University of Zurich[UZH]

Analysing Spatial Patterns of Coastal Cultural Ecosystem Services using Flickr and Wikipedia Data

GEO 511 Master's Thesis

Author
Nicolas Steinmann
15-718-885

Supervised by
Prof. Dr. Ross Purves

Faculty representative
Prof. Dr. Ross Purves

30.09.2022
Department of Geography, University of Zurich

Analysing Spatial Patterns of Coastal Cultural Ecosystem Services using Flickr and Wikipedia Data

A study of the Cultural Landscape at the East Coast of Great Britain

Nicolas Steinmann
15-718-885

Master Thesis
30 September 2022

Supervised by:
Prof. Dr. Ross S. Purves
Department of Geography
University of Zurich

Abstract

The world's coasts provide many ecosystem services that benefit human well-being. Coastal ecosystems are particularly important as a third of humanity lives within 100 kilometres of the coast. They supply provisioning, regulating, supporting and cultural services to sustain human livelihood. Cultural Ecosystem Services (CES) are immaterial services related to recreational, inspirational and social values that humans can benefit from. CES are very abundant at the coast and have increasingly been subject to research in the past years. In order to properly manage and protect ecosystems, it is important to be aware of which types of CES are provided and where they are located. To study the geographies of these services at the coast and inland, laborious PPGIS field studies have been conducted, that are usually limited in spatial and temporal scale. In recent years it became common to research CES with the analysis of Volunteered Geographic Data. The most popular source for these studies is the photo-sharing platform Flickr, as it offers easy and free access, contains abundant information about CES and allows users to geographically reference their content. The information conveyed by the image and its metadata (titles, tags, descriptions) have thus been used to differentiate and spatially analyse various types of CES.

Dependence on a single source of data introduces different kinds of bias that may misrepresent perspectives of certain demographics and their posting behaviour. Including a secondary data source for reference could help to complement the perspectives of Flickr data and visualise where these sources agree or disagree. Wikipedia offers a large repository of geolocated text data that could provide valuable information about CES across the landscape. This source has rarely been considered in research so far and thus this study aims to find out if Wikipedia is a suitable resource for spatial research on CES by comparing it to Flickr data.

The initial objective was to automatically classify Flickr posts with the CES class they represent, on a large scale along the entire East Coast of Britain. This was achieved using a Random Forest machine learning algorithm that uses the user-assigned tags to allot one of three CES classes selected for this study (*Landscape Appreciation*, *Historical Monuments* and *Nature Appreciation*). Using a limited set of variables, it was be demonstrated that a fairly accurate prediction can be achieved with this method. The main objective of this study is to compare the information content relating to CES and the spatial patterns between the Flickr classification and the Wikipedia article data set. The results were mixed and contained some uncertainties regarding data quality. Matching terms and concepts could be found

in Wikipedia and Flickr data and there was also some spatial overlap between the two data sources. There was also a correlation between Wikipedia articles containing relevant information for a CES and the amount of related Flickr posts in the vicinity. However, the large differences in sample size and the ambiguous spatial representation of Wikipedia articles introduce significant uncertainty to the results.

As a secondary objective, the Flickr classification was visually assessed for significant spatial patterns and the observations compared against established literature. The spatial patterns agree with related research, as the data points are often concentrated close to accessible (cities, close to roads) and touristic places (e.g. castles, old towns). A large part of the posts are also located very close to the coastline, most strikingly the posts referring to landscapes. The co-occurrence of CES, so called *bundles* were studied as well by measuring the spatial correlation between the classes. There was a significant correlation between the posts of the classes *Landscape Appreciation* and *Historical Monuments* to be found.

Acknowledgements

Here I would like to express my gratitude to all the people that supported me through the process of writing this thesis. I would like to thank all my friends and family for their support, motivation and occasional distraction during this whole time.

Special thanks goes to:

My supervisor Ross Purves, for his valuable advice, guidance and encouragement.

Laura Wysling, for proofreading and her relentless optimism.

Florian Sturzenegger, for annotating training data and pleasant company in the library.

Contents

List of Figures

List of Tables

Acronyms

API	Application Programming Interface
HCA	Hierarchical Clustering Algorithm
CES	Cultural Ecosystem Services
CICES	Common International Classification of Ecosystem Services
MEA	Millenium Ecosystem Assessment
NLP	Natural Language Processing
OSM	Open Street Map
POI	Point of Interest
PPGIS	Public Participation GIS
PUD	Photo User Days
RF	Random Forest
VGI	Volunteered Geographic Information

1 Introduction

1.1 Motivation

Since the inception of the Ecosystem Services concept in 1997 and its popularisation following the publishing of the Millenium Ecosystem Assessment, it has become a well-known paradigm to describe the usage of any natural resources as services ecosystems provide to sustain human life on earth (Costanza et al., 1997; Reid et al., 2005). Whereas this school of thought is usually concerned with material benefits, there is a whole class of Ecosystem Services that is defined by its non-materiality or as its core property. These are commonly referred to as Cultural Ecosystem Services and are often associated with more abstract concepts such as recreation, inspiration or sense of place that benefit human well-being (Cheng et al., 2019). Some scholars reject the simplistic definition of CES by their non-tangibility and reason that these services are produced in a reciprocal process between the ecosystem and humans that eventually produce them (Fish et al., 2016).

Few ecosystem provide more ecosystem services and CES than coastal ecosystems (Agardy et al., 2005; Brown & Hausner, 2017). This is noteworthy, as large and increasing proportion of the world population lives near the coast (CIESIN, 2012; Maul & Duedall, 2019). As many other ecosystems, the coastal ecosystems and their services are under threat all around the world. Driving factors for the deprecation of coastal CES include sea-level rise, tourism and urban expansion (Carranza et al., 2020; Smart et al., 2021; Taff et al., 2019). But not only can this theoretical framework be useful for conservation but can help with the planning and governance of urban green spaces (Guerrero et al., 2016), managing natural ecosystems (Clemente et al., 2019) and for touristic development (Ruskule et al., 2018).

There has been an interest to study the geographies of CES across different landscapes. In the past, this has usually been done by the means of PPGIS studies, where stakeholders are tasked to map CES in their surroundings (Plieninger et al., 2013; Sijtsma et al., 2019). With the rise of social media in the last two decade and the possibility to access its geographically referenced data, opened up new opportunities to study CES on a variety of spatial and temporal scales, with Flickr being one of the most popular data source for this task (Egarter Vigl et al., 2021; Ruiz-Frau et al., 2020; Santos Vieira et al., 2021). Flickr is a photo-sharing site that enables users to upload their pictures and allows other users to interact with the content. Flickr has more than 75 million users which on averge post around 3.5 million photos day, 4.5% of which are geotagged (Lopez et al., 2019). Apart from pictures, users can provide additional data to their posts such as descriptions, tags and a geographic reference which can all be accessed over an API (Flickr Development Team, 2014). Many studies use only one source of user-generated data at a time. However, some recognized that there is merit in using different data sources for different types of CES or to use another source to complement another (Havinga et al., 2020; Jenkins et al., 2016). In the context of CES research, one large user generated data source is often overlooked: The online encyclopedia project Wikipedia. It is currently the 7th most visited website world-wide (SimilarWeb, 2022) and has more than 55 million articles at the time, 6.5 million of which are in English. There is a community of 44 million users that contributed to Wikipedia (Wikipedia, 2022). This is an enormous repository of information with a wealth of geographic information and also provides geographic coordinates for many of its articles. Wikipedia will thus serve as a secondary data source in this study and its usefulness for providing accurate spatial information about CES shall be assessed therein.

The objectives of this thesis is to automatically classify a Flickr data set with a CES classification scheme and then to compare it against Wikipedia to find overlaps and gaps in the spatial occurrence of CES along the British East coast. This way, it should be determined if Wikipedia conveys the same information about CES in Wiki articles as in Flickr and if it correlates spatially to complement Flickr data and possibly increase the confidence in user-generated data. Apart from this, the automatic classification based on the Random Forest machine learning algorithm utilising user assigned tags shall be assessed for its accuracy. For the Flickr data there will also be a general analysis of spatial patterns and the occurrence of bundles in the study area to confirm whether patterns observed in literature are present in this study as well. This thesis is hoped to confirm that analysing user-generated data is indeed a valid tool for studying the spatial distribution of CES in coastal landscapes in a scalable and cost-efficient manner.

1.2 Research Questions

In the following, the research questions to be answered in this study are listed, divided into sub-question in order to provide a clear overview over the research objectives of this thesis.

Q1: Methodology:

- How well does the CES classification of Flickr posts using the RF method work?
  - How accurate is the classification?
  - Which factors influence data quality?
- Using Wikipedia as a complementary data source:
  - Do the two data sources agree/disagree?
  - Could Wikipedia data be used to complement Flickr data to map CES?

Q2: Cultural Ecosystem Services:

- Spatial Patterns of CES in Flickr data:
  - Do spatial patterns match those found in literature?
  - Are there bundles of CES that can be identified?

2 State of the Art

2.1 Cultural Ecosystem Services

Ecosystem Services

The thematic frame work of this study will be based on the concept of ecosystem services, more precisely on Cultural Ecosystem Services (CES). Even though this concept seems fairly popular and established in environmental sciences, it was conceived not too long ago. A precursor to this field arose in the 1980s amidst growing awareness and research interest into ecological issues leading to a transdisciplinary field referred to as *ecological economics* (Costanza et al., 2017). In the 1990s, the concept raised more awareness after a book was published including definitions and cases studies by Daily et al. (1997) and there was an initial attempt at valuing the total of global ecosystem services, which was estimated at multiples of the global GDP at the time (Costanza et al., 1997). Daily et al. (1997) defined Ecosystem Services as "the conditions and processes through which natural ecosystems, and the species that make them up, sustain and fulfill human life". At the same time Costanza et al. (1997) devised 17 categories of ecosystem services with their corresponding ecosystem functions. On the initiative of the Millenium Ecosystem Assessment (2005) (MEA), ecosystem services were divided into four broader categories, *Provisioning Services*, *Regulating Services*, *Cultural Services* and *Supporting Services*. Provisioning services provide humanity with goods and resources extracted from the ecosystem needed for humans to survive. Regulating services provide benefits such as water purification, storm protection and pest control. Cultural services enable humans to maintain recreational, aesthetic, educational values from an ecosystem. Supporting services refer to the basic functions that an ecosystem needs to maintain to *indirectly* provide ecosystem services to humanity, soil build-up and carbon cycling for instance (Costanza et al., 2017). The MEA had the objective to survey what implications ecosystem change has on human well-being, where Ecosystem Services are considered *linkages* between the two spheres (Millenium Ecosystem Assessment, 2005). This classification system has been adopted and developed by the Common International Classification of Ecosystem Services (CICES) to build a "hierarchical and science-based" classification scheme for Ecosystem Services (Costanza et al., 2017; Haines-Young & Potschin, 2012). The ecosystem services paradigm allows us to assess the dependencies between the natural environment and the feasibility and quality of human life. It has since become a very popular research framework and has been used increasingly in environmental sciences since the publication of the MEA (Bennett, 2017).

Cultural Ecosystem Services

Cultural Ecosystem Services are an integral part of ecosystem service classification since the very beginning of the concept (Costanza et al., 1997). The Millenium Ecosystem Assessment (2005) defines CES as "the non-material benefits people obtain from ecosystems through spiritual enrichment, cognitive development, reflection, recreation and aesthetic experiences" and commonly uses a classification system consisting of ten categories or classes (Table 1). Often, the Cultural Ecosystem Services are defined as being *non-material* in nature and at the same time acknowledging that defining concept is not straight-forward, as many *material* ecosystem benefits can have a *cultural* significance as well (e.g. sport fishing, hunting) (Haines-Young & Potschin, 2012). While CES are often the most compelling argument for ecosystem conservation and one of the biggest motivations for land ownership and management, they are disproportionately less considered in research compared to other ecosystem services (Hernández-Morcillo et al., 2013). The intangibility of CES is widely seen as one of the reasons these services receive less attention compared to provisioning and regulating services, as the benefits of CES are often subtle and manifest themselves in an indirect, intuitive manner and are more difficult to measure because of this (Cheng et al., 2019; Milcu et al., 2013). Within the different classes of CES there is a big discrepancy of representation in literature, as classes that are more tangible and thus easier to evaluate (e.g. recreation) are considered more frequently than classes that are more abstract (e.g. sense of place, inspiration) and thus a large part of studies focus on recreation and ecotourism (Cheng et al., 2019). Further obstacles that persist are the arduous way in which CES data typically has to be procured, mainly through questionnaires and interviews, and ambiguities and overlaps between the classes in a classification system. For example, recreational and inspirational benefits can be dependent on each other and thus are not easily distinguishable. Also there are different classification system (e.g. MEA, CICES) in use and the different service categories are not always congruent and transferable between the systems (Cheng et al., 2019). Even the defining property of intangibility or "non-materialness" has been challenged and a new framework to conceptualise CES has been proposed. One of the findings of Fish et al. (2016) is that CES should not be considered as a simple *subject-object* relation, where humans simply take what nature provides, but that there is a more complex relationship between the humans and the ecosystem. As humans engage with the ecosystem in cultural practice to procure cultural benefits from it, they shape the natural space with their activities. On the other hand, the ecosystem enables humans to engage in certain cultural practices in the first place. Mentioning intangibility or immateriality as the sole defining attribute of CES thus misrepresents the nature of how CES are produced which does involve some very tangible elements. This research helps to understand CES not simply as products that are consumed but as a multi-faceted system of flows and feedbacks between humans and nature.

Table 1: Classification of CES categories based on the 2005 Millenium Ecosystem Assessment report, in (Cheng et al., 2019).

Class	Concept
Cultural diversity	The diversity of ecosystems is one factor influencing the diversity of cultures.
Spiritual and religious values	Many religions attach spiritual and religious values to ecosystems or their components.
Knowledge systems	Ecosystems influence the types of knowledge systems developed by different cultures.
Educational values	Ecosystems and their components and processes provide the basis for both formal and informal education in many societies.
Inspiration	Ecosystems provide a rich source of inspiration for art, folklore, national symbols, architecture, and advertising.
Aesthetic values	Many people find beauty or aesthetic value in various aspects of ecosystems, as reflected in the support for parks, 'scenic drives,' and the selection of housing locations.
Social relations	Ecosystems influence the types of social relations that are established in particular cultures. Fishing societies, for example, differ in many respects in their social relations from nomadic herding or agricultural societies.
Sense of place	Many people value the 'sense of place' that is associated with recognized features of their environment, including aspects of the ecosystem.
Cultural heritage values	Many societies place high value on the maintenance of either historically important landscapes ('cultural landscapes') or culturally significant species.
Recreation and ecotourism	People often choose where to spend their leisure time based in part on the characteristics of the natural or cultivated landscapes in a particular area.

2.2 Coastal CES

Coastal ecosystems supply humanity with numerous benefits such as fishery and aquaculture production, climate regulation, flood protection and tourism among many others (Granek et al., 2010). Coastal systems provide disproportionately more ecosystem services than any other system, even if they are larger in area (Agardy et al., 2005). This is also true for CES in particular (Brown & Hausner, 2017). These findings bear even greater significance if one considers that more than a third of the global population lives within a 100 kilometre distance from the coastline with the global percentage expected to increase in the coming years in most parts of the world (CIESIN, 2012; Maul & Duedall, 2019). The drivers of change for CES have been identified to be of mainly economic, demographic and ecological nature. Economic drivers are processes such as coastal infrastructure, industrial fishing and aquaculture, while demographic drivers consist of phenomena such as rural depopulation or aging communities at the coast. Notable ecological drivers are habitat and biodiversity loss (Rodrigues Garcia et al., 2017). The various environmental pressures, such as climate change, pollution and habitat loss, threaten coastal ecosystems and the services they provide (Crain et al., 2009; Lu et al., 2018). These threats do also impact the provision of CES. Urban expansion can interfere with CES, for example, the sand dunes at the Mediterranean coast provide recreational benefits, but they have been degraded by urban sprawl along the coast line in the past decades which has an influence on the abundance and quality of the CES provided (Carranza et al., 2020). Tourism can also have a negative effect on many ecosystem services including CES, which can lead to a negative feedback on the attractiveness of a destination (Drius et al., 2019; Taff et al., 2019). Environmental changes can have an impact on CES at coasts too: Erosion of beaches and degradation of tidal areas would not only impact biodiversity and life supporting services but also cultural ones, such as scenic and recreational values (Brown & Hausner, 2017). Urban expansion can also work together with environmental effect such as sea-level rise in order to endanger the provision of CES at coast. This mechanism is referred to *Coastal Squeeze* and occurs when rising water levels and expanding urban spaces constrict natural spaces at the coast, leading to the loss of CES services (Smart et al., 2021). Environmental disasters, such as a 2019 oil-spill off the coast of Brazil, have been observed to lead to a immediate negative dynamic of the provision of CES in the affected areas (Azevedo et al., 2022). Coastal CES are thus threatened by environmental and climate change, which can have a major impact on the well-being of people, especially on the more vulnerable, such as disabled and indigenous communities (Kosanic & Petzold, 2020). The vulnerability of coastal areas that are home to a large part of the human population, further accentuates the importance of protecting coastal CES (Agardy et al., 2005).

As the importance of coastal CES is being acknowledged in the research community, the amount of relevant literature has increased in recent years. Rodrigues Garcia et al. (2017) reviewed a large collection of literature and found that literature on coastal CES follows the general trend of all coastal ecosystem services publications, and is seeing a substantial increase in recent years. Still, most publications were focused on provisioning and regulating services. For the CES-relevant literature assessed, most studies in coastal environments have been conducted in Western Europe, East Asia and North America. Another review agrees that the geographical distribution of these studies is heavily concentrated on the northern hemisphere with not a single case study situated in Africa (Martin et al., 2016; Rodrigues Garcia et al., 2017). The most researched CES classes were recreation and aesthetic followed by cultural heritage and identity (Brown & Fagerholm, 2015; Rodrigues Garcia et al., 2017). The apparent focus on recreation and leisure might exist because not all the classes have clear definitions, while recreation is more easily definable and has a counterpart in each conceptual framework (e.g. MEA, CICES) (Cheng et al., 2019; Hernández-Morcillo et al., 2013). These services are also more easily quantifiable and thus simpler to measure and valuate as the more abstract classes. As many studies are concerned with monetary valuation of CES, this is not unexpected (Milcu et al., 2013). On a similar note, recreation and aesthetics are considered more important to the global economy than other, less tangible ecosystem services (Rodrigues Garcia et al., 2017). A common finding was the recognition of synergies and trade-offs between CES and other ecosystem services that occur in bundles and are common in coastal landscapes (Hernández-Morcillo et al., 2013; Rodrigues Garcia et al., 2017). Recreation can be appreciated because of the aesthetic qualities or the historical significance of a site and thus the services a landscape provides can not be separated clearly into distinct categories and should be considered as bundles in synergy with other CES classes (Ahtiainen et al., 2019; Plieninger et al., 2013). In a coastal context, these synergies were found to typically occur with other CES classes while trade-offs were more common between CES and other services (Rodrigues Garcia et al., 2017). Categorising different CES is thus not as trivial as it appears at a first glance. When doing so, one needs to take into account that there are no discrete conceptual borders between classes and that ecosystem services often work in synergies and trade-offs with each other.

2.3 Spatial Analysis of coastal CES

CES in proximity to the coast have already been assessed in several case studies based on questionnaires. For example, this was done for the German North Sea coast to assess the potential impact of off-shore wind power plants on CES provision and to compare the relative importance of CES values between different nations around the Baltic Sea (Ahtiainen et al., 2019; Gee & Burkhard, 2010), which is also the most frequent method used in literature (Martin et al., 2016). While these studies give interesting insights into the perception of CES in a coastal setting, they don't explore the explicit geographies of the CES that have been insufficiently researched in the past (Brown & Fagerholm, 2015). Ruiz-Frau et al. (2013) went one step further and used low resolution maps to visualise coastal CES, but still used conventional questionnaires to gather the data.

As this study is concerned with analysing spatial patterns of CES, it is essential to discuss existing methods on how to assess the geography of CES. As established, CES are produced at the level of interaction between humans and the ecosystem (Fish et al., 2016). This makes the case for the use of Volunteered Geographic Information, where the perception of human beings serves as sensors in the field (Goodchild, 2007). Hence, data that originates from the *sense of place* of its users is rarely readily available from an official authority and needs to be collected first by the population that interacts with the landscape, describes it with its own words and provide a geographic reference (Purves et al., 2011). A popular tool to achieve this, are Public Participatory GIS (PPGIS) methods. This involves citizens providing geographically explicit information from their own local knowledge and experience which is quite popular for analysing the distribution of CES in particular (Brown & Fagerholm, 2015). Participatory approaches are deemed effective to identify CES and illustrate the diversity of its values and benefits in different places and ecosystems (Martin et al., 2016). It has indeed been a popular method for assessing CES at the coast and elsewhere. In fact, tied with interdisciplinary methods, it is the most used method to assess the spatial coverage of CES (Gliozzo et al., 2016). In the Wadden Sea region between Germany, the Netherlands and Denmark, places considered attractive were mapped by the community of *fans*. In this analysis the patterns and differences between the visitors (*fans*) home locations and CES at the Wadden Sea were assessed using an online GIS tool (Sijtsma et al., 2019). Cultural ecosystem services were also put on the map in the Oberlausitz region of Germany to analyse the perception of CES bundles and their geographic distribution in relation to landscape features (Plieninger et al., 2013). On a much larger scale, Brown and Hausner (2017) mapped cultural ecosystem values in various coastal areas around the world to discover spatial patterns that typically occur in proximity to the coastline and analysed the character and abundance of CES in various spatial contexts, such as closeness to roads and distance from the coast. A similar study used an online mapping tool for users to mark perceived cultural ecosystem values at the northern coast of Australia in different geographical contexts (distance to coast, land forms) (Kobryn et al., 2018). These two studies confirm that the closer to the coast, the higher the abundance of CES benefits becomes. Participatory GIS is seen as a viable method of assessing CES having the potential to include the perspectives of many different stakeholders (Depietri et al., 2021). If the analysis is web-based, one can reach broad audiences even far away from the study area (Sijtsma et al., 2019). Large scale assessments can be conducted this way albeit at the price of reduced quality and detail. A facilitated approach including face-to-face interviews and workshops with the stakeholders has the advantage, that the data collection can be more consistent and better monitored, though it is more labour intensive and thus costlier (Brown et al., 2012). The laborious nature of this method is the reason why participatory mapping studies are limited in spatial as well as temporal scale and can only provide a snapshot of the situation on the ground (Gee & Burkhard, 2010). A strength of PPGIS data collection and analysis, is the detailed demographic data that is typically procured during data collection, which allows to interpret the results in the context of different demographic factors (Depietri et al., 2021; Plieninger et al., 2013; Sijtsma et al., 2019). This also enables to recognize the demographic bias that might be present in the data (Brown & Hausner, 2017). This is a significant advantage over other popular methods, such as social media analysis, where reliable demographic information is hard to come by (Depietri et al., 2021).

2.4 CES in User Generated Content

Because of its limitations, spatially explicit studies of CES have long been constrained to a limited temporal and spatial scale, that can be expanded by utilising PPGIS methods at least in the spatial scale (Brown & Hausner, 2017; Depietri et al., 2021; Sijtsma et al., 2019). The rise of social media and the ability for users to spatially localize their content provides another opportunity to assess the interactions between humans and their environment (Gliozzo et al., 2016; Havinga et al., 2020). This interaction is recognized in the underlying theory on how CES are created and thus makes social media a valuable asset to observe CES in the environment (Fish et al., 2016). The intangibility that is often considered a challenge to CES assessment should be easier to grasp when the CES are *materialized* in social media post as the users leave digital traces of their cultural interactions with the landscape (Gliozzo et al., 2016). As of 2018 there were three billion active internet users, thus there potentially is a immense supply of geolocated data that could be used for the purpose of analysing CES in space (Egarter Vigl et al., 2021). The data that can be accessed is as diverse as it is plentiful. For CES it is very popular to work with image content that is typically automatically tagged by image-recognition tools that annotate image content (Cao et al., 2022; Egarter Vigl et al., 2021; Ruiz-Frau et al., 2020; Runge et al., 2020; Santos Vieira et al., 2021; Van Zanten et al., 2016). The most popular image source is the Flickr in these studies, while a few used Instagram or Panoramio for the image content. Even for photo-centric data sources there is usually an abundance of meta-data, including the geographic reference essential for studies like this. There is usually textual information that comes with each post, such as tags or titles that have been defined by the user (Gliozzo et al., 2016). Some studies use image tagging in combination with the tags or the user defined tags on their own to infer CES from the data (Fox et al., 2021; Hale et al., 2019). While other platforms, like Twitter, have also been used in recent studies, Flickr has been the most popular data source for CES research for its ease of access through APIs and the high diversity of information about the users preferences accessible through it (Gliozzo et al., 2016; Havinga et al., 2020). If text is used, sometimes it serves just to analyse the sentiment of the post and not for inferring the CES itself (Cao et al., 2022). While machine-tagging of photographs has become very popular, Hale et al. (2019) argue to also consider the *meta-data* that come with the data, as the tags captions and titles associated with the images provide a more complete picture of the user intent. In some cases, the location is used as the sole indicator of CES, without including images or further meta-data: Azevedo et al. (2022) for instance used the count of Flickr images at the Brazilian coast to measure the impact of an oil spill on CES provision. Other sources include crowd-sourced point and track data on hikes and bike rides on Strava and geolocated species observations recorded on bird-watching websites such as eBird. These were deemed especially useful to assess recreation- and nature-related CES provisions (Havinga et al., 2020). All these mentioned types of data are generally considered to be Volunteered Geographic Information (VGI), which uses technological systems to enable the general population to voluntarily contribute spatial data to a given data set (Cui et al., 2021).

There are several reasons why VGI data, in particular Flickr data, is seen as valuable for spatially explicit CES analysis. As mentioned before, this kind of data enables to observe human-nature interaction by collecting the *digital traces* users leave behind. Picture sharing usually occurs more often in locations that have a high cultural value to people. Thus these posts from Flickr and other image sharing pages have a dual purpose: To act as a receptor of CES and to report the properties of the CES observed (humans as sensors) (Gliozzo et al., 2016; Goodchild, 2007). It enables one not only to find out about the *what* and *where* but also how frequently the data points occur at a given location. Another advantage of using VGI data is, that data is created passively and independently from planning and management objectives, which is usually not the case when conducting a PPGIS study as an alternative (Depietri et al., 2021). A decisive advantage of using user generated content, is the large spatial and temporal scale that is possible while avoiding the high costs that come with doing the same with PPGIS methods (Brown & Reed, 2009; Depietri et al., 2021). Studies employing automated methods with crowd-sourced data are often on a sub-continental or even continental scale, something that would be very hard to achieve with field-based assessments alone (Runge et al., 2020; Van Zanten et al., 2016). But studies on a national or sub-national scale are common as well (Cao et al., 2022; Hale et al., 2019; Santos Vieira et al., 2021). As VGI data usually comes with a timestamp, it is possible to easily investigate the temporal patterns of CES conveniently without conducting several studies on the field (Depietri et al., 2021). For example one can examine seasonal as well as diurnal patterns depending on how precise the time stamps are in the data (Retka et al., 2019) It is even possible to conduct acute studies that can measure the temporal dynamics in context of an unforeseen event, as the oil spill example proves (Azevedo et al., 2022). Depending on the area surveyed, there is usually a large amount of data points present. The studies surveyed for this thesis usually had several thousands, to hundreds of thousands and up to several million spatial objects in their data set (Hale et al., 2019; Runge et al., 2020; Van Zanten et al., 2016). But the availability of data can vary from region to region. For Flickr there is a comparably good coverage for our study area in Britain, while the global south there is much less data available (Belyi et al., 2017).

In the existing research about CES provision with user generated data there have been different research objectives identified in the literature. Some studies are concerned with the properties of different data sources, usually Flickr, Instagram, Panoramio. The research objectives include comparisons of spatial distribution (Gliozzo et al., 2016; Van Zanten et al., 2016) or the content of the VGI data sources (Ruiz-Frau et al., 2020). Another common aspect that is studied frequently, is the connection between CES and *landscape values* or *landscape features* (Hale et al., 2019; Oteros-Rozas et al., 2018). Similarly, researching CES provision in protected areas and comparing it to non-protected ones is also a popular topic (Runge et al., 2020; Santos Vieira et al., 2021). Spatial patterns like these are generally well researched, while only few studies have been found that look at *bundles* of CES and how they work together spatially (Hale et al., 2019; Retka et al., 2019). This could be especially useful for learning the synergies and trade-offs between ecosystem services, which has not gained a lot of attention in coastal CES so far (Rodrigues Garcia et al., 2017). Many studies however, do not focus on how CES relate to environmental context but are rather interested in the technical aspects of automated classification of CES into distinct classes. A common objective is to use images and metadata to best capture user intent and usually employ a variety of machine-learning based methods to classify the data into distinct categories. Using image data only is the preferred method in many studies (Mouttaki & Erraiss, 2022; Runge et al., 2020; Santos Vieira et al., 2021), while some acknowledge that textual data can be useful for sentiment-analysis: Whether an experience is valued positively or negatively (Cao et al., 2022; Fox et al., 2021). However, some researchers have recognized that the abundance of textual metadata can be utilised to capture user intent better and reduce interpreter's bias in the training data (Hale et al., 2019; Ruiz-Frau et al., 2020).

As this thesis is concerned with studying CES in a *coastal setting*, precedent in literature is always useful. The closer one moves to coastal zones, CES generally become much more abundant (Brown & Hausner, 2017). Thus it is expected that there is also some significant research concerning mapping CES using VGI data. In recent years, more studies were published that surveyed the geographies of coastal CES in this way. One of the earlier studies is of Retka et al. (2019), where CES were studied at a strip of coast in Brazil with manually classified Flickr posts. A study comparing different data sources for their information content in select coastal tourist destinations (Ruiz-Frau et al., 2020). The same year, a large scale study, covering the entire Arctic Circle area was published (Runge et al., 2020). Another large scale study along the entire Brazilian coastline revealed spatial patterns in the context of protected areas (Santos Vieira et al., 2021). The latest three studies are from 2022 and include a paper about the Hong Kong coastal area (Cao et al., 2022), another study about the Brazilian coast line (Azevedo et al., 2022) and a pre-print in the Baltic region using a Convolutional Neural Network (CNN) (Mouttaki & Erraiss, 2022). Notable is the recent rise of coastal studies compared to general CES, which came earlier in the 2010s (Gliozzo et al., 2016; Van Zanten et al., 2016). Being underappreciated for a long time, coastal ecosystem thus received more attention in recent years.

For unstructured geo-text data like georeferenced Wikipedia articles, there has not been as much precedent for the purpose of mapping cultural ecosystem services. Jenkins et al. (2016) used Wikipedia articles in concert with Twitter posts to map *sense of place* across New York City. Wikipedia was chosen as a complimentary source as it offers a collective perception of a place, in contrast to Twitter post that reflect individual opinions and impressions. In combination with Flickr data, as intended in this very thesis, there was one study to be found. However, it did not semantically analyse the articles directly, but used its ontology of concepts to assign the image machine-tags to related topics (Egarter Vigl et al., 2021). However, this does not mean that Wikipedia data has not been considered for its geographical information value in the past. Research has been focused on recognizing geographic entities in the text and analysing the co-occurrence of place names (Hu, 2018). The study of Jenkins et al. (2016) indicates that the Wikipedia articles contain useful information for characterising a place. Owing to the scarce amount of literature, one can assume that there is a need for further investigation and research on this topic.

2.5 Machine Learning and CES

For the large spatial scale and high density of data in the VGI sources assessed, it is often not feasible to pursue a manual approach to analyse or classify the information provided. Fortunately there are various methods in data science to automate and expedite the otherwise tedious manual classification. Thus it is seen necessary to provide an overview over the various methods that have been used to classify CES or related indicators. However it also occurred that the data was classified manually (Retka et al., 2019) or by assigning the machine tags of the images directly to CES classes (Depietri et al., 2021; Runge et al., 2020). One of the examples found in literature is Santos Vieira et al. (2021), who used the Hierarchical Clustering Algorithm (HCA), which constructs a dendrogram based on a term-distance matrix between each photograph based on comparing the similarity (based on multidimensional "distance") for each possible word pair. In this case study the HCA has not been validated towards its accuracy and the output was simply assumed to be accurate enough for the cause. Another way of classifying social media posts based on their CES content, is a classification algorithm based on co-occurrence of tags or words in a single post. This method is called Graph Theory Network Analysis (GTNA) and has been explored in a CES-analysis of Twitter and Instagram hashtags in selected tourist locations (Ruiz-Frau et al., 2020). This method also works based on the natural clustering of the data but identifies a few broad thematical cluster from the tags. The advantage of this method is its emphasis on the relationship between tags, how often they co-occur and how interconnected they are. This analysis can be used to produce intuitive network graphs that can help to interpret the data more extensively. There is a further example of using multivariate-regression to determine the most influencial factors that correlate to a CES and using these factors to assign a CES class (Cao et al., 2022). Convolutional Neural Networks (CNN) were also among the methods tested, for example in a study along the coast of Lithuania which yielded good results in classifying multiple CES classes (Mouttaki & Erraiss, 2022). In a study by Chesnokova and Purves (2018) where environmental sound-scape (Geophony) indicators were extracted and classified from descriptions of the Geograph project in the United Kingdom. In this study, different variables containing the most common key-words and dictionaries of relevant terms where utilised to construct a Random Forest classifier to assign different classes of sound sources (e.g. Biophony, Geophony) to Geopgraph grid locations.

There are thus many ways how to automatically identify CES classes in VGI reliably, which leaves one with a lot of choice depending on the task at hand. Methods like these will be helpful for the following analysis as they can deal with a large amount of data that would be very labour intensive to classify manually.

2.6 Research Gaps

During the literature review, research gaps could be identified with open questions that remain to be answered. Some of which will be part of this thesis. The science of researching CES with user generated data is fairly recent, thus there are many opportunities to enhance the understanding of their geographies.

- Text driven analysis of CES: While there is a lot of precedent for machine tagging social media images, the textual meta data that comes with the dataset has often been neglected and could help understand the user intent better.
- Differential data sources: Most research focuses on using a single data source and thus representing a limited perspective of the user demographic. Including different text sources, such as Wikipedia articles, can include deviating perspectives on the perception of CES.
- Coastal CES are underrepresented in research: Even though CES are abundantly present in proximity to the sea shore and thus contribute to human well-being disproportionally, they have received less attention in the past than land-based CES provisions. Understanding the geographic patterns of CES enables to manage and protect these ecosystems better.
- The trade-offs and synergies between coastal CES and other ecosystem services have not been conclusively studied. Co-occurrence and clusters of different CES types and their interdependency in the spatial dimension needs to be considered to enable informed decision-making.

3 Research Area

For this thesis, the British North Sea Coast has been chosen as the study area. Initially, it was planned that the whole North Sea coast would serve as the study area, but this was soon reduced to the current dimension. For one, there are large differences in posting behaviour and frequency between different countries, this became apparent in initial trials and is also confirmed by literature (Belyi et al., 2017). Another reason is the large diversity of languages around the North Sea, which posed a great challenge for this analysis as it uses only text metadata from the Flickr posts. The British East Coast was chosen, as there was good data availability and the language the posts are written in is usually English and thus did not require further language skills or translation of the data. Conveniently, there was a large and recent data set available covering all of Britain, which made data access easier than sampling it with the Flickr API. The study area can not be made up of a single line and therefore a coastal zone needed to be defined. This zone was defined as a 1 kilometre buffer on both sides of the line. There is no official definition of how wide the coastal zone has to be. Distances at a similar magnitude were used in other studies (Brown & Hausner, 2017; Cao et al., 2022). The study of Santos Vieira et al. (2021) had even less tolerance at 50 metres on each side. During classification and initial experiments, the majority of data was in a coastal context, even with the one kilometre buffer, so its dimension seems appropriate for this study.

3.1 The British North Sea Coast

Situated at the eastern shore of Great Britain, the North Sea coastline stretches from Leathercote Point at the Channel all the way up to Dunnet Head in Scotland (International Hydrographic Organization, 1953). In the southern part, sandy beaches with dunes, tidal marshlands and soft-rock cliffs shape the landscape, while in Northern England and Scotland steep, hard-rock cliffs become more common (JNCC, 2003). Estuaries and firths, from the Thames estuary up to the Donnoch Firth, break the coastline and lead the sea water inland. While much of the coast is bordered by rural villages, and small to mid-sized cities, there are also larger population centres such as Tyne and Wear (Newcastle and Sunderland) and the Edinburgh metropolitan area. The coastal landscape provides many opportunities for recreational activities and is thus a centre of attraction for tourist as seaside towns had been developed into coastal resorts since the 19th century (Light & Chapman, 2022). And still today, coastal tourism provides 210'000 jobs in England and Wales alone, which makes the seaside tourism sector one of the biggest employers in the region (Beatty et al., 2014). The coast is also a habitat for the local marine and coastal wildlife and a significant part of the shore is within a Marine Protected Area

(NatureScot, 2022; The Wildlife Trusts, 2022). There are numerous National Nature Reserves (NNR) and Areas of Outstanding Natural Beauty (AONB) along the coast of both nations (Marsh, 2018; Natural England, 2020; Scotland's NNR, 2022). Thus, the study area is close to sites that are a valuable habitat for marine species and that also have an aesthetic quality attractive to visitors. Furthermore, the coast is a recreational area for the tourists and local inhabitants to enjoy. There are thus many different stakeholders with different interest that make this area suitable and interesting for studying Cultural Ecosystem services.

4 Data

4.1 Flickr Data

4.1.1 Data Retrieval

The Flickr data collection used in this analysis was provided by the Geocomputation Unit of the Department of Geography and has been used for a number of different projects before. It was compiled using the Flickr API and is more than 6.5 GB in size. It contains more than 27 million posts sampled in a grid of square bounding boxes covering the United Kingdom, with an upload date between the January 1st 2007 and December 31st 2016. The data was received in folders of text files containing the following data fields:

- Geographic coordinates (lon/lat)
- Post Title
- Photo URL
- Photo ID
- Username
- User ID
- Date/time taken
- Date/time uploaded
- Number of views
- User defined image tags

An example on how the metadata typically is presented online, including all the available text meta data (Username, Title, Description, Tags) is pictured in Figure 1. Note that the description is not included in the study data set but is accessible through the official Flickr API. Presumably to preserve memory while retrieving the data through the API, the data was saved in single text files, containing about 50 MB of data each. These files were put into folders named after the bounding box they were sampled in. To make the data less unwieldy and easier to work with, the text files were combined in a single CSV file for each bounding box.

4.1.2 Pre-Processing

In the pre-processing, the Flickr data was cropped to the study area and some initial filtering was conducted to assemble the final dataset that this study will work with. In a first step, the data was cropped to the study area around the British East Coast. This was done by using the one-kilometre buffer zone geometry to conduct a spatial join with the point-geometry data of the Flickr posts to select the points

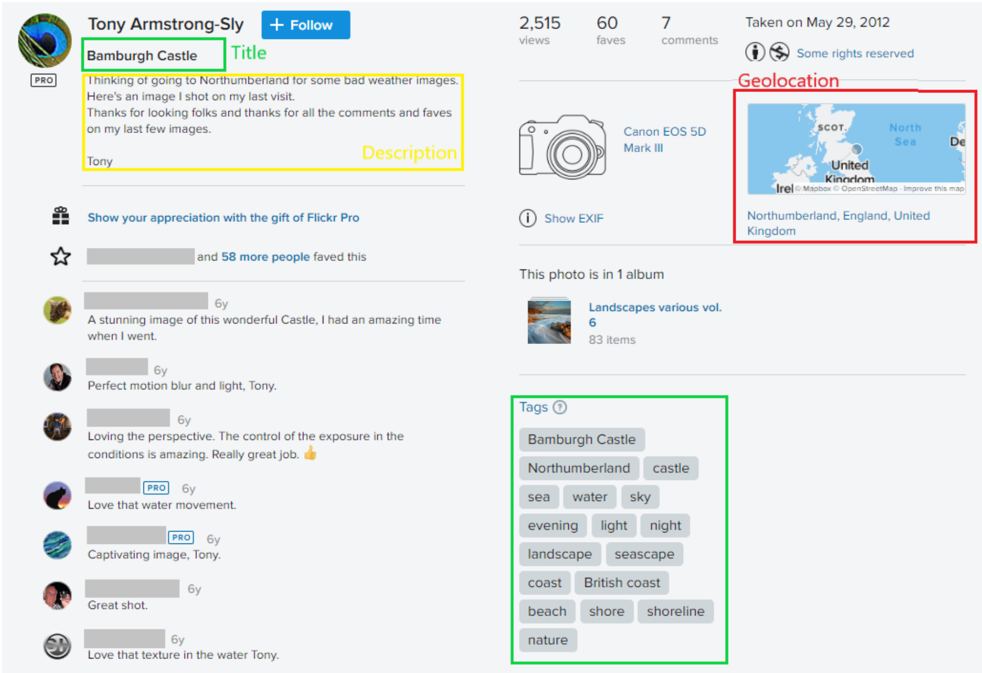


Figure 1: Overview over the available meta-data for an exemplary Flickr post picturing Bamburgh Castle, as displayed on the Flickr user interface.

within the buffer zone. This was repeated for each bounding-box CSV as to not run out of memory in the process. The cropped data from each CSV-file were then merged into a single one, containing all the posts along the entire coastline. With the cropping of the data complete, the data still needed to be tidied up. First, only photos taken between January 1st 2007 and December 31st 2016 were included, so that only posts actually taken during the sampling period were included. Correctly dated historical images were also present in the data set, but they provide little information about the contemporary state of CES presence in the study area and thus needed to be excluded as well. Second, posts with less than three tags were excluded, as a certain number of tags should be present in order for the post to be useful for this analysis.

In the next step, some of the bias inherent to the data was addressed. One source of bias in the data is the uneven distribution of posting activity between users. While there are many users that are not very active, the vast majority of pictures on Flickr is posted by a small minority of very active users (Purves et al., 2011). In the cropped raw data of this data set, the top five percent most active users (by number of posts) in the study area contributed more than two thirds of all posts (Fig. 3). Pictures from very inactive users that only posted two pictures or less in the area were first removed. These users are assumed to be *test-users*, posting with the intention of trying out Flickr features and do not intend to share their impressions like more active users (Runge et al., 2020). Furthermore, the less active users are expected to show a different tagging behavior which sets them apart from more prolific users (Purves et al., 2011). The posts of the most active users were not removed all together. Instead, each user, no matter how active, was allowed one randomly sampled post per day and all the other posts were rejected. This metric is called Photo-User-Day (PUD) in literature and deals with the problem of users uploading large collections of photos shot on a single day in the same place (e.g. sporting events, air shows) and thus are over-represented in the data set (Gosal et al., 2019). While this process reduced the data set to 10% of the original size, it did help to reduce the imbalance between active and less-active Flickr-Users. After processing, the top five percent of users contributed just about 45 percent of all posts and the cumulative graph is slightly steeper than before in the upper quantiles (Fig. 3).

The filtered data set thus contains 90'413 single Flickr posts, down from the cropped 1 kilometre buffer raw data that contained over 874'335 posts. Most of the data was lost during the Photo-User-Day sampling which suggests that users often upload multiple pictures taken in one day. For this analysis however, this is expected to be an appropriate amount of data to work with. When inspecting the data on a map it becomes apparent that there is a tile of data missing around the Thames Estuary in the very south of the coastline, visible as a rectangular cut-out from the point data (Fig. 2). As the raw data was retrieved divided into tiles as well, it is suspected that the tile was accidentally left out from the data set during retrieval, or was not included as it is located on the periphery of the sampling area. To make sure, there was no error in the unpacking of the data, the code was double-checked and run repeatedly which yielded the same results every time. However, this gap is considered only a minor nuisance because the large majority of the study area is still covered by the data.

4.1.3 Temporal Overview

The temporal distribution of the Flickr data shows a distinct seasonal pattern with a peak in the middle of the year before reaching a valley at the end of the year (Fig. 4). This seasonal pattern could be explained by the warmer weather in summer which leads to more people taking pictures outside. During the summer, people also tend to spend more time at the coast and this might also accentuate this dynamic. Over the whole sampling period, there is an increase of posting activity to be observed followed by a steady decline after 2014. Most likely, this is the effect of the popularity of the platform.

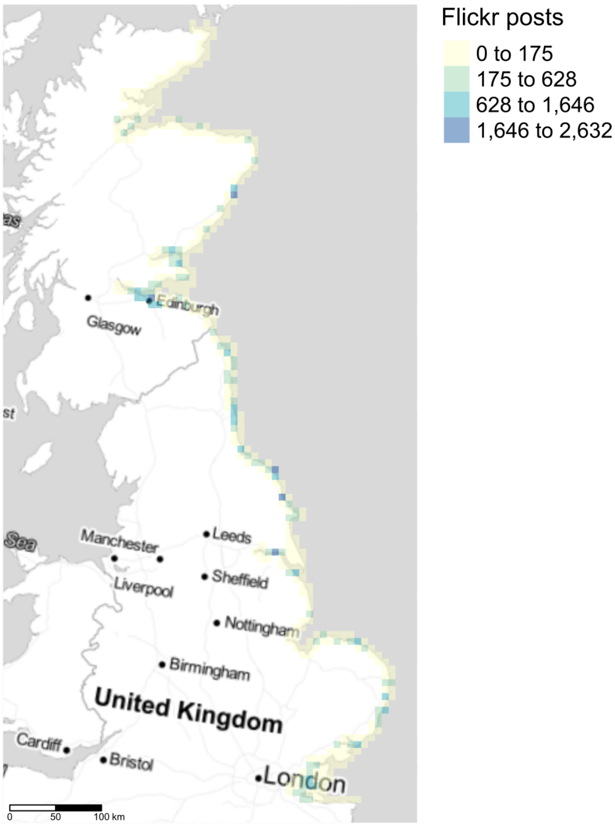


Figure 2: Flickr post distribution in the research area, aggregated into 7.5 kilometre grid cells.

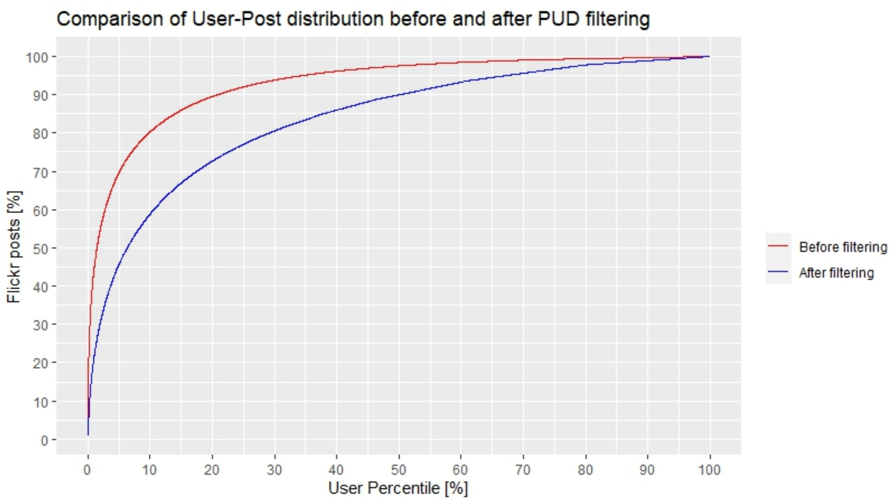


Figure 3: Cumulative distribution of Flickr posts by user quantile.

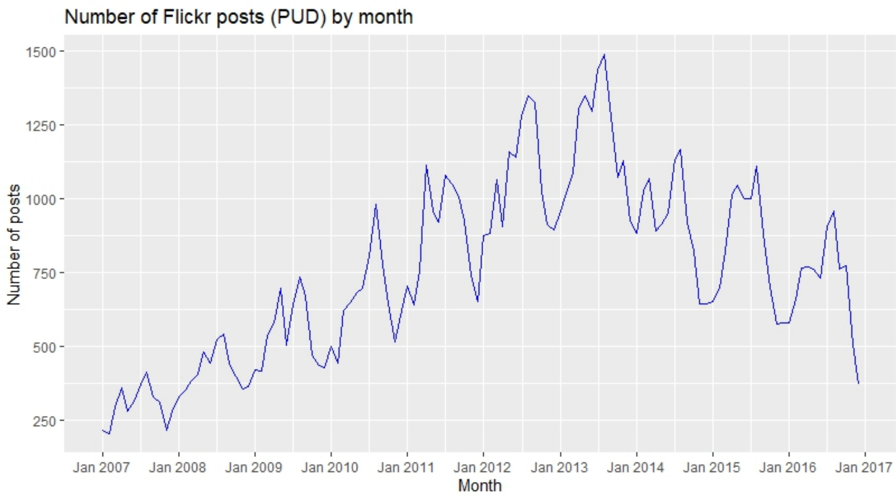


Figure 4: Temporal distribution of Flickr posts by month over the entire sampling period.

4.2 Wikipedia Data

4.2.1 Data Retrieval

The metadata for the geolocated Wikipedia articles were queried with the SPARQL-Wrapper Python library that allows a user to query the DBpedia database, which provides a defined database structure for Wikipedia-article metadata. SPARQL is a query language that resembles SQL (Structural Query Language) in many ways and is the usual choice for accessing DBpedia datasets (Lehmann et al., 2015). Only articles that were categorized as a *place* were included. Geolocated articles that refer to other locatable entities such as one-time events (e.g. battles, plane crashes) were excluded this way, as they are expected to provide little information about the contemporary presence of CES. Another essential requirement for the articles is that they need to be georeferenced with a lon/lat coordinate pair (only point data possible) in a defined bounding box containing the study area along the British East Coast. Furthermore, only articles in the English language were included in the query. The query used to access this data can be viewed in Figure 5.

The following additional metadata items were accessed for each queried entry:

- Article ID (unique)
- Article Title (english)
- Article Category 1st degree
- Article Category 2nd degree
- Article URL
- Number of Characters

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

SELECT DISTINCT ?id ?label ?link ?lat ?long ?cat_lab ?cat_lab2 ?nchar
WHERE  {

?uri a dbo:Place .
?uri rdfs:label ?label . FILTER(lang(?label) = 'en') .

?uri rdf:type ?cat . FILTER (?cat LIKE <http://dbpedia.org/ontology/%>).
?cat rdfs:subClassOf ?cat2 . FILTER (?cat2 LIKE <http://dbpedia.org/ontology/%> AND
                                     ! ?cat2 LIKE <http://dbpedia.org/ontology/Place> AND
                                     ! ?cat2 LIKE <http://dbpedia.org/ontology/Location>) .
?cat rdfs:label ?cat_lab . FILTER(lang(?cat_lab) = 'en')
?cat2 rdfs:label  ?cat_lab2 . FILTER(lang(?cat_lab2) = 'en')

?uri dbo:wikiPageID ?id .
?uri geo:lat ?lat .
?uri geo:long ?long .
?uri dbo:wikiPageLength ?nchar .
?uri prov:wasDerivedFrom ?link .

FILTER(?long >= -1.11 AND ?long <= -0.5 AND ?lat >= 54.165 AND ?lat <= 54.72)

}

ORDER BY DESC (?label)

LIMIT 10000
OFFSET 0
```

Figure 5: SPARQL query for accessing Wikipedia article objects from the DBpedia database.

While running and adjusting the query several times, the technical limitation of DBpedia became apparent. The queries were limited to return no more than 10'000 records, presumably implemented to prevent overburdening the database with excessively large query requests. This could be bypassed by running the same query multiple times and increasing the offset by another 10'000 each time it was run, until no records are returned. These records were each written to a CSV file and then appended into a single CSV containing all the records. This manual approach was viable in this case as the total amount of articles returned were manageable this way and it would need to be automatically iterated if many more articles needed to be accessed. During the initial plotting of the raw data as point vector data, there were several vertical strips devoid of data to be seen in the map, which could not be explained by the nature of the data and were likely the fault of the database query. This was remedied by adjusting the bounding box to the extent of the strips and appending the returned data to the larger, "striped" dataset. The raw data contains 24'315 pieces of Wikipedia article metadata.

4.2.2 Pre-Processing

These articles now needed to be further processed by spatially cropping the article metadata to the buffer polygon that defines the research area. Using the research area geometry directly to select the articles would be possible in SPARQL as well, however it is not trivial and was faster implemented by using QGIS for one single data set. A spatial join with a two-sided one and five kilometre buffers was conducted on the dataset. The five kilometre buffer was calculated as a fallback in case a larger study area for Wikipedia data was needed. This produced two datasets, one containing articles in the one kilometre buffer with 615 entries and the other with 1671 articles within a five kilometre buffer. It should be noted, that this dataset only contains the metadata for each Wikipedia article. The text content of the articles will be accessed with the provided URL and an HTML scraper when needed. As there is missing Flickr data around the Thames Estuary, this area was excluded from the Wikipedia data as well, to not skew the results of the spatial analysis by missing data. This reduced the total amount of geo- referenced Wikipedia articles in the two-sided one kilometre buffer to 553 in the 1 kilometre buffer (Fig. 6).

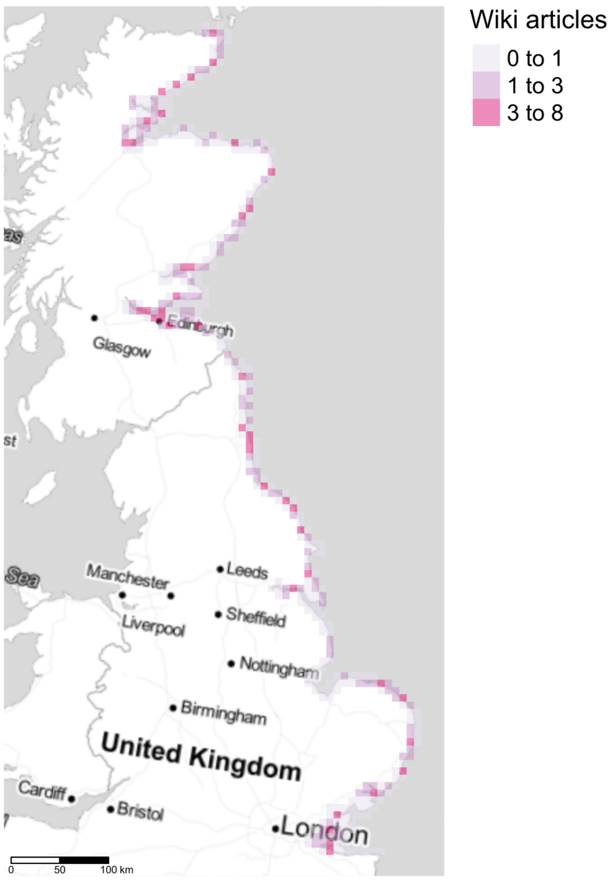


Figure 6: Wikipedia article distribution in the research area, aggregated in 7.5 kilometre grid cells.

5 Methods

5.1 Flickr Data

5.1.1 Training Data

In order to construct a supervised classification model, training data is required to provide the model with a number of independent variables and a dependent variable, in this case a CES class. The independent variables will be based on the user-assigned tags for each spatial object (Flickr post). The classification scheme of this study originates from the nine-class categorisation of Retka et al. (2019), but was adapted to meet the requirements of this study. Based on the experiences in early trials, the classes *Natural Structures and Monuments* and *Landscape Appreciation* were hard to distinguish, as similar terminology was used in the tags. Because of this, it was decided to combine the two classes under *Landscape Appreciation* as natural structures can always be considered elements of a landscape, which might explain the conceptual closeness of the two classes. Furthermore, to reduce complexity and sample size issues, the least represented CES classes were not considered and the three most frequent classes remained to be assessed in this analysis. The remaining classes (*Sport Recreation*, *Religious, Spiritual or Ceremonial Activities*, *Research and Education*, *Artistic or Cultural Expressions and Appreciation*) thus were to be labeled *Other* in the training data.

The training data for the CES-classfication of the Flickr posts was subsetted from the pre-processed data set by random selection. For the annotation of the training data the *LightTag* online annotation tool was used to assign a CES-class to each of the posts. On the online interface, the annotators could read the tags and open a low-resolution version of the photograph the tags were assigned to. The annotators were tasked to determine the *user intent* based on the tags and the image and assign one of the four CES-classes chosen for this analysis. Assigning only one class to each photograph is considered standard practice, albeit the possibility that different classes can be contained in the same post (Retka et al., 2019). In a short trial of 200 posts, the Cohen's Kappa coefficent was calculated. This measure is used to assess the agreement between two rounds of annotation to determine if the classes chosen are easily definable and unambiguous. Between two annotations by the same person three days apart, the Kappa was at 0.89, with 1 being full agreement. This is a very good level of agreement but is expected to be lower if more annotators were to be involved. The annotators were issued a list of precise annotation rules to help them choose the correct CES class (Figure 7). If the post in question represented another CES class or none at all, it was intended to be tagged as *Other*. If the image was missing, annotators had to assess the post based on the tags alone. If the tags were vague or inconclusive, the class had to be assigned by image alone. In the case of both the image being missing and the tags being vague, the annotators were instructed to classify as *Other*.

CES class	Description
Historical Monuments	Posts depicting historical manmade structures • Examples: old churches, castles, forts, ruins, statues, memorials, plaques, historic buildings, old harbours, disused industrial facilities and mines, light house, historic ships. • Historic building has to be the clear focus of the picture or the tags, if it is part of a wider landscape it should be classified as landscape appreciation.
Landscape Appreciation	Posts for which the main focus is on a view of the landscape • Can also include single natural structures that make up the landscape (cliffs, rivers, dunes etc.) • Wide scale views of man-made structures such as of a city/town or a harbour are also considered a landscape.
Nature Appreciation	Posts focusing on animals, plants or other living organisms • Has to have a clear focus on the creature: e.g. images/tags focusing on a cliff with birds flying around it belongs to landscape appreciation • Applies to wild animals mostly, pets (e.g. dog being walked), rather belong to social recreation, domestic animals in a field would count as nature appreciation • Animals and plants in an indoor setting (e.g. cat pictures) shall be classified as "other"
Social Recreation	Posts that represent groups of people in an informal or non-dedicated recreative (i.e. not sport) social environment in an **outdoor** coastal setting • People meeting other people on the coast and engaging in social activities
Other	Posts that represent **another CES** not in the scope of the study • Sports Recreation: any activity depicting people engaging in sports • Artistic expression: concerts, festivals, local culture, parties, folklore, exhibitions • Research and Education: people engaging in research or teaching activities Posts that **do not represent** a coastal CES • Not a coastal or maritime setting • Advertisements and Memes • Artwork • Plane/train/bus/ship-spotting: Focus on vehicle/infrastructure indicating technical terms such as registration number, model, manufacturer, name, line/route, company etc.

Figure 7: Annotator instructions for annotating the training data set with the selected CES classes.

For the final training data set, 4000 posts were manually classified by a single annotator. Out of all the posts classified, 1222 were found to represent *Landscape Appreciation*, 413 posts related to *Historical Monuments*, 403 to *Nature Appreciation*, while *Social Recreation* amounted to 264 posts and the remaining 1698 posts were classified as *Other*. Another 500 posts were classified by one additional annotator and shall be used as an additional set for a final round of validation to ensure the quality of the classification.

5.1.2 Random Forest Model

To classify the Flickr with CES classes, the Random Forest (RF) algorithm was chosen. This algorithm is related to the Decision Tree machine learning algorithm, as it also uses a Decision Tree for its core mechanism. A decision tree can be conceptualized as a cascade of nodes and branches. At each decision node, the observation is passed to one of two branches according to a test conducted in the node. This happens from the root to the leaf node, where the value of the target variable will be returned (T R, 2015). As the name implies, the Random Forest model grows a forest of decision trees and lets each observation run through every tree. The most common output value from all the trees will then be returned as the dependent variable. The trees are usually grown with the same algorithms Decision Tree models use, but the training data set as well as the variables tested are randomly subsetted in order to produce variations of trees. When training an algorithm, the number of trees grown can be specified, in the R package *randomForest* the default value is set to 500 trees (Biau & Scornet, 2016; Liaw & Wiener, 2002). The RF algorithm was conceived by Breiman, 2001 in this form and offers many advantages compared to simple decision trees and other machine learning methods in this use case. Advantages include that RF can be applied to a variety of prediction problems and can cope with small sample sizes and high-dimensional feature spaces while retaining a high accuracy (Biau & Scornet, 2016). Additionally, this method is easy to use as it essentially needs two parameters to adjust: The number of trees and the number of variables used at each node of the trees (Liaw & Wiener, 2002). Its versatility and simplicity makes this the preferred method for the task at hand, as it is easy to implement and adapt in order to optimize the algorithm.

Recent examples of the use of RF algorithms in GIS prove that this is a valid approach for assessing landscape qualities. The perks of being able to easily integrate different data types from various sources with RF has already been demonstrated by combining Flickr image- and metadata-tags and official data such as elevation models, landcover types and historic POIs to model the scenic quality of a landscape (Havinga et al., 2021a, 2021b). On a similar note, there has been extensive research on classifying sensory qualities of landscapes in Britain using unstructured text descriptions by combining most frequent tags with compiled dictionaries of relevant terms as additional features (Chesnokova et al., 2017; Chesnokova & Purves, 2018). The existing research provides precedent that the RF supervised machine learning model is suitable for the purposes of this study. Furthermore, the versatility and adaptability of the method allows for incremental improvement of the model by experimentation with different combinations of variables and features.

The tools chosen to implement the classifier model for this study were the *R* statistical programming language (version 4.2.0, 64bit) with the relevant software libraries for RF modelling. For RF packages, there is essentially a choice between two well-known packages, *randomForest* and *ranger*. The former is the standard implementation based on the original Fortran version by Breiman and Cutler while the latter is a newer, improved version by different developers (Liaw et al., 2022; Wright & Ziegler, 2017). The *ranger* library is substantially faster but uses much less memory compared to *randomForest* but performs just as well and uses the same simple syntax of the original library (Wright & Ziegler, 2017). This comparison made the *ranger* package the right choice because it conserves valuable memory resources and saves a lot of processing time during the implementation.

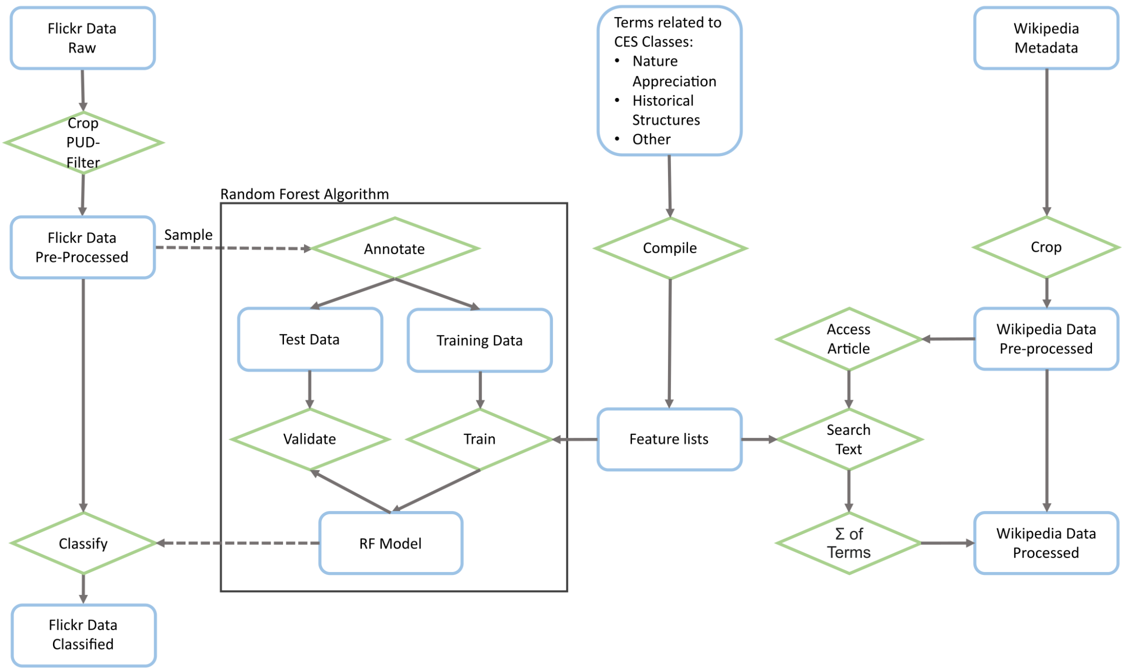


Figure 8: Process diagram of the processing and classification of Flickr and Wikipedia data used in this study.

5.1.3 Implementation & Validation

For the RF model to be able to classify the Flickr posts, the data needed to be converted into a format that the model could work with. The *ranger* package takes a data frame as an input and uses the columns as variables to build a multi-dimensional feature space. The tags are supplied as delimited chains of characters in the raw data and this is a format that the algorithm can not work with. Thus the tags selected as variables are represented by a field each and their occurrence in an object be indicated by a numerical binary value. Using every single word occurring in the tags would use an excessive amount of processing power and words rarely used in the training data are generally not helpful to a classification. As a consequence, the choice of words had to be limited. This was done by selecting the 500 most common terms in the training data set. In the word list to be used in the end, terms were then manually edited to exclude place names, user names, brand names of camera equipment, numbers and tags in a non-latin alphabet. Place names were excluded as they were not expected to help discriminate the classes very well and bias the classification to more popular places. User names were also excluded for a similar reason, as it was suspected that they may bias the classification results to users that tag many uploads with their user name. The unedited tag list was also used in the validation process to determine the effect the editing has on the accuracy of the algorithm. For each post, the tags had to be checked if they contain each word by iterating through the word list. If a word did occur in the tags, the respective field received the value of 1, else the value 0 was assigned. This resulted in a table of binary values with each post represented as rows and variables (tags) as columns. This matrix was then rejoined with the original data table of the training data. This transformation from text- to codified binary data enables the model to build the decision trees and classify the whole data set.

To improve the accuracy of the model, word lists of relevant terms for the CES classes *Nature Appreciation*, *Historical Monuments* and *Other* were created. Any post that contained tags that also occurred in one of the lists was assigned a positive binary value in the corresponding field. This method was successfully employed by (Chesnokova & Purves, 2018) to improve the accuracy of the RF algorithm, referred to these lists as *features*. For *Historical Monuments*, the feature contains different generic types of historically significant structures (e.g. castle, abbey, fort), adjectives referring to age (e.g. old, ancient, ruined) and time periods (e.g. medieval, victorian, worldwar) containing 62 terms in total. For *Nature Appreciation* the list consisted of English designations of common sea- and waterbird species (e.g. Puffin, Heron, Gannet, Duck), marine mammal species (e.g. Seal, Dolphin, Porpoise) and a selection of generic terms relating to flora and fauna (e.g. flower, nature, wildlife). This feature contains 82 terms. As species names are often occurring together with a subspecies designation (e.g. Arctic Tern, Common Gull) the matching mechanism utilised for this feature was a partial match instead of a perfect match. The *Other* feature mainly contains transport related terms (e.g. train, bus, plane, ship) because many posts in this class refer to different types of vehicles. Furthermore, selected words referring to CES classes not considered here were also added to the list. The *Other* feature contains 112 terms in total.

With the training data complete and appropriately formatted, the algorithm could now be trained and optimised. A simplified overview over the process is illustrated in the process diagram in Figure 8. The machine learning algorithm was to be optimised to maximize sensitivity and specificity while keeping complexity and processing power needed to a minimum. Sensitivity represents the fraction of posts that were correctly classified as belonging to a class in respect to the annotated ref-

Table 2: Specificity and Sensitivity of different word list and feature combinations used as variables for the Random Forest classifier.

Variable selection	Unredacted, no features	Unredacted, 3 features	Redacted, no features	Redacted, 3 features
n	896	896	757	757
Sensitivity	0.65	0.71	0.71	0.75
Specificity	0.76	0.81	0.81	0.85

erence (true-positive rate). The specificity describes the fraction of posts correctly identified as *not* belonging to a certain class as compared to the reference data (true-negative rate). The annotated sample was filtered to only contain data with tags that are contained in the redacted list of the 500 most frequent words at least one time per post. This reduced the initial sample set by about 24% from 4000 down to 3028. This sample was then randomly split into a training set containing 75% of the sample and a test data set containing the remaining 25%. After the model had been trained using the training data, the test data set was classified, compared to the reference data and the aforementioned accuracy measures were calculated for validation. The process started out with only the 500 most frequent terms as independent variables either with manual edit or without. Later, the three features were included in the independent variables. In the very first trials it became apparent, that the CES class *Social Recreation* had insufficient sensitivity (true-positive) rates ranging between five and 20 percent, which did not improve much during the optimisation process. For this reason, the posts of this class were reclassified to *Other*. It is suspected that this class performed badly because of its low sample size in the training data and the difficulty to discriminate it from other CES classes during annotation. For the optimisation, different combinations of word list and feature lists were assessed for their sensitivity and specificity. To find out what effect the choice of variables has on the accuracy measures, different combinations of variables were compared. In the comparison, it became apparent that including the three features did improve the sensitivity by about five percentage points in combination with either the edited or unedited list. The specificity (true-negative rate) did also improve between the unredacted list with no features and the redacted list with all the features by about 10 percentage points (Table 2). For the classification of the entire dataset, the redacted word list including the features was chosen, not only because it is the most accurate, but also because the sensitivity was more balanced between the classes in the model (Table 3).

The model chosen for the classification has an overall sensitivity of 0.75 (CI 0.72 - 0.78). The overall specificity is at 0.85. All the classes except *Historical Monuments* reached sensitivity value of around 0.8, the latter only reaching a true positive rate

Table 3: Accuracy measures for a training set by class, using an edited 500-word list and 3 features, n = 757.

	CES classes			
	Landscape Appreciation	Historical Monuments	Nature Appreciation	Other
n (reference)	274	77	78	328
n (prediction)	293	51	85	328
Sensitivity	0.79	0.47	0.81	0.77
Specificity	0.84	0.98	0.97	0.82

Table 4: Confusion table for the RF model in Table 3. Class key: 1 = Landscape Appreciation, 2 = Historical Monuments, 3 = Nature Appreciation, 4 = Other.

		Reference			
		Class 1	Class 2	Class 3	Class 4
Prediction	Class 1	216	16	7	54
	Class 2	5	36	1	9
	Class 3	9	0	63	13
	Class 4	44	25	7	252

of 0.48. The Specificity (true-negative rate) was satisfactory as well and ranges between 0.82 and 0.98 (Table 3). By looking at the confusion matrix generated for this implementation, the most frequent misclassification was attributed to *Other* for almost all classes. Between the remaining classes, the confusion is usually lower, only between *Landscape Appreciation* and *Historical Monuments* was some relatively frequent mismatch to be found (Fig. 4). Because the choice of variables was derived from the very training data that was used to build and verify the model, it had to be assessed if it yields similar results by running an independent training data set annotated by a second annotator through the model. As expected, the results were less accurate than with the original training data set, with a sensitivity of just 0.65 (CI 0.61 - 0.71), underperforming the first assessment with 0.65 in *Landscape Appreciation* and 0.28 in *Historical Monuments*, whereas the other classes did perform the same or slightly better than before.

5.2 Wikipedia Data

The text data sourced from Wikipedia articles in the research area was also to be analysed for its content relating to CES classes. Compared to the Flickr data set, the Wikipedia articles are more complex as the data is in natural language and does not follow a standardised structure. The Flickr tags used here consist of cleanly delimited terms and are thus arguably easier to work with as they do not require complex Natural Language Processing (NLP) operations to extract the relevant information from a text. Due to the complexity of the task and time constraints, it was decided to forego an extensive classification of Wikipedia articles akin to the RF algorithm used with the Flickr posts. Instead, it was decided to search the articles for the occurrence of terms relevant to the CES classes of interest. The number of relevant terms and their incidence in each article is used as a proxy for the information content referring to the corresponding CES class. The search terms used for this operation were taken from the feature word-lists employed in the Flickr classification. This was done to ensure that the results refer to similar concepts as the Flickr data, which should make them more comparable. These lists had to be adapted to exclude compound words and terms that where too ambiguous for analysing a continuous text. For example, the term *historic* could refer to past events in a place's history and could not exclusively be used as an attribute for a monument that still stands. Therefore, this term and similar terms were excluded from the list. As *History* is a very common chapter in Wikipedia articles, this change is expected to influence the outcome. The *Historical Monuments* feature-list thus mainly contains terms relating to buildings and structures. The list for *Nature Appreciation* was largely left untouched as most of the terms unambiguously relate to the presence of wildlife or protected areas in the vicinity of the article. There were 83 search terms related to *Nature Appreciation* and 36 terms referring to *Historical Monuments*. This analysis was limited to the two classes *Historical Monuments* and *Nature Appreciation* as they already used features in the Flickr classification and it was expected to find information about these CES in the Wikipedia as well as the Flickr data.

After establishing the search terms, the text data from Wikipedia was accessed with a HTML-scraper implemented in the R library *rvest* (Wickham, 2021). This operation yielded the raw text contained in each paragraph of the 553 articles. These text files then needed to be tokenized by NLP software, this means that the text had to be broken down into single words. The library used for this was *spacyr*, which is a wrapper around the Python package *SpaCy* (Benoit & Matsuo, 2020). By tokenizing a text, the NLP-tool not only separates the text into single word but also tags its parts of speech and converts nouns, adjectives and verbs into their basic form in a step called lemmatisation. The occurrence of these lemmas were counted for each Wikipedia article. Then this list was searched with the terms in the modified feature lists. If a relevant term occurs in the article, the frequency of the lemmas in the text was added to the total for each article. The more individual words occur in the text and the more often they are used therein, the higher is the total count of words. This number is thus the proxy for relevance to the CES classes in question. This step was done for each class separately, resulting in two columns with word counts. The logic behind this is the greater the variety and incidence of the terms, more relevant points of interest are expected around this location, which is also expected to be reflected by the amount of Flickr posts in the vicinity. This operation is illustrated on the right hand side of the process diagram (Fig. 8). As the amount of Flickr objects and totals of terms in Wikipedia articles are not equitable by absolute numbers, the Wikipedia articles were split into two groups, more- and less-relevant (to CES classes). This split was conducted based on a statistical threshold, by which the articles in the 75th percentile (top 25%) of word sums were deemed relevant to the respective CES class (*Historical Monuments* and *Nature Appreciation*), whereas the lower three-quarters are considered less-relevant, as they may still contain relevant terms. The frequency of Flickr posts of each class around the Wiki locations shall be compared between the two groups, expecting to find significant differences between them.

5.3 Spatial Analysis

5.3.1 Flickr Classification

The Flickr data set is to be spatially analysed for clusters by calculating the Pearson's ρ correlation measure between each possible CES class combination. The variable compared was the incidence of posts in congruent square grid cells of 2 kilometre mesh-width overlaid over the whole study area. This was done for all three classes assessed and summarised in Table 7. While this particular analysis aims to provide insights into CES bundles, there will also be a qualitative survey of the data by spot checking a map visualisation in QGIS to check for fine and coarse scale spatial patterns in the classified data set.

5.3.2 Flickr Incidence vs. Wikipedia Articles

To compare spatial co-occurrence of CES related information between Wikipedia- and Flickr-data, it was decided to compare the number of Flickr posts in a circumference around Wiki articles between articles that are relevant and less relevant to CES classes, expecting to find significantly more Flickr posts around the relevant articles. For this, circular buffers with a radius of two kilometres were calculated for the Wikipedia article locations. For each of the two classes (*Historical Monuments* and *Nature Appreciation*), the buffers were spatially joined with the Flickr point data and the Flickr posts of each class were counted for each buffer. This buffer radius was chosen because it covers the study area from east to west, even if the location is at the very edge of the perimeter. It was also determined visually, that these buffer dimensions minimise overlap while still reaching a sufficient coverage of the study area. According to the division into more- and less-relevant articles for the respective CES class, the Wiki article locations were divided into two samples. As the Flickr post incidence was not normally distributed across the buffer zones, it was decided to conduct a Whitney-Mann-U test as the Student t-test is only suited for normally distributed data. Akin to the t-test, the the Whitney-Mann-U test can also determine if there is a significant difference between samples but does so by comparing rank-sums instead of the sample means (University of Zurich, 2022).

5.3.3 Visual Comparison Flickr vs. Wikipedia

To assess the patterns of CES occurrence along the study area, a visual comparison was conducted between the Wikipedia and Flickr data. A map was generated where the Flickr posts and Wikipedia articles where aggregated into grid squares of ten kilometre side length. While ten kilometres seems excessive considering the study area is only about two kilometres wide, this was necessary because smaller tiles would have been barely visible on the map covering the East Coast in its full length at this scale. This was done for Wikipedia and Flickr data for both classes each. A map for each class and data source, juxtaposing the aggregated results, was created that allows one to compare the spatial patterns along the whole study area at a single glance (Figures 12 and 13). For the same purpose, an interactive map was created in QGIS, displaying the data as points overlaid on an OSM base map. The grid can also be shown on the map to make the hot spots in the data more visible. With this data overlay, it is easy to zoom in and out and look at individual data points as well as view the data from a distance. This visualisation will be useful to discuss and interpret the results in the following chapters.

6 Results and Interpretation

6.1 Flickr Classification

With the RF model trained and validated using the annotated training data, the entire Flickr data set was codified the same way as the training data and run through the model. Posts with zero matches between user tags and word lists were excluded from the data set beforehand, which left 66'870 data points to be classified by the algorithm. This amounts to around 74% of the original 90'413 posts. The classified posts are not equally distributed across all classes. The highest share of posts has the class *Landscape Appreciation* with close to 40% of all posts, while *Historical Monuments* constitute around 7.5% and *Nature Appreciation* close to 12% of the data set (Table 5). The class *Other* at over 40% is most represented in the data set, but it is made up of the remaining classes in the classification scheme of Retka et al. (2019) and non-related posts as explained in the annotation rules (Fig. 7). This reveals the thematic priority of Flickr users being overwhelmingly interested in photographing impressions of the landscape in their posts. If this reflects a real difference in popularity of CES provisions in reality should not be concluded from this, because the interest of the Flickr community might be skewed to photographing landscapes. Even if the numbers are not representative of the real popularity of CES, the locations of clusters of posts of a certain class can still tell us where certain points of attraction are and how they relate to others in space.

Between the classes there are very clear semantic differences to be seen when comparing word frequency clouds compiled from all the tags used in posts of each class (Fig. 9). The font size is proportional to the frequency of the terms, which helps judge their prominence amongst other terms quite well. The tags usually refer to concepts that are thematically related to the class. For *Historical Monuments*, nouns referring to historical structures were most popular (e.g castle, church, ruin) whereas *Nature Appreciation* had a high representation of words that relate to wildlife and habitats (e.g. bird, nature, garden). This is reassuring, as the majority of these terms were thematically related to the respective classes. Very striking in the class *Other* was the high presence of nouns relating to vehicles of all kinds, most frequently trains and buses. Flickr also seems to be a very popular tool for sharing photos in the train- and bus-spotting community, which was noticed during annotation of the training set as well. Some words that refer to technical concepts such as camera brands, lens parameters and geo-tagging features are also well represented in the data set, but are regarded of limited use for distinguishing the CES classes referenced because they are a popular tag in any class. Excluded from the word-clouds are tags referring to geographic entities like seaside settlements or broader regions, as this was done during model training as well. Broad geographic terms (e.g. greatbritain, scotland, uk, europe) were usually the most common tags in each class and would obscure the words that are thematically relevant in the cloud. This indicates that users frequently use the tag field to communicate the location the image refers to.

Overlaid on a OSM base-map (similar to Fig. 16 and 17), the Flickr data was spot-checked for the occurrence of the classes where they were expected to be present all along the study area. There are indeed comprehensible clusters to be spotted in these POIs. For instance, there are accumulations of data points close to scenic spots like beaches and cliffs which were mostly classified as *Landscape Appreciation.* Close to castle ruins, old town squares and churches there are usually groups of *Historical Monuments* posts close by. Parks, nature reserves and islands were often described by posts referring to nature and classed as *Nature Appreciation.* This spot check confirmed that the classification worked as intended, even for *Historical Monuments*, which has a low sensitivity (true-positive) value (Table 3), the clusters were clearly visible. However, during the visual checks there were also some inaccuracies to be found. For instance, around the town of Seahouses there was a cluster of nature related posts, even though the data mostly did not refer to the *Nature Appreciation* class. In the training data set, the term *seahouses* was not considered a place name during the manual editing of the tag lists by mistake. Some posts in the training data set classed as *Nature Appreciation* happened to include the tag and thus the algorithm mistakenly associated the term with this class even when there were no other nature-related terms within the tags. A similar effect could be observed in Gardenstown, Scotland. The term *garden* was included in the feature list that used partial matching for the terms during the algorithm training. This emphasizes the importance of excluding place names from the training data set as they can lead to misleading associations of terms and CES provisions. Because some place names can not be easily recognized as such in a list, automated Named Entity Recognition with a suitable gazetteer should be used on the data set to annotate and exclude these terms. These instances were though not very common and in general, the spatial distribution of classes is comprehensible. While data quality can still be improved, the data shows significant spatial patterns that could potentially yield compelling results in the subsequent spatial analysis.

Table 5: Absolute and relative distribution of classified Flickr posts between the three selected CES classes an *Other*

	Landscape Appreciation	Nature Appreciation	Historical Monuments	Other	Total n
n	25'309	7973	5141	28'447	66'870
%	37.85	11.92	7.69	42.54	100

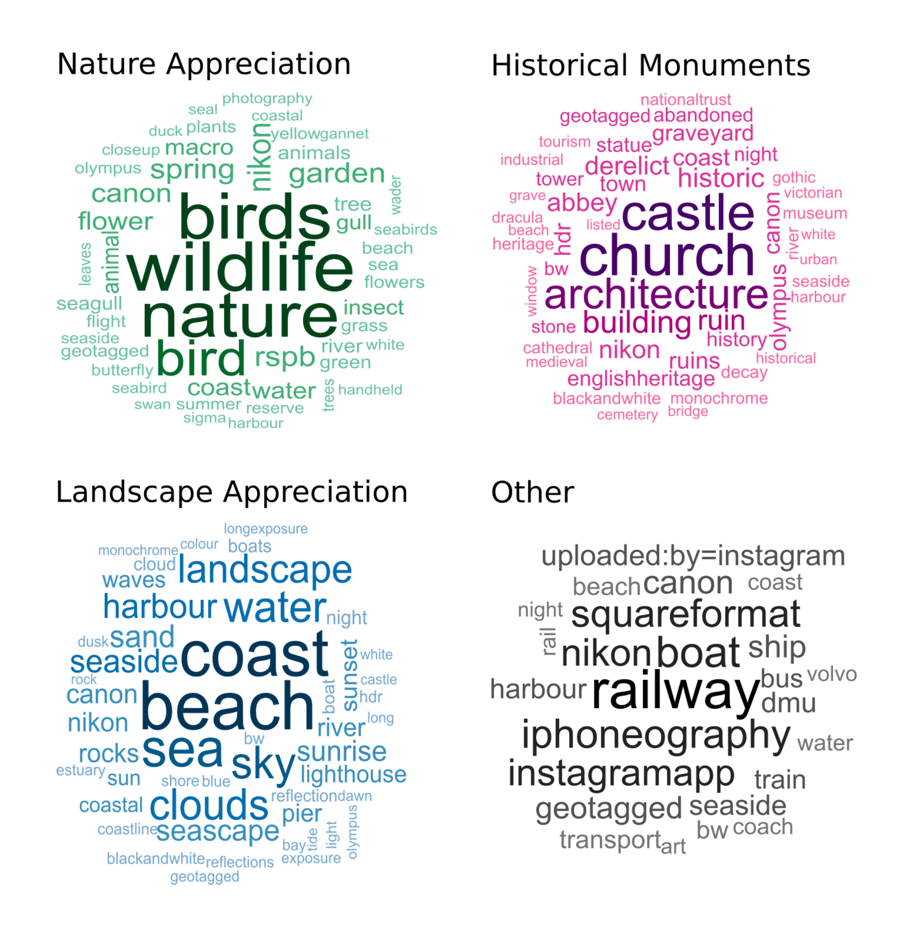


Figure 9: Wordclouds of the 50 most common tags in each CES class annotated in the Flickr classification process.

6.2 Wikipedia Classification

To determine the amount of information relating to the two CES selected in the Wikipedia data set, all the articles were searched for the terms that were compiled in the feature lists for *Nature Appreciation* and *Historical Monuments* in the Flickr classification process. The count of all the relevant terms in each article was summed up for each class as a proxy for information content. The articles in the 75th percentile of term count were deemed relevant whereas the others were considered less relevant to indicating CES. There are stark differences in the amount of relevant terms found in each of the articles, this is largely suspected to occur because there are also very large differences in the amount of characters used in Wiki articles. While a lot of articles are so called *stubs*, meaning that they were still missing content, only containing the most elementary information, for example in the article about *Pitcalnie* (Fig. 10), some articles have a high character count and provide a meticulously researched description of the place and its characteristics such as the article about the *Island of Stroma* (Fig. 11). For *Pitcalnie*, the only information provided are the various constituencies that the place is part of . It is clear that certain locations receive less attention than others. This does not necessarily indicate missing information, as there certainly are places where CES provisions are not very abundant. But as long as an article remains a *stub* there is no way to be certain. This is where an alternative data source like Flickr can become useful to confirm the presence or absence of CES in a certain location.

The spatial distribution of Wikipedia articles is closely connected to the settlements as a large part of the articles are usually related to towns or cities. This is similar to the spatial distribution of the Flickr posts, which also cluster around more populated areas. In the Wiki data, one can not find significant clusters at a small scale, as the amount of posts is many times lower compared to the Flickr data set. The typical distribution, with each town having only a single article related to it, casts doubts on the completeness of the data set. The *type* variable in the SPARQL query being set to *place* probably was not a definition that was broad enough to include other POIs that are also covered by georeferenced articles (Fig. 5). For example, the well known Forth Bridge and Bamburgh Castle which both have georeferenced articles, do not appear in the data set and are likely to contain information about historical monuments. This observation leads to the conclusion that the data set is most likely not complete and misses an unknown amount of posts that are potentially valuable for the analysis. Despite this, there is a relatively dense coverage of articles all along the coast (Fig. 6). While it is expected that there is at least some information about nature and historical structures in these articles, which mainly cover settlements and thus are vaguely spatially referenced (single coordinate), having all the objects of interest in the data set as well would be more spatially precise and increase the sample size. Even though the classification method was much simpler than the one used for classifying the Flickr posts, the Wikipedia data classification showed that the concepts relating to nature and places of historical value could be found in Wikipedia as well by using the same terminology used in the feature list for the Wikipedia classification. This indicates that there is information relating to the CES provision to be found in Wikipedia articles, at least for *Historical Monuments* and *Nature Appreciation*.

Pitcalnie

From Wikipedia, the free encyclopedia

Coordinates: 57.72311°N 3.99912°W

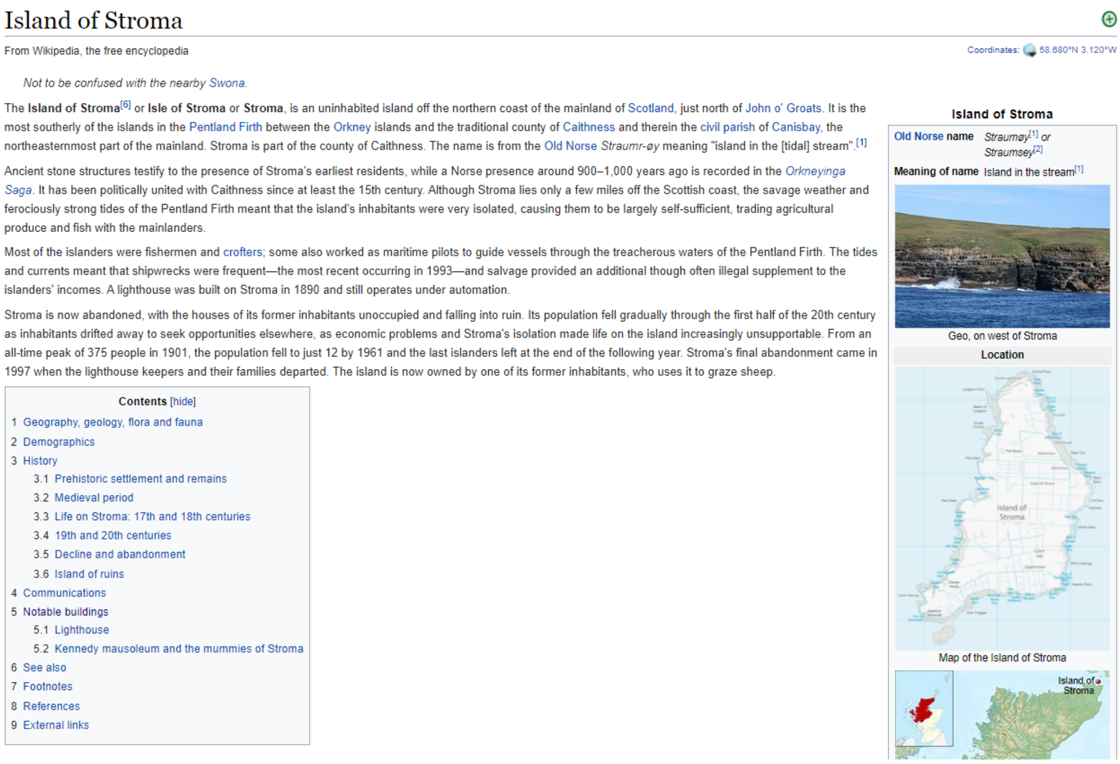
Pitcalnie (Scottish Gaelic: *Baile Chailnidh or Cuilt Eararaidh*) is a small hamlet on the north east corner of Nigg Bay in Ross-shire, Scottish Highlands and is in the Scottish council area of Highland.

This Highland location article is a stub. You can help Wikipedia by expanding it.

Pitcalnie
Scottish Gaelic: *Baile Chailnidh or Cuilt Eararaidh*

Location within the Highland council area

Figure 10: Header of the Wikipedia article about the hamlet of Pitcalnie: An example of a incomplete *stub* article as of 08-09-2022.



Island of Stroma

From Wikipedia, the free encyclopedia

Coordinates: 58.680°N 3.120°W

Not to be confused with the nearby Swona.

The **Island of Stroma**[6] or **Isle of Stroma** or **Stroma**, is an uninhabited island off the northern coast of the mainland of Scotland, just north of John o' Groats. It is the most southerly of the islands in the Pentland Firth between the Orkney islands and the traditional county of Caithness and therein the civil parish of Canisbay, the northeasternmost part of the mainland. Stroma is part of the county of Caithness. The name is from the Old Norse *Straumr-øy* meaning "island in the [tidal] stream".[1]

Ancient stone structures testify to the presence of Stroma's earliest residents, while a Norse presence around 900–1,000 years ago is recorded in the *Orkneyinga Saga*. It has been politically united with Caithness since at least the 15th century. Although Stroma lies only a few miles off the Scottish coast, the savage weather and ferociously strong tides of the Pentland Firth meant that the island's inhabitants were very isolated, causing them to be largely self-sufficient, trading agricultural produce and fish with the mainlanders.

Most of the islanders were fishermen and crofters; some also worked as maritime pilots to guide vessels through the treacherous waters of the Pentland Firth. The tides and currents meant that shipwrecks were frequent—the most recent occurring in 1993—and salvage provided an additional though often illegal supplement to the islanders' incomes. A lighthouse was built on Stroma in 1890 and still operates under automation.

Stroma is now abandoned, with the houses of its former inhabitants unoccupied and falling into ruin. Its population fell gradually through the first half of the 20th century as inhabitants drifted away to seek opportunities elsewhere, as economic problems and Stroma's isolation made life on the island increasingly unsupportable. From an all-time peak of 375 people in 1901, the population fell to just 12 by 1961 and the last islanders left at the end of the following year. Stroma's final abandonment came in 1997 when the lighthouse keepers and their families departed. The island is now owned by one of its former inhabitants, who uses it to graze sheep.

Contents [hide]

Island of Stroma

Old Norse name	*Straumøy*[1] or *Straumsey*[2]
Meaning of name	Island in the stream[1]

Geo, on west of Stroma

Location

Map of the Island of Stroma

Figure 11: Header and content table of the Wikipedia article about the Island of Stroma: An example of an extensive article as of 08-09-2022.

6.3 Spatial Analysis

6.3.1 Wikipedia vs. Flickr Visual Comparison

In a primarily visual comparison, the commonalities and differences between the Flickr and Wikipedia have been assessed. For this purpose, two comparison maps have been created putting the incidence of posts and articles relating to the CES classes *Nature Appreciation* and *Historical Monuments* side by side with the counts aggregated into ten-kilometre grid cells (Fig. 12 and 13). The grids were also added as a layer over an OSM base map for a more detailed cluster-scaled comparison.

For both classes, it can be determined that the distribution of posts and articles is strongly connected to populated areas with larger cities usually being associated with higher counts of data points. The areas around Edinburgh, Newcastle and the Thames Estuary are good example for a high count, whereas northern Scotland has a rather sparse occurrence of both Wikipedia articles and Flickr posts. By looking at the data more closely, the class *Historical Monuments* has a good overlap in the southern part. Where there are usually clusters of Flickr posts where relevant Wiki articles are present too. At the Norfolk coast however, there seems to be more information about historical structures in Wikipedia articles than in Flickr. Similarly, the North of Scotland looks to have more clusters of Wikipedia articles, while there is not a similar concentration of Flickr posts. There are also instances where there are clusters of Flickr posts but no matching information in the nearby Wikipedia article. This is the case for the city of Aberdeen for instance. This probably occurred because the main article for the city was not included in the data set as the article coordinate was not within one kilometre of the shore. This is one of the downsides of the geographic reference being limited to a single point coordinate instead of footprints, which would be more representative of areal entities such as cities. The posts of the *Nature Appreciation* class seemed to roughly match with the Wikipedia article distribution as well. In comparison to the the other class, northern Scotland is matching slightly better but still seems underrepresented by the Flickr data. There are also a few instances where Wikipedia articles are missing from the data set. For instance, there are virtually no articles relating to the various nature reserves in the data set. The RSPB (Royal Society for the Protection of Birds) Minsmere in Suffolk contains quite a lot of Flickr posts relating to nature but the corresponding Wikipedia article does not appear in the data set. Upon reading the article, there was a lot of relevant information concerning nature and habitats to be found therein. While this means that a lot of information is missing, the overlapping pattern in the existing data indicates that there is at least some information about close natural areas contained in articles that are nearby, such as the towns where these areas belong to. Similarly to the *Historical Monument* data, there is also a concentration of nature related posts in the vicinity of cities. This is assumed to be a function of the available green space (gardens, parks, riversides) and the general accessibility of these sites by a large amount of people which is also visible in the Flickr data.

While there is some overlap to be found in the spatial distribution of CES related content in Wikipedia and Flickr, there are still some uncertainties involved. The main problem is the large discrepancy between the sample sizes of Flickr posts and the Wikipedia articles across the research area. Doing a correlation analysis using Pearson's ρ of the article/post count between grid squares, would help to compare the correlation between the two data sources in a quantifiable manner. But as the density of Wiki articles is so much lower, an error of just one unit could have a large effect on a regression, where the range of values is only five for the Wiki articles. There is every reason to believe that the data is prone to errors. As many articles relate to a rather large areas but are only represented by a single point coordinate, there is a likelihood for a mismatch with the grid cell the related Flickr data points are located in. To decrease this spatial ambiguity, it would require a larger sample and/or document footprints to define the spatial relationships of the Wikipedia articles more precisely. As already established in the analysis of the Wikipedia data set, there is possibly a large amount of data missing that should be taken into account for further quantitative studies. Nevertheless, this visual analysis provides some clues that there is indeed a spatial relationship between Flickr and Wikipedia data that can contribute to locate CES provision along the coast. However, additional steps must be taken to improve data availability and geographic accuracy in future studies.

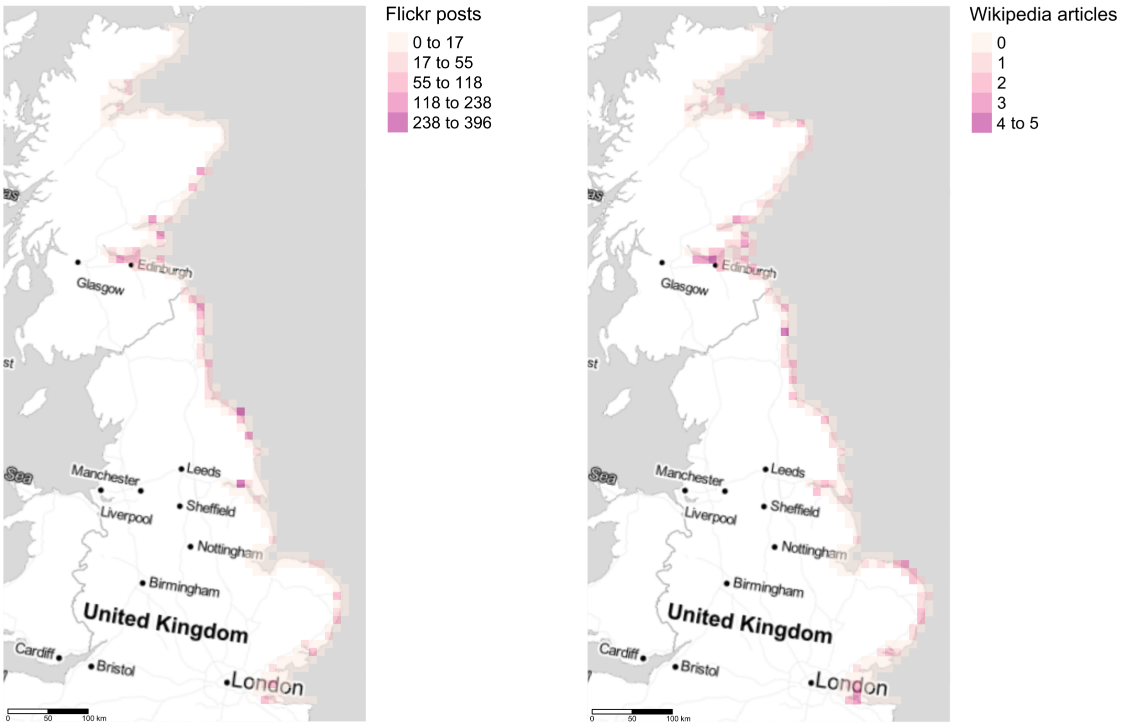


Figure 12: Post/Article incidence relating to the class *Historical Monuments* of Flickr (left) and Wikipedia (right), aggregated into 10 kilomtre grid cells.

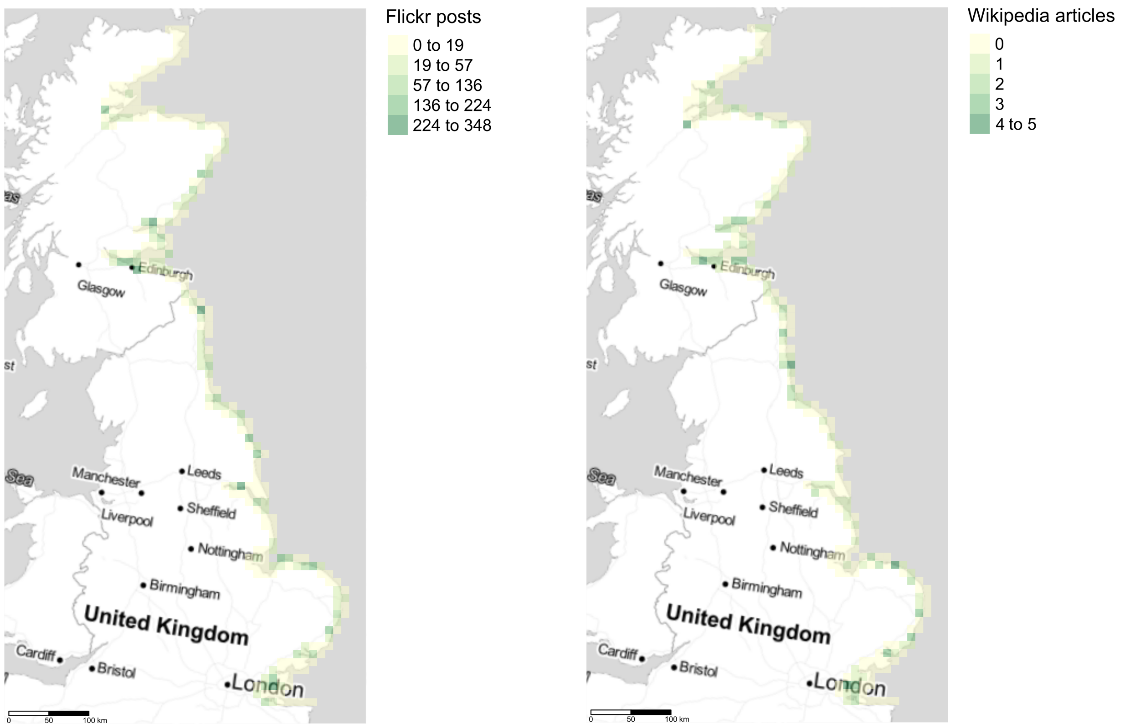


Figure 13: Post/Article incidence relating to the class *Nature Appreciation* of Flickr (left) and Wikipedia (right), aggregated into 10 kilomtre grid cells.

6.3.2 Wilcoxon-Rank-Sum Test

For each CES class a Wilcoxon-Rank-Sum test was conducted between the sample of Flickr posts in a two kilometre radius around a Wikipedia article locations that are relevant to the class (top quartile of most frequent keyword occurrence in the article) and ones that are less relevant (bottom 75%). There were 161 articles in the sample relevant to *Nature Appreciation* and 143 for *Historical Monuments* with the total number of articles at 553 (Table 6). It is possible that there is some overlap between the samples as one Wikipedia article can be relevant for both classes at the same time. This statistical test was chosen because it was not certain if the independent variables are normally distributed. For *Nature Appreciation* the null-hypothesis could be rejected, as there was a significant difference in post quantity at $\alpha < 0.05$ (p = 0.0005). For the class *Historical Monuments* the difference was also significant (p = $4.99e^{-12}$). There are thus significantly more Flickr posts around a geolocated Wikipedia article that is relevant to the respective CES class. This is true for both classes assessed here. The results are visualised as a violin plot for each CES class in (Fig. 15) for *Historical Monuments* and in (Fig. 14) for the class *Nature Appreciation*. The relevant figures for for the tests are listed in Table 6.

The correlation indicates that a high incidence of Flickr posts spatially coincides with a high information density in Wikipedia articles for articles relevant for the two classes assessed. This means that there is a spatial correlation between the information density of each data source. Thus, Wikipedia and Flickr tend to be spatially congruent as popular CES on Flickr are usually more likely to be also documented in Wikipedia. However, one should be careful to conclude that there are fewer CES provisions where there generally is a low incidence of both Wikipedia articles and Flickr posting. While agreement between two independent sources should be seen a promising sign of real CES provisions or the lack thereof, there could be bias at play which could lead to both sources overlooking certain locations, which are a popular hot spot in reality. This could be the case here too, as the Wikipedia data set is skewed to include populated places (see section 6.2), which usually correlates with more Flickr posts (Alivand & Hochmair, 2017). The outcome of this analysis should thus be considered with these caveats in mind. In fact, it does not explain any casualty between the variables compared and only shows that they are correlated.

Comparing the violin-graphs and the corresponding table (Table 6) against each other, one can see that for *Nature Appreciation* the median value of Flickr post incidence is twice as high in the *relevant* subset as in the *less-relevant one* (14 vs.

7). For the *Historical Monuments* this discrepancy is even more pronounced at a four-fold increase (12 vs. 3) (Table 6). Being close to a Wiki article relating to Historical Monuments has thus a significantly bigger effect on the incidence of Flickr posts relating to the same class as being close to a Wiki article relating to *Nature Appreciation*. The results for *Nature Appreciation* (Fig. 14) also have a higher variance which is also asymmetrical towards the upper end. The variance for *Historical Monuments* (Fig. 15) on the other is smaller and less skewed. Visually the distributions are matching in both graphs, the less-relevant violin-graphs both have a "flying saucer" shape with a distinctive spike at the median. The shape of the graphs for the *relevant* articles are more elongated and have smoother and less pronounced peaks and a visible neck that stretches out and slowly fades until only the outliers remain. The *less-relevant* graphs however, diminish very soon after their very pronounced peak. These graphs emphasize the difference that can be seen in numbers (Table 6) that there is a clear difference in the frequency and variance of posts in the different types of sites compared.

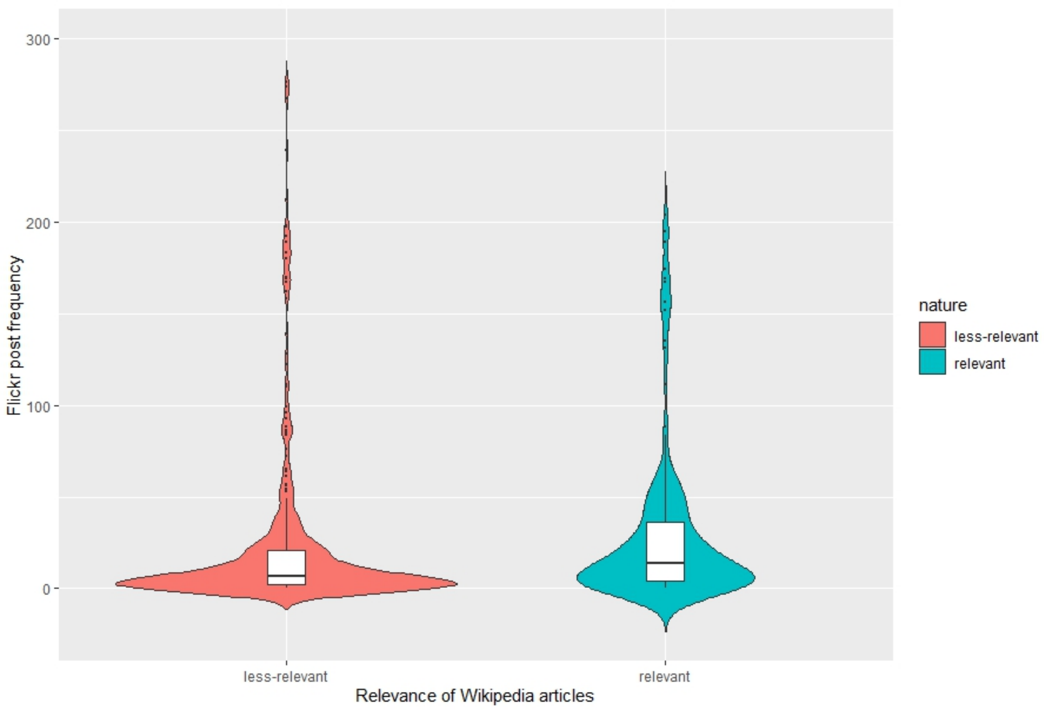


Figure 14: Violin plot comparing the quantity of Nature Appreciation Flickr posts 2 kilometres around a Wikipedia articles relevant to the class, against the count around less-relevant articles.

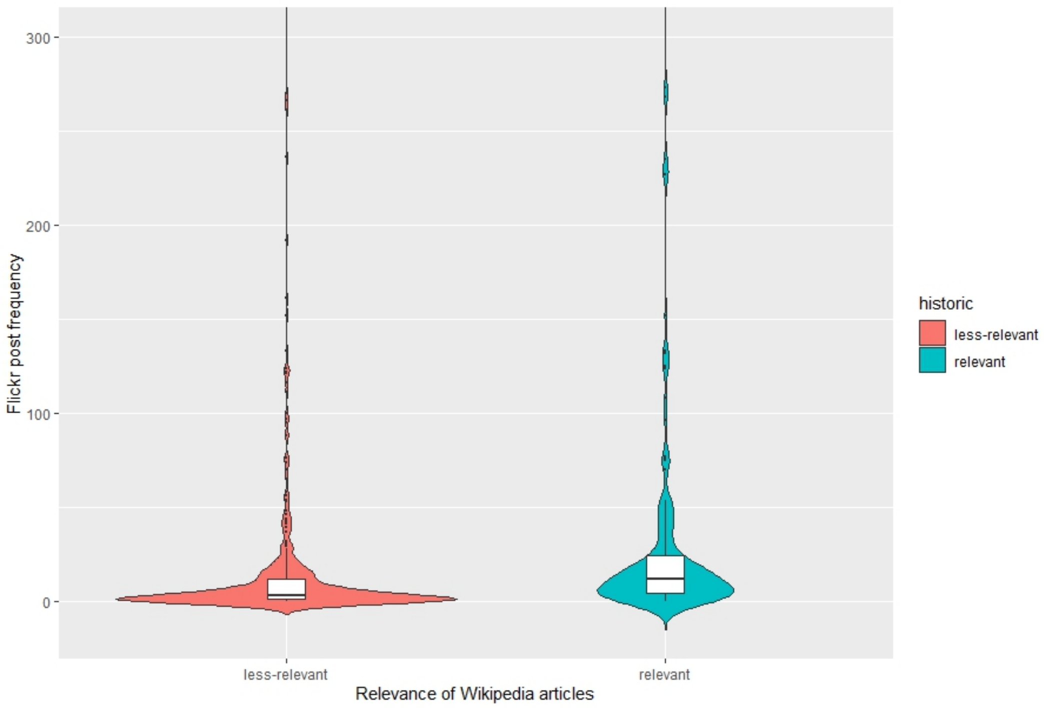


Figure 15: Violin plot comparing the quantity of Historical Monuments Flickr posts 2 kilometres around a Wikipedia article relevant to the class, against the count around less-relevant articles.

Table 6: Statistics for the median count of Flickr posts in a two kilometre radius around geolocated Wikipedia articles between posts with higher and lower topical relevance. n = 553, IQR = Inter-Quartile-Range.

	Nature Appreciation		Historical Monuments	
Relevance	relevant	less-relevant	relevant	less-relevant
n	161	392	143	410
Median	14	7	12	3
IQR	32	19	20	11

6.3.3 Pearson's ρ Cross-Correlation

This analysis was conducted to shed light on the presence of CES *bundles*, meaning the co-occurrence of CES provisions that could point towards possible synergies or trade-offs between them. This has been previously theorized and assessed in literature but remains a niche research subject (Retka et al., 2019; Rodrigues Garcia et al., 2017). The Pearson's ρ indicates how much the cell-counts of Flickr posts in a two-kilometre grid correlate between every possible combination of classes across the study area. A value of 1 indicates a perfect (and positive) correlation and a value of 0 signifies no correlation at all. This has been tested for all three assessed CES classes in this study and the results thereof are found in (Table 7). The posts of *Nature Appreciation* only correlated weakly with the other classes at a ρ value of around 0.5. Only *Landscape Appreciation* and *Historical Monuments* correlate more with a ρ of about 0.7 which is not very strong, but still significant. The incidence of these classes thus correlates better than with *Nature Appreciation* which could indicate that they form bundles across the landscape more often. This correlation is oftentimes visible on the map, as *Nature Appreciation* clusters are often spatially disjoint from the other classes, while the other two CES classes often occur together close to highly accessible places with scenic values (Fig. 16). The assumption here is that nature enthusiasts would seek out places that are undisturbed by human activity which is abundant in places which are considered valuable both for their scenery and historic heritage. During annotation of the training data, it was observed that *Historical Monuments* and *Landscape Appreciation* were sometimes harder to conceptually separate because historical structures were often part of landscape scaled images where the perceived aesthetic appeal of a monument contributed to the scenic value of the picture. This conceptual closeness could also explain why *Historical Monuments* had a comparably low accuracy during the validation of the RF algorithm (Table 3). There was significant miss-allocation from *Historical Monuments* to *Landscape Appreciation* (Table 4). This analysis indicates that bundles of CES classes are a common occurrence in the landscape which are worth investigating more thoroughly, possibly involving a more diverse set of CES classes.

Table 7: Pearson's ρ cross-correlation of Flickr post incidence in a 2 km square grid.

	Nature Appreciation	Historical Monuments	Landscape Appreciation
Nature Appreciation	1.000	-	-
Historical Monuments	0.521	1.000	-
Landscape Appreciation	0.551	0.722	1.000

6.3.4 Spatial Patterns in Flickr Data

For analysing the spatial distribution patterns was also a visual analysis and description of the classified data facilitated by a visualisation in QGIS on top of a OSM base map with the Flickr post location displayed as dots coloured according to the CES class assigned by the RF algorithm. For visual reference to the readers of this study, small scale maps of selected exemplary sites, namely Dundee and Bamburgh Castle, have been created as well (Fig. 16 and 17). For analysis, the map was spot checked all along the coast and the most conspicuous and visible patterns were noted in the process.

The patterns and clusters are assessed on a fine as well as a coarse scale. Fine scale relates to patterns within settlements and smaller regions, whereas the coarse scale analysis emphasizes general patterns true for the entire coastline. From afar it becomes apparent that posts cluster at certain locations. These locations are mainly populated places such as cities, villages and hamlets at the sea side. This observation is generally true for all three classes, however the *Nature Appreciation* also appears quite often in locations off the beaten track, presumably because people interested in it specifically seek out location where nature is abundant and mostly undisturbed by human civilisation. This can be observed in areas that have many natural habitats such as the Norfolk and Suffolk coast in the Southeast. But clusters are also to be found within cities as well, usually in green spaces like parks and along water bodies. Where the clustering is most noticeable is for the class *Historical Monuments*, which are very often very close or within settlements. Posts appreciating landscapes occur in a clustered manner too, but are usually less concentrated, as the objects that are mainly referred to (beaches, cliffs etc.) are spread out over a larger area compared to historically valued sites that are often spatially limited (churches, castles). This clustering pattern becomes more visible as one proceeds northwards beyond the Firth of Forth: The clusters are further apart close to the sparse settlements along the coast. The low population density and the more challenging topography of the Scottish coast are suspected to be the main contributors here. In stark contrast to this, along densely populated urban areas (e.g. Newcastle and Edinburgh) the clusters are visibly closer and almost form an unbroken line when viewed from afar. The main take away from the large scale observation is that there seems to be a clear connection between accessibility of a site and the incidence of Flickr posts as the incidence of posts becomes visibly lower the further it is distanced from human infrastructure.

Zooming in, the typical fine scale pattern become visible on the map. Most noticeable here is that many posts are very close to the coastline. While the buffer the posts were sampled in reaches one kilometre inland, most posts where within 100-500 metres of the coast. Only within larger towns, clusters could be spotted further away as there are usually other significant POIs distributed across the city further away from the shore. This was most conspicuous for the posts classified as *Landscape Appreciation*, as these are typically located right *at* the coastline as can be seen in Figure 16 as well as in 17. This is no surprise, as a large amount photos viewed during annotation and assigned to this class were snapshots of beaches, harbours and sunrises/sunsets at the horizon, usually taken directly from the shore. The geographic closeness of nature related posts to the coast is suspected because the abundance of birds is higher, which is a common motif that is often found close to water. Some exceptions to this are parks and other green spaces within cities, an example of this is the cluster at the botanical gardens west of Dundee (Fig. 17) Small uninhabited islands off the coast are also a common hot spot for nature enthusiast content as these islands are a habitat for marine birds and mammals, such as the Farne Islands right across Bamburgh Castle (Fig. 16).

The typical small scale patterns observed for the *Historical Monument* class are clusters that are confined to a POI and do not spread out across a larger area as the landscape related posts do. This can also be seen on the map around Bamburgh Castle where the castle on the main land and the sights on Lindisfarne Island (historic town, castle) are located. Of note is also that these objects are surrounded by *Landscape Appreciation* posts, possibly due to the contribution of historic sites to the landscape aesthetic. The Bamburgh map (Fig. 16) provides an example of how these two classes work together synergistically. While there are other beaches nearby which might have an aesthetic value and are also quite accessible, there are barely any posts relating to it on the map. This can be observed on Ross Beach which is located between Bamburgh castle and Lindisfarne. A similar connection can be observed in Whitby, where a ruin of an abbey is in a similarly exposed location which also is used as a backdrop for landscape photography as well. Clusters of landscape related posts also occur without a co-occurrence with historically relevant posts but the posts of the *Historical Monuments* class are on their own seldomly. This observation goes in line with the analysis concerning cross-correlation, where the incidence of posts between these two classes is higher than in any other cross-comparison (Table 7). Places where all three surveyed classes cluster together are typically larger towns and cities where different types of POIs are close together (e.g. parks, historic structures and view points) that provide different kinds of CES services and are also highly accessible. This can be observed in places like Dundee and Edinburgh as well as cities in the Tyne and Wear metropolitan area (e.g. Newcastle).

This visual analysis shows that the data generated by the RF algorithm shows distinct spatial patterns that differ between the various classes assessed and suggests that there are casual relationships between the environment and the occurrence of clusters of classified data. The subordinate finding is that there is a strong connection between post incidence and the accessibility of the site to the user. There were thus more post were more people could conveniently reach it, which is true for any CES class. Furthermore, it seems that the coastal area that contains the majority of posts usually does not extend 1000 metres inland and most posts are very close to the actual coastline.

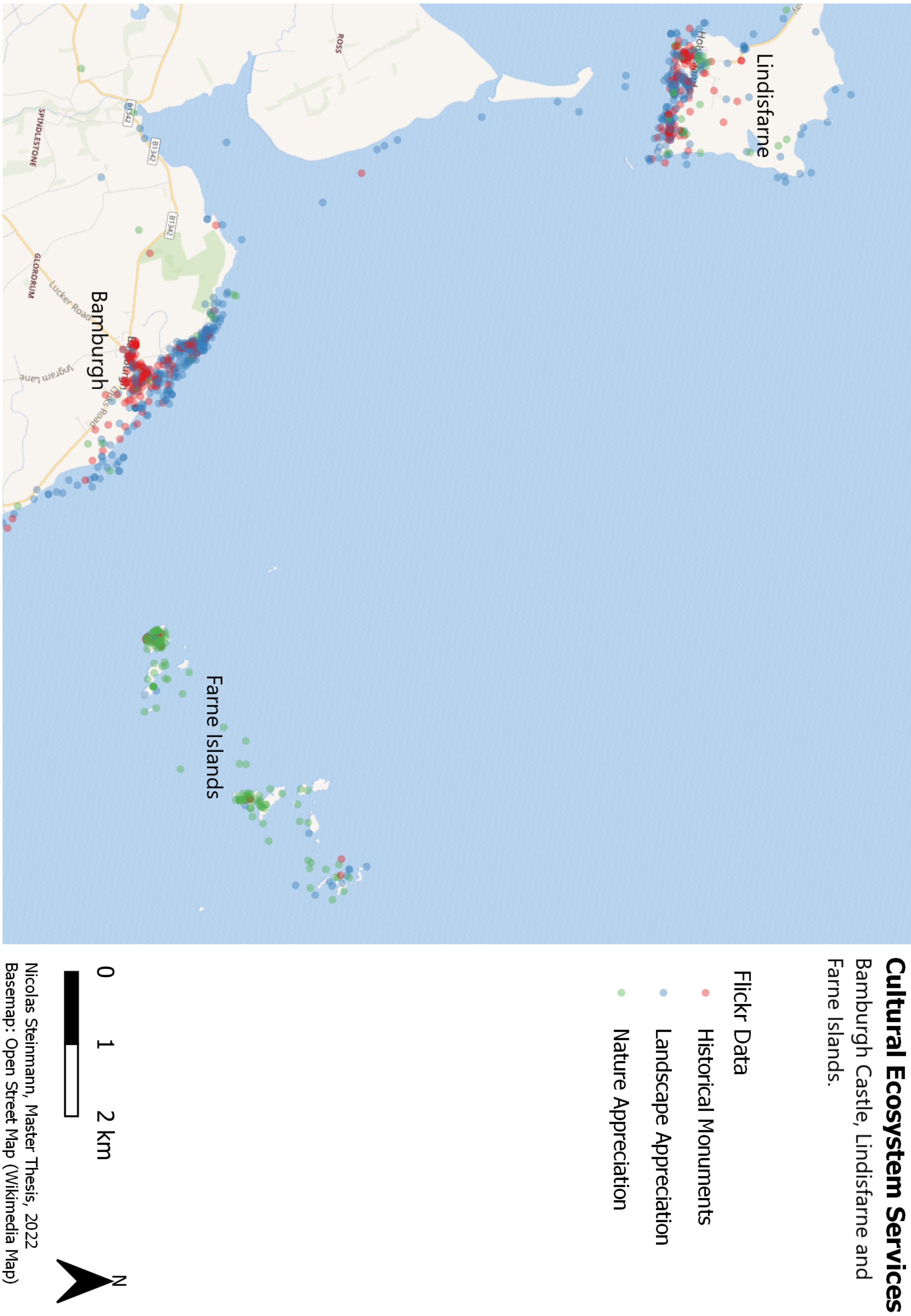


Figure 16: Fine scale map of the coastal area around Bamburgh.

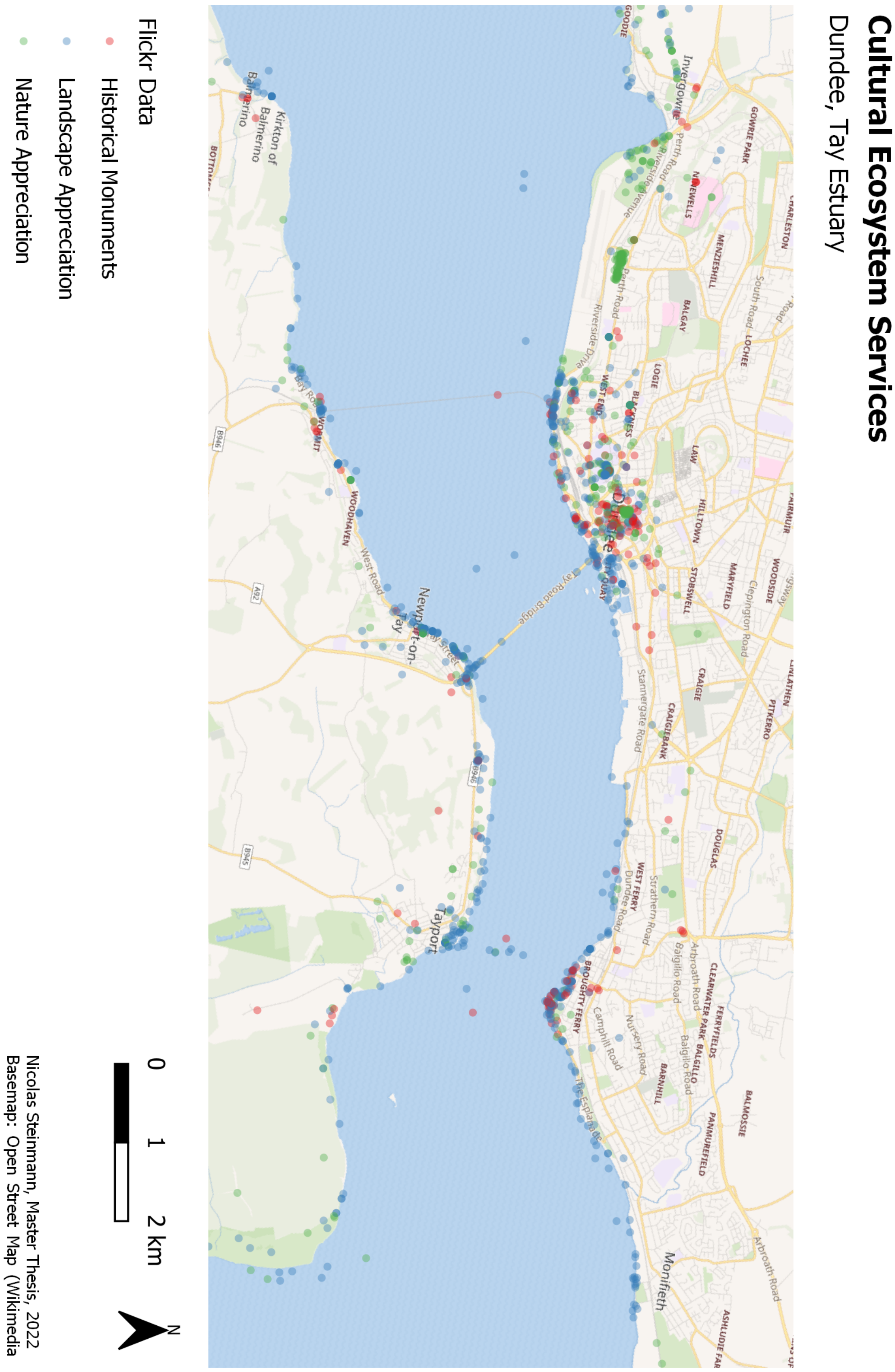


Figure 17: Fine scale map of the area around Dundee at the Tay Estuary.

7 Discussion

7.1 Research Question 1.1

How well does the CES classification of Flickr posts using the RF method work?

RF Model Accuracy

In order to answer this question, a set of around 70'000 individual Flickr posts were classified by a Random Forest (RF) algorithm, that was trained to recognize the CES class referenced in the post based on the user-defined tags provided in its metadata. For the training set, 4'000 individual posts were viewed and the best matching CES class was manually assigned. Some of the most numerous classes in the classification scheme were classified in the end: *Landscape Appreciation*, *Nature Appreciation* and *Historical Monuments*. All remaining classes and posts that did not relate to CES were classed as *Other*. The classification reached a fairly accurate result at a sensitivity of around 75%, ranging between 47% and 81% for the individual classes (Tables 2 and 3). The data set did pass visual inspection, displaying meaningful and geographically accurate distribution of CES clusters (see Section 6.3.4). The classification did exclude image data (e.g. machine-tags) on purpose, as there is a significant cost connected to using image tagging services and to demonstrate that the meta-data provided by the user on its own, is enough to allow for an accurate prediction. The usage of meta-data also allows one to better assume the user's perspective and represent their intent more genuinely (Depietri et al., 2021; Hale et al., 2019). The inclusion of feature lists as used by Chesnokova and Purves (2018) did help to improve the accuracy only by around 5%, they were still included to get the best possible result and reach more balanced sensitivity values between the classes. Editing the word list to get rid of the very prominent toponyms in the data also had a minor positive effect on the sensitivity values. The reasons for which are not completely understood, but it was observed in the data that the algorithm would occasionally strongly associate toponyms with a certain CES class and thus mismatch certain posts that have been tagged with said place name. Should the amount of data provided from the tags alone not suffice, there is even more textual data accessible through the Flickr API such as the post title and description (Fig. 1). Addition of image content tags would also be an option, which is a tried method in related literature (Cao et al., 2022; Santos Vieira et al., 2021). If that is not sufficient for the task at hand, the algorithm could be easily adapted to include data from further sources different various data types.

While the employment of the method has worked in general and produced a result that included plausible spatial patterns, there are still some shortcomings that should be addressed to improve the analysis of CES based on user-generated content like Flickr posts. To be discussed first is the fundamental paradigm of using conceptually delimited classes to characterise the CES represented in a post. There are various schemes to delimit the different classes of CES from each other with varying applications and fields of use (Ahtiainen et al., 2019; Cheng et al., 2019; Havinga et al., 2020). For this thesis, a classification scheme was chosen that was adapted to the marine/coastal environment and had also been used in a similar study before (Retka et al., 2019; Santos Vieira et al., 2021). While it worked reasonably well to identify the CES in this data set, during the annotation of the training data, there were some instances where multiple CES categories were equally present in a post. A detailed annotation guide was created to establish a number of rules to enforce consistency (Fig. 7) but this could not cover every imaginable case. The confusion matrix for the training data (Table 4) also indicates that the RF algorithm sometimes had trouble differentiating between classes. For example, the algorithm often had difficulties separating *Historical Monuments* from *Landscape Appreciation* and there is an even more significant mismatch between *Historical Monuments* and *Other* which hints that there is potential overlap with all the remaining classes which were grouped under *Other*. This apparent overlap or *fuzzyness* between the different concepts is in a stark contrast to the method applied here, where each data point is assigned only a single CES class. Assigning multiple CES classes to a single post is imaginable and could be achieved by using an algorithm that can give an estimate of how likely it is that a post represents any of the classes. However, the *ranger* R-package used here, does not return uncertainty measures for individual data points (Wright & Ziegler, 2017). As there are very distinct spatial clusters in the data, one could aggregate the data into these clusters (e.g. by using k-means spatial clustering) and thus determine places that provide multiple CES at once by analysing the relative occurrence of each class within the cluster. One could then identify common concepts in the tags that share importance for multiple classes at (e.g. landscape features) to assess in more depth. The *fuzziness* of the concept in reality does not have to contradict with the rather straightforward definition of CES as conceptually delimited. This concept has lots of similarities with CES bundles, which will be discussed in more detail later on.

To reduce the complexity of this study, the amount of CES classes to be assessed was limited. The three classes were chosen based on their relative prominence in a trial study of post annotation and the expectation that these CES would be mentioned in Wikipedia as well. In the classification scheme used, there remain another six classes that were grouped under *Other* in this study (Retka et al., 2019). These other classes were usually in so rare in the data that it was feared that the low sample would not lead to an accurate prediction. This was the case for the class *Social Recreation*, which was removed after it had a very low true-positive rate (see Chapter 5.1). This means that the training data set would have needed to be much larger to get an appropriate sample of these classes and hence would require more time to classify. However, this does not mean that this is impossible to achieve, one study reached very high accuracy classifying a data set with 6 CES classes. Though CNN were used with image-tags instead of user-defined tags as data (Mouttaki & Erraiss, 2022). Another solution would be to use different data sources for the various CES classes where these CES classes are better represented. Havinga et al. (2020) suggest the data sources most suitable for each class. For recreational activities for example, it is advised to use geolocated data from activity tracking software such as *Strava*, whereas bird-watching sites like *eBird* offer a wealth of information about wildlife. This would certainly provide a bigger sample size than just relying on Flickr alone.

Data Quality

When working with user-generated data it is especially important to think about the quality of the data that is used for any analysis. Because this is crowd sourced data, there is no authoritative quality control to ensure that the information is acccurate. For assessing VGI data there are multiple aspects of data quality. For this study it will be limited to *Positional Accuracy*, *Thematic Accuracy* and *Completeness* of the Flickr data set (Fonte et al., 2017). This will not be a full quantitative assessment of the data, but contributing factors will be discussed that might be present in the data based on observation and additional literature. Positional accuracy is definitely a relevant topic as there were some conspicuous signs observed during visual analysis. For example, many data points are located out at sea (Fig. 16). Upon inspecting the photographs, most were taken onshore. There are essentially two ways how geo-location works on Flickr, one way is by GPS coordinates recorded in the meta data of the camera, and the other is self-assigned referencing where the user puts a marker on a map (Hauff, 2013). GPS derived coordinates should be fairly accurate with an error between 10-20 metres (Zandbergen & Barbeau, 2011), self-assigned coordinates depend on the users judgement on how accurate the location needs to be. The study of Hauff (2013) found out that the positional accuracy of photos picturing tourist sights can vary considerably depending on how popular the location is. Popular locations reached accuracy on a par with GPS derived coordinates, while unpopular sights had an error of 47 to 167 metres. For self-assigned coordinates there is also the problem of *what* the location is perceived as: Is it either the location a picture was taken from, or the location of the main motif in a picture (e.g. a castle or bridge)? (Zielstra & Hochmair, 2013). Depending on the research question, this *inaccuracy* can actually be useful, as it adds in another layer of perception similar to the tagging of the post. If this accuracy is sufficient, also depends on the task at hand. For this study, which has a rather coarse scale, it is definitely accurate enough, but if one would do such an analysis on the scale of city blocks, these accuracy numbers could become an issue.

Thematic Accuracy is also quite important, as it determines how valid the information provided by the data is. As this data is usually collected on the internet, there is a limited ability to ensure that the users are competent to convey accurate information and act in good faith (Fonte et al., 2017). Fonte et al. (2017) propose indicators to evaluate data quality of VGI data: These are *Data-based*, *Demographic and socio-economic* and *Contributor* indicators. For assessing the thematic accuracy of the Flickr data set, a data-based indicator could be used, this can be achieved by comparing the results against another data set. In the case of this study, this was attempted with Wikipedia data. There was some thematical overlap to be seen (see Chapter 6). Because there were other data quality issues in the Wikipedia as well, this has to be taken with caution. These issues will be discussed in chapter 7.2.

Another valid concern is the *Completeness* of the data. VGI data sets often contain inherent biases that over-represent popular locations, or locations that are especially valuable to the individual user (Fonte et al., 2017). There is typically a correlation of Flickr post density with population density (Li et al., 2013). Hecht and Stephens (2014) and Lopez et al. (2019) state that urban perspectives are better represented per capita than rural ones. There are thus more posts, users and tags close to larger population centres. Furthermore, popular touristic places are usually more present in social media data as well (Depietri et al., 2021). For the purpose of this study, this heterogeneous spatial distribution is less of an issue. By concept, CES are created by the interaction of humans with nature (Fish et al., 2016). A higher incidence of data points is thus expected where there is also an increased human presence and hence potentially higher CES provision (Havinga et al., 2020). However, the lower user adoption rate and tag frequency in rural areas as well as the bias to more touristic places does influence data quality. The concentration of data in cities and around touristic landmarks could be readily observed in the Flickr data, for example at Bamburgh Castle and all the larger population centres in the study area (Fig. 12 and 16).

The dependence on the perspective of individual users makes it necessary to determine the demographic figures of a user base as *contributor indicators* to find out more about *who* prefers to use Flickr to post content. This is important as social media sources usually only represent a specific subset of the population (Havinga et al., 2020). As Flickr meta data does not include demographic data (e.g. age, gender, location), determining these biases is not a trivial task (Depietri et al., 2021). It has been established that the dominant age group posting on Flickr is between 35-44 years old and the gender distribution is biased as well, with 79% of users identifying as male (Hartmann et al., 2022). The Flickr user community thus is not representative of the whole society, as perspectives of the female demographic and certain age groups are less considered. To make these biases visible, it may also be useful to infer demographic data about the individual users from their content or meta data, which also raises privacy concerns (Lopez et al., 2019).

Another option to validate the thematic accuracy would be to introduce authoritative data sets to compare against and assess for differences in spatial distribution (Fonte et al., 2017). For this study these data-based indicators could be official listings of historical structures or boundaries of nature reserves. Including different VGI data sources that represent certain CES better than others, can also help to achieve a more extensive set of perspectives (Havinga et al., 2020). To correct for all these biases in the Flickr data set would be a complex task and would warrant a thesis on its own. The results should thus be viewed with these considerations in mind. However, one bias has been addressed at the very beginning. Typical for Flickr data, a minority of users contribute the majority of posts (Purves et al., 2011) and to equalize this, the data sets were PUD-sampled (see Chapter 4), which means that only one post per day for each user was kept in the data-set by random sampling. This did not lead to a completely balanced data set and there is still a minority of 5% posting almost half of the content, down from 70% in the unfiltered data set. Not excluding prolific users would massively over-represent their perspective and make the ones of more causal users count less towards the overall picture of CES distribution. During annotation it became clear that very active users often post large collections of photos from the same place or event taken in one day. Thus increasing the quantity of data does not necessarily lead to a better representation of the population.

7.2 Research Question 1.2

Could Wikipedia complement Flickr data for CES mapping?

To find out about the combined value of Wikipedia and Flickr data for mapping CES more confidently, Wikipedia meta data was accessed using the DBpedia data base. Running a spatial query to include locations in two-sided buffer one kilometer wide on each side, yielded a list of 553 entries to be included in the analysis. The articles were accessed using HTML scraper software via the article link and the retrieved text files tokenized using NLP methods. The documents were then searched for the terms relating to *Historical Monuments* and *Nature Appreciation* that were already used in the training of the RF algorithm for classifying the Flickr posts. With this it could be proven, that parts of the terminology used to classify Flickr posts can be useful to search Wikipedia for relevant information as well. For both classes, the relevance of the documents (Wikipedia articles) was quantified by counting the occurrence of the search terms and adding them up. Classified as *relevant* were the top 25% of articles with the highest sum of keyword occurrence. The Flickr post incidence in a two kilometre radius around each article was compared between relevant and less-relevant articles and tested for significant differences. A visual comparison was conducted as well to assess if and where spatial agreements or disagreements occur between the two data sources. The statistical analysis led to a signficant result, indicating that relevant Wikipedia articles also seem to be associated with a higher incidence of Flickr posts nearby and this could indicate that the two sources generally agree with each other, at least for the two classes tested. Visual analysis yielded a more ambivalent picture: Generally, there was an association to be observed but also some disagreements between the data sources. Because the Wikipedia data set is not as spatially precise and is a significantly smaller sample with a much coarser resolution, there are some uncertainties that persist.

The results of this analysis help to identify the limitations of using Wikipedia data and discuss ideas on how to improve the quality of the data in order to achieve more meaningful results. As was mentioned before, there is a semantic overlap between Flickr and Wikipedia, meaning that terms used to identify CES classes in Flickr also occur in Wikipedia posts that are geographically close. However, there might be terminology used on Wikipedia that is not shared with Flickr to describe CES and there is the possibility that one misses important information this way. It was attempted to create word clouds of the whole Wikipedia corpus used in this study to look into the semantics of Wikipedia articles more closely. Unfortunately, processing such a large corpus required excess memory resources to succeed. Studying the semantic content of the Wikipedia corpus in a more detailed manner would enable a more comprehensive understanding of the differences between the two data source's terminology. Another issue that needs to be addressed is the non-standardised data structure of Wikipedia articles. While Wikipedia articles are usually divided into chapters, there is no consistent structure of sections and subsections that is used across all the articles on Wikipedia with a related subject. This makes it more challenging to filter out chapters that are potentially misleading. As observed during the lecture of some Wikipedia articles, there is a lot of information expected about historical structures in chapters referring to history, but there could be the case that these buildings are long gone and have no relevance to the contemporary historic value of a place anymore. Using the HTML scraper, one is able to exclude these sections based on the paragraph title, but there is no apparent rule on how these sections are named and structured and shorter articles sometimes do not have any sections. In Flickr data however, there is a clear data structure which is standardized for every post. This makes it more accessible and easier to exclude certain data, which is much less convenient in full-text data like Wikipedia articles.

Concerning the spatial accuracy of Wikipedia data there is also an issue with how these articles are geographically referenced. Like the coordinates in the Flickr metadata, it is only possible to reference the article with a single point coordinate. While this is an appropriate representation for the Flickr post (spot where the photo is taken of/from) it is fairly vague for many Wikipedia articles, which often refer to larger footprints in reality, where a single point is thus not a sufficiently accurate representation of the true spatial extent. If the articles would be appropriately represented as a planar footprint, it would be easier and more precise to spatially relate Flickr posts with the article. To get access to the footprints, one option would be to source the data on OSM, where there is a Wikidata key or a direct link included in the tags for spatial features that have a an article associated with them. Ballatore and Bertolotto (2011) propose a method where the semantically rich DBpedia ontology can be combined with a spatially elaborate OSM data to be able to spatially query the DBpedia data set. While the purpose of this framework is not to source the Wikipedia article content directly but to provide ontological context to OSM data, the linking of these two data sets in this way could theoretically enable that. Having an accurate spatial representation for each feature would remove a lot of spatial ambiguity from the data set and would enable to conduct more precise spatial analyses.

The *Thematic Accuracy* of Wikipedia data also needs to be discussed in more detail. Wikipedia is an open source project where everyone can contribute and thus there are concerns about the truthfulness of the information contained in the articles. Even in academics, Wikipedia is often regarded as not very trustworthy, despite the number of errors being similar to peer-reviewed papers (Jemielniak, 2019). While everyone can contribute, writing and editing articles is usually a collaborative effort involving multiple authors with various experience levels. The guided collaboration by the editors, the possibility to propose articles for deletion or improvement and the promotion of high-quality articles all help to maintain a generally high quality that it is on-par with established works like the *Encyclopedia Britannica* (Cusinato et al., 2009). Because open source projects such as Wikipedia emphasize transparency, it is possible to analyse the collaborative process and how the editing teams work together on each article and this can serve as another indicator quality (Liu & Ram, 2018). Wikipedia has mechanisms in place that aim to ensure that the information retrieved is accurate, in contrast to Flickr. There is nothing stopping a Flickr user to assign wrong coordinates to a picture or enter misleading tags. Contributing inaccurate information to Wikipedia would be much less likely to go unnoticed because of the collaborative mechanisms mentioned. This integrated quality insurance is a reason why Wikipedia seems suitable as a complementary source for this thesis, as Wikipedia has the ambition to be truthful and is transparent about the processes that work to ensure this. Furthermore, Wikipedia articles establish a consensus on the fundamental characteristics of a place and thus can be viewed as an "expression of the collective perception", whereas social media data represent a more individualistic perspective (Jenkins et al., 2016).

Another significant quality issue is the lack of *Completeness* of the Wikipedia data set. As established in chapter 6, there is almost certainly a lot of data missing. The goal was to retrieve all of the articles in the study area to get as much information as possible about all the locatable articles, which would be a *complete* data set. For this reason, the *type* variable was set as *place*, which is a high-level class in the DBpedia ontology that should not only return settlements but also natural features and historic places (DBpedia, 2022). The data set retrieved did not reflect the large variety of classes that should be part of it based on the ontology mapping of DBpedia. It would have been sensible to include the class *ArchitecturalStructure* as well, as this would have included the missing articles about Bamburgh Castle and the Forth Bridge. But still, there are some buildings included in the data, for example the article about the Hull City Hall. The query for *Place* also did not include nature reserves such as the *RSPB Minsmere*, which is categorized as such in the DBpedia ontology mapping (DBpedia, 2022). It is not clear if there was an error in the query that inadvertently excluded certain articles, or if the database had trouble returning all the relevant articles in the bounding box. In the beginning, the density of articles was regarded as sufficient for the task and only after the analysis it became clear that a bigger sample would be beneficial to the significance of the results. For future studies, this experience should be seen as a cautionary tale and measures should be taken to ensure a more extensive data set is retrieved. For example, multiple specific queries could be run with low-level classes (e.g. architectural structures, natural areas, geomorphological features), that are relevant for the task at hand to ensure that different types of entities are represented in the data set.

When assessing the completeness of the data, one has to consider the various biases that over- or under-represent certain perspectives (Fonte et al., 2017). Compared to the Flickr data set, the Wikipedia articles where distributed more homogeneously along the coast, with each settlement typically being associated with one article. But in large populated areas the density of article locations is usually a bit higher (Fig. 6).

The user demographic of Wikipedia contributors is also strongly skewed. It has come to attention long ago that women are dramatically underrepresented as contributors to articles with less than 15% of authors identifying as female (Glott & Ghosh, 2010). Not only are there less female contributors included, women were under-represented in the talk pages, where collaboration takes place and contributions are discussed (Cabrera et al., 2018). This has effects on the content of the encyclopedia. For instance, biographies of women are less numerous and they are often portrayed differently in the articles than men (Wagner et al., 2021). Even if biographies are not relevant to this study, it illustrates what an influence such biases can have on the content. Regarding age, there is a strong tendency toward younger contributors: 50% of contributors are less than 22 years old with an average age of 22.5 years (Glott & Ghosh, 2010). When comparing these numbers to the Flickr figures (see section 7.1), Wikipedia contributors are much more likely to be male and younger than the Flickr user base. In terms of addressing bias and increasing representativeness, complementing Flickr with Wikipedia data would amplify the biases that already persist in the Flickr data set.

7.3 Research Question 2.1

Do spatial patterns match those found in literature?

For the second research questions, the Flickr data set was assessed for spatial patterns that help understand the characteristics and geographic distribution of CES better. For the first question, this was done purely visually on the basis of QGIS map layers on top of a OSM base map. This was done for the distribution of all data points and the differences in distribution between different classes were analysed as well. In this part the patterns observed will be discussed based on the literature found about preceding studies.

The most conspicuous spatial pattern is higher post density in populated areas. As established in chapter 6, there is usually a bias towards urban areas with rural perspectives usually underrepresented (Hecht & Stephens, 2014; Li et al., 2013).

This is also confirmed by the observations in this study. For instance, northern Scotland has significantly less posts than less remote places further south. In areas outside of settlements Gliozzo et al. (2016) identified three major factors that explain the spatial occurrence of hot-spots. First is the presence of "accessible views over landscapes". Accessibility by roads/paths has been recognized as one of the major factors for explaining clusters of CES (Brown & Hausner, 2017; Depietri et al., 2021). This can be observed in the data as well, data points usually do not stray far from settlements and roads and there often is a car park nearby too. Clusters typically occur where there is a view, may it be beaches, cliffs or piers (Fig. 16). This observation was also made by Ruskule et al. (2018), who found hot-spots around areas with scenic views and quaint geomorphological features. The only class that is found further from human infrastructure is *Nature Appreciation*. Clusters tend to occur in places that are undisturbed by humans, such as nature reserves and marshlands. This observation has been shared by Runge et al. (2020) above the arctic circle, where nature photographers also were moving further away from roads than other user types to take their pictures. Another spot that is remote and popular for nature enthusiasts are islands (Kobryn et al., 2018). This was very noticeable at the British East Coast too, for example on the Farne Islands off Bamburgh (Fig. 16), as well as Fife Island east off the Firth of Forth. The photographs in these places usually featured various sea birds like Puffins and Seagulls. The second factor for clustering of posts is the presence of historic human artifacts such as castles and ruins that attract a lot of visitors (Gliozzo et al., 2016). This does agree with the general observation that sea side castles are a major point of attraction. Touristic places are very common in the Flickr data set and are usually over-represented as posts of picturesque places attract more visitors, which leads to positive feedback effect on the number of posts in these areas (Depietri et al., 2021). The third factor identified was the presence of areas with an emphasis on biodiversity (Gliozzo et al., 2016). This was partially observed in the study area as well, but there was no consistent patten. Nature reserves did occasionally contain clusters of nature related posts, but there were just as many clusters close or within cities. A pattern that is immediately visible when looking at the data on a map is the very close distance from the shore, where most Flickr posts are located. The buffer distance of one kilometre seems to be very generous, as most posts come to lie not further than half this distance from the coast, especially for the posts classed as *Landscape Appreciation*. The coastline and the sea were a very popular motif found during annotation, so this does not surprise. In similar studies, this distance was usually shorter. Santos Vieira et al. (2021) used a buffer only 100 metres wide (50m inland, 50m offshore) while Cao et al. (2022) used one that reaches 500 metres inland based on the convenient distance to reach a beach by pedestrians. This dependence on distance to the coast is concurrent with the findings of other studies which also find a very sharp increase of CES provision at the coastline (Brown & Hausner, 2017; Kobryn et al., 2018). If this study were to be repeated, the buffer distance would thus be significantly shorter.

In the study area of this thesis there are thus similar spatial patterns to be found that as were observed in literature and thus confirms that Flickr data can be used as a valid indicator for CES provision at the British east coast. To be more certain about these patterns and to get more detailed insights into the spatial distribution of CES there are more quantitative studies needed. For example one could compare urban areas against rural areas or look at distance to human infrastructure (e.g. car parks, roads) or landscape features.

7.4 Research Question 2.2.

Are there bundles of CES that can be identified?

To answer this question, a more quantifiable approach has been used combined with a visual analysis of the data on a map. For the quantitative analysis, the posts were aggregated into 2 kilometre grid cells by class, then the grid counts of each class were compared with each other and a Pearson's ρ was calculated to quantify the cross correlation between the classes (Table 7). To find out about why these cross-correlations occur, it is necessary to look at the data and the spatial distributions to find out the reasons behind it.

There was a high degree of correlation between the classes *Historical Monuments* and *Landscape Appreciation* to be seen with a ρ of around 0.7. The correlation between *Nature Appreciation* and the other classes was not as strong at 0.5. A certain correlation was expected because the posts usually cluster in and around settlements. The increased correlation between *Landscape Appreciation* and *Historical Monuments* could be explained by the importance of historical structures being a popular element in pictures as well, as was already discussed in chapter 6. CES bundles in the coastal and marine environment are considered common and also increasingly being researched (Milcu et al., 2013; Rodrigues Garcia et al., 2017). Other studies did not find the same CES bundles that were found in this analysis (Mouttaki & Erraiss, 2022; Retka et al., 2019). However, Retka et al. (2019) did research in an area off the shore and for this reason their study might not be comparable. There could also be differences in posting behaviours and cultural differences at play that could explain the differing results between the studies. Furthermore, there is not a single standardized CES classification scheme used in the studies, which makes comparisons difficult. Studying CES bundles with this method seems to be viable and could be interesting for future research as way to assess the synergies and trade-offs that might exist between CES and other Ecosystem Services (Plieninger et al., 2013; Rodrigues Garcia et al., 2017).

8 Conclusion

The main goal of this thesis is to classify CES classes using Flickr meta data with a Random Forest machine learning algorithm and to spatially compare the information content for some of the CES classes with Wikipedia articles. As a secondary research objective, the Flickr data has been analysed for spatial patterns and bundles of CES in the research area at the East Coast of Great Britain.

The classification of the Flickr posts did work reasonably well, though only a subset of the classes in classification scheme were assessed, which were *Landscape Appreciation*, *Nature Appreciation* and *Historical Monuments*. Working with the user-assigned tags in combination of related word lists, so called *features*, was sufficient to reach a relatively accurate classification result of about 75% true-positive rate across all classes. It produced meaningful visual results on the map as well. During the implementation of this classifier, it was found that some posts represent different CES at once, whereas the classification method can only assign a single class. It appears, that CES classes are not always clearly delineated and can appear together at the same place. This underlines the importance of bundles for the spatial research of CES.

The feature lists used for the Flickr classification were used to search the Wikipedia articles with the terms contained therein and add their counts together for each article. This was done for the classes *Nature Appreciation* and *Historical Monuments* seperately. It showed that Wikipedia uses similar terminology as Flickr to describe places. It was found that the top 25% of articles containing the highest sums of terms were associated with significantly more Flickr posts in the vicinity of 2 kilometres, than the lower 75% for both classes. Spatially there was some overlap to be observed with Flickr data as well, matching less in the northern part of the study area. However, this result comes with various uncertainties to keep in mind. For one, the Wikipedia data set is missing articles that could have had relevance to the CES assessed. There were also issues with the spatial representation of Wikipedia articles, which does not allow for more complex geometries than point coordinates and thus introduces uncertainties into spatial relations between Flickr posts and Wiki articles. Wikipedia data can not be considered representative of the general population as the contributor demographic is rather young and male by a large majority, the same bias that is found in the Flickr user base. Using Wikipedia as a data source for mapping and assessing ecosystems in conjuncture with other data sources could potentially be viable, at least for the CES classes assessed in this analysis. However the issues mentioned need to be addressed to confirm this conclusion with more confidence.

The spatial patterns that could be visually observed in Flickr data mostly agreed with patterns found in literature about PPGIS and VGI studies at the coast and inland. The main observation is the uneven distribution of posts between urban and rural places, with cities being major hot spots. Outside of densely populated places, data points mainly cluster around places that offer views over landscapes (dunes, cliffs, piers, beaches) and picturesque human made objects like ruins and castles. Places with a significance for biodiversity were a major point of attraction as well. These locations are often close to clusters of Flickr post with the related CES class and thus confirm that the classification process returned accurate results. Other observations concluded that Flickr posts are mostly located close to human infrastructure. Furthermore, a large part of Flickr posts were located within 500 metres to the coastline, especially those classed as *Landscape Appreciation*. The agreement between literature and the classification of Flickr posts validates the use of Flickr as a data source for spatially analysing CES in a coastal context. However, the bias towards urban perspectives and popular touristic places needs to be considered to not disregard the perspective of rural inhabitants in the area.

Less conclusive results were achieved by analysing the bundles of CES occurring all along the coast. The Pearson's ρ coefficient was used to compare the spatial correlation between the counts of each CES class, aggregated in two-kilometre grid cells. *Landscape Appreciation* and *Historical Monuments* did spatially correlate more than any other combination. This did not agree with related literature, which did not indicate any significant correlation between those classes. However, by assessing the clusters visually and looking at photographs, it could be observed that historical structures are not only attractive because of their heritage value, but also because they are recognized as landscape feature as well. Thus, this might be the cause why these classes co-occur more often. Comparing bundles of CES in a coastal setting could be a promising research topic and would offer more insights if further CES classes were to be included as well.

References

Agardy, T., Alder, J., Dayton, P., Curran, S., Kitchingman, M., Wilson, M., & Saifullah, S. (2005). Coastal systems. *Ecosystems and Human Well-being: Current State and Trends, vol. 1* (pp. 513–549).

Ahtiainen, H., Liski, E., Pouta, E., Soini, K., Bertram, C., Rehdanz, K., Pakalniete, K., & Meyerhof, J. (2019). Cultural ecosystem services provided by the Baltic Sea marine environment. *Ambio*, *48*(11), 1350–1361. https://doi.org/10.1007/s13280-019-01239-1

Alivand, M., & Hochmair, H. H. (2017). Spatiotemporal analysis of photo contribution patterns to Panoramio and Flickr. *Cartography and Geographic Information Science*, *44*(2), 170–184. https://doi.org/10.1080/15230406.2016.1211489

Azevedo, A. K., Vieira, F. A., Guedes-Santos, J., Gaia, J. A., Pinheiro, B. R., Bragagnolo, C., Correia, R. A., Ladle, R. J., & Malhado, A. C. (2022). A big data approach to identify the loss of coastal cultural ecosystem services caused by the 2019 Brazilian oil spill disaster. *Anais da Academia Brasileira de Ciencias*, *94*, 1–11. https://doi.org/10.1590/0001-3765202220210397

Ballatore, A., & Bertolotto, M. (2011). Semantically Enriching VGI in Support of Implicit Feedback Analysis. *6574*, 78–93. https://doi.org/10.1007/978-3-642-19173-2_8

Beatty, C., Fothergill, S., & Gore, T. (2014). *The Seaside Tourist Industry in England and Wales Employment, economic output, location and trends.*

Belyi, A., Bojic, I., Sobolevsky, S., Sitko, I., Hawelka, B., Rudikova, L., Kurbatski, A., & Ratti, C. (2017). Global multi-layer network of human mobility. *International Journal of Geographical Information Science*, *31*(7), 1381–1402. https://doi.org/10.1080/13658816.2017.1301455

Bennett, E. M. (2017). Research Frontiers in Ecosystem Service Science. *Ecosystems*, *20*(1), 31–37. https://doi.org/10.1007/s10021-016-0049-0

Benoit, K., & Matsuo, A. (2020). *spacyr: Wrapper to the 'spaCy' 'NLP' Library* [R package version 1.2.1]. https://CRAN.R-project.org/package=spacyr

Biau, G., & Scornet, E. (2016). A random forest guided tour. *Test*, *25*(2), 197–227. https://doi.org/10.1007/s11749-016-0481-7

Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5–32.

Brown, G., & Fagerholm, N. (2015). Empirical PPGIS/PGIS mapping of ecosystem services: A review and evaluation. *Ecosystem Services*, *13*, 119–133. https://doi.org/10.1016/j.ecoser.2014.10.007

Brown, G., & Hausner, V. H. (2017). An empirical analysis of cultural ecosystem values in coastal landscapes. *Ocean Coastal Management*, *142*, 49–60. https://doi.org/10.1016/j.ocecoaman.2017.03.019

Brown, G., Montag, J. M., & Lyon, K. (2012). Public Participation GIS: A Method for Identifying Ecosystem Services. *Society and Natural Resources*, *25*(7), 633–651. https://doi.org/10.1080/08941920.2011.621511

Brown, G., & Reed, P. (2009). Public Participation GIS: A New Method for Use in National Forest Planning. *Forest Sci.*, *55*.

Cabrera, B., Ross, B., Dado, M., & Heisel, M. (2018). The gender gap in Wikipedia talk pages. *Twelfth International AAAI Conference on Web and Social Media.*

Cao, H., Wang, M., Su, S., & Kang, M. (2022). Explicit quantification of coastal cultural ecosystem services: A novel approach based on the content and sentimental analysis of social media. *Ecological Indicators*, *137*, 108756. https://doi.org/10.1016/j.ecolind.2022.108756

Carranza, M. L., Drius, M., Marzialetti, F., Malavasi, M., de Francesco, M. C., Acosta, A. T., & Stanisci, A. (2020). Urban expansion depletes cultural ecosystem services: an insight into a Mediterranean coastline. *Rendiconti Lincei*, *31*(1), 103–111. https://doi.org/10.1007/s12210-019-00866-w

Cheng, X., Van Damme, S., Li, L., & Uyttenhove, P. (2019). Evaluation of cultural ecosystem services: A review of methods. *37*(June 2018), 100925. https://doi.org/10.1016/j.ecoser.2019.100925

Chesnokova, O., Nowak, M., & Purves, R. S. (2017). A crowdsourced model of landscape preference. *Leibniz International Proceedings in Informatics, LIPIcs*, *86*(19), 1–13. https://doi.org/10.4230/LIPIcs.COSIT.2017.19

Chesnokova, O., & Purves, R. (2018). From image descriptions to perceived sounds and sources in landscape: Analyzing aural experience through text. *Applied Geography*, *93*(December), 103–111. https://doi.org/10.1016/j.apgeog.2018.02.014

CIESIN. (2012). National Aggregates of Geospatial Data Collection: Population, Landscape, And Climate Estimates, Version 3 (PLACE III).

Clemente, P., Calvache, M., Antunes, P., Santos, R., Cerdeira, J. O., & Martins, M. J. (2019). Combining social media photographs and species distribution models to map cultural ecosystem services: The case of a Natural Park in Portugal. *Ecological Indicators*, *96*, 59–68. https://doi.org/10.1016/j.ecolind.2018.08.043

Costanza, R., D'Arge, R., de Groot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., O'Neill, R. V., Paruelo, J., Raskin, R. G., Sutton, P., & van den Belt, M. (1997). The value of the world's ecosystem services and natural capital. *Nature*, *387*(6630), 253–260. https://doi.org/10.1038/387253a0

Costanza, R., de Groot, R., Braat, L., Kubiszewski, I., Fioramonti, L., Sutton, P., Farber, S., & Grasso, M. (2017). Twenty years of ecosystem services: How far have we come and how far do we still need to go? *Ecosystem Services*, *28*, 1–16. https://doi.org/10.1016/j.ecoser.2017.09.008

Crain, C. M., Halpern, B. S., Beck, M. W., & Kappel, C. V. (2009). Understanding and Managing Human Threats to the Coastal Marine Environment. *Annals of the New York Academy of Sciences*, *1162*(1), 39–62. https://doi.org/10.1111/j.1749-6632.2009.04496.x

Cui, N., Malleson, N., Houlden, V., & Comber, A. (2021). Using VGI and Social Media Data to Understand Urban Green Space: A Narrative Literature Review. *ISPRS International Journal of Geo-Information*, *10*(7). https://doi.org/10.3390/ijgi10070425

Cusinato, A., Della Mea, V., Di Salvatore, F., & Mizzaro, S. (2009). QuWi: Quality Control in Wikipedia, 27–34. https://doi.org/10.1145/1526993.1527001

Daily, G. C. et al. (1997). Introduction: what are ecosystem services. *Nature's services: Societal dependence on natural ecosystems*, *1*(1).

DBpedia. (2022). *DBpedia Ontology Classes*. Retrieved September 13, 2022, from http://mappings.dbpedia.org/server/ontology/classes/

Depietri, Y., Ghermandi, A., Campisi-Pinto, S., & Orenstein, D. E. (2021). Public participation GIS versus geolocated social media data to assess urban cultural ecosystem services: Instances of complementarity. *Ecosystem Services*, *50*(June), 101277. https://doi.org/10.1016/j.ecoser.2021.101277

Drius, M., Bongiorni, L., Depellegrin, D., Menegon, S., Pugnetti, A., & Stifter, S. (2019). Tackling challenges for Mediterranean sustainable coastal tourism: An ecosystem service perspective. *Science of the Total Environment*, *652*, 1302–1317. https://doi.org/10.1016/j.scitotenv.2018.10.121

Egarter Vigl, L., Marsoner, T., Giombini, V., Pecher, C., Simion, H., Stemle, E., Tasser, E., & Depellegrin, D. (2021). Harnessing artificial intelligence technology and social media data to support Cultural Ecosystem Service assessments. *People and Nature*, *3*(3), 673–685. https://doi.org/10.1002/pan3.10199

Fish, R., Church, A., & Winter, M. (2016). Conceptualising cultural ecosystem services: A novel framework for research and critical engagement. *Ecosystem Services*, *21*(September), 208–217. https://doi.org/10.1016/j.ecoser.2016.09.002

Flickr Development Team. (2014). *Flickr API Documentation.* Retrieved September 20, 2022, from https://www.flickr.com/services/api/

Fonte, C. C., Antoniou, V., Bastin, L., Estima, J., Arsanjani, J. J., Bayas, J.-C. L., See, L., & Vatseva, R. (2017). Assessing VGI Data Quality. *Mapping and the citizen sensor* (pp. 137–164). Ubiquity Press.

Fox, N., Graham, L. J., Eigenbrod, F., Bullock, J. M., & Parks, K. E. (2021). Enriching social media data allows a more robust representation of cultural ecosystem services. *Ecosystem Services*, *50*(June), 101328. https://doi.org/10.1016/j.ecoser.2021.101328

Gee, K., & Burkhard, B. (2010). Cultural ecosystem services in the context of offshore wind farming: A case study from the west coast of Schleswig-Holstein. *Ecological Complexity*, *7*(3), 349–358. https://doi.org/10.1016/j.ecocom.2010.02.008

Gliozzo, G., Pettorelli, N., & Muki Haklay, M. (2016). Using crowdsourced imagery to detect cultural ecosystem services: A case study in South Wales, UK. *Ecology and Society*, *21*(3). https://doi.org/10.5751/ES-08436-210306

Glott, R., & Ghosh, R. (2010). *Analysis of Wikipedia Survey Data.* Retrieved September 20, 2022, from https://www.merit.unu.edu/wp-content/uploads/2019/03/Wikipedia_Overview_15March2010-FINAL.pdf

Goodchild, M. (2007). Citizens as Sensors: The World of Volunteered Geography. *GeoJournal*, *69*, 211–221. https://doi.org/10.1007/s10708-007-9111-y

Gosal, A. S., Geijzendorffer, I. R., Václavík, T., Poulin, B., & Ziv, G. (2019). Using social media, machine learning and natural language processing to map multiple recreational beneficiaries. *Ecosystem Services*, *38*. https://doi.org/10.1016/j.ecoser.2019.100958

Granek, E. F., Polasky, S., Kappel, C. V., Reed, D. J., Stoms, D. M., Koch, E. W., Kennedy, C. J., Cramer, L. A., Hacker, S. D., Barbier, E. B., Aswani, S., Ruckelshaus, M., Perillo, G. M., Silliman, B. R., Muthiga, N., Bael, D., & Wolanski, E. (2010). Ecosystem services as a common language for coastal ecosystem-based management. *Conservation biology : the journal of the Society for Conservation Biology*, *24*(1), 207–216. https://doi.org/10.1111/j.1523-1739.2009.01355.x

Guerrero, P., Møller, M. S., Olafsson, A. S., & Snizek, B. (2016). Revealing Cultural Ecosystem Services through Instagram Images: The Potential of Social Media Volunteered Geographic Information for Urban Green Infrastructure Planning and Governance. *Urban Planning*, *1*(2), 1–17. https://doi.org/10.17645/up.v1i2.609

Haines-Young, R., & Potschin, M. (2012). *Common International Classification of Ecosystem Services (CICES, Version 4.1).* School of Geography, University of Nottingham. Nottingham.

Hale, R. L., Cook, E. M., & Beltrán, B. J. (2019). Cultural ecosystem services provided by rivers across diverse social-ecological landscapes: A social media analysis. *Ecological Indicators*, *107*(July), 105580. https://doi.org/10.1016/j.ecolind.2019.105580

Hartmann, M. C., Koblet, O., Baer, M. F., & Purves, R. S. (2022). Automated motif identification: Analysing Flickr images to identify popular viewpoints in Europe's protected areas. *Journal of Outdoor Recreation and Tourism*, *37*, 100479. https://doi.org/10.1016/j.jort.2021.100479

Hauff, C. (2013). A study on the accuracy of Flickr's geotag data. *SIGIR 2013 - Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1037–1040. https://doi.org/10.1145/2484028.2484154

Havinga, I., Bogaart, P. W., Hein, L., & Tuia, D. (2020). Defining and spatially modelling cultural ecosystem services using crowdsourced data. *Ecosystem Services*, *43*(March), 101091. https://doi.org/10.1016/j.ecoser.2020.101091

Havinga, I., Marcos, D., Bogaart, P. W., Hein, L., & Tuia, D. (2021a). Geo-Data for Mapping Scenic Beauty: Exploring the Potential of Remote Sensing and Social Media. *International Geoscience and Remote Sensing Symposium (IGARSS)*, *2021-July*, 188–191. https://doi.org/10.1109/IGARSS47720.2021.9553417

Havinga, I., Marcos, D., Bogaart, P. W., Hein, L., & Tuia, D. (2021b). Social media and deep learning capture the aesthetic quality of the landscape. *Scientific Reports*, *11*(1), 1–11. https://doi.org/10.1038/s41598-021-99282-0

Hecht, B., & Stephens, M. (2014). A Tale of Cities : Urban Biases in Volunteered Geographic Information. *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*, 197–205.

Hernández-Morcillo, M., Plieninger, T., & Bieling, C. (2013). An empirical review of cultural ecosystem service indicators. *Ecological Indicators*, *29*, 434–444. https://doi.org/10.1016/j.ecolind.2013.01.013

Hu, Y. (2018). Geo-text data and data-driven geospatial semantics. *Geography Compass*, *12*(11). https://doi.org/10.1111/gec3.12404

International Hydrographic Organization. (1953). Limits of Oceans and Seas.

Jemielniak, D. (2019). Wikipedia: Why is the common knowledge resource still neglected by academics? *GigaScience*, *8*(12).

Jenkins, A., Croitoru, A., Crooks, A. T., & Stefanidis, A. (2016). Crowdsourcing a Collective Sense of Place. (T. Preis, Ed.). *PloS one*, *11*(4), e0152932. https://doi.org/10.1371/journal.pone.0152932

JNCC. (2003). *Coastal geomorphology of Great Britain*. Joint Nature Conservation Committee.

Kobryn, H. T., Brown, G., Munro, J., & Moore, S. A. (2018). Cultural ecosystem values of the Kimberley coastline: An empirical analysis with implications for coastal and marine policy. *Ocean and Coastal Management*, *162*, 71–84. https://doi.org/10.1016/j.ocecoaman.2017.09.002

Kosanic, A., & Petzold, J. (2020). A systematic review of cultural ecosystem services and human wellbeing. *Ecosystem Services*, *45*, 101168. https://doi.org/10.1016/j.ecoser.2020.101168

Lehmann, J., Isele, R., Jakob, M., Jentzsch, A., Kontokostas, D., Mendes, P. N., Hellmann, S., Morsey, M., Van Kleef, P., Auer, S., & Bizer, C. (2015). DBpedia - A large-scale, multilingual knowledge base extracted from Wikipedia. *Semantic Web*, *6*(2), 167–195. https://doi.org/10.3233/SW-140134

Li, L., Goodchild, M. F., & Xu, B. (2013). Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr. *Cartography and Geographic Information Science*, *40*(2), 61–77. https://doi.org/10.1080/15230406.2013.777139

Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. *R News*, *2*(3), 18–22.

Liaw, A., Wiener, M., Breiman, L., & Cutler, A. (2022). *randomForest: Breiman and Cutler's Random Forests for Classification and Regression* [R package version 4.7-1.1.].

Light, D., & Chapman, A. (2022). The neglected heritage of the English seaside holiday. *Coastal Studies & Society*, *1*(1), 34–54. https://doi.org/10.1177/26349817211070456

Liu, J., & Ram, S. (2018). Using big data and network analysis to understand Wikipedia article quality. *Data Knowledge Engineering*, *115*, 80–93. https://doi.org/10.1016/j.datak.2018.02.004

Lopez, B. E., Magliocca, N. R., & Crooks, A. T. (2019). Challenges and Opportunities of Social Media Data for Socio-Environmental Systems Research. *Land*, *8*(7). https://doi.org/10.3390/land8070107

Lu, Y., Yuan, J., Lu, X., Su, C., Zhang, Y., Wang, C., Cao, X., Li, Q., Su, J., Ittekkot, V., Garbutt, R. A., Bush, S., Fletcher, S., Wagey, T., Kachur, A., & Sweijd, N. (2018). Major threats of pollution and climate change to global coastal ecosystems and enhanced management for sustainability. *Environmental Pollution*, *239*, 670–680. https://doi.org/10.1016/j.envpol.2018.04.016

Marsh, D. (2018). *Accounting for National Nature Reserves* (tech. rep.). Natural England. https://doi.org/10.13140/RG.2.2.26163.07209

Martin, C. L., Momtaz, S., Gaston, T., & Moltschaniwskyj, N. A. (2016). A systematic quantitative review of coastal and marine cultural ecosystem services: Current status and future research. *Marine Policy*, *74*, 25–32. https://doi.org/10.1016/j.marpol.2016.09.004

Maul, G. A., & Duedall, I. W. (2019). Demography of Coastal Populations. In C. W. Finkl & C. Makowski (Eds.), *Encyclopedia of coastal science* (pp. 692–700). Springer International Publishing. https://doi.org/10.1007/978-3-319-93806-6_115

Milcu, A. I., Hanspach, J., Abson, D., & Fischer, J. (2013). Cultural ecosystem services: A literature review and prospects for future research. *Ecology and Society*, *18*(3). https://doi.org/10.5751/ES-05790-180344

Millenium Ecosystem Assessment. (2005). *Ecosystems and human well-being*. Washington, DC.

Mouttaki, I., & Erraiss, M. (2022). *Classifying and mapping cultural ecosystem service using artificial intelligence and social media data* [preprint on webpage at researchsquare.com/article/rs-1765424/v1].

Natural England. (2020). *Areas of Outstanding Natural Beauty (England)*. Retrieved August 18, 2022, from https://naturalengland-defra.opendata.arcgis.com/datasets/Defra::areas-of-outstanding-natural-beauty-england/about

NatureScot. (2022). *Marine Protected Areas (MPA)*. Retrieved August 18, 2022, from https://www.nature.scot/professional-advice/protected-areas-and-species/protected-areas/marine-protected-areas-mpas

Oteros-Rozas, E., Martín-López, B., Fagerholm, N., Bieling, C., & Plieninger, T. (2018). Using social media photos to explore the relation between cultural ecosystem services and landscape features across five European sites. *Ecological Indicators*, *94*, 74–86. https://doi.org/10.1016/j.ecolind.2017.02.009

Plieninger, T., Dijks, S., Oteros-Rozas, E., & Bieling, C. (2013). Assessing, mapping, and quantifying cultural ecosystem services at community level. *Land Use Policy*, *33*, 118–129. https://doi.org/10.1016/j.landusepol.2012.12.013

Purves, R. S., Edwardes, A. J., & Wood, J. (2011). View of Describing place through user generated content. *First Monday*, *16*(9). https://journals.uic.edu/ojs/index.php/fm/article/view/3710/3035

Reid, W., Mooney, H., Cropper, A., Capistrano, D., Carpenter, S., Chopra, K., Dasgupta, P., Dietz, T., Duraiappah, A., Hassan, R., Kasperson, R., Leemans, R., May, R., Mcmichael, A., Pingali, P., Samper, C., Scholes, R., Watson, R., Zakri, A., & Zurek, M. (2005). *Millenium Ecosystem Assessment Synthesis Report*.

Retka, J., Jepson, P., Ladle, R. J., Malhado, A. C., Vieira, F. A., Normande, I. C., Souza, C. N., Bragagnolo, C., & Correia, R. A. (2019). Assessing cultural ecosystem services of a large marine protected area through social media photographs. *Ocean and Coastal Management*, *176*(July), 40–48. https://doi.org/10.1016/j.ocecoaman.2019.04.018

Rodrigues Garcia, J., Conides, A. J., Rodriguez Rivero, S., Raicevich, S., Pita, P., Kleisner, K. M., Pita, C., Lopes, P. F., Roldán Alonso, V., Ramos, S. S., Klaoudatos, D., Outeiro, L., Armstrong, C., Teneva, L., Stefanski, S., Böhnke-Henrichs, A., Kruse, M., Lillebø, A. I., Bennett, E. M., . . . Villasante, S. (2017). Marine and coastal cultural ecosystem services: Knowledge gaps and research priorities. *One Ecosystem*, *2*, e12290. https://doi.org/10.3897/oneeco.2.e12290

Ruiz-Frau, A., Hinz, H., Edwards-Jones, G., & Kaiser, M. (2013). Spatially explicit economic assessment of cultural ecosystem services: Non-extractive recreational uses of the coastal environment related to marine biodiversity. *Marine Policy*, *38*, 90–98. https://doi.org/10.1016/j.marpol.2012.05.023

Ruiz-Frau, A., Ospina-Alvarez, A., Villasante, S., Pita, P., Maya-Jariego, I., & de Juan, S. (2020). Using graph theory and social media data to assess cultural ecosystem services in coastal areas: Method development and application. *Ecosystem Services*, *45*(July), 101176. https://doi.org/10.1016/j.ecoser.2020.101176

Runge, C. A., Hausner, V. H., Daigle, R. M., & Monz, C. A. (2020). Pan-arctic analysis of cultural ecosystem services using social media and automated content analysis. *Environmental Research Communications*, *2*(7). https://doi.org/10.1088/2515-7620/ab9c33

Ruskule, A., Klepers, A., & Veidemane, K. (2018). Mapping and assessment of cultural ecosystem services of Latvian coastal areas. *One Ecosystem*, *3*. https://doi.org/10.3897/oneeco.3.e25499

Santos Vieira, F. A., Vinhas Santos, D. T., Bragagnolo, C., Campos-Silva, J. V., Henriques Correia, R. A., Jepson, P., Mendes Malhado, A. C., & Ladle, R. J. (2021). Social media data reveals multiple cultural services along the 8.500 kilometers of Brazilian coastline. *Ocean Coastal Management*, *214*(March), 105918. https://doi.org/10.1016/j.ocecoaman.2021.105918

Scotland's NNR. (2022). *Search by Map*. Retrieved August 18, 2022, from https://www.nnr.scot/Search_by_Map

Sijtsma, F. J., Mehnen, N., Angelstam, P., & Muñoz-Rojas, J. (2019). Multi-scale mapping of cultural ecosystem services in a socio-ecological landscape: A case study of the international Wadden Sea Region. *Landscape Ecology*, *34*(7), 1751–1768. https://doi.org/10.1007/s10980-019-00841-8

SimilarWeb. (2022). *Top Websites Ranking*. Retrieved September 18, 2022, from https://www.similarweb.com/top-websites/

Smart, L. S., Vukomanovic, J., Sills, E. O., & Sanchez, G. (2021). Cultural ecosystem services caught in a 'coastal squeeze' between sea level rise and urban expansion. *Global Environmental Change*, *66*(November 2020), 102209. https://doi.org/10.1016/j.gloenvcha.2020.102209

T R, P. (2015). A Comparative Study on Decision Tree and Random Forest Using R Tool. *IJARCCE*, *4*(1), 196–199. https://doi.org/10.17148/IJARCCE.2015.4142

Taff, B. D., Benfield, J., Miller, Z. D., D'Antonio, A., & Schwartz, F. (2019). The role of tourism impacts on cultural ecosystem services. *Environments - MDPI*, *6*(4), 1–12. https://doi.org/10.3390/environments6040043

The Wildlife Trusts. (2022). *Marine Protected Areas in England*. Retrieved August 18, 2022, from https://www.wildlifetrusts.org/marine-protected-areas/england

University of Zurich. (2022). *UZH Methodenberatung - Datenanalyse mit SPSS*. Retrieved July 7, 2022, from https://www.methodenberatung.uzh.ch/de/datenanalyse_spss.html

Van Zanten, B. T., Van Berkel, D. B., Meentemeyer, R. K., Smith, J. W., Tieskens, K. F., & Verburg, P. H. (2016). Continental-scale quantification of landscape values using social media data. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(46), 12974–12979. https://doi.org/10.1073/pnas.1614158113

Wagner, C., Garcia, D., Jadidi, M., & Strohmaier, M. (2021). It's a Man's Wikipedia? Assessing Gender Inequality in an Online Encyclopedia. *Proceedings of the International AAAI Conference on Web and Social Media*, *9*(1), 454–463.

Wickham, H. (2021). *rvest: Easily Harvest (Scrape) Web Pages* [R package version 1.0.2]. https://CRAN.R-project.org/package=rvest

Wikipedia. (2022). *Wikipedia: Size of Wikipedia*. Retrieved September 18, 2022, from en.wikipedia.org/wiki/Wikipedia:Size_of_Wikipedia

Wright, M. N., & Ziegler, A. (2017). Ranger: A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software*, *77*(1), 1–17. https://doi.org/10.18637/jss.v077.i01

Zandbergen, P. A., & Barbeau, S. J. (2011). Positional Accuracy of Assisted GPS Data from High-Sensitivity GPS-enabled Mobile Phones. *Journal of Navigation*, *64*(3), 381–399. https://doi.org/10.1017/S0373463311000051

Zielstra, D., & Hochmair, H. H. (2013). Positional accuracy analysis of Flickr and Panoramio images for selected world regions. *Journal of Spatial Science*, *58*(2), 251–273. https://doi.org/10.1080/14498596.2013.801331

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.

Nicolas Steinmann, September 2022