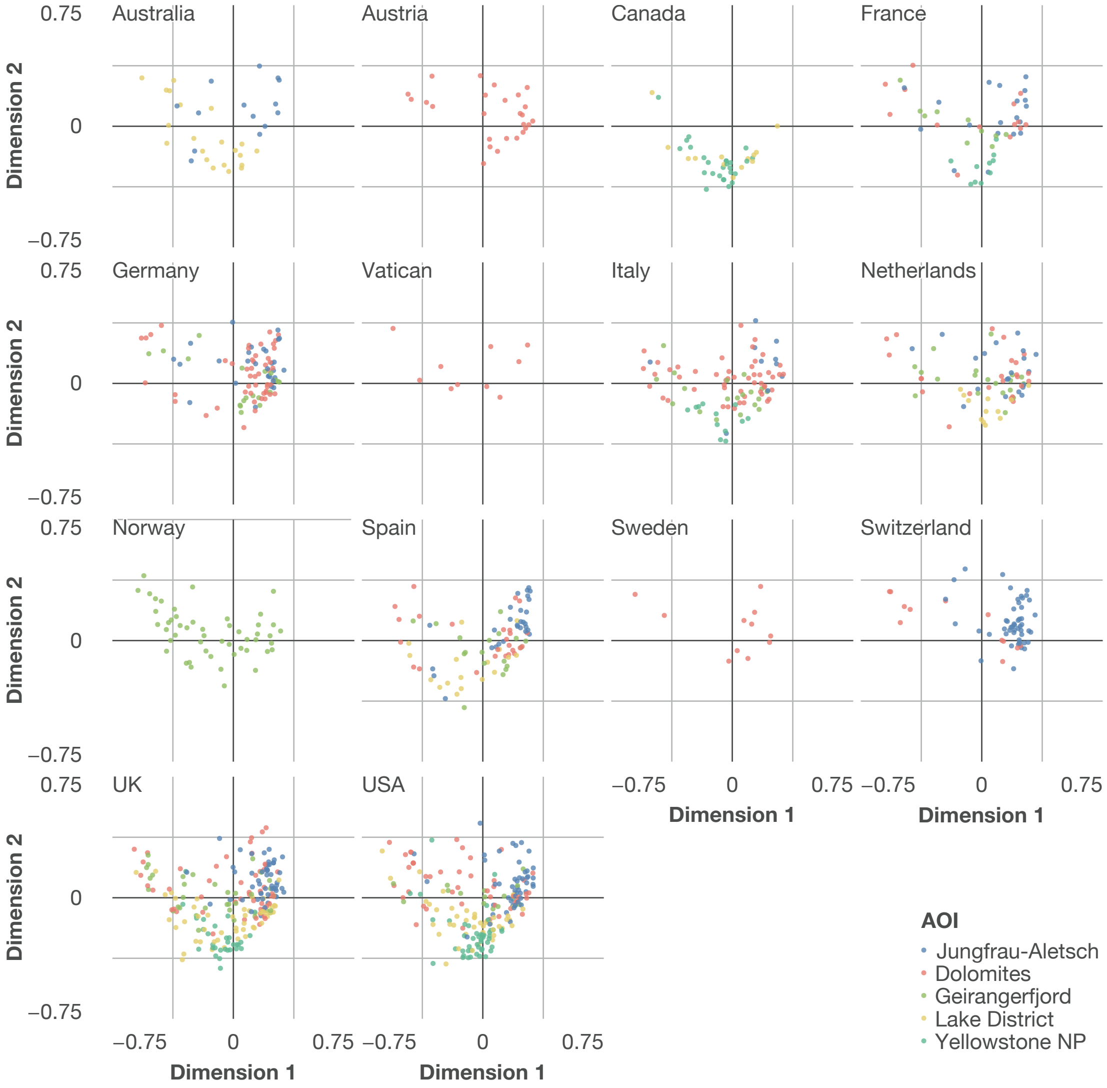
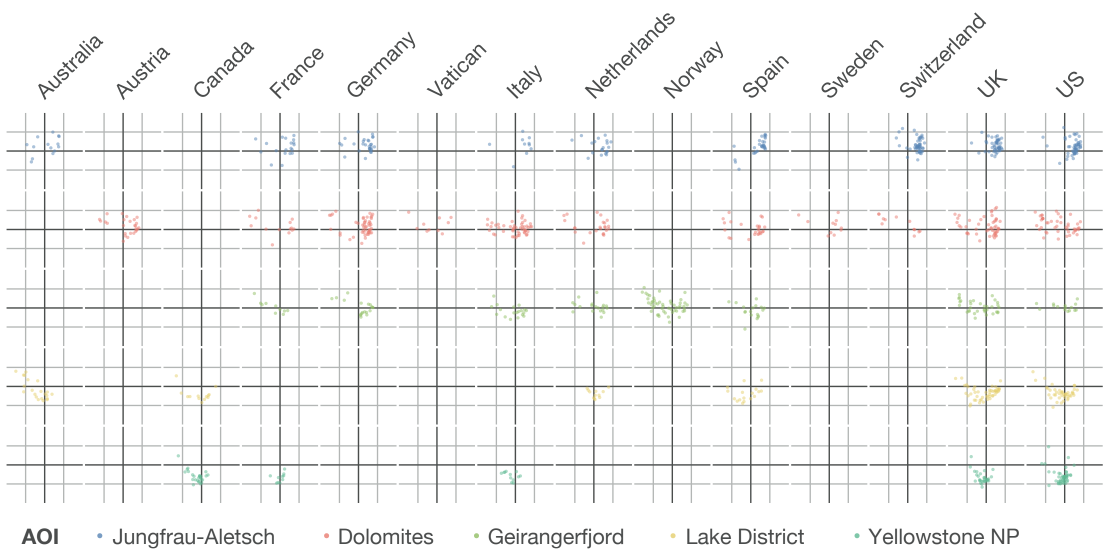


Figure 6.46: Result of MDS of users based on all 5 254 tags, facetted by AOI and coloured by DHC.


Australia
Austria
Canada
France
Germany
Vatican
Italy
Netherlands
Norway
Spain
Sweden
Switzerland
UK
USA
Dimension 1
Dimension 2
0.75
0
−0.75
AOI
Jungfrau-Aletsch
Dolomites
Geirangerfjord
Lake District
Yellowstone NP


Australia
Austria
Canada
France
Germany
Vatican
Italy
Netherlands
Norway
Spain
Sweden
Switzerland
UK
US
AOI
Jungfrau-Aletsch
Dolomites
Geirangerfjord
Lake District
Yellowstone NP

◀

Figure 6.47 (left page, top): Result of MDS of users based on the 157 revelant tags found in subset$_{analysed}$, facetted by DHC and coloured by AOI.

◀

Figure 6.48 (left page, bottom): Result of MDS of users based on all unique 1057 tags found in subset$_{analysed}$, facetted by DHC and AOI. The scales represent the two MDS dimensions, both ranging from –0.75 to 0.75.

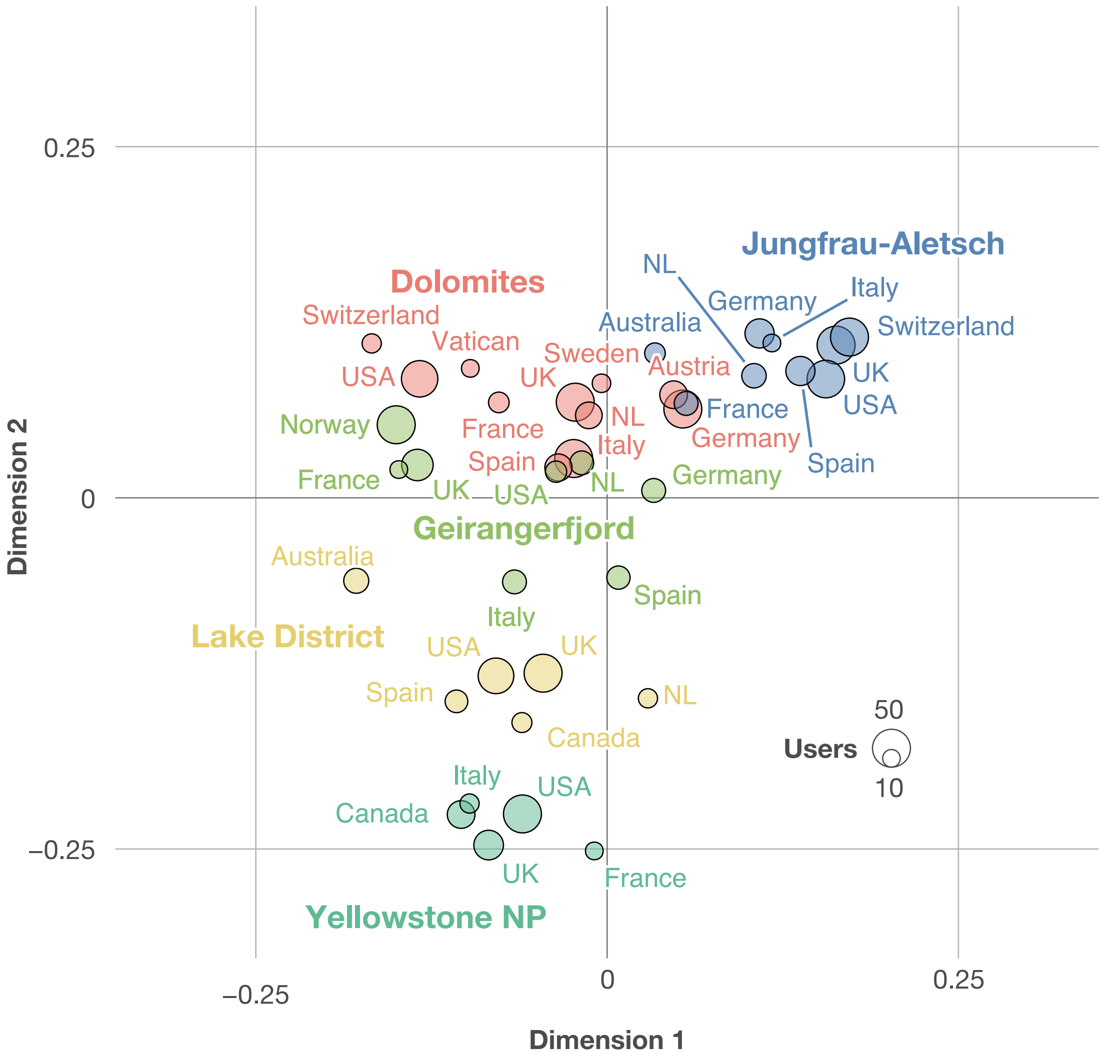


Figure 6.49: Results of MDS based on the 157 relevant tags and grouped to DHCs.

various DHCs. Deceptive because of the above-mentioned problem of the hidden variation within the DHCs. Nevertheless, the following observations are made:

- Larger samples tend to be near each other, while smaller samples are found in more ‹remote› positions.
- While the Jungfrau-Aletsch and the Dolomites show small distances, all the other AOIs are quite separated.
- DHCs in the Geirangerfjord area less densely clustered than in other AOIs.
- Yellowstone is the furthest away from the other AOIs. This was expected, as very similar results are found in the co-correlation of tags (see *6.3.2 Tag Structures*) as well as in the other MDS results (see *Figure 6.45*).

As the sample size is relatively small, the question remains of how much influence single users have on the results. A circumstance that is likely closely related to the observation, that larger samples tend to be near each other, at least per AOI. This can be expected, as larger numbers of users often also mean more images and tags. This would then lead to a more diverse tag collection that is quite resistant towards singe outliers (unexpected tags in a user's images). To get a better understanding, a sensitivity analysis is performed, that might enable a verified discussion of the suggested interpretation.

6.4.2 Sensitivity Analysis

The sensitivity analysis is performed with multiple random subsamples of users per DHC and AOI. Out of all available users in $\text{subset}_{\text{analysed}}$ ten random samples of 10 users per DHC and AOI are drawn. For each of these samples, the MDS results based on the 157 *relevant* tags are aggregated, as it was done for the results discussed in *6.4.1 Mapping Cultures*. With the chosen sampling strategy it is possible, that some of the samples are very similar while others consist of totally different users. The results are visualised in *Figure 6.50*. It gives an overview of the results of the 10 sample runs. Each data point represents one DHC as a summary of all its users found in the sample. Each run generates a distinct data point for each DHC that is found within an AOI. Therefore, each DHC is visualised 10 times, once for each run. The closer those 10 points lie to each other, the more robustness the specific DHC and AOI combination shows.

The overall pattern of the five AOIs is still very dominant, the borders have dissolved slightly, especially in the case of Jungfrau-Aletsch and the Dolomites as well as Geirangerfjord and the Lake District.

As mentioned, the numbers of users within each DHC and AOI varies from 10 to 50. The observation that larger samples tend to stay closer together is still unanswered. To find an explanation, *Figure 6.50* is taken apart into a separate figure for each DHC (*Figure 6.51*).

On a DHC basis (*Figure 6.51*) the strength of the variation between the ten samples becomes visible. On one hand, there are cases where the single samples all cluster around the aggregated data point. These have a low variation. A small variation means the DHC is relatively robust in this specific AOI. On the other hand, there are many cases where the points are scattered over a larger area, which is a clear sign of greater variation, meaning the samples

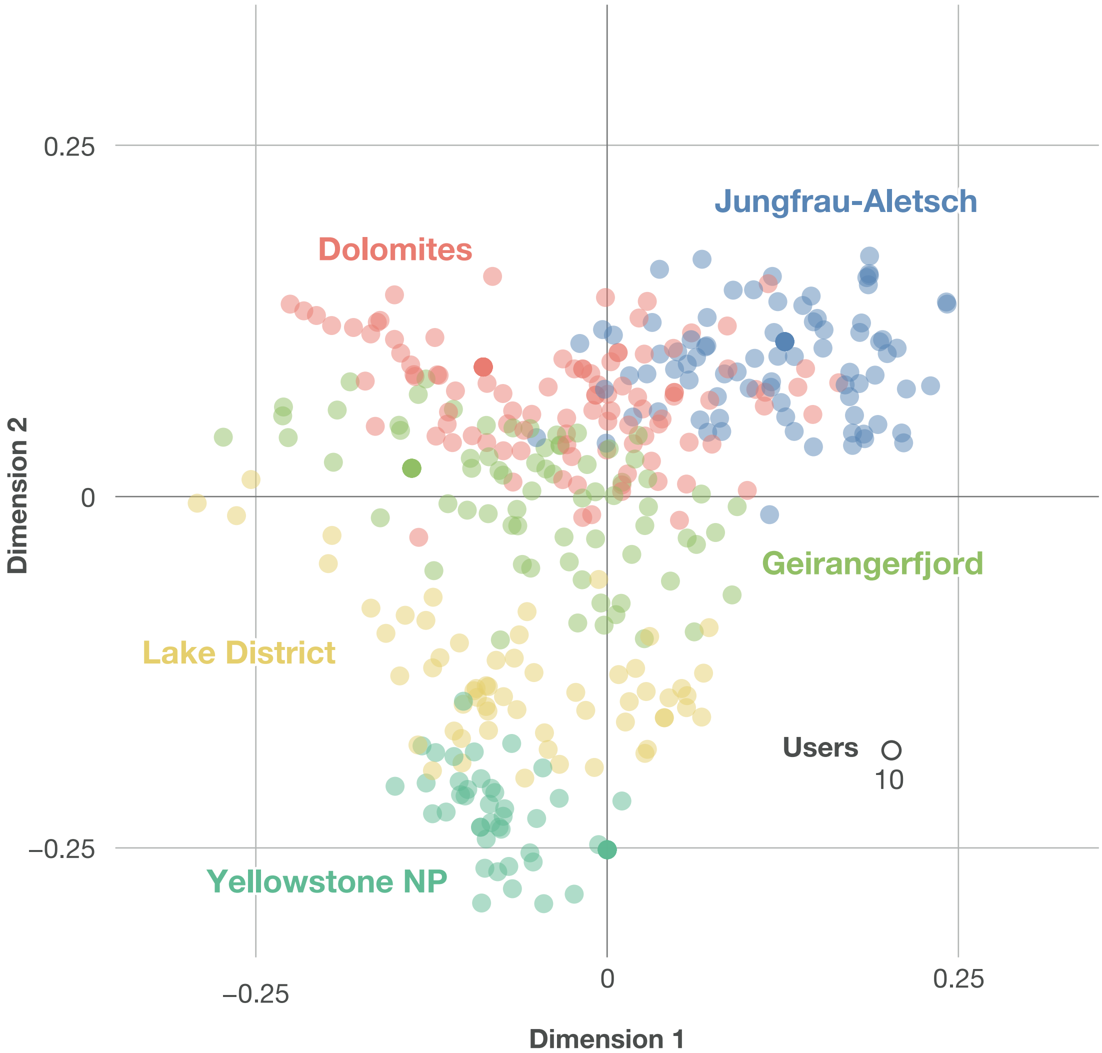


Figure 6.50: MDS based on the 157 relevant tags, grouped to the random draws of 10 users per DHC. Each data point thus corresponds to the averaged location of the collection of all the relevant tags found in the images of the 10 sampled users.

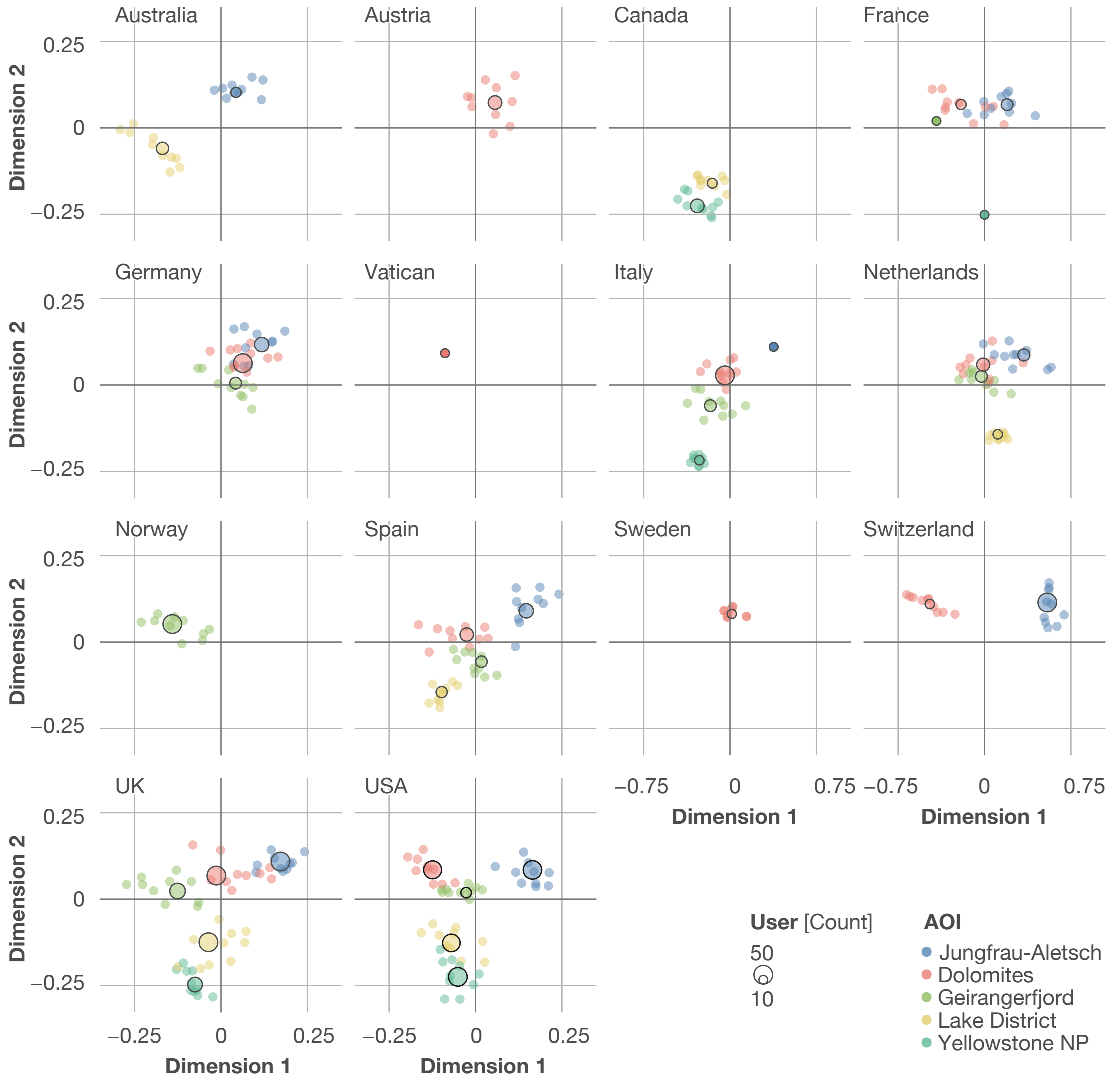


Figure 6.51: MDS based on the 157 relevant tags, grouped to the random draws of 10 users per DHC and facetted by DHC. Additionally, the DHC's position (black outline) as found in Figure 6.49 is indicated for reference.

are very sensitive to the selection of users. In other words, they are not very robust. But, there are exceptions as seen in the cases of Italy (Jungfrau-Aletsch), France (Geirangerfjord, Yellowstone NP), and Vatican City State (Dolomites). Here, the result of all ten sample runs is exactly the same. Clearly, this is not a coincidence. The answer is found in *Table 6.8* that gives an overview of all DHCs and their number of users, images, and (unique) tags by AOI. In all of the mentioned cases, there are exactly ten users available. Every random sample of ten users out of ten users will without naturally result in the same selection of users, which explains the non-existing variation. This problem, that is clearly bigger than the above discussed four special cases. The variations in robustness observed in *Figure 6.51* are – at least for many cases – caused by the available user sample size and not exactly the underlying similarities or differences in each DHC and AOI. Therefore, this sensitivity analysis is limited by not large enough samples for all combinations of DHC and AOI that would enable the creation of larger – in terms of user numbers – subsamples.

However, when comparing only those cases where close to 50 users per AOI are available (emphasized in *Table 6.8*), this limitation can be avoided or at least minimized. For these selected cases it means that if the available user size was the only determinant for variation, the results of the ten runs would yield very similar results regarding the robustness. For users located in the UK this does not hold entirely true, as the results for Jungfrau-Aletsch form a denser cluster than for example the Dolomites or the Lake District. While the larger number of tags – 1 184 are found in Jungfrau-Aletsch – might explain the larger variation for the Lake District (3 067 tags), this can certainly not be the case for the Dolomites (919 tags). At the same time, there is the case of the USA for which no differences are observable in the robustness of the MDS results. A clear pattern remains hidden.

Table 6.8: Overview of DHCs, user counts, image counts, tag counts and unique tag counts in subset$_{analysed}$ when reduced to the 157 relevant tags.

DHC	*AOI*	User Count	Image Count	Tag Count	Unique Tag Count
UK	*Jungfrau-Aletsch*	50	1 184	20 316	157
	Dolomites	50	919	15 854	157
	Lake District	50	3 067	47 547	157
	Geirangerfjord	34	366	5 858	156
	Yellowstone NP	30	595	6 155	150
USA	*Yellowstone NP*	50	3 077	37 705	157
	Jungfrau-Aletsch	49	932	18 401	157
	Dolomites	46	541	6 445	157
	Lake District	44	534	9 500	157
	Geirangerfjord	15	125	2 330	148
Germany	*Dolomites*	50	955	14 887	157
	Jungfrau-Aletsch	29	335	6 408	155
	Geirangerfjord	19	335	5 780	155
Italy	*Dolomites*	50	3 405	47 175	157
	Geirangerfjord	19	130	2 474	149
	Yellowstone NP	12	193	2 792	149
	Jungfrau-Aletsch	10	104	1 927	135

Spain	*Jungfrau-Aletsch*	29	203	3 734	152
	Dolomites	25	399	6 971	156
	Geirangerfjord	18	216	4 113	152
	Lake District	17	263	3 558	155
Netherlands	*Dolomites*	24	296	4 932	154
	Jungfrau-Aletsch	20	74	1 439	135
	Geirangerfjord	19	148	2 557	143
	Lake District	12	98	1 869	147
Switzerland	*Jungfrau-Aletsch*	50	2 052	37 525	157
	Dolomites	12	35	699	116
France	*Jungfrau-Aletsch*	19	356	7 011	155
	Dolomites	14	449	6 966	157
	Geirangerfjord	10	162	2 927	150
	Yellowstone NP	10	259	3 240	146
Norway	*Geirangerfjord*	50	1 031	14 036	157
Australia	*Lake District*	21	491	7 700	154
	Jungfrau-Aletsch	14	190	2 697	151
	Dolomites	9	103	1 841	152
Canada	*Yellowstone NP*	26	617	7 061	153
	Lake District	13	176	3 186	149
Austria	*Dolomites*	26	537	9 118	157
Sweden	*Dolomites*	11	337	5 149	150
Vatican	*Dolomites*	10	356	3 839	152

Chapter 7 | Discussion

The previous chapter introduced the results obtained based on the earlier outlined methodology. The objective of this chapter is, based on the theory introduced in *2 Theoretical Background* and results obtained and documented in 6 Results, to discuss the research questions as outlined in *3 Research Gaps and Objectives*. To reliably answer the research questions, it is necessary that the data supporting it is not or only negligibly biased by any of the pre-processing steps. Therefore, a prior discussion of the representativeness of the $\text{subsets}_{\text{analysed}}$ is as a basic requirement. In a next step the two research questions are individually answered based on the results. While RQ.1 can be fully answered and discussed, this is not entirely true for RQ.2. More data and further, more sophisticated analysis is needed to yield clearer answers.

7.1 Representativeness of Subsets

Various characteristics of the final datasets on which the further analysis is performed have been introduced, namely user count, image count, and images per user. By comparing *user and image counts* as well as their combination (image count per user) in $\text{subset}_{\text{analysed}}$ to the initial $\text{subset}_{\text{original}}$, its representativeness was examined. It is expected that with each processing step the number of users and images is reduced, a trend that is clearly shown in the results. Although not all AOIs show similar rates of reduction (see *Table 6.1*), in all cases the data is reduced by 50–90%. A clear tendency is seen that smaller AOIs show smaller reduction rates as the two limits (1) for maximum user per country and (2) image count per user (see *Figure 6.2*) are reached less often.

An interesting observation is made in the case of the *image count per user*. Users with 1 to 5 images are often underrepresented in the $\text{subset}_{\text{analysed}}$ while all other users – featuring more than five and up to 100 images – are systematically overrepresented. Due to the fact that the subsets feature many users in those two groups and the sampling strategy not considering the distribution of image counts per user, this distortion is still present. Nonetheless, it is expected, that the added bias is quite low in comparison to the underlying ones (e.g. the bias introduced by using Flickr data and all the prerequisites an image must have to be part of the initial dataset, YFCC100M). In terms of the *spatial characteristics*, it could be shown that for all AOIs the initial distribution could be quite well preserved. The NNI values remained mostly constant over all subsets, showing a negligible trend towards less clustered data towards the final subset$_{analysed}$. As the sampling strategy did not involve any spatial constraints, areas of low densities but maybe great interest (e.g. images along ring road in Yellowstone NP or train track in Jungfrau-Aletsch) have mostly vanished. A more in depth analysis would be needed to accurately assess the effects of this potentially introduced bias.

In conclusion, although a few (potential) biases were introduced with the creation of the various subsets, these subsets tend to represent the initial dataset (subset$_{original}$) to large parts.

7.2 Deriving Home Locations

RQ.1 **How and with what accuracy can the home location of Flickr users be derived?**
- How can Flickr user's home locations be derived?
- What is the accuracy of such a method?

Building on the assumption that a large number of images are taken within close proximity of a user's location, the single source of information used is the user's photostream available on Flickr.com. Random sampling of images with attached coordinates resulted in a collection of images for each user from which the main cluster was derived with DBScan. The median longitude and latitude of the images making up the largest cluster by image count is then defined as a user's derived home location (DHL). Reversed geocoding was used to generate place names for the found coordinate pair. The proposed approach to derive the home locations of Flickr users performed acceptably on a city level (~50% accuracy) and very well on a country level (~90% accuracy).

7.2.1 Interpretation of Results

The designed and implemented algorithm to derive users home locations based on their photostreams was successfully used to geo-reference around 4 000 users. Two methods, named as Median and DBScan, were compared. Interestingly, the home locations based on the Median method, although based on a highly simplified method, already showed a quite high agreement with the validation information (user location accessed from their online Flickr profiles). Yet, DBScan performed slightly better and therefore is the preferable method. When comparing the obtained accuracy on a city level (~50% accuracy) to the study results (80–90% accuracy) by Popescu & Grefenstette (2010), the featured solution is clearly inferior. However, there are two important differences: First, there is no prior creation of look up information necessary as it is the case in the approach proposed by Popescu & Grefenstette. The implemented method is applicable to any user, without any prior knowledge of the potential home locations if geo-referenced images for this user are available in their online photostream. Second, the here proposed method was specifically designed to derive a user's home country and not a location on a city-level, a degree of detail not specifically required by the introduced working definition of culture.

Validation
A sample of 150 users (30 randomly drawn from each of the five AOIs, with the sole prerequisite to feature a location in the users' profiles) was manually validated on city (~50%), country (~90%) and continent level (~100%). Accuracy does increase with each aggregation step and reaches nearly 100% for the continent level. While the higher accuracy on a continent level could be expected, as the spatial aggregation leaves more space for locational errors, the relatively high accuracy on a country and city level was unexpected.

7.2.2 Uncertainties and Limitations

In the process of deriving home locations/countries for the Flickr users found in $\text{subset}_{\text{original}}$, there are a few uncertainties and limitations involved relating to the applied methods as well as the findings resulting from their application.

Dependency on Online Flickr Profiles
Although the localisation does not need any pre-generated data, it depends on the existence of an online user profile. If the user profile is no longer available online, which is the case for 5% of the 4 992 users found in $\text{subset}_{\text{original}}$, the proposed method is not able to derive a home location.

Clustering Method
Regarding the clustering method used (DBScan), there are the following potential limitations: (a) selection of clustering algorithm, (b) the epsilon distance of 200 kilometres, and (c) the metric to select the home cluster. (a) DBScan seems to be an adequate, but most likely not the perfect solution. An algorithm that favours roundish clusters would be preferred, although it is not expected to substantially increase the already reached accuracy of 90% on a country level. (b) The epsilon distance, a threshold distance to decide up to which distance two points are matched to the same cluster, is crucial in the size and number of the resulting clusters. The chosen value of 200 kilometres is a trade-off in between small, focussed clusters and a large enough distance that for most users at least one cluster is found. The value of 200 kilometres was found in the process of optimizing the threshold with the goal to find a DHL

for as many users as possible, but still having reasonably sized clusters (maximum diameter of a few hundred Kilometres). A more detailed and structured experiment to derive the threshold value could be an interesting measure to gain a better understanding of the relationship of epsilon distance and DHL, as well as their resulting accuracy. (c) To find the cluster that contains the potential home location, a more elaborate metric than just the number of images could be used. A possible improvement would be an index of time span and/or regularity. For example, a metric that gives preference to longer time spans in which the images were taken as well as the number of events in the found time span could be used.

Conceptualisation of Home

The basic assumption that Flickr users will take most of their images in close proximity of their home is exactly what it is: an assumption. Although Hecht & Gergle (2010) have shown in a study based on Flickr users that 50% of them contribute local information, the question remains, what kind of content the other 50% upload. The validation shows that 50% of all user locations derived on a city level are correct. Maybe there is better ways of defining and deriving a user's home, approaches that take into account more than a random sample of geo-referenced images of a user's photostream. Additional information could for instance be the inclusion of the user locations found in the users' Flickr profiles (see Strauman, Çöltekin & Andrienko 2014) or image tags (see Kordopatis-Zilos, Papadopoulos & Kompatsiaris 2015).

7.2.3 Reflections

With respect to RQ.1, it can be concluded that based on a random sample of users' online photostream, their home locations can be derived. By clustering the sample and reverse-geocoding the cluster centre a location is found that can be used as an approximation of the user's home location. For around 75% of all users within the initial $\text{subset}_{\text{original}}$ (YFCC100M filtered to the five AOI and images only), a home location could be derived with the presented approach. The method's accuracy was measured on a city or country level. On a city level, 50% of the user-provided information within their profile and the derived home city were lying in very close proximity (less than 10 kilometres). On a country-level, the accuracy is around 90%. The achieved accuracy on a country level seems sufficient to use the derived home country (DHC) as a proxy for culture.

7.3 Cultural Differences in Landscape Preferences

RQ.2 **What, potentially measurable, differences can be found in the preferences of landscape elements between groups of people with similar cultural backgrounds?**

- To what extent can machine generated tags (in the case of Google Vision API) be used for the extraction of landscape elements?
- What are the differences between the tag collections on a derived home country level?

Having derived a home location, or more importantly a home country, as a proxy for each person's cultural background (see RQ.1), the elements in each of their images were extracted. This was achieved by using Vision API, an online machine learning service provided by Google. For each image 1–60 tags could be extracted. Due to the fact that the algorithm is trained on tags and training dataset that are kept secret, it would be false to speak of landscape elements, therefore the term tags is used. In order to have tags that are comparable over larger sets of DHCs and AOIs, their number was limited from 5 254 to a selection 157 tags, each of which is available in at least half of the all possible combinations of DHCs and AOIs. Nearly 60% of these tags could be matched with a landscape element class as defined by Conrad (2011), mostly to the category ‹natural landforms› (40% of all tags). Although co-correlations of terms are quite large for several thematic tag clusters, they varied with the AOIs. This finding strengthens the impression that tags are independent in the way they are extracted. This means that not every time ‹mountain› is labelled in an image, the closely related tags ‹ridge› and ‹summit› are extracted as well. Therefore, none of the 157 tags were excluded for further analysis as they all seem to add valuable and – at least not entirely – redundant information. Multi-dimensional scaling (MDS) was then used to reduce the 157 dimensions (tags) of 14 DHCs to two dimensions. As many of the intermediate results already showed, a comparison with AOI as a control variable is essential, as the overall trends are heavily influenced by the AOI the images were taken in (see *Figure 6.36*).

7.3.1 Interpretation of Results

It was shown that the Vision API generated tags that – at least in the context of the small number of samples presented – generated reasonable results. Using these tags as a proxy for landscape elements photographed, they were multi-dimensionally scaled (MDS) and aggregated to DHCs. In the following an overview is given on the visual validation of the results and the additionally performed sensitivity analysis.

Validation

The visual validation on the level of DHCs is based on clusters of DHCs which are quite distinctly separated between AOIs. Given the fact that the depicted MDS results always have the axis of largest variation as their horizontal component, it is important to mention that the same distance in a vertical direction, although equally long, is less strong. This said and combined with the fact that all data points occur within ± 0.25 units on either axis, the differences are not as large as they initially might seem. Nevertheless, DHCs in the Dolomites, the Jungfrau-Region as well as the Geirangerfjord AOIs lie relatively close together. The Lake District itself is found close to Geirangerfjord, most likely due to water as an important thematic focus (see *Figures 6.50* & *6.51* as well as cluster 4 in *Figure 6.36*). Yellowstone NP is found the furthest away from the centre, in a peripheral position. Although it would be exciting to now interpret the differences seen in single DHCs and their spatial relationships in the realm of the MDS, it would lead at best to fortunate coincidences. The reason for this cautious attitude towards these initial results lies in the findings of the performed sensitivity analysis.

Sensitivity Analysis

Due to the – after all the filtering procedures – small numbers of users per DHC, it is expected that single users can have a strong impact on the MDS results. Therefore, the robustness of the tag sets clustered by DHC was tested. For each of the ten runs another random sample of exactly 10 users per AOI and DHC was selected and its MDS results visualised (*Figure 6.51* and per DHC in *Figure 6.50*). The overall distribution by AOI remains relatively clear, with slightly more overlaps in the bordering areas. But what is now visible, especially in the individual visualisation for each DHC, is the circumstance that the per AOI results are not very robust. When interpreting the scatter plots, it is important to take into consideration the number of users found in each AOI. With small numbers of users to sample from, the ten samples will all end up being very similar, solely since – in the most extreme cases – up to 100% of the available users are found in each of the samples. This is the case for Italy (Jungfrau-Aletsch), France (Geirangerfjord, Yellowstone NP), and Vatican City State (Dolomites). Clearly, the comparison has to be made between the cases with many users per DHC and AOI. The pattern is mixed, but the tendency is clear: none of the DHCs seem very robust. Yet, and this is an important finding, the sensitivity observed – although introduced by the bias of single users – is most likely just highly visible due to the very small sample size (10 users per AOI and DHC). Therefore, a final assessment of the robustness or sensitivity of the final DHC and AOI is not possible. Still, there are trends found in the analysed dataset, such as for example the considerable larger proportion of tags related to transport infrastructure found in Asian countries. Of course, this might be simply owed to the fact that a large portion of their images was taken in the Jungfrau-Aletsch area. This is an area visited by a large proportion of Asian tourists for the widely known Jungfraujoch, which is mainly accessible by train. The overall assessment is that with larger samples of users per DHC and AOI, clearer results might be achievable.

7.3.2 Uncertainties and Limitations

As for the first research question, the uncertainties regarding RQ.2 are discussed and their handling is justified.

Assumption regarding landscape preferences

Most of RQ.2 builds on the bold assumption that people tend to photograph and document landscapes they prefer. Although the performed analysis did not result in a conclusive evaluation, due to sample sizes which are potentially too small, the assumption's validity needs to be discussed. Several studies (Gliozzo, Pettorelli & Haklay 2016; Seresinhe, Moat & Preis 2018; Strauman, Çöltekin & Andrienko) support the assumption that there is a connection in between landscape preferences and created photographs. However, what if this does not hold true for every location? The examined AOIs are all listed as UNESCO World Natural Heritage Sites and were explicitly chosen for this characteristic. As all five AOI have their very typical ‹vistas›, defined by popular imagery, people might mainly photograph exactly these traits of the landscape and not exactly what they would individually prefer in the first place.

Assumption regarding culture

The assumption that the country someone is living in can be used as the grouping variable for different cultures is equally daring. Nonetheless, the approach though is found in studies of various fields of research (Buijs, Elands & Langers 2009; Yu 1994). Further studies have used solely language as a proxy for culture (Majid et al. 2018). It remains unclear to what extent any of the two assumptions hold true, but it seems important to question such a simplifying understanding especially as current cultural research builds on a far more complex understanding of culture. Therefore, it might be inevitable to answer the question of cultural differences with a dataset that holds more information on the individual users than just their (derived) location. Another potential issue is the granularity of culture. It is a challenging task to balance the granularity between too wide and too narrow. If too wide, the potentially existing cultural differences will only be noticeable as noise; and if too narrow, the measured differences are inter-personal. Ultimately, it all lives or fails with the definition of culture and its appropriateness for the research questions in which the working definition thereof is used.

Landscape elements

From a theoretical background, one of the main visual cues on which different landscape preferences can be explained are landscape elements. As shown the clearest in *Figure 6.44*, there are many tags extracted by the Vision API that can be matched to a group of landscape elements. Yet, there are also 40% of tags (in the reduced set of 157 tags) that do not directly relate to a specific group. In these cases, the tags are still used, as most of them do describe a visual attribute (e.g. ‹house› or ‹national park›) that is helpful in getting more semantic understanding of a user's image and ultimately the tag set per DHC and AOI.

7.3.3 Reflections

In a first part, RQ.2 covers the quality of the tags extracted through the Vision API as well as their compatibility with landscape elements. It was shown that the tags tend to describe the image elements (which do not have to be landscape elements) quite well, at least for the few selected examples. Overall, the tag sets found for the images feature many unique tags. The subset of 157 tags was used to examine their distribution along the three levels of categorization: superordinate, basic, and subordinate level. It was shown that in comparison to the tags extracted through Places365 used by Seresinhe, Preis & Moat (2017), slightly more superordinate tags were found. These are descriptions that operate on a very general level (i.e. ‹coastal and oceanic landforms› or ‹wildlife›). A potential implication of this finding is that these broad tags prevent the detection of the – in case they do exist – subtle differences in between different cultures. A last restriction that comes with the usage of the Vision API is the circumstance that the complete tag set that could potentially be returned is unknown (see *5.6 Methodological Limitations*). Therefore, no statement can be made on the method's recall. The second part of RQ.2 focusses on the potential cultural differences in landscape preferences. For several reasons, no definitive conclusion can be drawn. There are trends visible in the data, but not enough to support a well-founded statement. There are a variety of potential reasons for this, sorted by strength of influence: (1) the amount of data used was too small, with larger sample sizes not only significant, but also relevant differences can be found. (2) The selected areas of interest, of which four are located in Europe and all of them are UNESCO World Natural Heritage Sites, are poorly chosen and with another sample there would have been much clearer results. (3) The chosen granularity of culture based on a country level is either too wide (better: sub regions), too narrow (better: aggregation to larger regions) or just the wrong unit (better: language regions). (4) Using a spatial proxy – in this case the location a user is assumed to live in – as a spatial proxy to culture is not sufficient. (5) The assumption that people photograph and upload to Flickr does not hold up or (6) limiting the perception of landscapes to visual cues masks larger parts of what could be understood as cultural differences. Several of these reasons would rule out Flickr as a potential data source. Many reasons have now been outlined for why no definitive conclusion can be drawn. It therefore seems that the ambiguity in results regarding cultural differences in landscape preferences described at Kaplan & Kaplan (1989) or (Hägerhäll 2018) remains.

Chapter 8 | Conclusion

This work set out to explore the usability of machine generated tags to examine cultural differences in landscape preferences. With the increasing availability of machine learning approaches for business and science, there is an ever-growing number of potential use cases. Image recognition is one of them and has been applied to the field of landscape perception research. The basic idea was to close the unveiled research gaps, namely (a) the usage of an automated process instead of the very commonly performed manually-coded content analysis; (b) the absence of large scale studies that compare broad samples of people from many different cultures with each other; (c) the contradicting results in the usage of landscape elements as a proxy to landscape preference; and (d) the limited knowledge in cultural differences in landscape preferences.

Insights

The main objective of this study was not simply to promote a solution that builds on machine learning, but to better understand the characteristics, limitations, and potential pitfalls of such an approach. To achieve this objective and as well address the identified research gaps, this study was two-fold. In a first step, a novel approach was developed to *geo-reference Flickr users* based on their online Flickr profile. For RQ.1, which focussed on the geolocation of Flickr users, a potential home location could be derived for around 75% of all users. A manual validation of the results shows an accuracy of 90% on a country-level. It can be concluded that the presented approach, in which the user's photostream is sampled for georeferenced images, worked well. Especially due to the advantage that even with relatively small samples of up to 30 images per user, the above reported accuracy was reached. Even on a city-level, an accuracy of 50% was achieved, although the method was specifically designed to yield useful results at a country level. In a second step, it was examined to what extent *cultural differences in landscape preferences* can be derived from Flickr imagery. Here, the prior extracted information on the users' home locations is used to group users by derived home countries (DHC). For five UNESCO World Natural Heritage Sites that served as areas of interest (AOI), all images found within the YFCC100M dataset were extracted and pre-processed. This included the creation of balanced samples regarding DHCs, AOIs, and user counts. For each image, its content was accessed through Vision API, which resulted in a tag sets for each image. The following analysis of these tag sets focused on two key research objectives: (1) To what extent can the Vision API be useful in the extraction of landscape elements from Flickr imagery? (2) What cultural differences can be found in the perception of these? Regarding the usefulness of the tags extracted by the Vision API, the general quality and quantity of returned tags is considered positive. On a sample of 157 *relevant* terms it could be shown that 60% of them can be unambiguously matched to known landscape element categories as found in literature. The results imply that the extracted tags have the potential to serve as a reliable source of on information on the landscape elements available in still imagery. The subsequent analysis of image tags in combination with the DHCs did not yield clear results. Although several trends could be found, such as the higher presence of transport infrastructure in images from users from Asian DHCs, no definitive conclusions can be drawn on any relevant differences in landscape perception. Besides a few other uncertainties, the relatively small sample sizes (50 users per AOI and DHC) are assumed to be the major limitation to the explanatory power of the found inter-country differences.

As an overarching conclusion, it remains crucial to question the origin of automatically extracted landscape elements, although they present themselves as promising. Moreover, due to the concealment of relevant information regarding the Vision API (i.e. training set and the complete set of potentially returned tags) on the part of Google, important questions remains open.

Future Work

Future work could apply the developed automatic procedures to other AOIs. This would generate additional data, which could be used to overcome the limitation of small user samples. More available users per DHC and AOI available could suffice to answer the still open question of cultural differences in landscape preferences. Methods including further image information (e.g. user-generated tags) could be used to extract the images that are taken within an AOI, instead of solely relying on the images coordinates. This would increase the number of images available for analysis. New data sources or previously neglected data sources that hold sufficient personal information to allow the incorporation of a more complex working definition of culture could increase the quality of the results. Secondly, it would be interesting to compare the results obtained with the Vision API to other object recognition applications as well as ‹ground truth› (i.e. by manually assigned tags). Comparing different applications could help identify their limitations and might give some idea of their recall. In general, it should be more deeply investigated how such automatically extracted tags differ from landscape elements as perceived and named by humans. Lastly, there is a general need of larger and more diverse samples in the research of cultural differences in landscape preferences. Diversity is not only present in the sense of cultures, but also senses (e.g. olfactory or auditory) and materials (e.g. landscape elements, descriptions, …). These factors can help to better understand how we perceive the landscape surrounding us. Such insights could prove valuable in future planning and nature management decisions.

Literature

Aaron, Jesse S. / Taylor, Aaron B. / Chew, Teng-Leong (2018): *Image co-localization. Co-occurrence versus correlation*. In: Journal of Cell Science 131(3), jcs211847.

Abdullah, Azizi / Veltkamp, Remco C. / Wiering, Marco A. (2009): *An Ensemble of Deep Support Vector Machines for Image Categorization*. In: 2009 International Conference of Soft Computing and Pattern Recognition. Malacca, Malaysia.

Adams, Ansel E. (1932): Retrospect. Nineteen-Thirty-One. In: Sierra Club Bulletin 17(1), pp. 1–11.

Adie, Bailey A. / Hall, C. Michael (2016): *Who visits World Heritage? A comparative analysis of three cultural sites*. In: Journal of Heritage Tourism, 12(1), pp. 67–80.

Alam, Cecep Nurul / Manaf, Khaerul / Atmadja, Aldy Rialdy / Aurum, Digital Khrisna (2016): *Implementation of haversine formula for counting event visitor in the radius based on Android application*. In: 4th International Conference on Cyber and IT Service Management. Bandung, Indonesia.

Antrop, Marc (2019): *A brief history of landscape research*. In: Howard, Peter / Thompson, Ian H. / Waterton, Emma (Editor) / Atha, Mick (Editor): The Routledge companion to landscape studies. Routledge, pp. 1–16.

Arnberger, Arne / Eder, Renate (2011): *Exploring the Heterogeneity of Rural Landscape Preferences. An Image-Based Latent Class Approach*. In: Landscape Research 36(1), pp. 19–40.

Appleton, Jay (1975): *The Experience of Landscape*. John Wiley & Sons.

Atha, Mick / Howard, Peter / Thompson, Ian / Waterton, Emma (2019): *Ways of knowing and being with landscapes. A beginning*. In: Howard, Peter / Thompson, Ian H. / Waterton, Emma (Editor) / Atha, Mick (Editor): The Routledge companion to landscape studies. Routledge, pp. xix–xxviii.

Begelman, Grigory / Keller, Philipp / Smadja, Frank (2006): *Automated Tag Clustering. Improving search and exploration in the tag space*. In: WWW2006, Edinburgh, UK.

Belyi, Alexander / Bojic, Iva / Sobolevsky, Stanislav / Sitko, Izabela / Hawelka, Bartosz / Rudikova, Lada / Kurbatski, Alexander / Ratti, Carlo (2016): *Global multi-layer network* *of human mobility*. In: Internation Journal of Geographical Information Science 7, pp. 1381–1402.

Bourassa, Steven (1991) *The Aesthetics of Landscape*. Belhaven Press.

Brook, Isis (2019): *Aesthetic appreciation of landscape*. In: Howard, Peter / Thompson, Ian H. / Waterton, Emma (Editor) / Atha, Mick (Editor): *The Routledge companion to landscape studies*. Routledge, pp. 39–50.

Bubalo, Martina / van Zanten, Boris T. / Verburg, Peter H. (2019): *Crowdsourcing geo-information on landscape perception and preferences. A review*. In: Landscape and Urban Planning 184, pp. 101–111.

Buijs, Arjen E. / Elands, Birgit H.M. / Langers, Fransje (2009): *No wilderness for immigrants. Cultural differences in images of nature and landscape preferences*. In: Landscape and Urban Planning 91, pp. 113–123.

Buller, David J. (2009): *Four fallacies of pop evolutionary psychology*. Scientific American (January).

Casalegno, Stefano / Inger, Richard / DeSilvey, Caitlin / Gaston, Kevin J. (2013): *Spatial Covariance between Aesthetic Value & Other Ecosystem Services*. In: PLoS ONE 8(6): e68437.

Chaovalit, Pimwadee / Zhou, Lina (2005): Movie review mining: A comparison between supervised and unsupervised classification approaches. In: Proceedings of the 38th Hawaii International Conference on System Sciences. Hawaii, USA.

Chen, Yan / Parkins, John R. / Sherren, Kate (2018): *Using geo-tagged Instagram posts to reveal landscape values around current and proposed hydroelectric dams and their reservoirs*. In: Landscape and Urban Planning 170, pp. 283–292.

Choi, Jaeyoung / Thomee, Bart / Larson, Martha (2017): *Practical Guide to Using the YFCC100M and MMCOMMONS on a Budget*. In: Newsletter ACM SIGMM Records 9(2), Article #5.

Clark, Philip J. / Evans, Francis C. (1954): *Distance to Nearest Neighbor as a Measure of Spatial Relationships in Populations*. In: Ecology 35(4), pp. 445–453.

Constantinides, Efthymios / Fountain, Stefan J. (2008): *Web 2.0. Conceptual foundations and marketing issues*. In: Journal of Direct, Data and Digital Marketing Practice 9, pp. 231–244.

Council of Europe (2000): *European Landscape Convention*. In: http://conventions.coe.int/Treaty/en/Treaties/Html/176.htm [16.04.19].

Crockford, Douglas (2017): *The JavaScript Object Notation (JSON) Data Interchange Format*. In: https://tools.ietf.org/pdf/rfc8259.pdf [03.04.20].

Da Rugna, Jérôme / Chareyron, Gael / Branchet, Berengere (2012): *Tourist behavior analysis through geotagged photographies: a method to identify the country of origin*. In:

IEEE 13th International Symposium on Computational Intelligence and Informatics (CINTI). Budapest, Romania.

Deng, Dong-Po / Chuang, Tyng-Ruey / Lemmens, Rob (2009): *Conceptualization of Place via Spatial Clustering and Co–occurrence Analysis*. In: ACM LBSN'09, Seattle WA, USA.

Depellegrin, Daniel / Blažauskas, Nerijus / Vigl, Lukas E. (2012): *Aesthetic value characterization of landscapes in coastal zones*. In: IEEE/OES Baltic International Symposium (BALTIC). Klaipeda, Lithuania.

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

Dutcher, Daniel D. / Finley, James C. / Luloff, A.E. / Johnson, Janet B. (2007): *Connectivity with nature as a measure of environmental values*. In: Environment and Behavior 39(4), pp. 474–493.

Encyclopaedia Britannica (2020): *Cirque. Geology*. In: https://www.britannica.com/science/cirque [20.02.20].

Falk, John. H. / Balling, John. D. (1982): *Development of Visual Preference for Natural Environments*. In: Environment and Behaviour 14(1), pp. 5–28.

Falk, John. H. / Balling, John D. (2010): *Evolutionary Influence on Human Landscape Preference*. In: Environment and Behaviour 42(4), pp. 479–493.

Fenton, Mark D. (1985): *Dimensions of meaning in the perception of natural settings and their relationship to aesthetic response*. In: Nasar, Jack L. [Ed.] (1988): *Environmental Aesthetics. Theory, Research, and Application*. Cambridge University Press.

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

Flickr (n.d. a): *API Documentation*. In: https://www.flickr.com/services/api/ [24.11.19].

Flickr (n.d. b): *Work at Flickr*. In: https://www.flickr.com/jobs/ [03.04.20].

GeoNames (2019): *Readme for GeoNames Gazetteer extract files*. In: http://download.geonames.org/export/dump/readme.txt [22.11.19].

Ghojogh, Benyamin / Crowley, Mark (2019): *The Theory behind overfitting, Cross Validation, Regularization, Bagging, and Boosting. Tutorial*. In: arXiv e-prints, 1905.12787.

Girardin, Fabien / Calabrese, Francesco / Dal Fiore, Filippo / Ratti, Carlo / Blat, Josep (2008): *Digital Footprinting. Uncovering Tourists with User-Generated Content*. In: IEEE Pervasive Computing 7(4), pp. 36–43.

Gliozzo, Gianfranco / Pettorelli, Nathalie / Haklay, Mordecahi M. (2016): *Using crowdsourced imagery to detect cultural ecosystem services. A case study in South Wales, UK*. In: Ecology and Society 21(3), 6.

Goodenough, Ward (1957): *Cultural Anthropology and Linguistics*. In: Garvin, P.L. (Editor): Report on the 7th Annual Round Table Meeting in Linguistics and Language Study, pp. 167–173.

Google Cloud (n.d.): *Vision AI*. In: https://cloud.google.com/vision/ [01.04.20].

Google Cloud (2017): *Release Notes. Label detection model update*. In: https://cloud.google.com/vision/docs/release-notes [28.03.20].

Google Cloud (2019): *Cloud Vision API Documentation. Label detection*. In: https://cloud.google.com/vision/docs/labels [24.11.19].

Google Cloud (2020): *Release Notes. Cloud Vision API will not return gendered labels such as 'man' and 'woman' after February 19, 2020*. In: https://cloud.google.com/vision/docs/release-notes [28.03.20].

Google Cloud Platform Support (2019): *In Trial Customer Inquiry #21375517*. Issued by the author.

Han, Ke-Tsung (2007): *Responses to six major terrestrial biomes in terms of scenic beauty, preference, and restoratives*. In: Environment and Behaviour 39(4), pp. 529–556.

Hausmann, Anna / Toivonen, Tuuli / Slotow / rob / Tenkanen, Henrikki / Moilanen, Atte / Heikinheimo, Vuokko / Di Mini, Enrico (2018): *Social Media Data Can Be Used to Understand Tourists' Preferences for Nature-Based Experiences in Protected Areas*. In: Conservation Letters, 11(1), pp. 1–10.

Hecht, Brent / Gergle, Darren (2010): *On the «Localness» of User-Generated Content*. In: CSCW 2010, February 6–10, 2010, Savannah GA, USA.

Heft, Harry / Nasar, Jack L. (2000): *Evaluating environmental scenes using dynamic versus static displays*. In: Environment and Behavior, 32(3), pp. 301–322.

Herzog, Thomas R. / Herbert, Eugene J. / Kaplan, Rachel / Crooks, C.L. (2000): *Cultural and Developmental Comparisons of Landscape Perceptions and Preferences*. In: *Environment and Behavior* 32(3), pp. 323–346.

Howard, Peter (2019): *Perceptual Lenses*. In: Howard, Peter / Thompson, Ian H. / Waterton, Emma (Editor) / Atha, Mick (Editor): *The Routledge companion to landscape studies*. Routledge, pp. 51–61.

Ignite Social Media (2012): *2012 Social Network Analysis Report – Demographic – Geographic and Search Data Revealed*. In: https://www.ignitesocialmedia.com/social-media-stats/2012-social-network-analysis-report/ [18.04.20].

Issue Tracker (2018): *Vision API topicality and score properties always contain the same value*. In: https://issuetracker.google.com/issues/117855698 [15.03.20].

Jongman, Rob H.G. (2002): *Landscape Planning for Biological Diversity*. In: Landscape Planning 27(2), pp. 187–195.

Jungfrau Region Tourismus AG (2017): *Geschäftsbericht*. In: https://bro.jungfrauregion.info/jrt/Geschaeftsbericht/17/mobile/ [09.03.20].

Kalkowski, Sebastian / Schulze, Christian / Dengel, Andreas / Borth, Damian (2015): *Real-time Analysis and Visualization of the YFCC100M Dataset*. In: ACM Multimeda Community-Organized Multimodal Mining: Opportunities for Novel Solutions (MMCOMMONS) Workshop. Brisbane, Australia.

Kaplan, Stephen (1988): *Perception and Landscape. Conceptions and Misconceptions*. In: Nasar, Jack L. [Ed.] (1988): Environmental Aesthetics. Theory, Research, and Application. Cambridge University Press, pp. 45–55.

Kaplan, Rachel / Herbert, Eugene J. (1988): *Familiarity and Preference. A cross-cultural Analysis*. In: Nasar, Jack L. [Ed.] (1988): *Environmental Aesthetics. Theory, Research, and Application*. Cambridge University Press, pp. 379–389.

Kaplan, Rachel / Kaplan, Stephen (1989): *The experience of Nature. A psychological Perspective*. University of Cambridge.

Kohsaka, Ryo / Flitner, Michael (2004): *Exploring forest aesthetics using forestry photo contests: Case studies examining Japanese and German public preferences*. In: Forest Policy and Economics, 6(3–4), pp. 289–299.

Kordopatis-Zilos, Giorgos / Papadopoulos, Symeon / Kompatsiaris, Yiannis (2015): *Geotagging Social Media Content with a Refined Language Modelling Approach*. In: Chau M. et al. (Editor): *PAISI 2015, LNCS 9074*. Springer, pp. 21–40.

Krell, Mario M. / Li, Yifan (n.d.): *MMCS*. In: http://search.mmcommons.org/ [05.02.20].

Kruskal, Joseph B. (1964): *Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis*. In: Psychometrika 29(1), pp. 1–27.

Leydesdorff, Loet / Vaughan, Liwen (2006): *Co-occurrence matrices and their applications in information science. Extending ACA to the Web environment*. In: Journal of the American Society for Information Science and Technology 57(12), pp. 1616–1628.

López-Guzmán, Tomás / Naranjo Tores, Mónica / Pérez Gálvez, Jesús C. / Carvache Franco, Wilmer (2017): *Segmentation and motivation of foreign tourists in world heritage sites. A case study, Quite (Ecuador)*. In: Current Issues in Tourism 22, pp. 1170–1189.

Lothian, Andrew (2017): *The Science of Scenery. How we see Scenic Beauty, what it is, why we love it, and how to measure and map it*. CreateSpace Independent Publishing Platform.

Majid, Asifa / Burenhult, Niclas / Stensmyr, Marcus / Valk, Josje de / Hansson, Bill S. (2018): *Olfactory Language and Abstraction across Cultures*. In: Philosophical Transactions B. 373(1752): 20170139.

Mark, David M. / Turk, Andrew G. / Burenhult, Niclas / Stea, David (2011): *Landscape in Language. An Introduction*. In: Mark, David M. / Turk, Andrew G. / Burenhult, Niclas / Stea, David [Ed.] (2011): *Landscape in Language. Transdiciplinary perspectives*. John Benjamins Publishing, pp. 1–24.

Maruti TechLabs (n.d.): *What is the Working of Image Recognition and How it is Used?* In: https://marutitech.com/working-image-recognition/ [04.04.20].

Menatti, Laura, & Casado da Rocha, Antonio (2016): *Landscape and health. Connecting psychology, aesthetics and philosophy through the concept of affordance*. In: Frontiers in Psychology 7, 571.

Menatti, Laura / Subiza-Pérez, Mikel / Villalpando-Flores, Arturo / Vozmediano, Laura (2019): *Place attachment and identification as predictors of expected landscape restorativeness*. In: Journal of Environmental Psychology 63, pp. 36–43.

Merriam Webster (2019): *Culture. Definition of culture*. In: https://www.merriam-webster.com/dictionary/culture [01.11.2019].

Mooney, Peter / Olteanu-Raimond, Anna-Maria / Juul, Niels C. / Touya, Guillaume / Alvanides, Seraphim / Kerle, Norman (2017): *Considerations of Privacy, Ethics and Legal Issues in Volunteered Geographic Information*. In: Foody, Giles / See, Lina / Fritz, Steffen / Mooney, Peter / Olteanu-Raimond, Ana-Maria / Fonte, Cidália C. (eds.): Mapping and the Citizen Sensor. Ubiquity Press, pp. 119–135.

Morris, Alix (2019): *The Decades-Long Effort to Protect the World's Largest Sheep*. In: https://www.smithsonianmag.com/science-nature/decades-long-effort-protect-worlds-largest-sheep-180972029/ [25.04.20].

Mousselly-Sergieh, Hatem / Egyed-Zsigmond, Elod / Gianini, Gabriele / Döller, Mario / Kosch, Harald (2013): *Tag Similarity in Folksonomies*. INFORSID 2013, May 2013. Paris, France. pp. 319–334.

Nasar, Jack L. (1988): *Editor's introduction*. In: Nasar, Jack L. (Editor): *Environmental Aesthetics. Theory, Research and Application*. Cambridge University Press, pp. 323–326.

Naqshbandi, Zoya / Fayaz, Sameer / Bhat, M.S. (2016): *A quantitative analysis of spatial organization of the urban centers in Kashmir Valley. A geographic information systems-based study using primacy index, rank-size rule, and nearest neighbor index*. In: Journal of Experimental Sciences 7, pp. 11–19.

Nguyen, Hanh T.H. / Wistuba, Martin / Schmidt-Thieme, Lars (2017): *Personalized Tag Recommendation for Images Using Deep Transfer Learning*. In: Ceci M., Hollmén J., Todorovski L., Vens C., Džeroski S. (Editors): *Machine Learning and Knowledge Discovery in Databases*. ECML PKDD 2017. Lecture Notes in Computer Science Vol. 10535. Springer, pp. 705–720.

Oteros-Rozas, Elisa / Martín-López, Berta / Fagerholm, Nora / Bieling, Claudia / Plieninger, Tobias (2018): *Using social media photos to explore the relation between cultural ecosystem services and landscape features across five European sites*. In: Ecological Indicators 94, pp. 74–86.

Pan, Steve / Lee, Jinsoo / Tsai, Henry (2014): *Travel photos. Motivations, image dimensions, and affective qualities of places*. In: Tourism Management, 40, pp. 59–69.

Penman, Richard (2018): *reverse_geocode source code*. In: https://bitbucket.org/richardpenman/reverse_geocode/src/default/__init__.py [22.11.19].

Perficient Digital Agency (2019): *Who Has the Best Image Recognition Engine?* In: https://www.perficientdigital.com/insights/our-research/image-recognition-accuracy-study [24.03.20].

Popescu, Adrian / Grefenstette, Gregory (2010): *Mining User Home Location and Gender from Flickr Tags*. In: Proceedings of the Fourth International AAAI Converence on Weblogs and Social Media. Washington D.C., USA.

Purcell, Terry / Peron, Erminielda / Berto, Rita (2001): *Why do preferences differ between scene types?* In: Environment and Behavior 33(1), pp. 93–106.

Qureshi, Salman / Breuste, Jürgen H. / Jim, C.Y. (2013): *Differential community and the perception of urban green spaces and their contents in the megacity of Karachi, Pakistan*. In: Urban Ecosystem 16, pp. 853–870.

Racicot, Aaron (2019): *Bboxfinder.com*. In: http://bboxfinder.com/ [21.05.19].

Russell, James A. (1988): *Affective Appraisals of Environments*. In: Nasar, Jack L. (Editor): Environmental Aesthetics. Theory, Research and Application. Cambridge University Press, pp. 120–129.

Sarlöv Herlin, Ingrid (2004): *New Challenges in the Field of Spatial Planning. Landscapes*. In: Landscape Research 29(4), pp. 399–411

Schedl, Markus (2019): *Deep Learning in Music Recommendation Systems*. In: Frontiers in Applied Mathematics and Statistics 5, pp. 44.

Schnädelbach Herbert (2004): *Zum Begriff der Moderne*. In: Schreiber, Hans-Peter (Editor): Biomedizin und Ethik. Birkhäuser, pp. 67–71.

Seresinhe, Chanuki I. / Preis, Tobias / Moat, Helen S. (2017): *Using deep learning to quantify the beauty of outdoor places*. In: Royal Society Open Science 4, 170170.

Seresinhe, Chanuki I. / Moat, Helen S. / Preis, Tobias (2018): *Quantifying scenic areas using crowdsourced data*. In: Environment and Planning B. Urban Analytics and City Science 45(3), pp. 567–582.

Shamma, David A. (2014): *One Hundred Million Creative Commons Flickr Images for Research*. In: https://yahooresearch.tumblr.com/post/89783581601/one-hundred-million-creative-commons-flickr-images [15.04.2019].

Shepitsen, Andriy / Gemmel, Jonathan / Mobasher, Bamshad / Burke, Robin (2008): *Personalized Recommendation in Social Tagging Systems Using Hierarchical Clustering*. In: RecSys'08. Lausanne, Switzerland.

Strauman, Ralph K. / Çöltekin, Arzu / Andrienko, Gennady (2014): *Towards (Re)Constructing Narratives from Georeferenced Photographs through Visual Analytics*. In: The Cartographic Journal, 51(2), pp. 152–165.

Tenerelli, Patrizia / Demšar, Urška / Luque, Sandra (2016): *Crowdsourcing indicators for cultural ecosystem services. A geographically weighted approach for mountain landscapes*. In: Ecological Indicators 64, pp. 237–248.

Tenerelli, Patrizia / Püffel, Catharina / Luque, Sandra (2017): *Spatial assessment of aesthetic services in a complex mountain region. Combining visual landscape properties with crowdsourced geographic information*. In: Landscape Ecology 23, pp. 1097–1115.

Thomee , Bart / Shamma, David A. / Friedland, Gerald / Elizalde, Benjamin / Ni, Karl / Poland, Douglas / Borth, Damian / Li, Li-Jia (2016): *YFCC100M: The New Data in Multimedia Research*. In: Communications of the ACM, 59(2), pp. 64–73.

Thompson, Catharina W. (2019): *Landscape perception and environmental psychology*. In: Howard, Peter / Thompson, Ian H. / Waterton, Emma (Editor) / Atha, Mick (Editor): The Routledge companion to landscape studies. Routledge, pp. 19–38.

Tieskens, Koen F. / Van Zanten, Boris T. / Schulp, Catharina J.E. / Verburg, Peter H. (2018): *Aesthetic appreciation of the cultural landscape through social media. An analysis of revealed preference in the Dutch river landscape*. In: Landscape and Urban Planning 177, pp. 128–137.

Tribot, Anne-Sophie / Deter, Julie / Mouquet, Nicolas (2018): *Integrating the aesthetic value of landscapes and biological diversity*. In: Proc. R. Soc. B 285, 20180971.

Tversky, Barbara / Hemenway, Kathleen (1983): *Categories of Environmental Scenes*. In: Cognitive Psychology 15(1), pp. 121–149.

Tylor, Edward B. (1871): *Primitive Culture. Research into the Development of Mythology, Philosophy, Religion, Art, and Custom. Murray, London*. In: https://reader.digitale-sammlungen.de/de/fs1/object/display/bsb11006440_00013.html [04.04.20].

Ulrich, Roger S. (1977): *Visual Landscape Preference. A Model and Application*. In: Man-Environment Systems 7(5), pp. 279–293.

Ulrich, Roger S. (1983): *Aesthetic and Affective Response to Natural Environment*. In: Altman, Irwin / Wohlwill, Joachim (Editor): *Human Behaviour and Environment*, Vol. 6: *Behaviour and Natural Environment*. Plenum, pp. 85–125.

UNESCO World Heritage Centre (2019): *World Heritage List*. In: http://whc.unesco.org/en/list/ [23.11.19].

[UNEP] United Nations Environment Programme (2017): *World Heritage Sites KMZ File*. In: https://www.unep-wcmc.org/resources-and-data/kml-file-of-world-heritage-sites [21.05.19].

[UZH] Universität Zürich (2018): *T-Test für abhängige Stichproben*. In: https://www.methodenberatung.uzh.ch/de/datenanalyse_spss/unterschiede/zentral/ttestabh.html [25.04.20].

Velarde, Maria D. / Fry, Gary / Tveit, Mari S. (2007): *Health effects of viewing landscapes. Landscape types in environmental psychology*. In: Urban Forestry & Urban Greening 6(4), pp. 199–212.

Wunsch-Vincent, Sacha / Vickery, Graham (2007): *Participative Web. User-Created Content*. Technical Report #JT03225396. Organisation for Economic Co-operation and

Development (OECD) Directorate for Science, Technology, and Industry; Working Party on the Information Economy.

Yahoo! Webscope (n.d.): *Image Data*. In: https://webscope.sandbox.yahoo.com/catalog.php?datatype=i [26.11.19].

Ye, Qiang / Zhang, Ziqiong / Law, Rob (2009): *Sentiment classification of online reviews to travel destinations by supervisedmachine learning approaches*. In: Expert Systems with Applications 36, pp. 6527–6535.

Yu, Kongjian (1995): *Cultural variations in landscape preferences. Comparisons among Chinese sub-groups and Western design experts*. In: Landscape and Urban Planning 32, pp. 107–126.

Zhang, Haipeng / Korayem, Mohammed / You, Erkang / Crandall, David J. (2012): *Beyond Co-occurrence. Discovering and Visualizing Tag Relationships from Geo-spatial and Temporal Similarities*. In: WSDM'12. Seattle WA, USA.

Zhou, Bolei / Lapedriza, Agata / Khosla, Aditya / Oliva, Aude / Torralba, Antonio (2017): *Places. A 10 million Image Database for Scene Recognition*. In: IEEE Transactions on Pattern Analysis and Machine Intelligence. IEEE Transactions on Pattern Analysis and Machine Intelligence 40(6), pp. 1452–1464.

Zierhofer, Wolfgang (2011): *Natur und Kultur als Konstruktion*. In: Gebhardt, Hans / Glaser, Rüdiger / Radtke, Ulrich / Reuber, Paul (Editor): *Geographie*. Spektrum, pp. 1080–1081.

Table A.1: Tag level classification for the features found in Seresinhe (2017, Figure 2) and the 157 relevant tags found in this study.

APPENDIX

A.1 Additional Tables

	Features found in Figure 2, Seresinhe 2017)	157 relevant tags in $subset_{analysed}$
superordinate level	vegetation	water feature, water resources, natural environment, wildlife, grass family, nonbuilding structure, vegetation, natural landscape, nature, infrastructure, biome, geology, coastal and oceanic landforms, body of water, mode of transport, art, furniture, tourist attraction, architecture, property, real estate, vehicle, formation, atmosphere, tourism, plant, meteorological phenomenon, terrain, adventure, recreation, geological phenomenon, atmospheric phenomenon, landscape, wilderness, leisure
basic-level	valley, coast, tundra, creek, mountain, cliff, river, glacier, waterfall, volcano, castle, church, ruin, pond, pasture, beach, yellow, cottage, hayfield, lagoon, saturation, boathouse, islet, butte, snowfield, viaduct, lighthouse, camping, raft, tower, rugged, harbor, aqueduct, boardwalk, orchard, snow, pier, lawn, swimming, badlands, marsh, blue, orange, trees, mansion, foliage, green, gas station, athletic field, amusement park, fire station, wind farm, hospital, excavation, playground, highway, kennel, campus, motel, racecourse, raceway, runway, hangar, greenhouse, street, crosswalk, junkyard, house, bridge, yard, slum, driveway, shed, park, oilrig, barn, airfield, pavilion, red, white, schoolhouse, synagogue, brown, grass, inn, farm, kasbah, clouds, warmth, grey	spring, watercourse, water, stream, sign, signage, rainforest, meadow, thouroughfare, road, forest, ravine, travel, vacation, loch, waterway, sound, sea, river, reflection, reservoir, bank, lake, mist, plateau, jungle, green, font, transport, tundra, plain, grass, grassland, pasture, light, night, style, monochrome, room, ice, moraine, massif, summit, slope, cirque, roof, home, building, house, window, facade, apartment, fjord, neighbourhood, town, city, cottage, estate, village, bedrock, outcrop, klippe, wadi, cliff, evening, wall, branch, leaf, metal, photography, rock, escarpment, tree, morning, sunlight, wood, fell, valley, ridge, world, flower, cloud, sky, hill, mountain, panorama, highland, horizon
subordinate level	mountain snowy, lake natural, mountain path, rock arch, Japanese garden, ice shelf, ski slope, formal garden, canal natural, forest path, desert road, forest broadleaf, ice floe, tree farm, desert sand, wheat field, desert vegetation, botanical garden, field road, topiary garden, forest road, golf course, rope bridge, field cultivated, picnic area, oast house, running water, construction site, parking lot, industrial area, roof garden, manufactured home, parking garage, water tower, general store, loading dock, apartment building, hunting lodge, railroad track, residential neighbourhood, bus station, dirt soil, colour variation, phone booth, natural light, corn field, open area, rice paddy	mountain river, state park, land lot, plant community, valdivian temp. rain forest, nature reserve, road surface, temp. coniferous forest, trop. and subtrop. coniferous forests, old-growth forest, temp. broadleaf and mixed forest, national park, rural area, stock photography, monochrome photography, black-and-white, glacial landform, glacial lake, residential area, cumulus, hill station, mountain pass, mount scenery, mountain range, flowering plant, wildflower, mountain village
unclear	natural, brightness, no horizon, driving, hiking	tarn, Lake District, Alps, arête

Table A.2: Classification of the 157 relevant tags according to the landscape element classes as defined by Conrad (2011).

Landscape Element Class	Tag
(A) Rural characteristics	nonbuilding structure, sign, signage, meadow, road, grassland, pasture, rural area
(B) Natural landforms	water feature, water resources, natural environment, grass family, vegetation, body of water, plant, terrain, spring, watercourse, water, stream, rainforest, forest, ravine, loch, waterway, sound, sea, river, bank, lake, plateau, jungle, tundra, plain, grass, grassland, ice, moraine, massif, summit, slope, cirque, fjord, bedrock, outcrop, klippe, wadi, cliff, wall, branch, leaf, rock, tree, wood, fell, valley, ridge, flower, hill, mountain, highland, mountain river, escarpment, plant community, valdivian temp. rain forest, temp. coniferous forest, trop. and subtrop. coniferous forests, old-growth forest, temp. broadleaf and mixed forest, glacial landform, glacial lake, mountain range, flowering plant, wildflower, arête
(C) Cultural features	—
(D) Specific locations	mount scenery, lake district, alps, tarn
(E) Intangible aspects	reflection, light, cumulus, mist, cloud, sky
(F) Visual aesthetic qualities	natural landscape, nature, coastal and oceanic landforms, atmosphere, meteorological phenomenon, geological phenomenon, atmospheric phenomenon, wilderness, panorama, horizon,
(–) No landscape element	wildlife, biome, geology, mode of transport, art, tourist attraction, property, real estate, vehicle, formation, tourism, adventure, recreation, landscape, leisure, thouroughfare, infrastructure, travel, vacation, reservoir, green, font, transport, night, style, monochrome, room, roof, home, building, house, window, facade, apartment, neighbourhood, town, city, cottage, estate, village, evening, metal, photography, morning, sunlight, world, state park, land lot, nature reserve, road surface, national park, stock photography, monochrome photography, black-and-white, residential area, hill station, mountain pass, mountain village, furniture, architecture

A.2 Metadata

A.2.1 Software/Scripts

R Programming Language and Environment

R	3.4.1	www.r-project.org/
R Studio	1.0.153	rstudio.com/

R Packages	*Version*	*URL*
sp	1.2-7	github.com/edzer/sp/
spatstat	1.52-1	www.spatstat.org
rpart	4.1-11	—
nlme	3.1-131	—
cowplot	0.9.2	github.com/wilkelab/cowplot
ggmap	3.0.0	github.com/dkahle/ggmap

viridis	0.4.0	github.com/sjmgarnier/viridis
viridisLite	0.3.0	github.com/sjmgarnier/viridisLite
RColorBrewer	1.1-2	—
slam	0.1-45	—
tm	0.7-6	tm.r-forge.r-project.org
NLP	0.2-0	—
geosphere	1.5-7	—
leaflet.extras	1.0.0	github.com/bhaskarvk/leaflet.extras
leaflet	2.0.0	rstudio.github.io/leaflet
scales	0.5.0.9000	github.com/hadley/scales
lubridate	1.7.4	lubridate.tidyverse.org
data.table	1.11.4	r-datatable.com
forcats	0.4.0	forcats.tidyverse.org
stringr	1.4.0	stringr.tidyverse.org
dplyr	0.8.3	dplyr.tidyverse.org
purrr	0.3.3	purrr.tidyverse.org
readr	1.3.1	readr.tidyverse.org
tidyr	1.0.0	tidyr.tidyverse.org
tibble	2.1.3	tibble.tidyverse.org
ggplot2	3.2.1	ggplot2.tidyverse.org
tidyverse	1.3.0	tidyverse.tidyverse.org
plyr	1.8.4	github.com/hadley/plyr
countrycode	1.1.0	github.com/vincentarelbundock/countrycode

Python Programming Language and Environment		
Python	3.6.8	python.org/
PyCharm	2016.2.3	www.jetbrains.com/pycharm/

External Python Libraries	*Version*	*URL*
pandas	0.24.2	pandas.pydata.org/
numpy	1.17.2	numpy.org/
urllib	1.24.2	docs.python.org/3.6/library/urllib.html
reverse_geocode	1.4	pypi.org/project/reverse-geocode/
scikit-learn	0.21.2	scikit-learn.org/stable/
flickrapi	2.4.0	pypi.org/project/flickrapi/
google-cloud-vision	0.37.0	pypi.org/project/google-cloud-vision/

A.2.2 Sources for Figure 6.31

Image Sources given row-wise, starting from top-left.

Ice: Image (http://www.flickr.com/photos/9376953@N07/11925943546/) by «kaveman743», licensed under CC BY-NC 2.0. / Image (http://www.flickr.com/photos/49503162874@N01/3253508798/) by «Weiko», licensed under CC BY-NC-SA 2.0. / Image (http://www.flickr.com/photos/40511613@N00/3923428383/) by «Herman Beun», licensed under CC BY-NC-ND 2.0. / Image (http://www.flickr.com/photos/11121568@N06/4221908700/) by «Alan Cleaver», licensed under CC BY 2.0. / Image (http://www.flickr.com/photos/39881231@N02/7776109448/) by «victorfe», licensed under CC BY-NC-SA 2.0. / Image (http://www.flickr.com/photos/76967548@N00/6097018932/) by «Jerry %26+Cara», licensed under CC BY 2.0. / Image (http://www.flickr.com/photos/37887244@N03/3803896665/) by «tallguyuk», licensed under CC BY-NC 2.0. / Image (http://www.flickr.com/photos/87893616@N00/130365694/) by «afloden», licensed under CC BY-NC 2.0. / Image (http://www.flickr.com/photos/57644250@N00/11584082666/) by «welshmackem», licensed under CC BY-NC-SA 2.0. / Image (http://www.flickr.com/photos/11060230@N03/6826865125/) by «Genthar», licensed under CC BY-NC-ND 2.0. / Image (http://www.flickr.com/photos/12940826@N02/5087119155/) by «dibaer», licensed under CC BY-SA 2.0. / Image (http://www.flickr.com/photos/12999062@N02/1448104561/) by «docpap», licensed under CC BY-NC-SA 2.0. / Image (http://www.flickr.com/photos/35888870@N07/5795487498/) by «Fairy Heart+%E2%99%A5», licensed under CC BY-SA 2.0. / Image (http://www.flickr.com/photos/39552752@N00/6695523833/) by «Graham Chastney», licensed under CC BY-NC 2.0. / Image (http://www.flickr.com/photos/10299779@N03/8566109512/) by «HBarrison», licensed under CC BY-SA 2.0.

Mountain River: Image (http://www.flickr.com/photos/40124504@N00/359356611/) by «Frapestaartje», licensed under CC BY 2.0. / Image (http://www.flickr.com/photos/22086443@N00/2494806113/) by «Neuro74», licensed under CC BY-NC-SA 2.0. / Image (http://www.flickr.com/photos/23975257@N00/2951655674/) by «erikjgreene», licensed under CC BY-NC 2.0. / Image (http://www.flickr.com/photos/34106830@N08/3177413962/) by «Manolo Blanco», licensed under CC BY-NC-SA 2.0. / Image (http://www.flickr.com/photos/85598619@N00/272631816/) by «RightIndex», licensed under CC BY 2.0. / Image (http://www.flickr.com/photos/32077994@N04/4325774227/) by «lagusa», licensed under CC BY-NC-ND 2.0. / Image (http://www.flickr.com/photos/93702253@N00/9778247454/) by «Mathew Knott», licensed under CC BY-NC-SA 2.0. / Image (http://www.flickr.com/photos/54851755@N00/6323320030/) by «m.prinke», licensed under CC BY-SA 2.0. / Image (http://www.flickr.com/photos/7755749@N05/3528215675/) by «Simon Collison», licensed under CC BY-NC-ND 2.0. / Image (http://www.flickr.com/photos/28278892@N03/2672123411/) by «jlk.1», licensed under CC BY-NC-SA 2.0. / Image (http://www.flickr.com/photos/44124400268@N01/4713166763/) by «Duncan Rawlinson.+Duncan.co», licensed under CC BY-NC 2.0. / Image (http://www.flickr.com/

photos/93389756@N06/8491149665/) by «Lars Hercher», licensed under CC BY-NC-SA 2.0. / Image (http://www.flickr.com/photos/68994272@N00/9026271283/) by «baba80», licensed under CC BY-NC 2.0. / Image (http://www.flickr.com/photos/75711708@N00/4505460356/) by «letscommunicate», licensed under CC BY-NC-ND 2.0. / Image (http://www.flickr.com/photos/8739828@N04/7974811577/) by «Dan Irizarry», licensed under CC BY-NC-SA 2.0.

Wilderness: Image (http://www.flickr.com/photos/29572618@N08/2893176765/) by «gali367», licensed under CC BY-NC 2.0. / Image (http://www.flickr.com/photos/54750593@N02/5250924589/) by «Saverio S», licensed under CC BY-NC-ND 2.0. / Image (http://www.flickr.com/photos/22147533@N03/2334174171/) by «H%C3%A5var og+Solveig», licensed under CC BY 2.0. / Image (http://www.flickr.com/photos/13898538@N05/3751741587/) by «csmramsden», licensed under CC BY-NC-SA 2.0. / Image (http://www.flickr.com/photos/44124370018@N01/4783448649/) by «daveynin», licensed under CC BY 2.0. / Image (http://www.flickr.com/photos/8414677@N03/3114391808/) by «IzuenGordelekua», licensed under CC BY-NC-SA 2.0. / Image (http://www.flickr.com/photos/7313591@N02/6064770210/) by «Renaud Camus», licensed under CC BY 2.0. / Image (http://www.flickr.com/photos/71088526@N00/5064047433/) by «darquati», licensed under CC BY-NC-SA 2.0. / Image (http://www.flickr.com/photos/15192926@N00/5872569412/) by «lostajy», licensed under CC BY 2.0. / Image (http://www.flickr.com/photos/78128495@N00/10293778406/) by «Tjflex2», licensed under CC BY-NC-ND 2.0. / Image (http://www.flickr.com/photos/43169383@N00/5005415704/) by «teo de+pap», licensed under CC BY-NC 2.0. / Image (http://www.flickr.com/photos/10888421@N00/7362182602/) by «Gnal», licensed under CC BY-NC-SA 2.0. / Image (http://www.flickr.com/photos/10857883@N05/8097205139/) by «John6536», licensed under CC BY-NC-ND 2.0. / Image (http://www.flickr.com/photos/13706945@N00/8474569934/) by «michaelday_bath», licensed under CC BY-NC 2.0. / Image (http://www.flickr.com/photos/8507625@N02/5923101787/) by «Steve Selwood», licensed under CC BY-NC-SA 2.0.

Wildlife: Image (http://www.flickr.com/photos/44124400268@N01/4713198269/) by «Duncan Rawlinson.+Duncan.co», licensed under CC BY-NC 2.0. / Image (http://www.flickr.com/photos/79474275@N00/8161486507/) by «RossoGialloBianco», licensed under CC BY-NC-ND 2.0. / Image (http://www.flickr.com/photos/23340342@N04/2784893128/) by «Roberto Hernando», licensed under CC BY-NC-ND 2.0. / Image (http://www.flickr.com/photos/55426027@N03/8114147953/) by «Peter G+Trimming», licensed under CC BY 2.0. / Image (http://www.flickr.com/photos/76516524@N00/4008259245/) by «sometimesong», licensed under CC BY-NC-ND 2.0. / Image (http://www.flickr.com/photos/39569656@N07/4345742135/) by «Ayrcan», licensed under CC BY-NC 2.0. / Image (http://www.flickr.com/photos/35098766@N08/4904250819/) by «yorksranter», licensed under CC BY-NC-SA 2.0. / Image (http://www.flickr.com/photos/29481407@N04/2755855881/) by «eNikodemus», licensed under CC BY-NC

2.0. / Image (http://www.flickr.com/photos/59455568@N00/4027499820/) by «Purple Heather», licensed under CC BY-NC-ND 2.0. / Image (http://www.flickr.com/photos/11060230@N03/6826854391/) by «Genthar», licensed under CC BY-NC-ND 2.0. / Image (http://www.flickr.com/photos/8414677@N03/2409336843/) by «IzuenGordelekua», licensed under CC BY-NC-SA 2.0. / Image (http://www.flickr.com/photos/37574471@N00/35409490/) by «Philou.cn», licensed under CC BY-SA 2.0. / Image (http://www.flickr.com/photos/89782291@N00/4850135888/) by «Brian Digital», licensed under CC BY-NC-SA 2.0. / Image (http://www.flickr.com/photos/25302607@N00/4382842445/) by «Joe Dunckley», licensed under CC BY-NC-SA 2.0. / Image (http://www.flickr.com/photos/19868111@N00/4106466536/) by «apolci», licensed under CC BY-NC-ND 2.0.

A.3 Icons

Icon	*URL*
location/landscape/gallery/timeline/ savannah/target	freepik.com [13.04.19]
database	flaticon.com/authors/those-icons [13.04.19]
photographer	flaticon.com/authors/smashicons [13.04.19]
deep learning	flaticon.com/authors/phatplus [13.04.19]
question/data analysis	flaticon.com/authors/eucalyp [01.05.19]
flowchart	flaticon.com/authors/linector [13.04.19]
social network	flaticon.com/authors/prettycons [13.04.19]
trees	flaticon.com/authors/monkik [13.04.19]
filter	flaticon.com/authors/kiranshastry [13.04.19]
statistical analysis	flaticon.com/authors/eucalyp [13.04.19]
landscape	flaticon.com/authors/xnimrodx [01.05.19]
influence spheres	flaticon.com/authors/geotatah [01.05.19]

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.

Tobias Frey, 28.04.2020